

Academic Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
225 Wyman Street, Waltham, MA 02451, USA

First edition 2014

Copyright © 2014 Elsevier Ltd. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at <http://elsevier.com/locate/permissions>, and selecting Obtaining permission to use Elsevier material.

Notice

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

Library of Congress Cataloging in Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-12-396500-4

For information on all Elsevier publications
visit our website at www.store.elsevier.com

Printed and bound in Poland.

14 15 16 17 10 9 8 7 6 5 4 3 2 1



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Authors Biography

CHAPTER 2

Mounir Ghogho received the Ph.D. degree in 1997 from the National Polytechnic Institute of Toulouse, France. He was an EPSRC Research Fellow with the University of Strathclyde, Glasgow (Scotland), from September 1997 to November 2001. Since December 2001, he has been a faculty member with the school of Electronic and Electrical Engineering at the University of Leeds (UK), where he is currently a Professor. He is also currently a Professor at the International University of Rabat (Morocco). He served as an Associate Editor of the IEEE Signal Processing Letters from 2001 to 2004, the IEEE Transactions on Signal Processing from 2005 to 2008, and the Elsevier Digital Signal Processing journal from 2011 to 2012. He served as a member of the IEEE Signal Processing Society SPCOM Technical Committee from 2005 to 2010, a member of IEEE Signal Processing Society SPTM Technical Committee from 2006 to 2011, and is currently a member of the IEEE Signal Processing Society SAM Technical Committee. He was the general co-chair of the eleventh IEEE workshop on Signal Processing for Advanced Wireless Communications (SPAWC'2010), the technical co-chair of the MIMO symposium of IWCNC 2007 and IWCNC 2008, and a technical area co-chair of Eusipco 2008, Eusipco 2009, and ISCCSP'05. He is the general co-chair of Eusipco 2013. He was the guest co-editor of special issues of the EURASIP Journal on Wireless Communications and Networking and the Elsevier Physical Communications Journal. His research interests are in signal processing and communication. He held invited scientist/professor positions at many institutions including the US Army Research Lab (USA), Telecom Paris-Tech (France), National Institute of Informatics (Japan), University Carlos 3rd of Madrid (Spain), ENSICA (France), Darmstadt Technical University (Germany), and Beijing University of Posts and Telecommunication (China). He was awarded the prestigious five-year Royal Academy of Engineering Research Fellowship in September 2000.



Philippe Ciblat was born in Paris, France, in 1973. He received the Engineering degree from Telecom ParisTech (formerly, Ecole Nationale Supérieure des Télécommunications—ENST) and the M.Sc. degree in automatic control and signal processing from the University of Paris-Sud, both in 1996, and the Ph.D. degree from University of Paris-Est at Marne-la-Vallée in 2000. He eventually received the HDR degree from the University of Paris-Est at Marne-la-Vallée in 2007. In 2001, he was a Postdoctoral Researcher with University of Louvain, Belgium. Then, he joined the Communications and Electronics Department at Télécom ParisTech, as an Associate Professor. Since 2011, he has been (full) Professor in the same institution. He also is the Head of Digital Communications Group. He is Technical program committee (TPC) member for several flagship conferences (ICASSP, EUSIPCO, GLOBECOM, etc). He served as Associate Editor for the IEEE Communications Letters from 2004 to 2007. He served as Associate Editor (2008–2012) and then Senior Area Editor (2011–2012) for the IEEE Transactions on Signal Processing. His research areas include



statistical signal processing (blind equalization, frequency estimation, and distributed estimation), signal processing for digital communications (synchronization), and resource allocation (multiple access technique optimization, power allocation).

Ananthram Swami received the B.Tech. degree from Indian Institute of Technology (IIT), Bombay; the M.S. degree from Rice University, Houston, TX, and the Ph.D. degree from the University of Southern California (USC), Los Angeles, all in electrical engineering. He has held positions with Unocal Corporation, USC, CS-3, and Malgudi Systems. He was a Statistical Consultant to the California Lottery, developed a Matlab-based toolbox for non-Gaussian signal processing, and has held visiting faculty positions at INP, Toulouse, France. He is with the US. Army Research Laboratory (ARL) where he is the ST for Network Science. His work is in the broad area of network science, with emphasis on wireless communication networks. He was the co-editor of “Wireless Sensor Networks: Signal Processing and Communications Perspectives” (New York: Wiley, 2007). He is a member of the IEEE SPS Technical Committee on Sensor Array and Multi-channel systems, and serves on the Senior Editorial Board of the IEEE Journal on Selected Topics in Signal Processing. He is an IEEE Fellow and an ARL Fellow. He has served as an Associate Editor for the IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, Signal Processing Magazine, the IEEE Transactions on Circuits and Systems II, the IEEE Transactions on Wireless Communications, and as Guest Editor for the IEEE Journal on Selected Areas in Communications. He was a tutorial speaker on “Networking Cognitive Radios for Dynamic Spectrum Access” at ICASSP 2008, DySpan 2008, MILCOM 2008, and ICC 2010. He received the best conference paper award at IEEE Trustcom 2008, and was co-Organizer and co-Chair of three IEEE workshops related to signal processing and communications, including IEEE SPAWC’10.



CHAPTER 3

Jitendra K. Tugnait received the B.Sc. (Hons.) degree in electronics and electrical communication engineering from the Punjab Engineering College, Chandigarh, India in 1971, the M.S. and the E.E. degrees from Syracuse University, Syracuse, NY and the Ph.D. degree from the University of Illinois, Urbana-Champaign in 1973, 1974, and 1978, respectively, all in electrical engineering.

From 1978 to 1982 he was an Assistant Professor of Electrical and Computer Engineering at the University of Iowa, Iowa City, IA. He was with the Long Range Research Division of the Exxon Production Research Company, Houston, TX, from June 1982 to September 1989. He joined the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, in September 1989 as a Professor. He currently holds the title of James B. Davis Professor. His current research interests are in statistical signal processing, wireless and wireline digital communications, cognitive radio networks, multiple sensor multiple target tracking, and stochastic systems analysis. He was elected Fellow of IEEE in 1994.



He is a past Associate Editor of the IEEE Transactions on Automatic Control, the IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, and the IEEE Transactions

on Wireless Communications. He is currently a Senior Area Editor of the IEEE Transactions on Signal Processing, and a Senior Editor of IEEE Wireless Communications Letters.

CHAPTER 4

Antoine Chevreuil (M'97) was born in 1971 in Caen, France. He received the M.Sc and the Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications in Paris (previously ENST, now Telecom Paris-Tech) respectively, in 1994 and 1997. In 1997, he obtained a grant from the Institut National de Recherche en Informatique et en Automatique (INRIA) and worked with Luc Vandendorpe in the TELE team in the University of Louvain-la-Neuve (Belgium).

From 1998 until 2009 he was Assistant Professor of Electrical Engineering in the University of Paris-East/Marne-la-Vallée. Since 2009, he is Professor in the Ecole Supérieure d'Ingénieurs en Electronique et Electrotechnique (ESIEE) located in Champs-sur-Marne (France).

He is a member of the laboratory UMR 8049 LIGM (Computing Department of the University of Paris-East). His present research interests are in statistical signal processing, and digital communications with a special emphasis on blind equalization, multi-user communication systems, and multicarrier modulations.



Philippe Loubaton (M'88) was born in 1958 in Villers Semeuse, France. He received the M.Sc and the Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1981 and 1988, respectively.

From 1982 to 1986, he was a member of the technical staff of Thomson-CSF/RGS, where he worked in digital communications. From 1986 to 1988, he worked with the Institut National des Télécommunications as an Assistant Professor of Electrical Engineering. In 1988, he joined the Ecole Nationale Supérieure des Télécommunications, Paris, France, working in the Signal Processing Department. Since 1995, he has been Professor of Electrical Engineering at Marne la Vallée University, Champs sur Marne, France. From 1996 to 2000, he was Director of the Laboratoire Système de Communication of Marne la Vallée University and is now a member of the Laboratoire Traitement et Communication de l'Information (CNRS/Ecole Nationale Supérieure des Télécommunications). His research interests are in statistical signal processing, and digital communications with a special emphasis on blind equalization, multi-user communication systems, and multi carrier modulations. His present research works concern the application of the theory of large random matrices to signal processing and digital telecommunications.



He is currently associate editor for the IEEE Transactions on Signal Processing and IEEE Communications Letters, and is a member of the IEEE Signal Processing for Communications technical committee.

CHAPTER 5

Paolo Banelli received the Laurea degree in electronics engineering and the Ph.D. degree in telecommunications from the University of Perugia, Perugia, Italy, in 1993 and 1998, respectively. In 2005, he was appointed Associate Professor at the Department of Electronic and Information Engineering (DIEI), University of Perugia, where he has been an Assistant Professor since 1998. In 2001, he joined the SpinComm group at the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, as a Visiting Researcher. His research interests mainly focus on signal processing for wireless communications, with emphasis on multicarrier transmissions, and more recently on signal processing for biomedical applications, with emphasis on electrocardiography and medical ultrasounds. In 2011 he has been elected as a member of the SP-COM Technical Committee of the IEEE Signal Processing Society. Since the end of 2012 he is serving as Associate Editor of the IEEE Transactions on Signal Processing, and the EURASIP Journal on Advances in Signal Processing. He was a General Co-Chair of the IEEE International Symposium on Signal Processing Advances for Wireless Communications 2009, he has been serving as a Reviewer for several technical journals, and as technical program committee member of leading international conferences on signal processing and telecommunications.



Luca Rugini was born in Perugia, Italy, in 1975. He received the Laurea degree in electronic engineering and the Ph.D. degree in telecommunications from the University of Perugia in 2000 and 2003, respectively. From February to July 2007, he visited Delft University of Technology, The Netherlands. He is currently an Assistant Professor with the Department of Electronic and Information Engineering at the University of Perugia. His research interests lie in the area of signal processing for multicarrier and spread-spectrum communications.



CHAPTER 6

Itsik Bergel received the B.Sc. degree in electrical engineering and the B.Sc. degree in physics from Ben Gurion University, Beer-Sheva, Israel, in 1993 and 1994, respectively, and the M.Sc. degree and Ph.D. in electrical engineering from the University of Tel Aviv, Tel-Aviv, Israel, in 2000 and 2005, respectively. From 2001 to 2003 he was a senior researcher at INTEL communications research laboratory in 2005 he did a postdoc research at the Dipartimento di Elettronica of Politecnico di Torino. He is currently a faculty member in the faculty of engineering at Bar-Ilan University, Ramat-Gan, Israel.



Amir Leshem (SM'06) received the B.Sc. (cum laude) in mathematics and physics, the M.Sc. (cum laude) in mathematics, and the Ph.D. in mathematics all from the Hebrew University, Jerusalem, Israel, in 1986, 1990, and 1998, respectively.

From 1998 to 2000 he was with Faculty of Information Technology and Systems, Delft University of Technology, The Netherlands, as a postdoctoral fellow working on algorithms for the reduction of terrestrial electromagnetic interference in radio-astronomical radio-telescope antenna arrays and signal processing for communication. From 2000 to 2003 he was director of advanced technologies with Metalink Broadband where he was responsible for research and development of new DSL and wireless MIMO modem technologies and served as a member of ITU-T SG15, ETSI TM06, NIPP-NAI, IEEE 802.3 and 802.11.



From 2000 to 2002 he was also a visiting researcher at Delft University of Technology. From 2003 to 2005 he was also the technical manager of the U-BROAD consortium developing technologies to provide 100 Mbps and beyond over copper lines.

In 2002 he was one of the founders of the faculty of engineering at Bar-Ilan university where he is a Professor and heads the Signal Processing track.

From 2008 to 2011 he was an associate editor of the IEEE Transactions on Signal Processing. Since 2010 he is member of the IEEE technical committee on Signal processing for communications. He was a leading guest editor for special issues on signal processing for astronomy and cosmology of IEEE SP magazine and IEEE Journal of Selected Topics in Signal Processing, and a guest editor of several other special issues on game theory and learning. He was also the technical co-chair of the IEEE SAM 2010.

His main research interests include multichannel wireless and wireline communication, applications of game theory to dynamic and adaptive spectrum management of communication networks, array and statistical signal processing with applications to multiple element sensor arrays and networks, wireless communications, radio-astronomical imaging and brain research, set theory, logic, and foundations of mathematics.

CHAPTER 7

Sergio Barbarossa (S'84, M'88, F'12) received the M.Sc. degree in 1984 and the Ph.D. degree in electrical engineering in 1988, both from the University of Rome “La Sapienza,” Rome, Italy. He has held positions as a Research Engineer with Selenia SpA (1984–1986) and with the Environmental Institute of Michigan (1988), as a Visiting Professor with the University of Virginia (1995 and 1997) and with the University of Minnesota (1999). He has taught short graduate courses at the Polytechnic University of Catalunya (2001 and 2009). Currently, he is a Full Professor with the University of Rome “La Sapienza.” His current research interests lie in the area of signal processing for self-organizing networks, bio-inspired signal processing, femtocell networks,



mobile cloud computing, graph theory, game theory, and distributed optimization algorithms. He is the author of a research monograph titled “Multiantenna Wireless Communication Systems.” He has been the scientific coordinator of the European projects WINSOC, on wireless sensor networks, and FREEDOM, on femtocell networks. He is currently the scientific coordinator of the European Project TROPIC, on cloud computing over small cell networks. He is also a principal investigator in the European Project SIMTISYS, on the radar monitoring of maritime traffic from satellites.

He has been nominated as an IEEE Fellow for his contributions to signal processing, sensor networks, and wireless communications. He received the 2010 EURASIP Technical Achievements Award for his contributions to synthetic aperture radar, sensor networks, and communication networks. He received the 2000 IEEE Best Paper Award from the IEEE Signal Processing Society. He is the coauthor of papers that received the Best Student Paper Award at ICASSP 2006, SPAWC 2010, EUSIPCO 2011, and CAMSAP 2011. From 1997 until 2003, he was a member of the IEEE Technical Committee for Signal Processing in Communications. He served as an Associate Editor for the IEEE Transactions on Signal processing for two terms (1998–2000 and 2004–2006). He is now a member of the IEEE Signal Processing Magazine Editorial Board. He has been the General Chairman of the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC) 2003 and he is the Technical Co-Chair of SPAWC 2013. He has been the Guest Editor for Special Issues on the IEEE Journal on Selected Areas in Communications, EURASIP *Journal of Applied Signal Processing*, EURASIP *Journal on Wireless Communications and Networking*, and the IEEE Signal Processing Magazine. In 2012, he was nominated IEEE Distinguished Lecturer from the Signal Processing Society.

Stefania Sardellitti (M’12) received the Dr. Eng. degree in Electronic Engineering from the University of Rome “La Sapienza,” Italy, in 1998 and the Ph.D. degree in Electrical and Information Engineering from the University of Cassino, Italy, in 2005. Since 2005 she is an appointed Professor of digital communications at the University of Cassino, Italy. She is currently a research assistant at the Department of Information, Electronics and Telecommunications, University of Rome, Sapienza, Italy. She has participated in the European project WINSOC, on wireless sensor networks and in the European project FREEDOM on femtocell networks. She is currently involved in the European TROPIC, on distributed computing, storage, and radio resource allocation over cooperative femtocells. Her research interests are in the area of statistical signal processing, in particular on multiple antenna and multiple access systems. Currently, her primary research activity is on the field of cognitive radios, femtocell networks, and wireless sensor networks, with emphasis on distributed decision.



Paolo Di Lorenzo (S’10-M’13) received the M.Sc. degree in 2008 and the Ph.D. in electrical engineering in 2012, both from University of Rome “La Sapienza,” Italy. He is currently a post-doctoral researcher in the Department of Information, Electronics and Telecommunications, University of Rome, “La Sapienza.” During 2010 he held a visiting research appointment in the

Department of Electrical Engineering, University of California at Los Angeles (UCLA). He has participated in the European research project FREEDOM on femtocell networks. He is currently involved in the European projects SIMTISYS, on moving target detection through satellite constellations, and TROPIC, on distributed computing, storage, and radio resource allocation over cooperative femtocells. His primary research interests are in statistical signal processing, distributed optimization algorithms for communication and sensor networks, graph theory, game theory, and adaptive filtering. He received three best student paper awards, respectively, at IEEE SPAWC'10, EURASIP EUSIPCO'11, and IEEE CAMSAP'11, for works in the area of signal processing for communications and synthetic aperture radar systems. He is recipient of the 2012 GTI (Italian national group on telecommunications and information theory) award for the Best Ph.D. Thesis in information technologies and communications.



CHAPTER 8

Mingyi Hong received his B.E. degree in Communications Engineering from Zhejiang University, China, in 2005, and his M.S. degree in Electrical Engineering from Stony Brook University in 2007, and Ph.D. degree in Systems Engineering from University of Virginia in 2011. He is currently a post-doctoral fellow with the Department of Electrical and Computer Engineering, University of Minnesota. His research interests are primarily in the fields of statistical signal processing, wireless communications, and optimization theory.



Zhi-Quan Luo received his B.Sc. degree in Applied Mathematics in 1984 from Peking University, Beijing, China. Subsequently, he was selected by a joint committee of the American Mathematical Society and the Society of Industrial and Applied Mathematics to pursue Ph.D study in the United States. After a one-year intensive training in mathematics and English at the Nankai Institute of Mathematics, Tianjin, China, he studied in the Operations Research Center and the Department of Electrical Engineering and Computer Science at MIT, where he received a Ph.D. degree in Operations Research in 1989. From 1989 to 2003, he held a faculty position with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, where he eventually became the department head and held a Canada Research Chair in Information Processing. Since April of 2003, he has been with the Department of Electrical and Computer Engineering at the University of Minnesota (Twin Cities) as a Full Professor and holds an endowed ADC Chair in digital technology. His research interests include optimization algorithms, signal processing, and digital communication.



He is a fellow of IEEE and SIAM, and serves as the chair of the IEEE Signal Processing Society Technical Committee on the Signal Processing for Communications (SPCOM). He is a recipient of the 2004 and 2009 IEEE Signal Processing Society Best Paper Awards, the 2010 Farkas Prize

from the INFRMS Optimization Society, the 2011 EURASIP Best Paper Award, and the 2011 ICC Best Paper Award. He has held editorial positions for several international journals including Journal of Optimization Theory and Applications, SIAM Journal on Optimization, Mathematics of Computation, and IEEE Transactions on Signal Processing. He currently serves as the Editor-in-Chief for the journal IEEE Transactions on Signal Processing.

CHAPTER 9

Seung-Jun Kim (SM'12) received his B.S. and M.S. degrees from Seoul National University in Seoul, Korea in 1996 and 1998, respectively, and his Ph.D. from the University of California at Santa Barbara in 2005, all in electrical engineering. From 2005 to 2008, he worked for NEC Laboratories America in Princeton, New Jersey, as a research staff member. Since 2008, he has been with the Department of Electrical and Computer Engineering at the University of Minnesota, where he is currently a Research Assistant Professor. His research interests lie in applying signal processing and optimization techniques to various domains including wireless communication and networking as well as smart power grids.



Emiliano Dall'Anese received the Laurea Triennale (B.Sc. Degree) and the Laurea Specialistica (M.Sc. Degree) in Telecommunications Engineering from the University of Padova, Italy, in 2005 and 2007, respectively, and the Ph.D in Information Engineering from the Department of Information Engineering (DEI), University of Padova, Italy, in 2011. From January 2009 to September 2010 he was a visiting scholar at the Department of Electrical and Computer Engineering, University of Minnesota, USA. Since January 2011, he has been a postdoctoral associate at the Department of Electrical and Computer Engineering and Digital Technology Center, University of Minnesota, USA.



His research interests lie in the areas of statistical signal processing, networking, and smart power systems. Current research includes optimal energy management in smart power distribution networks; robust, distributed, and sparsity-aware statistical inference; and, and optimization of wireless cognitive radio networks.

Juan Andrés Bazerque (S'06) received his B.Sc. degree in Electrical Engineering from Universidad de la República (UdelaR), Montevideo, Uruguay in 2003. Since August 2006 he has been a research assistant at the University of Minnesota (U of M), Minneapolis, where he received his M.Sc. in Electrical Engineering in August 2009, and his Ph.D. in May 2013. From 2000 to 2006 he was a teaching assistant with the Department of Mathematics and Statistics, and with the Department of Electrical Engineering (UdelaR). From 2003 to 2006 he worked as a telecommunications engineer at the Uruguayan company Unitel S.A. developing applications for Voice over IP. His broad research interests lie in the general areas of networking and signal processing. His current research focuses on distributed signal processing, cooperative wireless communications, compressive sampling,



sparsity-aware statistical modeling, and gene expression networks. He received the UoM's Distinguished M.Sc. Thesis Award in 2009, and the best student paper award at the second International Conference on Cognitive Radio Oriented Wireless Networks and Communication (CROWNCOM) 2007.

Ketan Rajawat received his B.Tech and M.Tech degrees in Electrical Engineering from Indian Institute of Technology Kanpur, in 2007; and his Ph.D. degree in Electrical and Computer Engineering from University of Minnesota, in 2012. Currently, he is an Assistant Professor in the Department of Electrical Engineering, IIT Kanpur. His research interests lie in the areas of Signal processing and Communication Networks. His current research focuses on cross-layer network optimization dynamic network monitoring.



Georgios B. Giannakis (Fellow'97) received his Diploma in Electrical Engineering from the Ntl. Tech. University of Athens, Greece, 1981. From 1982 to 1986 he was with the University of Southern California (USC), where he received his M.Sc. in Electrical Engineering, 1983, M.Sc. in Mathematics, 1986, and Ph.D. in Electrical Engineering, 1986. Since 1999 he has been a Professor with the University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE Department and serves as Director of the Digital Technology Center.



His general interests span the areas of communications, networking, and statistical signal processing - subjects on which he has published more than 350 journal papers, 580 conference papers, 20 book chapters, two edited books, and two research monographs (h-index 103). Current research focuses on sparsity and big data analytics, wireless cognitive radios, mobile ad hoc networks, renewable energy, power grid, gene-regulatory, and social networks. He is the (co-) inventor of 21 patents issued, and the (co-) recipient of 8 best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.

CHAPTER 11

Maria S. Greco graduated in Electronic Engineering in 1993 and received the Ph.D. degree in Telecommunication Engineering in 1998, from University of Pisa, Italy. From December 1997 to May 1998 she joined the Georgia Tech Research Institute, Atlanta, USA as a visiting research scholar where she carried on research activity in the field of radar detection in non-Gaussian background.



In 1993 she joined the Department of "Ingegneria dell'Informazione" of the University of Pisa, where she is Associate Professor since December 2011. She is IEEE fellow since January 2011 and she was co-recipient of the 2001 IEEE

Aerospace and Electronic Systems Society's Barry Carlton Award for Best Paper and recipient of the 2008 Fred Nathanson Young Engineer of the Year award for contributions to signal processing, estimation, and detection theory. She has been general-chair, technical chair, and technical committee member of many conferences. She was guest co-editor of the special issue of *the Journal of the IEEE Signal Processing Society on Special Topics in Signal Processing* on "Adaptive Waveform Design for Agile Sensing and Communication," published in June 2007 and was lead guest editor of the special issue of *International Journal of Navigation and Observation* on "Modeling and Processing of Radar Signals for Earth Observation" published in August 2008. She is Associate Editor of IET Proceedings—Sonar, Radar and Navigation, of the IEEE Transactions on Signal Processing, of the IEEE Aerospace and Electronic Systems Magazine, member of the Editorial Board of the Journal of Advances in Signal Processing (JASP), member of the IEEE Signal Processing Theory and Methods (SPTM) and Signal Array Processing (SAM) Technical Committees.

Her general interests are in the areas of statistical signal processing, estimation, and detection theory. In particular, her research interests include clutter models, spectral analysis, coherent and incoherent detection in non-Gaussian clutter, CFAR techniques, radar waveform diversity, and bistatic/multistatic radars. She co-authored two book chapters, more than 100 journal and conference papers.

Simon Watts graduated from the University of Oxford in 1971, obtained an M.Sc. from the University of Birmingham in 1972 and a Ph.D. from the CNAA in 1987. He is currently deputy Scientific Director and Technical Fellow in Thales UK and is also a Visiting Professor in the department of Electronic and Electrical Engineering at University College London. He joined Thales (then EMI Electronics) in 1967 and since then has worked on a wide range of radar and EW projects, with a particular research interest in maritime radar and sea clutter. He is author and co-author of over 50 journal and conference papers, a book on sea clutter and several patents. He was Chairman of the international radar conference RADAR-97 in Edinburgh UK. He received the IEE JJ Thomson Premium Award in 1987 and the IEE Mountbatten Premium Award in 1991. He serves on the IEEE AES Radar Systems Panel, is an Associate Editor for Radar for the IEEE Transactions AES and a member of the Editorial Board of IET Radar, Sonar and Navigation. He was appointed MBE in 1996 for services to the UK defence industry and is a Fellow of the Royal Academy of Engineering, Fellow of the IET, Fellow of the IMA, and Fellow of the IEEE.



CHAPTER 12

William L. Melvin is Director of the Sensors and Electromagnetic Applications Laboratory at the Georgia Tech Research Institute and an Adjunct Professor in Georgia Tech's Electrical and Computer Engineering Department. His research interests include all aspects of RF and acoustic sensor development. He has authored over 180 publications in his areas of research interest and holds three US patents on sensor technology.



Among his distinctions, he is the recent recipient of the 2006 IEEE AESS Young Engineer of the Year Award, the 2003 US Air Force Research Laboratory Reservist of the Year Award, and the 2002 US Air Force Materiel Command Engineering and Technical Management Reservist of the Year Award. He was chosen as an IEEE Fellow for his contributions to adaptive radar technology. Also, he is a member of the Board on Army Science and Technology through the National Academy of Science.

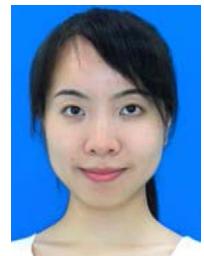
He received the Ph.D. in Electrical Engineering from Lehigh University, as well as the MSEE and BSEE degrees (with high honors) from this same institution, respectively.

CHAPTER 13

Qian He received her B.E. and Ph.D. degrees both with honors in Electronic Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2004 and 2010, respectively.

From 2007 to 2009, she was a visiting scholar in the Electrical and Computer Engineering Department, Lehigh University, Bethlehem, PA. From 2010 to 2011, she was a Postdoctoral Research Associate in the same department at Lehigh University. In 2010, she joined the faculty of the Electronic Engineering Department at UESTC, where she is currently an Associate Professor. Her current research interests include statistical signal processing, array signal processing, adaptive signal processing, and their applications in radar, communication, and smart grid systems.

She is a member of Sigma Xi. She received the 2010 Excellent Doctoral Dissertation Award of Sichuan Province and the 2012 Outstanding Faculty Member Award from the School of Electronic Engineering at UESTC. She is on the editorial board for the Journal of Communications and Information Sciences and for the Advances in Energy and Power Engineering. She is an affiliate member of the Sensor Array and Multichannel Technical Committee of the IEEE Signal Processing Society.



Yang Yang received the B.E. degree in Information Engineering from Xi'an Jiaotong University, Xi'an, China, in 2001, the M.E. degree in Electrical Engineering from the National University of Singapore in 2004, and the Ph.D. degree in Electrical Engineering from Lehigh University, Bethlehem, PA, in 2009.

From 2002 to 2004, he was with the Institute for Infocomm Research, Singapore, as a research scholar. During the summer of 2007, he was an intern at Bell Laboratories, Alcatel-Lucent, Crawford Hill, Holmdel, NJ. From 2009 to 2011, he was a Postdoctoral Research Associate at Lehigh University. He is currently with Broadcom Corporation, Matawan, NJ.



His research interests include signal processing, communications, and their applications to distributed sensor processing and wireless sensor networks.

Rick S. Blum received a B.S. in Electrical Engineering from the Pennsylvania State University in 1984 and his M.S. and Ph.D in Electrical Engineering from the University of Pennsylvania in 1987 and 1991.

From 1984 to 1991 he was a member of technical staff at General Electric Aerospace in Valley Forge, Pennsylvania and he graduated from GE's Advanced Course in Engineering. Since 1991, he has been with the Electrical and Computer Engineering Department at Lehigh University in Bethlehem, Pennsylvania where he is currently a Professor and holds the Robert W. Wieseman Chaired Research Professorship in Electrical Engineering. His research interests include signal processing for smart grid, communications, sensor networking, radar and sensor processing. He is on the editorial board for the Journal of Advances in Information Fusion of the International Society of Information Fusion. He was an associate editor for IEEE Transactions on Signal Processing and for IEEE Communications Letters. He has edited special issues for IEEE Transactions on Signal Processing, IEEE Journal of Selected Topics in Signal Processing and IEEE Journal on Selected Areas in Communications. He is a member of the SAM Technical Committee (TC) of the IEEE Signal Processing Society. He was a member of the Signal Processing for Communications TC of the IEEE Signal Processing Society and is a member of the Communications Theory TC of the IEEE Communication Society. He was on the awards Committee of the IEEE Communication Society.



He is a Fellow of the IEEE, an IEEE Third Millennium Medal winner, a member of Eta Kappa Nu and Sigma Xi, and holds several patents. He was awarded an ONR Young Investigator Award in 1997 and an NSF Research Initiation Award in 1992. His IEEE Fellow Citation "for scientific contributions to detection, data fusion and signal processing with multiple sensors" acknowledges contributions to the field of sensor networking.

CHAPTER 14

Joseph R. Guerci has over 29 years of experience in advanced technology research and development in government, industrial, and academic settings. His government service included a 7-year term with the Defense Advanced Research Projects Agency (DARPA) in which he held the positions of Program Manager, Deputy Office Director, and finally Director of the Special Projects Office (SPO). In these capacities, he was involved in the inception, research, development, execution, and ultimately transition of next generation multi-disciplinary defense technologies.



He is also a recognized leader in the research and development of next generation sensor and electronic warfare systems. In addition to authoring over 100 peer-reviewed articles, he has several book chapters and is the author of *Space-Time Adaptive Processing for Radar* (Artech House, 2003), and the recently published *Cognitive Radar: The Knowledge-Aided Fully Adaptive Approach*, (Artech House, 2010). He also recently received the Warren D. White Award from the Institute of Electrical and Electronics Engineers (IEEE) for "Excellence in Radar Adaptive Processing and Waveform Diversity," and the IEEE Waveform Diversity Person of the Year for 2010 for "For Scientific, Technical, and Executive Leadership Contributions in Making Waveform Diversity a Fielded Technology." He has also recently been appointed the General

Chair for the 2015 IEEE International Radar Conference, and is an elected member of the IEEE Aerospace and Electronic Systems Society (AESS) Board of Governors.

A graduate of Polytechnic University (now NYU School of Engineering) with a Ph.D.E.E (System Engineering), he has held adjunct professorships in engineering and applied mathematics at The City University of New York, Polytechnic University, The Cooper Union for Advancement of Art and Science, and Virginia Tech. Additionally, he has held senior engineer and scientist positions in industry and was Chief Technology Officer (CTO) for SAIC's \$2B+/ year Research, Development, Test & Evaluation (RDT&E) Group. A member of the IEEE Radar Systems panel, he is also a Fellow of the IEEE for "*Contributions to Advanced Radar Theory and its Embodiment in Real-World Systems*," holds eight US Patents, and is a member of several industrial, academic, and government advisory boards including the Air Force Scientific Advisory Board (SAB) for a 4-year term beginning in the Fall of 2013.

CHAPTER 15

X. Chen He was born in P.R. China, 1981. He received his B.Eng. and M.Eng. degrees, both in electronic engineering, from the University of Electronic Science and Technology of China (UESTC), ChengDu, SiChuan Prov., P.R. China in 2004 and 2007, respectively. He received his Ph.D. degree in electrical and computer engineering from McMaster University, Hamilton, Ontario, Canada, in 2012. Currently he is working as a Postdoctoral Researcher in the Electrical and Computer Engineering Department at McMaster University, Canada. His research interests include point processes, multitarget-multisensor state estimation and tracking, probability density estimation, and general stochastic signal processing using the stochastic differential equation (SDE).



R. Tharmarasa He received the B.Sc.Eng. degree in electronic and telecommunication engineering from University of Moratuwa, Sri Lanka in 2001, and the M.A.Sc. and Ph.D. degrees in electrical engineering from McMaster University, Canada in 2003 and 2007, respectively. From 2001 to 2002 he was an instructor in electronic and telecommunication engineering at the University of Moratuwa, Sri Lanka. During 2002–2007 he was a graduate student/research assistant in ECE department at the McMaster University, Canada. Currently he is working as a Research Associate in the Electrical and Computer Engineering Department at McMaster University, Canada. His research interests include target tracking, information fusion and sensor resource management.



T. Kirubarajan Thiagalingam Kirubarajan received the B.A. and M.A. degrees in electrical and information engineering from Cambridge University, England, in 1991 and 1993, and the M.S. and Ph.D. degrees in electrical engineering from the University of Connecticut, Storrs, in 1995 and 1998, respectively. Currently, he is a professor in the Electrical and Computer Engineering

Department at McMaster University, Hamilton, Ontario. He is also serving as an Adjunct Assistant Professor and the Associate Director of the Estimation and Signal Processing Research Laboratory at the University of Connecticut. His research interests are in estimation, target tracking, multisource information fusion, sensor resource management, signal detection and fault diagnosis. His research activities at McMaster University and at the University of Connecticut are supported by US Missile Defense Agency, US Office of Naval Research, NASA, Qualtech Systems, Inc., Raytheon Canada Ltd. and Defense Research Development Canada, Ottawa. In September 2001, Dr. Kirubarajan served in a DARPA expert panel on unattended surveillance, homeland defense and counterterrorism. He has also served as a consultant in these areas to a number of companies, including Motorola Corporation, Northrop-Grumman Corporation, Pacific-Sierra Research Corporation, Lockheed Martin Corporation, Qualtech Systems, Inc., Orincon Corporation and BAE systems. He has worked on the development of a number of engineering software programs, including BEARDAT for target localization from bearing and frequency measurements in clutter, FUSEDAT for fusion of multisensor data for tracking. He has also worked with Qualtech Systems, Inc., to develop an advanced fault diagnosis engine. He has published about 100 articles in areas of his research interests, in addition to one book on estimation, tracking and navigation and two edited volumes. He is a recipient of Ontario Premier's Research Excellence Award (2002).



CHAPTER 16

Hugh Griffiths holds the THALES/Royal Academy Chair of RF Sensors in the Department of Electronic and Electrical Engineering at University College London, England. From 2006 to 2008 he was Principal of the Defence Academy College of Management and Technology. He received the MA degree in Physics from Oxford University in 1975, then spent 3 years working in industry, before joining University College London, where he received the Ph.D. degree in 1986 and the DSc(Eng) degree in 2000. He served as Head of Department from 2001 to 2006.



His research interests include radar and sonar systems and signal processing (particularly synthetic aperture radar and bistatic and multistatic radar), and antenna measurement techniques. He has published over 400 papers and technical articles in the fields of radar, antennas, and sonar. In 1996 he received the IEEE AESS Nathanson Award (Radar Systems Panel Award). He has also received the URSI Young Scientist Award, the IERE Brabazon Premium and the IEE Mountbatten and Maxwell Premium Awards, and he received the NATO SET Panel Award in 2012 as one of the lecturers on the NATO SET-119 Lecture Series on "Waveform Diversity." He is a Fellow of the IET, Fellow of the IEEE, and in 1997 he was elected to Fellowship of the Royal Academy of Engineering. He serves as President of the IEEE Aerospace and Electronic Systems Society for 2012/2013.

CHAPTER 17

Moeness G. Amin received his Ph.D. degree in 1984 from University of Colorado, in Electrical Engineering. He has been on the Faculty of the Department of Electrical and Computer Engineering at Villanova University since 1985. In 2002, he became the Director of the Center for Advanced Communications, College of Engineering.

He is the Recipient of the 2009 Individual Technical Achievement Award from the European Association of Signal Processing, and the Recipient of the 2010 NATO Scientific Achievement Award. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), Fellow of the International Society of Optical Engineering; and a Fellow of the Institute of Engineering and Technology (IET). He is a Recipient of the IEEE Third Millennium Medal; Recipient of the Chief of Naval Research Challenge Award, 2010; Distinguished Lecturer of the IEEE Signal Processing Society, 2003–2004; Active Member of the Franklin Institute Committee on Science and the Arts; Recipient of Villanova University Outstanding Faculty Research Award, 1997; and the Recipient of the IEEE Philadelphia Section Award, 1997. He is a member of IEEE, SPIE, EURASIP, ION, Eta Kappa Nu, Sigma Xi, and Phi Kappa Phi.



He has over 450 journal and conference publications in the areas of Wireless Communications, Time-Frequency Analysis, Smart Antennas, Waveform Design and Diversity, Interference Cancellation in Broadband Communication Platforms, Anti-Jam GPS, Target Localization and Tracking, Direction Finding, Channel Diversity and Equalization, Ultrasound Imaging and Radar Signal Processing. He is a recipient of seven best paper awards. He currently serves on the Overview Board of the IEEE Transactions on Signal Processing. He also serves on the Editorial Board of the EURASIP Signal Processing Journal and the Editorial Board of the Signal Processing Magazine. He was a Plenary Speaker at ICASSP 2010. He was the Special Session Co-Chair of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing. He was the Technical Program Chair of the second IEEE International Symposium on Signal Processing and Information Technology, 2002. He was the General and Organization Chair of the IEEE Workshop on Statistical Signal and Array Processing, 2000. He was the General and Organization Chair of the IEEE International Symposium on Time-Frequency and Time-Scale Analysis, 1994. He was an Associate Editor of the IEEE Transactions on Signal Processing during 1996–1998. He was a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications during 1998–2002. He was a Member of the IEEE Signal Processing Society Technical Committee on Statistical Signal and Array Processing during 1995–1997.

He was the Guest Editor of the Journal of Franklin Institute September 2008 Special Issue on Advances in Indoor Radar Imaging. He was a Guest Editor of the IEEE Transactions on Geoscience and Remote Sensing May 2009 Special Issue on Remote Sensing of Building Interior, and a Guest Editor of the IET Signal Processing December 2009 Special Issue on Time-Frequency Approach to Radar Detection, Imaging, and Classification.

Fauzia Ahmad received her MS degree in Electrical Engineering in 1996, and Ph.D. degree in Electrical Engineering in 1997, both from the University of Pennsylvania, Philadelphia, PA.

From 1998 to 2000, she was an Assistant Professor in the College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan. During

2000–2001, she served as an Assistant Professor at Fizaia College of Information Technology, Pakistan. Since 2002, she has been with the Center for Advanced Communications, Villanova University, Villanova, PA, where she is now a Research Associate Professor and the Director of the Radar Imaging Laboratory.

She has over 110 journal and conference publications in the areas of radar imaging, radar signal processing, waveform design and diversity, compressive sensing, array signal processing, sensor networks, ultrasound imaging, and over-the-horizon radar.



CHAPTER 18

Stefan V. Baumgartner received the Dipl.-Ing. degree in Electrical Engineering and Communication Technology from the Graz University of Technology, Graz, Austria in 2004.

Since 2004, he has been with the Microwaves and Radar Institute (HR), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He is currently with the Radar Concepts Department, where his field of activity is the development of ground moving target indication and parameter estimation algorithms for future road-traffic-monitoring applications using multichannel air- and spaceborne synthetic aperture radars (SAR). His research interests include SAR along-track interferometry, time-frequency analysis, and other advanced signal and imaging processing techniques.



Gerhard Krieger received the Dipl.-Ing. (M.S.) and Dr.-Ing. (Ph.D.) degrees (with honors) in electrical and communication engineering from the Technical University of Munich, Germany, in 1992 and 1999, respectively.

From 1992 to 1999, he was with the Ludwig Maximilians University, Munich, where he conducted multidisciplinary research on neuronal modeling and nonlinear information processing in biological and technical vision systems. Since 1999, he has been with the Microwaves and Radar Institute (HR) of the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where he developed signal and image processing algorithms for a novel forward looking radar system employing digital beamforming on receive. From 2001 to 2007 he led the New Synthetic Aperture Radar (SAR) Missions Group which pioneered the development of advanced bistatic and multistatic radar systems as exemplified by the TanDEM-X mission, as well as innovative multichannel SAR techniques and algorithms for high-resolution wide-swath SAR imaging. Since 2008, he has been the Head of the new Radar Concepts Department of the Microwaves and Radar Institute, DLR, Oberpfaffenhofen, Germany. Gerhard Krieger has authored more than 50 peer-reviewed journal papers, 7 invited book chapters, about 300 conference papers, and 5 patent families. His current research interests focus on the development of multichannel radar techniques and algorithms for innovative Multiple-Input Multiple-Output (MIMO) SAR systems, the demonstration of novel interferometric and tomographic Earth observation applications, and the conceptual design of advanced bi- and multistatic radar missions.



He is IEEE Fellow and received several national and international awards, including the W.R.G. Baker Prize Paper Award from the IEEE Board of Directors and the Transactions Prize Paper Award of the IEEE Geoscience and Remote Sensing Society. In 2012, he and his colleagues were nominated for the German President's Award for Technology and Innovation. Since 2012, he has been Associate Editor of the IEEE Transactions on Geoscience and Remote Sensing.

CHAPTER 19

Marco Martorella received his Laurea degree (Bachelor+Masters) in Telecommunication Engineering in 1999 (cum laude) and his PhD in Remote Sensing in 2003, both at the University of Pisa. He is now an Associate Professor at the Department of Information Engineering of the University of Pisa where he lectures “Fundamentals of Radar” and “Digital Communications” and an external Professor at the University of Cape Town where he lectures “High Resolution and Imaging Radar” within the “Masters in Radar and Electronic Defence.” He is a regular visiting Professor at the University of Adelaide and at the University of Queensland in Australia. He is author of more than a 100 international journal and conference papers and three book chapters. He has presented several tutorials at international radar conferences and organized a special issue on Inverse Synthetic Aperture Radar for the Journal of Applied Signal Processing. He is a member of the IET Radar Sonar and Navigation Editorial Board, a senior member of the IEEE, and a member of AFCEA. He is also chair of the NATO SET-196 on “Multichannel/Multistatic radar imaging of non-cooperative targets.” He has been recipient of the 2008 Italy-Australia Award for young researchers, the 2010 Best Reviewer for the IEEE GRSL and the IEEE 2013 Fred Nathanson Memorial Radar Award. His research interests are mainly in the field of radar imaging, including passive, multi-channel, multistatic and polarimetric radar imaging.



CHAPTER 20

Gianfranco Fornaro received the M.S. degree (summa cum laude) in electronic engineering and the Ph.D. degree from the University of Naples “Federico II” in 1992 and 1997, respectively. Since 1993 he has been with Institute for Electromagnetic Sensing of the Environment (IREA) of the Italian National Research Council, Naples, where he is currently a senior researcher working in the area of Synthetic Aperture Radar (SAR) focusing, SAR interferometry, and SAR tomography. He has been Adjunct Professor in the area of communications in several Universities, currently at the University of Napoli “Parthenope.” He has been visiting scientist at the Politecnico of Milano and at the German Aerospace Establishment (DLR), also for within the Italy-Germany cooperation during the SIR-C/X-SAR mission. He has been United Nations consultant at the Istituto Tecnologico de Aeronautica (ITA) in Sao José dos Campos (Brazil) and at RESTEC (Tokyo). Since 2010 he has been Lecturer at the International radar/SAR Summer School of the Fraunhofer Institute; he has been also Convener, Tutorial Lecturer, and Chairman of sessions dedicated to SAR processing and



SAR interferometry in several international conferences. He has authored more than a hundred of papers (peer-review journals and proceedings of international conferences); in 2005 served as Editor of the “Advances in Interferometric SAR processing” special issue of the EURASIP Journal on Applied Signal Processing (JASP). He received the Mountbatten Premium by the Institution of Electrical Engineers (IEE) in 1997, the Institute of Electrical and Electronics Engineers (IEEE) Geoscience and 2011 Remote Sensing Letters Best Paper Award, and the Mention for Best 2011 Reviewer for the IEEE Transactions on Geoscience and Remote Sensing journal.

Vito Pascazio graduated summa cum laude in 1986 at the Università di Bari, Italy, in Electronic Engineering. In the same year he joined the Department of Electronic Engineering of the Università di Napoli Federico II, Italy, where, in 1990 he received the Ph.D. degree in Electronic Engineering and Computer Science. During the same year he was awarded of the Philip Morris Prize for Scientific and Technological Research. In 1990 he was first at the Research Institute on Electromagnetics and Electronic Devices (IRECE) of the Italian National Council of Research (CNR), Napoli, Italy, and then he joined the Università di Napoli Parthenope, Italy, where he is presently Full Professor of Telecommunications. In 1994–1995 he was visiting scientist at the Laboratoire des Signaux et Systèmes of the Ecole Supérieure d'Electricité (Supelec), Gif sur Yvette, France, and in 1998–1999 at the Université de Nice Sophia-Antipolis, France. He is presently also Director of National Laboratory of Multi-Media Communications of the Italian Inter- University Consortium of Telecommunications (CNIT), Napoli, Italy. His main scientific interests are in the fields of Remote Sensing, and Linear and Non-linear Estimation, with particular emphasis to image computing and processing and reconstruction of microwave and radar images. He is Associate Editor of International Journal of Navigation and Observation. He published more than 150 technical papers, and he is Senior Member of IEEE.



CHAPTER 21

Laurent Ferro-Famil received the Laurea degree in electronics systems and computer engineering, the M.S. degree in electronics, and the Ph.D. degree from the University of Nantes, Nantes, France, in 1996, 1996, and 2000, respectively.

In 2001 he became an Associate Professor and since 2011 he has been a Full Professor with the University of Rennes 1, Rennes, France, where he is currently the head of the Remote Sensing department, within the Institute of Electronics and Telecommunications of Rennes. His current activities in education are concerned with analog electronics, digital communications, microwave theory, signal processing, and polarimetric SAR remote sensing. He is particularly interested in polarimetric SAR signal statistical processing, radar polarimetry theory, and natural media remote sensing using multibaseline PolInSAR data, with application to classification, electromagnetic scattering modeling and physical parameter retrieval, time-frequency analysis, and 3-D reconstruction of environments using tomography.



Eric Pottier (M'95, SM'06, F'11) received the M.Sc. and Ph.D. in signal processing and telecommunication from the University of Rennes 1, respectively, in 1987 and 1990, and the Habilitation from the University of Nantes in 1998.

Since 1999, he has been a Full Professor at the University of Rennes 1, France, where he is currently the Director of the Institute of Electronics and Telecommunications of Rennes (I.E.T.R—CNRS UMR 6164). His current activities of research and education are centered in the topics of analog electronics, microwave theory, and radar imaging with emphasis in radar polarimetry. His research covers a wide spectrum of areas from radar image processing (SAR, ISAR), polarimetric scattering modeling, supervised/unsupervised polarimetric segmentation, and classification to fundamentals and basic theory of polarimetry.



He has published 10 chapters in books, more than 60 papers in refereed journals and presented more than 350 papers during International Conferences, Symposiums, and Workshops. He has presented advance courses and seminars on Radar Polarimetry to a wide range of organizations and events.

He was presented the Best Paper Award at the Third European Conference on Synthetic Aperture Radar (EUSAR2000) and received the 2007 IEEE GRS-S Letters Prize Paper Award.

He has published a book co-authored with Dr. Jong-Sen Lee: *Polarimetric Radar Imaging: From basics to applications*, CRC Press, Taylor & Francis editor, 397 pages, January 2009, ISBN: 978-1-4200-5497-2.

He is a recipient of the 2007 IEEE GRS-S Education Award “*In recognition of his significant educational contributions to Geoscience and Remote Sensing*.”

He has been elevated to IEEE Fellow (January 2011) with the accompanying citation: “*for contributions to polarimetric Synthetic Aperture Radar*.”

He is a recipient of the 2012 Einstein Professorship from the Chinese Academy of Science.

CHAPTER 22

Alfonso Farina (M'95–SM'98–F'00) received the doctor degree in electronic engineering from the University of Rome (I), Italy, in 1973.

In 1974, he joined Selenia, now SELEX Electronic Systems, where he has been a Manager since May 1988. He was Scientific Director in the Chief Technical Office. He was the Director of the Analysis of Integrated Systems Unit. He was also the Director of Engineering in the Large Business Systems Division. In 2012, he was the Chief Technology Officer of the Company (SELEX Sistemi Integrati) reporting directly to the President. Today he is Senior Advisor to CTO of SELEX ES. In his professional life, he has provided technical contributions to detection, signal, data, image processing and fusion for the main radar systems conceived, designed, and developed in the Company. He has provided leadership in many projects—also conducted in the international arena—in surveillance for ground and naval applications, in airborne early warning and in imaging radar. From 1979 to 1985, he has also been a Professor of radar techniques with the University of Naples; in 1985 he was appointed Associate Professor. He is the author of more than 500 peer-reviewed technical publications and the author of books and monographs: *Radar Data Processing* (Vol. 1 and 2) (translated in Russian and Chinese), (UK: Researches Studies Press, and New York: Wiley, 1985–1986); *Optimized Radar Processors*, (on behalf of IEE, London, UK: Peter Peregrinus, 1987); and *Antenna Based*



I Authors Biography

Signal Processing Techniques for Radar Systems, 1992. He wrote the chapter “ECCM Techniques” in the Radar Handbook (2nd ed., 1990, and 3rd ed., 2008), edited by Dr. M. I. Skolnik (NRL, USA).

He has been session chairman at many international radar conferences. In addition to lecturing at universities and research centers in Italy and abroad, he also frequently gives tutorials at the International Radar Conferences on signal, data, and image processing for radar; in particular on multisensor fusion, adaptive signal processing, space-time adaptive processing (STAP), and detection. In 1987, he received the Radar Systems Panel Award of IEEE Aerospace and Electronic Systems Society (AEES) for development of radar data processing techniques. He is the Italian representative of the International Radar Systems Panel of the IEEE AEES. He is VP for industry in the BoG of IEEE AEES. He has been the Italian industrial representative (Panel Member at Large) at the Sensor and Electronic Technology (SET) of Research Technology Organisation (RTO) of NATO. He has been on the BoD of the International Society for Information Fusion (ISIF). He has been the Executive Chair of the International Conference on Information Fusion (Fusion) 2006, Florence, Italy, July 10–13, 2006. He has been nominated Fellow of IEEE with the following citation: “For development and application of adaptive signal processing methods for radar systems.” Recently, he has been nominated international fellow of the Royal Academy of Engineering, UK; this fellowship was presented to him by HRH Prince Philip, the Duke of Edinburgh. He is a referee of numerous publications submitted to several journals of IEEE, IEE, Elsevier, etc. He has also cooperated with the Editorial Board of the IEE Electronics & Communication Engineering Journal. More recently, he has served as a member of the Editorial Board of Signal Processing (Elsevier) and has been Co-Guest Editor of its Special Issue on New Trends and Findings in Antenna Array Processing for Radar, September 2004. He is the corecipient of the following Best Paper Awards: entitled to B. Carlton of the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS for 2001 and 2003 and also of the International Conference on Fusion 2005. He has been the leader of the team that received the 2002 AMS CEO award for Innovation Technology. He has been the corecipient of the AMS Radar Division award for Innovation Technology in 2003. Moreover, he has been the corecipient of the 2004 AMS CEO Award for Innovation Technology.

He has been the leader of the team that won the 2004 First Prize Award for Innovation Technology of Finmeccanica, Italy. This award context has seen the submission of more than 320 projects. This award has been set for the first time in 2004. In September 7, 2006, he received the Annual European Group Technical Achievement Award 2006 by the European Association for Signal, Speech and Image Processing (EURASIP), with the citation: “For development and application of adaptive signal processing technique in practical radar systems.” In 2006 and 2009, he was a corecipient of the annual Innovation Technology award of SELEX Sistemi Integrati. He has been appointed member in the Editorial Boards of IET Radar, Sonar and Navigation and of Signal, Image, and Video Processing Journal (SIVP). He has been the General Chairman of the IEEE Radar Conference 2008, Rome, May 26–30, 2008. He is a Fellow of the Institution of Engineering and Technology (IET), UK. He has been recently nominated Fellow of EURASIP, the citation reads: “For contributions to radar system design, signal, data and image processing, data fusion and particularly for the development of innovative algorithms for deployment into practical radar systems.” He is also the recipient of the 2010 IEEE Dennis J. Picard Gold

Medal for Radar Technologies and Applications with the following citation: "For continuous, innovative, theoretical and practical contributions to radar systems and adaptive signal processing techniques."

Luciana Ortenzi received the Doctor Degree in Telecommunications Engineering in October 2001 from the University of Rome, "La Sapienza" and Ph.D. in Remote Sensing applications in March 2007. She joined AMS, now SELEX Electronic Systems, in February 2002, and since then she works in the area of Integrated System Analysis Group as a system analyst. Her areas of investigation are oriented toward adaptive signal processing, detection, estimation, and tracking filtering for radar systems development, integrated system design for homeland security. She has been the corecipient of the 2003 AMS MD Award for Innovation Technology for and the 2006 Innovation Technology Award of Selex-Sistemi Integrati.



Branko Ristic received Ph.D. degree in 1995 from Queensland University of Technology. He held various research/engineering positions in former Yugoslavia and Australia before joining in 1996 the Australian Defence Science and Technology Organization (DSTO). His research interests include nonlinear filtering, target tracking, sensor fusion, and reasoning under uncertainty.

He co-authored two books: Beyond the Kalman filter, (Artech House, 2004), and Particle filters for *random set models* (Springer, 2013, in press) and published over 60 journal papers. He is currently an Associate Editor of the IEEE Transactions on Signal Processing.



Alex Skvortsov He holds PhD degree from Moscow Tech University in Mathematical Physics. Dr Skvortsov has been working in Defence Science and Technology Organisation (DSTO) since 2005. His current role is the Science Team Leader in Hazard Modelling. His areas of research include atmospheric dispersion, hazard source backtracking, sensor systems and mathematical biology. He published more than 50 papers.



Introduction

Signal Processing at Your Fingertips!

Let us flash back to the 1970s when the editors-in-chief of this e-reference were graduate students. One of the time-honored traditions then was to visit the libraries several times a week to keep track of the latest research findings. After your advisor and teachers, the librarians were your best friends. We visited the engineering and mathematics libraries of our Universities every Friday afternoon and poured over the IEEE Transactions, Annals of Statistics, the Journal of Royal Statistical Society, Biometrika, and other journals so that we could keep track of the recent results published in these journals. Another ritual that was part of these outings was to take sufficient number of coins so that papers of interest could be xeroxed. As there was no Internet, one would often request copies of reprints from authors by mailing postcards and most authors would oblige. Our generation maintained thick folders of hard-copies of papers. Prof. Azriel Rosenfeld (one of RC's mentors) maintained a library of over 30,000 papers going back to the early 1950s!

Another fact to recall is that in the absence of Internet, research results were not so widely disseminated then and even if they were, there was a delay between when the results were published in technologically advanced western countries and when these results were known to scientists in third world countries. For example, till the late 1990s, scientists in US and most countries in Europe had a lead time of at least a year to 18 months since it took that much time for papers to appear in journals after submission. Add to this the time it took for the Transactions to go by surface mails to various libraries in the world. Scientists who lived and worked in the more prosperous countries were aware of the progress in their fields by visiting each other or attending conferences.

Let us race back to 21st century! We live and experience a world which is fast changing with rates unseen before in the human history. The era of Information and Knowledge societies had an impact on all aspects of our social as well as personal lives. In many ways, it has changed the way we experience and understand the world around us; that is, the way we learn. Such a change is much more obvious to the younger generation, which carries much less momentum from the past, compared to us, the older generation. A generation which has grew up in the Internet age, the age of Images and Video games, the age of IPAD and Kindle, the age of the fast exchange of information. These new technologies comprise a part of their "real" world, and Education and Learning can no more ignore this reality. Although many questions are still open for discussions among sociologists, one thing is certain. Electronic publishing and dissemination, embodying new technologies, is here to stay. This is the only way that effective pedagogic tools can be developed and used to assist the learning process from now on. Many kids in the early school or even preschool years have their own IPADs to access information in the Internet. When they grow up to study engineering, science, or medicine or law, we doubt if they ever will visit a library as they would by then expect all information to be available at their fingertips, literally!

Another consequence of this development is the leveling of the playing field. Many institutions in lesser developed countries could not afford to buy the IEEE Transactions and other journals of repute. Even if they did, given the time between submission and publication of papers in journals and the time it took for the Transactions to be sent over surface mails, scientists and engineers in lesser developed countries were behind by two years or so. Also, most libraries did not acquire the proceedings of conferences and so there was a huge gap in the awareness of what was going on in technologically advanced

countries. The lucky few who could visit US and some countries in Europe were able to keep up with the progress in these countries. This has changed. Anyone with an Internet connection can request or download papers from the sites of scientists. Thus there is a leveling of the playing field which will lead to more scientist and engineers being groomed all over the world.

The aim of Online Reference for Signal Processing project is to implement such a vision. We all know that asking any of our students to search for information, the first step for him/her will be to click on the web and possibly in the Wikipedia. This was the inspiration for our project. To develop a site, related to the Signal Processing, where a selected set of reviewed articles will become available at a first “click.” However, these articles are fully refereed and written by experts in the respected topic. Moreover, the authors will have the “luxury” to update their articles regularly, so that to keep up with the advances that take place as time evolves. This will have a double benefit. Such articles, besides the more classical material, will also convey the most recent results providing the students/researchers with up-to-date information. In addition, the authors will have the chance of making their article a more “permanent” source of reference, that keeps up its freshness in spite of the passing time.

The other major advantage is that authors have the chance to provide, alongside their chapters, any multimedia tool in order to clarify concepts as well as to demonstrate more vividly the performance of various methods, in addition to the static figures and tables. Such tools can be updated at the author’s will, building upon previous experience and comments. We do hope that, in future editions, this aspect of this project will be further enriched and strengthened.

In the previously stated context, the Online Reference in Signal Processing provides a revolutionary way of accessing, updating and interacting with online content. In particular, the Online Reference will be a living, highly structured, and searchable peer-reviewed electronic reference in signal/image/video Processing and related applications, using existing books and newly commissioned content, which gives tutorial overviews of the latest technologies and research, key equations, algorithms, applications, standards, code, core principles, and links to key Elsevier journal articles and abstracts of non-Elsevier journals.

The audience of the Online Reference in Signal Processing is intended to include practicing engineers in signal/image processing and applications, researchers, PhD students, post Docs, consultants, and policy makers in governments. In particular, the readers can be benefited in the following needs:

- To learn about new areas outside their own expertise.
- To understand how their area of research is connected to other areas outside their expertise.
- To learn how different areas are interconnected and impact on each other: the need for a “helicopter” perspective that shows the “wood for the trees.”
- To keep up-to-date with new technologies as they develop: what they are about, what is their potential, what are the research issues that need to be resolved, and how can they be used.
- To find the best and most appropriate journal papers and keeping up-to-date with the newest, best papers as they are written.
- To link principles to the new technologies.

The Signal Processing topics have been divided into a number of subtopics, which have also dictated the way the different articles have been compiled together. Each one of the subtopics has been coordinated by an AE (Associate Editor). In particular:

1. Signal Processing Theory (Prof. P. Diniz)
2. Machine Learning (Prof. J. Suykens)
3. DSP for Communications (Prof. N. Sidiropoulos)
4. Radar Signal Processing (Prof. F. Gini)
5. Statistical SP (Prof. A. Zoubir)
6. Array Signal Processing (Prof. M. Viberg)
7. Image Enhancement and Restoration (Prof. H. J. Trussell)
8. Image Analysis and Recognition (Prof. Anuj Srivastava)
9. Video Processing (other than compression), Tracking, Super Resolution, Motion Estimation, etc. (Prof. A. R. Chowdhury)
10. Hardware and Software for Signal Processing Applications (Prof. Ankur Srivastava)
11. Speech Processing/Audio Processing (Prof. P. Naylor)
12. Still Image Compression
13. Video Compression

We would like to thank all the Associate Editors for all the time and effort in inviting authors as well as coordinating the reviewing process. The Associate Editors have also provided succinct summaries of their areas.

The articles included in the current editions comprise the first phase of the project. In the second phase, besides the updates of the current articles, more articles will be included to further enrich the existing number of topics. Also, we envisage that, in the future editions, besides the scientific articles we are going to be able to include articles of historical value. Signal Processing has now reached an age that its history has to be traced back and written.

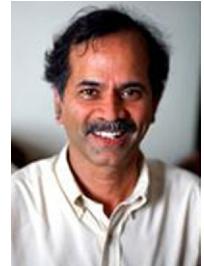
Last but not least, we would like to thank all the authors for their effort to contribute in this new and exciting project. We earnestly hope that in the area of Signal Processing, this reference will help level the playing field by highlighting the research progress made in a timely and accessible manner to anyone who has access to the Internet. With this effort the next breakthrough advances may be coming from all around the world.

The companion site for this work: <http://booksite.elsevier.com/9780124166165> includes multimedia files (Video/Audio) and MATLAB codes for selected chapters.

Rama Chellappa
Sergios Theodoridis

About the Editors

Rama Chellappa received the B.E. (Hons.) degree in Electronics and Communication Engineering from the University of Madras, India in 1975 and the M.E. (with Distinction) degree from the Indian Institute of Science, Bangalore, India in 1977. He received the M.S.E.E. and Ph.D. Degrees in Electrical Engineering from Purdue University, West Lafayette, IN, in 1978 and 1981, respectively. During 1981–1991, he was a faculty member in the department of EE-Systems at University of Southern California (USC). Since 1991, he has been a Professor of Electrical and Computer Engineering (ECE) and an affiliate Professor of Computer Science at University of Maryland (UMD), College Park. He is also affiliated with the Center for Automation Research, the Institute for Advanced Computer Studies (Permanent Member) and is serving as the Chair of the ECE department. In 2005, he was named a Minta Martin Professor of Engineering. His current research interests are face recognition, clustering and video summarization, 3D modeling from video, image and video-based recognition of objects, events and activities, dictionary-based inference, compressive sensing, domain adaptation and hyper spectral processing.



Prof. Chellappa received an NSF Presidential Young Investigator Award, four IBM Faculty Development Awards, an Excellence in Teaching Award from the School of Engineering at USC, and two paper awards from the International Association of Pattern Recognition (IAPR). He is a recipient of the K.S. Fu Prize from IAPR. He received the Society, Technical Achievement, and Meritorious Service Awards from the IEEE Signal Processing Society. He also received the Technical Achievement and Meritorious Service Awards from the IEEE Computer Society. At UMD, he was elected as a Distinguished Faculty Research Fellow, as a Distinguished Scholar-Teacher, received an Outstanding Innovator Award from the Office of Technology Commercialization, and an Outstanding GEMSTONE Mentor Award from the Honors College. He received the Outstanding Faculty Research Award and the Poole and Kent Teaching Award for Senior Faculty from the College of Engineering. In 2010, he was recognized as an Outstanding ECE by Purdue University. He is a Fellow of IEEE, IAPR, OSA, and AAAS. He holds four patents.

Prof. Chellappa served as the Editor-in-Chief of IEEE Transactions on Pattern Analysis and Machine Intelligence. He has served as a General and Technical Program Chair for several IEEE international and national conferences and workshops. He is a Golden Core Member of the IEEE Computer Society and served as a Distinguished Lecturer of the IEEE Signal Processing Society. Recently, he completed a two-year term as the President of the IEEE Biometrics Council.

Sergios Theodoridis is currently Professor of Signal Processing and Communications in the Department of Informatics and Telecommunications of the University of Athens. His research interests lie in the areas of Adaptive Algorithms and Communications, Machine Learning and Pattern Recognition, Signal Processing for Audio Processing and Retrieval. He is the co-editor of the book “Efficient Algorithms for Signal Processing and System Identification,” Prentice Hall 1993, the co-author of the best selling book “Pattern Recognition,” Academic Press, 4th ed. 2008, the co-author of the book “Introduction to Pattern Recognition: A MATLAB Approach,” Academic Press, 2009, and the co-author of three books in Greek, two of them for the Greek Open University. He is Editor-in-Chief for the Signal Processing Book Series, Academic Press and for the E-Reference Signal Processing, Elsevier.



He is the co-author of six papers that have received best paper awards including the 2009 IEEE Computational Intelligence Society Transactions on Neural Networks Outstanding paper Award. He has served as an IEEE Signal Processing Society Distinguished Lecturer. He was *Otto Monstead Guest Professor*, Technical University of Denmark, 2012, and holder of the *Excellence Chair*, Department of Signal Processing and Communications, University Carlos III, Madrid, Spain, 2011.

He was the General Chairman of EUSIPCO-98, the Technical Program co-Chair for ISCAS-2006 and ISCAS-2013, and co-Chairman and co-Founder of CIP-2008 and co-Chairman of CIP-2010. He has served as President of the European Association for Signal Processing (EURASIP) and as member of the Board of Governors for the IEEE CAS Society. He currently serves as member of the Board of Governors (Member-at-Large) of the IEEE SP Society.

He has served as a member of the Greek National Council for Research and Technology and he was Chairman of the SP advisory committee for the Edinburgh Research Partnership (ERP). He has served as Vice Chairman of the Greek Pedagogical Institute and he was for 4 years member of the Board of Directors of COSMOTE (the Greek mobile phone operating company). He is Fellow of IET, a Corresponding Fellow of the Royal Society of Edinburgh (RSE), a Fellow of EURASIP, and a Fellow of IEEE.

Section Editors

Section 1

Nicholas D. Sidiropoulos received the Diploma in Electrical Engineering from the Aristotelian University of Thessaloniki, Greece, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland—College Park, in 1988, 1990 and 1992, respectively. He has served as Assistant Professor in the Department of Electrical Engineering at the University of Virginia (1997–1999); Associate Professor in the Department of Electrical and Computer Engineering at the University of Minnesota—Minneapolis (2000–2002); Professor in the Department of Electronic and Computer Engineering at the Technical University of Crete, Chania—Crete, Greece (2002–2011); and Professor in the Department of Electrical and Computer Engineering at the University of Minnesota—Minneapolis (2011–). His research interests are in signal processing for communications, convex optimization, cross-layer resource allocation for wireless networks, and multiway analysis—i.e., linear algebra for data arrays indexed by three or more variables. His current research focuses primarily on signal and tensor analytics, with applications in cognitive radio, big data, and preference measurement. He received the NSF/CAREER award in 1998, and the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011. He served as IEEE SPS Distinguished Lecturer (2008–2009), and as Chair of the IEEE Signal Processing for Communications and Networking Technical Committee (2007–2008). He received the 2010 IEEE Signal Processing Society Meritorious Service Award. He is a Fellow of IEEE.



Section 2

Fulvio Gini (Fellow, IEEE) received the Doctor Engineer (cum laude) and the Research Doctor degrees in electronic engineering from the University of Pisa, Italy, in 1990 and 1995, respectively. In 1993 he joined the Department of Ingegneria dell'Informazione of the University of Pisa, where he became Associate Professor in 2000 and he is Full Professor since 2006. From July 1996 through January 1997, he was a visiting researcher at the Department of Electrical Engineering, University of Virginia, Charlottesville. He is an Associate Editor for the IEEE Transactions on Aerospace and Electronic Systems and for the Elsevier Signal Processing journal. He has been AE for the Transactions on Signal Processing (2000–2006) and a Member of the EURASIP JASP Editorial Board. He is the Editor-in-Chief of the Hindawi International Journal on Navigation and Observation (IJNO). He has been the Area Editor for the Special issues of the IEEE Signal Processing Magazine. He was co-recipient of the 2001 IEEE AES Society's Barry Carlton Award for Best Paper. He was recipient of the 2003 IEE Achievement Award for outstanding contribution in signal processing and of the 2003 IEEE AES Society Nathanson Award to the Young Engineer of the Year. He is a Member of the Signal Processing Theory and Methods (SPTM) Technical Committee (TC) of the IEEE Signal Processing Society and he has been a Member of the Sensor Array and Multichannel (SAM) TC. He is a Member of the Board of Directors (BoD) of the EURASIP



Society and the Award Chair since 2006 and he is the EURASIP President elected for the years 2013–2014. He was the Technical co-Chair of the 2006 EURASIP Signal and Image Processing Conference (EUSIPCO), Florence, Italy, September 2006, and the Technical co-Chair of the 2008 Radar Conference, Rome, Italy, May 2008. He was the General co-Chair of the second Workshop on Cognitive Information Processing (CIP2010) and the General co-Chair of the forthcoming ICASSP2014, to be held in Florence in May 2014. He was the guest co-editor of the special section of the Journal of the IEEE SP Society on Special Topics in Signal Processing on “Adaptive Waveform Design for Agile Sensing and Communication” (2007), guest editor of the special section of the IEEE Signal Processing Magazine on “Knowledge Based Systems for Adaptive Radar Detection, Tracking and Classification” (2006), guest co-editor of the two special issues of the EURASIP Signal Processing journal on “New trends and findings in antenna array processing for radar” (2004) and on “Advances in Sensor Array Processing (in memory of Alex Gershman)” (2013). He is co-editor and author of the book “Knowledge Based Radar Detection, Tracking and Classification” (2008) and of the book “Waveform Diversity and Design” (2012). His research interests include modeling and statistical analysis of radar clutter data, non-Gaussian signal detection and estimation, parameter estimation and data extraction from multichannel interferometric SAR data. He authored or co-authored 5 book chapters, more than 100 journal papers and about 120 conference papers.

Introduction to Signal Processing for Communications

Nicholas D. Sidiropoulos

Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA

2.01.1 Some history

Signal processing and communications have always been closely intertwined, owing in part to their common foundation on signals and systems, to the many distinguished researchers that “wore double hats” and had a pivotal role in cross-fertilizing the two disciplines, and to . . . mere necessity, since

- Analog and digital filters are critical components of any transmitter or receiver chain.
- Timing, carrier frequency, and phase offset estimation are bread-and-butter signal processing tasks.
- System identification tools from control theory and signal processing are important for channel estimation.

Beyond those core issues, there are many well-established aspects of communication theory and algorithms which were co-developed by signal processing researchers, such as adaptive equalization, blind equalization, adaptive beamforming, and signal intelligence. In the mid 1990s, a number of emerging developments, including

- multi-user communications, particularly multi-user detection;
- multi-channel communications, including multi-antenna systems and fractionally sampled equalization; and
- digital subscriber line (DSL) communications,

opened up new and exciting problems at the intersection of signal processing and communications, generating a wave and making a lasting impact to both disciplines. A decade later, wireless networking aspects were added, as ad hoc and multi-hop wireless communications, cross-layer network design, optimization, and resource allocation came into the picture. Sensor networks, cognitive radio networks, distributed detection, estimation, computation, and optimization are prominent trends in recent years. Nowadays, a significant percentage of contributions to the flagship signal processing journals are concerned with communications problems and applications—I quickly estimated 15–20% by perusing recent issues. Reciprocating, a good percentage of articles in the flagship communications journals is contributed by signal processing researchers.

Interestingly, whereas new areas and topics continuously emerge and fade away, much of what we do remains firmly anchored close to the heart of signal processing. Spectrum sensing, for example, which

is a key component of cognitive radio, is so near and dear to spectral analysis that we feel right at home with it.

Signal processing does not have a chartered application, so signal processing researchers often switch hats and move on to new application areas. But, as a good friend of mine once told me “You can chase after fashion, or wait for it to come back to what you like.” Well, he started from Ithaca! For the rest of us: happy travels in signal processing for communications and beyond!

2.01.2 Contents and contributors

When Sergios Theodoridis enlisted me as area editor for this section, I first mapped out the key titles needed, then called upon and enlisted some of the very best researchers in the area, renown not only for their research but also for their effectiveness in writing tutorials. I was delighted that they accepted my invitation. This section is largely the result of their diligent efforts, and I’m sure you will appreciate the breadth, depth, and overall quality of their work. I also owe a big thanks to the many colleagues who were kind enough to review the various contributions, providing valuable feedback and detailed, constructive comments. Many thanks are also due to Georgios Giannakis, Athanasios Liavas, Zhi-Quan (Tom) Luo, Xiaoli Ma, Evangelia Matskani, Shuichi Ohno, Ioannis Schizas, Brian Sadler, Erchin Serpedin, Slawomir Stanczak, Paschalis Tsiaflakis, Alle-Jan van der Veen, Arie Yeredor, Wei Yu, Pengfei Xia, Yan Xin, for their advice and feedback on various aspects and versions of the overall section.

2.01.2.1 Synchronization

Synchronization is the first task in any communication system, so let me begin my short guided tour from the contribution of Ghogho, Ciblat, and Swami [1], an impressively comprehensive tutorial on synchronization and associated performance analysis for a broad variety of communication systems. The clarity of exposition and right level of detail will make this article an indispensable tool in the hands of students and practitioners alike—I felt ready to program some of the methods as I was reading their article.

2.01.2.2 Channel estimation, equalization, precoding, and tracking

Channel estimation and equalization are the next step after (or intertwined with) synchronization. Channel estimation and equalization are needed to combat intersymbol interference due to multipath propagation in wireless and wireline systems. Precoding can be used as an alternative or complement to equalization, provided that channel state information is available at the transmitter. Following channel acquisition, channel tracking is used to lock on slow channel variations, at low complexity and signaling overhead. The article by Tugnait [2] is an impressive tour over two decades of research in the area, by one of the leading experts. It covers a broad spectrum of techniques, from maximum likelihood sequence estimation, linear, decision-feedback, and on to turbo equalization, training-based, semi-blind, and blind techniques, doubly-selective channels and basis expansion models, precoding (including the case of partial channel state information at the transmitter), and judicious simulations.

2.01.2.3 Blind signal separation

Much of the initial flurry of signal processing for communications research back in the mid- to late-1990s was motivated by blind signal separation and equalization problems in wireless communications, aiming to do without training in situations where the transmitter(s) are not cooperating (e.g., signal interception), or, in an attempt to boost the data rate. The article by Chevreuil and Loubaton [3] reviews important theoretical and algorithmic aspects of the blind signal separation and equalization problem with an eye toward applications in digital communications.

2.01.2.4 OFDM and multicarrier signal processing

The basic idea behind multicarrier communications was known many decades before Orthogonal Frequency Division Multiplexing (OFDM) and various other multicarrier communication modalities made their debut in practical systems. The key is that complex exponentials are eigenfunctions of linear time invariant systems, hence signaling using complex exponentials has the potential to circumvent inter-symbol interference. What made it all practical though was discrete-time baseband-equivalent processing using digital signal processing chipsets, and the use of circularity to restore exact orthogonality of the subcarriers. Banelli and Rugini [4] have spared no effort in bringing multicarrier “down to baseband,” with insightful illustrations providing valuable intuition, in a broad and technically deep tutorial that is sure to be appreciated by graduate students and practicing engineers alike.

2.01.2.5 Signal processing for vectored multichannel VDSL

Multicarrier modulation is now used across the board in all types of wireless systems, but is most prevalent in DSL. DSL systems have gone long ways in recent years, reaching previously unthinkable rates over legacy twisted pairs—owing in part to the push to get the back-end modems closer to the customer premises, and in part due to the use of multiple-input, multiple-output (MIMO) technologies—simultaneously signaling over and/or listening to multiple twisted pairs. The article by Bergel and Leshem [5] offers an insightful tutorial introduction to state-of-art MIMO DSL technology and the requisite signal processing behind it. Coming from researchers who have actually designed and built VDSL modems, the exposition strikes the right balance between theory and practice.

2.01.2.6 Distributed detection and estimation in wireless sensor networks

Wireless sensor networks (WSNs) have diverse applications in environmental, production, critical infrastructure, and radio spectrum monitoring, and new potential application domains are rapidly emerging, e.g., in monitoring renewable energy sources and the smart grid. The article by Barbarossa et al. [6] offers an intuitive, well-rounded, and technically rewarding tutorial overview of distributed and decentralized detection and estimation over WSNs, taking into account communication and energy considerations. The celebrated network *consensus* problem is revisited under realistic channel models, and basic and advanced distributed estimation and detection methodologies are explained with clear examples. The problem of matching the network topology to the given inference task is also considered.

2.01.2.7 Signal processing and optimal resource allocation for the interference channel

The article by Hong and Luo [7] is a comprehensive tour of recent developments in resource allocation and signal processing for the interference channel, coming from a leading research group in the area. As wireless network design and operation continuously pushes the technological boundaries, network optimization becomes increasingly important. As shown by Luo et al. however, most network optimization problems are very hard to solve. This represents a grand challenge and fertile ground for signal processing and communications research. Hong and Luo show how many of these problems can be (approximately) tackled using computationally efficient algorithms. The contents of this article have far broader use in many other signal processing application areas.

2.01.2.8 Advances in spectrum sensing and cross-layer design for cognitive radio networks

Last but not least is an article addressing what is currently perhaps the most actively researched topic in signal processing for communications: dynamic spectrum sensing and access using a network of collaborating sensors. The article by Giannakis et al. [8] overviews the concept of radio spectrum *cartography*, which has already drawn considerable attention. The interplay between spectrum sensing and spectrum access is further considered, leading to a simple and elegant comprehensive solution for spatio-temporal collaborative cognitive sensing and access. This article is indicative of the present focus and state of art of signal processing for communications research, by one of the leading groups in the area.

2.01.3 Outlook

No project of this type is ever complete, in part because one can always augment the fundamentals, and in part because new topics emerge and should be covered in due time. On the fundamentals front, a tutorial dedicated to multiple-input, multiple-output communications would be useful, albeit the material is partially covered in other contributions. Likewise, tutorials on channel modeling, cooperative communications, and a focused account of modern MIMO detection would be very welcome. Interference alignment has recently attracted much attention, and it brings together information theory, communications, signal processing, and optimization tools—a fertile ground for new innovations.

On the emerging topics front, research on signal processing and communication aspects of social networks is the new thing, together with signal processing tools for big data analytics. Which simply means that the E-reference in Signal Processing is looking forward to future editions!

References

- [1] M. Ghogho, P. Ciblat, Ananthram Swami, Synchronization, Elsevier Major Reference Works: E-reference Signal Processing, vol. 2.
- [2] J. Tugnait, Channel Estimation, Equalization, Precoding, and Tracking, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.

- [3] A. Chevrelui, P. Loubaton, *Blind Signal Separation for Digital Communication Data*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.
- [4] P. Banelli, L. Rugini, *OFDM and Multicarrier Signal Processing*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.
- [5] I. Bergel, A. Leshem, *Signal Processing for Vectored Multichannel VDSL*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.
- [6] S. Barbarossa, S. Sardellitti, P. Di Lorenzo, *Distributed Detection and Estimation in Wireless Sensor Networks*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.
- [7] M. Hong, Z.-Q. Luo, *Signal Processing and Optimal Resource Allocation for the Interference Channel*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.
- [8] Seung-Jun Kim, Emiliano DallAnese, Juan Andrés Bazerque, Ketan Rajawat, Georgios B. Giannakis, *Advances in Spectrum Sensing and Cross-Layer Design for Cognitive Radio Networks*, Elsevier Major Reference Works: E-reference, Signal Processing, vol. 2.

Synchronization

2

Mounir Ghogho^{*},[†], Philippe Ciblat[‡], and Ananthram Swami[§]

^{*}School of Electronic and Electrical Engineering, University of Leeds, UK

[†]International University of Rabat, Morocco

[‡]Telecom Paris-Tech, Paris, France

[§]US Army Research Lab, MD, USA

2.02.1 Introduction

Synchronization is fundamental to the proper operation of communication networks. Synchronization may be performed at various levels: from very coarse to very fine-grained; the required accuracy depends upon the application and the network environment, and hence can change dynamically. For example, coarse time synchronization suffices for many detection and tracking applications, but finer synchronization is required for high data-rate communication, distributed array processing, analog to digital conversion (ADC), slot synchronization, and cooperative communications.

According to Wikipedia, synchronization is timekeeping which requires the coordination of events to operate a system in unison. The familiar conductor of an orchestra serves to keep the orchestra in time. Systems operating with all their parts in synchrony are said to be synchronous or in sync. Some systems may be only approximately synchronized, or plesiochronous. For some applications relative offsets between events need to be determined, for others only the order of the events is important.

The aim of synchronization may be different in different applications, as illustrated next.

- Reliable coherent digital communication systems require accurate synchronization, particularly when the data rate or bandwidth is large. Synchronization consists of a number of tasks. First, carrier synchronization aims at generating a reference carrier with a phase that closely matches that of the data signal, thus compensating for the effects of oscillators mismatch and Doppler effects. This generated carrier is used at the receiver to perform coherent demodulation of the received signal to generate the baseband signal. The second task, referred to as symbol time synchronization, aims at synchronizing the receiver clock with the baseband data-symbol signal. If these tasks are carried out by multiplexing known symbols (also called training symbols), the synchronization method is referred to as data-aided; otherwise it is referred to as non-data-aided or blind. Non-data-aided synchronization techniques have been proposed in order to save bandwidth. However, there is a tradeoff between accuracy, implementation complexity and power consumption. Practical systems typically use some form of training (or pilot) symbols which may be prepended, appended, or embedded in the data packet. Training typically leads to low complexity receivers and good performance

with moderate sacrifice in rate. In addition to the above tasks, for most modern communication systems, frame and packet synchronization are also required to determine the start of the frame/packet.

- In energy constrained WSNs and MANETs, accurate clock synchronization facilitates energy-efficient Medium Access Control (MAC). Node geo-location also typically requires time synchronization.
- Many networked signal processing tasks depend critically upon a common time reference; examples include event detection and target tracking, localization, (multi-modal) sensor fusion, distributed array processing, and synchronization of distributed information caches.
- In robotics, synchronization and delay management are crucial components of networked control and actuation; poor synchronization can lead to control instabilities. Other applications in collaborative robotics, such as mapping or geo-location, also require a common time reference.
- Synchronization is crucial for distributed communications such as cooperative communications and relaying, and in network management (probing and monitoring).

Network synchronization deals with timekeeping ensuring coordination of events to operate a system in unison when universal clock and so time reference is not available. This type of synchronization is a well-studied topic with an extensive history. Typically, these works assumed high quality devices, availability of fine control of the network, extensive connectivity with little or no mutual interference, as well as known (or repeatable and measurable) propagation and processing delays [1]. These works usually lead to the development of synchronization protocols which describe the different operations that the network nodes have to do in order to be jointly synchronized. This type of synchronization is of great importance in wireless sensors networks (WSN) and mobile ad-hoc networks (MANET) where a coordinating entity with a time reference may not be available. Surveys of WSN synchronization protocols may be found in [1–6].

Notice that network synchronization is a very old problem. Indeed, a fascinating book by Peter Galison [7] describes the pioneering work of Einstein and Poincaré on finding common time references (circa 1909). Einstein, then a young obscure German physicist, was experimenting with measuring time using telegraph networks and with the coordination of clocks at train stations. The renowned mathematician Poincaré, president of the French Bureau of Longitude, was mapping time coordinates across continents. Sundials and watches had been in use for centuries and they were relatively inaccurate. However, travel and communications were slow (until at least the end of the nineteenth century), and those local time differences were of little importance. Synchronization is often encountered in nature. Strogatz [8] provides an elegant description of synchronization among fireflies in Malaysia, and circadian rhythms. Studies of the robustness of circadian clocks (e.g., via clock neurons in the superchiasmatic nucleus) indicate that individual neurons are sloppy timekeepers but synchronized neurons are precise clocks. Bio-inspired approaches have led to the development of algorithms for synchronizing communication networks (e.g., transmit beamforming, GPS) that exploit combinations of local and global signaling [9]. The process of synchronization, whether it be in nature or in engineered systems, can also be re-interpreted via more recent tools developed for distributed control [10], and opinion dynamics [11], since the goal is broadly to coordinate actions in order to reach a common objective.

The ensuing sections in this chapter will study the synchronization problem. Our main focus is on time and carrier synchronization when a reference is available (typically via a base station). In such a context,

we will consider both single and multi-carrier systems, both single and multi-user scenarios, and both flat-fading and non-flat-fading (frequency-selective) channels. In the last section, a brief state-of-the-art on network synchronization is provided.

The chapter is written at the beginner graduate level. The reader is assumed to have a working knowledge of digital communications (e.g., [12, 13]) and estimation theory (e.g., [14]).

2.02.2 Synchronization in flat fading channel

In this section, we assume that the multipath delay spread in the propagation channel does not lead to Inter-Symbol Interference (ISI). In practice, this holds if the symbol period is sufficiently larger than the delay spread of the channel. The frequency response of the channel is thus considered flat. From a practical point of view, such an assumption on the channel is satisfied in satellite communications (e.g., DVB-S), anisotropic transmissions (e.g., between two DVB-T transmitters), optical fiber communications (e.g., with pre-compensation of the static chromatic dispersion) [15, 16].

As the channel does not induce ISI, the OFDM modulation scheme does not provide advantages over single-carrier transmissions. Therefore, we will only focus on the single-carrier case in this Section. The transmitted signal (in baseband) [12, 13] is given by

$$x_a(t) = \sum_{n=0}^{N-1} s_n h_a(t - kT_s),$$

where

- $\{s_0, \dots, s_{N-1}\}$ are the transmitted symbols belonging to ASK, PSK or QAM constellations or any linearly precoded scheme, with N being the sequence length used for estimating the synchronization (sync) parameters;
- T_s is the symbol period which is assumed to be known throughout this chapter. For more information about the blind estimation of T_s , the reader is referred to [17];
- $h_a(t)$ is the shaping filter that classically is a square-root Nyquist filter. For example, in 3G systems, it is the square-root raised cosine filter with roll-off 0.35. The main property that we need further is that the filter is band-limited with bandwidth between $1/T_s$ and $2/T_s$.

For the flat fading channel, the received signal (in baseband) is described by

$$y_a(t) = x_a(t - \tau_0) e^{2\pi(f_0 t + \phi_0)} + b_a(t),$$

where

- the synchronization parameters are the symbol timing τ_0 , the (constant) phase ϕ_0 , and the carrier frequency offset (CFO) f_0 [18–20].
- $b_a(t)$ is the complex-valued circularly-symmetric white zero-mean Gaussian process with variance N_0 per real dimension (for more details about the assumption on the noise, see [13]). The concept of circularity is crucial and will be defined in detail later.

The received signal can be re-written as follows:

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n h_a(t - kT_s - \tau_0) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t).$$

We do not assume an *a-priori* distribution for the timing and phase parameters. Indeed, the timing can be assumed uniformly distributed across the interval $[0, T_s]$ since the receiver and the transmitter are not synchronized in time. Due to the channel propagation, the phase can take any value over $[0, 2\pi)$. In contrast, the CFO, which is due either to a mismatch between local oscillators or Doppler effects, lies in a pre-defined interval. To illustrate this, let us consider the worst case when only a cheap local oscillator is available. Its precision is about 40 ppm which leads to a CFO of 40 kHz at carrier frequency 1 GHz. Assuming a rather high vehicle speed of 360 km/h, the Doppler induced offset is upper-bounded by 333 Hz. Consequently, the CFO is much smaller than the typical signal bandwidth which is often equal to several MHz. Note that the main source of CFO is the local oscillator mismatch and not the Doppler effect. Indeed, the Doppler effect has a greater influence on the coherence time of the channel and thus on the statistical model of the channel (for instance, according to the Doppler value, the channel can be viewed as a fast flat fading channel or a slow flat fading channel). The nature of the flat fading channel (fast or slow) is crucial for designing properly the communication scheme (feedback link, diversity management, etc.) but not the synchronization step since, except in very infrequent cases, the synchronization step duration is always much smaller than the channel coherence time.

Before going further, we recap the optimal symbol detector when synchronization is perfect ($\tau_0 = 0$, $\phi_0 = 0$, and $f_0 = 0$) [12, 13, 18, 19]. If the information symbols are equally likely, the maximum likelihood detector is the optimal one in the sense of error probability minimization. Therefore, we have that

$$\{\hat{s}_n\}_{n=0}^{N-1} = \arg \max_{\{s_n\}_{n=0,\dots,N-1}} p(y(t)|\{s_n\}_{n=0,\dots,N-1}).$$

Due to the Gaussianity of the noise process, we have that [21]

$$p(y(t)|\{s_n\}_{n=0,\dots,N-1}) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h(t - nT_s)|^2 dt / 2N_0}.$$

Thus

$$\hat{s}_n = \arg \min_{\{s_n\}_{n=0,\dots,N-1}} \int_{\mathbb{R}} \left| y(t) - \sum_{n=0}^{N-1} s_n h(t - nT_s) \right|^2 dt.$$

We finally obtain that

$$\hat{s}_n = \arg \min_{\{s_n\}_{n=0,\dots,N-1}} J_N(s),$$

where

$$J_N(s) = \int_{\mathbb{R}} |y(t)|^2 dt + \sum_{n,n'=0}^{N-1} s_n s_{n'}^* \tilde{h}(n - n') - 2 \sum_{n=0}^{N-1} \Re\{s_n^* z(n)\},$$

where

- $z_a(t) = h_a(-t)^* \star y_a(t)$ is the continuous-time output of the so-called matched filter;
- $z(n) = z_a(nT_s)$ is the sampled version (at symbol rate) of the matched filter output;
- $\tilde{h}_a(t) = h_a(-t)^* \star h_a(t)$ is the continuous-time equivalent filter; and
- $\tilde{h}(n) = \tilde{h}_a(nT_s)$ is the discrete equivalent filter.

We remark that the optimality criterion depends on the received signal only through the discrete-time matched filter output. The introduction of the signal $z(n)$ can be also justified by following an alternative way: the received signal (under the perfect synchronization assumption) can be re-written as follows:

$$y_a(t) = \sum_n s_n \Phi_n(t) + b_a(t),$$

where $\Phi_n(t) = h_a(t - nT_s)$. As the useful information in $y_a(t)$ are the symbols s_n , the received signal can be split into two parts: the useful one associated with the signal subspace spanned by the functions $\Phi_n(t)$, and the other one generated by the space orthogonal to the signal subspace, the so-called ‘‘noise subspace.’’ Let $\tilde{z}(n) = \langle y_a(t) | \Phi_n(t) \rangle$ be the projected signal onto the signal subspace, where $\langle \bullet | \bullet \rangle$ denotes the canonical inner product. One can easily check that $\tilde{z}(n) = z(n)$. As the noise is Gaussian, its contribution to the signal subspace is independent of those to the orthogonal noise subspace. Therefore, without loss of optimality, one needs to only work with $\tilde{z}(n) = z(n)$ and can drop the projection onto the noise subspace.

Moreover, one can easily see that

$$z(n) = \sum_k \tilde{h}(k) s_{n-k} + \tilde{b}(n), \quad (2.1)$$

where $\tilde{b}(n) = h_a(-t)^* \star b_a(t)|_{t=nT_s}$ is circularly-symmetric Gaussian noise with zero mean and power spectral density $S_{\tilde{b}}(e^{2i\pi f}) = 2N_0 \tilde{h}(e^{2i\pi f}) = 2N_0 \sum_k \tilde{h}(k) e^{-2i\pi kf}$. The maximum likelihood (ML) criterion depends on the shaping and propagation filters only through the so-called discrete-time equivalent filter \tilde{h} . Therefore the system performance will only be driven by the filter \tilde{h} and the signal-to-noise ratio (SNR).

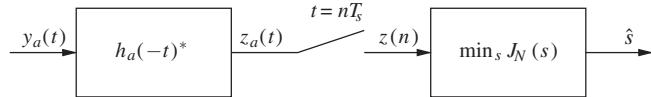
Now the second step of the optimal detector is to find the minimum of $J_N(s)$. When the shaping filter $h_a(t)$ is a square-root Nyquist filter, it is well known that $\tilde{h}_a(t)$ is the Nyquist filter and $\tilde{h}(n) = \delta_{0,n}$ where $\delta_{0,n}$ is the Kronecker index. Thus, the function $J_N(s)$ can be significantly simplified to

$$J_N(s) = \sum_{n=0}^{N-1} |z(n) - s_n|^2.$$

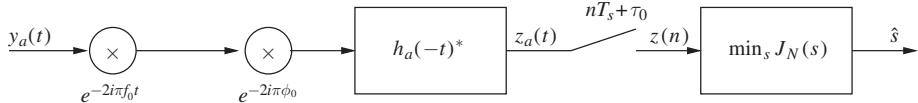
Consequently, the optimal detector is a symbol-by-symbol detector

$$\hat{s}_n = \arg \min_{s_n} |z(n) - s_n|^2$$

which is the so-called threshold detector. When the Nyquist condition is not satisfied (especially when the channel is not flat), the minimization of $J_N(s)$ is much harder and can be done via the famous

**FIGURE 2.1**

Optimal receiver structure (with perfect synchronization).

**FIGURE 2.2**

Optimal receiver structure (with known synchronization parameters).

Viterbi algorithm [22,23]. When the Viterbi algorithm is too complex (channel too long and/or high modulation size), suboptimal detectors, such as the zero-forcing (ZF), minimum mean-square error (MMSE), decision-feedback equalizer (DFE), can be used. For details, mathematical explanations and derivations, we refer the reader to [12,13]. The optimal receiver is summarized in Figure 2.1.

When the synchronization parameters are non-zero but known, the continuous-time received signal is given by

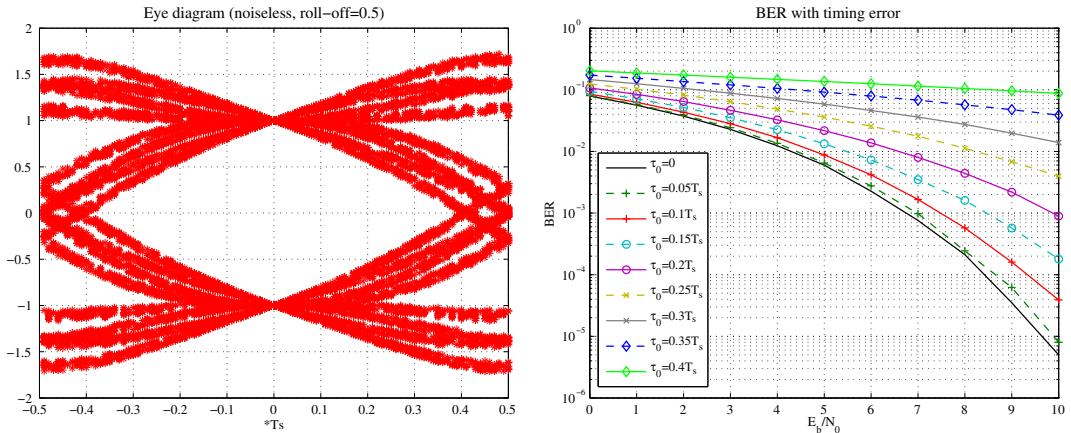
$$y_a(t) = \sum_n s_n \Psi_n(t - nT_s) + b_a(t)$$

with $\Psi_n(t) = h_a(t - nT_s - \tau_0) e^{2i\pi(f_0 t + \phi_0)}$. Once again, the optimal operation at the receiver side is to project $y_a(t)$ onto the function $\Psi_n(t)$. Therefore, the optimal receiver is now given in Figure 2.2.

Notice that the phase compensator can be located anywhere in Figure 2.2 since it commutes with the other operators. The CFO compensator can be located anywhere if and only if the CFO is small enough compared to the filter bandwidth¹. This last assumption is usually satisfied as mentioned earlier.

Let us now move on to some performance evaluation when the synchronization step is not carried out. In order to understand the influence of each parameter, we will consider the sync parameters one by one, assuming the remaining ones known. Let us first focus on the timing. In the flat fading context, the shaping filter is usually a square-root Nyquist and more precisely a square-root raised cosine filter with roll-off ρ . Under the perfect synchronization assumption, $z(n)$ is not distorted by Inter-Symbol Interference (ISI). But if the timing is not perfectly known, $z(n)$ will be affected by ISI. In Figure 2.3 (left), we display the eye diagram of $z_a(t)$ when $\tau_0 = 0$ and $\rho = 0.5$ with BPSK modulation. We remark that if the sampling operation is not done at a multiple of T_s , the eye will be less open and performance will be degraded. In Figure 2.3 (right), the bit-error rate (BER) versus SNR E_b/N_0 curves are shown

¹Let $h_1(t)$ and $h_2(t)$ be two filters of bandwidth $1/T_s$. Let us assume also that $f_0 T_s \ll 1$. We have that $h_1(t) \star (h_2(t) e^{2i\pi f_0 t}) = e^{2i\pi f_0 t} \int H_2(v) H_1(v + f_0) e^{2i\pi v t} dv \approx e^{2i\pi f_0 t} (h_1(t) \star h_2(t))$ where H_k is the Fourier transform of h_k . The approximation holds since f_0 is small compared to $1/T_s$. We thus conclude that the CFO operation can be permuted with the filtering operator.

**FIGURE 2.3**

Eye diagram (left) in the noiseless case and BER (right) versus E_b/N_0 in the presence of timing mis-synchronization.

for different values of τ_0 . Notice that performance degrades significantly when the timing error exceeds 10% of the symbol period.

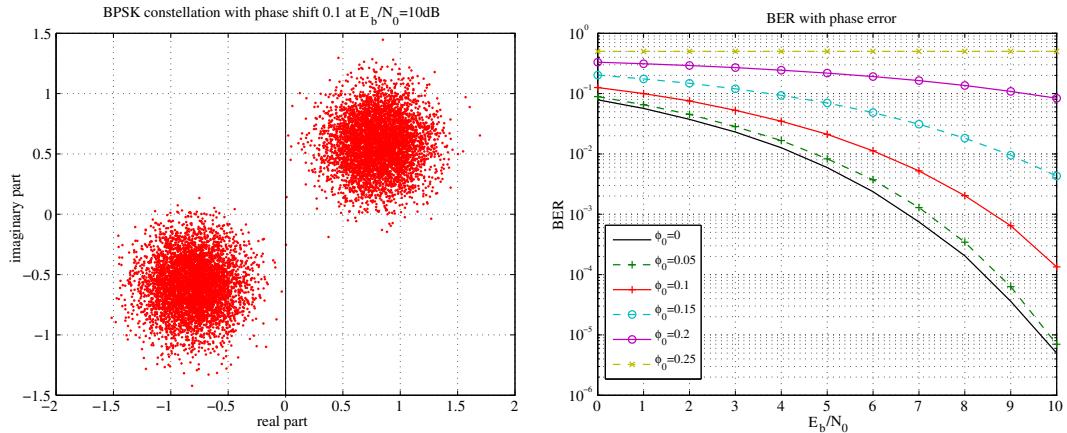
We next consider perfect timing synchronization, but without any phase synchronization. The shaping filter is once again a square-root raised cosine filter with roll-off $\rho = 0.5$. In Figure 2.4 (left), we plot the samples $z(n)$ when BPSK is employed and the phase shift is equal to 0.1 at $E_b/N_0 = 10$ dB. The constellation is thus rotated, which induces an increase of the BER. In Figure 2.4 (right), we display the BER versus E_b/N_0 for different values of the phase shift. Performance degradation is significant.

We next examine the influence of the CFO on performance. The same simulation set-up is used as previously. In Figure 2.5 (left), we plot the samples $z(n)$ when the CFO is equal to 0.01 at $E_b/N_0 = 30$ dB. We remark that the BPSK constellation is rotated with different rotation angles f_{0n} at each time index n which leads to a circle if the frame is long enough. Due to the noise, we observe a ring. In Figure 2.5 (right), we compute the BER versus E_b/N_0 for different values of CFO. The frame length in this example is 1000 data symbols.

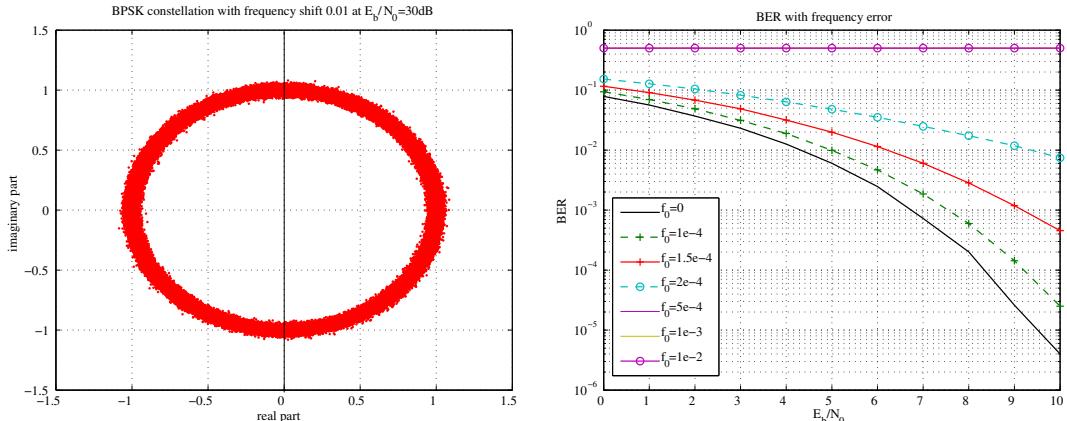
For each synchronization parameter, a loss in performance is observed, which may lead to communication failure. Therefore, we clearly need to add a synchronization step at the receiver side in order to keep the performance as high as possible.

As we just saw, synchronization parameters have a great influence on system performance; therefore, estimating these parameters is crucial. The following remarks are in order:

- Why should we estimate the timing parameter τ_0 ? Knowledge of τ_0 is crucial for choosing the correct sampling time when the continuous-time signal is converted to the (symbol rate) discrete-time signal (see Figure 2.3). The need for sampling at the symbol rate at the exact timing has its origin in the non-satisfaction of the Nyquist-Shannon sampling theorem. Therefore, by oversampling at least at the baud rate (but in practice twice the symbol rate), the discrete-time signal

**FIGURE 2.4**

Samples $z(n)$ before decision (left) when $\phi_0 = 0.1$ and $E_b/N_0 = 10$ dB and BER (right) versus E_b/N_0 in the presence of phase mis-synchronization.

**FIGURE 2.5**

Samples $z(n)$ before decision (left) when $f_0 = 0.01$ and $E_b/N_0 = 30$ dB and BER (right) versus E_b/N_0 in the presence of frequency mis-synchronization.

$\tilde{y}(n) = y_a(nT_s/2)$ contains all the information in $y_d(t)$. Then the ‘‘symbol timing’’ matching can be done via digital processing, namely, interpolation. More interestingly, by considering the bivariate process $\mathbf{y}(n) = [y_a(nT_s), y_a(nT_s + T_s/2)]^T$ which corresponds exactly to $\tilde{y}(n)$, a fractionally-spaced (FS) equalizer can be used to retrieve the information symbol. To implement this, one needs to know the filters $h_1(n) = h_a(nT_s)$ and $h_2(n) = h_a(nT_s + T_s/2)$. Thus, the symbol timing

estimation issue has been replaced by a pair of filter estimation. Notice that at least one filter (usually $h_2(n)$ does not satisfy the Nyquist criterion) is non-flat (i.e., frequency-selective) in its discrete-time version. Another way to cope with timing offset is to incorporate the timing into the channel by rewriting the received signal as:

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n \check{h}_a(t - kT_s) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t),$$

where

$$\check{h}_a(t) = h_a(t - \tau_0),$$

is the new equivalent channel. Now the channel is unknown since τ_0 is unknown. Therefore, once again the timing estimation issue boils down to a channel estimation issue.

- As the phase rotation is a linear time-invariant operation, it can be viewed as a filtering operator and thus incorporated into the filter. Therefore we have that

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n \mathring{h}_a(t - kT_s) \right) e^{2i\pi f_0 t} + b_a(t),$$

where

$$\mathring{h}_a(t) = h_a(t - \tau_0) e^{2i\pi \phi_0}$$

and the phase estimation issue can be avoided.

- The CFO cannot be viewed as the modification of a linear filter since the CFO is not a time-invariant transformation. Therefore, the CFO estimation issue cannot be avoided.

In the case of a flat fading channel, if channel estimation is used to synchronize in time and phase (in order to avoid the “real” synchronization steps), we do not use all of our knowledge about the signal, and in particular its Nyquist property. Nevertheless, if the synchronization step has to be avoided, the estimation issue deals with the joint frequency and (non-flat) channel estimation issue. This problem can be treated similarly to the frequency estimation issue in the flat channel context [24,25] and so will be omitted here. In the non-data-aided mode, the CFO can be estimated regardless of the channel, as shown in Section 2.02.2.2, and the channel can also be estimated by following the approaches given in the “Channel Estimation” E-reference.

The communication protocol permits the transmitter to explicitly adjust its signaling so as to facilitate receiver side synchronization. Therefore, there are two main classes of estimation problems:

- Data-Aided (DA)*²: the receiver knows a sequence of transmitted symbols s_n . These symbols belong to the so-called “training sequence.” Obviously, this leads to a loss in spectral efficiency since during the transmission of the training symbols, no information symbols are sent. Nevertheless this approach (which is very popular in civilian applications such as GSM, 3G, DVB-S) has several advantages since it enables us to dramatically simplify the design and the implementation of the

²Also called training approach, supervised mode.

synchronization parameter estimators as will be seen later. Moreover, good performance can usually be attained with only a few (training) symbols, and thus the cost in spectral efficiency remains quite low and acceptable.

- *Non-Data-aided (NDA)*³: the transmitter does not send any training sequences per se; thus the receiver does not have deterministic knowledge about a symbol sub-sequence. However, the receiver will have some structural or statistical information about the symbols such as the nature of the constellation, the correlation between the symbols, etc. Such a scenario obviously occurs in passive listening in security applications. In civilian applications, the NDA approach is sometimes useful for tracking parameter fluctuations or in broadcasting applications (e.g., TV). Indeed, the TV application differs completely from a peer-to-peer application (such as GSM between the mobile and the Base Station) since the TV receiver can be switched on at any moment and will not warn the broadcast transmitter that it is switched on (for otherwise the broadcast transmitter would always be “interrupted,” since there is no dedicated channel for learning the channel and synchronization parameters). So the broadcast transmitter will not transmit a training sequence when a TV receiver goes live in the network. To speed up the process, the TV receiver has to synchronize itself in a blind manner.

Some remarks before going further:

1. The main property used in this section compared to the non-flat fading channel case is that the channel here is assumed to be known. The fact that the filter is also a square-root Nyquist is not required but will greatly simplify some derivations and thus the algorithms.
2. The estimation procedure that we will develop can be used and adapted (sometimes) to the non-flat fading channel case.
3. The trade-off between the estimation quality (provided by training) and the remaining time for transmitting data can be studied within the framework of information theory (see [26] for more details).

2.02.2.1 DA case

Let $\hat{\theta}_N$ be an unbiased estimator of the sought parameter θ relying on N observation samples. Based on [21, 27], we know that the Mean Square Error (MSE), defined as $\mathbb{E}[\|\hat{\theta} - \theta\|^2]$ of any unbiased estimator of θ is lower-bounded by the so-called Cramer-Rao bound (CRB) which will be described mathematically in this chapter. An estimator whose MSE is equal to the CRB is called *efficient*. Efficient estimators do not exist for many estimation problems. Therefore, the notion of asymptotic efficiency has been introduced. This means that the ratio between the MSE and the CRB tends to 1 as the number of samples N goes to infinity. Under mild conditions (see e.g., [21, 27]), the Maximum-Likelihood estimator is asymptotically efficient and normally distributed.

In view of the asymptotic efficiency and normality, it is natural to consider the maximum-likelihood estimator first. If the ML estimator can be implemented, one often considers the problem to be solved. In contrast, if the ML estimator cannot be implemented (because of its complexity or sometimes even intractability), the estimation issue is open and other estimators have to be found. We hasten to add that

³Also called blind, unsupervised.

if an efficient estimator does not exist, then it may be possible to find better estimators than the ML. Second, the CRB may not be tight when the number of samples (or SNR) is low, and tighter bounds may (such as the Bhattacharya and Ziv-Zakai bounds) may need to be considered. We refer the reader to [27] for details.

So let us start with the introduction of the ML estimator for the joint synchronization parameters. The likelihood function can be written as follows

$$p(y(t)|\tau, \phi, f) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h_a(t-nT_s-\tau) e^{2i\pi(f t + \phi)}|^2 dt / 2N_0},$$

where the training sequence, $\{s_n\}$, is known. The ML estimator is given by

$$[\hat{\tau}_N, \hat{\phi}_N, \hat{f}_N] = \arg \min_{\tau, \phi, f} \int_{\mathbb{R}} \left| y(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau) e^{2i\pi(f t + \phi)} \right|^2 dt.$$

Setting the derivative to zero, and assuming that the CFO is small compared to the signal bandwidth, we obtain the following set of equations:

$$\begin{cases} \Re \left\{ \sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} e^{-2i\pi f T_s n} z'_\tau(n) \right\} = 0, \\ \Im \left\{ \sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} e^{-2i\pi f T_s n} z_\tau(n) \right\} = 0, \\ \Im \left\{ \sum_{n=0}^{N-1} s_n^* e^{-2i\pi\phi} n e^{-2i\pi f T_s n} z_\tau(n) \right\} = 0, \end{cases} \quad (2.2)$$

where

$$\begin{aligned} z_\tau(n) &= \int_{\mathbb{R}} y_a(t) h_a(t - nT_s - \tau)^* dt = h_a(-t)^* \star y_a(t) |_{t=nT_s+\tau} \\ z'_\tau(n) &= \int_{\mathbb{R}} y_a(t) h'_a(t - nT_s - \tau)^* dt = h'_a(-t)^* \star y_a(t) |_{t=nT_s+\tau} \end{aligned}$$

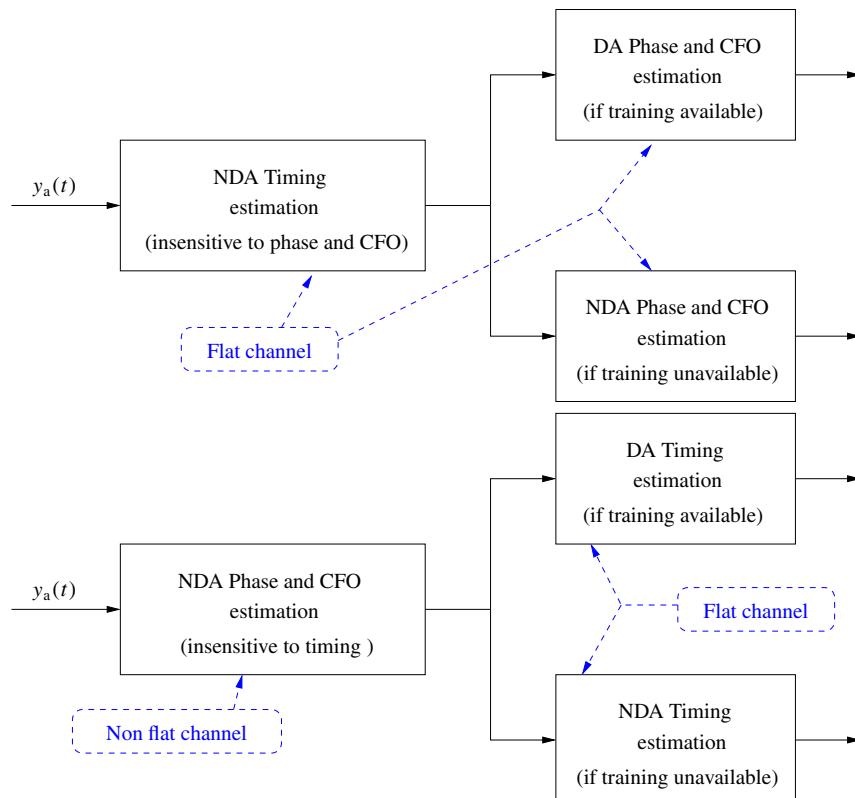
with $h'_a(t)$ is the derivative of $h_a(t)$.

A typical approach is variable elimination: we try to express a parameter in terms of the others, and eliminate it by using the previously-found expression. Here, this algebraic manipulation can be applied to the phase (as a function of the timing and the CFO), thanks to Eq. (2.2). But the timing and the CFO cannot then be written as explicit functions due to the non-linearity of these equations. Therefore, the joint (timing, phase, CFO) problem is intractable.

We remark that joint frequency and phase estimators can be developed when the timing is known. Indeed, as we will see, if the timing is known (and thus assumed to be zero without loss of generality), then the second and third equations can lead to practical estimators with reasonable computational complexity. Therefore, we split our problem into two different problems that will be treated separately: (i) the timing sub-task and (ii) the phase and CFO estimation sub-tasks.

The ordering of the two problems has a great influence on the nature of estimators that will be used, as we see next:

- *First scheme:* Timing is estimated first and then the phase and CFO are estimated, assuming that the timing estimate is perfect. As the DA timing cannot be derived in closed-form when the phase and CFO are unknown, we need to develop a NDA timing estimator that is insensitive to the actual phase and CFO values. The second part of this scheme deals with phase and CFO estimations issues. As the timing is now known, DA estimators can be developed. If training is not available, NDA estimators can also be considered.
- *Second scheme:* the phase and the CFO are estimated first; timing is then estimated, assuming that the phase and CFO estimates are perfect, and thus perfectly corrected. Once again, the phase and CFO estimators here have to be NDA and insensitive to timing error. As the timing can be wrong, the sampled filter can generate ISI. Therefore, we need to design joint phase and CFO NDA estimators that can work even when the (non-flat) fading channel is unknown. Assuming that phase and CFO are perfectly compensated, the timing estimator can be blind or aided by a training sequence.

**FIGURE 2.6**

Summary of the first scheme (top) and the second scheme (bottom).

The two schemes are summarized in Figure 2.6.

Consequently, we need to solve the following issues (even if training sequences are assumed available!):

- *Problem 1:* DA Phase and CFO estimation (when timing is known and so can be considered to be zero, without loss of generality (wlog)).
- *Problem 1':* DA Timing estimation (when Phase and CFO are known and so can be assumed to be zero wlog).
- *Problem 2:* NDA Phase and CFO estimation (insensitive to unknown timing and hence a non-flat channel).
- *Problem 2':* NDA Timing estimation (insensitive to unknown phase and CFO).

Notice that the first scheme (associated with Problems 1 and 2') is used frequently and advocated in a lot of practical systems. Hereafter, we first focus on Problem 1 and Problem 1' relying on training sequences.

2.02.2.1.1 DA timing estimation

In this subsection, we focus on the following estimation problem:

$$y_a(t) = \sum_{n=0}^{N-1} s_n h_a(t - nT_s - \tau_0) + b_a(t).$$

After simple algebraic manipulations (similar to those associated with the maximum-likelihood detector), the ML timing estimator can be expressed as:

$$\hat{\tau}_N = \arg \max_{\tau} \underbrace{\sum_{n=0}^{N-1} \Re\{s_n^* z_\tau(n)\}}_{J_N(\tau)}, \quad (2.3)$$

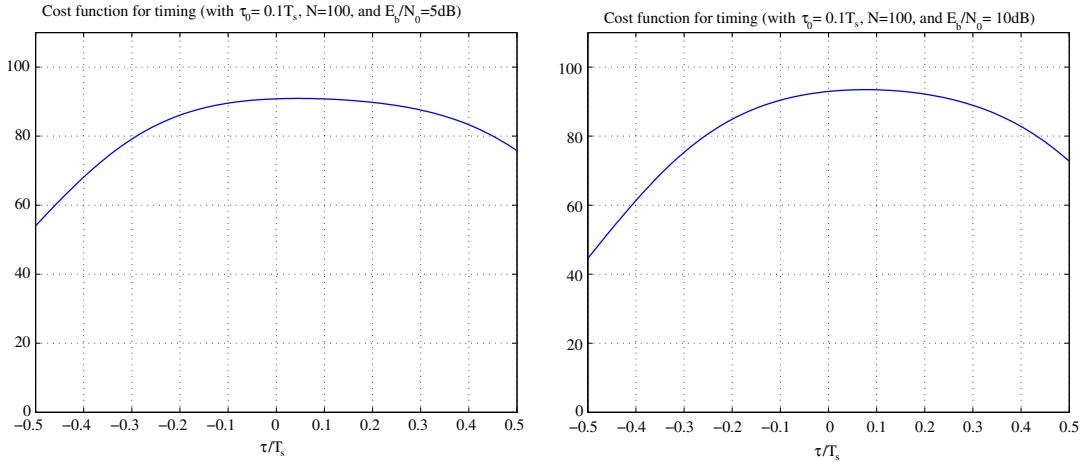
where we recall that $z_\tau(n)$ is the output of the matched filter sampled at $nT_s + \tau$.

The cost function $J_N(\cdot)$ of (2.3) is plotted in Figure 2.7 for $E_b/N_0 = 5$ dB (left) and $E_b/N_0 = 10$ dB (right) with $N = 100$. The shaping filter is a square-root raised cosine with roll-off $\rho = 0.5$. The value of the sought τ_0 is $0.1T_s$ and the x -axis has been normalized by T_s .

Observe that the cost function is concave around the true point. Therefore, one can proceed to find the maximum of $J_N(\cdot)$ in two steps

- A *coarse* search through a 1-D grid which provides a first estimate $\hat{\tau}_N^{(0)}$ of τ_0 .
- Once the coarse search has roughly localized the maximum, one can use a gradient-descent algorithm on the function $J_N(\cdot)$ initialized by $\hat{\tau}_N^{(0)}$, as follows

$$\begin{aligned} \hat{\tau}_N^{(m)} &= \hat{\tau}_N^{(m-1)} + \mu J'_N(\hat{\tau}_N^{(m-1)}) \\ &= \hat{\tau}_N^{(m-1)} + \mu \sum_{n=0}^{N-1} \Re \left\{ s_n^* \frac{\partial z_\tau(n)}{\partial \tau} \Big|_{\hat{\tau}_N^{(m-1)}} \right\}, \end{aligned} \quad (2.4)$$

**FIGURE 2.7**

Cost function $J_N(\tau)$ for $E_b/N_0 = 5$ dB (left) and $E_b/N_0 = 10$ dB (right) with $N = 100$ and $\tau_0 = 0.1T_s$.

where m is the iteration index of the gradient algorithm. The implementation of the derivative term is actually quite easy since its corresponds to the sampled version (at $\hat{\tau}_N^{(m-1)}$) of the output of the received signal passed through the matched filter associated with the derivative filter $h'_a(t)$. This can be done digitally (via interpolation) if the received signal $y_a(t)$ has been sampled at twice the symbol rate or faster.

In order to benchmark different estimators associated with the same estimation issue, it is useful to have a performance lower-bound on the MSE. The CRB is the most popular one because it is quite simple to derive (especially in DA mode) and tight (we remind that under mild conditions, the ML is asymptotically efficient). We remind below the CRB definition.

$$\text{CRB}(\tau) = \frac{1}{\mathbb{E} \left[\left(\frac{\partial \log p(y|\tau)}{\partial \tau} \Big|_{\tau_0} \right)^2 \right]} = - \frac{1}{\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \Big|_{\tau_0} \right]}.$$

After standard algebraic manipulations, we have

$$\log p(y|\tau) = \frac{1}{N_0} \Re \left\{ \sum_{n=0}^{N-1} s_n^* \int_{\mathbb{R}} y_a(t) h_a(t - nT_s - \tau)^* dt \right\} + \text{const.}$$

Thus, we obtain

$$\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} = \frac{1}{N_0} \Re \left\{ \sum_{n=0}^{N-1} s_n^* \int_{\mathbb{R}} y_a(t) h_a''(t - nT_s - \tau)^* dt \right\},$$

with $h_a''(t)$ the second derivative function of $h_a(t)$. By taking the mathematical expectation (and using the zero-mean property of the noise), we obtain

$$\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \right] = \frac{1}{N_0} \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} h_a(t - n' T_s) h_a''(t - n T_s - \tau)^* dt \right\}.$$

Using Parseval's identity, we finally obtain

$$\mathbb{E} \left[\frac{\partial^2 \log p(y|\tau)}{(\partial \tau)^2} \right] = -\frac{4\pi^2}{N_0} \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} f^2 |H(f)|^2 e^{2i\pi(n'-n)fT_s} df \right\},$$

where $H(f)$ is the Fourier transform of $h_a(t)$.

Consequently, the CRB for the timing estimation issue is given by

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 \Re \left\{ \sum_{n,n'=0}^{N-1} s_n^* s_{n'} \int_{\mathbb{R}} f^2 |H(f)|^2 e^{2i\pi(n'-n)fT_s} df \right\}}.$$

This expression shows the influence of the shaping filter through the integral of the square of $f H(f)$. But the influence of the training sequence (especially its size) is still quite unclear. To address this issue, we will model our training sequence as a realization of a random process. More precisely, s_n is a realization of pseudo-noise stationary process. In practice, the training sequence is generally generated through a shift register. Let us consider $r_s(m) = \mathbb{E}[s_{n+m} s_n^*]$ the autocorrelation function and the associated spectrum $S_s(e^{2i\pi f}) = \sum_m r_s(m) e^{-2i\pi m f}$. By using results on Cesaro sums, one can prove that

$$\frac{1}{N} \sum_{n,n'=0}^{N-1} \mathbb{E}[s_n s_n^*] e^{2i\pi(n'-n)fT_s} \xrightarrow{a.s.} S_s(e^{-2i\pi fT_s}), \quad \text{when } N \rightarrow \infty.$$

As a consequence, we have that

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 N \int_{\mathbb{R}} f^2 |H(f)|^2 S_s(e^{-2i\pi fT_s}) df}.$$

For channel estimation (when synchronization is perfectly done), a white training sequence (i.e., with flat spectrum) is optimal [28,29]. But a white sequence is not necessarily the best choice for the synchronization parameters.⁴ Further, the best training sequence correlation property may be different for different synchronization parameters. Since a white sequence always leads to reasonable performance, it provides a good trade-off. Therefore, a white training sequence is chosen in current real-life systems.

⁴For example, for the timing, the best sequence is one whose spectrum is $S_s(e^{2i\pi fT_s}) = \sigma_s^2 \delta((f - f_{\max})T_s)$ where $f_{\max} = \arg \max_f f^2 |H(f)|^2$.

When the training sequence is white, the CRB simplifies to

$$\text{CRB}(\tau) = \frac{N_0}{4\pi^2 E_s N \int_{\mathbb{R}} f^2 |H(f)|^2 df},$$

where E_s is the variance of s_n .

The CRB provides some insights about the behavior of the estimates.

- CRB is $\mathcal{O}(1/N)$. This is reasonable since the CRB associated with channel estimation (cf. “Channel estimation” E-reference) is also $\mathcal{O}(1/N)$. We recall that the timing could be incorporated into the channel estimation box without loss in performance. Thus, fortunately the CRB offers the same behavior.
- CRB is $\mathcal{O}(1/\text{SNR})$. Once again, this result is consistent with that one associated with channel estimation.
- The influence of the shaping filter can be analyzed, especially the influence of the roll-off factor. Recall that the system is most sensitive to timing error if the roll-off is small and thus if the occupied extra bandwidth is small. Unfortunately, the quality of estimation also decreases when the roll-off becomes small (see [18, 19] for more details).

So far, we have considered a blockwise approach. In the past, adaptive approaches (i.e., sample-by-sample) were of great interest due to buffer size limitation and computational complexity. Today, the adaptive approach is mainly useful for tracking parameter variations during data (and not training) transmission. The parameter update can be implemented under two different modes: (i) the NDA mode since it is carried out during data transmission and (ii) the Decision-Directed (DD) mode. In the NDA mode, different updates can be developed and usually are obtained through the adaptive version of the blockwise approaches described in Section 2.02.2.2. In the DD mode, different updates are obtained by calculating the adaptive version of the DA estimators and then replacing the training s_n with a decision on data symbol \hat{s}_n . Therefore, it is useful to develop an adaptive version of the previously-described DA ML timing estimator. Thus, instead of working block-by-block, we work sample-by-sample. The (stochastic) gradient-descent algorithm at time n can be derived from the blockwise version of the gradient descent algorithm (cf. Eq. (2.4)) by keeping only the derivative term associated with time n . We thus have

$$\hat{\tau}_n = \hat{\tau}_{n-1} + \mu \underbrace{\Re \left\{ s_n^* \frac{\partial z_\tau(n)}{\partial \tau} \Big|_{\hat{\tau}_N^{(m-1)}} \right\}}_{e(n)}, \quad (2.5)$$

where $\hat{\tau}_n$ is the estimated value of τ_0 at time n . Now the time index and the iteration number coincide. Such an adaptive algorithm can be fully analyzed by using the stochastic approximation tool [30], but this analysis is out of the scope of this chapter.

In DD mode, the symbol s_n which is unknown during data transmission has to be replaced with the (hard) decision of the symbol denoted by \hat{s}_n . DD algorithm can be also applied with soft decision on the symbol as was done in [31]. More details about the DD algorithm with soft decision will be given later, in the context of phase estimation. To perform well, the error probability should be small enough to limit the impact of error propagation. Therefore, DD mode is always implemented after an initialization

step (feasible thanks to the training sequence) in order to ensure low data detection error. In order to avoid the calculation of the derivative at each symbol period (notice that in the blockwise approach the number of iterations and thus of derivative calculations is much smaller than the number of samples), we estimate it with

$$\frac{\partial z_\tau(n)}{\partial \tau} \Big|_{\hat{\tau}_N^{(m-1)}} = \frac{z_{\hat{\tau}_N^{(m-1)} + \Delta}(n) - z_{\hat{\tau}_N^{(m-1)} - \Delta}(n)}{2\Delta},$$

where Δ is a design parameter that must be carefully adjusted. We have just described the very popular early-late (adaptive) estimator [32]. Notice that other update equations with *ad hoc* $e(n)$ have been proposed in the literature [33, 34] and often perform much better.

2.02.2.1.2 DA phase and CFO estimation

In this subsection, we assume that the timing is known and thus can be considered to be zero wlog. Therefore, our received signal model is

$$y_a(t) = \left(\sum_{n=0}^{N-1} s_n h_a(t - nT_s) \right) e^{2i\pi(f_0 t + \phi_0)} + b_a(t).$$

Once again the likelihood can be written as follows:

$$p(y(t)|\phi, f) \propto e^{-\int_{\mathbb{R}} |y_a(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s) e^{2i\pi(f t + \phi)}|^2 dt / 2N_0}.$$

Recall that the sequence $\{s_n\}$ is known. So the joint ML estimator for the phase and CFO takes the following form:

$$[\hat{\phi}_N, \hat{f}_N] = \arg \min_{\phi, f} \underbrace{\int_{\mathbb{R}} \left| y_a(t) - \sum_{n=0}^{N-1} s_n h_a(t - nT_s) e^{2i\pi(f t + \phi)} \right|^2 dt}_{J_N(\phi, f)}.$$

By assuming that the CFO is small compared to the bandwidth and by manipulating the square, we obtain

$$J_N(\phi, f) = -2\Re \left\{ \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi(f T_s n + \phi)} \right\} + \text{const.}, \quad (2.6)$$

where $z(n) = h_a(-t)^* \star y_a(t)|_{t=nT_s}$ and can be written as follows (under the small CFO assumption)

$$z(n) = s_n e^{2i\pi(f_0 T_s n + \phi_0)} + b(n), \quad (2.7)$$

where $b(n)$ is still a white noise process due to the square-root Nyquist property of the shaping filter.

By letting

$$\alpha_N(f) = \frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi f T_s n},$$

we have that

$$J_N(\phi, f) = -2N\Re\{\alpha_N(f)e^{-2i\pi\phi}\} + \text{const.}$$

Then, it is easy to check that the value of ϕ minimizing $J_N(\phi, f)$ for a given f , is

$$\hat{\phi}_N = \frac{1}{2\pi} \angle(\alpha_N(f)) = \frac{1}{2\pi} \arctan \left(\frac{\Im(\alpha_N(f))}{\Re(\alpha_N(f))} \right) \quad (2.8)$$

where \angle stands for the phase of a complex-valued number. Now by inserting Eq. (2.8) into Eq. (2.6), we can easily show that the frequency estimator is obtained by maximizing the modulus of $\alpha_N(f)$. Therefore, the joint DA ML phase and CFO estimates can be expressed as follows:

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi f T_s n} \right|^2 \quad \text{and} \quad \hat{\phi}_N = \frac{1}{2\pi} \arctan \left(\frac{\Im \left(\frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi \hat{f}_N T_s n} \right)}{\Re \left(\frac{1}{N} \sum_{n=0}^{N-1} s_n^* z(n) e^{-2i\pi \hat{f}_N T_s n} \right)} \right). \quad (2.9)$$

While the phase estimate is in closed-form, the CFO estimate still needs a maximization step. Actually the function to be maximized is a periodogram-like function. As in the case of timing, this (periodogram) maximization step may be carried out in two steps: the coarse step is done using an FFT of size N . The resulting frequency estimate enables us to initialize a gradient-descent algorithm around the true point.

Remark.

The estimators have been developed by developing the MLE based on the continuous-time received signal. Another way is as follows. The received signal can be viewed as

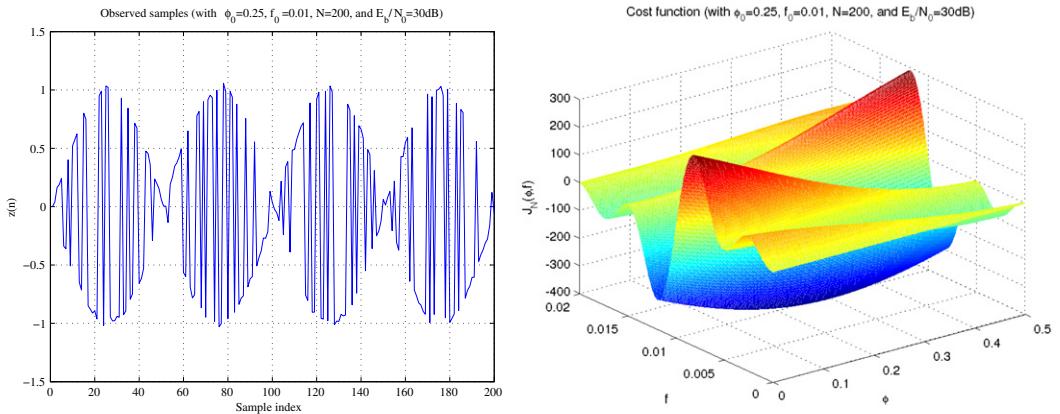
$$y_a(t) = \sum_{n=0}^{N-1} s_n e^{2i\pi(f_0 t + \phi_0)} \Phi_n(t) + b_a(t),$$

where $\Phi_n(t) = h_a(t - nT_s)$. As the shaping filter is a square-root Nyquist filter, the basis functions $\Phi_n(t)$ are orthogonal. In the absence of CFO, the useful part of the received signal is generated by $\Phi_n(t)$; hence, we can project the received signal onto these basis functions without loss of information on the data. Let $u(n) = \langle y_a(t) | \Phi_n(t) \rangle$ with $\langle \cdot | \cdot \rangle$ the inner product. One can check that $u(n) = z(n)$ where $z(n)$ is given in Eq. (2.7). Developing the ML estimator of the phase and CFO based on $z(n)$ will lead fortunately to the equations reported above. We can also prove that the CRBs obtained by both approaches are identical.

As an illustration, we plot in Figure 2.8 (left) a realization of 1000 samples of $z(n)$ when $\phi_0 = 0.1$ and $f_0 T_s = 0.01$ at $E_b/N_0 = 30$ dB with BSPK modulation. In Figure 2.8 (right) the corresponding cost function $J_N(\phi, f)$ has been displayed for $N = 100$. A peak can be observed around the true values of ϕ_0 and f_0 .

Let us now analyze the CRB associated with this estimation issue. In case of estimation of more than one parameter, the CRB is defined through a matrix as follows:

$$\text{CRB}(\phi, f) = F^{-1}(\phi, f),$$

**FIGURE 2.8**

Samples $z(n)$ (left) and corresponding cost function $(\phi, f) \mapsto J_N(\phi, f)$ (right).

where $F(\phi, f)$ is the Fisher information matrix whose components here are

$$F(\phi, f) = - \begin{bmatrix} \frac{\partial^2 \log p(y|\phi, f)}{(\partial\phi)^2} & \frac{\partial^2 \log p(y|\phi, f)}{\partial\phi\partial f} \\ \frac{\partial^2 \log p(y|\phi, f)}{\partial f\partial\phi} & \frac{\partial^2 \log p(y|\phi, f)}{(\partial f)^2} \end{bmatrix}.$$

After straightforward algebraic manipulations, we obtain the CRB associated with the phase, that is defined as the first element of the diagonal of $\text{CRB}(\phi, f)$, as [24]

$$\text{CRB}(\phi) = \frac{N_0 w_2}{4\pi^2 N(w_0 w_2 - w_1^2)},$$

where for any integer k ,

$$w_k = \frac{1}{N^{(k+1)}} \sum_{n=0}^{N-1} n^k |s_n|^2.$$

The CRB associated with the frequency is the second element of the diagonal of $\text{CRB}(\phi, f)$, and is given by

$$\text{CRB}(f) = \frac{N_0 w_0}{T_s^2 4\pi^2 N^3 (w_0 w_2 - w_1^2)}.$$

These expressions can be simplified if $N \gg 1$. We then obtain the so-called asymptotic CRB. One can prove that

$$w_k \xrightarrow{a.s.} \frac{\sigma_s^2}{k+1}.$$

Consequently, we have

$$\text{CRB}(\phi) \approx \frac{1}{\pi^2} \frac{N_0}{E_s} \frac{1}{N} \quad \text{and} \quad \text{CRB}(f) \approx \frac{3}{T_s^2 \pi^2} \frac{N_0}{E_s} \frac{1}{N^3}. \quad (2.10)$$

Notice that, in [25], the (large samples) MSEs of the ML have been calculated and are identical to those given by the asymptotic CRB. Consequently the ML is asymptotically efficient and thus we do not need to spend time to design other estimators since the ML is almost optimal and can be implemented easily in practice.

Thanks to these expressions, some insights about the estimation performance can be given:

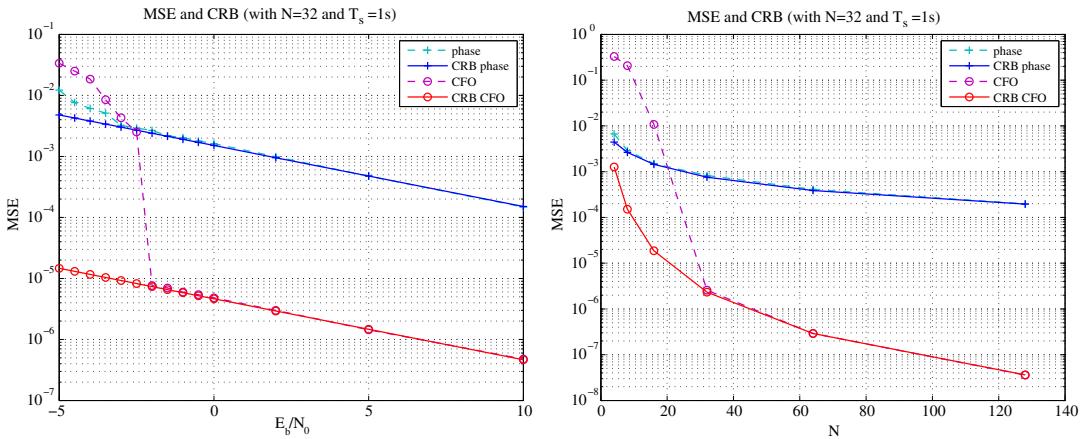
- MSE of the phase estimator is $\mathcal{O}(1/N)$ and $\mathcal{O}(1/\text{SNR})$. The behavior mimics that of the channel estimator. This is reasonable since the phase rotation can be viewed as a one-tap linear filtering operation.
- In contrast, the MSE for the CFO decreases much faster as $\mathcal{O}(1/N^3)$. The convergence speed seems to be very high. We will see that we need to have this high convergence speed for the system to operate properly. Let us consider a frame with a training sequence of length N_T followed by a data sequence of length N_D . For the system to operate well, the phase rotation due to the CFO should be kept as low as possible. At the end of the frame, the phase rotation (after correction) is $2\pi T_s (\hat{f}_N - f)(N_T + N_D)$ which is of order $\mathcal{O}((N_T + N_D)/N_T^{3/2})$. Assume a constant ratio β between N_T and N_D . Thus $\beta = N_T/N_D$ and corresponds to the loss in spectral efficiency caused by the training. Then the phase rotation is proportional to $1/\sqrt{N_D}$ and this tends to zero when the frame is large enough. Notice that if the frequency MSE was $\mathcal{O}(1/N^p)$ with $p \leq 2$, the system cannot perform well due to the unbounded phase rotation associated with CFO.

The influence of the noise is similar since the MSE is of order $\mathcal{O}(1/\text{SNR})$.

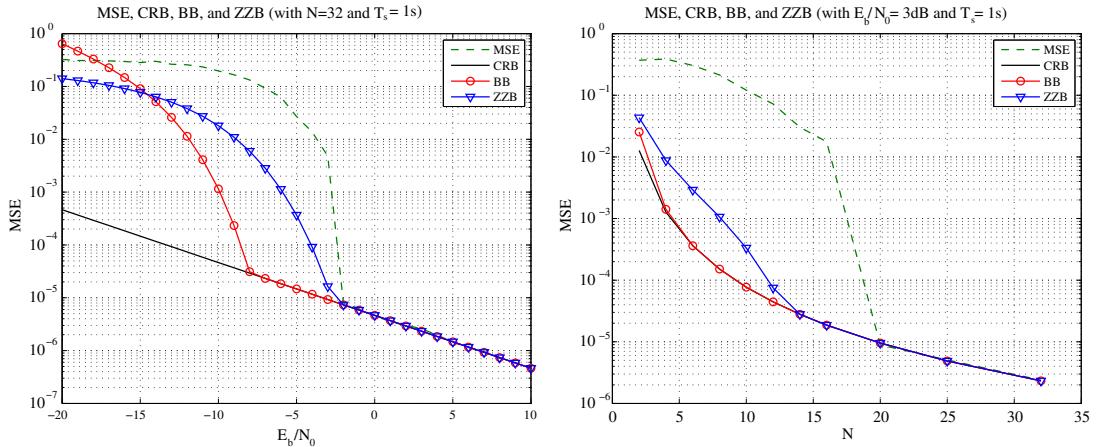
In Figure 2.9, we plot the MSE of the phase and CFO ML estimate and the corresponding CRB: versus E_b/N_0 (left) with $N = 32$, $T_s = 1$ s, and versus N (right) with $E_b/N_0 = 3$ dB and $T_s = 1$ s. The training sequence was BPSK modulated.

We first observe that when the SNR and the number of samples are high enough, ML estimate performance perfectly matches CRB as expected. We also observe a strange phenomenon at low SNR and/or when the number of samples is not large enough: there is a mismatch between the theoretical performance and the empirical one. Moreover the ML is no longer efficient. This is the so-called outliers effect [35–38]. It is associated with the failure of the first step of the periodogram maximization. This phenomenon has been analyzed in the literature, and modified expressions for MSE, taking into account this effect have been derived. The most interesting question is: can we fill up the gap between the ML and the CRB by using another estimator? The answer is no. To answer this, other lower bounds have been developed and analyzed. One can mention Barankin-like bounds [39–43], Bhattacharya-like bounds [44], and Ziv-Zakaï-like bounds [45–48]. A lot of work has been done on deriving such bounds for the harmonic retrieval problem, where it has been shown that the CRB was not tight at low SNR. Other bounds (especially the Ziv-Zakaï one) are actually very close to ML performance, so that it is hard to find better estimates.

In Figure 2.10, we plot CRB, the (first-order) Barankin Bound (BB), the Ziv-Zakaï Bound (ZZB) and the MSE of the ML estimator for CFO parameter versus E_b/N_0 (left) when $N = 32$ and versus

**FIGURE 2.9**

MSE and CRB for phase and CFO versus E_b/N_0 (left) and N (right).

**FIGURE 2.10**

MSE, CRB, BB, and ZZB for CFO versus E_b/N_0 (left) and N (right).

N (right) when $E_b/N_0 = 3$ dB. Notice that the threshold (from which the ML performance can be distinguished from the CRB) can be moved to the left by increasing the SNR (when N is fixed) or by increasing N (when the SNR is fixed) in order to obtain the target performance.

Let us now move on to the adaptive version of the ML. As the CFO can be viewed as phase variations (with a very specific structure), the adaptive algorithm which has the ability of tracking phase variation has been almost always developed under the assumption that only the phase is non-zero and that the

CFO is zero. Therefore, we work with the following discrete-time signal:

$$z(n) = s_n e^{2i\pi\phi} + b(n) \quad (2.11)$$

and the (blockwise) ML for the phase leads to the following cost function (cf. Eq. (2.6))

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \Re \left\{ s_n^* z(n) e^{-2i\pi\phi} \right\}.$$

Following an approach similar to that for timing, $\hat{\phi}_n$, the estimate of the phase at the n th iteration of the (stochastic) gradient algorithm, is updated as follows:

$$\begin{aligned} \hat{\phi}_n &= \hat{\phi}_{n-1} + \mu \frac{\partial \Re \left\{ s_n^* z(n) e^{-2i\pi\phi} \right\}}{\partial \phi} \Big|_{\hat{\phi}_{n-1}} \\ &= \hat{\phi}_{n-1} + \mu \underbrace{\Im \left\{ s_n^* z(n) e^{-2i\pi\hat{\phi}_{n-1}} \right\}}_{e(n)}. \end{aligned} \quad (2.12)$$

Thanks to this update equation, we can introduce the well-known digital Phase-Locked Loop (PLL) scheme to track the phase.

Note that $\hat{\phi}_n - \hat{\phi}_{n-1}$ can be written as $[1 - z^{-1}] \cdot \hat{\phi}_n$ which means that $\hat{\phi}_n$ is obtained after the phase error $e(n)$ is passed through a filter whose Z-transform is $z \mapsto 1/(1 - z^{-1})$. Therefore, the PLL scheme can be described as in Figure 2.11.

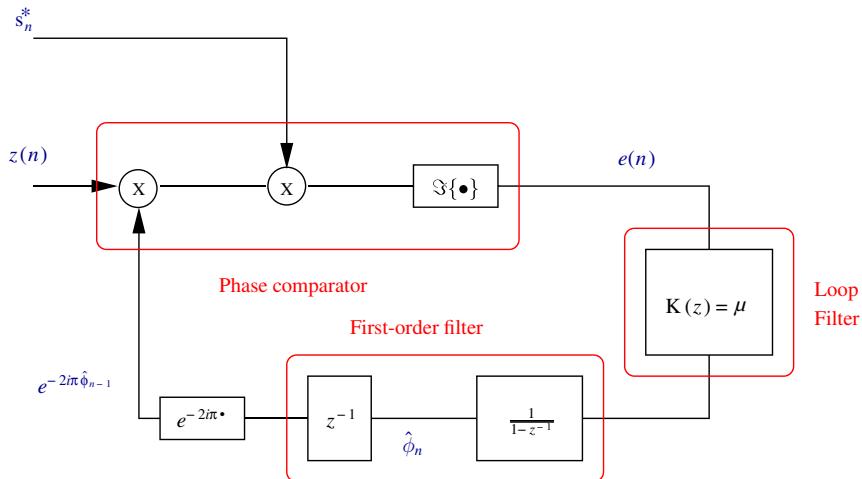
We would now like to inspect the convergence of the PLL first when the phase is fixed, and then when the phase is time-varying. To do this, we consider that ϕ_0 is fixed for the first 1000 samples $z(n)$, and then is time-varying for the next 1000 samples as follows:

$$\phi_{0,n} = \phi_{0,n-1} + w_n, \quad \text{for } n = 1001, \dots, 2000,$$

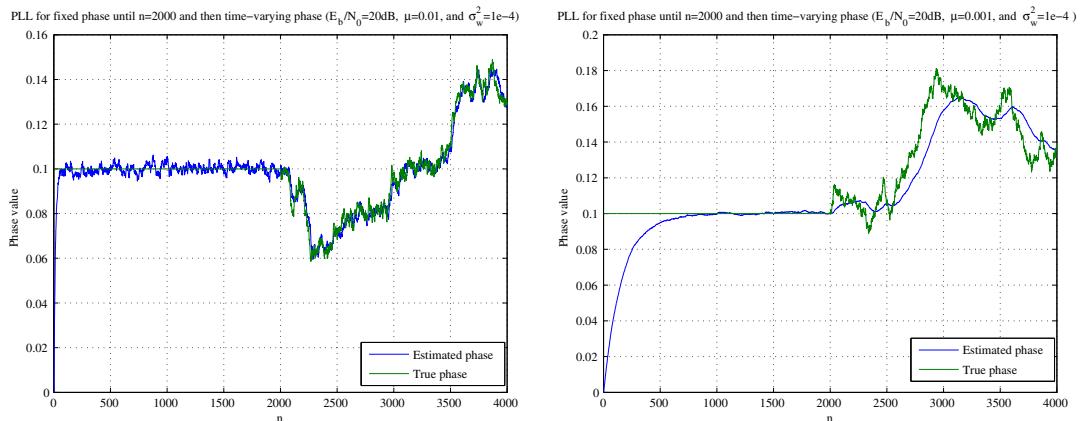
where $\phi_{0,1000} = \phi_0$ and w_n is a real-valued i.i.d. Gaussian process with zero-mean and variance σ_w^2 . In Figure 2.12, we display a realization of the estimated $\hat{\phi}_n$ and the true phase $\phi_{0,n}$ with $\sigma_w^2 = 10^{-4}$ and $E_b/N_0 = 20$ dB.

We remark that proper choice of μ is crucial for good PLL performance. If μ is high, the PLL will rapidly reach an interval around the true value but then will oscillate around the true point without converging. Moreover, a high value of μ enables us to efficiently track the phase variation/noise. In contrast, when μ is small, the convergence speed is low but the phase estimate does not oscillate very much around the true point. But the small value of μ prevents us from following the (too-fast) phase variation/noise.

Indeed, if the phase to be estimated is fixed, it is well known that it is best to consider a time-varying step size μ_n satisfying $\sum_n \mu_n = +\infty$ and $\sum_n \mu_n^2 < +\infty$ [30]. Therefore, an appropriate choice is $\mu_n = \mu/n$ when the phase is fixed. However, such a choice leads to very poor performance if the phase becomes time-varying. A lot of work has been done on designing μ_n to adapt to phase noise characteristics [30].

**FIGURE 2.11**

PLL scheme (in DA context).

**FIGURE 2.12**

Estimated phase and true phase versus the number of iterations/samples for $\mu = 0.01$ (left) and $\mu = 0.001$ (right).

2.02.2.2 NDA case

Here we will develop NDA estimators:

- for timing insensitive to phase and CFO (Problem 2'),
- for phase and frequency insensitive to timing (Problem 2).

As before, we will start with the ML estimators. As in DA mode, the global estimation issue (for timing, phase and frequency) is not tractable. Therefore, we will focus on

- NDA timing ML estimator when phase and CFO are known and thus zero wlog.
- NDA phase and CFO ML estimators when timing is known and thus zero wlog.

2.02.2.2.1 NDA ML timing estimator

Now the symbols s_n are unknown. We just assume that s_n belongs to a particular constellation with P states (P -PAM, P -PSK, P -QAM) and that each constellation point is equally-likely. We denote by $\{s^{(p)}\}_{p=0,\dots,P-1}$ the set of these constellation points. We also denote by $\mathbf{s}_N = [s_0, \dots, s_{N-1}]^T$ the set of transmitted data symbols. The likelihood takes a much more complicated form due to the need to average over all the potential symbols. Therefore, we have

$$p(y|\tau) = \int \cdots \int p(y|\tau, \mathbf{s}_N) p_s(\mathbf{s}_N) d\mathbf{s}_N,$$

where $p_s(\cdot)$ is the probability density function of \mathbf{s}_N . Obviously, we have

$$p(y|\tau, \mathbf{s}_N) \propto e^{-\int_{\mathbb{R}} |y(t) - \sum_{n=0}^{N-1} s_n h_a(t-nT_s-\tau)|^2 dt / 2N_0}$$

and

$$p_s(\mathbf{s}_N) = \prod_{n=0}^{N-1} \left(\frac{1}{P} \sum_{p=0}^{P-1} \delta(s_n - s^{(p)}) \right)$$

as s_n are i.i.d.

After simple but tedious algebraic manipulations (using the fact that the shaping filter is a square-root Nyquist filter), we obtain

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s^{(p)}|^2}{2N_0}} e^{\frac{\Re\{s^{(p)*} z_{\tau}(n)\}}{N_0}} \right).$$

When BPSK is employed, we obtain

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} \log \left(\cosh \left(\frac{\Re\{z_{\tau}(n)\}}{N_0} \right) \right).$$

At low SNR, the Taylor series expansion of $\log(\cosh(\cdot))$ can be used to yield

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} \left(\Re\{z_{\tau}(n)^2\} + |z_{\tau}(n)|^2 \right) \quad (\text{BPSK at low SNR}).$$

When QPSK is employed, similar derivations lead to

$$\hat{\tau}_N = \arg \max_{\tau} \sum_{n=0}^{N-1} |z_{\tau}(n)|^2 \quad (\text{QPSK at low SNR}).$$

We will see that the NDA ML approach for timing is very complicated except for BPSK and QPSK. To handle constellations with more states, Expectation-Maximization (EM) algorithm can be employed and will be explained later in the section devoted to Code-aided synchronization. Notice that in [49] EM algorithm is also implemented but in another way: indeed, the timing error is viewed as the nuisance parameter with an *a-priori* distribution while the data are viewed as the useful parameters to be detected. The authors thus attempt to extend the Maximum Likelihood Sequence Estimator to the case of timing error and are able to correct it through an iterative implementation.

Actually, this timing estimator is very sensitive to phase and CFO. Therefore, we do not continue its analysis in depth. Obviously, an adaptive version may be implemented and the early-late trick can be employed as well.

It is clear that in the NDA case, we need to develop a sub-optimal (non-ML) timing estimator which is insensitive to phase and CFO. We do so next.

2.02.2.2 NDA ML phase and CFO estimator

We now assume that the timing is known and thus zero wlog. Once again, we would like to characterize the NDA joint ML phase and CFO estimator.

By following the same reasoning as for the timing, we obtain that

$$[\hat{\phi}_N, \hat{f}_N] = \arg \max_{\phi, f} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s(p)|^2}{2N_0}} e^{\frac{\Re\{s(p)^* z(n)e^{-2i\pi(fT_s n + \phi)}\}}{N_0}} \right). \quad (2.13)$$

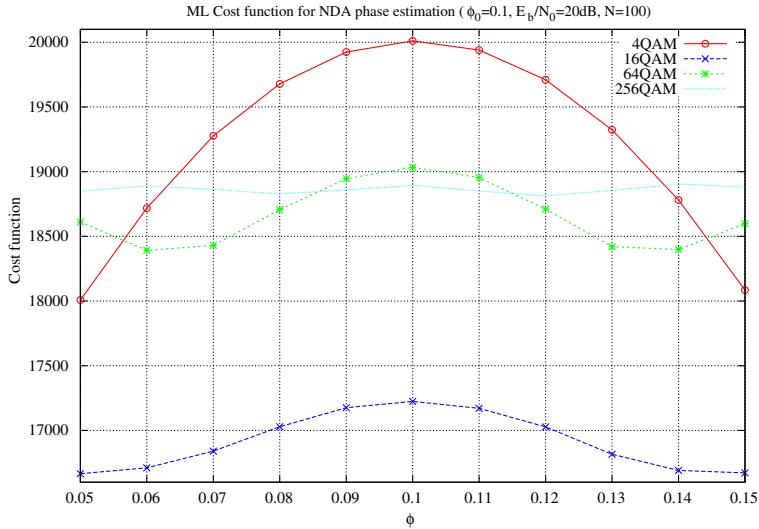
In contrast with the DA case, we are not able to write ϕ with respect to f . As a consequence, the maximization remains a 2-D search which is extremely time consuming. Therefore, we could focus on (i) CFO estimation issue when phase is known and (ii) phase estimation when CFO is known. The first scheme is clearly unrealistic, and thus will be omitted. In contrast, the second scheme is of interest, especially in the tracking regime.

We now assume that CFO is known and thus can be assumed to be zero wlog. Equation (2.13) can be simplified as follows:

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \log \left(\frac{1}{P} \sum_{p=0}^{P-1} e^{-\frac{|s(p)|^2}{2N_0}} e^{\frac{\Re\{s(p)^* z(n)e^{-2i\pi\phi}\}}{N_0}} \right). \quad (2.14)$$

In Figure 2.13, we have plotted the cost function given in Eq. (2.14) for various QAM constellations. Notice that the smaller the constellation, the sharper is the cost function. Therefore, it is easier to estimate phase for small constellation sizes. When a high-order constellation is used, the SNR must be high enough to ensure accurate synchronization (this is not necessarily a drawback since high-order modulation requires high SNR for detection) and the number of samples must be large enough as well (which clearly may become an issue). Figure 2.14 depicts the MSE of NDA ML based phase estimate versus SNR, N and the timing error respectively for BPSK constellation. The MSE decreases proportionally to $1/\text{SNR}$ and $1/N$. Moreover it is insensitive to timing error.

To overcome the implementation issues due to the highly complicated shape of the cost function, EM algorithm can be employed as in [50].

**FIGURE 2.13**

Cost function of NDA ML phase estimate (with $\phi_0 = 0.1$, $E_b/N_0 = 20$ dB, $N = 100$) for various P-QAM.

Once again, as with the NDA ML based timing estimate, the cost function in Eq. (2.14) can be simplified to when BPSK is used

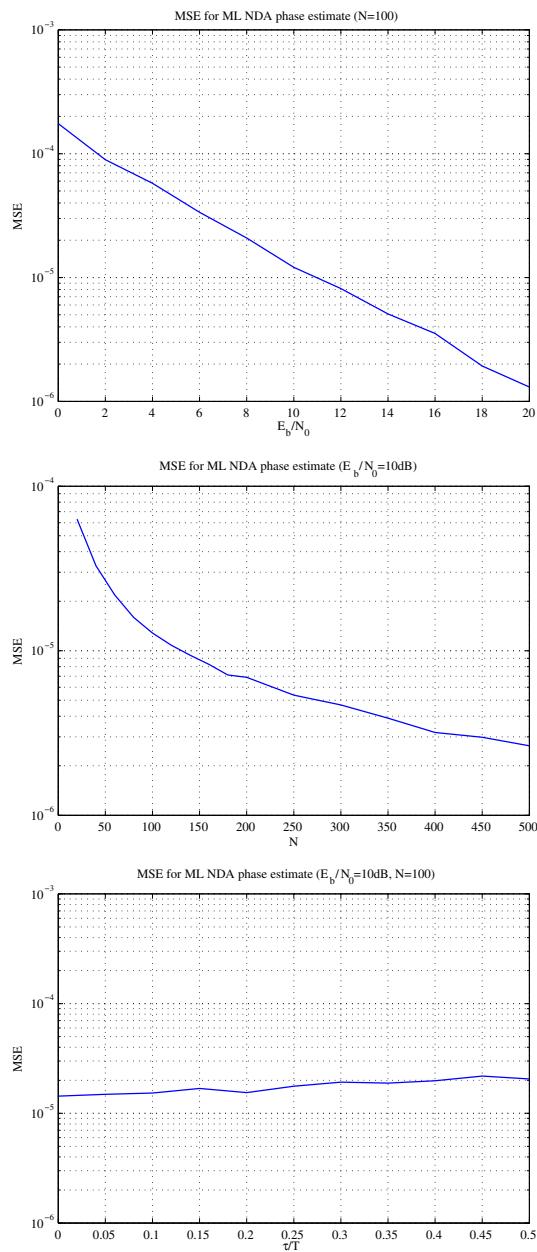
$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \log \left(\cosh \left(\frac{\Re\{z(n)e^{-2i\pi\phi}\}}{N_0} \right) \right) \quad (\text{BPSK}). \quad (2.15)$$

An exhaustive search to solve Eq. (2.15) would be quite time consuming. However, at low SNR, as with the timing case, we obtain that

$$\hat{\phi}_N = \arg \max_{\phi} \sum_{n=0}^{N-1} \Re\{z(n)^2 e^{-2i\pi 2\phi}\} \Leftrightarrow \hat{\phi}_N = \frac{1}{2} \angle \left(\sum_{n=0}^{N-1} z(n)^2 \right) \quad (\text{BPSK at low SNR}).$$

The NDA ML phase estimate is equivalent to the so-called square-power estimate at low SNR which can be easily implemented. Similar results can be obtained for any P-QAM or any P-PSK. Indeed, the NDA ML phase estimate is equivalent to the M th Power estimate with $M = 4$ for any P-QAM and $M = P$ for any P-PSK [51]. Further information about other phase estimators, with PSK modulation, can be found in [52].

Let us consider the BPSK case in more detail. We would like to implement Eq. (2.15) in an adaptive manner. Let $\hat{\phi}_n$ be the value of the estimated phase at time n (i.e., at the n th iteration). The update

**FIGURE 2.14**

MSE of NDA ML phase estimate with BPSK modulation versus E_b/N_0 (top), N (middle), and τ_0 (bottom).

equation is

$$\begin{aligned}\hat{\phi}_n &= \hat{\phi}_{n-1} + \mu \left. \frac{\partial \log \left(\cosh \left(\frac{\Re\{z(n)e^{-2i\pi\phi}\}}{N_0} \right) \right)}{\partial \phi} \right|_{\hat{\phi}_{n-1}} \\ &= \hat{\phi}_{n-1} + \mu \Im \left\{ \check{s}_n z(n) e^{-2i\pi\hat{\phi}_{n-1}} \right\}\end{aligned}$$

with

$$\check{s}_n = \tanh \left(\frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \right). \quad (2.16)$$

Notice that the update equation is very similar to that in the DA case (cf. Eq. (2.12)) except that s_n has been replaced with \check{s}_n . By looking carefully at \check{s}_n , we observe that it corresponds to the so-called soft decision under BPSK constellation assumption. If the soft decision is replaced with a hard decision, then we are in the Decision Directed context, and the estimator \check{s}_n becomes

$$\check{s}_n = \text{sign} \left(\frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \right) \quad (\text{Decision-Directed}).$$

If a decision is made (other than soft or hard), the term \check{s}_n can be written as

$$\check{s}_n = \frac{\Re\{z(n)e^{-2i\pi\hat{\phi}_{n-1}}\}}{N_0} \quad (\text{No data decision})$$

and corresponds to the famous Costas loop originally introduced for continuous-time amplitude-modulated signals [53, 54].

In conclusion, we recall that the joint NDA sync parameters estimation issue is intractable. To simplify it, we first focus on the timing. The NDA ML timing estimator is quite complicated (except for BPSK and QPSK) and very sensitive to phase and CFO. So it is useless in practice. We next consider phase and CFO estimation issues. Clearly the joint problem is still intractable. Therefore, we only focus on the phase estimation issue assuming CFO is known. When the constellation size is small enough, the NDA ML phase is implementable in blockwise or adaptive manner and is not too sensitive to timing error.

Consequently, in order to build practical system, we still need sub-optimal (non-ML) blind methods

- For estimating the timing without the knowledge of the phase and the CFO.
- For estimating the CFO without the knowledge of the timing and the phase.
- For estimating the phase when the constellation size becomes high. The insensitivity to the timing and the CFO is not required since the phase estimator is often carried out after timing and CFO correction.

2.02.2.2.3 Sub-optimal estimators

In this section, we will develop sub-optimal estimators for the sync parameters. To do that, we will analyze carefully some statistical properties of the received signal. Based on these properties, we will

be able to introduce some powerful estimators. The main statistical tools we need are briefly described below:

- *The Cyclostationarity tool:* For sake of simplicity, we will only define second-order cyclostationarity. For more details about any-order cyclostationarity, the reader may refer to [55, 56]. Let us consider $x(t)$ a (complex-valued) continuous-time process and its correlation function $(t, \tau) \mapsto r(t, \tau) = \mathbb{E}[x(t + \tau)x(t)^*]$. When the process is second-order stationary, then the function $(t, \tau) \mapsto r(t, \tau)$ is independent of t . In contrast, if the function $(t, \tau) \mapsto r(t, \tau)$ is periodic with respect to t (and the period is independent of τ), the process is said to be cyclostationary. Let T_c be the period of the correlation function. T_c is called the cyclic period. By considering a Fourier series expansion, we have

$$r(t, \tau) = \sum_k r^{(k)}(\tau) e^{2i\pi k/T_c},$$

where

- $f_k = k/T_c$ is the so-called k th cyclic frequency,
- $\tau \mapsto r^{(k)}(\tau)$ is the cyclic correlation (at cyclic frequency k),
- $S^{(k)}(e^{2i\pi f}) = \sum_{\tau} r^{(k)}(\tau) e^{-2i\pi f\tau}$ is the cyclic spectrum (at cyclic frequency k).

A similar analysis can be done with the pseudo-correlation $(t, \tau) \mapsto u(t, \tau) = \mathbb{E}[x(t + \tau)x(t)]$ if it is not identically zero.

- *The Non-Circularity tool:* let us consider a random (complex-valued) zero-mean scalar variable x . A variable is said to be circularly-symmetric (or simpler circular) if x and $\tilde{x} = e^{2i\pi\theta}x$ have the same distribution regardless of the rotation angle θ . Consequently the distribution is rotationally invariant. By assuming that the moment at any order exists, we have $\mathbb{E}[x^p x^{*q}] = \mathbb{E}[\tilde{x}^p \tilde{x}^{*q}]$ which implies that $\mathbb{E}[x^p x^{*q}] = 0$ as soon as $p \neq q$. Thus, when a variable is circularly-symmetric, only the moment at even orders may be non-zero. Moreover only the moments depending on a power of $|x|$ are non-zero. In digital communications, many variables are circular not for arbitrary orders but only until a certain order. If

$$\mathbb{E}[x^p x^{*q}] = 0$$

for $p \neq q$ and $p + q < M$, the random variable x is said to be circular until the $(M - 1)$ th order or equivalently to be non-circular from the M th order. For more details about the non-circularity, the reader may refer to [13, 51, 57].

The rest of this section is organized as follows: we focus on the timing estimator (based on the cyclostationary tool), then the CFO estimator (based on the cyclostationary or non-circularity tools), and the phase estimator (based on the non-circularity tool).

Timing estimation

We observe that the signal $y_a(t)$ is cyclostationary with period T_s [17]. Consequently the discrete-time (over-sampled) signal $\tilde{y}(n) = y_a(nT_s/Q)$ is also cyclostationary with period Q , i.e., the sequence $\mathbb{E}[\tilde{y}(n + m)\tilde{y}(n)^*] = \mathbb{E}[\tilde{y}(Q + n + m)\tilde{y}(Q + n)^*]$ is periodic for any integer m . Notice that if the continuous-time received signal is sampled at the symbol rate, $Q = 1$ and thus $y(n)$ is stationary. In contrast, if $Q \geq 2$, $\tilde{y}(n)$ is strictly cyclostationary.

First of all, assuming that the CFO is zero, it has been remarked in [58] that $r^{(k)}(m) = \alpha_{k,m} e^{-2i\pi k\tau/T_s}$ for $k \neq 0$, with $\alpha_{k,m}$ a known complex-valued constant depending on the shaping filter. Thanks to this relationship between $r^{(k)}(m)$ and τ , the following ad hoc estimator was proposed in [58]:

$$\hat{\tau}_N = -T_s \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}} \frac{1}{k} \angle \left(\hat{r}^{(k)}(m) \alpha_{k,m}^{-1} \right), \quad (2.17)$$

where \mathbb{K}^* is any set of integers, not including 0, and \mathbb{M} is any set of integers m such that $r^{(k)}(m) \neq 0$. The term $\hat{r}^{(k)}(m)$ is the empirical estimate of $r^{(k)}(m)$ based on N samples, and is given by

$$\hat{r}^{(k)}(m) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{y}(n+m) \tilde{y}(n)^* e^{-2i\pi nk/Q}. \quad (2.18)$$

The performance of such an empirical cyclic correlation estimate can be found in [17, 58–61].

Notice that the timing estimator is insensitive to the phase, and thus can be used before the phase estimation step. In contrast, when the CFO is not zero, it must be estimated first; the timing estimate is then modified as follows:

$$\hat{\tau}_N = -T_s \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}} \frac{1}{k} \angle \left(\hat{r}^{(k)}(m) \alpha_{k,m}^{-1} e^{-2i\pi m \hat{f}_N T_s / Q} \right), \quad (2.19)$$

where \hat{f}_N is the CFO estimate.

When $\mathbb{K} = \{1\}$, $\mathbb{M} = \{0\}$, and the conjugate operation on \tilde{y} is removed, we obtain the heuristic algorithm introduced in [62] which actually is well-suited for BPSK constellation. When $\mathbb{K} = \{1\}$ and \mathbb{M} is any singleton, we obtain the algorithm in [63]. With $\mathbb{K} = \{-1, 1\}$ and $\mathbb{M} = \{0\}$, the estimator reduces to that in [64]. This last algorithm has the great advantage of being insensitive to the CFO since, for $m = 0$, Eq. (2.17) is identical to Eq. (2.19).

All these timing estimators have been theoretically analyzed and closed-form expressions for the Mean-Square Error are given in [65].

The question now is: what is the best blind estimator for timing relying on the cyclic correlation. When $P = 2$, the answer can be found in [65] and corresponds to the weighted covariance matching principle [66, 67]. The gain in performance is only incremental at the expense of higher complexity. Therefore, the NDA timing estimator proposed in Eq. (2.17) (especially that in [64] which is insensitive to CFO) is a strong candidate for our problem.

CFO estimation

The cyclostationarity property can also be used for estimating the CFO blindly. Indeed, in [58], it has been remarked that $r^{(k)}(m) = \alpha_{k,m} e^{2i\pi k\tau/T_s} e^{2i\pi m f T_s / Q}$ with β_k a complex valued scalar which induces the following estimator:

$$\hat{f}_N = \frac{P}{4\pi T_s} \sum_{k \in \mathbb{K}^*} \sum_{m \in \mathbb{M}^*} \frac{1}{m} \angle \left(\hat{r}^{(k)}(m) \hat{r}^{(-k)}(m) \alpha_{k,m}^{-1} \alpha_{-k,m}^{-1} \right), \quad (2.20)$$

where \mathbb{M}^* corresponds to any set of integers that does not contain zero. This estimator is insensitive to timing and phase. Once again, by taking specific elements for the set \mathbb{K} and \mathbb{M} , we obtain the estimators introduced in [63, 64]. The theoretical analysis of the estimator has been done in [59].

Another approach to estimate the CFO independent of the timing and the phase is to use the non-circularity property of the received signal $y_a(t)$. Indeed, it is easy to check [51] that

$$\mathbb{E} \left[s_n^M \right] \neq 0$$

with

- $M = 2$ for P -PAM constellation,
- $M = P$ for P -PSK constellation,
- $M = 4$ for P -QAM constellation.

Consequently, the received signal is also non-circular of M th order. For the sake of simplicity, let us consider that the received signal has been sampled at the symbol rate after passing through the matched filter. Extension to the oversampled case is straightforward. Then

$$y(n) = \left(\sum_{\ell=0}^L g(\ell) s_{n-\ell} \right) e^{2i\pi f T_s n} + b(n),$$

where $g(n) = \tilde{h}_a(nT_s - \tau_0) e^{2i\pi \phi_0}$, $(L+1)$ is the channel length and $b(n)$ the additive white Gaussian noise. Notice that if the timing has been perfectly corrected and the shaping filter is a square-root Nyquist filter, we have $g(n) = \delta_{0,n} e^{2i\pi \phi_0}$ and thus $L = 0$. When the timing is not fully corrected, the filter $g(n)$ is no longer a one-tap filter and ISI occurs. Moreover as τ_0 and ϕ_0 are unknown, we have to assume that $g(n)$ is unknown as well.

We can rewrite $y(n)$ as

$$y(n) = a(n) e^{2i\pi f T_s n} + b(n) \quad \text{where} \quad a(n) = \sum_{\ell=0}^L g(\ell) s_{n-\ell}.$$

Our estimation problem is then equivalent to the estimation of a harmonic f disturbed by additive white Gaussian noise $b(n)$ and multiplicative noise $a(n)$. A lot of work has been done on the problem of harmonic retrieval in multiplicative and additive noise. We will summarize the essential results.

Let us first assume P -PAM constellation. Let $u_a(m) = \mathbb{E}[a(n+m)a(n)]$ be the pseudo-correlation at lag m . As PAM is employed, there exists at least one lag such that $u_a(m)$ is non-zero. Let $p_m(n) = y(n+m)y(n)$. We have

$$p_m(n) = u_a(m) e^{2i\pi(2fT_s)n} + e_m(n),$$

where $e_m(n)$ is a zero-mean process which can be viewed as noise. But this noise is not Gaussian, not white, and not stationary. Nevertheless, by working with $p_m(n)$, we now have to estimate a harmonic

$(2f)$ disturbed only by additive noise. The multiplicative noise has been removed. We recall that if $e(n)$ is Gaussian white and stationary, the ML based on $p_m(n)$ will lead to peak-picking the periodogram. Even if $e(n)$ does not satisfy these standard assumptions, it is still usual practice to estimate the frequency through periodogram peak-selection (even if it no longer has any link with the ML). Then, we have

$$\hat{f}_N = \arg \max_f \sum_m \left| \frac{1}{N} \sum_{n=0}^{N-1} p_m(n) e^{-2i\pi(2f)T_s n} \right|^2.$$

When $m = 0$, we have

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} y(n)^2 e^{-2i\pi(2f)T_s n} \right|^2,$$

which is the well-known (and classical) square-power estimator.

To analyze the theoretical performance of these estimators, the standard work on harmonic retrieval (which assumes the additive noise is Gaussian) cannot be applied [68–70]. However, the analysis has been done quite recently in [71]. Notice that the above approach holds if $a(n)$ is real-valued without assuming any specific structure [61, 72]. However, the previous estimator has limited impact since the PAM constellation is not spectrally efficient.

We now consider P -PSK and P -QAM constellations. Let M be the non-circularity order. Recall that $M = P$ for PSK and $M = 4$ for QAM. Let $q(n) = y(n)^M$. Then

$$q(n) = \mathbb{E}[q(n)]e^{2i\pi(Mf)n} + e'(n),$$

where $e'(n)$ is a zero-mean process that can be interpreted as additive noise. Once again one can carry out periodogram peak-picking for retrieving f . Thus, we have

$$\hat{f}_N = \arg \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} y(n)^M e^{2i\pi(Mf)T_s n} \right|^2. \quad (2.21)$$

This estimator has been introduced by [73] for PSK. Several extensions of these estimators may be introduced. For example, instead of relying only on $y(n)^M$, one could also work with $y(n)y(n+m_1)\cdots y(n+m_{M-1})$. This has not been done in the literature. One could also find the best non-linear transformation \mathcal{F} such that peak-picking the periodogram based on $\mathcal{F}(y(n))$ has some desired properties. One can find some results about optimization of \mathcal{F} in [74].

In Figures 2.15 and 2.16, we plot the theoretical and empirical performance of the M th power estimate for various QAM constellations, varying the SNR and the number of samples N , respectively.

For QAM, we observe a self-noise phenomenon since the performance has an error floor with respect to SNR. It is due to the fact that $s_n^4 \neq \mathbb{E}[s_n^4]$. A recent approach can be used to remove the self-noise (see [75]).

The outliers effect still occurs as in any approach based on peak-picking periodogram. The theoretical MSE taking into account the outliers has been analyzed in [38] for modulated signals and [35] for unmodulated signals.

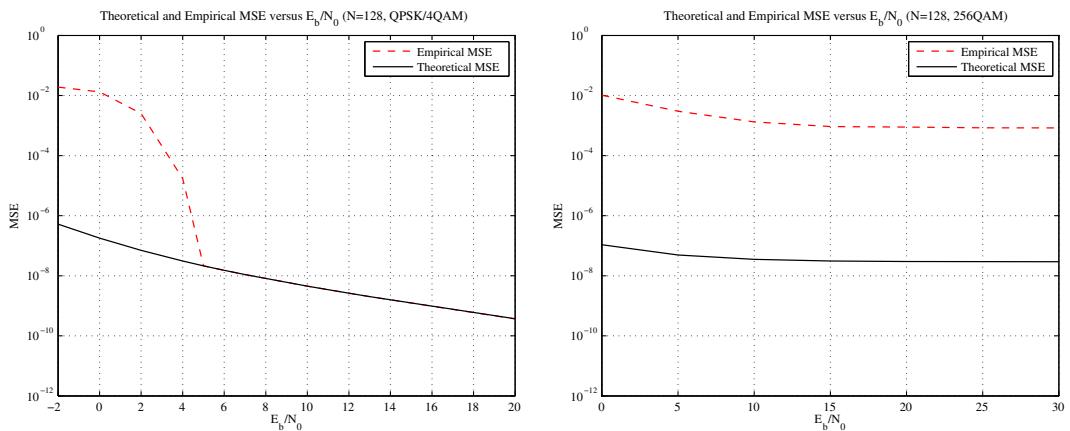


FIGURE 2.15

Theoretical and Empirical MSE of M th power CFO estimate versus E_b/N_0 for 4-QAM (left) and 256-QAM (right) with $N = 128$.

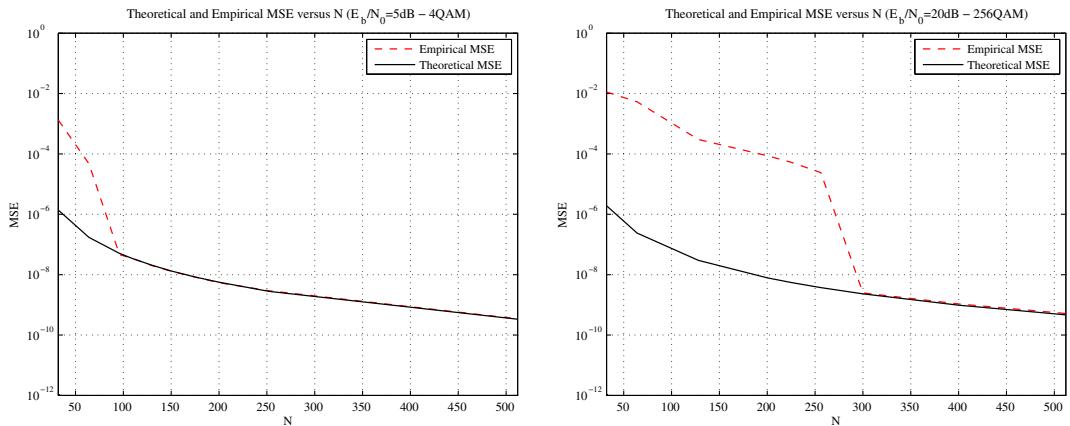


FIGURE 2.16

Theoretical and Empirical MSE of M th power CFO estimate versus N for 4-QAM at $E_b/N_0 = 5$ dB (left) and 256-QAM at $E_b/N_0 = 20$ dB (right).

Phase estimation

For estimating the phase, we will assume that the timing and CFO have already been correctly compensated. Then we focus on the sampled (at the symbol rate) output of the matched filter. Therefore, we have

$$z(n) = s_n e^{2i\pi\phi_0} + b(n).$$

The approach for estimating the phase will be similar to that for the estimation of CFO by using the non-circularity property of the constellation. Since

$$q(n) = y(n)^M = \mathbb{E}[s_n^M]e^{2i\pi(M\phi_0)} + e(n),$$

where $e(n)$ is a zero-mean process and $\mathbb{E}[s_n^M]$ is assumed to be known with phase ϕ_* . Then it is easy to build a simple estimator as follows:

$$\hat{\phi}_N = \frac{1}{M} \angle \left(\frac{1}{N} \sum_{n=0}^{N-1} y(n)^M \right) - \phi_*. \quad (2.22)$$

Obviously we have an ambiguity of $2\pi/M$ since the constellation is invariant to rotations of angle $2\pi k/M$, for any integer k . We recall that this estimate is close to the ML at low SNR.

This approach was introduced for PSK by [73] and for QAM by [51]. Notice that other less powerful estimates of the phase have been introduced in the literature. A deep theoretical analysis of the estimators can be found in [76].

2.02.2.2.4 CRB

In the DA case, we have observed that the derivations of CRB was not a difficult task. Moreover, we were able to obtain rather nice closed-form expressions that enable us to provide some interesting insights. In the NDA case, in contrast, the CRB expressions in closed-form are much more difficult to be achieved due to the presence of the unknown data sequence. We thus work in the framework of the CRB with nuisance parameters. Obviously the nuisance parameters here are the data. Why is it difficult? To calculate the CRB, we need at least to have a closed-form expression for the likelihood. As noted earlier in the discussion of blind ML estimators, a closed-form expression of the likelihood is very difficult to obtain.

The main idea is to define a variant of the CRB assuming a specific property for the data sequence (the nuisance parameters). In order to be useful, these CRB variations should be easier to derive and should provide some insights. Typically, these CRB variants will not be as tight as the classical CRB, and hence they will be optimistic.

A lot of work has been done on the derivation of the CRB for synchronization parameters and/or channel estimators [77–87]. We will summarize the main principle in deriving CRB with nuisance parameters. Applications to phase and CFO estimation are provided with more details.

Let us start with the definition of several CRBs:

- *True-TCRB* (also called Unconditional-UCRB or Stochastic-SCRB): The data sequence is viewed as a random process with a known distribution (a product of sum of Dirac distributions). Let θ be the synchronization parameter vector [83–85, 87]. The true Fisher Information Matrix is defined as

$$F_t(\theta) = -\mathbb{E} \left[\frac{\partial^2 \log(p(y|\theta))}{\partial \theta (\partial \theta)^T} \right]$$

with the true likelihood

$$p(y, \theta) = \mathbb{E}_{\mathbf{s}} [p(y|\theta, \mathbf{s})] = \int \cdots \int p(y, |\theta, \mathbf{s}) p(\mathbf{s}) d\mathbf{s}.$$

Then the true CRB is obtained as

$$\text{TCRB}(\theta) = F_t(\theta)^{-1}.$$

The main drawback is that it is usually intractable to express this in closed-form. Some expressions are given in [84] but it is not really in closed-form due to the huge number of sums and products.

- *Conditional-CCRB* (also called Deterministic-DCRB): The data sequence is assumed to be of interest and is added to the parameters to be estimated. Notice that the structure of the data sequence (i.e., data belong to a specific discrete constellation) is not taken into account. If we would like to take this structure into account, we should calculate the CRB when strong constraints on the parameters have to be satisfied. Such a problem is very hard [88, 89] and thus we are far away from our objective of simplifying the CRB derivations. The difficulty can be partially overcome by considering *Bayesian* CRBs which use *a-priori* distributions on the data [27]. Let us come back to the CCRB [77]. Let F_c be the Fisher information matrix when \hat{s}_N is given. We have

$$F_c(\theta) = -\mathbb{E} \left[\frac{\partial^2 \log(p(y|\theta, \hat{s}_N))}{\partial \theta (\partial \theta)^T} \right],$$

where \hat{s}_N is obtained by maximizing over s (by omitting all the constraints on s) the conditioned likelihood $s \mapsto p(y|\theta, s)$. Then the conditioned CRB is obtained as

$$\text{CCRB}(\theta) = F_c(\theta)^{-1}.$$

Usually the asymptotic version is used i.e., the number of samples is assumed to be large.

- *Gaussian-GCRB*: The true CRB is very difficult to compute due to the distribution of s . To overcome this problem, one can assume that s is Gaussian distributed (even though this is not true). The GCRB is thus obtained by using the same definition of the true CRB but by assuming a Gaussian data sequence. Obviously this GCRB is not generally a bound since it implicitly assumes that the higher-order cumulants are zero. For example, if the data is PSK and QAM, we have seen that the higher-order moments are very important through the M th power estimate and thus the GCRB does not capture this information. Nevertheless the GCRB has two advantages: it is quite easy to derive in closed-form, and it is a bound on the MSE for all the estimates based on second-order statistics (even if the data sequence is not Gaussian) [79, 80, 86].
- *Modified-MCRB*: Data are assumed to be known in order to calculate the Fisher Information Matrix. In order to have a Fisher Information Matrix (FIM) independent of data, the data-dependent FIM is averaged over the data. Therefore, we have

$$F_m(\theta) = -\mathbb{E}_s \left[\mathbb{E}_y \left[\frac{\partial^2 \log(p(y|\theta, s))}{\partial \theta (\partial \theta)^T} \right] \right].$$

Then the MCRB [78, 81, 82] is defined as

$$\text{MCRB}(\theta) = F_m(\theta)^{-1}.$$

In terms of derivations, the main difference from the TCRB is that the average over the data is outside the log (for MCRB) and inside the log for the TCRB. This quite small difference leads to tractable

MCRB and usually intractable TCRB. Indeed, the log will be removed by the exp in $p(y)$ due to the Gaussian additive noise. In MCRB, log will directly remove the exp although it is not the case for TCRB, thus leading to very hard calculations.

In the sequel, we provide some general results on the links between the various CRBs. This relationship is independent of the estimation problem:

- At low SNR, the TCRB is much more tractable by doing a Taylor series expansion of e^x for small x [83, 85, 87].
- At any SNR, the MCRB is a bound but sometimes too optimistic and thus not tight enough.
- At any SNR, we have $\text{TCRB} \geq \text{MCRB}$ and $\text{CCRB} \geq \text{MCRB}$ [77].
- At high SNR, if the data sequence belongs to a discrete set, we have $\text{TCRB}/\text{MCRB} \rightarrow 1$. So in our context where data belong to a finite constellation, this property holds and thus the MCRB is very useful as soon as the SNR is large enough [83].

Let us now focus on our specific estimation problem: for timing estimation (when phase and CFO are perfectly corrected), there already exists a very good tutorial [20, 77]. Therefore, we focus on phase and CFO estimation (when timing is known and thus zero wlog). Then we work with the (symbol-rate) sampled output of the matched filter. Thus our signal model is

$$y(n) = s_n e^{2i\pi(f_0 T_s n + \phi_0)} + b(n).$$

Let us focus on a “toy” example to highlight some interesting results. We first assume that s_n can be decomposed as follows

$$s_n = \sigma_R s_n^{(R)} + i\sigma_I s_n^{(I)},$$

where $s_n^{(R)}$ and $s_n^{(I)}$ are two unit-variance white BPSK sequences independent of each other; $\sigma_R^2 = (1+u)/2$ and $\sigma_I^2 = (1-u)/2$ with $u \in [0, 1]$ such that the correlation $\mathbb{E}[|s_n|^2] = 1$ and $\mathbb{E}[s_n^2] = u$. Notice that if $u = 1$, then s_n becomes a standard BPSK. And if $u = 0$, s_n is no longer a (second-order) non-circular process. Even though both phase and CFO have to be estimated, we only provide the expressions for CRB associated with the CFO estimate.

- *TCRB derivations:* As the nuisance parameter is discrete, we know that at high SNR the TCRB is equivalent to the MCRB; the latter can be calculated very easily by averaging the FIM given in Eq. (2.10). Therefore, we have

$$\text{TCRB}(f)|_{\text{high SNR}} = \frac{3N_0}{\pi^2 T_s^2 N^3}.$$

At low SNR [83], a derivation based on the Taylor series expansion of $x \mapsto e^x$ leads to

$$\text{TCRB}(f)|_{\text{low SNR}} = \frac{3N_0^2}{\pi^2 u^2 T_s^2 N^3}.$$

We observe that the non-circularity power (induced by u) has a great impact at low SNR where the performance is proportional to SNR^2/u^2 .

- *CCRB derivations:* As we deal with the (symbol rate) sampled signal $y(n)$, the CCRB corresponds to an under-determined estimation problem: we have N samples to estimate $N + 2$ parameters, and no additional constraints. As a consequence, the CCRB is not finite. To overcome the problem, we need to work with the oversampled version of the received signal. This is out of the scope of our toy example. For more details, please refer to [77, 90].
- *GCRB derivations:* The expressions can be found in [86] and are reported below

$$\text{GCRB}(f) = \frac{3(1 - u^2 + 4N_0 + 4N_0^2)}{4\pi^2 u^2 T_s^2 N^3} \quad (\text{for } N \text{ large enough}).$$

Once again the non-circularity power (representing by u) has a great impact which verifies that the non-circularity property is an important tool in blind estimation of frequency and phase. We can even show that when $u \neq 0$ (and especially for $u = 1$, i.e., BPSK), the square-power estimate MSE is identical to GCRB. Therefore, the square-power estimator for non-circular white multiplicative noise is the best second order estimator.

If $u = 0$, s_n is not (second-order) non-circular anymore. The second order statistics reduce to the correlation $r_y(m) = r_s(m)e^{2i\pi f_0 T_s m} + 2N_0\delta(m)$. If s_n is white, we do not have information about f and GCRB will go to infinity. In contrast, if s_n is colored, we can easily build an estimate as follows: $\hat{f}_N = (\angle(\hat{r}_y(m)) - \angle(r_s(m)))/m$ where $\hat{r}_y(m)$ is the empirical estimate of $r_y(m)$ and $r_s(m)$ is known. The frequency is thus viewed as the phase of the correlation function. GCRB is given by [80]

$$\text{GCRB}(f)_{|\text{circular case}} = \frac{1}{N \int_0^1 \left(\frac{S'_s(e^{2i\pi f})}{S_s(e^{2i\pi f}) + 2N_0} \right)^2 df} \quad (\text{for } N \text{ large enough}),$$

where $S_s(e^{2i\pi f})$ is the power spectrum of s_n and $S'_s(e^{2i\pi f})$ its derivative function. We note that GCRB decreases as $1/N$ which is the convergence speed associated with the phase estimation issue. So, clearly, to have faster convergence, we need to use high-order statistics of s_n when it is second-order circular.

- *MCRB derivations:* We just have to average the FIM and it is the term given in Eq. (2.10).

$$\text{MCRB}(f) = \frac{3N_0}{\pi^2 T_s^3 N^3}.$$

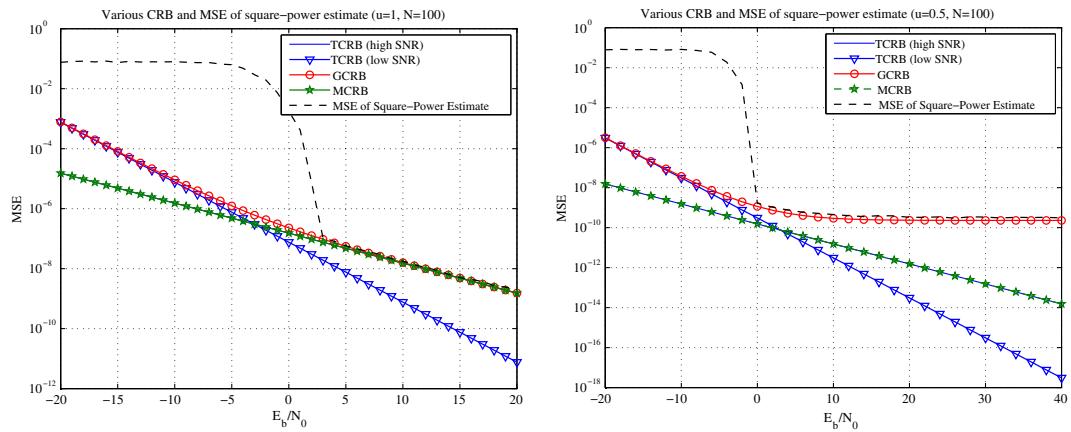
The MCRB does not capture the influence of non-circularity; it is not tight enough and is too optimistic except at high SNR.

Let us now assume that s_n is an i.i.d. sequence drawn from a PSK or QAM constellation. GCRB and MCRB are not modified at all although the process is no longer (second-order) non-circular. In contrast, TCRB (at low SNR) is completely different and is given [83] by

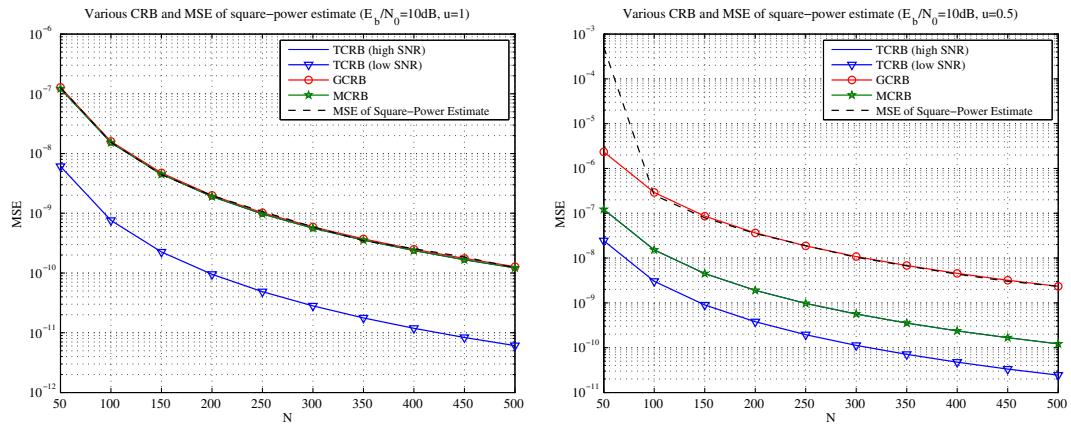
$$\text{TCRB}_{|\text{low SNR, } M\text{th order non.circ}} = \mathcal{O}(1/\text{SNR}^M).$$

Once again, the non-circularity tool is fundamental for estimating the CFO and the phase. TCRB has similar behavior as the MSE of the M th power (if outlier effects are not taken into account) at low SNR.

In Figures 2.17–2.19, we plot various CRBs and empirical MSE of the M th power estimate versus SNR, N and u , respectively, for the process s_n used in the toy example.

**FIGURE 2.17**

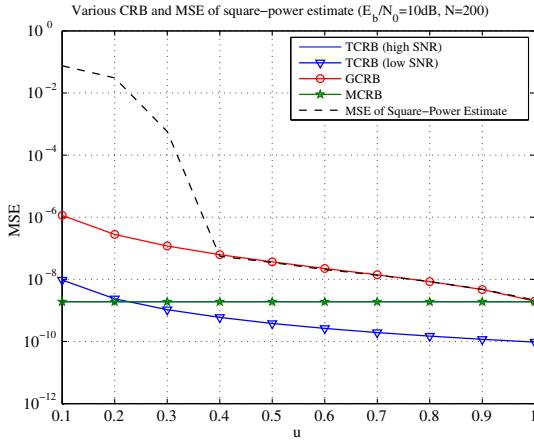
Various CRBs and MSE versus E_b/N_0 for $u = 1$ (left) and for $u = 0.5$ (right).

**FIGURE 2.18**

Various CRBs and MSE versus N for $u = 1$ (left) and for $u = 0.5$ (right).

2.02.2.2.5 Code-Aided synchronization

When the SNR is low, the NDA estimators may offer poor performance with realistic number of samples. Until now, we have not exploited the usual structure of the data. Indeed, in order to obtain the targeted BER in current systems, channel coding is used. There are two ways of using the channel structure to improve the synchronization step:

**FIGURE 2.19**

Various CRBs and MSE versus u ($E_B/N_0 = 10$ dB and $N = 200$).

- *DD with hard/soft decision:* The hard decision has already been introduced in this tutorial when hard decision was done after the channel decoding. Hard decision can be replaced with soft decision. As soft decision is usually needed for iterative decoding, we can implement jointly the decoding of turbo or LDPC and the synchronization which leads to the so-called turbo-synchronization concept. More details will be given below.
- *LLR maximization:* The performance of the system will be better if the sync parameters are well chosen. Therefore, sync estimators relying on cost function depending on reliable functions can be developed [91].

Here, we will focus on DD with soft decision. The use of soft decision is really interesting at low SNR when synchronization is very difficult. Usually, at low SNR, the channel coding requires the use of turbo-codes or LDPC, namely, of iterative coding. Therefore the next synchronization way is usually called “turbo-synchronization” developed by Vandendorpe-Luisse, and others. A nice tutorial treatment is given in [31]. Here, we briefly provide the main points. For simplicity of exposition, we focus only on phase estimation (assuming the other synchronization parameters are known).

Let $\hat{\phi}^{(n)}$ be the phase estimator at the n th iteration. The EM algorithm has two basic steps:

$$\text{E-Step : } Q(\phi, \hat{\phi}^{(n)}) = \int p(\tilde{y}|y, \hat{\phi}^{(n)}) \log(p(\tilde{y}|\phi)) d\tilde{y},$$

$$\text{M-Step : } \hat{\phi}^{(n+1)} = \arg \max_{\phi} Q(\phi, \hat{\phi}^{(n)}),$$

where

- \tilde{y} is the *complete* set of data.
- y is the *incomplete* set of data.

By considering the complete set of data as the received signal and the transmitted symbols, i.e., $\tilde{y} = [y, \mathbf{s}_N]$, one can obtain that

$$Q(\phi, \hat{\phi}^{(n)}) = \Re \left\{ \sum_{n=0}^{N-1} \check{s}_{n, \hat{\phi}^{(n)}} z(n) e^{-2i\pi\phi} \right\}, \quad (2.23)$$

where

$$\check{s}_{n, \hat{\phi}^{(n)}} = \sum_{p=0}^{P-1} s^{(p)} p(s_n = s^{(p)} | y, \hat{\phi}^{(n)}). \quad (2.24)$$

Thus, the M-step leads to the following solution:

$$\hat{\phi}^{(n+1)} = \angle \left(\sum_{n=0}^{N-1} \check{s}_{n, \hat{\phi}^{(n)}} z(n) \right).$$

We thus remark that the phase estimator is very similar to DA or DD estimator but now the symbol is neither known, nor decided, but is replaced with the mean of the a posteriori distribution.

The previous EM approach can be considered with or without coding. The performance will just be different because the a posteriori mean will be more or less accurate compared to the true transmitted symbol (see [49] for non-coded case). In order to connect some previous results, let us consider the non-coded BPSK case. We then have

$$\check{s}_{n, \hat{\phi}^{(n)}} = p(s_n = 1 | y, \hat{\phi}^{(n)}) - p(s_n = -1 | y, \hat{\phi}^{(n)}), \quad (2.25)$$

$$= \tanh \left(\frac{1}{2} \log \left(\frac{p(s_n = 1 | y, \hat{\phi}^{(n)})}{p(s_n = -1 | y, \hat{\phi}^{(n)})} \right) \right), \quad (2.26)$$

$$= \tanh \left(\frac{\Re \left\{ z(n) e^{-2i\pi\hat{\phi}^{(n)}} \right\}}{N_0} \right). \quad (2.27)$$

Equation (2.26) corresponds to the LLR of the symbol, and Eq. (2.27) corresponds to the standard soft decision on the BPSK symbol. We remark that the EM approach (which is strongly connected to ML) leads “fortunately” to the equation of the true ML developed for the non-coded BPSK in Eq. (2.16) even though the iteration does not mean the same thing. In the EM approach, we iterate inside a data block, whereas in Eq. (2.16) we iterate at each ongoing sample. The non-coded BPSK based example shows that there exists a strong link between EM and LLR. Nevertheless, this link cannot be extended directly to coded system (whatever the constellation). For more details, we refer the reader to [31].

In terms of complexity, we need to compute Eq. (2.24). For (coded) BPSK, the BCJR algorithm can be used. For other constellations, usually approximations are done [92, 93]. Notice that we need the soft decision directed approach only at low SNR when the constellation size is in practice small enough which reduces the computational load for obtaining Eq. (2.24).

Table 2.1 Some Algorithms Associated with Each Problem Statement

Problem	Algorithms
1 (DA phase and CFO estimation with perfect timing)	(2.9)
1' (DA Timing estimation with perfect carrier recovery)	(2.3)
2 (NDA Phase and CFO estimation insensitive to timing)	(2.20, 2.21, 2.22)
2' (NDA Timing estimation insensitive to phase and CFO)	(2.19) if $\mathbb{M} = \{0\}$

2.02.2.6 Summary

In Table 2.1, we summarize the links between the various estimators that we have discussed and the assumptions under which they work. We also indicate whether each of the problems (1,1',2,2') has a solution.

2.02.3 Synchronization for non-flat fading channels

When the channel is frequency-selective, the orthogonal frequency-division multiplexing (OFDM) modulation scheme is the standard of choice. OFDM has been widely employed in various commercial applications that include wireless local area networks (IEEE 802.11a/g/n and HIPERLAN/2), wireless metropolitan area networks (WMAN/WiMax, IEEE 802.16), terrestrial digital audio broadcasting (DAB) and terrestrial digital video broadcasting (DVB) systems in Europe, Multimedia Mobile Access Communications (MMAC) in Japan. The popularity of OFDM stems from its ability to transform a wideband frequency-selective channel to a set of parallel flat-fading narrowband channels, which substantially simplifies the channel equalization problem. Because of the time-frequency granularity that it offers, OFDM appears to be a natural solution when the available spectrum is not contiguous, for overlay systems, and to cope with issues such as narrowband jamming. In the multiuser context, this granularity also accommodates variable quality-of-service (QoS) requirements and bursty data. A noticeable example of this multiuser application is the combination of OFDM with frequency-division multiple access (FDMA) protocol, i.e., orthogonal frequency-division multiple access (OFDMA), which has become part of the IEEE 802.16 standards for WMAN. Since subcarriers are allocated to distinct users in a non-overlapping manner, one attractive feature of OFDMA is its capability to mitigate the effects of multiple-access interference (MAI). Another appealing feature of OFDMA is dynamic subcarrier assignment which enables it to optimally allocate system resources such as transmission power and spectrum.

Despite the above-mentioned appealing features, the synchronization task turns out to be a critical issue for OFDM based systems. The synchronization problems of OFDM based systems include timing and frequency synchronization. Timing and frequency offsets come from two sources. One source is the local oscillator frequency mismatch between the transmitter and the receiver, and the other is the Doppler spread due to the relative motion between the transmitter and the receiver. Both timing and frequency synchronization errors introduce extra interference to OFDM systems and result in performance degradation. In addition, timing synchronization may affect the performance of channel

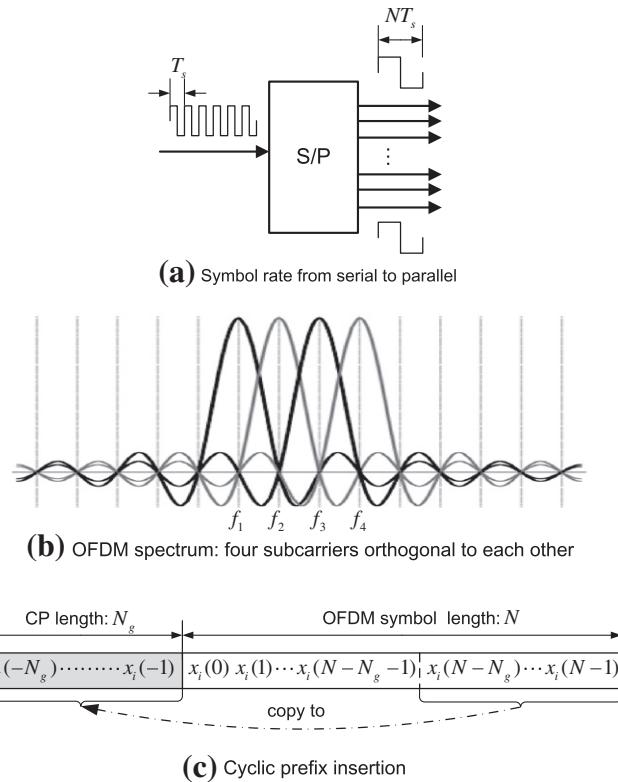
estimation [94]. Therefore, effective synchronization is a key to improve the performance of an OFDM based system [95].

2.02.3.1 Signal model and preliminaries

The operational principle of an OFDM system is that the available bandwidth is divided into a large number of subchannels, over each of which the wireless channel can be considered non-dispersive or flat fading. The original data stream at rate R is split into N parallel data streams, each at rate R/N . The symbol duration, T_s , for these parallel data streams is therefore increased by a factor of N , i.e., $T = NT_s$ as shown in Figure 2.20a. Conceptually, each of the data streams modulates a carrier with a different frequency and the resulting signals are transmitted simultaneously. The carriers for each subchannels are made orthogonal to one another, allowing them to be spaced very close together with no overhead. This is shown in Figure 2.20b for four carriers. Correspondingly, the receiver consists of N parallel receiver paths. Due to the increased symbol duration, the intersymbol interference (ISI) over each channel is reduced to $\lceil \tau_{\max}/(NT_s) \rceil$ symbols. Thus, an advantage of OFDM is that, for frequency-selective fading channels, the OFDM symbols are less affected by channel fades than are single-carrier transmitted symbols. This is due to the increased symbol duration in an OFDM system. While many symbols during a channel fade might be lost in a single-carrier system, the symbols of an OFDM system can still be correctly detected since only a fraction of each symbol might be affected by the fade. On the other hand, if the channel is time selective, i.e., the channel impulse response varies significantly within the OFDM symbol period, then the channel matrix is no longer Toeplitz and conventional OFDM would fail.

Since multicarrier modulation is based on a block transmission scheme, measures have to be taken to avoid or compensate for interblock interference (IBI), which contributes to the overall ISI. OFDM systems can be categorized by the way they handle IBI. In the most popular systems, a guard time is inserted between consecutive OFDM symbols in the form of a cyclic prefix (CP); i.e., the tail of the OFDM symbol is prefixed as shown in Figure 2.20c. The length of the CP is chosen to be larger than the expected delay spread; after proper time synchronization, the receiver discards the CP and thus the IBI is eliminated. Time guarding by zero padding the OFDM symbols has also been proposed in [96, 97]. The issue here is one of turning the transmitter on and off and increased receiver complexity versus the increased signal-to-noise ratio (SNR) and decreased symbol error rate (SER). Comparisons between cyclic-prefixing and zero-padding OFDM systems may be found in [98]. In this chapter, we focus on CP based OFDM systems only.

The choice of the OFDM parameters is a trade-off between various, often conflicting requirements. The length of the CP is dictated by the delay spread of the channel. Introduction of the CP entails a reduction in rate (or wasted bandwidth), as well as an SNR loss; to minimize these inefficiencies, the number of subcarriers, N , should be large. However, a large number of subcarriers induces high implementation complexity, increased sensitivity to frequency offset and phase noise (since the subcarriers get closer to each other as N increases), and increased peak-to-average power ratios (PAPRs). N is dictated by concerns regarding practical FFT sizes as well as the coherence time of the channel. We will not address the issue of practical choice of OFDM parameters here; we refer the reader to [99] and references therein. In this chapter, we address the crucial issue of timing and frequency offsets estimation.

**FIGURE 2.20**

The operational principle of an OFDM system.

2.02.3.1.1 Subcarrier allocation strategies

Conventionally, all subcarriers are allocated to one specific user in single user OFDM systems. For multiuser OFDM or OFDMA systems, subcarrier allocation strategies are needed. Let N and K_u denote the total number of subcarriers and maximum number of active users, respectively. The current number of users M is limited to K_u , i.e., $M \leq K_u$. Let I_i and \mathcal{I}_i indicate the number and indices of subcarriers allocated to the i th user, respectively; we have that

$$\sum_{i=1}^M I_i \leq N, \quad (2.28)$$

$$\mathcal{I}_i \cap \mathcal{I}_j = \emptyset, \quad i \neq j. \quad (2.29)$$

As shown in Figure 2.21, generally, there are three subcarrier allocation schemes. For illustration purposes, we set $N = 32$, $M = 3$, $K_u = 4$ and $I_i = 8$ in Figure 2.21. The *subband carrier allocation*

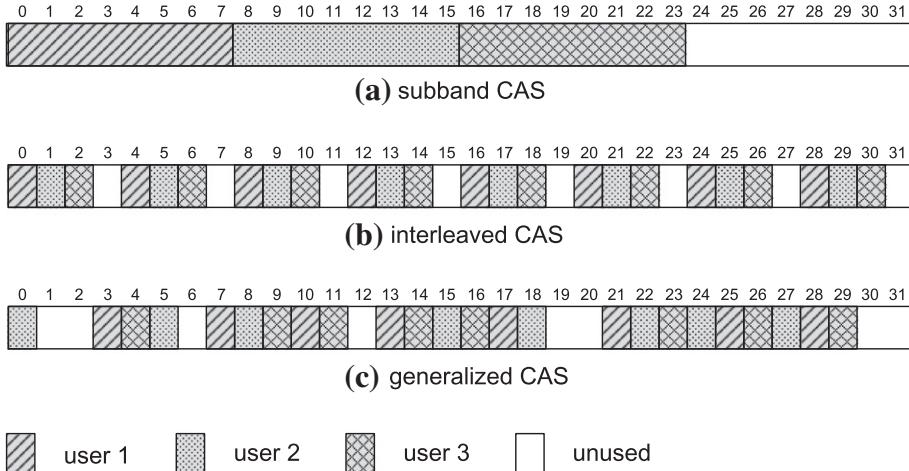
**FIGURE 2.21**

Illustration of subcarrier allocation schemes.

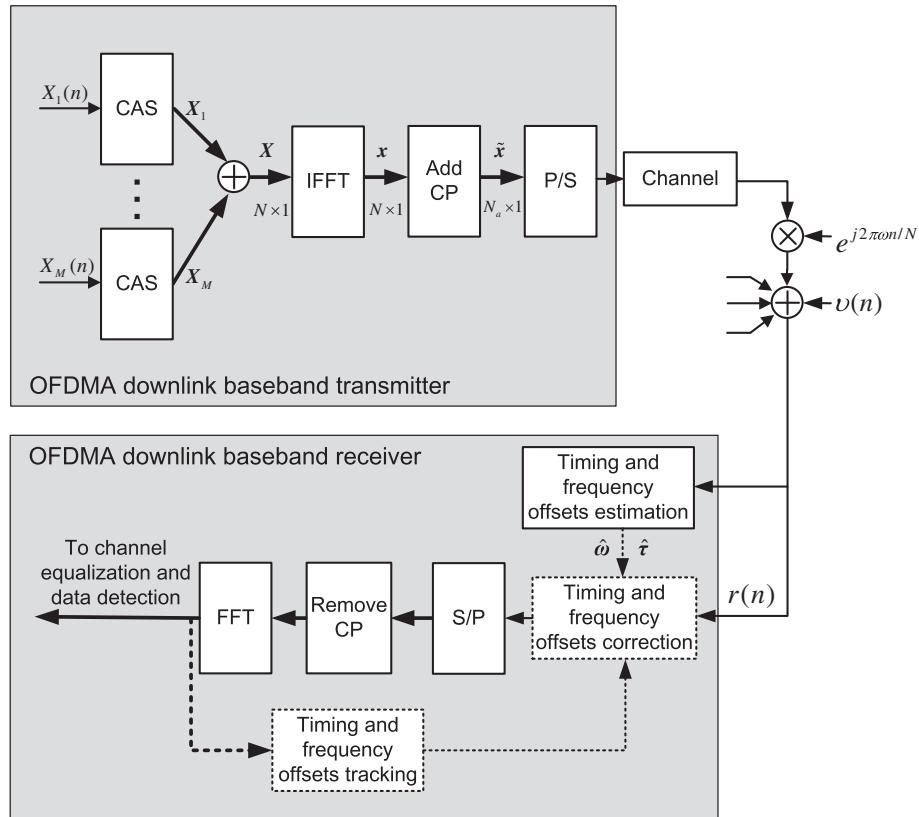
scheme (CAS) is shown in Figure 2.21a. A group of I_i adjacent subcarriers is assigned to the i th user in the subband CAS so that the signal of each user can be separated easily at the base station (BS) through a filter bank. However, subband CAS prevents the possibility of optimally exploiting the channel diversity. A deep fade might hit a substantial number of subcarriers of a given user if they are close together [100]. To reserve some multipath diversity, *interleaved CAS* shown in Figure 2.21b can be adopted for an uplink OFDMA system. The assigned subcarriers of the i th user are equi-spaced with a distance K_u in interleaved CAS. More dynamic resource allocation and flexibility can be achieved by employing *generalized CAS*, where no strict association between subcarriers and users is required, as illustrated in Figure 2.21c. However, generalized CAS will increase the synchronization complexity significantly as shown later.

2.02.3.1.2 OFDM transmission

The discrete-time block diagram of a standard downlink OFDMA transmission system is depicted in Figure 2.22. According to the CAS employed, I_i data symbols for each user and $N - \sum_{i=1}^M I_i$ zeros are assembled into one OFDM symbol as

$$X(k) = \begin{cases} X_i(k), & \text{if } k \in \mathcal{I}_i, \\ 0, & \text{otherwise.} \end{cases} \quad (2.30)$$

After the inverse DFT modulation, each OFDM block is preceded by a CP whose duration is usually longer than the delay spread of the propagation channel, so that IBI can be eliminated at the receiver

**FIGURE 2.22**

Discrete-time complex baseband representation of downlink OFDMA transmission.

without affecting the orthogonality of the subcarriers. The time-domain samples after CP insertion can be expressed as

$$\tilde{x}(n) = \begin{cases} x(N+n), & -N_g \leq n \leq -1, \\ x(n), & 0 \leq n \leq N-1, \end{cases} \quad (2.31)$$

where $x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi nk/N}$. The ($N_a = N + N_g$) samples of each block are then pulse shaped, upconverted to the carrier frequency, and transmitted sequentially through the channel.

In this chapter, we model the frequency-selective channel as a finite impulse response (FIR) filter with channel impulse response (CIR) $h = [h(0), \dots, h(L-1)]^T$, where L is the channel order and is determined by the maximum channel delay spread and data sampling rate. In practice, the system is usually designed such that $L \leq N_g \leq N$. We assume that the CIR is time invariant over $N_T \geq 1$ consecutive symbol blocks, but could vary from one set of N_T blocks to the next.

At the receiver, the signal is downconverted to baseband and sampled at the rate of N_a samples per extended OFDM symbol. We will index these received samples by $[-N_g, \dots, N - 1]$. Discarding the samples with indices $n = -N_g, \dots, -1$ is known as discarding the cyclic prefix. In a perfectly synchronized system, the received signal can be written as

$$r(n) = \sum_{l=0}^{L-1} h(l)\tilde{x}(n-l) + v(n) \quad (2.32)$$

for $n = 0, \dots, N - 1$; here $v(n)$ is complex-valued additive white Gaussian noise (AWGN) with variance σ_v^2 . Recall that with the insertion of CP, collected samples $\{r(n)\}_{n=0}^{N-1}$ can be expressed as

$$r = \mathcal{H}x + v, \quad (2.33)$$

where \mathcal{H} is an $(N \times N)$ circulant matrix whose first column is $[h(0), h(1), \dots, h(L-1), 0, \dots, 0]^T$. The circulant matrix \mathcal{H} can be written as $\mathcal{H} = \mathbf{F}^H \mathbf{H} \mathbf{F}$, where \mathbf{F} is the DFT matrix with $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$ and $\mathbf{H} = \text{diag}\{H(0), H(1), \dots, H(N-1)\}$ with

$$H(k) = \sum_{l=0}^{L-1} h(l)e^{-j2\pi kl/N}. \quad (2.34)$$

Hence, after performing DFT, the output R can be expressed as

$$R = \mathbf{H}X + \Upsilon, \quad (2.35)$$

where $X = [X(0), \dots, X(N-1)]^T$ and $\Upsilon = [\Upsilon(0), \dots, \Upsilon(N-1)]^T$ is again AWGN with covariance matrix $\sigma_\Upsilon^2 \mathbf{I}$. Since \mathbf{H} is a diagonal matrix, Eq. (2.35) indicates that the effect of the frequency-selective channel on the OFDM signal is completely captured by scalar multiplications of the data symbols by the frequency responses of the channel at the subcarrier frequencies. Further, demodulation at the receiver does not color the additive noise. If none of the channel zeros coincide with an activated subcarrier, maximum likelihood detection of the symbols is straightforward. Zero-forcing and minimum mean square error (MMSE) equalisers can also be applied on a per-carrier basis.

As we mentioned previously, the signal model (2.35) is only valid for an ideal timing and frequency synchronized system. However, in practical systems, Doppler shifts and instable oscillators result in a carrier frequency offset (CFO) f_0 between the received carrier and the local sinusoids used for signal demodulation. In addition, unknown transmission timing and propagation delay cause the DFT window to be placed in a wrong position at the receiver. This results in a timing error, denoted by t_d , which must properly be compensated to avoid severe performance degradation. Since fractional (normalized to sampling period) timing offsets can be absorbed into the channel, it is a common practice to model the timing offsets as a multiple of the sampling period. Letting ω (a real number) and τ denote the CFO normalized to the subcarrier spacing and the integer part of the timing offset normalized to the sample period, respectively, i.e., $\omega = Nf_0T_s$ and $\tau = \lfloor t_d/T_s \rfloor$, in the presence of timing and frequency offsets, then Eq. (2.32) becomes

$$r(n) = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)\tilde{x}(n-l-\tau) + v(n). \quad (2.36)$$

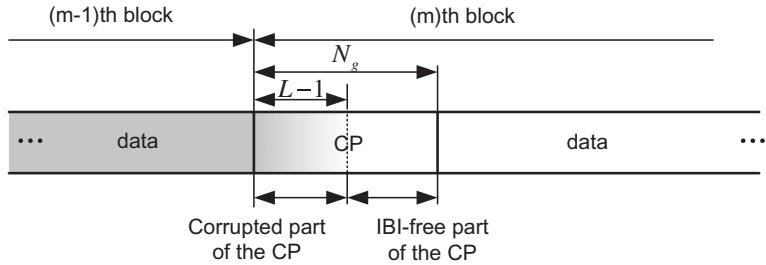
**FIGURE 2.23**

Illustration of the effect of multipath dispersion on CP.

For single timing and frequency offset, the timing and frequency offsets can be estimated and corrected as shown in Figure 2.22. However, it is a tough task to estimate and compensate multiple timing and frequency offsets. Next, we will review the effects of timing and frequency offsets on the performance of OFDM systems.

2.02.3.1.3 Effects of timing errors on system performance

Due to the multipath dispersion, the tail of each received block extends over the first $L - 1$ samples of the successive OFDM block as shown in Figure 2.23. By inserting a CP which consists of more than $L - 1$ samples, the interference from the previous OFDM block can be readily removed by properly determining the starting point of the OFDM symbol. However, the inaccuracy of timing offset estimation will cause performance degradation. To quantify the effect of timing errors on system performance, we assume perfect frequency synchronization here, i.e., $\omega = 0$.

Since the length of CP is (assumed to be) always larger than the maximum channel delay spread, and using the time-shift property of the Fourier transform, we find that the timing error $\Delta\tau = \hat{\tau} - \tau$ within interval $[-N_g + L - 1, 0]$ only causes a linear phase rotation across the subcarriers as

$$R(k) = e^{j2\pi k \Delta\tau / N} H(k) X(k) + \Upsilon(k). \quad (2.37)$$

The effect of this timing error can be readily compensated by the channel equaliser. On the other hand, if the timing error is outside interval $[-N_g + L - 1, 0]$, samples from adjacent OFDM blocks not only cause IBI, but also result in a loss of orthogonality among subcarriers which generates inter-carrier interference (ICI). A comprehensive mathematical analysis of the effects of timing errors is discussed in [94, 101]. In this case, the received signal after DFT can be written as

$$R(k) = \frac{\alpha_d}{N}(k) e^{j2\pi k \Delta\tau / N} H(k) X(k) + \gamma(k) + \Upsilon(k), \quad (2.38)$$

where

$$\alpha_d(k) = \begin{cases} (1 - e^{j2\pi k d / N}) / (1 - e^{j2\pi k / N}), & k \neq 0, \\ N - d, & k = 0, \end{cases} \quad (2.39)$$

$$d = \begin{cases} \Delta\tau, & \text{if } \Delta\tau > 0, \\ \max\{L - 1 - (N_g + \Delta\tau), 0\}, & \text{if } \Delta\tau \leq 0, \end{cases} \quad (2.40)$$

and $\gamma(k)$ is the combination of IBI and ICI which is defined as in [101]. Following the same lines as derived in [101], the signal-to-interference (SIR) in the presence of timing errors can be expressed as

$$\text{SIR} = \frac{(N-d)^2}{d(2N-d) - 2\frac{N-d}{\sigma_h^2}\beta(d)}, \quad (2.41)$$

where $\sigma_h^2 = \sum_{l=0}^{L-1} \sigma_{h(l)}^2$ and

$$\beta(d) = \begin{cases} \sum_{m=0}^{d-1} \sum_{l=m+1}^{L-1} \sigma_{h(l)}^2, & \Delta\tau > 0, \\ \sum_{m=0}^{d-1} \sum_{l=0}^{N_g + \Delta\tau + m - 1} \sigma_{h(l)}^2, & \Delta\tau \leq 0. \end{cases} \quad (2.42)$$

Figure 2.24 shows the SIR versus timing error $\Delta\tau$ for $N = 64$. Both the exponential power delay profile, i.e., $E\{|h(l)|^2\}_{l=0}^{L-1} = C \exp(-0.2l)$ where C is a scalar factor that ensures that the total energy of the channel taps is normalized to unity, and the uniform power delay profile, i.e., $E\{|h(l)|^2\}_{l=0}^{L-1} = 1/L$ profile, have been tested. The length of the CP and CIR are set to 16 and 8, respectively, i.e.,

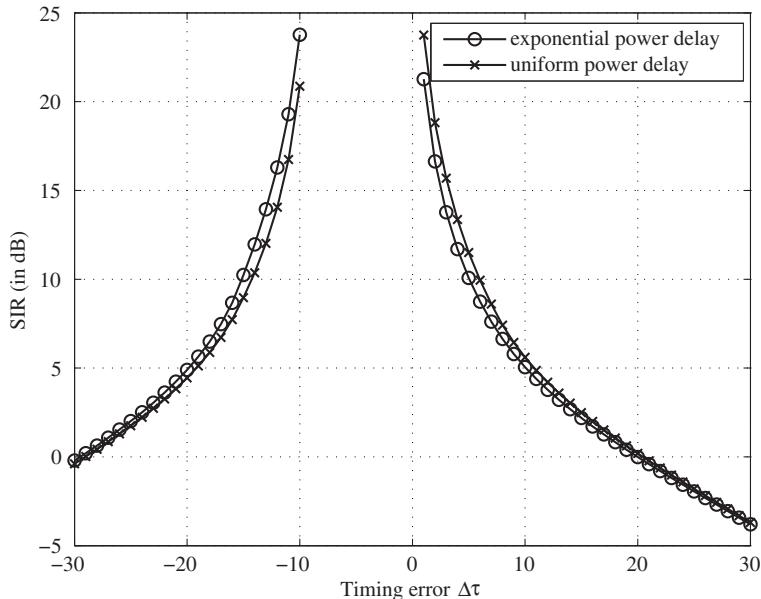


FIGURE 2.24

SIR versus timing error $\Delta\tau$.

$N_g = 16$ and $L = 8$. As discussed previously, the timing error within interval $[-N_g + L - 1, 0]$, which is $[-9, 0]$ for our simulation setup, will not cause ICI and IBI. The timing error outside the interval $[-9, 0]$ results in a significant SIR loss, especially for positive timing errors. To keep the SIR degradation within a tolerable level, accurate timing offset estimation is necessary.

2.02.3.1.4 Effects of carrier frequency offsets on system performance

As we mentioned in a previous section, CFO is caused by Doppler shifts and mismatched oscillators. In general, the CFO can be several times the subchannel spacing. Thus, the CFO is usually divided into an integer part and a fractional part by normalizing to the subcarrier spacing. Assuming perfect timing synchronization and using the frequency-shift property of the Fourier transform, the received signal under an integer valued CFO ω can be expressed as

$$R(k) = H((k - \omega)_{\text{mod } N}) X((k - \omega)_{\text{mod } N}) + \Upsilon(k). \quad (2.43)$$

It can be seen from the above equation that the integer valued CFO causes a circular shift of the transmitted symbols, but does not cause ICI; i.e., the orthogonality of the subcarriers is maintained. The fractional part, however, causes ICI. Assuming that CFO ω is a fractional value, the received signal can be written as

$$R(k) = \sum_{n=0}^{N-1} H(n) X(n) f(\omega + n - k) + \Upsilon(k), \quad (2.44)$$

where

$$f(n) = \frac{\sin(\pi n)}{N \sin(\pi n/N)} e^{j\pi n(N-1)/N}. \quad (2.45)$$

Equation (2.44) can be re-written as

$$R(k) = H(k) X(k) f(\omega) + \gamma(k) + \Upsilon(k), \quad (2.46)$$

where $\gamma(k) = \sum_{n=0, n \neq k}^{N-1} H(n) X(n) f(\omega + n - k)$ is the zero-mean ICI term with power $\sigma_\gamma^2(\omega)$.

After some manipulations as shown in [102], we have

$$\sigma_\gamma^2(\omega) = \sigma_x^2 (1 - |f(\omega)|^2).$$

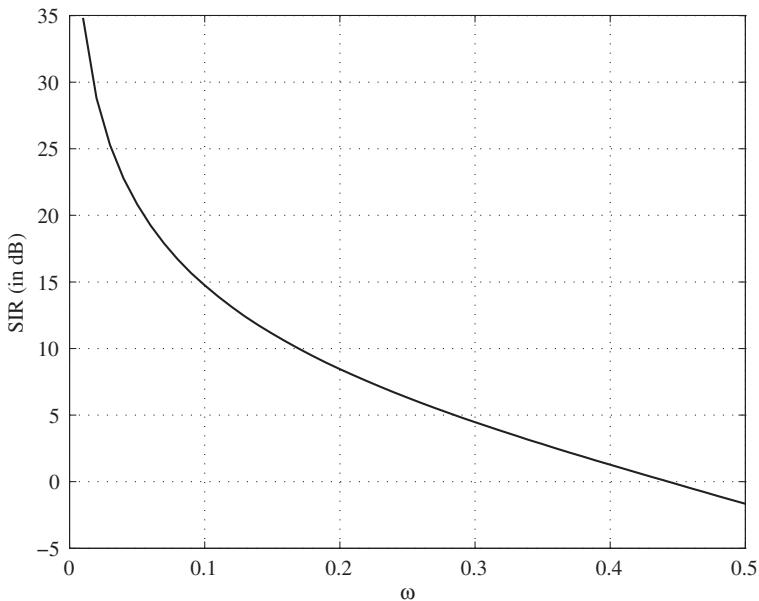
Thus, the SIR can be expressed as

$$\text{SIR} = \frac{|f(\omega)|^2}{1 - |f(\omega)|^2}. \quad (2.47)$$

The SIR versus CFO ω for $N = 64$ is shown in Figure 2.25. Notice that the SIR decreases rapidly as the CFO is increased. Again, to keep the SIR degradation to a tolerable level, effective CFO estimation and compensation methods are required. More precise techniques for computing the SNR loss due to CFO can be found in [103].

2.02.3.2 Downlink OFDMA

As shown in Figure 2.26, the BS broadcasts training sequences followed by data blocks to the potential users and each user operates independently in downlink OFDMA transmission. Thus, the synchronization problem in downlink OFDMA is similar to that of single-user OFDM systems. Generally,

**FIGURE 2.25**

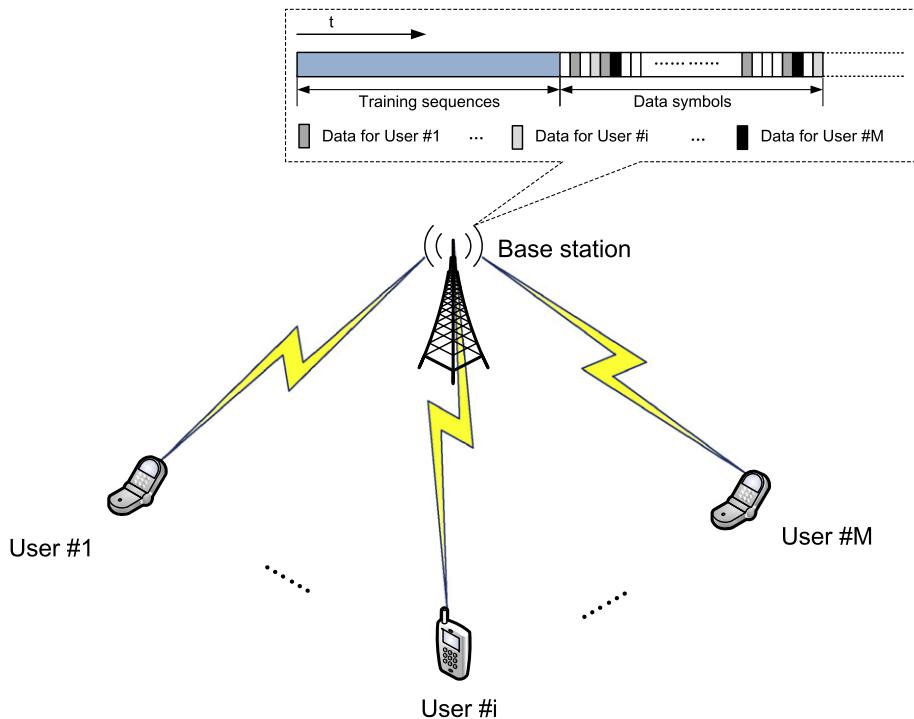
SIR versus fractional carrier frequency offset ω .

synchronization can be divided into a coarse acquisition phase and a fine tracking phase. In this section, we provide a brief survey of existing synchronization techniques in the downlink OFDMA scenario.

2.02.3.2.1 Coarse synchronization methods

The coarse synchronization task typically has two sub-tasks, i.e., finding the start of an OFDM frame over an approximate range of sample values and aligning the local oscillator of the receiver to the received carrier frequency. Coarse timing acquisition is usually the first task performed in the synchronization procedure. In practice, the CFO is assumed to be completely unknown at this stage. Hence, the orthogonality of the subcarriers may not be retained to provide a useful post-FFT signal. Consequently, coarse timing acquisition is obtained in the time domain. In coarse frequency synchronization, the usual approach is to decompose the CFO into a fractional part plus an integer part. Pre-FFT or post-FFT methods may be adopted to estimate the CFO.

Depending on the system requirements, coarse timing and frequency acquisition can be carried out by exploiting either the repeated cyclic prefix [104–106] or specially designed training sequences (preambles) [107–110]. Exploiting the correlation of CP, CP-based algorithms can work blindly without the overhead of an explicit training sequence. However, as standardized in many commercial systems, reliable coarse acquisition methods for frequency-selective channels are based on a training sequence with a repetitive structure. The motivation behind this idea is that the repetitive property is preserved

**FIGURE 2.26**

Downlink OFDMA representation.

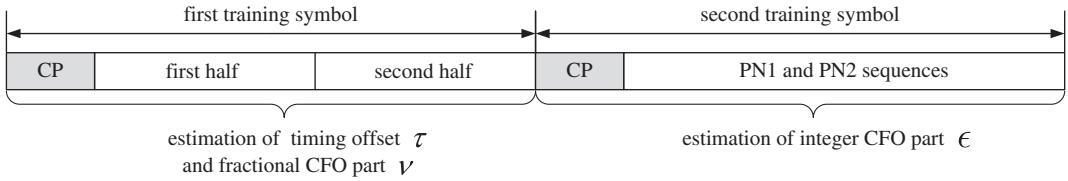
after propagation through a multipath channel, except for a phase rotation due to the CFO. In this chapter, we consider training-sequence based synchronization in the following.

Second-order statistics based methods

There are basically two methods for training-sequence based coarse synchronization, i.e., first-order statistics based and second-order statistics based methods. The latter class was first proposed by Moose in [111] and further studied by Schmidl and Cox (S&C) in [107], where two identical slots with length of $N/2$ were used in the first training symbol as shown in Figure 2.27. The CFO normalized to the subcarrier spacing is decomposed into two parts as

$$\omega = \nu + 2\varepsilon, \quad (2.48)$$

where $\nu \in (-1, 1]$ and ε is an integer. To generate the repetitive-structure of the first training symbol, we can simply transmit a pseudonoise (PN) sequence on the even subcarriers, while zeros are used on the odd subcarriers. For the second training symbol, a PN sequence PN1 is transmitted on the odd subcarriers which may be employed in channel estimation; and a differentially-modulated PN sequence PN2 deployed on the even subcarriers is used for the integer CFO estimation. Let $\{r(n)\}_{n=0}^{N-1}$ denote the

**FIGURE 2.27**

Training symbols of S&C algorithm.

received signals, we have

$$r(n) = z(n) + v(n), \quad (2.49)$$

where

$$z(n) = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)x(n - \tau - l), \quad (2.50)$$

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N/2-1} X(2k)e^{-j4\pi kn/N}, \quad (2.51)$$

where τ , an integer, is the normalized nonfractional part of the timing offset, ω is the CFO normalized to the subcarrier spacing, $x(n)$ and $X(k)$ are respectively the transmitted time domain and frequency domain training sequences, $h(l)$ is the l th tap of the CIR and $v(n)$ is an AWGN with variance σ_v^2 . It can be easily verified that the first and second halves of the received signal can be expressed as

$$r(n) = z(n) + v(n), \quad \tau \leq n \leq \tau + \frac{N}{2} - 1, \quad (2.52)$$

$$r\left(n + \frac{N}{2}\right) = e^{j\pi\nu} z(n) + v\left(n + \frac{N}{2}\right), \quad \tau \leq n \leq \tau + \frac{N}{2} - 1. \quad (2.53)$$

Exploiting the correlation between first and second halves, the S&C timing estimator can be expressed as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \Lambda(\tilde{\tau}), \quad (2.54)$$

where

$$\Lambda(\tilde{\tau}) = \frac{\left| \sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} r^*(n)r\left(n + \frac{N}{2}\right) \right|^2}{\left(\sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} |r\left(n + \frac{N}{2}\right)|^2 \right)^2}.$$

Moreover, assuming perfect timing synchronization, the estimate of fractional part CFO ν can be obtained as

$$\hat{\nu} = \frac{1}{\pi} \arg \left\{ \sum_{n=\tau}^{\tau+N/2-1} r^*(n)r\left(n + \frac{N}{2}\right) \right\}. \quad (2.55)$$

In practice, the timing offset τ in Eq. (2.55) can be replaced by its estimated value $\hat{\tau}$ given in eq. (2.54). If the normalized CFO can be guaranteed to be less than 1, the second training symbol would not be needed. Otherwise, the second training symbol and a post-FFT method can be adopted to estimate the integer part of the CFO, ε , as we describe next.

After compensating the fractional offset by multiplying the two training symbols by $e^{-j2\pi\hat{v}n/N}$, the FFT output of two training symbols, denoted as $R_1(k)$ and $R_2(k)$, can be expressed as

$$R_1(k) = Z_1(k) + W_1(k) \quad (2.56)$$

$$R_2(k) = Z_2(k) + W_2(k) \quad (2.57)$$

for $k = 0, \dots, N - 1$ and

$$Z_1(k) = H((k - 2\varepsilon)_{\text{mod } N})X_1((k - 2\varepsilon)_{\text{mod } N}) \quad (2.58)$$

$$Z_2(k) = e^{j\phi}H((k - 2\varepsilon)_{\text{mod } N})X_2((k - 2\varepsilon)_{\text{mod } N}), \quad (2.59)$$

where $\phi = 4\pi\varepsilon(N + N_g)/N$. Let $d(k) = \frac{X_2(k)}{X_1(k)}$ represent the differentially-modulated PN sequence on the even subcarriers of the second training symbol, we have

$$Z_2(k) = e^{j\phi}d((k - 2\varepsilon)_{\text{mod } N})Z_1(k) \quad (2.60)$$

for even k . Thus, the estimator of the integer part of the CFO, ε , can be expressed as

$$\hat{\varepsilon} = \arg \max_{\tilde{\varepsilon}} \Psi(\tilde{\varepsilon}), \quad (2.61)$$

where

$$\Psi(\tilde{\varepsilon}) = \frac{\left| \sum_{k \text{ even}} R_1^*(k)d^*((k - 2\tilde{\varepsilon})_{\text{mod } N})R_2(k) \right|^2}{2 \left(\sum_{k \text{ even}} |R_2(k)|^2 \right)^2}, \quad (2.62)$$

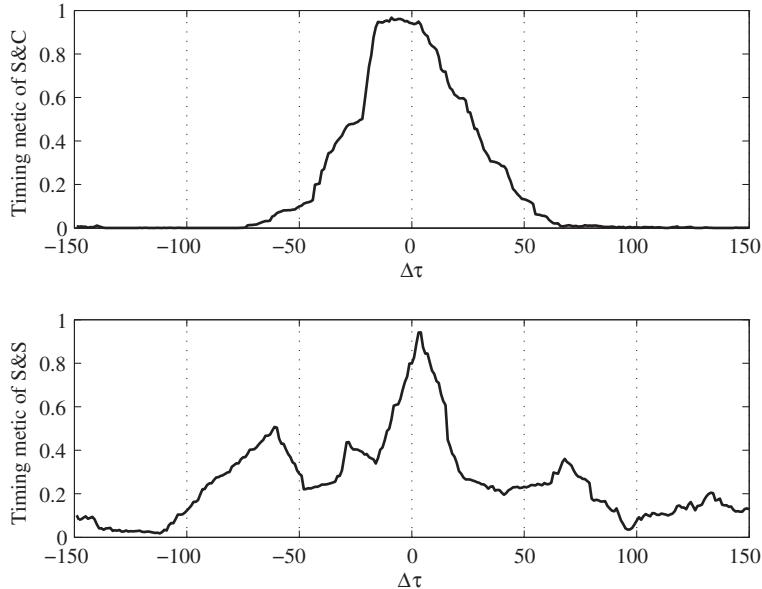
and integer $\tilde{\varepsilon}$ varies over the range of possible frequency offsets. Then, the frequency offset estimate would be $\hat{\omega} = \hat{v} + 2\hat{\varepsilon}$.

The S&C timing estimator (2.54) is expected to capture a peak when the correlation window is perfectly aligned with the received training sequence. Unfortunately, as shown in Figure 2.28, the timing metric of the S&C estimator exhibits a “plateau” which reduces the acquisition accuracy significantly. To obtain a steeper timing metric trajectory, many training patterns were proposed in [108, 109]. In [109], a training of the form $[\mathbf{B}, \mathbf{B}, -\mathbf{B}, \mathbf{B}]$ in time domain was proposed by Shi and Serpedin (S&S). Arranging the received N samples $[r(\tilde{\tau}) \dots r(\tilde{\tau} + N - 1)]^T$ into four parts as $\{r_i(\tilde{\tau}) = [r(iN/4 + \tilde{\tau}) \dots r((i+1)N/4 + \tilde{\tau} - 1)]^T\}_{i=0}^3$, the S&S timing estimator can be expressed as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \Lambda(\tilde{\tau}), \quad (2.63)$$

where

$$\Lambda(\tilde{\tau}) = \frac{\sum_{i=0}^2 |P_i(\tilde{\tau})|}{\frac{3}{2} \sum_{i=0}^3 \|r_i(\tilde{\tau})\|^2},$$

**FIGURE 2.28**

Timing metrics for S&C and S&S estimators, $N = 128$, $N_g = 16$, $L = 16$, SNR=20 dB.

and

$$\begin{aligned} P_0(\tilde{\tau}) &= r_0^H(\tilde{\tau})r_1(\tilde{\tau}) - r_1^H(\tilde{\tau})r_2(\tilde{\tau}) - r_2^H(\tilde{\tau})r_3(\tilde{\tau}), \\ P_1(\tilde{\tau}) &= r_1^H(\tilde{\tau})r_3(\tilde{\tau}) - r_0^H(\tilde{\tau})r_2(\tilde{\tau}), \\ P_2(\tilde{\tau}) &= r_0^H(\tilde{\tau})r_3(\tilde{\tau}). \end{aligned}$$

Since the training symbol is divided into four parts, we find that the CFO causes a phase shift of $\pi\omega/2$ in each part for a flat fading channel. Thus, the CFO estimator of S&S algorithm for practical systems can be expressed as

$$\hat{\omega} = \frac{2}{\pi} \arg \{P_0(\hat{\tau})\}. \quad (2.64)$$

Compared to the S&C CFO estimator (2.55), we can see that the acquisition range of (2.64) increased to $[-2, 2]$.

Both timing metrics of the S&C and S&S timing estimators, $\Lambda(\tilde{\tau})$, are illustrated in Figure 2.28, where $\Delta\tau = \tilde{\tau} - \tau$. The results are obtained under the exponential power delay profile channel introduced previously and the SNR is defined as σ_x^2/σ_v^2 . We can see that the “plateau” present in the S&C estimator is significantly reduced in the S&S estimator. As pointed out in [108], a steeper timing metric trajectory can be obtained by increasing the number of repetitive slots.

Although the correlation method adopted in S&C and S&S estimators has low computational complexity, those estimators will exhibit a floor effect since they are still based on second-order statistics of the received signal. As shown in [112], much more accurate timing and frequency estimation can be achieved by using the first-order statistics, at the expense of a slight increase in implementation complexity.

First-order statistics based methods

Using the signal model (2.49), the mean of the received signal is given by

$$\mathbb{E}\{r(n)\} = e^{j2\pi\omega n/N} \sum_{l=0}^{L-1} h(l)x(n - \tau - l), n = \tau - N_g, \dots, \tau + N + L - 1. \quad (2.65)$$

Let $r_{\tilde{\tau}} = [r(\tilde{\tau}), \dots, r(\tilde{\tau} + N - 1)]^T$ and $\mathbb{E}\{r_{\tilde{\tau}}\} \triangleq \mu = \Gamma(\omega)\mathbf{X}h$, where $\Gamma(\omega) = \text{diag}\{1, \dots, e^{j2\pi\omega(N-1)/N}\}$. \mathbf{X} is a circulant matrix whose first column is x .

The variance of the received signal is

$$\text{var}\{r(n)\} = \begin{cases} \sigma_v^2, & n = 0, \dots, \tau + N - 1, \\ > \sigma_v^2, & n > \tau + N - 1. \end{cases} \quad (2.66)$$

Thus, the variance of the received signal is minimum during the noise-only period, which precedes reception of the frame, and during the reception of the training block. During data reception, it is equal to $\sigma_v^2 + \|h\|^2$, where we assume that data symbols have unit power, without loss of generality. Figure 2.29 illustrates these observation. If μ were known, τ could be estimated by minimizing the Euclidean distance between $r_{\tilde{\tau}}$ and μ over $\tilde{\tau}$. Since this is not the case and in order to avoid the noise-only period, Ghogho and Swami [112] proposed two estimators which are obtained by minimizing the following modified versions of the nonlinear least-squares (NLLS) criterion:

$$C_1(\tilde{\tau}, \tilde{\omega}, \tilde{h}) = \frac{\|r_{\tilde{\tau}} - \tilde{\mu}\|^2}{\|r_{\tilde{\tau}}\|^2} \quad (2.67)$$

and

$$C_2(\tilde{\tau}, \tilde{\omega}, \tilde{h}) = \|r_{\tilde{\tau}} - \tilde{\mu}\|^2 - \|r_{\tilde{\tau}}\|^2 \quad (2.68)$$

over $\tilde{\tau}$, $\tilde{\omega}$ and \tilde{h} , where $\tilde{\mu}$ is obtained as in μ after replacing h and ω by \tilde{h} and $\tilde{\omega}$, respectively.

Both the normalization factor in (2.67) and the second term of the RHS of (2.68) guarantee the uniqueness of the solution, i.e., avoid the noise-only (i.e., $\tilde{h} = 0$) solution. Indeed, in the noise-only period, the minima of $E\{C_1\}$ and $E\{C_2\}$, which are obtained with $\tilde{h} = 0$, are (approximately) one and zero, respectively. During data reception, the minima of $E\{C_1\}$ and $E\{C_2\}$ are again obtained with $\tilde{h} = 0$ and are also approximately equal to one and zero, respectively. During the reception of the training sequence, the minima of the $E\{C_1\}$ and $E\{C_2\}$ are (approximately) $1/(1 + \text{SNR})$ and $-\|h\|^2$, respectively, where SNR is defined as $\text{SNR} = \|r_{\tau}\|^2/(N\sigma_v^2)$. These minima are smaller than those obtained in the noise-only and data transmission periods. Hence, when the processed signal contains the received preamble, the minima of $E\{C_1\}$ and $E\{C_2\}$ are achieved if and only if $\tilde{\tau} = \tau$ and $\tilde{\mu} = \mu$. Figure 2.29 illustrates some of these results.

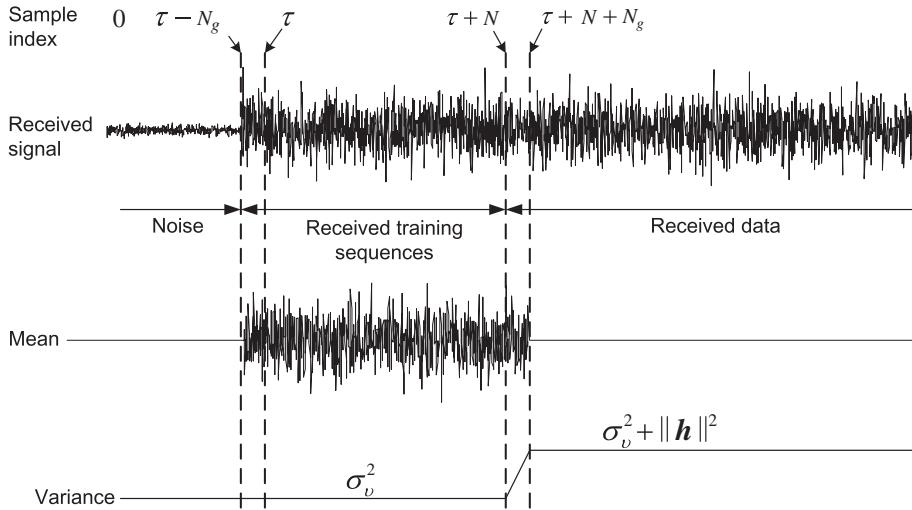


FIGURE 2.29

Real part of one realization of the received signal as well as the corresponding instantaneous mean and variance.

Since the statistical expectation of C_1 and C_2 are unknown, only estimates of the unknown parameters can be obtained by minimizing C_1 and C_2 themselves. The obtained estimators were referred to as modified nonlinear least squares estimators (MNLLS) in [112], since they combine the NLLS estimation method and detection. The above optimization problems can be simplified by noting that C_1 and C_2 are quadratic in \tilde{h} . Thus, closed-form expressions for \hat{h}_1 and \hat{h}_2 can be obtained as

$$\hat{h}_i = \mathbf{X}^\dagger \Gamma^H(\hat{\omega}_i) r_{\hat{\tau}_i}, \quad i = 1, 2. \quad (2.69)$$

Substituting the above estimates for \tilde{h} in C_1 and C_2 , the equivalent criterion is to maximize

$$C_i''(\tilde{\tau}, \tilde{\omega}) = g_i(\tilde{\tau}) r_{\tilde{\tau}}^H \Gamma(\tilde{\omega}) \Pi_X \Gamma^H(\tilde{\omega}) r_{\tilde{\tau}}, \quad i = 1, 2, \quad (2.70)$$

where $g_1(\tilde{\tau}) = \|r_{\tilde{\tau}}\|^{-2}$ and $g_2(\tilde{\tau}) = 1$, and $\Pi_X = \mathbf{X}(\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H$, which is a fixed matrix and can thus be precomputed and stored at the receiver.

The above optimization problems are two-dimensional. Although they are discrete in one dimension, they are still computationally challenging to solve. To reduce the computational complexity, timing acquisition using C_i'' is performed by ignoring the CFO-related terms. It was shown that a coarse but closed-form estimate of the CFO for each timing offset candidate is good enough to (nearly) obtain $\hat{\tau}_1$ and $\hat{\tau}_2$. To obtain the CFO estimate, the repetitive structure of the training block and the second-order statistic-based method in [107] were adopted to estimate the fractional part of ω , i.e., v . The estimate of the integer part of ω , \hat{e} and timing offset $\hat{\tau}_i$ are given by

$$\{\hat{\tau}_i, \hat{e}_i\} = \arg \max_{\tilde{\tau}, \tilde{e}} g_i(\tilde{\tau}) r_{\tilde{\tau}}^H \Gamma(\tilde{e} + \tilde{v}) \Pi_X \Gamma^H(\tilde{e} + \tilde{v}) r_{\tilde{\tau}}, \quad (2.71)$$

where candidate values for $\tilde{\varepsilon}$ are in $(-Q + 1, Q - 1)$, and \tilde{v} is given by [107]

$$\tilde{v} = \frac{1}{\pi} \arg \left\{ \sum_{n=\tilde{\tau}}^{\tilde{\tau}+N/2-1} r^*(n)r(n+N/2) \right\}. \quad (2.72)$$

Although the optimization problem (2.71) is two-dimensional, it is discrete and the possible values of ε dictated by Q may be small in practice. It is worth pointing out that the optimization problems can be reduced to one-dimensional problems if the preamble is made of Q repetitive slots, since in this case, a closed-form estimate of ω can be obtained (see [113, 114]). However, the performance of timing acquisition in this case becomes similar to that of existing methods, unless $N \gg 2Q(L + 1)$.

Once τ and ε are estimated, estimates of ω are obtained as $\hat{\omega}_i = \hat{v} + \hat{\varepsilon}$, where \hat{v} is given by (2.72) with $\tilde{\tau} = \hat{\tau}$. A more accurate estimate can be obtained by maximizing $C_i''(\hat{\tau}, \tilde{\omega})$ over $\tilde{\omega}$ after initialising with $\hat{\omega}_i$. Since this results in the optimization of a continuous-valued variable, ω , it may not be appropriate in practice, especially for the downlink. Moreover, simulations have shown that the performance improvement is not significant.

Finally, as a product of the above synchronization method, estimates of the channel can be obtained from (2.69) after replacing τ and ω by the above estimates.

To further reduce the complexity of the computation of the cost functions C_i'' , the projection matrix Π_X can be replaced by $(1/N)\mathbf{X}\mathbf{X}^H$ obtained by approximating $\mathbf{X}^H\mathbf{X}$ by $N\mathbf{I}$, since $N/2 > L + 1$, and using the law of large numbers. Using this, C_2'' is obtained as the squared L_2 -norm of a vector and C_1'' is its normalized version, and the corresponding estimates of τ_i and ε_i are given by

$$\{\check{\tau}_i, \check{\varepsilon}_i\} = \arg \max_{\tilde{\tau}, \tilde{\varepsilon} \in (-Q/2, Q/2)} N^{-1} g_i(\tilde{\tau}) \|\mathbf{X}^H \Gamma^H (\tilde{v} + \tilde{\varepsilon}) r_{\tilde{\tau}}\|^2. \quad (2.73)$$

Note that the computation of $\mathbf{X}^H y$ requires only changing the sign of some of the elements of the vector y and additions. Also, note that the computation of the simplified version of C_2'' is simpler than that of C_1'' since in the former, $\|r_{\tilde{\tau}}\|^2$ needs to be computed. However, the normalization in C_1'' is desirable for reducing the complexity of the search procedure for real-time implementation.

2.02.3.2.2 Fine synchronization methods

In certain applications, due to unstable oscillators or high-mobility environment, the coarse estimates may be inadequate for the entire frame. Thus, timing and frequency tracking are required to compensate the short-term variations produced by oscillator drifts and/or time-varying Doppler shifts.

Fine timing synchronization can be achieved through correlation either in the time domain [115, 116] or in the frequency domain [117–120]. A timing tracking scheme based on time-domain PN-sequence correlation and a weighted time correlation scheme exploiting the redundancy in both the cyclic prefix and available pilot symbols were shown to provide better performance than repetitive-structured OFDM symbols in [115, 116], respectively. Frequency-domain based schemes model the timing error as part of the CIR vector. This is motivated by the fact that the estimated CIR shifts cyclicly with respect to the FFT-window. Since the frequency-domain based schemes can resolve channel multipaths effectively, they generally provide better performance than the time-domain schemes in multipath fading channels. There are several approaches to update the coarse timing estimate in frequency-domain based tracking

algorithms. One method is to locate the peak position of the estimated CIR \hat{h} [118]. Many modified peak-finding versions were studied in [108], to take into account both the effects of noise, as well as the fact that the first peak may not be the largest peak in the CIR. Another method which maximizes the energy window of the channel estimation has been investigated in [105, 117]. Exploiting the timing information embedded in pilot-aided channel estimation, timing can be estimated without a specific training sequence, as was shown in [121, 122].

Similar to fine timing synchronization, frequency tracking can be performed either in the time domain [104, 123] or in the frequency domain [124, 125]. In [104, 123], the correlation between the CP and the last N_c samples of each block is exploited to estimate the residual frequency offset. The residual CFO was tracked using the temporal correlation in the data-aided post-FFT stage and the frequency domain channel estimate was adopted to deduce the weights for a weighted-least-squares CFO estimator in [124, 125], respectively. Several blind CFO estimation methods, e.g., [126–136], also can be employed during the tracking step without increasing synchronization overhead.

2.02.3.2.3 Simulation results for downlink OFDMA systems

We consider an downlink OFDMA system with a total of $N = 128$ subcarriers. In the simulation, we assume that the CIR length is $L = 16$, and a cyclic prefix of length $N_g = 16$. The channel taps $h(l)$ are uncorrelated zero-mean Gaussian random variables with exponential power delay profile $E\{|h(l)|^2\}_{l=0}^{L-1} = C \exp(-0.2l)$, where C is a scalar factor that ensures that the total energy of the

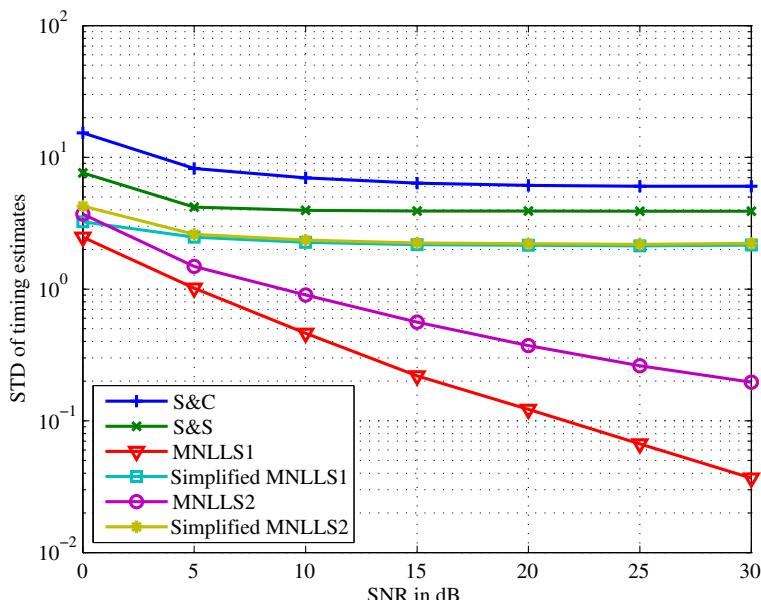


FIGURE 2.30

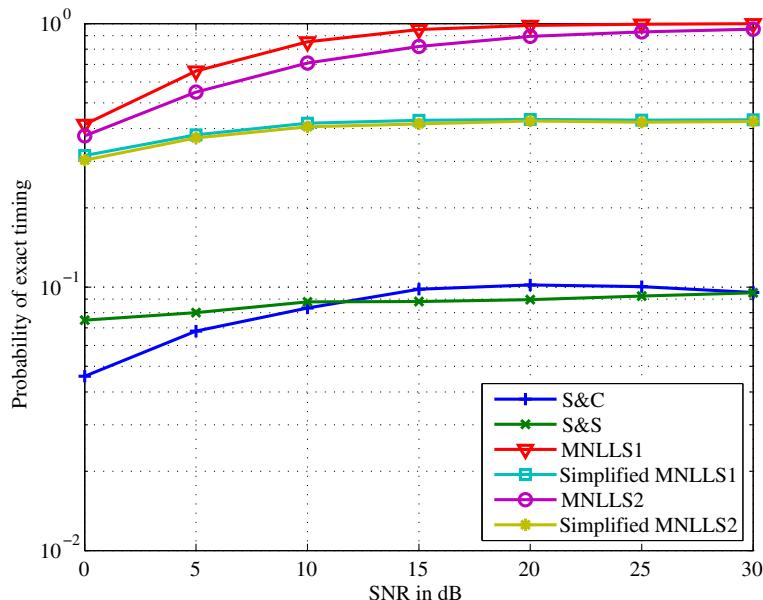
Standard deviation of timing estimates for downlink OFDMA.

channel taps is normalized to unity. Correspondingly, the SNR of received signal is equal to σ_x^2/σ_v^2 , where σ_x^2 is the power allocated to each subcarrier. Since only the fractional part of the CFO can be estimated by the S&S algorithm in [109], we generate ω randomly from the interval $[-0.5, 0.5]$ and generate a new random channel for each Monte Carlo run. Moreover, as the integer part of the CFO is zero, the second training symbol shown in Figure 2.27 is unnecessary and we set $Q = 1$ for the first-order statistics based methods.

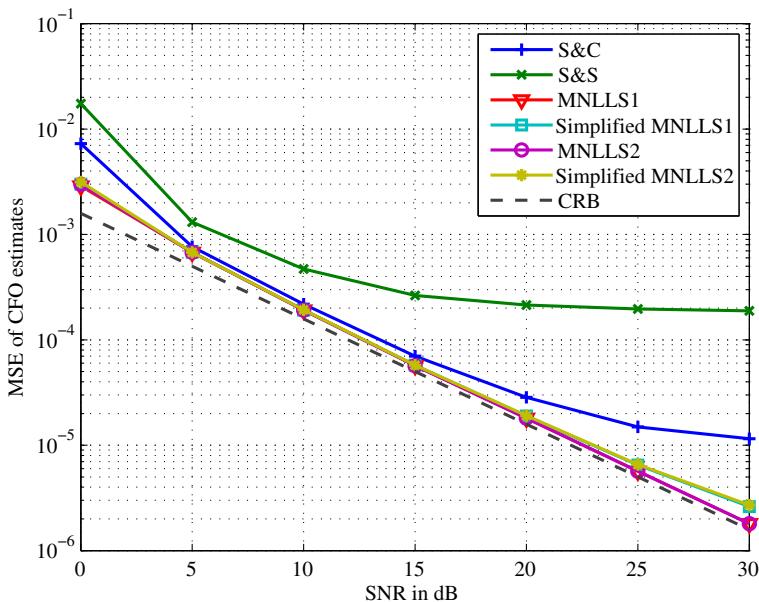
The results of timing and frequency estimates are calculated using 20,000 Monte Carlo runs. Figures 2.30–2.32 show that the first-order statistics based methods significantly outperform the second-order statistics based methods in terms of timing and CFO estimation. For timing estimation, MNLLS1 outperforms MNLLS2. The simplified MNLLS algorithms yield the same timing estimation performance. In Figure 2.31, the probability of exact timing refers to the probability that the associated algorithm identifies τ without error. For comparison, in Figure 2.32, we also show the Cramér-Rao bound (CRB) for the case where timing is perfect. Using the method in [137], the CRB is found to be

$$\text{var}\{\hat{\omega}\} \geq \text{CRB}(\hat{\omega}) = \frac{2\sigma_v^2}{N\pi^2\sigma_x^2}. \quad (2.74)$$

Notice that the performance of first-order statistics based methods is close to CRB. The simplified MNLLS methods have the same CFO estimation performance merits as the MNLLS methods. From the simulation results, we can see that the S&S algorithm achieves more accurate timing estimation

**FIGURE 2.31**

Probability of exact timing acquisition for downlink OFDMA.

**FIGURE 2.32**

Mean square error of CFO estimates for downlink OFDMA.

performance than the S&C algorithm at the price of a decrease in the CFO estimation performance. As we discussed previously, the gains of first-order statistics based methods come along with the increasing complexity. For real-time implementation, it is important to set a threshold on the synchronization criterion so that complexity can be reduced by pruning the set of timing candidates. Details of implementation issues of first-order statistics based methods can be found in [112].

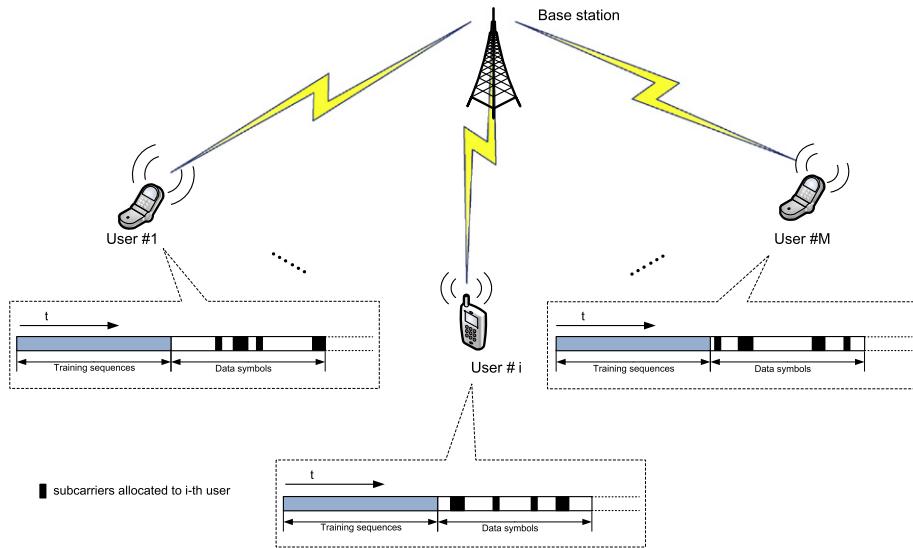
2.02.4 Multiuser synchronization

2.02.4.1 Uplink signal model and synchronization policy

We consider an uplink OFDMA system where M active users simultaneously communicate with the BS as shown in Figure 2.33. The users' data streams are assembled into OFDM symbols according to the CAS employed.

Let $r(n)$ denote the signal received at the BS; we have that

$$r(n) = \sum_{i=1}^M z_{(i)}(n) + v(n), \quad (2.75)$$

**FIGURE 2.33**

OFDMA uplink representation.

where $z_{(i)}(n)$ is the signal transmitted from the i th user and can be expressed as

$$z_{(i)}(n) = e^{j2\pi\omega_i n/N} \sum_{l=0}^{L-1} h_i(l) x_i(n - l - \tau_i), \quad (2.76)$$

and

$$x_i(n) = \frac{1}{\sqrt{N}} \sum_{k \in \mathcal{I}_i} X(k) e^{-j2\pi kn/N}. \quad (2.77)$$

It can be seen from Eq. (2.75) that the received signal at BS is the combination of the signal from all active users. Thus, the uplink synchronization is a multi-parameter estimation procedure. To guarantee that the residual synchronization errors of the uplink transmission are much smaller than that of a completely asynchronous system, before uplink transmission, an initial synchronization is performed during downlink procedure. The timing offsets among users in the uplink are mainly due to different propagation distances between users and BS. The frequency offsets between users and BS are caused by the Doppler spread and/or the instability of local oscillators. Generally, after the synchronization performed via downlink transmission, the CFO can be guaranteed to be in a small range. In this chapter, we assume the frequency offset is smaller than half the subcarrier spacing.

To combat the residual synchronization errors simply and directly, a method based on downlink control channel is suggested in [100,138], where the synchronization parameters are estimated at the BS and adjustment is performed at the user side based on the information derived from feedback channel.

A similar idea is adopted in IEEE 802.16e [139] standard to accomplish the synchronization task. By using interference cancellation or multiuser detection algorithms, e.g., [140–143] and references therein, the effects of multiple frequency offsets also can be mitigated at the BS at the price of increased receiver complexity. In this chapter, we focus only on the estimation of the timing and frequency offsets at the BS. Depending on the CAS employed, the synchronization task in uplink OFDMA can be categorized into three cases as explained next.

2.02.4.2 Synchronization with subband CAS

The ML estimation of timing and frequency offsets for subband CAS uplink OFDMA was first studied in [138], where the users' signals are separated by a bank of band-pass filters at BS as shown in Figure 2.34. After signal separation, the timing and frequency offsets can be estimated independently for each active user, which is similar to the downlink OFDMA case. Since perfect signal separation is impractical, the filtered i th subband signal $r_{(i)}(n)$ can be written as

$$r_{(i)}(n) = z_{(i)}(n) + \Phi_{(i)}(n) + v_{(i)}(n), \quad (2.78)$$

where the $\Phi_{(i)}(n)$ and $v_{(i)}(n)$ denote the inter-carrier interference due to imperfect separation and noise corresponding to i th subband signal. Possible ways of reducing the interference term $\Phi_{(i)}(n)$ include the adoption of higher order band-pass filters or increasing the number of guard carriers used between adjacent frequency bands. The timing and frequency offsets estimators in [138] exploit the redundancy of CP; the estimates for the i th user can be expressed as:

$$\hat{\tau}_i = \arg \max_{\tilde{\tau}_i} \{ \Lambda_i(\tilde{\tau}_i) - \rho_i C_i(\tilde{\tau}_i) \}, \quad (2.79)$$

$$\hat{\omega}_i = \frac{1}{2\pi} \arg \{ \Lambda_i(\hat{\tau}_i) \}, \quad (2.80)$$

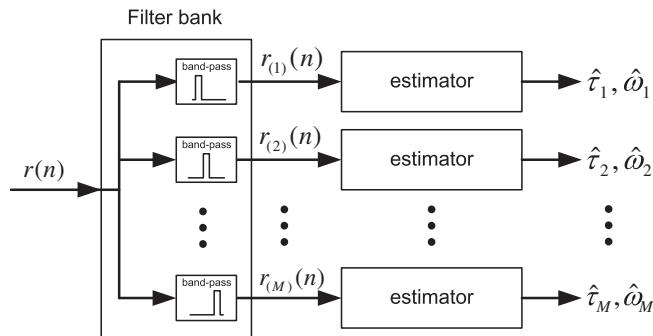


FIGURE 2.34

Timing and frequency offsets estimation for subband CAS uplink OFDMA.

where

$$\Lambda_i(\tilde{\tau}_i) = \sum_{n=\tilde{\tau}_i-N_g}^{\tilde{\tau}_i-1} r_{(i)}^*(n)r_{(i)}(n+N), \quad (2.81)$$

$$C_i(\tilde{\tau}_i) = \frac{1}{2} \sum_{n=\tilde{\tau}_i-N_g}^{\tilde{\tau}_i-1} (|r_{(i)}(n)|^2 + |r_{(i)}(n+N)|^2), \quad (2.82)$$

and $\rho_i = \sigma_x^2/(\sigma_x^2 + \sigma_v^2)$. The estimates in (2.79) and (2.80) are one-shot estimators in the sense that the estimates are based on the observation of a single OFDM symbol. More accurate estimates can be obtained by averaging the cost function over Q successive OFDM blocks as

$$\hat{\tau}_i = \arg \max_{\tilde{\tau}_i} \left\{ \sum_{q=0}^{Q-1} (\Lambda_i(\tilde{\tau}_i + q(N + N_g)) - \rho_i C_i(\tilde{\tau}_i + q(N + N_g))) \right\}, \quad (2.83)$$

$$\hat{\omega}_i = \frac{1}{2\pi} \arg \left\{ \sum_{q=0}^{Q-1} \Lambda_i(\hat{\tau}_i + q(N + N_g)) \right\}. \quad (2.84)$$

An alternative blind scheme to obtain estimates of timing and frequency offsets for subband CAS can be found in [144]. As pointed out in [145], the subband CAS offers the possibility of separating signals from different users through a simple filter bank even in a completely asynchronous scenario with arbitrarily large timing errors. Synchronization algorithms for the downlink OFDMA can be easily extended to subband-based OFDMA systems. On the other hand, grouping subcarriers together makes systems vulnerable to frequency-selective fading. The adoption of an interleaved CAS can provide users with some form of frequency diversity at the expense of slightly increasing the complexity of synchronization.

2.02.4.3 Synchronization with interleaved CAS

Interleaved subcarrier allocation minimizes the distances between subcarriers assigned to different users; hence, in the presence of frequency synchronization errors, signals from different users will overlap in the time domain and interfere with each other in the frequency domain due to loss of orthogonality [146]. Thus, it is a challenging task to separate the multiple user signals compared to the subband CAS. However, advanced signal-processing algorithms, e.g., subspace decomposition based methods, can be employed to reduce the synchronization complexity of interleaved systems.

Subspace-based CFO estimation algorithms are studied in [146] and later in [147, 148] for the uplink OFDMA systems; the key is to exploit the periodic structure of the interleaved transmission. generalized subspace-based CFO estimation also have been studied by [114, 127, 130, 132] for single user OFDM systems. Let $\mathcal{I}_i = \{\eta_i + pK_u\}_{p=0}^{P-1}$ denote the indices set of the subcarriers allocated to i th user, where η_i is an integer in the interval $[0, K_u - 1]$ and $\eta_i \neq \eta_j$ if $i \neq j$, K_u is the maximum number of users and $P = N/K_u$. We assume that the total number of subcarriers N is an integer multiple of K_u in this

chapter. The signal from i th user, given in Eq. (2.76), can be re-written as

$$z_{(i)}(n) = \alpha e^{j2\pi(\omega_i + \eta_i)n/N} \sum_{p=0}^{P-1} H_i(\eta_i + pK_u) X_i(\eta_i + pK_u) e^{j2\pi p(n-\tau_i)/P} \quad (2.85)$$

for $n = 0, \dots, N - 1$ and $\alpha = e^{-j2\pi\eta_i\tau_i/N}$. It can be found readily that

$$z_{(i)}(n + \mu P) = e^{j\mu\theta_i} z_{(i)}(n) \quad (2.86)$$

for $n = 0, \dots, P - 1$, $\mu = 0, \dots, K_u - 1$ and $\theta_i = 2\pi(\omega_i + \eta_i)/K_u$. From Eq. (2.75), we get

$$(n + \mu P) = \sum_{i=1}^M e^{j\mu\theta_i} z_{(i)}(n) + v(n + \mu P) \quad (2.87)$$

$$n = 0, \dots, P - 1, \mu = 0, \dots, K_u - 1.$$

We arrange the $\{r(n)\}_{n=0}^{N-1}$ samples into a $K_u \times P$ matrix

$$\mathbf{R} = \begin{bmatrix} r(0) & \cdots & r(P-1) \\ r(P) & \cdots & r(2P-1) \\ \vdots & \ddots & \vdots \\ r((K_u-1)P) & \cdots & r(N-1) \end{bmatrix}_{K_u \times P}. \quad (2.88)$$

Letting R_p denote the p th column of matrix \mathbf{R} , we have that

$$R_p = \mathbf{G}z_p + v_p, \quad p = 0, \dots, P - 1, \quad (2.89)$$

where $z_p = [z_{(1)}(p) z_{(2)}(p) \cdots z_{(M)}(p)]^T$ and \mathbf{G} is a $(K_u \times M)$ matrix given by

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \cdots & e^{j\theta_M} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(K_u-1)\theta_1} & e^{j(K_u-1)\theta_2} & \cdots & e^{j(K_u-1)\theta_M} \end{bmatrix}_{K_u \times M}.$$

Once we express the received samples as in Eq. (2.87), CFO estimation can be carried out using a signal subspace decomposition approach. The dimension of the null subspace is dictated by the number of null subcarriers, which is equal to $(K_u - M)P$. The main idea behind the low-complexity CFO estimation algorithms studied in [146–148] is to estimate the θ_i s, which are distinct from each other if $|\omega_i| < 0.5$. Indeed, the θ_i s in Eq. (2.87) cause phase shifts to identical P -sample long segments. Hence, the subspace approach can in a way be seen as an extension of the repetitive-slot CFO estimation approach (see e.g., [110, 113]) to the case of multiple CFO estimation. Employing the estimation of signal parameters via rotational invariance technique (ESPRIT) algorithm, the multiple CFOs can be estimated using the following steps:

Step 1. Arrange the received signal $\{r(n)\}_{n=0}^{N-1}$ into matrix \mathbf{R} .

Step 2. The covariance matrix $\Omega = E \left\{ R_p R_p^H \right\}$ of R_p is estimated by

$$\widehat{\Omega} = \frac{1}{P} \mathbf{R} \mathbf{R}^H.$$

Step 3. Compute singular value decomposition (SVD) of $\widehat{\Omega}$:

$$\widehat{\Omega} = [\mathbf{U}_s \mathbf{U}_z] \begin{bmatrix} \Sigma_s & \mathbf{0} \\ \mathbf{0} & \Sigma_z \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_z^H \end{bmatrix}, \quad (2.90)$$

where \mathbf{U}_s is a $K_u \times M$ matrix composed of M eigenvectors corresponding to the M largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ and \mathbf{U}_z is a $K_u \times (K_u - M)$ matrix composed of $K_u - M$ eigenvectors corresponding to the rest eigenvalues $\lambda_{M+1} \geq \dots \geq \lambda_{K_u}$.

Step 4. Let \mathbf{U}_{s1} denote the first $(K_u - 1)$ rows of \mathbf{U}_s and \mathbf{U}_{s2} denote the last $(K_u - 1)$ rows of \mathbf{U}_s . The θ_i 's are estimated as

$$\hat{\theta}_i = \angle(\beta_i), \quad (2.91)$$

where $\angle(\cdot)$ denotes the angle of the complex number and $\{\beta_i\}_{i=0}^{M-1}$ are the eigenvalues of

$$\Xi = (\mathbf{U}_{s1}^H \mathbf{U}_{s1})^{-1} \mathbf{U}_{s1}^H \mathbf{U}_{s2}.$$

Step 5. After estimating the $\hat{\theta}_{i,s}$, the estimate of CFO ω_i can be computed as

$$\hat{\omega}_i = \frac{K_u \hat{\theta}_i}{2\pi} - \eta_i, \quad i = 1, \dots, M. \quad (2.92)$$

Another subspace-based method, the spectral multiple signal classification (MUSIC) algorithm can also be applied to estimate the multiple CFOs by replacing Step 4 above by

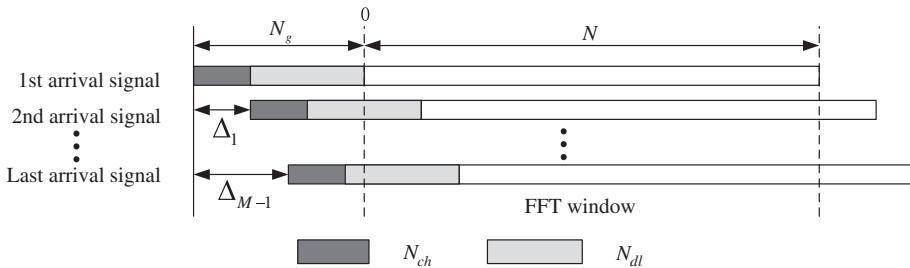
Step 4. Find the M largest peaks of following metric:

$$\Lambda(\tilde{\theta}) = \frac{1}{|a^H(\tilde{\theta}) \mathbf{U}_z \mathbf{U}_z^H a(\tilde{\theta})|^2}, \quad (2.93)$$

where $a(\tilde{\theta}) = [1, e^{j\tilde{\theta}}, \dots, e^{j(K_u-1)\tilde{\theta}}]^T$.

As compared in [148], the ESPRIT based estimation algorithm outperforms the Spectral MUSIC based estimation algorithm at low SNR region. Moreover, ESPRIT algorithm avoids the search operation required by Spectral MUSIC. Thus, the ESPRIT based estimation algorithm is preferred in practice. As shown in [148], the estimation error of CFOs can be reduced by using more than one OFDM block.

To apply the subspace-based CFO estimation algorithm, we have to find the starting point of the receive signal first. As argued in [146], the CFO estimation algorithm is applicable to a quasi-synchronous system. The starting point is determined by the downlink procedure and the effect of

**FIGURE 2.35**

CP for quasi-synchronous uplink OFDMA uplink.

timing offsets due to propagation delay can be removed by introducing a long CP. As shown in Figure 2.35, the CP is composed of two parts $N_g = N_{ch} + N_{dl}$, where N_{ch} is the portion of the CP for accommodating channel delay spreads, while the additional N_{dl} samples are intended for accommodating different timing offsets among users. To completely remove the inter-block interference, a necessary condition is that $N_{ch} \geq L$ and $N_{dl} \geq \Delta_{M-1}$, where L is the maximum channel delay spread and Δ_{M-1} is the maximum timing offset among users. In this case, the extra overhead, i.e., N_{dl} , will be increased by an increasing Δ_{M-1} . To reduce the overhead, accurate knowledge of the timing offset of each user is necessary to align all user signals at BS. As shown in [145, 149], a possible way to estimate the multiple timing offsets is to estimate the timing offsets together with the channel responses. Similar to fine timing estimation in downlink OFDMA, the maximum energy criterion for timing offset estimation can be expressed as:

$$\tilde{\tau}_i = \arg \max_{\tilde{\tau}_i} \left\{ \sum_{l=\tilde{\tau}_i}^{\tilde{\tau}_i+L-1} |\hat{h}'_i(l)|^2 \right\}, \quad (2.94)$$

where $\hat{h}'_i(l)$ is the l th entry of \hat{h}'_i and $h'_i = [0_{\tau_i} h_i 0_{L_{ex}-\tau_i-L}]^T$ is the extended channel vector with length L_{ex} . We can set $L_{ex} = N_g$ for simplicity.

Alternative timing estimation algorithms can be found in [150, 151]. However, the introduced iterative approaches make the algorithms much more complicated compared to the maximum energy criterion discussed above.

2.02.4.4 Synchronization with generalized CAS

As mentioned before, since there is no rigid constraint between subcarriers and users in generalized CAS, this subcarrier allocation scheme is more flexible than the subband or interleaved CAS. The BS can assign the best subcarriers which are currently available to a user according to the users' channel condition. Thus, the generalized CAS can improve the systems performance significantly. On the other hand, lack of constraint among subcarriers makes the synchronization task even more challenging than that of interleaved CAS.

The joint ML estimation of timing and frequency offsets for the generalized CAS was first studied by Morelli [100] and a suboptimal solution was proposed based on repetitively transmitted training

symbols. However, it assumes that only one new user enters the network at each time, which may be too strict in practical applications. Alternative ML-based synchronization schemes for the generalized CAS are described in [149, 152], where iterative alternating projection and space-alternating generalized expectation-maximization (SAGE) algorithms are employed to reduce the complexity of ML estimation, respectively. Similar to the subspace-based algorithm for interleaved CAS, the iterative-based algorithms studied in [149, 152] are only applicable to a quasi-synchronous system.

Under the quasi-synchronous assumption as shown in Figure 2.35, after removing the CP, the IBI free received signal expressed in (2.75) can be re-written in matrix form as

$$r = \sum_{i=1}^M \Gamma(\omega_i) \mathbf{A}_i h'_i + v, \quad (2.95)$$

or equivalently

$$r = \sum_{i=1}^M \Gamma(\omega_i) \mathbf{D}_i(\tau_i) h_i + v, \quad (2.96)$$

where

$$r = [r(0), \dots, r(N-1)]^T, \quad (2.97)$$

$$\Gamma(\omega_i) = \text{diag}\{1, e^{j2\pi\omega_i/N}, \dots, e^{j2\pi\omega_i(N-1)/N}\}, \quad (2.98)$$

$$[\mathbf{A}_i]_{m,n} = [x_i]_{(m-n)\bmod N}, \quad 0 \leq m \leq N-1, \quad n = 0 \leq n \leq N_g - 1, \quad (2.99)$$

$$[\mathbf{D}_i(\tau_i)]_{m,n} = [x_i]_{(m-n-\tau_i)\bmod N}, \quad 0 \leq m \leq N-1, \quad n = 0 \leq n \leq L-1, \quad (2.100)$$

$$h_i = [h_i(0), \dots, h_i(L-1)]^T, \quad (2.101)$$

$$h'_i = \left[0_{\tau_i \times 1}^T h_i^T 0_{(N_g-L-\tau_i) \times 1}^T \right]^T, \quad (2.102)$$

where $[\mathbf{A}_i]_{m,n}$ denotes the (m, n) th entry of matrix \mathbf{A}_i and $[x_i]_m$ represents the m th entry of vector x_i . Rewrite (2.95) as

$$r = \mathbf{B}(\omega) h' + v, \quad (2.103)$$

where $\mathbf{B}(\omega) = [\Gamma(\omega_1) \mathbf{A}_1 \Gamma(\omega_2) \mathbf{A}_2 \cdots \Gamma(\omega_M) \mathbf{A}_M]$ and $h' = [(h'_1)^T \cdots (h'_M)^T]^T$. The log-likelihood function for the frequency offsets ω and extended equivalent channel h' can be expressed as

$$\Lambda(\tilde{\omega}, \tilde{h}') = -N \ln \left(\pi \sigma_v^2 \right) - \frac{1}{\sigma_v^2} \|r - \mathbf{B}(\tilde{\omega}) \tilde{h}'\|^2, \quad (2.104)$$

where $\tilde{\omega}$ and \tilde{h}' are trial values of ω and h' respectively. Thus, the joint ML estimates of ω and h' can be obtained as

$$\hat{\omega} = \arg \max_{\tilde{\omega}} \left\{ \|\Pi_B(\tilde{\omega}) r\|^2 \right\}, \quad (2.105)$$

$$\hat{h}' = \left(\mathbf{B}^H(\hat{\omega}) \mathbf{B}(\hat{\omega}) \right)^{-1} \mathbf{B}^H(\hat{\omega}) r, \quad (2.106)$$

where $\Pi_B(\tilde{\omega}) = \mathbf{B}(\tilde{\omega}) (\mathbf{B}^H(\tilde{\omega})\mathbf{B}(\tilde{\omega}))^{-1} \mathbf{B}^H(\tilde{\omega})$. The maximization in (2.105) requires a grid-search over the multidimensional domain spanned by $\tilde{\omega}$, which is too cumbersome in practice [149]. A simple way to reduce the complexity is to use the iterative alternating projection method studied in [149].

Let $\hat{\omega}_i^{(k)}$ denote the estimate of ω_i at the k th iteration and define the $M - 1$ dimensional vector $\hat{\omega}_i^{(k)}$ as

$$\hat{\omega}_i^{(k)} = [\hat{\omega}_1^{(k+1)}, \dots, \hat{\omega}_{i-1}^{(k+1)}, \hat{\omega}_i^{(k)}, \hat{\omega}_{i+1}^{(k)}, \dots, \hat{\omega}_M^{(k)}]^T. \quad (2.107)$$

At the i th step of the $(k + 1)$ th iteration, the estimate of ω_i is updated by the alternating projection frequency estimator (APFE) as

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\tilde{\omega}_i} \left\{ \left\| \Pi_B \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) \mathbf{r} \right\|^2 \right\}. \quad (2.108)$$

Exploiting the structure of $\mathbf{B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$, the estimator (2.108) can be further simplified as

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\tilde{\omega}_i} \left\{ \left\| \Pi_{C_B} \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) \mathbf{r} \right\|^2 \right\}, \quad (2.109)$$

where

$$\Pi_{C_B} \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) = \mathbf{C}_B \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) \left(\mathbf{C}_B^H \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) \mathbf{C}_B \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) \right)^{-1} \mathbf{C}_B^H \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right), \quad (2.110)$$

$$\mathbf{C}_B^H \left(\tilde{\omega}_i, \hat{\omega}_i^{(k)} \right) = \mathbf{I}_N - \Pi_C \left(\hat{\omega}_i^{(k)} \right) \Gamma(\tilde{\omega}_i) \mathbf{A}_i, \quad (2.111)$$

$$\Pi_C \left(\hat{\omega}_i^{(k)} \right) = \mathbf{C} \left(\hat{\omega}_i^{(k)} \right) \left(\mathbf{C}^H \left(\hat{\omega}_i^{(k)} \right) \mathbf{C} \left(\hat{\omega}_i^{(k)} \right) \right)^{-1} \mathbf{C}^H \left(\hat{\omega}_i^{(k)} \right), \quad (2.112)$$

$$\mathbf{C} \left(\hat{\omega}_i^{(k)} \right) = \left[\Gamma \left(\hat{\omega}_1^{(k+1)} \right) \mathbf{A}_1 \cdots \Gamma \left(\hat{\omega}_{i-1}^{(k+1)} \right) \mathbf{A}_{i-1} \Gamma \left(\hat{\omega}_{i+1}^{(k)} \right) \mathbf{A}_{i+1} \cdots \Gamma \left(\hat{\omega}_M^{(k)} \right) \mathbf{A}_M \right]. \quad (2.113)$$

Computing $\Pi_{C_B}(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$ only requires the inversion of a $N_g \times N_g$ matrix, which is significantly less complex than computing $\Pi_B(\tilde{\omega}_i, \hat{\omega}_i^{(k)})$. From (2.109), we see that the M -D search required by the ML estimator (2.105) is split into a series of 1-D maximization problems, and is thus much more effective than the original maximization problem. After obtaining the frequency estimates $\hat{\omega}$, the ML estimates of timing offsets can be derived from Eq. (2.96) as

$$\hat{\tau} = \arg \max_{\tilde{\tau}} \left\{ \left\| \Pi_\Psi(\hat{\omega}, \tilde{\tau}) \mathbf{r} \right\| \right\}, \quad (2.114)$$

where

$$\Pi_\Psi(\hat{\omega}, \tilde{\tau}) = \Psi(\hat{\omega}, \tilde{\tau}) \left(\Psi^H(\hat{\omega}, \tilde{\tau}) \Psi(\hat{\omega}, \tilde{\tau}) \right)^{-1} \Psi^H(\hat{\omega}, \tilde{\tau}), \quad (2.115)$$

$$\Psi(\hat{\omega}, \tilde{\tau}) = [\Gamma(\hat{\omega}_1) \mathbf{D}_1(\tilde{\tau}_1) \cdots \Gamma(\hat{\omega}_M) \mathbf{D}_M(\tilde{\tau}_M)]. \quad (2.116)$$

Similar to the problem in (2.105), the maximization problem in (2.114) can be efficiently solved by resorting to iterative alternating projection methods; the resulting estimator is referred to as alternating projection timing estimator (APTE) in [149]. Since the timing and frequency estimators introduced above are iterative, initial estimates of ω and τ , referred to as $\hat{\omega}^{(0)}$ and $\hat{\tau}^{(0)}$ respectively, are required. A simple way to initialise the estimates of CFOs is to use the expected value of ω_i , i.e., $\hat{\omega}_i^{(0)} = 0$. Alternatively, $\hat{\omega}^{(0)}$ can be taken as the output of the frequency estimator proposed in [153]. The initial estimates $\hat{\tau}^{(0)}$ can be obtained by first estimating h'_i according to Eq. (2.106) and then exploiting the specific structure of \hat{h}'_i . The index of the first significant element of \hat{h}'_i is taken as $\hat{\tau}_i^{(0)}$.

Another alternative approach to avoid the multidimensional search in ML based estimator is employing the SAGE algorithm as in [151, 152]. Since the estimates of timing offsets can be obtained via (2.94), we assume that the timing offsets are estimated first for SAGE algorithm. From Eq. (2.96), the i th cycle of the k th iteration of the SAGE algorithm can be performed as follows:

1. *Expectation step:* Compute

$$y_i^{(k)} = r - \sum_{j=1}^{i-1} \hat{r}_j^{(k)} - \sum_{j=i+1}^M \hat{r}_j^{(k-1)}. \quad (2.117)$$

2. *Maximization step:* The likelihood function for the unknown frequency offset ω_i and channel h_i can be expressed as

$$\Lambda(\tilde{\omega}_i, \tilde{h}_i) = -N \ln \left(\pi \sigma_v^2 \right) - \frac{1}{\sigma_v^2} \| y_i^{(k)} - \Gamma(\tilde{\omega}_i) \mathbf{D}_i(\hat{\tau}_i) \tilde{h}_i \|^2. \quad (2.118)$$

Thus, the joint ML estimates of ω_i and h_i can be written as

$$\hat{\omega}_i^{(k)} = \arg \max_{\tilde{\omega}_i} \left\{ \| \Pi_W(\tilde{\omega}_i) y_i^{(k)} \|^2 \right\}, \quad (2.119)$$

$$\hat{h}_i^{(k)} = \left(\mathbf{D}_i^H(\hat{\tau}_i) \mathbf{D}_i(\hat{\tau}_i) \right)^{-1} \mathbf{D}_i^H(\hat{\tau}_i) \Gamma^H(\hat{\omega}_i^{(k)}) y_i^{(k)}, \quad (2.120)$$

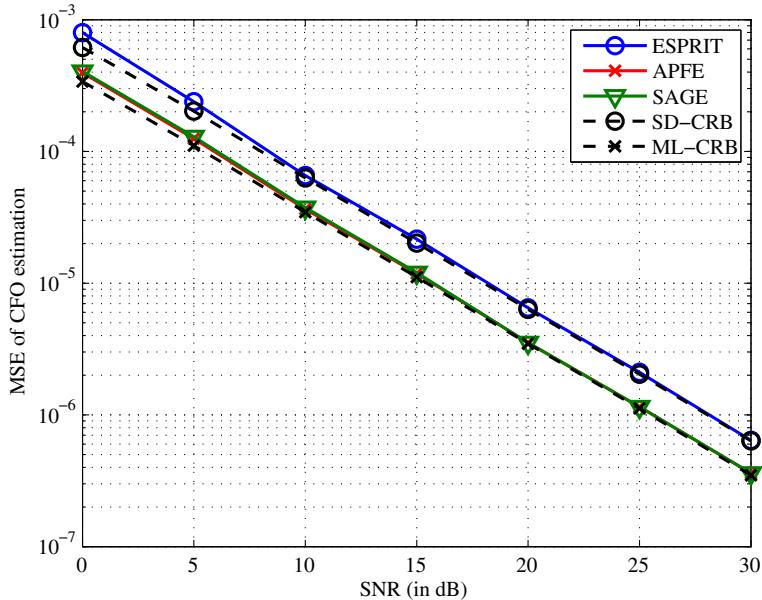
where $\Pi_W = \Gamma(\tilde{\omega}_i) \mathbf{D}_i(\hat{\tau}_i) \left(\mathbf{D}_i^H(\hat{\tau}_i) \mathbf{D}_i(\hat{\tau}_i) \right)^{-1} \mathbf{D}_i^H(\hat{\tau}_i) \Gamma^H(\tilde{\omega}_i)$. After obtaining the estimates $\hat{\omega}_i^{(k)}$ and $\hat{h}_i^{(k)}$ which is utilized in the Expectation step of the next cycle or iteration can be updated as

$$\hat{r}_i^{(k)} = \Gamma \left(\hat{\omega}_i^{(k)} \right) \mathbf{D}_i(\hat{\tau}_i) \hat{h}_i^{(k)}. \quad (2.121)$$

Again, the initial estimates of CFOs for the SAGE algorithm can be obtained via iterative alternating projection methods. Moreover, we can see from Eq. (2.121) that initial channel estimates are required for the SAGE algorithm while no such requirement is needed for the iterative alternating projection methods. Inaccurate channel estimates will deteriorate the SAGE performance significantly. With the aim of obtaining a tradeoff between performance and complexity, several non-ML based multiple CFOs estimators were proposed in [154, 155].

2.02.4.5 Simulation results for uplink OFDMA systems

The synchronization of uplink OFDMA systems depends on the subcarrier allocation schemes. Here, we compare the synchronization performance of subspace based and ML based methods, based on

**FIGURE 2.36**

Mean square error of CFO estimates for uplink OFDMA without timing offsets.

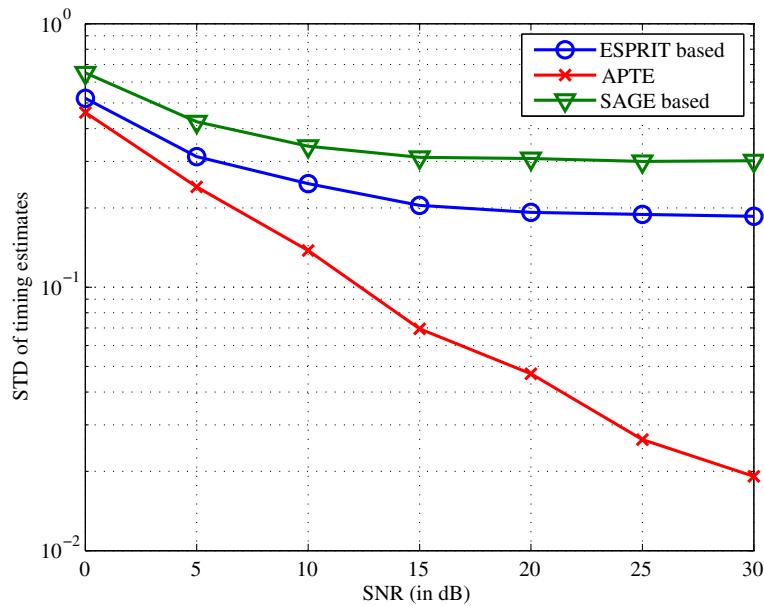
interleaved CAS. The total number of subcarriers is set to 512 with a CP of length 32, i.e., $N = 512$ and $N_g = 32$. The maximum and active numbers of users are set to 4 and 3, respectively, i.e., $K_u = 4$ and $M = 3$. The channel with the exponential power delay profile, introduced previously, is considered. The maximum channel delay spread L is equal to 16 and the channels for different users are assumed uncorrelated. We set the CFOs of the three users to -0.1 , -0.2 and 0.3 , respectively.

Figure 2.36 shows the CFO estimation performance without timing offsets. From Eq. (2.95), the CRB of ML based CFO estimation algorithm is given in [149] as

$$\text{var}\{\hat{\omega}_i\} \geq \text{CRB}\hat{\omega}_i = \frac{N^2\sigma_v^2}{8\pi^2} \left[\left(\Re \left\{ \Psi^H \Pi_{\mathbf{B}}^\perp \Psi \right\} \right) \right]_{i,i}, \quad (2.122)$$

where $\Psi = [\Psi_1, \dots, \Psi_M]$, $\Psi_i = \mathbf{W}\Gamma(\omega_i)\mathbf{A}_i h'_i$ with $\mathbf{W} = \text{diag}\{0, 1, \dots, N - 1\}$; $\Pi_{\mathbf{B}}^\perp = \mathbf{I} - \Pi_{\mathbf{B}}(\omega)$. The CRB corresponding to the subspace decomposition based CFO estimation algorithm can be found in [95]. From Figure 2.36, we can find that ML based algorithms, i.e., APFE and SAGE, have identical performances and provide nearly 3-dB gains over ESPRIT algorithm. This can be ascribed to the fact that subspace decomposition based CFO estimation is actually a non-data aided method. Both ML and subspace decomposition based CFO estimation algorithms can achieve their ‘‘corresponding CRBs.’’

Under the quasi-synchronous assumption, we assume that the first arriving signal is known perfectly by the BS and the timing offsets of other two user’s signals are normally distributed in an interval $[0, 16]$. Figure 2.37 shows the timing estimation performance of ESPRIT and SAGE based estimators,

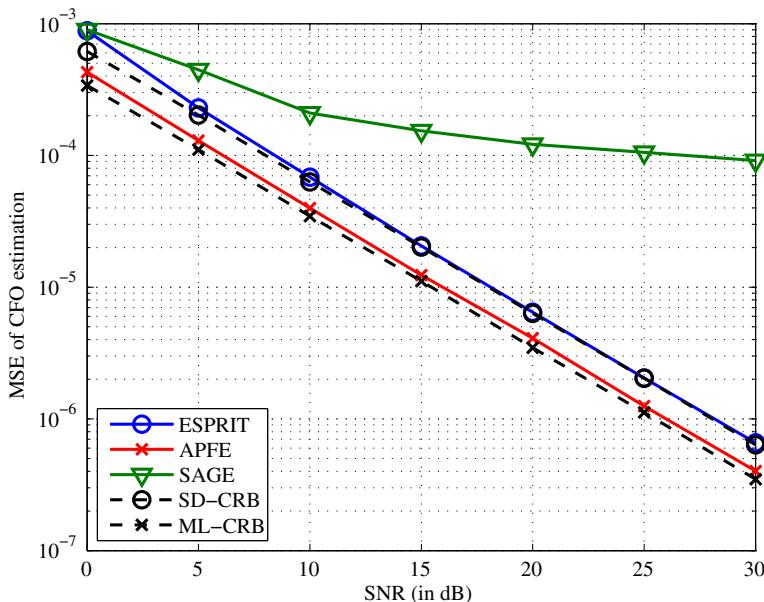
**FIGURE 2.37**

Standard deviation of timing estimates for uplink OFDMA.

and the ML estimator (2.114), i.e., APTE. As we can see from the figure, the ML estimator outperforms the maximum channel energy based estimators significantly at the price of increased computational complexity. Since APFE and ESPRIT based CFO estimators can work without timing offset information, we estimate CFO prior to timing offset estimation. On the other hand, timing offsets information is crucial to the SAGE based CFO estimator. Thus we perform timing offset estimation before SAGE based CFO estimation. From Figure 2.37, we see that ESPRIT based timing estimator outperforms SAGE. This can be ascribed to the fact that employing the CFO estimates obtained in the ESPRIT based CFO estimator improves timing estimation. Figure 2.38 shows the CFO estimation performance with timing offsets estimation. We see that the performance of SAGE based CFO estimator is affected significantly by the inaccuracy of timing estimation. On the other hand, both ESPRIT and APFE estimators are robust to timing inaccuracies.

2.02.5 Network synchronization

In this section, we discuss the challenges in the network time synchronization problem and the performance metrics of interest. We provide a brief overview of current devices and their limitations, and describe common clock models. We then describe a taxonomy of network time sync protocols and provide some examples.

**FIGURE 2.38**

Mean square error of CFO estimates for uplink OFDMA with timing offsets.

As stated in the Introduction to this chapter, network synchronization is a well-studied topic with an extensive history, especially for wired networks, e.g., see [156, 157], and references therein. Typically, these works assumed high quality devices, availability of fine control of the network, extensive connectivity with little or no mutual interference, as well as often assuming known (or repeatable and measurable) propagation and processing delays [1]. Surveys of WSN sync protocols may be found in the papers by Sivrikaya and Yener [2], Johannessen [3], Sundararaman et al. [1, 4–6].

The network synchronization problem is to ensure that all nodes in the network operate on a common clock, i.e., have a common time reference.

Challenges in synchronization in WSN stem from several sources, broadly related to the transmitter, the propagation channel, and the receiver:

1. Channel conditions (such as fading, shadowing, interference) lead to time-varying connectivity even for static nodes: scatterers move in any case; mobility adds its own challenges.
2. The devices are cheap and clocks drift, often erratically, due to fluctuations in ambient temperature, and with age. The time difference between two clocks may be fixed (a fixed offset) or may vary with time (due to clock oscillator frequency drifts).
3. Queuing and processing delays are variable (thus rendering it difficult to use standard protocols such as NTP). There is variability in the time it takes for a packet to go from the application layer to the MAC layer, variable delays within the MAC layer (the major source of error), in packet generation

and transmission at the PHY layer. There are similar variations at the receiver, including inaccuracy in detecting packet arrival.

4. Variabilities in propagation time due to non-line of sight issues, and non-reciprocity of the channel. Typically, the propagation time is negligible compared with the queueing and processing delays.
5. Communication rates are variable in a large network; if a node is involved infrequently in regular communications, then “heartbeat” signals may be essential to keep the node in quasi-sync with the rest of the network, and thus connected to the network.
6. Protocols must be scalable to a large number of nodes, and must deal with heterogeneity of nodes.
7. These devices are often battery powered so that energy is a finite resource, and energy consumption directly affects node lifetime.
8. Given energy constraints, WSN nodes must exploit external assets which may have more relaxed energy constraints; e.g., basestations, unmanned aerial vehicles (UAVs), various broadcast beacons.

Metrics: How should one evaluate the performance of a network sync protocol? When GPS is available, a reasonable metric could be the bias compared to the “true” time. But access to GPS can be difficult, particularly indoors, under forest canopy, and in other challenging conditions. Often, it suffices that the nodes converge on some common time reference (regardless of whether it is “true”). Some commonly used metrics are:

- *Synchronization accuracy:* Worst case (or average case) pair wise error between any one-hop neighbors
- *Energy efficiency:* The number of packet transmissions and receptions necessary to achieve sync, and the rate and frequency of messages that need to be exchanged to maintain sync.
- *Synchronization convergence time:* The time taken for all nodes (or a given percentage of nodes) to be in sync with one-hop neighbors.
- *Fault tolerance:* The robustness of sync schemes under (intermittent) failure of (critical) nodes and/or links; robustness to (slow) time variations in clock parameters and clock jitter.
- *Scalability with network size:* Does the sync-error increase with size? Does convergence time increase (only) with the diameter of the network? Or other aspects of topology such as degree distribution?
- *Impact of stochastic channel conditions:* How well does the protocol perform in the presence of stochastic channel conditions (congestion, mobility, duty cycling, queueing delays, propagation delays, processing times)?
- *Engineering design:* Is the protocol simple versus complex?

From the above questions, it is clear that a given protocol offers a set of alternatives in this rich tradespace.

2.02.5.1 Clock models

Let $T_k(t)$ denote the local time at node k , where t denotes the “true” time; often, we will drop the node index k . Formally, *clock drift* is defined as

$$\rho(t) = \frac{dT(t)}{dt} - 1.$$

Table 2.2 Comparisons of Clock Characteristics

	Accuracy (PPM)	Power	Lifetime in hours AA battery
GPS	10^{-8} – 10^{-11}	180 mW	16.7 h
Chip-Scale Atomic Clock	10^{-11}	30 mW	100 h
MCXO	3×10^{-8}	75 mW	40 h
TCXO	6×10^{-6}	6 mW	21 days
Watch clock	200×10^{-6}	1 μW	342 years

A reasonable assumption is that the drifts are bounded, and the clocks do not run backwards, which translate to

$$|\rho(t)| < \rho_{\max}; \rho(t) > -1.$$

A Taylor expansion of the local clock time $T(t)$ wrt the global clock t yields

$$T_k(t) = \alpha_k + \beta_k t + \gamma_k t^2 + \dots, \quad (2.123)$$

wherein α is the *offset* and β the *skew*. The quadratic term, denoted by γ , is typically used only to test for departures from the linear model. Skew has been modeled as an AR process in [158] (see Table 2.2).

2.02.5.2 Net sync protocols

With the above background, we can broadly classify net sync protocols into four broad categories:

1. Broadcast protocols: Based on the notion of broadcast, possibly over a hierarchical tree topology.
2. Distributed synchronization: builds consensus on clock parameters in a peer-to-peer setting.
3. Unilateral sync to an external (broadcast) reference clock.

Other classifications are possible depending upon the viewpoint: e.g., client (sensor node) initiated versus server (gateway) initiated.

Many WSN synchronization protocols have been proposed. Popular ones include the Reference Broadcast Systems (RBS) [159], a time-stamp transformation approach based on bounded offsets [160], the Tiny/Mini-Sync (TMS) protocol [161], Timing-sync Protocol for Sensor Networks (TPSN) [162], Lightweight Tree-based Synchronization (LTS) [163], and the networked Control Time Protocol (CTP) [164]. Probabilistic approaches were considered in [165], and refined and extended in [166]. Bounds under various assumptions are derived in [167–169] and citeNoh-Serpedin.

Among the above protocols, several primary themes emerge. One natural and common notion is the use of time-stamps: time-stamp a packet with the transmitter's clock, time-stamp the reception time, use these stamps to estimate the round trip time (RTT), which is then used to synchronize the two clocks. RTT is often highly variable, and often has a heavy-tailed distribution, which naturally calls for the use of robust estimation techniques. Reliability increases as the number of such exchanges increases, but with a concomitant increase in delay and energy, and sometimes (more than linear) complexity.

A second recurring theme is that the estimation of relative clock offset and skew can be cast as a linear estimation problem.

2.02.5.2.1 Unilateral synchronization

Assume that an external source broadcasts time-stamped messages at “true” times $T(i)$, which are received by a node at $R(i)$ on its local clock. Then from the clock model considered earlier in (2.123), we have

$$R(i) = \alpha + \beta T(i) + \varepsilon(i), \quad i = 1, \dots, I, \quad (2.124)$$

where n is the number of observations. Here $\varepsilon(i)$ represents the modeling errors. Let $\mathbf{1}$ be $I \times 1$ vector of ones, $\mathbf{t} = [T(1), \dots, T(I)]'$, $\mathbf{r} = [R(1), \dots, R(I)]'$, $\mathbf{Z} = [\mathbf{1}, \mathbf{t}]$. Then linear regression yields

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (\mathbf{Z}' \mathbf{Z}^{-1} \mathbf{Z}' \mathbf{r})$$

which is the best linear estimate if the $\varepsilon(i)$ are zero-mean. Note that in this unilateral scheme, one cannot account separately for the propagation delay, or a non-zero-mean ε : both are absorbed into the offset. Thus clocks that are at relatively different distances from the broadcast source will not be in sync with each other.

2.02.5.2.2 Pairwise synchronization

As in the classical Network Time Protocol (NTP) [170], clock offset can be estimated by exchanging time-stamped messages and computing the round trip time (RTT). If the queuing delays are exponentially distributed with the same mean delay then the MLE of the offset is given by the minimum of the observed delays [171]; if the mean delays vary from node to node, then the bootstrap-bias correction method of [172] can be used. Further details may be found in [173].

In these pair-wise protocols, a node “B” synchronizes with a node “A” which is treated as the reference node. Let α and β denote the relative offset and skew of node B wrt node A. Nodes exchange multiple time-stamped messages as follows. Nodes A sends a time-stamped message to node B at its local time $T_{A,k}$, where k is the round number. Node B receives it at its local time $R_{B,k}$. At $T_{B,k}$, it sends a time-stamped message which includes $R_{B,k}$; this is received by node A at $R_{A,k}$. K such rounds of message exchange take place. Let δ denote the fixed portion of the delay and $\varepsilon_{AB,k}$, $\varepsilon_{BA,k}$, the variable portion; delay here includes propagation delay, processing delay and queueing delay. The time-stamps are related via:

$$R_{B,k} = (T_{A,k} + \delta + \varepsilon_{AB,k})\beta + \alpha, \quad (2.125)$$

$$T_{B,k} = (R_{A,k} - \delta - \varepsilon_{BA,k})\beta + \alpha. \quad (2.126)$$

Assuming that the delays are independent and exponentially distributed, it is shown in [171, 173] that the MLE of α is given by

$$\hat{\alpha} = \frac{1}{2} \left[\min_k (R_{B,k} - T_{A,k}) - \min_k (R_{A,k} - T_{B,k}) \right].$$

Let $i = \arg \min_k (R_{A,k} - T_{A,k})$ and $j = \arg \min_{k,k \neq i} (R_{A,k} - T_{A,k})$, which are the first two order statistics of $R_{A,k} - T_{A,k}$. Then the proposed estimator of the skew in [173] is

$$\hat{\beta} = \frac{R_{B,i} + T_{B,i} - R_{B,j} - T_{B,j}}{R_{A,i} + T_{A,i} - R_{A,j} - T_{A,j}}.$$

Once α, β have been estimated, it is easy to estimate the propagation delay as well. These algorithms have been shown to be robust to other delay models.

In the RBS protocol [159], a beacon node transmits a reference packet (as above), but the K receiving nodes exchange time-of-receipt to estimate offsets and skews. Consider (2.124), now indexed by the receiving node's id:

$$R_i(n) = \alpha_i + \beta_i T(n) + \varepsilon_i(n). \quad (2.127)$$

Defining $\Delta_{ij} := \alpha_i - \beta_{ij}\alpha_j$ and $\beta_{ij} := \frac{\beta_i}{\beta_j}$, as the relative offset and skew, one can eliminate $T(n)$ to obtain

$$R_i(n) = \beta_{ij} R_j(n) + \Delta_{ij} + \varepsilon_{ij}(n).$$

Notes i and j can estimate the relative skew and offset via linear regression.

The tree-based sync protocol in [162] is similar to the above algorithms in the estimation part. Here a root node broadcasts a beacon. Its set of 1-hop neighbors (i.e., those who hear the root directly) are called level 1 nodes. Level 1 nodes, in turn, relay the beacon to level-2 nodes, and so on. Nodes at level i synchronize to a parent node at level $i - 1$. The relative skew is assumed to be unity, so the focus is on estimating offsets. Consider a pair of parent-child nodes; one can write the packet reception times as

$$\begin{aligned} R_{B,k} &= T_{A,k} + \alpha + \delta + \varepsilon_{AB,k}, \\ R_{A,k} &= T_{B,k} - \alpha + \delta + \varepsilon_{BA,k}. \end{aligned}$$

After K messages, one can estimate

$$\begin{aligned} \hat{\alpha} &= \frac{1}{2K} \sum_{k=1}^K (R_{B,k} - R_{A,k} + T_{B,k} - T_{A,k}), \\ \hat{\delta} &= \frac{1}{2K} \sum_{k=1}^K (R_{B,k} + R_{A,k} - T_{B,k} - T_{A,k}). \end{aligned}$$

Under the assumption that the noise terms, the ε s are zero-mean and independent, one obtains the following equation for the variance of the estimators:

$$\text{var}(\hat{\beta}) = \text{var}(\hat{\delta}) = \frac{\sigma_{12}^2 + \sigma_{21}^2}{K}.$$

The TPSN protocol is easy to implement. However, it assumes unit skew. Tine-Sync and Mini-Sync are variations that can cope with skew [161].

Another variation was proposed in [1]. The basic idea is that if node A transmits a pair of time-stamped messages ρ units apart on its own clock, then node B should be able to estimate the relative skew.

2.02.5.3 Distributed clock sync

The notion of distributed consensus, i.e., distributed agreement on a parameter (vector) via repeated exchange of messages has recently become popular, and several protocols have been proposed. Here, the idea is that there is no single root node and hence no single point of failure. These protocols seek to achieve global consensus via local exchange of information. They rely upon the broadcast nature of the wireless slink.

2.02.5.3.1 Bio-inspired approaches

The seminal work of Miroollo and Strogatz [174] introduced the basic idea that a population of “integrate-and-fire” oscillators will start firing simultaneously after a finite time starting from all almost any initial condition. Empirical evidence for this is the well-cited firefly sync [8] and circadian sync [9]. The work was extended to multi-hop and time-varying topologies in [175] who made use of results from Algebraic Graph theory. These results have been recently refined, corrected and extended by [176]. The Miroollo-Strogatz model has been exploited in [177] for network time sync. Another consensus-based approach is the so-called diffusion-based approach of [178].

2.02.5.3.2 Consensus-type approaches

We consider next one example of a consensus-based approach [179]. This scheme assumes the presence of a reference node (i.e., a root node whose clock is assumed to be correct, and which does not update its clock). Multiple consistent reference nodes may be present as well. For simplicity assume that all nodes have unit skew, and node i has offset α_i wrt the reference node. Using one of the pairwise schemes described earlier, nodes can obtain an estimate of the relative offset

$$y_{i,j} = \alpha_i - \alpha_j + \varepsilon_{i,j},$$

where $\varepsilon_{i,j}$ is the measurement noise which we model as zero-mean and with variance $\sigma_{i,j}^2$. A node i would have such an estimate for each $j \in \mathcal{N}_i$, the set of neighbors that it can hear. At every iteration, a node receives the current estimates of its neighbors which are used to update its own estimate. Node i 's estimate of its offset α_i in the k th iteration is given by

$$\hat{\alpha}_i(k) = \left(\sum_{j \in \mathcal{N}_i} \frac{1}{\sigma_{i,j}^2} \right)^{-1} \sum_{j \in \mathcal{N}_i} \frac{1}{\sigma_{i,j}^2} (\hat{\alpha}_j(k-1) + y_{ij}).$$

The convergence of this algorithm was studied in [179].

Convergence depends upon the topology, the coupling and the measurement statistics. We can associate a graph $G = (V, E)$ with this problem; here V is the set of nodes, and E the set of edges representing one-hop links. The energy (convergence time) spent for achieving consensus is proportional to $1/\lambda_2(L)$ where L is the graph Laplacian and $\lambda_2(\cdot)$ is the second-largest eigenvalue. The algebraic network connectivity (the so-called Fiedler value), $\lambda_2(L)$ can be increased by increasing the transmit power, i.e., making the graph more connected. However, this also increases the total energy consumption, since the total power consumed by the network is proportional to $p_T/\lambda_2(L(p_T))$, where $L(p_T)$ is the Laplacian of the graph corresponding to transmit power p_T . A question then is: does a global trade-off exist between local transmit power, convergence rate and network topology? Analysis in [180] indicates that when the path loss is high, the optimal topology tends to be sparse, with few connections.

2.02.6 Conclusion

Synchronization is a critical element of ensuring performance of communication systems, such as cellular networks, wireless sensor networks and mobile ad hoc networks, networked signal processing, controls and robotic systems.

In this chapter, our focus was largely on time and frequency synchronization. We first reviewed point-to-point synchronization over a flat-fading link. We studied both the data-aided as well as blind scenarios; we developed the maximum-likelihood estimator, derived the Cramer-Rao bound; investigated sub-optimal estimators and characterized their performance. We then studied the synchronization problem in the context of a frequency-selective channel, focusing on OFDMA systems. We reviewed several estimators and compared their performance with the Cramer-Rao bound. We then turned our attention to multi-user synchronization, again developing estimators and establishing performance bounds. The last section provided a review of the network time synchronization problem, commonly used protocols and performance metrics. We reviewed both broadcast protocols as well as distributed schemes.

Relevant Theory: Signal Processing Theory and Statistical Signal Processing

See [Vol. 1, Chapter 3](#) Discrete-Time Signals and Systems

See [Vol. 1, Chapter 4](#) Random Signals and Stochastic Processes

See [Vol. 1, Chapter 11](#) Parametric Estimation

See [Vol. 3, Chapter 4](#) Bayesian Computational Methods in Signal Processing

See [Vol. 3, Chapter 8](#) Performance Analysis and Bounds

References

- [1] B. Sadler, A. Swami, Synchronization in sensor networks: an overview, in: IEEE MILCOM 2006, 2006.
- [2] F. Sivrikaya, B. Yener, Time synchronization in sensor networks: a survey, IEEE Network 18 (4) (2004) 45–50.
- [3] S. Johannessen, Time synchronization in a local area network, IEEE Control Syst. Mag. 24 (2) (2004) 61–69.
- [4] B. Sundararaman, U. Buy, A. Kshemkalyani, Clock Synchronization for Wireless Sensor Networks: A Survey, Ad Hoc Networks, Elsevier, vol. 3 (3) 2005, pp. 281–323.
- [5] Y.R. Faizulkhakov, Time synchronization methods for wireless sensor networks: a survey, Program. Comput. Soft. 33 (4) (2007) 214–226 (SpringerLink).
- [6] Y.C. Wu, Q. Chaudhari, E. Serpedin, Clock synchronization in wireless sensor networks: an overview, IEEE Signal Process. Mag. 28 (1) 2011.
- [7] P. Galison, Einstein’s Clocks, Poincaré’s Maps: Empires of Time, Norton, 2003.
- [8] S. Strogatz, SYNC: The Emerging Science of Spontaneous Order, Hyperion, Hyperion Press, 2003.
- [9] Y. Wang, F. J. Doyle III, On influences of global and local cues on the rate of synchronization of oscillator networks, Automatica 47 (6) (2011) 1236–1242.
- [10] A. Olshevsky, J. Tsitsiklis, Convergence speed in distributed consensus and averaging, SIAM Review 53 (4) (2011) 747–772.
- [11] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, J. Tsitsiklis, Convergence in multiagent coordination, consensus, and flocking, in: Proceedings of the Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC’05), Seville, Spain, 2005.

- [12] J. Proakis, Digital Communications, Mc-Graw-Hill, 2001.
- [13] D. Tse, P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.
- [14] S. Kay, Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory, Prentice Hall, 1993.
- [15] E. Tozer, Broadcast Engineer's Reference Book, Focal Press, 2004.
- [16] S. Savory, Digital filters for coherent optical receivers, *Opt. Express* 16 (2) (2008) 804–817.
- [17] P. Ciblat, P. Loubaton, E. Serpedin, G. Giannakis, Asymptotic analysis of blind cyclic correlation based symbol rate estimation, *IEEE Trans. Inform. Theory* 48 (7) (2002) 1922–1934.
- [18] U. Mengali, A. d'Andrea, Synchronization Techniques for Digital Receivers, Plenum Press, 1997.
- [19] H. Meyr, M. Moeneclaey, S. Fechtel, Digital Communications Receivers: Synchronization, Channel Estimation, and Signal Processing, J. Wiley & Sons, 1998.
- [20] C. Georghiades, E. Serpedin, Synchronization, The Handbook of Communications, CRC Press, 2002.
- [21] B. Porat, Digital Signal Processing of Random Signals, Prentice Hall, 1994.
- [22] A. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Trans. Inform. Theory* (1967).
- [23] G. Forney, The viterbi algorithm, *Proc. IEEE*, 1973.
- [24] M. Morelli, U. Mengali, Carrier frequency estimation for transmissions over selective channels, *IEEE Trans. Commun.* 48 (9) (2000) 1580–1589.
- [25] P. Ciblat, L. Vandendorpe, On the maximum-likelihood based data-aided frequency offset and channel estimates, in: EURASIP European Signal Processing Conference (EUSIPCO), vol. 1, Toulouse, France, September 2002, pp. 627–630.
- [26] B. Hassibi, B. Hochwald, How much training is needed in multiple-antenna wireless links? *IEEE Trans. Inform. Theory* 49 (4) (2003) 951–963.
- [27] H.V. Trees, Detection, Estimation, and Modulation Theory, J. Wiley & Sons, 2003.
- [28] S. Crozier, D. Falconer, S. Mahmoud, Least sum of squared errors channel estimation, *IEE Proc. F Radar Signal Process.* 138 (4) (1991) 371–378.
- [29] C. Tellambura, M. Parker, Y. Guo, S. Sheperd, S. Barton, Optimal sequences for channel estimation using discrete Fourier transform techniques, *IEEE Trans. Commun.* 47 (2) (1999) 230–237.
- [30] A. Benveniste, M. Métivier, P. Priouret, Adaptive Algorithms and Stochastic Approximations, Springer-Verlag, 1990.
- [31] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, L. Vandendorpe, Code-aided turbo synchronization, *Proc. IEEE* 95 (6) (2007) 1255–1271.
- [32] W. Lindsey, M. Simon, Telecommunications Systems Engineering, Prentice Hall, Englewood Cliffs, 1972.
- [33] F. Gardner, Demodulator Reference recovery techniques suitable for digital implementation, ESTEC Contrat No. 6847/86/NL/DG, European Space Agency, 1988.
- [34] K. Mueller, M. Mueller, Timing recovery in digital synchronous data receivers, *IEEE Trans. Commun.* 24 (5) (1976) 516–531.
- [35] D. Rife, R. Boorstyn, Single-tone parameter estimation from discrete-time observations, *IEEE Trans. Inform. Theory* 20 (5) (1974) 591–598.
- [36] R.M. Aulay, L. Seidman, A useful form of Barankin lower bound and its application to PPM threshold analysis, *IEEE Trans. Inform. Theory* 15 (2) (1969) 273–279.
- [37] L. Atallah, J.-P. Barbot, P. Larzabal, SNR threshold indicators in data-aided phase and frequency synchronization for QAM, *IEEE Signal Process. Lett.* 11 (8) (2004) 652–654.
- [38] P. Ciblat, M. Ghogho, Blind NLLS carrier frequency-offset estimation for QAM, PSK, and PAM modulations: performance at low SNR, *IEEE Trans. Commun.* 54 (10) (2006) 1725–1730.
- [39] E. Barankin, Locally best unbiased estimates, *Ann. Math. Statist.* 20 (1949) 447–501.
- [40] L. Knockaert, The Barankin bound and threshold behavior in frequency estimation, *IEEE Trans. Signal Process.* 45 (1997) 2398–2401.

- [41] R.M. Aulay, E. Hofstetter, Barankin bounds on parameter estimation, *IEEE Trans. Inform. Theory* 17 (6) (1971) 669–676.
- [42] I. Reuven, H. Messer, A Barankin-type lower bound on the estimation error of hybrid parameter vector, *IEEE Trans. Inform. Theory* 43 (3) (1997) 1084–1093.
- [43] I. Reuven, H. Messer, On the effect of nuisance parameters on the threshold SNR value of the Barankin bound, *IEEE Trans. Signal Process.* 47 (1999) 523–527.
- [44] P. Forster, P. Larzabal, On lower bounds for deterministic parameter estimation, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, Orlando, FL, May 2002, pp. 1137–1140.
- [45] J. Ziv, M. Zakai, Some lower bounds on Signal Processing estimation, *IEEE Trans. Inform. Theory* 15 (3) (1969) 386–391.
- [46] K. Bell, Y. Steinberg, Y. Ephraim, H.V. Trees, Extended Ziv-Zakai lower bound for vector parameter estimation, *IEEE Trans. Inform. Theory* 43 (2) (1997) 624–637.
- [47] A. Weiss, E. Weinstein, Fundamental limitations in passive time-delay estimation—Part I: Narrow-band systems, *IEEE Trans. Acoust. Speech Signal Process.* 31 (2) (1983) 472–486.
- [48] P. Ciblat, M. Ghogho, Ziv-zakai bound for harmonic retrieval in multiplicative and additive Gaussian noise, in: *IEEE Workshop on Statistical Signal Processing (SSP)*, Bordeaux, France, July 2005.
- [49] C. Georghiades, D. Snyder, The expectation-maximization algorithm for symbol unsynchronized sequence detection, *IEEE Trans. Commun.* 39 (1) (1991) 54–61.
- [50] E. Panagirci, C. Georghiades, Joint ML timing and phase estimation in OFDM systems using the EM algorithm, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, July 2000, pp. 2949–2952.
- [51] M. Moeneclaey, G. de Jonghe, ML-oriented nda carrier synchronization for general rotationally symmetric signal constellations, *IEEE Trans. Commun.* 42 (8) (1994) 2531–2533.
- [52] M. Morelli, U. Mengali, Feedforward frequency estimation for PSK: a tutorial review, *Eur. Trans. Telecommun.* 9 (2) (1998) 103–115.
- [53] J. Costas, Synchronous communications, *Proc. IRE* 44 (12) (1956) 1713–1718.
- [54] D. Taylor, Introduction TP synchronous communications, *Proc. IEEE* 90 (8) (2002) 1459–1460.
- [55] W. Gardner, A. Napolitano, L. Paura, Cyclostationarity: half a century of research, *Signal Process.* 86 (4) (2006) 639–697.
- [56] G.B. Giannakis, F. Panduru, E. Serpedin, I. Sari, A bibliography on cyclostationary signal processing, *Signal Process.* 85 (12) (2005) 2233–2303.
- [57] B. Picinbono, P. Chevalier, Widely linear estimation with complex data, *IEEE Trans. Signal Process.* 43 (8) (1995) 2030–2033.
- [58] F. Gini, G. Giannakis, Frequency offset and symbol timing recovery in flat-fading channels: a cyclostationary approach, *IEEE Trans. Commun.* 46 (3) (1998) 400–411.
- [59] Y. Wang, P. Ciblat, E. Serpedin, P. Loubaton, Performance analysis of a class of non-data aided carrier frequency offset and symbol timing delay estimators for flat-fading channels, *IEEE Trans. Signal Process.* 50 (9) (2002) 2295–2305.
- [60] A. Dandawate, G. Giannakis, Asymptotic theory of mixed time averages and k th-order cyclic moment and cumulant statistics, *IEEE Trans. Inform. Theory* 41 (1) (1995) 216–232.
- [61] G. Zhou, G. Giannakis, Harmonics in multiplicative and additive noise: performance analysis of cyclic estimators, *IEEE Trans. Signal Process.* 43 (6) (1995) 1445–1460.
- [62] M. Oerder, H. Meyr, Digital filter ans square timing recovery, *IEEE Trans. Commun.* 36 (5) (1988) 605–612.
- [63] K. Scott, E. Olasz, Simultaneous clock phase and frequency offset estimation, *IEEE Trans. Commun.* 43 (7) (1995) 2263–2270.

- [64] M. Ghogho, A. Swami, T. Durrani, On blind carrier recovery in time-selective fading channel, in: Asilomar Conference on Signals, Systems and Computers, vol. 1, 1999, pp. 243–247.
- [65] Y. Wang, E. Serpedin, P. Ciblat, Blind feedforward cyclostationarity-based timing estimation for linear modulations, *IEEE Trans. Wireless Commun.* 3 (3) (2004) 709–715.
- [66] J. Delmas, Asymptotic performance of second-order algorithms, *IEEE Trans. Signal Process.* 50 (1) (2002) 49–57.
- [67] B. Ottersten, P. Stoica, R. Roy, Covariance matching estimation techniques for array signal processing applications, *Digital Signal Process.* 8 (1998) 185–210.
- [68] E. Hannan, Non linear time series regression, *J. Appl. Probab.* 8 (1972) 767–780.
- [69] E. Hannan, The estimation of frequency, *J. Appl. Probab.* 10 (1973) 510–519.
- [70] T. Hasan, Non linear time series regression for a class of amplitude modulated cosinusoids, *J. Time Ser. Anal.* 3 (2) (1982) 109–122.
- [71] P. Ciblat, P. Loubaton, E. Serpedin, G. Giannakis, Performance analysis of blind carrier frequency offset estimators for non-circular transmissions through frequency-selective channels, *IEEE Trans. Signal Process.* 50 (1) (2002) 130–140.
- [72] O. Besson, P. Stoica, Nonlinear least-squares approach to frequency estimation and detection for sinusoidal signals with arbitrary envelope, *Digital Signal Process.* 9 (1) (1999) 45–56.
- [73] A. Viterbi, A. Viterbi, Non-linear estimation of PSK-modulated carrier phase with application to burst digital transmissions, *IEEE Trans. Inform. Theory* 29 (July 1983) 543–551.
- [74] Y. Wang, E. Serpedin, P. Ciblat, Optimal blind nonlinear least-squares carrier phase and frequency offset estimation for general QAM modulations, *IEEE Trans. Wireless Commun.* 2 (2003) 1040–1054.
- [75] S. Colonnese, S. Rinauro, G. Panci, G. Scarano, Gain-control-free blind carrier frequency offset acquisition for QAM constellations, *IEEE Trans. Signal Process.* 58 (2010) 349–361.
- [76] E. Serpedin, P. Ciblat, G.B. Giannakis, P. Loubaton, Performance analysis of blind carrier phase estimators for general QAM constellations, *IEEE Trans. Signal Process.* 49 (2001) 1816–1823.
- [77] G. Vazquez, J. Riba, *Signal Processing Advances for Wireless Mobile Communications*, vol. 2, Englewood Cliffs, Prentice-Hall, 2000, (Chapter 9 Non-Data-Aided Digital Synchronization).
- [78] A. d'Andrea, U. Mengali, R. Reggiannini, The modified Cramer-Rao bound and its application to synchronization problems, *IEEE Trans. Commun.* 42 (2–4) (1994) 1391–1399.
- [79] G. Zhou, G. Giannakis, Harmonics in Gaussian multiplicative and additive white noise: Cramer-Rao bounds, *IEEE Trans. Signal Process.* 43 (1996) 1217–1231.
- [80] M. Ghogho, A. Swami, T. Durrani, Frequency estimation in the presence of Doppler spread: performance analysis, *IEEE Trans. Signal Process.* 49 (4) (2001) 777–789.
- [81] F. Gini, R. Reggiannini, U. Mengali, The modified Cramer-Rao Bound in vector parameter estimation, *IEEE Trans. Commun.* 46 (1) (1998) 52–60.
- [82] F. Gini, M. Luise, R. Reggiannini, Cramer-Rao bounds in the parametric estimation of fading radiotransmission channels, *IEEE Trans. Commun.* 46 (1998) 1390–1398.
- [83] M. Moeneclaey, On the true and the modified Cramer-Rao Bounds for the estimation of a scalar parameter in the presence of nuisance parameters, *IEEE Trans. Commun.* 46 (11) (1998) 1536–1544.
- [84] F. Rice, B. Cowley, B. Moran, M. Rice, Cramer-Rao lower bound for QAM phase and frequency estimation, *IEEE Trans. Commun.* 49 (2001) 1582–1591.
- [85] H. Steendam, M. Moeneclaey, Low-SNR limit of the Cramer-Rao Bound for estimating the carrier phase and frequency of a PAM, PSK, or QAM waveform, *IEEE Commun. Lett.* 5 (5) (2001) 218–220.
- [86] P. Ciblat, M. Ghogho, P. Larzabal, P. Forster, Harmonic retrieval in the presence of non-circular Gaussian multiplicative noise: Performance bounds, *EURASIP Signal Process.* 85 (2005) 737–749.
- [87] N. Noels, H. Steendam, M. Moeneclaey, The true Cramer-Rao bound for carrier frequency estimation from a PSK signal, *IEEE Trans. Commun.* 52 (2004) 834–844.

- [88] P. Stoica, B. Ng, On the Cramer-Rao bound under parametric constraints, *IEEE Signal Process. Lett.* 5 (1998) 177–179.
- [89] Z. Ben-Haim, Y. Eldar, On the constrained Cramer-Rao bound with singular fisher information matrix, *IEEE Signal Process. Lett.* 16 (2009) 453–456.
- [90] G. Riba, G. Vazquez, S. Calvo, Conditional maximum likelihood frequency estimation for offset modulation, in: *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, vol. 6, May 1998, pp. 3425–3428.
- [91] R. Imad, S. Houcke, M. Ghogho, Blind estimation of the phase and carrier frequency offset for LDPC coded systems, *EURASIP J. Adv. Signal Process.* 2010 (2010).
- [92] T. Heskes, O. Zoeter, W. Wiegerinck, *Advances in Neural Information Processing*. MIT Press, 2004 (Chapter Approximate Expectation-Maximization).
- [93] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, L. Vandendorpe, Turbo-synchronization: an EM approach, in: *IEEE International Conference on Communications (ICC)*, 2003, pp. 2933–2937.
- [94] M. Speth, S. Fechtel, G. Fock, H. Meyr, Optimum receiver design for wireless broad-band systems using OFDM—Part I, *IEEE Trans. Commun.* 47 (11) (1999) 1668–1677.
- [95] Q. Huang, M. Ghogho, Synchronization in OFDM-based systems, in: F. Bader, N. Zorba (Eds.), *OFDM/OFDMA Wireless Communication Systems*, Nova Publishers, 2011.
- [96] G.B. Giannakis, Filterbanks for blind channel identification and equalization, *IEEE Signal Process. Lett.* 4 (6) (1997) 184–187.
- [97] A. Scaglione, G.B. Giannakis, S. Barbarossa, Redundant filter-bank precoders and equalizers. Part I. Unification and optimal designs and Part II. Blind channel estimation, synchronization and direct equalization, *IEEE Trans. Signal Process.* 47 (7) (1999) 1988–2022.
- [98] B. Muquet, Z. Wang, G. Giannakis, M. de Courville, P. Duhamel, Cyclic prefixing or zero padding for wireless multicarrier transmissions?, *IEEE Trans. Commun.* 50 (12) (2002) 2136–2148.
- [99] L. Hanzo, M. Muenster, B.-J. Choi, T. Keller, *OFDM and MC-CDMA for broadband multi-user communications*, John Wiley & Sons, New York, 2003.
- [100] M. Morelli, Timing and frequency synchronization for the uplink of an OFDMA system, *IEEE Trans. Commun.* 52 (2) (2004) 296–306.
- [101] Y. Mostofi, D.C. Cox, Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system, *IEEE Trans. Commun.* 54 (2) (2006) 226–230.
- [102] X. Ma, H. Kobayashi, S. Schwartz, Effect of frequency offset on BER of OFDM and single carrier systems, in: *Proceedings of the IEEE International Symposium on Personal Indoor and Mobile Radio Communication*, vol. 3, 7–10 September 2003, pp. 2239–2243.
- [103] K. Sathananthan, C. Tellambura, Probability of error calculation of OFDM systems with frequency offset, *IEEE Trans. Commun.* 49 (11) (2001) 1884–1888.
- [104] J.J. Van de Beek, M. Sandell, P.O. Borjesson, ML estimation of timing and frequency offset in OFDM systems, *IEEE Trans. Signal Process.* 45 (6) (1997) 1800–1805.
- [105] M. Speth, S. Fechtel, G. Fock, H. Meyr, Optimum receiver design for OFDM-based broadband transmission—Part II: a case study, *IEEE Trans. Commun.* 49 (4) (2001) 571–578.
- [106] D. Lee, K. Cheun, Coarse symbol synchronization algorithms for OFDM systems in multipath channels, *IEEE Trans. Commun.* 6 (10) (2002) 446–448.
- [107] T.M. Schmidl, D.C. Cox, Robust frequency and timing synchronization for OFDM, *IEEE Trans. Commun.* 45 (12) (1997) 1613–1621.
- [108] H. Minn, V.K. Bhargava, K.B. Letaief, A robust timing and frequency synchronization for OFDM systems, *IEEE Trans. Wireless Commun.* 2 (4) (2003) 822–839.

- [109] K. Shi, E. Serpedin, Coarse frame and carrier synchronization of OFDM systems: a new metric and comparison, *IEEE Trans. Wireless Commun.* 3 (4) (2004) 1271–1284.
- [110] M. Ghogho, P. Ciblat, A. Swami, P. Bianchi, Training design for repetitive-slot-based CFO estimation in OFDM, *IEEE Trans. Signal Process.* 57 (12) (2009) 4958–4964.
- [111] P. Moose, A technique for orthogonal frequency division multiplexing frequency offset correction, *IEEE Trans. Commun.* 42 (10) (1994) 2908–2914.
- [112] M. Ghogho, A. Swami, Frame and frequency acquisition for OFDM, *IEEE Signal Process. Lett.* 15 (2008) 605–608.
- [113] M. Morelli, U. Mengali, An improved frequency offset estimator for OFDM applications, *IEEE Trans. Commun.* 3 (3) (1999) 75–77.
- [114] M. Ghogho, A. Swami, Unified framework for a class of frequency-offset estimation techniques for OFDM, in: *IEEE ICASSP 2004*, May 2004.
- [115] F. Tufvesson, O. Edfors, M. Faulkner, Time and frequency synchronization for OFDM using PN-sequence preambles, in: *IEEE Vehicular Technology Conference*, vol. 4, September 1999, pp. 2203–2207.
- [116] D. Landstrom, S.K. Wilson, J.J. Beek, P. Odling, P.O. Borjesson, Symbol time offset estimation in coherent OFDM systems, *IEEE Trans. Commun.* 50 (4) (2002) 545–549.
- [117] M. Speth, F. Classen, H. Meyr, Frame synchronization of OFDM systems in frequency selective fading channels, in: *IEEE Vehicular Technology Conference 1997*, 1997, pp. 1807–1811.
- [118] B. Yang, K.B. Letaief, R.S. Cheng, Z. Cao, Timing recovery for OFDM transmission, *IEEE J. Sel. Areas Commun.* 18 (11) (2000) 2278–2291.
- [119] N. Chen, M. Tanaka, R. Heaton, OFDM timing synchronisation under multi-path channels, in: *IEEE Vehicular Technology Conference*, vol. 1, April 2003, pp. 378–382.
- [120] H. Zhou, Y.-F. Huang, A maximum likelihood fine timing estimation for wireless OFDM systems, *IEEE Trans. Broadcast.* 55 (1) (2009) 31–41.
- [121] Y. Mostofi, D.C. Cox, A robust timing synchronization design in OFDM systems? Part I: low-mobility cases, *IEEE Trans. Wireless Commun.* 6 (12) (2007) 4329–4339.
- [122] Y. Mostofi, D.C. Cox, A robust timing synchronization design in OFDM systems—Part II: high-mobility cases, *IEEE Trans. Wireless Commun.* 6 (12) (2007) 4340–4348.
- [123] F. Daffara, O. Adami, A novel carrier recovery technique for orthogonal multicarrier systems, *Eur. Trans. Telecommun.* 7 (1996) 323–334.
- [124] F. Daffara, A. Chouly, Maximum likelihood frequency detectors for orthogonal multicarrier systems, in: *IEEE International Conference on Communications 1993*, pp. 766–771.
- [125] M. Morelli, A. Andrea, U. Mengali, Feedback frequency synchronization for OFDM applications, *IEEE Trans. Commun.* 5 (1) (2001) 28–30.
- [126] H. Liu, U. Tureli, A high-efficiency carrier estimator for OFDM communications, *IEEE Trans. Commun.* 2 (4) (1998) 104–106.
- [127] U. Tureli, H. Liu, M.D. Zoltowski, OFDM blind carrier offset estimation: ESPRIT, *IEEE Trans. Commun.* 48 (9) (2000) 1459–1461.
- [128] G. Santella, A frequency and symbol synchronization system for OFDM signals: architecture and simulation results, *IEEE Trans. Veh. Technol.* 49 (1) (2000) 245–275.
- [129] X. Ma, C. Tepedelenlioglu, G.B. Giannakis, S. Barbarossa, Non-data-aided carrier offset estimator for OFDM with null subcarriers, *IEEE J. Sel. Areas Commun.* 19 (12) (2001) 2504–2515.
- [130] M. Ghogho, A. Swami, G.B. Giannakis, Optimized null-subcarrier selection for CFO estimation in OFDM over frequency selective fading channels, in: *IEEE Globecom 2001*, vol. 48, November 2001, pp. 202–206.
- [131] H. Bolcskei, Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems, *IEEE Trans. Commun.* 48 (6) (2001) 988–999.

- [132] U. Tureli, D. Kivanc, H. Liu, Experimental and analytical studies on high-resolution OFDM carrier frequency offset estimator, *IEEE Trans. Veh. Technol.* 50 (2001) 629–643.
- [133] M. Ghogho, A. Swami, Semi-blind frequency offset synchronization for OFDM, in: *IEEE ICASSP 2002*, vol. 3, May 2002, pp. 2333–2336.
- [134] M. Ghogho, A. Swami, Blind frequency-offset estimator for OFDM systems transmitting constant-modulus symbols, *IEEE Trans. Commun.* 6 (8) (2002) 343–345.
- [135] F. Yang, K.H. Li, K.C. Teh, A carrier frequency offset estimator with minimum output variance for OFDM systems, *IEEE Trans. Commun.* 8 (11) (2004) 677–679.
- [136] Y.W. Yao, G.B. Giannakis, Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems, *IEEE Trans. Commun.* 53 (1) (2005) 173–183.
- [137] D. Rife, R. Boorstyn, Single-tone parameter estimation from discrete-time observations, *IEEE Trans. Inform. Theory* 20 (5) (1974) 591–598.
- [138] J. Van de Beek, P. Borjesson, M. Boucheret, D. Landstrom, J. Arenas, P. Odling, C. Ostberg, M. Wahlqvist, S. Wilson, A time and frequency synchronization scheme for multiuser OFDM, *IEEE J. Sel. Areas Commun.* 17 (11) (1999) 1900–1914.
- [139] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004), vol. 29, 2009, pp. C1–2004.
- [140] D. Huang, K.B. Letaief, An interference-cancellation scheme for carrier frequency offsets correction in OFDMA systems, *IEEE Trans. Commun.* 53 (7) (2005) 1155–1165.
- [141] D. Marabissi, R. Fantacci, S. Papini, Robust multiuser interference cancellation for OFDM systems with frequency offset, *IEEE Trans. Wireless Commun.* 5 (11) (2006) 3068–3076.
- [142] S. Manohar, D. Sreedhar, V. Tikiya, A. Chockalingam, Cancellation of multiuser interference due to carrier frequency offsets in uplink OFDMA, *IEEE Trans. Wireless Commun.* 6 (7) (2007) 2560–2571.
- [143] S.-W. Hou, C.C. Ko, Multiple-access interference suppression for interleaved OFDMA system uplink, *IEEE Trans. Veh. Technol.* 57 (1) (2008) 194–205.
- [144] S. Barbarossa, M. Pompili, G.B. Giannakis, Channel-independent synchronization of orthogonal frequency division multiple access systems, *IEEE J. Sel. Areas Commun.* 20 (2) (2002) 474–486.
- [145] M. Morelli, C.-C.J. Kuo, M.-O. Pun, Synchronization techniques for orthogonal frequency division multiple access (OFDMA): a tutorial review, *Proc. IEEE* 95 (7) (2007) 1394–1427.
- [146] Z. Cao, U. Tureli, Y. Yao, Deterministic multiuser carrier-frequency offset estimation for interleaved OFDMA uplink, *IEEE Trans. Commun.* 52 (9) (2004) 1585–1594.
- [147] D.S. Radovic, M.M. Eric, Performance of subspace based multi-user CFO estimation for interleaved OFDMA uplink, in: *Proc. Int. Conf. Computer as a Tool*, vol. 2, Belgrade, Serbia and Montenegro, November 2005, pp. 1602–1605.
- [148] J. Lee, S. Lee, K.-J. Bang, S. Cha, D. Hong, Carrier frequency offset estimation using ESPRIT for interleaved OFDMA uplink systems, *IEEE Trans. Veh. Technol.* 56 (5) (2007) 3227–3231.
- [149] M.-O. Pun, M. Morelli, C.-C. Kuo, Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions, *IEEE Trans. Commun.* 54 (4) (2006) 726–736.
- [150] X.Y. Fu, Y.H. Li, H. Minn, A new ranging method for OFDMA systems, *IEEE Trans. Commun.* 6 (2) (2007) 659–669.
- [151] J.-H. Lee, S.-C. Kim, Time and frequency synchronization for OFDMA uplink system using the SAGE algorithm, *IEEE Trans. Wireless Commun.* 6 (4) (2007) 1176–1181.
- [152] M.-O. Pun, M. Morelli, C.-C. Kuo, Iterative detection and frequency synchronization for OFDMA uplink transmissions, *IEEE Trans. Wireless Commun.* 6 (2) (2007) 629–639.
- [153] M. Morelli, U. Mengali, Carrier-frequency estimation for transmissions over selective channels, *IEEE Trans. Commun.* 48 (9) (2000) 1580–1589.

- [154] S. Sezginer, P. Bianchi, Asymptotically efficient reduced complexity frequency offset and channel estimators for uplink MIMO-OFDMA systems, *IEEE Trans. Signal Process.* 56 (3) (2008) 964–979.
- [155] X.N. Zeng, A. Ghayeb, Joint CFO and channel estimation for OFDMA uplink: an application of the variable projection method, *IEEE Trans. Wireless Commun.* 8 (5) (2009) 2306–2311.
- [156] W.C. Lindsey, F. Ghazvinian, W.C. Hagmann, K. Dessouky, Network synchronization, *Proc. IEEE* 73 (10) (1985) 1445–1467.
- [157] S. Bregni, A historical perspective on telecommunications network synchronization, *IEEE Commun. Mag.* (1988) 150–166.
- [158] B. Hamilton, X. Ma, Q. Zhao, J. Xu, ACES: adaptive clock estimation and synchronization using kalman filtering, in: *Proc. MobiCom'08*, 2008.
- [159] J. Elson, L. Girod, D. Estrin, Fine-grained network time synchronization using reference broadcasts, in: *Proc. Fifth ACM SIGOPS Operating Syst. Review*, December 2002, pp. 147–163.
- [160] K. Romer, Time synchronization in ad hoc networks, in: *Proc. 2nd ACM Int. Symp. Mobile Ad Hoc Networking Comp.*, 2002, pp. 173–182.
- [161] M. Sichitiu, C. Veerarittiphan, Simple, accurate time synchronization for wireless sensor networks, in: *Proc. IEEE Wireless Commun. Networking Conf. (WCNC)*, 2003.
- [162] S. Ganeriwal, R. Kumar, M.B. Srivastava, Timing-sync protocol for sensor networks, in: *Proc. First Intl. Conf. Embedded Networked Sensor Syst. (SenSys-03)*, 2003.
- [163] J. van Greunen, J. Rabaey, Lightweight time synchronization for sensor networks, in: *Proc. WSNA'03*, 2003.
- [164] S. Graham, P.R. Kumar, Time in general-purpose control systems: the control time protocol and an experimental evaluation, in: *Proc. 43rd IEEE Conf. Decision Control*, 2004, pp. 4004–4009.
- [165] F. Cristian, A probabilistic approach to distributed clock synchronization, *Distrib. Comput.* 3 (1989) 146–158.
- [166] K. Arvind, Probabilistic clock synchronization in distributed systems, *IEEE Trans. Parallel Distrib. Syst.* 5 (5) (1994) 474–487.
- [167] D. Dolev, J. Halpern, H.R. Strong, On the possibility and impossibility of achieving clock synchronization, in: *Proc. 16th Annual ACM Symp. Theory Comput.* 1984.
- [168] J. Lundelius, N. Lynch, An upper and lower bound for clock synchronization, *Inform. Control* 62 (1984) 190–204.
- [169] S. Biaz, J.L. Welch, Closed form bounds for clock synchronization under simple uncertainty assumptions, *IPL: Inform. Process. Lett.* 80 (3) (2001) 151–157.
- [170] D.L. Mills, Internet time synchronization: the network time protocol, *IEEE Trans. Commun.* 39 (10) (1991) 1482–1493.
- [171] D.R. Jeske, On maximum-likelihood estimation of clock offset, *IEEE Trans. Commun.* 53 (1) (2005) 53–54.
- [172] D.R. Jeske, A. Sampath, Estimation of clock offset using bootstrap bias-correction techniques, *Technometrics* 45 (3) (2003) 256–261.
- [173] Q. Chaudhari, E. Serpedin, K. Qaraqe, On maximum likelihood estimation of clock offset and skew in networks with exponential distribution, *IEEE Trans. Signal Process.* 56 (4) 2008 1685–1697.
- [174] Mirollo, S. Strogatz, Synchronization of pulse-coupled biological oscillators, *SIAM J. Appl. Math.* 50 (6) (1990) 1645–1662.
- [175] D. Lucarelli, I.-J. Wang, Decentralized synchronization protocols with nearest neighbor communications, in: *Proc. ACM SenSys*, 2004.
- [176] E. Mallada, A. Tang, Synchronization of phase-coupled oscillators with arbitrary topology, in: *Proc. American Control Conference (ACC'10)*, 2010.
- [177] Y.-W. Hong, A. Scaglione, A scalable synchronization protocol for large sensor networks and its applications, *IEEE J. Sel. Areas Commun.* 23 (5) (2005) 1085–1099.
- [178] Q. Li, D. Rus, Global clock synchronization in sensor networks, *IEEE Trans. Comput.* 55 (2) (2006) 214–226.

- [179] P. Barooah, J. Hespanha, A. Swami, On the effect of asymmetric communication on distributed time synchronization, in: Proc. 46th IEEE Conf. Decision Control (CDC'07), 2007.
- [180] S. Sardellitti, S. Barbarossa, A. Swami, Optimal topology control and power allocation for minimum energy consumption in consensus networks, *IEEE Trans. Signal Process.* 60 (1) (2012) 383–399.

Channel Estimation, Equalization, Precoding, and Tracking

3

Jitendra K. Tugnait

Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA

2.03.1 Introduction

Multipath propagation results in a received signal that is a superposition of several delayed and scaled copies of the transmitted signal giving rise to frequency-selective fading. It leads to intersymbol interference (ISI) at the receiver which, in turn, may lead to high error rates in symbol detection. Equalizers are designed to compensate for these channel distortions. One may directly design an equalizer given the received signal, or one may first estimate the channel impulse response and then design an equalizer based on the estimated channel. After some processing (matched filtering, for instance), the continuous-time received signals are sampled at the baud (symbol) or higher (fractional) rate before processing them for channel estimation and/or equalization. It is therefore convenient to work with a baseband-equivalent discrete-time channel model. Depending upon the sampling rate, one has either a single-input single-output (SISO) (baud rate sampling), or a single-input multiple-output (SIMO) (fractional sampling), complex discrete-time equivalent baseband channel. Knowledge of the channel response can be advantageously used at the transmitter to precode the information sequence to be transmitted so as to simplify the equalizer complexity at the receiver and to mitigate interference in MIMO systems.

In this chapter, we present an overview/review of various approaches to channel estimation and equalization for wireless mobile systems with a brief discussion of channel precoding techniques. Since approaches to channel estimation depend upon the underlying channel model, we also review various approaches to channel modeling. In Section 2.03.2 we present the relevant channel models including time-variant and time-invariant models. In Section 2.03.3 various channel estimation methods are discussed. In Section 2.03.4 equalization approaches are reviewed and in Section 2.03.5 a brief discussion of precoding approaches is presented. In Section 2.03.6 we discuss tracking to adapt to time-varying channels.

Notation 1. Superscripts H , $*$, T , and \dagger denote the complex conjugate transpose, complex conjugation, transpose, and Moore-Penrose pseudo-inverse operations, respectively. The function $\delta(l)$ is the Kronecker delta function with $\delta(l) = 1$ if $l = 0$, $= 0$ otherwise, and \mathbf{I}_N is the $N \times N$ identity matrix. The symbol \otimes denotes the Kronecker product, and $\text{tr}(\mathbf{A})$ is the trace of a square matrix \mathbf{A} . The (n, m) th entry of a matrix \mathbf{C} is denoted by $[\mathbf{C}]_{n,m}$.

2.03.2 Channel models

2.03.2.1 Time-variant (doubly selective) channels

Consider a time-varying (e.g., mobile wireless) channel (linear system) with complex baseband, continuous-time, received signal $r(t)$ and transmitted complex baseband, continuous-time information signal $s(t)$ (with symbol interval T_s seconds) related by [1]

$$r(t) = \int_{-\infty}^{\infty} \tilde{h}(t; \tau) \tilde{s}(t - \tau) d\tau + w(t), \quad (3.1)$$

where $\tilde{h}(t; \tau)$ is the time-varying impulse response of the channel denoting the response of the channel at time t to a unit impulse input at time $t - \tau$ and $w(t)$ is the additive noise (typically white Gaussian). A delay-Doppler spread function $H(f; \tau)$ is defined as the Fourier transform of $\tilde{h}(t; \tau)$ with respect to t [1,2]

$$H(f; \tau) = \int_{-\infty}^{\infty} \tilde{h}(t; \tau) e^{-j2\pi f t} dt. \quad (3.2)$$

If $|H(f; \tau)| \approx 0$ for $|\tau| > \tau_d$, then τ_d is called the (multipath) delay-spread of the channel. If $|H(f; \tau)| \approx 0$ for $|f| > f_d$, then f_d is called the Doppler spread of the channel. In order to capture the complexity of the physical interactions characterizing the transmission through a real channel, $\tilde{h}(t; \tau)$ is typically modeled as a two-dimensional zero-mean random process. If $\tilde{h}(t; \tau)$ is wide-sense stationary in variable t , and $\tilde{h}(t; \tau_1)$ is uncorrelated with $\tilde{h}(t; \tau_2)$ for $\tau_1 \neq \tau_2$ and any t , one obtains the well-known wide-sense stationary uncorrelated scattering (WSSUS) channel [1,2, Chapter 14].

2.03.2.1.1 Tapped delay line model

We now consider a discrete-time channel model. If a linear modulation scheme is used, the baseband transmitted signal can be represented as

$$\tilde{s}(t) = \sum_{k=-\infty}^{\infty} s(k) p_T(t - kT_s), \quad (3.3)$$

where $\{s(k)\}$ is the information sequence and $p_T(t)$ is the transmit (lowpass) filter (typically a root raised cosine filter). Therefore, the baseband signal at the receiver is given by

$$r(t) = \sum_{k=-\infty}^{\infty} s(k) \int_{-\infty}^{\infty} \tilde{h}(t; \alpha) p_T(t - kT_s - \alpha) d\alpha + w(t). \quad (3.4)$$

After filtering with a receive filter with impulse response $p_R(t)$, the received baseband signal is given by

$$\tilde{y}(t) = \sum_{k=-\infty}^{\infty} s(k) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_R(t - \beta) \tilde{h}(\beta; \alpha) p_T(\beta - kT_s - \alpha) d\alpha d\beta + \tilde{v}(t), \quad (3.5)$$

where $\tilde{v}(t) = \int p_R(\tau)w(t - \tau)d\tau$. If the continuous-time signal $\tilde{y}(t)$ is sampled at once every T_s sec., we obtain the discrete-time sequence

$$y(n) := \tilde{y}(t)|_{t=nT_s} = \sum_{k=-\infty}^{\infty} s(k)h(n; n-k) + v(n) = \sum_{l=-\infty}^{\infty} h(n; l)s(n-l) + v(n), \quad (3.6)$$

where $h(n; l)$ is the (effective) channel response at time n to a unit impulse input at time $n - l$ and

$$h(n; n - k) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_R(nT_s - \beta) \tilde{h}(\beta; \alpha) p_T(\beta - kT_s - \alpha) d\alpha d\beta. \quad (3.7)$$

Note that the noise sequence $\{v(n)\}$ in (3.6) is no longer necessarily white; it can be whitened by further time-invariant linear filtering (see [1]). Henceforth, we assume that a whitening filter has been applied to $y(n)$, but with an abuse of notation, we will still use (3.6). For a causal system, $h(n; l) = 0$ for $l < 0$ ($\forall n$) and for a finite length channel of maximum length $T_s L$, $h(n; l) = 0$ for $l > L$ ($\forall n$). In this case we modify (3.6) as (recall the noise whitening filter)

$$y(n) = \sum_{l=0}^L h(n; l)s(n-l) + v(n). \quad (3.8)$$

The model (3.8) represents a time- and frequency-selective linear channel. A tapped delay line structure for this model is shown in Figure 3.1. For a slowly (compared to the baud-rate) time-varying system, one often simplifies (3.8) to a time-invariant system as

$$y(n) = \sum_{l=0}^L h(l)s(n-l) + v(n), \quad (3.9)$$

where $h(l) = h(0; l)$ is the time-invariant channel response to a unit impulse input at time 0. The model (3.9) represents a frequency-selective linear channel with no time selectivity. It is a widely used model for receiver design.

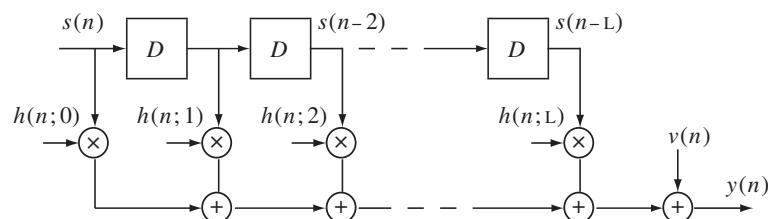


FIGURE 3.1

Tapped delay line model of frequency- and time-selective channel with finite impulse response. D represents a unit (symbol duration) delay.

Suppose that $h(n; l) = h(n)\delta(l)$. Then we have the time-selective and frequency-nonselective channel whose output is given by

$$y(n) = h(n)s(n) + v(n). \quad (3.10)$$

Finally, a time-nonselective and frequency-nonselective channel is modeled as

$$y(n) = hs(n) + v(n), \quad (3.11)$$

where h is a random variable (or a constant).

2.03.2.1.2 Autoregressive (AR) models

It is possible to accurately represent a wide-sense stationary uncorrelated scattering (WSSUS) channel by a large order AR model; see [3–5] and references therein. Let

$$\tilde{\mathbf{h}}(n) := [h(n; 0) \ h(n; 1) \ \cdots \ h(n; L)]^T, \quad (3.12)$$

where $\tilde{\mathbf{h}}(n)$ is $(L + 1) \times 1$ vector. Then a p th order AR model, AR(p), for $\tilde{\mathbf{h}}(n)$ is given by

$$\tilde{\mathbf{h}}(n) = \sum_{i=1}^p \mathbf{A}_i \tilde{\mathbf{h}}(n-i) + \mathbf{G}_0 \tilde{\mathbf{w}}(n), \quad (3.13)$$

where \mathbf{A}_i s are the $(L + 1) \times (L + 1)$ AR coefficient matrices, \mathbf{G}_0 is also $(L + 1) \times (L + 1)$ and independent and identically distributed (i.i.d.), $(L + 1) \times 1$ driving noise $\tilde{\mathbf{w}}(n)$ is zero-mean with identity covariance matrix. Suppose that we know the correlation function $\mathbf{R}_h(m) = E\{\tilde{\mathbf{h}}(n+m)\tilde{\mathbf{h}}^H(n)\}$ for lags $m = 0, 1, \dots, p$. The following Yule-Walker equation holds for (3.13) [6]:

$$\mathbf{R}_h(m) = \sum_{i=0}^p \mathbf{A}_i \mathbf{R}_h(m-i) + \mathbf{G}_0 \mathbf{G}_0^H \delta(m). \quad (3.14)$$

Using (3.14) for $m = 1, 2, \dots, p$, and the fact that $\mathbf{R}_h(-m) = \mathbf{R}_h^H(m)$, one can estimate \mathbf{A}_i s. Using the estimated \mathbf{A}_i s and (3.14) for $m = 0$, one can find $\mathbf{G}_0 \mathbf{G}_0^H$, from which one can find (non-unique) \mathbf{G}_0 by computing its “square root” [7, p. 358]. In [3] high values of p (several tens) have been used for channel simulation, whereas in [4,5] only AR(1) or AR(2) models have been used where the objective is channel estimation and related issues. An AR(1) model is given by

$$h(n; l) = \alpha_c h(n-1; l) + w_c(n), \quad (3.15)$$

where α_c is the AR coefficient, and the driving noise $w_c(n)$ is zero-mean complex Gaussian with variance σ_{wc}^2 and statistically independent of $h(n-1; l)$. Assume that $h(n; l)$ is also zero-mean, complex Gaussian with variance σ_h^2 . Then [8]

$$\alpha_c = \frac{1}{\sigma_h^2} E\{h(n; l)h^*(n-1; l)\}, \quad \sigma_{wc}^2 = \sigma_h^2(1 - |\alpha_c|^2). \quad (3.16)$$

2.03.2.1.3 Basis expansion models

Basis expansion models (BEMs) have also been widely investigated to represent doubly selective channels in wireless applications [9–13], where the time-varying taps are expressed as superpositions of time-varying basis functions in modeling Doppler effects, weighted by time-invariant coefficients. Candidate basis functions include complex exponential (Fourier) functions [10, 11], polynomials [9], and discrete prolate spheroidal sequences [13], etc. In contrast to AR models that describe temporal variation on a symbol-by-symbol update basis, a BEM depicts the evolution of the channel over a period (block) of time. Intuitively, the coefficients of the BEM approximation should evolve much more slowly in time than the channel, and hence are more convenient to track in a fast fading environment.

Suppose that we include the effects of transmit and receive filters in the time-variant impulse response $\tilde{h}(t; \tau)$ in (3.1). Suppose that this channel has a delay-spread τ_d and a Doppler spread f_d . Consider the k th block of data consisting of an observation window of T_B symbols where the baud-rate data samples in the block are indexed as $n = \bar{n}_k, \bar{n}_k + 1, \dots, \bar{n}_k + T_B - 1, \bar{n}_k := (k-1)T_B$. If $2f_d\tau_d < 1$ (underspread channel), the complex exponential basis expansion model (CE-BEM) representation of $h(n; l)$ in (3.8) is given by

$$h(n; l) = \sum_{q=1}^Q h_q(l) e^{j\omega_q n}, \quad n = \bar{n}_k, \bar{n}_k + 1, \dots, \bar{n}_k + T_B - 1, \quad (3.17)$$

where one chooses ($l = 0, 1, \dots, L$, and K is an integer)

$$T := KT_B, \quad K \geq 1, \quad Q \geq 2\lceil f_d TT_s \rceil + 1, \quad (3.18)$$

$$\omega_q := \frac{2\pi}{T} [q - (Q+1)/2], \quad q = 1, 2, \dots, Q, \quad L := \lfloor \tau_d/T_s \rfloor. \quad (3.19)$$

The BEM coefficients $h_q(l)$ s remain invariant during this block, but are allowed to change at the next block, and the Fourier basis functions $\{e^{j\omega_q n}\}$ ($q = 1, 2, \dots, Q$) are common for each block. If the delay spread τ_d and the Doppler spread f_d of the channel (or at least their upper-bounds) are known, one can infer the basis functions of the CE-BEM [11]. Treating the basis functions as known, estimation of a time-varying process is reduced to estimating the invariant coefficients over a block of length T_B symbols. Note that the BEM period is $T = KT_B$ whereas the block size is T_B symbols. If $K > 1$ (e.g., $K = 2$ or $K = 3$), then the Doppler spectrum is said to be over-sampled [14] compared to the case $K = 1$ where the Doppler spectrum is said to be critically sampled. In [10, 11] only $K = 1$ (henceforth called CE-BEM) is considered whereas [14] considers $K \geq 2$ (henceforth called over-sampled CE-BEM).

Equation (3.17) applies to single-input single-output systems—one user and one receiver with symbol-rate sampling. It is easily modified to handle multiuser, multiple transmit and receive antennas, and higher than symbol rate sampling (multiple samples per symbol)—the basic representation remains essentially unchanged.

The representation $h(n; l)$ in (3.17) is a special case of a more general representation

$$h(n; l) = \sum_{q=1}^Q h_q(l) \phi_q(n), \quad (3.20)$$

where $\{\phi_q(n)\}_{q=1}^Q$ are a set of orthogonal basis functions (over the time interval under consideration). Examples include wavelet-based expansions as in [15], polynomial bases as in [9] and other possibilities [16]. In discrete prolate spheroidal BEM (DPS-BEM), the i th DPS vector

$$\mathbf{u}_i := [u_i(0), \dots, u_i(T_B - 1)]^T$$

(called Slepian sequence in [13], which is a time-windowed (infinite) DPS sequence) is the i th eigenvector of a matrix \mathbf{C} [17]:

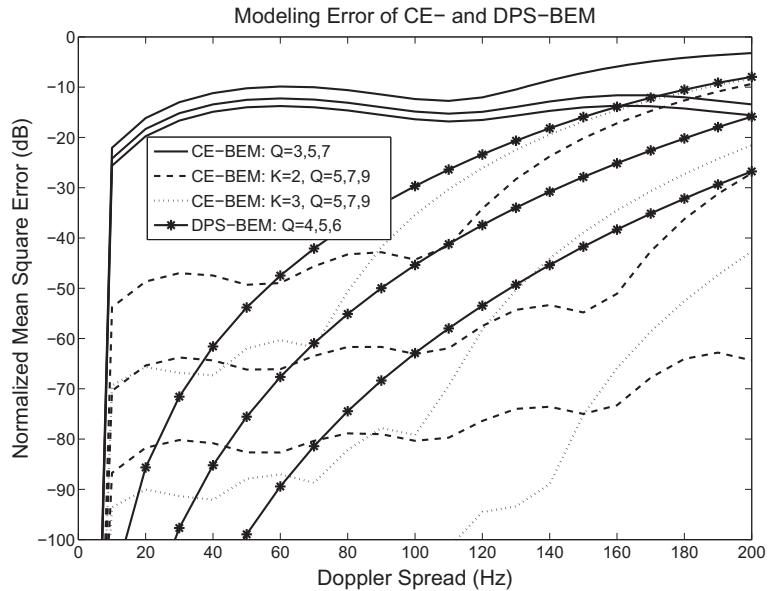
$$\mathbf{C}\mathbf{u}_i = \lambda_i \mathbf{u}_i, \quad \text{where } [\mathbf{C}]_{n,m} = \frac{\sin [2\pi(n-m)f_d T_s]}{[\pi(n-m)]}$$

is the (n, m) th entry of \mathbf{C} and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{T_B}$ are the eigenvalues of \mathbf{C} . The Slepian sequences $\{u_q(n)\}$ are orthonormal over the finite time interval $[0, T_B - 1]$. The modeling error of the CE-BEM can result in a noticeable floor in BER curves [18]. The polynomial basis functions are neither time-limited nor band-limited and their square bias varies heavily over the range of Doppler spread considered in [13]. DPS sequences are a good alternative as a basis set to approximate bandlimited channels alleviating the spectral leakage of CE-BEM [13]. The (infinite) DPS sequences have their maximum energy concentration in an interval with length T while being bandlimited to $[-f_d T_s, f_d T_s]$, where $u_1(n)$ is the unique sequence that is bandlimited and most time-concentrated, $u_2(n)$ is the next sequence having maximum energy concentration among the DPS sequences orthogonal to $u_1(n)$, and so on [17].

Figure 3.2 shows the channel modeling errors resulting from (critically sampled) CE-BEM, DPS-BEM and oversampled CE-BEM ($K = 2$ or 3) when the underlying channel is a one-tap time-selective channel following Jakes' spectrum. The results are based on Monte Carlo averaging over 1000 runs with $T_B = 400$, $T_s = 25 \mu\text{s}$ and varying Doppler spreads. (The results were obtained following the procedure in [13].) For a fixed value of Q , DPS-BEM provides the best fit whereas CE-BEM (no oversampling) yields minor improvements with increasing Q . On the other hand, the basis functions in oversampled CE-BEM are not mutually orthogonal leading to “analytical” difficulties. There exists a vast literature based on CE-BEM (no oversampling) where it is assumed that physical channel is accurately described by CE-BEM both for analysis and simulations; see e.g., [11, 19–24].

2.03.2.2 Time-invariant channels

For a baud (symbol)-rate sampled system, the equivalent baseband channel model is given by (3.9) which is a single-input single-output (SISO) complex discrete-time baseband-equivalent channel model. The output sequence $\{y(n)\}$ in (3.9) is discrete-time stationary. When there is excess channel bandwidth [bandwidth $> \frac{1}{2} \times (\text{baud rate})$], baud rate sampling is below the Nyquist rate leading to aliasing and depending upon the symbol timing phase, in certain cases, causing deep spectral notches in sampled, aliased channel transfer function [25]. Linear equalizers designed on the basis of the baud-rate sampled channel response, are quite sensitive to symbol timing errors. Initially, in the trained case, fractional sampling was investigated to robustify the equalizer performance against timing errors. The model (3.9) does not apply to fractionally-spaced samples, i.e., when the sampling interval is a fraction of the symbol duration. The fractionally sampled digital communication signal is a cyclostationary signal [26] which may be represented as a vector stationary sequence using a time series representation (TSR)

**FIGURE 3.2**

Channel modeling error for one-tap Jakes' channel, $T_B = 400$, $T_s = 25 \mu\text{s}$.

([26] Section 12.6). Suppose that we sample at N -times the baud rate with signal samples spaced T_s/N sec. apart where T_s is the symbol duration. Then a TSR for the sampled signal is given by

$$y_i(n) = \sum_{l=0}^L h_i(l)s(n-l) + v_i(n); \quad (i = 1, 2, \dots, N), \quad (3.21)$$

where now we have N samples every symbol period, indexed by i . Notice, however, that the information sequence $s(n)$ is still one “sample” per symbol. It is assumed that the signal incident at the receiver is first passed through a receive filter whose transfer function equals the square root of a raised cosine pulse, and that the receive filter is matched to the transmit filter. The noise sequence in (3.21) is the result of the fractional rate sampling of a filtered continuous-time white Gaussian noise process. Therefore, the sampled noise sequence is white at the symbol rate, but correlated at the fractional rate. Stack N consecutive received samples in the n th symbol duration to form a N -vector $\mathbf{y}(n)$ satisfying

$$\mathbf{y}(n) = \sum_{l=0}^L \mathbf{h}(l)s(n-l) + \mathbf{v}(n), \quad (3.22)$$

where $\mathbf{h}(n)$ is the vector impulse response of the SIMO equivalent channel model given by

$$\mathbf{h}(n) = [h_1(n) \ h_2(n) \ \dots \ h_N(n)]^T, \quad (3.23)$$

and $\mathbf{y}(n)$ and $\mathbf{v}(n)$ are defined similarly.

2.03.2.3 MIMO channels

A general MIMO channel model with K inputs (users, antennas, ...) and N outputs (receivers, antennas, ...) can be formulated as in (3.48) (given later); however, it lacks “physical” parameters (such as antenna spacing and arrangement). Representative works on MIMO channel models that incorporate “propagation” effects include [27–31] and references therein. The Kronecker model of [30] is a popular analytical model for spatially-correlated MIMO channels. It models the correlation at the receiver and at the transmitter independently, neglecting the statistical interdependence of both link ends. Improvements upon this model include the virtual channel representation model of [29] and the stochastic model of [31] where joint correlation at both link ends have been considered and experimentally validated. These models are suitable for wireless local area networks (WLANs). For wireless personal area networks (WPANs), a different MIMO channel model has been proposed in [28] to account for irregular (nonuniform) antenna arrangements and other deviations.

2.03.3 Channel estimation

We first consider three types of channel estimators within the framework of maximizing the likelihood function. [Unless otherwise noted the underlying channel model is given by the time-invariant model (3.22).] In general, one of the most effective and popular parameter estimation algorithms is the maximum likelihood (ML) method.

Let us consider the N -vector channel model given in (3.22). Suppose that we have collected M samples of the observation $\tilde{\mathbf{y}} = [\mathbf{y}^T(T_B - 1), \dots, \mathbf{y}^T(0)]^T$. We then have the following linear model

$$\begin{aligned}\tilde{\mathbf{y}} &= \begin{pmatrix} s(T_B - 1)\mathbf{I}_N & s(T_B - 2)\mathbf{I}_N & \cdots & s(T_B - L - 1)\mathbf{I}_N \\ \vdots & \text{Block} & \text{Hankel} & \text{Matrix} \\ s(0)\mathbf{I}_N & s(-1)\mathbf{I}_N & \cdots & s(-L)\mathbf{I}_N \end{pmatrix} \begin{pmatrix} \mathbf{h}(0) \\ \vdots \\ \mathbf{h}(L) \end{pmatrix} + \begin{pmatrix} \mathbf{v}(T_B - 1) \\ \vdots \\ \mathbf{v}(0) \end{pmatrix} \\ &= \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}} + \tilde{\mathbf{v}},\end{aligned}\quad (3.24)$$

where \mathbf{I}_N is an $N \times N$ identity matrix, \mathbf{s} and $\tilde{\mathbf{v}}$ are vectors consisting of samples of the input sequence $\{s(n)\}$ and noise $\{\mathbf{v}(n)\}$, respectively, $\tilde{\mathbf{h}}$ is the vector of the channel parameters, and a block Hankel matrix has identical block entries on its block antidiagonals. Let $\boldsymbol{\theta}$ be the vector of unknown parameters that may include the channel parameters $\tilde{\mathbf{h}}$ and possibly the entire or part of the input vector \mathbf{s} . Given the probability space that describes jointly the noise vector $\tilde{\mathbf{v}}$ and possibly the input data vector \mathbf{s} , we can then obtain, in principle, the probability density function (pdf) of the observation $\tilde{\mathbf{y}}$. As a function of the unknown parameter $\boldsymbol{\theta}$, the pdf of the observation $f(\tilde{\mathbf{y}} | \boldsymbol{\theta})$ is referred to as the *likelihood function*. The maximum likelihood estimator is defined by the following optimization

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} f(\tilde{\mathbf{y}} | \boldsymbol{\theta}), \quad (3.25)$$

where Θ defines the domain of the optimization.

While the ML estimator is conceptually simple, and it usually has good performance when the sample size is sufficiently large, the implementation of ML estimator is sometimes computationally intensive. Furthermore, the optimization of the likelihood function in (3.25) is often hampered by the existence of local maxima. Therefore, it is desirable that effective initialization techniques are used in conjunction with the ML estimation.

2.03.3.1 Training-based channel estimation

The training-based channel estimation assumes the availability of the input vector \mathbf{s} (as training symbols) and its corresponding observation vector $\tilde{\mathbf{y}}$. When the noise samples are zero mean, white Gaussian, i.e., $\tilde{\mathbf{v}}$ is a zero mean, Gaussian random vector with covariance $\sigma_v^2 \mathbf{I}_{T_B N}$, the ML estimator defined in (3.25), with $\boldsymbol{\theta} = \mathbf{h}$, is given by

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} \|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}}\|^2 = \mathcal{T}^\dagger(\mathbf{s})\tilde{\mathbf{h}} \quad (3.26)$$

where $\mathcal{T}^\dagger(\mathbf{s})$ is the Moore-Penrose pseudo-inverse of the $\mathcal{T}(\mathbf{s})$ defined in (3.24). This is also the classical linear least-squares estimator which can be implemented recursively, and it turns out to be the best (in terms of having minimum mean square error) among all unbiased estimators and it is the most efficient in the sense that it achieves the Cramer-Rao lower bound. Various adaptive implementations can be found in [1].

2.03.3.1.1 Time-variant channels

In case of general time-varying channels represented by (3.6), a simple generalization of [32] (see also [11]) is to use a periodic Kronecker delta function sequence with period \bar{P} as training:

$$s(n) = \sum_m \delta(n - m\bar{P}). \quad (3.27)$$

With (3.27) as input to model (3.6), one obtains

$$y(n) = \sum_m h(n; n - m\bar{P}) + v(n), \quad (3.28)$$

so that if $\bar{P} > L$, we have for $0 \leq i \leq L$,

$$y(k\bar{P} + i) = h(k\bar{P} + i; i) + v(k\bar{P} + i). \quad (3.29)$$

Therefore, one may take the estimate of $h(k\bar{P}; i)$ as

$$\hat{h}(k\bar{P}; i) = y(k\bar{P} + i) = h(k\bar{P} + i; i) + v(k\bar{P} + i). \quad (3.30)$$

For time samples between $k\bar{P}$ (k an integer), linear interpolation may be used to obtain channel estimates.

If we use the BEM representation (3.20), then we directly estimate the time-invariant parameters $h_q(l)$ s. From (3.6) and (3.20) we have

$$y(n) = \sum_{l=0}^L \sum_{q=1}^Q h_q(l) \phi_q(n) s(n-l) + v(n) \quad (3.31)$$

$$= \sum_{l=0}^L s(n-l) \underbrace{\begin{bmatrix} \phi_1(n) & \cdots & \phi_Q(n) \end{bmatrix}}_{\boldsymbol{\phi}(n)} \underbrace{\begin{bmatrix} h_1(l) \\ \vdots \\ h_Q(l) \end{bmatrix}}_{\mathbf{h}(l)} + v(n). \quad (3.32)$$

Collecting T_B samples of the observations $\mathbf{y} = [y(T_B - 1), \dots, y(0)]^T$ we have the linear model

$$\mathbf{y} = \begin{pmatrix} s(T_B - 1)\boldsymbol{\phi}(T_B - 1) & \cdots & s(T_B - L - 1)\boldsymbol{\phi}(T_B - 1) \\ \vdots & & \vdots \\ s(0)\boldsymbol{\phi}(0) & \cdots & s(-L)\boldsymbol{\phi}(0) \end{pmatrix} \begin{pmatrix} \mathbf{h}(0) \\ \vdots \\ \mathbf{h}(L) \end{pmatrix} + \begin{pmatrix} v(T_B - 1) \\ \vdots \\ v(0) \end{pmatrix}. \quad (3.33)$$

Now we have a model similar to (3.24) with a solution similar to (3.26).

2.03.3.2 Blind channel estimation

Here no training symbols are available (or exploited). Blind techniques can not resolve phase ambiguity in the channel estimate.

2.03.3.2.1 Combined channel and symbol estimation

The simultaneous estimation of the input vector and the channel is general ill-posed; however, utilization of qualitative information about the channel and the input can help alleviate this deficiency. To this end, we consider two different types of maximum likelihood techniques based on different models of the input sequence.

Stochastic Maximum Likelihood Estimation. While the input vector \mathbf{s} is unknown, it may be modeled as a random vector with a known distribution. In such a case, the likelihood function of the unknown parameter $\boldsymbol{\theta} = \tilde{\mathbf{h}}$ (cf. (3.24)) can be obtained by

$$f(\tilde{\mathbf{y}} | \tilde{\mathbf{h}}) = \int f(\tilde{\mathbf{y}} | \mathbf{s}, \tilde{\mathbf{h}}) f(\mathbf{s}) d\mathbf{s}, \quad (3.34)$$

where $f(\mathbf{s})$ is the marginal pdf of the input vector and $f(\tilde{\mathbf{y}} | \mathbf{s}, \tilde{\mathbf{h}})$ is the likelihood function when the input is known. Assume, for example, that the input data symbol $s(n)$ takes, with equal probability, a finite number (\bar{K}) of values. Consequently, the input data vector \mathbf{s} also takes values from the signal set $\{\mathbf{s}_1, \dots, \mathbf{s}_{\bar{K}}\}$. The likelihood function of the channel parameters is then given by

$$f(\tilde{\mathbf{y}} | \tilde{\mathbf{h}}) = \sum_{i=1}^{\bar{K}} f(\tilde{\mathbf{y}} | \mathbf{s}_i, \tilde{\mathbf{h}}) \text{Prob}(\mathbf{s} = \mathbf{s}_i) = c \sum_{i=1}^{\bar{K}} \exp \left\{ -\frac{\|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s}_i)\tilde{\mathbf{h}}\|^2}{2\sigma^2} \right\}, \quad (3.35)$$

where c is a constant, $\|\tilde{\mathbf{y}}\|^2 := \tilde{\mathbf{y}}^H \tilde{\mathbf{y}}$, and the stochastic maximum likelihood estimator is given by

$$\hat{\mathbf{h}} = \arg \min_{\tilde{\mathbf{h}}} \sum_{i=1}^K \exp \left\{ -\frac{\|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s}_i)\tilde{\mathbf{h}}\|^2}{2\sigma^2} \right\}. \quad (3.36)$$

The maximization of the likelihood function defined in (3.34) is in general difficult because $f(\tilde{\mathbf{y}} | \theta)$ is non-convex. The Expectation-Maximization (EM) algorithm can be applied to transform the complicated optimization to a sequence of quadratic optimizations. Kaled and Vallet [33] first applied the EM algorithm to the equalization of communication channels with input sequence having finite alphabet property. By using a *Hidden Markov Model* (HMM), they developed a batch (off-line) procedure that includes the so-called forward and backward recursions. The complexity of this algorithm increases exponentially with the channel memory.

To relax the memory requirements and facilitate channel tracking, “on-line” sequential approaches have been proposed in [34] for input with finite alphabet properties under a HMM formulation. Given the appropriate regularity conditions and a good initialization guess, it can be shown that these algorithms converge to the true channel value.

Deterministic Maximum Likelihood Estimation. The deterministic ML approach assumes no statistical model for the input sequence $\{s(n)\}$. In other words, both the channel vector $\tilde{\mathbf{h}}$ and the input source vector \mathbf{s} are parameters to be estimated. When the noise is zero-mean Gaussian with covariance $\sigma_v^2 \mathbf{I}_{TBN}$, the ML estimates can be obtained by the nonlinear least squares optimization

$$\{\hat{\mathbf{h}}, \hat{\mathbf{s}}\} = \arg \min \|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}}\|^2. \quad (3.37)$$

The joint minimization of the likelihood function with respect to both the channel and the source parameter spaces is difficult. Fortunately, the observation vector $\tilde{\mathbf{y}}$ is linear in both the channel and the input parameters individually. In particular, we have

$$\tilde{\mathbf{y}} = \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}} + \tilde{\mathbf{v}} = \mathcal{F}(\tilde{\mathbf{h}})\mathbf{s} + \tilde{\mathbf{v}}, \quad (3.38)$$

where

$$\mathcal{F}(\tilde{\mathbf{h}}) = \begin{pmatrix} \mathbf{h}(0) & \cdots & \mathbf{h}(L) \\ \ddots & \ddots & \ddots \\ \mathbf{h}(0) & \cdots & \mathbf{h}(L) \end{pmatrix} \quad (3.39)$$

is the so-called filtering matrix. We therefore have a separable nonlinear least squares problem that can be solved sequentially

$$\{\hat{\mathbf{h}}, \hat{\mathbf{s}}\} = \arg \min_{\mathbf{s}} \{\min_{\tilde{\mathbf{h}}} \|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}}\|^2\} = \arg \min_{\tilde{\mathbf{h}}} \{\min_{\mathbf{s}} \|\tilde{\mathbf{y}} - \mathcal{F}(\tilde{\mathbf{h}})\mathbf{s}\|^2\}. \quad (3.40)$$

If we are only interested in estimating the channel, the above minimization can be rewritten as

$$\hat{\mathbf{h}} = \arg \min_{\tilde{\mathbf{h}}} \| \underbrace{(\mathbf{I} - \mathcal{F}(\tilde{\mathbf{h}})\mathcal{F}^\dagger(\tilde{\mathbf{h}})) \tilde{\mathbf{y}} \|_2^2} = \arg \min_{\tilde{\mathbf{h}}} \|\mathcal{P}(\tilde{\mathbf{h}})\tilde{\mathbf{y}}\|^2, \quad (3.41)$$

where $\mathcal{P}(\tilde{\mathbf{h}})$ is a projection transform of $\tilde{\mathbf{y}}$ into the orthogonal complement of the range space of $\mathcal{F}(\tilde{\mathbf{h}})$, or the noise subspace of the observation, and $\mathcal{F}^\dagger(\tilde{\mathbf{h}})$ denotes the pseudo-inverse of $\mathcal{F}(\tilde{\mathbf{h}})$. Discussions of algorithms of this type can be found in [35].

Similar to the HMM for statistical maximum likelihood approach, the finite alphabet properties of the input sequence can also be incorporated into the deterministic maximum likelihood methods. These algorithms, first proposed by Seshadri [36] and Ghosh and Weber [37], iterate between estimates of the channel and the input. At iteration k , with an initial guess of the channel $\tilde{\mathbf{h}}^{(k)}$, the algorithm estimates the input sequence $\mathbf{s}^{(k)}$ and the channel $\tilde{\mathbf{h}}^{(k+1)}$ for the next iteration by

$$\mathbf{s}^{(k)} = \arg \min_{\mathbf{s} \in \mathcal{S}} \|Y - \mathcal{F}(\tilde{\mathbf{h}}^{(k)})\mathbf{s}\|^2, \quad (3.42)$$

$$\tilde{\mathbf{h}}^{(k+1)} = \arg \min_{\tilde{\mathbf{h}}} \|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s}^{(k)})\tilde{\mathbf{h}}\|^2, \quad (3.43)$$

where \mathcal{S} is the (discrete) domain of \mathbf{s} . The optimization in (3.43) is a linear least squares problem whereas the optimization in (3.42) can be achieved by using the Viterbi algorithm [1]. Seshadri [36] presented blind trellis search techniques. Reduced-state sequence estimation was proposed in [37]. Raheli et al. proposed a per-survivor processing technique in [38]. The convergence of such approaches is not guaranteed in general. Interesting examples have been provided in [39] where two different combinations of $\tilde{\mathbf{h}}$ and \mathbf{s} lead to the same cost $\|\tilde{\mathbf{y}} - \mathcal{T}(\mathbf{s})\tilde{\mathbf{h}}\|^2$.

2.03.3.2.2 The methods of moments

Although the ML channel estimator discussed in Section 2.03.3.2.1 usually provides better performance, the computation complexity and the existence of local optima are the two major difficulties. Therefore, “simpler” approaches have also been investigated.

SISO Channel Estimation. For baud-rate data, second-order statistics of the data do not carry enough information to allow estimation of the channel impulse response as a typical channel is nonminimum-phase. On the other hand, higher order statistics (in particular, fourth-order cumulants) of the baud-rate (or fractional rate) data can be exploited to yield the channel estimates to within a scale factor. Given the mathematical model (3.9), there are two broad classes of direct approaches to channel estimation, the distinguishing feature among them being the choice of the optimization criterion. All of the approaches involve (more or less) a least-squares error measure. The error definition differs, however, as follows:

- *Fitting error:* Match the model-based higher-order (typically fourth-order) statistics to the estimated (data-based) statistics in a least-squares sense to estimate the channel impulse response, as in [40, 41], for example. This approach allows consideration of noisy observations. In general, it results in a nonlinear optimization problem. It requires availability of a good initial guess to prevent convergence to a local minimum. It yields estimates of the channel impulse response.
- *Equation error:* It is based on minimizing an “equation error” in some equation which is satisfied ideally. The approaches of [42, 43] (among others) fall in this category. In general, this class of approaches results in a closed-form solution for the channel impulse response so that a global extremum is always guaranteed provided that the channel length (order) is known. These approaches may also provide good initial guesses for the nonlinear fitting error approaches. Quite a few of these approaches fail if the channel length is unknown.

Further details may be found in [44] and references therein.

SIMO Channel Estimation. Here we will concentrate upon second-order statistical methods. For single-input multiple-output vector channels the autocorrelation function of the observation is sufficient for the identification of the channel impulse response up to an unknown constant [45, 46], provided that the various subchannels have no common zeros. This observation led to a number of techniques under both statistical and deterministic assumptions of the input sequence [35]. By exploiting the multichannel aspects of the channel (e.g., crosscorrelation among the outputs of various subchannels), many of these techniques lead to a constrained quadratic optimization

$$\hat{\mathbf{h}} = \arg \min_{\|\tilde{\mathbf{h}}\|=1} \tilde{\mathbf{h}}^H Q(\tilde{\mathbf{y}}) \tilde{\mathbf{h}}, \quad (3.44)$$

where $Q(\tilde{\mathbf{y}})$ is a positive definite matrix constructed from the observation. Asymptotically (either as the sample size increases to infinity or the noise variance approaches to zero), these estimates converge to true channel parameters.

2.03.3.3 Semi-blind approaches

Semi-blind approaches utilize a combination of training-based and blind approaches. Here we present a brief discussion about the idea and refer the reader to the survey [47] for details. The objective of semi-blind channel estimation (and equalization) is to exploit the information used by blind methods as well as the information exploited by the training-based methods. Semi-blind channel estimation assumes additional knowledge of the input sequence. Specifically, part of the input data vector is known. Both the statistical and deterministic maximum likelihood estimators remain the same except that the likelihood function needs to be modified to incorporate the knowledge of the input. However, semi-blind channel estimation may offer significant performance improvement over either the blind or the training based methods as demonstrated in the evaluation of Cramer-Rao lower bound in [47].

There are many generalizations of blind channel estimation techniques to incorporate known symbols. In [48], Tsatsanis and Cirpan extended the approach of Kaleh and Vallet by restricting the transition of hidden Markov model. In [49], the knowledge of the known symbol is used to avoid the local maxima in the maximization of the likelihood function. A popular approach is to combine the objective function used to derive blind channel estimator with the least squares cost in the training-based channel estimation. For example, a weighted linear combination of the cost for blind channel estimator and that for the training based estimator can be used [50, 51].

2.03.3.4 Superimposed training-based approaches

In the superimposed training (hidden pilots) based approach, one takes

$$s(n) = b(n) + c(n), \quad (3.45)$$

where $\{b(n)\}$ is the information sequence and $\{c(n)\}$ is a non-random periodic training (pilot) sequence. Exploitation of the periodicity of $\{c(n)\}$ allows identification of the channel without allocating any explicit time slots for training, unlike traditional training methods. There is no loss in data transmission rate. On the other hand, some useful power is wasted in superimposed training which could have otherwise been allocated to the information sequence. This lowers the effective signal-to-noise ratio (SNR) for the information sequence and affects the bit error rate (BER) at the receiver.

Superimposed training-based approaches have been discussed in [52–54] for SISO systems. A block transmission method has been proposed in [55,56] where a data-dependent component is added to the superimposed training such that interference due to data (information sequence) is greatly reduced in channel estimation at the receiver. This method is applicable to time-invariant channels only and it requires “data-blocking” for block transmissions and insertion of a cyclic prefix in each data block. Its extension to a class of time-variant channels is given in [57]; see also [58]. The UTRA specification for 3G systems [59] allows to adopt a spread pilot (superimposed) sequence in the base station’s common pilot channel, which is suitable for downlinks. Periodic superimposed training for channel estimation via first-order statistics for SISO systems have been discussed in [60–63]. In [64] performance bounds for training and superimposed training-based semiblind SISO channel estimation for time-varying flat fading channels have been discussed.

2.03.3.5 MIMO channel estimation

All of the channel estimation approaches described earlier apply to MIMO channels; however, efficacy of the approaches depends upon the underlying analytical model used. For MIMO channel estimation in correlated fading environments, Chen and su [27] presents two analytical MIMO channel models and low-complexity iterative channel estimation methods based on these models, exploiting the optimal training sequences proposed in [65,66]. Further details may be found in these papers and references therein.

2.03.4 Equalization

A communication channel is typically modeled as a linear system whose output is corrupted by additive noise. Equalizers are designed to compensate for channel distortions as well as noise. One may directly design an equalizer given the received signal, or one may first estimate the channel impulse response and then design an equalizer based on the estimated channel. The structure of the equalizer is dictated by channel models, computational complexity and possible exchange of information with a channel (error correction) decoder.

2.03.4.1 Linear equalization

The most common channel equalizer structure is a linear transversal filter. Given model (3.9) for the baud-rate sampled received signal, the linear transversal equalizer output $\hat{s}(k)$ is an estimate of $s(k)$ given by

$$\hat{s}(k) = \sum_{n=-N_e}^{N_e} c(n)y(k-n), \quad (3.46)$$

where $\{c(n)\}_{n=-N_e}^{n=N_e}$ are the $(2N_e + 1)$ -tap weight (equalizer) coefficients of the $(2N_e + 1)$ -tap equalizer. Two criteria have found widespread use in optimizing the equalizer coefficients: peak distortion criterion and mean-square error (MSE) criterion. Under the MSE criterion one chooses $c(n)$ s to minimize $E\{|\hat{s}(k) - s(k)|^2\}$. Linear equalizers designed on the basis of the baud-rate sampled received signal are quite sensitive to symbol timing errors [25]. Therefore, fractionally-spaced linear equalizers (typically

with twice the baud-rate sampling: oversampling by a factor of two) are quite widely used to mitigate sensitivity to symbol timing errors. A fractionally spaced equalizer (FSE) in the linear transversal structure has the output

$$\hat{s}(k) = \sum_{n=-N_e}^{N_e} \left(\sum_{i=1}^N c_i(n) y_i(k-n) \right), \quad (3.47)$$

where $\{c_i(n)\}_{n=-N_e}^{N_e}$ are the $(2N_e + 1)$ tap weight coefficients of the i th sub-equalizer. Note that the FSE outputs data at the symbol rate. Various criteria and cost functions exist to design the linear equalizers in both batch and recursive (adaptive) form [1, Chapters 10, 14].

For time-varying channels the equalizer coefficients $c(n)$ or $c_i(n)$ in (3.46) and (3.47), respectively, are also functions of k . An alternative is to use either a BEM-based equalizer with time-invariant equalizer coefficients coupled with time-varying exponential basis function [67] or a Kalman fixed-lag smoother (Kalman Detector) [68]. To illustrate the Kalman detector, consider a multi-input multi-output (MIMO) channel with K inputs (users, antennas, ...) and N outputs (receivers, antennas, ...); one can adapt this easily to single-input single-output (SISO) or single-input multi-output (SIMO) systems. Let $\{s_k(n)\}$ denote k th user's information sequence that is input to the time-varying channel with discrete-time response $\{\mathbf{h}_k(n; l)\}$ (channel response for the k th user at time instance n to a unit input at time instance $n - l$). We assume that $\{s_k(n)\}$ s are mutually independent and identically distributed (i.i.d.) with zero mean and variance $E\{s_k(n) s_k^*(n)\} = \sigma_{s_k}^2 = \sigma_s^2$ for $k = 1, 2, \dots, K$. Then, at symbol-rate sampling, the noisy N -column channel output vector is given by ($n = 0, 1, \dots$)

$$\mathbf{y}(n) = \sum_{k=1}^K \sum_{l=0}^L \mathbf{h}_k(n; l) s_k(n-l) + \mathbf{v}(n), \quad (3.48)$$

where the N -column vector $\mathbf{v}(n)$ is zero-mean, white, uncorrelated with $s_k(n)$, complex Gaussian noise, with the autocorrelation $E\{\mathbf{v}(n + \tau) \mathbf{v}^H(n)\} = \sigma_v^2 \mathbf{I}_N \delta(\tau)$. Define

$$\begin{aligned} \mathbf{s}(n) &:= [s_1(n) \ s_2(n) \ \cdots \ s_K(n)]^T \\ \mathbf{H}(n; l) &:= [\mathbf{h}_1(n; l) \ \mathbf{h}_2(n; l) \ \cdots \ \mathbf{h}_K(n; l)]. \end{aligned}$$

Then we may rewrite (3.48) as

$$\mathbf{y}(n) = \sum_{l=0}^L \mathbf{H}(n; l) \mathbf{s}(n-l) + \mathbf{v}(n). \quad (3.49)$$

2.03.4.1.1 Kalman detector (KD)

The Kalman filter, together with a quantizer, acts as the symbol detector at the receiver end. The state and the measurement equations are given by

$$\mathbf{s}_d(n) = \Phi \mathbf{s}_d(n-1) + \boldsymbol{\Gamma} \bar{\mathbf{s}}(n) + \boldsymbol{\Gamma} \tilde{\mathbf{s}}(n), \quad (3.50)$$

$$\mathbf{y}(n) = \tilde{\mathbf{H}}_d(n) \mathbf{s}_d(n) + \mathbf{v}(n), \quad (3.51)$$

with the following definitions

$$\begin{aligned}\mathbf{s}_d(n) &:= [\mathbf{s}^T(n) \ \mathbf{s}^T(n-1) \ \cdots \ \mathbf{s}^T(n-d)]^T, \\ \bar{\mathbf{s}}(n) &:= E\{\mathbf{s}(n)\}, \quad \tilde{\mathbf{s}}(n) := \mathbf{s}(n) - \bar{\mathbf{s}}(n), \\ \boldsymbol{\Phi} &:= \begin{bmatrix} \mathbf{0}_d^T & 0 \\ \mathbf{I}_d & \mathbf{0}_d \end{bmatrix} \otimes \mathbf{I}_K, \quad \boldsymbol{\Gamma} := [1 \ \mathbf{0}_d^T]^T \otimes \mathbf{I}_K, \\ \tilde{\mathbf{H}}_d(n) &:= [\hat{\mathbf{H}}(n; 0) \ \hat{\mathbf{H}}(n; 1) \ \cdots \ \hat{\mathbf{H}}(n; L) \ \mathbf{0}_{N \times K(d-L)}]\end{aligned}$$

where $\mathbf{s}(n)$ is K -column vector of symbols (data or training), $\hat{\mathbf{H}}(n; l)$ is the known or estimated $N \times K$ channel matrix and integer $d \geq L$ (it will also be the equalization delay). Assume data symbols are zero-mean and white. If $s(n)$ is a data symbol, we have $\bar{s}(n) = 0$, $\tilde{s}(n) = s(n)$ and $\sigma_s^2(n) = \sigma_s^2$; if $s(n)$ is a training symbol, $\bar{s}(n) = s(n)$, $\tilde{s}(n) = 0$ and $\sigma_s^2(n) = 0$.

Kalman filtering for the system described by (3.50) and (3.51) is initialized with

$$\hat{\mathbf{s}}_d(-1|-1) = \mathbf{0}_{K(d+1)} \text{ and } \mathbf{P}(-1|-1) = \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T,$$

where $\hat{\mathbf{s}}_d(k|m)$ denotes the estimate of $\mathbf{s}_k(p)$ given the observations $\{\mathbf{y}(n)\}_{n=0}^m$, and $\mathbf{P}(k|m)$ denotes the error covariance matrix of $\hat{\mathbf{s}}_d(k|m)$, defined as

$$\mathbf{P}(k|m) := E\{[\hat{\mathbf{s}}_d(k|m) - \mathbf{s}_d(k)][\hat{\mathbf{s}}_d(k|m) - \mathbf{s}_d(k)]^H\}.$$

Then recursive filtering (for $n = 0, 1, \dots$) is applied via the following steps:

1. Time update:

$$\begin{aligned}\hat{\mathbf{s}}_d(n | n-1) &= \boldsymbol{\Phi} \hat{\mathbf{s}}_d(n-1 | n-1) + \boldsymbol{\Gamma} \bar{\mathbf{s}}(n), \\ \mathbf{P}(n | n-1) &= \boldsymbol{\Phi} P(n-1 | n-1) \boldsymbol{\Phi}^T + \sigma_s^2(n) \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T;\end{aligned}$$

2. Kalman gain:

$$\begin{aligned}\mathbf{P}_\eta(n) &= \sigma_v^2 \mathbf{I}_N + \tilde{\mathbf{H}}_d(n) \mathbf{P}(n | n-1) \tilde{\mathbf{H}}_d^H(n), \\ \mathbf{K}(n) &= \mathbf{P}(n | n-1) \tilde{\mathbf{H}}_d^H(n) \mathbf{P}_\eta^{-1}(n);\end{aligned}$$

3. Measurement update:

$$\begin{aligned}\hat{\mathbf{s}}_d(n | n) &= \hat{\mathbf{s}}_d(n | n-1) + \mathbf{K}(n) \left\{ \mathbf{y}(n) - \tilde{\mathbf{H}}_d(n) \hat{\mathbf{s}}_d(n | n-1) \right\}, \\ \mathbf{P}(n | n) &= \left\{ \mathbf{I}_{K(d+1)} - \mathbf{K}(n) \tilde{\mathbf{H}}_d(n) \right\} \mathbf{P}(n | n-1).\end{aligned}$$

The estimated state vector is given by

$$\hat{\mathbf{s}}_d(n | n) = [\hat{\mathbf{s}}^T(n | n) \ \hat{\mathbf{s}}^T(n-1 | n) \ \cdots \ \hat{\mathbf{s}}^T(n-d | n)]^T$$

and we extract its last (K -column vector) term $\hat{\mathbf{s}}(n-d | n)$ as the desired equalized output for K -users with equalization delay d . Finally, we hard-quantize $\hat{\mathbf{s}}(n-d | n)$ to acquire the detected symbols.

When using the estimated channel, one may rewrite the received signal (3.48) as

$$\mathbf{y}(n) = \sum_{l=0}^L \hat{\mathbf{h}}(n; l) \mathbf{s}(n-l) + \underbrace{\sum_{l=0}^L [\mathbf{h}(n; l) - \hat{\mathbf{h}}(n; l)] \mathbf{s}(n-l)}_{=: \tilde{\mathbf{v}}(n)} + \mathbf{v}(n), \quad (3.52)$$

where the “effective” noise is $\tilde{\mathbf{v}}(n)$ instead of $\mathbf{v}(n)$. In order to compensate for this channel estimation error, as a first-order approximation, one may wish to take the variance of $\tilde{\mathbf{v}}(n)$ in (3.52) to be larger than σ_v^2 , the variance of $\mathbf{v}(n)$.

2.03.4.2 Decision feedback equalization

Linear equalizers do not perform well when the underlying channels have deep spectral nulls in the passband. Several nonlinear equalizers have been developed to deal with such channels. One of them *Decision Feedback Equalizer* (DFE) is a nonlinear equalizer that employs previously detected symbols to eliminate the ISI due to the previously detected symbols on the current symbol to be detected. The use of the previously detected symbols makes the equalizer output a nonlinear function of the data. DFE can be symbol-spaced or fractionally spaced.

In [69, 70] MMSE design of finite-length DFEs have been considered for time-invariant channels. Their approach extends trivially to time-varying channels. We now discuss application of their approach to model (3.48).

The DFE structure is shown in Figure 3.3 to equalize the delayed symbols $\mathbf{s}(n-d)$, with the feed-forward (FF) and feed-back (FB) filters. Since each measurement $\mathbf{y}(n)$ contains inter-symbol-interference (ISI) caused by prior symbols, DFE is designed to reduce ISI and to recover $\mathbf{s}(n)$ using FIR filters. The FF filter takes current and prior measurements \mathbf{y}_f as its input to get information correlated with ISI and to remove its effect.

Stack the inputs of the FF filter with l_f taps at time n into a “tall” vector

$$\mathbf{y}_f(n) := [\mathbf{y}^T(n) \ \mathbf{y}^T(n-1) \ \cdots \ \mathbf{y}^T(n-l_f+1)]^T,$$

where $\mathbf{y}(n)$ is N -column vector and also define $\mathbf{v}_f(n)$ likewise. Then, the received signal is given as

$$\mathbf{y}_f(n) = \mathbf{H}(n)\mathbf{s}_f(n) + \mathbf{v}_f(n), \quad (3.53)$$

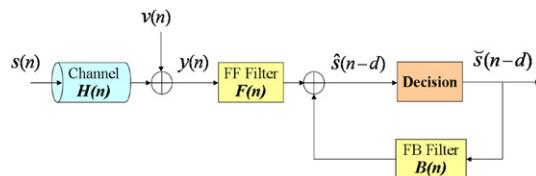


FIGURE 3.3

Decision-feedback equalizer (DFE).

with

$$\mathbf{H}(n) := \begin{bmatrix} \mathbf{h}(n; 0) & \cdots & \mathbf{h}(n; L) \\ \ddots & \ddots & \ddots \\ \mathbf{h}(n - l_f + 1; 0) & \cdots & \mathbf{h}(n - l_f + 1; L) \end{bmatrix}$$

$$\mathbf{s}_f(n) := [\mathbf{s}^T(n) \ \mathbf{s}^T(n - 1) \ \cdots \ \mathbf{s}^T(n - l_f - L + 1)]^T$$

where $\mathbf{h}(n; l)$ is $N \times K$ matrix channel response at time n to a unit input at time $n - l$ and $\mathbf{s}(n)$ is K -column vector. As shown in Figure 3.3, the input to the FB filter comes from the decision output, denoted by $\check{\mathbf{s}}_f(n - d)$. The FB filter uses prior symbol decisions to cancel the trailing ISI by mapping the estimate $\hat{\mathbf{s}}_f(n - d)$ to the closest point in the symbol constellation. We define the input vector of FB filter with l_b taps as

$$\mathbf{s}_b(n) := [\check{\mathbf{s}}^T(n - d) \ \check{\mathbf{s}}^T(n - d - 1) \ \cdots \ \check{\mathbf{s}}^T(n - d - l_b)]^T.$$

The estimate of the information symbol, $\hat{\mathbf{s}}(n - d)$ is obtained by combining the outputs of FF and FB filters and can be written at time n with delay d as

$$\hat{\mathbf{s}}(n - d) = \sum_{i=0}^{l_f-1} \mathbf{F}_i^T(n) \mathbf{y}(n - i) - \sum_{j=1}^{l_b} \mathbf{B}_j(n) \check{\mathbf{s}}(n - d - j), \quad (3.54)$$

where $\mathbf{F}_i(n)$ s (that are $N \times K$ matrices) and $\mathbf{B}_j(n)$ s (that are $K \times K$ matrices) are the taps of FF and FB time-varying filters at time n , and $\check{\mathbf{s}}(n - d - j)$ is the hard decision of $\hat{\mathbf{s}}(n - d - j)$. The estimate $\hat{\mathbf{s}}(n - d)$ is also fed into the quantizer to obtain the symbol decision $\check{\mathbf{s}}(n - d)$. Let $\mathbf{F}(n)$ and $\mathbf{B}(n)$ denote the vectors of time-varying taps of FF and FB filters,

$$\mathbf{F}(n) := [\mathbf{F}_0^T(n) \ \mathbf{F}_1^T(n) \ \cdots \ \mathbf{F}_{l_f-1}^T(n)]^T,$$

$$\mathbf{B}(n) := [\mathbf{I} \ \mathbf{B}_1(n) \ \mathbf{B}_2(n) \ \cdots \ \mathbf{B}_{l_b}(n)]^T,$$

then the error signal is given by

$$\check{\mathbf{s}}(n - d) = \mathbf{s}(n - d) - \hat{\mathbf{s}}(n - d) = \mathbf{B}(n) \mathbf{s}_b(n) - \mathbf{F}(n) \mathbf{y}_f(n). \quad (3.55)$$

Assuming the decisions $\{\check{\mathbf{s}}(n)\}$ are correct and equal to $\{\mathbf{s}(n)\}$, we can solve a nonlinear optimization problem which minimizes the variance of the error signal in (3.55),

$$\min_{\{\mathbf{F}(n), \mathbf{B}(n)\}} E \left\{ |\mathbf{B}(n) \mathbf{s}_b(n) - \mathbf{F}(n) \mathbf{y}_f(n)|^2 \right\}. \quad (3.56)$$

Solve a standard linear least-mean-squares estimation problem over $\mathbf{F}(n)$ with $\mathbf{B}(n)$ fixed and then we have a constrained optimization problem; note that the leading entry of $\mathbf{B}(n)$ is the identity matrix. Therefore, the FF and the FB time-varying filters of the MMSE-DFE are given by [69]

$$\mathbf{B}_{\text{MMSE}}(n) = \mathbf{R}_\delta^{-1} \boldsymbol{\Psi} (\boldsymbol{\Psi}^T \mathbf{R}_\delta^{-1} \boldsymbol{\Psi})^{-1}, \quad (3.57)$$

$$\mathbf{F}_{\text{MMSE}}(n) = \mathbf{R}_{yy}^{-1}(n) \mathbf{R}_{sy}^H(n) \mathbf{B}_{\text{MMSE}}(n), \quad (3.58)$$

where

$$\begin{aligned}\Psi &:= [1 \ 0 \ 0 \ \cdots \ 0]^T \otimes \mathbf{I}_K, \\ \mathbf{R}_\delta &:= \mathbf{R}_{ss}(n) - \mathbf{R}_{sy}(n)\mathbf{R}_{yy}^{-1}(n)\mathbf{R}_{sy}^H(n) \\ &= \Phi \left[\frac{1}{\sigma_v^2} \mathbf{H}(n)\mathbf{H}^H(n) + \frac{1}{\sigma_s^2} \mathbf{I}_{Nl_f} \right]^{-1} \Phi^H.\end{aligned}$$

By the assumption that $\{\mathbf{s}(n)\}$ are independent and identically distributed (i.i.d.) with variance σ_s^2 , and based on (3.53), we have

$$\begin{aligned}\mathbf{R}_{ss}(n) &:= E \left\{ \mathbf{s}_b(n)\mathbf{s}_b^H(n) \right\} = \sigma_s^2 \mathbf{I}_{(l_b+1)}, \\ \mathbf{R}_{sy}(n) &:= E \left\{ \mathbf{s}_b(n)\mathbf{y}_f^H(n) \right\} = \sigma_s^2 \Phi H^H(n), \\ \mathbf{R}_{yy}(n) &:= E \left\{ \mathbf{y}_f(n)\mathbf{y}_f^H(n) \right\} = \sigma_s^2 \mathbf{H}(n)\mathbf{H}^H(n) + \sigma_v^2 \mathbf{I}_{Nl_f},\end{aligned}$$

where

$$\mathbf{s}_b(n) = \Phi \mathbf{s}_f(n), \quad \Phi := \left[\mathbf{0}_{(l_b+1) \times d} \ \mathbf{I}_{l_b+1} \ \mathbf{0}_{(l_b+1) \times (l_f+L-d-l_b-1)} \right] \otimes \mathbf{I}_K.$$

Using (3.57) and (3.58) in (3.54), we have the symbol estimate $\{\hat{\mathbf{s}}(n-d)\}$.

2.03.4.3 Maximum likelihood sequence detection

Maximum Likelihood Sequence Detector (MLSD) estimates the information sequence to maximize the joint probability of the received sequence conditioned on the information sequence. It is sequence estimator compared to the linear equalizers and DFE which are symbol-by-symbol detectors. For a scalar system a detailed discussion may be found in [1]; for MIMO channels see [71, Section 7.8]. The optimal equalization methods for minimizing sequence error rate or the bit error rate (BER) are based on MAP (maximum *a posteriori*) estimation, which turns into maximum likelihood (ML) estimation when the transmitted symbols are equally likely. For instance, see Viterbi algorithm (VA) [1, 72, 73] for ML sequence estimation and BCJR algorithm [74] for MAP sequence estimation. The MAP/ML-based solutions often suffer from high computational load for channels with long memory or large constellation sizes. For model (3.9), the ML estimate of the channel input sequence $s(n)$ based on a sequence of channel output $\{y(n)\}_{n=1}^{T_B}$ and given knowledge of the channel, can be obtained by maximizing the likelihood function, or equivalently, by minimizing

$$\sum_{n=1}^{T_B} \left| y(n) - \sum_{l=0}^N h(l)s(n-l) \right|^2. \quad (3.59)$$

If the symbol alphabet size is M , then the Viterbi algorithm can be implemented by denoting M^L states as all possible L -tuples of $(s(n), s(n-1), \dots, s(n-L+1))$. The trellis is determined by the symbol alphabet \mathcal{S} while the metrics of the Viterbi algorithm depend upon the channel.

2.03.4.4 Turbo equalization

In turbo equalization one exploits the information obtained from a channel decoder to improve equalization. A practical digital communication system has a forward error correction (FEC) channel encoder at the transmitter which adds redundancies to the information symbols before transmitting the encoded sequence over the channel. At the receiver, compensation for channel distortions is the task of equalizers while subsequent recovery of the data symbols from the equalized (and quantized) symbols making use of the FEC encoding is the task for the channel decoder. Typically these two tasks are considered separately (to reduce computational complexity) with limited interaction between the two [75]. An optimal joint processing of the equalization and decoding steps is usually impossible due to complexity considerations. A number of iterative receiver algorithms repeat the equalization and decoding tasks on the same set of received data, where feedback information from the decoder is incorporated into the equalization process. This method, called turbo equalization, was originally developed for concatenated convolutional codes (turbo code [76]) and is now adapted to various communication problems.

Communicating soft information probability distribution between the equalizer and the decoder, instead of hard information (symbol estimates only), improves the BER performance but usually requires more complex decoding algorithms. State-of-the-art systems for a variety of communication channels employ convolutional codes and ML equalizers together with an interleaver after the encoder and a deinterleaver before the decoder [75]. Interleaving shuffles symbols within a given time frame or block of data and thus decorrelates error events introduced by the equalizer between neighboring symbols. The MAP/ML-based solutions often suffer from high computational load for channels with long memory or large constellation sizes (expensive equalizer) or convolutional codes with long memory (expensive decoder). This situation is exacerbated by the need to perform equalization and decoding several times for each block of data. A major research issue is thus the complexity reduction of such iterative algorithms.

2.03.4.4.1 Principle of turbo equalization

Consider a simple transmitter where a sequence of data $\mathbf{b}(n') = [b^1(n'), b^2(n'), \dots, b^{k_0}(n')] \in \{1, 0\}^{k_0}$ is encoded to the code symbols $\mathbf{c}(n') = [c^1(n'), c^2(n'), \dots, c^{n_0}(n')] \in \{1, 0\}^{n_0}$ with a code rate $R_c = k_0/n_0$, which is interleaved in a block and then mapped into binary phase shift keying (BPSK) symbols. Figure 3.4 depicts the receiver structure for turbo equalization [77], where both the soft-input soft-output (SISO_f: we use subscript f to distinguish between single-input single-output abbreviated as SISO and soft-input soft-output abbreviated as SISO_f) equalizer and SISO_f decoder are of the MAP type. The extrinsic log-likelihood ratios (LLRs), $\mathbf{L}_e\{\mathbf{c}(\cdot)\}$ are transferred iteratively between the equalizer

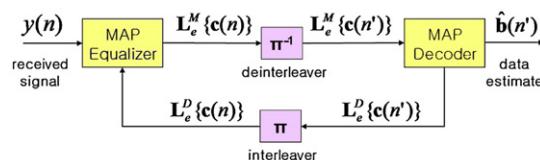


FIGURE 3.4

A turbo equalization receiver.

and decoder. [The subscript “ e ” for representing “extrinsic,” the superscript “ E ” and “ D ” for output of “Equalizer” and “Decoder” respectively]. The MAP equalizer computes the *a posteriori* probabilities (APPs) given T_r received symbols, $P\{c^i(n) = b \mid \mathbf{y}(l), 1 \leq l \leq T_r\}, b \in \{1, 0\}$ and generates the extrinsic LLR as (*a posteriori* LLR – *a priori* LLR)

$$L_e^E\{c^i(n)\} := \ln \frac{P\{c^i(n) = 1 \mid \mathbf{y}(l), 0 \leq l < T_r\}}{P\{c^i(n) = 0 \mid \mathbf{y}(l), 0 \leq l < T_r\}} - \underbrace{\ln \frac{P\{c^i(n) = 1\}}{P\{c^i(n) = 0\}}}_{=: L\{c^i(n)\}}. \quad (3.60)$$

The *a priori* LLR of the MAP equalizer, $L\{c^i(n)\}$ is provided by the interleaved output of the MAP decoder at the previous iteration but $L\{c^i(n)\} = 0$ for the first iteration. The extrinsic LLRs $\mathbf{L}_e^E\{\mathbf{c}(n)\}$ produced by the MAP demodulator is sent to the MAP decoder as the *a priori* LLRs for channel decoding. Based on the *a priori* LLRs and the channel code constraints, the MAP decoder computes the APPs $P\{c^i(n') = b \mid \mathbf{L}_e^E\{\mathbf{c}(l)\}, 1 \leq l \leq T_r\}, b \in \{1, 0\}$ and generates the extrinsic LLR as

$$L_e^D\{c^i(n')\} := \ln \frac{P\{c^i(n') = 1 \mid \mathbf{L}\{\mathbf{c}(l)\}, 0 \leq l < T_r\}}{P\{c^i(n') = 0 \mid \mathbf{L}\{\mathbf{c}(l)\}, 0 \leq l < T_r\}} - \underbrace{\ln \frac{P\{c^i(n') = 1\}}{P\{c^i(n') = 0\}}}_{=: L\{c^i(n')\}}. \quad (3.61)$$

The extrinsic output of the MAP decoder is fed back to the MAP equalizer iteratively. Note that (3.60) and (3.61) are valid only if the *a posteriori* outputs are independent of the *a priori* inputs for the equalizer and the decoder. Assuming ideal interleaver between the equalizer and the decoder, we can apply the turbo principle and the correct ordering of the LLRs $\mathbf{L}\{\mathbf{c}(n')\} = \pi^{-1}[\mathbf{L}_e^E\{\mathbf{c}(n)\}]$ and $\mathbf{L}\{\mathbf{c}(n)\} = \pi[\mathbf{L}_e^D\{\mathbf{c}(n')\}]$, which are input to the MAP decoder and the MAP equalizer respectively. The MAP decoder also compute the *a posteriori* probabilities of the input data bit and then the data estimate $\hat{\mathbf{b}}(n') = [\hat{b}^1(n'), \hat{b}^2(n'), \dots, \hat{b}^{k_0}(n')]$ as

$$\hat{b}^i(n') = \arg \max_{b \in \{1, 0\}} P\{b^i(n') = b \mid \mathbf{L}\{\mathbf{c}(l)\}, 1 \leq l \leq T_r\}. \quad (3.62)$$

By combining a MAP equalizer and a MAP decoder, and exchanging probabilistic information about data symbols iteratively, turbo equalization usually can achieve close-to-optimal performance but much lower complexity [77, 78]. In [79], a turbo-equalization-like system using linear equalizers based on soft interference cancellation and linear minimum mean-square error (MMSE) filtering is proposed as part of a multiuser detector for CDMA. Based on this work, a variety of SISO_f equalizers employing linear MMSE and decision feedback equalization (DFE) are proposed in [75, 80, 81].

2.03.4.4.2 Turbo equalization for doubly-selective channels

For doubly-selective channels, an adaptive SISO_f equalizer has been presented in [5], using extended Kalman filter (EKF) to incorporate channel estimation into the equalization process. This adaptive soft nonlinear Kalman equalizer takes the soft decisions of data symbols from the SISO_f decoder as its *a priori* information, and performs equalization process iteratively. With such an approach, the proposed scheme jointly optimizes the estimates of the channel and data symbols in each iteration. This avoids the common drawback in separate channel estimation and equalization/detection approach in that the

correlation between channel estimate and data symbol decision is considered. The complexity of the method of [5] is comparable to that of the turbo equalizers using linear filters [82–84], and is usually much lower than that of the ML/MAP based joint channel estimation and data detection schemes.

Based on the turbo equalization approach proposed in [5] and CE-BEM, an adaptive turbo equalizer with nonlinear Kalman filtering has been proposed in [85]. It is discussed in more detail in later under adaptive channel estimation and equalization.

2.03.5 Precoding

Thus far we have discussed channel estimation and equalization at the receiver. The basic idea behind channel precoding is that if the channel state information (CSI) is known at the transmit side, one can move the “equalizer” to the transmitter thereby simplifying the equalizer at the receiver and/or minimizing noise enhancement at the receiver [1, 86–88]. In the SISO case precoding helps in ISI mitigation. In the MIMO case the objective is both ISI cancellation and multiuser/multiantenna interference mitigation. The book [87] provides a recent comprehensive review of recent advances, and [89, 90] are good review articles for SISO and MIMO channels, respectively.

2.03.5.1 Precoding for SISO channels

Precoding the information symbols at the transmitter with full knowledge of CSI to mitigate ISI was first proposed in [86, 88]. Since then it has been generalized to a wide variety of scenarios [87, 89, 90]. In the original works of [86, 88], real-valued M ary PAM (pulse amplitude modulation) signal set was considered which is what we do here. The information alphabet is taken from the set $\mathcal{A} := \{\pm 1, \pm 3, \dots, \pm(M - 1)\}$. Consider the channel given by (3.9). Unlike (3.9), the information symbol $s(n)$ is precoded to $x(n)$ chosen from \mathcal{A} as follows

$$x(n) = s(n) - \sum_{l=1}^L \tilde{h}(l)x(n-l) \text{ modulo } 2M \quad \text{where} \quad \tilde{h}(l) = h(l)/h(0). \quad (3.63)$$

The transmitter transmits the precoded sequence $\{x(n)\}$ instead of $\{s(n)\}$ over the channel represented by (3.9). At the receiver one has $y(n) = \sum_{l=0}^L h(l)x(n-l) + v(n)$. By (3.63) there exists a unique integer $b(n)$ such that

$$x(n) = s(n) - \sum_{l=1}^L \tilde{h}(l)x(n-l) + 2Mb(n) \in (-M, M]. \quad (3.64)$$

Using z -transform of the sequences in (3.64) we have

$$X(z) = S(z) - \left(\frac{H(z)}{h(0)} - 1 \right) X(z) + 2MB(z). \quad (3.65)$$

Therefore at the receiver we have the z -transform of the received sequence $\{y(n)\}$ as

$$Y(z) = X(z)H(z) + V(z) = h(0)(S(z) + 2MB(z)) + V(z). \quad (3.66)$$

Thus there is no ISI in (3.66) unlike (3.9).

2.03.5.2 Precoding for MIMO channels

This is a very active area of current research [87, 90] and our overview will be quite brief. The basic idea here is to map a block of information symbols \underline{s}_i into a larger block of data \underline{x}_i via some linear transformation (precoding) $\underline{x}_i = \mathbf{F}\underline{s}_i$, and the precoded symbol block \underline{x}_i is then modulated and transmitted over the channel. At the receiver one designs a decoder \mathbf{G} to operate on the noisy received signal block $\underline{y}_i = \underline{\mathbf{H}}\underline{x}_i + \underline{\mathbf{v}}_i$ ($\underline{\mathbf{H}}$ represents the channel effect) to yield the symbol block estimate $\hat{\underline{s}}_i = \mathbf{G}\underline{\mathbf{H}}\underline{x}_i + \mathbf{G}\underline{\mathbf{v}}_i$. The structure of \mathbf{F} , the redundancy added per block and the design criteria for the choice of precoder-decoder pair \mathbf{F} and \mathbf{G} together with the underlying channel matrix $\underline{\mathbf{H}}$ dictate the resulting performance of the system.

Consider (3.49) with K users (transmit antennas), N receive antennas and input of precoded symbols $\underline{\mathbf{x}}(n)$ instead of information symbols $\underline{\mathbf{s}}(n)$. Following [91], for some integer $M > 0$, let $P = M + L$ and define the blocks

$$\underline{\mathbf{x}}_i = \text{vec}([\mathbf{x}(iP), \mathbf{x}(iP+1), \dots, \mathbf{x}(iP+P-1)]), \quad (3.67)$$

$$\underline{\mathbf{y}}_i = \text{vec}([\mathbf{y}(iP+L), \mathbf{y}(iP+1), \dots, \mathbf{y}(iP+P-1)]), \quad (3.68)$$

where $\underline{\mathbf{x}}_i$ is $KP \times 1$ and $\underline{\mathbf{y}}_i$ is $NM \times 1$. In (3.68) the first L vectors have been deleted to cancel interblock interference (IBI); an alternative is to zero-pad the tail of every block \underline{s}_i [91]. From (3.49), (3.67), and (3.68), one can deduce

$$\underline{\mathbf{y}}_i = \underline{\mathbf{H}}\underline{\mathbf{x}}_i + \underline{\mathbf{v}}_i, \quad (3.69)$$

where $\underline{\mathbf{H}}$ is an $NM \times KP$ block-banded matrix and it becomes a block Toeplitz matrix for time-invariant channels. A block of N_s information symbols \underline{s}_i is precoded as

$$\underline{\mathbf{x}}_i = \mathbf{F}\underline{s}_i \quad \text{where } N_s \leq \min(KP, NM). \quad (3.70)$$

In [91] under additive white Gaussian noise, various designs of precoder-decoder pairs \mathbf{F}, \mathbf{G} are considered. Let

$$\text{MSE}(\mathbf{F}, \mathbf{G}) = E\{\|\hat{\underline{s}}_i - \underline{s}_i\|^2 | \mathbf{F}, \mathbf{G}\}. \quad (3.71)$$

For given \mathbf{F} and \mathbf{H} , decoder \mathbf{G} is chosen to minimize $\text{MSE}(\mathbf{F}, \mathbf{G})$. Then under a transmit power or similar constraint, \mathbf{F} is chosen to minimize a function of $\text{MSE}(\mathbf{F}, \mathbf{G}_{\text{opt}})$.

The joint linear precoder/decoder design is in general a complicated non-convex problem [87]. The linear precoder/decoder optimization decouples the MIMO channel into parallel subchannels if the criterion is the minimization of the weighted sum of MSEs of all subchannels. Note also that MSE is not the only criterion for precoder design; other criteria include maximization of SNR, maximization of information rate, and minimization of bit error probability [87].

2.03.5.3 Precoding with partial or no CSI at transmitter

When the channel is time-varying, the assumption of knowledge of CSI at the transmitter is not entirely justified. Then an appropriate approach is to design precoder/decoder on the basis of the statistical knowledge of the CSI or resort to blind methods. For precoder designs using the first- and second-order statistics of the channels at the transmitter, see [92] and references therein. Precoding can also facilitate blind channel estimation and equalization in the absence of any training; see [93, 94] and references therein.

A fairly comprehensive review of some of these issues and techniques may be found in [87, 90].

2.03.6 Tracking

2.03.6.1 Adaptive channel estimation for slowly varying channels

When the channel characteristics vary slowly with time, recursive implementations of the “traditional” (linear or DFE) equalizers aided with initial transmission of a training sequence work well [1, Chapter 11]. The equalization parameters are often updated through the MMSE criterion. This requires that a known channel input sequence be transmitted initially. Adaptive channel equalizers begin adaptation with the assistance of a known training sequence transmitted during the initial training stage by the transmitter. Since the input signal is available, adaptive algorithms can be used to adjust the equalizer parameters by minimizing the MSE between the equalizer output and the known channel input with an equalization delay $s(n - d)$. After training, equalizer parameters should be sufficiently close to the desired settings such that much of the ISI is removed. As the channel input can now be correctly recovered from the equalizer output through a decision device (hard quantizer), the second (operational) stage can begin. In the operational stage, the receivers typically switch to decision-directed mode where the equalized signal is sent to a symbol detector and the detected symbols are used as a (pseudo-)training sequence to update equalizer coefficients. Baud-rate linear transversal equalizers, FSE and DFE all can be updated in this way. During either stage, the equalizer parameters can be determined using the well-known recursive least-squares (RLS) or least mean square (LMS) algorithms [1, 70].

In the absence of training, blind equalizers may be employed; they tend to be much slower in convergence and tracking.

2.03.6.2 Block-adaptive channel estimation using CE-BEM

Here we summarize the time-multiplexed training approach of [11]. In [11] each transmitted block of symbols $\{s(n)\}_{n=0}^{T_B-1}$ is segmented into \bar{P} subblocks of time-multiplexed training and information symbols. Each subblock is of equal length l_b symbols with l_d information symbols and l_t training symbols ($l_b = l_d + l_t$). If \mathbf{s} denotes a column-vector composed of $\{s(n)\}_{n=0}^{T_B-1}$, then \mathbf{s} is arranged as

$$\mathbf{s} := \begin{bmatrix} \mathbf{b}_0^T & \mathbf{c}_0^T & \mathbf{b}_1^T & \mathbf{c}_1^T & \cdots & \mathbf{b}_{\bar{P}-1}^T & \mathbf{c}_{\bar{P}-1}^T \end{bmatrix}^T, \quad (3.72)$$

where \mathbf{b}_p ($p = 0, 1, \dots, \bar{P} - 1$) is a column of l_d information symbols and \mathbf{c}_p is a column of l_t training symbols. We clearly have $T_B = \bar{P}l_b$. Given (3.8) and CE-BEM (3.17), Ma et al. [11] has shown that (3.72) is an optimum structure for $K = 1$ with $l_t = 2L + 1$, $\bar{P} \geq Q$ and

$$\mathbf{c}_p := [\mathbf{0}_L^T \ \gamma \ \mathbf{0}_L^T]^T, \quad \gamma > 0. \quad (3.73)$$

Thus, given a transmission block of size T_B , $(2L + 1)\bar{P}$ symbols have to be devoted to training and the remaining $T_B - (2L + 1)\bar{P}$ are available for information symbols.

Let $n_p := pl_b + l_d + L$ ($p = 0, 1, \dots, \bar{P} - 1$) denote the location of (nonzero) γ 's in the optimum \mathbf{c}_p 's in the P subblocks. Then by design, received signal (assuming timing synchronization)

$$y(n_p + l) = \gamma h(n_p + l; l) + v(n_p + l) \quad (3.74)$$

for $l = 0, 1, \dots, L$. Using (3.17) in these $y(n_p + l)$ s, one can uniquely obtain $h_q(l)$ s via a least-squares approach. The channel estimates are given by the CE-BEM (3.17) using the estimated BEM coefficients.

2.03.6.3 Adaptive channel estimation via subblock tracking

Suppose that we collect the received signal over a time interval of \tilde{T} symbols. We wish to estimate the time-variant channel using a channel model and time-multiplexed training (such as that discussed in Section 2.03.6.2 and [11]), and subsequently using the estimated channel, estimate the information symbols. For CE-BEM, if we choose \tilde{T} as the block size, then in general Q value will be very high requiring estimation of a large number of parameters, thereby degrading the channel estimation performance. If we divide \tilde{T} into blocks of size T_B , and then fit CE-BEM block by block, we need smaller Q ; however, estimation of $h_q(l)$'s is now based on a shorter observation size of T_B symbols which might also degrade channel estimation performance. Thus one has to strike a balance between estimation variance and block size. Such considerations do not apply to the AR channel model fitting. [95] proposes a novel subblock tracking approach to CE-BEM channel estimation where one updates estimates of $h_q(l)$ s every subblock based on all of the past training symbols (see Figure 3.5).

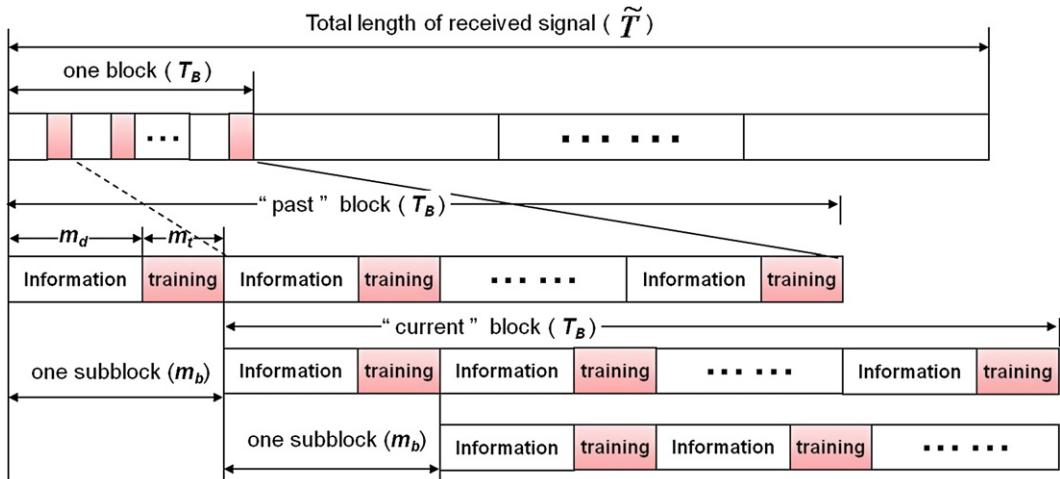


FIGURE 3.5

Subblock-wise channel estimation: overlapping blocks and subblocks, where one block comprises several subblocks and each subblock has an information session followed by a training session.

By exploiting the invariance of the coefficients of CE-BEM over each block, hence, each of the \bar{P} subblocks per block of length T_B symbols, one seeks subblock-wise tracking of the BEM coefficients of the doubly-selective channel. Consider two overlapping blocks that differ by just one subblock: blocks with $n = m, m + 1, \dots, m + T_B - 1$ where $m = m_0$ for the “past” block and $m = m_0 + m_b$ for the “current” block. If the two blocks overlap so significantly, one would expect the BEM coefficients to vary only a little from the past block to the current overlapping block. Therefore, rather than estimate $h_q(l)$ s anew with every non-overlapping block as in Section 2.03.6.2 and [11], one can track the BEM coefficients subblock by subblock using a first-order AR model for their variations. Stack the channel coefficients in (3.17) into vectors

$$\mathbf{h}_l := [h_1(l) \ h_2(l) \ \dots \ h_Q(l)]^T, \quad (3.75)$$

$$\mathbf{h} := [\mathbf{h}_0^T \ \mathbf{h}_1^T \ \dots \ \mathbf{h}_{\bar{P}}^T]^T, \quad (3.76)$$

of size Q and $M := Q(\bar{P} + 1)$ respectively. The coefficient vector in (3.76) for the p th subblock ($p = 0, 1, \dots$) will be denoted by $\mathbf{h}(p)$. We assume that the BEM coefficients over each subblock are Markovian: a simplified model is given by the first-order AR process, i.e.,

$$\mathbf{h}(p) = \alpha \mathbf{h}(p-1) + \mathbf{w}(p), \quad (3.77)$$

where α is the AR coefficient, and the driving noise vector $\mathbf{w}(p)$ is zero-mean complex Gaussian with variance $\sigma_w^2 \mathbf{I}_M$ and statistically independent of $\mathbf{h}(p-1)$. If the channel is stationary and coefficients $h_q(l)$ s are independent (as assumed in [11]), then it follows from (3.77) that $\sigma_w^2 = \sigma_h^2(1 - |\alpha|^2)/Q$ with $\sigma_h^2 := E\{h(n; l)h^*(n; l)\}$. Since the coefficients evolve slowly, we have $\alpha \approx 1$ (but $\alpha < 1$ for tracking). Under this formulation we do not have a “strict” definition of the block size T_B because, although we still use (3.17) for any n , we allow $h_q(l)$ s to change subblock by subblock based on the training symbols.

Define $\mathbf{e}(n) := [e^{-j\omega_1 n} \ e^{-j\omega_2 n} \ \dots \ e^{-j\omega_Q n}]^T$. If at time n the p th subblock is being received, by (3.6), (3.17–3.19) and (3.75–3.76), the received signal can be written as

$$y(n) = \mathbf{s}^T(n) [\mathbf{I}_{L+1} \otimes \mathbf{e}(n)]^H \mathbf{h}(p) + v(n), \quad (3.78)$$

where $\mathbf{s}(n) := [s(n) \ s(n-1) \ \dots \ s(n-L)]^T$. Treating (3.77) and (3.78) as the state and the measurement equations respectively, Kalman filtering can be applied to track the coefficient vector $\mathbf{h}(p)$ for each subblock; further details are in [95]. Alternatively, one can devise RLS-based approaches for subblock-based tracking one does not need any prior models for BEM coefficients; further details are in [95].

2.03.6.4 Symbol-adaptive joint channel estimation and data detection

Representative approaches in this category are [4, 5] and references therein. A Gauss-Markov model for channel variations (typically an autoregressive model) is coupled with a state-space model for received data to form an augmented state-space model with nonlinear measurement equation. This results in a nonlinear state estimation problem. In [4] a finite-length minimum mean-square error (MMSE) DFE is used during non-data-aided periods to generate hard decisions. Reference [5] presents a low complexity

turbo equalization receiver for coded signals where a nonlinear Kalman filtering based adaptive equalizer (using Extended Kalman Filter (EKF)) is coupled with a soft-in soft-out decoder. These approaches work well so long as the channel does not fade too fast.

2.03.6.4.1 Turbo equalization using EKF and CE-BEM

Now we describe an extension of [5] as reported in [85] where a CE-BEM-based subblock tracking model (with one sample long subblock) is used. Kin and Tuguit [85] considers two overlapping blocks (each of T_B symbols) that differ by just one symbol: the “past” block beginning at time n_0 and the “present” block beginning at time $n_0 + 1$. Since the two blocks overlap so significantly, one would expect the BEM coefficients to vary only “a little” from the past block to the present overlapping one. One can track the BEM coefficients (rather than the channel tap gains) symbol-by-symbol using a first-order AR model for their variations. Kin and Tuguit [85] use (3.17) for all times n , not just the particular block of size T_B symbols, by allowing the coefficients $h_q(l)$ s to change with time. Note that model (3.17) is periodic with period KT_B whereas the channel is by no means periodic. So long as the effective “memory” of the Kalman filter used later is less than the model period T_p , there are no deleterious effects due to the use of (3.17) for all time.

Bit-Interleaved Coded Modulation (BICM). We consider a BICM transmitter (as in [96]) for a doubly-selective fading channel as shown in Figure 3.6. A sequence of independent data vector $b(n') = [b^1(n'), b^2(n'), \dots, b^{k_0}(n')] \in \{1, 0\}^{k_0}$ are fed into a convolutional encoder with a code rate $R_c = k_0/n_0$. The coded output $\mathbf{c}(n') = [c^1(n'), c^2(n'), \dots, c^{n_0}(n')] \in \{1, 0\}^{n_0}$ is passed through a bit-wise random interleaver π , generating the interleaved coded bit sequence $\mathbf{c}(n) = \pi[\mathbf{c}(n')]$. The binary coded bits are then mapped to a signal sequence $d(n)$ over a 2-dimensional signal constellation \mathcal{X} of cardinality $\mathcal{M} = 2^m$ by a \mathcal{M} -ary modulator with an one-to-one binary map $\mu : \{1, 0\}^m \rightarrow \mathcal{X}$. In this section, we only consider the case of phase-shift keying (PSK) or quadrature amplitude modulation (QAM) with the average energy of the constellation \mathcal{X} to be unity. That is, the signal $d(n)$ drawn from \mathcal{X} has mean $E[d(n)] = 0$ and variance $E[|d(n)|^2] = 1$. After modulation, we periodically insert short training sequences into the data symbol sequence. The training symbols $t(n)$, which are known to the receiver, are randomly drawn from the signal constellation \mathcal{X} with equal probabilities. The symbol $\{s(n)\}$ will be used to denote the symbol sequence after training $\{t(n)\}$ insertion into data symbol sequence $\{d(n)\}$.

Receiver structure. A turbo equalization structure, as depicted in Figure 3.7, is employed in the receiver, as in [5] except that [5] uses symbol-wise AR models. The adaptive SISO_f equalizer is embedded into the iterative decoding (ID) process of the BICM transmission system (BICM-ID) [96]. In each decoding iteration, the equalizer takes the training symbols and the soft decision information about

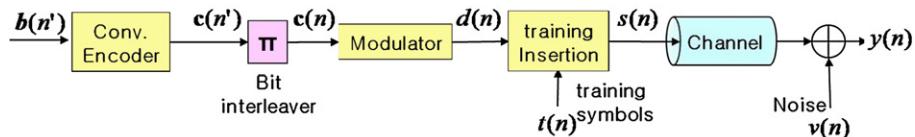


FIGURE 3.6

Bit-interleaved coded modulation system model for doubly-selective fading channel.

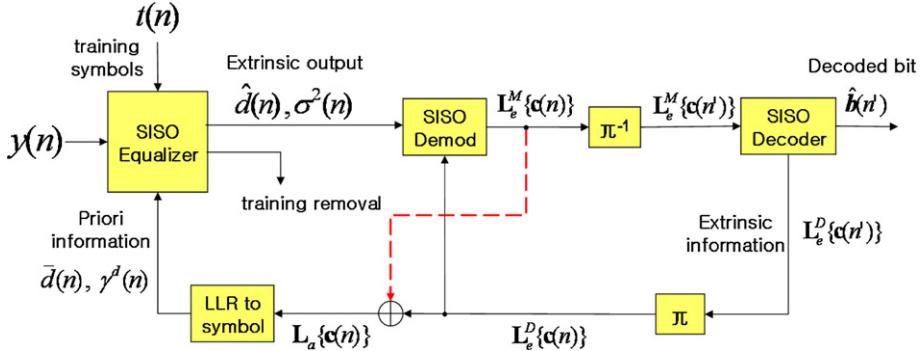


FIGURE 3.7

Turbo equalization receiver. Following [5, 97, 98] and contrary to the original turbo principle, *a posteriori* LLR $\mathbf{L}_a\{\mathbf{c}(n)\} = \mathbf{L}_e^M\{\mathbf{c}(n)\} + \mathbf{L}_e^D\{\mathbf{c}(n)\}$ instead of the extrinsic LLR $\mathbf{L}_e^D\{\mathbf{c}(n)\}$ can be input to the LLR-to-symbol block. Inclusion of $\mathbf{L}_e^M\{\mathbf{c}(n)\}$ to create *a posteriori* LLR is shown via dashed line. For our proposed approach we follow [5, 97, 98].

data symbols supplied by the SISO_f decoder from the previous iteration as its *a priori* information to perform joint adaptive channel estimation and equalization. The equalizer produces the soft-valued extrinsic estimate of the data symbols, which are independent of their *a priori* information. The output of the equalizer is an updated sequence of soft estimates $\hat{d}(n)$ and its error variance $\sigma^2(n)$. Using the adaptive SISO_f equalizer described later, we have extrinsic information for the data symbols $d(n)$. The training symbols are removed at the SISO_f equalizer output and the iterative process that follows is only for data symbols. The SISO_f equalizer based on the CE-BEM is described later. The SISO_f demodulator follows [96] whereas the SISO_f decoder follows the MAP decoding algorithm (“BCJR”) [99] Section 6.2.

Adaptive SISO_f nonlinear Kalman equalizer. Using a symbol-wise AR-model for channel variations, an adaptive SISO_f equalizer using fixed-lag EKF was presented in [5] for joint channel estimation and equalization where their correlation was (implicitly) considered. The CE-BEM model-based SISO_f nonlinear Kalman equalizer for turbo equalization follows a similar approach. We will perform equalization with a delay $\delta > 0$. Define a parameter

$$\bar{\delta} := \max\{\delta + 1, L + 1\} \quad (3.79)$$

and the data vector

$$\mathbf{z}(n) := [s(n) \ s(n - 1) \ \dots \ s(n - \bar{\delta} + 1)]^T. \quad (3.80)$$

In order to apply (extended) Kalman filtering to joint channel estimation and equalization, we stack $\mathbf{h}(n)$ (which is given by (3.77) after replacing p therein with n to reflect the fact our subblock size now is one sample) and data vector $\mathbf{z}(n)$ together into a $J \times 1$ state vector $\mathbf{x}(n)$ at time n as

$$\mathbf{x}(n) := [\mathbf{z}^T(n) \ \mathbf{h}^T(n)]^T, \quad J := \bar{\delta} + Q(L + 1). \quad (3.81)$$

As in [5] (and others), we consider the symbol sequence $\{s(n)\}$ as a stochastic process so as to utilize the soft decisions on the data symbols generated in the iterative decoding process as its *a priori* information. We can express $s(n)$ as $s(n) = \bar{s}(n) + \tilde{s}(n)$ where $\bar{s}(n) = E[s(n)]$ and $\tilde{s}(n)$ is approximated as a zero-mean uncorrelated sequence such that $E[\tilde{s}(n)\tilde{s}^*(n+j)] = \gamma(n)\delta(j)$, assuming an ideal interleaver. Note that $\bar{s}(n)$ and $\gamma(n)$ are provided via the *a priori* information. We have $\bar{s}(n) = \bar{d}(n)$ and $\gamma(n) = \gamma^d(n)$ for a data symbol $d(n)$ (where $\bar{d}(n)$ and $\gamma^d(n)$ are obtained from the extrinsic LLRs of the SISO_f decoder), while $\bar{s}(n) = t(n)$ and $\gamma(n) = 0$ for a training symbol $t(n)$. Using $\mathbf{x}(n)$, the state equation turns out to be

$$\mathbf{x}(n) = \mathcal{T}\mathbf{x}(n-1) + \mathbf{e}_0\bar{s}(n) + \mathbf{u}(n), \quad (3.82)$$

where

$$\mathcal{T} = \begin{bmatrix} \Phi & \mathbf{0}_{\bar{\delta} \times Q(L+1)} \\ \mathbf{0}_{Q(L+1) \times \bar{\delta}} & \mathbf{F} \end{bmatrix}_{J \times J}, \quad \mathbf{F} = \alpha \mathbf{I}_{Q(L+1)}, \quad (3.83)$$

$$\Phi = \begin{bmatrix} \mathbf{0}_{1 \times (\bar{\delta}-1)} & \mathbf{0}_{1 \times 1} \\ \mathbf{I}_{(\bar{\delta}-1)} & \mathbf{0}_{(\bar{\delta}-1) \times 1} \end{bmatrix}_{\bar{\delta} \times \bar{\delta}}, \quad \mathbf{e}_0 = [1 \ \mathbf{0}_{1 \times (J-1)}]^T, \quad (3.84)$$

the vector

$$\mathbf{u}(n) := \left[\mathbf{e}_{\bar{\delta}}^T \bar{s}(n) \ \mathbf{w}^T(n) \right]^T \quad (3.85)$$

is zero-mean uncorrelated process noise where $\mathbf{e}_{\bar{\delta}} = [1 \ 0_{1 \times (\bar{\delta}-1)}]^T$, $\mathbf{w}(n)$ is given in (3.77) and

$$\mathbf{Q}(n) := E[\mathbf{u}(n)\mathbf{u}^H(n)] = \tilde{\mathbf{Q}} + \gamma(n)\mathbf{e}_0\mathbf{e}_0^T, \quad \tilde{\mathbf{Q}} := \begin{bmatrix} \mathbf{0}_{\bar{\delta} \times \bar{\delta}} & \mathbf{0}_{\bar{\delta} \times Q(L+1)} \\ \mathbf{0}_{Q(L+1) \times \bar{\delta}} & \sigma_w^2 \mathbf{I}_{Q(L+1)} \end{bmatrix}_{J \times J}. \quad (3.86)$$

The channel output $y(n)$ in (3.6) can be rewritten by CE-BEM given in (3.17) as

$$y(n) = \mathbf{s}^T(n) [\mathbf{I}_{(L+1)} \otimes \mathbf{e}(n)]^H \mathbf{h}(n) + v(n), \quad (3.87)$$

where $\mathbf{s}(n) = [s(n) \ s(n-1) \ \dots \ s(n-L)]^T$ and $\mathbf{e}(n) := [e^{-j\omega_1 n} \ e^{-j\omega_2 n} \ \dots \ e^{-j\omega_Q n}]^T$. Using the state vector that comprises the information symbols and channel coefficients, the measurement equation can be given as

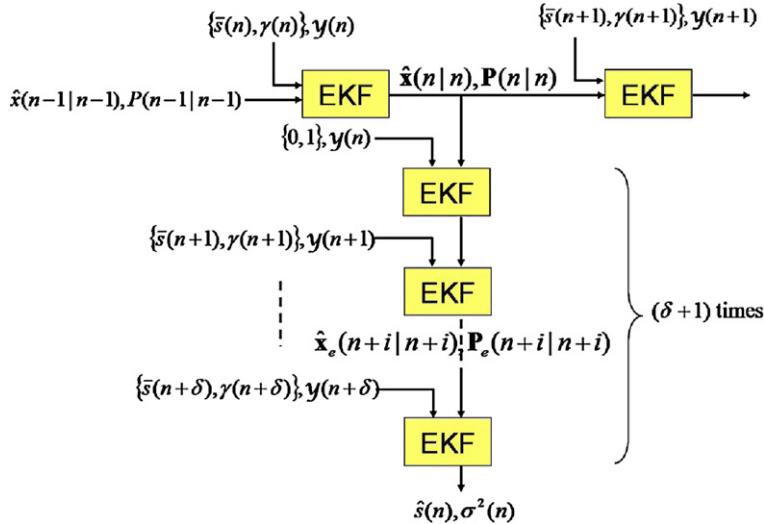
$$y(n) = f[\mathbf{x}(n)] + v(n), \quad (3.88)$$

where

$$f[\mathbf{x}(n)] := \mathbf{x}^T(n) \underbrace{[\mathbf{I}_{(L+1)} \ \mathbf{0}_{(L+1) \times (J-L-1)}]^T [\mathbf{I}_{(L+1)} \otimes \mathcal{E}(n)]^H [\mathbf{0}_{[Q(L+1)] \times \bar{\delta}} \ \mathbf{I}_{Q(L+1)}]}_{=: \mathbf{D}} \mathbf{x}(n). \quad (3.89)$$

With (3.82) and (3.88) as the state and measurement equations, respectively, nonlinear Kalman filtering via EKF is applied to track $\mathbf{x}(n)$ for joint channel estimation and equalization.

Structure of adaptive soft-input soft-output equalizer. The fixed-lag EKF takes soft inputs and generates a delayed *a posteriori* estimate for $s(n)$. In order to generate extrinsic estimate independent

**FIGURE 3.8**

Structure of the adaptive SISO_f equalizer proposed in [5].

of the *a priori* information $\{\bar{s}(n), \gamma(n)\}$, a “comb” structure in conjunction with the EKF in Figure 3.8 is used for the SISO_f equalization, just as in [5]. At each time n , the vertical branch composed of $(\delta + 1)$ EKFs produce the extrinsic estimate $\hat{s}(n)$, while the horizontal branch keeps updating the *a posteriori* estimate $\hat{x}(n | n)$ and its error covariance $\mathbf{P}(n | n)$; here $\hat{x}(p|m)$ denotes the estimate of $\mathbf{x}(p)$ given the observations $\{y(0), y(1), \dots, y(m)\}$, and $\mathbf{P}(p|m)$ denotes the error covariance matrix of $\hat{x}(p|m)$. The first vertical EKF has an input $\{0, 1\}$ in place of $\{\bar{s}(n), \gamma(n)\}$ to exclude the effect of the *a priori* information. Let $\hat{x}_e(n+i | n+i)$ and $\mathbf{P}_e(n+i | n+i)$ denote the state estimate and its error covariance matrix, respectively, generated by the $(i+1)$ th vertical filtering branch. Then the extrinsic estimate $\hat{s}(n)$ of $s(n)$ and its error variance $\sigma^2(n)$ are given by

$$\hat{s}(n) = \text{6th component of vector } \hat{x}_e(n+\delta | n+\delta), \quad (3.90)$$

$$\sigma^2(n) = (\delta, \delta)\text{th component of matrix } \mathbf{P}_e(n+\delta | n+\delta). \quad (3.91)$$

Note that the extrinsic outputs $\hat{s}(n)$ and $\sigma^2(n)$ are computed for data symbol $d(n)$, not for training symbol $t(n)$, and then used in the later parts of the turbo-equalization receiver (see Figure 3.7). Further details regarding generation of extrinsic estimates can be found in [5].

Simulation examples. A random time- and frequency-selective Rayleigh fading channel is considered. We assume $h(n; l)$ is zero-mean, complex Gaussian, and white with autocorrelation σ_h^2 . We take $L = 2$ (3 taps) and $\sigma_h^2 = 1/(L+1)$. For different l s, $h(n; l)$ s are mutually independent and satisfy Jakes’ model. To this end, we simulate each single tap following [100] (with a correction in the Appendix of [13]). We consider a normalized Doppler spread $f_d T_s$ from 0.001 to 0.01. The additive noise is zero-mean complex white Gaussian. The (receiver) SNR refers to the average energy per symbol over one-sided noise spectral density. In the simulations, we use a four-state convolutional code of

rate $R_c = 1/2$ with octal generators (5, 7). The information block size is set to 3000 bits ($T_i = 3000$) leading to a coded block size of 6000 bits, and the interleaver size is equal to the coded block size. In the modulator, the QPSK constellation with Gray mapping is used, which gives $\mathcal{M} = 4$ and a block size of 3000 symbols. After modulation, training symbol sequences of length l_p are inserted in front of every l_s data symbols, leading to a sequence of length $T_r = 3750$ when $l_p = 5$ and $l_s = 20$ (20% training overhead).

We compared the following schemes:

- The approach of [83] that uses the linear MMSE equalizer (e.g., [80]) coupled with modified RLS channel estimation, where we set the linear filter length = 3 (6 pre-cursor taps and 3 post-cursor taps are used). This scheme is denoted by “TE-LE.”
- The AR(p) model-based scheme in [5]. The AR(p) model is fitted using [3] to Jakes’ spectrum with $f_d T_s = 0.01$ (the maximum anticipated normalized Doppler spread), denoted by “TE-AR5” for AR(5) model and “TE-AR9” for AR(9) model.
- The proposed BEM-based turbo equalization schemes, where we consider BEM period $T = 200$ and 400 respectively, so that $Q = 5$ and 9, respectively, by (3.19). For the channel BEM coefficients, we take the AR-coefficient in (3.17) as $\alpha = 0.996$ for $T = 200$ and $\alpha = 0.998$ for $T = 400$. This scheme is denoted by “TE-BEM(200)” for $T = 200$ and “TE-BEM(400)” for $T = 400$.

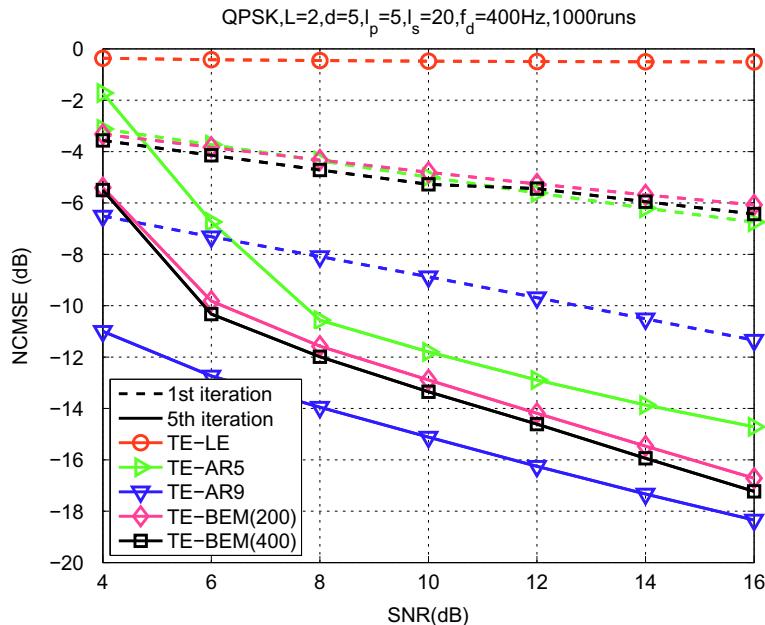


FIGURE 3.9

Turbo equalization: performance comparison for SNRs under $f_d T_s = 0.01$, $l_p = 5$, $l_s = 20$ (20% training overhead).

- The turbo equalizer based on the fixed-lag Kalman filter with perfect knowledge of the true channel, denoted by “TrueCH.”
- The turbo equalizer based on the optimum trellis-based MAP (BJCR) method [101] with perfect knowledge of the true channel, denoted by “Opt-MAP-TrueCH.”

We evaluate the performances of various schemes by considering their normalized channel mean square error (NCMSE) and their bit error rates (BER). The NCMSE is defined as

$$\text{NCMSE} := \frac{\sum_{i=1}^{M_r} \sum_{n=0}^{T_N-1} \sum_{l=0}^L \left\| \hat{h}^{(i)}(n; l) - h^{(i)}(n; l) \right\|^2}{\sum_{i=1}^{M_r} \sum_{n=0}^{T_N-1} \sum_{l=0}^L \| h^{(i)}(n; l) \|^2},$$

where $h^{(i)}(n; l)$ is the true channel and $\hat{h}^{(i)}(n; l)$ is the estimated channel at the i th Monte Carlo run, among total M_r runs. The BERs are evaluated by employing the equalization delay $\delta = 5$, using the decoded information symbol sequences at the turbo-equalization receiver. All the simulation results are based on 1000 runs.

In Figures 3.9 and 3.10, the performance of all the above schemes, under normalized Doppler spread $f_d T_s = 0.01$, are compared for different SNRs. In Figures 3.11 and 3.12 those schemes are compared over varying Doppler spread $f_d s$, under $\text{SNR} = 10 \text{ dB}$. It is clear from these figures that since the

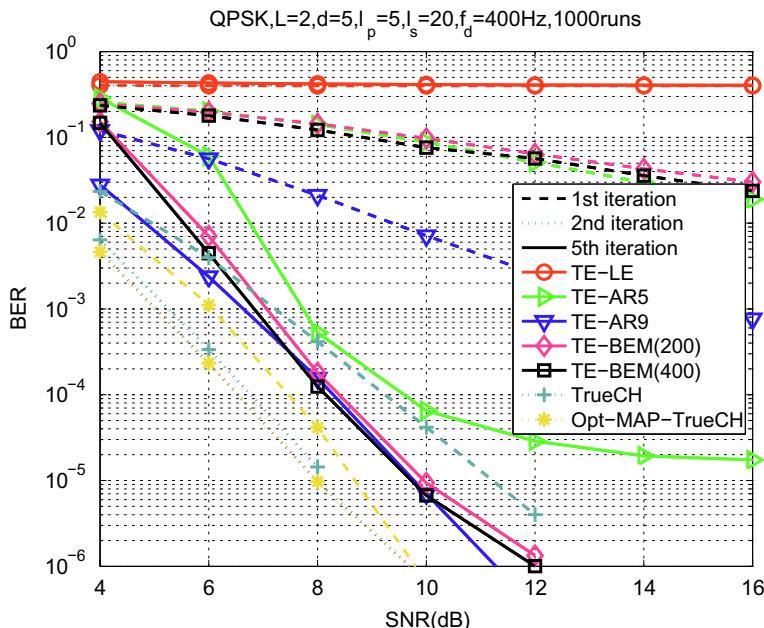


FIGURE 3.10

Turbo equalization: performance comparison for SNRs under $f_d T_s = 0.01$, $I_p = 5$, $I_s = 20$ (20% training overhead).

channel variations are well captured by the BEM coefficients, TE-BEM-based approach yields good performance even for “low” SNRs and over a wide range of Doppler spreads. Note that TE-BEM with larger block parameter $T_p = 400$ has a (slightly) better performance than with the smaller parameter $T_p = 200$. The BER of TE-BEM varies only “slightly” with increasing normalized Doppler spread implying that its performance is not sensitive to the actual Doppler spread. Therefore, we do not have to know the exact Doppler spread of the channel—an upper bound on it is sufficient in practice. The performance of TE-AR5 is significantly worse than that of TE-BEM(200) (the two approaches have comparable computational complexity) in Figure 3.10 with increasing SNR for a fixed $f_d T_s = 0.01$, and is slightly worse in Figure 3.12 for a fixed SNR of 10 dB and varying Doppler spreads. On the other hand, while the performance of TE-AR9 is slightly better than that of TE-BEM(400) (the two approaches have comparable computational complexity) in Figure 3.10 with increasing SNR for a fixed $f_d T_s = 0.01$, it is significantly worse in Figure 3.12 for a fixed SNR of 10 dB and varying Doppler spreads. While increasing the BEM period T_p improves performance, increasing the AR model order does not necessarily do so: we get inconsistent performance. A possible reason is that, as noted in [3], AR model fitting to a given correlation function can be numerically ill-conditioned for “large” model orders.

In Figures 3.10 and 3.12 the scheme TE-LE refers to the approach of [83] that uses the linear MMSE equalizer (e.g., [80]) coupled with modified RLS channel estimation. It is seen that this approach only works for normalized Doppler spread values of ≤ 0.002 . In Figure 3.10 we also present the

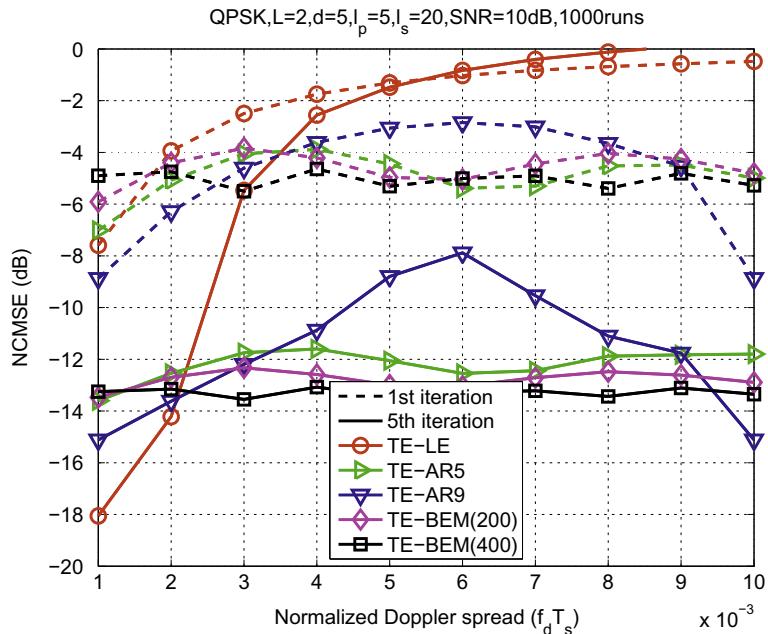
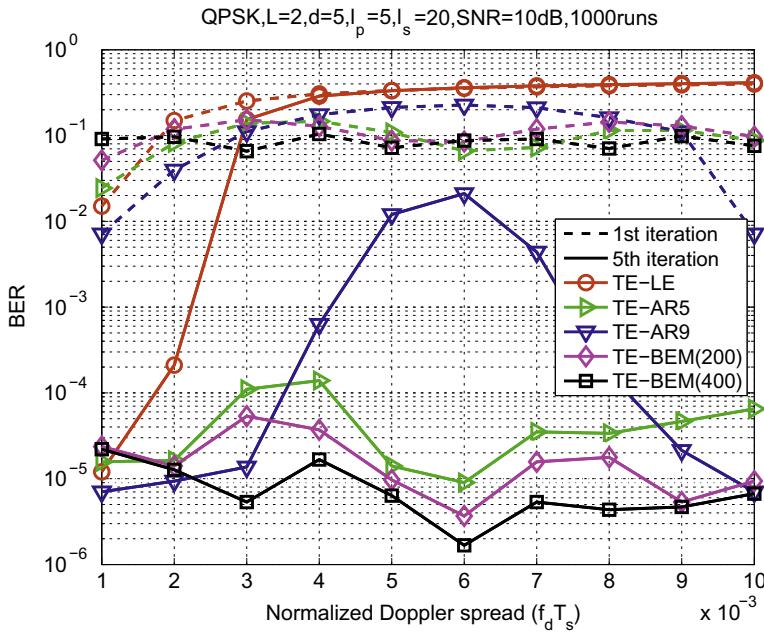


FIGURE 3.11

Turbo equalization: performance comparison for normalized Doppler spread ($f_d T_s$)_s under SNR = 10 dB, $I_p = 5$, $I_s = 20$ (20% training overhead).

**FIGURE 3.12**

Turbo equalization: performance comparison for normalized Doppler spread ($f_d T_s$) under SNR = 10 dB, $I_p = 5$, $I_s = 20$ (20% training overhead).

performance of the turbo equalizer based on the fixed-lag Kalman filter with knowledge of the true channel (curves with plus sign marker and labeled “TrueCH”) in order to illustrate the effectiveness of the proposed channel estimation approach; as there was little improvement beyond the second iteration, we only show the second iterative result with dotted curve labeled “TrueCH.” It is seen that there is a slightly more than 2 dB SNR penalty due to channel estimation. As has been noted in the literature, the Kalman filter based equalization is a sub-optimum equalizer compared to the trellis-based MAP (BCJR) equalizer [101]. In Figure 3.10 we present the performance of the turbo equalizer based on the optimum BCJR method with knowledge of the true channel (curves with asterisk marker and labeled “Opt-MAP-TrueCH”) in order to illustrate loss in performance due to suboptimality of the Kalman equalizer; as there was little improvement beyond the second iteration, we only show the second iterative result with dotted curve labeled “Opt-MAP-TrueCH”. It is seen that while there is a large difference in performance initially (see 1st iteration results for “TrueCH” and “Opt-MAP-TrueCH” where both are dashed curves with plus sign and asterisk markers, respectively), just one turbo iteration yields very close performance (see the two dotted curves). That is, at least for this example, performance loss in using Kalman equalizer instead of the BCJR equalizer is quite negligible.

2.03.7 Conclusion

A review of various approaches to channel estimation and equalization for communications systems was presented. Emphasis was on linear baseband equivalent models with a tapped delay line

structure, and both time-invariant and time-variant (doubly-selective) models were discussed. Also emphasis was on basis expansion modeling for time-variant channels where the basis functions are related to the physical parameters of the channel (such as Doppler and delay spreads). Channel modeling was followed by a discussion of various approaches to channel estimation including training-based approaches, blind approaches, semi-blind approaches and superimposed training based approaches. Channel estimation approaches were followed by a discussion of channel equalization approaches including turbo-equalization for time-varying channels. A brief discussion of precoding was presented when the channel state information is available at the transmitter. We concluded the chapter with a discussion of channel tracking and combined data detection and channel tracking for time-varying channels. Channel tracking can be at block level suitable for block transmissions, or symbol-by-symbol level suitable for serial transmissions. Some of the approaches were illustrated via simulations.

Relevant Theory: Signal Processing Theory

See [Vol. 1, Chapter 11](#) Parametric Estimation

See [Vol. 1, Chapter 12](#) Adaptive Filters

References

- [1] J. Proakis, Digital Communications, fourth ed., McGraw-Hill, New York, NY, 2001.
- [2] P. Bello, Characterization of randomly time-variant channels, *IEEE Trans. Comm. Syst.* 11 (1963) 360–393.
- [3] K.E. Baddour, N.C. Beaulieu, Autoregressive modeling for fading channel simulation, *IEEE Trans. Wireless Commun.* 4 (2005) 1650–1662.
- [4] C. Komninos, C. Fragouli, A. Sayed, R. Wesel, Multi-input multi-output fading channel tracking and equalization using Kalman estimation, *IEEE Trans. Signal Process.* 50 (5) (2002) 1065–1076.
- [5] X. Li, T. Wong, Turbo equalization with nonlinear Kalman filtering for time-varying frequency-selective fading channels, *IEEE Trans. Wireless Commun.* 6 (2007) 691–700.
- [6] S.M. Kay, Modern Spectral Analysis: Theory and Application, Prentice Hall, Englewoods Cliffs, NJ, 1988.
- [7] P. Stoica, R. Moses, Spectral Analysis of Signals, Pearson Prentice Hall, Upper Saddle River, NJ, 2005.
- [8] Z. Liu, X. Ma, G. Giannakis, Space-time coding and Kalman filtering for time-selective fading channels, *IEEE Trans. Commun.* 50 (2002) 183–186.
- [9] D. Borah, B. Hart, Receiver structures for time-varying frequency-selective fading channels, *IEEE J. Sel. Areas Commun.* 17 (1999) 1863–1875.
- [10] G. Giannakis, C. Tepedelenlioğlu, Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels, *Proc. IEEE* 86 (1998) 1969–1986.
- [11] X. Ma, G. Giannakis, S. Ohno, Optimal training for block transmissions over doubly selective wireless fading channels, *IEEE Trans. Signal Process.* 51 (2003) 1351–1366.
- [12] A. Sayeed, B. Aazhang, Joint multipath-doppler diversity in mobile wireless communications, *IEEE Trans. Commun.* 47 (1999) 123–132.
- [13] T. Zemen, C. Mecklenbräuker, Time-variant channel estimation using discrete prolate spheroidal sequences, *IEEE Trans. Signal Process.* 53 (2005) 3597–3607.
- [14] G. Leus, On the estimation of rapidly time-varying channels, in: Proceedings of the European Signal Processing, Vienna, Austria, 2004, pp. 2227–2230.
- [15] M. Martone, Wavelet-based separating kernels for array processing of cellular DS/CDMA signals in fast fading, *IEEE Trans. Commun.* 48 (2000) 979–995.

- [16] M. Niedzwiecki, Identification of Time-Varying Processes, Wiley, New York, NY, 2000.
- [17] D. Slepian, Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case, *Bell Syst. Tech. J.* 57 (1978) 1371–1430.
- [18] S. He, J. Tugnait, On bias-variance trade-off in superimposed training-based doubly selective channel estimation, in: Proceedings of the 2006 Conference on Information Systems and Sciences, Princeton University, NJ, 2006.
- [19] M.-A. Baissas, A. Sayeed, Pilot-based estimation of time-varying multipath channels for coherent CDMA receivers, *IEEE Trans. Signal Process.* 50 (2002) 2037–2049.
- [20] A. Kannu, P. Schniter, Capacity analysis of MMSE pilot-aided transmission for doubly selective channels, in: Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless, Communications, 2005, pp. 801–805.
- [21] A. Kannu, P. Schniter, MSE-optimal training for linear time-varying channels, in: Proceedings of the 2005 IEEE Conference on Acoustics, Speech and Signal Processing, Philadelphia, PA, 2005, pp. 789–792.
- [22] A. Kannu, P. Schniter, Minimum mean-squared error pilot-aided transmission for MIMO doubly selective channels, in: Proceedings of the 2006 Conference on Information Systems & Sciences, Princeton University, NJ, 2006.
- [23] G. Leus, S. Zhou, G. Giannakis, Orthogonal multiple access over time- and frequency-selective channels, *IEEE Trans. Inform. theory* 49 (2003) 1942–1950.
- [24] A. Sayeed, A. Sendonaris, B. Aazhang, Multiuser detection in fast-fading multipath environment, *IEEE J. Sel. Areas Commun.* 16 (1998) 1691–1701.
- [25] R. Gitlin, S. Weinstein, Fractionally-spaced equalization: an improved digital transversal equalizer, *Bell Syst. Tech. J.* 60 (1981) 275–296.
- [26] W. Gardner, Introduction to Random Processes with Applications to Signals and Systems, McGraw-Hill, New York, NY, 1989.
- [27] Y.-C. Chen, Y. Su, MIMO channel estimation in correlated fading environments, *IEEE Trans. Wireless Commun.* 9 (3) (2010) 1108–1119.
- [28] J. Karedal, P. Almers, A. Johansson, F. Tufvesson, A. Molisch, A MIMO channel model for wireless personal area networks, *IEEE Trans. Wireless Commun.* 9 (1) (2010) 245–255.
- [29] A. Sayeed, Deconstructing multiantenna fading channels, *IEEE Trans. Signal Process.* 50 (10) (2002) 2563–2579.
- [30] D. Shiu, G. Foschini, M. Gans, J. Kahn, Fading correlation and its effect on the capacity of multielement antenna systems, *IEEE Trans. Commun.* 48 (3) (2000) 502–513.
- [31] W. Weichselberger, M. Herdin, H. Ozcelik, E. Bonek, A stochastic MIMO channel model with joint correlation of both link ends, *IEEE Trans. Wireless Commun.* 5 (3) (2006) 90–100.
- [32] J. Cavers, Pilot symbol assisted modulation and differential detection in fading and delay spread, *IEEE Trans. Commun.* 43 (1995) 2206–2212.
- [33] G. Kaleh, R. Vallet, Joint parameter estimation and symbol detection for linear or nonlinear unknown dispersive channels, *IEEE Trans. Commun.* 42 (1994) 2406–2413.
- [34] V. Krishnamurthy, J. Moore, On-line estimation of hidden Markov model parameters based on Kullback-Leibler information measure, *IEEE Trans. Signal Process.* 41 (1993) 2557–2573.
- [35] L. Tong, S. Perreau, Multichannel blind channel estimation: from subspace to maximum likelihood methods, *Proc. IEEE* 86 (1998) 1951–1968.
- [36] N. Seshadri, Joint data and channel estimation using blind trellis search techniques, *IEEE Trans. Commun.* 42 (1994) 1000–1011.
- [37] M. Ghosh, C. Weber, Maximum-likelihood blind equalization, *Optical Eng.* 31 (1992) 1224–1228.
- [38] R. Raheli, A. Polydoros, C. Tzou, Per-survivor processing: A general approach to MLSE in uncertain environments, *IEEE Trans. Commun.* 43 (1995) 354–364.

- [39] K. Chugg, Blind acquisition characteristics of PSP-based sequence detectors, *IEEE J. Sel. Areas Commun.* 16 (1998) 1518–1529.
- [40] J. Tugnait, Blind estimation and equalization of digital communication FIR channels using cumulant matching, *IEEE Trans. Commun.* 43 (1995) 1240–1245.
- [41] J. Tugnait, Blind equalization and estimation of FIR communications channels using fractional sampling, *IEEE Trans. Commun.* 44 (1996) 324–336.
- [42] D. Hatzinakos, C. Nikias, Blind equalization using a tricepstrum based algorithm, *IEEE Trans. Commun.* 39 (1991) 669–681.
- [43] J. Vidal, J. Fonollosa, Adaptive blind equalization using weighted cumulant slices, *Int. J. Adapt. Control Signal Process.* 10 (1996) 213–238.
- [44] J. Tugnait, Channel estimation and equalization using higher-order statistics, in: G. Giannakis, Y. Hua, P. Stoica, L. Tong (Eds), *Signal Processing Advances in Wireless and Mobile Communications, Trends in Channel Estimation and Equalization*, vol. I, Prentice Hall, Upper Saddle River, NJ, 2001, pp. 1–39.
- [45] L. Tong, G. Xu, T. Kailath, A new approach to blind identification and equalization of multipath channels, *IEEE Trans. Inform. Theory* 40 (1994) 340–349.
- [46] J. Tugnait, On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics, *IEEE Trans. Inform. Theory* 41 (1995b) 308–311.
- [47] E. de Carvalho, D. Slock, Stochastic maximum likelihood methods for semi-blind channel estimation, in: G. Giannakis, Y. Hua, P. Stoica, L. Tong (Eds.), *Signal Processing Advances in Wireless and Mobile Communications*, vol. I, Prentice Hall, Upper Saddle River, NJ, 2001 (Chapter 7).
- [48] H. Cirpan, M. Tsatsanis, Stochastic maximum likelihood methods for semi-blind channel estimation, *IEEE Signal Proc. Lett.* 5 (1998) 21–24.
- [49] J. Laurila, K. Kopsa, E. Bonek, Semi-blind signal estimation for smart antennas using subspace tracking. In *Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communications*, Annapolis, MD, 1999.
- [50] A. Gorokhov, P. Loubaton, Semi-blind second order identification of convolutive channels, in: *Proceedings of the 1997 IEEE Conference on Acoustics, Speech and Signal Processing*, Munich, Germany, 1997, pp. 3905–3908.
- [51] S. Lasaulce, P. Loubaton, E. Moulines, A semi-blind channel estimation technique based on second-order blind method for CDMA systems, *IEEE Trans. Signal Process.* 51 (2003) 1894–1904.
- [52] P. Hoher, F. Tufvesson, Channel estimation with superimposed pilot sequence, in: *Proceedings of the IEEE Global Communications Conference*, Rio de Janeiro, Brazil, 1999, pp. 2162–66.
- [53] T. Holden, K. Feher, A spread spectrum based system technique for synchronization of digital mobile communication system, *IEEE Trans. Broadcast.* 36 (1990) 185–194.
- [54] F. Mazzenga, Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence, *IEEE Trans. Broadcast.* 46 (2000) 170–176.
- [55] M. Ghogho, D. McLernon, E. Alamdea-Hernandez, A. Swami, Channel estimation and symbol detection for block transmission using data-dependent superimposed training, *IEEE Signal Proc. Lett.* 12 (2005) 226–229.
- [56] M. Ghogho, D. McLernon, E. Alamdea-Hernandez, A. Swami, SISO and MIMO channel estimation and symbol detection using data-dependent superimposed training. in: *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, PA, Philadelphia, 2005, pp. III-461–III-464..
- [57] J. Tugnait, S. He, Doubly-selective channel estimation using data-dependent superimposed training and exponential bases models, in: *Proceedings of the Conf. on Information Systems and Sciences*, Princeton University, NJ, 2006.
- [58] J. Tugnait, S. He, Multiuser/MIMO doubly selective fading channel estimation using superimposed training and Slepian sequences, *IEEE Trans. Veh. Technol.* 59 (2010) 1341–1354.

- [59] H. Holma, A.E. Toskala, second ed., WCDMA for UMTS: Radio Access for Third Generation Mobile Communications, Wiley, New York, NY, 2002.
- [60] A. Orozco-Lugo, M. Lara, D. McLernon, Channel estimation using implicit training, *IEEE Trans. Signal Proc.* 52 (2004) 240–254.
- [61] J. Tugnait, W. Luo, On channel estimation using superimposed training and first-order statistics, in: Proceedings of the, IEEE Conference on Acoustics, Speech and Signal Processing, Hong Kong, 2003.
- [62] J. Tugnait, W. Luo, On channel estimation using superimposed training and first-order statistics, *IEEE Commun. Lett.* 8 (2003) 413–415.
- [63] G. Zhou, M. Viberg, T. McKelvey, A first-order statistical method for channel estimation, *IEEE Signal Process. Lett.* 10 (2003) 57–60.
- [64] M. Dong, L. Tong, B. Sadler, Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels, in: Proceedings of the 2003 IEEE Workshop on Signal Processing Advances in Wireless Communications, Rome, Italy. IEEE, 2003, pp. 472–476.
- [65] M. Biguesh, A. Gershman, Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals, *IEEE Trans. Signal Proc.* 54 (3) (2006) 884–893.
- [66] X. Ma, L. Yang, G. Giannakis, Optimal training for MIMO frequency-selective fading channels, *IEEE Trans. Wireless Commun.* 4 (2) (2005) 453–466.
- [67] I. Barhumi, G. Leus, M. Moonen, Time-varying FIR equalization for doubly selective channels, *IEEE Trans. Wireless Commun.* 4 (1) (2005) 202–214.
- [68] L. Song, J. Tugnait, Doubly-selective fading channel equalization: A comparison of the Kalman filter approach with the basis expansion model-based equalizers, *IEEE Trans. Wireless Commun.* 8 (1) (2009) 60–65.
- [69] N. Al-Dhahir, J.M. Cioffi, MMSE decision-feedback equalizers: finite-length results, *IEEE Trans. Inf. Theory* 41 (4) (1995) 961–975.
- [70] A.H. Sayed, Fundamentals of Adaptive Filtering, Wiley, New York, NY, 2003.
- [71] G.L. Stüber, Principles of Mobile Communication, second ed., Kluwer, Boston, MA, 2002.
- [72] G. Forney, Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference, *IEEE Trans. Inform. Theory* 18 (1972) 363–378.
- [73] S. Lin, J.J. Costello, Error Control Coding, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [74] L. Bahl, J. Cocke, F. Jelinek, J. Raviv, Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Trans. Inform. Theory* 20 (1974) 284–287.
- [75] M. Tuchler, A. Singer, Turbo equalization: an overview, *IEEE Trans. Inform. theory* 70 (2) (2011) 920–952.
- [76] C. Berrou, A. Glavieux, P. Thitimajshima, Near Shannon limit error-correcting coding and decoding: turbo codes, in: Proceedings of IEEE Int. Conf. on Commun., Geneva, Switzerland, 1993.
- [77] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, A. Glavieux, Iterative correction of intersymbol interference: turbo-equalization, *Eur. Trans. Telecommun.* 6 (1995) 507–511.
- [78] A. Anastassopoulos, K. Chugg, Iterative equalization/decoding for TCM for frequency-selective channels, in: Proceedings of the 1997 Asilomar Conf. on Signals, Systems, Computers, Pacific Grove, CA, 1997, pp. 177–181.
- [79] X. Wang, H. Poor, Iterative (turbo) soft interference cancellation and decoding for coded CDMA, *IEEE Trans. Commun.* 47 (7) (1999) 1046–1061.
- [80] M. Tuchler, Minimum mean squared error equalization using a priori information, *IEEE Trans. Signal Process.* 50 (3) (2002) 673–683.
- [81] M. Tuchler, Turbo equalization: principles and new results, *IEEE Trans. Commun.* 50 (5) (2002) 754–767.
- [82] M. Nissilä, S. Pasupathy, Adaptive Bayesian and EM-based detectors for frequency-selective fading channels, *IEEE Trans. Commun.* 51 (2003) 1325–1336.

- [83] R. Otnes, M. Tuchler, Iterative channel estimation for turbo equalization for time-varying frequency-selective channels, *IEEE Trans. Wireless Commun.* 3 (2004) 1918–1923.
- [84] S. Song, A. Singer, K. Sung, Soft input channel estimation for turbo equalization, *IEEE Trans. Signal Process.* 52 (2004) 2885–2894.
- [85] H. Kim, J. Tugnait, Turbo equalization for doubly-selective fading channels using nonlinear Kalman filtering and basis expansion models, *IEEE Trans. Wireless Commun.* 9 (2010) 2076–87.
- [86] H. Harashima, H. Miyakawa, Matched-transmission technique for channels with intersymbol interference, *IEEE Trans. Commun.* 20 (1972) 774–780.
- [87] J.C.-C. Kuo, S.-O. Tsai, L. Tadjpour, Y.-H. Chang, *Precoding Techniques for Digital Communication Systems*, Springer, New York, NY, 2008.
- [88] M. Tomlinson, New automatic equalizer employing modulo arithmetic, *Electron. Lett.* 7 (1971) 138–139.
- [89] G. Forney, M. Eyuboglu, Combined equalization and coding using precoding, *IEEE Commun. Mag.* 29 (1991) 25–34.
- [90] M. Vu, A. Paulraj, MIMO wireless linear precoding, *IEEE Signal Proc. Mag.* 24 (5) (2007) 86–105.
- [91] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, H. Sampath, Optimal designs for space-time linear precoders and decoders, *IEEE Trans. Signal Process.* 50 (2002) 1051–1063.
- [92] T. Liu, J.-K. Zhang, K. Wong, Optimal precoder design for correlated MIMO communication systems using zero-forcing decision feedback equalization, *IEEE Trans. Signal Proc.* 57 (2009) 3600–3612.
- [93] W. Su, W. Zheng, L. Chai, Precoder design and blind channel identification of MIMO channels, *IEEE Trans. Signal Process.* 54 (7) (2006) 2725–2736.
- [94] Z. Wang, G. Giannakis, Block precoding for MUI/ISI-resilient generalized multicarrier CDMA with multirate capabilities, *IEEE Trans. Commun.* 49 (11) (2001) 2016–2027.
- [95] J. Tugnait, S. He, H. Kim, Doubly-selective channel estimation using exponential basis models and subblock tracking, *IEEE Trans. Signal Process.* 58 (2010) 1275–1289.
- [96] X. Li, J. Ritcey, Bit-interleaved coded modulation with iterative decoding, in: Proceedings of IEEE International Conference Communication, 1999, pp. 858–864.
- [97] R. Lopes, J. Berry, The soft-feedback equalizer for turbo equalization of highly dispersive channels, *IEEE Trans. Commun.* 54 (5) (2006) 783–788.
- [98] F. Vogelbruch, S. Haar, Low complexity turbo equalization based on soft feedback interference cancellation, *IEEE Commun. Lett.* 9 (7) (2005) 586–588.
- [99] X. Wang, H. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception*, Prentice Hall, Upper Saddle River, NJ, 2004.
- [100] Y. Zheng, C. Xiao, Simulation models with correct statistical properties for Rayleigh fading channels, *IEEE Trans. Commun.* 51 (2003) 920–928.
- [101] R. Koetter, A. Singer, M. Tuchler, Turbo equalization, *IEEE Signal Proc. Mag.* 21 (2004) 67–80.

Blind Signal Separation for Digital Communication Data

4

Antoine Chevreuil^{*} and Philippe Loubaton[†]

^{*}ESIEE Paris/UMR 8049 LIGM, 2 bd Blaise Pascal BP 99, 93162 Noisy-le-grand cedex, France

[†]Université de Paris-Est Marne-la-Vallée, UMR 8049 LIGM, 5 bd Descartes, 77454 Marne-la-Vallée cedex 2, France

2.04.1 Introduction

2.04.1.1 Generalities on blind source separation

The goal of blind source separation is to retrieve the components of a mixture of independent signals when no a priori information is available on the mixing matrix. This question was introduced in the eighties in the pioneering works of Jutten and Herault [1]. Since then, Blind Source Separation (BSS), also called independent component analysis, was developed by many research teams in the context of various applicative contexts. The purpose of this chapter is to present BSS methods that have been developed in the past in the context of digital communications. In this case, K digital communication devices sharing the same band of frequencies transmit simultaneously K signals. The receiver is equipped with $M \geq K$ antennas, and has to retrieve a part of (or even all) the transmitted signals. The use of BSS techniques appears to be relevant when the receiver has no a priori knowledge on the channels between the transmitters and the receiver. As many digital communication systems use training sequences which allow one to estimate the channels at the receiver side, blind source separation is in general not a very useful tool. However, it appears to be of particular interest in contexts such as spectrum monitoring or passive listening in which it is necessary to characterize unknown transmitters (estimation of technical parameters such as the carrier frequency, symbol rate, symbol constellation,...) interfering in a certain bandwidth. For this, it is reasonable to try to firstly retrieve the transmitted signals, and then to analyze each of them in order to characterize the system it has been generated by. In this chapter, we provide a comprehensive introduction to the blind separation techniques that can be used to achieve the first step.

In order to explain the specificity of the problems we address in the following, we first recall what are the most classical BSS methodologies. The observation is a discrete-time M -variate signal $\mathbf{y}(n)$ defined as $\mathbf{y}(n) = \mathbf{H}\mathbf{s}(n)$ where the components of the K -dimensional ($K \leq M$) time series $\mathbf{s}(n) = (s_1(n), \dots, s_K(n))^T$ represent K signals which are statistically independent. The signal \mathbf{y} is thus an *instantaneous* mixture of the K independent source signals $(s_k)_{k=1,\dots,K}$ in the sense that $\mathbf{y}(n)$ only depends on the value of \mathbf{s} at time n . The signal \mathbf{y} is said to be a convolutive mixture of the K independent source signals $(s_k)_{k=1,\dots,K}$ if $\mathbf{y}(n) = \sum_\ell \mathbf{H}_\ell \mathbf{s}(n - \ell) = [\mathbf{H}(z)]\mathbf{s}(n)$ where $\mathbf{H}(z)$ represents the transfer function $\mathbf{H}(z) = \sum_\ell \mathbf{H}_\ell z^{-\ell}$. For the sake of simplicity, we just consider the context of instantaneous mixtures *in this introductory section*. The goal of blind source separation is to retrieve the signals $(s_k)_{k=1,\dots,K}$ from the sole knowledge of the observations. Fundamental results of Darmois

(see e.g., [2]) show that if the source signals are non Gaussian, then it is possible to achieve the separation of the sources by adapting a matrix \mathbf{G} in such a way that the components of $\mathbf{r}(n) = \mathbf{Gy}(n)$ are statistically independent. For this, it has been shown that it is sufficient to optimize over \mathbf{G} a function, usually called a contrast function, that can be expressed in terms of certain moments of the joint probability distribution of $\mathbf{r}(n)$. A number of successful contrast functions have been derived in the case where the signal $(s_k)_{k=1,\dots,K}$ are stationary sequences [2–5]. However, it will be explained below that in the context of digital communications, the signals $(s_k)_{k=1,\dots,K}$ are *not stationary, but cyclostationary*, in the sense that their statistical properties are almost periodic function of the time index. For example, for each k , the sequence $n \mapsto \mathbb{E}[|s_k(n)|^2]$ appears to be a superposition of sinusoids whose frequencies, called cyclic frequencies, depend on the symbol rate of transmitter k , and are therefore unknown at the receiver side. The cyclostationarity of the $(s_k)_{k=1,\dots,K}$ induces specific methodological difficulties that are not relevant in other applications of blind source separation.

2.04.1.2 Illustration of the potential of BSS techniques for communication signals

The example we provide is purely academical. We consider the transmission of two BSPK sequences modulated with a Nyquist raised-cosine filter (see Section 2.04.2.1) whose symbol period is $T = 1$ and roll-off factor is fixed to 50%. The energy per symbol equals \mathcal{E}_1 for the first source and \mathcal{E}_2 for the second source. The receiver has $M = 2$ antennas and the channel between the source no i and the antenna no j is a delay times a real constant $h_{i,j}$. After sampling at $T/3$, the noiseless model of the received data is $\mathbf{Y}(n) = \mathbf{Hs}(n)$ as specified in the introduction, where the component (i, j) of \mathbf{H} is $h_{i,j}$ (more details are provided in Section 2.04.2.3). An additive white noise Gaussian corrupts the model whose variance is $\frac{N_0}{T_e}$ where $\frac{\mathcal{E}_1}{N_0}$ is fixed to 100 dB (we purposely fixed the noise level to a low value in order to show results that can be graphically interpreted). Moreover, $\frac{\mathcal{E}_1}{\mathcal{E}_2} = 3$ dB.

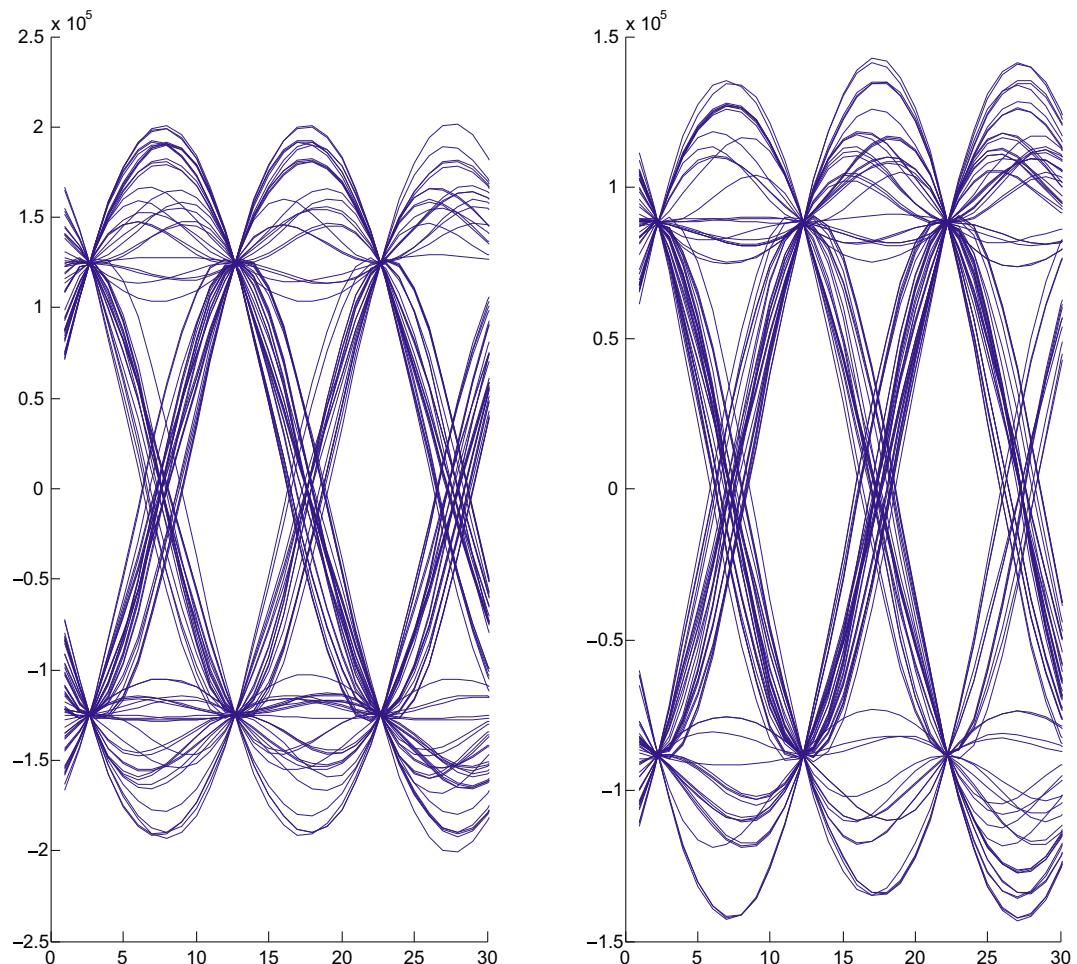
Suppose in a first step that the channel is ideal such that the mixing matrix \mathbf{H} is the identity matrix. We may have a look at the eye diagrams of the two components of the received data. We obtain Figure 4.1. This is almost a perfectly opened eye since the noise is negligible. We may also have a look at a 2D-histogram of the data. Notice that the components of $\mathbf{X}(n)$ are not stationary. We hence down-sample these data by a factor 3 in order to have stationary data. We plot the 2D-histogram: see Figure 4.2. As the two components are independent, their joint probability density function (pdf) is separable which seems to be the case in view of the figure.

Let us now consider the case of the channel matrix:

$$\mathbf{H} = \begin{pmatrix} 1 & 0.7 \\ 0.5 & 1 \end{pmatrix}.$$

We obtain Figures 4.3 and 4.4 respectively for the eye diagrams and the 2D-histograms. Clearly the channels are severe and close the eyes. Moreover, the pdf is obviously not separable, which attests to the non independency of the two components of $\mathbf{X}(n)$.

We run the JADE algorithm (see Sections 2.04.3.5 and 2.04.3.8) on the data (the observation duration is fixed to 1000 symbols): we obtain a 2×2 matrix \mathbf{G} such that, theoretically at least, \mathbf{GH} should be diagonal. We form the data $\mathbf{Y}(n) = \mathbf{GX}(n)$ and plot Figures 4.5 and 4.6. The eyes have been opened and

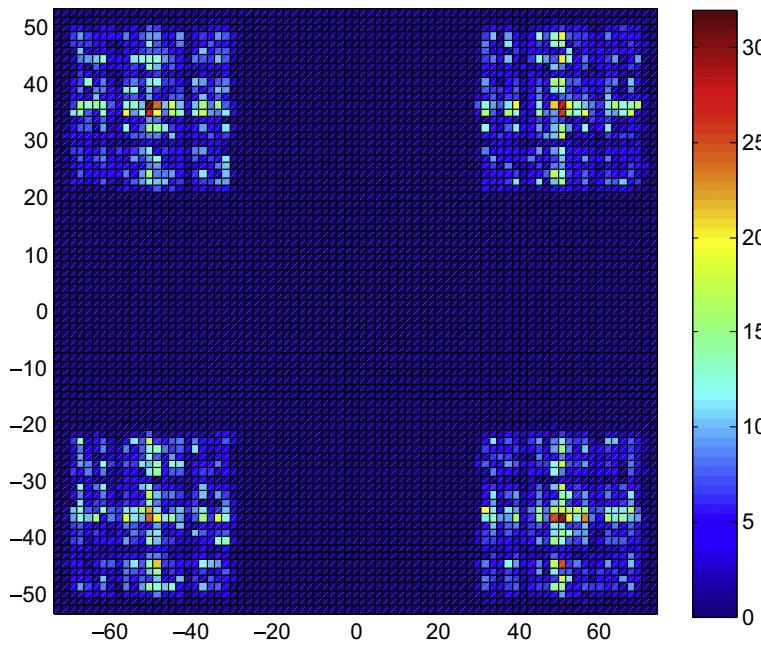
**FIGURE 4.1**

Eye diagram. $\frac{E_1}{N_0} = 100$ dB. After the JADE algorithm.

the joint pdf is hence separable. This is not a surprise since we have computed the resulting matrix \mathbf{GH} :

$$\mathbf{GH} \approx \begin{pmatrix} 0.8439 & -0.0129 \\ -0.0039 & 1.1829 \end{pmatrix}$$

which is close to a diagonal matrix. We need to explain why BSS has been successfully achieved in this simple example and why it can also be achieved in much more difficult contexts.

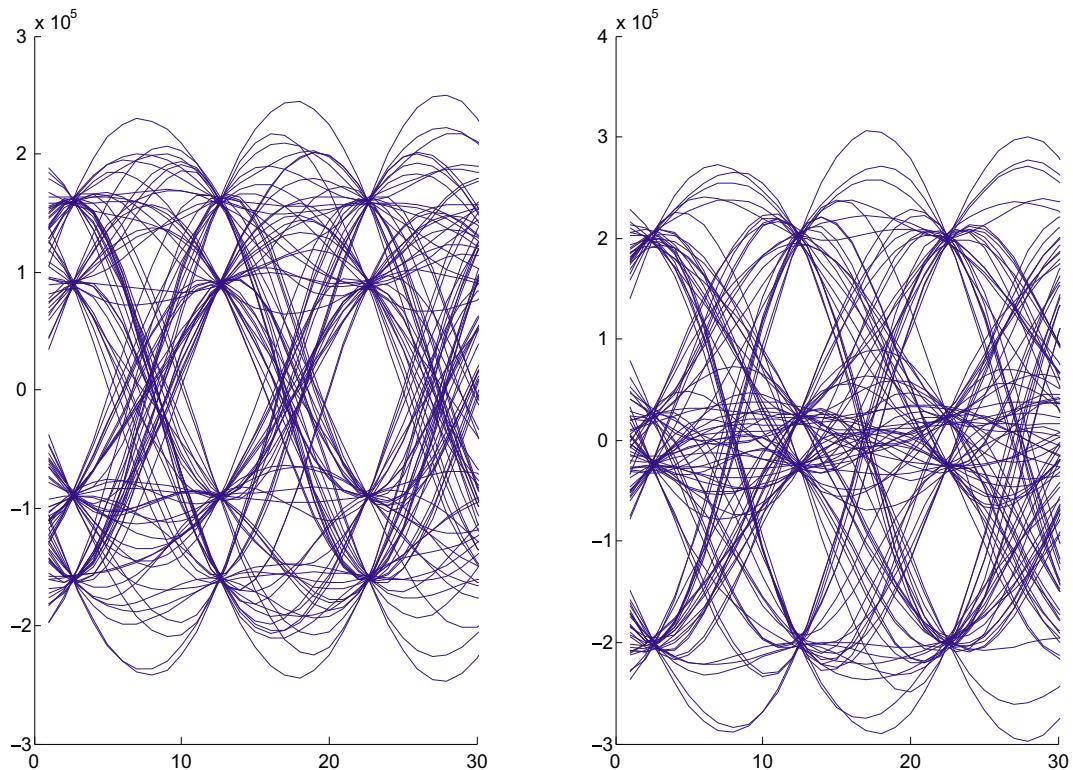
**FIGURE 4.2**

2D-histogram. $\frac{\mathcal{E}_1}{N_0} = 100 \text{dB}$. Ideal channels.

2.04.1.3 Organisation of the paper

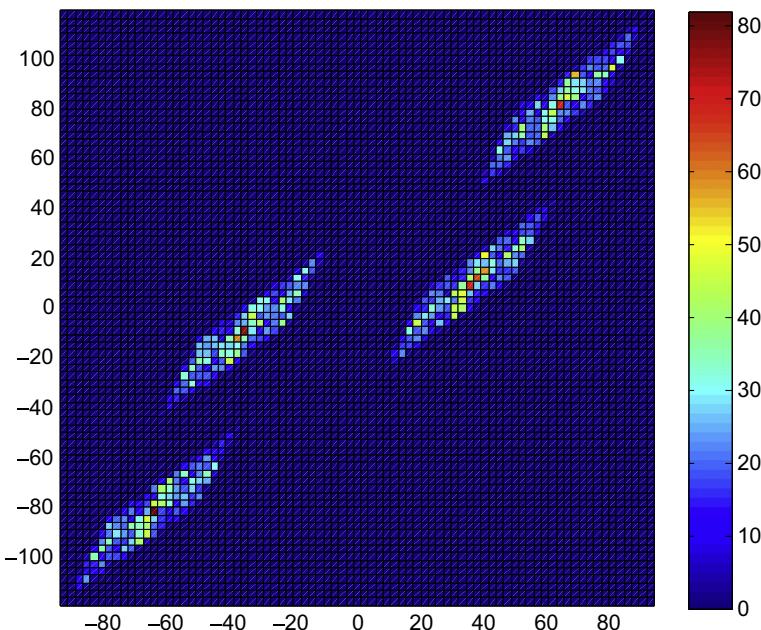
This chapter is organized as follows. In Section 2.04.2, we provide the model of the signals which are supposed to be linear modulations of symbols (Section 2.04.2.1). We discuss the statistics of the sampled versions of the transmitted sources in Section 2.04.2.2: in general, a sampled version is cyclo-stationary and we provide the basic tools and notation used along the paper. The model of the received data is specified in Section 2.04.2.3: (1) If the propagation channel between each transmitter and the receiver is a single path channel, the received signal is an instantaneous mixture of the transmitted signals; (2) if at least one of the propagation channel is a multipath channel, the mixture appears to be convolutive. Besides, we discuss the assumptions under which the received data are stationary. In general, however, the data are cyclo-stationary with unknown cyclic frequencies.

The case of instantaneous mixtures is addressed in Section 2.04.3. When the sources are independent and identically distributed (i.i.d.) (this case is discussed in Section 2.04.2.3), and that strong a priori information on the constellations are known, it is possible to provide algebraic solutions to the BSS problem, e.g., the Iterative Least Squares Projections (ILSP) algorithm or the Algebraic Constant Modulus Algorithms (ACMA): these methods are explained in Section 2.04.3.2. In Section 2.04.3.3, we consider the case of second-order methods (one of the advantages of these latter is that they are robust to the cyclo-stationarities, hence can be applied to general scenarios): the outlines of one of

**FIGURE 4.3**

Eye diagrams. $\frac{\mathcal{E}_1}{N_0} = 100$ dB. Channel \mathbf{H} . Left: first component. Right: second component.

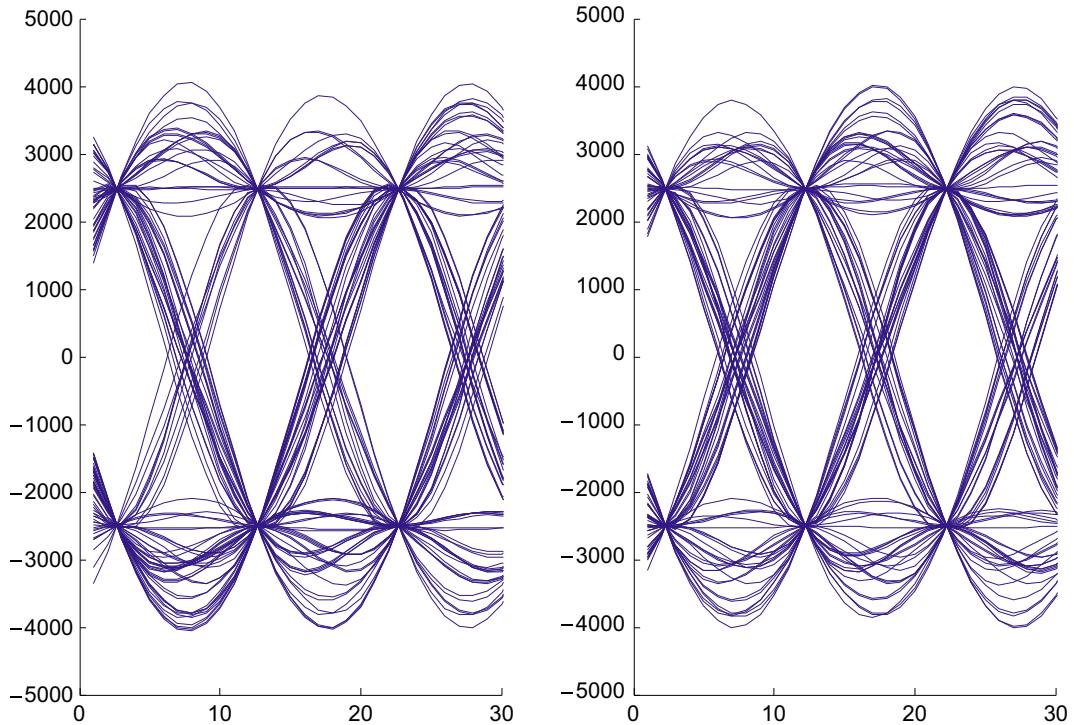
the most popular approach, the Second-Order Blind Identification (SOBI) algorithm, which consists in estimating the mixing matrix from the autocorrelation function of the received signal. This approach is conceptually simple, and the corresponding scheme allows one to identify the mixture and hence, to separate the source signals. That SOBI is rarely considered for BSS of digital communication signals is explained. The subsections that follow cope with BSS methods based on fourth-order cumulants. They are called “direct” BSS methods since they provide estimates of the sources with no prior estimation of the unknown channel matrix. For pedagogical and historical reasons, we firstly cope with the very particular case of stationary signals. One-by-one methods based are explained (Section 2.04.3.4) and are shown to be convergent; the associated deflation procedure is introduced and an improvement is presented. Global methods (also called *joint separating methods*) aim at separating jointly the K sources: they are depicted in Section 2.04.3.5; these approaches are based on the minimization of well chosen contrast functions over the set of $K \times K$ unitary matrices: the famous Joint Approximate Diagonalization of Eigenmatrices (JADE) algorithm is presented, since it represents a touchstone in the domain of BSS. When the sources are cyclo-stationary, which is really the interesting point for the

**FIGURE 4.4**

2D-histogram. $\frac{\mathcal{E}_1}{N_0} = 100$ dB. Channel \mathbf{H} .

context of this paper, the preceding “stationary” methods (one-by-one and global) are again considered. The following problem is addressed: do the convergence result still hold when the algorithms are fed by cyclo-stationary data instead of stationary ones? Sufficient conditions are shown to assure the convergence: semi-analytical computations (Section 2.04.3.9) prove that the conditions in question hold true.

In Sections 2.04.4 and 2.04.5, the case of convolutive mixtures is addressed. In certain particular scenarios, e.g., sparse channels, the gap between the instantaneous case and the convolutive one can be bridged quite directly (Section 2.04.4). More precisely, if the delays of the various multipaths are sufficiently spread out on the one hand and if, on the other hand, the number of antennas of the receiver is large enough, it is still possible to formulate the source separation problem as the separation of a certain instantaneous mixture. If these conditions do not hold, we face a real convolutive mixture, i.e., the received data are the output of a Multi-Input/Multi-Output (MIMO) unknown filter driven by jointly independent (cyclo-) stationary sources. Due to their historical and theoretical importance, we present algebraic methods (Section 2.04.5.1) when the data are stationary. Under this latter assumption, the identification of the unknown transfer function can be achieved using standard methods using the Moving Average (MA) or Auto-Regressive (AR) properties: see Section 2.04.5.2. The famous subspace method, introduced in Section 2.04.5.3, is based on second-order moments and can be used for general cyclo-stationary data; its inherent numerical problems are discussed. In Section 2.04.5.4, global direct

**FIGURE 4.5**

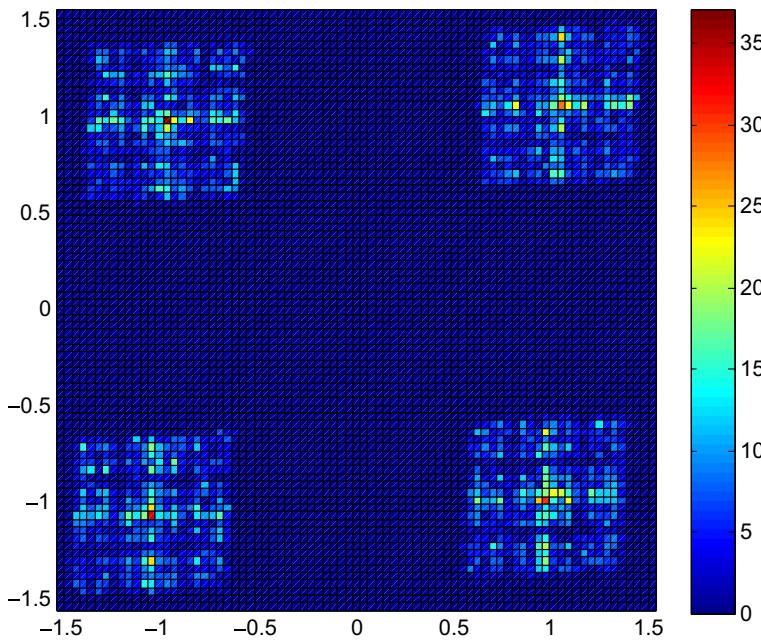
Eye diagram. $\frac{\mathcal{E}_1}{N_0} = 100$ dB. After the JADE algorithm.

methods are evoked (temporal domain and frequency domain) for stationary data. In Section 2.04.5.5, the case of one-to-one methods previously introduced in Section 2.04.3.4 is extended to the convolutive case and positive results for BSS are provided. The results are further extended for the cyclo-stationary case in Section 2.04.5.6 where convergence results are shown.

In Section 2.04.7, we discuss several points that have not been developed in the core of the paper. Further bibliographic entries are provided.

2.04.2 Signals

We have specified in the Introduction that the domain of source separation is not restricted to the context of telecommunication signals. In the following, however, most of the results apply specifically to digital telecommunication signals.

**FIGURE 4.6**

2D-histogram. $\frac{\mathcal{E}_1}{N_0} = 100$ dB. After the JADE algorithm.

2.04.2.1 Source signals. Basic assumptions

We assume that K digital telecommunication devices simultaneously transmit information in the same band of frequencies. For $t \in \mathbb{R}$, we denote by $s_{a,k}(t)$ the complex envelope of the k th transmitted signal (“the k th source”). The subscript “ a ” in $s_{a,k}(t)$ underlines that the signal is “analog”. Throughout this contribution, $s_{a,k}(t)$ is supposed to stem from a linear modulation of a sequence of symbols. The model is hence:

$$\forall t \in \mathbb{R} \quad s_{a,k}(t) = \sum_n d_k(n) c_{a,k}(t - nT_k). \quad (4.1)$$

In this latter equation, $d_k(n)$ is a sequence of symbols belonging to a certain constellation. The function $c_{a,k}$ is a shaping function and T_k is the duration of a symbol $d_k(n)$. We denote by f_k the carrier frequency associated with the k th source. Along this contribution, the following assumptions and notation are adopted.

Assumptions on the source signals: For a given index k , the sequence $d_k(n)$ is assumed to be independent and identically distributed (i.i.d.). We assume that it has zero mean $E[d_k(n)] = 0$. With no restriction at all, the normalization $E[|d_k(n)|^2] = 1$ holds. We also suppose that it is second-order complex-circular in the sense that

$$E[d_k(n)^2] = 0. \quad (4.2)$$

It is undoubtedly a restriction to impose the condition (4.2) especially in the telecommunication context of this paper; indeed, the BPSK modulation for instance does not verify (4.2). Some points on the extensions to general non circular mixtures are provided in Section 2.04.7.1.

The Kurtosis of the symbol $d_k(n)$, is defined as

$$\kappa(d_k) = \text{cum}(d_k(n), d_k(n)^*, d_k(n), d_k(n)^*),$$

where the fourth-order cumulant of a complex-valued random variable X is defined when it makes sense, as $\text{cum}(X, X^*, X, X^*) = E[|X|^4] - 2(E[|X|^2])^2 - |E[X^2]|^2$: see for instance [6]. Here, by the circularity Assumption 4.2, we have

$$\kappa(s_k) = E[|d_k(n)|^4] - 2(E[|d_k(n)|^2])^2.$$

We assume that we have, for any index k :

$$\kappa(d_k) < 0. \quad (4.3)$$

This inequality is given as an assumption; it has more the flavor of a result, since we do not know complex-circular constellations such that (4.3) is not satisfied.

We may now come to the key assumption: the sources d_1, d_2, \dots, d_K are mutually independent.

Concerning the shaping filter, $c_{a,k}$ we suppose that $c_{a,k}$ is a square-root raised cosine with excess bandwidth (also called *roll-off factor*) $0 \leq \gamma_k \leq 1$.

2.04.2.2 Cyclo-stationarity of a source

In this short paragraph, we drop the index of the source and $s_a, T, c_a, d(n)$ refer to respectively $s_{a,k}, T_k, c_{a,k}, d_k(n)$. Thanks to Eq. (4.1), it is quite obvious that $s_a(t)$ and $s_a(t + T)$ are similarly distributed since $d_k(n)$ is i.i.d. This simple reasoning applies to any vector $(s_a(t_1), \dots, s_a(t_m))$ whose distribution equals this of $(s_a(t_1 + T), \dots, s_a(t_m + T))$. This shows that the process s_a is cyclo-stationary in the strict sense with period T . In particular, its second and fourth-order moments evolve as T -periodic functions of the time. Let us focus on the second-order moments: $t \mapsto E[s_a(t + \tau)s_a(t)^*]$ is hence a periodic function with period T . We let its Fourier expansion be

$$E[s_a(t + \tau)s_a(t)^*] = \sum_{m \in \mathbb{Z}} R_{s_a}^{(m/T)}(\tau) e^{i2\pi m t/T}, \quad (4.4)$$

where $R_{s_a}^{(m/T)}(\tau)$ is called “cyclo-correlation” of s_a at cyclic frequency m/T and time lag τ . We have the reverse formula:

$$R_{s_a}^{(m/T)}(\tau) = \frac{1}{T} \int_0^T E[s_a(t + \tau)s_a(t)^*] e^{-i2\pi m t/T} dt$$

or

$$R_{s_a}^{(m/T)}(\tau) = \lim_{\Delta \rightarrow \infty} \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} E[s_a(t + \tau)s_a(t)^*] e^{-i2\pi m t/T} dt.$$

Generally, we may introduce the cyclo-correlation at any cyclic frequency α :

$$R_{s_a}^{(\alpha)}(\tau) = \lim_{\Delta \rightarrow \infty} \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} E[s_a(t + \tau)s_a(t)^*] e^{-i2\pi \alpha t} dt.$$

In the case of s_a given by Eq. (4.1), $R_{s_a}^{(\alpha)}(\tau)$ is identically zero for cyclic frequencies α that are not multiples of $1/T$. In passing, we have the following symmetry:

$$R_{s_a}^{(\alpha)}(\tau) = R_{s_a}^{(-\alpha)}(-\tau)^*. \quad (4.5)$$

Let us inspect a bit further a main specificity of linear modulations on the Fourier expansion of Eq. (4.4). In this respect, denote by $f \mapsto S_{s_a}^{(\alpha)}(f)$ the Fourier transform of the cyclo-correlation function $\tau \mapsto R_{s_a}^{(\alpha)}(\tau)$. We have, after elementary calculus (see [7,8]):

$$S_{s_a}^{(m/T)}(f) = \frac{1}{T} \hat{c}_a(f) \hat{c}_a(f - m/T)^*,$$

where $\hat{c}_a(f) = \int_{\mathbb{R}} c_a(t) e^{-i2\pi ft} dt$ is the Fourier transform of c_a . This formula is visibly a generalization of the so-called Benett Equality (see [9] Section 4.4.1) that gives the power spectral density of s_a : indeed, in the above equation, if one takes $m = 0$, we obtain $S_{s_a}^{(0)}(f) = \frac{1}{T} |\hat{c}_a(f)|^2$ which is the power spectral density. An important consequence was underlined in [8]:

Lemma 1. *For any excess bandwidth factor γ such that $\gamma \leq 1$, we have:*

$$\forall |m| > 1 \quad R_{s_a}^{(m/T)}(\tau) \text{ is uniformly zero.}$$

In other words, the cyclic frequencies of $s_a(t)$ given by Eq. (4.1) belong to the set $\{0, \pm\frac{1}{T}\}$.

Proof. The proof is obvious since the support of $\hat{c}_a(f)$ is $[-\frac{1+\gamma}{2T}, \frac{1+\gamma}{2T}]$ with $\gamma < 1$ hence the supports of $\hat{c}_a(f)$ and $\hat{c}_a(f - m/T)^*$ do not overlap except if $m = 0, \pm 1$. \square

We deduce from what precedes some consequences on the second-order statistics of a sampled version of a source. In this respect, we denote by T_e any sampling period; the discrete-time signal associated with $s_a(t)$ is hence $s(n) = s_a(nT_e)$ for $n \in \mathbb{Z}$. Thanks to Lemma 1, the expansion (4.4) may be re-written as:

$$\forall \ell \in \mathbb{Z} \quad E[s(n + \ell)s(n)^*] = \sum_{m \in \{-1, 0, 1\}} R_{s_a}^{(m/T)}(\ell T_e) e^{i2\pi nm\alpha}, \quad (4.6)$$

where we let α be $\alpha = T_e/T$. We distinguish between three cases:

1. If $\alpha = T_e/T$ is a integer, the three terms of the r.h.s. of (4.6) all aggregate in a single term, making the function $E[s(n + \ell)s(n)^*]$ not depend on the time index n . This is not surprising since the condition $T_e = pT$ where p is a non-null integer corresponds to a strict-sense stationary signal (see the polyphase decomposition in [10]). In particular, if $T_e = T$, we have:

$$s(n) = [c_0(z)]d(n), \quad (4.7)$$

- where $c_0(z) = \sum_n c_a(nT)z^{-n}$. In the following, we will not study the case $T_e = pT$ with $p \geq 2$.
2. If $\alpha \neq 0$ modulo $1/2$: unless α is rational, it cannot be said that, as a function of n , $E[s(n + \ell)s(n)^*]$ is periodic: it is called an almost periodic function [11] and $s(n)$ is hence called almost periodically correlated, having $0, \pm\alpha$ as cyclic-frequencies. We introduce

$$R_s^{(\alpha)}(\ell) = \langle E[s(n + \ell)s(n)^*] e^{-i2\pi\alpha n} \rangle_n, \quad (4.8)$$

where the operator $\langle \cdot \rangle$ is the time-averaging one, i.e., for a complex-valued (deterministic) series $u(n)$

$$\langle u(n) \rangle_n = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} u(n)$$

when the limit makes sense. As the three cyclic-frequencies in the expansion (4.6) are distinct modulo 1 then we obtain

$$\forall m \in \{0, \pm 1\} \quad R_s^{(mT_e/T)}(\ell) = R_{s_a}^{(m/T)}(\ell T_e),$$

which reminds us of the Shannon sampling theorem.

3. If $\alpha = \frac{1}{2}$ it turns out that, similarly to the previous case, the discrete-time source is cyclo-stationary, having $0, \pm \frac{1}{2}$ as cyclic-frequencies. Moreover $R_s^{(0)}(\ell) = R_{s_a}^{(0)}(\ell T_e)$ and $R_s^{(1/2)}(\ell) = R_{s_a}^{(-1/2)}(\ell) = R_{s_a}^{(1/T)}(\ell T_e) + R_{s_a}^{(-1/T)}(\ell T_e)$.

2.04.2.3 Received signals

The receiver is equipped with M antennas, the number of antennas being as big as the number of sources, i.e., $M \geq K$ (see section 2.04.2.3.2). We denote by

$$\mathbf{y}_a(t) = \begin{pmatrix} y_a^{(1)}(t) \\ \vdots \\ y_a^{(m)}(t) \end{pmatrix}$$

the complex envelope of the received $M \times 1$ vector computed at a frequency of demodulation denoted by f_0 . We consider that $\mathbf{y}_a(t)$ obeys the linear model

$$\mathbf{y}_a(t) = \sum_{k=1}^K \mathbf{y}_{a,k}(t),$$

where $\mathbf{y}_{a,k}(t)$ is the contribution of the k th source to the observation. We further assume that $\mathbf{y}_{a,k}(t)$ stems from delayed/attenuated versions of $s_{a,k}(t)$. In this respect, we may write $y_{a,k}^{(m)}(t)$, the component of $\mathbf{y}_{a,k}(t)$ associated with the m th sensor, as:

$$y_{a,k}^{(m)}(t) = \sum_{\ell=1}^{L_k} \lambda_{k,m,\ell} e^{-i2\pi f_0 \tau_{k,m,\ell}} \tilde{s}_{a,k}(t - \tau_{k,m,\ell}), \quad (4.9)$$

where the index ℓ represents the path index, L_k the number of paths associated with the source no k , $\lambda_{k,m,\ell}$ an attenuation factor and $\tau_{k,m,\ell}$ the delay of the propagation along the path no ℓ between the source no k and the sensor no m . In this latter equation, $\tilde{s}_{a,k}(t)$ is the complex envelope of the modulated signal $s_{a,k}(t)e^{i2\pi f_k t}$ at the demodulation frequency, i.e., $\tilde{s}_{a,k}(t) = s_{a,k}(t)e^{i2\pi(f_k - f_0)t}$.

2.04.2.3.1 Models of the sampled data

We distinguish between two cases:

- Instantaneous mixture:* This scenario holds when the signal $\mathbf{y}_{a,k}(t)$ evolves in a linear space of dimension 1, that is when the components $\tilde{s}_{a,k}(t - \tau_{k,m,\ell})$ in (4.9) do not depend neither on ℓ nor on m . This holds when there exists τ_k such that $|\tau_{k,m,\ell} - \tau_k| \ll T_k$ for all indices m, ℓ . This happens, for instance, when there is a single path ($L_k = 1$) and the transmitted signal is narrow-band ($f_k \gg 1/T_k$). In this case we have

$$\mathbf{y}_a(t) = \sum_{k=1}^K \mathbf{h}_{1,k} s_{a,k}(t - \tau_k) e^{j2\pi(f_k - f_0)t}.$$

More compactly, this gives

$$\forall t \in \mathbb{R} \quad \mathbf{y}_a(t) = \mathbf{H} \mathbf{s}_a(t),$$

where $\mathbf{H} = (\mathbf{h}_{1,1}, \dots, \mathbf{h}_{1,K})$ is a $M \times K$ mixing matrix, and

$$\mathbf{s}_a(t) = (s_{a,1}(t - \tau_1) e^{j2\pi(f_1 - f_0)t}, \dots, s_{a,K}(t - \tau_K) e^{j2\pi(f_K - f_0)t})^T. \quad (4.10)$$

If T_e is the sampling period of the receiver, it is supposed that all the components of the data are low-pass filtered in the sampling band (the matched-filter cannot be considered since the shaping filters are not supposed to be known to the receiver). Finally, the (noiseless model) of the data is

$$\text{General instantaneous model: } \forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \mathbf{H} \mathbf{s}(n), \quad (4.11)$$

where

$$\mathbf{s}(n) = (s_{a,1}(nT_e - \tau_1) e^{j2\pi(f_1 - f_0)nT_e}, \dots, s_{a,K}(nT_e - \tau_K) e^{j2\pi(f_K - f_0)nT_e})^T. \quad (4.12)$$

Generally speaking, any of the components of the source vector $\mathbf{s}(n)$ is cyclo-stationary (see Section 2.04.2.2) hence the model given by (4.11) is a cyclo-stationary one. For simplification, let us suppose in the following that $f_k = f_0$ for all the indices k (this point is discussed in Section 2.04.7). As the original theory of BSS assumed stationary data, we inspect under which conditions the above model can be stationary. A necessary and sufficient condition is that all the components of $\mathbf{s}(n)$ be stationary. As discussed previously, this can happen when all the symbol periods are equal to, say, T and if the sampling period $T_e = T$. Under these conditions, we even have:

$$s_{a,k}(t - \tau_k) = [c_{\tau_k,k}(z)]d_k(n),$$

where we have set for any delay τ

$$c_{\tau,k}(z) = \sum_{\ell} c_{a,k}(\ell T_k - \tau) z^{-\ell}. \quad (4.13)$$

The stationary model can be written as:

$$\text{Stationary instantaneous model: } \forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \mathbf{H} \mathbf{s}(n), \quad (4.14)$$

where we have set

$$\mathbf{s}(n) = \begin{pmatrix} [c_{\tau_1,1}(z)] & d_1(n) \\ \vdots & \vdots \\ [c_{\tau_K,K}(z)] & d_K(n) \end{pmatrix}. \quad (4.15)$$

(note: the notation might be confusing since $\mathbf{s}(n)$ was already defined in (4.12): in the following sections, the context is always specified which prevents the confusion) In the literature, it is sometimes required that the sources be i.i.d. In the context of this paper, this i.i.d. condition is fulfilled when the filters $c_{\tau_k,k}(z)$ all have the form of a constant times a delay: in short, this happens when (1) all the transmitted symbols are synchronized (2) the receiver runs a matched filter (square-root Nyquist), and (3) the symbol synchronization is performed at the receiver. In this case we have:

$$\text{i.i.d. instantaneous model: } \forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \mathbf{H} \begin{pmatrix} d_1(n) \\ \vdots \\ d_K(n) \end{pmatrix}. \quad (4.16)$$

The reader may find this set of condition very restrictive in real scenarios. It is indeed; however, the developments of BSS are based on the stationary assumption. Moreover, many interesting methods exploit the i.i.d. condition.

2. *Convulsive mixture*: This is the general case when multi-paths affect the propagation. We provide the discrete-time version of Eq. (4.9). Let us begin by the general case. In this respect, we assume that the sampling period T_e verifies the Shannon sampling condition, i.e.,

$$\frac{1}{T_e} > \max_k \left(\frac{1 + \gamma_k}{T_k} + |f_k - f_0| \right).$$

This is a non-restrictive condition whatever the scenario: a crude prior spectral analysis of the data is simply needed. Provided this condition, the discrete-time signal $s_{a,k}(nT_e - \tau_{k,m,\ell})$, for any indices k, m, ℓ , is a filtered version of $(s_{a,k}(nT_e))_{n \in \mathbb{Z}}$. It is hence easy to deduce that the sampled data $\mathbf{y}(n)$ follows the equation:

$$\text{General convulsive model: } \forall n \in \mathbb{Z} \quad \mathbf{y}(n) = [\mathbf{H}(z)]\mathbf{s}(n), \quad (4.17)$$

where $\mathbf{s}(n) = (s_{a,1}(nT_e)e^{j2\pi(f_1-f_0)nT_e}, \dots, s_{a,K}(nT_e)e^{j2\pi(f_K-f_0)nT_e})$ is a vector of mutually independent sources and $\mathbf{H}(z)$ is certain the $M \times K$ transfer function whose k th column is the digital channel between the k th source and the receiver: it depends on the parameters $(\lambda_{k,m,\ell}, \tau_{k,m,\ell})_{\ell=1,\dots,L_k, m=1,\dots,M}$. The above general model is, in general, cyclo-stationary. For simplification, we assume ion the following that $f_k = f_0$ for all indices k . Similarly to the case of instantaneous mixtures, it is instructive to find conditions under which the data are stationary. This occurs when the symbol periods T_k all coincide with a certain T , and when the sampling period T_e equals T . Under all these conditions, $\mathbf{y}_k(n)$, the contribution of the k th source to the mixture, can be written as

$$\mathbf{y}_k(n) = \begin{pmatrix} \sum_{\ell} \lambda_{k,1,\ell} [c_{\tau_{k,1,\ell,k}(z)}] d_k(n) \\ \vdots \\ \sum_{\ell} \lambda_{k,M,\ell} [c_{\tau_{k,M,\ell,k}(z)}] d_k(n) \end{pmatrix} = [\mathbf{E}_k(z)] d_k(n),$$

hence, setting $\mathbf{E}(z) = (\mathbf{E}_1(z), \dots, \mathbf{E}_K(z))$, it yields

$$\text{Stationary convolutive model: } \forall n \in \mathbb{Z} \quad \mathbf{y}(n) = [\mathbf{E}(z)]\mathbf{d}(n), \quad (4.18)$$

where $\mathbf{d}(n) = (d_1(n), \dots, d_K(n))^T$ and $\mathbf{E}(z)$ is a certain $M \times K$ unknown filter matrix. This shows that $\mathbf{E}(z)$ depends on the shaping filters, the steering vectors associated with the paths and their corresponding delays.

2.04.2.3.2 Assumptions on the channels

In this paper, we consider over-determined mixtures, that is: mixtures such that the number of sensors exceeds the number of sources ($M \geq K$). This condition is necessary in order to retrieve the vector $\mathbf{s}(n)$ —see model (4.11) (respectively (4.17))—from the data $\mathbf{y}(n)$ by means of a $K \times M$ constant matrix (respectively a $K \times M$ filter). This has to be specified.

For instantaneous mixtures, the following condition holds:

$$\text{Assumption (instantaneous mixtures): } \text{rank}(\mathbf{H}) = K. \quad (4.19)$$

Under this assumption, there exist $K \times M$ matrices \mathbf{G} such that $\mathbf{G}\mathbf{y}(n) = \mathbf{s}(n)$.

For convolutive mixtures, it is conventional to assume that the components of $\mathbf{H}(z)$ are polynomials in z^{-1} (this is an approximation that is justified since the shaping filters $c_{a,k}(t)$ are rapidly vanishing when $|t| \rightarrow \infty$). We further assume that

$$\text{Assumption (convolutive cyclostationary mixtures): } \forall z \neq 0 \quad \text{rank}(\mathbf{H}(z)) = K. \quad (4.20)$$

Under this condition, there exist polynomial matrices $\mathbf{G}(z)$ such that $[\mathbf{G}(z)]\mathbf{y}(n) = \mathbf{s}(n)$: see for instance [12, 13]. The same kind of assumption holds in the stationary case—see the model (4.18): namely,

$$\text{Assumption (convolutive stationary mixtures): } \forall z \neq 0 \quad \text{rank}(\mathbf{E}(z)) = K. \quad (4.21)$$

At this level, we would like to point out a curiosity. In this respect, we assume further that the excess bandwidth factors of one source—say the first one—equals zero. As the choice $T_e = T_1$ satisfies the Shannon sampling condition, we may write $s_{a,1}(nT_1 - \tau) = [\phi_\tau(z)]s_{a,1}(nT_1)$ where

$$\phi_\tau(z) = \sum_k \frac{T_1}{\pi(kT_1 - \tau)} \sin\left(\frac{\pi(kT - \tau)}{T_1}\right) z^{-k}.$$

As $s_{a,1}(nT_1) = [c_{0,1}(z)]d_1(n)$, the first column $\mathbf{E}_1(z)$ of $\mathbf{E}(z)$ can be factored as $c_{0,1}(z) \tilde{\mathbf{E}}_1(z)$. In particular, after the standard FIR approximations, it yields that the condition given in (4.21) is not fulfilled.

2.04.3 Instantaneous mixtures

The model of the data $\mathbf{y}(n)$ is given by (4.11). The mixing matrix \mathbf{H} is unknown. BSS can be achieved either by estimating \mathbf{H} —this is the point of Section 2.04.3.3—or by computing directly estimates of the sources (up to indeterminacies).

2.04.3.1 Indeterminacies

It is always possible to consider that the sources have equal and normalized power. Indeed, as $s_1(n) = \rho_1 \frac{s_1(n)}{\rho_1}$ where $\rho_1 = \sqrt{\langle E[|s_1(n)|^2] \rangle}$ is the square-root of the power of the first source, we suggest to scale the first column of \mathbf{H} by ρ_1 . Repeating this process for all the sources, we have constructed a new matrix $\tilde{\mathbf{H}}$. Eventually denoting $\tilde{\mathbf{s}}(n) = \left(\frac{s_1(n)}{\rho_1}, \frac{s_2(n)}{\rho_2}, \dots, \frac{s_K(n)}{\rho_K} \right)^T$, we obviously show that the data alternatively writes

$$\forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \tilde{\mathbf{H}} \tilde{\mathbf{s}}(n).$$

Though apparently innocent, this remark gives precious a priori indications. First of all, it says that the model (4.22) is not uniquely defined. As a consequence, it is always possible to consider, without restricting the model, that the sources have equal power equal to one—this precisely corresponds to the above defined $\tilde{\mathbf{s}}(n)$; specifically, we will assume in the following that

$$\left\langle E \left[\mathbf{s}(n) \mathbf{s}(n)^H \right] \right\rangle = \mathbf{I}_K \quad (4.22)$$

This shows it is beyond a reasonable expectation to retrieve the sources with no scaling ambiguities. Similarly, if \mathbf{P} is a permutation matrix, $\mathbf{y}(n) = (\mathbf{H}\mathbf{P}^{-1})(\mathbf{P}\mathbf{s}(n))$, underlining the non-uniqueness of the model.

With no further assumptions on the sources, the ultimate result that can be achieved is: retrieve the sources up to unknown *complex scaling factors* (scaling and phase ambiguities) and a *permutation*.

2.04.3.2 Algebraic methods (i.i.d. scenario)

The model of the data is given by (4.16). We may collect the N available data in a $M \times N$ matrix \mathcal{Y} , we have: $\mathcal{Y} = \mathbf{H}\mathcal{D}$ where $\mathcal{D} = (\mathbf{d}(0), \dots, \mathbf{d}(N-1))$. As any entry of \mathcal{D} corresponds to a symbol, associated specificities (e.g., finite alphabet constellations or modulus one symbols) are a priori relations the receiver can make use of. As far as the identifiability is concerned, it is proven in [14] (Lemma 1) that the above factorization is essentially unique for modulus one symbols, at least if the number of snapshots N verifies $N \geq 2K$ (which is the case in practical contexts). By essentially unique, we mean that the rows of \mathcal{D} may be permuted and/or multiplied by modulus one constants.

Talwar et al. [15, 16] propose iterative algorithms that assume known the alphabets of the symbols. Call $\hat{\mathbf{H}}^{(\ell)}$ an estimate of \mathbf{H} at the iteration no ℓ . The Iterative Least Square with Projection (ILSP) is:

1. Take any full rank $\hat{\mathbf{H}}^{(0)}$ for iteration $\ell = 0$
2. $\ell \leftarrow \ell + 1$
 - $\hat{\mathcal{D}}^{(\ell)} = (\hat{\mathbf{H}}^{(\ell-1)})^\sharp \mathcal{Y}$ where $(.)^\sharp$ denotes the pseudo-inverse.
 - $\hat{\mathcal{D}}^{(\ell)} \leftarrow$ projection of each component of $\hat{\mathcal{D}}^{(\ell)}$ on the corresponding alphabet.
 - $\hat{\mathbf{H}}^{(\ell)} = \mathcal{Y}(\hat{\mathcal{D}}^{(\ell)})^\sharp$.

Similar projection-based algorithms that rather take into account the constant modulus property of the entries of \mathcal{D} have been considered [17, 18]: similarly to the IMSP algorithm, no results on the convergence can be given (how many samples are required? are there local minima the algorithm could be trapped in?). Van der Veen et al. [14] proposes a non-iterative algorithm, called the Algebraic CMA

(ACMA): the ACMA provides exactly “the” solution (up to the above mentioned ambiguities) of the factorization of \mathcal{Y} —at least if the number of data N exceeds K^2 . It is based on a joint diagonalization of a pencil of K matrices.

Certain BSS methods for convolutive mixtures need, as a final step, to run such algorithms (see e.g., Section 2.04.5.1).

2.04.3.3 Second-order based identification (general cyclo-stationary case)

In this section, we address the “indirect” BSS; by this terminology, we mean that the BSS is achieved in two steps. The first step consists in estimating the unknown mixing matrix \mathbf{H} by, say, $\hat{\mathbf{H}}$. In a second step, the proper separation is carried out. If $\hat{\mathbf{H}}$ is an accurate estimate, then $\hat{\mathbf{s}}(n) = \hat{\mathbf{H}}^\dagger \mathbf{y}(n)$ is the natural estimate of the source vector. In general, however, noise is present (estimation noise and additive noise in the observed signals) and other strategies have to be considered: this aspect is not addressed in this paper.

The first point to be addressed in this section is the *pre-whitening* of the data. We suppose that $M = K$ (notice: in the non-square case, a principal component analysis is processed). In this respect, we consider the auto-correlation matrix of the data $\mathbf{R}_y^{(0)}(0) = \langle E[\mathbf{y}(n)\mathbf{y}(n)^H] \rangle$ can be written (we recall that the sources are assumed to have equal normalized powers as discussed previously):

$$\mathbf{R}_y^{(0)}(0) = \mathbf{H}\mathbf{H}^H.$$

Since \mathbf{H} is full rank, the above matrix is positive definite and we form the new data

$$\mathbf{x}(n) = (\mathbf{R}_y^{(0)}(0))^{-1/2} \mathbf{y}(n).$$

We have:

$$\mathbf{x}(n) = \mathbf{U}\mathbf{s}(n),$$

where $\mathbf{U} = (\mathbf{H}\mathbf{H}^H)^{-1/2} \mathbf{H}$ is a unitary matrix.

The second point concerns the estimation of the unitary matrix \mathbf{U} . The data $\mathbf{x}(n)$ is cyclo-stationary. As the cyclic-frequencies are not always directly accessible, the identification of the unknown mixing matrix \mathbf{U} is done by solely considering the statistics

$$R_x^{(0)}(\ell) = \langle \mathbf{x}(n + \ell) \mathbf{x}(n)^H \rangle,$$

which can be expressed as

$$R_x^{(0)}(\ell) = \mathbf{U} R_s^{(0)}(\ell) \mathbf{U}^H.$$

This says that the normal matrix $R_x^{(0)}(\ell)$, for any index ℓ , is diagonalized in the orthonormal basis formed by the columns of \mathbf{U} . For $\ell = 0$, this gives $R_x^{(0)}(0) = \mathbf{U} \mathbf{U}^H = \mathbf{I}$ and this is clearly not sufficient to identify \mathbf{U} ! On the contrary, consider that the spectra of the sources $S_{s_k}^{(0)}(\nu) = \sum_\ell R_{s_k}^{(0)}(\ell) e^{-i2\pi\ell\nu}$ are all different at least for a frequency ν . For any unitary matrix \mathbf{V} , the matrix $\mathbf{V}^H R_x^{(0)}(\ell) \mathbf{V}$ is diagonal for every indices ℓ if and only if the columns of \mathbf{V} equal these of \mathbf{U} up to a modulus one factor and a permutation. This remark was done in [19] and an algorithm (SOBI) was deduced based on a joint diagonalization technique [20].

The reader has noticed the suboptimality of the above method when the mixture is cyclo-stationary. The exploited statistics are only the $R_{\mathbf{x}}^{(0)}(\ell)$ for certain indices ℓ . In [21], it is suggested to take advantage of the cyclic-statistics of the mixture. In this respect, notice that for any α cyclic frequency of the mixture, we have

$$R_{\mathbf{x}}^{(\alpha)}(\ell) = \mathbf{U} R_{\mathbf{s}}^{(\alpha)}(\ell) \mathbf{U}^H$$

hence these “new” statistics could be added in the pencil of matrices to be jointly diagonalized. This theoretical appeal is attenuated by the fact that the non null-components of $R_{\mathbf{s}}^{(\alpha)}(\ell)$ are numerically inconsistent.

At this level, we should emphasize that the statistics $R_{\mathbf{x}}^{(\alpha)}(\ell)$ for any α (zero or not) are not accessible to the receiver and should be replaced by the empirical estimate denoted by $\hat{R}_{\mathbf{x}}^{(\alpha)}(\ell)$ and defined for $\ell \geq 0$ as

$$\hat{R}_{\mathbf{x}}^{(\alpha)}(\ell) = \frac{1}{N} \sum_{n=0}^{N-\ell-1} \mathbf{x}(n + \ell) \mathbf{x}(n)^H e^{-i2\pi\alpha n}, \quad (4.23)$$

where N is the number of snapshots. This estimate is a consistent estimate of the matrix $R_{\mathbf{x}}^{(\alpha)}(\ell)$. In an ideal scenario where the model $\mathbf{x}(n) = \mathbf{U}\mathbf{s}(n)$ holds true, it is remarkable that $\hat{R}_{\mathbf{x}}^{(\alpha)}(\ell) = \mathbf{U}\hat{R}_{\mathbf{s}}^{(\alpha)}(\ell)\mathbf{U}^H$ and the joint diagonalization of the estimated statistics should provide the exact mixing matrix: the algorithm is called *deterministic*. In a realistic context, however, the data are perturbed by an additive noise term: in this case, the above factorization does not hold true anymore and the joint diagonalization is an *approximate* joint diagonalization.

In practice, despite its attractivity, SOBI is seldom used to achieve BSS of digital communication data. Indeed, the condition that there are no two sources whose spectra are identical (up to a multiplicative constant) does not make sense most of the time. Indeed, the transmitted symbols are generally white sequences whose shaping functions are close from to one another. As the spectra are numerically similar, the joint diagonalization approach is bound to suffer from numerical problems.

2.04.3.4 Iterative BSS (stationary case)

As was specified, the stationary scenario assumes that for all indices $k : T_k = T$, i.e., all the baud-rates are equal, and $T_e = T$. Under these very specific circumstances, the model (4.14) involves a source vector $\mathbf{s}(n)$ whose components are **stationary** and **mutually independent**. We insist on the fact that the components of the source vector are not the i.i.d. symbol sequences but linear processes generated by these symbol sequences as indicated by Eq. (4.15). BSS aims at estimating the sources, not the symbol sequences. Hence, BSS may be seen as a preliminary step before the estimation of the symbols.

Contrary to other methods, no pre-processing of the data is necessary (PCA, pre-whitening).

In this section, we firstly design methods able to recover one of the sources (or a scaled version). In a second step, we present the so-called deflation that allows one to run the extraction of another source from a deflated mixture where the contribution of the first estimated source has been removed. The convergence is established: after K such steps, the K sources are expected to be estimated. Convergence properties are discussed.

2.04.3.4.1 Estimation of one source: theoretical considerations

Thanks to the mixing matrix \mathbf{H} having full-rank—see condition (4.19)—we know that, for any source index k , there exist column vectors \mathbf{g}_k such that

$$\mathbf{g}_k^T \mathbf{y}(n) = s_k(n).$$

Denoting $r_k(n) = \mathbf{g}_k^T \mathbf{y}(n)$, we may call this new signal as the reconstructed source since it involves only one source. This new signal is obtained after a so-called spatial filtering of the data. Of course, it is not possible to compute \mathbf{g}_k since \mathbf{H} is not accessible. A possible approach hence consists in adapting a spatial filter \mathbf{g} that makes

$$r(n) = \mathbf{g}^T \mathbf{y}(n) \quad (4.24)$$

resemble one of the sources. This will be done by considering particular statistics of the signal $r(n)$. We may write this signal under this form

$$r(n) = \sum_{k=1}^K f_k s_k(n), \quad (4.25)$$

where the taps f_k are the components of the vector

$$\mathbf{f} = \mathbf{g}^T \mathbf{H}. \quad (4.26)$$

The term $f_k s_k(n)$ in $r(n)$ represents the contribution of the k th source to the reconstructed signal $r(n)$. As may be easily understood, we aim at finding a “good” \mathbf{g} , i.e., such that \mathbf{f} is a vector having a single non-null component.

Definition 2. A vector \mathbf{f} is said to be separating if all its components are null except one.

Evidently, the signal $r(n)$ involves a single source if and only if the composite vector \mathbf{f} is separating.

We may inspect higher-order statistics and particularly the fourth-order ones. It has been proposed to consider the fourth-order cumulant (see Section 2.04.3.6 for theoretical justifications):

$$\kappa(r) = \text{cum}(r(n), r(n)^*, r(n), r(n)^*).$$

In this respect, we may introduce the following function, called *normalized (fourth-order) cumulant*:

$$\Upsilon(r) = \frac{\kappa(r)}{(E[|r(n)|^2])^2}. \quad (4.27)$$

Thanks to the definition of the cumulants, we have: $\kappa(r) = E[|r(n)|^4] - 2(E[|r(n)|^2])^2 - |E[r(n)^2]|^2$. Now, the circularity assumption of the symbol sequences (4.2) implies the circularity of the sources, hence

$$E[r(n)^2] = 0.$$

We re-express $\Upsilon(r)$ as a function of the moments of $r(n)$:

$$\Upsilon(r) = \frac{E[|r(n)|^4]}{(E[|r(n)|^2])^2} - 2. \quad (4.28)$$

We have the result:

Proposition 3. *As, by assumption, the Kurtosis of the sources, $\kappa(d_k)$, are strictly negative, the function $\Upsilon(r)$ achieves its minimum at a separating vector. Moreover, the separating vector in question has its single non-null element located at an index k_0 such that $\Upsilon(s_{k_0}) = \min_{k=1,\dots,K} \Upsilon(s_k)$.*

Proof. On the one hand, the mixture $r(n)$ is a linear mixture of independent random variables. The multi-linearity of the cumulants [6] gives:

$$\kappa(r) = \sum_{k=1}^K |f_k|^4 \kappa(s_k).$$

After noticing that $E[|r(n)|^2] = \sum_k |f_k|^2$ we arrive at the expansion:

$$\Upsilon(r) = \frac{\sum_{k=1}^K \kappa(s_k) |f_k|^4}{\left(\sum_{k=1}^K |f_k|^2\right)^2}. \quad (4.29)$$

On the other hand, $s_k(n)$ is a linear process generated by the i.i.d. symbol sequence $d_k(n)$: see Eq. (4.7). As was supposed (or noticed) in the Introduction, $\kappa(d_k) < 0$ hence

$$\kappa(s_k) < 0$$

for any index k . The sources are sometimes referred to as *platykurtic sources*. Denote by $\kappa_{\min} = \min(\kappa(s_1), \dots, \kappa(s_K))$. Thanks to the above result, $\kappa_{\min} < 0$. Hence $\kappa(r) \geq -|\kappa_{\min}| \sum_k |f_k|^4$. Besides, we recall that $(\sum_k |f_k|^2)^2 \geq \sum_k |f_k|^4$ with equality if and only if the coefficients $|f_k|$ are all null except one, i.e., if and only if the vector \mathbf{f} is separating.

We insist on the fact that the assumption that one of the $\kappa(s_k)$ is strictly negative is fundamental. Imagine on the contrary that, for all the indices k , $\kappa(d_k) > 0$. Then $\sum_{k=1}^K \kappa(s_k) |f_k|^4 \geq \min_k \kappa(s_k) \left(\sum_{k=1}^K |f_k|^4\right)$. As

$$\left(\frac{1}{K} \sum_{k=1}^K |f_k|^4\right) \geq \left(\frac{1}{K} \sum_{k=1}^K |f_k|^2\right)^2$$

with equality iff all the $|f_k|$ are equal, this implies that the argument minima of $\Upsilon(r)$ are not separating (on the contrary, the coefficients f_k equally weigh the sources). \square

As a remark, it is instructive, though superfluous in this paper since the digital communication symbols have negative Kurtosis, to address the optimization of $\Upsilon(r)$ for general distributions of the $\kappa(s_k)$: the reader may find the details in [5].

One may inspect the minimum minimorum of $\Upsilon(r)$ over all the possible constellations. The Jensen inequality (see [22] p. 80) gives: $E[(\varphi|r(n)|^2)] \geq \varphi(E[|r(n)|^2])$ for a convex mapping φ ; the equality is achieved when $|r(n)| = 1$. Taking $\varphi(x) = x^2$, we obtain

$$\Upsilon(r) \geq -1$$

and the equality is achieved when $r(n)$ has unit modulus. Of course, this can only happen if one of the sources has a modulus equal to one or, there exists an index k such that $\Upsilon(s_k) = -1$. This does not happen in general, but this remark shows that the minimization of $\Upsilon(r)$ tends to make $r(n)$ resemble as much as possible a constant modulus sequence. We inspect this point a bit further. A way to measure the distance of $r(n)$ to the modulus one is simply to consider

$$\Upsilon_{CM}(r) = E[(|r(n)|^2 - 1)^2]. \quad (4.30)$$

This function was originally considered for deconvolution problems [23, 24] and then for source separation problems ([25–27] for instance).

We may bridge the gap between $\Upsilon(r)$ and $\Upsilon_{CM}(r)$:

Proposition 4. Define $\Upsilon_{CM,\inf} = \inf_r \Upsilon_{CM}(r)$ and $\Upsilon_{\inf} = \inf_r \Upsilon(r)$. The minimization of $\Upsilon_{CM}(r)$ is linked to the minimization of $\Upsilon(r)$ in the sense that:

$$\Upsilon_{CM,\inf} = 1 - \frac{1}{2 + \Upsilon_{\inf}},$$

if \mathbf{f} achieves to minimize Υ then $\sqrt{\frac{1}{2 + \Upsilon_{\inf}}} \frac{1}{\|\mathbf{f}\|} \mathbf{f}$ is a minimizer of Υ_{CM} . Conversely, if \mathbf{f} achieves to minimize Υ_{CM} , then $\rho \mathbf{f}$ minimizes Υ for any $\rho \neq 0$.

Proof. For any \mathbf{f} , we have: $\Upsilon_{CM}(r) = E[|r(n)|^4] - 2E[|r(n)|^2] + 1$. Thanks to the expression of $\Upsilon(r)$ in (4.28), it is always true that: $E[|r(n)|^4] \geq (2 + \Upsilon_{\inf})(E[|r(n)|^2])^2$. We hence have:

$$\Upsilon_{CM}(r) \geq (2 + \Upsilon_{\inf})(E[|r(n)|^2])^2 - 2E[|r(n)|^2] + 1.$$

We set $\rho = E[|r(n)|^2]$. The second-order polynomial $\rho \mapsto (2 + \Upsilon_{\inf})\rho^2 - 2\rho + 1$ has minimal value $1 - \frac{1}{2 + \Upsilon_{\inf}}$ for $\rho = \frac{1}{2 + \Upsilon_{\inf}}$. We deduce the inequality: $\Upsilon_{CM}(r) \geq 1 - \frac{1}{2 + \Upsilon_{\inf}}$. If \mathbf{f} reaches the infimum of $\Upsilon(r)$ then, evidently, the choice $\sqrt{\frac{1}{2 + \Upsilon_{\inf}}} \frac{1}{\|\mathbf{f}\|} \mathbf{f}$ makes $\Upsilon_{CM}(r) = 1 - \frac{1}{2 + \Upsilon_{\inf}}$. Hence $1 - \frac{1}{2 + \Upsilon_{\inf}}$ is the minimum of $\Upsilon_{CM}(r)$. Conversely, for any $r(n)$, by definition, we have: $E[|r(n)|^4] - 2E[|r(n)|^2] + 1 \geq \Upsilon_{CM,\inf}$. In this inequality, substitute $\sqrt{\rho}r(n)$ for any positive ρ . We have:

$$\rho^2 E[|r(n)|^4] - 2\rho E[|r(n)|^2] + 1 \geq \Upsilon_{CM,\inf}.$$

This is in particular true for $\rho = \frac{E[|r(n)|^2]}{E[|r(n)|^4]}$ hence showing that $-\frac{(E[|r(n)|^2])^2}{E[|r(n)|^4]} + 1 \geq \Upsilon_{CM,\inf}$ or

$$\Upsilon(r) \geq 2 + \frac{1}{1 - \Upsilon_{CM,\inf}}.$$

The case of equality in the latter equation occurs when \mathbf{f} is any non-null scaled version of a minimizer of Υ_{CM} \square

As is explained in the next section, the search of a global minimum of $\Upsilon(r)$ or $\Upsilon_{CM}(r)$ is done according to a gradient method. It is well known that such an algorithm may be stuck in a local minimum of the function to be minimized. We have the result (see [5], Lemma 1):

Lemma 5. *Fix K real constants $\beta_k < 0$. The local minima of the function $\mathbf{f} \mapsto \sum_{k=1}^K \beta_k |f_k|^4$ over the unit sphere $\sum_{k=1}^K |f_k|^2 = 1$ are the separating vectors (of unit norm).*

Thanks to the expansion (4.29), and the fact that for all the sources $\beta_k = \kappa(s_k) < 0$, the local minima of $\Upsilon(r)$ over the unit sphere are the separating vectors (of unit norm). After simple topological considerations, it can even be deduced that:

Proposition 6. *Any local minimum of Υ is separating.*

As far as the function Υ_{CM} is concerned, we have:

Proposition 7. *Any local minimum of Υ_{CM} is separating.*

Proof. We consider the arguments given in [28]. The idea consists in writing \mathbf{f} in its polar form $\mathbf{f} = \rho \tilde{\mathbf{f}}$ where $\rho = \|\mathbf{f}\|$ and $\|\tilde{\mathbf{f}}\| = 1$. After setting $\tilde{r}(n) = \frac{1}{\|\mathbf{f}\|} r(n)$ the normalized version of $r(n)$, we have: $\Upsilon_{CM}(r) = \rho^4 E[|\tilde{r}(n)|^4] - 2\rho^2 E[|\tilde{r}(n)|^2] + 1$. Write this function $\psi(\rho, \tilde{\mathbf{f}})$. Necessarily, for a stationary point of Υ_{CM} the derivative of ψ w.r.t. ρ is zero. This gives: $\rho = 0$ or $\rho(\tilde{\mathbf{f}}) = \frac{E[|\tilde{r}(n)|^2]}{E[|\tilde{r}(n)|^4]}$.

The case $\rho = 0$ can be shown to correspond to a local maximum [26]. This says that a local minimum of Υ_{CM} is a local minimum of $\psi(\rho(\tilde{\mathbf{f}}), \tilde{\mathbf{f}})$. Now, this latter function is:

$$\psi(\rho(\tilde{\mathbf{f}}), \tilde{\mathbf{f}}) = 1 - \frac{1}{\Upsilon(\tilde{r})}.$$

We deduce that such a local minimum is also a local minimum of $\Upsilon(r)$ on the unit sphere. Thanks to Lemma 5 we deduce that the local minimum in question is separating. \square

2.04.3.4.2 Estimation of one source: practical aspects

Basic algorithms: Two problems arise when one focuses on the implementation of the results presented so forth: the first one concerns the estimation of the cost functions $\Upsilon(r)$ or $\Upsilon_{CM}(r)$, the second one is to choose a method able to find the argument minima of these estimated functions.

The two functions we have considered involve second and fourth-order moments of the signal $r(n)$. As the number of available data is finite—say, we observe $\mathbf{y}(n)$ for $n = 0, \dots, N-1$ —it is not possible to compute any of the moments of $r(n)$. However, a version of the law of large numbers allows one to consider estimates of the moments:

Lemma 8. *For $p = 2, 4$, we have, with probability one:*

$$\frac{1}{N} \sum_{n=0}^{N-1} |r(n)|^p \longrightarrow E[|r(n)|^p].$$

We are in position to estimate both functions $\Upsilon(r)$ and $\Upsilon_{CM}(r)$ respectively by

$$\hat{\Upsilon}(r) = \frac{\frac{1}{N} \sum_{n=0}^{N-1} |r(n)|^4}{\left(\frac{1}{N} \sum_{n=0}^{N-1} |r(n)|^2\right)^2} - 2, \quad (4.31)$$

$$\hat{\Upsilon}_{CM}(r) = \frac{1}{N} \sum_{n=0}^{N-1} (|r(n)|^2 - 1)^2. \quad (4.32)$$

Indeed, we have the result:

Proposition 9. $\hat{\Upsilon}_{CM}(r) \rightarrow \Upsilon_{CM}(r)$ and $\hat{\Upsilon}(r) \rightarrow \Upsilon(r)$ with probability one.

The functions $\hat{\Upsilon}_{CM}(r)$ and $\hat{\Upsilon}(r)$ to be minimized are non-convex, and the associated machinery cannot be considered. The functions, however, are regular w.r.t. the parameter \mathbf{g} . Hence, we choose to seek the argument minima by means of a gradient method. For instance, consider the minimization of Υ_{CM} . The notation for the gradient of $\hat{\Upsilon}_{CM}$ calculated at the point \mathbf{g} being $\nabla \hat{\Upsilon}_{CM}(r)$, the gradient algorithm, for a fixed $\mu > 0$, can be written as:

1. choose an initial vector $\mathbf{g}^{(0)}$ and compute $r^{(0)}(n) = \mathbf{g}^{(0)T} \mathbf{y}(n)$ for all the available data;
2. at the m th step: compute $\mathbf{g}^{(m)} = \mathbf{g}^{(m-1)} + \mu \nabla \hat{\Upsilon}_{CM}(r^{(m-1)})$ and the associated updated signal $r^{(m)}(n) = \mathbf{g}^{(m)T} \mathbf{y}(n)$;
3. redo the above step until the convergence is reached.

The same algorithm could be written for the minimization of $\hat{\Upsilon}$. However, the fact this latter function is homogeneous may involve numerical problems (the vector \mathbf{g} is not bounded). This is why the projected gradient algorithm is preferred: it consists in normalizing at each iteration of the algorithm the updated signal $r^{(m)}(n)$, i.e., projecting the current parameter $\mathbf{g}^{(m)}$ on the set

$$\left\{ \mathbf{g} \mid \mathbf{g}^T \left(\frac{1}{N} \sum_{n=0}^{N-1} E[\mathbf{y}(n)\mathbf{y}(n)^H] \right) \mathbf{g} = 1 \right\}.$$

Whatever the considered cost function, the parameter μ controls the performance. The next section faces the problem of choosing μ .

Refinement: choosing a locally optimal μ . For simplicity, the minimization of Υ_{CM} is addressed. The same idea may be considered for the minimization of Υ by means of the projected gradient. In order to boost the speed of convergence, it has been proposed to change μ at each step of the algorithm: the parameter μ is chosen such that the value of the function evaluated at the point $\mathbf{g}^{(m)}$ is minimum. It is easily seen that the function $\mu \mapsto \hat{\Upsilon}_{CM}(r^{(m)})$ is a polynomial of degree four. The minimum is hence easily (numerically) computed.

Robustness of the algorithms to the presence of local minima: It is well-known that such a gradient algorithm may be trapped in a local minimum: this, in general, is a clear limitation to the use of such an algorithm. Of course, it is not possible to say much on the local minima of the estimated functions $\hat{\Upsilon}(r)$ and $\hat{\Upsilon}_{CM}(r)$. However, Propositions 6 and 7 indicate that, asymptotically, if the algorithms are trapped in a local minimum, this does not impact the performance since this local minimum is precisely separating. This remark certainly explains why the algorithms show very good performance.

2.04.3.4.3 The deflation step

The algorithms depicted above provide a way to retrieve one of the sources. Of course, we aim at estimating all the sources. An idea hence consists in running again the previous algorithm. However, it is not possible to guarantee that the second extracted source is not the first extracted one. In the literature, three methods have been presented that overcome this major problem.

In the first one [29] it is proposed to penalize the cost function $\Upsilon(r)$ or $\Upsilon_{CM}(r)$ by adding to them a positive term that gives a measure of decorrelation between the current signal $r(n)$ and the previously extracted source. It is simple to show that, indeed, the global minimum is achieved if and only if the $r(n)$ is an other source. However, this approach has been noticed to show poor performance. The reason is that the extended cost function, contrary to the original, has many local minima that do not correspond to separating solutions. The algorithms is known to be trapped in such local minima and, in this case, the provided solution is not an estimate of one of the remaining sources.

The second one is algebraic: the idea is to estimate the subspace associated with the first estimated source and to run the minimization of $\Upsilon(r)$ or $\Upsilon_{CM}(r)$ on the orthogonal complement of the subspace in question: see [5].

The third is the most popular for source separation [30,31]: it consists in deflating the mixture by subtracting an estimation of the contribution of the extracted source and then to redo the minimization of $\Upsilon(r)$ or $\Upsilon_{CM}(r)$. Ideally, the “new” mixture should not involve the source that has been extracted and the minimization hence allows one to estimate another source. We provide some details.

Thanks to the previous results, we may suppose that we have $r^{(1)}(n) \approx \alpha s_1(n)$ where α is an unknown scaling. We have arbitrarily considered that the extracted source was the one numbered “1”: this has of course no impact on the generality. The contribution of the first source in the mixture $\mathbf{y}(n)$ has the form $\mathbf{h}_1 s_1(n)$ where \mathbf{h}_1 is the first column of the mixing matrix. We adopt a least square approach: the contribution of the first source is estimated as $\hat{\mathbf{h}}^{(1)}$ where this vector is defined as the minimizer of

$$\mathbf{h} \mapsto \sum_{n=0}^{N-1} \|\mathbf{y}(n) - \mathbf{h} r^{(1)}(n)\|^2.$$

Then the “deflated mixture” to be considered is

$$\mathbf{y}^{(2)}(n) = \mathbf{y}(n) - \hat{\mathbf{h}}^{(1)} r^{(1)}(n).$$

Ideally, the deflated mixture should not involve the first source. Hence running the Constant Modulus algorithm on this mixture should provide an estimate $r^{(2)}(n)$ of another source—say $s_2(n)$. The deflation is done again: this time $\hat{\mathbf{h}}^{(2)} r^{(2)}(n)$ is an estimation of the contribution of the second source. The deflated mixture is

$$\mathbf{y}^{(3)}(n) = \mathbf{y}^{(2)}(n) - \hat{\mathbf{h}}^{(2)} r^{(2)}(n).$$

And so forth until all the source are estimated. Notice that, asymptotically (when $N \rightarrow \infty$) the deflation procedure is convergent: in K steps the K sources are estimated.

2.04.3.4.4 Improving the deflation

Though its inherent advantages (simplicity, convergence of the algorithm of extraction), the above approach is supposed to suffer from the K deflation steps. Indeed, the deflation is expected to increase,

step after step, the noise level, impinging dramatically the extraction of the “last” source. This aspect has already been addressed and partially got round: we shortly address the re-initialization procedure introduced in [32].

Consider the extraction of the “second” source and apply the deflation technique. The source extraction algorithm is run on the deflated mixture and is likely to provide a spatial filter $\mathbf{g}^{(2)}$. We have, up to a scaling factor:

$$\mathbf{g}^{(2)T} \left(\underbrace{\mathbf{y}(n) - \hat{\mathbf{h}}^{(1)} \mathbf{g}^{(1)T} \mathbf{y}(n)}_{\text{deflated mixture}} \right) \approx s_2(n),$$

which provides the approximation:

$$\mathbf{g}^T \mathbf{y}(n) \approx s_2(n),$$

where $\mathbf{g}^T = \mathbf{g}^{(2)T} (\mathbf{I}_M - \hat{\mathbf{h}}^{(1)} \mathbf{g}^{(1)T})$. We hence have computed a spatial filter \mathbf{g} that is close to a separating filter w.r.t. the initial mixture. The idea is hence the following: run the algorithm of minimization on the initial mixture, taking \mathbf{g} as an initial point. As \mathbf{g} is close to a filter that is a local minimum of the function to minimize (see Propositions 6, 7), the computed spatial filter hence obtained after convergence is likely to separate s_2 from the initial mixture. This procedure can be iterated: at each step, the separation is processed on the initial $\mathbf{y}(n)$ and not on a deflated mixtures. Though simple, this procedure considerably enhances the performance.

2.04.3.4.5 Extensions

Many contributions in BSS consider such functions as

$$\varphi(\kappa(r)) \quad (4.33)$$

over the unit sphere $E[|r(n)|^2] = 1$ where $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is any continuous function on \mathbb{R} such that $\varphi(0) = 0$ and φ is strictly monotone over \mathbb{R}^+ and \mathbb{R}^- . The most common choices are $\varphi_1(x) = x^2$ and $\varphi_2(x) = |x|$ and $\varphi_3(x) = -x$. It quite simple to prove the following result:

Proposition 10. *Provided that $\varphi(\kappa(r_k)) > 0$ for at least one source, then the local maxima of (4.33) are separating.*

2.04.3.5 Global BSS (stationary mixture)

In this section, we present global methods, i.e., methods that “invert” the system in one shot. Assuming that $K = M$, it is possible to linearly transform the data such as in Section 2.04.3.3. The “new” data can be written as

$$\mathbf{x}(n) = \mathbf{U} \mathbf{s}(n),$$

where \mathbf{U} is a unitary $K \times K$ matrix. A global BSS method hence aims at determining a *unitary* matrix \mathbf{V} such that the components of

$$\mathbf{r}(n) = \mathbf{V} \mathbf{x}(n)$$

correspond to the sources up to modulus one scalings and a permutation. In this respect, we suggest to take profit of certain results of Section 2.04.3.4.

2.04.3.5.1 First result

Denote by $r_k(n)$ the k th component of $\mathbf{r}(n)$. Due to the pre-whitening, we have: $E[|r_k|^2] = 1$, hence $\Upsilon(r_k) = \kappa(r_k)$. This later can be seen as a function of the k th row of the matrix \mathbf{V} . We have shown that $\Upsilon(r_k)$ is minimum if r_k corresponds to one of the sources up to a modulus one scaling, i.e., if the k th row of $\mathbf{F} = \mathbf{V}\mathbf{U}$ is separating, its non-zero component being located at an index corresponding to the sources that have the smallest Kurtosis. The idea is hence to form the function

$$\Psi(\mathbf{r}) = \sum_{k=1}^K \kappa(r_k). \quad (4.34)$$

Obviously, we have $\Psi(\mathbf{r}) \geq K \min_{k=1,\dots,K} \kappa(s_k)$. Conversely, this lower bound cannot be achieved in general: assume for instance that $\min_{k=1,\dots,K} \kappa(s_k)$ is reached *once*: say, the first source. The above lower bound is achieved only if \mathbf{F} is the matrix having non-zero components on the first column only. This of course violates the constraint that \mathbf{F} is unitary. We have the following tight result:

Proposition 11. *As $\kappa(s_k) < 0$ for all the sources, we have, for any unitary \mathbf{F} :*

$$\Psi(\mathbf{r}) \geq \Psi(\mathbf{s}) = \sum_{k=1}^K \kappa(s_k). \quad (4.35)$$

Moreover, the inequality is an equality if and only if $\mathbf{F} = \mathbf{V}\mathbf{U}$ essentially equals the identity matrix.

Proof. By “essentially equal to,” we mean that \mathbf{F} is a diagonal matrix with modulus entries, whose columns are permuted. The proof of this result follows the proof of Proposition 3 and was suggested by Common in [2]. We indeed, express $\Psi(\mathbf{r})$ as

$$\Psi(\mathbf{r}) = \sum_{i=1}^K \sum_{k=1}^K \kappa(s_k) |f_{i,k}|^4 = - \sum_{i=1}^K \sum_{k=1}^K (|\kappa(s_k)|^{1/4} |f_{i,k}|)^4.$$

On the other hand, we have the inequality:

$$\sum_{k=1}^K \left(|\kappa(s_k)|^{1/4} |f_{i,k}| \right)^4 \leq \left(\sum_{k=1}^K |\kappa(s_k)|^{1/2} |f_{i,k}|^2 \right)^2$$

implying that

$$-\Psi(\mathbf{r}) \leq \sum_{i=1}^K \left(\sum_{k=1}^K |\kappa(s_k)|^{1/2} |f_{i,k}|^2 \right)^2. \quad (4.36)$$

We now denote by \mathbb{F} the matrix whose component (i, k) is $|f_{i,k}|^2$ and by $\mathbf{b}^{(1/2)} = (|\kappa(s_1)|^{1/2}, \dots, |\kappa(s_K)|^{1/2})^T$ so that the r.h.s. of (4.36) is simply $\|\mathbb{F} \mathbf{b}^{(1/2)}\|_2^2$. As $\mathbf{F} \mathbf{F}^H = \mathbf{I}_K$ we deduce that the sum of the elements of any row/column of \mathbb{F} is equal to one: \mathbb{F} is called doubly stochastic. As a consequence of the Birkhoff theorem (see [33] Chapter 2), \mathbb{F} can be seen as the convex sum of

permutation matrices, i.e., $\mathbf{F} = \sum_{j=1}^K \lambda_j \mathbf{P}_j$ where, for any index j , $\lambda_j \geq 0$ and $\sum_{j=1}^K \lambda_j = 1$. We deduce that

$$\begin{aligned}\|\mathbf{F} \mathbf{b}^{(1/2)}\|_2^2 &\leq \left(\sum_{j=1}^K \lambda_j \|\mathbf{P}_j \mathbf{b}^{(1/2)}\|_2 \right)^2 = \left(\sum_{j=1}^K \lambda_j \|\mathbf{b}^{(1/2)}\|_2 \right)^2 \\ &= \sum_{j=1}^K |\kappa(s_k)|.\end{aligned}$$

This proves Eq. (4.35). Let us inspect the case of equality. If equality occurs, the inequality (4.36) is necessarily an equality. Hence we have, for any indices i and $k_1 \neq k_2$

$$\kappa_4(s_{k_1})\kappa_4(s_{k_2}) f_{i,k_1} f_{i,k_2} = 0.$$

If all the numbers $\kappa(s_k)$ are strictly negative, hence non-null, the above conditions implies that, for any index i , the vector $(f_{i,1}, \dots, f_{i,K})$, i.e., the i th row of the matrix \mathbf{F} , has at most one non zero component. On the other hand, the matrix \mathbf{F} is unitary. This imposes that \mathbf{F} essentially equals the identity matrix. \square

In practice, the function $\Psi(\mathbf{r})$ given by Eq. (4.34) cannot be computed. As the data are supposed to be complex-circular at the second-order, we have

$$\kappa(r_k) = E[|r_k(n)|^4] - 2.$$

A consistent estimate of $\Psi(\mathbf{r})$ is hence

$$\hat{\Psi}(\mathbf{r}) = \sum_{k=1}^K \left(\frac{1}{N} \sum_{n=1}^N |r_k(n)|^4 - 2 \right). \quad (4.37)$$

The minimization of *this* function is carried out over the space of unitary matrices. This can be done by using a Jacobi-like algorithm: unitary matrices are parametrized by means of the Given angles. The reader may find details in [2]. Notice that no results concerning the convergence of the algorithm can be said, since it has not been shown that the local minima of the “true” function $\Psi(\mathbf{r})$ are “good ones”, i.e., achieve the BSS.

2.04.3.5.2 Generalization: notion of contrast function

We introduce the function

$$\Psi_\varphi(\mathbf{r}) = \sum_{k=1}^K \varphi(\kappa(r_k)), \quad (4.38)$$

where $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a function to be specified. Rather than minimizing Ψ_φ , we address its maximization. If the maximum is attained when (and only when) BSS is achieved, Ψ_φ is called a contrast function. As seen previously, the choice $\varphi = -\text{Id}$ makes Ψ_φ be a contrast function. We have, more generally:

Proposition 12. If φ is a convex function on \mathbb{R} such that $\varphi(0) = 0$ and if, for any source index k the condition $\varphi(\kappa(s_k)) > 0$ is fulfilled, then Ψ_φ is a contrast function.

Proof. For any index i , we have: $\sum_k |f_{i,k}|^4 \leq (\sum_k |f_{i,k}|^2)^2$. As \mathbf{F} is unitary, we deduce that $\sum_k |f_{i,k}|^2 = 1$. Hence $\sum_k |f_{i,k}|^4 \leq 1$. We set $\rho = 1 - \sum_k |f_{i,k}|^4$. The Jensen inequality gives here:

$$\varphi\left(\sum_k \kappa(s_k)|f_{i,k}|^4 + 0(1-\rho)\right) \leq \sum_k \varphi(\kappa(s_k))|f_{i,k}|^4 + \varphi(0)(1-\rho).$$

As $\varphi(0) = 1$, we deduce that

$$\Psi_\varphi(\mathbf{r}) \leq \sum_i \sum_k \varphi(\kappa(s_k))|f_{i,k}|^4.$$

By assumption all the $\varphi(\kappa(s_k)) > 0$ hence the same argument as the one given in the proof of Proposition 11 shows that the maximum is $\sum_k \varphi(\kappa(s_k))$ and that this maximum is reached if and only if the matrix \mathbf{F} is essentially the identity. \square

The standard choice for φ is $\varphi(x) = x^2$. In this case, it is enlightening to notice that maximizing Ψ_φ is equivalent to minimizing:

$$\sum_{i,j,k,\ell \text{ non all equal}} (\text{cum}(r_i, r_j^*, r_k, r_\ell^*))^2 \quad (4.39)$$

which is clearly a measure of independence (up to the fourth-order).

2.04.3.5.3 A popular algorithm: JADE

If the indices j, k are fixed, it may be noticed that the matrix $\mathbf{M}_{j,k}$ whose entry i, ℓ is given by $\text{cum}(x_i, x_j^*, x_k, x_\ell^*)$ admits the factorization

$$\mathbf{M}_{j,k} = \mathbf{U}\mathbf{D}_{j,k}\mathbf{U}^*,$$

where $\mathbf{D}_{j,k}$ is diagonal; the entry (ℓ, ℓ) is $u_{j,\ell}^* u_{k,\ell} \kappa(s_\ell)$. Otherwise stated, \mathbf{U} diagonalizes the normal matrix $\mathbf{M}_{j,k}$ whatever the indices j, k . Introducing the off(.) operator that sums all the entries of a matrix except the ones located on the main diagonal, this says that

$$\sum_{j,k} \text{off}(\mathbf{V}\mathbf{M}_{j,k}\mathbf{V}^*)^2 \quad (4.40)$$

is minimum, equal to zero, when $\mathbf{F} = \mathbf{V}\mathbf{U}$ essentially equals the identity matrix. Conversely, it can be shown that (4.40) can be written as (4.39). Hence a minimizer of the function given in (4.40) is a maximizer of $\Psi_\varphi(\mathbf{r})$ with $\varphi(x) = x^2$. Now, Proposition 12 with the fact that, for any index k , $(\kappa(s_k))^2 > 0$ proves that the maximizers of $\Psi_\varphi(\mathbf{r})$ are such that \mathbf{F} is essentially equal to the identity matrix. This trick of algebra allows one to achieve the maximization of Ψ_φ thanks to a joint diagonalization of a set of normal matrices. It has been proposed by Cardoso [4]. The algorithm associated with this approach is called JADE. It is very popular since efficient algorithms of joint diagonalization are known [20].

Notice that the pencil of matrices is a pencil of normal matrices hence the matrix \mathbf{U} is searched in the set of unitary matrices. Recently, Yeredor et al. [34] suggest to relax the unitary constraint. These authors even suggest to skip the pre-whitening of the data: they argue that the pre-whitening may limit the attainable performance as Cardoso pointed it out in [35]. With no whitening of the data, the matrices of cumulants are still jointly diagonalized; if the channel matrix \mathbf{H} is square, the columns of this latter form a basis for the diagonalization. The converse is not clear on the one hand and, on the other hand, the case of tall matrices \mathbf{H} remains to be addressed.

2.04.3.6 Generalizations

We have presented *ad hoc* BSS methods, whose theoretical foundations are solid and whose good performance is well-known. In the literature, however, many other methods can be found. They stem from considerations of information theory. We provide some key ideas and related bibliographical references.

After pre-whitening, we recall that the received data is $\mathbf{x} = \mathbf{U} \mathbf{s}$ (we have dropped the time index): on the one hand, \mathbf{U} is unitary and on the other hand, the component of the random variables in \mathbf{s} are mutually independent. The idea of independent Component Analysis (ICA) is hence to exhibit matrices \mathbf{V} such that the components of $\mathbf{r} = \mathbf{V} \mathbf{x}$ are “as much independent as possible”. This independency may be measured by the Kullback-Leibler divergence between the distribution of \mathbf{r} and this of the product of the marginals: this is the mutual information $\mathcal{I}(\mathbf{r})$. It is well-known that $\mathcal{I}(\mathbf{r}) \geq 0$ with equality iff the components of \mathbf{r} are independent.

Besides, it has been underlined by Comon in [2] that the Darmois theorem states the fact: as the sources are non-Gaussian, the fact that the components of \mathbf{r} are pair-wise independent (which is naturally the case if they are mutually independent) implies that $\mathbf{F} = \mathbf{V} \mathbf{U}$ is essentially equal to the identity matrix. Hence, the minimization of the mutual information of $\mathcal{I}(\mathbf{r})$ is legitimate. This induces of course no BSS algorithm, since the mutual information cannot be simply estimated.

Now, Comon underlines in the seminal paper [2] that $\mathcal{I}(\mathbf{r}) = \mathcal{J}(\mathbf{r}) - \sum_{k=1}^K \mathcal{J}(r_k)$ where $\mathcal{J}(\mathbf{r})$ is the negentropy of the vector \mathbf{r} , i.e., $\mathcal{J}(\mathbf{r}) = \mathcal{H}(\mathbf{r}_{\text{gaussian}}) - \mathcal{H}(\mathbf{r})$: here, $\mathcal{H}(\mathbf{r})$ is the differential entropy of \mathbf{r} and $\mathcal{H}(\mathbf{r}_{\text{gaussian}})$ the differential entropy of the Gaussian vector whose mean and covariance matrix are those of \mathbf{r} . The negentropy shows the nice property of invariance w.r.t. any invertible change of variables: hence $\mathcal{J}(\mathbf{r})$ does not depend on \mathbf{F} . The independence is then obtained by maximizing

$$\sum_{k=1}^K \mathcal{J}(r_k). \quad (4.41)$$

Notice that $\mathcal{J}(r_k) \geq 0$ with equality when r_k is Gaussian. Hence the maximization in question tends to maximize the distance of the reconstructed source \mathbf{r} to the Gaussian case.

On the other hand, this shows that, in order to achieve independency, it suffices to consider the maximization of a sum of functions, each of which simply depends on a component of the reconstructed source vector. Evidently, the maximization of the function $\mathcal{J}(r_1)$ under the constraint that the first row of \mathbf{F} has norm one, is achieved when r_1 coincides with one of the sources up to a modulus one factor. In this respect, this remark provides a justification of the iterative methods proposed in Section 2.04.3.4, even if the function considered are not the negentropy. As far as the function $\Psi_\phi(\mathbf{r})$ is concerned, notice that it has the form (4.41).

Comon proposes an approximation of the negentropy, based on the Edgeworth expansion of the probability density functions of the random variables r_k . Thanks to the circularity assumption, this approximation is $\mathcal{J}(r_k) \approx \frac{1}{48}\kappa(r_k)^2$. This calls for an important remark: the function $\Psi_\phi(\mathbf{r})$ given in Eq. (4.38) with $\phi(x) = x^2$ is hence closely connected to a measure of independence—this confirms a remark previously done.

Hyvärinen departs from functions based on the cumulants. In [36], it is suggested to consider a wider class of functions which are not directly related to cumulants. In short, the new functions to be maximized are more robust to outliers. The price to be paid is the weaker results concerning the separation: for instance, the precious results concerning the separability of the local maxima do not hold. An efficient algorithm has given rise to the popular method called FastICA. In the original paper [36], the sources are real valued which is not the case in this paper. An extension to complex-valued sources can be found in [3].

2.04.3.7 Iterative BSS (general cyclo-stationary case)

We specified that the assumption of stationarity of the sources, as required in the previous sections, is somewhat restrictive in a realistic scenario of telecommunication. Indeed, the stationarity implicitly assumes that all the sources have the same symbol period and that the data are sampled at a period equal to the symbol period. In general—think for instance of a passive listening context—the sources have different baud-rates. We denote the symbol periods of the K sources by T_1, \dots, T_K . If T_e is the sampling period—a priori different of any of the symbol periods—we have deduce from Section 2.04.2 that, for any index k , the source $s_k(n)$ is cyclostationary. In particular, the second moment $E[|s_k(n)|^2]$ varies with the time-index n . More specifically, we have

$$E[|s_k(n)|^2] = \sum_{\alpha \in \mathcal{I}_k} R_{s_k}^{(\alpha)} e^{i2\pi\alpha_k n},$$

where

$$\mathcal{I}_k = \left\{ 0, \pm \frac{T_e}{T_k} \right\}$$

is the set of the second-order cyclic frequencies of s_k and the Fourier coefficients $R_{s_k}^{(\alpha)}$ are given by Eq. (4.8) (we have dropped the time-lag in $R_{s_k}^{(\alpha)}(\ell)$ since, in the sequel, no other time delay than $\ell = 0$ is considered). In this section, the channel is supposed to be memoryless, as in all the Section 2.04.3. Hence, the model given by Eq. (4.11) still holds: the components of the source vector $\mathbf{s}(n)$ are effectively mutually independent, but are not individually stationary. As far as the normalization of the sources is concerned, it may also be assumed: in the cyclo-stationary context of this section, this means that

$$\langle E[\mathbf{s}(n)\mathbf{s}(n)^H] \rangle_n = \mathbf{I}_K. \quad (4.42)$$

In the following, we provide assumptions on the sources that guarantee that the algorithms of source separation depicted in Section 2.04.3.4 still converge to desirable solutions, i.e., allow one to separate the sources. In other words, the algorithms previously considered are run as if the data were stationary: this

means that nothing has to be changed in any part of the stationary BSS algorithms. Surprisingly, the fact that the data are not stationary is *shown* not to impact the convergence to a good (separating) solution.

The algorithms encountered for stationary data (see Section 2.04.3.4) are designed to minimize either the function $\hat{\Upsilon}(r)$ given by Eq. (4.31) or the Godard function $\hat{\Upsilon}_{CM}(r)$ given by Eq. (4.32). Let us recall that the signal $r(n)$ is the output of the variable spatial filter \mathbf{g} , i.e.,

$$\begin{aligned} r(n) &= \mathbf{g}^T \mathbf{y}(n), \\ &= \mathbf{f}^T \mathbf{s}(n), \end{aligned}$$

where $\mathbf{f}^T = \mathbf{g}^T \mathbf{H}$. As in the stationary case, we analyze the argument minima of the theoretical associated function. The first point to be addressed is hence: to which functions $\hat{\Upsilon}(r)$ and $\hat{\Upsilon}_{CM}(r)$ converge? Due to the non-stationarity of the model, the function $\hat{\Upsilon}(r)$ for instance does not converge to $\Upsilon(r)$ as given in Eq. (4.30): this cannot be the case since the latter function depends on the time-lag n .

There is a version of the law of the large numbers for non-stationary data. Lemma 8 can be written in this context under this form:

Lemma 13. *For $p = 2, 4$, we have, with probability one:*

$$\frac{1}{N} \sum_{n=0}^{N-1} |r(n)|^p \longrightarrow \langle E[|r(n)|^p] \rangle.$$

We deduce that $\hat{\Upsilon}_{CM}(r) \longrightarrow \left\langle E \left[(|r(n)|^2 - 1)^2 \right] \right\rangle$. We define this limit as (the superscript “c” means “cyclo-stationary”):

$$\Upsilon_{CM}^{(c)}(r) = \left\langle E \left[(|r(n)|^2 - 1)^2 \right] \right\rangle_n. \quad (4.43)$$

For the same reason, we have $\hat{\Upsilon}(r) \longrightarrow \Upsilon^{(c)}(r)$ where

$$\Upsilon^{(c)}(r) = \frac{\langle E [|r(n)|^4] \rangle_n}{(\langle E [|r(n)|^2] \rangle_n)^2} - 2. \quad (4.44)$$

Notice that, in the case of stationary data, $\Upsilon^{(c)}(r)$ (respectively $\Upsilon_{CM}^{(c)}(r)$) equals $\Upsilon(r)$ (respectively $\Upsilon_{CM}(r)$).

We first address the minimization of $\Upsilon^{(c)}(r)$. Once done, we deduce results concerning the minimization of $\Upsilon_{CM}^{(c)}(r)$.

We consider a prior expansion of the moments involved in $\Upsilon^{(c)}(r)$. The following equality always holds true (we recall that the signals are all complex-circular at the second-order):

$$E[|r(n)|^4] = \kappa(r(n)) + 2(E[|r(n)|^2])^2. \quad (4.45)$$

The multi-linearity of the cumulant gives:

$$\langle \kappa(r(n)) \rangle_n = \sum_{k=1}^K \eta(s_k) |f_k|^4,$$

where we let for any source $s = s_k$ the number $\eta(s)$ be

$$\eta(s) = \langle \kappa(s(n)) \rangle_n. \quad (4.46)$$

The output of the spatial filter, $r(n)$, is obviously a linear combination of the sources s_1, \dots, s_K : $r(n) = \sum_{k=1}^K f_k s_k(n)$. As s_k admits $\mathcal{I}_k = \{0, \pm\alpha_k\}$ as the set of its second-order cyclic frequencies, we deduce that $r(n)$ is cyclo-stationary and its second-order cyclo-frequencies are in the set

$$\mathcal{I} = \bigcup_{k=1}^K \mathcal{I}_k.$$

This means that $E[|r(n)|^2] = \sum_{\alpha \in \mathcal{I}} R_r^{(\alpha)} e^{j2\pi\alpha n}$. As a consequence, the term $\left\langle (E [|r(n)|^2])^2 \right\rangle_n$, may be computed thanks to the Parseval equality. We have indeed

$$\left\langle |E [|r(n)|^2]|^2 \right\rangle_n = \sum_{\alpha \in \mathcal{I}} |R_r^{(\alpha)}|^2.$$

On the other hand, the sources are mutually decorrelated hence

$$\begin{aligned} R_r^{(\alpha)} &= \sum_{k=1}^K R_{f_k s_k}^{(\alpha)} \\ &= \sum_{k=1}^K |f_k|^2 R_{s_k}^{(\alpha)}. \end{aligned}$$

By definition, $\mathcal{I}^* = \mathcal{I} \setminus \{0\}$ is the set of all the non-null cyclic frequencies of the mixture. After isolating the term associated with the cyclic frequency $\alpha = 0$, we may hence write that

$$\begin{aligned} \left\langle (E [|r(n)|^2])^2 \right\rangle_n &= (R_r^{(0)})^2 + \sum_{\alpha \in \mathcal{I}^*} \left| \sum_{k=1}^K |f_k|^2 R_{s_k}^{(\alpha)} \right|^2 \\ &= (R_r^{(0)})^2 + \sum_{k=1}^K |f_k|^4 \left(\sum_{\alpha \in \mathcal{I}_k^*} |R_{s_k}^{(\alpha)}|^2 \right) + \frac{1}{2} \sum_{k_1 \neq k_2} |f_{k_1}|^2 |f_{k_2}|^2 \varepsilon(s_{k_1}, s_{k_2}) \\ &= (R_r^{(0)})^2 + 2 \sum_{k=1}^K |f_k|^4 |R_{s_k}^{(\alpha_k)}|^2 + \frac{1}{2} \sum_{k_1 \neq k_2} |f_{k_1}|^2 |f_{k_2}|^2 \varepsilon(s_{k_1}, s_{k_2}), \end{aligned}$$

where

$$\varepsilon(s_{k_1}, s_{k_2}) = 2 \sum_{\alpha \in \mathcal{I}^*} R_{s_{k_1}}^{(\alpha)} R_{s_{k_2}}^{(\alpha)*}. \quad (4.47)$$

We have used the symmetry:

$$R_{s_k}^{(-\alpha_k)} = R_{s_k}^{(\alpha_k)*}. \quad (4.48)$$

The set \mathcal{I}^* is simply $\{\pm\alpha_k \mid k = 1, \dots, K\}$. Denoting by \mathcal{I}_+^* the set $\{\alpha_k \mid k = 1, \dots, K\}$ we have, due to (4.48), the alternative expression of ε that underlines that ε is real:

$$\varepsilon(s_{k_1}, s_{k_2}) = 4\operatorname{Re} \left(\sum_{\alpha \in \mathcal{I}_+^*} R_{s_{k_1}}^{(\alpha)} R_{s_{k_2}}^{(\alpha)*} \right). \quad (4.49)$$

We hence have the following expression:

$$\Upsilon^{(c)}(r) = \frac{\sum_{k=1}^K \zeta(s_k) |f_k|^4 + \sum_{k_1 \neq k_2} |f_{k_1}|^2 |f_{k_2}|^2 \varepsilon(s_{k_1}, s_{k_2})}{\left(\sum_{k=1}^K |f_k|^2 \right)^2} \quad (4.50)$$

where we have let, for any of the sources $s = s_k$ whose positive cyclic frequency is α the number $\zeta(s)$ be

$$\zeta(s) = \eta(s) + 4 \left| R_s^{(\alpha)} \right|^2. \quad (4.51)$$

It is to be noticed that $\zeta(s)$ can also be expressed as

$$\zeta(s) = \left\langle E \left[|s(n)|^4 \right] \right\rangle_n - 2. \quad (4.52)$$

We are in position to discuss the minimization of this function. Two cases have to be distinguished.

2.04.3.7.1 Case of different baud-rates

If the cyclic frequencies $\pm \frac{T_e}{T_k}$ are all different modulo 1, the coefficients $\varepsilon(s_{k_1}, s_{k_2})$ all verify:

$$\forall k_1 \neq k_2 \quad \varepsilon(s_{k_1}, s_{k_2}) = 0. \quad (4.53)$$

Indeed, it suffices to consider the expression of Eq. (4.47): let α_k be the positive cyclo-frequency of the source no k : $\alpha_k = \frac{T_e}{T_k}$. In Eq. (4.47), there are at most two terms: α is either α_{k_1} or $-\alpha_{k_1} = 1 - \alpha_{k_1}$ modulo 1. Take for instance $\alpha = \alpha_{k_1}$. The associated term is non null if $\alpha = \alpha_{k_2}$ or $\alpha = -\alpha_{k_2} = 1 - \alpha_{k_2}$: if $\pm\alpha_{k_1}$ and $\pm\alpha_{k_2}$ are different modulo 1, this cannot occur. The same reasoning applies for $\alpha = -\alpha_{k_1}$. In other words, $\varepsilon(s_{k_1}, s_{k_2}) = 0$.

Remark. It is desirable to bridge the gap between this condition on the disparity of the cyclo-frequencies and practical considerations. We insist on the fact that the above condition is not equivalent to condition that the symbol frequencies are different. Indeed, consider $\alpha_1 = 1/8$ and $\alpha_2 = 7/8$. Then $-\alpha_2 = \alpha_1$ modulo 1 and the condition is not fulfilled whereas the symbol periods are different. However, if the sampling frequency is big enough such that for all indices k : $\alpha_k < 1/2$ then the condition on the disparity of the cyclic frequencies simply means: all the symbol periods are different. In the following, we systematically assume the condition

$$\alpha_k < \frac{1}{2}. \quad (4.54)$$

Proposition 3 can be written in the non-stationary context:

Proposition 14. *If the cyclic frequencies $\pm \frac{T_e}{T_k}$ are all different modulo 1, then the cost function $\Upsilon^{(c)}(r)$ achieves its minimum at a separating vector if and only if one of the $\zeta(s_k)$ is strictly negative. Moreover, the separating vector in question has its single non-null element located at an index k_0 such that $\zeta(s_{k_0}) = \min_{k=1,\dots,K} \zeta(s_k)$.*

We hence consider the following assumption

$$\mathcal{A}_1 \quad \text{for any of source } s \text{ of the mixture } \zeta(s) < 0 \quad (4.55)$$

that is analyzed in Section 2.04.3.9.

Together with Proposition 14, this assumption \mathcal{A}_1 allows one to claim that the minimization of the estimated function $\hat{\Upsilon}(r)$ is legitimate as far as the extraction of a source is concerned. Moreover, the result of Lemma 5 still holds: indeed, the expression of $\Upsilon^{(c)}(r)$ given by (4.50) simplifies as

$$\Upsilon^{(c)}(r) = \frac{\sum_{k=1}^K \zeta(s_k) |f_k|^4}{\left(\sum_{k=1}^K |f_k|^2\right)^2}$$

which is formally the same as the function $\Upsilon(r)$ considered in the stationary case—see Eq. (4.29). The local minima of $\Upsilon^{(c)}(r)$ are separating, which makes the minimization algorithm based on a gradient method robust to the presence of local minima.

We now face the minimization of the function $\Upsilon_{CM}^{(c)}(r)$. The question is whether the non-stationarity of the data changes such a result as the one given in Proposition 7 or not. The careful reader will be assured that indeed Proposition 7 is true in the non-stationary case. Moreover, the result of Proposition 4 is unchanged in the cyclo-stationary case. We can claim:

Proposition 15. *If the cyclic frequencies $\pm \frac{T_e}{T_k}$ are all different modulo 1, and if Assumption \mathcal{A}_1 in (4.55) holds, then the local minima of the cost functions $\Upsilon^{(c)}(r)$ and $\Upsilon_{CM}^{(c)}$ are separating.*

2.04.3.7.2 General case

Contrary to the case where the cyclic-frequencies are different, the Eq. (4.53) does not hold in general, and the results given in the stationary context concerning the minimization of $\Upsilon^{(c)}(r)$ cannot be directly recast. However, numerical considerations may help to give similar results. In this respect, the reader should understand that the quantities $\varepsilon(s_{k_1}, s_{k_2})$ are either zero or “small” since they involve cyclo-correlation at non-null frequencies. This is specific of telecommunication signals; the reason is due to the fact that the excess bandwidth of a transmitted signal is small (see [37]). We will discuss this fact further on, but for the moment, we simply consider the following assumption for any of the sources, denoted by s :

$$\mathcal{A}_2 \quad \text{for any source } s \text{ of the mixture } |R_s^{(\alpha)}| < \frac{1}{2} \sqrt{|\zeta(s)|}. \quad (4.56)$$

We discuss this assumption further in Section 2.04.3.9. For the moment, we suppose that Assumptions \mathcal{A}_1 and \mathcal{A}_2 both hold.

Consider two distinct indices k_1 and k_2 ; for sake of simplicity, let them be $k_1 = 1$ and $k_2 = 2$. Notice that if the sources numbered 1 and 2 are such that their associated (single) cyclic frequencies are different modulo 1 then $\varepsilon(s_1, s_2) = 0$. Suppose on the contrary now that the cyclic frequencies are equal modulo 1 (thanks to (4.54) this happens if and only if the two baud-rates are the same). In this case, the summation (4.49) has only one term and $\varepsilon(s_1, s_2) = 4\mathcal{R}(R_{s_1}^{(\alpha)} R_{s_2}^{(\alpha)*})$. Assumption \mathcal{A}_2 given in (4.56) holds, it follows *in both cases* that $|\varepsilon(s_1, s_2)| < \sqrt{|\zeta(s_1)| |\zeta(s_2)|}$. Whatever the indices $k_1 \neq k_2$ may be, Assumption \mathcal{A}_2 implies

$$\forall k_1 \neq k_2 \quad |\varepsilon(s_{k_1}, s_{k_2})| < \sqrt{|\zeta(s_{k_1}) \zeta(s_{k_2})|}. \quad (4.57)$$

We may consider the expression of $\Upsilon^{(c)}(r)$ given in Eq. (4.50). We have, thanks to (4.57):

$$\sum_{k=1}^K \zeta(s_k) |f_k|^4 + \sum_{k_1 \neq k_2}^K |f_{k_1}|^2 |f_{k_2}|^2 \varepsilon(s_{k_1}, s_{k_2}) \geq - \sum_{k=1}^K |\zeta(s_k)| |f_k|^4 - \sum_{k_1 \neq k_2} |f_{k_1}|^2 |f_{k_2}|^2 \sqrt{|\zeta(s_{k_1}) \zeta(s_{k_2})|}. \quad (4.58)$$

As the r.h.s of the latter equation is simply $-\left(\sum_k \sqrt{|\zeta(s_k)|} |f_k|^2\right)^2$, this shows that

$$\Upsilon^{(c)}(r) \geq - \left(\frac{\sum_k \sqrt{|\zeta(s_k)|} |f_k|^2}{\sum_k |f_k|^2} \right)^2 \geq - \max_k |\zeta(s_k)|.$$

If Assumption \mathcal{A}_1 holds, this lower bound is evidently reached. It remains to inspect the cases of equality for this lower-bound. In this respect, we suppose that $\Upsilon^{(c)}(r) = \zeta_{\min}$ where $\zeta_{\min} = \min_k \zeta(s_k)$. For convenience, we assume that $\sum_{k=1}^P |f_k|^2 = 1$. The case of equality implies that (4.58) is an equality. It implies that

$$\sum_{k_1 \neq k_2} |f_{k_1}|^2 |f_{k_2}|^2 \left(\varepsilon(s_{k_1}, s_{k_2}) + \sqrt{|\zeta(s_{k_1}) \zeta(s_{k_2})|} \right) = 0.$$

Now, the inequality \mathcal{A}_2 implies that $\varepsilon(s_{k_1}, s_{k_2}) + \sqrt{|\zeta(s_{k_1}) \zeta(s_{k_2})|} > 0$. Necessarily, we must have that, for every couples (k_1, k_2) such that $k_1 \neq k_2$, $f_{k_1} f_{k_2} = 0$. Otherwise stated: the vector \mathbf{f} has a single non-null component, hence is separating.

We have shown the result:

Proposition 16. *If the Assumptions \mathcal{A}_1 and \mathcal{A}_2 hold, then the cost function $\Upsilon^{(c)}(r)$ and $\Upsilon_{CM}^{(c)}$ achieve their minimum at a separating vector.*

2.04.3.8 Global BSS (general cyclo-stationary case)

Again in this section, the global source separation method depicted in Section 2.04.3.5 is considered. Its convergence was specifically shown for stationary sources. We show that the convergence of the algorithm to a separating matrix is not affected when the data are cyclo-stationary. This key-result was first provided in [38]; a condition of separability of JADE was given, which was rather difficult to interpret especially when the number of sources is greater than 3. We show the result quite differently, following [39].

Notice that the pre-whitening of the observed data is still possible even when the data are cyclo-stationary: the algorithm is even not changed at all. Hence, we begin by considering the estimate $\hat{\Psi}(\mathbf{r})$ given by Eq. (4.37). Obviously, this estimate can not converge to the function of the cumulants given by Eq. (4.34) since the terms in this equation depend on the time lag. Nevertheless, the limit as the number of snapshots grows can be expressed as

$$\Psi^{(c)}(\mathbf{r}) = \sum_{i=1}^K \left(\left\langle E \left[|r_i(n)|^4 \right] \right\rangle_n - 2 \right).$$

Thanks to the algebra already done along Section 2.04.3.7, we may directly write

$$\Psi^{(c)}(\mathbf{r}) = \sum_{i=1}^K \left(\sum_{k=1}^K \zeta(s_k) |f_{i,k}|^4 + \sum_{k_1 \neq k_2} |f_{i,k_1}|^2 |f_{i,k_2}|^2 \varepsilon(s_{k_1}, s_{k_2}) \right).$$

It remains little to do as soon as Assumptions \mathcal{A}_1 and \mathcal{A}_2 hold, since the following inequality holds:

$$\Psi^{(c)}(\mathbf{r}) \geq - \sum_{i=1}^K \left(\sum_{k=1}^K \sqrt{\|\zeta(s_k)\|} |f_{i,k}|^2 \right)^2 = \|\mathbb{F}\mathbf{z}^{(1/2)}\|_2^2,$$

where $\mathbf{z}^{(1/2)} = (|\zeta|(s_1), \dots, |\zeta|(s_K))^T$. The same argument as the one given in Section 2.04.3.5 (Birkhoff theorem) proves that $\Psi^{(c)}(\mathbf{r}) \geq - \sum_{k=1}^K |\zeta(s_k)|$. As the numbers $\zeta(s_k)$, thanks to Assumption \mathcal{A}_1 , are strictly negative, we directly show that $- \sum_{k=1}^K |\zeta(s_k)|$ is the infimum of $\Psi^{(c)}(\mathbf{r})$ and this infimum is reached for unitary matrices \mathbf{F} that are essentially equal to the identity matrix. This proves the following proposition, that is the sister of Proposition 11:

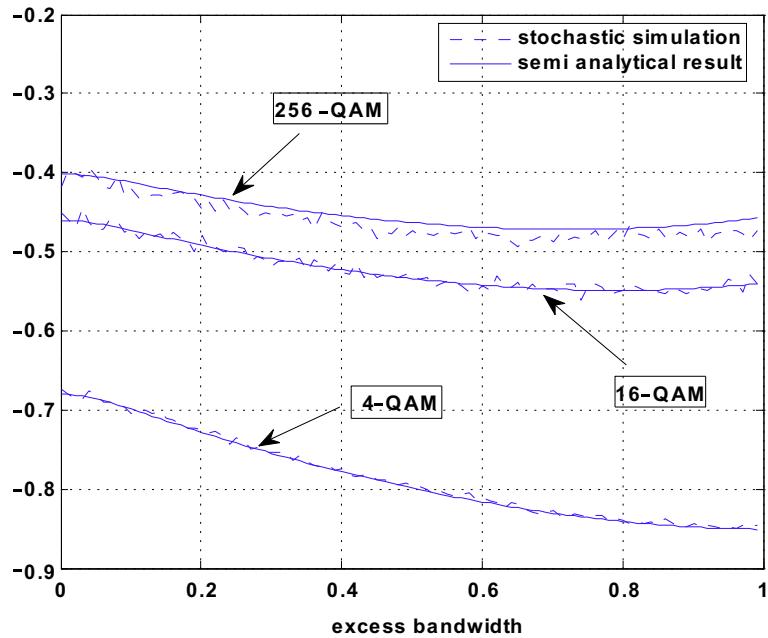
Proposition 17. *If Assumptions \mathcal{A}_1 and \mathcal{A}_2 hold, we have, for any unitary \mathbf{F} :*

$$\Psi^{(c)}(\mathbf{r}) \geq \Psi^{(c)}(\mathbf{s}) = \sum_{k=1}^K \zeta(s_k).$$

Moreover, the inequality is an equality if and only if \mathbf{F} essentially equals the identity matrix.

2.04.3.9 Validity of assumptions \mathcal{A}_1 and \mathcal{A}_2 : semi-analytical considerations

In essence, we have shown that the cyclo-stationarity of the data does not affect neither the one-by-one methods of Section 2.04.3.7, nor the global method depicted in Section 2.04.3.5. This positive answer to our question however requires to inspect the validity of Assumptions \mathcal{A}_1 and \mathcal{A}_2 . We first discuss \mathcal{A}_1 . We recall that, in a stationary environment, $\zeta(s)$ is simply the fourth-order cumulant of the (normalized) source. As this latter is a filtered version of an i.i.d. sequence of symbols having a strictly negative Kurtosis, it is straight-forward that $\zeta(s) < 0$. In a cyclo-stationary environment, the mentioned argument does not hold anymore. In this case, indeed, $\zeta(s)$ is given by (4.51) and it is not possible to conclude directly that $\zeta(s) < 0$. However, these statistics can be shown to express as integrals of the

**FIGURE 4.7**

Validity of \mathcal{A}_1 .

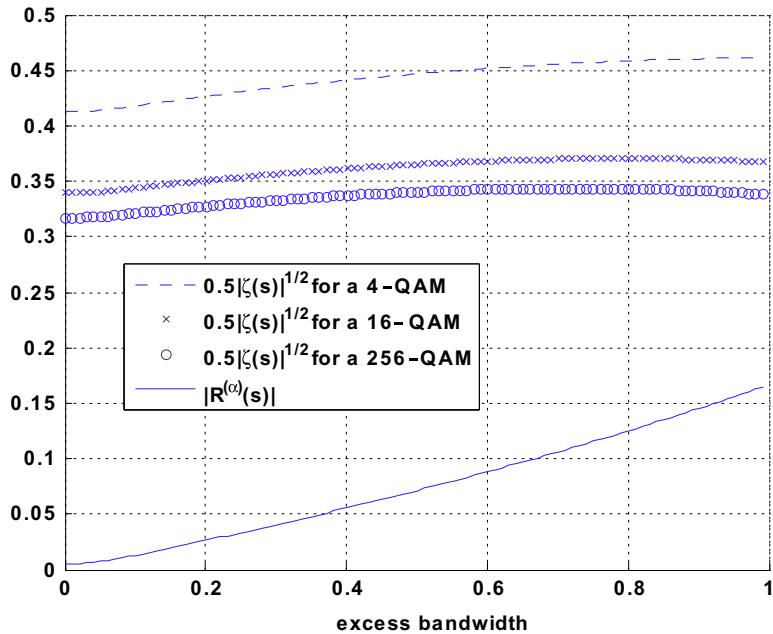
shaping filter $c_a(t)$, at least for a generic sampling frequency (see [40]). More precisely, except for four sampling frequencies that are irrelevant to consider, it can be shown quite simply that

$$R_s^{(\alpha)} = \frac{1}{T} \int |c_a(t)|^2 e^{-i2\pi t/T} dt$$

and

$$\zeta(s) = \kappa(d) \frac{1}{T} \int |c_a(t)|^4 dt + 4|R_s^{(\alpha)}|^2,$$

where c_a , as specified in the introduction, is a normalized square-root raised-cosine filter. In [40], it is shown rigorously that $R_s^{(\alpha)}$ and $\zeta(s)$ do not depend on the symbol period T . It is hence possible to compute numerically $\zeta(s)$ as a function of the excess bandwidth, while considering a few values of $\kappa(d)$ corresponding to typical modulations. The reader may find details on the computation of these numbers in the above reference. The results are reported in Figure 4.7. In order to validate that the above formulas are correct, we also have plotted the estimate of $\zeta(s)$ obtained by stochastic simulations with respect to Eq. (4.52) (we have generated sequences of 10,000 symbols according to the modulation). For all the values of the excess bandwidth, the numerical results let us claim that \mathcal{A}_1 is true.

**FIGURE 4.8**

Validity of \mathcal{A}_2 .

At a first sight, Assumption \mathcal{A}_2 looks quite audacious; indeed it is. As far as we know, it is not possible, for the same reasons as those mentioned above, to prove this result analytically. Resorting again to semi-analytical considerations, we compute the numbers $\frac{1}{2}\sqrt{|\zeta(s)|}$ and $|R_s^{(\alpha)}|$ for different modulations and excess-bandwidth factors. We have obtained the results that can be seen in Figure 4.8. These computations indicate that \mathcal{A}_2 can be claimed to hold.

2.04.4 Convulsive mixtures: case of sparse channels

We now face the problem of blind source separation when the channel is not memoryless. The contribution of a given source to received data is not simply the delayed source up to an unknown constant, but a filtered version of the source. In this section, we specify the model and we explain how it is possible to connect quite directly the previous results (on instantaneous mixtures) to this model: this requires a strong assumption on the delays as will be explained. We refer to [41] (Chapter 17) for more details.

The convulsive effect stems from the presence of multiple paths as specified in Section 2.04.2.1. Consider a source $s_{a,k}(t)$; the contribution of this source on the received signal (m th sensor) is given by (4.9). In most of the high-rate digital communication systems, the narrow-band assumption may be

taken for granted. By narrow-band signal, we mean that the carrier frequency is big enough to consider the signal as monochromatic. More specifically, this means that the bandwidth of the signal in baseband is much smaller than the carrier frequency, i.e.,

$$\frac{1}{T_k} \ll f_k.$$

On the other hand, for any delay τ such that $\tau \ll T_k$, we have the approximation: $s_{a,k}(t - \tau) \approx s_{a,k}(t)$. In order to take the full benefit of the antenna array, the distance between two consecutive antennas should be of the order of half the wave-length. This says that the delay of propagation between two consecutive antennas is of the order of $\frac{1}{2f_k} \ll T_k$. As a consequence, the contribution of the k th source and ℓ th path to the mixture is a rank one signal and can be written as $\mathbf{h}_{k,\ell}s_a(t - \tau_{k,\ell})$ where $\mathbf{h}_{k,\ell}$ is called *steering vector*.

We assume that the associated delays are sorted such that $\tau_{1,k} < \tau_{2,k} < \dots$. We consider the case when the delays are “sufficiently spread out.” In this respect, the reader should notice that it is possible to use the fact that the shaping function $c_{a,k}$ numerically vanishes: $c_{a,k}(t)$ can be numerically neglected if $|t| > P_k T_k$ where the integer P_k depends on the roll-off. This implies that

$$s_{a,k}(t) \approx \sum_{\ell=\left\lfloor \frac{t}{T_k} \right\rfloor - P_k}^{\left\lfloor \frac{t}{T_k} \right\rfloor + P_k} d_k(\ell) c_{a,k}(t - \ell T_k).$$

Consider that the above approximation is an equality. As a consequence, if $\tau_{2,k} - \tau_{1,k} > 2P_k T_k$ then the random variables $s_{a,k}(t - \tau_{1,k})$ and $s_{a,k}(t - \tau_{2,k})$ do not involve the same symbols hence they are independent. Two consecutive paths can hence be treated as *independent sources*. Suppose that all the successive delays are separated by more than $2P_k T_k$, this property holding true for all the sources. As L_k denotes the number of paths associated with the k th source, we may write the observed data according to Eq. (4.11). This time, the “source vector” $\mathbf{s}_a(t)$ is defined as

$$\left(\underbrace{s_{a,1}(t - \tau_{1,1}), \dots, s_{a,1}(t - \tau_{1,L_1})}_{L_1 \text{“sources”}}, \dots, \underbrace{s_{a,K}(t - \tau_{1,K}), \dots, s_{a,K}(t - \tau_{L_K,K})}_{L_K \text{“sources”}} \right)^T.$$

The $\tilde{K} = L_1 + L_2 + \dots + L_K \geq K$ “sources” in the vector $\mathbf{s}_a(t)$ are (approximately) mutually independent. The advantage of this formulation is that all the results given in Section 2.04.3.7 remain true. Nevertheless, we want to stress the drawbacks of this approach.

1. A fundamental requirement is that the number of sensors, M , is greater than the number of sources. We recall that this necessary condition comes from the requirement that the mixing matrix \mathbf{H} should have full -column rank. This gives here

$$M \geq \tilde{K}.$$

This condition may be limiting as soon as the number of paths per communication source is big.

2. The source separation methods all require that the components of $\mathbf{s}_a(t)$ be mutually independent. Now, the delays are spread out in very specific conditions; this particular condition occurs

in long-range communication channels (ionospheric transmissions): we refer to the reference [41] (Chapter 17) for details. On the contrary, this assumption on the distribution of the delays associated with a given source is scarcely fulfilled in many cases (urban GSM channels for instance).

3. The complexity of the instantaneous (one-by-one or global) methods directly increases with the number of sources.
4. The algorithms of source separation, ideally, should provide, up to constants, all the components of $\mathbf{s}_a(t)$. For instance, L_1 of these “sources” among the \tilde{K} components involve the sequence of symbols $d_1(n)$. As the ultimate goal is to eventually estimate these transmitted symbols, a recombination of the L_1 “sources” has to be computed. Notice that the sources are generally not ordered in any way. The association of the reconstructed paths can be processed after lagged correlation between the reconstructed sources. This is clearly not a child’s play.

For practical considerations, we refer to Chapter 17 of reference [41] where a comparison between instantaneous (described in this section) and convulsive approaches described in the next section are done.

2.04.5 Convulsive mixtures

We now face the case of the general multi-path channels; no such condition as the sparsity of the channels is assumed to hold.

2.04.5.1 Identifying the symbols: algebraic methods (stationary data)

In this section, the model of the data is stationary and is given by (4.18). We have justified in the introduction that it is legitimate to approximate the MIMO filter $\mathbf{E}(z)$ by a polynomial. We denote by L its order, i.e.,

$\mathbf{E}(z) = \sum_{\ell=0}^L \mathbf{E}_\ell z^{-\ell}$. The approaches that we want to introduce exploit algebraic properties of the model (convolution) and of the source signal $\mathbf{d}(n)$. Van der Veen et al. [42] suggest the following method.

Consider the vector

$$\mathbf{y}_P(n) = (\mathbf{y}(n)^T, \mathbf{y}(n-1)^T, \dots, \mathbf{y}(n-(P-1))^T)^T \quad (4.59)$$

for $P \geq L$. We remark that

$$\mathbf{y}_P(n) = \mathcal{T}_P(\mathbf{E})\mathbf{d}_{P+L-1}(n),$$

where $\mathbf{d}_{P+L-1}(n)$ is defined in the same manner as $\mathbf{y}_P(n)$ and where matrix $\mathcal{T}_P(\mathbf{E})$ is the $MP \times K(L+P)$ block-Toeplitz matrix given by

$$\mathcal{T}_P(\mathbf{E}) = \begin{pmatrix} \mathbf{E}_0 & \dots & \mathbf{E}_L & 0 & \dots & 0 \\ 0 & \mathbf{E}_0 & \dots & \mathbf{E}_L & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{E}_0 & \dots & \mathbf{E}_L \end{pmatrix}, \quad (4.60)$$

where we have set $\mathbf{E} = (\mathbf{E}_0, \dots, \mathbf{E}_L)$. Denote by \mathcal{Y}_P the $MP \times (N - P - L)$ matrix (N is the number of snapshots) in which all the data $\mathbf{y}(n)$ from $n = L + P$ to $n = N$ are collected ($P \geq L$): the n th column of \mathcal{Y}_P is $\mathbf{y}_P(n)$ —see (4.59). It yields $\mathcal{Y}_P = \mathcal{T}_P(\mathbf{E}) \mathcal{D}_{P+L,N}$ where $\mathcal{D}_{P+L,N}$ is the block-Toeplitz matrix defined by

$$\mathcal{D}_{P+L,N} = \begin{pmatrix} \mathbf{d}(P+L-1) & \mathbf{d}(P+L) & & \mathbf{d}(N) \\ \mathbf{d}(P+L-2) & \mathbf{d}(P+L-1) & \ddots & \\ \vdots & & \ddots & \\ \mathbf{d}(0) & \mathbf{d}(1) & & \mathbf{d}(N-P-L+1) \end{pmatrix}.$$

It is known (see [13, 43]) that the assumption given in (4.21) involves that $\mathcal{T}_P(\mathbf{E})$ is a full column rank matrix. We deduce that the row space of \mathcal{Y}_P equals the row space of $\mathcal{D}_{P+L,N}$. An idea consists in finding the matrices $\hat{\mathcal{D}}_{P+L,N}$ having the Toeplitz structure of $\mathcal{D}_{P+L,N}$ such that their row space is prescribed. Notice that an ambiguity arises since the estimated $\hat{\mathbf{d}}(n)$ of $\mathbf{d}(n)$ coincide up to an invertible matrix. This latter can be removed by considering one of the algebraic methods (instantaneous mixtures) evoked in Section 2.04.3.2 in which a priori information on the symbols is exploited (finite alphabets or constant modulus modulations).

2.04.5.2 Estimation of the channels: MA/AR structures (stationary data)

Again in this section, the data are stationary and follow the model (4.18), so that $\mathbf{y}(n)$ appears to be a moving average model driven by non-Gaussian i.i.d. sequences. A number of blind identification methods of MA models using higher statistics have been derived, and could be used in the present context (see e.g., [44]). However, the corresponding algorithms show poor performance.

If $\mathbf{E}(z)$ is irreducible—see (4.21)—there exists a left polynomial inverse of $\mathbf{E}(z)$ [43]—say $\mathbf{G}(z)$. This implies that $[\mathbf{G}(z)]\mathbf{y}(n) = \mathbf{d}(n)$ where $\mathbf{d}(n)$ is i.i.d., which says that $\mathbf{y}(n)$ is an AR model. This can be used in order to identify the matrix $\mathbf{E}(z)$ thanks to a linear prediction approach, and hence to retrieve the symbol sequences [45] up to a constant $K \times K$ matrix. We note however that the irreducibility of $\mathbf{E}(z)$ does not hold when the excess band-width are all zero due to the factorization evoked at the end of Section 2.04.3.2. This tends to indicate a certain lack of robustness of the linear prediction method when the excess bandwidth factors are small.

2.04.5.3 Estimation of the channels: subspace methods (cyclo-stationary data)

In the previous sections, the main assumption is that the data are stationary. Here, we rather consider the general model (4.17). We consider here the general cyclo-stationary model.

In order to achieve BSS, it is suggested here to identify the transfer function $\mathbf{H}(z)$, and then to evaluate one of its left inverse in order to retrieve the source signals (this kind of approaches is sometimes called *indirect*). In this respect, $\mathbf{H}(z)$ is usually modeled as a FIR causal filter, i.e., $\mathbf{H}(z) = \sum_{\ell=0}^L \mathbf{H}_\ell z^{-\ell}$. It has been shown in [43] that if $M > K$ and $\mathcal{T}_P(\mathbf{H})$ is full column rank, then $\mathbf{H}(z)$ can be identified from the column space of $\mathcal{T}_P(\mathbf{H})$ up to a constant $K \times K$ matrix which can itself be estimated using any instantaneous mixture blind source separation method. In practice, the column space of $\mathcal{T}_P(\mathbf{H})$ is usually estimated by means of the eigenvalue/eigenvector decomposition of an estimate of the covariance matrix

of vector $\mathbf{y}_P(n)$ given in (4.59). We should perhaps specify this point. Indeed, we have

$$\begin{aligned}\hat{R}_{\mathbf{y}_P}(0) &= \frac{1}{N} \sum_{n=L+P}^N \mathbf{y}_P(n) \mathbf{y}_P(n)^H \\ &= \mathcal{T}_P(\mathbf{H}) \hat{\Gamma} \mathcal{T}_P(\mathbf{H})^H,\end{aligned}$$

where $\hat{\Gamma} = \frac{1}{N} \sum_{n=L+P}^N \mathbf{s}_{P+L-1}(n) \mathbf{s}_{P+L-1}(n)^H$. If this latter is full rank, the column space of $\hat{R}_{\mathbf{y}_P}(0)$ coincides with the column space of $\mathcal{T}_P(\mathbf{H})$. However, it should be emphasized that it is not always legitimate to assume that $\hat{\Gamma}$ is full rank except when all the sources occupy all the band of frequencies $[-\frac{1}{2T_e}, \frac{1}{2T_e}]$. On the contrary, the rank of $\hat{\Gamma}$ is expected to fall. We do not provide the details: the reader should compute the limit of $\hat{\Gamma}$ as $N \rightarrow \infty$ and show that the limit Γ is a block Toeplitz matrix whose block (p, q) has the expression:

$$\Gamma_{p,q} = \int_0^1 e^{-i2\pi(p-q)v} S_s^{(0)}(e^{i2\pi v}) dv,$$

where $S_s^{(0)}(e^{i2\pi v})$ is the (diagonal) power spectral density of the vector $\mathbf{s}(n)$. As the sampling period verifies the Shannon sampling condition, some of the entries of $S_s^{(0)}(e^{i2\pi v})$ are band-limited, which prevents the rank of Γ from being full (see the works of Slepian and Pollak on the prolate spheroidal wave functions).

Further refinements are proposed in [12]. Notice that this subspace method, although apparently quite appealing, performs poorly as soon as the matrix $\mathcal{T}_P(\mathbf{H})$ is ill-conditioned, which, in practice, is quite often the case.

2.04.5.4 Global BSS approaches

2.04.5.4.1 Temporal approaches: extensions of the Comon contrast function (stationary data)

We now focus on the direct and global BSS methods, i.e., methods that allow one to compute a global separator in one shot (up to indeterminacies that will be specified). In this respect, we intend to provide extensions of the approaches given for the instantaneous context—see Section 2.04.3.5. For simplification, assume the (restrictive) stationary case. As specified in the previous paragraph, the model of the data is not stricto sensu $[\mathbf{H}(z)]\mathbf{s}(n)$ as in Eq. (4.17) since the Shannon sampling condition does not hold, but rather (4.18).

On the one hand, a direct extension of a contrast as defined by Comon can be done (see Section 8.4.2 of [41]). The first step to be considered is the decorrelation of the data (the prewhitening), i.e., a filter matrix $\mathbf{Q}(z)$ is computed such that $\mathbf{x}(n) = [\mathbf{Q}(z)]\mathbf{y}(n)$ is decorrelated both spatially and temporally. This is equivalent to computing $\mathbf{Q}(z)$ such that $\mathbf{U}(z) = \mathbf{Q}(z)\mathbf{E}(z)$ is para-unitary, i.e., $\mathbf{U}(z)$ verifies for any frequency $v \in [-\frac{1}{2}, \frac{1}{2}]$: $\mathbf{U}(e^{i2\pi v})\mathbf{U}(e^{i2\pi v})^H = \mathbf{I}_K$. The reader may find details of this procedure in [46, 47]. The reconstructed vector of the sources may hence be searched as $\mathbf{r}(n) = [\mathbf{V}(z)]\mathbf{x}(n)$ where

$\mathbf{V}(z)$ is a para-unitary matrix. It is possible to consider, as in the instantaneous case, the function

$$\Psi(\mathbf{r}) = \sum_{k=1}^N \varphi(\kappa(r_k)),$$

where φ shows the same properties as in Proposition 12. It can be simply shown that Ψ , as a function of the para-unitary matrix $\mathbf{V}(z)$, is a contrast, i.e., achieves its maximum for separating matrices $\mathbf{U}(z)$: see [48]. Theoretically appealing (at least in a stationary environment), this solution calls for implicit prerequisites, namely the pre-whitening and the optimization over the set of para-unitary matrices. Concerning this latter point, solutions have been given (see [10, 49]); Comon et al. also suggested to consider a subset of the set of the para-unitary matrices [50]. The solutions are not simple. Besides, the algorithms might be trapped in local maxima.

2.04.5.4.2 Frequency-domain approaches

In the case of stationary data (same restrictive context as in the above subsection), it is possible to recast the results of the instantaneous case after processing the discrete Fourier transform of the data: the separation is processed at each frequency. The difficulty is that the indeterminacies change from a frequency to the other which makes the approach quite difficult. The reader may find references in [51].

Assuming now the general cyclo-stationary model given by (4.17), one may focus on second-order methods. The power spectrum of the data, defined as the discrete Fourier transform of the correlation function at the null cyclic frequency $(R_y^{(0)}(\ell))_{\ell \in \mathbb{Z}}$ can be written as

$$S_y^{(0)}(e^{j2\pi\nu}) = \mathbf{H}(e^{j2\pi\nu}) S_s^{(0)}(e^{j2\pi\nu}) \mathbf{H}(e^{j2\pi\nu})^H, \quad (4.61)$$

where $S_s^{(0)}(e^{j2\pi\nu})$ is the $K \times K$ power spectrum of the source vector given by (4.10). The components of \mathbf{s} being jointly independent, hence decorrelated, the matrices $S_s^{(0)}(e^{j2\pi\nu})$ are diagonal. As far as the identifiability of the unknown $\mathbf{H}(z)$ based on the relation (4.61) is addressed, it is shown in [52] that the conditions required are that (4.20) holds on the one hand and, on the other hand, that spectral diversity occurs namely that the entries of $S_s^{(0)}(e^{j2\pi\nu})$ are all distinct for every frequency. In [53, 54] it is shown that the matrices $\mathbf{G}(e^{j2\pi\nu})$ such that $\mathbf{G}(e^{j2\pi\nu})^H S_y^{(0)}(e^{j2\pi\nu}) \mathbf{G}(e^{j2\pi\nu})$ is diagonal for all the frequencies ν are separating matrices. Criteria measuring the closeness to diagonal matrices have been proposed. See also [55] and the references therein. Again, the diversity of the spectra is not always pertinent for digital communication contexts.

2.04.5.5 Iterative BSS (stationary case)

Instead of considering a simple spatial filtering of the data as indicated in Eq. (4.24) we rather process a spatio-temporal filtering as depicted below:

$$r(n) = [\mathbf{g}(z)^T]^T \mathbf{y}(n). \quad (4.62)$$

The “reconstructed source” $r(n)$ can be expanded as

$$r(n) = \sum_{k=1}^K \underbrace{[f_k(z)] d_k(n)}_{r_k(n)}, \quad (4.63)$$

where the $f_k(z)$ are components of the global filter

$$\mathbf{f}(z) = \mathbf{g}(z)^T \mathbf{E}(z). \quad (4.64)$$

A key trick for the following results is the normalization step: we might write $r_k(n) = \|f_k\|\tilde{r}_k$ where $E[|\tilde{r}_k|^2] = 1$ and

$$\|f_k\|^2 = \int_0^1 |f_k(e^{i2\pi v})|^2 dv. \quad (4.65)$$

Notice that the separation is achieved if and only if the real-valued vector $(\|f_1\|, \dots, \|f_K\|)^T$ is separating in the sense given in Section 2.04.3.4. As a consequence, when the separation is achieved, the “reconstructed source” is one of the sources up to a filter with unit norm.

With no modification as compared to the method given for the instantaneous case in 2.04.3.4, we consider the optimization of the function $\Upsilon(r)$ as given in Eq. (4.27)—or equivalently (4.28). For sake of simplicity, we keep the notation $\Upsilon_{CM}^{(c)}(r)$ and $\Upsilon^{(c)}(r)$ even if it should be understood that the functions in question depend on the filters $f_1(z), \dots, f_K(z)$. No extra computation is needed as compared to the instantaneous case: indeed, it suffices to substitute the “new source” $\tilde{r}_k(n)$ to the actual source $s_k(n)$: one arrives at the expression:

$$\Upsilon(r) = \frac{\sum_{k=1}^K \kappa(\tilde{r}_k) \|f_k\|^4}{\left(\sum_{k=1}^K \|f_k\|^2\right)^2}.$$

This time, the cumulants $\kappa(\tilde{r}_k)$ are not constant but depend on the norm-one filter $\tilde{f}_k(z)$. Anyway, $\tilde{r}_k(n)$ is a linear process generated by the symbol sequence $d_k(n)$: indeed, $\tilde{r}_k(n) = [\tilde{f}_k(z)]s_k(n)$ and $s_k(n)$ has the form given by Eq. (4.7). As such, $\kappa(\tilde{r}_k) < 0$ since by assumption $\kappa(d_k) < 0$. We let $\kappa_{\min,k}$ be:

$$\kappa_{\min,k} = \inf_{\|\tilde{f}_k\|=1} \kappa([\tilde{f}_k(z)]s_k(n))$$

and $\kappa_{\min} = \min_k \kappa_{\min,k}$. We obviously have: $\kappa_{\min,k} < 0$ and $\kappa_{\min} < 0$. As a consequence, the following inequality holds:

$$\Upsilon(r) \geq \kappa_{\min} \frac{\sum_{k=1}^K \|f_k\|^4}{\left(\sum_{k=1}^K \|f_k\|^2\right)^2} \geq \kappa_{\min}.$$

Moreover, the equality holds if and only if

1. $\|f_k\| = \|f_{k_0}\| \delta(k - k_0)$.
2. $\tilde{f}_{k_0}(z)$ is such that $\kappa_{\min} = \kappa_{\min,k_0} = \kappa(\tilde{r}_{k_0}(n))$ where $\tilde{r}_{k_0}(n) = [\tilde{f}_{k_0}(z)]s_{k_0}(n)$.

We suggest an qualitative interesting remark as far as the reconstructed signal $r(n) = \|f_{k_0}\|\tilde{r}_{k_0}(n)$. Indeed, this filter minimizes the Kurtosis of $\tilde{r}_{k_0}(n)$. As $\tilde{r}_{k_0}(n)$ is a filtered version of the non-Gaussian i.i.d. sequence $(d_{k_0}(n))_{n \in \mathbb{Z}}$, it is shown in [56] that its minimum is reached when $\tilde{r}_{k_0}(n)$ is $e^{i\theta} d_{k_0}(n - n_0)$ where θ is an unknown phase and n_0 an uncontrolled delay. The reader might show that such a result also holds for the function $\Upsilon_{CM}(r) = E[(|r(n)|^2 - 1)^2]$.

Ultimately, any local minimum of $\Upsilon(r)$ or $\Upsilon_{CM}(r)$ can be shown to be separating [31, 57].

2.04.5.6 Iterative BSS (general cyclo-stationary)

Similarly to the instantaneous case (Section 2.04.3.7), we recast the approach of the previous section for cyclo-stationary data. The received data is $\mathbf{y}(n) = [\mathbf{H}(z)]\mathbf{s}(n)$, the reconstructed source $r(n)$ is still given by (4.62). We may expand this signal as

$$r(n) = \sum_{k=1}^K \underbrace{[f_k(z)]s_k(n)}_{r_k(n)}. \quad (4.66)$$

This time the global filter $\mathbf{f}(z)$ is

$$\mathbf{f}(z) = \mathbf{g}^T(z)\mathbf{H}(z).$$

As far as the normalization of $r(n)$ is concerned, we still have $r(n) = \sum_k \|f_k\| \tilde{r}_k(n)$ but $\|f_k\|$ does not have the expression (4.65) since $s_k(n)$ is not i.i.d. with unit power but has a non-constant power spectral density $S_k^{(0)}(e^{j2\pi\nu})$:

$$\|f_k\|^2 = \int_0^1 \left| f_k(e^{j2\pi\nu}) \right|^2 S_k^{(0)}(e^{j2\pi\nu}) d\nu. \quad (4.67)$$

As a consequence, we consider the minimization of the functions $\Upsilon_{CM}^{(c)}(r)$ and $\Upsilon^{(c)}(r)$ whose definition is given by Eqs. (4.43) and 4.44. Similarly to the instantaneous case, the minimization of $\Upsilon_{CM}^{(c)}(r)$ and $\Upsilon^{(c)}(r)$ are equivalent problems in the sense of Proposition 4. Here, the functions in question depend not only on the positive coefficients $\|f_1\|, \dots, \|f_K\|$ but also on the norm-one filters $\tilde{f}_1(z), \dots, \tilde{f}_K(z)$. We recall that $\tilde{r}_k(n) = [\tilde{f}_k(z)] s_k(n)$ where $\|\tilde{f}_k\| = 1$; then, with practically no effort, we deduce from Section 2.04.3.7 that

$$\Upsilon^{(c)}(r) = \frac{\sum_{k=1}^K \zeta(\tilde{r}_k) \|f_k\|^4 + \sum_{k_1 \neq k_2} \|f_{k_1}\|^2 \|f_{k_2}\|^2 \varepsilon(\tilde{r}_{k_1}, \tilde{r}_{k_2})}{\left(\sum_{k=1}^K \|f_k\|^2 \right)^2}, \quad (4.68)$$

where

$$\varepsilon(\tilde{r}_{k_1}, \tilde{r}_{k_2}) = 4\text{Re} \left(\sum_{\alpha \in \mathcal{I}_+^*} R_{\tilde{r}_{k_1}}^{(\alpha)} R_{\tilde{r}_{k_2}}^{(\alpha)*} \right), \quad (4.69)$$

$$\begin{aligned} \zeta(\tilde{r}_k) &= \eta(\tilde{r}_k) + 4 \left| R_{\tilde{r}_k}^{(\alpha)} \right|^2, \\ \eta(\tilde{r}_k) &= \langle \kappa(\tilde{r}_k(n)) \rangle. \end{aligned} \quad (4.70)$$

It is to be noticed that $\zeta(\tilde{r}_k)$ can also be expressed as

$$\zeta(\tilde{r}_k) = \left\langle E \left[|\tilde{r}_k(n)|^4 \right] \right\rangle_n - 2. \quad (4.71)$$

Similarly to the instantaneous case, where the filters $\tilde{f}_k(z)$ are reduced to being 1, the minimization of $\Upsilon^{(c)}(r)$ is considerably easier when the cross-terms in (4.68) vanish, i.e., when for all the couples

(k_1, k_2) with $k_1 \neq k_2$ we have $\varepsilon(\tilde{r}_{k_1}, \tilde{r}_{k_2}) = 0$. Indeed, in this case,

$$\Upsilon^{(c)}(r) = \frac{\sum_{k=1}^K \zeta(\tilde{r}_k) \|f_k\|^4}{\left(\sum_{k=1}^K \|f_k\|^2\right)^2}$$

which resembles the expression (4.29) except that here, the numbers $\zeta(\tilde{r}_k)$ depend on the parameters the minimization is run over. In the following we hence restrict the further analysis to this case. We recall that $\varepsilon(\tilde{r}_{k_1}, \tilde{r}_{k_2}) = 0$ as soon as the cyclic frequencies $\alpha_k = \pm \frac{T_e}{T_k}$ are all different modulo one. This condition is fulfilled when all the baud rates are different and the sampling frequency is high enough, i.e., $\frac{1}{T_e} > \max_k \frac{2}{T_k}$.

We define $\zeta_{\min,k} = \inf_{\|\tilde{f}_k\|=1} \zeta([\tilde{f}_k(z)]s_k(n))$.

Proposition 18. *If the cyclic frequencies $\pm \frac{T_e}{T_k}$ are all different modulo 1 then the cost function $\Upsilon^{(c)}(r)$ achieves its minimum at a separating vector $\mathbf{f}_*(z)$ if and only if at least one of the $\zeta_{\min,k}$ is strictly negative. Moreover, the \mathbf{f}_* in question has its single non-null element located at an index k_0 such that $\zeta_{\min,k_0} = \min_{k=1,\dots,K} \zeta_{\min,k}$.*

Proof. If $\zeta_{\min,k} < 0$ for a certain index k , then $\zeta_{\min} < 0$ where we have defined $\zeta_{\min} = \min_\ell \zeta_{\min,\ell}$. Evidently, $\sum_{\ell=1}^K \zeta(\tilde{r}_\ell) \|f_\ell\|^4 \geq \sum_{\ell=1}^K \zeta_{\min,\ell} \|f_\ell\|^4 \geq \zeta_{\min} \sum_{\ell=1}^K \|f_\ell\|^4$. This shows that $\Upsilon^{(c)}(r) \geq \zeta_{\min}$ and this lower bound is attained for any vector $\mathbf{f}_*(z)$ whose components are all zero except at an index k_0 such that $\zeta_{\min,k_0} = \zeta_{\min}$ and $\zeta([\tilde{f}_{k_0}(z)]s_{k_0}(n)) = \zeta_{\min,k_0}$. Conversely, if $\zeta_{\min,k} \geq 0$ whatever k , the lower bound is attained for non separating filters (see the proof of Proposition 3). \square

Contrary to the stationary case, it is not possible to say much about the minimizing filter $\tilde{f}_{k_0}(z)$ and hence about the reconstructed signal $r(n)$. However, the residual filter $\tilde{f}_{k_0}(z)$ minimizes the $\zeta(\tilde{r}_{k_0})$ hence tends to make the modulus of $\tilde{r}_k(n)$ the most constant as possible.

At this point, we have to analyze the condition: $\zeta_{\min,k} < 0$. Recall that this condition is the adaptation to the convulsive case of the assumption \mathcal{A}_1 . This latter was proven to hold true by means of semi-analytical considerations. Curiously, it is possible to prove that the condition $\zeta_{\min,k} < 0$ holds; no numerical computation is needed. We have:

Proposition 19. *As the Kurtosis of the symbols $\kappa(d_k)$ are all strictly negative, we have*

$$\zeta_{\min,k} = \inf_{\|\tilde{f}_k\|=1} \zeta([\tilde{f}_k(z)]s_k(n)) < 0.$$

Proof. More solid arguments than the ones we give here are provided in [40]. The band of frequencies of the sampled version $s_k(n)$ of the source number k is the interval $[-\frac{1+\gamma_k}{2}\alpha_k, \frac{1+\gamma_k}{2}\alpha_k] \subset [-\frac{1}{2}, \frac{1}{2}]$. The infimum to be computed is hence over the set of unit-norm filters $\tilde{f}_k(z)$ belonging to $\mathcal{F}(\gamma_k)$ where $\mathcal{F}(\gamma_k)$ is the set of digital filters whose transfer function is limited to the band $[-\frac{1+\gamma_k}{2}\alpha_k, \frac{1+\gamma_k}{2}\alpha_k]$. Naturally, $\gamma'_k < \gamma_k$ implies that $\mathcal{F}(\gamma'_k) \subset \mathcal{F}(\gamma_k)$. Taking $\gamma'_k = 0$, we deduce the following inequality:

$$\zeta_{\min,k} = \inf_{\|\tilde{f}_k\|=1, \tilde{f}_k(z) \in \mathcal{F}(\gamma_k)} \zeta([\tilde{f}_k(z)]s_k(n)) \leq \inf_{\|\tilde{f}_k\|=1, \tilde{f}_k(z) \in \mathcal{F}(0)} \zeta([\tilde{f}_k(z)]s_k(n)).$$

On the other hand, we recall that $\zeta([\tilde{f}_k(z)]s_k(n)) = \zeta(\tilde{r}_k) = \langle \kappa(\tilde{r}_k(n)) \rangle + 4 \left| R_{\tilde{r}_k}^{(\alpha_k)} \right|^2$. Thanks to Section 2.04.2.2, we know that $R_{\tilde{r}_k}^{(\alpha_k)} = 0$ when the excess bandwidth factor is zero. Hence, for any unit norm filter $\tilde{f}_k(z) \in \mathcal{F}(0)$, $\zeta(\tilde{r}_k) = \langle \kappa(\tilde{r}_k(n)) \rangle$. We deduce that, for such a filter, $\zeta(\tilde{r}_k) < 0$. \square

When all the baud-rates are equal, the minimization of $\Upsilon_{CM}^{(c)}(r)$ is much more difficult and sufficient conditions on the sources have been set forth that assure that the minimizers of $\Upsilon_{CM}^{(c)}(r)$ are separating. We refer to [40].

However, the general case remains to be addressed (i.e., for any distribution of the baud-rates): we conjecture that the global minimum is separating and even that any local minimum is also separating. These conjectures come from intensive simulation experiments.

2.04.6 Simulation

This section does not aim at making a benchmark of all the previous methods. It rather intends to show the pertinence of BSS in digital communication contexts.

We first present the environment. We have considered a mixture of $K = 3$ sources; the modulations are QPSK (two sources) and 16-QAM (one source). The symbol periods are all equal to T and $1/T = 277$ kHz (which is the rate of GSM); the carrier frequency is 1 GHz. This latter is assumed known to the receiver so that there are no frequency offset in the source vectors. The excess bandwidth factors all equal $\gamma = 0.5$. As far as the antenna array is concerned, we have simulated a circular array of $M = 5$ sensors distanced from one another by half a wavelength, i.e., 30 cm. The sampling period T_e is fixed to $T_e = T/1.6$ so that the Shannon sampling condition is fulfilled.

The propagation channels are multi path and affected by a Rayleigh fading. An arbitrary path—say number ℓ is characterized by its delay (propagation between the source and a sensor of reference), its elevation, azimuth and attenuation. We consider the ETSI channels BUx, TUx, HTx, RAx. For each experiment, the arrival angles of a path are randomly chosen in $[-\pi/2, \pi/2]$ for the elevation and $[-\pi, \pi]$ for the azimuth. The different complex amplitudes on each path are also randomly chosen for each experiment.

The received signal is corrupted by a white, additive complex gaussian noise with power spectral density N_0 . The received signals are low-pass filtered in the band $[-\frac{1}{2T_e}, \frac{1}{2T_e}]$. The three sources have the same energy per symbol \mathcal{E}_s . This latter is chosen so that the \mathcal{E}_s/N_0 of the 16-QAM source would be associated with a bit error rate of 10^{-3} if the channel were a single-path channel and if there were only one antenna.

We have considered three algorithms for the separation. In order to have a reference, we have computed the Wiener solution (the MMSE separator): this is naturally a non-blind algorithm. In this respect, we have supposed that the receiver has access to the transmitted symbols. The two BSS algorithms we have considered are the iterative CMA (see Section 2.04.5.6) and the global method JADE. Once the separation is (supposedly) achieved, we need to equalize each source in order to compute the bit error rate. This extra-step is done after re-sampling the three “source estimates” at the true sampling rate (supposed at this level to be known). The re-sampled signals are expected to be filtered versions

Table 4.1 Some Simulation Results

	2000			1000			500		
	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4
channel BUx									
MMSE	100%	68.8%	100%	100%	55.9%	100%	100%	23%	100%
CMA	100%	70.5%	100%	100%	20.6%	99.8%	95.9%	0.1 %	97.6%
JADE	58.4	1.3%	60.7%	55.5%	0.7%	57.7%	48.8%	0.1%	51.7%
channel TUx	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4
MMSE	100%	80.2%	100 %	100%	69.8%	100%	100%	35.3%	100%
CMA	99.7%	84.3%	99.3%	99.3%	47.5%	99.2%	98.5%	0.7%	98.9%
JADE	94.9%	29.9%	95.1%	93.9%	18.1%	94.4%	93.1%	3%	92.9%
channel RAx	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4
MMSE	99.8%	91.9%	99.6%	99.8%	88.8%	99.6%	99.7%	70.4%	99.6%
CMA	99.1%	84.7%	99.2%	99%	50.2%	99.3%	98.1%	0.6%	98.6%
JADE	99.7%	89.2%	99.6%	99.6%	79.7%	99.6%	99.6%	24.5%	99.5%
channel HTx	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4	QAM-4	QAM-16	QAM-4
MMSE	99.4%	36.2%	99.6%	99.4%	26.8%	99.7%	99.4%	9%	99.5%
CMA	92%	13%	93.2%	90.8%	1.8%	91.6%	83.5%	0%	82.1%
JADE	86.7%	13.1%	87.4%	85.7%	6.3%	86.3%	83.6%	1.1%	84.1%

of the symbols. A blind equalization algorithm is run (the CMA) in order to compute estimates of the transmitted symbols. The scaling ambiguity is removed non-blindly.

The performance index for the tested algorithms and for a given source is the following: for each experiment, we inspect if the bit error rate (for the transmitted symbols) is less than 10^{-2} . Averaging the process on 1000 independent trials allows us to provide a percentage of success.

Three observation durations were considered: $500T$, $1000T$, and $2000T$.

The performance of the algorithms can be seen in Table 4.1. The performance of the MMSE (non-blind) provides, in some sense, an ultimate bound. In this respect, one may notice that the channels BUx and HTx are by far the “most difficult” channels since they are associated with the weakest performance of the MMSE. For these two channels, the multipath effects are severe and this explains why the JADE algorithm, designed for instantaneous mixtures, performs poorly. On the contrary, the CMA shows good performance as far as the extraction of the two modulus-one sources is concerned, even if the observation duration is small (500 symbols): the performance index is above 80%. However, the 16-QAM source is associated with a miserable performance: the CMA as a BSS algorithm is not to be incriminated, since the two other sources are correctly equalized, hence the 16-QAM is itself correctly separated. Hence the bad performance is due to the equalization algorithm (again the CMA run this time on the re-sampled extracted signal): this is a well-known fact that the CMA equalizes the non-modulus one modulation with difficulties.

As far as the TUx and RAx channels are concerned, they are associated with less severe multi-path effects: this explains why the JADE algorithm performs well—even better than the CMA for the RAx channel.

2.04.7 Extensions and further readings

2.04.7.1 Case of non-circular sources

Along this paper, the data we considered as circular. This assumption is not crucial for second-order or algebraic methods. However, the presence of a non-circular source in the mixture considerably affects the higher-than-second-order methods. For instance, the fourth-order cumulant $\kappa(r(n))$ if $r(n)$ is the output of a separator, does not have the same expression as when all the sources are circular. It can be even been shown that the separation is not always achieved when two sources are non-circular.

The interested reader might find results and references in the following works: [38, 58–60].

2.04.7.2 Exploiting the non-stationarity

Cyclo-stationarity is a main statistical feature of the mixture that has long been thought of as a benefit for source separation. For instantaneous BSS using second-order moments, see [61, 62]: the idea is that the mixing matrix is constant during the observation, while the second-order statistics of the sources vary. An idea is hence to cut the observation interval in subintervals. Recalling the SOBI approach (see Section 2.04.3.3), the receiver may compute the correlation matrices for the u th interval: $R_y^{(u)}(\ell) = \mathbf{H}R_s^{(u)}(\ell)\mathbf{H}^H$. As the sources are non-stationary, the diagonal matrices $R_s^{(u)}(\ell)$ vary with u hence the pencil of matrices to be jointly diagonalized has more elements and the conditions of identifiability are weaker, hence the algorithm is more robust. We refer to the work of Pham [63] for the Maximum Likelihood approach. The reader might be interested by a work of Wang et al. [64].

In the case of digital communication signals, the cyclo-stationarity is not strong enough in order to consider such approaches: we have pointed out that the power of a source $E[s_a(t + \tau)s_a(t)^*]$ has very small variations since the cyclic-spectra at the cyclic-frequencies $\pm\frac{1}{T}$ are numerically small due to the spectral limitation of the shaping functions. On the one hand, the strength of cyclo-stationarity is too weak to be exploited. On the other hand, it cannot be neglected in the computations (for instance in the expression of the fourth-order cumulants).

2.04.7.3 Presence of additive noise

In this paper, we have considered a noise-free model. Many references may be found where the impact of the noise on the BSS methods is analyzed. Among others, we would like to cite the work of Cardoso concerning the performance analysis a class of BSS algorithms who have a the so-called “invariance” feature [35]. Concerning the CMA when noise is present: the reader may have a look at the work of Fijalkow et al. [29] for the use of the CMA as an equalizer algorithm and the proximity of a solution in the presence of noise to a Minimum Mean Square Error (MMSE) equalizer; the case of BSS is analyzed in [25, 65] where it is shown that the local minima of the CM cost function are “not far” from MMSE separator. A systematic analysis is provided in Leshem et al. [66] where both the presence of noise and the effect of a finite number of samples are considered.

2.04.8 Conclusion

In this paper, we have given some methods for achieving the BSS in the context of telecommunication. We have focused our attention on contrast functions (joint or deflation-based approaches), and particularly the contrast functions depending on fourth-order statistics of the data. These approaches fit the blind problems evoked in the introduction (spectrum monitoring) since they are associated with convergent and performance algorithms. When the channels involve multi-path effects and no a priori information on the distribution of the delays is available, the deflation-based algorithms such as the CMA or the minimization of the normalized Kurtosis are good candidates for the BSS.

Relevant Theory: Signal Processing Theory and Statistical Signal Processing

See [Vol. 1, Chapter 4 Random Signals and Stochastic Processes](#)

See [Vol. 3, Chapter 3 Non-stationary Signal Analysis](#)

References

- [1] C. Jutten, J. Hérault, Blind separation of sources, Part I—An adaptative algorithm based on neuromimetic architecture, *Signal Process.* 24 (1991) 1–10.
- [2] P. Comon, Independent component analysis, a new concept? *Signal Process.* 36 (3) (1994) 287–314 (special issue on higher-order statistics).
- [3] E. Bingham, H. Hyvärinen, A fast fixed-point algorithm for independent component analysis of complex valued signals, *Int. J. Neural Syst.* 10 (1) (2000) 1–8.
- [4] J.F. Cardoso, A. Souloumiac, Blind beamforming for non gaussian signals, *IEEE Proc. F* 140 (6) (1993) 362–370.
- [5] N. Delfosse, P. Loubaton, Adaptative blind separation of independant sources: a deflation approach, *Signal Process.* 45 (1995) 59–83.
- [6] Boaz Porat, *A Course in Digital Signal Processing*, first ed., Wiley, October 1996.
- [7] W.A. Gardner, Spectral correlation of modulated signals: Part I—Analog modulations, *IEEE Trans. Commun.* 35 (1987) 584–594.
- [8] Antonio Napolitano, Cyclic higher-order statistics: Input/output relations for discrete- and continuous-time MIMO linear almost-periodically time-variant systems, *Signal Process.* 42 (2) (1995) 147–166.
- [9] John Proakis, *Digital Communications*. McGraw-Hill Science/Engineering/Math, fourth ed., August 2000.
- [10] P.P. Vaidyanathan, *Multirate Systems And Filter Banks*, first ed., Prentice Hall, 1992.
- [11] C. Corduneanu, *Almost Periodic Functions*, second ed., American Mathematical Society, 1989.
- [12] A. Gorokhov, P. Loubaton, Subspace-based techniques for blind separation of mixtures with temporally correlated sources, *IEEE Trans. Circ. Syst.* 44 (9) (1997) 813–820.
- [13] Yujiro Inouye, Ruey-Wen Liu, A system-theoretic foundation for blind equalization of an FIR MIMO channel systemchannel system, *IEEE Trans. Circ. Syst.—I* 49 (4) (2002) 425–436.
- [14] A.J. Van Der Veen, A. Paulraj, An analytical constant modulus algorithm, *IEEE Trans. Signal Process.* 5 (44) (1996) 1136–1155.
- [15] S. Talwar, M. Viberg, A. Paulraj, Blind estimation of multiple co-channel digital signals using an antenna array, *Signal Process. Lett.* 1 (2) (1994) 29–31.

- [16] S. Talwar, M. Viberg, A. Paulraj, Blind separation of synchronous co-channel digital signals using an antenna array—Part I: Algorithms, *IEEE Trans. Signal Process.* 44 (5) (1996) 1184–1197.
- [17] B.G. Agee, The least-squares CMA: a new technique for rapid correction of constant modulus signals, in: *Proceedings of the ICASSP*, Tokyo, 1986, pp. 953–956.
- [18] R.P. Gooch, J.P. Lundell, The CM array: an adaptive beamformer for constant modulus signals, in: *Proceedings of the ICASSP*, Tokyo, Japan, 1986, pp. 2523–2526.
- [19] A. Belouchrani, K. Abed-Meraim, J-F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, *IEEE Trans. Signal Process.* 45 (2) (1997) 434–444.
- [20] A. Bunse-Gerstner, R. Byers, V. Mehrmann, Numerical methods for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.* 14 (4) (1993) 927–949.
- [21] K. Abed-Meraim, Y. Xiang, H. Manton, Y. Hua, Blind source separation using second-order cyclostationary statistics, *IEEE Trans. Signal Process.* 49 (4) (2001) 694–701.
- [22] Patrick Billingsley, *Probability and Measure*, third ed., Wiley-Interscience, 1995.
- [23] D N Godard, Self-recovering equalization and carrier tracking in two-dimensionnal data communication systems, *IEEE Trans. Commun.* 28 (11) (1980) 1867–1875.
- [24] Y. Sato, A method of self-recovering equalization for multilevel amplitude-modulation systems, *IEEE Trans. Commun.* 23 (6) (1975) 679–682.
- [25] D. Liu, L. Tong, An analysis of constant modulus algorithm for array signal processing, *Signal Process.* 73 (1999) 81–104.
- [26] J.R. Treichler, B.G. Agee, A new approach to multipath correction of constant modulus signals, in: *IEEE Trans. Acoust. Speech, Signal Process.* vol. 31, 1983, pp. 459–472.
- [27] J.R. Treichler, M.G. Larimore, New processing techniques based on constant modulus adaptive algorithm, *IEEE Trans. Acoust. Speech Signal Process.* 33 (8) (1985) 420–431.
- [28] P.A. Regalia, On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization, *Signal Process.* 73 (1999) 185–190.
- [29] I. Fijalkow, A. Touzni, J.R. Treichler, Fractionally spaced equalization using CMA: robustness to channel noise and lack of disparity, *IEEE Trans. Signal Process.* 46 (1) (1998) 227–231.
- [30] C. Simon, Ph Loubaton, C. Jutten, Separation of a class of convolutive mixtures: a contrast function approach, *Signal Process.* 81 (2001) 883–887.
- [31] J.K. Tugnait, Identification and deconvolution of multi-channel non-gaussian processes using higher-order statistics and inverse filter criteria, *IEEE Trans. Signal Process.* 45 (3) (1997) 658–672.
- [32] J.K. Tugnait, Adaptive blind separation of convolutive mixtures of independent linear signals, *Signal Process.* 73 (1999) 139–152.
- [33] Albert W. Marshall, Ingram Olkin, Barry Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer, September, 2010.
- [34] A. Yeredor, Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation, *IEEE Trans. Signal Process.* 50 (7) (2002) 1545–1553.
- [35] J.-F. Cardoso, On the performance of orthogonal source separation algorithms, in: *EUSIPCO*, Edinburgh, 1994, pp. 776–779.
- [36] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Trans. Neural Networks* 10 (3) (1999) 626–634.
- [37] L. Mazet, P. Loubaton, Cyclic correlation based symbol rate estimation, in: *proc. Asilomar Conference on Signals, Systems, and Computers*, 1999, pp. 1008–1012.
- [38] A. Ferréol, P. Chevalier, On the behavior of current second and higher order blind source separation methods for cyclostationary sources, *IEEE Trans. Signal Process.* 48 (2002) 990 (6 (Erratum: 50, N°4) 1712–1725).
- [39] P. Jallon, A. Chevreuil, Separation of instantaneous mixtures of cyclostationary sources, *Signal Process.* 87 (2007) 2718–2732.

- [40] Pierre Jallon, Antoine Chevreuil, Philippe Loubaton, Separation of digital communication mixtures with the CMA: case of unknown symbol rates, *Signal Process.* (2010) 2633–2647.
- [41] Pierre Comon, Christian Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [42] A.J. Van Der Veen, S. Talwar, A. Paulraj, A subspace approach to blind space-time signal processing for wireless communication systems, *IEEE Trans. Signal Process.* 45 (1) (1997) 173–190.
- [43] K. Abed-Meraim, P. Loubaton, E. Moulines, A subspace algorithm for certain blind identification problems, *IEEE Trans. Inform. Theory* 43 (2) (1997) 499–511.
- [44] A. Swami, G.B. Giannakis, S. Shamsunder, Multichannel ARMA processes, *Trans. Signal Process.* 42 (4) (1994) 898–913.
- [45] A. Gorokhov, Ph Loubaton, Blind identification of MIMO-FIR systems: a generalized linear prediction approach, *Signal Process.* 73 (1999) 104–124.
- [46] A. Belouchrani, A. Cichocki, Robust whitening procedure in blind separation context, *Electron. Lett.* 36 (24) (2000) 2050–2051.
- [47] Peter J. Brockwell, Richard A. Davis, *Time Series: Theory and Methods*, second ed., Springer, 1991.
- [48] P. Comon, Contrasts for multichannel blind deconvolution, *IEEE Signal Process. Lett.* 3 (209–211) (1996).
- [49] John G. McWhirter, Paul D. Baxter, Tom Cooper, Soydan Redif, Joanne Foster, An EVD algorithm for Para-Hermitian polynomial matrices, *IEEE Trans. Signal Process.* 55 (5) (2007) 2158–2169.
- [50] P. Comon, L. Rota, Blind separation of independant sources from convolutive mixtures, *IEICE Trans. Fund. Electron. Commun. Comput. Sci.* E86-A (3) (2003) 542–549.
- [51] W. Wang, S. Sanei, J.A. Chambers, Penalty function-based joint diagonalization approach for convolutive blind separation of non-stationary sources, *IEEE Trans. Signal Process.* 53 (5) (2005) 1654–1669.
- [52] Yingbo Hua, Jitendra K. Tugnait, Blind identifiability of FIR-MIMO systems with colored input using second order statistics 7 (12) (2000) 348–350.
- [53] Mitsuru Kawamoto, Yujiro Inouye, Blind deconvolution of MIMO-FIR systems with colored inputs using second-order statistics, *IEICE Trans. Fund. Electron. Commun. Comput. Sci.* E86-A (3) (2003) 597–604.
- [54] Mitsuru Kawamoto, Yujiro Inouye, Blind separation of multiple convolved colored signals using second-order statistics, in: *Proceedings of the ICA'03*, Nara, Japan, 2003.
- [55] K. Sabri, M. El Badaoui, F. Guillet, A. Adib, D. Aboutajdine, A frequency domain-based approach for blind MIMO system identification using second-order cyclic statistics, *Signal Process.* 89 (1) (2009) 77–86.
- [56] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of non-minimum phase systems, *IEEE Trans. Inform. Theory* 36 (1990) 312–321.
- [57] P. Loubaton, P.A. Regalia, Blind deconvolution of multivariate signals: a deflation approach, in: *International Conference on Communications (ICC)*, Geneva, Switzerland, 1993.
- [58] E. Florian, A. Chevreuil, P. Loubaton, Blind source separation of convolutive mixtures of non circular linearly modulated signals with unknown baud rates, *Signal Process.* 92 (2012) (715–726).
- [59] M. Novey, T. Adali, On extending the complex fastica algorithm to non-circular sources, *IEEE Trans. Signal Process.* 56 (5) (2008) 2148–2153 .
- [60] H. Zhang, L. Li, W. Li, Independent vector analysis for convolutive blind non-circular source separation, *Signal Process.* 92 (9) (2012) 2275–2283.
- [61] M.K. Tsatsanis, C. Kweon, Source separation using second-order statistics: identifiability conditions and algorithms, in: *Asilomar Conference on Signals, Systems and Computers 1998*, pp. 1574–1578.
- [62] A. Souloumiac, Blind source detection and separation using second-order nonstationarity, in: *ICASSP*, 1995, pp. 1912–1915.
- [63] D-T. Pham, J-F. Cardoso, Blind separation of instantaneous mixtures of nonstationary sources, *IEEE Trans. Signal Process.* 49 (9) 2001 1837–1848.

- [64] W. Wang, M.G. Jafari, S. Sanei, J.A. Chambers, Blind separation of convolutive mixtures of cyclo-stationary signals, *Int. J. Adapt. Control Signal Process.* 18 (3) (2004) 279–298.
- [65] L. Tong H. Zeng, C. Johnson, An analysis of constant modulus receivers, *IEEE Trans. Signal Process.*, 47 (1999) 2990–2999.
- [66] A. Leshem, A.-J. van der Veen, Blind source separation: the location of local minima in the case of finitely many samples, *IEEE Trans. Signal Process.* 59 (2008) 4340–4353.

OFDM and Multicarrier Signal Processing

5

Paolo Banelli and Luca Rugini

Dipartimento di Ingegneria, Università degli Studi di Perugia, Perugia, Italy

2.05.1 Introduction

Orthogonal Frequency-Division Multiplexing (OFDM) is indubitably a milestone in wired and wireless communications, which pervades most of the telecommunication standards developed in the last two decades, such as DAB, DVB-T/H, WiFi (IEEE 802.11a/n), WiMAX (IEEE 802.16e), UMTS-LTE, ADSL, etc. [71–73, 104, 105, 107, 237]. However, the OFDM principle dates back to the early 1960s for military communication systems [69, 268]. More generally, the roots of OFDM are at the end of the nineteenth century when frequency division multiplexing (FDM) was investigated to increase the efficiency of telegraph systems [209], and successively was widely employed during the twentieth century in analog telephone lines [209]. Thus, OFDM is the typical example of a technology, whose idea has evolved theoretically for a long time, and that had to wait for mature electronic and software technologies to be implemented, at reasonable costs, for mass-market applications. The interested readers can find in [256] an excellent historical perspective of OFDM evolution, and a recent survey of OFDM literature in [103]. Additional information about OFDM and multicarrier techniques can be found in tutorial articles and books such as [12, 29, 53, 120, 136, 145, 251, 269].

The success of OFDM with respect to classical single-carrier (SC) communications is mainly due to its capability to enable wideband communications in time-dispersive (frequency-selective) channels with low-complexity channel equalization, and high spectral efficiency. The simpler equalization with respect to SC is somehow borrowed from classical FDM, where the whole data stream is split in multiple sub-streams with lower data-rate (bandwidth), each one modulated by a different carrier frequency. This way, each sub-stream experiences an almost frequency-flat channel, which requests a simple equalization.

Additionally, OFDM doubles the spectral efficiency with respect to FDM by employing properly separated (e.g., orthogonal) carriers, which let the spectrum of each sub-stream overlap with one another [48]. By exploiting orthogonality, the multiple sub-streams can be separated at the receiver side by a bank of correlation-based receivers, rather than by the bandpass filtering of FDM, where the spectra of different sub-streams cannot overlap. Although the orthogonality principle is quite intuitive and mathematically simple, its implementation in the analog OFDM systems of the 1960s was requesting a bank of perfectly tuned oscillators, both at the transmitter and the receiver side, making OFDM implementation costly and limited to a small number of parallel carriers [268]. Actually, at the beginning

of the 1970s, Weinstein and Ebert recognized that OFDM signals can be generated and demodulated by Fourier-based synthesis [257], foreseeing the substitution of the bank of analog oscillators by digital Discrete Fourier Transform (DFT) processing together with digital-to-analog (D/A) [and analog-to-digital (A/D)] conversion. The introduction of a cyclic-prefix (CP) between each transmitted symbol has been another crucial step towards easy OFDM implementation in frequency-selective channels, as firstly proposed in [179]. Thanks to the CP, the almost frequency-flat channel experienced by each sub-stream in the continuous domain, is turned in a perfectly flat channel in the discrete (e.g., DFT) domain. Thus, by standard DFT processing, OFDM is capable to convert the transmission of multiple symbols through a frequency-selective channel, in a set of independent parallel transmissions through a set of frequency-flat channels, which can be easily and independently equalized. When at the end of the 1980s, the progress of electronic technology (A/D and D/A converters, CPU clock frequencies, memories, etc.) made real-time digital signal processor (DSP) a reality, the concurrent increasing request for high data-rate transmissions in frequency-selective channels suggested that the time for OFDM “*has come*” [29]. At the beginning of the 1990s European broadcasting manufacturers and operators were among the first to recognize the potentiality of OFDM, by producing DAB and DVB-T as the first wireless communication standards for digital audio and video broadcasting [71, 72, 184].

Similarly, in the same decades, telephone companies and operators recognized to OFDM [or discrete multi-tone (DMT)], the capability to convey high data-rate bit-streams by classical twisted-pairs wires, practically solving the so called last-mile problem for widespread connection of user homes with high-speed digital communications. Indeed, according to the water-filling principle, the use of wires for digital communications permitted also to exploit the almost optimality of OFDM from the capacity point of view, as predicted by the famous Shannon’s paper in 1948 [212]. The first pioneering work of Cioffi and his co-workers on the subject [58], found successively full application in the ADSL standard [107].

The smart use of multiple antennas is another important milestone of wireless communications in the last two decades, while historically multiple antennas were only employed as a source of diversity at the receiver side. Conversely, it has been recognized that multiple transmit and receive antennas can boost the capacity of wireless links by multiple-input multiple-output (MIMO) systems [79], and improve coverage and connection quality by space-time (ST) coding [229]. For instance, the Alamouti ST-coding [5], currently adopted in several communications standards [such as [105]], is a remarkable example of a quite simple idea, which greatly improves the performance of a wireless communication link by employing only two transmit antennas and one, or more, receive antennas.

Anyway, both MIMO and ST-coding approaches have been conceived, and are quite easy to deploy, for SC communications on frequency-flat channels. Thus, the capability to convert frequency-selective channels in a set of parallel frequency-flat channels made OFDM the perfect tool to enable both MIMO and ST-coding techniques. However, this topic is out of the scope of this paper, which for space constraint will address only the fundamental signal processing aspects of OFDM and multicarrier (MC) communications. The interested readers are redirected to [84, 88, 178], and references therein, for an overview of MIMO and ST-coding, and to [37, 111, 263], for their use in OFDM-based systems. For a thorough introduction and deep survey of MIMO techniques, the interested readers are redirected to the E-Reference contribution of Davidson [66].

2.05.2 Mathematical background

The paper is written targeting an electrical engineering student with adequate background on Fourier transforms (continuous and discrete), signal and systems theory, linear algebra basics (vector, matrices, eigenvalues, eigenvector, singular value decomposition, etc.), basics of probability and estimation theory, and basics of digital communications. In order to better motivate why OFDM emerged and obscured classical SC communications in the last decades, the background on classical SC digital communications is summarized in the next paragraph, which is also useful to clarify similarities and differences between SC and OFDM systems.

2.05.3 Single carrier background

The baseband equivalent $x(t)$ of a radio-frequency (RF) modulated signal

$$x_{\text{RF}}(t) = \text{Re}\{x(t)e^{j2\pi f_c t}\}, \quad (5.1)$$

which conveys a stream of digital information symbols $\{d_i \in \mathcal{D} \subset \mathbb{C}\}_{i \in \mathbb{N}}$ by the RF carrier f_c , has the general expression [183]

$$x(t) = \sum_{i=-\infty}^{+\infty} \varphi(t, d_i; \phi_i). \quad (5.2)$$

Equation (5.2) highlights that in general the waveform $\varphi(\cdot)$ associated to the i th symbol d_i may depend on a state variable $\varphi_i = \varphi(d_{i-1}, d_{i-2}, \dots)$ that takes into account some memory, i.e., dependence, on the previously emitted symbols, as it happens for instance for continuous-phase modulations [183]. Focusing on the simpler scenario of linear modulations without memory, (5.2) may be rewritten as

$$x(t) = \sum_{i=-\infty}^{+\infty} d_i p(t - iT_s), \quad (5.3)$$

where T_s is the symbol period, $f_s = 1/T_s$ is the symbol rate, and $p(t) \in \mathbb{R}$ is the pulse-shaping filter that imposes the spectral occupancy.¹ Specifically, for a given information stream, the spectrum of the transmitted signal is expressed by

$$X(f) = P(f) \sum_{i=-\infty}^{+\infty} d_i e^{-j2\pi f iT_s} = P(f) D_{T_s}(f), \quad (5.4)$$

where $X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f t} dt$ is the Fourier transform (FT) of $x(t)$, and $D_{T_s}(f)$ is the discrete-time FT (DTFT) of the transmitted sequence $\{d_i\}_{i \in \mathbb{N}}$. Thus, the information associated to each

¹A more general expression could consider different pulse-shaping filters for the real and imaginary parts of $\{d_i\}$. This is not common in practical applications, and it is not considered herein for notation simplicity. See [28] for further details.

symbol d_i spreads by $D_{T_s}(f)$ throughout all the frequencies, it is mixed with the information associated to the other symbols $\{d_j\}_{j \neq i}$, and it is shaped (and band-limited) by $P(f)$.

The signal $y(t)$ received on an additive white Gaussian noise (AWGN) channel is expressed by

$$y(t) = kx(t - t_o) + w(t), \quad (5.5)$$

where k and t_o represent attenuation and delay, respectively, and $w(t)$ stands for the thermal receiver noise. Thus, the AWGN channel is characterized by the propagation of the transmitted signal through a linear system with impulse response $h_c(t) = k\delta(t - t_o)$, where $\delta(t)$ is the Dirac delta function. In order to recover the transmitted signal, and assuming $k = 1$ and $t_o = 0$ without restriction of generality, the receiver processes $y(t)$ by a filter $p_r(t)$, which is responsible to maximize the signal-to-noise power-ratio (SNR), and to avoid inter-symbol interference (ISI) [183]. The SNR is maximized by a matched filter (MF), e.g., $p_r(t) = p^*(-t)$ [183], leading to a filtered signal $y_f(t)$ expressed by

$$\begin{aligned} y_f(t) &= \sum_{i=-\infty}^{+\infty} d_i p(t - iT_s) * p^*(-t) + n(t) * p^*(-t) \\ &= \sum_{i=-\infty}^{+\infty} d_i R_{pp}(t - iT_s) + w_f(t). \end{aligned} \quad (5.6)$$

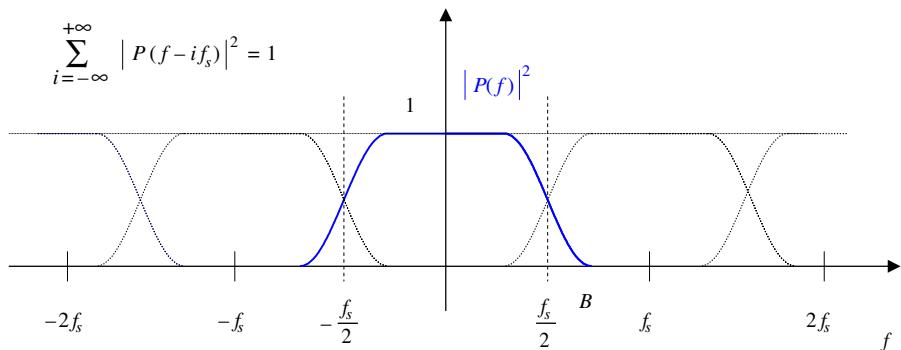
In (5.6), $*$ stands for the linear convolution operator, $*$ is the complex-conjugate operator, $R_{pp}(t) = p(t) * p^*(-t)$ is the pulse-shape autocorrelation function, and $w_f(t)$ is the filtered noise. Thus, the MF is also known as correlator receiver [183]. If the pulse shaper $p(t)$ is designed with unit energy $R_{pp}(0) = 1$ and such that $\{R_{pp}(iT_s) = 0\}_{i \neq 0}$, the sample y_k extracted at $t_k = kT_s$ is expressed by

$$\begin{aligned} y_k &= y_f(kT_s) = R_{pp}(0)d_k + \sum_{i \neq k} d_i R_{pp}((k - i)T_s) + w_k \\ &= d_k + w_k, \end{aligned} \quad (5.7)$$

do not experience any ISI from adjacent symbols, and optimal reception can be performed on a symbol by symbol basis [183]. The ISI-free constraint on $R_{pp}(t)$ imposes a constraint on its FT $\mathcal{F}[R_{pp}(t)] = |P(f)|^2$, which has to be shaped such that [183]

$$\sum_{i=-\infty}^{+\infty} |P(f - i/T_s)|^2 = R_{pp}(0). \quad (5.8)$$

The ISI-free constraint in (5.8) can be satisfied if $|P(f)|$ has an anti-symmetrical shape with respect to the frequency $1/(2T_s)$ and $P(f) = 0$ when $|f| > 1/T_s$ (see Figure 5.1). This is known as the Nyquist criterion, and suggests that the positive (half) bandwidth B associated with a linear digital modulation is $B \in [f_s/2, f_s]$, where $f_s = 1/T_s$. A typical prototype filter that respects such constraints is the root raised-cosine filter [183], whose shape is shown in Figure 5.1 together with the graphical interpretation of (5.8).

**FIGURE 5.1**

Shape of $|P(f)|^2$ for a root raised-cosine filter and ISI-free condition in the frequency-domain.

However, in most of wireless and wired communication systems, the propagation channel is not AWGN. Indeed, because of multipath propagation in wireless communications [222], and of frequency-dependent attenuation in wired communications [183], the propagation channel is modeled by a frequency-selective channel with a time-dispersive impulse response $h_c(t) \neq k\delta(t - t_o)$.

A typical baseband representation for a linear time-invariant (LTI) channel, with $P + 1$ path is summarized by its time-dispersive impulse response

$$h_c(t) = \sum_{p=0}^P \alpha_p \delta(t - \tau_p), \quad (5.9)$$

where α_p and τ_p are the attenuation and delay, respectively, associated with the p th path, and

$$H_c(f) = \sum_{p=0}^P \alpha_p e^{-j2\pi f \tau_p} \quad (5.10)$$

is the frequency-selective channel spectrum, as shown in Figure 5.6.

In the presence of a time-dispersive (frequency-selective) channel, the received signal at the MF output is expressed by

$$y_f(t) = \sum_{i=-\infty}^{+\infty} d_i h(t - iT_s) + w_f(t), \quad (5.11)$$

where $h(t) = p(t) * h_c(t) * p_r(t) = R_{pp}(t) * h_c(t)$ is the overall pulse shape that is perceived at the receiver side. The effect of the channel in the frequency-domain is captured by

$$Y_f(f) = P^*(f) H_c(f) X(f) = H(f) D_{T_s}(f), \quad (5.12)$$

where $H(f) = H_c(f) |P(f)|^2$ replaces the role of $|P(f)|^2$ in (5.8).

Thus, if $H_c(f)$ is not constant in the frequency support where $P(f) \neq 0$, e.g., if $H_c(f) \neq k_o$ when $|f| \leq B$, the frequency selectivity of the physical channel destroys the ISI-free design of Figure 5.1. This means that, in the presence of multipath (frequency-selectivity), the time dispersion of the channel $h_c(t)$ destroys the ISI-free design, because $h(iT_s) \neq 0$, and the received samples at the MF output become

$$y_k = d_k + \sum_{i \neq k} d_i h_{k-i} + w_k, \quad (5.13)$$

where $h_i = h(iT_s)$ is the equivalent discrete-time impulse response of the overall channel, which is generally modeled as a finite impulse response (FIR) filter of order L .

The presence of ISI from adjacent symbols requests both the estimation of the channel coefficients $\{h_i\}_{i=0,\dots,L}$ and more complex detection techniques. For equiprobable transmitted symbols, the optimal detection choice consists in a maximum likelihood (ML) receiver that can exploit the Viterbi algorithm [183], whose complexity grows exponentially with the discrete-time channel length $L + 1$. Alternative and simpler detection schemes, characterized by (sometimes significant) bit-error-rate (BER) performance degradation, first try to counteract the effect of the channel by a linear filter (equalizer), and successively resort to a single-symbol decision as in (5.7). The simplest choice for the linear equalizer is represented by a transversal FIR filter, whose coefficients (taps) can be determined according to different criteria [zero-forcing, minimum mean-squared error (MMSE), etc. [183]]. The channel equalization complexity for each transmitted symbol d_k , without considering the channel estimation and the equalizer design, is linearly proportional to the number of taps L_{eq} of the FIR equalizer. Typically the FIR equalizer suffers from noise amplification and gives acceptable BER performance, when $L_{eq} \in [3L, 5L]$. Thus, we can conclude that the simplest linear equalization in SC is characterized by a complexity for transmitted symbol that grows linearly with the channel order L , rather than exponentially as for ML equalizers.

Another possibility to contrast the effect of the ISI, and avoid the noise amplification associated with linear equalization is to use decision feedback equalization, which is a nonlinear data-aided approach that represents a trade-off between BER performance and computational complexity [183].

It is important to observe that any physical channel is characterized by its own frequency selectivity and coherence bandwidth B_c , i.e., a band $|f| \leq B_c$ where $H_c(f) \approx k_o$ is almost constant. If the frequency support of the transmitted signal is within the channel coherence bandwidth, i.e., if the bandwidth of $P(f)$ is $B \leq B_c$, the channel will not significantly alter the ISI-free design of Eq. (5.8) and Figure 5.1, which will be almost respected by $H(f) = H_c(f)|P(f)|^2 \approx k_0|P(f)|^2$. On the contrary, the effect of ISI becomes critical if the system symbol rate f_s increases, leading to a proportional increase of the system bandwidth B (see Figure 5.1), which may become greater than the coherence bandwidth B_c . In the discrete-time domain, the increase of the symbol rate would correspond to an increased number $L + 1$ of channel coefficients $h_i \neq 0$ [222], and consequently an increased complexity for the channel equalizer.

Just to give an idea, a wireless SC system in a urban environment is typically characterized by a multipath channel $h_c(t)$ with a maximum delay spread $\Delta\tau = \tau_P - \tau_0 \in [7, 20] \mu\text{s}$ [60]: if the requested system capacity can be granted by a bandwidth $W = 2B = 10 \text{ MHz}$, this means that the channel length is of the order $L \approx \Delta\tau/T_s \approx \Delta\tau W \in [70 \div 200]$, and a linear transversal FIR equalizers with at least $[210 \div 600]$ taps is request to obtain acceptable BER performance. Needless to say that in this scenario it is impossible to think about ML equalization/detection, whose complexity is exponential in L .

Thus, the ever increasing demand for high data-rate communications in the last decades has posed tremendous challenges to channel estimation and equalization of classical SC wideband communications. This was the main motivation, even if not the single one, to look for an alternative communication scheme that could more easily handle the propagation through frequency-selective (multipath) channels.

2.05.4 The path to OFDM

This section is dedicated to show how the OFDM principle can naturally evolve from SC communications, through the classical concept of FDM.

2.05.4.1 Frequency division multiplexing

When the information stream $\{d_i\}$ is characterized by a very-high symbol rate f_s , a natural way to deal with the frequency selectivity of the channel is to split the data stream into M multiple parallel sub-streams. This way, each sub-stream is characterized by a reduced symbol rate $f'_s = f_s/M$ and, consequently, can be transmitted by a baseband signal $x^{(m)}(t)$ expressed by

$$x^{(m)}(t) = \sum_{k=-\infty}^{+\infty} d_{m,k} p_m(t - kMT_s), \quad (5.14)$$

where $d_{m,k} = d_{kM+m}$ represents a serial-to-parallel (S/P) conversion, and $p_m(t)$ is the pulse-shaping waveform of the m th sub-stream. The simplest way to separate M data streams at RF was historically offered by FDM [29], where all the streams are characterized by the same bandwidth $W_m = W/M = 2B/M$, and are modulated at RF on different (sub-) carriers $f_{RF,m} = f_c + (m - M/2)\Delta f$. The subcarrier separation $\Delta f > W/M$ grants that each sub-stream can be separated from the others by simple bandpass filtering at the receiver side (see $H_m(f)$ in Figure 5.2), followed by a classical MF receiver for each sub-stream.

Thus, taking into account (5.1), the equivalent baseband model of the overall transmitted stream is simply represented by the superposition of the M parallel baseband streams, as expressed by

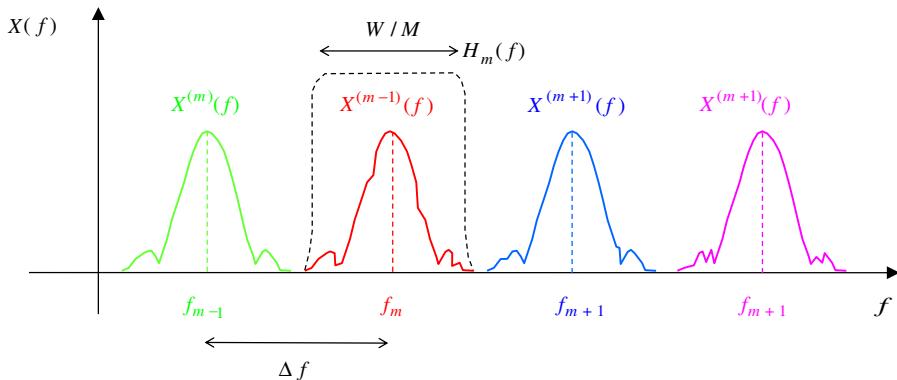
$$x(t) = \sum_{m=0}^{M-1} x^{(m)}(t) = \sum_{k=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} p_m(t - kMT_s), \quad (5.15)$$

where

$$p_m(t) = p(t)e^{j2\pi f_m t}, \quad (5.16)$$

and $f_m = (m - M/2)\Delta f$, thus producing a *multicarrier* (MC) communication system.

As for SC communications, the role of $p(t)$ is twofold, i.e., to shape the spectrum $X^{(m)}(f)$ and to guarantee ISI-free reception by the receiver MF of each single sub-stream. The role of $e^{j2\pi f_m t}$ is to shift in frequency each single sub-stream, such that the sub-streams are almost separated, i.e., $X^{(m)}(f)X^{(n)}(f) \simeq 0$: this way, each single sub-stream can be easily separated by bandpass filtering (see also Figure 5.2). In practice, in order to filter each sub-stream with physically realizable filters, it is also necessary to introduce a frequency guard-band B_g between adjacent frequency channels and,

**FIGURE 5.2**

The concept of frequency division multiplexing.

consequently, $\Delta f = 2\frac{B}{M} + B_g$. Thus, the overall RF spectral occupancy is $2B + (M - 1)B_g$ and FDM sacrifices spectral efficiency with respect to the SC solution, whose spectrum support is $2B$ (see Figure 5.1).

2.05.4.2 Orthogonal multiplexing principle

The multiple-stream expression in (5.14) and (5.15) may have a much broader interpretation than simple FDM. On one hand, in order to allow for a simple receiver architecture, e.g., a parallel of M SC-like receivers (see Figure 5.3), the pulse shaper $p_m(t)$ has to guarantee the maximum SNR by MF and the ISI-free properties of each single sub-stream, as in FDM. On the other hand, additionally, the pulse shaper $p_m(t)$ should guarantee the absence of interference from the other (superimposed) $M - 1$ sub-streams.

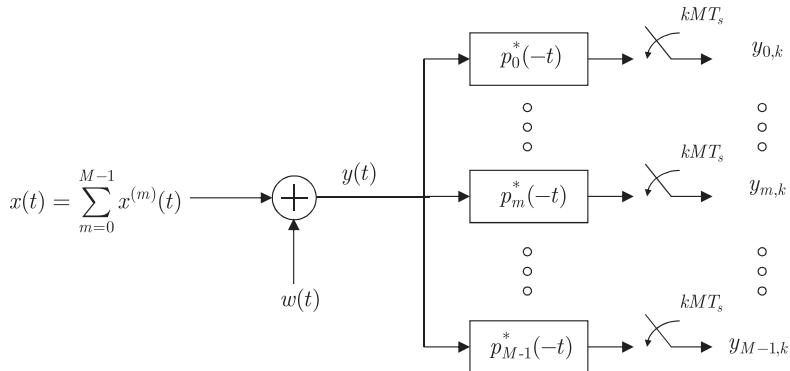
By means of (5.6), the output of the m th MF is

$$y_f^{(m)}(t) = \sum_{l=-\infty}^{+\infty} d_{m,l} R_{p_m p_m}(t - lMT_s) + \sum_{j \neq m} \sum_{l=-\infty}^{+\infty} d_{j,l} R_{p_m p_j}(t - lMT_s) + n_f^{(m)}(t), \quad (5.17)$$

which, assuming the ISI-free design for $p_m(t)$ and a sampling rate $f'_s = (1/MT_s)$, corresponds to a discrete-time signal expressed by

$$y_{m,k} = y_f^{(m)}(kMT_s) = R_{p_m p_m}(0)d_{m,k} + \sum_{j \neq m} \sum_{l=-\infty}^{+\infty} d_{j,k} R_{p_m p_j}((k - l)MT_s) + n_{m,k}. \quad (5.18)$$

Equation (5.18) highlights that the interference caused by the symbols $\{d_{j,k}\}_{j \neq m}$ of the other streams $j \neq m$ can be eliminated if, and only if, the pulse-shaper cross-correlation $R_{p_m p_j}(t) = p_m(t) * p_j^*(-t)$ is characterized by equispaced samples $R_{p_m p_j}(iMT_s) = 0$.

**FIGURE 5.3**

System architecture with a bank of parallel MF at the receiver side.

For practical considerations, such as the transmission of a symbol in a limited time, it is natural to consider pulse shapers with a limited duration (time support) T_u . To this end, let us express the m th pulse shaper as

$$p_m(t) = c_m(t)\text{rect}_{T_u}(t - T_u/2), \quad (5.19)$$

where $\text{rect}_T(t) = 1$ when $|t| < T/2$ and $\text{rect}_T(t) = 0$ elsewhere, and $c_m(t)$ contains the shape of the pulse within the limited time support. The finite time support T_u leads also to autocorrelation and cross-correlation functions with limited time-support $2T_u$, i.e., $R_{p_m p_j}(t) = 0$ when $|t| \geq T_u$, for $j = 0, \dots, m, \dots, M-1$. Thus, choosing $T_u \leq MT_s$ would easily guarantee that $R_{p_m p_j}(iMT_s) = 0$ when $i \neq 0$. This allows for the elimination of the ISI caused by those symbols (of all the M substreams) transmitted in the interval $l \neq k$, where k is the interval index of the target symbol. Thus, for time-limited pulse shapers with unit energy $R_{p_m p_m}(0) = 1$ and duration $T_u = MT_s$, (5.18) reduces to

$$y_{m,k} = d_{m,k} + \sum_{j \neq m} d_{j,k} R_{p_m p_j}(0) + n_{m,k}, \quad (5.20)$$

where the m th stream suffers only from the interference caused by those symbols (of the $M-1$ substreams with $j \neq m$) transmitted during the same k th symbol interval. In order to avoid such an interference, it is necessary to design the M pulses such that

$$R_{p_m p_j}(0) = \int_{-\infty}^{+\infty} p_m^*(t)p_j(t)dt = \int_0^{T_u} c_m^*(t)c_j(t)dt = 0, \quad (5.21)$$

which means that the pulse shapers have to be orthogonal over the finite support T_u . Therefore, for any orthogonal and unit-energy design of the pulse shapers, the output of the m th MF branch in Figure 5.3, is expressed by

$$y_{m,k} = d_{m,k} + n_{m,k}, \quad m = 0, \dots, M-1. \quad (5.22)$$

There are wide classes of functions that are orthogonal over a finite interval, and every choice would be equivalent in this respect. By such a design, the system in Figure 5.3 is nothing else than a classical (short) code-division multiple-access (CDMA) receiver [183], where $c_m(t)$ is the signature waveform (i.e., the pulse-shaped code) associated with the m th sub-stream. The interested reader can refer to [247] for further details on CDMA principles and systems, and Section 2.05.10.2 for a comparison of MC and CDMA philosophies.

2.05.4.3 Orthogonal frequency-division multiplexing principle

Noteworthy, the exponential functions $e^{j2\pi f_m t}$ involved in the FDM design are candidate functions for $c_m(t)$ in (5.21), if the frequencies $\{f_m\}$ are carefully selected. Indeed, in order to approximate any regular function on a finite time-interval, the *Fourier* series [173] exploits orthogonal exponential functions whose frequencies are proportional to the fundamental frequency $f_1 = 1/T_u$, e.g., $f_m = m f_s/M$. This way the baseband spectrum would exploit only positive frequencies, and in order to be consistent with (5.1) where f_c is the central RF spectrum frequency, it would be necessary to use $f_m = m f_s/M - \frac{M-1}{2} \frac{f_s}{M}$. However, this choice implies only a fixed frequency shift of the spectrum, which is inessential. Thus, in the rest of the paper the frequency shift $\frac{M-1}{2} \frac{f_s}{M}$ is ignored, which can also be interpreted as assuming $f_c + \frac{M-1}{2} \frac{f_s}{M}$ as the RF central frequency.

Consequently, a multicarrier design that guarantees an easy sub-stream separation exploits

$$p_m(t) = \frac{1}{\sqrt{MT_s}} e^{j2\pi \frac{m}{MT_s} t} \text{rect}_{MT_s}(t - MT_s/2) \quad (5.23)$$

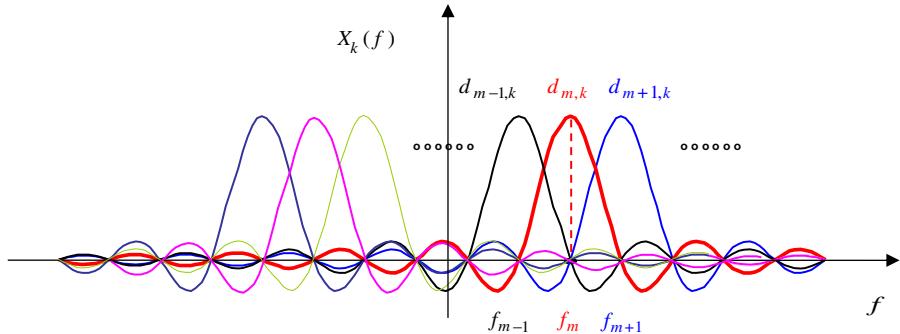
and the overall transmitted signal is expressed by

$$x(t) = \sum_{m=0}^{M-1} x^{(m)}(t) = \frac{1}{\sqrt{MT_s}} \sum_{k=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{MT_s} (t - kMT_s)} \text{rect}_{MT_s}(t - (k + 1/2)MT_s). \quad (5.24)$$

Equation (5.23) is analogous to (5.16) for FDM, where $p(t) = \frac{1}{\sqrt{MT_s}} \text{rect}_{MT_s}(t - MT_s/2)$: this fact highlights that the system designed with (5.23) and (5.24) is a special FDM system that exploits (baseband) orthogonal frequencies $f_m = m f_s/M$ and, as such, is typically called OFDM.

However, the philosophy of OFDM is noticeably different from FDM, as can be noted from the following observations:

- a. The ISI-free condition for each sub-stream is granted by a pulse shaper $p(t)$ with a finite time-support $T_u = MT_s$.
- b. The rectangular pulse shaper $p(t)$ induces, for each sub-stream, a baseband spectrum with $|P(f)| = |\sqrt{T_u} \text{sinc}[\pi T_u f]|$. Thus, each sub-stream is not band-limited, although a high percentage of its energy is mainly concentrated into its main lobe, i.e., on an RF frequency support equal to $2/T_u = 2 f_s/M$.
- c. The spectra $P^{(m)}(f) = P(f - f_m)$ of different sub-streams, separated by $f_m - f_j = (m - j)/T_u$, are not disjoint in the frequency-domain and they significantly overlap with the adjacent ones (see Figure 5.4).

**FIGURE 5.4**

Spectrum of a single OFDM block (k is the block index).

We now focus on the signal $x_k(t)$ associated with a block of M parallel symbols transmitted during a single OFDM interval of duration $T_u = M T_s$, expressed by

$$x_k(t) = \frac{1}{\sqrt{T_u}} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{T_u} t} \text{rect}_{T_u}(t - (k + 1/2)T_u). \quad (5.25)$$

The signal $x_k(t)$ in (5.25), or equivalently the data $\{d_{m,k}\}_{m=0,\dots,M-1}$ in (5.25), is usually denoted as the k th OFDM symbol or the k th OFDM block. The spectrum associated with the single OFDM block in (5.25) is

$$X_k(f) = \sqrt{T_u} \sum_{m=0}^{M-1} d_{m,k} \text{sinc} [\pi T_u (f - f_m)] e^{-j\phi_k(f - f_m)}, \quad (5.26)$$

where $\phi_k(f) = \pi(2k+1)T_u f$ is a linear-phase term that takes into account the delay associated with the k th OFDM block. Observing the shape of $X_k(f)$ in Figure 5.4 (where the phase $\phi_k(f) = \pi(2k+1)T_u f$ is not considered for simplicity), it is evident that the transmission of an OFDM symbol corresponds to the parallel transmission of M data symbols $\{d_{m,k}\}_{m=0,\dots,M-1}$, each one modulating the amplitude of a pulse-shaper spectrum $P(f) = \sqrt{T_u} \text{sinc} [\pi T_u f] e^{j\phi_k(f)}$. From the same figure, it is also evident that the middle points of the M overlapping spectra are equispaced in the frequency-domain, by $\Delta f = f_s/M$. Thus, also the data overlap (interfere) in the frequency-domain, except in the equispaced locations of the zeros of each sinc-like spectrum where

$$X_k(f_m) = \sqrt{T_u} d_{m,k}, \quad m = 0, \dots, M - 1. \quad (5.27)$$

Consequently, as Figure 5.4 clarifies, OFDM performs in the frequency-domain a dual operation of SC transmissions in the time-domain: the equispaced zero-location property of $|P(f)|^2$ at frequencies $f_m = mf'_s$ is the dual of the equispaced zero-location property of $R_{pp}(t)$ that grants ISI-free transmission at $t_k = kT_s$ in SC systems. In this case, $f_m = mf'_s$ are the points in the frequency-domain where the symbols $d_{m,k}$ belonging to different subcarriers do not interfere with each other. In the *multicarrier* literature

[29, 59, 94, 196, 246], this kind of interference is called intercarrier interference (ICI), or interchannel interference: OFDM is characterized by an ICI-free design for $P(f)$.

Figure 5.4 suggests that in order to recover the data $\{d_{m,k}\}_{m=0,\dots,M-1}$ transmitted during the k th OFDM block, it would be sufficient to sample its spectrum at the frequencies $\{f_m\}_{m=0,\dots,M-1}$. Indeed, the signal received in an AWGN channel during the k th OFDM block is

$$y_k(t) = x_k(t) + n_k(t) = \left[\frac{1}{\sqrt{T_u}} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{T_u} t} + n(t) \right] \text{rect}_{T_u}(t - T_u/2) \quad (5.28)$$

and the output $y_{m,k}$ of the m th receiver branch in (5.22) is expressed by

$$\begin{aligned} y_{m,k} &= \frac{1}{\sqrt{T_u}} \int_0^{T_u} y_k(t) e^{-j2\pi f_m t} dt = \frac{1}{\sqrt{T_u}} \int_{-\infty}^{+\infty} y_k(t) e^{-j2\pi f_m t} dt \\ &= \frac{1}{\sqrt{T_u}} Y_k(f_m) = \frac{1}{\sqrt{T_u}} [X_k(f_m) + N_k(f_m)] = d_{m,k} + n_{m,k}, \end{aligned} \quad (5.29)$$

which confirms the intuition that in the OFDM case the optimal parallel MF architecture in Figure 5.3 actually performs an equispaced (frequency-domain) sampling of the received spectrum.

2.05.4.4 DFT-based implementation of OFDM

By the Nyquist-Shannon sampling theorem [213], it is well known that an equispaced spectrum sampling of a band-limited signal can be accomplished by the DFT of its time-domain samples, collected with a sampling frequency $F_s \geq 2B$ in order to avoid aliasing. Actually, the spectrum $X(f)$ in Figure 5.4 is not rigorously band-limited, because of the everlasting tails of the sinc(\cdot) function associated to the rectangular pulse shaper $p(t)$ in (5.19). However, if the sampling frequency is chosen such that $F_s = p f_s = p M (f_s/M)$, where p is a positive integer, by the Nyquist sampling theorem the sampled signal

$$x_k^{(s)}(t) = \sum_{n=-\infty}^{+\infty} x_k(n/F_s) \delta(t - n/F_s) = \sum_{n=0}^{M-1} x_k(n/F_s) \delta(t - n/F_s), \quad (5.30)$$

has a spectrum that is the sum of the replicas of the original spectrum, as summarized by

$$\begin{aligned} X_k^{(s)}(f) &= \mathcal{F} \left\{ x_k^{(s)}(t) \right\} = F_s \sum_{l=-\infty}^{+\infty} X_k(f - lF_s) \\ &= F_s \sqrt{T_u} \sum_{l=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} \text{sinc} \left[\pi T_u \left(f - (m + lpM) \frac{f_s}{M} \right) \right] e^{-j\phi_k(f - f_m - lF_s)}. \end{aligned} \quad (5.31)$$

Due to the fact that also the spectral aliases are centered at integer multiples of the OFDM symbol frequency f_s/M , also (some of) the zeros of their sinc-like functions will coincide with the orthogonal frequencies $\{f_m = m f_s/M\}_{m=0,\dots,M-1}$. Thus, aliasing is null on the frequency grid $\{m/T_u\}_{m=0,\dots,M-1}$ where the spectrum is sampled in order to recover the transmitted data $\{d_{m,k}\}_{m=0,\dots,M-1}$. In formulas,

when $p = 1$, i.e., when the signal is sampled at the receiver side by $F_s = f_s$, and ignoring the presence of the noise, we have

$$\begin{aligned} X_k^{(s)}(f_m) &= \sqrt{\frac{M}{T_s}} \sum_{l=-\infty}^{+\infty} \sum_{m'=0}^{M-1} d_{m,k} \operatorname{sinc} [\pi(m - m' + lM)] e^{-j\varphi_k((m-m'+lM)f_s)} \\ &= F_s X_k(f_m) = \sqrt{\frac{M}{T_s}} d_{m,k}, \quad m = 0, \dots, M - 1. \end{aligned} \quad (5.32)$$

The last equation highlights that the receiver can recover the transmitted data $\{d_{m,k}\}_{m=0,\dots,M-1}$ by sampling the spectrum of the (discrete-time) received signal. In the absence of noise, sampling the spectrum of the discrete-time received signal corresponds to compute

$$\begin{aligned} d_{m,k} &= \sqrt{\frac{T_s}{M}} X_k^{(s)}(f_m) = \sqrt{\frac{T_s}{M}} \int_{-\infty}^{+\infty} x_k^{(s)}(t) e^{-j2\pi f_m t} dt \\ &= \sqrt{\frac{T_s}{M}} \sum_{n=0}^{M-1} x_k(nT_s) \int_{-\infty}^{+\infty} \delta(t - nT_s) e^{-j2\pi f_m t} dt \\ &= \sqrt{\frac{T_s}{M}} \sum_{n=0}^{M-1} x_{n,k} e^{-j2\pi f_m T_s n} = \sqrt{\frac{T_s}{M}} \sum_{n=0}^{M-1} x_{n,k} e^{-j\frac{2\pi}{M} mn}, \quad m = 0, \dots, M - 1, \end{aligned} \quad (5.33)$$

which, by the last equality [173], corresponds to compute the DFT of the discrete-time sequence $\{x_{n,k} = x_k(nT_s)\}_{n=0,\dots,M-1}$, scaled by $\sqrt{T_s/M}$.

Consequently, also the inverse is true by exploiting the properties of DFT, i.e., the transmitted discrete-time symbols $\{x_{n,k}\}_{n=0,\dots,M-1}$ can be obtained by the inverse DFT (IDFT) of the spectrum samples $X_{m,k} = f_s X_k(f_m) = \sqrt{\frac{M}{T_s}} d_{m,k}$, scaled by $\sqrt{M/T_s}$. This can be also easily verified by direct substitution in (5.25), which leads to

$$\begin{aligned} x_{n,k} &= x_k(nT_s) = \frac{1}{\sqrt{MT_s}} \sum_{n=0}^{M-1} d_{m,k} e^{j\frac{2\pi}{M} mn} \\ &= \sqrt{\frac{M}{T_s}} \left[\frac{1}{M} \sum_{n=0}^{M-1} d_{m,k} e^{j\frac{2\pi}{M} mn} \right], \quad n = 0, \dots, M - 1. \end{aligned} \quad (5.34)$$

Equations (5.33) and (5.34) highlight that an efficient digital implementation of OFDM is possible by exploiting (inverse) fast Fourier transform (FFT) algorithms for (inverse) DFT computations [173]. This observation was first recognized in [257] and it is at the base of the renewed interest in OFDM between 1980 and 1990, when the sampling frequencies of A/D and D/A converters, as well as the clock frequencies of integrated circuits (IC) and memories, have started to become compatible with a real-time implementation. It was indeed at the end of that decade that the first mass-market commercial OFDM standards were deployed, namely digital audio broadcasting (DAB) [71] and digital video broadcasting-terrestrial (DVB-T) [72].

2.05.5 OFDM in frequency-selective multipath channels

Note that, the architecture in Figure 5.5 has solved only the problem of efficiently separating M parallel streams in AWGN channels. In the presence of frequency-selective (multipath) channels, each of the M orthogonal single (sub-) carrier systems would suffer from the classical problem of ISI. However, the reduction of the symbol rate of each stream to f_s/M leads to consequent reduction, by a factor M , of the length of the discrete-time impulse response $h_m[n]$ associated with each of the M parallel systems. Thus, the OFDM system could be equalized by M parallel SC equalizers. As we already mentioned at the end of Section 2.05.1, the complexity grows linearly and exponentially with the channel length for time-domain linear and ML equalization, respectively [183]: thus we can immediately conclude that the equalization complexity of OFDM is significantly reduced with respect to a SC system with the same bandwidth. Furthermore, we will show that the smart use of time guards further reduces OFDM equalization, by a frequency-domain approach.

In order to identify how to recover the transmitted data also in the presence of a frequency-selective channel, let us investigate the expression for the received signal in case of a typical time-invariant multipath propagation, as shown in Figure 5.6 and summarized by

$$h(t) = h_c(t) * g(t) = \sum_{p=0}^P \alpha_p g(t - \tau_p), \quad (5.35)$$

where $h_c(t)$ has been defined in (5.9), and $g(t)$ captures the low-pass filtering included in the A/D and D/A converters, as well as the baseband (BB) equivalent of the band-pass filtering stages in the RF-to-BB and the BB-to-RF conversion.

To better understand the effect of the channel, and the potential resistance of OFDM to multipath propagation, first it is convenient to consider the spectrum of the received signal, which is given by

$$Y_k(f) = H(f)X_k(f) + W(f). \quad (5.36)$$

Focusing on a single OFDM symbol, the useful part of the spectrum (i.e., ignoring the noise) is expressed by

$$H(f)X_k(f) = T_u \sum_{m=0}^{M-1} d_{m,k} H(f) \text{sinc} [\pi T_u (f - f_m)] e^{-j\phi_k(f - f_m)}, \quad (5.37)$$

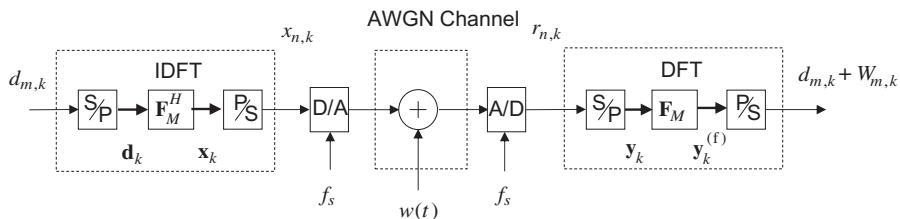
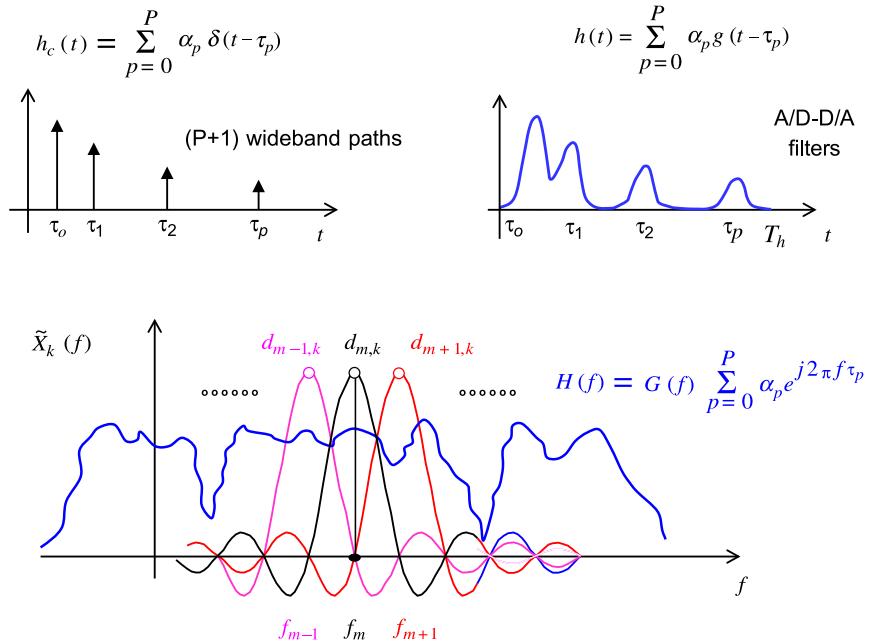


FIGURE 5.5

Discrete-time equivalent model of an OFDM system in AWGN (without CP).


FIGURE 5.6

Frequency selective channel induced by multipath propagation.

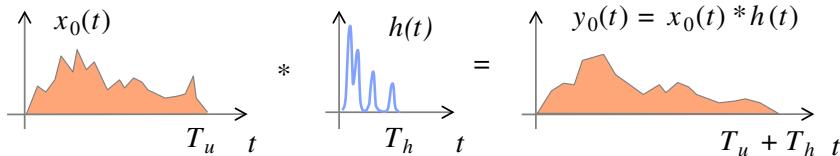
which clarifies that a frequency-selective channel $H(f)$ does not destroy the orthogonality of the transmitted data on the equispaced frequency grid $\{f_m = m f_s / M\}_{m=0, \dots, M-1}$, as shown in Figure 5.1, and summarized by

$$\frac{H(f_m)X_0(f_m)}{T_u} = H(f_m)d_{m,0}, \quad m = 0, \dots, M-1. \quad (5.38)$$

Thus, the receiver can effectively separate and detect the data transmitted during a single OFDM block, still by sampling the OFDM spectrum of the received signal $y_k(t) = x_k(t) * h(t) + n_k(t)$. However, in time-dispersive channels, because of the convolution with the channel impulse response (CIR) (5.35), the duration of each single received OFDM block will be longer than $T_u = M T_s$. If the CIR $h(t)$ has a time support (duration) T_h , the received OFDM block $y_k(t)$ will have duration $T_r = T_u + T_h$, as shown in Figure 5.7.

Thus, sampling the OFDM spectrum of the 0th block requests to compute

$$\begin{aligned} Y_0(f_m) &= T_s H(f_m) d_{m,0} + N_m \\ &= \int_{-\infty}^{+\infty} y_0(t) e^{-j2\pi f_m t} dt = \int_0^{T_u+T_h} y_0(t) e^{-j2\pi f_m t} dt, \end{aligned} \quad (5.39)$$

**FIGURE 5.7**

Effect of a time-dispersive channel on a single OFDM block.

which can be obtained by the DTFT of the sequence $y_{n,0} = y_0(nT_s)$, evaluated at $f = f_m$, as expressed by

$$Y_0(f_m) = \sum_{n=0}^{M+L-1} y_{n,0} e^{-j\frac{2\pi}{M}nm}. \quad (5.40)$$

Note that (5.40) is quite similar to a DFT processing: the only difference is that the summation index n exceeds the DFT period $[0, M - 1]$. However, because of the periodicity of the discrete Fourier exponentials, i.e., $e^{-j\frac{2\pi}{M}mn} = e^{-j\frac{2\pi}{M}m(n)_{\text{mod } M}}$, it is evident that

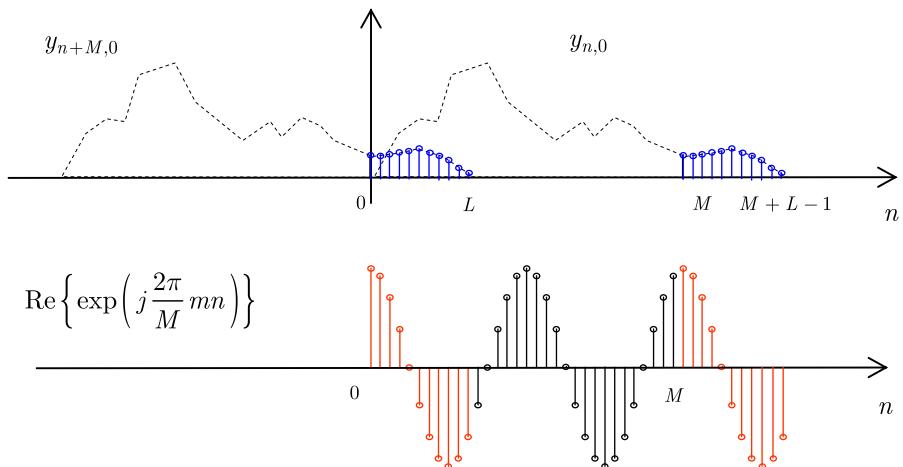
$$Y_0(f_m) = \sum_{n=0}^{M-1} [y_{n,0} + y_{n+M,0}] e^{-j\frac{2\pi}{M}nm}. \quad (5.41)$$

Equation (5.41) highlights that the sampling can still be performed by an M -point DFT processing, on a new M -length sequence $\tilde{y}_{n,0} = y_{n,0} + y_{n+M,0}$ obtained adding to the received sequence $y_{n,0}$ its replica anticipated by M discrete-time indexes, by an overlap-and-add strategy [173]. This way, the tail of $y_{n,0}$ that exceeds the transmitted length M is added at the beginning of the original sequence, as shown in Figure 5.8, making (5.41) equivalent to (5.40).

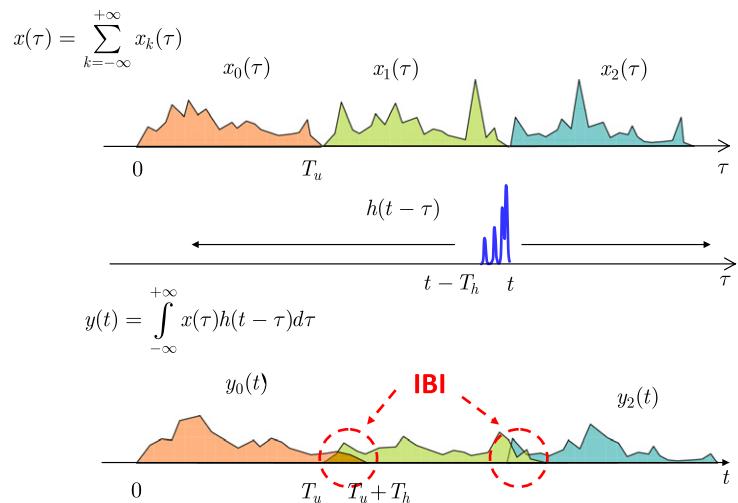
2.05.5.1 Zero-padded OFDM (ZP-OFDM)

Equation (5.38), and consequently (5.41), has shown a simple receiving strategy that preserves orthogonality among the subcarriers at the receiver side, also in the presence of a time-dispersive channel. This strategy, which has not been outlined in any of the historical and classical papers on OFDM [29, 59, 257], is an alternative to the CP-OFDM transmission that will be described shortly. However, note that (5.41) alone leads to a simple DFT-based receiver in frequency-selective channels, for a single OFDM block. Actually, the transmission of consecutive OFDM blocks through a time-dispersive channel would introduce ISI between successive blocks, typically called inter-block interference (IBI), as shown in Figure 5.9.

This means that (5.41) avoids the ISI between different subcarriers (i.e., the ICI), but each (sub-) SC signal $y^{(m)}(t)$ that constitutes the overall OFDM signal still suffers from the ISI between adjacent symbol periods of duration $T_u = MT_s$ (i.e., IBI). Anyway, by designing $T_u \gg T_h$, i.e., by employing a high number M of carriers, the IBI impairs only a small fraction of the symbol period T_u . Consequently, the performance degradation induced on each (sub-) SC system would be much lower than for a single SC system with the same data rate, whose symbol duration is $T_s = T_u/M$. Obviously, there exist methods to

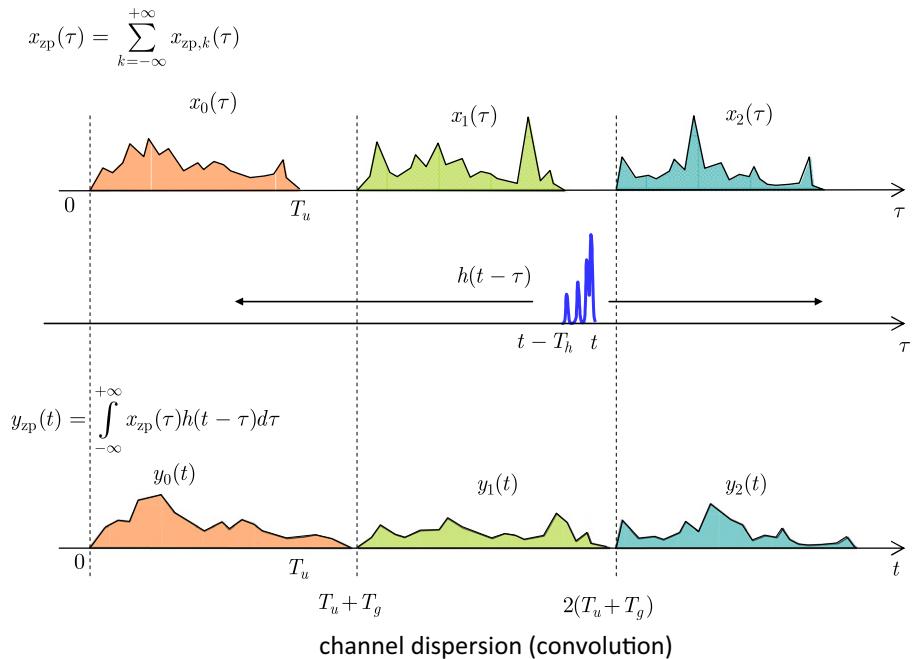
**FIGURE 5.8**

The cyclic processing induced by frequency-domain sampling.

**FIGURE 5.9**

IBI induced by a time-dispersive channel.

compensate for such IBI (i.e., equalization algorithms), as we will detail later on; however, if the goal is a very low-complexity receiver, it would be easy to deterministically avoid IBI by designing an OFDM system that inserts a silent period $T_g \geq T_h$ (i.e., a time-guard) between consecutive OFDM symbols.

**FIGURE 5.10**

ZP to avoid IBI.

This approach, shown in Figure 5.10, is summarized by

$$x_{zp}(t) = \sum_{m=0}^{M-1} x_{zp}^{(m)}(t) = \sum_{k=-\infty}^{+\infty} x_{zp,k}(t) = \sum_{k=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{T_u}(t-kT)} \text{rect}_{T_u}(t - (k + 1/2)T), \quad (5.42)$$

where $T = MT_s + T_g$, is the new duration of each OFDM block, including the time guard T_g filled with a signal equal to zero. As a consequence, a multicarrier system based on (5.42) is typically called zero-padded (ZP)-OFDM [251] or trailing-zeros (TZ) OFDM [166].

The insertion of a null symbol, with duration greater than the channel delay spread T_h , grants an IBI-free design and consequently the possibility to recover the data $\{d_{m,k}\}_{m=0, \dots, M-1}$ transmitted during each OFDM block, by the simple per-subcarrier receiver architecture subsumed by (5.41). However, note that this simple receiver is not the optimal one for ZP-OFDM from the BER performance point-of-view [251], as it will be clarified later on.

2.05.5.2 Cyclic-prefix OFDM (CP-OFDM)

A careful observation of the simple receiver for ZP-OFDM, induced by (5.41) and shown in Figure 5.8, suggests that the superposition to each *received* OFDM symbol of L samples from its

anticipated replica (tail) could be induced also at the transmitter side. Indeed, by some abuse of notation, it could be observed that

$$y_k(t + MT_s) = x_k(t + MT_s) * h(t) + w_k(t), \quad (5.43)$$

i.e., an anticipated replica of the signal received during the k th OFDM symbol can be obtained by transmitting, through the same channel, a replica of the original signal $x_k(t)$ anticipated by M samples, i.e., anticipated by the useful symbol period $T_u = MT_s$. Moreover, only the last $L + 1$ samples of $y_k[n + M]$ are necessary in (5.41) to enable the simple per-subcarrier receiver: therefore it is necessary to transmit only the last LT_s seconds of the anticipated signal $x(t + T_u)$. This strategy leads to the so called CP-OFDM, shown in Figure 5.11, which is nowadays the most used approach in wireless communications standards [71, 72, 104, 105, 237].

Thanks to the periodicity of the IDFT, mathematically the insertion of the CP is simply represented by a longer rectangular pulse shaper with duration $T = T_u + T_g$, which is anticipated by T_g , as summarized by

$$x_{\text{cp}}(t) = \sum_{m=0}^{M-1} x_{\text{cp}}^{(m)}(t) = \sum_{k=-\infty}^{+\infty} x_{\text{cp},k}(t) = \sum_{k=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{T_u}(t-kT)} \text{rect}_T(t + T_g - (k + 1/2)T). \quad (5.44)$$

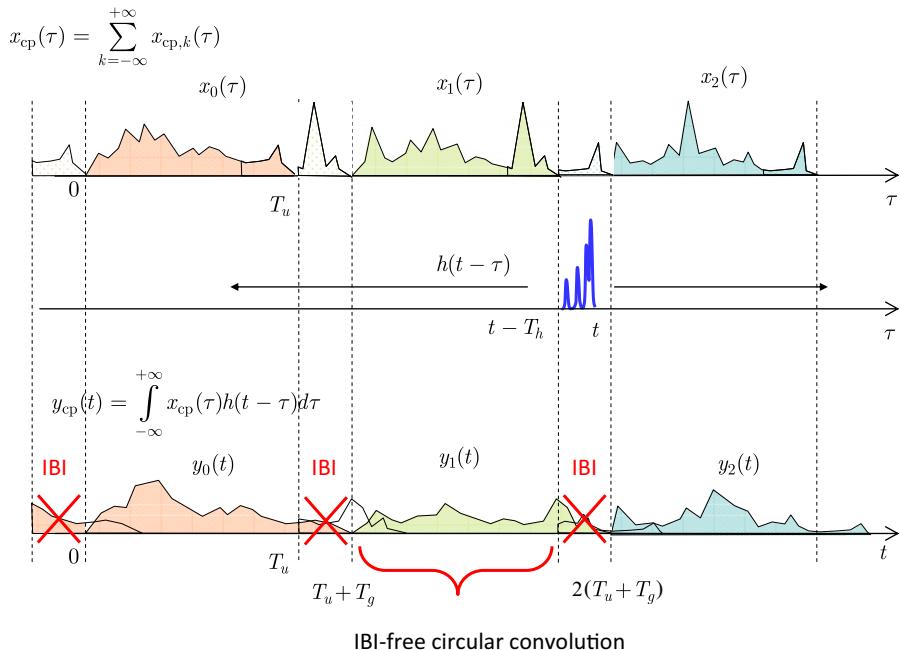


FIGURE 5.11

CP to avoid IBI and induce circular convolution.

Figure 5.11 shows that the time guard $T_g \geq T_h$ is exploited for the insertion of the CP, rather than to insert zeros as in ZP-OFDM. This CP insertion induces at the receiver side both the desired overlap during the useful time period T_u , and an undesired IBI during the time guards. However, the IBI can just be removed because the signal received during the useful period $T_u = MT_s$ has the same structure of Figure 5.8 and, consequently, grants the simple separation of the M data flows by frequency-domain sampling. This separation is obtained by DFT processing through

$$\begin{aligned} Y_k^{(\text{cp})}(f_m) &= T_s \sum_{n=0}^{M-1} y_{n,k}^{(\text{cp})} e^{-j\frac{2\pi}{M}nm} \\ &= T_s \sum_{n=0}^{M-1} [x_{n,k}^{(\text{cp})} * h_n + w_{n,k}] e^{-j\frac{2\pi}{M}nm} \\ &= MT_s H(f_m) d_{m,k} + W_{m,k}. \end{aligned} \quad (5.45)$$

Equation (5.45) highlights that, similarly to ZP-OFDM in (5.39), CP-OFDM performs a separation of the data in M orthogonal data-flows, where channel equalization reduces to a simple per-subcarrier equalization (PSE) of the flat channel $H(f_m)$. Noteworthy, CP-OFDM performs the orthogonalization by overlapping only the transmitted signal and not the noise $w_{n,k}$, as it conversely happens for ZP-OFDM in (5.41) and Figure 5.8. Thus, PSE of CP-OFDM has a noise performance advantage with respect to PSE of ZP-OFDM. However, it should be taken in mind that PSE is not the optimal equalization strategy for ZP-OFDM, as it will be clarified later on.

An easy interpretation of CP-OFDM is that the CP induces a circular convolution of the transmitted signal with the CIR, instead of the classical linear convolution. Thus, passing to the digital domain, and after removal of the time guards, the discrete-time received signal during the k th OFDM symbol is expressed by

$$y_{n,k} = h_n *_M x_{n,k} + w_{n,k}, \quad n = 0, \dots, M-1, \quad (5.46)$$

where $h_n = \frac{1}{T_s} h(nT_s)$ is the discrete-time equivalent of the overall CIR, and $*_M$ stands for the M -point discrete circular convolution operator [173]. By well-known DFT properties, the DFT of a time-domain circular convolution of two sequences corresponds to the product of their DFTs, as expressed by

$$\begin{aligned} Y_{m,k} &= H_m X_{m,k} + W_{m,k} \\ &= H_m d_{m,k} + W_{m,k}, \quad m = 0, \dots, M-1, \end{aligned} \quad (5.47)$$

which enables the easy and optimal PSE and detection [251]

$$\tilde{d}_{m,k} = \frac{Y_{m,k}}{H_m} = d_{m,k} + \frac{W_{m,k}}{H_m}, \quad m = 0, \dots, M-1, \quad (5.48)$$

$$\hat{d}_{m,k} = \arg \min_{d \in \mathcal{D}} \left\{ |d - \tilde{d}_{m,k}| \right\}, \quad m = 0, \dots, M-1. \quad (5.49)$$

The capability to convert any system affected by a frequency-selective channel into a set of M independent flat-fading channels by the insertion of a CP longer than the channel delay spread, is the main

motivation for the OFDM success. The main price to be paid is a waste of power efficiency, by a factor equal to $L_g/(L_g + M)$, due to the insertion of L_g time samples employed for the CP. Additionally, for both CP-OFDM and ZP-OFDM, the CP or ZP insertion to grant IBI-free transmission induces also a data-rate loss, again of a factor $L_g/(L_g + M)$.

2.05.6 A vector-matrix representation for OFDM

Figure 5.5 shows the discrete-time equivalent model of an OFDM system that exploiting a serial-to-parallel (S/P) converter, selects a block of M data symbols every MT_s seconds, then generates the M discrete-time samples to be transmitted by IDFT processing, and finally, after a parallel-to-serial (P/S) conversion, sends the M samples to a D/A converter. In a vector-matrix notation equivalent to (5.34), the OFDM symbol $\mathbf{x}_k = [x_{0,k}, \dots, x_{M-1,k}]^T$ transmitted during the k th period is obtained by

$$\mathbf{x}_k = \mathbf{F}_M^H \mathbf{d}_k, \quad (5.50)$$

where $\mathbf{d}_k = [d_{0,k}, \dots, d_{M-1,k}]^T$ is a zero-mean uncorrelated data vector with covariance $\mathbf{R}_d = \sigma_d^2 \mathbf{I}_M$, and $[\mathbf{F}_M]_{n+1,m+1} = \frac{1}{\sqrt{M}} \exp(-j \frac{2\pi}{M} mn)$ is the unitary DFT matrix of size M . For notational convenience, in (5.50) we have normalized the transmitted samples with respect to $\sqrt{T_s}$. In AWGN channels, the received vector $\mathbf{y}_k = [y_{0,k}, \dots, y_{M-1,k}]^T$ is simply expressed by

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{w}_k, \quad (5.51)$$

where $\mathbf{w}_k = [w_{0,k}, \dots, w_{M-1,k}]^T$, is an i.i.d. zero-mean Gaussian vector with covariance $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_M$, whose elements $w_{m,k} = w_{lp}((kM + m)T_s)$ are the samples of the noise $w_{lp}(t)$ generated by low-pass filtering of $w(t)$ in the A/D converter. Thus, the transmitted data \mathbf{d}_k can be recovered by a DFT processing that generates the soft estimate $\mathbf{y}_k^{(f)}$ of the data vector \mathbf{d}_k , as expressed by

$$\begin{aligned} \mathbf{y}_k^{(f)} &= \mathbf{F}_M \mathbf{y}_k = \mathbf{F}_M (\mathbf{F}_M^H \mathbf{d}_k + \mathbf{w}_k) \\ &= \mathbf{d}_k + \mathbf{w}_k^{(f)}, \end{aligned} \quad (5.52)$$

where $[\mathbf{y}_k^{(f)}]_m = Y_{m,k} = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} y_{n,k} e^{-j \frac{2\pi}{M} mn}$ is the m th sample of the unitary DFT of the received vector \mathbf{y}_k , and $\mathbf{w}_k^{(f)}$ is a vector containing the frequency-domain samples of the noise, which is still AWGN.

The presence of an LTI multipath channel induces a dispersion in time of the symbols transmitted during the k th block, as expressed by

$$r_{n,k} = \sum_{l=0}^L h_l s_{n-l,k}, \quad n = 0, \dots, M + L - 1, \quad (5.53)$$

where L is the CIR order, and $s_{n,k} = x_{n,k}$ for OFDM transmissions. In a vector-matrix notation, this time-domain spreading is captured by the $N \times (N - L)$ Toeplitz convolution matrix of the channel $\mathbf{H}_t^{(N)}$, which is expressed by

$$\mathbf{H}_t^{(N)} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & h_1 & \ddots & 0 \\ h_L & \vdots & \ddots & h_0 \\ 0 & h_L & \vdots & h_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & h_L \end{bmatrix}, \quad (5.54)$$

such that a transmitted vector \mathbf{s}_k of size M , generates a received vector $\mathbf{r}_k^{(N)}$, of size $N = M + L$, as expressed by

$$\mathbf{r}_k^{(N)} = \mathbf{H}_t^{(N)} \mathbf{s}_k. \quad (5.55)$$

However, because of the channel delay spread L , the received blocks associated to different OFDM blocks overlap in time, thereby generating IBI (see Figure 5.9). As detailed in Section 2.05.3, a simple way to contrast IBI is the introduction of a time guard between OFDM blocks, with a duration $L_g \geq L$, which can be exploited to insert either zeros as in ZP-OFDM, or the CP as in CP-OFDM. Thus, as shown in Figure 5.12, each transmitted block \mathbf{s}_k , of size $M + L_g$, is expressed by

$$\mathbf{s}_k = \mathbf{T} \mathbf{x}_k, \quad (5.56)$$

where \mathbf{x}_k is the M -size OFDM block before the insertion of the time-guard, $\mathbf{T} = \mathbf{T}_{\text{cp}} = [\mathbf{P}_{\text{cp}}^T \mathbf{I}_M]^T$ with $\mathbf{P}_{\text{cp}} = [\mathbf{0}_{L_g \times (M-L_g)} \mathbf{I}_{L_g}]$ is the $(M + L_g) \times M$ CP insertion matrix for CP-OFDM, and

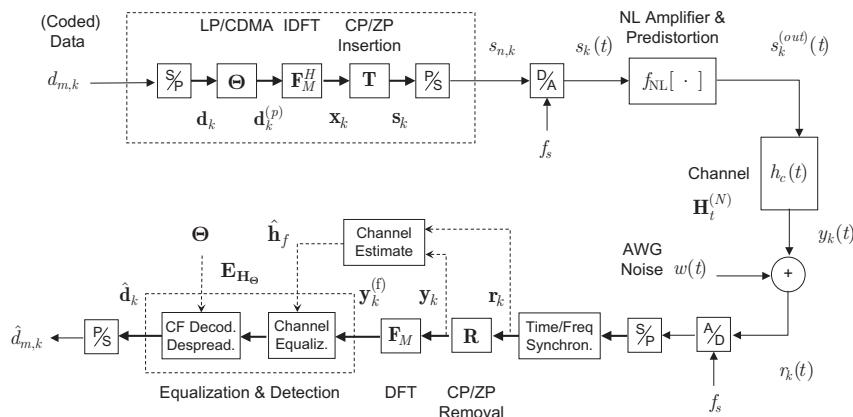


FIGURE 5.12

Baseband equivalent system model of a LP-OFDM/MC-CDMA system in a multipath fading channel.

$\mathbf{T} = \mathbf{T}_{\text{zp}} = [\mathbf{I}_M \ \mathbf{0}_{L_g}]^T$ is the $(M + L_g) \times M$ ZP insertion matrix for ZP-OFDM. Assuming for simplicity that $L_g = L$, and that the delay spread L of the discrete-time channel is shorter than the OFDM block length M , the received samples associated to the k th transmitted symbol can be collected in a vector \mathbf{r}_k of size $N + L = M + 2L$, as expressed by

$$\begin{aligned}\mathbf{r}_k &= \mathbf{S}_a \mathbf{r}_{k-1}^{(N+L)} + \mathbf{r}_k^{(N+L)} + \mathbf{S}_d \mathbf{r}_{k+1}^{(N+L)} + \tilde{\mathbf{w}}_k \\ &= \mathbf{S}_a \mathbf{H}_t^{(N+L)} \mathbf{s}_{k-1} + \mathbf{H}_t^{(N+L)} \mathbf{s}_k + \mathbf{S}_d \mathbf{H}_t^{(N+L)} \mathbf{s}_{k+1} + \tilde{\mathbf{w}}_k \\ &= \mathbf{S}_a \mathbf{H}_t^{(N+L)} \mathbf{T} \mathbf{x}_{k-1} + \mathbf{H}_t^{(N+L)} \mathbf{T} \mathbf{x}_k + \mathbf{S}_d \mathbf{H}_t^{(N+L)} \mathbf{T} \mathbf{x}_{k+1} + \tilde{\mathbf{w}}_k \\ &= \mathbf{H}_{-1} \mathbf{x}_{k-1} + \mathbf{H}_0 \mathbf{x}_k + \mathbf{H}_1 \mathbf{x}_{k+1} + \tilde{\mathbf{w}}_k.\end{aligned}\tag{5.57}$$

In (5.57) $\mathbf{S}_a = \begin{bmatrix} \mathbf{0}_{L \times N} & \mathbf{I}_L \\ \mathbf{0}_{N \times N+L} & \end{bmatrix}$ is the $(N + L) \times (N + L)$ matrix that shifts up a vector, inserting zeros in the empty positions, to mimic the time advance associated with the transmission of consecutive OFDM blocks, $\mathbf{S}_d = \begin{bmatrix} \mathbf{0}_{N \times N+L} \\ \mathbf{I}_L \quad \mathbf{0}_{L \times N} \end{bmatrix}$ is the $(N + L) \times (N + L)$ matrix that shifts down a vector, inserting zeros in the empty positions, to represent the time delay, $\mathbf{H}_{-1} = \mathbf{S}_a \mathbf{H}_t^{(N+L)} \mathbf{T}$, $\mathbf{H}_0 = \mathbf{H}_t^{(N+L)} \mathbf{T}$, $\mathbf{H}_1 = \mathbf{S}_d \mathbf{H}_t^{(N+L)} \mathbf{T}$, and $\tilde{\mathbf{w}}_k$ is the AWGN term.

2.05.6.1 Vector-matrix representation of CP-OFDM

In CP-OFDM, it is straightforward to derive that the structure of the three channel matrices $\mathbf{H}_{-1}^{(\text{cp})} = \mathbf{S}_a \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{cp}}$, $\mathbf{H}_0^{(\text{cp})} = \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{cp}}$, $\mathbf{H}_1^{(\text{cp})} = \mathbf{S}_d \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{cp}}$, is expressed by (5.58). From (5.57) and (5.58), it is clear that the vector \mathbf{r}_k , of size $N + L = M + 2L$, suffers from IBI caused by the previous (successive) block only in the first (last) L samples of \mathbf{r}_k (see also Figure 5.11). Thus, a simple way to remove the IBI is to select only the central $M = N - L$ samples of \mathbf{r}_k : this corresponds to remove the CP and to ignore the last L samples induced by the channel delay spread. Note that the last L samples of \mathbf{r}_k overlap with the CP of \mathbf{r}_{k+1} , which has to be removed as well, because we adopted a receiving window (of length $N + L$) that overlaps for adjacent blocks.

$$\begin{aligned}\mathbf{H}_0^{(\text{cp})} &= \left[\begin{array}{cccccc} 0 & \cdots & 0 & h_0 & 0 & 0 \\ \vdots & & \vdots & \ddots & 0 & \\ 0 & \cdots & 0 & h_{L-1} & \cdots & h_0 \\ \hline h_0 & 0 & \cdots & h_L & \cdots & h_1 \\ \vdots & h_0 & 0 & \ddots & \vdots & \\ \vdots & \vdots & \ddots & \ddots & & h_L \\ h_L & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & h_L & \vdots & \vdots & \ddots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots & h_0 \\ \hline 0 & 0 & \ddots & \vdots & \vdots & \vdots \\ \vdots & 0 & \ddots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & h_L \end{array} \right] \quad \mathbf{H}_{-1}^{(\text{cp})} = \left[\begin{array}{cccccc} 0 & \cdots & 0 & h_L & h_{L-1} & \cdots & h_1 \\ \vdots & & \vdots & \ddots & h_L & \ddots & \vdots \\ \vdots & & \vdots & \ddots & \ddots & \ddots & h_{L-1} \\ \hline 0 & \cdots & \cdots & \cdots & \cdots & 0 & h_L \\ \hline \mathbf{0}_{N \times M} \end{array} \right] \\ \mathbf{H}_{+1}^{(\text{cp})} &= \left[\begin{array}{cccccc} \mathbf{0}_{N \times M} \\ \hline 0 & \cdots & 0 & h_0 & 0 & \cdots & 0 \\ \vdots & & \vdots & h_1 & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \ddots & \ddots & h_0 & 0 \\ \hline 0 & \cdots & 0 & h_{L-1} & \cdots & h_1 & h_0 \end{array} \right]\end{aligned}\tag{5.58}$$

Thus, the received vector typically used for the detection of a CP-OFDM symbol is expressed by

$$\begin{aligned}\mathbf{y}_k^{(\text{cp})} &= [\mathbf{R}_{\text{cp}} \mathbf{0}_{M \times L}] \mathbf{r}_k \\ &= [\mathbf{R}_{\text{cp}} \mathbf{0}_{M \times L}] \mathbf{H}_0^{(\text{cp})} \mathbf{x}_k + [\mathbf{R}_{\text{cp}} \mathbf{0}_{M \times L}] \mathbf{w}_k,\end{aligned}\quad (5.59)$$

where $\mathbf{R}_{\text{cp}} = [\mathbf{0}_{M \times L} \mathbf{I}_M]$ is the CP removal matrix defined in [251] for non-overlapping receiving windows. Summarizing, the overall effect of inserting a CP by \mathbf{T}_{cp} , removing it by \mathbf{R}_{cp} , and discarding the last L samples, leads to an overall observation equation equal to

$$\begin{aligned}\mathbf{y}_k^{(\text{cp})} &= \mathbf{H}_t^{(\text{cp})} \mathbf{x}_k + \mathbf{w}_k \\ &= \mathbf{H}_t^{(\text{cp})} \mathbf{F}_M^H \mathbf{d}_k + \mathbf{w}_k,\end{aligned}\quad (5.60)$$

where $\mathbf{H}_t^{(\text{cp})}$ corresponds to the M central rows of $\mathbf{H}_0^{(\text{cp})}$ in (5.58), as expressed by

$$\begin{aligned}\mathbf{H}_t^{(\text{cp})} &= [\mathbf{R}_{\text{cp}} \mathbf{0}_{N \times L}] \mathbf{H}_0^{(\text{cp})} \\ &= [\mathbf{R}_{\text{cp}} \mathbf{0}_{N \times L}] \mathbf{H}_t^{(N)} \mathbf{T}_{\text{cp}} = \mathbf{R}_{\text{cp}} \mathbf{H}_t^{(M)} \mathbf{T}_{\text{cp}}.\end{aligned}\quad (5.61)$$

The matrix $\mathbf{H}_t^{(\text{cp})}$ is circulant, its first column contains the CIR $\mathbf{h}_t = [h_0, h_1, \dots, h_L, 0, \dots, 0]^T$, and the last equality in (5.61) highlights the equivalence with the notation in [251]. A circulant matrix can be always diagonalized by DFT matrices [89], as expressed by

$$\mathbf{H}_t^{(\text{cp})} = \mathbf{F}_M^H \text{diag}(\mathbf{h}_f) \mathbf{F}_M, \quad (5.62)$$

where $\mathbf{h}_f = \sqrt{M} \mathbf{F}_M \mathbf{h}_t$ is the frequency transfer function of the discrete-time channel, whose $(m+1)$ th element is expressed by (see Figure 5.6)

$$H_m = [\mathbf{h}_f]_m = \sum_{n=0}^{M-1} h_n e^{-j \frac{2\pi}{M} mn} = H(f_m), \quad (5.63)$$

for perfectly band-limited channel models $H(f)$.

Consequently, taking into account (5.50)

$$\mathbf{y}_k^{(\text{cp})} = \mathbf{F}_M^H \text{diag}(\mathbf{h}_f) \mathbf{d}_k + \mathbf{w}_k, \quad (5.64)$$

and applying a DFT on the received vector after the CP removal, we obtain

$$\mathbf{y}_k^{(f)} = \mathbf{F}_M \mathbf{y}_k^{(\text{cp})} = \text{diag}(\mathbf{h}_f) \mathbf{d}_k + \mathbf{w}_k^{(f)}. \quad (5.65)$$

Equation (5.65) highlights that the data transmitted on different subcarriers do not interfere with each other, i.e., there is no ICI, and that the transmitted data are simply scaled by the diagonal frequency-domain channel matrix $\mathbf{H}_f^{(\text{cp})} = \text{diag}(\mathbf{h}_f)$. Thus, for a Gaussian white noise $\mathbf{w}_k^{(f)} = \mathbf{F}_M \mathbf{w}_k$, the optimal

reception based on (5.65) is equivalent to the ML reception based on (5.60), as expressed by [117]

$$\begin{aligned}\hat{\mathbf{d}}_k &= \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left\| \mathbf{y}_k^{(\text{cp})} - \mathbf{H}_t^{(\text{cp})} \mathbf{F}_M^H \mathbf{d} \right\|^2 \right] \\ &= \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left\| \mathbf{y}_k^{(f)} - \text{diag}(\mathbf{h}_f) \mathbf{d} \right\|^2 \right].\end{aligned}\quad (5.66)$$

Since the frequency-domain channel matrix is diagonal, (5.66) can be separated in M independent ML decisions

$$\begin{aligned}\hat{d}_{m,k} &= \left[\hat{\mathbf{d}}_k \right]_m = \arg \min_{d \in \mathcal{D}} \left[|y_{k,m} - H_m d| \right] \\ &= \arg \min_{d \in \mathcal{D}} \left[|y_{k,m}/H_m - d| \right].\end{aligned}\quad (5.67)$$

Thus, the ML detector for a CP-OFDM system (that discards the CP) is a per-subcarrier zero-forcing equalizer, followed by a threshold device that selects from the alphabet \mathcal{D} the constellation symbol that is closest to the equalizer output $y_{k,m}/H_m$.

2.05.6.2 A note on the CP philosophy

A careful observation of (5.60) and (5.62) reveals that the easy equalization of OFDM is granted by the CP insertion and removal (which produce a circulant channel convolution) rather than by the use of orthogonal subcarriers. Thus, also SC communications could exploit this easy equalization property, by inserting a CP of length L every M symbols. This type of block transmission is typically called CP-SC [199, 251]. For CP-SC systems, the k th transmitted block is

$$\mathbf{s}_{\text{sc},k}^{(\text{cp})} = \mathbf{T}_{\text{cp}} \mathbf{d}_k, \quad (5.68)$$

and, after CP removal by \mathbf{R}_{cp} and DFT processing, the observation vector in the frequency-domain at the receiver side is

$$\begin{aligned}\mathbf{y}_{\text{sc},k}^{(f)} &= \mathbf{F}_M \mathbf{R}_{\text{cp}} \mathbf{H}_t^{(M)} \mathbf{T}_{\text{cp}} \mathbf{d}_k + \mathbf{w}_k^{(f)} \\ &= \mathbf{F}_M \mathbf{H}_t^{(\text{cp})} \mathbf{d}_k + \mathbf{w}_k^{(f)} \\ &= \text{diag}(\mathbf{h}_f) \mathbf{F}_M \mathbf{d}_k + \mathbf{w}_k^{(f)}.\end{aligned}\quad (5.69)$$

Equation (5.69) also allows for an easy non-ML detection, which consists in a diagonal equalization in the frequency-domain (M complex divisions), followed by an IDFT processing and by M scalar data detections, as expressed by

$$\hat{d}_{m,k} = \arg \max_{d \in \mathcal{D}} \left[\left[\tilde{\mathbf{d}}_k \right]_m - d \right], \quad m = 0, \dots, M-1, \quad (5.70)$$

where

$$\begin{aligned}\tilde{\mathbf{d}}_k &= \mathbf{F}_M^H \text{diag}^{-1}(\mathbf{h}_f) \mathbf{y}_{\text{sc},k}^{(f)} \\ &= \mathbf{d}_k + \mathbf{F}_M^H \text{diag}^{-1}(\mathbf{h}_f) \mathbf{F}_M \mathbf{w}_k \\ &= \mathbf{d}_k + M^{-1/2} \text{circ}(\mathbf{F}_M^H \mathbf{h}_f^{-1}) \mathbf{w}_k,\end{aligned}\quad (5.71)$$

where $\text{circ}(\mathbf{x})$ is the circulant matrix with \mathbf{x} as first column. Note that (5.70) is not the joint ML detector of \mathbf{d}_k in (5.69), and that the least-squares (LS) estimator (5.71) is not the linear minimum variance estimator for \mathbf{d}_k in (5.69). As firstly suggested in [199], and subsequently analyzed in [251, 253, 254], the CP-SC transmission is the equivalent of the CP-OFDM transmission of Figure 5.5, where also the IDFT processing is moved at the receiver side. Noteworthy, SC block transmission could induce a circulant convolution, and exploit diagonal frequency-domain equalization, also by substituting the CP with a known symbol [46] that can be further exploited for channel estimation and synchronization purposes.

2.05.6.3 Vector-matrix representation of ZP-OFDM

For ZP-OFDM, the three channel matrices $\mathbf{H}_{-1}^{(\text{zp})} = \mathbf{S}_a \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{zp}}$, $\mathbf{H}_0^{(\text{zp})} = \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{zp}}$ and $\mathbf{H}_1^{(\text{zp})} = \mathbf{S}_d \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{zp}}$ in (5.57) are expressed by (5.72). Thanks to the insertion of zeros among the blocks, the received vector \mathbf{r}_k does not contain any IBI (see also Figure 5.10). The useful information about the data \mathbf{x}_k transmitted during the k th block is contained in the first N samples.

$$\mathbf{H}_0^{(\text{zp})} = \begin{bmatrix} h_0 & 0 & \cdots & \cdots & 0 \\ \vdots & h_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ h_L & \vdots & \vdots & \ddots & 0 \\ 0 & h_L & \vdots & \vdots & h_0 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & h_L \\ \hline & \mathbf{0}_{L \times M} \end{bmatrix} \quad \mathbf{H}_{-1}^{(\text{zp})} = \mathbf{0}_{N+L \times M}$$

$$\mathbf{H}_{+1}^{(\text{zp})} = \begin{bmatrix} & & & \mathbf{0}_{N \times M} & & & \\ h_0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ h_1 & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & h_0 & 0 & \vdots & \ddots & \vdots \\ h_{L-1} & \cdots & h_1 & h_0 & 0 & \cdots & 0 \end{bmatrix} \quad (5.72)$$

Therefore, in order to recover the data, it is sufficient to employ a receiving window of size $N = M + L$ to collect

$$\begin{aligned} \mathbf{y}_k^{(\text{zp})} &= [\mathbf{R}_{\text{zp}} \mathbf{0}_{N \times L}] \mathbf{r}_k \\ &= [\mathbf{R}_{\text{zp}} \mathbf{0}_{N \times L}] \mathbf{H}_0^{(\text{zp})} \mathbf{x}_k + [\mathbf{R}_{\text{zp}} \mathbf{0}_{N \times L}] \tilde{\mathbf{w}}_k \\ &= [\mathbf{R}_{\text{zp}} \mathbf{0}_{N \times L}] \mathbf{H}_t^{(N+L)} \mathbf{T}_{\text{zp}} \mathbf{x}_k + \mathbf{w}_k \\ &= \mathbf{R}_{\text{zp}} \tilde{\mathbf{H}}_t^{(N)} \mathbf{T}_{\text{zp}} \mathbf{x}_k + \mathbf{w}_k \\ &= \mathbf{H}_t^{(\text{zp})} \mathbf{x}_k + \mathbf{w}_k, \end{aligned} \quad (5.73)$$

where the receiving matrix for ZP-OFDM is simply expressed by $\mathbf{R}_{\text{zp}} = \mathbf{I}_N$, and $\tilde{\mathbf{H}}_t^{(N)}$ contains the first N rows of $\mathbf{H}_t^{(N+L)}$. In (5.73), the last-but-one equality highlights the equivalence with the notation used in [251]. Summarizing, the ZP-OFDM transmission through a time-dispersive channel is affected by the banded Toeplitz channel matrix $\mathbf{H}_t^{(\text{zp})}$, contained in the first N rows of $\mathbf{H}_0^{(\text{zp})}$ in (5.72), which has always

full rank. In this case, the channel matrix is not circulant, and the received vector has a size $N = M + L$. Thus, a receiver that employs a pure DFT operation should use a DFT matrix \mathbf{F}_{M+L}^H of size $M + L$, which would correspond to project the received vector on a set of discrete frequencies different from those used at the transmitter side, which does not make any sense. The intuition would suggest to use, at the receiver side, the same discrete frequencies $\mathbf{f}_m = M^{-1/2} \left[1, \dots, e^{-j\frac{2\pi}{M}mn}, \dots, e^{-j\frac{2\pi}{M}m(M-1)} \right]^T$ used at the transmitter (which are periodic), by extending them on the longer time-support $N = M + L$. This corresponds to use a set of basis functions contained in the columns of the matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{F}_M & \mathbf{F}_M^{(1:L)} \end{bmatrix}^H, \quad (5.74)$$

where $\mathbf{F}_M^{(1:L)} = \mathbf{F}_M [\mathbf{I}_L \ \mathbf{0}_{L \times M-L}]^T$ is the $M \times L$ matrix containing the first L columns of \mathbf{F}_M , to extend each basis vector. Thus, projecting the received vector onto the extended basis, we obtain

$$\begin{aligned} \tilde{\mathbf{y}}_k^{(f)} &= \mathbf{B}^H \mathbf{y}_k^{(\text{zp})} = [\mathbf{F}_M \ \mathbf{F}_M [\mathbf{I}_L \ \mathbf{0}_{L \times M-L}]^T] \mathbf{y}_k^{(\text{zp})} \\ &= \mathbf{F}_M \begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_M & \mathbf{0}_{M-L \times L} \end{bmatrix} \mathbf{y}_k^{(\text{zp})} = \mathbf{F}_M \left([\mathbf{I}_M \ \mathbf{0}_{M \times L}] + \begin{bmatrix} \mathbf{I}_L \\ \mathbf{0}_{M \times M} & \mathbf{0}_{M-L \times L} \end{bmatrix} \right) \mathbf{y}_k^{(\text{zp})} \\ &= \mathbf{F}_M (\mathbf{y}_k^{(\text{s})} + \mathbf{y}_k^{(\text{a})}), \end{aligned} \quad (5.75)$$

where $\mathbf{y}_k^{(\text{s})} = [\mathbf{I}_M \ \mathbf{0}_{M \times L}] \mathbf{y}_k^{(\text{zp})}$ is a shortened version of $\mathbf{y}_k^{(\text{zp})}$ to its first M values, and $\mathbf{y}_k^{(\text{a})}$ is an M vector containing an anticipated version of $\mathbf{y}_k^{(\text{zp})}$, with its last L samples in the first L positions, and zeros elsewhere. Thus, (5.75) is the vector-matrix equivalent of (5.40) and (5.41), which enabled in ZP-OFDM the simple PSE and decoding of the data transmitted on each separate subcarrier \mathbf{f}_m . Indeed, plugging (5.73) in (5.75), we obtain

$$\begin{aligned} \tilde{\mathbf{y}}_k^{(f)} &= \mathbf{F}_M \begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_M & \mathbf{0}_{M-L \times L} \end{bmatrix} \tilde{\mathbf{H}}_t^{(N)} \mathbf{T}_{\text{zp}} \mathbf{x}_k + \mathbf{F}_M (\mathbf{w}_k + \mathbf{w}_k^{(\text{a})}) \\ &= \mathbf{F}_M \begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_M & \mathbf{0}_{M-L \times L} \end{bmatrix} \tilde{\mathbf{H}}_t^{(N)} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \mathbf{x}_k + \mathbf{F}_M (\mathbf{w}_k + \mathbf{w}_k^{(\text{a})}). \end{aligned} \quad (5.76)$$

By direct substitution, it is straightforward to prove that

$$\begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_M & \mathbf{0}_{M-L \times L} \end{bmatrix} \tilde{\mathbf{H}}_t^{(N)} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} = \mathbf{H}_t^{(\text{cp})}, \quad (5.77)$$

and consequently

$$\begin{aligned}\tilde{\mathbf{y}}_k^{(f)} &= \mathbf{F}_M \mathbf{H}_t^{(\text{cp})} \mathbf{F}_M^H \mathbf{d}_k + \mathbf{F}_M \left(\mathbf{w}_k + \mathbf{w}_k^{(\text{a})} \right) \\ &= \text{diag}(\mathbf{h}_f) \mathbf{d}_k + \mathbf{w}_k^{(f)} + \mathbf{e}_k^{(\text{zp})} \\ &= \text{diag}(\mathbf{h}_f) \mathbf{d}_k + \mathbf{w}_k^{(\text{zp})}.\end{aligned}\quad (5.78)$$

Note that, except for the extra noise term $\mathbf{e}_k^{(\text{zp})}$ induced by the anticipated replica $\mathbf{y}_k^{(\text{a})}$, (5.78) is equivalent to the receiving Eq. (5.65) for CP-OFDM, and consequently simple PSE is possible also in this case. However, PSE and detection by (5.67) would not result in the ML estimator of \mathbf{d}_k in (5.78), because $\mathbf{e}_k^{(\text{zp})}$ is characterized by a (non-diagonal) circulant correlation matrix, expressed by

$$\begin{aligned}\mathbf{R}_e &= \sigma_w^2 \mathbf{F}_M \text{diag} \left([\mathbf{1}_L \ \mathbf{0}_{1 \times M-L}] \right) \mathbf{F}_M^H \\ &= \sigma_w^2 M^{-1/2} \text{circ} \left(\mathbf{F}_M [\mathbf{1}_L \ \mathbf{0}_{1 \times M-L}] \right) = \sigma_w^2 M^{-1/2} \text{circ} \left(\sum_{m=0}^{L-1} \mathbf{f}_m \right),\end{aligned}$$

which introduces a color on the overall noise $\mathbf{w}_k^{(\text{zp})}$ [117]. Indeed, in this case, the ML estimator of \mathbf{d}_k in (5.78), with a perfect channel knowledge, would request to compute

$$\begin{aligned}\hat{\mathbf{d}}_k &= \arg \max_{\mathbf{d} \in \mathcal{D}^M} \left[f_{\tilde{\mathbf{y}}_k^{(f)}} \left(\tilde{\mathbf{y}}_k^{(f)}; \mathbf{d} \right) \right] \\ &= \arg \max_{\mathbf{d} \in \mathcal{D}^M} \left[f_{\mathbf{w}_k^{(\text{zp})}} \left(\tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right) \right] \\ &= \arg \max_{\mathbf{d} \in \mathcal{D}^M} \left[\exp \left(\left(\tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right)^H \mathbf{R}_{\mathbf{w}_k^{(\text{zp})}} \left(\tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right) \right) \right] \\ &= \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left(\tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right)^H \mathbf{R}_{\mathbf{w}_k^{(\text{zp})}} \left(\tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right) \right]\end{aligned}\quad (5.79)$$

that reduces to

$$\hat{\mathbf{d}}_k = \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left\| \tilde{\mathbf{y}}_k^{(f)} - \mathbf{H}_f^{(\text{cp})} \mathbf{d} \right\|^2 \right] \quad (5.80)$$

only if the noise is white and uncorrelated, that is if $\mathbf{R}_{\mathbf{w}_k^{(\text{zp})}}$ is a scaled identity matrix, which is not this case since $\mathbf{R}_{\mathbf{w}_k^{(\text{zp})}} = \sigma_w^2 \mathbf{I}_M + \mathbf{R}_e$.

2.05.7 Symbol error rate (SER) performance analysis

The SER analysis for ML detection of CP-OFDM is quite simple, because of the frequency-domain channel diagonalization induced by the CP. Indeed, (5.65) and (5.45) show that the data transmitted on each separate subcarrier are impaired by a flat-fading channel, as expressed by

$$y_{m,k} = H_m d_{m,k} + W_{m,k}. \quad (5.81)$$

If the discrete-time channel taps $\{h_l\}_{l=0,\dots,L}$ are zero-mean Gaussian distributed, then it is straightforward to verify that H_m in (5.63) is also zero-mean Gaussian, $|H_m|$ is Rayleigh distributed with a probability density function (pdf) expressed by $f_{|H_m|}(|H_m|) = (|H_m|/\alpha_m^2) e^{-|H_m|^2/2\alpha_m^2}$, with $|H_m| > 0$, and

$$2\alpha_m^2 = E \left\{ |H_m|^2 \right\} = \sum_{l=0}^L \sum_{i=0}^L [\mathbf{R}_{\mathbf{h}_t}]_{l+1,i+1} e^{j \frac{2\pi}{M}(l-i)m}, \quad (5.82)$$

where $\mathbf{R}_{\mathbf{h}_t} = E \{ \mathbf{h}_t \mathbf{h}_t^H \}$ is the autocorrelation matrix of the discrete-time channel, and $W_{m,k}$ is the AWGN in the frequency-domain. Thus, the symbol error rate (SER) performance on a frequency-flat Rayleigh fading channel for each subcarrier is expressed by [183]

$$\text{SER}_m = P\{\hat{d}_{k,m} \neq d_{d,m}\} = \int_0^{+\infty} \alpha_Q Q \left(\sqrt{\beta_Q \frac{|H_m|^2 \sigma_d^2}{\sigma_w^2}} \right) f_{|H_m|}(|H_m|) d |H_m|, \quad (5.83)$$

where α_Q and β_Q depend on the complex alphabet \mathcal{D} of $d_{m,k}$. Defining the average SNR for each subcarrier as $\text{SNR}_m = E \{ |H_m|^2 \} \sigma_d^2 / \sigma_w^2$, if $d_{m,k}$ is drawn from a QPSK alphabet, (5.83) leads to [183]

$$\text{SER}_m = \frac{1}{2} \left(1 - \sqrt{\frac{\text{SNR}_m}{2 + \text{SNR}_m}} \right), \quad (5.84)$$

and the total SER is the average of (5.84) over all the subcarriers, as expressed by

$$\text{SER} = \frac{1}{M} \sum_{m=0}^{M-1} \text{SER}_m. \quad (5.85)$$

Equation (5.84), together with (5.82), highlights that the SER performance depends on the sum of the second order moments of the time-domain channel paths rather than on the number L of the paths. For instance, in the simpler case when the channel paths are independents (i.e., $[\mathbf{R}_{\mathbf{h}_t}]_{l,j} = \sigma_l^2 \delta[l-j]$), (5.82) becomes

$$E \left\{ |H_m|^2 \right\} = \sum_{l=0}^L \sigma_l^2, \quad (5.86)$$

which highlights the two following observations:

- i. SNR_m , and consequently SER_m , is the same for all the subcarriers, and $\text{SER} = \text{SER}_m$.
- ii. The SER depends only on the total power $P_h = \sum_{l=0}^L \sigma_l^2$ of the multipath channel and does not depend on the power distribution among the different paths (i.e., does not depend on the power-delay profile of the channel). Thus, the performance of CP-OFDM in a multipath Rayleigh fading channel would be the same performance in a single-path Rayleigh fading channel, where all the power is concentrated in the single path.

The last observation indicates that an uncoded CP-OFDM system is not capable to exploit the potential diversity offered by a multipath fading channel. Indeed, while it is highly improbable that all the L time-domain paths fade to zero simultaneously, the single frequency-domain path associated to each data in (5.81) has a greater probability to fade towards zero, making the data $d_{m,k}$ unrecoverable also in the presence of a low noise value $W_{m,k}$. Actually, each single path of the CIR is a potential source of *diversity*,² because the receiver collects multiple (independent) copies of the transmitted signal. This is similar to the space diversity offered by multi-antenna receiving systems [183]. However, the CP insertion and removal, which leads to the frequency-domain channel diagonalization, destroys the multipath diversity and sacrifices uncoded SER performance for a simpler per-subcarrier equalization. Such loss of diversity for uncoded SER performance can be recovered by channel coding (see Section 2.05.6), or by linear precoding (see Section 2.05.7), which spreads each uncoded data over all (or several) different subcarriers [167]. The SER analysis for ML decoding of ZP-OFDM is not as simple as for CP-OFDM, because the ML detection based on (5.73) leads to

$$\hat{\mathbf{d}}_k = \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left\| \mathbf{y}_k^{(zp)} - \mathbf{H}_t^{(zp)} \mathbf{F}_M^H \mathbf{d} \right\|^2 \right], \quad (5.87)$$

which does not have an equivalent (diagonal) formulation in the frequency-domain, because of the banded Toeplitz structure of $\mathbf{H}_t^{(zp)}$. Consequently, differently from CP-OFDM, the ML detector for ZP-OFDM is not based on PSE. SER and diversity analysis for ML problems like in (5.87) are typically addressed by the probability that a given transmitted data vector $\mathbf{d}_a \in \mathcal{D}^M$ is confused at the receiver side with another possible transmitted vector $\mathbf{d}_b \in \mathcal{D}^M$. This leads to the so-called pairwise-error probability (PEP) [218, 219, 229, 236], which is defined as

$$\text{PEP}_{a \rightarrow b} = P\{\hat{\mathbf{d}}_k = \mathbf{d}_b | \mathbf{d}_k = \mathbf{d}_a\}, \quad (5.88)$$

and dominates the SER. As detailed in [250], the minimum PEP for all the possible couples $\{\mathbf{d}_a, \mathbf{d}_b\}$ is granted if the observation matrix $\Theta_{\mathbf{H}}^{(zp)} = \mathbf{H}_t^{(zp)} \mathbf{F}_M^H$ has a full column rank for any channel realization. This is equivalent to state that, in the absence of noise, is not possible to confuse \mathbf{d}_a with \mathbf{d}_b at the receiver, or equivalently, to the symbol detectability condition [251]

$$\Theta_{\mathbf{H}}^{(zp)}(\mathbf{d}_a - \mathbf{d}_b) \neq \mathbf{0}_{M \times 1}, \quad \forall \mathbf{h}_t, \quad \forall \mathbf{d}_a \neq \mathbf{d}_b. \quad (5.89)$$

Actually, (5.89) is granted because both \mathbf{F}_M^H and $\mathbf{H}_t^{(zp)}$ are always full column rank, and consequently $\text{rank}\{\Theta_{\mathbf{H}}^{(zp)}\} = M$. The same analysis for CP-OFDM highlights that the observation matrix in (5.66) is $\Theta_{\mathbf{H}}^{(cp)} = \text{diag}(\mathbf{h}_f)$, which is not full rank for those channel realizations \mathbf{h}_f that contain a zero in one of its elements. Indeed, in CP-OFDM, if the discrete-time channel \mathbf{h}_t induces a frequency-domain channel $\mathbf{h}_f = \sqrt{M} \mathbf{F}_M \mathbf{h}_t$ with a zero on the $(m+1)$ th subcarrier, i.e., $[\mathbf{h}_f]_m = H_m = 0$, the data $d_{m,k}$ on that subcarrier would be not recoverable even in the absence of noise, as clarified by (5.81).

²The diversity d captured by a communication system is defined as the negative slope of the SER curve plotted (on a logarithmic scale) versus $(\text{SNR})_{\text{dB}}$, for $(\text{SNR})_{\text{dB}} \rightarrow +\infty$. In frequency-flat Rayleigh fading channels, the diversity is $d = 1$. A discrete-time channel with $L+1$ independent paths potentially offers a maximum diversity $L+1$. See [148, 229] for a rigorous definition of diversity in block transmissions.

Thus, at the receiver side, those data vectors \mathbf{d}_b that differ from \mathbf{d}_a only on the $(m + 1)$ th element would be undistinguishable from \mathbf{d}_a , causing a degradation of the average PEP and SER.

The price paid by ZP-OFDM with respect to CP-OFDM, for better uncoded SER performance with ML detection, is the higher detection complexity induced by (5.87), which grows exponentially with the vector size M and the cardinality of the data alphabet \mathcal{D} . Conversely, if ZP-OFDM is detected by exploiting the overlap-and-add (OLA) approach of (5.75), the detection complexity is significantly reduced. Indeed, the OLA operation of (5.75) induces a circulant channel at the receiver side: in this case, a simpler PSE can be employed by compensating for the $\text{diag}(\mathbf{h}_f)$ term in (5.78). Therefore, OLA-based per-subcarrier detection of ZP-OFDM presents a decoding complexity $O(M \log M)$ that is similar to CP-OFDM. In addition, using OLA-based per-subcarrier detection of ZP-OFDM, also the SER performance would be somewhat similar to CP-OFDM (see Figure 5.13). Indeed, with respect to CP-OFDM, OLA-based ZP-OFDM presents a noise power penalty caused by the extra noise term $\mathbf{e}_k^{(zp)}$

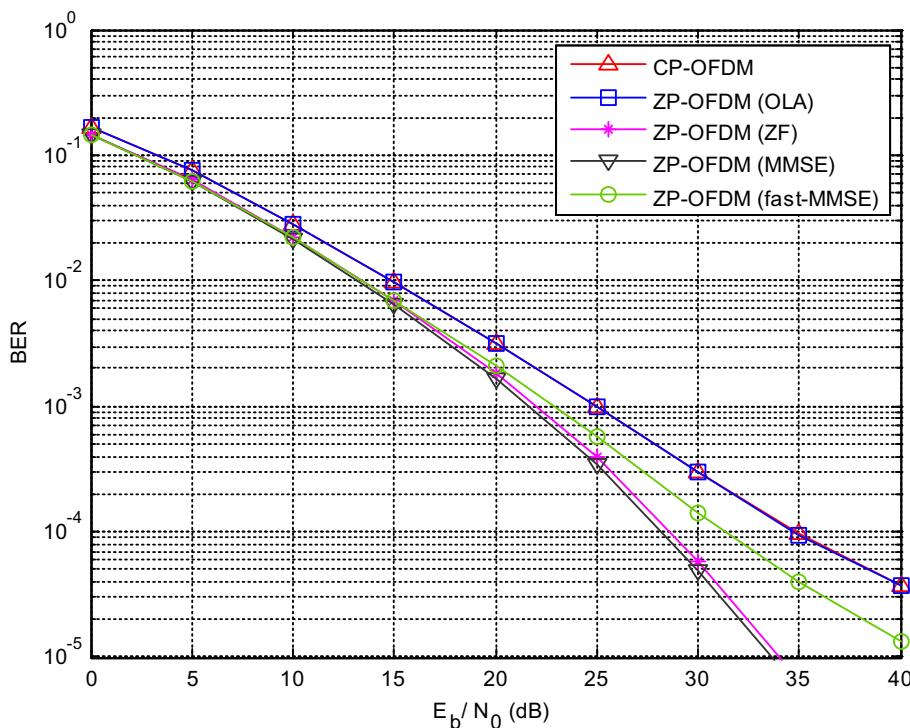


FIGURE 5.13

Performance comparison between CP-OFDM and different ZP-OFDM detectors in a multipath Rayleigh fading channel (Uncoded BER, QPSK, $M = 64$, $L = 16$, Channel A of HIPERLAN/2). This figure has been generated using the MATLAB script [zp_vs_cp.m](#).

in (5.78), and an equivalent advantage on the average (useful) signal power due to the absence of CP transmission. Each of these two effects compensates with one another, as shown in Figure 5.13.

2.05.7.1 Suboptimal linear equalization/detection for ZP-OFDM

As detailed in [167], other suboptimal receivers for ZP-OFDM are possible, allowing for different performance-complexity trade-offs. For instance, linear receivers produce a soft estimate of the transmitted data by

$$\tilde{d}_k = \mathbf{E}_{\Theta_{\mathbf{H}}} \mathbf{y}_k^{(\text{zp})}, \quad (5.90)$$

where the equalization/detection matrix can be designed according to a LS or MMSE approach [117], as expressed by

$$\mathbf{E}_{\Theta_{\mathbf{H}}}^{(\text{ls})} = (\Theta_{\mathbf{H}}^H \Theta_{\mathbf{H}})^{-1} \Theta_{\mathbf{H}}^H = \mathbf{F}_M \left(\mathbf{H}_t^{(\text{zp})H} \mathbf{H}_t^{(\text{zp})} \right)^{-1} \mathbf{H}_t^{(\text{zp})H}, \quad (5.91)$$

and

$$\mathbf{E}_{\Theta_{\mathbf{H}}}^{(\text{mmse})} = \left(\Theta_{\mathbf{H}}^H \Theta_{\mathbf{H}} + \frac{1}{\text{SNR}} \mathbf{I} \right)^{-1} \Theta_{\mathbf{H}}^H = \mathbf{F}_M \left(\mathbf{H}_t^{(\text{zp})H} \mathbf{H}_t^{(\text{zp})} + \frac{1}{\text{SNR}} \mathbf{I} \right)^{-1} \mathbf{H}_t^{(\text{zp})H}. \quad (5.92)$$

It is clear from (5.91) and (5.92), that this approach consists in a LS and MMSE equalization of the tall and Toeplitz channel matrix $\mathbf{H}_t^{(\text{zp})}$ in (5.73), followed by a frequency-domain per subcarrier detection.

These LS and MMSE approaches, denoted in Figure 5.13 with ZF and MMSE, respectively, give a performance gain with respect to OLA-based detection, at the price of increased complexity, which is $O((M + L)^2)$.

Another low-complexity suboptimal approach, denoted with fast-MMSE in Figure 5.13, relies on a sliding window of dimension $M + 2L$ that includes two ZP parts: the ZP part that precedes, and the ZP part that follows, the data vector of size M . In this case, the first ZP part, of size L , can be interpreted as the CP of the remaining vector, of size $M + L$, because the last part of the remaining vector is the second ZP part. As a consequence, the detection approach can be similar to CP-OFDM: the first ZP part is discarded, and the FFT of size $M + L$ is applied on the remaining vector of size $M + L$. The complexity of this fast-MMSE approach is $O((M + L) \log(M + L))$.

2.05.8 Coded—OFDM

As highlighted in Section 2.05.5, and summarized by (5.81) and (5.83), CP-OFDM is not capable to exploit the potential diversity offered by a multipath fading channel, whose taps fade in an independent fashion. This is a direct consequence of transmitting each single data $d_{m,k}$ onto a single subcarrier: if the m th frequency channel H_m fades to zero, there is no way to recover $d_{m,k}$, because $d_{m,k}$ is not contained in any other *diversely* received sample. Thus, in highly frequency-selective channels, the SER performance is significantly degraded because of the loss of the data $d_{m,k}$ on those subcarriers that are impaired by an almost null channel transfer function. For instance, in multipath Rayleigh fading conditions, the SER performance is equivalent to that of a single-carrier system impaired by a Rayleigh flat-fading channel, with diversity order equal to one.

Possible counter-measures to channel fading are typically known as diversity techniques and historically emerged in telecommunications by providing the receiver with multiple copies of the same information. These multiple copies of the same information are transmitted in the frequency-domain [or in the space (antenna)-domain or in the time-domain], and possibly are impaired by independent fading distortions [183]. The correct combination of multiple information manifests in an increase of the average SNR, a better SNR statistic, and consequently a SER performance that tends to the SER in AWGN channels, when the number of independent information copies is high enough [183]. Anyway, a pure diversity technique cannot improve performance over the SER threshold represented by the AWGN scenario.

Actually, Forward Error Correcting codes (FECs), or channel codes, were historically designed to contrast the errors in AWGN scenarios, by introducing algebraically-structured redundancy [183]. This form of redundancy, introduced at the transmitter side, distributes the information data on almost all the subcarriers: consequently, by channel decoding at the receiver side, it is possible to exploit this redundancy to jointly contrast the errors induced by fading channels and collect part of the channel diversity. Some examples of popular FEC schemes [183] exploited in standard Coded-OFDM (C-OFDM) systems include block-codes (e.g., Reed-Solomon codes), convolutional codes (CC), and concatenated codes, which were exploited since the beginning of the OFDM era in commercial systems such as DAB [71], DVB-T/S [72], WiFi [104], WiMAX [105], while BCH codes, turbo codes and low-density parity-check (LDPC) codes [183] have been more recently exploited in DVB-T2/S2 [73] and UMTS-LTE [237].

The basic idea to use FEC to contrast fading is to make the errors statistic induced by fading looks similar to the error statistic introduced by the noise in AWGN channels: this is obtained by equipping the OFDM system with appropriate interleavers, which scramble the coded data through the subcarriers to collect the channel frequency diversity, and through consecutive OFDM blocks to collect the channel time diversity. The price for such a joint resistance to fading and AWGN is the increase of decoding delay, because the interleaver depth may span several OFDM blocks in order to be effective. Thus, especially for time-sensitive communications, it could be preferable to have some sort of diversity resistance that does not require coding over several consecutive OFDM blocks.

This possibility has been explored more recently, by exploiting the possibility to transmit (possibly redundant) linear combinations of the data $d_{m,k}$ by the so called linearly precoded-OFDM (see Section 2.05.9), or by antenna-diversity through space-time coding (see [66, 84, 88, 178] ad references therein for further details).

2.05.9 Linearly precoded-OFDM (LP-OFDM)

FEC are typically designed to contrast AWGN channels and therefore are designed in order to maximize the Euclidean distance among codewords in order to be resistant to an additive distortion. This Euclidean maximization is not the optimal strategy in a fading channel, where the Hamming distance is the key element to minimize the PEP among codewords [229, 250], in order to be resistant to a multiplicative Distortion.

In an OFDM system, which conveys information on parallel carriers, the transmitter could spread a group of data $\{d_{m,k}\}$ over several of these carriers, by exploiting a linear precoding (LP) matrix Θ . The aim is to introduce frequency diversity and potentially be capable to recover all the transmitted data

(exploiting the finite alphabet of $d_{m,k}$) even if one of the carriers fades to zero. Thus, an LP transmitter generates a (potentially redundant) vector $\mathbf{d}_k^{(p)}$ by performing a linear combination of the original data, as shown in Figure 5.12 and expressed by

$$\mathbf{d}_k^{(p)} = \Theta \mathbf{d}_k, \quad (5.93)$$

where $\Theta = [\theta_0, \dots, \theta_{M-1}]$, is the (possibly redundant) precoding matrix³ of size $M' \times M$. Thus, with LP, a CP-OFDM system transmits $\mathbf{d}_k^{(p)}$ through a multipath fading channel \mathbf{h}_t , and (5.65) becomes

$$\mathbf{y}_k^{(f)} = \text{diag}(\mathbf{h}_f) \Theta \mathbf{d}_k + \mathbf{w}_k^{(f)}, \quad (5.94)$$

where the number of subcarriers (i.e., the size of $\mathbf{y}_k^{(f)}$ and \mathbf{h}_f) is $M' \geq M$. From (5.94), ML detection requests to compute

$$\hat{\mathbf{d}}_k = \arg \min_{\mathbf{d} \in \mathcal{D}^M} \left[\left\| \mathbf{y}_k^{(f)} - \text{diag}(\mathbf{h}_f) \Theta \mathbf{d}_k \right\|^2 \right], \quad (5.95)$$

which involves the non-diagonal channel-precoding matrix $\mathbf{H}_\Theta = \text{diag}(\mathbf{h}_f) \Theta$. Thus, the easy PSE (and easy ML detection) of OFDM is lost, but the PEP (and SER) performance can be boosted by a proper design of the precoding matrix Θ .

As clarified in [172], in order to exploit all the diversity d offered by a multipath fading channel, the overall observation matrix should guarantee that in the absence of noise any couple of transmitted data vectors are observable at the receiver side with a minimum Hamming distance greater than or equal to d . To this end, it is sufficient (but not necessary) to guarantee the symbol detectability condition, i.e., that \mathbf{H}_Θ has full column rank. Note that the discrete-time frequency response \mathbf{h}_f of a time-domain channel \mathbf{h}_t with $L+1$ non-zero paths may have at most L zeros: therefore, $\text{diag}(\mathbf{h}_f)$ can cancel-out up to L rows of Θ . Thus, it would be enough to redundantly design Θ with $M' = M + L$ linearly independent rows, such that any set of M rows of Θ are linearly independent. If this design condition is verified, the row cancellation in $\text{diag}(\mathbf{h}_f) \Theta$ still guarantee that $\mathbf{H}_\Theta = \text{diag}(\mathbf{h}_f) \Theta$ has (full-column) rank M . Thus, by the M out of M' observations in (5.94) that are different from zero, it would be always possible in the absence of noise to distinguish any transmitted data vector \mathbf{d}_a from any other \mathbf{d}_b . A possible choice for such precoding matrices are Vandermonde matrices where [251]

$$[\Theta]_{i,j} = \rho_i^{j-1}, \quad \text{with } \rho_i \neq \rho_l \text{ and } \rho_i \in \mathbb{C}, \quad (5.96)$$

or cosine matrices with

$$[\Theta]_{i,j} = \cos((j + 0.5)\phi_i), \quad \text{where } \phi_i \pm \phi_l \neq 2k\pi, \quad \forall i \neq l, \quad \forall k \in \mathbb{Z}. \quad (5.97)$$

When $\rho_i = e^{j\frac{2\pi}{M}i}$ and $M' = M$, the Vandermonde matrix becomes the DFT matrix $\Theta = \mathbf{F}_M$, while, when $\phi_i = i\pi/N$, Θ represents the discrete cosine transform (DCT) matrix. Moreover, note that the

³Note that by (5.93) it is possible to describe also classical block channel coding techniques developed for AWGN channels, with some minor differences. While for LP the precoded data is constructed on the complex field, i.e., $\Theta \mathbf{d}_k \in \mathbb{C}^{M'}$, in classical binary block codes $\Theta \mathbf{d}_k \in \{-1, 1\}^{M'}$ for binary PSK, such that the coded vectors can be represented on a lattice, $[\Theta]_{i,j}, [\mathbf{d}_k]_i \in GF(2)$, and the redundancy inserted by Θ is typically exploited to increase the minimum Euclidean distance among codewords.

precoding matrix should be normalized such that all the columns have a unit norm, i.e., $\|\boldsymbol{\theta}_m\|^2 = 1$ [250] in order to not decrease the minimum Euclidean distance between the precoded vectors and consequently penalize the SER performance in AWGN channels where $\text{diag}(\mathbf{h}_f)^H = h_o \mathbf{I}_{M'}$.

An interesting choice for the redundant precoding matrix is $\boldsymbol{\Theta} = (\mathbf{F}_{M+L}^{(1:M)})^H$, i.e., the Vandermonde matrix obtained by selecting the first M columns of an IDFT matrix of size $M + L$. By this choice, the precoded vector to be transmitted, is expressed by

$$\mathbf{d}_k^{(p)} = \mathbf{F}_{M+L} \left(\mathbf{F}_{M+L}^{(1:L)} \right)^H \mathbf{d}_k = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \mathbf{d}_k = \mathbf{T}_{zp} \mathbf{d}_k, \quad (5.98)$$

which is nothing else than a ZP-SC system, which does not need the use of a CP. Thus, the redundancy introduced by ZP in a block SC system grants not only IBI suppression, but also symbol detectability and, consequently, the capability by ML decoding to exploit the maximum diversity offered by the channel [253, 254].

In practice, LP-OFDM systems (including ZP-SC), can capture (part of) the diversity [236] also by exploiting suboptimal linear equalization/detection schemes by producing the soft estimate

$$\tilde{d}_k = \mathbf{E}_{\mathbf{H}_{\Theta}} \mathbf{y}_k^{(f)}, \quad (5.99)$$

where the equalization/detection matrix $\mathbf{E}_{\mathbf{H}_{\Theta}}$ can be computed by an LS or MMSE approach [117], by substituting $\boldsymbol{\Theta}_H$ with \mathbf{H}_{Θ} in (5.91) and (5.92), respectively. Further details can be found in [250].

However, the sufficient condition $\{\text{diag}(\mathbf{h}_f) \boldsymbol{\Theta}\} = M$ guarantees that any couple of vectors $\mathbf{d}_a \neq \mathbf{d}_b$ could be distinguished at the receiver side in the absence of noise, without exploiting the fact that the elements of \mathbf{d}_k belong to a finite alphabet \mathcal{D} . Thus, by exploiting the structure imposed on the possible data vectors by the finite alphabet and algebraic number theory, it is possible to design also non-redundant (square) precoding matrices $\boldsymbol{\Theta}$ that are capable to capture all the diversity offered by the channel, with $M' = M$. As detailed in [148], such a matrix is proved to always exist: it is a square Vandermonde matrix, whose elements are expressed by

$$\rho_i = \frac{e^{j \frac{2\pi}{P} [P(i-1)/M + p]}}{\sqrt{M}}, \quad \text{such that} \quad \begin{aligned} \Phi(P) &= 2kM, \quad k \in \mathbb{N}^+, \\ p &\in [1, P/M) \cap \mathbb{N}^+, \\ \text{MCD}([P(i-1)/M + p], P) &= 1, \end{aligned} \quad (5.100)$$

where $\Phi(P)$ is the Euler's totient function, which identifies the number of integers that are lower, and relative prime, with respect to P . Although for a given size M the solution of (5.100) is not unique, for the special but important case where M is a power of two, the precoding matrix can be always written as

$$\boldsymbol{\Theta} = \mathbf{F}_M \text{diag}([1, \rho_1, \dots, \rho_1^{M-1}]), \quad (5.101)$$

which highlights that the precoding matrix is unitary, i.e., $\boldsymbol{\Theta} \boldsymbol{\Theta}^H = \mathbf{I}_M$. Moreover, (5.101) practically states that the diversity-optimum CP-aided block transmission, is a CP-SC system equipped by a simple diagonal non-redundant precoder, as expressed by the transmitted data vector

$$\mathbf{x}_k^{(\text{cp-lp-sc})} = \mathbf{T}_{\text{cp}} \mathbf{F}_M^H \mathbf{d}_k^{(p)} = \mathbf{T}_{\text{cp}} \mathbf{F}_M^H \boldsymbol{\Theta} \mathbf{d}_k = \mathbf{T}_{\text{cp}} \text{diag}([1, \rho_1, \dots, \rho_1^{M-1}]) \mathbf{d}_k. \quad (5.102)$$

It is important to highlight that the non redundant precoding based on (5.100) or (5.101) is not only capable to capture all the diversity, but can also maximize the *coding gain*⁴ in SER performance, by selecting the lowest p that satisfies (5.100). For further details on optimal constellation rotations, the interested readers are redirected to [24, 32, 87].

The price to be paid for the exploitation of the channel diversity is the computational complexity of ML decoding, which grows exponentially with the size M of the data vector \mathbf{d}_k . Quasi-ML decoding performance can be obtained by algorithms such as sphere decoding [250] characterized by reduced complexity for moderate precoder sizes and many SNR values [108].

However, it can be observed that, in order to not sacrifice too much the system efficiency, the CP or ZP length is in general $L + 1 \ll M$, and consequently, in order to collect the diversity $d \leq L + 1$ offered by the channel, it would be enough to spread each symbol on a number of subcarriers $Q_p = M/G_p \geq L + 1$, where G_p is an integer greater than 1. Therefore, instead of a large precoder of size M equal to the number of subcarriers, many precoders can be applied to different groups of $Q_p = M/G_p$ subcarriers. This corresponds to an overall precoded matrix that can be written as

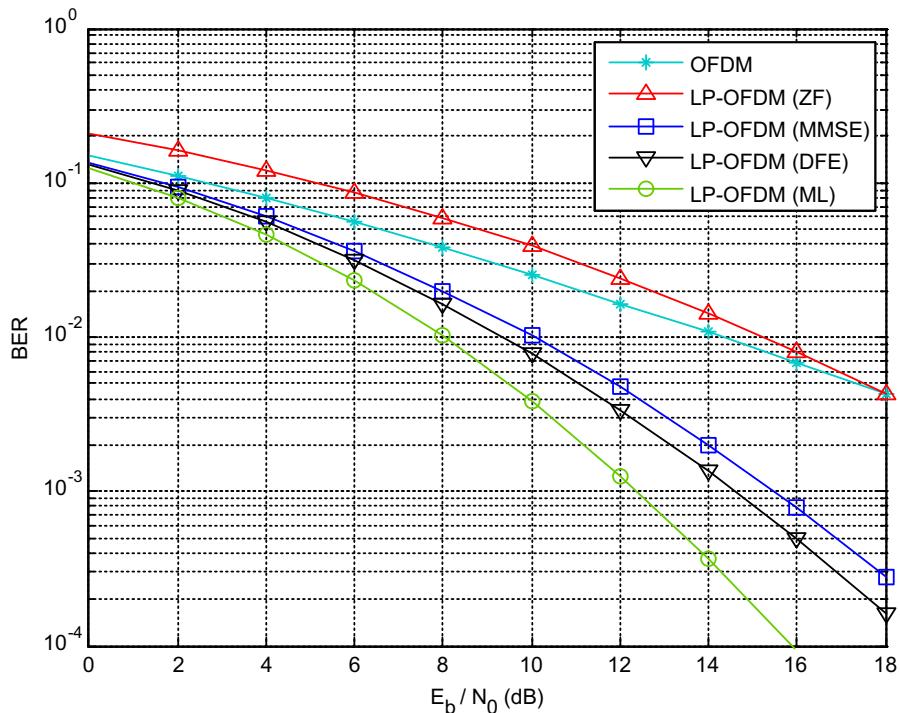
$$\Theta = \mathbf{P} \begin{bmatrix} \Theta_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Theta_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \Theta_{G_p-1} \end{bmatrix}, \quad (5.103)$$

where \mathbf{P} is a permutation matrix that establishes how the M subcarriers are grouped into the G_p groups, and $\{\Theta_g = \Theta_0\}_{g=0,\dots,G_p-1}$ is a $Q_p \times Q_p$ non-redundant precoding matrix designed according to (5.101). Since the benefit in SER performance tends to reduce for increasing diversity [183], suboptimal designs based on (5.103) with $Q_p = 3, 4 < d \leq L + 1$ can be employed: this way, the ML decoding in (5.95) leads to G_p parallel ML decoding problems for the data observed (and transmitted) on the G_p orthogonal sets, with a complexity $\mathcal{O}(G_p |\mathcal{D}|^{Q_p})$ instead of $\mathcal{O}(|\mathcal{D}|^{Q_p G_p})$. The parallel ML decoding is expressed by

$$\hat{\mathbf{d}}_{k,g} = \arg \min_{\mathbf{d}_g \in \mathcal{D}^{Q_p}} \left[\left\| \mathbf{y}_{k,g}^{(f)} - \text{diag}(\mathbf{h}_{f,g}) \Theta_0 \mathbf{d}_{k,g} \right\|^2 \right], \quad g = 0, \dots, G_p - 1. \quad (5.104)$$

The data sub-vectors $\mathbf{d}_{k,g}$, as well as the other quantities in (5.104), are just the sub-vectors selected from the corresponding ones in (5.95), by means of (5.103). This approach has been proposed in [148], where it is called grouped linear precoding (GLP). In order to collect the highest possible diversity for all the subgroups, and thus optimize the average SER, the best strategy is to design the permutation matrix \mathbf{P} such that each subgroups of data $\mathbf{d}_{k,g}$ is assigned to a different set of equispaced subcarriers, which, being separated, are most likely different (more *diverse*) than contiguous subcarriers [148]. In the context of non-redundant GLP-OFDM, the performance of ML and other suboptimal equalization/detection schemes, such as ZF and MMSE, is shown in Figure 5.14.

⁴While the *diversity gain* is a measure of the (negative) slope of the SER (or PEP) curve versus $(\text{SNR})_{\text{dB}}$, in high SNR regimes, the *coding gain* (for a fixed *diversity gain*) measures the $(\text{SNR})_{\text{dB}}$ shift to the left of the SER, with respect to a precoder that does not provide any coding gain. While the *diversity gain* is imposed by the rank of the overall channel observation matrix, the *coding gain* is established by the determinant of the observation matrix [148, 229].

**FIGURE 5.14**

Performance comparison among different detection criteria for GLP-OFDM (Uncoded BER, BPSK, $M = 16$, $G_P = 2$, $Q_P = 8$, $L = 2$, multipath channel with uniform power-delay profile). This figure has been generated using the MATLAB script `lp_ofdm.m`.

Linear precoding approaches for multicarrier transmissions have been discussed also in [38], which compares linear precoders based on rotated (Fourier or Walsh-Hadamard) transforms, and in [146], which derives the conditions for minimum-SER precoding design for a linear detector.

A fundamental contribution to redundant LP can be found in [201, 202]. The first of these two companion papers generalizes OFDM, DMT, CDMA, and TDMA communications, and establishes sufficient conditions for a precoded transmission scheme to perfectly equalize any FIR channel (channel identifiability), independently of its zeros locations. Additionally, a joint ZF design of the transmitter and receiver filterbanks is proposed, under MMSE and maximum SNR criteria. The second paper, by a ZP approach, derives also blind channel estimators, block synchronizers, and direct-equalizers for the proposed filterbank structure. Note that, also multiuser MC systems, such those addressed in the next section, fall within the general filterbank formulation in [201, 202].

As a final remark, it should be noted that LP and FEC are not alternative with one another, but they are both valuable tools to protect an OFDM system from errors induced by a fading channel. As already explained, FECs have been historically introduced to protect digital communications from AWGN, they introduce redundancy, and can be effective also to combat channel fading by properly scrambling and

interleaving the coded data. Conversely, LP techniques are designed to exploit the channel diversity (if it is present), cannot improve performance in AWGN, and they do not necessarily request redundancy. Thus, a proper design could combine LP with FEC, in order to collect the channel diversity (if any) to make the equivalent channel almost AWGN, and then exploit FEC mainly to protect from AWGN. In severe multipath fading channels, this approach can guarantee SER performance equivalent to those of a pure FEC-OFDM system, with lower redundancy, lower complexity, and lower decoding delay. Readers interested on a comparison of LP with FEC, and on their joint design, are redirected to [252–254].

2.05.10 Multiuser multicarrier systems

Section 2.05.2 has clarified that OFDM can be seen as a special case of CDMA, where M parallel data flows are simultaneously transmitted in time and distinguished by the use of different codes $c_m(t)$, as expressed by

$$x(t) = \sum_{m=0}^{M-1} x^{(m)}(t) = \sum_{k=-\infty}^{+\infty} \sum_{m=0}^{M-1} d_{m,k} c_m(t - kT) \text{rect}_{\Delta T}(t - (k + 1/2)T), \quad (5.105)$$

where T is the symbol period and ΔT is the symbol duration, less than or equal to the symbol period. For instance, a CDMA system with M parallel transmissions, can be represented during the useful k th symbol by a transmitted data vector $\mathbf{x}_k^{(c)} = [x(kT), \dots, x(kT + (M-1)T_s)]^T$ expressed by

$$\mathbf{x}_k^{(c)} = \sum_{m=0}^{M-1} \mathbf{c}_m d_{m,k} = \mathbf{C} \mathbf{d}_k, \quad (5.106)$$

where $\mathbf{c}_m = [c_{m,0}, \dots, c_{m,M-1}]^T$, with $c_{m,n} = c_m(nT_s)$, is the m th spreading sequence, also known as spreading code, and $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_{M-1}]$ is the code matrix. Additionally, to handle IBI as in OFDM transmissions, each transmitted block in (5.105) could be separated in time by $T = T_u + T_h$, leading to a CP-CDMA when $\Delta T = T$, or to a ZP-CDMA when $\Delta T = T_u$. Further considerations on CDMA are left for the end of this section where CDMA systems are briefly compared with multiuser MC systems.

In the case of OFDM, the codes are $c_m(t) = e^{j2\pi f_m t}$, which are orthogonal on the useful duration T_u by (21), $\mathbf{C} = \mathbf{F}_M^H$, and the overall transmitted block is

$$\mathbf{T} \mathbf{x}_k^{(c)} = \mathbf{T} \sum_{m=0}^{M-1} \mathbf{f}_m d_{m,k} = \mathbf{T} \mathbf{F}_M^H \mathbf{d}_k, \quad (5.107)$$

with $\mathbf{T} = \mathbf{T}_{\text{cp}}$ for CP-OFDM and $\mathbf{T} = \mathbf{T}_{\text{zp}}$ for ZP-OFDM. When orthogonal frequencies \mathbf{f}_m are employed to distinguish different users, the system is identified as an OFDM access (OFDMA) systems, which has the nice property to preserve users' orthogonality in LTI channels. Moreover, different data rates can be easily handled in OFDMA by reserving different groups of subcarriers to different users, as it happens for instance in WiMAX and LTE [105,237].

Researchers have historically tried to combine the easy equalization of OFDM, with desirable properties of classical CDMA systems, such as resistance to narrowband interference, and capability to exploit diversity in multipath fading channels [183, 247].

A possibility is to use a code in the frequency-domain, to spread a single data $d_{u,k}$ on all the subcarriers, such that, after spreading, the data vector for the u th user is $\mathbf{d}_{u,k}^{(c)} = \mathbf{c}_u d_{u,k}$, and the overall data vector for all the U users is expressed by

$$\mathbf{d}_k^{(c)} = \sum_{u=0}^{U-1} \mathbf{d}_{u,k}^{(c)} = \sum_{u=0}^{U-1} \mathbf{c}_u d_{u,k} = \mathbf{C} \mathbf{d}_k, \quad (5.108)$$

which is transmitted by OFDM through

$$\begin{aligned} \mathbf{x}_k &= \mathbf{T} \mathbf{F}_M^H \mathbf{d}_k^{(c)} = \mathbf{T} \mathbf{F}_M^H \mathbf{C} \mathbf{d}_k = \mathbf{T} \sum_{u=0}^{U-1} \mathbf{F}_M^H \mathbf{c}_u d_{u,k} \\ &= \mathbf{T} \sum_{u=0}^{U-1} \mathbf{c}_u^{(f)} d_{u,k} = \mathbf{T} \mathbf{C}^{(f)} \mathbf{d}_k, \end{aligned} \quad (5.109)$$

and where the number of orthogonal codes (users) is $U \leq M$.

The transmission mode with $\mathbf{T} = \mathbf{T}_{cp}$ is typically called multicarrier CDMA (MC-CDMA) [95, 251], whose implementation principle is shown in Figure 5.15.

Note that the last line in (5.109) highlights that MC-CDMA corresponds to a CP-CDMA system, which transmits the inverse spectra $\mathbf{c}_u^{(f)} = \mathbf{F}_M^H \mathbf{c}_u$ of the original codes \mathbf{c}_u , plus a CP to handle IBI and to guarantee easy frequency-domain equalization. This way, MC-CDMA grants both easy PSE and multipath diversity exploitation, since each data symbol is transmitted (spread) by the code matrix \mathbf{C} on several (potentially all) subcarriers. Indeed, the frequency-domain spreading operation expressed

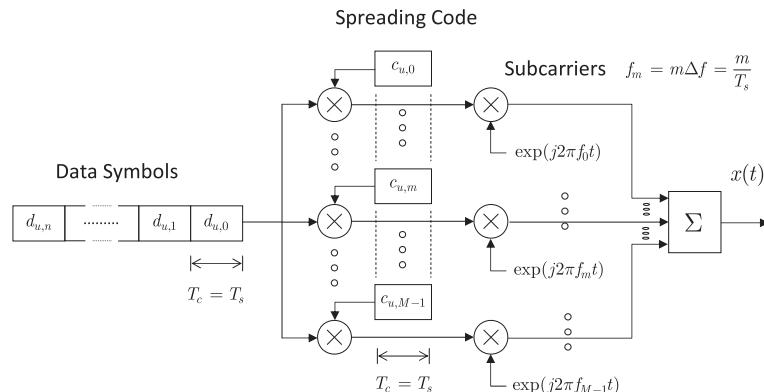


FIGURE 5.15

An MC-CDMA system.

in (5.109) is performed on all the M subcarriers, and therefore tries to achieve a diversity gain of M . However, the frequency-domain channel coefficients on the M subcarriers are correlated, because they are obtained by DFT of $L + 1$ nonzero time-domain channel coefficients in (5.63), with $L + 1 \ll M$. Therefore, the maximum achievable diversity gain is only $L + 1$. For this reason, similarly to GLP in (5.103), multicarrier CDMA systems can employ an alternative frequency-domain spreading operation that uses only $L + 1$ subcarriers out of M . In order to minimize the correlation, the $L + 1$ subcarriers are chosen as maximally separated, i.e., equispaced. Therefore, using spreading sequences of length $L + 1$, we can accommodate up to $L + 1$ users into the $L + 1$ subcarriers. If we assume that $M/(L + 1)$ is integer, there are $M/(L + 1)$ different groups of equispaced subcarriers, and therefore $M/(L + 1)$ groups of $L + 1$ users can be accommodated. This scheme is usually known as group-orthogonal MC-CDMA [42], because each group behaves as a reduced-dimension MC-CDMA systems, while the different groups are orthogonal in the frequency-domain, similarly to OFDMA.

The above mentioned group approach is also suitable for multirate transmissions. For instance, the $L + 1$ subcarriers of a given group, and all the $L + 1$ associated spreading codes, can be assigned to the same user. In this case, only $M/(L + 1)$ users can be accommodated, but the data rate of each user can be $L + 1$ times higher than conventional MC-CDMA. This multirate scheme is basically a combination of linear precoding and OFDMA, since users, which are orthogonal in the frequency-domain, can apply a linear precoder onto the $L + 1$ assigned subcarriers, as detailed in Section 2.05.7. Alternatively, multirate transmissions may be obtained by multicode MC-CDMA, as detailed in [187].

Another possibility to mix CDMA with OFDM is to apply the CDMA spreading principle to each of the M independent SC transmissions embedded in OFDM, as shown in Figure 5.16. Indeed, by (5.105), each SC data flow can accommodate $U \leq G$ different users by expressing

$$x^{(m)}(t) = \sum_{u=0}^{U-1} x_u^{(m)}(t) = \sum_{u=0}^{U-1} \sum_{k=-\infty}^{+\infty} d_{m,k}^{(u)} c_m^{(u)}(t - kT), \quad (5.110)$$

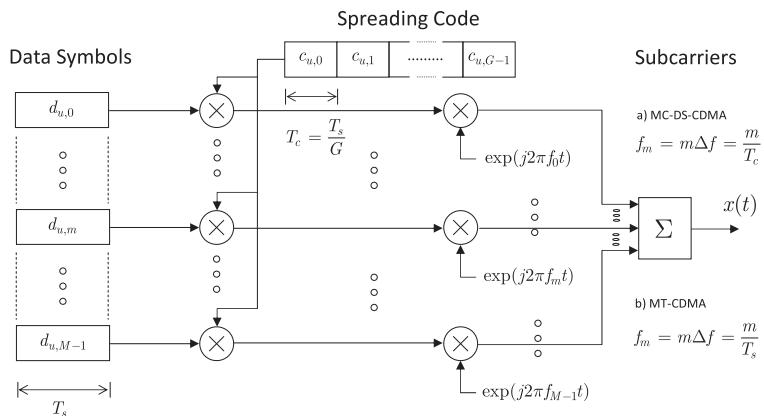


FIGURE 5.16

MC-DS-CDMA and MT-CDMA systems.

where, ignoring the potential presence of time guards to prevent IBI, the pulse shaper $c_m^{(u)}(t)$ associated to each data is spread by the chip sequence $\{c_{u,g}\}_{g=0,\dots,G-1}$ through the signal $\sum_{g=0}^{G-1} c_{u,g} \text{rect}_{T_c}(t - gT_c)$, as expressed by

$$c_m^{(u)}(t) \text{rect}_{T_G}(t - T_G/2) = e^{j2\pi f_m t} \sum_{g=0}^{G-1} c_{u,g} \text{rect}_{T_c}(t - gT_c - T_c/2), \quad (5.111)$$

where the code duration is $T_G = GT_c$. The time spreading of each waveform in (5.111) induces a spectrum

$$C_m^{(u)}(f) = \text{sinc}(\pi T_c(f - f_m)) \sum_{g=0}^{G-1} c_{u,g} e^{-i2\pi(f-f_m)gT_c}, \quad (5.112)$$

and an overall spectrum support on the m th subcarrier expressed by

$$X_k^{(m)}(f) = e^{-j2\pi f k T} \sum_{u=0}^{U-1} d_{m,k}^{(u)} C_m^{(u)}(f), \quad (5.113)$$

which is centered on f_m and has equispaced zeros on the frequency grid $\{f_z = f_m \pm i/T_c\}_{i \in N^+}$.

Thus, similarly to classical OFDM, the spectra on different subcarriers will preserve orthogonality if the subcarriers are chosen such that the subcarrier separations are multiple of $1/T_c$, e.g., when $f_m = m/T_c = mG/T_G$. This transmission mode is typically called multicarrier direct-sequence CDMA (MC-DS-CDMA) [95]: this corresponds to an OFDM system that transmits a chip of duration T_c in each OFDM block, i.e., in MC-DS-CDMA the OFDM block duration is T_c , and the symbol duration is GT_c (see Figure 5.16). Thus, assuming a fixed number M of subcarriers, if we want to compare MC-DS-CDMA (with block duration T_c) with respect to single-user OFDM (with block duration T_u), two cases are of interest. First, if each MC-DS-CDMA user wants to preserve the same data rate with respect to single-user OFDM systems, the spreading operation imposes a shorter OFDM block duration $T_c = T_u/G = MT_s/G$, and, consequently, a bandwidth overexpansion of a factor G . Alternatively, if each MC-DS-CDMA user wants to preserve the same bandwidth with respect to single-user OFDM systems, the spreading operation leads to a data rate reduction of a factor $1/G$, which is a direct consequence of the OFDM block duration $T_c = T_u$ and of the increased symbol duration $GT_c = GT_u$. However, it should be emphasized that CDMA systems allows for a maximum of G orthogonal users: therefore, in the second case, the aggregate data rate of the G MC-DS-CDMA users would be equal to that of single-user OFDM systems.

Moreover, in MC-DS-CDMA, in order to prevent IBI (i.e., inter-chip interference) and grant easy equalization in LTI channels, CP or ZP has to be inserted between each OFDM block (i.e., between each chip). In a vector-matrix notation, this corresponds to a coded (row) vector for each subcarrier expressed by

$$\mathbf{s}_{m,k}^{(u)T} = d_{m,k}^{(u)} \mathbf{c}_u^T, \quad (5.114)$$

which leads to an overall spread data-matrix for each user

$$\mathbf{D}_k^{(u)} = \left[\mathbf{d}_{0,k}^{(u)}, \dots, \mathbf{d}_{G-1,k}^{(u)} \right] = \left[\mathbf{s}_{0,k}^{(u)}, \dots, \mathbf{s}_{M-1,k}^{(u)} \right]^T = \text{diag}(\mathbf{s}_k^{(u)}) (\mathbf{1}_M \otimes \mathbf{c}_u^T), \quad (5.115)$$

whose columns are sequentially transmitted by OFDM. In (5.115), \otimes stands for the Kronecker product. The vector transmitted during each chip period is expressed by

$$\mathbf{d}_{g,k}^{(u)} = c_{u,g} \mathbf{d}_k^{(u)} \quad (5.116)$$

and the overall transmitted OFDM chip, i.e., OFDM block, is

$$\sum_{u=0}^{U-1} \mathbf{T} \mathbf{F}_M^H c_{u,g} \mathbf{d}_k^{(u)} = \mathbf{T} \mathbf{F}_M^H \sum_{u=0}^{U-1} c_{u,g} \mathbf{d}_k^{(u)} = \mathbf{T} \mathbf{F}_M^H \mathbf{D}_k \left[\mathbf{C}^T \right]_{:,g}, \quad (5.117)$$

where $\mathbf{D}_k = \left[\mathbf{d}_k^{(0)}, \dots, \mathbf{d}_k^{(G-1)} \right]$ is the column-wise data matrix containing the data vector associated to each user, and $\left[\mathbf{C}^T \right]_{:,g}$ is the column vector containing the g th chips of all the users, i.e., the g th row of the code matrix \mathbf{C} . Equation (5.117) highlights that MC-DS-CDMA grants the easy separation (and equalization) of the m th flow on the M subcarriers by preserving the subcarrier orthogonality through a frequency-selective channel, and the separation among different users by orthogonal design of the code matrix, which requests to collect all the G chips associated with a single data matrix \mathbf{D}_k .

Alternatively, as also shown in Figure 5.16, it is possible to renounce to easy equalization and separation of subcarriers by resorting to the so-called multi-tone CDMA (MT-CDMA) [95]. This system corresponds to choose subcarriers that are not orthogonal on the chip duration, and which would be orthogonal (in the absence of spreading) on the symbol duration $T_G = G T_c$, that is, $f_m = m/T_G = m/(G T_c)$. Thus, the spectra in (5.112), for different m , are not centered on the zeros of the other spectra as shown in Figure 5.17.

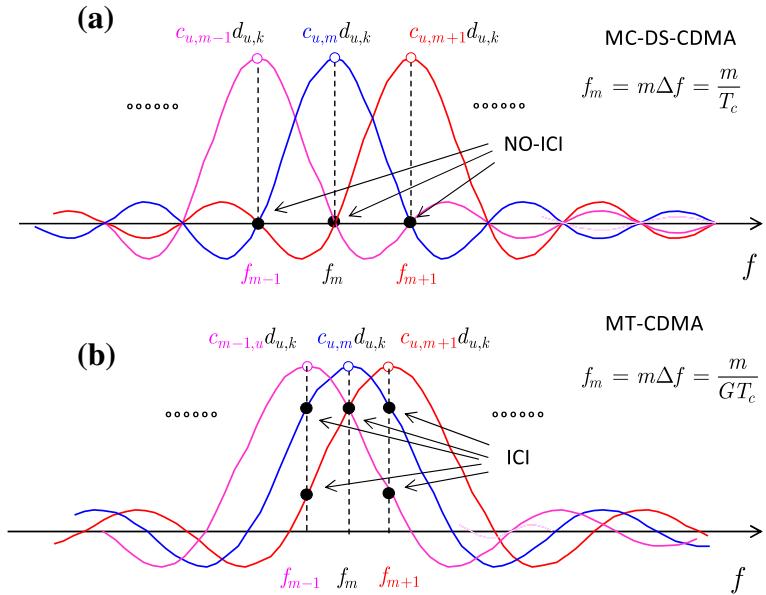
This way, ICI would emerge at the receiver side, independently of the presence or absence of CP or ZP. For this reason, MT-CDMA systems typically avoid the use of a CP [244]. MT-CDMA corresponds to an OFDM system, without CP, on the time interval T_G , followed by a chip spreading (bandwidth expansion) to distinguish different users [244]. This signal could be generated by a classical OFDM modulator at the symbol rate, followed by G up-sampling to describe the signal at the chip-rate and by the spreading operation, which is just a chip-by chip multiplication. Due to the fact that a time-domain up-sampling can be easily generated by DFT processing, through a zero-filled spectrum [173], a possible matrix representation for the overall data vector transmitted by the u th user is

$$\text{diag}(\mathbf{1}_M \otimes \mathbf{c}_u) \mathbf{F}_{GM}^H \begin{bmatrix} \mathbf{d}_{k,up}^{(u)} \\ \mathbf{0}_{(G-1)M} \\ \mathbf{d}_{k,low}^{(u)} \end{bmatrix} = \text{diag}(\mathbf{1}_M \otimes \mathbf{c}_u) \tilde{\mathbf{F}}_{GM}^H \mathbf{d}_k^{(u)}, \quad (5.118)$$

where $\tilde{\mathbf{F}}_{GM}$ is a selection of the DFT matrix \mathbf{F}_{GM} , containing its first and last $M/2$ rows, and $\mathbf{d}_{k,up}^{(u)}$ and $\mathbf{d}_{k,low}^{(u)}$ are two vectors of size $M/2$ obtained by splitting $\mathbf{d}_k^{(u)}$ into two parts.

2.05.10.1 Equalization and data recovery for multiuser multicarrier systems

In multicarrier multiuser systems, different data recovery algorithms may be employed, depending on the type of system (e.g., MC-CDMA, MC-DS-CDMA, or MT-CDMA), on the type of communication

**FIGURE 5.17**

Spectrum of (a) MC-DS-CDMA $f_m = m/T_c = mG/T_G$, (b) MT-CDMA $f_m = m/T_G = m/(GT_c)$.

(e.g., downlink or uplink), and on the type of channel (e.g., linear or nonlinear, time invariant or time varying, with or without multipath, and so on). In this work, as an example, we consider MC-CDMA downlink communications subject to time-invariant multipath channels. We also assume channel knowledge at the receiver side. In this case, after CP removal and FFT processing, by means of (5.65) and (5.108) the frequency-domain received vector can be expressed as

$$\mathbf{y}_k^{(f)} = \text{diag}(\mathbf{h}_f)\mathbf{d}_k^{(c)} + \mathbf{w}_k^{(f)} = \text{diag}(\mathbf{h}_f)\mathbf{C}\mathbf{d}_k + \mathbf{w}_k^{(f)}, \quad (5.119)$$

where $\mathbf{y}_k^{(f)}$ represents the k th received block, with size equal to the number of subcarriers M , \mathbf{h}_f is the frequency-domain channel, \mathbf{C} is the code matrix defined after (5.106), with size $M \times G$, where G is the number of users, \mathbf{d}_k is the transmitted data vector of all the G users, and $\mathbf{w}_k^{(f)}$ is the AWGN. Note the similarity of (5.119) with (5.94): indeed, when the number of users is $G = M$, MC-CDMA becomes a non-redundant LP-OFDM system with $\Theta = \mathbf{C}$, and therefore the detection techniques suitable for LP-OFDM can be employed (see Figure 5.14). On the other hand, when $G < M$, in general the user of interest does not know which codes have been assigned to the other users, and hence the full knowledge of \mathbf{C} cannot be exploited. Similarly to the suboptimal equalization/detection schemes of ZP-OFDM and LP-OFDM, summarized by (5.90) and (5.99), linear detection schemes can be used also for MC-CDMA, to produce a soft estimate

$$\tilde{\mathbf{d}}_k = \mathbf{E}_{\mathbf{H}\mathbf{C}}\mathbf{y}_k^{(f)}, \quad (5.120)$$

and where $\mathbf{E}_{\mathbf{H}\mathbf{C}}$ is obtained either by (5.91) or (5.92), by substituting $\mathbf{H}\mathbf{C} = \text{diag}(\mathbf{h}_f)\mathbf{C}$ to $\Theta_{\mathbf{H}}$.

Usually, for MC-CDMA downlink channels, the u th user is interested only to its transmitted data, and consequently performs the data estimation by means of a linear detector that employs only the code \mathbf{c}_u [95]. In this case, when the u th user has no knowledge about the codes of the other users, the joint equalization/despreadening in (5.120) can be split in two, as expressed by

$$\begin{aligned}\hat{d}_{u,k} &= \mathbf{c}_u^H \mathbf{E}_h \mathbf{y}_k^{(f)} \\ &= \mathbf{c}_u^H \mathbf{E}_h \text{diag}(\mathbf{h}_f) \mathbf{C} \mathbf{d}_k + \mathbf{c}_u^H \mathbf{E}_h \mathbf{w}_k^{(f)},\end{aligned}\quad (5.121)$$

where \mathbf{E}_h is a diagonal matrix that depends on the frequency-domain channel. As clearly expressed by (5.121), \mathbf{E}_h performs the equalization task, and \mathbf{c}_u performs the despreading operation.

For the matrix \mathbf{E}_h , different choices are possible [95]. The choice

$$\mathbf{E}_h^{(\text{orc})} = (\text{diag}(\mathbf{h}_f))^{-1}, \quad (5.122)$$

known as orthogonal restoring combining (ORC), perfectly compensates for the multipath channel \mathbf{h}_f in (5.119) and restores the orthogonality between users. Therefore, ORC eliminates the inter-user interference, also known as multiple-access interference (MAI). However, the drawback of this ZF approach is a noise enhancement that arises when one or more elements of \mathbf{h}_f have a low modulus: indeed, in this case, the corresponding elements of $\mathbf{E}_h^{(\text{orc})}$ have a high modulus, and hence the noise term $\mathbf{w}_k^{(f)}$ in (5.121) is enhanced. To reduce the noise enhancement, some subcarriers may be excluded from the detection: this ORC variant is known as controlled equalization [95]. An alternative choice for \mathbf{E} is the maximum ratio combining (MRC), as expressed by

$$\mathbf{E}_h^{(\text{mrc})} = (\text{diag}(\mathbf{h}_f))^H. \quad (5.123)$$

By inserting (5.123) into (5.121), it is clear that the MRC approach weights the subcarriers proportionally to their channel amplitudes, and compensate for the different phase offsets of the frequency-domain channel. Therefore, the MRC approach is a sort of frequency-domain matched filter, which maximizes the SNR when there is only a single active user. However, in the presence of multiple users, the amount of MAI may be significant [95]. A third approach, known as equal-gain combining (EGC), tries to avoid both excessive MAI and significant noise boosting, by means of

$$\mathbf{E}_h^{(\text{egc})} = (\text{diag}(e^{j\angle \mathbf{h}_f}))^H, \quad (5.124)$$

where $e^{j\angle \mathbf{h}_f}$ is the vector that contains the phases of the elements of \mathbf{h}_f . In general, when there are many active users, EGC provides better performance than ORC and MRC [95]. Another possible approach is the MMSE combining (MMSEC), which also tries to balance residual MAI and noise at the detector output. For MMSEC, the matrix \mathbf{E}_h in (5.121) is expressed by [95, 117]

$$\mathbf{E}_h^{(\text{mmse})} = (\text{diag}(\mathbf{h}_f))^H \left(G \text{diag}(\mathbf{h}_f) (\text{diag}(\mathbf{h}_f))^H + \sigma_w^2 \mathbf{I}_M \right)^{-1}. \quad (5.125)$$

Note that the MMSEC expression (5.125) requires also the knowledge of the number of active users G and AWGN power σ_w^2 . Specifically, at high SNR, σ_w^2 in (5.125) becomes negligible and therefore

MMSEC tends to ORC (5.122), thereby minimizing the MAI at the detector output. From (5.125), it is also clear that noise enhancement is avoided, because, when some elements of \mathbf{h}_f have a low modulus, $\mathbf{E}_{\mathbf{h}}$ does not contain elements with high modulus. For these reasons, usually the MMSEC detector outperforms ORC, MRC, and also EGC [95].

If the spreading codes of the other users are known to the user of interest, many other detection strategies are possible [95], including those used for LP-OFDM. For instance, (soft or hard) interference cancellation approaches, or quasi-ML detection techniques, may be pursued. However, we remind that the knowledge of the spreading codes of all users is more reasonable in uplink rather than in downlink. Indeed, in uplink communications, the base station has to detect the signals of all users, and therefore may perform a joint detection of the multiuser signals.

For MC-DS-CDMA, the detection strategies are usually different than for MC-CDMA. Specifically, in time-invariant frequency-selective channels, because of the time-domain spreading of MC-DS-CDMA, the use of orthogonal codes preserves user orthogonality. Hence, the despreading operation is performed in the time-domain, as in DS-CDMA [191] whereas the channel compensation is performed in the frequency-domain, as in OFDMA. However, because of the absence of frequency-domain spreading, pure MC-DS-CDMA does not collect frequency diversity, and therefore presents a performance loss

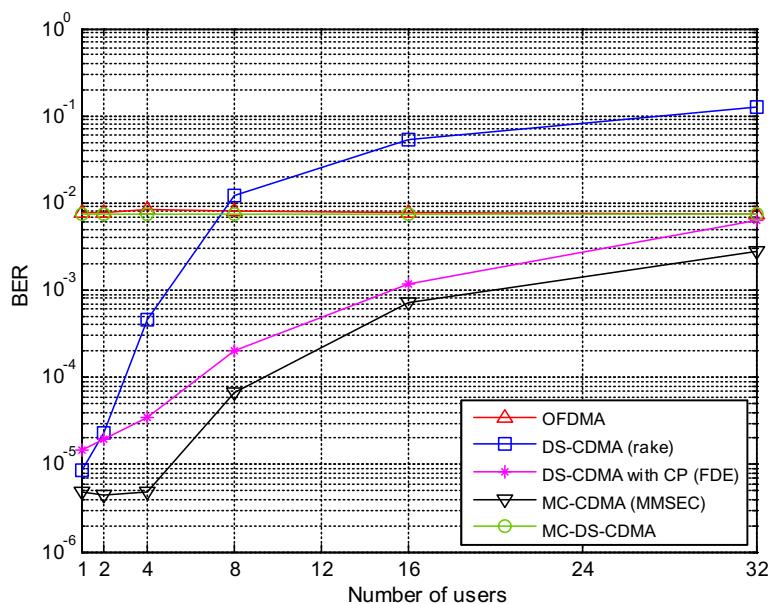


FIGURE 5.18

Performance comparison among OFDMA, DS-CDMA, MC-CDMA, and MC-DS-CDMA (Uncoded BER, QPSK, $E_b/N_0 = 6$ dB, $M = 32$, $G = 32$ and Walsh-Hadamard spreading codes for both MC-CDMA and MC-DS-CDMA, $G = 31$ and Gold spreading codes for DS-CDMA, $L = 4$, multipath channel with uniform power-delay profile).

with respect to MC-CDMA (see Figure 5.18). To avoid this performance loss, MC-DS-CDMA must incorporate additional features, such as in [127, 225], and [259]. On the other hand, MT-CDMA suffers from ICI, and therefore also presents a performance loss with respect to MC-CDMA, especially when the number of users is high (see [95]).

2.05.10.2 A note on CDMA versus MC communications

It has been clarified in Section 2.05.2, that according to the orthogonal design in (5.21) OFDM can be seen as a special case of a CDMA system, where each code is an exponential function as in (5.23). This is indeed the principle underlying OFDMA, where each (set of) carrier(s) is assigned to a specific user. Differently, in a classical short-sequence CDMA system the pulse assigned to each user is expressed by

$$p_m(t) = \sum_{g=1}^G c_{m,g} \text{rect}_{T_c}(t - gT_c - T_c/2) \quad (5.126)$$

and the orthogonality condition among two different users in (5.21) simply becomes

$$\int_{-\infty}^{+\infty} p_j(t) p_m^*(t) dt = T_c \sum_{g=0}^{G-1} c_{j,g} c_{m,g}^* = 0, \quad (5.127)$$

which is guaranteed if the two user codes \mathbf{c}_j and \mathbf{c}_m are orthogonal vectors, such as the Walsh-Hadamard codes [247]. However, the exponential codes used by OFDM and MC, are the only codes that preserve orthogonality through a frequency-selective channel, because they are eigenvectors of linear time-invariant systems, as explained in detail in Section 2.05.3.

Thus, a CDMA system in frequency-selective channels typically suffers significant BER performance degradation with respect to a MC system, because of the lost of codes orthogonality, i.e., the multiuser interference (MUI) that is introduced. This MUI can be contrasted by employing multiuser receivers [247] that anyway request the knowledge of the codes of all the users, which may be impractical in the downlink of cellular systems. Moreover, CDMA systems are particularly sensitive to MUI introduced by asynchronous reception (delay) of the users' codes, as it happens in the uplink of cellular systems. This is the reason why typically CDMA systems employ non-orthogonal codes (which would not be preserved anyway) such as the Gold codes, which conversely are designed in order to minimize the MUI in the presence of asynchronism among users [247].

The other difference, and probably the most important one, among classical CDMA systems and MC system is the insertion of the CP [or ZP, or a known-symbol padding (KP)], which greatly simplifies channels estimation and equalization in the frequency-domain, as deeply explained in the previous sections. Actually, it has been shown that, likewise in OFDM, also CP-SC systems can effectively diagonalize the channel in the frequency-domain. In this view, it is further possible to observe in Figure 5.12 that a classical SC-CDMA system equipped with a CP, is equivalent to a LP-OFDM system that employs $\Theta = \mathbf{F}_M \mathbf{C}$.

Thus, for CP-CDMA the signal received in the frequency-domain would be

$$\mathbf{y}_k^{(f)} = \text{diag}(\mathbf{h}_f) \mathbf{F}_M \mathbf{C} \mathbf{d}_k + \mathbf{w}_k^{(f)}, \quad (5.128)$$

and a simple (diagonal) frequency-domain equalization is possible by means of the ORC or MMSE principles in (5.122) and (5.125), respectively. Note that, for CP-CDMA the use of orthogonal codes could be preferred: indeed, the possible different delays between different users in a cellular uplink can be handled by a proper design of the CP (or ZP) duration, i.e., by granting that the CP length is greater than the sum of the maximum channel length with the maximum user delay. Thus, if the codes are orthonormal, after the simple diagonal equalization by \mathbf{E}_h , the data belonging to user m can be recovered by a simple per-user decision on the soft estimate

$$\tilde{d}_{m,k} = \mathbf{c}_m^H \mathbf{F}_M^H \mathbf{E}_h \mathbf{y}_k^{(f)}, \quad (5.129)$$

whose computational complexity is dominated by the IDFT processing, i.e., is of the order of $M \log M$.

For instance, in the case of ORC equalization (5.129) becomes

$$\tilde{d}_{m,k} = \mathbf{c}_m^H \mathbf{F}_M^H \text{diag}(\mathbf{h}_f)^{-1} \mathbf{y}_k^{(f)} = \mathbf{c}_m^H \mathbf{C} \mathbf{d}_k + \mathbf{c}_m^H \mathbf{F}_M^H \text{diag}(\mathbf{h}_f)^{-1} \mathbf{w}_k^{(f)} = d_{m,k} + e_{m,k}. \quad (5.130)$$

Anyway, the error $e_{m,k}$ may suffer noise enhancement introduced by $\text{diag}(\mathbf{h}_f)^{-1}$ when the values of \mathbf{h}_f are close to zero: consequently $\mathbf{E}_h^{(\text{mmse})}$ in (5.125) is in general a preferable choice for the equalizer.

2.05.11 OFDM and multicarrier communications in nonlinear systems

For the reader convenience, we report again the analytical expression (5.25) of the OFDM signal $x_k(t)$ transmitted during the useful period T_u of the k th symbol

$$x_k(t) = \frac{1}{\sqrt{T_u}} \sum_{m=0}^{M-1} d_{m,k} e^{j2\pi \frac{m}{T_u} t} \text{rect}_{T_u}(t - (k + 1/2)T_u), \quad (5.131)$$

which reveals that the transmitted signal is characterized by a highly variable envelope $R_k(t) = |x_k(t)|$, because at any different time instant $t \in [kT_u, (k+1)T_u]$ the M exponentials in (5.131) may add in phase or out-of-phase, depending on the phases of the transmitted data symbols $\{d_{m,k}\}_{m=0,\dots,M-1}$. Therefore, the OFDM signal is particularly sensitive to any baseband nonlinear device, such as A/D and D/A converters with limited dynamic [63, 171], and RF distortions, such those introduced by nonlinear (NL) high power amplifiers (HPA). Generally, also RF nonlinearities without memory can be summarized by their equivalent (possibly complex) BB nonlinearity [30]. For instance, a bandpass NL HPA, whose input is

$$x_{\text{RF}}^{(\text{in})}(t) = R(t) \cos(2\pi f_o t + \theta(t)),$$

has an output that can be expressed by

$$x_{\text{RF}}^{(\text{out})}(t) = A[R(t)] \cos(2\pi f_o t + \theta(t) + \Phi[R(t)]),$$

where $A[\cdot]$ and $\Phi[\cdot]$ are the so called AM/AM and AM/PM distortion functions [30, 118]. Thus, the BB equivalent of the OFDM signal $x_k(t) = R_k(t)e^{j\theta_k(t)}$ that passes through a (band-pass) NL HPA is expressed by

$$x_k^{(\text{out})}(t) = f_{\text{NL}}[x_k(t)] = G[R_k(t)] e^{j\theta_k(t)} = A[R_k(t)] e^{j[\Phi[R_k(t)] + \theta_k(t)]}, \quad (5.132)$$

which highlights how the NL function is generally complex (e.g., when AM/PM is present), and that the NL function operates on the OFDM signal envelope $R_k(t)$. Alternatively, if the nonlinearity is induced by an A/D conversion stage, it generally operates separately on the real and imaginary components of the OFDM BB signal, as expressed by

$$x_k^{(\text{out})}(t) = f_{\text{NL}}[x_k(t)] = Q_{\text{NL}}[\text{Re}\{x_k(t)\}] + j Q_{\text{NL}}[\text{Im}\{x_k(t)\}], \quad (5.133)$$

where $Q_{\text{NL}}[\cdot]$ is a real-valued NL function representing the A/D input-output characteristic.

Whichever is the situation, the overall NL distortion [$f_{\text{NL}}[\cdot]$ in (5.132) or $Q_{\text{NL}}[\cdot]$ in (5.133)] induces the intermodulation [65] of each subcarrier in (5.131) with the others, thus producing new spectral components that are partially superimposed to the OFDM tones, and partially spread out on the adjacent bands. The former intermodulation products induce a sort of ICI and, consequently, SER performance degradation, while the out-of-band intermodulation products are a source of potential interference for users or services on the adjacent bands.

In order to quantify the interference power generated by NL distortions, and their effect on the SER performance, it would be possible to resort to the classical literature for the characterization of the intermodulation products of FDM systems in satellite transponders [23, 80, 216]. However, an alternative and quite accurate analysis is possible when the number M of the OFDM subcarriers is sufficiently high, to make the OFDM signal well modeled as a complex Gaussian process, by central limit Theorem (CLT) arguments [15, 64, 171]. In this case, it is possible to exploit the Bussgang Theorem [177] and its complex extensions in [64, 158] to express the NL outputs as a scaled replica of the input plus an uncorrelated distortion term $d_k^{(\text{NL})}(t)$, as expressed in

$$x_k^{(\text{out})}(t) = \alpha(t)x_k^{(\text{in})}(t) + d_k^{(\text{NL})}(t), \quad (5.134)$$

where the scaling coefficient $\alpha(t)$ is expressed by

$$\alpha(t) = \frac{E\left\{x_k^{(\text{out})}(t)\left(x_k^{(\text{in})}(t)\right)^*\right\}}{E\left\{\left|x_k^{(\text{in})}(t)\right|^2\right\}}. \quad (5.135)$$

Actually, under fairly mild conditions [64], which hold true for classical OFDM, the coefficient $\alpha(t)$ is constant, as expressed by

$$\alpha(t) \simeq \alpha_o = E\left\{f_{\text{NL}}\left[x_k^{(\text{in})}(0)\right]\left(x_k^{(\text{in})}(0)\right)^*\right\} / 2\sigma_x^2, \quad (5.136)$$

and its computation in closed form is available for a wide class of NL distorting function, either expressed by (5.132) or (5.133) (see [15, 64, 171], and the references therein).

The uncorrelatedness of the distortion $d_k^{(\text{NL})}(t)$ with the OFDM input $x_k^{(\text{in})}(t)$ lets to express the NL output autocorrelation function as

$$R_{xx}^{(\text{out})}(\tau) = |\alpha_o|^2 R_{xx}^{(\text{in})}(\tau) + R_{dd}(\tau), \quad (5.137)$$

where $R_{dd}(\tau) = E\{f_{NL}^*[x_k^{(in)}(t)]f_{NL}[x_k^{(in)}(t + \tau)]\}$ is the autocorrelation function of the NL distortion. The Gaussian approximation of the OFDM signal $x_k^{(in)}(t)$ lets to compute $R_{dd}(\tau)$ in closed form for a wide set of NL distortions. The general result states that [65]

$$R_{xx}^{(out)}(\tau) = \sum_{n=0}^{\infty} c_n \left[R_{xx}^{(in)}(\tau)/2\sigma_x^2 \right]^{2n+1} = \frac{c_0}{2\sigma_x^2} R_{xx}^{(in)}(\tau) + \sum_{n=1}^{\infty} c_n \left[R_{xx}^{(in)}(\tau)/2\sigma_x^2 \right]^{2n+1}, \quad (5.138)$$

where $|\alpha_0|^2 = c_0/2\sigma_x^2$, the terms of the series for $n \geq 1$ represent the expansion of $R_{dd}(\tau)$, and the coefficients $c_n = f(2\sigma_x^2, f_{NL}[\cdot])$ depend on both the OFDM input power $2\sigma_x^2$ and the NL shape $f_{NL}[\cdot]$. Thus, the power spectral density (PSD) $S_{xx}^{(out)}(f) = \mathcal{F}\{R_{xx}^{(out)}(\tau)\}$ of the NL output is expressed by

$$S_{xx}^{(out)}(f) = |\alpha_0|^2 S_{xx}^{(in)}(f) + S_{dd}(f) = \frac{c_0}{2\sigma_x^2} S_{xx}^{(in)}(f) + \sum_{n=1}^{\infty} \frac{c_n}{(2\sigma_x^2)^{2n+1}} \left[S_{xx}^{(in)}(f) \underbrace{* \dots *}_{2n} S_{xx}^{(in)}(f) \right], \quad (5.139)$$

where $S_{xx}^{(in)}(f) = \mathcal{F}\{R_{xx}^{(in)}(\tau)\}$ is the input PSD, and $\underbrace{* \dots *}_{2n}$ stands for $2n$ successive convolution integrals of $S_{xx}^{(in)}(f)$ by itself, which induces the NL distortion PSD.

When the NL distortion $f_{NL}[r]$ is expressed by (5.132), the coefficients $\{c_n\}$ are generally expressed by [15, 30]

$$c_n = \frac{1}{2\sigma_x^2} \frac{1}{n+1} \left| \int_0^\infty f_{NL}[r] \frac{r^2}{\sigma_x^2} e^{-\frac{r^2}{2\sigma_x^2}} L_n^{(1)} \left(\frac{r^2}{2\sigma_x^2} \right) dr \right|^2, \quad (5.140)$$

where $L_n^{(1)}(x)$ is the Laguerre polynomial of the first type and order n [2], and (5.140) leads to closed form expressions for a wide class on NL distorting functions (see [15, 30]). Similar results are also available for NL distortion $f_{NL}[r]$ expressed by (5.133) (see, [64, 65, 171, 186] and references therein).

For instance, if the NL in (5.132) clips the signal envelope only when the input envelope overpasses a given threshold A (i.e., the NL is an ideally predistorted amplifier [15]), as expressed by

$$x_k^{(out)}(t) = f_{NL}[x_k(t)] = \begin{cases} G_{hpa} R_k(t) e^{j\theta_k(t)}, & 0 \leq R_k(t) < A, \\ G_{hpa} A e^{j\theta_k(t)}, & R_k(t) \geq A, \end{cases} \quad (5.141)$$

where G_{hpa} is the gain of the predistorted amplifier in the linear region, then

$$c_n = G_{hpa}^2 \frac{2\sigma_x^2}{n! (n+1)!} \left[\frac{(2n)! \sqrt{2\pi}}{n! 2^{n+1}} \frac{A}{\sigma_x} + \sum_{k=0}^{\infty} \frac{(n+k+1)!}{k! (k+2)!} \frac{(-1)^{k+1}}{(2k+3)} \left(\frac{A^2}{2\sigma_x^2} \right)^{k+2} \right], \\ n = 0, 1, \dots, \infty. \quad (5.142)$$

By some manipulations, expression (5.142) can be further simplified to

$$\begin{aligned}
 c_0 &= G_{\text{hpa}}^2 \sigma_x^2 \left[1 - e^{-\frac{A^2}{2\sigma_x^2}} + \frac{1}{2} \sqrt{\frac{\pi}{2}} \frac{A}{\sigma_x} \operatorname{erfc}\left(\frac{A}{\sqrt{2}\sigma_x}\right) \right]^2, \\
 c_1 &= G_{\text{hpa}}^2 \sigma_x^2 \left[\frac{1}{2} \frac{A^2}{2\sigma_x^2} e^{-\frac{A^2}{2\sigma_x^2}} + \frac{1}{4} \sqrt{\frac{\pi}{2}} \frac{A}{\sigma_x} \operatorname{erfc}\left(\frac{A}{\sqrt{2}\sigma_x}\right) \right]^2, \\
 c_n &= \frac{G_{\text{hpa}}^2 2\sigma_x^2}{n! (n+1)!} \left[\left(\frac{(2n)!}{2^{2n} n!} - \sum_{i=0}^{n-2} \delta_{i,n} \left(\frac{A^2}{2\sigma_x^2} \right)^{i+1} \right) \frac{A^2}{2\sigma_x^2} e^{-\frac{A^2}{2\sigma_x^2}} + \frac{(2n)!}{2^{2n+1} n!} \sqrt{\frac{\pi}{2}} \frac{A}{\sigma_x} \operatorname{erfc}\left(\frac{A}{\sqrt{2}\sigma_x}\right) \right]^2, \\
 n &= 2, \dots, \infty,
 \end{aligned} \tag{5.143}$$

which avoids the infinite series, and where $\delta_{i,n}$ are opportune coefficients that can be computed recursively [15].

Otherwise, if $f_{\text{NL}}[\cdot]$ does not amplify the signal (e.g., $G_{\text{hpa}} = 1$) and separately clips to A the real and imaginary components of the OFDM signal by means of (5.133), than the coefficients are expressed by [215]

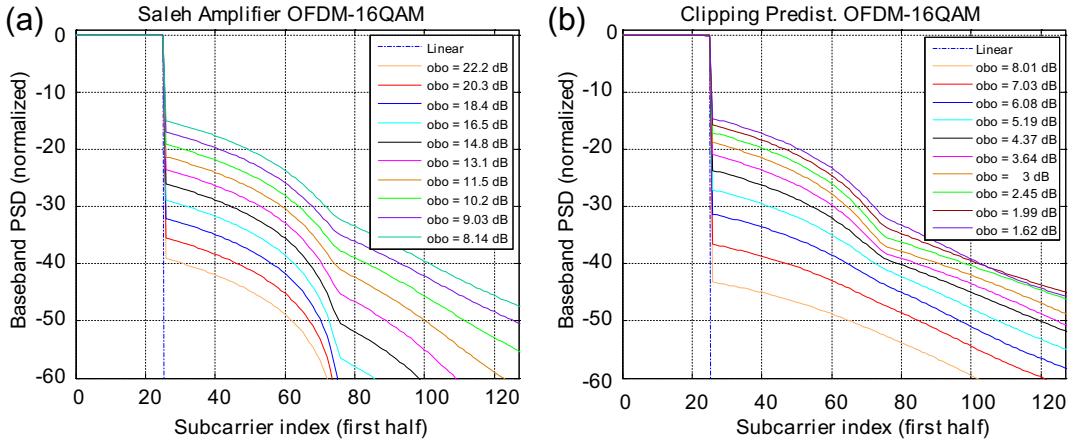
$$\begin{aligned}
 c_0 &= 2\sigma_x^2 \left[\operatorname{erf}\left(\frac{A}{\sqrt{2}\sigma_x}\right) \right]^2, \\
 c_n &= \frac{1}{\pi} \frac{2\sigma_x^2}{4^{n-1} (2n+1)!} H_{2n-1}^2 \left(\frac{A}{\sqrt{2}\sigma_x} \right) e^{-\frac{A^2}{\sigma_x^2}}, \quad n \geq 1,
 \end{aligned} \tag{5.144}$$

where $H_n(x)$ is the Hermite polynomial of order n [2].

In the following, we report the normalized PSD for the output of an HPA, that has been perfectly predistorted [118] up to its saturation input power, whose residual AM/AM distortion is summarized by the envelope clipper in (5.141), and the coefficients $\{c_n\}$ in (5.139) are expressed by (5.142) or (5.143). Note that Figure 5.19 has been obtained using some simplifying assumptions: (a) the input PSD $S_{xx}^{(\text{in})}(f)$ is approximated by a rectangular shape, which is acceptable for OFDM modulations; (b) the bandwidth has been normalized with respect to the subcarrier separation; (c) there is no CP; (d) the input signal is Gaussian distributed, as it happens by CLT when OFDM employs a high number M of carriers. Further results on the performance of practical (i.e., not ideal) HPA predistorters can be found in [6, 16, 67, 114].

The PSD curves in Figure 5.19 are drawn for different ratios among the average input power $2\sigma_x^2$ and the input saturation power A^2 , as captured by the amplifier input power back-off (ibo), defined by⁵ $\text{ibo} = A^2/(2\sigma_x^2)$. It can be noted that a higher ibo corresponds to a lower distortion and, consequently, a lower intermodulation spectral regrowth. However, a lower ibo also induces a lower HPA output power and, consequently, a lower coverage and BER performance degradation. In order to quantify the average

⁵Note that, in general, the coefficients $\{c_n\}$ in (5.140) depends on $\text{ibo} = A^2/(2\sigma_x^2)$, as confirmed by the specific cases of (5.142), (5.143), and (5.144).


FIGURE 5.19

PSD at the output of (a) Saleh-amplifier and (b) Ideal-predictor for different obos.

output power, it is useful to define the output power back-off (obo) that, by CLT arguments [15], is accurately approximated by

$$\text{obo} = G_{\text{hpa}}^2 A^2 / E \left\{ |x_k^{(\text{out})}(t)|^2 \right\} \approx \text{ibo} / (1 - e^{-\text{ibo}}). \quad (5.145)$$

In order to quantify the performance degradation induced by the HPA NL distortion, it is useful to exploit (5.134)–(5.136), which let to express the signal received through an LTI frequency-selective channel $h(t)$, during the k th block, by

$$y_k(t) = x_k^{(\text{out})}(t) * h(t) + w_k(t) = \alpha_0 x_k^{(\text{in})}(t) * h(t) + d_k^{(\text{NL})}(t) * h(t) + w_k(t). \quad (5.146)$$

The NL distortion noise $d_k^{(\text{NL})}(t)$ (which has a CP-based structure like $x_k^{(\text{in})}(t)$), is first filtered by the propagation channel and successively adds-up with the thermal receiver noise $w_k(t)$. OFDM data modulation and detection are based on the signal spectrum $X_k^{(\text{in})}(f)$ and, by means of (5.45)–(5.47), after CP removal and DFT processing, the data received on the m th subcarrier is expressed by

$$Y_{m,k} = \alpha_0 H_m X_{m,k} + H_m D_{m,k}^{(\text{NL})} + W_{m,k}, \quad m = 0, \dots, M-1. \quad (5.147)$$

The term $D_{m,k}^{(\text{NL})} = \sum_{n=0}^{M-1} d_k^{(\text{NL})}(n T_s) e^{-j \frac{2\pi}{M} mn}$ represents the NL distortion noise spectrum on the m th subcarrier. Note that, although $d_k^{(\text{NL})}(t)$ is not Gaussian distributed, if the number of subcarriers is sufficiently high, by CLT arguments $D_{m,k}$ is almost Gaussian distributed. Actually, it is also necessary that $d_k^{(\text{NL})}(n T_s) \neq 0$, which means that the number of distorted samples in a single OFDM block should be high enough, as it happens for instance in (5.141) when the ibo is not too high, or equivalently, when the obo is significantly different from the obo.

Under this assumption, conditionally on a fixed channel realization H_m , the data $X_{m,k}$ transmitted on the m th subcarrier during the k th block is impaired by an overall additive noise $W'_{m,k} = H_m D_{m,k}^{(\text{NL})} + W_{m,k}$, which is Gaussian distributed, with power $\sigma_{m,k}^2 = |H_{m,k}|^2 E \left\{ |D_{m,k}^{(\text{NL})}|^2 \right\} + \sigma_w^2$, similarly to (5.81). Thus, the optimum detector employs a PSE and, similarly to (5.49) the ML detector is expressed by

$$\hat{d}_{m,k} = \arg \min_{d \in \mathcal{D}} \left\{ |d - Y_{m,k}/(\alpha_o H_m)| \right\}, \quad m = 0, \dots, M-1. \quad (5.148)$$

The BER performance associated with (5.148), conditioned on a fixed channel H_m , is the classical one under AWGN assumption [183], as expressed by

$$\text{SER}_{m|H_m} = P\{\hat{d}_{k,m} \neq d_{d,m}|H_m\} = \alpha_Q Q\left(\beta_Q \text{SNR}_{m|H_m}\right), \quad (5.149)$$

where $\text{SNR}_{m|H_m}$ is the conditional SNR on the m th subcarrier, as expressed by

$$\text{SNR}_{m|H_m} = \frac{|\alpha_0|^2 |H_m|^2 \sigma_d^2}{\sigma_{m,k}^2}, \quad (5.150)$$

and α_Q and β_Q are two constellation-dependent parameters. By observing that the power of the distortion noise on the m th subcarrier is represented by the sample of the NL distortion PSD $S_{dd}(f)$ on the m th subcarrier [15]

$$\sigma_{m,k}^2 = |H_m|^2 E \left\{ |D_{m,k}^{(\text{NL})}|^2 \right\} + \sigma_w^2 = |H_m|^2 S_{dd}(m/MT_s) + \sigma_w^2, \quad (5.151)$$

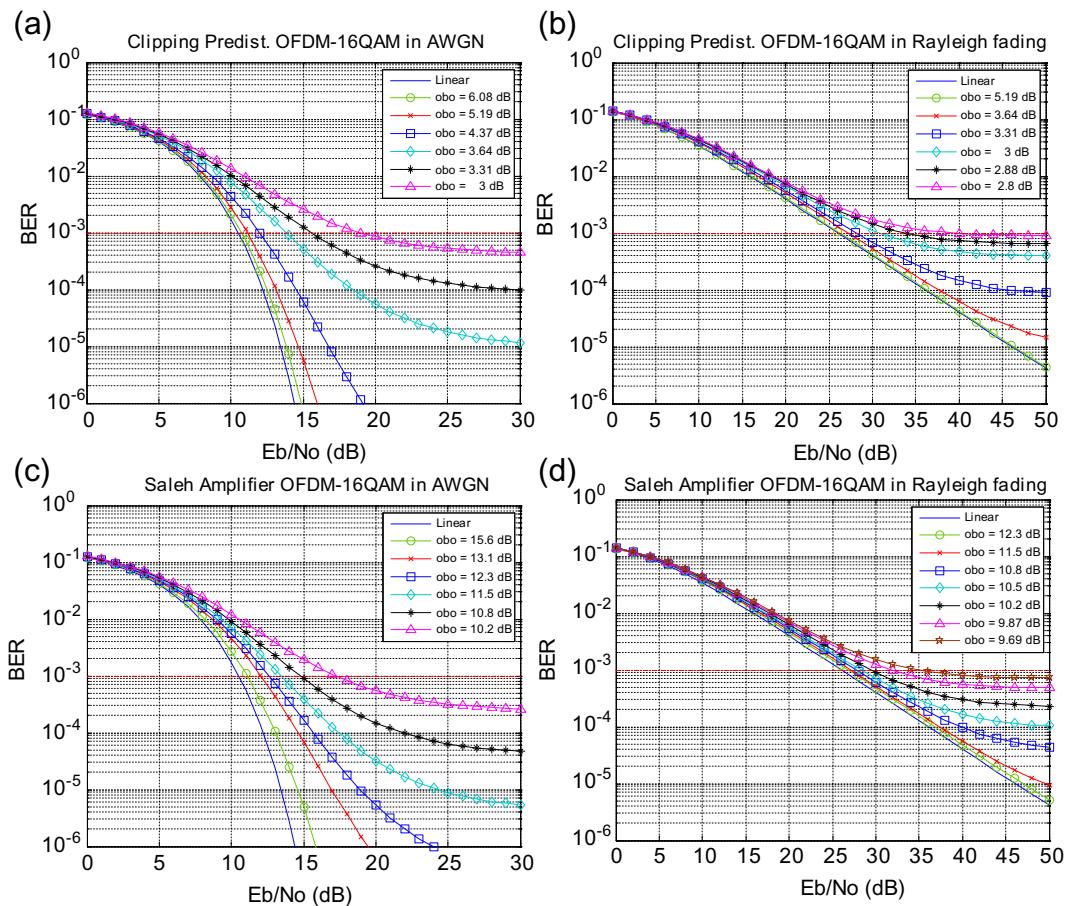
it is clear that $\text{SER}_{m|H_m}$ can be computed in closed form for several NL distortions by plugging in $\text{SNR}_{m|H_m}$ the values of α_0 and $\sigma_{m,k}^2$ obtained by (5.139) and (5.140).

If the channel $h(t)$ is not fixed from block to block, but it (slowly and randomly) changes from block to block, then the conditional $\text{SER}_{m|H_m}$ should be averaged on the $\text{SNR}_{m|H_m}$ statistic, which actually depends on the pdf of $|H_m|$ and, similarly to (5.83), is expressed by

$$\text{SER}_m = P\{\hat{d}_{k,m} \neq d_{d,m}\} = \int_0^{+\infty} \alpha_Q Q \left(\sqrt{\beta_Q \frac{|\alpha_0|^2 |H_m|^2 \sigma_d^2}{S_{dd}(m/MT_s) |H_m|^2 + \sigma_w^2}} \right) f_{|H_m|}(|H_m|) d |H_m|. \quad (5.152)$$

When the channel $h(t)$ in (5.35) is zero-mean Gaussian distributed, then $|H_m|$ in (5.63) is Rayleigh distributed and the integral in (5.152) can be computed in closed form by resorting to generalized hypergeometric functions, as detailed in [189].

Figure 5.20 shows the uncoded SER performance obtained by (5.152) for an OFDM system impaired by a Saleh-Amplifier [197] and the improvement in performance obtained by an ideal predistortion

**FIGURE 5.20**

SER of 16-QAM as a function of the obo, for (c,d) Saleh-amplifier and (a,b) Ideal-predistortion in (a,c) AWGN and (b,d) Rayleigh fading channels.

strategy (e.g., envelope clipping). As can be observed in Figure 5.20, the SER performance degrades for decreasing obos. However low values of the obo is appealing for HPA efficiency and also for higher transmitted power and, consequently, coverage. Actually, only a part of the higher transmitted power would benefit the useful signal: indeed, as highlighted by (5.134) and (5.147), part of it will increase α_0 and another part will increase the NL distortion noise power σ_d^2 . From the performance point of view there exists an optimum obo that minimizes the SER. Closed form expressions for such an optimum obo as a function of the specific data constellation, specific non-linearity, channel statistic and receiver noise power σ_n^2 are not easy to derive, and involve minimization of (5.152) with respect to obo. An analogous approach defines another cost function to be minimized with respect to the obo, which is

called total degradation (TD) [116], and it is defined as

$$\text{TD}_{\text{dB}} = \left[\text{SNR}_{\text{dB}}^{(\text{obo})}(\text{SER}_t) - \text{SNR}_{\text{dB}}^{(\text{lin})}(\text{SER}_t) \right] + \text{obo}_{\text{dB}}. \quad (5.153)$$

The TD is the sum of two power penalty terms: one is the excess SNR for a given obo that is requested at the receiver side (due to the insertion of distortion noise) to obtain the same target SER (SER_t) that is obtained in a linear scenario; the other, which is the obo, is the transmitter power penalty induced by backing-off the maximum HPA output power.

Note that the minimum SER and minimum TD criteria are quite similar, but not mathematically equivalent. Obviously also the TD-wise optimum obo depends on the channel statistic, on the specific nonlinearity, and the data constellations that are used [14, 17], as shown in Figure 5.21.

As a last comment it should be observed that Figure 5.21 shows the TD curves obtained by the SER analysis in (5.152). This analysis is based on the assumption that $D_{m,k}^{(\text{NL})}$ in (5.147) is Gaussian distributed and independent through the different subcarriers, which lead to the easy PSE and ML detection in (5.148). Actually, the independence hypothesis is not really true due to the NL distortion introduced on each time-domain sample distributes by DFT across all the subcarriers in the frequency-domain. Thus, if the receiver has knowledge of the NL distortion $f_{\text{NL}}[\cdot]$ that has distorted the signal, it should exploit this information to end-up with a different (non-per subcarrier) ML detector of the transmitted data \mathbf{d}_k . This fact has been investigated in [52, 122, 234] and lets to significantly improve the SER in a wide set of scenarios, by iterative data-aided receivers coupled with an ideal predistortion strategy. Obviously, these smart joint-ML receivers will end-up with different (lower) SER and consequently a different (lower) optimum obo.

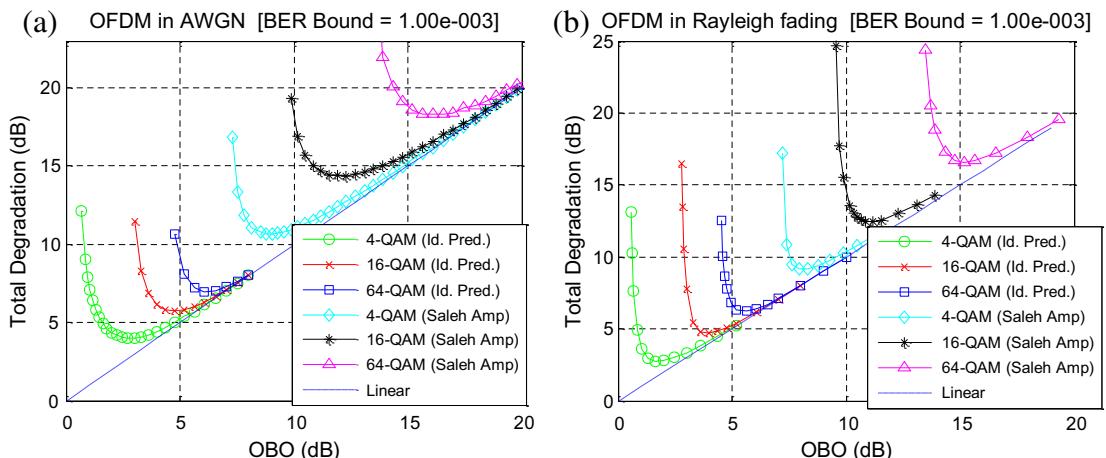


FIGURE 5.21

Total degradation of OFDM systems for Saleh-amplifier and Ideal-predistortion in (a) AWGN and (b) Rayleigh fading channels.

2.05.11.1 PAPR reduction methods

One of the main parameters that characterize the variability of the envelope of $x_k(t)$ in (5.131) is the peak-to-average power ratio (PAPR) of the transmitted signal $x_k(t)$. The baseband PAPR is the ratio between the maximum instantaneous peak power and the average power of the transmitted signal, as expressed by [93]

$$\text{PAPR}_{x_k(t)} = \frac{\max_{kT_u \leq t < (k+1)T_u} \{|x_k(t)|^2\}}{\frac{1}{T_u} \int_{kT_u}^{(k+1)T_u} |x_k(t)|^2 dt}. \quad (5.154)$$

For constant-envelope communications in the time domain, such as frequency modulation (FM), the PAPR is equal to 1, which is the lowest possible value. However, since the OFDM signal in (5.131) is obtained by summing many different subcarrier signals (up to M), its PAPR can be very high. For OFDM with PSK, the PAPR can be as high as the number of subcarriers M , while the PAPR value can be on the order of $3M$ for OFDM signals with QAM [112]. Therefore, significant distortion may be introduced by NL devices, such as HPA when the obo is kept low for power-efficiency reasons. In order to diminish the NL distortion, the PAPR could be reduced by means of suitable techniques that we will briefly describe in the following paragraphs. Since the PAPR is a random variable, PAPR reduction techniques reduce the probability that the PAPR exceeds a fixed threshold [93, 112].

Instead of the continuous-time PAPR (5.154), usually PAPR reduction techniques target the reduction of the discrete-time PAPR, expressed by

$$\text{PAPR}_{x_k[n]} = \frac{\max_{kO_s M \leq n < (k+1)O_s M} \{|x_k[n]|^2\}}{E \{ |x_k[n]|^2 \}}, \quad (5.155)$$

where $x_k[n]$ is a discrete-time version of $x_k(t)$ obtained by sampling at a rate $O_s M / T_u$. In (5.155), O_s is the oversampling rate with respect to the OFDM chip-rate M / T_u , where chip-rate sampling means that the number of samples for each OFDM block duration T_u (excluding the CP) is equal to the number of subcarriers M . Indeed, when $O_s \geq 4$, the discrete-time PAPR (5.155) is a good approximation of the continuous-time PAPR (5.154), and therefore the reduction of the discrete-time PAPR produces a corresponding reduction of the continuous-time PAPR [112]. The discrete-time signal $x_k[n]$ in (5.155) can be easily expressed by

$$x_k[n] = \frac{1}{\sqrt{T_u}} \sum_{m=0}^{O_s M - 1} d_{m,k}^{(a)} e^{j2\pi \frac{m}{O_s M} n}, \quad (5.156)$$

where $d_{m,k}^{(a)}$ is the augmented data sequence of the k th OFDM block, of size $O_s M$, which is obtained by zero padding the original data sequence $d_{m,k}$ of size M .

Among the PAPR reduction techniques, the simplest method is equivalent to an ideally predistorted HPA, which behaves as a clipper of the envelope of the HPA input signal. As clearly explained by (5.145), the clipping method reduces the PAPR as long as the obo is reduced. The main drawback of the clipping method is the distortion introduced at low obo. As explained in the previous discussion, the NL distortion caused by clipping falls both into the signal band, thereby worsening the SER performance, and outside the signal band, thereby producing adjacent channel interference (ACI). The ACI can be

reduced by filtering, but filtering may augment the PAPR. As a consequence, in order to reduce the PAPR, usually a sequence of clipping and filtering operations is performed [8, 93, 112].

A second way to reduce the PAPR is by means of coding techniques [113]. For instance, a convenient codeword design may exclude those codewords that would increase the PAPR, thereby selecting only those codewords that keep the PAPR below a given threshold. However, both the code design and the decoding procedure have an inherent complexity that makes this option suitable only for low values of M [93, 112]. In addition, coding techniques also reduce the data rate for a fixed bandwidth, or equivalently require a bandwidth expansion for a fixed data rate.

A third PAPR reduction approach is based on deliberate introduction of phase offsets into the frequency-domain data symbols, before the IDFT. The underlying idea is to choose those phase offsets that, together with the data-phases, avoids that (almost) all the subcarriers can add in phase during the OFDM block duration. For instance, the technique known as partial transmit sequences (PTS) splits the OFDM block into G_{PTS} sub-blocks: then, each sub-block is multiplied by a phase offset to be chosen from a given set of allowed phase offsets, where Q_{PTS} is the number of possible phase offsets. Basically, the PTS technique evaluates the PAPR that would be obtained for all the $(Q_{\text{PTS}})^{G_{\text{PTS}}-1}$ possible combinations of phase offsets, and then selects the phase offsets that produce the lowest PAPR [93, 112, 169]. Alternatively, the technique known as selected mapping (SLM) multiplies the frequency-domain data block, of size M , with U_{SLM} sequences of known phase offsets, where the length of each phase offset sequence is of course equal to M . Then, among the U_{SLM} sequences, SLM chooses the one that yields the best PAPR reduction [26, 93, 112]. The main disadvantage of PTS and SLM techniques is the computational complexity required for finding the best phase-offset sequence. Moreover, both PTS and SLM have a data-rate loss caused by the transmission of side information, which is necessary at the receiver in order to compensate for the phase offsets [93, 112].

A fourth method for PAPR reduction exploits multiple interleavers [245]. Basically, before the IDFT, the frequency-domain data symbols are scrambled in U_{int} different ways. Since the resulting signals, after the IDFT, have different PAPR values, the transmitter can select the signal with lowest PAPR. Similarly to PTS and SLM, a drawback of the interleaving approach is the data-rate loss caused by the side information transmitted in order to inform the receiver about which deinterleaver should be used [93].

A fifth option for PAPR reduction is known as tone reservation (TR) [233]. In this case, some predetermined subcarriers are not used for data symbols but are reserved for PAPR reduction purposes. Basically, TR adds a data-dependent signal onto the reserved subcarriers, in order to minimize the peak of the time-domain signal. The signal transmitted onto the reserved subcarriers is usually chosen by convex optimization techniques [93]. Obviously, the subcarriers reserved for peak reduction are orthogonal to the data subcarriers, and hence the TR method does not produce any distortion of the data signal. From the receiver complexity viewpoint, the TR technique is one of the simplest methods, since the receiver can simply ignore the subcarriers used for PAPR reduction. However, the TR technique has two drawbacks: a data-rate loss caused by the reserved subcarriers, which are unused for the data symbols, and a power loss caused by the signal transmitted on the reserved subcarriers [93, 112].

Similarly to TR, the tone injection (TI) method adds a data-dependent signal onto some specific subcarriers in order to reduce the PAPR [233]. However, differently from TR, the TI method superimposes the added signal onto the data subcarriers, in order to avoid the data-rate loss. To avoid signal distortion, the constellation size is redundantly increased, so that multiple points of the constellation correspond

to the same information symbol: given the information symbol, the choice of the specific constellation point is performed aiming at reducing the PAPR. Unfortunately, also the TI method suffers from power loss, because of the increased constellation size [93]. A technique closely related to TI is called active constellation extension (ACE) [129]. In this case, the PAPR is reduced by allowing some change of the points of the constellations: the PAPR reduction is obtained by conveniently moving away the outer points of the constellations, in a data-dependent and subcarrier-dependent way [93]. The ACE method does not complicate the receiver, because the detection thresholds can remain the same. However, in this case the detection will not be ML. Besides, the dynamic change of the outer constellation points generates a power loss [93].

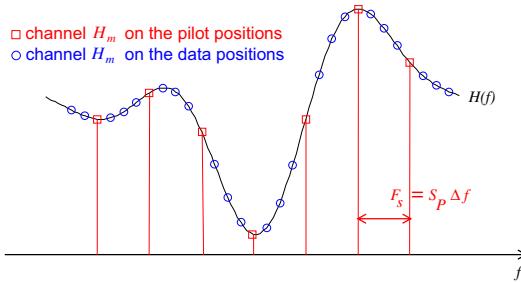
A further PAPR reduction method makes use of companding after the IDFT [112, 249]. Similarly to clipping, the companding method applies an instantaneous nonlinear function to the time-domain OFDM signal. However, the companding technique smoothly modifies the OFDM signal in order to expand the weak amplitudes and compress the high amplitudes, which obviously reduces the PAPR. The nonlinear function can be designed using the probability density function of OFDM signals, which is approximately Gaussian, in order to target a predetermined pdf of the compounded signal [112]. Using a monotone nonlinear function, the receiver can recover the original signal by applying inverse companding. The main drawbacks of the companding approach are two: the ACI generated by the nonlinear function, and the performance loss caused by inverse companding, which modifies the noise statistics.

The last PAPR reduction technique we are going to mention is frequency-domain linear precoding, by means of (5.93). Actually, the PAPR-optimum non-redundant linear precoder is $\Theta = \mathbf{F}_M$: indeed, applying a DFT precoder before the IDFT, the aggregate effect of DFT and IDFT vanishes, so the resulting signal is CP-SC, which has the minimum discrete PAPR (potentially imposed by not constant modulus data) and much lower PAPR than the corresponding CP-OFDM [199]. Note moreover, that the use of (5.101) would guarantee not only the minimum PAPR but also a diversity optimal design. Clearly, this PAPR reduction method changes the type of transmission, which is not OFDM anymore. By the way, precoding moves the complexity toward the receiver, which now needs both DFT and IDFT for frequency-domain equalization (and more complex ML decoding for the diversity-optimum design).

2.05.12 Estimation of LTI channels for OFDM

From (5.64) and (5.78), it is clear that ML detection of OFDM signals requires an estimate of the frequency-domain channel \mathbf{h}_f , or equivalently an estimate of the discrete-time CIR $\mathbf{h}_t = M^{-1/2} \mathbf{F}_M^H \mathbf{h}_f$. Since the frequency-domain discrete-time channel \mathbf{h}_f is the sampled version of the channel frequency response $H(f)$ of the continuous-time channel, channel estimation tries to reconstruct the continuous-time channel on equispaced frequency locations. Intuitively, if the channel frequency response $H(f)$ is known for some specific frequencies, the whole function $H(f)$ could be reconstructed by means of interpolation, as depicted in Figure 5.22.

For instance, let us assume that $H(f)$ is known for all the frequency locations separated by multiples of $F_s = S_P \Delta f$, where $\Delta f = 1/T_u = 1/(MT_s)$ is the OFDM subcarrier separation, and S_P is a positive integer greater than 1. In this case, the whole $H(f)$ can be reconstructed without error, provided that F_s is chosen in such a way to respect the Shannon sampling theorem. Specifically, in order to avoid

**FIGURE 5.22**

Pilot-based estimation of frequency-selective channels.

aliasing of the CIR in the time-lag domain, $1/F_s$ must be greater than or equal to $(L + 1)T_s$, which is the maximum duration of the continuous-time CIR. This leads to

$$S_P \leq \frac{M}{L + 1}. \quad (5.157)$$

Despite blind methods are also possible [94], commonly channel estimation techniques make use of pilot symbols for channel sounding and consequent interpolation [174]. The pilot symbols can be placed in the time-domain, in the frequency-domain, or in both domains [174]. Usually, for packet transmissions, such as WLAN, time-domain pilot symbols are preferred. For instance, a known pilot symbol can be inserted at the beginning of the frame, and can be used also for synchronization purposes. In this case, channel estimation can be performed using the same methods employed for single-carrier systems. Therefore, herein we omit the mathematical description of channel estimation methods based on time-domain pilot symbols. When the CIR is slowly time varying, in order to update or refine the channel estimate, the pilot symbol can be repeated after some OFDM data blocks. Alternatively to repetition in the time-domain, few subcarriers can be reserved for the frequency-domain allocation of additional pilot symbols, commonly known as pilot tones.

On the other hand, for continuous transmissions, frequency-domain or joint time-frequency-domain pilot symbols are usually adopted. For instance, in DVB-T, there are two sets of pilot tones, denoted as continuous pilot tones and scattered pilot tones [72]. The continuous pilot tones are dedicated subcarriers used only for pilot symbols. The scattered pilot tones are frequency-domain pilot symbols whose frequency location changes on a per OFDM-block basis, and therefore the pilot pattern is designed in the time-frequency-domain.

Frequency-domain pilot symbols seem a natural choice for OFDM systems, because the subcarrier orthogonality keeps the pilot tones separated from the data symbols after the LTI channel. In the following, we describe some channel estimation methods assuming frequency-domain pilot tones.

The main assumptions are:

- a. At the transmitter, P out of M subcarrier are reserved for pilot symbols, while the remaining $D = M - P$ are used for data symbols. We assume $M = PS_P$, where the spacing S_P between adjacent pilot tones is integer. We also assume that the 0th subcarrier is a pilot subcarrier, and that

pilot tones are equispaced. We denote with \mathbf{a}_k the frequency-domain transmitted vector, obtained by interlacing the pilot vector \mathbf{p}_k of size P with the data vector \mathbf{d}_k of size D . The number of pilot symbols is greater than or equal to the number of discrete-time channel paths, that is, $P \geq L + 1$. Pilot symbols are taken from a PSK constellation, i.e., $|[\mathbf{p}_k]_m| = 1$.

- b.** The transmission format is CP-OFDM.
- c.** The multipath channel is assumed time invariant, at least for the duration of an OFDM block, and characterized by Rayleigh fading.
- d.** At the receiver, the number of discrete-time channel paths $L + 1$ is known, and the channel estimation is performed using only a single OFDM block.

The transmitted vector, after IFFT processing and before CP insertion, similarly to (5.50) can be expressed as

$$\mathbf{x}_k = \mathbf{F}_M^H \mathbf{a}_k, \quad (5.158)$$

while the received vector, after CP removal and FFT processing, similarly to (5.65) can be expressed as

$$\mathbf{y}_k^{(f)} = \text{diag}(\mathbf{h}_f) \mathbf{a}_k + \mathbf{w}_k^{(f)} = \text{diag}(\mathbf{a}_k) \mathbf{h}_f + \mathbf{w}_k^{(f)}. \quad (5.159)$$

The receiver now selects the subvector of $\mathbf{y}_k^{(f)}$ that corresponds to the pilot positions by applying the $P \times M$ pilot selection matrix $\mathbf{S}_P = \mathbf{I}_P \otimes [1, \mathbf{0}_{1 \times S_P - 1}]$, as expressed by

$$\mathbf{y}_{k,P}^{(f)} = \mathbf{S}_P \mathbf{y}_k^{(f)} = \text{diag}(\mathbf{p}_k) \mathbf{S}_P \mathbf{h}_f + \mathbf{S}_P \mathbf{w}_k^{(f)}, \quad (5.160)$$

and then compensates for the pilot symbols, as expressed by

$$\begin{aligned} \tilde{\mathbf{h}}_{k,P}^{(f)} &= (\text{diag}(\mathbf{p}_k))^{-1} \mathbf{y}_{k,P}^{(f)} = \mathbf{S}_P \mathbf{h}_f + \tilde{\mathbf{w}}_{k,P}^{(f)} \\ &= \sqrt{M} \mathbf{S}_P \mathbf{F}_M \mathbf{h}_t + \tilde{\mathbf{w}}_{k,P}^{(f)}, \end{aligned} \quad (5.161)$$

where $\tilde{\mathbf{w}}_{k,P}^{(f)} = (\text{diag}(\mathbf{p}_k))^{-1} \mathbf{S}_P \mathbf{w}_k^{(f)}$ is an AWGN term. By defining $\mathbf{F}_M^{(1:L+1)} = \mathbf{F}_M [\mathbf{I}_{L+1} \mathbf{0}_{L+1 \times M-L-1}]^T$ as the $M \times L + 1$ matrix containing the first $L + 1$ columns of \mathbf{F}_M , and observing that $\mathbf{F}_M \mathbf{h}_t = \mathbf{F}_M^{(1:L+1)} \mathbf{h}_t^{(1:L+1)}$, where $\mathbf{h}_t^{(1:L+1)} = [h_0, h_1, \dots, h_L]^T$, (5.160) becomes

$$\tilde{\mathbf{h}}_k^{(f,P)} = \mathbf{B}_P \mathbf{h}_t^{(1:L+1)} + \tilde{\mathbf{w}}_k^{(f,P)}, \quad (5.162)$$

where $\mathbf{B}_P = \sqrt{M} \mathbf{S}_P \mathbf{F}_M^{(1:L+1)}$ is a $P \times L + 1$ matrix. As a consequence, the ML estimate of $\mathbf{h}_t^{(1:L+1)}$ is expressed by [117, 164]

$$\hat{\mathbf{h}}_{\text{ML},t}^{(1:L+1)} = (\mathbf{B}_P^H \mathbf{B}_P)^{-1} \mathbf{B}_P^H \tilde{\mathbf{h}}_k^{(f,P)}, \quad (5.163)$$

and the ML estimate of \mathbf{h}_f is obtained by projecting $\hat{\mathbf{h}}_{\text{ML},t}^{(1:L+1)}$ in the frequency-domain, as expressed by [164]

$$\hat{\mathbf{h}}_{\text{ML},f} = \mathbf{F}_M^{(1:L+1)} \hat{\mathbf{h}}_{\text{ML},t}^{(1:L+1)} = \mathbf{F}_M \left[\left(\hat{\mathbf{h}}_{\text{ML},t}^{(1:L+1)} \right)^T, \mathbf{0}_{1 \times M-L-1} \right]^T. \quad (5.164)$$

Note that ML estimation does not require any knowledge about the noise power. It can be shown that ML estimation achieves the Cramér-Rao lower bound (CRLB) [164], which is expressed by

$$\text{MSE}(H_m) = \frac{L+1}{P} \sigma_w^2. \quad (5.165)$$

If additional information is available at the receiver, such as the channel covariance matrix $\mathbf{C}_{\tilde{\mathbf{h}}_t^{(1:L+1)}}$ and the noise power $\sigma_{\tilde{w}_{k,P}^{(f)}}^2$, the mean-squared error (MSE) of the channel estimate can be further reduced by a Bayesian MMSE estimation approach [117], as expressed by [164]

$$\hat{\mathbf{h}}_{\text{MMSE},t}^{(1:L+1)} = \left(\mathbf{B}_P^H \mathbf{B}_P + \sigma_{\tilde{w}_{k,P}^{(f)}}^2 \mathbf{C}_{\tilde{\mathbf{h}}_t^{(1:L+1)}}^{-1} \right)^{-1} \mathbf{B}_P^H \tilde{\mathbf{h}}_{k,P}^{(f)}. \quad (5.166)$$

Note that, with respect to ML estimation, MMSE estimation is characterized by greater computational complexity, because the matrix inverse in (5.166) depends on the noise power $\sigma_{\tilde{w}_{k,P}^{(f)}}^2$. However, the complexity order for both ML and MMSE methods is $O(P(L+1) + M \log_2 M)$.

A low-complexity alternative to ML and MMSE estimation is linear interpolation in the frequency-domain of the channel estimates over two adjacent pilot tones, obtained as

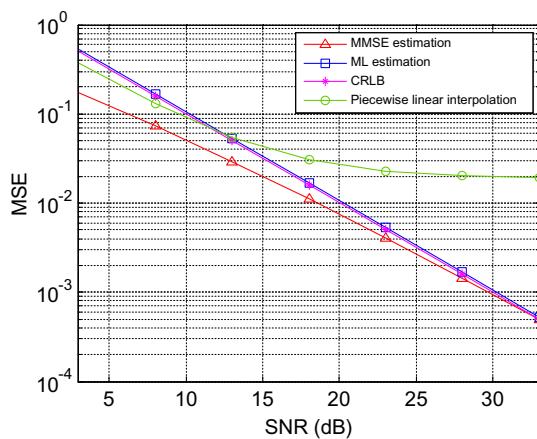
$$\hat{\mathbf{h}}_{\text{lin},f} = \mathbf{L} \left(\tilde{\mathbf{h}}_{k,P}^{(f)} \otimes [1, \mathbf{0}_{1 \times S_P-1}]^T \right), \quad (5.167)$$

where \mathbf{L} is an $M \times M$ symmetric banded circulant nonnegative matrix with elements

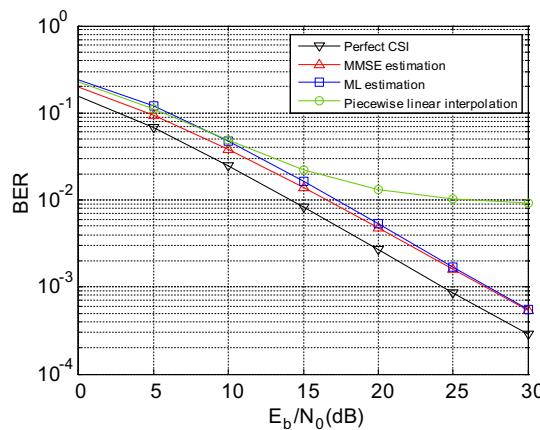
$$[\mathbf{L}]_{m,m+n} = \begin{cases} 1 - \frac{|n|}{S_P}, & 0 \leq |n| \leq S_P - 1, \\ 0, & S_P \leq |n| \leq M - S_P, \\ 1 - \frac{M - |n|}{S_P}, & M - S_P + 1 \leq |n| \leq M - 1. \end{cases} \quad (5.168)$$

Figures 5.23 and 5.24 compare the performance of the above mentioned estimators: ML, MMSE and piecewise linear interpolation. The equations above may be generalized to multicarrier systems with guard frequency bands and non-equispaced pilot subcarriers. In general, the MSE of the channel depends not only on the estimation method, but also on the subcarrier index. Obviously, the MSE is reduced for those subcarriers that are close to the pilot subcarriers. In addition, the MSE tends to increase for the edge subcarriers, i.e., for those subcarriers close to the guard frequency bands. This distortion, sometimes referred to as leakage effect, can be reduced by leakage suppression [210] or by reducing the pilot spacing for the edge subcarriers [164].

The equations above may be generalized also to the estimation of slowly time-varying channels, assuming that the CIR stays constant for one OFDM block, but changes from block to block. In this case, frequency-domain channel estimation may be repeated after a certain number of OFDM blocks. The channel in between two successive estimates may be recovered by time-domain interpolation techniques. This approach, based on successive one-dimensional interpolations, has been adopted in [102] and [132]. Alternatively, joint interpolation/estimation in the time-frequency-domain may be performed, such as the Wiener filtering approaches proposed in [98] and [211].

**FIGURE 5.23**

Channel MSE performance comparison among different channel estimation techniques ($M = 256$, $P = L + 1 = 16$, multipath channel with an exponentially decaying power-delay profile). This figure has been generated using the MATLAB script [perf_chann_estim.m](#).

**FIGURE 5.24**

BER performance comparison among different channel estimation techniques (Uncoded BER, QPSK, $M = 256$, $P = L + 1 = 16$, multipath channel with an exponentially decaying power-delay profile). This figure has been generated using the MATLAB script [perf_chann_estim.m](#).

Adaptive estimators have been proposed in [198, 207], and [203].

In addition to pilot-based techniques, joint pilot-data-aided channel estimators have been proposed by many authors (see the references in [94] and [174]). Joint pilot-data-aided techniques, also known as decision-directed channel estimation algorithms and iterative (turbo) processing methods, make use of estimated data symbols as “virtual pilots” in order to refine the channel estimate and further improve the reliability of the data estimate itself.

2.05.13 OFDM propagation in rapidly time-varying channels

In the presence of a mobile transmitter or receiver, the (multipath) channel becomes time varying, that is, its CIR changes with time. A simple physical interpretation is offered by the Doppler effect, which affects a sinusoid transmitted between two equipments in relative motion.

If a sinusoidal signal $c_T(t) = A \sin(2\pi f_o t + \phi_0)$ is transmitted in a wireless channel, and the receiver is in relative motion with velocity $v = v_\theta / \cos(\theta_0)$ and angle θ_0 with respect to the direction of the wave propagation, the received signal is expressed by [97]

$$c_R(t) = \beta_0 c_T((1 + v_\theta/c)t - \tau_0) = \beta_0 A \sin(2\pi f'_o t + \phi_0 - \Delta\phi_0), \quad (5.169)$$

where c is the speed of the traveling wave, $\Delta\phi_0 = 2\pi f_o \tau_0$, $\beta_0 \in \mathbb{R}$ is an attenuation coefficient, τ_0 is the propagation delay in the absence of movement, and the perceived frequency f'_o at the receiver side is expressed by

$$f'_o = f_o \pm f_D \cos(\theta_0) = f_o \pm f_o \frac{v}{c} \cos(\theta_0). \quad (5.170)$$

In (5.170), f_D is the maximum Doppler frequency experienced when the two equipments are moving towards ($+f_D$), or away ($-f_D$) from each other. Equation (5.170) establishes that the receiver perceives a Doppler shift $f_{d0} = f_D \cos(\theta_0)$ with respect to the nominal frequency f_o . If the transmitted carrier reaches the receiver through different paths, the receiver perceives a different frequency shift f_{di} for each path, depending on the different angle of arrival of each path with respect to the direction of motion (see Figure 5.25).

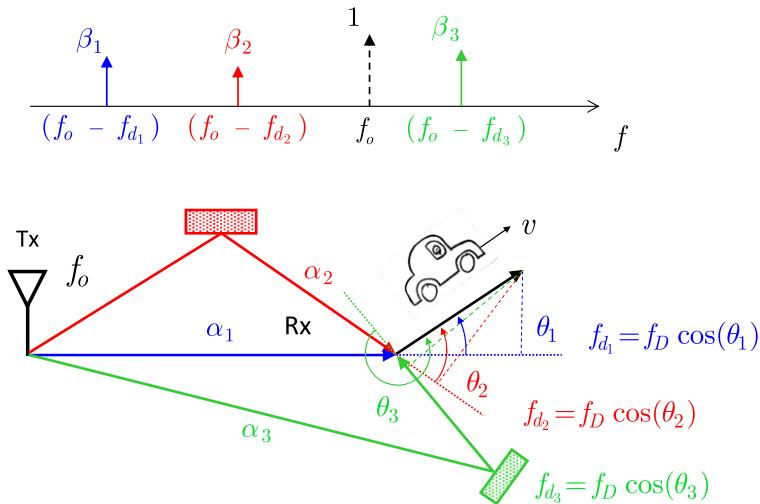
If the transmitted waveform $x_T^{(rf)}(t) = x_T(t)c_T(t)$ is amplitude modulated by a signal $x_T(t)$, whose bandwidth $B \ll f_0$, i.e., the signal is almost constant during each period of the carrier, the signal received through a single path can be approximated as

$$x_R^{(rf)}(t) = x \left(\left(1 \pm \frac{v_\theta}{c} \right) t - \tau_0 \right) c_R(t) \simeq \beta_0 x_T(t - \tau_0) A \sin(2\pi f'_o t + \phi_0 - \Delta\phi_0), \quad (5.171)$$

where $v_\theta = v \cos(\theta)$ is the component of the velocity in the direction of the wave propagation. The BB equivalent of right-hand side of (5.171) is expressed by

$$x_R(t) = x_T(t - \tau_0) \beta_o e^{-j\Delta\phi_0} e^{j2\pi f_{d0} t} = \alpha_0 e^{j2\pi f_{d0} t} x_T(t - \tau_0), \quad (5.172)$$

where $\alpha_0 = \beta_0 e^{-j\Delta\phi_0} \in \mathbb{C}$ is the complex attenuation coefficient in a time invariant set-up, i.e., when $v_\theta = 0$ and $f_{d0} = 0$. Note that the approximation in (5.171) means that the change in time-scale can be safely ignored when $v_\theta \ll c$, for instance for electromagnetic propagation, where c is the speed of

**FIGURE 5.25**

Multiple Doppler shifts for a moving receiver.

light, while this is not the case for ultrasound communications in underwater acoustic, where c is the speed of sound [97].

It is straightforward to verify that (5.172) is obtained by the time-varying convolution operator $*_t$, summarized by

$$x_R(t) = h_0(t, \tau) *_t x_T(t) = \int_{-\infty}^{+\infty} h_0(t, \tau) x_T(t - \tau) d\tau, \quad (5.173)$$

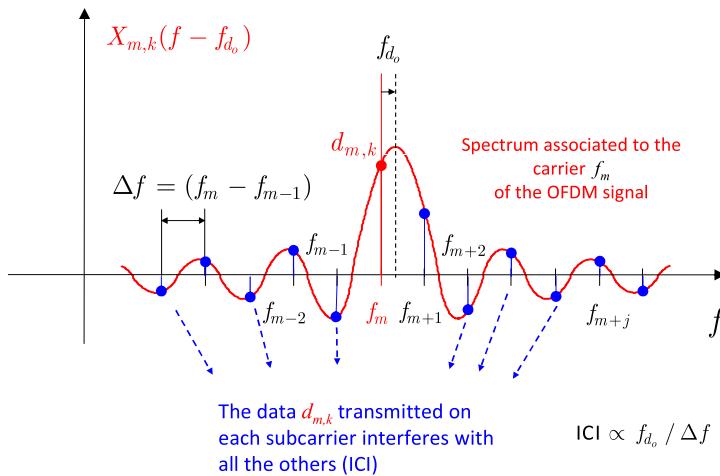
where

$$h_0(t, \tau) = \alpha_0 e^{j2\pi f_{d0} t} \delta(\tau - \tau_0) = \alpha_0(t) \delta(\tau - \tau_0) \quad (5.174)$$

is the time-varying CIR associated to a single-path propagation channel with propagation delay τ_0 . By means of (5.172) and (5.173), the effect of a single time-varying path on the transmitted signal is a delay, and the multiplication for a time-varying signal $\alpha_0(t) = \alpha_0 e^{j2\pi f_{d0} t}$. Thus, the effect on its spectrum is a linear phase shift, plus a convolution with the spectrum $\mathcal{F}\{\alpha_0(t)\} = \alpha_0 \delta(f - f_{d0})$, which causes a frequency shift f_{d0} , as expressed by

$$X_R(f) = X_T(f - f_{d0}) e^{-j2\pi(f - f_{d0})\tau_0}. \quad (5.175)$$

Thus, if $x_T(t) = x_k(t)$ is the (useful) signal transmitted during the k th OFDM symbol, and expressed by (5.25), a single-path linear time-varying (LTV) channel as in (5.174) would cause a shift at the receiver side of the frequency spectrum $X_k(f)$ in (5.26) and Figure 5.4. If left uncompensated, this frequency shift would introduce ICI, because the zeros of the spectrum of each subcarrier are shifted from f_m to $f_m - f_{d0}$, while the DFT processing at the receiver produces a sampling at the original frequencies f_m , as clarified in Figure 5.26. Note that this is the same effect introduced in AWGN channels, when

**FIGURE 5.26**

ICI introduced by a Doppler shift or carrier frequency offset (CFO).

there is a carrier frequency offset (CFO) between the receiver RF local oscillator, tuned at a frequency $f'_o = f_o - f_{d_0}$, and the transmitter RF local oscillator tuned at f_o .

Obviously, a CFO or a single-path LTV channel can be compensated at the receiver side by a single-tap time-varying equalizer, by simply multiplying the received signal $x_R(t)$ in (5.172) by $\alpha_0^{-1} e^{-j2\pi f_{d_0} t}$. However, the RF propagation generally experiences N_p different paths that emerge with the same delay τ_p , and consequently a more general expression for the time-varying CIR is

$$\begin{aligned} h_c(t, \tau) &= \sum_{p=0}^{P-1} h_p(t, \tau) = \sum_{p=0}^{P-1} \alpha_p(t) \delta(\tau - \tau_p) \\ &= \sum_{p=0}^{P-1} \sum_{n=0}^{N_p-1} \alpha_{n,p} e^{j2\pi f_{d_{p,n}} t} \delta(\tau - \tau_p), \end{aligned} \quad (5.176)$$

where $\alpha_{n,p}$ and $f_{d_{n,p}}$ are the complex attenuation and the Doppler shift, respectively, associated to the n th path that experiences the delay τ_p , and $\alpha_p(t) = \sum_{n=0}^{N_p-1} \alpha_{n,p} e^{j2\pi f_{d_{p,n}} t}$ is the time-varying gain associated to each path. The spectrum of $\alpha_p(t)$ would be consequently composed by several frequencies, as expressed by $\mathcal{F}\{\alpha_p(t)\} = \sum_{n=0}^{N_p-1} \alpha_{n,p} \delta(f - f_{d_{p,n}})$, and the received signal would be composed by several shifted replicas of the transmitted signal, as expressed by

$$\begin{aligned} X_R(f) &= \sum_{p=0}^{P-1} \left(X_T(f) e^{-j2\pi f \tau_p} \right) * \mathcal{F}\{\alpha_p(t)\} \\ &= \sum_{p=0}^{P-1} \sum_{n=0}^{N_p-1} \alpha_{n,p} X_T(f - f_{d_{p,n}}) e^{-j2\pi(f - f_{d_{p,n}}) \tau_p}. \end{aligned} \quad (5.177)$$

Equation (5.177) reveals that the sum of the P convolutions introduces ICI, which needs for a more complex mitigation technique rather than the simple frequency-shift compensation used for $P = 1$ [see (5.175) and discussion after (5.175)]. More generally, also the directions of the user movement may change with time, making each path characterized by a time-varying angle $\theta_{p,n}(t)$ and time-varying frequencies $f_{d_{p,n}}(t)$. Consequently, $\alpha_p(t) = \sum_{n=0}^{N_p-1} \alpha_{n,p} e^{-j2\pi f_{d_{p,n}}(t)t}$ is characterized by a continuous spectrum $A_p(f)$ that destroys the sinc shape associated to each subcarrier of $X_k(f)$ in (5.26), by changing the zeros and leading to ICI, as expressed by

$$\begin{aligned} X_R(f) &= \sum_{p=0}^{P-1} \left(X_T(f) e^{-j2\pi f \tau_p} \right) * A_p(f) \\ &= T_u \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} d_{m,k} (X_{m,k}(f) e^{-j2\pi f \tau_p}) * A_p(f) \\ &= T_u \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} d_{m,k} \int_{-\infty}^{+\infty} \text{sinc} [\pi T_u (\nu - f_m)] A_p(f - \nu) e^{-j[\phi_k(\nu - f_m) + 2\pi \nu \tau_p]} d\nu. \end{aligned} \quad (5.178)$$

The ICI introduced by (5.178) is graphically shown in Figure 5.27, focusing on the effect of a single time-varying tap on the spectrum $X_{m,k}(f)$ of a single subcarrier.

Thus, assuming that the path delays τ_p remain constant with time, when the OFDM signal passes through the LTV channel expressed in first line of (5.176), in the presence of AWGN the received signal

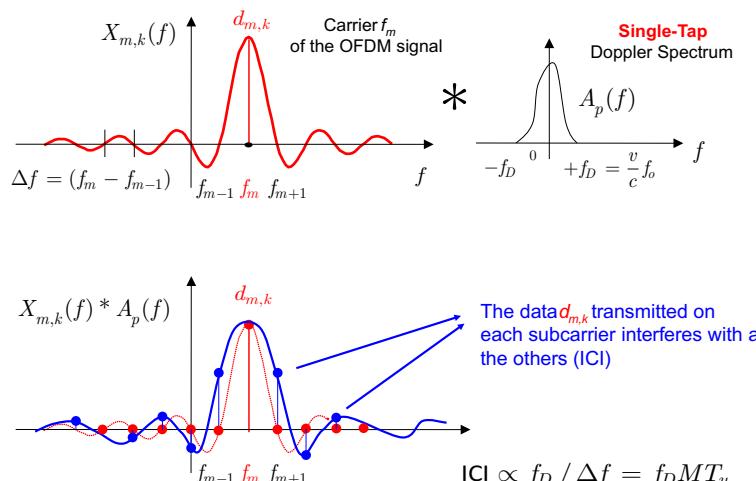


FIGURE 5.27

ICI introduced by Doppler spread.

can be expressed as [27]

$$\begin{aligned}
 y(t) &= \int_{-\infty}^{+\infty} h_c(t, \tau) x(t - \tau) d\tau + w(t) \\
 &= \sum_{p=0}^P \int_{-\infty}^{+\infty} \alpha_p(t) \delta(\tau - \tau_p) x(t - \tau) d\tau + w(t) \\
 &= \sum_{p=0}^P \alpha_p(t) x(t - \tau_p) + w(t),
 \end{aligned} \tag{5.179}$$

which, after filtering [matched to $p(t)$], becomes

$$\begin{aligned}
 y_f(t) &= y(t) * p^*(-t) = \sum_{p=0}^P \int_{-\infty}^{+\infty} \alpha_p(t) x(t - \tau_p) p^*(\tau - t) d\tau + w(t) * p^*(-t) \\
 &\simeq \sum_{p=0}^P \alpha_p(t) [x(t - \tau_p) * p^*(-t)] + w_f(t).
 \end{aligned} \tag{5.180}$$

The last equality in (5.180) is a good approximation under the hypothesis that the attenuation $\alpha_p(t)$ of each path is virtually constant for a time interval with duration T_s , which is the time support of the pulse $p(t)$. Thus, if the transmitted signal is expressed by means of its digital samples $x_n = x(nT_s)$, the received signal is expressed by

$$\begin{aligned}
 y_f(t) &\simeq \sum_{p=0}^P \alpha_p(t) \sum_{i=-\infty}^{+\infty} x_i p(t - \tau_p - iT_s) * p^*(-t) + w_f(t) \\
 &\simeq \sum_{p=0}^P \alpha_p(t) \sum_{i=-\infty}^{+\infty} x_i R_{pp}(t - \tau_p - iT_s) + w_f(t),
 \end{aligned} \tag{5.181}$$

which after uniform sampling with period T_s , leads to a received signal $y_n = y_f(nT_s)$ expressed by

$$\begin{aligned}
 y_n &= \sum_{i=-\infty}^{+\infty} x_i \sum_{p=0}^P \alpha_p(iT_s) R_{pp}((n-i)T_s - \tau_p) + w_f(nT_s) \\
 &= \sum_{i=-\infty}^{+\infty} x_i \tilde{h}_{i,n-i} + w_n = \sum_{l=-\infty}^{+\infty} h_{n,l} x_{n-l} + w_n,
 \end{aligned} \tag{5.182}$$

where $h_{n,l} = \tilde{h}_{n-l,l} = \sum_{p=0}^P \alpha_p(n-l) R_{pp}(lT_s - \tau_p)$ is the discrete-time equivalent time-varying CIR.

2.05.13.1 Carrier frequency offset (CFO)

As briefly anticipated after (5.175), a special case of LTV channel is obtained in the presence of a CFO f_{d_o} between the transmitter and the receiver oscillators, i.e., in the absence of (perfect) frequency

synchronization. In this case, each multipath component is affected by the same time-varying coefficient, and therefore the discrete-time channel becomes $h_{n,l} = e^{j2\pi\varepsilon n/M} \sum_{p=0}^P \alpha_p R_{pp}(lT_s - \tau_p)$, where $\varepsilon = f_{d_o} T_u = f_{d_o}/\Delta f$ represents the CFO normalized to the subcarrier spacing $\Delta f = 1/T_u = 1/(MT_s)$. Thus, instead of (5.60), the k th received vector after CP removal is expressed by [188]

$$\begin{aligned}\mathbf{y}_k^{(\text{cp})} &= e^{j2\pi\varepsilon(kN+L_g)/M} \text{diag}(\mathbf{e}) \mathbf{H}_t^{(\text{cp})} \mathbf{x}_k + \mathbf{w}_k \\ &= e^{j2\pi\varepsilon(kN+L_g)/M} \text{diag}(\mathbf{e}) \mathbf{H}_t^{(\text{cp})} \mathbf{F}_M^H \mathbf{d}_k + \mathbf{w}_k,\end{aligned}\quad (5.183)$$

where $\mathbf{e} = [1, e^{j2\pi\varepsilon/M}, \dots, e^{j2\pi\varepsilon(M-1)/M}]^T$ and $N = M + L_g$. By applying the DFT on the received vector $\mathbf{y}_k^{(\text{cp})}$, (5.183) becomes

$$\mathbf{y}_k^{(f)} = \mathbf{F}_M \mathbf{y}_k^{(\text{cp})} = e^{j2\pi\varepsilon(kN+L_g)/M} \mathbf{E} \text{diag}(\mathbf{h}_f) \mathbf{d}_k + \mathbf{w}_k^{(f)}, \quad (5.184)$$

where $\mathbf{E} = \mathbf{F}_M \text{diag}(\mathbf{e}) \mathbf{F}_M^H$ is circulant, with (m, n) th element expressed by

$$[\mathbf{E}]_{m,n} = \frac{\sin(\pi((n-m)_{\text{mod } M} + \varepsilon))}{M \sin(\pi((n-m)_{\text{mod } M} + \varepsilon)/M)} e^{j\pi \frac{M-1}{M} ((n-m)_{\text{mod } M} + \varepsilon)}. \quad (5.185)$$

As a result of (5.184) and (5.185), a CFO produces three effects on the received signal, described in the following. First, since \mathbf{E} is not diagonal, $\mathbf{y}_k^{(f)}$ contains interference among the data elements of the vector \mathbf{d}_k . This is the ICI, introduced by the frequency-shift of each subcarrier spectrum, which is no longer sampled on its zero-crossing locations, as shown in Figure 5.26. Therefore, in order to compensate this first effect, equalization or ICI cancellation is necessary. Second, since the elements on the main diagonal of \mathbf{E} have modulus lower than one, the energy of the useful signal is reduced by a factor $\text{sinc}^2(\pi\varepsilon)$, as also clarified by Figure 5.26. Note that \mathbf{E} is unitary, and hence this energy reduction is exactly balanced by the energy of the ICI. Also this second effect can be compensated using equalization or ICI cancellation techniques. If the ICI is ignored at the equalizer, such as the conventional equalizer of (5.67), the signal-to-interference ratio (SIR) on the estimated symbol is [185]

$$\text{SIR} = \frac{\text{sinc}^2(\pi\varepsilon)}{1 - \text{sinc}^2(\pi\varepsilon)} \approx \frac{3}{(\pi\varepsilon)^2} - 1, \quad (5.186)$$

where the last approximation is reasonable for $\varepsilon < 0.15$. Third, due to the complex exponential term $e^{j2\pi\varepsilon(kN+L_g)/M}$ in (5.184), the CFO also introduces a phase shift that depends on the OFDM block index. This phase shift can be easily estimated (and successively compensated) by using pilot subcarriers or, equivalently, the estimation (and equalization) can be included in the (block-by-block) channel estimation techniques described in Section 2.05.12. The estimation of the CFO is the frequency synchronization problem for OFDM, and as such it will be addressed in Section 2.05.15.

2.05.13.2 LTV channel equalization

Omitting the OFDM symbol index for notational compactness, in the case of a general LTV channel, after discarding the CP the input-output relation expressed by (5.182) can be rewritten in vector-matrix

form as

$$\begin{aligned}\mathbf{y}^{(\text{cp})} &= \mathbf{H}_t^{(\text{cp})} \mathbf{x} + \mathbf{w} \\ &= \mathbf{H}_t^{(\text{cp})} \mathbf{F}_M^H \mathbf{d} + \mathbf{w},\end{aligned}\quad (5.187)$$

where $\mathbf{y}^{(\text{cp})}$, \mathbf{d} , and \mathbf{w} are vectors of size M that represent the received time-domain elements, the transmitted frequency-domain symbols, and the received AWGN, respectively, for a generic OFDM block, and $\mathbf{H}_t^{(\text{cp})}$ stands for the $M \times M$ time-domain system matrix. From (5.188), it is evident that each diagonal of $\mathbf{H}_t^{(\text{cp})}$ contains different elements, i.e., the discrete-time evolution of each channel tap. In other words, differently from the time-invariant channel case, $\mathbf{H}_t^{(\text{cp})}$ in (5.188) is not circulant and therefore cannot be diagonalized using FFT matrices.

$$\mathbf{H}_t^{(\text{cp})} = \begin{bmatrix} h_{0,0} & 0 & \cdots & 0 & h_{0,L} & \cdots & h_{0,1} \\ h_{1,1} & h_{1,0} & 0 & \ddots & 0 & \ddots & \vdots \\ \vdots & h_{2,1} & \ddots & \ddots & \ddots & \ddots & h_{L-1,L} \\ h_{L,L} & \vdots & \ddots & \ddots & \ddots & \vdots & 0 \\ 0 & h_{L+1,L} & \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & h_{M-2,0} & 0 \\ 0 & \cdots & 0 & h_{M-1,L} & \cdots & h_{M-1,1} & h_{M-1,0} \end{bmatrix}. \quad (5.188)$$

By applying an FFT we obtain

$$\mathbf{y}^{(f)} = \mathbf{F}_M \mathbf{y}^{(\text{cp})} = \mathbf{H}_f \mathbf{d} + \mathbf{w}^{(f)}, \quad (5.189)$$

where \mathbf{H}_f is not diagonal. Similarly to the CFO case, also a general LTV channel introduces ICI [144] onto the frequency-domain received vector $\mathbf{y}^{(f)}$, because the data symbol d_m transmitted on the m th subcarrier is received also on the subcarriers with index different than m , as shown in Figure 5.27.

In order to reduce the ICI, different equalization techniques are possible [41, 56, 90, 97, 204, 193]. For instance, a zero-forcing (ZF) equalizer produces the data estimate

$$\hat{\mathbf{d}}_{\text{ZF}} = \mathbf{H}_f^{-1} \mathbf{y}^{(f)} = \mathbf{d} + \mathbf{H}_f^{-1} \mathbf{w}^{(f)}. \quad (5.190)$$

This means that, in general, channel estimation is necessary for the availability of \mathbf{H}_f . When the channel is rapidly time-varying, channel estimation must be repeated for each block. An example of channel estimation for LTV channels will be described later on. Moreover, ZF equalization requires a matrix inversion whose complexity is $O(M^3)$ per block. Therefore, when the number of subcarrier is large, the pure ZF approach of (5.190) is impractical. In this case, the equalizer may exploit some specific structured models for the LTV channel in order to enable fast inversion algorithms of \mathbf{H}_f . By observing Figure 5.27, because of the sinc(\cdot) shape of each subcarrier spectrum, it is clear that most of the ICI is introduced in the nearby subcarriers. Equivalently, by observing the signal received on a single subcarrier, it is reasonable to assume that most of the ICI comes from the nearby subcarriers and, consequently, \mathbf{H}_f is modeled as banded and can be inverted with complexity $O(M \log M)$ per block. In addition, the frequency-domain channel matrix \mathbf{H}_f is usually ill-conditioned [89], and therefore

the noise enhancement produced by $\mathbf{H}_f^{-1}\mathbf{w}^{(f)}$ may be excessive. To reduce the noise enhancement introduced by ZF equalization, classical linear MMSE equalization may be applied, as expressed by

$$\hat{\mathbf{d}}_{\text{MMSE}} = \left(\mathbf{H}_f^H \mathbf{H}_f + (\sigma_w^2 / \sigma_d^2) \mathbf{I}_M \right)^{-1} \mathbf{H}_f^H \mathbf{y}^{(f)}, \quad (5.191)$$

which, analogously to the ZF solution, can be implemented with reduced complexity by exploiting the almost banded structure of \mathbf{H}_f [41, 194, 204, 193]. A BER performance comparison of some different ZF and MMSE alternatives is shown in Figure 5.28.

Alternatively, nonlinear techniques may be employed, using decision feedback methods for ICI cancellation or any other tree-search approach. An example of the performance advantages provided by DFE methods with respect to linear MMSE approaches is shown in Figure 5.29.

Iterative or turbo approaches that can incorporate channel decoding have also been proposed [75, 97, 204]. When the channel statistics are known at the receiver, the equalizer may also exploit this knowledge by applying a pre-processing stage with specially designed windows [304]. Nonlinear and statistic-aided methods usually outperform linear equalization, at the price of somewhat increased complexity [97]. A performance comparison, for some potential NL equalization alternatives is shown in Figure 5.30.

2.05.13.3 LTV channel estimation

The estimation of a multipath LTV channel is a non trivial task, because each channel path changes in time. Thus, during the transmission of an M -length OFDM block, there are $M(L + 1)$ unknowns to be estimated in the discrete-time-domain, which are contained in the circulant channel matrix $\mathbf{H}_t^{(\text{cp})}$ in (5.188). However, these unknowns are not independent from each other, especially those that represent

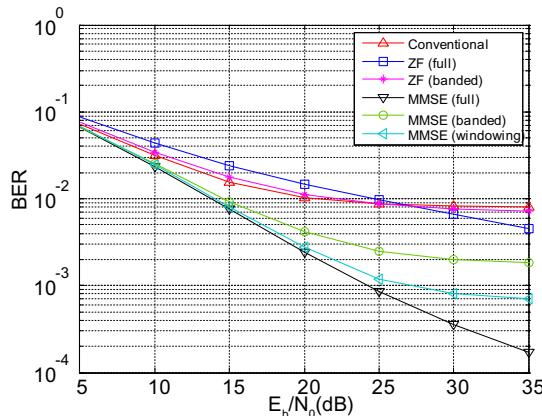
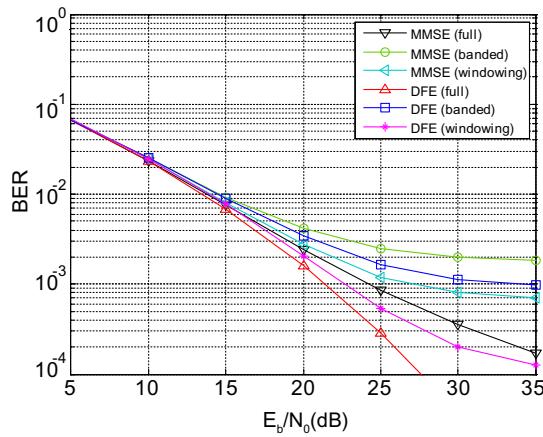
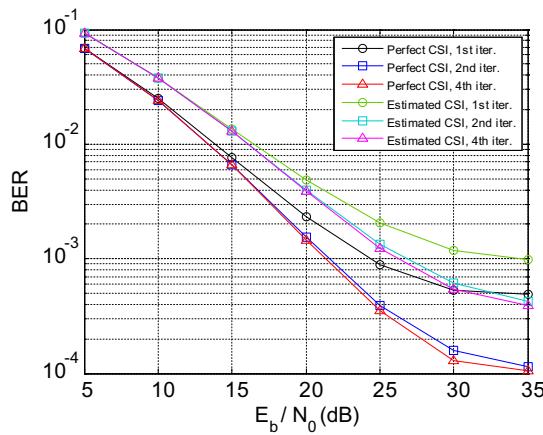


FIGURE 5.28

BER performance comparison among ZF and MMSE equalizers for LTV channels [Uncoded BER, QPSK, $M = 128$ subcarriers with $M_{\text{act}} = 96$ consecutive data subcarriers and $M_{\text{vir}} = M - M_{\text{act}} = 32$ zero subcarriers, $D_b = 2$ for banded and window-aided equalizers, MBAE-SOE window [195], $L = 7$, Multipath channel with uniform power-delay profile, $\max \{f_{dp,n}\} / \Delta f = 0.1$, perfect channel state information (CSI)]. This figure has been generated using the MATLAB script `perf_ltv_equaliz.m`.

**FIGURE 5.29**

BER performance comparison among MMSE and DFE equalizers for LTV channels [Uncoded BER, QPSK, $M = 128$ subcarriers with $M_{act} = 96$ consecutive data subcarriers and $M_{vir} = M - M_{act} = 32$ zero subcarriers, $D_b = 2$ for banded and window-aided equalizers, MBAE-SOE window [195], $L = 7$, Multipath channel with uniform power-delay profile, $\max \{f_{d_{p,n}}\} / \Delta f = 0.1$, perfect channel state information (CSI)]. This figure has been generated using the MATLAB script [perf_ltv_equaliz.m](#).

**FIGURE 5.30**

BER performance of turbo MMSE equalizers for LTV channels [Uncoded BER, QPSK, $M = 128$ subcarriers with $M_{act} = 107$ non-consecutive data subcarriers and $P = M - M_{act} = 21$ pilot subcarriers, whose $M_{vir} = 14$ are zero and $P - M_{vir} = 7$ are nonzero, $D_b = 2$, MBAE-SOE window [195], $L = 5$, Multipath channel with uniform power-delay profile, $\max \{f_{d_{p,n}}\} / \Delta f = 0.1$].

the time-evolution of a single delay path. Thus, the structure in the problem can be exploited to recast it in an equivalent estimation problem, with a reduced number of unknowns. For instance, by approximating the quasi-banded frequency-domain channel matrix \mathbf{H}_f in (5.189) as perfectly banded with only D_b super and lower diagonals different from zeros, the number of unknowns are $M(2D_b + 1)$, which is convenient when $D_b < L/2$. Actually, each diagonal in \mathbf{H}_f represents a discrete Doppler frequency shift induced by the LTV channel [97, 204], and assuming \mathbf{H}_f as perfectly banded corresponds to assume that each LTV channel tap can be expressed by a linear combination of $2D_b + 1$ discrete-time (Fourier) exponentials, as expressed by

$$h_{n,l} = \sum_{q=0}^{2D_b} c_{q,l} e^{j \frac{2\pi}{M}(q-D_b)n}, \quad n \in [0, M-1], \quad l \in [0, L]. \quad (5.192)$$

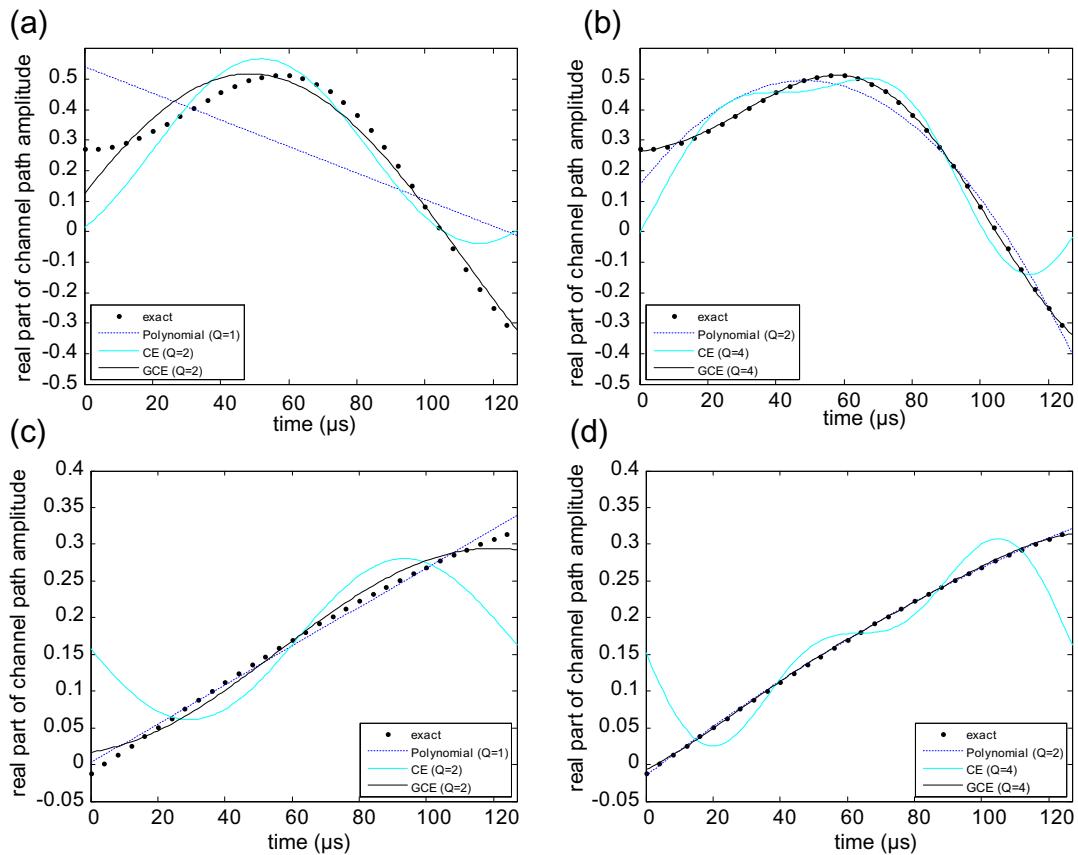
Thus, the M unknowns that characterize the time-evolution of each channel tap are turned in $2D_b + 1$ unknowns by the coefficients $c_{q,l}$, while the time evolution is captured by a known set of complex exponential functions, which are the same for all the $L + 1$ taps. Globally the problem can be casted as the estimation of $(L + 1)(2D_b + 1)$ unknowns, where typically $2D_b + 1 \leq 5$ is enough to accurately represent the time-evolution of each channel tap, or equivalently to model the ICI only on $2D_b$ adjacent subcarriers, by \mathbf{H}_f . More generally, the time evolution of the paths can be captured by other basis functions, by the so called basis expansion model (BEM), expressed by

$$h_{n,l} = \sum_{q=0}^Q c_{q,l} u_q[n], \quad n \in [0, M-1], \quad l \in [0, L], \quad (5.193)$$

where $Q + 1$ is the number of BEM functions, and $u_q[n]$ is the q th BEM function. Basically, the BEM functions $\{u_q[n]\}$, are the same for different OFDM blocks and different channel taps. Several choices have been investigated, such as power polynomials [33], complex exponential (CE) [239], Legendre polynomials [99], generalized complex exponentials (GCE) [139], discrete prolate spheroidal (DPS) functions [266], discrete Karhunen-Loève (DKL) [235]. For instance, Figure 5.31 shows the time evolution of a single channel path $h_{n,l}$, and its LS fitting obtained by either Fourier or Legendre polynomial functions.

On the other hand, the unknown BEM coefficients $\{c_{q,l}\}$ are time invariant within (at least) the considered OFDM block, but change from block to block (or from multiple blocks to multiple blocks), and therefore must be re-estimated, or updated, every OFDM block (or multiple blocks).

As already discussed for LTI channel estimation, based on the transmission of training pilots, several options are available, with pilots either in the frequency, or in the time-domain, or both. Preamble-based pilot design may be preferable in WLAN applications characterized by burst transmissions, while frequency-domain pilots interleaved with data within each OFDM block may be preferable for continuous transmission, such as in broadcasting applications (DVB-T/H, DAB). In the following, we describe a frequency-domain pilot-based LTV channel estimation technique, which exploits the structured BEM model for the channel time variation. However, the BEM structure can be similarly exploited also by preamble-based approaches in the time-domain, as detailed for instance in [231] for SC block transmissions. Specifically we make the following assumptions:

**FIGURE 5.31**

LTV channel tap gain for normalized Doppler spread (a,b) $\max \{f_{d,n}\} / \Delta f = 100\%$; (c,d) $\max \{f_{d,n}\} / \Delta f = 20\%$.

- The transmission format is CP-OFDM.
- At the transmitter, P out of M subcarrier are reserved for pilot symbols, while the remaining $D = M - P$ are used for data symbols. We assume that the P pilots are arranged in $L_P + 1$ clusters, with $P = (2Q_P + 1)(L_P + 1)$, where each pilot cluster is a vector of size $2Q_P + 1$ with p on the central position and Q_P zeros on each side. We also assume that $L_P \geq L$. We denote with **a** the frequency-domain transmitted vector, obtained by interlacing the pilot clusters with the data clusters. For convenience, we denote with **p** the pilot vector, which is zero also on the data positions, and with **d** the data vector, which is zero on the pilot positions, so that **a** = **p** + **d**.

We assume that the time variation of each discrete-time channel path can be represented by a BEM, as expressed by (5.193), and the BEM coefficients $\{c_{q,l}\}$ estimated every OFDM block.

- At the receiver, the channel estimation is performed using only a single OFDM block.

Using the BEM, it can be shown that [97, 230]

$$\mathbf{H}_f = \sum_{l=0}^L \sum_{q=0}^Q c_{q,l} \mathbf{U}_q \mathbf{Z}_l, \quad (5.194)$$

where $\mathbf{U}_q = \mathbf{F}_M \text{diag}(\mathbf{u}_q) \mathbf{F}_M^H$ is the circulant matrix that takes into account for the discrete Doppler spectrum associated with the q th BEM-function vector $\mathbf{u}_q = [u_q[0], \dots, u_q[M-1]]^T$, and $\mathbf{Z}_l = \text{diag}(\mathbf{z}_l)$ with $\mathbf{z}_l = [1, e^{j2\pi l/M}, \dots, e^{j2\pi l(M-1)/M}]^T$. For instance, choosing $Q = 2D_b$ even and complex exponential BEM functions $u_q[n] = e^{j2\pi(q-Q/2)n/M}$ as in (5.192), the matrix \mathbf{U}_q in (5.194) is a cyclic-shift matrix of $q - Q/2$ positions: in this case, the ICI on a certain subcarrier comes only from the Q closest subcarriers ($Q/2$ on each side), and a choice $Q_P \geq Q$ is sufficient to avoid ICI [230]. After some algebra, the matrix (5.194) may be rewritten as

$$\mathbf{H}_f = \mathbf{G}(\mathbf{c} \otimes \mathbf{I}_M), \quad (5.195)$$

where $\mathbf{G} = [\mathbf{U}_0 \mathbf{Z}_0, \dots, \mathbf{U}_Q \mathbf{Z}_0, \dots, \mathbf{U}_0 \mathbf{Z}_L, \dots, \mathbf{U}_Q \mathbf{Z}_L]$ and $\mathbf{c} = [c_{0,0}, \dots, c_{Q,0}, \dots, c_{0,L}, \dots, c_{Q,L}]^T$. Combining (5.195) with $\mathbf{y}^{(f)} = \mathbf{H}_f \mathbf{a} + \mathbf{w}^{(f)}$, we obtain

$$\mathbf{y}^{(f)} = \mathbf{G}(\mathbf{c} \otimes \mathbf{I}_M) \mathbf{a} + \mathbf{w}^{(f)} = \mathbf{G}(\mathbf{I}_{(Q+1)(L+1)} \otimes \mathbf{a}) \mathbf{c} + \mathbf{w}^{(f)}, \quad (5.196)$$

which clearly highlights the linear relation between the unknown coefficients in \mathbf{c} and the received frequency-domain vector $\mathbf{y}^{(f)}$. By selecting the $(Q_P + 1)(L_P + 1)$ subcarriers that correspond to the $L_P + 1$ pilot clusters, we obtain

$$\mathbf{y}_P^{(f)} = \mathbf{S}_P \mathbf{y}^{(f)} = \mathbf{B}_P \mathbf{c} + \mathbf{i} + \mathbf{S}_P \mathbf{w}^{(f)}, \quad (5.197)$$

where \mathbf{S}_P is the $(Q_P + 1)(L_P + 1) \times M$ selection matrix, $\mathbf{B}_P = \mathbf{S}_P \mathbf{G}(\mathbf{I}_{(Q+1)(L+1)} \otimes \mathbf{p})$, and $\mathbf{i} = \mathbf{S}_P \mathbf{G}(\mathbf{I}_{(Q+1)(L+1)} \otimes \mathbf{d}) \mathbf{c}$ is the residual ICI caused by the data \mathbf{d} . For a complex exponential BEM with $Q_P \geq Q$, it can be shown that the residual ICI is virtually absent, that is [230]

$$\mathbf{y}_P^{(f)} = \mathbf{B}_P \mathbf{c} + \mathbf{S}_P \mathbf{w}^{(f)}. \quad (5.198)$$

Therefore, an estimate $\hat{\mathbf{c}} = [\hat{c}_{0,0}, \dots, \hat{c}_{Q,0}, \dots, \hat{c}_{0,L}, \dots, \hat{c}_{Q,L}]^T$ of the channel coefficients may be obtained by LS (or linear MMSE) estimation by

$$\hat{\mathbf{c}} = \mathbf{G}_{\text{eq}} \mathbf{y}_P^{(f)}, \quad (5.199)$$

where \mathbf{G}_{eq} is equivalent to (5.190) [or to (5.191)] using \mathbf{B}_P for \mathbf{H}_f and $\mathbf{y}_P^{(f)}$ for $\mathbf{y}^{(f)}$ [117]. Then, by (5.193) an estimate of the LTV channel $\hat{h}_{n,l}$ is obtained as

$$\hat{h}_{n,l} = \sum_{q=0}^Q \hat{c}_{q,l} u_q[n], \quad (5.200)$$

and, by (5.194) and (5.195), an estimate of the frequency-domain channel matrix is obtained as

$$\hat{\mathbf{H}}_f = \sum_{l=0}^L \sum_{q=0}^Q \hat{c}_{q,l} \mathbf{U}_q \mathbf{Z}_l = \mathbf{G}(\hat{\mathbf{c}} \otimes \mathbf{I}_M). \quad (5.201)$$

When the interference \mathbf{i} in (5.197) is not negligible, iterative approaches that make use of (hard or soft) ICI cancellation may be employed [74, 230].

Figure 5.32a and c shows the MSE performance of the proposed approach, for different choices of the basis functions and different maximum Doppler spread, for several channel estimators such as the MMSE, LS, and best linear unbiased estimator [117]. Figure 5.32b and d show the corresponding BER performance for QPSK modulation and the banded version of the linear MMSE equalizer in (5.191) [230, 193].

The BEM channel estimation approach described by (5.192)–(5.201) is appropriate when the LTV channel is wide-sense stationary (WSS), that is, when the statistics of the LTV channel, such as the Doppler power spectral density, do not change with time. In some cases, non-WSS channels have to be estimated, such as when the mobile receiver changes its speed. Indeed, a speed change entails a change of the maximum Doppler frequency. In these cases, adaptive approaches are more suitable for channel (i.e., BEM coefficients) tracking purposes. Examples of these techniques have been presented in [18, 19, 44, 123, 168, 223].

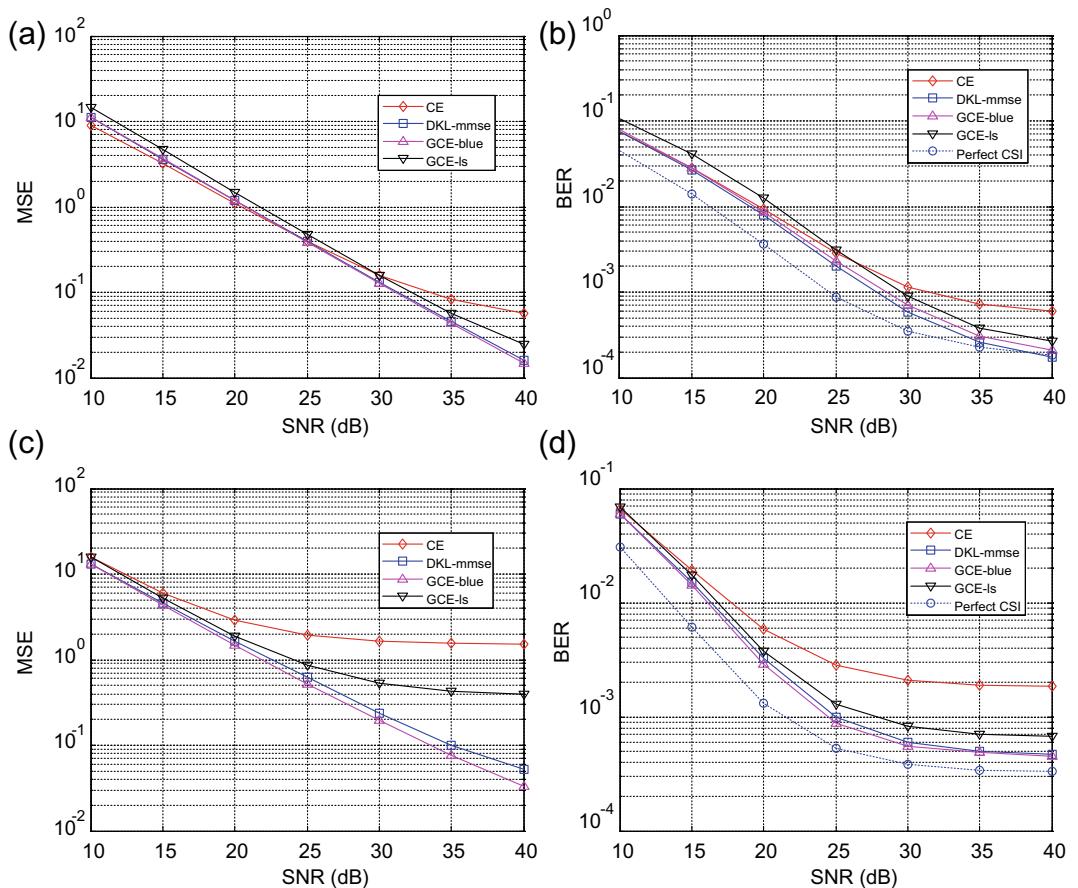
Readers interested in further hints on channel estimation and tracking are redirected to the E-Reference contribution in [241].

2.05.14 Design of OFDM systems: practical guidelines and commercial standards

This section aims at providing general guidelines for the design of the typical parameters of an OFDM system, i.e., how to choose the number M of subcarriers, the subcarrier separation Δf , the CP length T_g , the OFDM symbol period T_u . We also highlight the interplay among the different parameters, because $T_u = M T_s$, where T_s is the sampling period, $T_g = L_g T_s$, where L_g is the number of samples in the guard period (CP, ZP, or KP), and $\Delta f = 1/T_u$. We also define the bandwidth of the OFDM system as $W = M \Delta f = 1/T_s$, which is equal to the (gross) symbol rate f_s . Often, the bandwidth requirement is imposed, either by the physical medium constraints or by the symbol rate requirements, and consequently the sampling period is fixed by $T_s = 1/W$.

As explained in Section 2.05.3, historically the main motivation for OFDM has been the need to efficiently contrast frequency-selective (multipath) channels. Assuming that the maximum delay of the multipath channel is T_h , in order to avoid IBI between successive OFDM blocks, the guard time should be $T_g \geq T_h$. Since the guard time reduces the information rate, it is convenient to select its minimum value, i.e., $T_g = T_h$. In this case, the number of samples in the guard period $L_g = T_g/T_s$ is equal to the discrete-time channel order $L = T_h/T_s$. Therefore, we use L in order to denote both the channel order and the number of samples of the guard period.

Thus, for a given frequency bandwidth that is available or it is necessary to guarantee specific capacity requirements, the increase of the number M of subcarriers reduces the system sensitivity to the

**FIGURE 5.32**

Channel MSE and BER versus SNR. $Q = 4$. (a)–(b) $\max \{f_{d_{p,n}}\} / \Delta f = 0.2$, (c)–(d) $\max \{f_{d_{p,n}}\} / \Delta f = 1.0$.

frequency selectivity of multipath channels (see Figure 5.6). However, note that this intuitive statement loses part of its significance when a guard-time (by CP, ZP, or KP) is inserted between successive OFDM symbols. Indeed, as proved by (5.47) and (5.65) the CP (and similarly the ZP or the KP) is capable to turn the time-domain multipath channel into a set of parallel and orthogonal single-path frequency-domain channels, independently of the number of subcarriers. Consequently, the use of a high number of subcarriers is attractive in order to maximize the OFDM system efficiency, whereas the resistance to the channel selectivity is caused by the presence of the CP rather than the use of a high number of subcarriers. In this respect, the efficiency of the OFDM system that does not transmit information during the CP (ZP, or KP) is inversely proportional to the CP versus the OFDM symbol duration, as expressed by

$$\eta_{\text{ofdm}} = \frac{T_u}{T_u + T_g} = \frac{M}{M + L} \xrightarrow{M \gg L} 1. \quad (5.202)$$

If we call $B_c = 1/T_h = 1/T_g$ the channel coherence bandwidth (other definitions are possible for the coherence bandwidth [222]), (5.202) shows that $T_g \ll T_u$ in order to obtain an adequate efficiency: in this case, the subcarrier separation should be

$$\Delta f = 1/T_u \ll 1/T_g = B_c. \quad (5.203)$$

Equivalently, in order to maximize the overall efficiency, the system designer is tempted to use a number M of subcarriers much greater than the discrete-time channel length $L = T_g/T_s$, which is imposed by the environment (indoor/outdoor, urban/rural area, etc.) and by the system bandwidth $W = 1/T_s$.

However, there are also good reasons to not increase M over a certain limit. Indeed, higher values of M induces a longer OFDM symbol duration $T_u = 1/\Delta f$. Very long symbol durations may be not desirable for many reasons. The symbols transmitted on each subcarrier are detected every OFDM symbol period $T_u + T_g$: the consequent detection delay with respect to SC communications may become unacceptable for real-time communications if $T_u = MT_s$ is too long. Most importantly, a very long OFDM symbol duration T_u may be not desirable in mobile communications, where the channel is not only frequency-selective but also time-varying (i.e., doubly-selective). In this last case, the assumptions that lead us to the easy per-subcarrier equalization for time-invariant multipath channels are no longer true, as detailed in Section 2.05.10. In other words, the CP-enabled subcarriers lose their orthogonality in time-varying (mobile) channels, due to the Doppler spread that modifies the sinc(\cdot) shaped spectrum of the data transmitted on each subcarrier, introducing ICI (see Figure 5.27). Thus, if the designer prefers to avoid the use of complex time-varying equalization schemes, it is necessary to design the OFDM symbol duration $T_u \ll T_c$, where T_c is the so called coherence time [222], within which the channel can be safely considered as time-invariant. Since the channel coherence time T_c is inversely proportional to the maximum Doppler spread f_D induced by the time-varying channel [222], the time-variation (thus the ICI) can be safely ignored when the subcarrier separation is $\Delta f = 1/T_u \gg f_D$ (see Figure 5.27). Actually, the fact that closer subcarriers are more sensitive to Doppler effects is also quite intuitive: as a practical rule of thumb, the BER performance degradation induced by a time varying channel is negligible when the maximum Doppler spread of the channel is within $1 \div 2\%$ of the subcarrier separation, that is $f_D/\Delta f = f_D T_u = f_D N T_s \leq 0.01 \div 0.02$.

Furthermore, note that a frequency synchronization error induced at the receiver by an oscillator drift f_D has similar effects (see Section 2.05.11) and the same considerations hold true to make the OFDM system resistant to such inaccuracy. Another reason for avoiding long symbol durations, or equivalently a high number of subcarriers, is the FFT complexity, which scales as $O(M \log_2 M)$ per block: as a consequence, the complexity per data sample (logarithmically) increases with M .

Obviously, the OFDM designer would like to find a good compromise between the two opposite requirements on the OFDM block duration, in order to design an OFDM system that is both efficient, in time-dispersive (frequency-selective) channels, and resistant, in the time-varying (frequency-dispersive) channels induced by equipment mobility or oscillators drifts. Anyway, in order to make this possible, the OFDM symbol duration has to satisfy both the previous inequalities, that is

$$\frac{1}{B_c} < T_u < \frac{1}{f_D}. \quad (5.204)$$

This is obviously possible only when the channel is characterized by

$$f_D/B_c = f_D T_c < 1, \quad (5.205)$$

which is known as the *underspread-channel* condition [97].

Table 5.1 summarizes the main OFDM parameters employed in popular standards such as DVB-T, DVB-T2, IEEE 802.11a/g, IEEE 802.16e, LTE and ADSL [72, 73, 104, 105, 107, 237]. Note that DVB-T and DVB-T2 have a fixed sampling frequency f_s and hence a fixed bandwidth W , while the number of subcarriers M depends on the transmission mode. Therefore, when the number of subcarriers M increases, the useful block duration $T_u = M/f_s$ increases and the subcarrier separation $\Delta f = 1/T_u$ decreases. As a consequence, DVB-T/T2 modes with many subcarriers are very sensitive to time-varying channels and Doppler effects. On the other hand, IEEE.802.16e and LTE have a fixed subcarrier separation and hence their sensitivity to Doppler effects is the same for all the transmission modes: in this case, the useful block duration $T_u = 1/\Delta f$ is fixed and the normalized Doppler spread is $f_D/\Delta f < 2\%$. For IEEE.802.16e and LTE standards, the increase of the number of subcarriers M implies an increased sampling frequency $f_s = M/T_u$ and hence an increased bandwidth W . Note that IEEE 802.11a/g standards have a fixed sampling frequency and a fixed subcarrier separation, which implies a single transmission mode with $M = 64$. For IEEE 802.11a/g standards, the mobile speed is very low (e.g., $v = 5 \text{ km/h}$) and the subcarrier separation is quite great (i.e., $\Delta f = 312.5 \text{ kHz}$), and therefore the Doppler effects are practically absent. Thus, in the IEEE 802.11p version for car-to-car communications, it has been necessary to increase resistance to longer channel multipath (by a longer CP duration T_g) rather than to Doppler spread: actually, IEEE 802.11p uses also a closer subcarrier separation (longer block duration $T_u = 1/\Delta f$) in order to preserve a quite high efficiency ratio $T_u/(T_u + T_g)$.

Note that the ADSL standard, for wired communications, is quite sensitive to multipath channels: when the CP length $L = 32$ is insufficient to deal with the multipath length of the channels, channel shortening techniques may be employed.

2.05.15 OFDM time and frequency synchronization

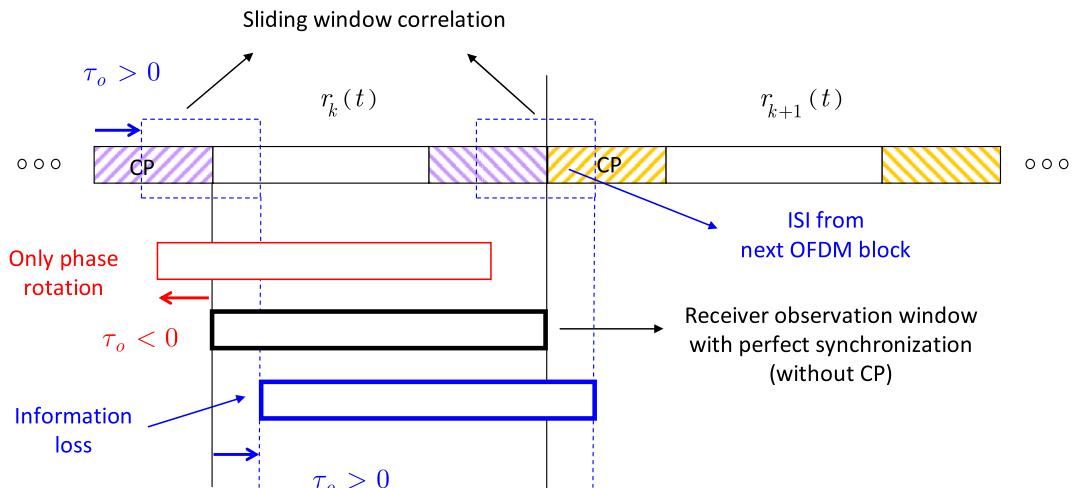
Synchronization is an important task for communication receivers of MC block transmissions. Time synchronization algorithms aim at identifying the correct starting point of an OFDM block, whereas frequency synchronization algorithms try to minimize the CFO (discussed in Section 2.05.12) between the transmitter oscillator and the receiver oscillator.

The lack of perfect synchronization, either in time or frequency, induces a dual effect on the received signal, with a phase rotation in frequency or time, respectively. These effects are summarized by (5.175) and (5.172), respectively, interpreting τ_0 in (5.175) as the time error, i.e., $\tau_0 = \tau_{\text{off}}/f_s$, where τ_{off} is the discrete-time offset between receiver and transmitter, and interpreting f_{d0} in (5.172) as the frequency synchronization error, i.e., $f_{d0} = \varepsilon f_s/M$. Note that a discrete-time offset $\tau_{\text{off}} > 0$ in (5.172) introduces not only the phase rotation in (5.175) on each subcarrier but also a partial loss of information on the received OFDM block and ISI from the next OFDM block, as clarified by Figure 5.33. The partial loss of time support for the transmitted data causes both an attenuation of the useful data and the introduction of ICI among subcarriers, which degrades the perceived SNR and the SER, as quantified in [226] for AWGN scenarios.

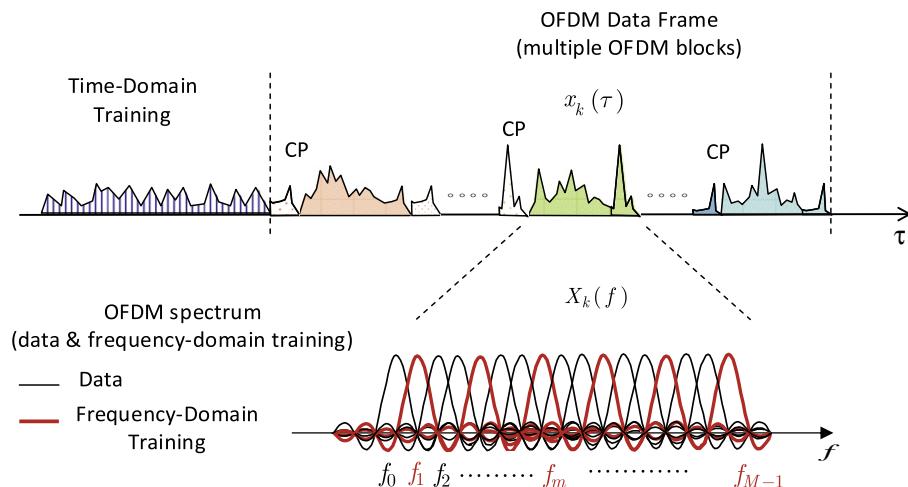
There is a large amount of possible methods that exploit the structure induced by (5.175) and (5.172) to estimate τ_{off} and ε , performing time and frequency synchronization in OFDM by their compensation. These methods may be based either on known training (pilot) signals, placed in the time- or frequency-domain, or on frequency-domain data symbols, or on both training and data (see Figure 5.34). However,

Table 5.1 Main Parameters for Some OFDM-Based Standards

Standard (Mode, Parameters)	FFT Size	Bandwidth	Gross Bit Rate (16-QAM)	Subcarrier Separation	Sampling Rate	Useful Block Duration	CP Size	CP Duration	Carrier Frequency	Speed	Normalized Doppler Spread
	M	W (MHz)	(Mb/s)	Δf (kHz)	f_s (MHz)	T_u (μs)	L	T_g (μs)	f_c (MHz)	v (km/h)	$f_D/\Delta f$
DVB-T (2k, CP 1/16)	2048	7.61	28.66	4.46	9.14	224	128	14	702	100	0.0146
DVB-T (8k, CP 1/16)	8192	7.61	28.64	1.12	9.14	896	512	56	702	100	0.0582
DVB-T (1k, CP 1/16)	1024	7.62	28.67	8.93	9.14	112	64	7	702	100	0.0073
DVB-T (16k, CP 1/16)	16,384	7.61	28.64	0.56	9.14	1792	1024	112	702	100	0.1165
IEEE 802.11a (CP 1/4)	64	16.56	52	312.5	20	3.2	16	0.8	5200	5	7.7E-05
IEEE 802.11 g (CP 1/4)	64	16.56	52	312.5	20	3.2	16	0.8	2412	5	3.6E-05
IEEE 802.11e (1.25 MHz, CP 1/8)	128	1.25	3.27	10.94	1.4	91.43	16	11.43	3500	60	0.0178
IEEE 802.11e (20 MHz, CP 1/8)	2048	20	65.33	10.94	22.4	91.43	256	11.43	3500	60	0.0178
LTE downlink (5 MHz, CP 1/4)	512	5	14.4	15	7.68	66.67	128	16.67	2140	100	0.0132
LTE downlink (20 MHz, CP 1/4)	2048	20	57.6	15	30.72	66.67	512	16.67	2140	100	0.0132
ADSL (Downstream)	512	1.10	4.12	4.31	2.21	231.88	32	14.49	0.55	0	0

**FIGURE 5.33**

OFDM with time offset at the receiver side.

**FIGURE 5.34**

Time-domain and frequency-domain training for OFDM synchronization.

a suitable synchronization method should take into account the type of transmission, i.e., burst or continuous. In (burst) packet-based OFDM transmissions, such as WLAN, the acquisition time should be kept to a minimum. In this case, it is convenient to transmit a specially designed training sequence, possibly placed in the time-domain, at the beginning of the whole frame, in order to speed up the

acquisition of time-frequency synchronization parameters. On the other hand, in continuous OFDM transmissions, such as broadcasting, it is meaningful to exploit repeated (periodic) transmission of training sequences, in order to track also slow frequency drifts of the oscillators. However, this approach produces a data-rate reduction. Therefore, for continuous OFDM transmissions, (blind) data-aided methods based on the CP presence or on statistical average over data symbols may be more convenient than training-based methods.

The remainder of this section briefly presents two joint time and frequency synchronization algorithms for OFDM. The first algorithm, proposed in [208], uses a training sequence at the beginning of the OFDM frame, and is well suited to burst transmissions. The second algorithm, proposed in [243], exploits the redundancy generated by the CP, and is well suited to continuous transmissions. Although there exist many other algorithms that outperform those described in this section, the algorithms of [208] and [243] are important because they establish the basic principles for OFDM synchronization and are often considered as a benchmark for performance comparison.

The time synchronization idea behind the Schmidl and Cox algorithm (SCA) is to detect the training sequence by using a time-domain correlation-maximization approach [208]. In AWGN, the receiver could perform a sliding-window correlation between the time-domain noisy received signal and the training sequence template, similarly to a sliding-window MF. However, in multipath channels, this correlation approach is not convenient, because the training sequence inside the received signal is modified by the LTI channel. Therefore, in multipath channels, the received signal should be correlated with a multipath-affected training sequence. Unfortunately, before time synchronization has been achieved, the multipath CIR is still unknown, and hence the multipath-affected template is not available. The SCA overcomes this problem by using a training sequence with two identical halves in the time-domain. Since both halves are affected in the same way by an LTI channel, half of the training sequence may be used as a (noisy) template in order to detect the other half of the training sequence. Let us denote with L_T the number of samples of each half sequence. Since a CP is inserted, the total length of the whole training sequence is $2L_T + L_g$. By denoting with r_n the generic received sample in the time-domain, the SCA constructs the sliding-window correlation function

$$P(\tau) = \sum_{n=\tau}^{\tau+L_T-1} r_n^* r_{n+L_T}, \quad (5.206)$$

where τ denotes a candidate time offset, and then estimates the time offset τ_{off} by selecting the value $\hat{\tau}_{\text{off}}$ that maximizes the metric $|P(\tau)|^2 / R(\tau)$, where $R(\tau)$ is a power normalization factor defined as

$$R(\tau) = \sum_{n=\tau}^{\tau+L_T-1} |r_{n+L_T}|^2. \quad (5.207)$$

After time synchronization has been performed, the SCA compensates for the CFO as follows. Let us denote with w_n the generic noise sample in the time-domain at the receiver, and $r'_n = r_n - w_n$, such that $r_n = r'_n + w_n$. It can be shown that, in case of perfect time synchronization (i.e., $\hat{\tau}_{\text{off}} = \tau_{\text{off}}$), it holds true that [161, 208]

$$r'_{n+L_T} = r'_n e^{j2\pi\varepsilon L_T/M}. \quad (5.208)$$

Therefore, at the correct time offset τ_{off} , and in the absence of noise, by (5.208) the sliding-window correlation function $P(\tau)$ in (5.206) would produce

$$P(\tau_{\text{off}}) = \sum_{n=\tau_{\text{off}}}^{\tau_{\text{off}}+L_T-1} (r'_n)^* r'_{n+L_T} = E e^{j2\pi\varepsilon L_T/M}, \quad (5.209)$$

where $E = \sum_{n=\tau_{\text{off}}}^{\tau_{\text{off}}+L_T-1} |r'_n|^2$ is the energy of half of the training sequence. As a consequence, the CFO can be estimated as

$$\hat{\varepsilon} = \frac{M}{2\pi L_T} \text{angle}(P(\hat{\tau}_{\text{off}})) = \frac{M}{2\pi L_T} \tan^{-1} \left(\frac{\text{Im}\{P(\hat{\tau}_{\text{off}})\}}{\text{Re}\{P(\hat{\tau}_{\text{off}})\}} \right). \quad (5.210)$$

The CFO estimator (5.210) is valid only if the CFO ε is within the CFO acquisition range $|\varepsilon| < \frac{M}{2L_T}$. Therefore, a longer training sequence reduces the frequency offset acquisition range. However, by (5.206), a longer training sequence also improves the time offset estimation accuracy, because the average (5.206) is more accurate, at the expense of an increased overhead (reduced data rate). In [161], the half-length of the training sequence is chosen as $L_T = M$, and therefore the CFO acquisition range is half of the subcarrier spacing. However, using $L_T = M$ permits to replace the two identical halves of a known training sequence with an unknown OFDM data symbol repeated twice [161], thereby reducing the overhead. In [208], a second training sequence is inserted in the time-domain, after the two identical halves of the first training sequence. The aim of this second sequence is to increase the CFO acquisition range.

For joint time-frequency synchronization and channel estimation purposes, the IEEE 802.11 WLAN standard [104] uses a mixed approach with both time-domain training and frequency-domain pilots. The time-domain training is inserted at the beginning of each OFDM packet: there are 10 identical short training symbols followed by 2 identical long training symbols with a CP. In addition, in each subsequent OFDM symbol, 4 pilot tones are inserted for CFO correction and tracking purposes.

The time synchronization idea behind the algorithm of van de Beek et al. is to detect the CP, again by using a time-domain correlation-maximization approach [243], as shown in Figure 5.33. Indeed, in the absence of both noise and multipath, the L_g samples of the CP are identical to the last L_g samples of the OFDM block: hence, similarly to the SCA, there are two identical sequences, separated in time by M samples, where M is the DFT size. The algorithm of van de Beek et al. evaluates the sliding-window correlation function

$$P_2(\tau) = \sum_{n=\tau}^{\tau+L_g-1} r_n r_{n+M}^*, \quad (5.211)$$

and the power term

$$R_2(\tau) = \frac{1}{2} \sum_{n=\tau}^{\tau+L_g-1} (|r_n|^2 + |r_{n+M}|^2), \quad (5.212)$$

and then estimates the time offset τ_{off} by selecting the value $\hat{\tau}_{\text{off}}$ such that

$$\hat{\tau}_{\text{off}} = \arg \max_{\tau} [M_{\text{CP}}(\tau)] = \arg \max_{\tau} [|P_2(\tau)| - \rho R_2(\tau)], \quad (5.213)$$

which is the ML estimates in AWGN scenarios, and where $\rho = \text{SNR}/(1 + \text{SNR})$.

Since the CP is included in every OFDM block, the functions in (5.211) and (5.212) may be averaged over multiple OFDM blocks, thereby improving the time offset estimation performance at the expense of increased complexity and increased delay.

Actually, as previously clarified, OFDM transmissions are conceived for non-AWGN channels, i.e., to easily handle the frequency selectivity of multipath channels, which request some further refinements to the CP-based time synchronization approach. Indeed, in the presence of multipath, the CP of the current OFDM block is affected by the samples of the previous OFDM block, while the last L_g samples of the current OFDM block are affected by other samples of the current OFDM block. Therefore, at the correct time offset τ_{off} , the two sequences (CP and last L_g samples of the current OFDM block) are no longer identical. This problem may be solved when the CIR order L is strictly lower than the CP length L_g , because, in the absence of noise, the last $L_g - L$ samples of the CP are still identical to the last $L_g - L$ samples of the OFDM block. As a consequence, in multipath channels, it may be convenient to replace the first summation index τ in (5.211) and (5.212) with $\tau + L$. In this case, the time synchronization performance would be reduced, because the noise averages in (5.211) and (5.212) are performed over a lower number of samples ($L_g - L$ instead of L_g).

Further improvements to the CP-based approach, are possible by exploiting also the pilot tones that are typically inserted in the frequency-domain for channel estimation. For instance, [133] proposes to estimate τ_{off} by

$$\hat{\tau}_{\text{off}} = \arg \max_{\tau} [\rho M_{\text{CP}}(\tau) + (1 - \rho) M_{\text{FP}}(\tau)], \quad (5.214)$$

where

$$M_{\text{FP}}(\tau) = (1 + \rho) \left| \sum_{n=\tau}^{\tau+L_g-1} r_n^* m_{n-\tau} \right| - \rho \left| \sum_{n=\tau}^{\tau+L_g-1} (r_n + r_{n+N})^* m_{n-\tau} \right|$$

represents the frequency pilot (FP) contribution to the ML estimates, and m_n is the time-domain sequence corresponding to the FP. Further improvements [165] can be obtained by a decision-aided approach, which refines the initial (coarse) estimate of $\hat{\tau}_{\text{off}}$ by minimizing a cost function, as expressed by

$$\hat{\tau}_{\text{off}}^{(\text{ref})} = \arg \min_{\tau} \left[\sum_{m=0}^{M-1} \left| R_m(\tau) / \hat{H}_m(\tau) - \tilde{X}_m(\tau) \right|^2 \right], \quad \hat{\tau}_{\text{off}} - \Delta_{\tau} \leq \tau \leq \hat{\tau}_{\text{off}} + \Delta_{\tau}, \quad (5.215)$$

where $\hat{H}_m(\tau)$ and $\tilde{X}_m(\tau)$ are the (frequency-domain) channel and data estimates on the m th carrier, respectively, which are obtained for a specific τ in the neighborhood Δ_{τ} of the coarse estimate $\hat{\tau}_{\text{off}}$.

After time synchronization has been performed, the algorithm in [243] compensates for the CFO similarly to the SCA algorithm in (5.210), as expressed by

$$\hat{\varepsilon} = -\frac{1}{2\pi} \text{angle}(P_2(\hat{\tau}_{\text{off}})) = -\frac{1}{2\pi} \tan^{-1} \left(\frac{\text{Im}\{P_2(\hat{\tau}_{\text{off}})\}}{\text{Re}\{P_2(\hat{\tau}_{\text{off}})\}} \right). \quad (5.216)$$

Another possibility, for CFO estimation and frequency synchronization, can exploit the redundancy in the frequency-domain represented by the insertion of null subcarriers at the edges of the OFDM spectrum, which are employed as guard frequency bands from adjacent RF channels. Indeed, while in

the presence of perfect synchronization the receiver should perceive only noise on those subcarriers, the presence of a CFO induces by ICI (see Figure 5.26) a leakage of the data on those guard frequency bands. Thus, it is possible to perform frequency synchronization by minimizing the energy leakage on those subcarriers, by simple gradient-based algorithms as proposed in [242] and further analyzed and discussed in [20] and [153]. Interestingly, this simple approach is robust to frequency-selective channels and can be applied also in multiuser OFDM communications, with minor modifications. In a multiuser set-up, (i.e., OFDMA) where different subcarriers are assigned to different users, the CFO estimation and synchronization for downlink transmissions can be generally borrowed from classical OFDM, due to the fact that each user perceives a single (although different) CFO (and time offset) with respect to the signal transmitted by the base-station (BS). Conversely, in the OFDMA uplink, the BS perceives the super-position of the signal transmitted by each user, each one characterized by a specific CFO and time-offset. Consequently, the estimation problem is more complex and noteworthy synchronization cannot be performed at the receiver side, but only by proper feedback to each transmitter of its own CFO and time offset, which have to be pre-compensated for. In order to simplify the estimation problem, the best strategy is to design different, and orthogonal, training pilots for each user, in such way, the BS can perform a separate estimation for each CFO (and time offset) [162]. The interested readers are redirected to [163] for a thorough survey on this subject. Readers interested in a thorough survey of signal processing algorithms for synchronization are redirected to the E-Reference contribution in [85].

2.05.16 Capacity for OFDM transmissions: bit and power-loading

Multicarrier systems are particularly attractive also due to the straightforward possibility to boost the system capacity when a frequency-selective channel is known at the transmitter, as theoretically motivated by the Shannon water-filling principle [61, 213]. Indeed, OFDM transforms a single-input single-output (SISO) time-invariant multipath channel in a set of M orthogonal frequency-flat channels, as summarized by (5.65), which is rewritten in the following for the reader convenience

$$\mathbf{y}_k^{(f)} = \text{diag}(\mathbf{h}_f)\mathbf{d}_k + \mathbf{w}_k^{(f)}. \quad (5.217)$$

Analogously to the classical expression of the capacity of a SISO flat-fading AWGN channel, the capacity of the diagonal MIMO system in (5.217) with an AWGN term $\mathbf{w}_k^{(f)}$ with covariance matrix $E\{\mathbf{w}_k^{(f)}\mathbf{w}_k^{(f)H}\} = \sigma_w^2 \mathbf{I}_M$, is achievable by a Gaussian distributed data vector \mathbf{d}_k and is expressed by [61]

$$C_{\text{OFDM}} = \log_2 \left| \mathbf{I}_M + \frac{1}{\sigma_w^2} \mathbf{D}_H \Sigma_d \mathbf{D}_H^H \right| \quad [\text{bits/channel use}], \quad (5.218)$$

where $D_H = \text{diag}(\mathbf{h}_f)$, $\Sigma_d = E\{\mathbf{d}_k \mathbf{d}_k^H\}$ is the data vector covariance, and $|\mathbf{A}|$ stands for the determinant of the matrix \mathbf{A} . Analogously, the capacity of any other system equipped with a CP, which transmits a vector \mathbf{x}_k of (possibly precoded) data is expressed by

$$C_{\text{CP}} = \log_2 \left| \mathbf{I}_M + \frac{1}{\sigma_w^2} \mathbf{H}_{\text{cp}} \Sigma_x \mathbf{H}_{\text{cp}}^H \right|,$$

where $\Sigma_x = E\{\mathbf{x}_k \mathbf{x}_k^H\}$. Bearing in mind the Hadamard inequality $|\mathbf{A}| \leq \prod_{i=1}^M |\mathbf{A}|_{i,i}$ (with equality if and only if $\mathbf{A} = \text{diag}(\mathbf{a})$), and reminding that $\mathbf{H}_{\text{cp}} = \mathbf{F}_M^H \mathbf{D}_H \mathbf{F}_M$, and $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$, it is straightforward to prove that

$$\left| \mathbf{I}_M + \frac{1}{\sigma_w^2} \mathbf{H}_{\text{cp}} \Sigma_x \mathbf{H}_{\text{cp}} \right| = \left| \mathbf{I}_M + \frac{1}{\sigma_w^2} \mathbf{D}_H \mathbf{F}_M \Sigma_x \mathbf{F}_M^H \mathbf{D}_H \right|.$$

Consequently, in order to make the overall argument a diagonal matrix and maximize C_{CP} , it is necessary that $\Sigma_x = \mathbf{F}_M^H \mathbf{D}_p \mathbf{F}_M$, with $\mathbf{D}_p = \text{diag}([P_0, \dots, P_{M-1}])$. This is actually obtained when the transmitted vector is expressed by $\mathbf{x}_k = \mathbf{F}_M^H \mathbf{d}_k = \mathbf{F}_M^H \mathbf{D}_p^{1/2} \tilde{\mathbf{d}}_k$, i.e., when the system employs CP-OFDM, and transmits uncorrelated data $\mathbf{d}_k = \mathbf{D}_p^{1/2} \tilde{\mathbf{d}}_k$, where $\Sigma_d = E\{\mathbf{d}_k \mathbf{d}_k^H\} = \mathbf{D}_p$ contains on the diagonal the power load on each subcarrier, that is $[\mathbf{D}_p]_{m,m} = E\{|d_{k,m}|^2\} = P_m$, and where $P_T = \sum_{m=0}^{M-1} P_m = \text{trace}(\mathbf{D}_p)$ is the power budget at the transmitter side. For further details on capacity-optimal block transmissions, the interested readers are redirected to [200].

Thus, it has been established that OFDM is the optimum choice for CP-based transmissions, and that the transmission of uncorrelated data lets to express (5.218) as

$$C_{\text{OFDM}} = \log_2 \left| \mathbf{I}_M + \frac{1}{\sigma_w^2} \mathbf{D}_{|H|}^2 \mathbf{D}_p \right| = \log_2 \prod_{m=1}^M \left(1 + |H_m|^2 P_m / \sigma_w^2 \right) = \sum_{i=1}^M \log_2 (1 + \text{SNR}_m), \quad (5.219)$$

i.e., as intuitive, the capacity of an OFDM system is the sum of the capacity associated to each (also statistically) orthogonal channel that experiences an $\text{SNR}_m = |H_m|^2 P_m / \sigma_w^2$.

The power loading problem consists on the optimal distribution of the power budget $P_T = \sum_{m=0}^{M-1} P_m$ among the different subcarriers such that the overall capacity is maximized. Mathematically, this is expressed by

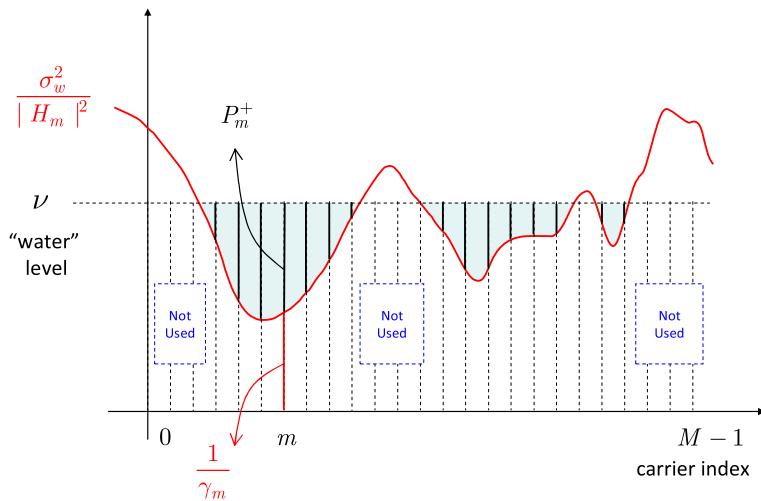
$$\arg \max_{P_m} \left[\sum_{m=0}^{M-1} \log_2 (1 + P_m \gamma_m) \right], \quad \text{s.t.} \quad \sum_{m=0}^{M-1} P_m \leq P_T, \\ , \quad \text{s.t.} \quad P_m \geq 0, \quad \forall m, \quad (5.220)$$

where $\gamma_m = |H_m|^2 / \sigma_w^2$ represents the SNR at the receiver side for a unitary transmitted power on each subcarrier. Thus the OFDM power loading is a convex optimization problem with convex constraints that admits a unique solution (computable by Lagrange multipliers), which, according to the Kuhn-Tucker conditions is expressed by [61, 213]

$$P_m^+ = (\nu - 1/\gamma_m)^+ = \begin{cases} \nu - 1/\gamma_m, & \nu > 1/\gamma_m, \\ 0, & \nu < 1/\gamma_m, \end{cases} \quad (5.221)$$

where $\lambda = -1/\nu$ is the Lagrange multiplier and, consequently, is determined by the power constraint such that $\sum_{m=0}^{M-1} (\nu - 1/\gamma_m)^+ \leq P_T$.

The solution (5.221) for maximum capacity power-loading is depicted in Figure 5.35, which clarifies why such a distribution is called the water-filling (WF) (or water-pouring) principle. Indeed, the total available power (water) distributes itself on each subcarrier as it would distribute the water poured on

**FIGURE 5.35**

OFDM subcarrier power loading according to the water-filling principle.

a vessel whose floor is represented by the inverse of the “channel” SNR γ_m . As expected, most of the power is given to the carrier with the best (e.g., highest) channel $|H_m|$. However, since the increase in capacity $C_m = \log_2 (1 + P_m \gamma_m)$ saturates for increasing power, after a certain amount of power has been “poured” to the first subcarrier, it is convenient to distribute the power also to other subcarriers, and so forth.

In some standards, such as ADSL [107], there is a limit to the maximum power \bar{P}_m that can be assigned to a single subcarrier, which means that the transmitted signal should respect a power spectral emission mask $\{\bar{P}_m\}_{m=0,\dots,M-1}$, which is mainly enforced to control the interference to coexisting services and systems. Consequently, the optimization problem in (5.220) has a more stringent constraint $0 \leq P_m \leq \bar{P}_m$, whose solution can be analogously derived and leads to the so-called “cap-limited” WF, which means that the vessel where to pour the power resembles a “pipe” which has not only the floor $1/\gamma_m$ shown in Figure 5.35 but also the cap $1/\gamma_m + \bar{P}_m$. Consequently, the poured water (power) cannot overpass the cap and the part of water of the classical WF solution that would be over the cap redistributes inside the “pipe”: the water level increases and more power (up to the cap) is assigned to other subcarriers (possibly also not used by the classical WF). For more details, see [176] and references therein.

A dual problem to the capacity (rate) maximization is the minimization of the total transmitted power subject to a minimum capacity constraint that, due to

$$P_m = \left(2^{C_m} - 1\right) / \gamma_m \quad (5.222)$$

is expressed by

$$\arg \min_{C_m} \left[\sum_{m=0}^{M-1} \left(2^{C_m} - 1 \right) / \gamma_m \right], \text{ s.t. } \sum_{m=0}^{M-1} C_m \geq C_{\min}, \\ \text{ s.t. } C_m \geq 0, \quad \forall m. \quad (5.223)$$

Also the solution of (5.223) can be derived by the Lagrange approach, which applying the Kuhn-Tucker conditions similarly leads to

$$C_m^+ = (\chi - \log_2 (1/\gamma_m))^+, \quad (5.224)$$

where $\lambda = -1/\chi$ is the Lagrange multiplier determined by the capacity constraint such that $\sum_{m=0}^{M-1} (\chi - \log_2 (1/\gamma_m))^+ = C_{\min}$. By substituting (5.224) in (5.222), it is evident that the optimal distribution among the subcarriers of the requested capacity C_{\min} , corresponds to a WF principle on the associated power, as expressed by

$$P_m^+ = \left(2^{C_m^+} - 1 \right) / \gamma_m = (2^\chi - 1/\gamma_m)^+. \quad (5.225)$$

Actually (5.225) is exactly the solution we would obtain by solving a capacity maximization problem with $P_T = \sum_{m=0}^{M-1} P_m^+$, and where 2^χ is the water level.

Note that the optimal solutions in (5.221) and (5.224) have been obtained for \mathbf{d}_k belonging to Gaussian distributed alphabets, and represent information theoretic bounds obtainable by ideal infinite-length coding and arbitrary small error probability at the receiver side [183]. However, any digital communication system transmits data \mathbf{d}_k that belong to finite alphabets (e.g., M-PSK, M-QAM, etc.) and suffers a certain SER. This fact can be modeled as an SNR penalty Γ_m with respect to an ideal system, which means that each subcarrier is capable to transmit with a rate per channel use

$$R_m = \log_2 (1 + P_m \gamma_m / \Gamma_m) < C_m. \quad (5.226)$$

The rationale behind (5.226) is to associate the rate penalty $R_m < C_m$ to a BER penalty suffered by the practical system with respect to an ideal one. Indeed, if we focus on an M-QAM constellation on each subcarrier, it is known that the SER can be approximated as [183]

$$\text{SER} \simeq 4Q \left(\sqrt{\frac{3\text{SNR}_m}{M_m - 1}} \right) = 4Q \left(\sqrt{\frac{3P_m \gamma_m}{2^{b_m} - 1}} \right), \quad (5.227)$$

which inverted with respect to b_m gives

$$b_m \simeq \log_2 \left(1 + \frac{P_m \gamma_m}{\frac{1}{3} Q^{-1}(\text{SER}/4)} \right). \quad (5.228)$$

Actually, (5.228) represents the approximated (non-integer) number of bits it is possible to load on the m th subcarrier for a target SER and, interestingly, it can be interpreted as (5.226) with $\Gamma_m = f_{\text{M-QAM}}(\text{SER}) = \frac{1}{3} Q^{-1}(\text{SER}/4)$. More generally, (5.226) can be used with $\Gamma_m = f(\text{SER})/G_c$, where $f(\text{SER})$ depends on the specific modulation and G_c represents the SNR gain granted by the

channel coding that is employed. Once Γ_m is known for a specific modulation/coding scheme, a practical system such as ADSL [107] may be interested to the best way to allocate (load) the power P_m and the bits b_m (e.g., determine the constellation size $M_m = 2^{b_m}$ to be used) on the m th subcarrier, in order to maximize the sum rate $\sum_{m=0}^{M-1} b_m$ for a given power budget $P_T = \sum_{m=0}^{M-1} P_m$, and target (uncoded) SER. Thus, in view of (5.226) and (5.228), the problem can be formalized by (5.220), substituting γ_m with $\gamma'_m = \gamma_m / \Gamma_m$, and the solutions are

$$P_m^+ = (\nu - \Gamma_m / \gamma_m)^+ = \begin{cases} \nu - \Gamma_m / \gamma_m, & \nu > \Gamma_m / \gamma_m, \\ 0, & \nu < \Gamma_m / \gamma_m, \end{cases} \quad (5.229)$$

$$b_m = \left\lfloor \log_2 \left(1 + \frac{\gamma_m}{\Gamma_m} P_m^+ \right) \right\rfloor = \left\lfloor \log_2 \left(\frac{\gamma_m}{\Gamma_M} \nu \right) \right\rfloor, \quad (5.230)$$

where $\lfloor \cdot \rfloor$ is the rounding operator to obtain an integer number of bits, and ν is determined by $\sum_{m=0}^{M-1} (\nu - \Gamma_m / \gamma_m)^+ = P_T$. Similarly (5.224) and (5.225) can be adapted to solve the dual problem to distribute a given number of bits $B_T = \sum_{m=0}^{M-1} b_m$ on the different subcarriers in order to employ the minimum amount of power $\sum_{m=0}^{M-1} P_m$, for a target SER.

The highest complexity of all the water-filling approaches is represented by the identification of the water level (i.e., ν in Figure 5.35), or equivalently of the Lagrange multiplier. Different strategies have been proposed in the literature: see [130], and [176] and the references therein.

In the rate-maximization case, the water level is determined by the constraint equation $\sum_{m=0}^{M-1} (\nu - 1/\gamma_m)^+ = P_T$, (using γ_m / Γ_m to include the system margin [57]). Consequently, indicating with $\mathcal{I}_A = \{m \in [0, M-1] : (\nu - 1/\gamma_m) > 0\}$ the set of subcarrier indexes that are effectively loaded with a power $P_m > 0$, it is straightforward to derive that the water level is expressed by

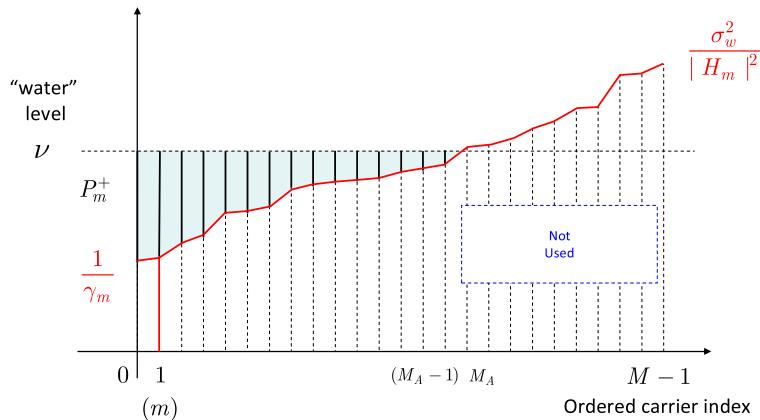
$$\nu(M_A) = \frac{P_T}{M_A} + \frac{1}{M_A} \sum_{m=0}^{M_A-1} 1/\gamma_m, \quad (5.231)$$

where $M_A = |\mathcal{I}_A|$ is the number of active subcarriers. Thus, the problem consists in the identification of those subcarriers that participate to the water-filling. It is obvious that the best subcarrier, i.e., the subcarrier with the highest channel SNR γ_m is included in \mathcal{I}_A , and probably also the second-best subcarrier, and so forth, while maybe the power is not enough to include the worst subcarrier with the lowest γ_m . This observation has inspired a bunch of similar algorithms, which are based on sorting the subcarrier indexes in increasing (or decreasing) γ_m , as shown in Figure 5.36, which considers the same (ordered) WF problem of Figure 5.35.

Figure 5.36 highlights that the water level ν has to be such that $1/\gamma_{M_A-1} \leq \nu < 1/\gamma_{M_A}$: this observation together with (5.231) suggests to identify the (sorted) subcarriers that participate to the water pouring by searching the only integer in $[0, M-1]$ that satisfies both

$$\nu(M_A) \geq 1/\gamma_{M_A-1} \quad \text{and} \quad \nu(M_A + 1) < 1/\gamma_{M_A}. \quad (5.232)$$

Once M_A [and $\nu(M_A)$] is known, the bit and power loading that maximize the sum-rate can be performed by (5.229) and (5.230). Similar consideration can be drawn for the bit and power loading that minimizes the power consumption for a target sum-rate.

**FIGURE 5.36**

Water filling with subcarrier ordered by their channel SNR γ_m .

It should be observed that the solution in (5.230) is suboptimal because the bit-loading problem is actually an integer optimization problem. Conversely, it has been solved by a relaxation approach [35] to the real optimization problem in (5.220) by ignoring the further constraint $\log_2(1 + P_m \gamma_m) \in \mathbb{Z}_0^+$, and rounding the real solution to the closest integer. The true integer optimization problem can be solved by greedy algorithms [11], which are iterative algorithms that (i) define a cost function associated to a unitary increment/decrement of the possible (integer) solution to the bit-loading problem, (ii) at every step moves towards the solution (among the ones at a unit step distance) characterized by the lowest value of the cost function, (iii) they never trace back to an already explored solution. By observing (5.226), and denoting with $\Delta P_m^{(+)}(R_m)$ the power cost to add a bit to R_m , it is possible to write

$$R_m + 1 = \log_2 \left(1 + (P_m + \Delta P_m^{(+)}) \gamma_m / \Gamma_m \right), \quad (5.233)$$

which solved with respect to $\Delta P_m^{(+)}$ together with (5.226) leads to

$$\Delta P_m^{(+)} = \frac{2^{R_m+1} - 1}{\gamma_m / \Gamma_m} - P_i = \frac{2^{R_m}}{\gamma_m / \Gamma_m}. \quad (5.234)$$

Analogously, the power saving $\Delta P_m^{(-)}(R_m)$ granted by reducing R_m to $R_m - 1$ is

$$\Delta P_m^{(-)} = \frac{2^{R_m-1} - 1}{\gamma_m / \Gamma_m} - P_i = \frac{2^{R_m-1}}{\gamma_m / \Gamma_m}. \quad (5.235)$$

Note that (5.234) and (5.235) are monotone convex cost and revenue functions, respectively. Consequently, when integer and convex optimization problems [such as (5.220) with the additional constraint $\log_2(1 + P_m \gamma_m) \in \mathbb{Z}_0^+$] are solved by means of iterative greedy algorithms based on the functions

(5.234) and (5.235), the convergence to a unique and optimal solution is guaranteed. Thus, a simple (bit-filling) algorithm that exploits (5.234) is that one that iteratively assigns an incremental bit to the subcarrier that requests the lower incremental cost in term of power, until the power budget is finished. This algorithm is summarized by the following pseudo code.

1. Set $k = 0$, $P_s = 0$, $b_m = 0$, and $P_m = 0$, $m = 0, \dots, M - 1$
2. compute $\Delta P_m^{(+)} = 2^{b_m} / (\gamma_m / \Gamma_m)$, $m = 0, \dots, M - 1$
3. set $m^* = \arg \min_m [\Delta P_m^{(+)})];$
4. $P_s = P_s + \Delta P_{m^*}^{(+)}$;
5. if $P_s \leq P_T$
6. $b_{m^*} = b_{m^*} + 1$;
7. $P_{m^*} = P_{m^*} + \Delta P_{m^*}^{(+)}$;
8. go to Step 2
9. end

The previous algorithm, as intuitive, starts assigning the first bit to the subcarrier m with the best channel (highest $\gamma_m = |H_m|^2 / \sigma_w^2$ assuming the SNR margin $\Gamma_m = \Gamma$ is the same for all the subcarriers). However, due to the fact that the rate (capacity) does not increase linearly with the power, it could cost less power to assign the second bit to another subcarrier (i.e., the second-best subcarrier) and so forth. Similar greedy algorithms, which exploit (5.235), are called bit-removal algorithms, and the interested readers are redirected to [176] and references therein, for further details.

Instead of rate maximization or transmitted power minimization approaches, other approaches may be considered, such as the SNR maximization approach of [78]. Readers interested in more advanced algorithms for ADSL power- and bit-loading are redirected to the E-Reference contribution in [137].

Finally, it should be observed that the algorithms discussed so far address the bit and power loading optimization for single-user (SU) scenarios. The multiuser (MU) case, which is more complicated, has been investigated, for instance, in [258] and [124]. An effective approach to solve the MU bit and power loading problem, is to split it in multiple SU optimization problems, which are run sequentially and iteratively, each one looking for the optimal solution for a specific user and assuming the power allocated to the other users as interference that adds to the Gaussian noise. It has been proven for instance that the MU water-filling problem can be accurately solved this way, with a guaranteed and fast convergence to the unique optimal solution. About this point, the interested readers may find detailed information in [47, 109, 205, 206, 214, 261, 262]. Another survey to inspire optimal allocation strategies is the E-Reference contribution of [151].

2.05.17 Further research topics and directions

In the following we shortly comment a non-exhaustive list of research topics that we did not consider in the previous sections: these topics have attracted a significant research effort in the recent years and we believe they may continue to be appealing in the near-future.

2.05.17.1 Power and bit loading with CSI information feedback

The availability of CSI information at the transmitter side can rely either on the use of symmetric time-domain duplexing between two transceivers, or on channel feedback from the receiver to the transmitter. In this second view, a critical factor is the coherence time of the channel which has to be significantly greater than the round-trip propagation/processing delay in order for the fed-back CSI to be not outdated. In this framework, there is substantial trade-off between the amount of the CSI information that is fed-back from the receiver (capacity loss) and the increase in rate this CSI is capable to guarantee (capacity gain) [54, 150].

For instance, in MIMO-OFDM, a possible way to reduce the amount of feed-back information is the use of codebooks which are predefined, and the receiver only select the best one to be used for the actual channel and accordingly informs the transmitter [55, 121]. Limited feedback schemes for OFDMA have been investigated in [3, 50], and [131].

2.05.17.2 Time-varying channel estimation and equalization for MIMO, and underwater communications

The estimation, equalization and linear precoding for OFDM and MC time-varying channels have been quite investigated topics in the recent past [41, 56, 74, 75, 90, 110, 152, 181, 193, 204, 230, 265, 304]. Although some results have been extended to MIMO scenarios [125, 192, 224, 227], this is an area where improvements are still possible, for instance, to design efficient sounding techniques coupled with low complexity implementations of channel estimation, tracking, and equalization [1, 49, 83].

Moreover, OFDM and MC communications, have been recently considered as a viable technology for underwater communications and sensor networks [140, 182, 291], where the ultrasound waves are typically the preferred choice for medium-range distances, in shallow-water environments [147]. Notably, wireless underwater communications present harder challenges with respect to air communications: the propagation channel is characterized by a very rich multi-path (especially in shallow-water), the Doppler effect is exacerbated by the reduced speed of ultrasound with respect to electromagnetic waves, and further (time-varying) distortions are also introduced by the movements of the propagation medium, which is not observed in electromagnetic communications. Thus, base-band signal processing algorithms developed for classical air-based OFDM and MC systems should be adapted, or completely redesigned, to take into account for more sophisticated system models. For instance, as already highlighted is Section 2.05.11, the reduced speed of ultrasound propagation does not allow to ignore the time-scale distortion in (5.171) which induces time-varying delays $\tau_l(t)$ in the channel impulse response and non uniform Doppler shifts [142, 143]. Actually, the time-scale change could be different for each single path, depending on its direction of arrival. Although some possible solution to deal with [115] and compensate (e.g., equalize) such different channels have been recently proposed in [100, 240, 255, 260] this is still a quite active research field including its MIMO implementation [45, 180, 126, 141].

2.05.17.3 OFDM-based cognitive radio applications

The cognitive radio paradigm [96, 160] is one of the most investigated topics in the last decade, due to the potentiality for secondary opportunistic networks to exploit the unused resources (in time, frequency,

space, codes, etc.) of primary legacy networks. Due to the fact that OFDM is the underlying modulation of several primary users' networks (i.e., DVB-T/H, WiFi, WiMAX, UMTS-LTE. etc.), the secondary users networks can exploit the knowledge of the OFDM signal structure (e.g., multitones, CP, almost Gaussian distribution) to design efficient statistical inference (spectrum sensing) on the presence or absence of a primary users to potentially exploit vacant resources [9, 10, 13, 138]. OFDMA, due to its flexibility to allocate frequencies to a specific user is one of the natural candidates to deploy a secondary network that accommodate to its users those frequencies that are not in use in the primary network.

If the secondary network is equipped with multiple-antennas further degrees of freedom are available to possibly cancel or minimize the interference towards the primary users. This approach falls within the more general topic of interference alignment [39, 40, 70, 155], which is recently revitalizing research efforts on MIMO-based resource management for multi-user and multi-cell communications. Anyway, in order to exploit the potentiality of cognitive-radio communications, it is of fundamental importance that the network has knowledge both of the spatial distribution of the electromagnetic fields, as well as of the channel gains each couple of node in the primary and secondary networks is linked with. Due to the sparsity of the users in the coverage area, channel sounding and estimation techniques can exploit quite recent results coming from compressive and sparse sampling theories [31, 43, 159], as for instance in [7, 25, 119, 267]. Noteworthy, these approaches can also be exploited for channel estimation of OFDM communications when the channel is sparse [232]. Readers interested in thorough survey for spectrum sensing in cognitive radio is redirected to the E-Reference contribution of [86].

2.05.17.4 Filterbank multicarriers

Another variant of multicarrier communications that recently attracted a significant research effort is filter-bank multicarrier (FBMC) [4, 22, 36, 51, 76, 91, 128, 156, 220]. The basic idea is that the data on each subcarrier can be (differently) filtered, in such a way to exhibits a spectrum different from the sinc-shaped one of classical-OFDM. Moreover, also the time support of each symbol may extend to more than a single multicarrier block (possibly introducing controlled ISI/IBI) and the use of CP is no longer fundamental and can be avoided. The higher degrees of freedom available can be exploited by properly designing the transmitter (and receiver) filterbanks in order to fulfill specific requests [34, 157, 238]. For instance, it is possible to reduce the spectral leakage of each subcarrier spectrum on adjacent subcarriers making FBMC less sensitive to CFO problems [81, 149] and also an interesting option for OFDMA-like implementation of secondary networks in cognitive-radio-systems [77]. The price to pay for the flexibility of FBMC is an increased implementation complexity, especially, at the receiver side, where channel estimation and equalization are not trivial [101, 106], and for multiuser and MIMO implementations, where still some research effort is required to make FBMC an attractive and practical modulation for future wireless standards.

2.05.17.5 OFDM-based relay communications

Relay-based communications is another topic that is attracting the interests of the wireless research community because of the potential increase in capacity, reliability, and coverage area of communication networks [134, 170]. In addition to classical relaying strategies, the use of relays as virtual antennas

lead to the so-called distributed ST-coding approaches [135], and naturally ask for an OFDM-based implementation in order to turn frequency-selective fading channels into frequency-flat-fading channels.

The multi-hop relay system of [248] aim at maximizing the system capacity using joint subcarrier matching and power allocation with separate power constraints. Different power allocation techniques are proposed: the optimal joint subcarrier matching and power allocation, and a reduced-complexity suboptimal scheme. Simulation results in [248] have shown that the capacity of the optimal scheme approaches the upper bound of the system capacity, and that the capacity of the suboptimal scheme is also quite close. For comparison purposes, the low-complexity one-to-one subcarrier matching technique is also included in [248]. Resource allocation algorithms for cooperative OFDM systems have been proposed also in [62, 154, 175].

The power allocation algorithms in [92] target the two-hop MIMO-OFDM case. Both frequency and space domains are considered, in order to maximize the instantaneous rate when channel state information at the transmitter is available. Two approaches are investigated: separate power-allocation optimization with individual per-node transmit-power constraints, and joint power-allocation optimization with joint transmit power constraint. The results of [92] also show that alternately repeating the separate optimization converges, and improves the achievable rate of the considered link.

In [68], a cooperative OFDM system with precoding is proposed, with the aim of collecting the available multipath diversity. In this case, the diversity order is given by the summation of the channel length in the direct link and the minimum of channel lengths in each relaying link. Two relaying strategies are compared: although the per-subcarrier basis strategy performs better, also the all-subcarrier basis strategy gives a performance improvement with respect to OFDM without cooperation. Additionally, relay selection methods are included in [68]. The work in [221] proposes a diversity-achieving protocol that makes use of limited feedback. Analytical studies about the performance of OFDM with amplify-and-forward relays have been carried out, for instance, in [228].

The ST cooperative design of [217] demonstrates the overall performance advantage of a cooperative OFDM system over an OFDM system without cooperation, not only under ideal conditions, but also in the presence of synchronization and channel estimation errors. Channel estimation is investigated also in OFDM systems with two-way relay networks, where the relays allow for bidirectional communications [82]. A two-phase channel-estimation training protocol is proposed, in order to maintain the compatibility with the two-phase data transmission scheme. First, a block-based training is studied, for the estimation of the cascaded source-relay-source channels. Second, a pilot-tone-based training is considered in [82], for the direct estimation of the individual channels between sources and relay. CFO mitigation is critical also in OFDM-based cooperative transmissions, especially when the number of cooperative transmitters or relays is large. In [142], the CP is exploited to mitigate the CFO. Depending on the CP length, the CFO error can be reduced or even eliminated. In [190], focusing on the relay-destination links, two pilot-aided algorithms for the joint estimation of the CFOs and the CIRs are proposed. The first algorithm uses a BEM to turn the joint estimation problem into the estimation of an LTV channel, from which CFOs and CIRs are extracted using ESPRIT. The second algorithm exploits a specific pilot structure and estimates the CFOs by directly applying ESPRIT on the observed signal. Channel estimation in the presence of CFO has been investigated also in [264]. Readers interested in signal processing algorithms for (distributed) cooperative systems are redirected to the E-Reference contribution in [21].

Relevant Theory: Signal Processing Theory

See [Vol. 1, Chapter 2](#) Continuous-Time Signals and Systems

See [Vol. 1, Chapter 3](#) Discrete-Time Signals and Systems

See [Vol. 1, Chapter 9](#) Discrete Multi-Scale Transforms in Signal Processing

See [Vol. 1, Chapter 7](#) Multirate Signal Processing for Software Radio Architectures

References

- [1] N. Aboutorab, W. Hardjawana, B. Vucetic, A new iterative Doppler-assisted channel estimation joint with parallel ICI cancellation for high-mobility MIMO-OFDM systems, *IEEE Trans. Veh. Technol.* 61 (4) (2012) 1577–1589.
- [2] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 10th printing, Dover Publications, New York, 1972.
- [3] R. Agarwal, V. Majjigi, Z. Han, R. Vannithamby, J. Cioffi, Low complexity resource allocation with opportunistic feedback over downlink OFDMA networks, *IEEE J. Sel. Areas Commun.* 26 (8) (2008) 1462–1472.
- [4] A.N. Akansu, P. Duhamel, X. Lin, M. de Courville, Orthogonal transmultiplexers in communication: a review, *IEEE Trans Signal Process.* 46 (4) (1998) 979–995.
- [5] S.M. Alamouti, A simple transmit diversity technique for wireless communications, *J. Sel. Areas Commun.* 16 (8) (1998) 1451–1458.
- [6] S. Andreoli, H.G. McClure, P. Banelli, S. Cacopardi, Digital linearizer for RF amplifiers, *IEEE Trans. Broadcast.* 43 (1) (1997) 12–19.
- [7] D. Angelosante, G.B. Giannakis, N.D. Sidiropoulos, Estimating multiple frequency-hopping signal parameters via sparse linear regression, *IEEE Trans. Signal Process.* 58 (10) (2010) 5044–5056.
- [8] J. Armstrong, Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering, *Electron. Lett.* 38 (8) (2002) 246–247.
- [9] E. Axell, E.G. Larsson, Optimal and sub-optimal spectrum sensing of OFDM signals in known and unknown noise variance, *IEEE J. Sel. Areas Commun.* 29 (2) (2011) 290–304.
- [10] E. Axell, G. Leus, E.G. Larsson, H.V. Poor, Spectrum sensing for cognitive radio: state-of-the-art and recent advances, *IEEE Signal Process. Mag.* 29 (3) (2012) 101–116.
- [11] E. Baccarelli, M. Biagi, Optimal integer bit-loading for multicarrier ADSL systems subject to spectral-compatibility limits, *Signal Process.* 84 (4) (2004) 729–741.
- [12] A.R.S. Bahai, B.R. Saltzberg, M. Ergen, *Multi-Carrier Digital Communications*, second ed., Springer, Boston, 2004.
- [13] G. Bafna, P. Banelli, L. Rugini, Three-stage centralized spectrum sensing of OFDM signals, in: *IEEE Int. Work. Signal Process. Adv. Wireless Commun.*, San Francisco, CA, June, 2011, pp. 26–29.
- [14] P. Banelli, Theoretical analysis and performance of OFDM signals in nonlinear fading channels, *IEEE Trans. Wireless Commun.* 2 (2) (2003) 284–293.
- [15] P. Banelli, S. Cacopardi, Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels, *IEEE Trans. Commun.* 48 (3) (2000) 430–441.
- [16] P. Banelli, G. Baruffa, Mixed BB-IF predistortion of OFDM signals in non-linear channels, *IEEE Trans. Broadcast.* 47 (2) (2001) 137–146.
- [17] P. Banelli, G. Baruffa, S. Cacopardi, Effects of HPA non linearity on frequency multiplexed OFDM signals, *IEEE Trans. Broadcast.* 47 (2) (2001) 123–136.

- [18] P. Banelli, R.C. Cannizzaro, L. Rugini, Data-aided Kalman tracking for channel estimation in Doppler-affected OFDM systems, in: IEEE Int. Conf. Acoust. Speech Signal Process., Honolulu, HI, April 15–20, 2007.
- [19] P. Banelli, L. Rugini, An H-infinity filtering approach for robust tracking of OFDM doubly-selective channels, in: IEEE Int. Work. Signal Process. Adv. Wireless Commun., Marrakech, Morocco, June 2010, pp. 20–23.
- [20] S. Barbarossa, M. Pompili, G.B. Giannakis, Channel-independent synchronization of orthogonal frequency division multiple access systems, *IEEE J. Sel. Areas Commun.* 20 (2) (2002) 474–486.
- [21] S. Barbarossa, Distributed estimation and detection with applications in wireless communication and sensing networks, E-Reference Signal Processing, Elsevier, 2013.
- [22] A. Barbieri, D. Fertonani, G. Colavolpe, Time-frequency packing for linear modulations: Spectral efficiency and practical detection schemes, *IEEE Trans. Commun.* 57 (10) (2009) 2951–2959.
- [23] G. Baruffa, G. Reali, A fast algorithm to find generic odd and even order intermodulation products, *IEEE Trans. Wireless Commun.* 6 (10) (2007) 3749–3759.
- [24] E. Bayer-Fluckiger, F. Oggier, E. Viterbo, New algebraic constructions of rotated Z^n -lattice constellations for the Rayleigh fading channel, *IEEE Trans. Inform. Theory* 50 (4) (2004) 702–714.
- [25] J.A. Bazerque, G.B. Giannakis, Distributed spectrum sensing for cognitive radio networks by exploiting sparsity, *IEEE Trans. Signal Process.* 58 (3) (2010) 1847–1862.
- [26] R.W. Bäuml, R.F.H. Fisher, J.B. Huber, Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping, *Electron. Lett.* 32 (22) (1996) 2056–2057.
- [27] P. Bello, Characterization of randomly time-variant linear channels, *IEEE Trans. Commun. Syst.* 11 (4) (1963) 360–393.
- [28] S. Benedetto, E. Biglieri, Principles of Digital Transmission: With Wireless Applications, Kluwer Academic / Plenum Publishers, New York, 1999.
- [29] J.A.C. Bingham, Multicarrier modulation for data transmission: an idea whose time has come, *IEEE Commun. Mag.* 28 (5) (1990) 5–14.
- [30] N. Blachman, Band-pass nonlinearities, *IEEE Trans. Inform. Theory* 10 (2) (1964) 162–164.
- [31] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, L. Coulot, Sparse sampling of signal innovations, *IEEE Signal Process. Mag.* 25 (2) (2008) 31–40.
- [32] J. Boutros, E. Viterbo, Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel, *IEEE Trans. Inform. Theory* 44 (4) (1998) 1453–1467.
- [33] D.K. Borah, B.T. Hart, Frequency-selective fading channel estimation with a polynomial time-varying channel model, *IEEE Trans. Commun.* 47 (6) (1999) 862–873.
- [34] B. Borna, T.N. Davidson, Efficient design of FMT systems, *IEEE Trans. Commun.* 54 (5) (2006) 794–797.
- [35] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, Cambridge, UK, 2004.
- [36] H. Bölskei, Orthogonal frequency division multiplexing based on offset QAM, in: H.G. Feichtinger, T. Strohmer (Eds.), Advances in Gabor Analysis, Birkhäuser, Boston, 2003.
- [37] H. Bölskei, MIMO-OFDM wireless systems: Basics, perspectives, and challenges, *IEEE Wireless Commun.* 13 (4) (2006) 31–37.
- [38] A. Bury, J. Egle, J. Lindner, Diversity comparison of spreading transforms for multicarrier spread spectrum transmission, *IEEE Trans. Commun.* 51 (5) (2003) 774–781.
- [39] V.R. Cadambe, S.A. Jafar, Interference alignment and degrees of freedom of the K-user interference channel, *IEEE Trans. Inform. Theory* 54 (8) (2008) 3425–3441.
- [40] V.R. Cadambe, S.A. Jafar, S. Shamai, Interference alignment on the deterministic channel and application to fully connected Gaussian interference networks, *IEEE Trans. Inform. Theory* 55 (1) (2009) 269–274.
- [41] X. Cai, G.B. Giannakis, Bounding performance and suppressing intercarrier interference in wireless mobile OFDM, *IEEE Trans. Commun.* 51 (12) (2003) 2047–2056.

- [42] X. Cai, S. Zhou, G.B. Giannakis, Group-orthogonal multicarrier CDMA, *IEEE Trans. Commun.* 52 (1) (2004) 90–99.
- [43] E.J. Candes, M.B. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.* 25 (2) (2008) 21–30.
- [44] C.R. Cannizzaro, P. Banelli, G. Leus, Adaptive channel estimation for OFDM systems with Doppler spreads, in: *IEEE Int. Work. Signal Process. Adv. Wireless Commun.*, Cannes, France, July 2006, pp. 2–5.
- [45] P.C. Carrascosa, M. Stojanovic, Adaptive channel estimation and data detection for underwater acoustic MIMO-OFDM systems, *IEEE J. Ocean. Eng.* 35 (3) (2010) 635–646.
- [46] R. Cendrillon, M. Moonen, Efficient equalizers for single and multi-carrier environments with known symbol padding, in: *IEEE Int. Symp. Signal Process. Appl.*, Kuala Lumpur, Malaysia, August 13–16, 2001.
- [47] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, Optimal multiuser spectrum balancing for digital subscriber lines, *IEEE Trans. Commun.* 54 (5) (2006) 922–933.
- [48] R.W. Chang, Synthesis of band-limited orthogonal signals for multichannel data transmission, *Bell Syst. Tech. J.* 45 (10) (1966) 1775–1796.
- [49] B.-S. Chen, C.-Y. Yang, W.-J. Liao, Robust fast time-varying multipath fading channel estimation and equalization for MIMO-OFDM systems via a fuzzy method, *IEEE Trans. Veh. Technol.* 61 (4) (2012) 1599–1609.
- [50] J. Chen, R. Berry, M. Honig, Limited feedback schemes for downlink OFDMA based on sub-channel groups, *IEEE J. Sel. Areas Commun.* 26 (8) (2008) 1451–1461.
- [51] G. Cherubini, E. Eleftheriou, S. Ölcer, J.M. Cioffi, Filter bank modulation techniques for very high speed digital subscriber lines, *IEEE Commun. Mag.* 38 (5) (2000) 98–104.
- [52] H. Chen, A.M. Haimovich, Iterative estimation and cancellation of clipping noise for OFDM signals, *IEEE Commun. Lett.* 7 (7) (2003) 305–307.
- [53] T.D. Chiueh, P.Y. Tsai, *OFDM Baseband Receiver Design for Wireless Communications*, Wiley, Singapore, 2007.
- [54] E.H. Choi, W. Choi, J.G. Andrews, B.F. Womack, Power loading using order mapping in OFDM systems with limited feedback, *IEEE Signal Process. Lett.* 15 (2008) 545–548.
- [55] J. Choi, R.W. Heath, Interpolation based transmit beamforming for MIMO-OFDM with limited feedback, *IEEE Trans. Signal Process.* 53 (11) (2005) 4125–4135.
- [56] Y.-S. Choi, P.J. Voltz, F.A. Cassara, On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels, *IEEE Trans. Commun.* 49 (8) (2001) 1375–1387.
- [57] P.S. Chow, J.M. Cioffi, J.A.C. Bingham, A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels, *IEEE Trans. Commun.* 43 (2/3/4) (1995) 773–775.
- [58] P.S. Chow, J.C. Tu, J.M. Cioffi, Performance evaluation of a multichannel transceiver system for ADSL and VHDSL services, *IEEE J. Sel. Areas Commun.* 9 (6) (1991) 909–919.
- [59] L.J. Cimini, Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing, *IEEE Trans. Commun.* 33 (7) (1985) 665–675.
- [60] COST 207, Digital land mobile radio communications, Office for Official Publications of the European Communities, Final report, Luxembourg, 1989.
- [61] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
- [62] W. Dang, M. Tao, H. Mu, J. Huang, Subcarrier-pair based resource allocation for cooperative multi-relay OFDM systems, *IEEE Trans. Wireless Commun.* 9 (5) (2010) 1640–1649.
- [63] D. Dardari, Joint clip and quantization effects characterization in OFDM receivers, *IEEE Trans. Circ. Syst. I, Reg. Papers* 53 (8) (2006) 1741–1748.
- [64] D. Dardari, V. Tralli, A. Vaccari, A theoretical characterization of nonlinear distortion effects in OFDM systems, *IEEE Trans. Commun.* 48 (10) (2000) 1755–1764.

- [65] W.B. Davenport, W.L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.
- [66] T. Davidson, The art of communicating over MIMO channels: contemporary perspectives, E-Reference Signal Processing, Elsevier, 2013.
- [67] L. Ding, G.T. Zhou, D.R. Morgan, Z. Ma, J.S. Kenney, J. Kim, C.R. Giardina, A robust digital baseband predistorter constructed using memory polynomials, *IEEE Trans. Commun.* 52 (1) (2004) 159–165.
- [68] Y. Ding, M. Uysal, Amplify-and-forward cooperative OFDM with multiple-relays: performance analysis and relay selection methods, *IEEE Trans. Wireless Commun.* 8 (10) (2009) 4963–4968.
- [69] M.L. Doelz, E.T. Heald, D.L. Martin, Binary data transmission techniques for linear systems, *Proc. IRE* 45 (1957) 656–661.
- [70] O. El Ayach, S.W. Peters, R.W. Heath, The feasibility of interference alignment over measured MIMO-OFDM channels, *IEEE Trans. Veh. Technol.* 59 (9) (2010) 4309–4321.
- [71] ETSI European Telecommunications Standards Institute, Digital audio broadcasting (DAB): data broadcasting—MPEG-2 TS streaming, ETSI TS 102 427, V1.1.1, Sophia Antipolis, France, July 2005.
- [72] ETSI European Telecommunications Standards Institute, Digital video broadcasting (DVB): framing structure channel coding and modulation for digital terrestrial television, ETSI EN 300 744, V1.5.1, Sophia Antipolis, France, November 2004.
- [73] ETSI, European Telecommunications Standards Institute, Digital Video Broadcasting (DVB), Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2), ETSI EN 302 755, V1.1.1, Sophia Antipolis, France, September 2009.
- [74] K. Fang, L. Rugini, G. Leus, Block transmissions over doubly selective channels: iterative channel estimation and turbo equalization, *EURASIP J. Adv. Signal Process.* (2010) 1–13 Article ID 974652.
- [75] K. Fang, L. Rugini, G. Leus, Low-complexity block turbo equalization for OFDM systems in time-varying channels, *IEEE Trans. Signal Process.* 56 (11) (2008) 5555–5566.
- [76] B. Farhang-Boroujeny, OFDM versus filter bank multicarrier, *IEEE Signal Process. Mag.* 28 (3) (2011) 92–112.
- [77] B. Farhang-Boroujeny, R. Kempter, Multicarrier communication techniques for spectrum sensing and communication in cognitive radios, *IEEE Commun. Mag.* 46 (4) (2008) 80–85.
- [78] R.F.H. Fischer, J.B. Huber, A new loading algorithm for discrete multitone transmission, in: IEEE Global Telecommun. Conf., London, UK, November 18–22, 1996.
- [79] G.J. Foschini, M.J. Gans, On limits of wireless communications in a fading environment when using multiple antennas, *Wireless Pers. Commun.* 6 (3) (1998) 311–335.
- [80] J.C. Fuenzalida, O. Shimbo, W.L. Cook, Time-domain analysis of intermodulation effects caused by nonlinear amplifiers, *COMSAT Tech. Rev.* 3 (1) (1973) 89–143.
- [81] T. Fusco, A. Petrella, M. Tanda, Data-aided symbol timing and CFO synchronization for filter bank multicarrier systems, *IEEE Trans. Wireless Commun.* 8 (5) (2009) 2705–2715.
- [82] F. Gao, R. Zhang, Y.-C. Liang, Channel estimation for OFDM modulated two-way relay networks, *IEEE Trans. Signal Process.* 57 (11) (2009) 4443–4455.
- [83] J. Gao, H. Liu, Low-complexity MAP channel estimation for mobile MIMO-OFDM systems, *IEEE Trans. Wireless Commun.* 7 (3) (2008) 774–780.
- [84] D. Gesbert, M. Shafi, D. Shiu, P.J. Smith, A. Naguib, From theory to practice: An overview of MIMO space-time coded wireless systems, *IEEE J. Sel. Areas Commun.* 21 (3) (2003) 281–302.
- [85] M. Ghogho, Synchronization and localization, E-Reference Signal Processing, Elsevier, 2013.
- [86] G.B. Giannakis, Spectrum Sensing for Cognitive Radio, E-reference Signal Processing, Elsevier, 2013.
- [87] X. Giraud, E. Boutillet, J.C. Belfiore, Algebraic tools to build modulation schemes for fading channels, *IEEE Trans. Inform. Theory* 43 (3) (1997) 938–952.

- [88] A. Goldsmith, S.A. Jafar, N. Jindal, S. Vishwanath, Capacity limits of MIMO channels, *IEEE J. Sel. Areas Commun.* 21 (5) (2003) 684–702.
- [89] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., Johns Hopkins University Press, London, UK, 1996.
- [90] A. Gorokhov, J.-P. Linnartz, Robust OFDM receivers for dispersive time-varying channels: equalization and channel acquisition, *IEEE Trans. Commun.* 52 (4) (2004) 572–583.
- [91] R. Haas, J.-C. Belfiore, A time-frequency well-localized pulse for multiple carrier transmission, *Wireless Pers. Commun.* 5 (1) (1997) 1–18.
- [92] I. Hammerstrom, A. Wittneben, Power allocation schemes for amplify-and-forward MIMO-OFDM relay links, *IEEE Trans. Wireless Commun.* 6 (8) (2007) 2798–2802.
- [93] S.H. Han, J.H. Lee, An overview of peak-to-average power ratio reduction techniques for multicarrier transmission, *IEEE Wireless Commun.* 12 (2) (2005) 56–65.
- [94] L. Hanzo, M. Münster, B.J. Choi, T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*, IEEE Press/John Wiley & Sons, Chichester, UK, 2003.
- [95] S. Hara, R. Prasad, Overview of multicarrier CDMA, *IEEE Commun. Mag.* 35 (12) (1997) 126–133.
- [96] S. Haykin, Cognitive radio: brain-empowered wireless communications, *IEEE J. Sel. Areas Commun.* 23 (2) (2005) 201–220.
- [97] F. Hlawatsch, G. Matz, *Wireless Communications Over Rapidly Time-Varying Channels*, Academic Press, Oxford, UK, 2011.
- [98] P. Hoeher, S. Kaiser, P. Robertson, Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering, in: *IEEE Int. Conf. Acoust. Speech Signal Process.*, Munich, Germany, April 21–24, 1997.
- [99] T. Hrycak, S. Das, G. Matz, H.G. Feichtinger, Practical estimation of rapidly varying channels for OFDM systems, *IEEE Trans. Commun.* 59 (11) (2011) 3040–3048.
- [100] J. Huang, S. Zhou, J. Huang, C.R. Berger, P. Willett, Progressive inter-carrier interference equalization for OFDM transmission over time-varying underwater acoustic channels, *IEEE J. Sel. Topics Signal Process.* 5 (8) (2011) 1524–1536.
- [101] T. Hunziker, D. Dahlhaus, Iterative detection for multicarrier transmission employing time-frequency concentrated pulses, *IEEE Trans. Commun.* 51 (4) (2003) 641–651.
- [102] A.A. Hutter, R. Hasholzner, J.S. Hammerschmidt, Channel estimation for mobile OFDM systems, in: *IEEE Veh. Technol. Conf. Fall*, Amsterdam, The Netherlands, September 19–22, 1999.
- [103] T. Hwang, C. Yang, G. Wu, S. Li, G.Y. Li, OFDM and its wireless applications: a survey, *IEEE Veh. Technol.* 58 (4) (2009) 1673–1694.
- [104] IEEE, Institute of Electrical and Electronics Engineers, Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: high-speed physical layer in the 5 GHz band, *IEEE Std 802.11a-1999*, New York, NY, December 1999.
- [105] IEEE, Institute of Electrical and Electronics Engineers, Part 16: air interface for fixed and mobile broadband wireless access systems, *IEEE Std 802.16-2004*, New York, NY, October 2004.
- [106] T. Ihlainen, A. Ikhlef, J. Louveaux, M. Renfors, Channel equalization for multi-antenna FBMC/OQAM receivers, *IEEE Trans. Veh. Technol.* 60 (5) (2011) 2070–2085.
- [107] ITU, International Telecommunication Union, Series G: transmission systems and media, digital systems and networks, Asymmetric digital subscriber line (ADSL) transceivers, *ITU-T Recommendation G.992.1*, Geneva, Switzerland, June 1999.
- [108] J. Jaldén, B. Ottersten, On the complexity of sphere decoding in digital communications, *IEEE Trans. Signal Process.* 53 (4) (2005) 1474–1484.
- [109] J. Jang, K.B. Lee, Transmit power adaptation for multiuser OFDM systems, *IEEE J. Sel. Areas Commun.* 21 (2) (2003) 171–178.

- [110] W.G. Jeon, K.H. Chang, Y.S. Cho, An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels, *IEEE Trans. Commun.* 47 (1) (1999) 27–32.
- [111] M. Jiang, L. Hanzo, Multiuser MIMO-OFDM for next-generation wireless systems, *Proc. IEEE* 95 (7) (2007) 1430–1469.
- [112] T. Jiang, Y. Wu, An overview: peak-to-average power ratio reduction techniques for OFDM signals, *IEEE Trans. Broadcast.* 54 (2) (2008) 257–268.
- [113] A.E. Jones, T.A. Wilkinson, S.K. Barton, Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission scheme, *Electron. Lett.* 30 (22) (1994) 2098–2099.
- [114] H.W. Kang, Y.S. Cho, D.H. Youn, On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter, *IEEE Trans. Commun.* 47 (4) (1999) 522–526.
- [115] T. Kang, R.A. Iltis, Iterative carrier frequency offset and channel estimation for underwater acoustic OFDM systems, *IEEE J. Sel. Areas Commun.* 26 (9) (2008) 1650–1661.
- [116] G. Karam, H. Sari, Data predistortion techniques using intersymbol interpolation, *IEEE Trans. Commun.* 38 (10) (1990) 1716–1723.
- [117] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, 1993.
- [118] A. Kaye, D. George, M. Eric, Analysis and compensation of bandpass nonlinearities for communications, *IEEE Trans. Commun.* 20 (5) (1972) 965–972.
- [119] V. Kekatos, G.B. Giannakis, From sparse signals to sparse residuals for robust sensing, *IEEE Trans. Signal Process.* 59 (7) (2011) 3355–3368.
- [120] T. Keller, L. Hanzo, Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications, *Proc. IEEE* 88 (5) (2000) 611–640.
- [121] N. Khaled, B. Mondal, G. Leus, R.W. Heath, F. Petre, Interpolation-based multi-mode precoding for MIMO-OFDM systems with limited feedback, *IEEE Trans. Wireless Commun.* 6 (3) (2007) 1003–1013.
- [122] D. Kim, G.L. Stüber, Clipping noise mitigation for OFDM by decision-aided reconstruction, *IEEE Commun. Lett.* 3 (1) (1999) 4–6.
- [123] H. Kim, J.K. Tugnait, Turbo equalization for doubly-selective fading channels using nonlinear Kalman filtering and basis expansion models, *IEEE Trans. Wireless Commun.* 9 (6) (2010) 2076–2087.
- [124] I. Kim, I.-S. Park, Y.H. Lee, Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM, *IEEE Trans. Veh. Technol.* 55 (4) (2006) 1195–1207.
- [125] J.-G. Kim, J.-T. Lim, MAP-based channel estimation for MIMO-OFDM over fast Rayleigh fading channels, *IEEE Trans. Veh. Technol.* 57 (3) (2008) 1963–1968.
- [126] S. Kim, Angle-domain frequency-selective sparse channel estimation for underwater MIMO-OFDM systems, *IEEE Commun. Lett.* 16 (5) (2012) 685–687.
- [127] S. Kondo, B. Milstein, Performance of multicarrier DS CDMA systems, *IEEE Trans. Commun.* 44 (2) (1996) 238–246.
- [128] W. Kozek, A.F. Molisch, Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels, *IEEE J. Sel. Areas Commun.* 16 (8) (1998) 1579–1589.
- [129] B.S. Krongold, D.L. Jones, PAR reduction in OFDM via active constellation extension, *IEEE Trans. Broadcast.* 49 (3) (2003) 258–268.
- [130] B.S. Krongold, K. Ramchandran, D.L. Jones, Computationally efficient optimal power allocation algorithms for multicarrier communication systems, *IEEE Trans. Commun.* 48 (1) (2000) 23–27.
- [131] A. Kuhne, A. Klein, Throughput analysis of multi-user OFDMA-systems using imperfect CQI feedback and diversity techniques, *IEEE J. Sel. Areas Commun.* 26 (8) (2008) 1440–1450.
- [132] I.-W. Lai, T.-D. Chiueh, One-dimensional interpolation based channel estimation for mobile DVB-H reception, in: *IEEE Int. Symp. Circ. Syst.*, Kos, Greece, May 2006, pp. 21–24.

- [133] D. Landström, S.K. Wilson, J.-J. van de Beek, P. Ödлин, P.O. Börjesson, Symbol time offset estimation in coherent OFDM systems, *IEEE Trans. Commun.* 50 (4) (2002) 545–549.
- [134] J.N. Laneman, D.N.C. Tse, G.W. Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behaviour, *IEEE Trans. Inform. Theory* 50 (12) (2004) 3062–3080.
- [135] J.N. Laneman, G.W. Wornell, Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks, *IEEE Trans. Inform. Theory* 49 (10) (2003) 2415–2425.
- [136] B. Le Floch, M. Alard, C. Berrou, Coded orthogonal frequency division multiplex, *Proc. IEEE* 83 (6) (1995) 982–996.
- [137] A. Leshem, I. Bergel, Signal processing for digital subscriber lines, *E-Reference Signal Processing*, Elsevier, 2013.
- [138] B.K. Letaief, W. Zhang, Cooperative communications for cognitive radio networks, *Proc. IEEE* 97 (5) (2009) 878–893.
- [139] G. Leus, On the estimation of rapidly time-varying channels, in: *European Signal Process. Conf.*, Vienna, Austria, September 6–10, 2004.
- [140] G. Leus, P. van Walree, Multiband OFDM for covert acoustic communications, *IEEE J. Sel. Areas Commun.* 26 (9) (2008) 1662–1673.
- [141] B. Li, J. Huang, S. Zhou, K. Ball, M. Stojanovic, L. Freitag, P. Willett, MIMO-OFDM for high-rate underwater acoustic communications, *IEEE J. Ocean. Eng.* 34 (4) (2009) 634–644.
- [142] B. Li, S. Zhou, M. Stojanovic, L. Freitag, P. Willett, Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts, *IEEE J. Ocean. Eng.* 33 (2) (2008) 198–209.
- [143] X. Li, F. Ng, T. Han, Carrier frequency offset mitigation in asynchronous cooperative OFDM transmissions, *IEEE Trans. Signal Process.* 56 (2) (2008) 675–685.
- [144] Y. Li, L.J. Cimini, Bounds on the interchannel interference of OFDM in time-varying impairments, *IEEE Trans. Commun.* 49 (3) (2001) 401–404.
- [145] Y. Li, G.L. Stüber, *Orthogonal Frequency Division Multiplexing for Wireless Communications*, Springer, New York, 2006.
- [146] Y.-P. Lin, S.-M. Phoong, BER minimized OFDM systems with channel independent precoders, *IEEE Trans. Signal Process.* 51 (9) (2003) 2369–2380.
- [147] L. Liu, S. Zhou, J.H. Cui, The prospects and problems of wireless communication for underwater sensor networks, *Wireless Commun. Mobile Comput.* 8 (8) (2008) 977–994.
- [148] Z. Liu, Y. Xin, G.B. Giannakis, Linear constellation precoding for OFDM with maximum multipath diversity and coding gains, *IEEE Trans. Commun.* 51 (3) (2003) 416–427.
- [149] V. Lottici, M. Luise, C. Saccomando, F. Spalla, Blind carrier frequency tracking for filterbank multicarrier wireless communications, *IEEE Trans. Commun.* 53 (10) (2005) 1762–1772.
- [150] D.J. Love, R.W. Heath, OFDM power loading using limited feedback, *IEEE Trans. Veh. Technol.* 54 (5) (2005) 1773–1780.
- [151] Z.-Q. Luo, Signal processing and optimal resource allocation for the interference channel, *E-Reference Signal Processing*, Elsevier, 2013.
- [152] X. Ma, G.B. Giannakis, S. Ohno, Optimal training for block transmissions over doubly selective wireless fading channels, *IEEE Trans. Signal Process.* 51 (5) (2003) 1351–1366.
- [153] X. Ma, C. Tepedelenlioglu, G.B. Giannakis, S. Barbarossa, Non-data-aided carrier offset estimators for OFDM with null subcarriers: identifiability, algorithms, and performance, *IEEE J. Sel. Areas Commun.* 19 (12) (2001) 2504–2515.
- [154] Y. Ma, N. Yi, R. Tafazolli, Bit and power loading for OFDM-based three-node relaying communications, *IEEE Trans. Signal Process.* 56 (7) (2008) 3236–3247.
- [155] M.A. Maddah-Ali, A.S. Motahari, A.K. Khandani, Communication over MIMO X channels: interference alignment, decomposition, and performance analysis, *IEEE Trans. Inform. Theory* 54 (8) (2008) 3457–3470.

- [156] G. Matz, D. Schafhuber, K. Gröchenig, M. Hartmann, F. Hlawatsch, Analysis, optimization, and implementation of low-interference wireless multicarrier systems, *IEEE Trans. Wireless Commun.* 6 (5) (2007) 1921–1931.
- [157] Y. Medjahdi, M. Terre, D. Le Ruyet, D. Roviras, A. Dziri, Performance analysis in the downlink of asynchronous OFDM/FBMC based multi-cellular networks, *IEEE Trans. Wireless Commun.* 10 (8) (2011) 2630–2639.
- [158] J. Minkoff, The role of AM-to-PM conversion in memoryless nonlinear systems, *IEEE Trans. Commun.* 33 (2) (1985) 139–144.
- [159] M. Mishali, Y.C. Eldar, From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals, *IEEE Sel. Topics Signal Process.* 4 (2) (2010) 375–391.
- [160] J. Mitola, G.Q. Maguire, Cognitive radio: making software radios more personal, *IEEE Pers. Commun.* 6 (4) (1999) 13–18.
- [161] P.H. Moose, A technique for orthogonal frequency division multiplexing frequency offset correction, *IEEE Trans. Commun.* 42 (10) (1994) 2908–2914.
- [162] M. Morelli, Timing and frequency synchronization for the uplink of an OFDMA system, *IEEE Trans. Commun.* 52 (2) (2004) 296–306.
- [163] M. Morelli, C.-C.J. Kuo, M.-O. Pun, Synchronization techniques for orthogonal frequency division multiple access (OFDMA): a tutorial review, *Proc. IEEE* 95 (7) (2007) 1394–1427.
- [164] M. Morelli, U. Mengali, A comparison of pilot-aided channel estimation methods for OFDM systems, *IEEE Trans. Signal Process.* 49 (12) (2001) 3065–3073.
- [165] Y. Mostofi, D.C. Cox, Timing synchronization in high mobility OFDM systems, in: *IEEE Int. Conf. Commun.*, Paris, France, June 20–24, 2004.
- [166] B. Muquet, M. de Courville, P. Duhamel, G. Giannakis, OFDM with trailing zeros versus OFDM with cyclic prefix: links, comparisons and application to the HiperLAN/2 system, in: *IEEE Int. Conf. Commun.*, New Orleans, LA, June 2000, pp. 18–22.
- [167] B. Muquet, Z. Wang, G.B. Giannakis, M. de Courville, P. Duhamel, Cyclic prefixing or zero padding for wireless multicarrier transmissions?, *IEEE Trans. Commun.* 50 (12) (2002) 2136–2148.
- [168] K. Muralidhar, K.H. Li, A low-complexity Kalman approach for channel estimation in doubly-selective OFDM systems, *IEEE Signal Process. Lett.* 16 (7) (2009) 632–635.
- [169] S.H. Müller, J.B. Huber, OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences, *Electron. Lett.* 33 (5) (1997) 368–369.
- [170] A. Nosratinia, T.E. Hunter, A. Hedayat, Cooperative communication in wireless networks, *IEEE Commun. Mag.* 42 (10) (2004) 74–80.
- [171] H. Ochiai, H. Imai, Performance analysis of deliberately clipped OFDM signals, *IEEE Trans. Commun.* 50 (1) (2002) 89–101.
- [172] S. Ohno, G.B. Giannakis, Optimal training and redundant precoding for block transmissions with application to wireless OFDM, *IEEE Trans. Commun.* 50 (12) (2002) 2113–2123.
- [173] A.V. Oppenheim, R.W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, 1989.
- [174] M.K. Ozdemir, H. Arslan, Channel estimation for wireless OFDM systems, *IEEE Commun. Surv. Tutor.* 9 (2) (2007) 18–48.
- [175] A.Y. Panah, R.W. Heath, Single-user and multicast OFDM power loading with nonregenerative relaying, *IEEE Trans. Veh. Technol.* 58 (9) (2009) 4890–4902.
- [176] N. Papandreou, T. Antonakopoulos, Bit and power allocation in constrained multicarrier systems: the single-user case, *EURASIP J. Adv. Signal Process.* (2008) 1–14 Article ID 643081.
- [177] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, third ed., McGraw-Hill, New York, 1991.

- [178] A.J. Paulraj, D.A. Gore, R.U. Nabar, H. Bölskei, An overview of MIMO communications -A key to gigabit wireless, *Proc. IEEE* 92 (2) (2004) 198–218.
- [179] R.Peled, A. Ruiz, Frequency domain data transmission using reduced computational complexity algorithms, in: *IEEE Int. Conf. Acoust. Speech Signal Process.*, Denver, CO., April 9–11, 1980.
- [180] K. Pelekanakis, A.B. Bagheroer, Exploiting space-time-frequency diversity with MIMO-OFDM for under-water acoustic communications, *IEEE J. Ocean. Eng.* 36 (4) (2011) 502–513.
- [181] M. Poggioni, L. Rugini, P. Banelli, DVB-T/H and T-DMB: Physical layer performance comparison in fast mobile channels, *IEEE Trans. Broadcast.* 55 (4) (2009) 719–730.
- [182] C. Polprasert, J.A. Ritcey, M. Stojanovic, Capacity of OFDM systems over fading underwater acoustic channels, *IEEE J. Ocean. Eng.* 36 (4) (2011) 514–524.
- [183] J.G. Proakis, M. Salehi, *Digital Communications*, fifth ed., McGraw-Hill, New York, 2008.
- [184] U. Reimers, Digital video broadcasting, *IEEE Commun. Mag.* 36 (6) (1998) 104–110.
- [185] P. Robertson, S. Kaiser, The effects of Doppler spreads in OFDM(A) mobile radio systems, in: *IEEE Veh. Technol. Conf. Fall*, Amsterdam, The Netherlands, September 19–22, 1999.
- [186] H.E. Rowe, Memoryless nonlinearities with Gaussian inputs: elementary results, *Bell Syst. Tech. J.* 61 (7) (1982) 1519–1525.
- [187] L. Rugini, Linear equalization for multicode MC-CDMA downlink channels, *IEEE Commun. Lett.* 16 (9) (2012) 1353–1356.
- [188] L. Rugini, P. Banelli, BER of OFDM systems impaired by carrier frequency offset in multipath fading channels, *IEEE Trans. Wireless Commun.* 4 (5) (2005) 2279–2288.
- [189] L. Rugini, P. Banelli, Probability of error of linearly modulated signals with Gaussian cochannel interference in maximally correlated Rayleigh fading channels, *EURASIP J. Wireless Commun. Network* (2010) 1–13 ID 193183.
- [190] L. Rugini, P. Banelli, Pilot-aided estimation of carrier frequency offsets and channel impulse responses for OFDM cooperative communications, in: *IEEE Int. Work. Signal Process Adv. Wireless Commun.*, Cesme, Turkey, June 2012, pp. 17–20.
- [191] L. Rugini, P. Banelli, S. Cacopardi, A full-rank regularization technique for MMSE detection in multiuser CDMA systems, *IEEE Commun. Lett.* 9 (1) (2005) 34–36.
- [192] L. Rugini, P. Banelli, K. Fang, G. Leus, Enhanced turbo MMSE equalization for MIMO-OFDM over rapidly time-varying frequency-selective channels, in: *IEEE Int. Work. Signal Process Adv. Wireless Commun.*, Perugia, Italy, June, 2009, pp. 21–24.
- [193] L. Rugini, P. Banelli, G. Leus, Block DFE and windowing for Doppler-affected OFDM systems, *IEEE Int. Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2005)*, New York, USA, June 5–8, 2005, pp. 470–474.
- [194] L. Rugini, P. Banelli, G. Leus, Simple equalization of time-varying channels for OFDM, *IEEE Commun. Lett.* 9 (7) (2005) 619–621.
- [195] L. Rugini, P. Banelli, G. Leus, Low-complexity banded equalizers for OFDM systems in Doppler spread channels, *EURASIP J. Appl. Signal Process.* (2006) 1–13 ID 67404.
- [196] M. Russell, G.L. Stüber, Interchannel interference analysis for OFDM in a mobile environment, in: *IEEE Veh. Technol. Conf.*, Chicago, IL, July 1995, pp. 25–28.
- [197] A.A.M. Saleh, Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers, *IEEE Trans. Commun.* 29 (11) (1981) 1715–1720.
- [198] F. Sanzi, J. Speidel, An adaptive two-dimensional channel estimator for wireless OFDM with application to mobile DVB-T, *IEEE Trans. Broadcast.* 46 (2) (2000) 128–133.
- [199] H. Sari, G. Karam, I. Jeanclaude, Transmission techniques for digital terrestrial TV broadcasting, *IEEE Commun. Mag.* 33 (2) (1995) 100–109.

- [200] A. Scaglione, S. Barbarossa, G.B. Giannakis, Filterbank transceivers optimizing information rate in block transmissions over dispersive channels, *IEEE Trans. Inform. Theory* 45 (3) (1999) 1019–1032.
- [201] A. Scaglione, G.B. Giannakis, S. Barbarossa, Redundant filterbank precoders and equalizers—Part I: unification and optimal designs, *IEEE Trans. Signal Process.* 47 (7) (1999) 1988–2006.
- [202] A. Scaglione, G.B. Giannakis, S. Barbarossa, Redundant filterbank precoders and equalizers—Part II: Blind channel estimation, synchronization, and direct equalization, *IEEE Trans. Signal Process.* 47 (7) (1999) 2007–2022.
- [203] D. Schafhuber, G. Matz, F. Hlawatsch, Adaptive Wiener filters for time-varying channel estimation in wireless OFDM systems, in: *IEEE Int. Conf. Acoust. Speech Signal Process.*, Hong Kong, April 6–10, 2003.
- [204] P. Schniter, Low-complexity equalization of OFDM in doubly selective channels, *IEEE Trans. Signal Process.* 52 (4) (2004) 1002–1011.
- [205] G. Scutari, D.P. Palomar, S. Barbarossa, Asynchronous iterative water-filling for Gaussian frequency-selective interference channels, *IEEE Trans. Inform. Theory* 54 (7) (2008) 2868–2878.
- [206] G. Scutari, D.P. Palomar, S. Barbarossa, The MIMO iterative water-filling algorithm, *IEEE Trans. Signal Process.* 57 (5) (2009) 1917–1935.
- [207] D. Schafhuber, G. Matz, MMSE and adaptive prediction of time-varying channels for OFDM systems, *IEEE Trans. Wireless Commun.* 4 (2) (2005) 593–602.
- [208] T.M. Schmidl, D.C. Cox, Robust frequency and timing synchronization for OFDM, *IEEE Trans. Commun.* 45 (12) (1997) 1613–1621.
- [209] M. Schwartz, The origins of carrier multiplexing: Major George Owen Squier and AT&T, *IEEE Commun. Mag.* 46 (5) (2008) 20–24.
- [210] J. Seo, S. Jang, J. Yang, W. Jeon, D.K. Kim, Analysis of pilot-aided channel estimation with optimal leakage suppression for OFDM systems, *IEEE Commun. Lett.* 14 (9) (2010) 809–811.
- [211] C. Sgraja, J. Lindner, Estimation of rapid time-variant channels for OFDM using Wiener filtering, in: *IEEE Int. Conf. Commun.*, Anchorage, AK, May 2003, pp. 11–15.
- [212] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (3–4) (1948) 379–423, 623–656.
- [213] C.E. Shannon, Communication in the presence of noise, *Proceedings of the Institute of Radio Engineers* 37 (1) (1949) 10–21, Reprinted in, *Proc. IEEE* 86 (2) (1998) 447–457.
- [214] Z. Shen, J.G. Andrews, B.L. Evans, Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints, *IEEE Trans. Wireless Commun.* 4 (6) (2005) 2726–2737.
- [215] Q. Shi, OFDM in bandpass nonlinearity, *IEEE Trans. Consum. Electron.* 42 (3) (1996) 253–258.
- [216] O. Shimbo, Effects of intermodulation, AM-PM conversion, and additive noise in multicarrier TWT systems, *Proc. IEEE* 59 (2) (1971) 230–238.
- [217] O.-S. Shin, A.M. Chan, H.T. Kung, V. Tarokh, Design of an OFDM cooperative space-time diversity system, *IEEE Trans. Veh. Technol.* 56 (4) (2007) 2203–2215.
- [218] M.K. Simon, M.-S. Alouini, *Digital Communication over Fading Channels*, second ed., Wiley-IEEE Press, Hoboken, 2005.
- [219] M.K. Simon, H. Jafarkhani, Performance evaluation of super-orthogonal space-time trellis codes using a moment generating function-based approach, *IEEE Trans. Signal Process.* 51 (11) (2003) 2739–2751.
- [220] P. Siohan, C. Siclet, N. Lacaille, Analysis and design of OFDM/OQAM systems based on filterbank theory, *IEEE Trans. Signal Process.* 50 (5) (2002) 1170–1183.
- [221] W. Siriwongpairat, A. Sadek, K.J.R. Liu, Cooperative communications protocol for multiuser OFDM networks, *IEEE Trans. Wireless Commun.* 7 (7) (2008) 2430–2435.
- [222] B. Sklar, Rayleigh fading channels in mobile digital communication systems—Part I: Characterization, *IEEE Commun. Mag.* 35 (7) (1997) 90–100.

- [223] L. Song, J.K. Tugnait, Doubly-selective fading channel equalization: a comparison of the Kalman filter approach with the basis expansion model-based equalizers, *IEEE Trans. Wireless Commun.* 8 (1) (2009) 60–65.
- [224] W.-G. Song, J.-T. Lim, Channel estimation and signal detection for MIMO-OFDM with time varying channels, *IEEE Commun. Lett.* 10 (7) (2006) 540–542.
- [225] E.A. Sourour, M. Nakagawa, Performance of orthogonal multicarrier CDMA in a multipath fading channel, *IEEE Trans. Commun.* 44 (3) (1996) 356–367.
- [226] M. Speth, S.A. Fechtel, G. Fock, H. Meyr, Optimum receiver design for wireless broad-band systems using OFDM—Part I, *IEEE Trans. Commun.* 47 (11) (1999) 1668–1677.
- [227] A. Stamoulis, S.N. Diggavi, N. Al-Dhahir, Intercarrier interference in MIMO OFDM, *IEEE Trans. Signal Process.* 50 (10) (2002) 2451–2464.
- [228] H.A. Suraweera, J. Armstrong, Performance of OFDM-based dual-hop amplify-and-forward relaying, *IEEE Commun. Lett.* 11 (9) (2007) 726–728.
- [229] V. Tarokh, N. Seshadri, A.R. Calderbank, Space-time codes for high data rate wireless communication: performance criterion and code construction, *IEEE Trans. Inform. Theory* 44 (2) (1998) 744–765.
- [230] Z. Tang, R.C. Cannizzaro, G. Leus, P. Banelli, Pilot-assisted time-varying channel estimation for OFDM systems, *IEEE Trans. Signal Process.* 55 (5) (2007) 2226–2238.
- [231] Z. Tang, G. Leus, A novel receiver architecture for single-carrier transmission over time-varying channels, *IEEE J. Sel. Areas Commun.* 26 (2) (2008) 366–377.
- [232] G. Tauböck, F. Hlawatsch, D. Eiwen, H. Rauhut, Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing, *IEEE J. Sel. Topics Signal Process.* 4 (2) (2010) 255–271.
- [233] J. Tellado, Peak to Average Power Reduction for Multicarrier Modulation, Ph.D. Thesis, Stanford University, Stanford, CA, 2000.
- [234] J. Tellado, L.M.C. Hoo, J.M. Cioffi, Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding, *IEEE Trans. Commun.* 51 (2) (2003) 218–228.
- [235] K.D. Teo, S. Ohno, Optimal MMSE finite parameter model for doubly-selective channels, in: IEEE Global Telecommun. Conf., St. Louis, MO, November 28–December 2, 2005.
- [236] C. Tepedelenlioglu, Maximum multipath diversity with linear equalization in precoded OFDM systems, *IEEE Trans. Infom. Theory* 50 (1) (2004) 232–235.
- [237] 3GPP 3rd Generation Partnership Project, Technical specification group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA): Physical channels and modulation, 3GPP TS 36.211 V8.5.0, Sophia Antipolis, France, December 2008.
- [238] A.M. Tonello, Performance limits for filtered multitone modulation in fading channels, *IEEE Trans. Wireless Commun.* 4 (5) (2005) 2121–2135.
- [239] M.K. Tsatsanis, G.B. Giannakis, Modeling and equalization of rapidly fading channels, *Int. J. Adapt. Control Signal Process.* 10 (1996) 159–176.
- [240] K. Tu, D. Fertonani, T.M. Duman, M. Stojanovic, J.G. Proakis, P. Hursky, Mitigation of intercarrier interference for OFDM over time-varying underwater acoustic channels, *IEEE J. Ocean. Eng.* 36 (2) (2011) 156–171.
- [241] J. Tugnait, Channel estimation, tracking, equalization, and precoding, E-Reference Signal Processing, Elsevier, 2013.
- [242] U. Tureli, H. Liu, M.D. Zoltowski, A high efficiency carrier estimator for OFDM communications, in: IEEE Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, November 2–5, (1997).
- [243] J.-J. van de Beek, M. Sandell, P.O. Börjesson, ML estimation of time and frequency offset in OFDM systems, *IEEE Trans. Signal Process.* 45 (7) (1997) 1800–1805.

- [244] L. Vandendorpe, Multitone spread spectrum multiple access communications system in a multipath Rician fading channel, *IEEE Trans. Veh. Technol.* 44 (2) (1995) 327–337.
- [245] P. Van Eetvelt, G. Wade, M. Tomlinson, Peak to average power reduction for OFDM schemes by selective scrambling, *Electron. Lett.* 32 (21) (1996) 1963–1964.
- [246] R. van Nee, R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House, Boston, 2000.
- [247] S. Verdú, *Multiuser Detection*, Cambridge University Press, New York, 1998.
- [248] W. Wang, R. Wu, Capacity maximization for OFDM two-hop relay system with separate power constraints, *IEEE Trans. Veh. Technol.* 58 (9) (2009) 4943–4954.
- [249] X.B. Wang, T.T. Tjhung, C.S. Ng, Reduction of peak-to-average power ratio of OFDM system using a companding technique, *IEEE Trans. Broadcast.* 45 (3) (1999) 303–307.
- [250] Z. Wang, G.B. Giannakis, Complex-field coding for OFDM over fading wireless channels, *IEEE Trans. Inform. Theory* 49 (3) (2003) 707–720.
- [251] Z. Wang, G.B. Giannakis, Wireless multicarrier communications: Where Fourier meets Shannon, *IEEE Signal Process. Mag.* 17 (3) (2000) 29–48.
- [252] Z. Wang, G.B. Giannakis, Linearly precoded or coded OFDM against wireless channel fades?, in: *IEEE Int. Work. Signal Process. Adv. Wireless Commun.*, Taiwan, March 20–23, 2001.
- [253] Z. Wang, X. Ma, G.B. Giannakis, OFDM or single-carrier block transmissions?, *IEEE Trans. Commun.* 52 (3) (2004) 380–394.
- [254] Z. Wang, S. Zhou, G.B. Giannakis, Joint coding-precoding with low-complexity turbo-decoding, *IEEE Trans. Wireless Commun.* 3 (3) (2004) 832–842.
- [255] Z. Wang, S. Zhou, G.B. Giannakis, C.R. Berger, J. Huang, Frequency-domain oversampling for zero-padded OFDM in underwater acoustic communications, *IEEE J. Ocean. Eng.* 37 (1) (2012) 14–24.
- [256] S.B. Weinstein, The history of orthogonal frequency-division multiplexing, *IEEE Commun. Mag.* 47 (11) (2009) 26–35.
- [257] S.B. Weinstein, P.M. Ebert, Data transmission for frequency-division multiplexing using the discrete Fourier transform, *IEEE Trans. Commun. Technol.* 19 (5) (1971) 628–634.
- [258] C.Y. Wong, R.S. Cheng, K.B. Letaief, R.D. Murch, Multiuser OFDM with adaptive subcarrier, bit, and power allocation, *IEEE J. Sel. Areas Commun.* 17 (10) (1999) 1747–1758.
- [259] L.-L. Yang, L. Hanzo, Performance of generalized multicarrier DS-CDMA over Nakagami-m fading channels, *IEEE Trans. Commun.* 50 (6) (2002) 956–966.
- [260] S. Yerramalli, U. Mitra, Optimal resampling of OFDM signals for multiscale-multilag underwater acoustic channels, *IEEE J. Ocean. Eng.* 36 (1) (2011) 126–138.
- [261] W. Yu, G. Ginis, J.M. Cioffi, Distributed multiuser power control for digital subscriber lines, *IEEE J. Sel. Areas Commun.* 20 (5) (2002) 1105–1115.
- [262] W. Yu, W. Rhee, S. Boyd, J.M. Cioffi, Iterative water-filling for Gaussian vector multiple-access channels, *IEEE Trans. Inform. Theory* 50 (1) (2004) 145–152.
- [263] W. Zhang, X.-G. Xia, K.B. Letaief, Space-time/frequency coding for MIMO-OFDM in next generation broadband wireless systems, *IEEE Wireless Commun.* 14 (3) (2007) 32–43.
- [264] Z. Zhang, W. Zhang, C. Tellambura, Cooperative OFDM channel estimation in the presence of frequency offsets, *IEEE Trans. Veh. Technol.* 58 (7) (2009) 3447–3459.
- [265] M. Zhao, Z. Shi, M.C. Reed, Iterative turbo channel estimation for OFDM system over rapid dispersive fading channel, *IEEE Trans. Wireless Commun.* 7 (8) (2008) 3174–3184.
- [266] T. Zemen, C.F. Mecklenbräuker, Time-variant channel estimation using discrete prolate spheroidal sequences, *IEEE Trans. Signal Process.* 53 (9) (2005) 3597–3607.
- [267] H. Zhu, G. Leus, G.B. Giannakis, Sparsity-cognizant total least-squares for perturbed compressive sampling, *IEEE Trans. Signal Process.* 59 (5) (2011) 2002–2016.

- [268] M.S. Zimmerman, A.L. Kirsch, The AN/GSC-10 (KATHRYN) variable rate data modem for HF radio, *IEEE Trans. Commun. Technol.* 15 (2) (1967) 197–204.
- [269] W.Y. Zou, Y. Wu, COFDM: an overview, *IEEE Trans. Broadcast.* 41 (1) (1995) 1–8.

Additional References

- [270] N. Al-Dhahir, J.M. Cioffi, Optimum finite-length equalization for multicarrier transceivers, *IEEE Trans. Commun.* 44 (1) (1996) 56–64.
- [271] J. Armstrong, Analysis of new and existing methods of reducing intercarrier interference due to carrier frequency offset in OFDM, *IEEE Trans. Commun.* 47 (3) (1999) 365–369.
- [272] G. Arslan, B.L. Evans, S. Kiaei, Equalization for discrete multitone transceivers to maximize bit rate, *IEEE Trans. Signal Process.* 49 (12) (2001) 3123–3135.
- [273] I. Barhumi, G. Leus, M. Moonen, Optimal training design for MIMO OFDM systems in mobile wireless channels, *IEEE Trans. Signal Process.* 51 (6) (2003) 1615–1624.
- [274] N. Benvenuto, R. Dinis, D. Falconer, S. Tomasin, Single carrier modulation with nonlinear frequency domain equalization: an idea whose time has come-again, *Proc. IEEE* 98 (1) (2010) 69–96.
- [275] H. Bölcseki, Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems, *IEEE Trans. Commun.* 49 (6) (2001) 988–999.
- [276] H. Bölcseki, D. Gesbert, A.J. Paulraj, On the capacity of OFDM-based spatial multiplexing systems, *IEEE Trans. Commun.* 50 (2) (2002) 225–234.
- [277] H.-H. Chen, J.-F. Yeh, N. Suehiro, A multicarrier CDMA architecture based on orthogonal complementary codes for new generations of wideband wireless communications, *IEEE Commun. Mag.* 39 (10) (2001) 126–135.
- [278] G. Cherubini, E. Eleftheriou, S. Ölcer, Filtered multitone modulation for very high-speed digital subscriber lines, *IEEE J. Sel. Areas Commun.* 20 (5) (2002) 1016–1028.
- [279] L.J. Cimini, N.R. Sollenberger, Peak-to-average power ratio reduction of an OFDM signal using partial transmit sequences, *IEEE Commun. Lett.* 4 (3) (2000) 86–88.
- [280] S. Coleri, M. Ergen, A. Puri, A. Bahai, Channel estimation techniques based on pilot arrangement in OFDM systems, *IEEE Trans. Broadcast.* 48 (3) (2002) 223–229.
- [281] O. Edfors, M. Sandell, J.-J. van de Beek, S.K. Wilson, P.O. Börjesson, OFDM channel estimation by singular value decomposition, *IEEE Trans. Commun.* 46 (7) (1998) 931–939.
- [282] M. Ghosh, Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems, *IEEE Trans. Commun.* 44 (2) (1996) 145–147.
- [283] B. Hirosaki, An orthogonally multiplexed QAM system using the discrete Fourier transform, *IEEE Trans. Commun.* 29 (7) (1981) 982–989.
- [284] D. Kivanc, G. Li, H. Liu, Computationally efficient bandwidth allocation and power control for OFDMA, *IEEE Trans. Wireless Commun.* 2 (6) (2003) 1150–1158.
- [285] Y. Li, Pilot-symbol-aided channel estimation for OFDM in wireless systems, *IEEE Trans. Veh. Technol.* 49 (4) (2000) 1207–1215.
- [286] Y. Li, Simplified channel estimation for OFDM systems with multiple transmit antennas, *IEEE Trans. Wireless Commun.* 1 (1) (2002) 67–75.
- [287] Y. Li, L.J. Cimini Jr, N.R. Sollenberger, Robust channel estimation for OFDM systems with rapid dispersive fading channels, *IEEE Trans. Commun.* 46 (7) (1998) 902–915.
- [288] Y. Li, N. Seshadri, S. Ariyavisitakul, Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels, *IEEE J. Sel. Areas Commun.* 17 (3) (1999) 461–471.

- [289] H. Liu, U. Tureli, A high-efficiency carrier estimator for OFDM communications, *IEEE Commun. Lett.* 2 (4) (1998) 104–106.
- [290] Z. Liu, Y. Xin, G.B. Giannakis, Space-time-frequency coded OFDM over frequency-selective fading channels, *IEEE Trans. Signal Process.* 50 (10) (2002) 2465–2476.
- [291] S. Mason, C. Berger, S. Zhou, P. Willett, Detection, synchronization, and Doppler scale estimation with multicarrier waveforms in underwater acoustic communication, *IEEE J. Sel. Areas Commun.* 26 (9) (2008) 1638–1649.
- [292] F. Meshkati, M. Chiang, H.V. Poor, S.C. Schwartz, A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems, *IEEE J. Sel. Areas Commun.* 24 (6) (2006) 1115–1129.
- [293] H. Minn, V.K. Bhargava, K.B. Letaief, A robust timing and frequency synchronization for OFDM systems, *IEEE Trans. Wireless Commun.* 2 (4) (2003) 822–839.
- [294] H. Minn, M. Zeng, V.K. Bhargava, On timing offset estimation for OFDM systems, *IEEE Commun. Lett.* 4 (7) (2000) 242–244.
- [295] M. Morelli, U. Mengali, An improved frequency offset estimator for OFDM applications, *IEEE Commun. Lett.* 3 (3) (1999) 75–77.
- [296] B. Natarajan, C.R. Nassar, S. Shattil, M. Michelini, Z. Wu, High-performance MC-CDMA via carrier interferometry codes, *IEEE Trans. Veh. Technol.* 50 (6) (2001) 1344–1353.
- [297] R. Negi, J. Cioffi, Pilot tone selection for channel estimation in a mobile OFDM system, *IEEE Trans. Consum. Electron.* 44 (3) (1998) 1122–1128.
- [298] S. Ohno, G.B. Giannakis, Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels, *IEEE Trans. Inform. Theory* 50 (9) (2004) 2138–2145.
- [299] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization, *IEEE Trans. Signal Process.* 51 (9) (2003) 2381–2401.
- [300] M. Poggioni, L. Rugini, P. Banelli, A novel simulation model for coded OFDM in Doppler scenarios, *IEEE Trans. Veh. Technol.* 57 (5) (2008) 2969–2980.
- [301] T. Pollet, M. Van Bladel, M. Moeneclaey, BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise, *IEEE Trans. Commun.* 43 (2/3/4) (1995) 191–193.
- [302] Y. Rong, X. Tang, Y. Hua, A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems, *IEEE Trans. Signal Process.* 57 (12) (2009) 4837–4851.
- [303] L. Rugini, P. Banelli, Joint impact of frequency synchronization errors and intermodulation distortion on the performance of multicarrier DS-CDMA systems, *EURASIP J. Appl. Signal Process.* 2005 (5) (2005) 730–742.
- [304] L. Rugini, P. Banelli, G.B. Giannakis, Local ML detection for multicarrier DS-CDMA downlink systems with grouped linear precoding, *IEEE Trans. Wireless Commun.* 5 (2) (2006) 306–311.
- [305] B.R. Saltzberg, Performance of an efficient parallel data transmission system, *IEEE Trans. Commun. Technol.* 15 (6) (1967) 805–811.
- [306] O. Simeone, Y. Bar-Ness, U. Spagnolini, Pilot-based channel estimation for OFDM systems by tracking the delay-subspace, *IEEE Trans. Wireless Commun.* 3 (1) (2004) 315–325.
- [307] G.L. Stüber, J.R. Barry, S.W. McLaughlin, Y. Li, M.A. Ingram, T.G. Pratt, Broadband MIMO-OFDM wireless communications, *Proc. IEEE* 92 (2) (2004) 271–294.
- [308] V. Tarokh, H. Jafarkhani, On the computation and reduction of the peak-to-average power ratio in multicarrier communications, *IEEE Trans. Commun.* 48 (1) (2000) 37–44.
- [309] L. Tomba, On the effect of Wiener phase noise in OFDM systems, *IEEE Trans. Commun.* 46 (5) (1998) 580–583.
- [310] U. Tureli, H. Liu, M.D. Zoltowski, OFDM blind carrier offset estimation: ESPRIT, *IEEE Trans. Commun.* 48 (9) (2000) 1459–1461.

- [311] L.-L. Yang, L. Hanzo, Multicarrier DS-CDMA: a multiple access scheme for ubiquitous broadband wireless communications, *IEEE Commun. Mag.* 41 (10) (2003) 116–124.
- [312] W. Yu, R. Lui, Dual methods for nonconvex spectrum optimization of multicarrier systems, *IEEE Trans. Commun.* 54 (7) (2006) 1310–1322.
- [313] Y. Zhao, S.-G. Häggman, Intercarrier interference self-cancellation scheme for OFDM mobile communication systems, *IEEE Trans. Commun.* 49 (7) (2001) 1185–1191.
- [314] S. Zhou, G.B. Giannakis, Finite-alphabet based channel estimation for OFDM and related multicarrier systems, *IEEE Trans. Commun.* 49 (8) (2001) 1402–1414.

Signal Processing for Vectored Multichannel VDSL

6

Itsik Bergel and Amir Leshem

Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel

2.06.1 Introduction

Digital subscriber line (DSL) is currently the most widespread technique for delivering broadband internet to homes worldwide. By the end of 2010, over 300 million customers were served with DSL, a number representing approximately 60% of global broadband customers. Furthermore, the telephone network infrastructure is still capable of providing the ever increasing demand for larger bandwidth. Data services to home customers began with the use of dialup modems over the public switched telephone network (PSTN). The rates supported by these modems increased from 300 bps in the first commercial modem introduced in 1962 to 56 kbps in the v.92 standardized in 1998. To provide these increasing data rates, more advanced communication and signal processing techniques have been developed. From FSK in the 300 bps mode to a system using advanced precoding, shaping and trellis coding in the v.34 and the newer modems. However, it was clear from the outset that signaling over the 4 kHz voiceband is an unnecessary limitation. In 1988 the UN telecommunication organization CCITT (which was later renamed ITU) defined the Integrated Services Data Networks (ISDN), which support symmetric rates of up to 128 kbps using 256 kHz. At the same time as the development of the ISDN another standard known as the high speed digital subscriber line (HDSL) was developed with the goal of providing operators with a replacement for the expensive T1 (in the US) and E1 (in Europe) lines used to carry 24 or 32 voice channels. The idea was to use improved communication techniques so that the data could be transmitted over regular telephone lines (commonly termed “twisted pairs”) instead of the special lines used previously. HDSL supported rates of up to 2.048 Mbps over two twisted pairs using a 2B1Q baseband modulation (similar to the ISDN modulation) and channel compensation techniques such as a decision feedback equalizer (DFE) and reduced state Viterbi. However transmission was un-coded. A good overview of the HDSL technology is given in [1]. Since transmission over two twisted pairs is costly, a new standard HDSL2 was developed in the US and later extended to the Single Pair High bit rate DSL (SHDSL) by the ITU-T. Concomitantly, an Asymmetric DSL (ADSL) was defined [2] mostly for residential broadband services. ADSL was the first system to use multicarrier modulation, dubbed Discrete Multi-Tone (DMT). In contrast to OFDM, the DMT modulation uses a different bit loading on each carrier and real baseband signaling. Riding on the success of ADSL, the VDSL standard was defined [3]. The major difference between VDSL and ADSL is the larger bandwidth used by the former. This allows for much higher upstream rates, but requires shorter lines. Therefore, the use of the VDSL was delayed until optical network units were deployed in street cabinets and basements.

VDSL provided data rates of up to 50 Mbps in the downstream and up to 25 Mbps in the downstream in typical loops of 300 m and above. A good overview of single line DSL technology can be found in books by Starr et al. [4] and Bingham [5].

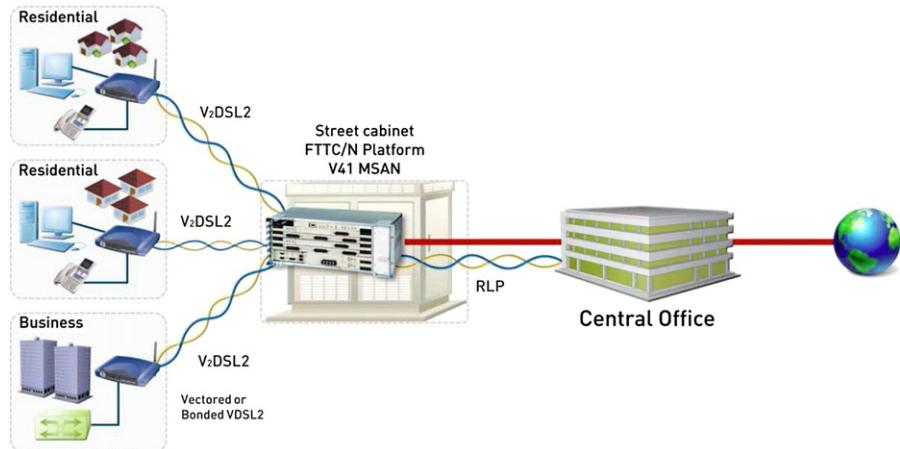
To understand the ways data rates can be increased beyond VDSL rates, and achieve 100 Mbps symmetric rates over typical twisted pairs, the channel characteristics of the twisted pair channel as well as the crosstalk coupling between twisted pairs in a typical binder of twisted pairs need further explanation. The characterization of the twisted pair channel and its capacity dates back many years. A good example is the paper by Foschini [6] that described the binder using a multiport model. Later, the statistical characteristics of crosstalk at low frequencies using log normal distributions was proposed by Adie and Gibbs [7] and by Lin [8,9]. The models for the average and the 1% worst case were proposed by Werner [10] and refined by Kerpez [11,12]. Kalet and Shamai [13] and Aslanis and Cioffi [14] studied the capacity of single channel DSL systems using the above models. Later Karipidis et al. [15,16] verified these models for frequencies up to 30 MHz, and studied the impact of crosstalk on the capacity of multichannel DSL systems.

These studies make it clear that crosstalk is the limiting factor for increasing the rate of DSL systems. To increase the data rates provided by single line VDSL, the crosstalk between twisted pairs has to be eliminated. This prompted the need for multichannel processing, in which the signals of several modems are jointly processed to overcome the crosstalk between the different twisted pairs. Such systems are often termed “vectored” systems, as the signals of the relevant modems can be ordered in a vector form that enables the use of matrix manipulations and linear algebra tools.

Unlike standard MIMO techniques, crosstalk cancellation in DSL systems can only be done at the network side, where all twisted pairs originate. Lechleider was one of the first to realize that crosstalk cancellation might be an advantage in the context of two pair HDSL systems [17]. Ginis and Cioffi [18] proposed a single sided cancellation in the downstream direction using a multidimensional Tomlinson-Harashima precoder. It was quickly observed that the loss in linear precoding is quite insubstantial, due to the diagonal dominance of the DSL channel [19,20]. In these papers they showed that both upstream and downstream crosstalk cancellation can be done using a linear zero forcing multichannel transmitter (upstream) or receiver (downstream). At the same time, Leshem and Li [21–23] demonstrated that the linear precoder can be simplified even further using an approximated power series expansion. This completely reduced the complexity of the required matrix inversion. Cendrillon et al. [24,25] analyzed the case of partial precoding where only some modems are coordinated.

An alternative adaptive precoder design was proposed by Louvaux and van der Veen [26]. Another simplified version of the adaptive precoder was suggested independently by Louvaux and van der Veen [27] and by Bergel and Leshem [28]. Bergel and Leshem provided a detailed analytic analysis. This analysis showed that there are advantages in not precoding the signals of modems that do not gain much from this precoding, since this can cause significant delay in the convergence of the adaptive precoder. Binyamini et al. [29] generalized this analysis to the case of partial precoding, where similarly to [24] only some of the modems’ signals are precoded.

Along with the advances in precoding techniques, initial prototypes were built and an ITU standard G.993.5 was developed that defined the requirements for vectored VDSL systems. Recently, commercial systems have been developed. These systems use the vectored capabilities to achieve reliable rates at much larger distances, or to provide much higher data rates for the same line lengths. One of the early prototypes was developed by an Israeli consortium called iSMART, which developed

**FIGURE 6.1**

Typical V41 ONU deployment.

techniques for vectoring. The outcome of this research was the V41 vectoring system developed by ECI Telecom,¹ which is depicted in Figure 6.1. A typical vectored VDSL system consists of an ONU, which is typically deployed in a street cabinet or in the basement of high-rise buildings and multiple customer premises equipment units (CPE). The ONU is connected to the operator's network through a high bandwidth optical link, and contains a multiuser vectored VDSL2 transceiver (V₂VDSL2) [30], which jointly receives the information from all CPEs and precodes the information transmitted to all CPEs such that all crosstalk between the lines is removed.

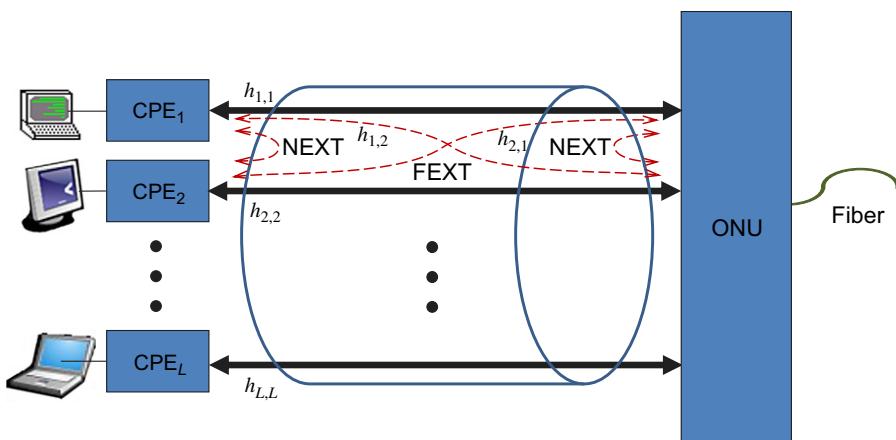
The V41 system can increase the coverage of 50 Mbps service from 400 m to 800 m over 0.4 mm copper twisted pairs. This results in a factor 4 reduction in the ONUs required to provide a universal 50 Mbps service. Similarly, rates of 85 Mbps can be provided at a distance of 400 m.

The purpose of this chapter is to present the mathematical principles behind the current-day massive multi-user systems, serving simultaneously 50–200 customers and canceling interference between all these customers. It is remarkable that similar systems in the wireless domain serve 4–8 customers simultaneously using the same spectrum. The main reason for the success of massive vectoring in the DSL world is that the loss due to zero forcing precoding is marginal because of a channel property called diagonal dominance and the ease of tracking the channels, which are relatively stationary compared to wireless channels.

2.06.2 System model

A typical DSL channel is composed of many users, each connected through a twisted copper wire pair to an Optical Network Unit (ONU), as depicted in Figure 6.2. Many of these twisted pairs are enclosed

¹We thank Ido Shargil from ECI Telecom for providing the information on the v41 platform.

**FIGURE 6.2**

The DSL channel structure.

(at least part of the way) in the same binder. In this work we focus on a group of users that are served through the same binder.

Each user's modem (termed also customer premises equipment, CPE) aims to transmit data to the network and receive data from the network over its twisted pair. The connection to the network is established through an optical network unit (ONU) which is connected to the network backbone through an optical fiber. The ONU can be located at the telephone switch (also termed central office). However, in recent years, in order to enable higher data rates, the ONU is typically located closer to the users, in street cabinets or even in the users' building. A shorter distance between the CPE and ONU will result in lower signal attenuation, which can lead to a higher signal to noise ratio (SNR) and higher data rates in the absence of interference. But, as the typical SNR increases, the main limiting factor of DSL systems is inter-user interference.

This inter-user interference results from electromagnetic coupling in the binder, and is divided into two types as shown in Figure 6.2. The crosstalk from signals that originate from the same end as the affected receiver is termed near end crosstalk (NEXT). In VDSL, the impact of NEXT is suppressed by employing frequency-division duplexing (FDD) and transmitters synchronization. The crosstalk from signals that originate from the opposite side of the binder is termed far end crosstalk (FEXT). FEXT is typically the main limiting factor on the performance of DSL systems.

The downstream and upstream of the DSL channel share the same channel, but are rather different in nature. The difference comes from the fact that multichannel processing can be applied only at the ONU. The upstream describes the transmission of data from the CPEs and reception in the ONU. This is a multiple access channel, in which the ONU decodes the data transmitted from multiple CPEs. The downstream describes the transmission of data from the ONU and reception in the CPEs. This is a broadcast channel, in which the ONU transmits data to multiple CPEs while minimizing the interference between the transmitted signals.

Since the upstream and downstream share the same channel, in this section we describe this shared channel. We therefore refer to transmitters and receivers, which will later be translated into CPEs and ONU according to the context. Note that regardless of the actual ONU internal structure, in the following we refer to the ONU as a set of cooperating transmitters/receivers, where each transmitter/receiver is connected to a single twisted pair.

The channel received at the i th receiver is:

$$y_i[n] = \sum_{j=0}^{L-1} \sum_{m=n-\nu+1}^n h_{i,j}^t(m) x_j[n-m] + w_i[n], \quad (6.1)$$

where L is the number of transmitters, $x_j[n]$ is the sample transmitted by transmitter j at time n , and $w_i[n]$ is the sampled noise. In this model, the sampled noise term can include in addition to thermal noise also interference from various sources such as radio transmissions, neighboring DSL systems and more. Nevertheless, it is generally assumed that the noise samples are independent and identically distributed (i.i.d) Gaussian random variables with zero mean and variance of $E[|w_i[n]|^2] = \sigma^2$.

The impulse response of the channel between the j th transmitter and the i th receiver is denoted by $h_{i,j}^t(m)$ and ν is the maximal length of the channel impulse response (also termed the channel memory).² Note that each receiver is affected by all the transmitters in the system. Thus, $h_{i,i}^t(m)$ describes the channels for the direct signals, and $h_{i,j}^t(m)$ for $i \neq j$ describes the crosstalk channels. Typically, the actual lengths of the various channel impulse responses are different. In the following we take ν to denote the maximal channel memory handled by the system, and assume that all channel impulse responses are zero padded to a length of ν .

Without loss of generality, in the following we focus on the transmission and reception of symbol number 0, i.e., the block of N input samples, $\mathbf{x}_j^t = [x_j[0], \dots, x_j[N-1]]^T$. DSL systems achieve reception that is free of inter symbol interference (ISI) using the discrete Fourier transform (DFT) and a cyclic prefix (CP).

The DFT processing of multiple user signals requires the synchronized reception of all transmissions. Such a synchronization is easily achieved at the downstream, as all transmission originate in the same location. In the upstream, the synchronized reception of all signals transmitted from different CPEs needs more attention. This synchronization is achieved by shifting the transmission time of each CPE according to the delay between the specific CPE and the ONU (using timing feedback from the ONU). This procedure is analogous to the problem of synchronized uplink transmissions in cellular networks. The synchronization timing requirements can be further relaxed using the “Zipper” technique (see [31, 32]).

The $\nu - 1$ samples prior to the transmission of each block are devoted to the CP using:

$$x_j[-n] = x_j[N-n] \quad (6.2)$$

²We use the notation \bullet^t to indicate time domain quantities. In the following we will also use the notations \bullet^f and \bullet^F to indicate frequency domain quantities.

for $n = 1, 2, \dots, v - 1$. Using (6.2) in (6.1), the vector of output symbols $\mathbf{y}_j^t = [y_j[0], \dots, y_j[N - 1]]^T$ can be described by:

$$\mathbf{y}_j^t = \sum_{j=0}^{L-1} \mathbf{H}_{i,j}^t \mathbf{x}_j^t + \mathbf{w}_i^t, \quad (6.3)$$

where $\mathbf{w}_i^t = [w_i[0], \dots, w_i[N - 1]]^T$ is the vector of noise samples, and the $\mathbf{H}_{i,j}^t$ is a circulant channel matrix in which the (a, b) element is:

$$\mathbf{H}_{i,j}^t(a, b) = \begin{cases} h_{i,j}^t(a - b \bmod N) & 0 \leq a - b \bmod N \leq v - 1, \\ 0 & \text{otherwise.} \end{cases} \quad (6.4)$$

More graphically, the circulant channel matrix has the form:

$$\mathbf{H}_{i,j}^t = \begin{bmatrix} h_{i,j}^t(0) & 0 & \cdots & 0 & h_{i,j}^t(v-1) & \cdots & h_{i,j}^t(1) \\ h_{i,j}^t(1) & h_{i,j}^t(0) & \ddots & \vdots & 0 & \ddots & \vdots \\ \vdots & h_{i,j}^t(1) & \ddots & 0 & \ddots & 0 & h_{i,j}^t(v-1) \\ h_{i,j}^t(v-1) & \vdots & \ddots & h_{i,j}^t(0) & 0 & \ddots & 0 \\ 0 & h_{i,j}^t(v-1) & \ddots & h_{i,j}^t(1) & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & h_{i,j}^t(0) & 0 \\ 0 & \cdots & 0 & h_{i,j}^t(v-1) & \cdots & h_{i,j}^t(1) & h_{i,j}^t(0) \end{bmatrix}. \quad (6.5)$$

In DSL the data processing is performed in the frequency domain using the discrete Fourier transform (DFT). The normalized DFT matrix is given by:

$$\mathbf{Q}_{\text{DFT}}(a, b) = \frac{1}{\sqrt{N}} e^{-j \frac{2\pi}{N} ab} \quad (6.6)$$

and satisfies $\mathbf{Q}_{\text{DFT}}^{-1} = \mathbf{Q}_{\text{DFT}}^H$. The transmitted samples of the j th user are generated using an inverse DFT (IDFT) operation: $\mathbf{x}_j^t = \mathbf{Q}_{\text{DFT}}^H \mathbf{x}_j^F$. Then, taking the DFT of the output vector, (6.3), results in the frequency domain channel output:

$$\mathbf{y}_i^F = \mathbf{Q}_{\text{DFT}} \mathbf{y}_i^t = \sum_{j=0}^{L-1} \mathbf{H}_{i,j}^F \mathbf{x}_j^F + \mathbf{w}_i^F, \quad (6.7)$$

where

$$\mathbf{H}_{i,j}^F = \mathbf{Q}_{\text{DFT}} \mathbf{H}_{i,j}^t \mathbf{Q}_{\text{DFT}}^H \quad (6.8)$$

is the frequency domain channel matrix between transmitter j and receiver i , and $\mathbf{w}_i^F = \mathbf{Q}_{\text{DFT}} \mathbf{w}_i^t$ is the frequency domain Gaussian noise with zero mean and covariance matrix of $E[\mathbf{w}_i^F (\mathbf{w}_i^F)^H] = \sigma^2 \cdot I$.

Recalling that the time domain channel matrix, (6.4), is a circulant matrix, the frequency domain channel matrix $\mathbf{H}_{i,j}^F$ is a diagonal matrix,³ and hence the received signal matrix is free of ISI.

Although the DSL signal is not degraded by ISI, inspecting the sum in (6.7) one can see that it is highly affected by another type of interference, namely, inter user interference that results from FEXT. To better characterize and resolve the FEXT, it is convenient to restate the input-output relation, grouping all users for each frequency bin separately. Defining $\mathbf{y}_f = [y_0^F(f), \dots, y_{L-1}^F(f)]^T$, the output sample at the f th frequency bin for all users can be described by:

$$\mathbf{y}_f = \mathbf{H}_f \mathbf{x}_f + \mathbf{w}_f, \quad (6.9)$$

where $\mathbf{x}_f = [x_0^F(f), \dots, x_{L-1}^F(f)]^T$ is the vector of samples transmitted by all users, $\mathbf{w}_f = [w_0^F(f), \dots, w_{L-1}^F(f)]^T$ is the vector of noise measured by all users and the combined channel matrix is given by:

$$H_f(i, j) = H_{i,j}^F(f, f). \quad (6.10)$$

In most cases, the processing and analysis of DSL signals is performed for a single frequency bin at a time. Thus, the relation described by Eq. (6.9) is very useful. In many cases we will even drop the frequency bin indices (f) where the specific frequency bin number is not crucial.

The power constraint is also stated in the frequency domain in the form of a spectrum power mask. In the following we assume that all users follow the same spectrum mask, and the power constraint is given by:

$$E[|x_f(j)|^2] \leq p_f. \quad (6.11)$$

2.06.2.1 Channel matrix structure

As the FEXT results from electromagnetic coupling and not from a direct wire connection, in most frequencies the FEXT is significantly smaller from the direct signal. Thus, it has been noted [19] that upstream VDSL channel matrices are column-wise diagonally dominant (CWDD), while downstream matrices are row-wise diagonally dominant (RWDD). This diagonal dominance channel matrix structure is one of the main features that distinguish DSL systems from other MIMO systems. Hence, most DSL specific research has used this feature in some way or another.

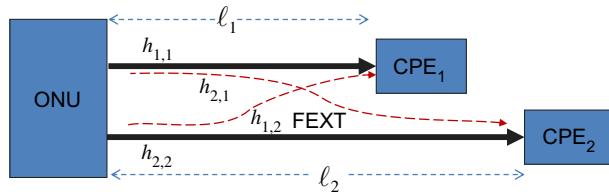
For simplicity we will first describe the RWDD structure of the downstream channel and then comment on the upstream channel. Basically, in a RWDD matrix, the diagonal element dominates all other elements in the row. The weakest definition states that a matrix is weakly RWDD if

$$|H(i, i)| > |H(i, j)| \quad (6.12)$$

for $j \neq i$. One can also quantify how much is the matrix RWDD using:

$$\beta_{D,1} = \max_{i,j} \frac{|H(i, j)|}{|H(i, i)|}. \quad (6.13)$$

³Recalling that the eigenvalue decomposition of a circulant matrix contains the DFT and IDFT matrices, and a diagonal frequency domain channel matrix.

**FIGURE 6.3**

Mixed length downstream DSL.

A matrix is weakly RWDD if $\beta_{D,1} < 1$, and as $\beta_{D,1}$ decreases the matrix is said to be “more” RWDD.

An alternative measure compares the diagonal element to the sum of the absolute values of the elements in the rest of the row:

$$\beta_{\text{row}} = \max_i \frac{\sum_{j \neq i} |H(i, j)|}{|H(i, i)|}. \quad (6.14)$$

A matrix that satisfies $\beta_{\text{row}} < 1$ is said to be strongly RWDD.

If all pairs in the binder have the same length, the magnitude of the diagonal elements in the matrix is nearly identical, and the matrix can be termed simply diagonally dominant. The differentiation between RWDD and CWDD is important mostly in mixed length scenarios. Figure 6.3 depicts a 2-user mixed length downstream scenario with a short link of length ℓ_1 and a long link of length ℓ_2 . Note that the FEXT is only generated in the joint part (of length ℓ_1). Thus, the FEXT into the short link is equivalent to the FEXT in the case of equal link binder, and it is significantly smaller than the direct link signal. For the longer link, the FEXT contribution can be split into a cascade of two systems: in the first system there is a FEXT contribution into the ℓ_1 initial segment of the link, and in the second system this FEXT is further attenuated by a link of length $\ell_2 - \ell_1$. Hence its mean response is [10]:

$$K_{\text{FEXT}} f^2 \ell_1 IL(\ell_1, f) IL(\ell_2 - \ell_1, f), \quad (6.15)$$

where $IL(\ell, f)$ is the typical loop insertion loss at length ℓ and frequency f and K_{FEXT} is a constant that depends on the type of cable. Note that from the equal length case we already concluded that $K_{\text{FEXT}} f^2 \ell_1 \ll IL(\ell_1, f)$. Recalling that

$$IL(\ell_1, f) IL(\ell_2 - \ell_1, f) = IL(\ell_2, f), \quad (6.16)$$

the cascade of the two LTI systems shows that

$$K_{\text{FEXT}} f^2 \ell_1 IL(\ell_2, f) \ll IL(\ell_2, f). \quad (6.17)$$

From (6.17) we can conclude that an even more accurate representation of the channel matrix structure is given by [22]:

$$\mathbf{H} = \mathbf{D}(I + \mathbf{E}), \quad (6.18)$$

where the matrix \mathbf{E} has zeros on its diagonal and all its elements are (significantly) smaller than one. This type of definition turns out to be important for the convergence analysis of adaptive precoders.

The corresponding measure is:

$$\beta_{D,\max} = \max \left\{ \max_i \frac{\sum_{j \neq i} |H(i, j)|}{|H(i, i)|}, \max_j \frac{\sum_{i \neq j} |H(i, j)|}{|H(i, i)|} \right\}. \quad (6.19)$$

Note that although the sum in the first maximization in (6.19) is taken over rows and the sum in the second maximization is taken over columns, both are measures of RWDD. This is because in both cases each matrix element, $H(i, j)$, is divided by the diagonal element in its own row, $H(i, i)$. Thus the $\beta_{D,\max}$ serves as another indication that no element in the matrix is significant compared to the diagonal element in its row.

As for the upstream, one only needs to observe that due to the channel reciprocity, the upstream channel matrix at each frequency is the transpose of the upstream channel matrix at the same frequency. DSL systems work in FDD mode and hence no frequency is used simultaneously for both upstream and downstream. However, for channel characterization purposes the channel reciprocity holds, and we can simply apply the above definitions ((6.13), (6.14) and (6.19)) to the transposed channel matrix. Thus, we define:

$$\beta_{U,1} = \max_{i,j} \frac{|H(j, i)|}{|H(i, i)|}, \quad (6.20)$$

$$\beta_{\text{col}} = \max_i \frac{\sum_{j \neq i} |H(j, i)|}{|H(i, i)|} \quad (6.21)$$

and

$$\beta_{U,\max} = \max \left\{ \max_i \frac{\sum_{j \neq i} |H(j, i)|}{|H(i, i)|}, \max_j \frac{\sum_{i \neq j} |H(j, i)|}{|H(i, i)|} \right\}. \quad (6.22)$$

A matrix is weakly CWDD if $\beta_{D,1} < 1$, and strongly CWDD if $\beta_{\text{row}} < 1$. The channel decomposition of (6.18) is replaced from the upstream with: [22]:

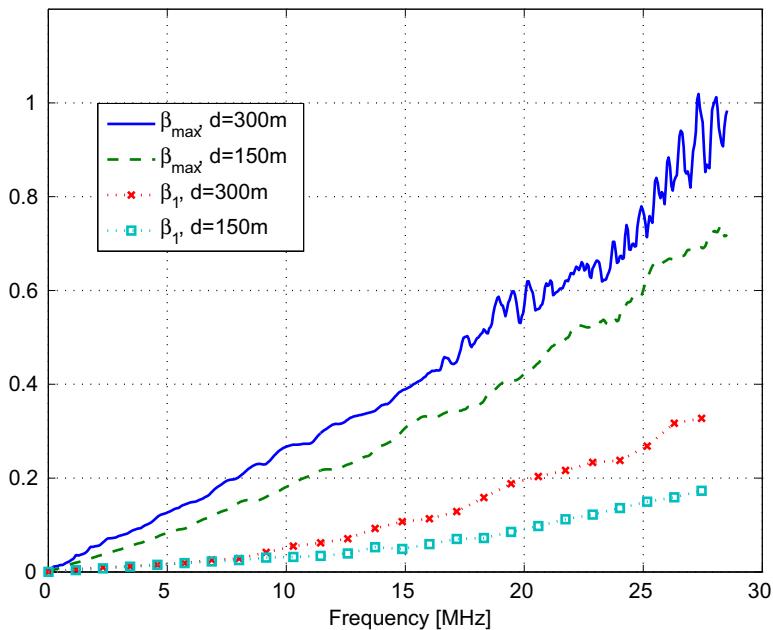
$$\mathbf{H} = (I + \mathbf{G})\mathbf{D}, \quad (6.23)$$

where the matrix \mathbf{G} has the same characteristics as the matrix \mathbf{E} , i.e., it has a zero on its diagonal and all its elements are (significantly) smaller than one.

To demonstrate the diagonal dominance of the channel matrices, Figure 6.4 depicts the values of $\beta_{D,1}$ and $\beta_{D,\max}$ as a function of frequency for binder lengths of 150–300 m. It is clear that, at both lengths, the channel matrices are strongly diagonal dominant for practically all frequencies. In this plot we do not show β_{row} , as it was nearly identical to $\beta_{D,\max}$ for all frequencies at both binder links.

Note on figures: All the numerical results in this paper rely on channel measurements carried out by France Telecom at different lengths over a binder with 28 twisted pairs.⁴ All simulations presented here were carried out under the assumption that the whole bandwidth of the VDSL system is completely dedicated to the downstream or completely dedicated to the upstream. The transmission SNR is set to $\eta = p/\sigma^2 = 80$ dB (corresponding to the typical values of $p = -60 + 10 \log_{10} B$ dBm and $\sigma^2 = -14010 \log_{10} B$ dBm, where $1/B$ is the symbol length).

⁴The authors thank M. Ouzzif, R. Tarafi, H. Marriott, and F. Gauthier of France Telecom R&D, who conducted the VDSL channel measurements under the auspices of the U-BROAD project.

**FIGURE 6.4**

Row-wise diagonal dominance measures: $\beta_{D,1}$ and $\beta_{D,\max}$ as a function of frequency for two binder lengths.

2.06.3 Downstream transmission

2.06.3.1 Nonlinear precoding

In the downstream (or in its more general term: the broadcast channel) the receiving modems are located in different customer premises and cannot cooperate to achieve FEXT cancellation. On the other hand, the transmitting modems are co-located at the ONU, and hence FEXT cancellation is feasible using joint signal processing of the transmitted symbols.

Ginis and Cioffi [18] considered the use of a QR decomposition followed by a Tomlinson-Harashima like modulo operation [33,34]. Considering the LQ decomposition of the channel matrix we can write:

$$\mathbf{H} = \mathbf{L}\mathbf{Q}^H, \quad (6.24)$$

where \mathbf{Q} is a unitary matrix and \mathbf{L} is a lower triangular matrix. An even more convenient representation is the QR decomposition of the conjugate transpose of the channel matrix, given by $\mathbf{H}^H = \mathbf{Q}\mathbf{R}$ where $\mathbf{R} = \mathbf{L}^H$ is an upper triangular matrix.

Let \mathbf{u} denote the modulated symbols that need to be transmitted in the f th frequency bin by all users. In the following we consistently assume that the modulated symbols of all users in the same subcarrier

are iid random variables (which is reasonable given our assumption that all users comply with the same spectral mask). We further assume that $E[u(i)] = 0$ and $E[|u(i)|^2] = p$.

The modulated symbols are first precoded to reduce interference due to the matrix \mathbf{R} :

$$\begin{aligned} x'(0) &= u(0), \\ x'(1) &= \Gamma c_1 \left\{ u(1) - \frac{R^*(0, 1)}{R^*(1, 1)} x'(0) \right\}, \\ &\vdots \\ x'(L-1) &= \Gamma c_{L-1} \left\{ u(L-1) - \sum_{i=0}^{L-2} \frac{R^*(i, L-1)}{R^*(L-1, L-1)} x'(i) \right\}, \end{aligned} \quad (6.25)$$

where the modulo operation is defined by:

$$\Gamma_{c_i}\{u\} = \text{Re}(c_i) \overline{\text{mod}} \left(\frac{\text{Re}(u)}{\text{Re}(c_i)} \right) + j \cdot \text{Im}(c_i) \overline{\text{mod}} \left(\frac{\text{Im}(u)}{\text{Im}(c_i)} \right) \quad (6.26)$$

the index c_i gives the constellation size for user i , i.e., $\text{Re}(c_i)$ is the number of possible modulation values for the real part of $u(i)$ multiplied by the distance between modulation points, $\text{Im}(c_i)$ is the number of possible modulation values for the imaginary part of $u(i)$ multiplied by the distance between modulation points. The operation $\overline{\text{mod}}()$ is a modulo type operation that returns values in the range $[-0.5, 0.5]$, i.e., $\overline{\text{mod}}(u) = (u + 0.5 \bmod 1) - 0.5$.

Figure 6.5 illustrates this precoding with a modulo operation for the case of QPSK modulation. The possible modulation values are depicted by circles. The allowed transmission range is depicted by the middle square, and its modulo equivalents are depicted by dashed squares. The transmitter wishes to transmit the top right modulation value ($u(j)$). The solid arrow represents the subtraction of interference from previous users, and hence, the square mark represents the value to be transmitted in order to achieve full interference cancellation without the modulo operation. The dashed arrow represents the modulo operation, and the final resulting precoded value, $x'(j)$, is represented by the x-mark.

In order to achieve interference-free reception, the symbols, \mathbf{x}' , still need to be rotated using the matrix \mathbf{Q} . This rotation is given by:

$$\mathbf{x} = \mathbf{Q}\mathbf{x}'. \quad (6.27)$$

It is easy to verify that without the modulo operation, this scheme already achieves complete FEXT cancellation. The role of the modulo operation is to reduce the power increase due to the FEXT cancellation (just as in the classic Tomlinson-Harashima setting [33, 34]).

To reciprocate the modulo operation, each receiver performs the modulo operation:

$$\hat{u}(i) = \Gamma_{c_i} \left\{ \frac{y(i)}{R^*(i, i)} \right\}. \quad (6.28)$$

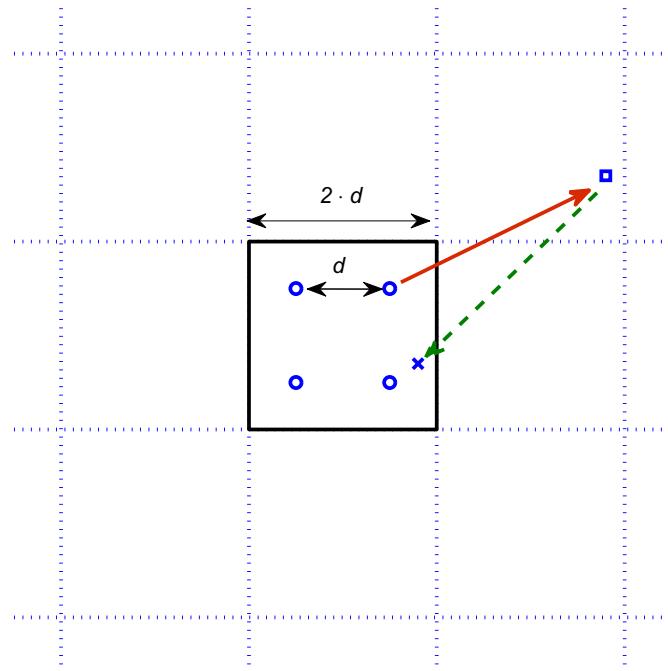
**FIGURE 6.5**

Illustration of the TH modulo operation.

To verify the FEXT cancellation, we substitute (6.9), (6.24), (6.25), and (6.27) in (6.28) and get:

$$\begin{aligned}
 \hat{u}(i) &= \Gamma_{c_i} \left\{ \frac{\sum_{j=0}^i R^*(j, i)x'(j) + w(i)}{R^*(i, i)} \right\} \\
 &= \Gamma_{c_i} \left\{ x'(i) + \sum_{j=0}^{i-1} \frac{R^*(j, i)}{R^*(i, i)} x'(j) + \frac{w(i)}{R^*(i, i)} \right\} \\
 &= \Gamma_{c_i} \left\{ \Gamma_{c_i} \left\{ u(i) - \sum_{j=0}^{i-1} \frac{R^*(j, i)}{R^*(i, i)} x'(j) \right\} + \sum_{j=0}^{i-1} \frac{R^*(j, i)}{R^*(i, i)} x'(j) + \frac{w(i)}{R^*(i, i)} \right\} \\
 &= \Gamma_{c_i} \left\{ u(i) + \frac{w(i)}{R^*(i, i)} \right\}, \tag{6.29}
 \end{aligned}$$

where the last line results from $\Gamma_{c_i}\{\Gamma_{c_i}\{a\} + b\} = \Gamma_{c_i}\{a + b\}$. It is further shown in [18] that the noise increase as well as the power increase are negligible in DSL systems, hence resulting in a near optimal FEXT free communication scheme.

The above method can be interpreted as a suboptimal implementation of the zero-forcing “dirty-paper” precoding scheme proposed in [35] and studied in [36]. An improvement of the scheme was proposed in [37], where Tomlinson-Harashima precoding was replaced by more efficient trellis precoding schemes. The broadcast channel has been further studied in many works, mostly in the context of wireless MIMO systems. However, for DSL systems, most of the focus has shifted to linear precoders that combine very good performance and lower implementation complexity.

2.06.3.2 Linear precoding

2.06.3.2.1 The zero forcing precoder

Linear precoders have often been considered for low complexity interference cancellation over broadcast channels. A linear precoder transmits a signal which is a linear combination of the modulated data symbols. Denoting the precoding matrix by \mathbf{F} , the transmitted vector for all users is given by:

$$\mathbf{x} = \mathbf{Fu}. \quad (6.30)$$

The linear precoder must satisfy the system power constraint. Assuming that all users have the same power constraint in each frequency (same spectral mask) and that the modulated data symbols \mathbf{u} satisfies the power constraint, the constraint on the precoder is given by:

$$\sum_{i=0}^{L-1} |F(j, i)|^2 \leq 1 \quad \forall j. \quad (6.31)$$

Linear precoders attracted considerable attention in DSL systems mostly following the work of Cendrillon et al. [38]. In that work, the authors showed that the zero forcing (ZF) precoder achieves near optimal performance in most DSL channels. Cendrillon et al. adopted the ZF linear precoder given by:

$$\mathbf{F} = \gamma \mathbf{H}^{-1} \mathbf{D}, \quad (6.32)$$

where \mathbf{D} is the diagonal matrix with the same diagonal elements as \mathbf{H} and γ is a constant chosen to satisfy the power constraint (6.31). (In that work they termed this precoder the *diagonalizing precoder*, to distinguish it from the more trivial ZF precoder: $\mathbf{F} = \mathbf{H}^{-1}$.) Obviously, substituting (6.32) and (6.30) in (6.9) results in FEXT free reception:

$$\mathbf{y} = \gamma \mathbf{Du} \quad (6.33)$$

and the achievable data rate by user i in any specific subcarrier is given by:

$$R = \Delta_F \cdot \log_2 \left(1 + \Gamma^{-1} \gamma^2 |H(i, i)|^2 \eta \right), \quad (6.34)$$

where Γ is the gap between the used precoder performance and the actual capacity (commonly termed the Shannon gap [39]), and Δ_F is the symbol rate.

The main question is how large can γ be. In their work Cendrillon et al. [38] showed that γ can typically be set very close to 1, and hence the ZF precoder achieves near optimal performance. The key to this near optimality is the structure of the DSL channel matrix, and in particular the row-wise diagonally dominant (RWDD) channel matrix. To prove this, they presented an upper bound on the

capacity and a lower bound on the achievable data rates using the ZF linear precoder. Here we present the simpler (but generally tighter) bounds of Bergel and Leshem [40], which rely on the strong diagonal dominance measures. The upper bound is described by the following theorem.

Theorem 1. *The data rate achievable by user i in any specific subcarrier of the downstream of DSL systems is upper bounded by:*

$$C \leq \Delta_F \cdot \log_2 \left(1 + \Gamma^{-1} \eta |H(i, i)|^2 (1 + \beta_{\text{row}})^2 \right). \quad (6.35)$$

Proof. Considering only the i th receiver. Taking into account the channel structure, (6.9), and the power constraint on the precoder, (6.31), the maximal signal power for this user is achieved using a precoder in which only the i th column is non-zero, and the absolute value of all elements in this column is 1. Using such a precoder, the maximal signal power for the i th user is: $(\sum_{j=0}^{L-1} |H(i, j)|)^2 \cdot p$. Calculating the achievable user rate, and maximizing over all users results in (6.35), and completes the proof of the theorem.

On the other side we have:

Theorem 2 [40]. *The data rate achievable by user i in any specific subcarrier when the ONU applies a ZF precoder is lower bounded by:*

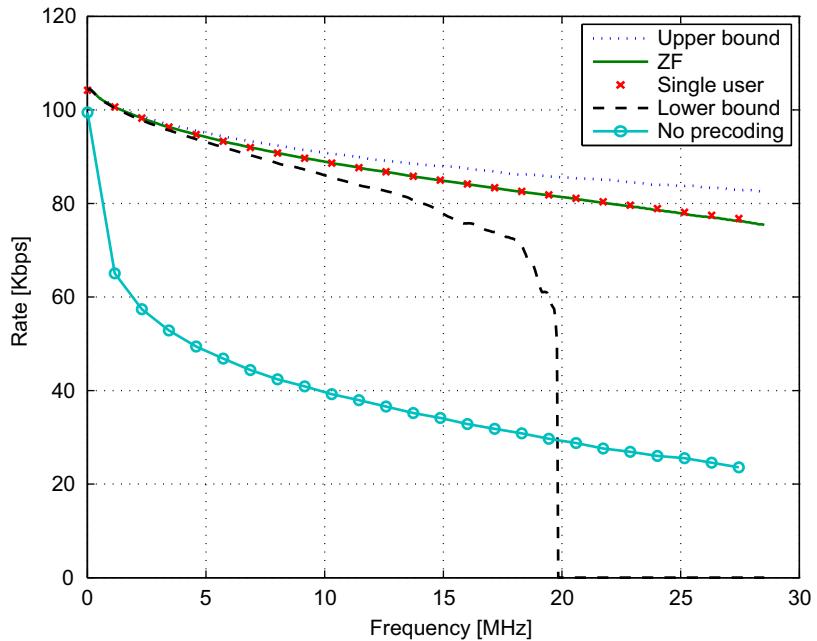
$$R \geq \Delta_F \cdot \log_2 \left(1 + \Gamma^{-1} f(\beta_{D,\max}) |H(i, i)|^2 \eta \right), \quad (6.36)$$

where

$$f(\beta_{D,\max}) = \max \left\{ 0, 1 - 2\beta_{D,\max} - \beta_{D,\max}^2 \right\}. \quad (6.37)$$

Figure 6.6 depicts the achievable performance using a ZF precoder and compares it to the upper and lower bounds. In this figure, as well as in all numerical rates evaluation in what follows, we assume a 0 dB Shannon gap, i.e., $\Gamma = 1$. Note that the bounds are tight for low frequencies. For such frequencies $\beta_{D,\max}$ is very small, and comparing (6.35) and (6.36) the bound tightness for small enough $\beta_{D,\max}$ is obvious. For higher frequencies the lower bound becomes useless when $\beta_{D,\max} > \sqrt{2} - 1$. Yet, the actual ZF performance is still very close to the upper bound for all frequencies. For reference, the figure also depicts the achievable rates without precoding, and the achievable rates in the single user case, i.e., when only a single twisted pair in the binder is used. It shows that ZF performance nearly matches the single user performance and is significantly higher than the no precoding scheme. However, although the single user performance seems as worthy target, it is important to note that this curve is not a performance bound. This is because we cannot rule out the possibility that the inter-user connections will be used to increase user capacity. In particular, the upper bound corresponds to the case where all the inter-user connections for one of the users are used to increase this user's capacity. Thus, the importance of the upper bound is in showing that even in the most optimistic case, the achievable rate cannot exceed the ZF rates by more than 10%.

Figure 6.7 shows that the Cendrillon et al. [38] bounds were ahead of their time. The figure depicts the actual achievable rate in the case that all of the VDSL bandwidth is dedicated to the downstream vs.

**FIGURE 6.6**

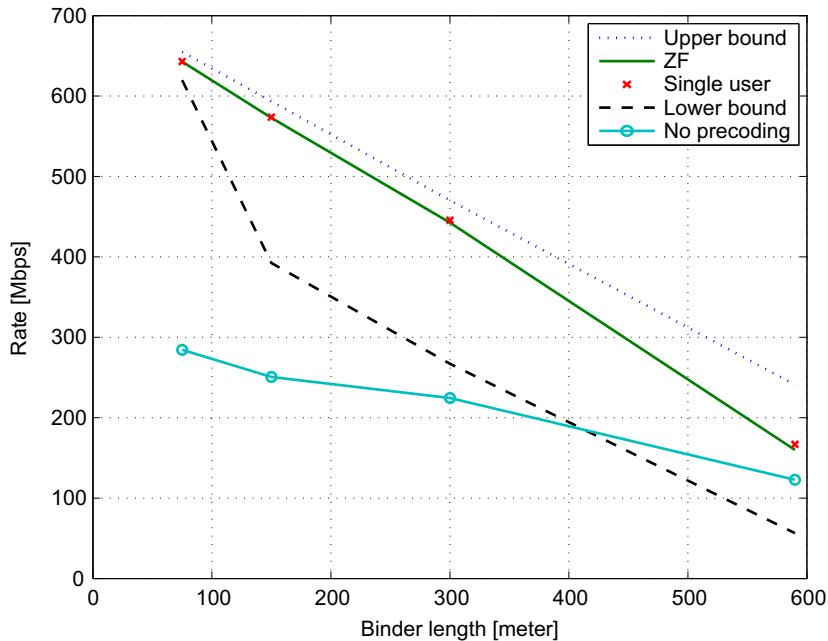
Average achievable user rate over a single subcarrier and bounds, as a function of the subcarrier frequency, at a 150 m binder.

the distance between the ONU and the CPEs. The figure also presents the upper bound of Theorem 3 and the lower bound of Theorem 2. It is clear that the bounds are tighter and the achievable data rate is higher for shorter link distances. In recent years, as the optical fibers deployment has expanded, the lengths of VDSL links have continually decreased. Hence, although the bounds presented by Cendrillon et al. were not tight for the typical link length at the time, nowadays, and certainly in the future, the typical link length is much shorter, and the bounds can be considered tight.

Nevertheless, it should be pointed out that the final conclusions of Cendrillon et al. [38] hold for all practical link distances. Comparing the achievable data rates of the ZF precoder to the upper bound shows that the precoder is indeed close to optimal. Furthermore, the performance of the ZF precoder is nearly identical to the more reasonable target of the single user performance (i.e., a system with a single user, transmitting over only one twisted pair).

2.06.3.2.2 Reduced complexity ZF precoding

The use of linear precoders reduces the transmitter implementation complexity, and does not require any modification of the CPE (the downstream receivers). But as DSL technology achieves ever higher data rates, the implementation complexity of linear receivers is still too high. The implementation complexity

**FIGURE 6.7**

Average achievable user rate over the whole bandwidth vs. binder length.

is composed of the computation of the precoding matrix, and the computation of the precoded signals using the precoding matrix.

To reduce the complexity of the first part, Leshem and Li [22] suggested computing the precoder matrix using a first or second order approximation of the channel matrix. As stated above, the γ power normalization constant in (6.32) is very close to 1. Neglecting this power normalization constant, the ZF precoder is given by:

$$\mathbf{F}_{\text{ZF}} = \mathbf{H}^{-1} \mathbf{D}. \quad (6.38)$$

Using the notation of (6.18) the ZF precoder can be rewritten as:

$$\mathbf{F}_{\text{ZF}} = (\mathbf{I} + \mathbf{E})^{-1}. \quad (6.39)$$

As shown in [28], if $\beta_{\mathbf{D},\max} < 1$, then the maximal singular value of \mathbf{E} is smaller than one, and hence the ZF precoder, (6.39), can be represented by a converging Taylor expansion:

$$\mathbf{F}_{\text{ZF}} = \mathbf{I} - \mathbf{E} + \cdots + (-1)^\ell \mathbf{E}^\ell + \cdots. \quad (6.40)$$

Hence, they suggested using a first or second order Taylor approximation of (6.40), which eliminates the need for matrix inversion.

The first order approximation is given by:

$$\mathbf{F}_{\text{ZF}1} = \mathbf{I} - \mathbf{E} = 2 \cdot \mathbf{I} - \mathbf{D}^{-1}\mathbf{H}. \quad (6.41)$$

The complexity reduction is significant in that the matrix inverse operation in (6.38) is replaced by the inverse of a diagonal matrix, which requires only L single element inversions. For the cases in which the first order approximation is not accurate enough, Leshem and Li suggested using the second order approximation, given by:

$$\mathbf{F}_{\text{ZF}2} = \mathbf{I} - \mathbf{E} + \mathbf{E}^2 = \mathbf{I} + (\mathbf{I} - \mathbf{D}^{-1}\mathbf{H})(2 \cdot \mathbf{I} - \mathbf{D}^{-1}\mathbf{H}). \quad (6.42)$$

See [22] for a detailed analysis of this implementation approach and the complexity reduction factor.

They also provided a lower bound on the performance of their proposed precoders. Interestingly, their bounds require only a form of weak RWDD, i.e., the bounds hold even in cases where the Taylor expansion in (6.40) might not converge. For example, for the first order ZF precoder they give the following bound:

Theorem 3 [22]. *If the channel matrix is weakly RWDD, and satisfies $(L - 1)\beta_{\text{D},1}^2 < 1$, then the data rate achievable by user i in any specific subcarrier of the downstream of a DSL system that applies the first order precoder given in (6.41) is lower bounded by:*

$$R \geq \Delta_F \cdot \log_2 \left(1 + \frac{\eta(1 - (L - 1)\beta_{\text{D},1}^2)|H(i,i)|^2}{\Gamma(1 + \eta(L - 2)^2(L - 1)\beta_{\text{D},1}^4|H(i,i)|^2)} \right). \quad (6.43)$$

Figure 6.8 shows the performance of the approximated precoders over real channel measurements. As can be seen, the first order approximation is good when the FEXT is low (in particular at low frequencies). For higher frequencies, the first order approximation precoder still achieves reasonable precoding at a link distance of 300 m (a capacity loss of around 10%). As link distances get shorter, the SNR increases and the FEXT cancellation of the first order approximation is not sufficient, as can be seen for the 75 m binder in the figure. On the other hand, the second order approximation precoder achieves near optimal performance in all the tested scenarios.

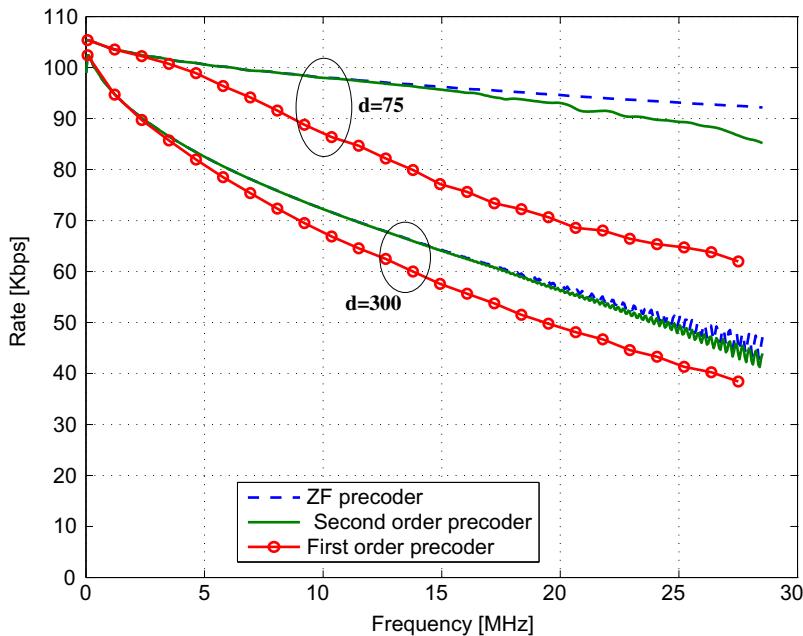
However, the computational complexity of the matrix inversion is not the only problem. With 4000 symbols per second, and thousands of subcarriers, even the application of the precoder (i.e., the multiplication $\mathbf{F} \cdot \mathbf{u}$) requires too many computations. This implementation complexity has been researched from two points of view. First, the number of bits per sample needs to be considered.

2.06.3.2.3 Quantization of word length for ZF precoding

Sayag et al. [41] studied the effect of word length on the performance of ZF precoders. Their main theorem formulates the capacity loss in a system that uses d bits for the quantization of the real part and d bits for the quantization of the imaginary part of the precoding matrix. Their main theorem is:

Theorem 4. *If the number of quantization bits of the ZF precoder satisfies $d \geq \frac{1}{2} + \log_2(1 + \beta_{\text{row}})$, then the transmission rate loss of the i th user at any specific subcarrier due to quantization is upper bounded by:*

$$\log_2(1 + 2(L - 1)(1 + \beta_{\text{row}})^2 2^{-2d} |H(i,i)|^2 \eta) - 2 \log_2(1 - \sqrt{2}(1 + \beta_{\text{row}}) 2^{-d}). \quad (6.44)$$

**FIGURE 6.8**

Average achievable user rate over a single subcarrier as a function of the subcarrier frequency, for the ZF precoder, and its first and second order approximation. The figure shows results for a 75 m binder and for a 300 m binder.

In their work they further studied the rate loss using the Werner channel model [10], and show, for example, that a quantization of 14 bits for the real part and 14 bits for the imaginary part is sufficient to guarantee a rate loss of no more than 1% for links of 300 m or more. Note that shorter links experience higher signal levels, and hence the FEXT cancellation requirements are stricter (thus requiring more quantization bits).

2.06.3.2.4 Partial ZF precoding

Another way to reduce the complexity of FEXT cancellation is to cancel only some of the interference sources. Writing explicitly the computation of each transmitted symbol, from (6.30) we have:

$$x(i) = \sum_{j=0}^{L-1} F(i, j)u(j). \quad (6.45)$$

Thus, every zero element in the precoder matrix, \mathbf{F} , saves one multiplication. Cendrillon et al. [25] studied the performance of ZF DSL precoders with various constraints on the number of non-zero elements in \mathbf{F} .

Given the set of interference sources to be canceled for each user, the evaluation of the optimal precoder is not trivial. Cendrillon et al. suggested an adjustment of the ZF criterion for this case, by reducing the number of constraints. If the total number of non-zero elements in the precoding matrix \mathbf{F} is r_T , one can solve a system of r_T linear equations, i.e., force r_T zeros in the resulting effective channel, \mathbf{HF} . Cendrillon et al. showed how to solve this ZF problem when the zero elements of \mathbf{HF} are in the same locations as the zero elements in the precoding matrix, \mathbf{F} . The description given here is slightly more general than the original, to include also the equations needed for partial adaptive precoding presented in Section 2.06.3.3.

Let $\mathcal{G}_j = \{g_j^1 \dots g_j^{r_j}\} \subseteq \{1, 2 \dots L\}$ be the set of receivers that need to cancel the FEXT originating from user j . Denote the set size by $0 \leq r_j = |\mathcal{G}_j| \leq L$.⁵ We define the ZF partial precoder as:

$$\mathbf{F}_P = \hat{\mathbf{F}}_P + I, \quad (6.46)$$

where the matrix $\hat{\mathbf{F}}_P$ satisfies $\hat{F}_P(i, j) = 0$ for any $i \notin \mathcal{G}_j$. The elements with non-zero values in $\hat{\mathbf{F}}_P$ are constructed according to the generalization of the ZF principle, which completely cancels the FEXT generated by the j th user to all users $i \in \mathcal{G}_j$. To simplify the precoder analysis we define the $(r_j \times L)$ selection matrices, Γ_j :

$$\Gamma_j(m, i) = \begin{cases} 1 & i = g_j^m \\ 0 & \text{otherwise} \end{cases}. \quad (6.47)$$

Applying the partial ZF criterion, the j th column of the precoder is constructed to satisfy:

$$\Gamma_j \mathbf{H} \mathbf{F}_P^{(j)} = \Gamma_j \mathbf{D}^{(j)}, \quad (6.48)$$

where $\mathbf{A}^{(j)}$ is the j th column of the matrix \mathbf{A} . Note that $\hat{\mathbf{F}}_P^{(j)}$ has zeros in the rows that do not belong to the set \mathcal{G}_j ; thus, the non-zero elements of $\hat{\mathbf{F}}_P^{(j)}$ are given by $\Gamma_j \hat{\mathbf{F}}_P^{(j)}$, and:

$$\hat{\mathbf{F}}_P^{(j)} = \Gamma_j^T \Gamma_j \hat{\mathbf{F}}_P^{(j)}. \quad (6.49)$$

Substituting (6.46) and (6.49), in (6.48) we have:

$$\Gamma_j \mathbf{D}^{(j)} = \Gamma_j \mathbf{H}^{(j)} + \Gamma_j \mathbf{H} \Gamma_j^T \Gamma_j \hat{\mathbf{F}}_P^{(j)}, \quad (6.50)$$

and the non-zero elements of $\hat{\mathbf{F}}_P^{(j)}$ are given by:

$$\Gamma_j \hat{\mathbf{F}}_P^{(j)} = (\Gamma_j \mathbf{H} \Gamma_j^T)^{-1} \Gamma_j (\mathbf{D}^{(j)} - \mathbf{H}^{(j)}). \quad (6.51)$$

Note that $\Gamma_j \mathbf{H} \Gamma_j^T$ is a non singular matrix formed by the elements of the channel matrix that are located in the rows and columns that belong to \mathcal{G}_j (i.e., by $H(m, i) : m, i \in \mathcal{G}_j$). Partial precoder results with

⁵Note that we allow the set \mathcal{G}_j to be empty. On the other hand, Cendrillon et al. [25] required \mathcal{G}_j to always include at least the element j . As can be seen from (6.46), setting $\mathcal{G}_j = \mathcal{G}_j \cup j$ does not increase the implementation complexity. Thus, $\mathcal{G}_j = \emptyset$ is required only in the context of adaptive precoding.

residual FEXT whose powers (for all users) are given by the diagonal of:

$$\mathbf{Z} = p \sum_{j=1}^L \left(\mathbf{H}\mathbf{F}_P^{(j)} - \mathbf{D}^{(j)} \right) \left(\mathbf{H}\mathbf{F}_P^{(j)} - \mathbf{D}^{(j)} \right)^H. \quad (6.52)$$

However, the utilization of a partial precoder requires the transmitter to first obtain a good enough channel estimation, and then apply a good enough selection algorithm. Cendrillon et al. [25] considered several user selection algorithms. They tested a binder with 8 users in which users at 900 m could increase their rate by 72% using complete FEXT cancellation. They showed that by proper user selection, a system can achieve 41% of this rate increase with only 20% of the implementation complexity. They also considered a FEXT cancellation over only part of the DMT subcarriers, and showed that a scheme that performs both user and tone selection can perform even better and achieve 77% of the rate increase with 20% of the implementation complexity. (Note however for shorter link lengths that FEXT is more dominant and the gains with the same implementation complexity will be smaller). Figure 6.11 below also depicts the performance of a partial precoder for a 28 user binder over a length of 300 m. Note that r_T represents the number of multiplications per precoding operation, averaged over all subcarriers. Again a significant part of the rate gain can be achieved with a relatively small part of the complexity. Vangorp et al. [42] extended this work, and presented a low complexity algorithm for optimal resource allocation given limited complexity for the partial FEXT cancellation.

Uplink partial FEXT cancellation was also considered in the same framework [24]. We will not describe this here in more detail, since the analysis and results are quite similar. It should be noted that for the upstream Pandey et al. [43] also considered a partial MMSE receiver. But as demonstrated in Section 2.06.4.1, the MMSE receiver does not gain much in most common scenarios.

2.06.3.3 Adaptive precoding

The use of low complexity linear precoders made FEXT cancellation feasible in high rate DSL systems. However, recall that in the downstream the transmitter at the ONU cannot directly measure the channel matrix. Hence, the calculation of the precoding matrix must rely on feedback from the receivers. Two main types of feedback have been considered: channel estimation feedback and signal error feedback.

Channel estimation feedback is based on the transmission of orthogonal (synchronized) pilot symbols from all transmitters simultaneously, and an estimation of a row of the channel matrix by each receiver. This estimated row is then transmitted (through the upstream) back to the transmitter, which uses it to construct the full channel matrix \mathbf{H} and to calculate the precoding matrix \mathbf{F} . This approach mostly uses standard (and not DSL specific) techniques, and hence has mostly been discussed in implementation oriented publications.

Signal error feedback is a more DSL specific method, and is based on the feedback of a quantized version of the error signal measured by each receiver. The error signal measured by the i th receiver is:

$$\epsilon(i) = y(i) - H(i, i)u(i). \quad (6.53)$$

The ONU collects all error signals feedback into a vector:

$$\boldsymbol{\epsilon} = [\epsilon(0), \dots, \epsilon(L-1)]^T = \mathbf{y} - \mathbf{Du} = (\mathbf{H}\mathbf{F} - \mathbf{D})\mathbf{u} + \mathbf{w}. \quad (6.54)$$

This error signal vector can be used by the transmitter to adapt a precoder that attempts to minimize the error signal energy.

However, because the error signal is measured by the receiver, it needs to be quantized and sent back to the receiver. We assume hereafter that the error signal is first normalized and then quantized. Each receiver normalizes its error signal by dividing it with the direct channel gain, $H(i, i)$. Thus, the quantized error signal $\hat{\epsilon}_n$ can be written as:

$$\begin{aligned}\hat{\epsilon} &= \mathbf{D}^{-1}(\mathbf{y} - \mathbf{Du}) + \mathbf{z} \\ &= (\mathbf{D}^{-1}\mathbf{HF} - I)\mathbf{u} + \mathbf{D}^{-1}\mathbf{w} + \mathbf{z},\end{aligned}\quad (6.55)$$

and \mathbf{z} is the quantization error. Note that \mathbf{z} can be statistically dependent on ϵ .

Adaptive precoding for DSL was first suggested by Louveaux and van der Veen [26]. They considered an adaptive precoder designed to minimize the norm $\|\epsilon\|^2$. Taking the derivative with respect to the precoding matrix, using the instantaneous correlation matrix and taking some approximations, they suggested using the update equation:

$$\mathbf{F}_{n+1} = \mathbf{F}_n - \alpha \mathbf{F}_n^{-T} \mathbf{D} \left[\mathbf{y}_n^* \mathbf{u}_n^T \right]_{nd}, \quad (6.56)$$

where the subscript indicates the time (i.e., \mathbf{F}_n , \mathbf{y}_n , and \mathbf{u}_n denotes the value at time n of the precoding matrix the received symbol and modulated data, respectively), α is the precoder update constant, and $[\bullet]_{nd}$ denotes the operation of zeroing all diagonal elements of a matrix. They also considered the use of a quantized version of the received signal, but gave no details. This adaptive precoder was studied only by simulations.

A simpler version of an adaptive precoder was proposed independently by Louveaux and van der Veen [27] and by Bergel and Leshem [28]. Their suggested precoder update is given by⁶:

$$\mathbf{F}_{n+1} = \mathbf{F}_n - \alpha \hat{\epsilon}_n \mathbf{u}_n^H. \quad (6.57)$$

This adaptive precoder has much lower implementation complexity than (6.56), as it does not include multiplication by the inverse matrix. This precoder is also simpler in terms of analysis. In [28] Bergel and Leshem considered the simplified case of rich enough quantization, such that the quantization error, \mathbf{z} is statistically independent on the error signal ϵ . For this simplified case they provided both convergence bounds and a steady state error analysis. The convergence is guaranteed by the following theorem:

Theorem 5 [28]. *A sufficient condition for the precoder given by (6.57) to converge is that:*

$$\beta_{D,max} < 1 \quad \text{and} \quad \alpha < \frac{2}{Kp} \cdot \frac{1}{1 + \beta_{D,max}}, \quad (6.58)$$

where $K = L - 1 + E [|u(i)|^4] / p^2$ (i.e., $E[\mathbf{u}\mathbf{u}^H \mathbf{u}\mathbf{u}^H] = Kp^2 I$). If (6.58) is satisfied, then after sufficiently long convergence time, the Frobenius norm of the difference between the zero forcing precoder

⁶The version by Louveaux and van der Veen only takes the non-diagonal elements of the update matrix. This version is not described here as its analysis is more complicated.

$(\mathbf{F}_{ZF} = \mathbf{H}^{-1}\mathbf{D})$ and the adaptive precoder is upper bounded by:

$$\lim_{n \rightarrow \infty} \text{Tr} \left\{ (\mathbf{F}_n - \mathbf{F}_{ZF})(\mathbf{F}_n - \mathbf{F}_{ZF})^H \right\} \leq \frac{\alpha L \sum_{i=0}^{L-1} \left(\sigma_z^2 + \frac{\sigma^2}{|H(i,i)|^2} \right)}{2 - \alpha K p (1 + \beta_{D,\max}^2) - 2|1 - \alpha K p| \beta_{D,\max}}, \quad (6.59)$$

where σ_z^2 is the variance of the quantization noise \mathbf{z} .

As can be seen in Figure 6.4 the strong RWDD condition in the theorem ($\beta_{D,\max} < 1$) is satisfied practically for all frequencies for link distances of up to 300 m. Thus, the precoder is very robust, and can converge in all channels. As $\beta_{D,\max}$ is typically not known in advance, a practical approach would be to take its maximal value when choosing the update constant, and use $\alpha \leq 1/Kp$. In such case, (6.59) simplifies to:

$$\lim_{n \rightarrow \infty} \text{Tr} \left\{ (\mathbf{F}_n - \mathbf{F}_{ZF})(\mathbf{F}_n - \mathbf{F}_{ZF})^H \right\} \leq \frac{\alpha L \sum_{i=0}^{L-1} \left(\sigma_z^2 + \frac{\sigma^2}{|H(i,i)|^2} \right)}{2(1 - \beta_{D,\max}) - \alpha K p (1 - \beta_{D,\max})^2}. \quad (6.60)$$

The convergence bound is important to guarantee the converge of the precoder. However, once the precoder has converged to a steady state, (6.59) does not provide a convenient performance measure, as it relates to the precoder error and not to the final performance. In the same work they also provide a more convenient steady state approximation. For large enough n (so that the precoder reaches steady state), the covariance matrix of the error signal is approximated by:

$$E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H] \simeq \left(1 + \frac{\alpha L p}{2 - \alpha K p} \right) \sigma^2 \cdot I + \frac{\alpha L p}{2 - \alpha K p} \sigma_z^2 \mathbf{D}\mathbf{D}^H, \quad (6.61)$$

where $\sigma_z^2 = E[z(i)^2]$ is the variance of the quantization noise. The covariance matrix in (6.61) is composed of two terms. The first is the error increase due to the use of the iterative precoder (note that the error signal covariance matrix with no FEXT is I). This error increase is negligible if $\alpha L p \ll 1$. The second term is due to the quantization error, and does not depend on the noise power.

Nevertheless, the simplifying assumption of rich quantization [28] is impractical, because such a quantization would cause a high feedback rate, and overload the upstream. Louveaux and van der Veen [44] presented an approximate analysis for the case of phase feedback only, under the assumption that the signal of each modem can be analyzed independently (which holds for very low $\beta_{D,\max}$). In a recent work, Binyamini and Bergel [45] analyzed the same precoder with low rate quantization. In particular they considered a low bit rate dithered [46] uniform quantization only of the phase of the error signal. This work showed that the precoder will converge for any update constant, α , and for any quantization level (even 1 bit per sample) as long as the channel is strongly RWDD and $\beta_{D,\max} < 1$.

Their quantization scheme uses a sequence of pseudo random phase variables, $\phi_n(i)$, which are designed to behave as iid random variables uniformly distributed over $[0, 2\pi)$ (in the what follows we assume that this is their actual distribution). This sequence is known in advance both at the transmitter and at all receivers. The error signal of the i th user at time n , $\epsilon_n(i)$, is first rotated by multiplying it with $e^{j\phi_n(i)}$, and then quantized to one of the values $e^{j\frac{k2\pi}{M}}$, where $k = 0, 1, \dots, M-1$, and M is the number of quantization values. At the transmitter, the actual quantization value is obtained by multiplying the

received quantization level by $e^{-j\frac{k2\pi}{M}}$. Thus, the quantized error signal can be written as:

$$\hat{\epsilon}_n(i) = e^{j\theta_n(i)} \frac{H^{-1}(i, i) \cdot \epsilon_n(i)}{|H^{-1}(i, i) \cdot \epsilon_n(i)|}, \quad (6.62)$$

and $\theta_n(i)$ is the phase quantization error. Note that the dithering operation ensures that the phase quantization error is an iid sequence, and each element is uniformly distributed over $[-\pi/M, \pi/M]$. In [45] the authors show that the dithering has a negligible effect on system performance, and is required mostly for analysis purposes, due to the convenient distribution of the quantization error. The convergence of a precoder that employs a dithered phase quantization scheme is guaranteed by:

Theorem 6 [45]. *Considering an adaptive precoder of the form of (6.57) which uses dithered phase quantization as described in (6.62). If $\beta_{D,\max} < 1$, the precoder converges, and after a sufficiently long convergence time, the averaged in time expectation on the absolute error measured by the i th receiver is upper bounded by:*

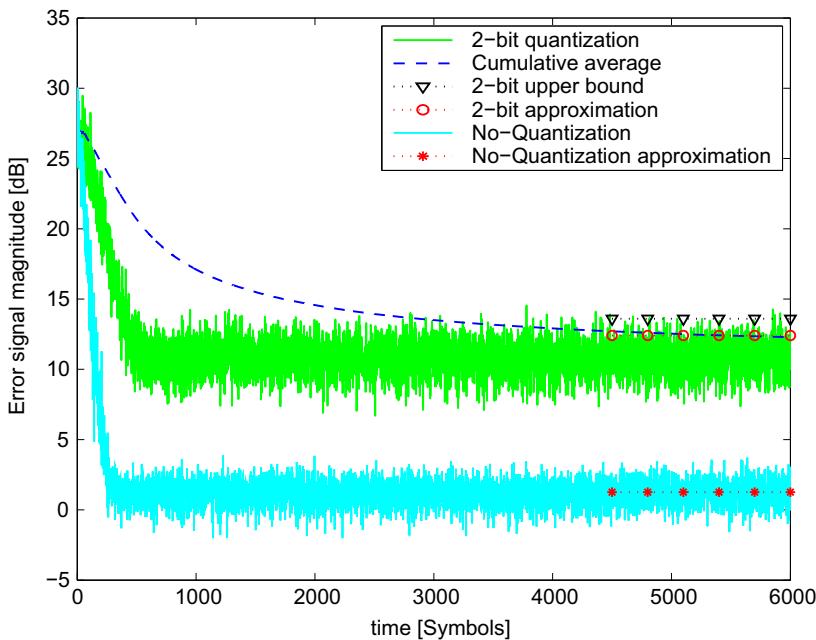
$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{n=1}^T E[|\epsilon_n(i)|] \leq \frac{\alpha L p (1 + \beta_{D,\max})^2 |H(i, i)|}{2(1 - \beta_{D,\max}) \cdot \frac{M}{\pi} \sin \frac{\pi}{M}} + \frac{\sqrt{\frac{\pi}{4}} \cdot \sigma}{(1 - \beta_{D,\max})}. \quad (6.63)$$

The theorem shows that the error signal is bounded by a noise increase of $1 - \beta_{D,\max}$ plus an additional term which is linear with the precoder update constant, α . Thus, it is feasible to approach FEXT free performance by choosing small enough precoder update, up to the noise increase term.

Naturally such low rate quantization results in lower performance than the rich quantization in Theorem 5 (i.e., slower convergence and/or higher steady state errors). Nevertheless, unlike Theorem 5, the phase quantization scheme in Theorem 6 achieves convergence for every value of the update constant. Note however that the bound (6.63) is not very tight. The authors observed that a better approximation of the actual error signal magnitude is obtained by substituting $\beta_{D,\max} = 0$ in (6.63) instead of the actual value of $\beta_{D,\max}$.

Figure 6.9 shows the (averaged over all users) normalized error signal magnitude of two adaptive precoders vs. the time (in symbols), over the subcarrier at a frequency of 5.78 MHz. The lower curve shows the magnitude of the error signal, divided by σ , in the precoder of (6.57) with no quantization and with $\alpha = 1/(2Lp)$ (about half the maximal allowed value). The dotted curve with stars shows the steady state approximation in (6.61) for the no quantization case (i.e., $10 \log_{10} (1 + \frac{\alpha L p}{2 - \alpha K p})$). The higher curve depicts the magnitude of the error signal, divided by σ when the adaptive precoder employs the low rate quantization of Theorem 6 with $M = 2$ (1-bit feedback) and $\alpha = 0.002/L\sqrt{p}$ (note that the measured user gain at this frequency is 0.21). As can be seen, the precoder converges at about the same time as the no quantization precoder, but with much higher error signal magnitude. The dashed line shows the cumulative averaged in time error signal divided by σ : $20 \log_{10} \left(\frac{1}{LT} \sum_{n=1}^T \sum_{i=0}^{L-1} |\epsilon_n(i)| / \sigma \right)$, and the dotted line with triangles depicts the steady state bound of (6.63). The dotted line with circles shows the steady state approximation (calculated by substituting $\beta_{D,\max} = 0$ in (6.63)). As can be seen, the precoders converge reasonably fast, and the suggested approximations are quite good.

To further demonstrate the accuracy of the suggested steady state error magnitude approximations, Figure 6.10 shows the average steady state error magnitude and compares it to the approximations

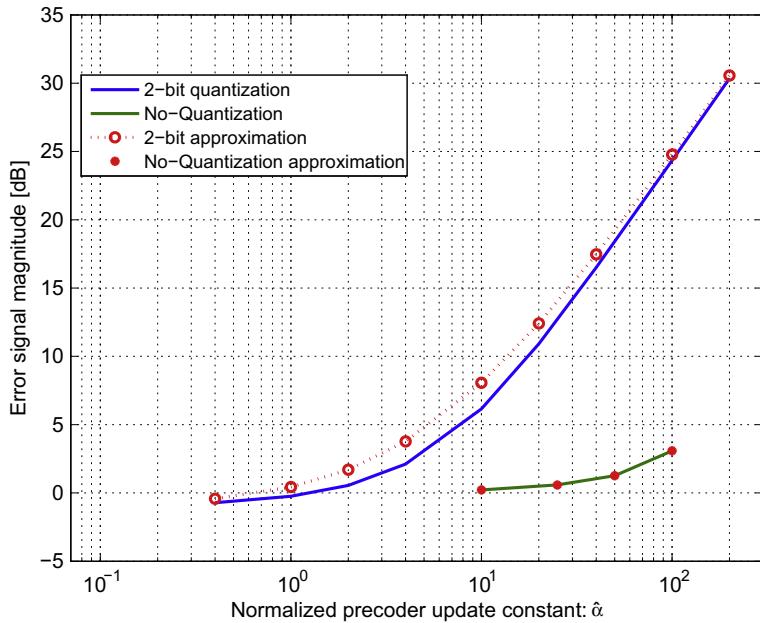
**FIGURE 6.9**

Error signal magnitude vs. time for two adaptive precoders.

of (6.61) and (6.63) vs. the normalized precoder update constant. (Note that in order to match the approximation expressions, the average is taken over the absolute value of the error signal magnitude for the low rate precoder, while for the non quantized precoder the average is taken over the square magnitude). The normalized precoder update constant is given by $\hat{\alpha} = \alpha L p$ for the no quantization case, and by $\hat{\alpha} = 100\alpha L \sqrt{p}$ for the low rate quantization precoder. One can see that both approximations are quite good. In particular, the low rate quantization precoder approximation holds even for very high values of α . This demonstrates the claim of Theorem 6 that the precoder converges for every value of α . Note however that very high values of α are not desirable, as they can cause even a FEXT increase compared to the original no precoding scheme.

The situation is slightly more complicated in mixed length scenarios, where the different pairs in the binder reach destinations at different distances. From (6.59) it can be seen that the precoder convergence is dominated by the channel gain of the user with the longest (weakest) link. On the other hand, in many cases, these weak users are noise limited and not FEXT limited, and hence FEXT cancellation is not useful for them. For this reason, it was suggested in [28] not to implement FEXT cancellation for such users, based on an SNR threshold test. (Note that the FEXT from these users to other users is still canceled.) Such an approach was shown to significantly shorten the convergence times in mixed length scenarios.

Binyamini et al. [29] observed that partial FEXT cancellation is also desirable to reduce implementation complexity. In this work the partial FEXT cancellation described above (in which some users do

**FIGURE 6.10**

Steady state error signal magnitude vs. normalized precoder update constant. The normalized precoder update constant is given by $\hat{\alpha} = \alpha L p$ for the no quantization case (marked by stars), and by $\hat{\alpha} = 100\alpha L \sqrt{p}$ for the low rate quantization precoder (marked by circles).

not employ FEXT cancellation at all) is termed partial FEXT cancellation of type I. They also define a partial FEXT cancellation of type II, where (for some users) only the FEXT from a subgroup of the users is canceled. This precoder takes advantage of the implementation complexity gain described by Cendrillon et al. [25] and in Section 2.06.3.2.4.

They suggested a simple adaption of the precoder in (6.57), which does not update the precoder at the non-canceled terms. The mathematical formulation (as well as the analysis) of this type of precoder is more complicated, as its update equation cannot be described in a matrix form. Instead, the precoder update for the j th column is given by:

$$\mathbf{F}_{n+1}^{(j)} = \mathbf{F}_n^{(j)} - \alpha \boldsymbol{\Gamma}_j^T \boldsymbol{\Gamma}_j \hat{\boldsymbol{\epsilon}}_n \mathbf{u}_n^H, \quad (6.64)$$

where the matrix $\boldsymbol{\Gamma}_j$ describes the selected users for FEXT cancellation as described in Section 2.06.3.2.4. Note that the set of L update Eq. (6.64), $j = 1, \dots, L$, are actually coupled, because the error signal at the $n + 1$ th symbol depends on all the columns of the precoding matrix that were updated in the n th symbol.

Although the analysis is more complicated, Binyamini et al. [29] showed that it complies with the same sufficient convergence condition as in the full update precoder (Theorem 5). More specifically:

Theorem 7 [29]. Assuming that the quantization noise is statistically independent of the error signal, a sufficient condition for the partial precoder given by (6.64) to converge is that:

$$\beta_{D,\max} < 1 \text{ and } \alpha < \frac{2}{Kp} \cdot \frac{1}{1 + \beta_{D,\max}}. \quad (6.65)$$

If (6.65) is satisfied, then after sufficiently long convergence time, the Frobenius norm of the difference between the precoder \mathbf{F}_n and the non-adaptive partial precoder \mathbf{F}_P that satisfies (6.48) and (6.49) is upper bounded by:

$$\lim_{n \rightarrow \infty} \text{Tr} \left\{ (\mathbf{F}_n - \mathbf{F}_P)(\mathbf{F}_n - \mathbf{F}_P)^H \right\} \leq \frac{\alpha \sum_{j=0}^{L-1} \sum_{i \in \mathcal{G}_j} \left(\sigma_z^2 + \frac{\sigma^2 + Z(i,i)}{|H(i,i)|^2} \right)}{2 - \alpha Kp(1 + \beta_{D,\max}^2) - 2|1 - \alpha Kp|\beta_{D,\max}}, \quad (6.66)$$

where σ_z^2 is the variance of the quantization noise \mathbf{z} and $Z(i, i)$ is the residual FEXT of the non-adaptive partial FEXT precoder, i.e., the i th element on the diagonal of the covariance residual FEXT covariance matrix, \mathbf{Z} given in (6.52).

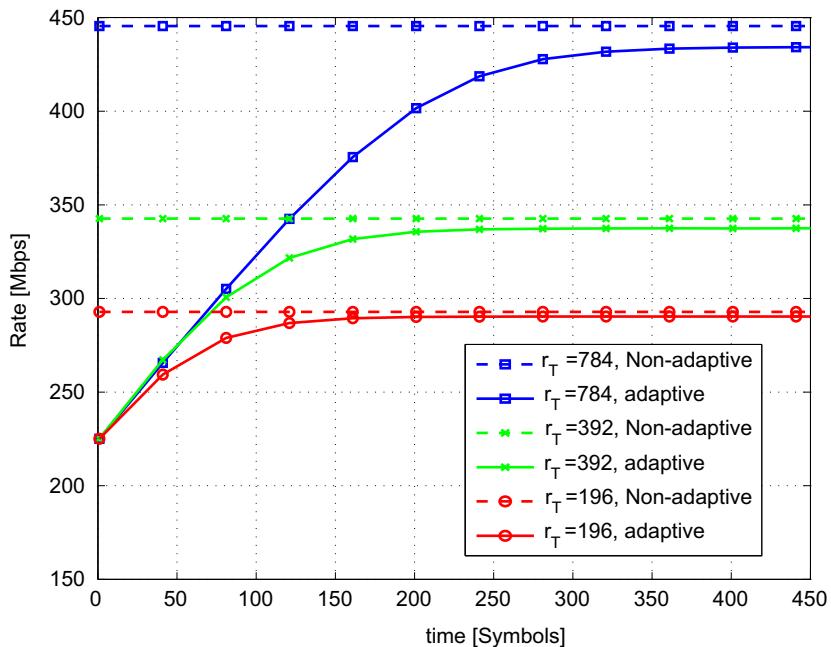
In the same work, they also provide a steady state approximation which is demonstrated to be quite accurate. This approximation is given by:

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{L-1} E[|\epsilon_n(j)|^2] \approx \left(\frac{\alpha p}{2 - \alpha Kp} \frac{r_T}{L} + 1 \right) \sum_{j=0}^{L-1} \left(\sigma^2 + Z(j, j) \right) + \frac{\alpha p \sigma_z^2}{2 - \alpha Kp} \frac{r_T}{L} \sum_{j=0}^{L-1} |H(i, i)|^2, \quad (6.67)$$

where $r_T = \sum_j r_j$ is the total number of elements used in the partial adaptive precoder. Note that the sum over $(1 + Z(j, j))$ is the power of the noise plus residual FEXT in the non-adaptive partial precoder of Section 2.06.3.2.4. Thus, the term $\frac{\alpha p}{2 - \alpha Kp} \frac{r_T}{L}$ represents the increase in the error signal power due to the adaptive precoding. The second term in the right hand side of (6.67) represents the effect of the error signal quantization noise, σ_z^2 .

The performance of the partial adaptive precoder is depicted in Figure 6.11, with a comparison to the non adaptive partial precoder described in Section 2.06.3.2.4. The figure depicts the average user rates (when all the bandwidth is allocated to the downstream) using partial precoding as a function of time, for various complexity levels. The case of $r_T = 28^2 = 784$ corresponds to full FEXT cancellation (i.e., in this case all the precoding matrix is updated). The performance in this case are identical to the performance described above. In the two other cases $r_T < 784$ represents the number of non-zero elements in the precoding matrix, averaged over all sub-carriers. It is clear that the adaptive partial precoder converges quite fast, and achieves rates close to those of the non-adaptive precoder. Note also that the difference between the adaptive and non-adaptive precoders is smaller for smaller values of r_T . This behavior is consistent with the approximation in (6.67), which shows that the additional FEXT due to the adaptive precoding is linear with r_T .

It is worth noting that the selection of the important terms for FEXT cancellation has not been studied so far in the context of adaptive precoding. This is an open issue because adaptive precoding systems do

**FIGURE 6.11**

Average achievable user rate over the whole bandwidth at a 300 m 28 user binder, using a adaptive and non-adaptive partial precoders.

not employ channel estimation, and only use the error signal feedback to adapt the precoder. Thus, such systems are required to select the important FEXT terms based on this error signal feedback. This can be done for example by testing long term correlations between the error signals and the transmitted symbols. Alternatively, one can try to use the precoder adaption as a learning mechanism, adapt the precoder for some test elements and then use the resulting precoder to decide on the significance of each element. Further research is required to define such methods and characterize their performance.

An alternative to the adaptive precoding scheme using error signal feedback is adaptive precoding using SNR measurement feedback, as presented by Whiting et al. [47]. The major advantage of this approach is that it requires only SNR feedback reports, which are inherent to all VDSL modems. Thus, adaptive FEXT cancellation can be applied without any change in the CPE. But, SNR reports have no phase information, and hence a more complex algorithm is required to learn the correct direction for precoder adaptation. The updated algorithm is composed of several stages in which the various lines are perturbed in various directions. Then the received SNR reports are compiled into an adaption of the precoder, and the process starts over again. Whiting et al. [47] provided convergence bounds and analytic expressions for the convergence time and expected performance. Comparing to the adaptive precoding using error signal described above, one should keep in mind that SNR feedback adaptive

precoding does not require any change in the CPE. On the other hand, its main disadvantage is its convergence time, which is several orders of magnitude slower (hundreds of symbols, i.e., less than a second using error signal feedback compared to minutes and even hours using SNR feedback).

2.06.4 Upstream transmission

2.06.4.1 Linear receiver

In the upstream, the receiving modems are colocated at the ONU. This allows for joint reception of the signals transmitted on the different twisted pairs. As was the case for the downstream, upstream performance has also been shown to be nearly optimized with a ZF receiver [48]. Using

$$\hat{\mathbf{y}} = \mathbf{H}^{-1}\mathbf{y} = \mathbf{x} + \mathbf{H}^{-1}\mathbf{w}, \quad (6.68)$$

and recalling that the upstream channel matrix is CWDD, Cendrillon et al. [48] presented bounds that guarantee that ZF precoder performance is very close to the maximal achievable performance. Their upper bound is given by:

Theorem 8 [48]. *The data rate achievable by user i in any specific subcarrier of the upstream of DSL systems is upper bounded by:*

$$C \leq \Delta_F \cdot \log_2 \left(1 + \Gamma^{-1} \eta \sigma^{-2} |H(i, i)|^2 \left(1 + (L - 1) \beta_{U,1}^2 \right) \right). \quad (6.69)$$

For the lower bound, we give here the bound from [40], which is slightly better than the original bound in [48]:

Theorem 9 [40]. *The data rate achievable by user i in any specific subcarrier of the upstream of DSL systems when the ONU applies a ZF receiver is lower bounded by:*

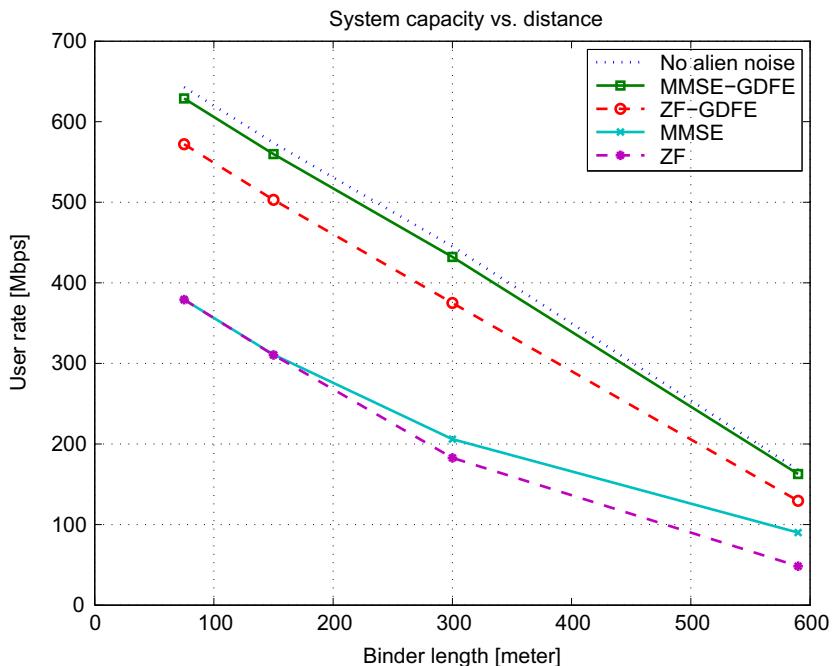
$$R \geq \Delta_F \cdot \log_2 \left(1 + \Gamma^{-1} \eta |H(i, i)|^2 \left(1 - 2\beta_{U,\max} - \beta_{U,\max}^2 \right) \right). \quad (6.70)$$

In some cases, the system is subject to interference from alien sources. This interference can be highly correlated between the different twisted pairs in the binder. Unlike the downstream, in the upstream the receiver can take advantage of this correlation and reduce the effect of alien interference. The ZF receiver does not take advantage of this correlation, and hence can suffer a dramatic decrease in capacity, as seen for example in Figure 6.12. To better visualize why the ZF receiver can be far from optimal in this case, one can consider a noise whitening approach:

$$\tilde{\mathbf{y}} = \Sigma^{-1/2}\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{w}}, \quad (6.71)$$

where Σ is the covariance matrix of \mathbf{w} , which includes both noise and interference. The resulting channel model is quite similar to the initial model in that the resulting noise is white ($E[\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H] = I$). However, Chen et al. [49, 50] observed that in some cases, the resulting effective channel matrix, $\tilde{\mathbf{H}}$, is no longer diagonal dominant. Thus, Theorem 9 does not help, and ZF performance can be far from optimal.⁷

⁷Note that in most cases, vectored VDSL systems coordinate all pairs in the binder. In such cases, the alien noise is usually negligible.

**FIGURE 6.12**

Average achievable user rate over the whole bandwidth in the upstream vs. binder length, for various receiver structures.

Pandey et al. [43] suggested that a MMSE linear receiver can gain from the alien noise correlation. They showed that for an 1200 m link in the presence of strong alien noise an MMSE receiver can achieve rates 5 times better than the rates achievable by ZF, and more than half of the actual capacity. However, for shorter link lengths, the signal power becomes more significant. Hence the MMSE receiver focuses on FEXT cancellation, and its performance is very close to the performance of the ZF receiver (see Figure 6.12).

2.06.4.2 Decision feedback equalizer

For cases in which the linear processing discussed above is too far from optimal, Chen et al. [50] proposed using a ZF generalized decision feedback equalizer (GDFE). The authors demonstrate that the ZF-GDFE is a simplified version of the MMSE-GDFE, which was shown to achieve optimal capacity in the Gaussian multiple access channel [51]. In the following we do not describe the MMSE-GDFE, as this general technique was developed mostly for wireless channels and has no particular characteristics in DSL systems.

In the ZF-GDFE scheme, the receiver applies a QR decomposition of the channel $\tilde{\mathbf{H}}$ of (6.71):

$$\tilde{\mathbf{H}} = \mathbf{Q}\mathbf{R}, \quad (6.72)$$

where \mathbf{Q} is a unitary matrix and \mathbf{R} is an upper triangular matrix. Rotating the input by:

$$\mathbf{Q}^H \tilde{\mathbf{y}} = \mathbf{Q}^H \Sigma^{-1/2} \mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{Q}^H \tilde{\mathbf{w}} \quad (6.73)$$

results in a simpler structure, which is convenient for GDFE. Due to the upper triangular structure of \mathbf{R} , the last user in the model of (6.73) receives the transmitted signal with no interference. This user is decoded first.⁸ Then, the effect of the decoded user is subtracted from the received signal, and the second user can be decoded with no interference. This scheme continues, until the effect of users 1 through $L - 1$ is subtracted, and user number 0 is decoded with no interference.

Note that in this description user number 0 has the highest priority, as the effect of all users is subtracted before its decoding. In order to optimize system performance, user priorities must be controlled. This is done by reordering the users before the detection. In addition, one can control the power allocation between the different users. Such an optimization is described in [50] (see also [52] for such an optimization for the MMSE-GDFE receiver).

On the other hand Leshem and Zehavi [53] noted that in DSL systems, the spectral mask power constraint is more dominant. Thus, the optimization with respect to the transmission powers is much simpler than in the problem considered by [50]. When the spectral mask power constraint is dominant, the optimal power allocation simply uses the maximal allowed power for each user at each subcarrier.

As for the user priorities and ordering, it turns out that in most cases there exists a trivial ordering that results in near optimal performance. For example, in the simulation of Figure 6.12, the ZF-GDFE receiver does not optimize user priorities, and selects user ordering in a random manner. The figure depicts the performance of several receiver types over various link lengths. The simulation includes 28 users and alien noise which is 30 dB stronger than the white noise. Without the alien noise, all receivers achieve nearly the same performance which is very close to channel capacity. In the presence of alien noise, the ZF receiver performs very poorly. The MMSE receiver performs somewhat better, but this advantage only holds for the longer link distances. As explained above, for short links the performance of the MMSE receiver is nearly identical to the performance of the ZF receiver. On the other hand, the ZF-GDFE receiver performs quite well, and can achieve 90% of the channel capacity. Note that the channel capacity (achievable using the MMSE-GDFE receiver) is only slightly affected by the presence of the alien noise.

It should be noted that complete FEXT cancellation may not always be feasible in near future practical systems. One factor is the complexity of MMSE-GDFE receivers that may prevent complete FEXT cancellation in the presence of alien noise. Another factor is the large amount of deployed CPEs that do not support vectoring. Thus, in the coming years we expect to see many systems that are partially vectored and partially non-vectored. In such cases it is important to combine FEXT cancellation techniques with dynamic spectrum management (DSM) techniques (see for example [54, 55]).

⁸Decoding can mean the decoding of the complete code word for this user. But due to complexity constraints, a more typical approach just takes hard decisions on the symbol value.

2.06.5 Conclusion

This chapter discussed the properties of the DSL multi-access and broadcast channels. It provided the basic system and channel models and presented the main solutions for both upstream and downstream. It is clear that a massive deployment of G.993.5 will take place in the near future. These systems perform multi-user MIMO-DMT transmission and reception using an order of 4000 tones and providing rates up to 100 Mbps. This massive vectoring is extremely important for reducing the need for fiber to the home deployment, and will enable aggregate rates larger than 5 Gbps for 64 users over distances of up to 800 m. Vectored DSL systems provide an excellent illustration of the fact that fitting signal processing and communication techniques to the specific properties of the channel at hand results in excellent performance at moderate computational cost.

Relevant Theory: Signal Processing Theory

See Vol. 1, Chapter 9 Discrete Multi-Scale Transforms in Signal Processing

References

- [1] J. Lechleider, Coordinated transmission for two-pair digital subscriber lines, *IEEE J. Sel. Areas Commun.* 9 (6) (1991) 920–930.
- [2] Symmetric Digital Subscriber Line (ADSL) Transceivers, ITU-T Recommendation G.992.1, SERIES G: Transmission Systems and Media, Digital Systems and Networks, 1999.
- [3] ITU-T Recommendation G.993.1-2004: Very High Speed Digital Subscriber Line. SERIES G: Transmission Systems and Media, Digital Systems and Networks, October 2004.
- [4] T. Starr, J.M. Cioffi, P. Silverman, *Understanding Digital Subscribe Line Technology*, Prencie-Hall, Upper Saddle River, NJ, 1999.
- [5] J.A.C. Bingham, *ADSL, VDSL, and Multicarrier Modulation*, John Wiley & Sons, Inc., 2000.
- [6] G. Foschini, Crosstalk in outside plant cable systems, *Bell Syst. Tech. J.* 50 (7) (1971) 2421–2450.
- [7] R. Addie, A. Gibbs, The covariance of near end crosstalk and its application to pcm system engineering in multipair cable, *IEEE Trans. Commun.* 27 (2) (1979) 469–477.
- [8] S. Lin, Statistical behaviour of multipair crosstalk, *Bell Syst. Tech. J.* 59 (6) (1980) 955–974.
- [9] S. Lin, Statistical behaviour of multipair crosstalk with dominant components, *Bell Syst. Tech. J.* 59 (6) (1980).
- [10] J. Werner, The HDSL environment, *IEEE J. Sel. Areas Commun.* 9 (6) (1991) 785–800.
- [11] K. Kerpez, Near-end crosstalk is almost gaussian, *IEEE Trans. Commun.* 41 (5) (1993) 670–672.
- [12] K. Kerpez, Models for the number of next disturbers and next loss, Tech. Rep., 1999, t1E1.4/99-471 (Online). Available: <ftp://ftp.t1.org/pub/t1e1/e1.4/DIR99/9e144710.pdf>.
- [13] I. Kalet, S. Shamai (Shitz), On the capacity of a twisted pair: Gaussian model, *IEEE Trans. Commun.* 38(3) (1990) 379–383.
- [14] J. Aslanis, J. Cioffi, Achievable information rates on digital subscriber loops: limiting information rates with crosstalk, *IEEE Trans. Commun.* 40 (2) (1992) 361–372.
- [15] E. Karipidis, N. Sidiropoulos, A. Leshem, Y. Li, Experimental evaluation of capacity statistics for short VDSL loops, *IEEE Trans. Commun.* 53 (7) (2005) 1119–1122.
- [16] E. Karipidis, N. Sidiropoulos, A. Leshem, Y. Li, R. Tarafi, M. Ouzzif, Crosstalk models for short VDSL2 lines from measured 30 MHz data, *EURASIP J. Appl. Signal Process.* 2006 (1) (2006) 90–90.

- [17] J. Lechleider, Coordinated transmission for two-pair digital subscriber lines, *IEEE J. Sel. Areas Commun.* 9 (6) (1991) 920–930.
- [18] G. Ginis, J. Cioffi, Vectored transmission for digital subscriber line systems, *IEEE J. Sel. Areas Commun.* 20 (5) (2002) 1085–1104.
- [19] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, G. Ginis, Improved linear crosstalk precompensation for DSL, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, Montreal, Canada, May, 2004, p. 4.
- [20] R. Cendrillon, M. Moonen, T. Bostoen, G. Ginis, The linear zero-forcing crosstalk canceller is near-optimal in DSL channels, in: *IEEE Global Communications Conference*, TX, Dallas, 2004, p. 5.
- [21] A. Leshem, Y. Li, A low complexity coordinated fext cancellation for VDSL, in: *Proceedings of IEEE ICECS 2004*, December 2004, pp. 338–341.
- [22] A. Leshem, Y. Li, A low complexity linear precoding technique for next generation VDSL downstream transmission over copper, *IEEE Trans. Signal Process.* 55 (11) (2007) 5527–5534.
- [23] Y. Li, A. Leshem, A lower bound on the performance of simplified linear precoding for vectored DSL, *Int. J. Pure Appl. Math.* 43 (1) (2008).
- [24] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, P. Vandaele, Partial crosstalk cancellation for upstream VDSL, *EURASIP J. Appl. Signal Process.* 2004 (2004) 1520–1535.
- [25] R. Cendrillon, G. Ginis, M. Moonen, K. Van Acker, Partial crosstalk precompensation in downstream VDSL, *Signal Process.* 84 (11) (2004) 2005–2019.
- [26] J. Louveaux, A. van der Veen, Adaptive DSL crosstalk precancellation design using low-rate feedback from end users, *IEEE Signal Process. Lett.* 13 (11) (2006) 665–668.
- [27] J. Louveaux, A.-J. van der Veen, Adaptive Precoding for Downstream Crosstalk Precancellation in DSL Systems Using Sign-Error Feedback, *IEEE Trans. Signal Process.* 58 (6) (2010) 3173–3179.
- [28] I. Bergel, A. Leshem, Convergence Analysis of Downstream VDSL Adaptive Multichannel Partial FEXT Cancellation, *IEEE Trans. Commun.* 58 (10) (2010) 3021–3027.
- [29] I. Binyamin, I. Bergel, A. Leshem, Arbitrary partial FEXT cancellation in adaptive precoding for multichannel downstream VDSL, *IEEE Trans. Signal Process.* 60 (11) (2012) 5754–5763.
- [30] ITU-T Recommendation G.993.5-2010, Self-FEXT Cancellation (Vectoring), 2010.
- [31] F. Sjöberg, M. Isaksson, R. Nilsson, P. Ödling, S. Wilson, P. Börjesson, Zipper: a duplex method for VDSL based on DMT, *IEEE Trans. Commun.* 47 (8) (1999) 1245–1252.
- [32] D. Mestdagh, M. Isaksson, P. Ödling, Zipper VDSL: a solution for robust duplex communication over telephone lines, *IEEE Commun. Mag.* 38 (5) (2000) 90–96.
- [33] M. Tomlinson, New automatic equaliser employing modulo arithmetic, *Electron. Lett.* 7 (5) (1971) 138–139.
- [34] H. Harashima, H. Miyakawa, Matched-transmission technique for channels with intersymbol interference, *IEEE Trans. Commun.* 20 (4) (1972) 774–780.
- [35] G. Caire, S. Shamai, On achievable rates in a multi-antenna broadcast downlink, in: *Proc. 38th Annu. Allerton Conf. Communication, Control and Computing*, 2000, pp. 1188–1193.
- [36] G. Caire, S. Shamai, On the achievable throughput of a multiantenna gaussian broadcast channel, *IEEE Trans. Inform. Theory* 49 (7) (2003) 1691–1706.
- [37] W. Yu, J. Cioffi, Trellis precoding for the broadcast channel, in: *IEEE Global Telecommunications Conference*, 2001, GLOBECOM’01, vol. 2, 2001, pp. 1344–1348.
- [38] R. Cendrillon, G. Ginis, E. Van den Bogaert, M. Moonen, A near-optimal linear crosstalk precoder for downstream VDSL, *IEEE Trans. Commun.* 55 (5) (2007) 860–863.
- [39] J. Cioffi, G. Dudevoir, M. Eyuboglu, G. Forney Jr., MMSE decision-feedback equalizers and coding. II. Coding results, *IEEE Trans. Commun.* 43 (10) (1995) 2595–2604.
- [40] I. Bergel, A. Leshem, The performance of zero forcing DSL systems, *IEEE Signal Process. Lett.* 20 (5) (2013) 527–530.

- [41] E. Sayag, A. Leshem, N.D. Sidiropoulos, Finite word length effects on transmission rate in zero forcing linear precoding for multichannel DSL, *IEEE Trans. Signal Process.* 57 (4) (2009) 1469–1482.
- [42] J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, G. Ysebaert, Dual decomposition approach to partial crosstalk cancellation in a multiuser DMT-xDSL environment, *EURASIP J. Adv. Signal Process.* 2007 (2007) 1–11.
- [43] P. Pandey, M. Moonen, L. Deneire, MMSE-based partial crosstalk cancellation for upstream VDSL, in: *IEEE International Conference on Communications (ICC)*, 2010.
- [44] J. Louveaux, A. van der Veen, Adaptive precoding for downstream crosstalk precancellation in DSL systems using sign-error feedback, *IEEE Trans. Signal Process.* 58 (6) (2010) 3173–3179.
- [45] Binyaminini, I. Bergel, Adaptive precoder using sign error feedback for FEXT cancellation in multichannel downstream VDSL, *IEEE Trans. Signal Process.* 61 (9) (2013) 2383–2393.
- [46] M. Gray, T.G. Stockham, Dithered quantizers, *IEEE Trans. Inform. Theory* 39 (3) 1993.
- [47] P. Whiting, G. Kramer, C. Nuzman, A. Ashikhmin, A. van Wijngaarden, M. Živković, Analysis of inverse crosstalk channel estimation using SNR feedback, *IEEE Trans. Signal Process.* 59 (2011) 1102–1115.
- [48] R. Cendrillon, G. Ginis, E. Van den Bogaert, M. Moonen, A near-optimal linear crosstalk canceler for upstream VDSL, *IEEE Trans. Signal Process.* 54 (8) (2006) 3136–3146.
- [49] C. Chen, K. Seong, R. Zhang, J. Cioffi, Optimized transmission for upstream vectored DSL systems using zero-forcing generalized decision-feedback equalizers, in: *Proc. IEEE Globecom*, 2006.
- [50] C. Chen, K. Seong, R. Zhang, J. Cioffi, Optimized resource allocation for upstream vectored DSL systems with zero-forcing generalized decision feedback equalizer, *IEEE J. Sel. Top. Signal Process.* 1 (4) (2007) 686–699.
- [51] M. Varanasi, T. Guess, Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the gaussian multiple-access channel, in: *Conference Record of the 31st Asilomar Conference on Signals, Systems and Computers*, vol. 2, 1997, pp. 1405–1409.
- [52] P. Tsiaflakis, J. Vangorp, J. Verlinden, M. Moonen, Multiple access channel optimal spectrum balancing for upstream DSL transmission, *IEEE Commun. Lett.* 11 (4) (2007) 398–300.
- [53] A. Leshem, E. Zehavi, Rate control for PSD limited multiple access systems through linear programming, in: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3220–3223.
- [54] V. Chan, W. Yu, Joint multiuser detection and optimal spectrum balancing for digital subscriber lines, *EURASIP J. Appl. Signal Process.* 2006 (2006) 1–13.
- [55] A. Chowdhery, J. Cioffi, Dynamic spectrum management for upstream mixtures of vectored and non-vectored DSL, in: *IEEE Global Telecommunications Conference (GLOBECOM)*.

Distributed Detection and Estimation in Wireless Sensor Networks*

7

Sergio Barbarossa, Stefania Sardellitti, and Paolo Di Lorenzo

Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy

2.07.1 Introduction

Wireless sensor networks (WSN) are receiving a lot of attention from both the theoretical and application sides, in view of the many applications spanning from environmental monitoring, as a tool to control physical parameters such as temperature, vibration, pressure, or pollutant concentration, to the monitoring of civil infrastructures, such as roads, bridges, buildings, etc. [1]. Some new areas of applications are emerging rapidly and have great potentials. A field that is gaining more and more interest is the use of WSNs as a support for *smart grids*. In such a case, a WSN is useful to: (i) monitor and predict energy production from renewable sources of energy such as wind or solar energy; (ii) monitor energy consumption; (iii) detect anomalies in the network. A further area of increasing interest is *vehicular sensor networks*. In such a case, the vehicles are nodes of an ad hoc network. The sensors onboard the vehicle can measure speed and position of the vehicle and forward this information to nearby vehicles or to the road side units (RSU). This information enables the construction of dynamic spatial traffic maps, which can be exploited to reroute traffic in case of accidents or to minimize energy consumption. A relatively recent and interesting application of WSNs is *cognitive radio (CR)*. In such a case, opportunistic (or secondary) users are allowed to access temporally unoccupied spectrum holes, under the constraint of not interfering with licensed (primary) users, and to release the channels as soon as they are requested by licensed users. The basic step enabling this dynamic access is sensing. The problem is that if sensing is carried out at a single location, it might be severely degraded by shadowing phenomena: If the sensor is in shadowed area, it might miss the presence of a primary user and then transmit by mistake over occupied slots, thus generating an undue interference. To overcome shadowing, it is useful to resort to a WSN whose nodes sense the channels and exchange information with each other in order to mitigate the effect of local shadowing phenomena. The goal of the WSN in such an application is to build a spatial map of channel occupancy. An opportunistic user willing to access radio resources within a confined region could then interrogate the closest sensor of a WSN and get a reliable information about which channels are temporarily available and when this utilization has to be stopped. The plethora of applications raises a series of challenging technical issues, which may be seen as sources of opportunities for engineers. Probably the first most important question concerns *energy supply*. In many applications, in

*This work has been partially supported by the European project TROPIC, nr. ICT-318784.

fact, the sensors are battery-operated and it may be difficult or costly to recharge the batteries or to substitute them. As a consequence, energy consumption is a basic constraint that should be properly taken into account. A second major concern is *reliability* of the whole system. In many cases, to allow for an economy of scale, the single sensors are devices with limited accuracy and computational capabilities. Nevertheless, the decision taken by the network as a whole must be very reliable, because it might affect crucial issues like security, safety, etc. The question is then how to build a reliable system out of the combination of many potentially unreliable nodes. Nature exhibits many examples of such systems. Human beings are capable of solving very sophisticated tasks and yet they are essentially built around basic unreliable chemical reactions occurring within cells whose lifetime is typically much smaller than the lifetime of a human being. Clearly, engineering is still far away from approaching the skills of living systems, but important inspirations can be gained by observing biological systems. Two particular features possessed by biological systems are self-organization and self-healing capabilities. Introducing these capabilities within a sensor network is the way to tackle the problem of building a reliable system out of the cooperation of many potentially unreliable units. In particular, self-organization is a key tool to enable the network to reconfigure itself, in terms of acquisition and transfer of information from the sensing nodes to the control centers, responsible for taking decisions, launching alarms or activating actuators aimed to counteract adverse phenomena. The network architecture plays a fundamental role in terms of reliability of the whole system. In conventional WSNs, there is typically one or a few sink nodes that collect the observations taken by the sensor nodes and process them, in a centralized fashion, to produce the desired decision about the observed phenomenon. This architecture arises a number of critical issues, such as: (a) potential congestion around the sink nodes; (b) vulnerability of the whole network to attacks or failure of sink nodes; (c) efficiency of the communication links established to send data from the sensor nodes to the sink. For all these reasons, a desirable characteristic of a WSN is to be designed in such a way that decisions are taken in a decentralized manner. Ideally, every node should be able, in principle, to achieve the final decision, thanks to the exchange of information with the other nodes, either directly or through multiple hops. In this way, vulnerability would be strongly reduced and the system would satisfy a scalability property. In practice, it is not necessary to make every single node to be able to take decisions as reliably as in a centralized system. But what is important to emphasize is that proper interaction among the nodes may help to improve reliability of single nodes, reduce vulnerability and congestion events, and make a better usage of radio resource capabilities. This last issue points indeed to one of the distinctive features of decentralized decision systems, namely the fact that *sensing and communicating are strictly intertwined* with each other and a proper system design must consider them jointly. The first important constraint inducing a strict link between sensing and communicating is that the transmission of the measurements collected by the nodes to the decision points occurs over realistic channels, utilizing standard communication protocols. For example, adopting common digital communication systems, the data gathered by the sensors need to be quantized and encoded before transmission. In principle, the number of bits used in each sensor should depend on the accuracy of the data acquisition on that sensor. At the same time, the number of bits transmitted per each channel use is upper bounded by the channel capacity, which depends on the transmit power and on the channel between sensor and sink node. This suggests that the number of bits to be used in each node for data quantization should be made dependent on both sensor accuracy and transmission channel. A further important consequence of the network architecture and of the resulting flow of information from peripheral sensing nodes to central decision nodes is the latency with which a global decision can be

taken. In a centralized decision system, the flow of information proceeds from the sensing nodes to the central control nodes, usually through multiple hops. The control node collects all the data, it carries out the computations, and takes a decision. Conversely, in a decentralized decision system, there is typically an iterated exchange of data among the nodes. This determines an increase of the time necessary to reach a decision. Furthermore, an iterated exchange of data implies an iterated energy consumption. Since in WSNs energy consumption is a fundamental concern, all the means to minimize the overall energy consumption necessary to reach a decision within a maximum latency are welcome. At a very fundamental level, we will see how an efficient design of the network requires a global cross layer design where the physical and the routing layers take explicitly into account the specific application for which the network has been built.

This chapter is organized as follows. In Section 2.07.2, we provide a general framework aimed to show how an efficient design of a sensor network requires a joint organization of in-network processing and communication. We show how the organization of the flow of information from the sensing nodes to the decision centers should depend not only on the WSN topology, but also on the statistical model of the observation. Finally, we briefly recall some fundamental information theoretical issues showing how in a multi-terminal decision network source and channel coding are strictly related to each other. In Section 2.07.3, we introduce the graph model as the formal tool to describe the interaction among the nodes. Then, we illustrate the so called consensus algorithm as a basic tool to reach globally optimal decisions through a decentralized approach. Since the interaction among the nodes occurs through a wireless channel, we also consider the impact of realistic channel models on consensus algorithm and show how consensus algorithms can be made robust against channel impairments. In Section 2.07.4, we address the distributed estimation problem. We show first an entirely decentralized approach, where observations and estimations are performed without the intervention of a fusion center. In such a case, we show how to achieve a globally optimal estimation through the local exchange of information among nearby nodes. Then, we consider the case where the estimation is performed at a decision center. In such a case, we show how to allocate quantization bits and transmit powers in the links between the sensing nodes and the fusion center, in order to accommodate the requirement on the maximum estimation variance, under a constraint on the global transmit power. In Section 2.07.5, we extend the approach to the detection problem. Also in this case, we consider the entirely distributed approach, where every node is enabled to achieve a globally optimal decision, and the case where the decision is taken at a central control node. In such a case, we show how to allocate coding bits and transmit power in order to maximize the detection probability, under constraints on the false alarm rate and the global transmit power. Then, in Section 2.07.6, we generalize consensus algorithms illustrating a distributed procedure that does not force all the nodes to reach a common value, as in consensus algorithms, but rather to converge to the projection of the overall observation vector onto a signal subspace. This algorithm is especially useful, for example, when it is required to smooth out the effect of noise, but without destroying valuable information present in the spatial variation of the useful signal. In wireless sensor networks, a special concern is energy consumption. We address this issue in Section 2.07.7, where we show how to optimize the network topology in order to minimize the energy necessary to achieve a global consensus. We show how to convert this, in principle, combinatorial problem, into a convex problem with minimal performance losses. Finally, in Section 2.07.8, we address the problem of matching the topology of the observation network to the graph describing the statistical dependencies among the observed variables. Finally, in Section 2.07.9, we draw some conclusions and we try to highlight some open problems and possible future developments.

2.07.2 General framework

The distinguishing feature of a decentralized detection or estimation system is that the measurements are gathered by a multiplicity of sensors dispersed over space, while the decision about what is being sensed is taken at one or a few fusion centers or sink nodes. The information gathered by the sensors has then to propagate from the peripheral nodes to the central control nodes. The challenge coming from this set-up is that in a WSN, information propagates through wireless channels, which are inherently broadcast, affected by fading and prone to interference. Installing a WSN requires then to set up a proper medium access control protocol (MAC) able to handle the communications among the nodes, in order to avoid interference and to ensure that the information reaches the final destination in a reliable manner. But what is decidedly specific of a WSN is that the sensing and communication aspects are strictly related to each other. In designing the MAC of a WSN, there are some fundamental aspects that distinguish a WSN from a typical telecommunication (TLC) network. The main difference stems from the analysis of goal and constraints of these two kinds of networks. A TLC network must make sure that every source packet reaches the final destination, perhaps through retransmission in case of errors or packet drop, irrespective of the packet content. In a WSN, what is really important is that the decision about what is being sensed be taken in the most reliable way, without necessarily implying the successful delivery of all source packets. Moreover, one of the major constraints in WSNs is energy consumption, because the nodes are typically battery operated and recharging the batteries is sometimes troublesome, especially when the nodes are installed in hard to reach places. Conversely, in a TLC network, energy provision is of course important, but it is not the central issue. At the same time, the trend in TLC networks is to support higher and higher data rates to accommodate for ever more demanding applications, while the data rates typically required in most WSNs are not so high. These considerations suggest that an efficient design of a WSN should take into account the application layer directly. This means, for example, that it is not really necessary that every packet sent by a sensor node reaches the final destination. What is important is only that the correct decision is taken in a reliable manner, possibly with low latency and low energy consumption. This enables data aggregation or in-network processing to avoid unnecessary data transmissions. It is then important to formulate this change of perspective in a formal way to envisage ad hoc information transmission and processing techniques.

2.07.2.1 Computing while communicating

In a very general setting, taking a decision based on the data collected by the sensors can be interpreted as computing a function of these data. Let us denote by x_i , with $i = 1, \dots, N$, the measurements collected by the i th node of the network, and by $f(\mathbf{x}) = f(x_1, \dots, x_N)$ the function to be computed. The straightforward approach for computing this function consists in sending all the measurements x_i to a fusion center through a proper communication network and then implement the computation of $f(\mathbf{x})$ at the fusion center. However, if $f(\mathbf{x})$ possesses a structure, it may be possible to take advantage of such a structure to better organize the flow of data from the sensing nodes to the fusion center. The idea of mingling computations and communications to make an efficient use of the radio resources, depending on the properties that the function $f(\mathbf{x})$ might possess, was proposed in [2]. Here, we will first recall the main results of [2]. Then, we will show how the interplay between computation and communication will be further affected by the structure of the probabilistic model underlying the observations.

To exploit the structure of the function $f(x_1, \dots, x_N)$ to be computed, it is necessary to define some relevant structural properties. One important property is *divisibility*. Let \mathcal{C} be a subset of $\{1, 2, \dots, N\}$ and let $\pi := \{C_1, \dots, C_s\}$ be a partition of \mathcal{C} . We denote by \mathbf{x}_{C_i} the vector composed by the set of measurements collected by the nodes whose indices belong to C_i . A function $f(x_1, \dots, x_N)$ is said to be divisible if, for any $\mathcal{C} \subset \{1, 2, \dots, N\}$ and any partition π , there exists a function $g^{(\pi)}$ such that

$$f(\mathbf{x}_C) = g^{(\pi)}(f(\mathbf{x}_{C_1}), f(\mathbf{x}_{C_2}), \dots, f(\mathbf{x}_{C_s})). \quad (7.1)$$

In words, (7.1) represents a sort of “divide and conquer” property: A function $f(\mathbf{x})$ is divisible if it is possible to split its computation into partial computations over subsets of data and then recombine the partial results to yield the desired outcome.

Let us suppose now that the N sensing nodes are randomly distributed over a circle of radius R . We assume a simple propagation model, such that two nodes are able to send information to each other in a reliable way if their distance is less than a coverage radius $r_0(N)$. At the same time, the interference between two links is considered negligible if the interfering transmitter is at a distance greater than $\alpha r_0(N)$ from the receiver, where α is chosen according to the propagation model. For any random deployment of the nodes, the choice of $r_0(N)$ induces a network topology, such that there is a link between two nodes if their distance is less than $r_0(N)$. The resulting graph having the nodes as vertices and the edges as links, is a random graph, because the positions of the nodes are random. This kind of graph is known as a *Random Geometric Graph* (RGG).¹ To make an efficient use of the radio resources, it is useful to take $r_0(N)$ as small as possible, to save local transmit power and make possible the reuse of radio resources, either frequency or time slots. However, $r_0(N)$ should not be too small to loose connectivity. In other words, we do not want the network to split in subnetworks that do not interact with each other. Since the node location is random, network connectivity can only be guaranteed in probability. It has been proved in [3] that, if $r_0(N)$ is chosen as follows:

$$r_0(N) = R \sqrt{\frac{\log N + c(N)}{\pi N}} \quad (7.2)$$

with $c(N)$ going to infinity, as N goes to infinity, the resulting RGG is asymptotically connected with high probability, as N goes to infinity. For instance, if we take $c(N) = (\pi - 1) \log N$, the coverage radius can be expressed simply as

$$r_0(N) = R \sqrt{\frac{\log N}{N}}. \quad (7.3)$$

A further property of a node is the number of neighbors of that node. For an undirected graph, the number of neighbors of a node is known as the degree of the node. Denoting by $d(N)$ the degree of an RGG with N nodes, it was proved in [4] that, choosing the coverage radius as in (7.2), $d(N)$ is (asymptotically) upper bounded by a function that behaves as $\log N$. More specifically,

$$\lim_{N \rightarrow \infty} \mathbb{P}\{d(N) \leq c \log N\} = 1. \quad (7.4)$$

¹A basic review of graph properties is reported in Appendix A.

In [4] it was established an interesting link between the properties of the function $f(\mathbf{x})$ to be computed by the network and the topology of the communication network. In particular, assuming as usual that the measurements are quantized in order to produce a value belonging to a finite alphabet, let us denote by $\mathcal{R}(f, N)$ the range of $f(\mathbf{x})$ and by $|\mathcal{R}(f, N)|$ the cardinality of $\mathcal{R}(f, N)$. In [4], it was proved that, under the following assumptions:

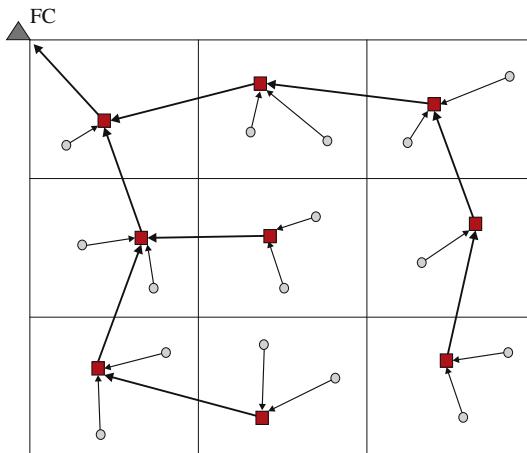
- A.1.** $f(\mathbf{x})$ is divisible;
- A.2.** the network is connected;
- A.3.** the degree of each node is chosen as $d(N) \leq k_1 \log |\mathcal{R}(f, N)|$;

then, the rate for computing $f(\mathbf{x})$ scales with N as

$$R(N) \geq \frac{c_1}{\log |\mathcal{R}(f, N)|}. \quad (7.5)$$

This is an important result that has practical consequences. It states, in fact, that, whenever $\log |\mathcal{R}(f, N)|$ scales with a law that increases more slowly than N , we can have an increase of efficiency if we organize the local computation and the flow of partial results properly. For instance, if the sensors communicate to the sink node through a Time Division Multiplexing Access (TDMA) scheme, with a standard approach it is necessary to allocate N time slots to send all the data to the sink node. Conversely, Eq. (7.5) suggests that, to compute the function $f(\mathbf{x})$, it is sufficient to allocate $\log |\mathcal{R}(f, N)|/c_1$ slots. The same result would apply in a Frequency Division Multiplexing Access (FDMA) scheme, simply reverting the role of time slots and frequency subchannels. This is indeed a paradigm shift, because it suggests that an efficient radio resource allocation in a WSN should depend on the cardinality of $\mathcal{R}(f, N)$. This implies a sort of cross-layer approach that involves physical, MAC and application layers jointly. The next question is how to devise an access protocol that enables such an efficient design. To this regard, the theorem proved in [4] contains a constructive proof, which suggests how to organize the flow of information from the sensing nodes to the control center. In particular, the strategy consists in making a tessellation of the area monitored by the sensor network, similarly to a cellular network, as pictorially described in Figure 7.1. Furthermore, the information flows from the peripheral nodes to the fusion center through a tree-like graph, having the fusion center as its root. In each cell, the nodes (circles) identify a node as the relay node (square). The relay node collects data from the nodes within its own cell and from relay nodes of its leaves, performs local computations and communicates the result to the parent relay nodes, with the goal of propagating these partial results towards the root (sink node). To handle interference, a graph coloring scheme is used to avoid interference among adjacent cells. This allows spatial reuse of radio resources, e.g., frequency or time slots, which can be used in parallel without generating an appreciable interference. The communication structure is conceptually similar to a cellular network, with the important difference that now the flow of information is directly related to the computational task. A few examples are useful to better grasp the possibilities of this approach.

Data uploading: Suppose it is necessary to convey all the data to the sink node. If each observed vector belongs to an alphabet \mathcal{X} , with cardinality $|\mathcal{X}|$, the cardinality of the whole data set is $|\mathcal{R}(f, N)| = |\mathcal{X}|^N$. Hence, $\log |\mathcal{R}(f, N)| = N \log |\mathcal{X}|$. This means that, according to (7.5), the capacity of the network scales as $1/N$. This is a rather disappointing result, as it shows that there is no real benefit with respect

**FIGURE 7.1**

Hierarchical organization of information flow from peripheral nodes to fusion center.

to the simplest communication case one could envisage: The nodes have to split the available bandwidth into a number of sub-bands equal to the number of nodes, with a consequent rate reduction per node.

Decision based on the histogram of the measurements: Let us suppose now that the decision to be taken at the control node can be based on the histogram of the data collected by the nodes, with no information loss. In this case, the function $f(\mathbf{x})$ is the histogram. It can be verified that the histogram is a divisible function. Furthermore, the cardinality of the histogram is

$$|\mathcal{R}(f, N)| = \binom{N + |\mathcal{X}| - 1}{|\mathcal{X}| - 1}. \quad (7.6)$$

Furthermore, it can be shown that

$$(N/|\mathcal{X}|)^{|\mathcal{X}|} \leq \binom{N + |\mathcal{X}| - 1}{|\mathcal{X}| - 1} \leq (N + 1)^{|\mathcal{X}|}. \quad (7.7)$$

Hence, in this case $\log |\mathcal{R}(f, N)|$ behaves as $\log N$ and then the rate $R(N)$ in (7.5) scales as $1/\log N$. This is indeed an interesting result, showing that if the decision can be based on the histogram of the data, rather than on each single measurement, adopting the right communication scheme, the rate per node behaves as $1/\log N$, rather than $1/N$, with a rate gain $N/\log N$, which increases as the number of nodes increases.

Symmetric functions: Let us consider now the case where $f(\mathbf{x})$ is a symmetric function. We recall that a function $f(\mathbf{x})$ is symmetric if it is invariant to permutations of its arguments, i.e., $f(\mathbf{x}) = f(\Pi\mathbf{x})$ for any permutation matrix Π and any argument vector \mathbf{x} . This property reflects the so called *data-centric* view, where what is important is the measurement *per se*, and not which node has taken which measurement.

Examples of symmetric functions include the mean, median, maximum/minimum, histogram, and so on. The key property of symmetric functions is that it can be shown that they depend on the argument \mathbf{x} only through the histogram of \mathbf{x} . Hence, the computation of symmetric functions is a particular case of the example examined before. Thus, the rate scales again as $1/\log N$.

2.07.2.2 Impact of observation model

Having recalled that the efficient design of a WSN requires an information flow that depends on the scope of the network, more specifically, on the structural properties of the function to be computed by the network, it is now time to be more specific on the decision tasks that are typical of WSNs, namely detection and estimation. Let us consider for example the simple hypothesis testing problem. In such a case, an ideal centralized detector having error-free access to the measurements collected by the nodes, should compute the likelihood ratio and compare it with a suitable threshold [5]. We denote with \mathcal{H}_0 and \mathcal{H}_1 the two alternative hypotheses, i.e., absence or presence of the event of interest, and with \mathbf{x}_i the set of measurements collected by node i . If we indicate with $p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_i)$ the joint probability density function of the whole set of observed data, under the hypothesis \mathcal{H}_i , with $i = 0, 1$, the likelihood ratio test amounts to comparing the likelihood ratio (LR) with a threshold γ , and decide for \mathcal{H}_1 if the threshold is exceeded or for \mathcal{H}_0 , otherwise. In formulas

$$\Lambda(\mathbf{x}) := \Lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_1)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \gamma. \quad (7.8)$$

The LR test (LRT) is optimal under a Bayes or a Neyman-Pearson criterion, the only difference being that the threshold γ assumes different values in the two cases [5]. In principle, to implement the LRT at the fusion center, every node should send its observation vector \mathbf{x}_i to the fusion center, through a proper MAC protocol. The fusion center, after having collected all the data, should then implement the LRT, as indicated in (7.8). However, the computation of the LR in (7.8) does not necessarily imply the transmission of the single vectors \mathbf{x}_i . Conversely, according to the theory recalled above, the transmission strategy should depend on the structural properties of the LR function, if any. Let us see how to exploit the structure of the LR function in two cases of practical interest.

2.07.2.2.1 Statistically independent observations

Let us start assuming that the observations taken by different sensors are statistically independent, conditioned to each hypothesis. This is an assumption valid in many cases. Under such an assumption, the LR can be factorized as follows

$$\Lambda(\mathbf{x}) := \frac{\prod_{n=1}^N p(\mathbf{x}_n; \mathcal{H}_1)}{\prod_{n=1}^N p(\mathbf{x}_n; \mathcal{H}_0)} := \prod_{n=1}^N \Lambda_n(\mathbf{x}_n) \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \gamma, \quad (7.9)$$

where $\Lambda_n(\mathbf{x}_n) = p(\mathbf{x}_n; \mathcal{H}_1)/p(\mathbf{x}_n; \mathcal{H}_0)$ denotes the local LR at the n th node. In this case, the global function $\Lambda(\mathbf{x})$ in (7.9) possesses a clear structure: It is factorizable in the product of the local LR functions. Then, since a factorizable function is divisible, it is possible to implement the efficient mechanisms described in the previous section to achieve an efficient design. The network nodes should

cluster as in Figure 7.1. Every relay node should compute the local LR, multiply it to the data received from the relays pertaining to the lower clusters and send the partial result to the relay of the upper cluster, until the result reaches the fusion center. The efficiency comes from the fact that many transmissions can occur in parallel, exploiting spatial reuse of radio resources. This result suggests also that the proper source encoding to be implemented at each sensor node consists in the computation of the local LR.

2.07.2.2.2 Markov observations

The previous result is appealing, but it pertains to the simple situation where the observations are statistically independent, conditioned to the hypotheses. In some circumstances, however, this assumption is unjustified. This is the case, for example, when the sensors monitor a field of spatially correlated values, like a temperature or atmospheric pressure field. In such cases, nearby nodes sense correlated values and then the statistical independence assumption is no longer valid. It is then of interest, in such cases, to check whether the statistical properties of the observations can still induce a structure on the function to be computed that can be exploited to improve network efficiency.

There is indeed a broad class of observation models where the joint pdf cannot be factorized into the product of the individual pdf's pertaining to each node, but it can still be factorized into functions of subsets of variables. This is the case of Bayes networks or Markov random fields. Here we will recall the basic properties of these models, as relevant to our problem. The interested reader can refer to many excellent books, like, for example, [6] or [7].

In the Bayes network's case, the statistical dependency among the random variables is described by an acyclic directed graph, whose vertices represent the random variables, while the edges represent local conditional probabilities. In particular, given a node x_i , whose parent nodes are identified by the set of indices $\text{pa}(i)$, the joint probability density function (pdf) of a Bayes network can be written as

$$p(x_1, \dots, x_N) = \prod_{i=1}^N p(x_i / \mathbf{x}_{\text{pa}(i)}), \quad (7.10)$$

where $\mathbf{x}_{\text{pa}(i)}$ collects all the variables corresponding to the parents of node i . If a node in (7.10) does not have parents, the corresponding probability is unconditional.

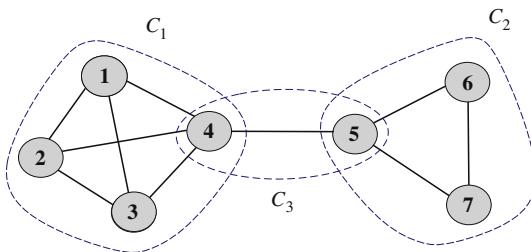
Alternatively, a Markov random field is represented through an undirected graph. More specifically, a Markov network consists of:

1. An undirected graph $G = (V, E)$, where each vertex $v \in V$ represents a random variable and each edge $\{u, v\} \in E$ represents statistical dependency between the random variables u and v ;
2. A set of potential (or compatibility) functions $\psi_c(\mathbf{x}_c)$ (also called clique potentials), that associate a non-negative number to the cliques² of G .

Let us denote by \mathcal{C} the set of all cliques present in the graph. The random vector \mathbf{x} is Markovian if its joint pdf admits the following factorization

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{c \in \mathcal{C}} \psi_c(\mathbf{x}_c), \quad (7.11)$$

²A clique is a subset of nodes which are fully connected and maximal, i.e., no additional node can be added to the subset so that the subset remains fully connected.

**FIGURE 7.2**

Example of Markov graph.

where \mathbf{x}_c denotes the vector of variables belonging to the clique c . The functions $\psi_c(\mathbf{x}_c)$ are called *compatibility functions*. The term Z is simply a normalization factor necessary to guarantee that $p(\mathbf{x})$ is a valid pdf. A node p is conditionally independent of another node q in the Markov network, given some set S of nodes, if every path from p to q passes through a node in S . Hence, representing a set of random variables by drawing the correspondent Markov graph is a meaningful pictorial way to identify the conditional dependencies occurring across the random variables. As an example, let us consider the graph reported in Figure 7.2. The graph represents conditional independencies among seven random variables. The variables are grouped into 3 cliques. In this case, for example, we can say that nodes 1–4 are statistically independent of nodes 6 and 7, conditioned to the knowledge of node 5. In this example, the joint pdf can be written as follows

$$p(\mathbf{x}) = \frac{1}{Z} \psi_1(x_1, x_2, x_3, x_4) \psi_2(x_5, x_6, x_7) \psi_3(x_4, x_5). \quad (7.12)$$

If the product in (7.11) is strictly positive for any \mathbf{x} , we can introduce the functions

$$V_c(\mathbf{x}_c) = -\log \psi_c(\mathbf{x}_c) \quad (7.13)$$

so that (7.11) can be rewritten in exponential form as

$$p(\mathbf{x}) = \frac{1}{Z} \exp \left(-\sum_{c \in \mathcal{C}} V_c(\mathbf{x}_c) \right). \quad (7.14)$$

This distribution is known, in physics, as the Gibbs (or Boltzman) distribution with interaction *potentials* $V_c(\mathbf{x}_c)$ and *energy* $\sum_{c \in \mathcal{C}} V_c(\mathbf{x}_c)$.

The independence graph conveys the key probabilistic information through absent edges: If nodes i and j are not neighbors, the random variables x_i and x_j are statistically independent, conditioned to the other variables. This is the so called *pairwise Markov property*. Given a subset $a \subset V$ of vertices, $p(\mathbf{x})$ factorizes as

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{c: c \cap a \neq \emptyset} \psi_c(\mathbf{x}_c) \prod_{c: c \cap a = \emptyset} \psi_c(\mathbf{x}_c), \quad (7.15)$$

where the second factor does not depend on a . As a consequence, denoting by $S - a$ the set of all nodes except the nodes in a and by \mathcal{N}_a the set of neighbors of the nodes in a , $p(\mathbf{x}_a / \mathbf{x}_{S-a})$ reduces to $p(\mathbf{x}_a / \mathcal{N}_a)$. Furthermore,

$$p(\mathbf{x}_a / \mathcal{N}_a) = \frac{1}{Z_a} \prod_{c:c \cap a \neq \emptyset} \psi_c(\mathbf{x}_c) = \frac{1}{Z_a} \exp \left(- \sum_{c:c \cap a \neq \emptyset} V_c(\mathbf{x}_c) \right). \quad (7.16)$$

This property states that the joint pdf factorizes in terms that contain only variables whose vertices are neighbors.

An important example of jointly Markov random variables is the Gaussian Markov Random Field (GMRF), characterized by having a pdf expressed as in (7.14), with the additional property that the energy function is a quadratic function of the variables. In particular, a vector \mathbf{x} of random variables is a GMRF if its joint pdf can be written as

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x}-\boldsymbol{\mu})} = \sqrt{\frac{|\mathbf{A}|}{(2\pi)^N}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{A}(\mathbf{x}-\boldsymbol{\mu})}, \quad (7.17)$$

where $\boldsymbol{\mu} = \mathbb{E}\{\mathbf{x}\}$ is the expected value of \mathbf{x} , $\mathbf{C} = \mathbb{E}\{(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T\}$ is the covariance matrix of \mathbf{x} and $\mathbf{A} = \mathbf{C}^{-1}$ is the so called *precision* matrix. In this case, the *Markovianity* of \mathbf{x} manifests itself through the *sparsity* of the precision matrix. As a particular case of (7.16), the coefficient a_{ij} of \mathbf{A} is different from zero if and only if nodes i and j are neighbors.

Having recalled the main properties of GMRF's, let us now go back to the problem of organizing the flow of information in a WSN aimed at deciding between two alternative hypotheses of GMRF. Let us consider for example the decision about the two alternative hypotheses:

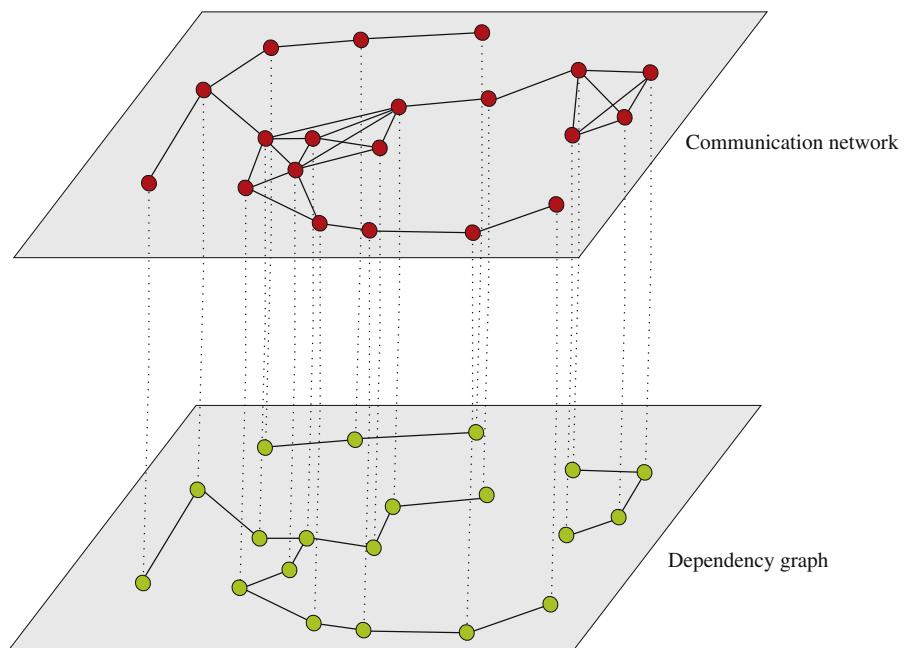
$$\mathcal{H}_0 : \mathbf{x} \sim p(\mathbf{x}; \mathcal{H}_0) = \frac{1}{Z_0} \prod_{c \in \mathcal{C}} \psi_c(\mathbf{x}_c; \mathcal{H}_0), \quad (7.18)$$

$$\mathcal{H}_1 : \mathbf{x} \sim p(\mathbf{x}; \mathcal{H}_1) = \frac{1}{Z_1} \prod_{c' \in \mathcal{C}'} \psi_{c'}(\mathbf{x}_{c'}; \mathcal{H}_1), \quad (7.19)$$

where the sets of cliques involved in the two cases are, in general, different. The factorizations in (7.18, 7.19) suggest how to implement the computation of the LRT:

1. Each cluster in the WSN should be composed of the nodes associated to the random variables pertaining to the same clique in the statistical dependency graph;
2. The observations gathered by the nodes pertaining to a clique c are locally encoded into the clique potential $\psi_c(\mathbf{x}_c; \mathcal{H}_i)$. This is the value that has to be transmitted by each cluster towards upper layers or to the FC;
3. As in Figure 7.1, each relay in the lowest layer compute the local potentials and forward these results to the upper layers. The relays of the intermediate clusters receive the partial results from the lower clusters, multiply these values by the local potential and forward the results to the relay of the upper cluster, until reaching the FC.

In general, different grouping may occur depending on the hypothesis. This organization represents a generalization of the distributed computation observed in the conditionally independent case, where the groups are simply singletons, i.e., sets composed by exactly one element. In that case, the clustering among nodes is only instrumental to the communication purposes, i.e., to enable spatial reuse of radio resources. In the more general Markovian case, the organization of the communication network in clusters (cells) should take into account, *jointly*, the grouping suggested by the cliques of the underlying dependency graph and the spatial grouping of nodes to enable concurrent transmission over the same radio resources without incurring in undesired interference. To visualize this general perspective, it is useful to have in mind two superimposed graphs, as depicted in Figure 7.3: the communication graph (top), whose vertices are the network nodes while the edges are the radio links; the dependency graph (bottom), whose vertices represent random variables, while the arcs represent statistical dependencies. Each communication cluster should incorporate at least one clique. Furthermore, in each cluster there is a relay node that is responsible for the exchange of data with nearby clusters. The whole communication network has a hierarchical tree-structure. Each node in the tree is a relay node belonging to a cluster. This node collects the measurements from the nodes belonging to its cluster, computes the potential (or the product of potentials if more cliques belong to the same cluster) and forwards this value to its relay parents. While we have depicted the two graphs as superimposed in Figure 7.3, it is useful to clarify that the nodes of the communication network are located in space and their relative position is well defined in a

**FIGURE 7.3**

Superposition of communication layer (top) over a Markov statistical dependency graph (bottom).

metric space. Conversely, the nodes of the Markov graph represent random variables for which there is no well defined notion of distance or, even if we define one, it is a notion that in general does not have a correspondence with distance in space. In other words, while the neighborhood of nodes in the top graph has to do with the concept of spatial distance among the nodes, the neighborhood of the nodes in the Markov graph has only to do with statistical dependencies. Nevertheless, it is also true that in the observation of physical entities like a temperature field, for example, it is reasonable to expect higher correlation among nearby (in the spatial sense) nodes (variables). An example of GMRF where the statistical dependencies incorporate the spatial distances was suggested in [8]. In summary, the previous considerations suggest that an efficient design of the communication network topology should keep into account the structure, if any, of dependency graph describing the observed variables. At the same time, the design of the network topology should keep into account physical constraints like the power consumption necessary to maintain the links with sufficient reliability (i.e., to insure the sufficient signal-to-noise ratio at the receiver). This is indeed an interesting line of research: How to match the network topology to the dependency graph, under physical constraints dictated by energy consumption, delay, etc. Some works have already addressed this issue. For example, in [9] the authors addressed the problem of implementing data fusion policies with minimal energy consumption, assuming a Markov random field observation model, and established the scaling laws for optimal and suboptimal fusion policies. An efficient message-passing algorithm taking into account the communication network constraints was recently proposed in [10].

2.07.2.3 Fundamental information-theoretical issues

In this section, we recall very briefly some of the fundamental information-theoretic limits of multi-terminal decision networks. We will not go into the details of this challenging fundamental problem. The interested reader can refer to [11] and the references therein. In a WSN, each sensor is observing a physical phenomenon, which can be regarded as a source of information, and the goal of the network is to take decisions about what is being sensed. In some cases, the decision is taken by a fusion center; in others, the decision is distributed across the nodes. In general, the data gathered by the nodes has to travel through realistic channels, prone to additive noise, channel fading and interference. This requires source and channel coding. In a point-to-point communication, when there is only one sensor transmitting data to the fusion center, the encoding of the data gathered by the sensor follows well known rules. In particular, the observation is first time-sampled and each sample is encoded in a finite number, let us say R , of bits per symbol. This converts an analog source of information into a digital source. In this analog-to-digital (AD) conversion, there is usually a distortion that can be properly quantified. More precisely, the source coding rate R depends on the constraint on the mean-square distortion level D . At the same time, given a constraint on the power budget (cost) P available at the transmit side, the maximum rate that can be transmitted with arbitrarily low error probability is the channel capacity $C(P)$, which depends on the transmit power constraint. A rate-distortion pair (D, P) is achievable if and only if

$$R(D) \leq C(P). \quad (7.20)$$

The source-channel coding separation theorem [12] states that the encoding operation necessary to transmit information through a noisy channel can be split, without loss of optimality, into the cascade of two successive *independent* operations: (i) source coding, where each symbol emitted by the source is encoded in a finite number of bits per symbol; (ii) channel coding, where a string of k bits are encoded into a codeword of length n bits, to make the codeword error probability arbitrarily low. This theorem has

been a milestone in digital communications, as it allows system designers to concentrate, separately, on source coding and channel coding techniques, with no loss of optimality. However, when we move from the point-to-point link to the multipoint-to-multipoint case, there is no equivalent of the source-channel coding separation theorem. This means that in the multi-terminal setting, splitting coding into source and channel coding does not come without a cost, anymore. Rephrasing the source/channel coding theorem in the multi-terminal context, denoting by $\mathcal{R}(D)$ the *rate region*, comprising all the source codes that satisfy the distortion constraint D , and by $\mathcal{C}(P)$ the *capacity region*, containing all the transmission rates satisfying the transmit power constraint P , a pair (D, P) is achievable if

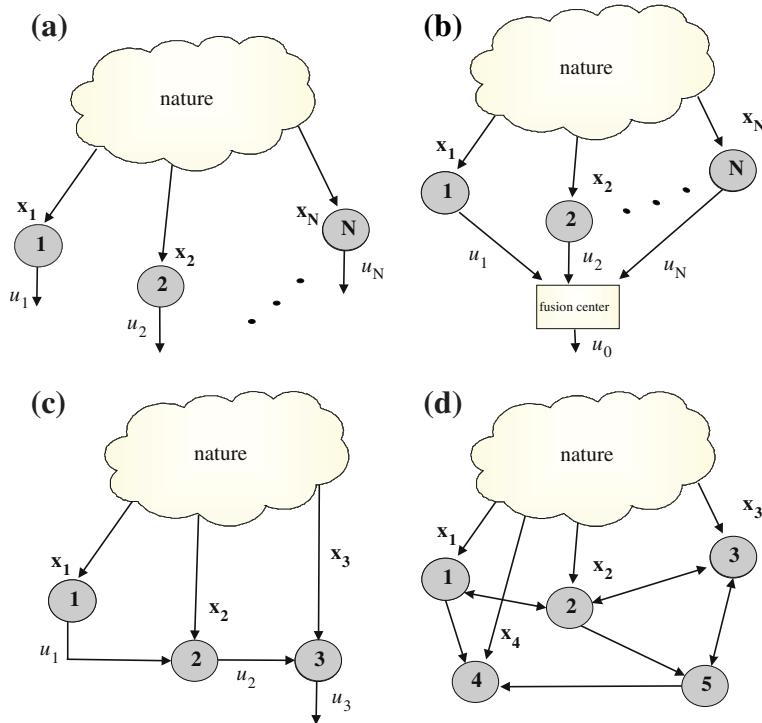
$$\mathcal{R}(D) \cap \mathcal{C}(P) \neq \emptyset. \quad (7.21)$$

However, Eq. (7.21) is no longer a necessary condition, meaning that there may exist a code that achieves the prescribed distortion D at a power cost P , which cannot be split into a source compression encoder *followed* by a channel encoder. In general, in the multiterminal case, a *joint* source/channel encoding is necessary. This suggests, from a fundamental theoretical perspective, that, again, in a distributed WSN local processing and communication have to be considered jointly.

2.07.2.4 Possible architectures

Alternative networks architectures may be envisaged depending on how the nodes take decision and exchange information with each other. A few examples are shown in Figure 7.4 where there is a set of N nodes observing a given phenomenon, denoted as “nature” for simplicity. The measurements made by node i are collected into the vector \mathbf{x}_i , with $i = 1, \dots, N$. In Figure 7.4a, each node takes an individual decision, which is represented by the variable $u_i : u_i = 1$ if node i decides for the presence of the event, otherwise $u_i = 0$. More generally, u_i could also be the result of a local source encoder, whose aim is to reduce the redundancy present in the observed data. The simplest case is sketched in Figure 7.4a, where a set of nodes observes a state of nature and each node takes a decision. Even if this is certainly the simplest form of monitoring, if the local decisions are taken according to a global optimality criterion, even in the case of statistically independent observations, the local decisions are coupled in a non trivial form. The next step, in terms of complexity, is to combine all the observations collected by the sensing nodes in a centralized node, called fusion center or sink node. This strategy is depicted in the architecture of Figure 7.4b. In such a case, each node takes a local decision and sends this information to the fusion center, which combines the local decision according to a globally optimum criterion. What is important, in a practical setting, is that the limitations occurring in the transmission of information from the sensing nodes to the fusion center are properly taken into account. An alternative approach is reported in Figure 7.4c, where node 1 takes a local decision and it notifies node 2 about this decision. Node 2, on its turn, based on the decision of node 1 and on its own measurements as well, takes a second decision, and so on. A further generalization occurs in the example of Figure 7.4d, where the nodes take local decisions and exchange information with the other nodes. In such a case, there is no fusion center and the final decision can be taken, in principle, by every node.

Besides the architecture describing the flow of information through the network, a key aspect concerns the constraint imposed by the communication links. Realistic channels are in fact affected by noise, fading, delays, and so on. Hence, a globally optimal design must incorporate the decision and

**FIGURE 7.4**

Alternative communication architectures between peripheral nodes and fusion center.

communication aspects jointly in a common context. The first step in this global design passes through a formal description of the interaction among the nodes.

2.07.3 Graphical models and consensus algorithm

The proper way to describe the interactions among the network nodes is to introduce the graph model of the network. Let us consider a network composed of N sensors. The flow of information across the sensing nodes implementing some form of distributed computation can be properly described by introducing a graph model whose vertices are the sensors and there is an edge between two nodes if they exchange information with each other.³ Let us denote the graph as $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ where \mathcal{V} denotes the set of N vertices (nodes) v_i and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges $e_{ij}(v_i, v_j)$. The most powerful tool to grasp the properties of a graph is *algebraic graph theory* [13], which is based on the description of the graph through appropriate matrices, whose definition we recall here below. Let $\mathbf{A} \in \mathbb{R}^{N \times N}$ be the *adjacency matrix* of the graph \mathcal{G} , whose elements a_{ij} represent the weights associated to each edge with $a_{ij} > 0$ if $e_{ij} \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. According to this notation and assuming no self-loops, i.e., $a_{ii} = 0$, $\forall i = 1, \dots, N$,

³We refer the reader to Appendix 2.07.9 for a review of the basic notations and properties of graphs.

the out-degree of node v_i is defined as $\deg_{\text{out}}(v_i) = \sum_{j=1}^N a_{ji}$. Similarly, the in-degree of node v_i is $\deg_{\text{in}}(v_i) = \sum_{j=1}^N a_{ij}$. The *degree matrix* \mathbf{D} is defined as the diagonal matrix whose i th diagonal entry is $d_{ii} = \deg_{\text{in}}(v_i)$. Let \mathcal{N}_i denote the set of neighbors of node i , so that $|\mathcal{N}_i| = \deg(v_i)$.⁴ The *Laplacian* matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$ of the graph \mathcal{G} is defined as $\mathbf{L} := \mathbf{D} - \mathbf{A}$. Some properties of the Laplacian will be extensively used in our distributed algorithms to be presented later on and then it is useful to recall them.

Properties of the Laplacian matrix

- P.1.** \mathbf{L} has, by construction, a null eigenvalue with associated eigenvector the vector $\mathbf{1}$ composed by all ones. This property can be easily checked verifying that $\mathbf{L}\mathbf{1} = \mathbf{0}$ since by construction, $\sum_{j=1}^N a_{ij} = d_{ii}$
- P.2.** The multiplicity of the null eigenvalue is equal to the number of connected components of the graph. Hence, the null eigenvalue is *simple* (it has multiplicity one) if and only if the graph is connected.
- P.3.** If we associate a state variable x_i to each node of the graph, if the graph is undirected, the disagreement between the values assumed by the variables is a quadratic form built on the Laplacian [13]:

$$J(\mathbf{x}) := \frac{1}{4} \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} a_{ij} (x_i - x_j)^2 = \frac{1}{2} \mathbf{x}^T \mathbf{L} \mathbf{x}, \quad (7.22)$$

where $\mathbf{x} = [x_1, \dots, x_N]^T$ denotes the network state vector.

2.07.3.1 Consensus algorithm

Given a set of measurements $x_i(0)$, for $i = 1, \dots, N$, collected by the network nodes, the goal of consensus algorithm is to minimize the disagreement among the nodes. This can be useful, for example, when the nodes are measuring some common variable and their measurement is affected by error. The scope of the interaction among the nodes is to reduce the effect of errors on the final estimate. In fact, consensus is one of fundamental tools to design distributed decision algorithms that satisfy a global optimality principle, as corroborated by many works on distributed optimization, see, e.g., [14–20]. We recall now the consensus algorithm as this will form the basis of the distributed estimation and detection algorithms developed in the ensuing sections. Let us consider, for simplicity, the case where the nodes are measuring a temperature and the goal is to find the average temperature. In this case, reaching a consensus over the average temperature can be seen as the minimization of the disagreement, as defined in (7.22), between the states $x_i(0)$ associated to the nodes.

The minimization of the disagreement can be obtained by using a simple gradient-descent algorithm. More specifically, using a continuous-time system, the minimum of (7.22) can be achieved by running the following dynamical system [15]

$$\dot{\mathbf{x}}(t) = -\mathbf{L}\mathbf{x}(t), \quad (7.23)$$

initialized with $\mathbf{x}(0) = \mathbf{x}_0$, where \mathbf{x}_0 is the vector containing all the initial measurements collected by the network nodes. This means that the state of each node evolves in time according to the first order differential equation

$$\dot{x}_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij} (x_j(t) - x_i(t)), \quad (7.24)$$

⁴By $|\cdot|$ we denote the cardinality of the set.

where \mathcal{N}_i indicates the set of neighbors of node i . Hence, every node updates its own state only by interacting with its neighbors.

Equation (7.23) assumes the form of a diffusion equation. Let us consider for example the evolution of a diffusing physical quantity $\psi(z; t)$ as a function of the spatial variable z and of time t ($\psi(z; t)$ could represent, for instance, the heat distribution), the diffusion equation assumes the form

$$\frac{\partial \psi(z; t)}{\partial t} = D \frac{\partial^2 \psi(z; t)}{\partial z^2}, \quad (7.25)$$

where D is the diffusion coefficient. If we discretize the space variable and approximate the second order derivative with a discrete-time second order difference, the diffusion Eq. (7.25) can be written as in (7.23), where the Laplacian matrix represents the discrete version of the Laplacian operator. This conceptual link between consensus equation and diffusion equation has been exploited in [21] to derive a fast consensus algorithm, mimicking the effect of advection. The interesting result derived in [21] is that to speed up the consensus (diffusion) process, it is necessary to use a *directed* graph, with time-varying adjacency matrix coefficients a_{ij} .

The solution of (7.23) is given by

$$\mathbf{x}(t) = \exp(-\mathbf{L}t)\mathbf{x}(0). \quad (7.26)$$

In the case analyzed so far, since the consensus algorithm has been deduced from the minimization of the disagreement and the disagreement has been defined for undirected graphs, the matrix \mathbf{L} is symmetric. Hence, its eigenvalues are real. The convergence of (7.26) is guaranteed because all the eigenvalues of \mathbf{L} are non-negative, by construction. If the graph is connected, according to property P.2, the eigenvalue zero has multiplicity one. Furthermore, the eigenvector associated to the zero eigenvalue is the vector $\mathbf{1}$. Hence, the system (7.23) converges to the consensus state:

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \frac{1}{N} \mathbf{1} \mathbf{1}^T \mathbf{x}(0). \quad (7.27)$$

This means that every node converges to the average value of the measurements collected by the whole network, i.e.,

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{N} \sum_{i=1}^N x_i(0) = x^*. \quad (7.28)$$

The convergence rate of system (7.24) is lower bounded by the slowest decaying mode of the dynamical system (7.23), i.e., by the second smallest eigenvalue of \mathbf{L} , $\lambda_2(\mathbf{L})$, also known as the *algebraic connectivity* of the graph [22]. More specifically, if the graph is connected or, equivalently, if $\lambda_2(\mathbf{L}) > 0$, then the dynamical system (7.23) converges to consensus exponentially [15], i.e., $\|\mathbf{x}(t) - x^* \mathbf{1}\| \leq \|\mathbf{x}(0) - x^* \mathbf{1}\| O(e^{-rt})$ with $r = \lambda_2(\mathbf{L})$.

In some applications, the nodes are required to converge to a *weighted* consensus, rather than average consensus. This can be achieved with a slight modification of the consensus algorithm. If we premultiply the left side of (7.24) by a positive coefficient c_i , the resulting equation

$$c_i \dot{x}_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij} (x_j(t) - x_i(t)) \quad (7.29)$$

converges to the weighted average

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{\sum_{i=1}^N c_i x_i(0)}{\sum_{i=1}^N c_i}. \quad (7.30)$$

This property will be used in deriving distributed estimation mechanisms in the next section.

Alternatively, the minimization of (7.22) can be achieved in discrete-time through the following iterative algorithm

$$\mathbf{x}[k+1] = \mathbf{x}[k] - \epsilon \mathbf{L} \mathbf{x}[k] := \mathbf{W} \mathbf{x}[k], \quad (7.31)$$

where we have introduced the so called *transition* matrix $\mathbf{W} = \mathbf{I} - \epsilon \mathbf{L}$. Also in this case, the discrete time equation is initialized with the measurements taken by the sensor nodes at time 0, i.e., $\mathbf{x}[0] := \mathbf{x}_0$. This time, to guarantee convergence of the system (7.31), we need to choose the coefficient ϵ properly. More specifically, the discrete time Eq. (7.31) converges if the eigenvalues of \mathbf{W} are bounded between -1 and 1 . This can be seen very easily considering that reiterating (7.31) k times, we get

$$\mathbf{x}[k] = \mathbf{W}^k \mathbf{x}[0]. \quad (7.32)$$

Let us denote by \mathbf{u}_k the eigenvectors of \mathbf{W} associated to the eigenvalues $\lambda_k(\mathbf{W})$, with $k = 1, \dots, N$. The eigenvalues of \mathbf{W} are real and we consider them ordered in increasing sense, so that $\lambda_N(\mathbf{W}) \geq \lambda_{N-1}(\mathbf{W}) \geq \dots \geq \lambda_1(\mathbf{W})$. Hence, the evolution of system (7.32) can be written as

$$\mathbf{x}[k] = \sum_{n=1}^N \lambda_n^k(\mathbf{W}) \mathbf{u}_n \mathbf{u}_n^T \mathbf{x}[0]. \quad (7.33)$$

The matrix \mathbf{W} has an eigenvector equal to $\mathbf{1}/\sqrt{N}$, associated to the eigenvalue 1 by construction. In fact, $\mathbf{W}\mathbf{1}/\sqrt{N} = \mathbf{1}/\sqrt{N} - \epsilon \mathbf{L}\mathbf{1}/\sqrt{N} = \mathbf{1}/\sqrt{N}$. If the graph is connected, the eigenvalue 1 of \mathbf{W} has multiplicity one. Furthermore, if ϵ is chosen such that $\epsilon(\mathbf{L}) < 2/\lambda_N(\mathbf{L})$, all other eigenvalues are less than 1. Hence, for a connected graph, the system (7.33) converges to

$$\lim_{k \rightarrow \infty} \mathbf{x}[k] = \frac{1}{N} \mathbf{1} \mathbf{1}^T \mathbf{x}[0]. \quad (7.34)$$

Again, this corresponds to having every node converging to the average consensus.

The consensus algorithm can be extended to the case of directed graphs. This case is indeed much richer of possibilities than the undirected case, because the consensus value ends up to depend more strictly on the graph topology. In the directed case, in fact, \mathbf{L} is an asymmetric matrix. The most important difference is that the graph connectivity turns out to depend on the orientation of the edges. Furthermore, each eigenvalue of \mathbf{L} gives rise to a pair of left and right eigenvectors which do not coincide with each other. These differences affect the final consensus state and induce different forms of consensus, as shown below.

The convergence of the system in (7.31) can be proved by exploiting the properties of non-negative matrices. A nonnegative matrix is row (or column) stochastic if all its row (or column) sums are equal to one. Furthermore if the graph associated to the network is strongly connected, i.e., the zero eigenvalue associated to \mathbf{L} has multiplicity one (see Appendix 2.07.9), \mathbf{W} is called an *irreducible* matrix. An irreducible stochastic matrix is primitive if it has only one eigenvalue with maximum modulus. Primitive nonnegative matrices, often named *Perron matrices*, satisfy the Perron-Frobenius theorem [23].

Theorem 1. Let γ_l and γ_r , respectively, the left and right eigenvectors associated to the unit eigenvalue of the primitive nonnegative matrix W , i.e., $W\gamma_r = \gamma_r$ and $\gamma_l^T W = \gamma_l^T$ with $\gamma_r^T \gamma_l = 1$, then $\lim_{k \rightarrow \infty} W^k = \gamma_r \gamma_l^T$.

Let us now apply to a sensor network modeled by the graph \mathcal{G} , with adjacency matrix A , the distributed consensus algorithm

$$x_i[k+1] = x_i[k] - \epsilon \sum_{j \in \mathcal{N}_i} a_{ij} (x_i[k] - x_j[k]) \quad (7.35)$$

with $0 < \epsilon < 1/d_{\max}$.

Interestingly, different forms of consensus can be achieved in a directed graph, depending on the graph connectivity properties [17]:

- a. If the graph is strongly connected, the dynamical system in (7.35) converges to a weighted consensus, for any initial state vector $x[0]$, i.e.,

$$\lim_{k \rightarrow \infty} W^k x[0] = x^* = \mathbf{1} \gamma_l^T x[0], \quad (7.36)$$

where $\gamma_l(i) > 0, \forall i$, and $\sum_{i=1}^N \gamma_l(i) = 1$. In this case, since the graph is strongly connected, W is an irreducible matrix. Then, applying Gershgorin theorem [23], it can be deduced that there exists a single eigenvalue $\mu_1(W) = 1$ with maximum modulus. Then W is a primitive nonnegative matrix and from Theorem 1 the convergence in (7.36) is straightforward. In this case, every node contributes to the final consensus value. Furthermore, the consensus value is a weighted combination of the initial observations, where the weights are the entries of the left eigenvector associated to the null eigenvalue of L (or the unit eigenvalue of W).

- b. If the digraph is strongly connected and balanced, i.e., $\mathbf{1}^T L = \mathbf{0}$ and $L\mathbf{1} = \mathbf{0}$, the systems achieves an average consensus or $x^* = \frac{\mathbf{1}\mathbf{1}^T}{N} x[0]$. In fact, for balanced graphs, W is a double stochastic matrix with $\gamma_l = \gamma_r = \mathbf{1}/\sqrt{N}$;
- c. If the digraph \mathcal{G} is weakly connected (WC), but not strongly connected, and it contains a forest with K strongly connected root components, the graph splits in K disjoint clusters $\mathcal{C}_1, \dots, \mathcal{C}_K \subseteq \{1, \dots, N\}$,⁵ and all the nodes pertaining to each cluster converge to the consensus values

$$x_q^* = \frac{\sum_{i \in \mathcal{C}_k} \gamma_i x_i[0]}{\sum_{i \in \mathcal{C}_k} \gamma_i}, \quad \forall q \in \mathcal{C}_k, \quad k = 1, \dots, K. \quad (7.37)$$

In words, there is no single consensus, in this case, but there is a local consensus within each cluster. Different clusters typically converge to different consensus values.

- d. If the digraph \mathcal{G} is composed of a single spanning tree, every node converges to the value assumed by the root node.

As far as the convergence rate, instead, in [15] it has been shown for undirected connected graphs that the dynamical system in (7.31) converges exponentially to the average consensus with a rate at least equal

⁵In general, the clusters $\mathcal{C}_1, \dots, \mathcal{C}_K$ are not a partition of the set of nodes $\{1, \dots, N\}$.

to $\mu_2(\mathbf{W}) = 1 - \epsilon\lambda_2(\mathbf{L})$ where $\mu_2(\mathbf{W})$ is the second largest eigenvalue of the Perron matrix \mathbf{W} . In fact by defining the disagreement vector $\boldsymbol{\delta} = \mathbf{x} - \mathbf{x}^*$, it can be easily verified [15] that $\boldsymbol{\delta}$ evolves according to the disagreement dynamic given by $\boldsymbol{\delta}[k+1] = \mathbf{W}\boldsymbol{\delta}[k]$. Hence $\psi[k] := \boldsymbol{\delta}[k]^T\boldsymbol{\delta}[k]$ represents a candidate Lyapunov function for the disagreement dynamics so that

$$\psi[k+1] = \boldsymbol{\delta}[k+1]^T\boldsymbol{\delta}[k+1] = \|\mathbf{W}\boldsymbol{\delta}[k]\|^2 \leq \mu_2(\mathbf{W})^2\|\boldsymbol{\delta}[k]\|^2 = \mu_2(\mathbf{W})^2\psi[k] \quad (7.38)$$

with $0 < \mu_2(\mathbf{W}) < 1$ since \mathbf{W} is a symmetric and primitive matrix. As a consequence the algorithm converges exponentially to consensus with *a rate at least equal to $\mu_2(\mathbf{W})$* .

2.07.3.2 Consensus algorithms over realistic channels

So far, we have recalled the basic properties of consensus algorithm assuming that the exchange of information across the nodes occurs with no errors. In this section we study what happens to consensus algorithms when the communications among the nodes are affected by quantization errors, noise, packet drops, etc. The problem of consensus protocols affected by stochastic disturbance has been considered in a series of previous papers [24–29]. In [24], the authors use a decreasing sequence of weights to prove the convergence of consensus protocols to an agreement space in the presence of additive noise under a fixed network topology. The works in [25, 26] consider consensus algorithms in the presence of link failures, which are modeled as i.i.d. Laplacian matrices of a directed graph. The papers present necessary and sufficient conditions for consensus exploiting the ergodicity of products of stochastic matrices. A distributed consensus algorithm in which the nodes utilize probabilistically quantized information to communicate with each other was proposed in [27]. As a result, the expected value of the consensus is equal to the average of the original sensor data. A stochastic approximation approach was followed in [28], which considered a stochastic consensus problem in a strongly connected directed graph where each agent has noisy measurements of its neighboring states. Finally, the study of a consensus protocol that is affected by both additive channel noise and a random topology was considered in [29]. The resulting algorithm relates to controlled Markov processes and the convergence analysis relies on stochastic approximation techniques.

In the study of consensus mechanisms over realistic channels, we consider the following sources of randomness:

1. **Node positions:** The first randomness is related to the spatial positions of the nodes, which are in general unknown. We model the spatial distribution of nodes as a random geometric graph composed of N nodes. In graph theory, a random geometric graph (RGG) is a random undirected graph drawn on a bounded region, eg. the unit disk, generated by:
 - a. Placing vertices at random uniformly and independently on the region,
 - b. Connecting two vertices, u, v if and only if the distance between them is inside a threshold radius r_0 , i.e., $d(u, v) \leq r_0$.

Several probabilistic results are known about RGG's. In particular, as shown in [3], if N nodes are placed in a disc of unit area in \mathbb{R}^2 and each node transmits with a power scaling with N as in (7.2), the resulting network is asymptotically connected with probability one, as $N \rightarrow \infty$.

- 2. Random link failures model:** In a realistic communication scenario, the packets exchanged among sensors may be received with errors, because of channel fading or noise. The retransmission of erroneous packets can be incorporated into the system, but packet retransmission introduces a nontrivial additional complexity in decentralized implementations and, most important, it introduces an unknown delay and delay jitter. It is then of interest to examine simple protocols where erroneous packets are simply dropped. Random packet dropping can be taken into account by modeling the coefficient a_{ij} describing the network topology as random variables that assume the value 1 or 0, if the packet is correctly delivered or not, respectively. In this case, the Laplacian varies with time as a sequence of i.i.d. matrices $\{\mathbf{L}[k]\}$, which can be written, without any loss of generality, as

$$\mathbf{L}[k] = \bar{\mathbf{L}} + \tilde{\mathbf{L}}[k], \quad (7.39)$$

where $\bar{\mathbf{L}}$ denotes the mean matrix and $\tilde{\mathbf{L}}[k]$ are i.i.d. perturbations around the mean. We do not make any assumptions about the link failure model. Although the link failures and the Laplacians are independent over time, during the same iteration, the link failures can still be spatially correlated. It is important to remark that we do not require the random instantiations $G[k]$ of the graph be connected for all k . We only require the graph to be connected on average. This condition is captured by requiring $\lambda_2(\bar{\mathbf{L}}) > 0$.

- 3. Dithered quantization:** We assume that each node encodes the message to be exchanged with the other nodes using a uniform quantizer, with a finite number of bits n_b , defined by the following vector mapping, $\mathbf{q}(\cdot) : \mathbb{R}^L \rightarrow Q^L$,

$$\mathbf{q}(\mathbf{y}) = [b_1\Delta, \dots, b_L\Delta]^T = \mathbf{y} + \mathbf{e}_q(\mathbf{y}), \quad (7.40)$$

where the entries of the vector \mathbf{y} , the quantization step $\Delta > 0$, and the error \mathbf{e}_q satisfy

$$(b_m - 1/2)\Delta \leq y_m \leq (b_m + 1/2)\Delta, \quad 1 \leq m \leq L, \quad (7.41)$$

$$-\Delta/2\mathbf{1}_L \leq \mathbf{e}_q(\mathbf{y}) \leq \Delta/2\mathbf{1}_L, \quad \text{for all } \mathbf{y}. \quad (7.42)$$

The quantization alphabet is

$$Q^L = \{[b_1\Delta, \dots, b_L\Delta]^T | b_m \in \mathbb{Z}, \forall m\}. \quad (7.43)$$

Conditioned on the input, the quantization error $\mathbf{e}_q(\mathbf{y})$ is deterministic. This induces a correlation among the quantization errors resulting at different nodes and different times, which may affect the convergence properties of the distributed algorithm. To avoid undesired error correlations, we introduce dithering, as in [30,31]. In particular, the dither added to randomize the quantization effects satisfies a special condition, namely the Schuchman conditions, as in subtractively dithered systems, [32]. Then, at every time instant k , adding to each component $y_m[k]$ a dither sequence $\{d_m[k]\}_{k \geq 0}$ of i.i.d. uniformly distributed random variables on $[-\Delta/2, \Delta/2]$ independent of the input sequence, the resultant error sequence $\{e_m[k]\}_{k \geq 0}$ becomes

$$e_m[k] = q(y_m[k] + d_m[k]) - (y_m[k] + d_m[k]). \quad (7.44)$$

The sequence $\{e_m[k]\}_{k \geq 0}$ is now an i.i.d. sequence of uniformly distributed random variables on $[-\Delta/2, \Delta/2]$, which is independent of the input sequence.

The convergence of consensus algorithm in the presence of random disturbance can be proved by exploiting results from supermartingale theory [33]. In an ideal communication case, by selecting the step-size of the algorithm to be sufficiently small (smaller than $2/\lambda_N(\mathbf{L})$, where $\lambda_N(\mathbf{L})$ is the maximum eigenvalue of the Laplacian matrix of the graph), the discrete-time consensus algorithm will asymptotically converge to the agreement subspace. However, in a realistic communication scenario, the links among the sensors may fail randomly and the exchanged data is corrupted by quantization noise. Under these nonideal conditions, the consensus algorithm needs to be properly adjusted to guarantee convergence. A discrete time consensus algorithm that accounts for random link failures and dithered quantization noise can be written as:

$$\mathbf{x}_i[k+1] = \mathbf{x}_i[k] + \alpha[k] \sum_{j=1}^N a_{ij}[k] (\mathbf{q}(\mathbf{x}_j[k] + \mathbf{d}_{ij}[k]) - \mathbf{x}_i[k]), \quad i = 1, \dots, N, \quad (7.45)$$

where $\alpha[k]$ is a positive iteration dependent step-size, and $\mathbf{d}_{ij}[k]$ is the dithered quantization vector. Now, exploiting the feature of subtractively dithered systems in (7.44), the previous expression can be recast as:

$$\mathbf{x}_i[k+1] = \mathbf{x}_i[k] + \alpha[k] \sum_{j=1}^N a_{ij}[k] (\mathbf{x}_j[k] - \mathbf{x}_i[k] + \mathbf{d}_{ij}[k] + \mathbf{e}_{ij}[k]), \quad i = 1, \dots, N. \quad (7.46)$$

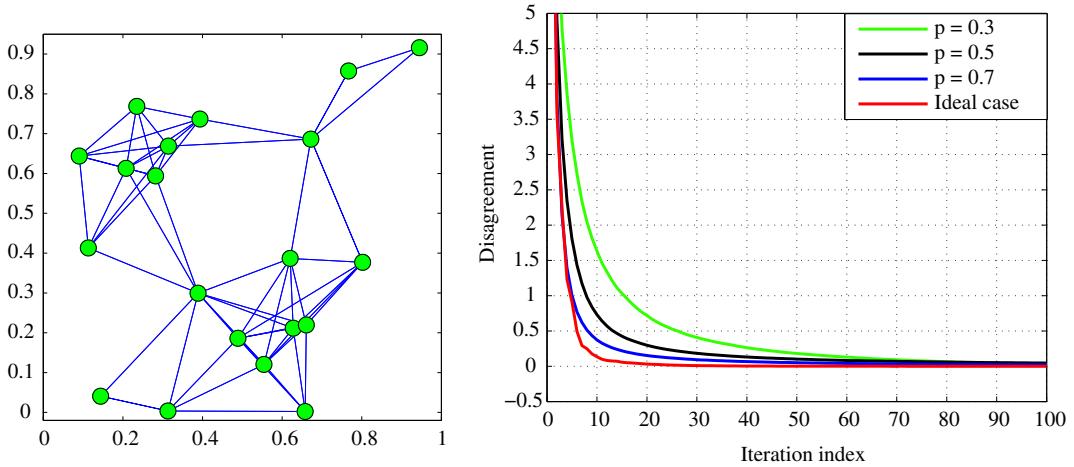
Starting from some initial value, $\mathbf{x}_i[0] \in \mathbb{R}^L$, each node generates via (7.46) a sequence of state variables, $\{\mathbf{x}_i[k]\}_{k \geq 0}$. The value $\mathbf{x}_i[k+1]$ at the i th node at time $k+1$ is a function of: its previous state $\mathbf{x}_i[k]$ and the quantized states correctly received at time k by the neighboring sensors. As described previously, the data are subtractively dithered-quantized, so that the quantized data received by the i th sensor from the j th sensor at time k is $\mathbf{q}(\mathbf{x}_j[k] + \mathbf{d}_{ij}[k])$. It then follows that the quantization error $\mathbf{e}_{ij}[k]$ is a random vector, whose components are i.i.d., uniformly distributed on $[-\Delta/2, \Delta/2]$, and independent of $\mathbf{x}_j[k]$.

One way to guarantee convergence of the previous system is to use a positive iteration-dependent step size $\alpha[k]$ satisfying [24, 29]

$$\lim_{k \rightarrow \infty} \alpha[k] = 0, \quad \sum_{k=0}^{\infty} \alpha[k] = \infty, \quad \sum_{k=0}^{\infty} \alpha^2[k] < \infty. \quad (7.47)$$

Exploiting results from stochastic approximation theory, this choice drives the noise variance to zero while guaranteeing the convergence to the consensus subspace.

A numerical example is useful to show the robustness of consensus algorithm in the presence of link failures and quantization noise. We consider a connected network composed of 20 nodes as depicted on the left side of Figure 7.5. The initial value of the state variable at each node is randomly chosen in the interval $[0, 1]$. At the k th iteration of the updating rule (7.45), each node communicates to its neighbors its current state, i.e., a scalar $x_i[k]$. Because of fading and additive noise, a communication link among two neighbors has a certain probability p to be established correctly. The values to be exchanged are (dither) quantized with 6 bits. The iteration-dependent step size is chosen as $\alpha[k] = \alpha_0/k$, with $\alpha_0 = 1.5/\lambda_N(\mathbf{L})$, in order to satisfy (7.47). The right side of Figure 7.5 shows the average behavior of the disagreement among the sensors in the network, versus the iteration index, for different values of the probability p to establish a communication link correctly. The result is averaged over 100 independent realizations. The ideal case corresponds to $p = 1$ and it is shown as a benchmark. As we can notice

**FIGURE 7.5**

Network (left). Disagreement vs. time index (right), for different probabilities of correct packet reception.

from the right side of Figure 7.5, even in the presence of random disturbances, an agreement is always reached by the network for any value of p . The only effect of the random link failures is to slow down the convergence process, without altering the final value of the global potential function. This proves the robustness of the algorithm.

2.07.4 Distributed estimation

Having introduced all the tools necessary to study distributed estimation and detection mechanisms, let us now start with the estimation problem. This problem has been the subject of an extensive literature, see, e.g., [34–43]. Most of the algorithms proposed in these works propose a mix of local estimation and consensus among neighbor nodes to improve upon the performance of the local estimators. In a first class of methods, like [35,36] for example, the nodes collect all the data first, perform local estimation and then interact iteratively with their neighbors. In alternative methods, the nodes keep interacting with each other while collecting new measurements or, in general, receiving new information, like in [37,38,42]. These two classes of methods can be seen as assigning different time scales to the local estimation and consensus steps. Indeed, it can be proved that a proper combination of local estimation and consensus can bring the whole network to a globally optimal estimate, provided that the graph describing the interaction among the nodes is connected. This approach was pursued, for example in [35], where the so called bridge nodes fulfilled the scope of enforcing local consensus. In the following, we will show how alternative formulations of the globally optimal estimation problem naturally lead to a different mix of the local estimation and the consensus steps, without the need to introduce any node having a special role nor enforcing different time scales a priori.

Let us denote with $\boldsymbol{\theta} \in \mathbb{R}^M$ the parameter vector to be estimated. In some cases, there is no prior information about $\boldsymbol{\theta}$. In other cases, $\boldsymbol{\theta}$ is known to belong to a given set \mathcal{C} : For instance, its entries are known

to be positive or to belong to a finite interval of known limits, and so on. In some applications, $\boldsymbol{\theta}$ may be the outcome of a random variable described by a known pdf $p_{\Theta}(\boldsymbol{\theta})$. Let us denote by \mathbf{x}_i the measurement vector collected by node i and by $\mathbf{x} := [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T$ the whole set of data collected by all the nodes. In the two cases of interest, the estimation can be obtained as the solution of the following problems:

Arbitrary case

$$\max_{\boldsymbol{\theta}} p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta}) \quad (7.48)$$

$$\text{s.t. } \boldsymbol{\theta} \in \mathcal{C} \quad (7.49)$$

where $p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta})$ is the joint pdf of vector \mathbf{x} , for a given arbitrary vector $\boldsymbol{\theta}$, or

Random case

$$\max_{\boldsymbol{\theta}} p_{X/\Theta}(\mathbf{x}/\boldsymbol{\theta}) p_{\Theta}(\boldsymbol{\theta}), \quad (7.50)$$

where $p_{\Theta}(\boldsymbol{\theta})$ is (known) prior pdf of the parameter vector and $p_{X/\Theta}(\mathbf{x}/\boldsymbol{\theta})$ is the pdf of \mathbf{x} conditioned to $\boldsymbol{\theta}$.

In general, it is not necessary to reconstruct the whole joint pdf $p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta})$ (or $p_{X/\Theta}(\mathbf{x}/\boldsymbol{\theta})$) to obtain the optimal estimate. Let us consider, for example, the case where the pdf can be factorized as

$$p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta}) = g[\mathbf{T}(\mathbf{x}), \boldsymbol{\theta}] h(\mathbf{x}), \quad (7.51)$$

where $g(\cdot, \cdot)$ depends on \mathbf{x} only through $\mathbf{T}(\mathbf{x})$, whereas $h(\cdot)$ does not depend on $\boldsymbol{\theta}$. The function $\mathbf{T}(\mathbf{x})$ is called a *sufficient statistic* for $\boldsymbol{\theta}$ [44]. In general, the sufficient statistic $\mathbf{T}(\mathbf{x})$ is a vector, as it may be constituted by a set of functions.

If (7.51) holds true, all is necessary to estimate $\boldsymbol{\theta}$ is not really $p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta})$, but only $g[\mathbf{T}(\mathbf{x}), \boldsymbol{\theta}]$. This means that any sensor able to evaluate $\mathbf{T}(\mathbf{x})$ through an interaction with the other sensors is able to find out the optimal parameter vector $\boldsymbol{\theta}$ as the vector that maximizes $g[\mathbf{T}(\mathbf{x}), \boldsymbol{\theta}]$.

A simple (yet common) example is given by the so called *exponential family* of pdf

$$p(\mathbf{x}; \boldsymbol{\theta}) = \exp[A(\boldsymbol{\theta})B(\mathbf{x}) + C(\mathbf{x}) + D(\boldsymbol{\theta})]. \quad (7.52)$$

Examples of random variables described by this class include the Gaussian, Rayleigh, and exponential pdf's. Hence, this is a rather common model. Let us assume now that the observations \mathbf{x}_i collected by different nodes are statistically independent and identically distributed (i.i.d.), according to (7.52). It is easy to check, simply applying the definition in (7.51), that a sufficient statistic in such a case is the scalar function:

$$\mathbf{T}(\mathbf{x}) = \sum_{i=1}^N B(\mathbf{x}_i). \quad (7.53)$$

This structure suggests that a simple distributed way to enable every node in the network to estimate the vector $\boldsymbol{\theta}$ locally, without loss of optimality with respect to the centralized approach, is to run a consensus algorithm, where the initial state of every node is set equal to $B(\mathbf{x}_i)$. At convergence, if the network is connected, every node has a state equal to the consensus value, i.e., $T(\mathbf{x})/N$. This enables every node to implement the optimal estimation by simply interacting with its neighbors to achieve a consensus. The only necessary condition for this simple method to work properly is that the network be connected. This is indeed a very simple example illustrating how consensus can be a fundamental step in deriving

an optimal estimation through a purely decentralized approach relying only upon the exchange of data among neighbors.

In the next two sections, we will analyze in more details the purely distributed case (with no fusion center) where the global estimation can be carried out in any node and the centralized case, where the final estimation is taken at the fusion center.

2.07.4.1 Decentralized observations with decentralized estimation

In the following we analyze different observation models and illustrate alternative distributed estimation algorithms. We will start with the conditionally independent case and then we will generalize the approach to a conditionally dependent model.

2.07.4.1.1 Conditionally independent observations

A case amenable for finding distributed solutions is given by the situations where the observations collected by different sensors are conditionally independent. In such a case, the joint pdf $p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta})$ can be factorized as follows

$$p_{X;\Theta}(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^N p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta}), \quad (7.54)$$

where $p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta})$ is the pdf of the vector \mathbf{x}_i observed by node i . Taking the log of this expression, the optimization problem can be cast, equivalently, as

$$\max_{\boldsymbol{\theta}} \sum_{i=1}^N \log p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta}). \quad (7.55)$$

Even if the objective function to be maximized is written as a sum of functions depending each on a local observation vector, the solution of the previous problem still requires a centralized approach because the vector $\boldsymbol{\theta}$ to be estimated is common to all the terms. A possible way to find a distributed solution to the problem in (7.55) consists in introducing an instrumental common variable \mathbf{z} and rewriting the previous problem in the following form

$$\begin{aligned} & \min_{\boldsymbol{\theta}_i} - \sum_{i=1}^N \log p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) \\ & \text{s.t. } \boldsymbol{\theta}_i = \mathbf{z}, \quad i = 1, 2, \dots, N. \end{aligned} \quad (7.56)$$

This is a constrained problem, whose Lagrangian is

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{z}) := \sum_{i=1}^N \left[-\log p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \boldsymbol{\lambda}_i^T (\boldsymbol{\theta}_i - \mathbf{z}) \right], \quad (7.57)$$

where $\boldsymbol{\lambda}_i$ are the vectors whose entries are the Lagrange multipliers associated to the equality constraints in (7.56). In many cases, it is useful to introduce the so called *augmented Lagrangian* [45]:

$$L_\rho(\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{z}) := \sum_{i=1}^N \left[-\log p_{X_i;\Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \boldsymbol{\lambda}_i^T (\boldsymbol{\theta}_i - \mathbf{z}) + \frac{\rho}{2} \|\boldsymbol{\theta}_i - \mathbf{z}\|_2^2 \right], \quad (7.58)$$

where ρ is a *penalty* parameter. Minimizing the augmented Lagrangian leads to the same solution as minimizing the original Lagrangian because any feasible vector satisfying the linear constraint yields a zero penalty. Nevertheless, there are some benefits in working with the augmented Lagrangian, namely: i) the objective function is differentiable under milder conditions than with the original Lagrangian; ii) convergence can be achieved without requiring strict convexity of the objective function (see [45] for more insight into the augmented Lagrangian method).

If the pdf's involved in (7.58) are log-concave functions of θ , the problem in (7.58) is strongly convex and then it admits a unique solution and there are efficient algorithms to compute the solution. Here, we are interested in deriving decentralized solutions.

A possible method to find a distributed solution of the problem in (7.58) is the *alternating direction method of multipliers (ADMM)* [45]. An excellent recent review of ADMM and its applications is [46]. The application of ADMM to distributed estimation problems was proposed in [35]. The method used in [35] relied on the introduction of the so called *bridge* nodes. Here, we will describe methods that do not require the introduction of any special class of nodes (in principle, every node has the same functionality as any other node). This is useful to simplify the estimation method as well as network design and management.

The ADMM algorithm applied to solve (7.58) works through the following steps:

$$\begin{aligned}\boldsymbol{\theta}_i[k+1] &= \arg \min_{\boldsymbol{\theta}_i} \left\{ -\log p_{X_i; \Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \boldsymbol{\lambda}_i^T[k](\boldsymbol{\theta}_i - \mathbf{z}[k]) + \frac{\rho}{2} \|\boldsymbol{\theta}_i - \mathbf{z}[k]\|_2^2 \right\}, \\ \mathbf{z}[k+1] &= \arg \min_{\mathbf{z}} \sum_{i=1}^N \left\{ \boldsymbol{\lambda}_i^T[k](\boldsymbol{\theta}_i[k+1] - \mathbf{z}) + \frac{\rho}{2} \|\boldsymbol{\theta}_i[k+1] - \mathbf{z}\|_2^2 \right\}, \\ \boldsymbol{\lambda}_i[k+1] &= \boldsymbol{\lambda}_i[k] + \rho (\boldsymbol{\theta}_i[k+1] - \mathbf{z}[k+1]).\end{aligned}\tag{7.59}$$

The first and second steps aim at minimizing the primal function (i.e., the augmented Lagrangian) over the unknown variables $\boldsymbol{\theta}$ and \mathbf{z} , for a given value of the Lagrange multipliers' vectors $\boldsymbol{\lambda}_i$, as computed in the previous iteration.

The third step is a dual variable update, whose goal is to maximize the dual function, as in the dual ascent method. We recall that, in our case, the dual function is defined as

$$g(\boldsymbol{\lambda}) = \inf_{\boldsymbol{\theta}, \mathbf{z}} L_\rho(\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{z}).\tag{7.60}$$

In ADMM, the dual ascent step uses a gradient ascent approach to update $\boldsymbol{\lambda}$ in order to maximize $g(\boldsymbol{\lambda})$, for a given value of vectors $\boldsymbol{\theta}_i$ and \mathbf{z} , with the important difference that the step size used to compute the update is exactly the penalty coefficient ρ .

In our case, the second step can be computed in closed form as follows

$$\boldsymbol{\theta}_i[k+1] = \arg \min_{\boldsymbol{\theta}_i} \left\{ -\log p_{X_i; \Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \boldsymbol{\lambda}_i^T[k](\boldsymbol{\theta}_i - \mathbf{z}[k]) + \frac{\rho}{2} \|\boldsymbol{\theta}_i - \mathbf{z}[k]\|_2^2 \right\},\tag{7.61}$$

$$\mathbf{z}[k+1] = \frac{1}{N} \sum_{i=1}^N \left(\boldsymbol{\theta}_i[k+1] + \frac{1}{\rho} \boldsymbol{\lambda}_i[k] \right),\tag{7.62}$$

$$\boldsymbol{\lambda}_i[k+1] = \boldsymbol{\lambda}_i[k] + \rho (\boldsymbol{\theta}_i[k+1] - \mathbf{z}[k+1]).\tag{7.63}$$

From this formulation, we can see that the first and third steps can be run in parallel, over each node. The only step that requires an exchange of values among the nodes is the second step that requires the computation of an average value. But, as we know from previous section, the average value can be computed through a distributed consensus algorithm. The only condition for the convergence of consensus algorithm to the average value is that the graph representing the links among the nodes is connected.

The step in (7.62) can be further simplified as follows. Let us denote with \bar{x} the averaging operation across the nodes, i.e.,

$$\bar{x} := \frac{1}{N} \sum_{i=1}^N x_i. \quad (7.64)$$

Using this notation, the z -update can be written as

$$z[k+1] = \overline{\theta[k+1]} + \frac{1}{\rho} \lambda[k]. \quad (7.65)$$

Similarly, averaging over the λ -update yields

$$\overline{\lambda[k+1]} = \overline{\lambda[k]} + \rho \left(\overline{\theta[k+1]} - z[k+1] \right). \quad (7.66)$$

Substituting (7.65) in (7.66), it is easy to check that, after the first iteration, $\overline{\lambda[k+1]} = 0$. Hence, using $z[k] = \overline{\theta[k]}$, the overall algorithm proceeds as indicated in Table 7.1.

The convergence criterion used in the steps of the algorithm is based on the relative absolute difference at two successive iterations: Given a sequence $y[k]$, the algorithm stops when $\|y[k+1] - y[k]\| / \|y[k]\| \leq \epsilon$, with ϵ a small positive value.

Equations 7.67–7.68 give rise to an interesting interpretation: the primal update (first equation) aims at implementing a local optimization, with a penalty related to the disagreement between the local

Table 7.1 Algorithm A.1

A.1

STEP 1: Set $k = 0$, ϵ equal to a small positive value and initialize $\theta_i[0]$, $\lambda_i[0]$, $\forall i$, and z randomly;

STEP 2: Compute $\theta_i[1]$, $\forall i$ using (7.61);

STEP 3: Run consensus over $\theta_i[1]$ and $\lambda_i[0]$ to get $\overline{\theta[1]}$ and $\overline{\lambda[0]}$;

STEP 4: Set $z[1] = \overline{\theta[1]} + \frac{1}{\rho} \lambda[0]$;

STEP 5: Compute $\lambda_i[1]$, $\forall i$, using (7.63);

STEP 6: Set $k = 1$;

STEP 7: Repeat until convergence

$$\theta_i[k+1] = \arg \min_{\theta_i} \left\{ -\log p_{X_i; \Theta}(\mathbf{x}_i; \theta_i) + \lambda_i^T[k](\theta_i - \overline{\theta[k]}) + \frac{\rho}{2} \|\theta_i - \overline{\theta[k]}\|_2^2 \right\} \quad (7.67)$$

Run consensus over $\theta_i[k+1]$ until convergence;

$$\lambda_i[k+1] = \lambda_i[k] + \rho \left(\theta_i[k+1] - \overline{\theta[k+1]} \right) \quad (7.68)$$

Set $k = k + 1$, if convergence criterion is satisfied stop, otherwise go to step 7.

solution and the global one; the dual update (second equation) aims at driving all the local solutions to converge to a common (consensus) value, which coincides with the globally optimal solution.

The straightforward implementation of (7.67) and (7.68) requires running, at each step k of the ADMM algorithm, a consensus algorithm. A possible alternative approach can be envisaged by reformulating the optimization problem as follows:

$$\begin{aligned} \min_{\boldsymbol{\theta}_i} \quad & \left\{ -\sum_{i=1}^N \log p_{X_i; \Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \boldsymbol{\lambda}_{ij}^T (\boldsymbol{\theta}_j - \boldsymbol{\theta}_i) + \frac{\rho}{2} \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \|\boldsymbol{\theta}_j - \boldsymbol{\theta}_i\|^2 \right\} \\ \text{s.t.} \quad & \boldsymbol{\theta}_j = \boldsymbol{\theta}_i; \forall j \in \mathcal{N}_i; \quad i = 1, 2, \dots, N, \end{aligned} \quad (7.69)$$

where \mathcal{N}_i denotes the set of node i 's neighbors. To make more clear the interaction among the nodes, it is useful to introduce the graph notation, as in previous section. Using the adjacency matrix \mathbf{A} , the previous problem can be rewritten as follows:

$$\begin{aligned} \min_{\boldsymbol{\theta}_i} \quad & \left\{ -\sum_{i=1}^N \log p_{X_i; \Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \sum_{i=1}^N \sum_{j=1}^N a_{ij} \boldsymbol{\lambda}_{ij}^T (\boldsymbol{\theta}_j - \boldsymbol{\theta}_i) + \frac{\rho}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} \|\boldsymbol{\theta}_j - \boldsymbol{\theta}_i\|^2 \right\} \\ \text{s.t.} \quad & \boldsymbol{\theta}_j = \boldsymbol{\theta}_i; \forall j \in \mathcal{N}_i; \quad i = 1, 2, \dots, N. \end{aligned} \quad (7.70)$$

This formulation does not require the introduction of the instrumental variable \mathbf{z} . We keep enforcing the constraint that all the local estimates $\boldsymbol{\theta}_i$ converge to the same value. However, the penalty is now formulated as the disagreement between the local estimates. From consensus algorithm, we know that nulling the disagreement is equivalent to forcing all the vectors $\boldsymbol{\theta}_i$ to reach the same value if the graph describing the interactions among the nodes is connected. Hence, if the network is connected, at convergence, the disagreement goes to zero and there is no bias resulting from the introduction of the disagreement penalty.

The formulation in (7.70) is more amenable for an implementation that does not require, at any step of the algorithm, the convergence of consensus algorithms. In fact, applying ADMM to the solution of (7.70) yields the algorithm described in Table 7.2.

Some examples of applications are useful to grasp the main features of these algorithms.

Table 7.2 Algorithm A.2

A.2

STEP 1: Set $k = 0$, and initialize $\boldsymbol{\theta}_i[0], \boldsymbol{\lambda}_{ij}[0], \forall i, j \in \mathcal{N}_i$;

STEP 2: Repeat until convergence

$$\begin{aligned} \boldsymbol{\theta}_i[k+1] = \arg \min_{\boldsymbol{\theta}_i} \quad & \left\{ -\sum_{i=1}^N \log p_{X_i; \Theta}(\mathbf{x}_i; \boldsymbol{\theta}_i) + \sum_{i=1}^N \sum_{j=1}^N a_{ij} \boldsymbol{\lambda}_{ij}^T[k] (\boldsymbol{\theta}_j[k] - \boldsymbol{\theta}_i) \right. \\ & \left. + \frac{\rho}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} \|\boldsymbol{\theta}_j[k] - \boldsymbol{\theta}_i\|^2 \right\} \end{aligned} \quad (7.71)$$

$$\boldsymbol{\lambda}_{ij}[k+1] = \boldsymbol{\lambda}_{ij}[k] + \rho a_{ij} (\boldsymbol{\theta}_j[k+1] - \boldsymbol{\theta}_i[k+1]) \quad (7.72)$$

Set $k = k + 1$, if convergence criterion is satisfied stop, otherwise go to step 2.

2.07.4.1.2 Distributed ML estimation under Gaussian noise

Let us consider the common situation where the measured vector $\mathbf{x}_i \in \mathbb{R}^Q$ is related to the parameter vector $\boldsymbol{\theta} \in \mathbb{R}^M$, with $Q \geq M$, through a linear observation model, as:

$$\mathbf{x}_i = \mathbf{A}_i \boldsymbol{\theta} + \mathbf{v}_i, \quad i = 1, \dots, N \quad (7.73)$$

where $\mathbf{A}_i \in \mathbb{R}^{Q \times M}$ and \mathbf{v}_i is a vector of jointly Gaussian random variables with zero mean and covariance matrix \mathbf{C}_i , i.e., $\mathbf{v}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_i)$.

In such a case, algorithm A.1 in (7.67) simplifies as the first step of (7.67) can be expressed in closed form

$$\begin{aligned} \boldsymbol{\theta}_i[k+1] &= \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i + \rho \mathbf{I} \right)^{-1} \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i - \boldsymbol{\lambda}_i[k] + \rho \overline{\boldsymbol{\theta}[k]} \right), \\ \boldsymbol{\lambda}_i[k+1] &= \boldsymbol{\lambda}_i[k] + \rho \left(\boldsymbol{\theta}_i[k+1] - \overline{\boldsymbol{\theta}[k+1]} \right). \end{aligned} \quad (7.74)$$

The two updates can be computed in parallel by all the nodes, after having computed the average values through the consensus algorithm.

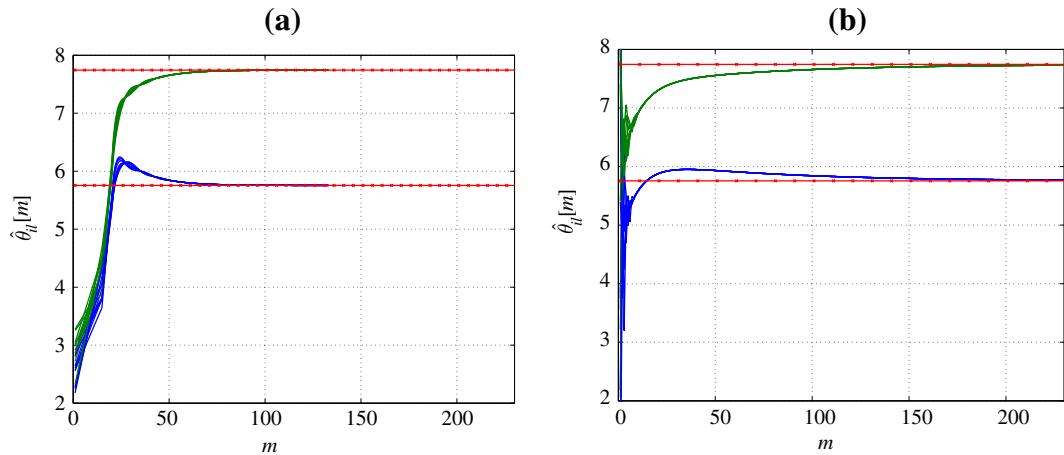
Alternatively, algorithm A.2 becomes

$$\begin{aligned} \boldsymbol{\theta}_i[k+1] &= \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i + 2\rho \sum_{j=1}^N a_{ij} \mathbf{I} \right)^{-1} \\ &\quad \times \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i + \sum_{j=1}^N a_{ij} (\boldsymbol{\lambda}_{ij}[k] - \boldsymbol{\lambda}_{ji}[k]) + 2\rho \sum_{j=1}^N a_{ij} \boldsymbol{\theta}_j[k] \right), \\ \boldsymbol{\lambda}_{ij}[k+1] &= \boldsymbol{\lambda}_{ij}[k] + \rho a_{ij} (\boldsymbol{\theta}_j[k+1] - \boldsymbol{\theta}_i[k+1]). \end{aligned} \quad (7.75)$$

In this case, there is no need of running the consensus algorithm for every iteration. Some numerical results are useful to compare the methods. As an example, we considered a connected network composed of $N = 10$ sensors. We set $\rho = 30$ and assumed an observation vector of size $Q = 30$. In Figure 7.6 we report the estimates $\hat{\boldsymbol{\theta}}_{i,l}$, for $l = 1, 2$, versus the iteration index m , for the two algorithms A.1 (left plot) and A.2 (right plot). The iteration index m includes also the iterations necessary for the consensus algorithm to converge within a prescribed accuracy (in this case, we stopped the consensus algorithm as soon as the absolute difference between two consecutive updates is below of 10^{-2} for all the nodes). In both figures, we report, as a benchmark, the maximum likelihood estimate (horizontal x-marked line) achievable by a centralized node that knows all the observation vectors and all the model parameters, i.e., $\mathbf{A}_i, \mathbf{C}_i, \forall i$. From Figure 7.6, we can see that the estimates obtained with both methods converge to the optimal ML estimates.

In the specific case where the observation model is as in (7.73), with additive Gaussian noise, and the noise vectors pertaining to different sensors are mutually uncorrelated, the globally optimal ML estimate is

$$\hat{\boldsymbol{\theta}}_{ML} = \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i \right). \quad (7.76)$$

**FIGURE 7.6**

Per node parameter estimation versus the iteration index m using algorithms **A.1** (left) and **A.2** (right).

Table 7.3 Algorithm A.3**A.3**

STEP 1: Set $k = 0$, and initialize $\boldsymbol{\theta}_i[0] = (\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i)^{-1} (\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i)$;

STEP 2: Repeat until convergence

$$\boldsymbol{\theta}_i[k+1] = \boldsymbol{\theta}_i[k] + \epsilon (\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i)^{-1} \sum_{j=1}^N a_{ij} (\boldsymbol{\theta}_j[k] - \boldsymbol{\theta}_i[k]) \quad (7.77)$$

Set $k = k + 1$, if convergence criterion is satisfied stop, otherwise go to step 2.

This formula is a vector weighted sum of the observations. Recalling that consensus algorithms, if properly initialized, can be made to converge to a weighted sum of the initial states, we can use the consensus algorithm directly to compute the globally optimal ML estimate through a totally distributed mechanism. In particular, in this case, the consensus algorithm proceeds as in Table 7.3.

Using again the basic properties of consensus algorithm, if the graph is connected and the step size ϵ is sufficiently small, the iterations in 7.77 converge to the globally optimal estimate (7.76).

2.07.4.1.3 *Distributed Bayesian estimation under Gaussian noise and Laplacian prior*

Let us consider now the case where the parameter vector is a random vector with known prior probability density function. Following a Bayesian approach, as in (7.50), the practical difference is that in such

a case the objective function must include a term depending on the prior probability. Let us consider, for instance, the interesting case where the observation is Gaussian, as in the previous example, and the prior pdf is Laplacian, i.e.,

$$p_{\Theta}(\boldsymbol{\theta}) = \mu \exp(-\mu \|\boldsymbol{\theta}\|_1) \quad (7.78)$$

with $\mu > 0$, where $\|\mathbf{x}\|_1$ denotes the l_1 norm of vector \mathbf{x} . In this case, the problem to be solved is the following

$$\min_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{A}_i \boldsymbol{\theta}\|_{C_i^{-1}}^2 + \mu \|\boldsymbol{\theta}\|_1 \right\}, \quad (7.79)$$

where $\|\mathbf{x}\|_A^2$ denotes the weighted l_2 norm of \mathbf{x} , i.e., $\|\mathbf{x}\|_A^2 := \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{2}$.

Interestingly, this formulation coincides with the formulation resulting from having no prior pdf, but incorporating an l_1 norm in order to drive the solution towards a sparse vector. This is the so called *least-absolute shrinkage and selection operator (lasso)* method [47]. A distributed algorithm to solve a linear regression problem with sparsity constraint was proposed in [48]. Here we provide a similar approach, with the important difference that, in each iteration, the update is computed in closed form. A decentralized solution can be found by reformulating the problem as follows

$$\begin{aligned} \min_{\boldsymbol{\theta}_i} \quad & \left\{ \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{A}_i \boldsymbol{\theta}_i\|_{C_i^{-1}}^2 + \frac{\rho}{2} \sum_{i=1}^N \|\boldsymbol{\theta}_i - \mathbf{z}\|^2 + \mu \|\mathbf{z}\|_1 \right\}, \\ \text{s.t.} \quad & \boldsymbol{\theta}_i = \mathbf{z}, \quad i = 1, \dots, N. \end{aligned} \quad (7.80)$$

Using the ADMM approach, the algorithm proceeds through the following updates

$$\begin{aligned} \boldsymbol{\theta}_i[k+1] &= \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i + \rho \mathbf{I} \right)^{-1} \left(\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i - \boldsymbol{\lambda}_i[k] + \rho \mathbf{z}[k] \right), \\ \mathbf{z}[k+1] &= \arg \min_{\mathbf{z}} \left\{ \mu \|\mathbf{z}\|_1 + \frac{\rho}{2} \sum_{i=1}^N \|\boldsymbol{\theta}_i[k+1] - \mathbf{z}\|^2 + \sum_{i=1}^N \boldsymbol{\lambda}_i^T (\boldsymbol{\theta}_i[k+1] - \mathbf{z}) \right\} \\ \boldsymbol{\lambda}_i[k+1] &= \boldsymbol{\lambda}_i[k] + \rho (\boldsymbol{\theta}_i[k+1] - \overline{\mathbf{z}[k+1]}). \end{aligned} \quad (7.81)$$

The second equation can also be expressed in closed form. Moreover, defining the vector threshold function $t_{\mu}(\mathbf{x})$ as the vector whose entries are obtained by applying the scalar thresholding function $t_{\mu}(x)$ to each element of vector \mathbf{x} , where

$$t_{\mu}(x) = \begin{cases} x - \mu, & x > \mu \\ 0, & -\mu \leq x \leq \mu \\ x + \mu, & x < -\mu \end{cases} \quad (7.82)$$

the overall algorithm is as in Table 7.4.

As a numerical example, in Figure 7.7 we report the behavior of the estimated variable obtained using Algorithm A.4 versus the iteration index m , which includes the convergence times of two consensus algorithms in the equation 7.84.

Table 7.4 Algorithm A.4**A.4**

STEP 1: Set $k = 0$, and initialize $\theta_i[0]$, $\lambda_i[0]$, $\forall i$, and $z[0]$ randomly;

STEP 2: Repeat until convergence

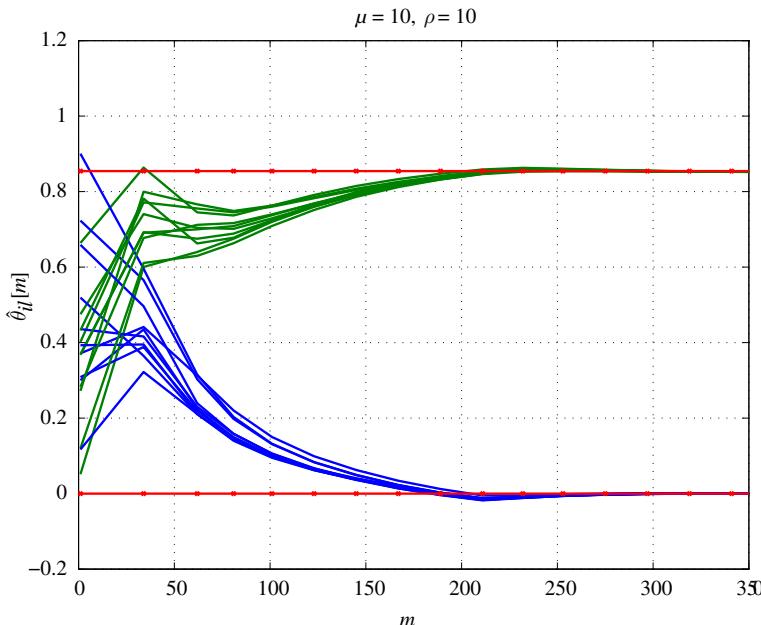
$$\theta_i[k+1] = (\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{A}_i + \rho I)^{-1} (\mathbf{A}_i^T \mathbf{C}_i^{-1} \mathbf{x}_i - \lambda_i[k] + \rho z[k]) \quad (7.83)$$

Run consensus over $\theta_i[k+1]$ and $\lambda_i[k]$ to get $\bar{\theta}[k+1]$ and $\bar{\lambda}[k]$ until ϵ -convergence;

$$z[k+1] = \frac{1}{\rho N} t_\mu (N \bar{\lambda}[k] + \rho N \bar{\theta}[k+1]) \quad (7.84)$$

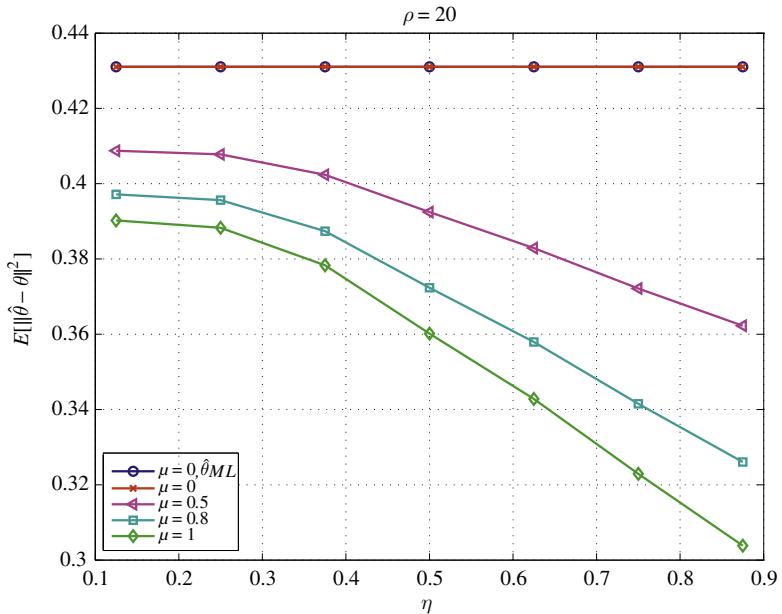
$$\lambda_i[k+1] = \lambda_i[k] + \rho (\theta_i[k+1] - z[k+1]) \quad (7.85)$$

Set $k = k + 1$, if convergence criterion is satisfied stop, otherwise go to step 2.

**FIGURE 7.7**

Per node estimated variable versus the iteration index m for distributed Bayesian estimation using the ADMM approach.

The example refers to a network of $N = 10$ nodes, using $\rho = \mu = 10$. The horizontal x-marked line represents the centralized optimal solution. The parameter vector of this example has two components, one of which has been set to zero to test the capability to recover the sparsity. We can notice from Figure 7.7 that, as expected, the algorithm converges to the globally optimal values.

**FIGURE 7.8**

Mean square estimation error versus the fraction η of null entries, for different μ values.

To show the impact of the penalty coefficient μ on the sparsity of the estimated vector, in Figure 7.8 we have reported the average mean square estimation error versus the coefficient η , defined as the fraction of zeros entries in the vector θ to be estimated, for different values of μ . It can be noted from Figure 7.8 that for $\mu = 0$ the optimal solution is independent by η and it coincides with the optimal (centralized) ML solution. Furthermore, we can observe that, as μ and η increase, the average estimation error decreases thanks to the recovering sparsity property of the ADMM approach with the lasso constraint.

2.07.4.1.4 *Distributed recursive least square estimation with sparsity constraint*

In some applications, the parameters to be estimated may be changing over time. In these cases, it is more advisable to adopt recursive procedure rather than the batch approach followed until now. We show now how to obtain a distributed recursive least square (RLS) estimation incorporating a sparsity constraint.

Let us assume a linear observation model

$$\mathbf{x}_i(l) = \mathbf{A}_i(l)\boldsymbol{\theta} + \mathbf{v}_i(l), \quad (7.86)$$

where $\mathbf{x}_i(l)$ denotes the observation taken by node i at time l , $\mathbf{A}_i(l)$ is a known, possibly time-varying, mixing matrix and $\mathbf{v}_i(l)$ is the observation noise, supposed to have zero mean and covariance matrix $\mathbf{C}_i(l)$.

In RLS estimation with a sparsity constraint, the goal is to find the parameter vector $\boldsymbol{\theta}$, at each time instant n , that minimizes the following objective function

$$\sum_{i=1}^N \sum_{l=1}^n \beta^{n-l} \|\mathbf{x}_i(l) - \mathbf{A}_i(l)\boldsymbol{\theta}\|_{C_i^{-1}(l)}^2 + \mu \|\boldsymbol{\theta}\|_1, \quad (7.87)$$

where $0 < \beta \leq 1$ is a forgetting factor used to weight more the most recent observations with respect to the older ones. The coefficient μ weights the importance of the sparsity constraint.

Proceeding as in the previous examples, a distributed solution can be found by formulating the problem, at each time n , as a constrained problem incorporating an instrumental variable \mathbf{z} to force all the nodes to converge to a common estimate. The problem can be made explicit as

$$\begin{aligned} \min_{\boldsymbol{\theta}_i} \quad & \sum_{i=1}^N \sum_{l=1}^n \beta^{n-l} \|\mathbf{x}_i(l) - \mathbf{A}_i(l)\boldsymbol{\theta}_i\|_{C_i^{-1}(l)}^2 + \frac{\rho}{2} \sum_{i=1}^N \|\boldsymbol{\theta}_i - \mathbf{z}\|^2 + \mu \|\mathbf{z}\|_1, \\ \text{s.t.} \quad & \boldsymbol{\theta}_i = \mathbf{z}, \quad i = 1, \dots, N. \end{aligned} \quad (7.88)$$

Again, the solution can be achieved by applying ADMM and the result is given by the algorithm described in Table 7.5

As before, the only step requiring the interaction among the nodes is a consensus algorithm to be run to compute the averages appearing in equation 7.89.

To test the convergence of Algorithm A.5, we considered a possible application to cooperative sensing for cognitive radio. We assumed the presence of a macro base station transmitting using a multicarrier scheme. We considered for simplicity of representation four channels, but the method can be easily extended to a larger number of channels. The sensing nodes aim to recover the activity of the macro transmitter, represented by a vector $\boldsymbol{\theta}$ composed of four entries, one for each channel. To improve the accuracy of the local estimation, the sensors cooperate with each other by running Algorithm A.5. At some time, the activity level switches from on to off or viceversa. As an example, in Figure 7.9 we report

Table 7.5 Algorithm A.5

A.5

STEP 1: Set $n = 0$ and $k = 0$, and initialize $\boldsymbol{\theta}_i[0, 0]$, $\lambda_i[0, 0]$, $\forall i$, and $\mathbf{z}[0, 0]$ randomly;

STEP 2: Repeat until convergence over index k

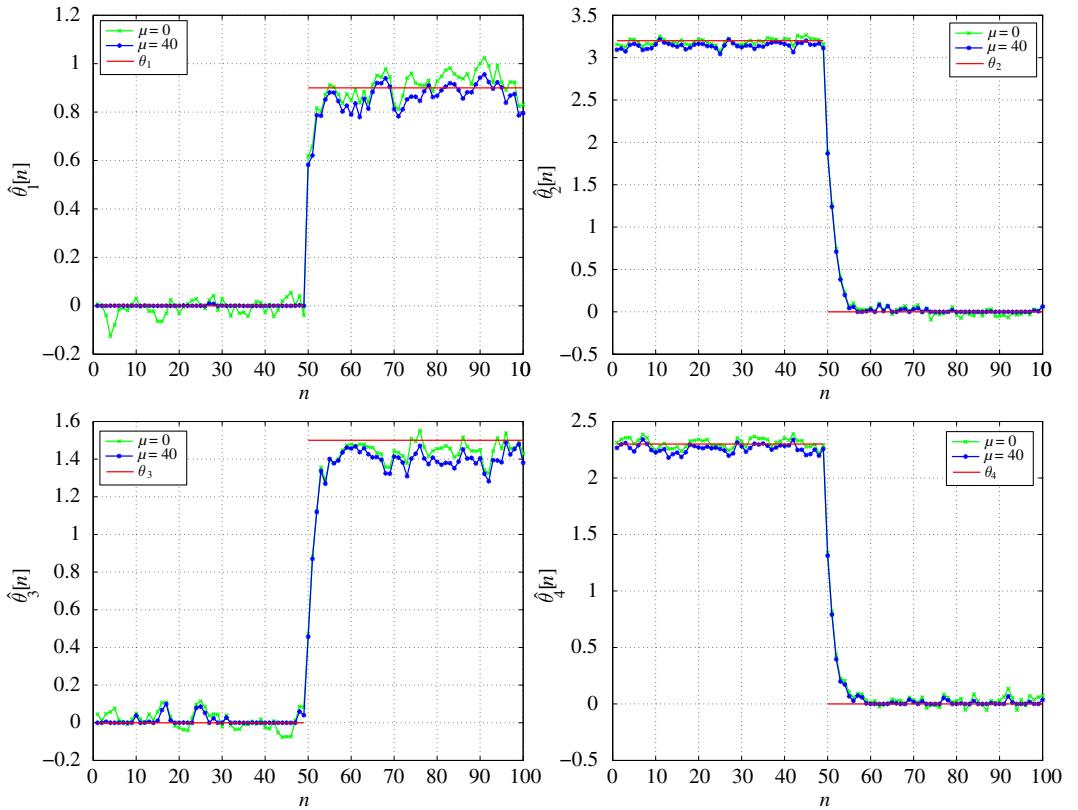
$$\boldsymbol{\theta}_i[k+1, n] = \left(\sum_{l=1}^n \beta^{n-l} \mathbf{A}_i^T(l) \mathbf{C}_i^{-1}(l) \mathbf{A}_i(l) + \rho I \right)^{-1} \left(\sum_{l=1}^n \beta^{n-l} \mathbf{A}_i^T(l) \mathbf{C}_i^{-1}(l) \mathbf{x}_i(l) - \lambda_i[k, n] + \rho \mathbf{z}[k, n] \right)$$

Run consensus over $\boldsymbol{\theta}_i[k+1, n]$ and $\lambda_i[k, n]$ to get $\overline{\boldsymbol{\theta}[k+1, n]}$ and $\overline{\lambda[k, n]}$ until convergence;

$$\mathbf{z}[k+1, n] = \frac{1}{\rho N} t_\mu \left(N \overline{\lambda[k, n]} + \rho N \overline{\boldsymbol{\theta}[k+1, n]} \right) \quad (7.89)$$

$$\lambda_i[k+1, n] = \lambda_i[k, n] + \rho \left(\boldsymbol{\theta}_i[k+1, n] - \overline{\boldsymbol{\theta}[k+1, n]} \right)$$

Set $k = k + 1$, if convergence criterion is satisfied set $n = n + 1$ and go to step 2, otherwise go to step 2.

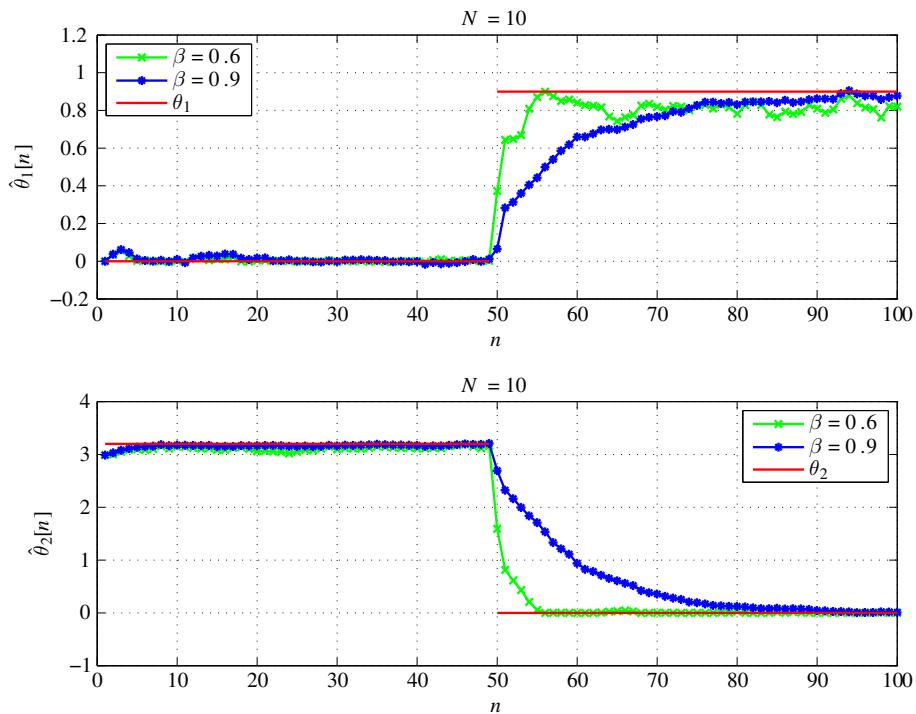
**FIGURE 7.9**

Estimated parameters versus the number of current observations n for the RLS algorithm, assuming $N = 10$, $\beta = 0.6$, and $\rho = 40$.

the four parameters to be estimated, indicated by the unmarked lines. At time $n = 50$, the parameters switch to test the tracking capability of the proposed method.

In Figure 7.9 we draw also the estimated parameters $\hat{\theta}_l$ for $l = 1, \dots, 4$ versus the current observation index n . We used $\beta = 0.6$, $\rho = 40$ and two values of the sparsity coefficient: $\mu = 0$ and $\mu = 40$. We can notice from Figure 7.9 that the method is able to track the true parameters. It is also interesting to see that, as the penalty coefficient μ increases, the zero coefficients are estimated with greater accuracy. Conversely, the positive coefficients are recovered with a slightly larger bias.

To evaluate the impact of the forgetting factor β on the accuracy and tracking capability of the distributed RLS method, in Figure 7.10 we reported the estimated parameters. using $\beta = 0.6$ and $\beta = 0.9$, having set $\rho = \mu = 40$. It can be noted that, as β increases, the larger memory of the filter yields more accurate estimates. At the same time, having a larger memory implies slower time to reaction to the parameter switch, as evidenced in Figure 7.10.

**FIGURE 7.10**

Parameter estimation versus the number of observations n of the recursive least square estimation using the ADMM approach, considering $N = 10$ and two different values of β .

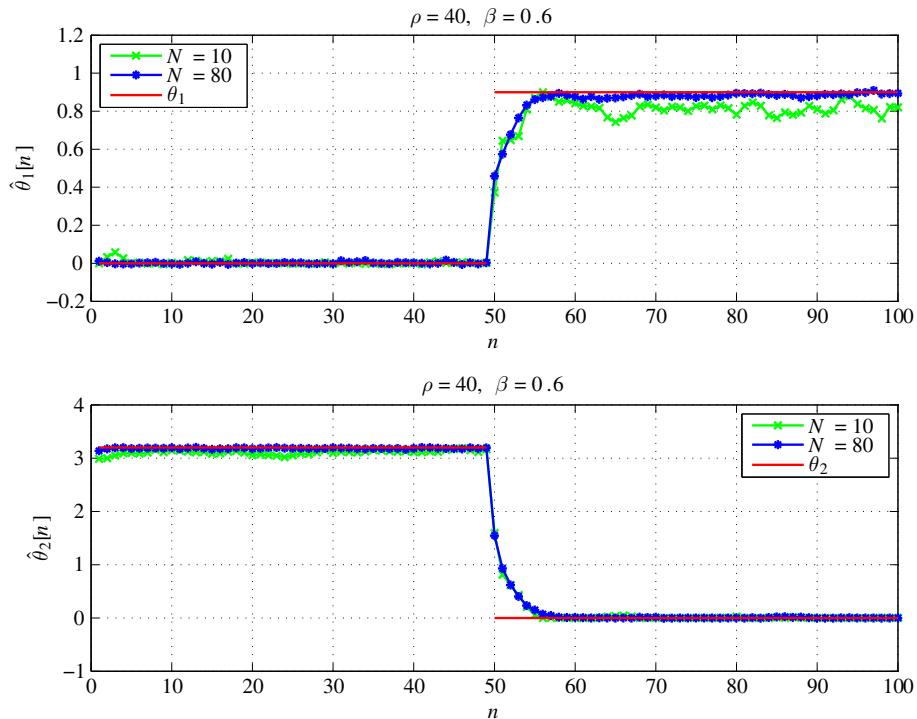
For any given forgetting factor β , the only possibility to improve the estimation accuracy is to have more nodes sensing a common macro base station. As an example, in Figure 7.11 we report the behavior of the estimates obtained with different number of nodes, for a forgetting factor $\beta = 0.6$. We can notice that, as expected, increasing the number of nodes, the estimation accuracy increases as well. This reveals a trade-off between forgetting factor (time memory) and number of nodes involved in cooperative sensing.

2.07.4.1.5 *Distributed parameter estimation in spatially correlated observations*

So far, we have analyzed the case of conditionally independent observations. Let us consider now the case where the observation noise is spatially correlated. More specifically, we assume here the following observation model, for each sensor

$$x_i = \theta + v_i, \quad i = 1, \dots, N, \quad (7.90)$$

where the noise variables v_i are jointly Gaussian with zero mean and covariance matrix C , or precision matrix $A = C^{-1}$. Furthermore, we assume that v is a Gaussian Markov random field, so that the precision matrix is typically a sparse matrix. The joint pdf of the observation vector can then be written

**FIGURE 7.11**

Parameter estimation versus the number of observations n of the recursive least square estimation using the ADMM approach, $\beta = 0.6$, $\mu = 40$, and two different values of N .

as in (7.17), i.e.,

$$p(\mathbf{x}; \theta) = \sqrt{\frac{|\mathbf{A}|}{(2\pi)^N}} \exp\left[-\frac{1}{2}(\mathbf{x} - \theta\mathbf{1})^T \mathbf{A}(\mathbf{x} - \theta\mathbf{1})\right] := \sqrt{\frac{|\mathbf{A}|}{(2\pi)^N}} \exp[-V(\mathbf{x})], \quad (7.91)$$

where $V(\mathbf{x})$ can be rewritten as follows

$$V(\mathbf{x}) = \sum_{i=1}^N \phi_i(\mathbf{x}_i; \theta) \quad (7.92)$$

with $\mathbf{x}_i = [x_i, \{x_j\}_{j \in \mathcal{N}_i, j > i}]^T$, and

$$\phi_i(\mathbf{x}_i; \theta) := \frac{1}{2}a_{ii}(x_i - \theta)^2 + \sum_{j \in \mathcal{N}_i, j > i} a_{ij}(x_j - \theta)(x_i - \theta). \quad (7.93)$$

As in the previous cases, also here a decentralized solution can be reached by formulating the problem as the minimization of the augmented Lagrangian

$$L_\rho(\boldsymbol{\theta}, \boldsymbol{\lambda}, z) := \sum_{i=1}^N \left\{ \phi_i(\mathbf{x}_i; \boldsymbol{\theta}) + \lambda_i(\theta_i - z) + \frac{\rho}{2}(\theta_i - z)^2 \right\}, \quad (7.94)$$

subject to $\theta_i = z$. Applying the ADMM algorithm to this case, we get the following algorithm

$$\begin{aligned} \theta_i[k+1] &= \arg \min_{\theta} \left[\phi_i(\mathbf{x}_i; \theta_i) + \lambda_i(\theta_i - z) + \frac{\rho}{2}(\theta_i - z)^2 \right], \\ z[k+1] &= \frac{1}{N} \sum_{i=1}^N \left(\theta_i[k+1] + \frac{1}{\rho} \lambda_i[k] \right), \\ \lambda_i[k+1] &= \lambda_i[k] + \rho (\theta_i[k+1] - z[k+1]). \end{aligned} \quad (7.95)$$

It is important to notice that, in this case, even if the global problem concerning the minimization of the augmented Lagrangian in (7.94) is certainly convex, the local problem in (7.95) is not necessarily convex because there is no guarantee that the term $\phi_i(\mathbf{x}_i; \boldsymbol{\theta})$ is a positive definite function. Nevertheless, the quadratic penalty present in (7.95) can make every local problem in (7.95) convex. At the same time, at convergence the penalty goes to zero and thus it does not induce any undesired bias on the final result.

The first step in (7.95) can be made explicit, so that the algorithm assumes the form described in Table 7.6.

As an example, in Figure 7.12 we report the estimation versus the cumulative iteration index m that includes the consensus steps and the iterations over k . The results refer to a connected network with $N = 5$ nodes; ρ has been chosen equal to 10 to guarantee that every local problem is convex. It can be noticed from Figure 7.12 that the distributed solution converges to the optimal centralized solution (horizontal x-marked line).

Table 7.6 Algorithm A.6

A.6

STEP 1: Set $k = 0$, and initialize $\theta_i[0]$, $\lambda_i[0]$, $\forall i$, and $z[0]$ randomly;

STEP 2: Repeat until convergence

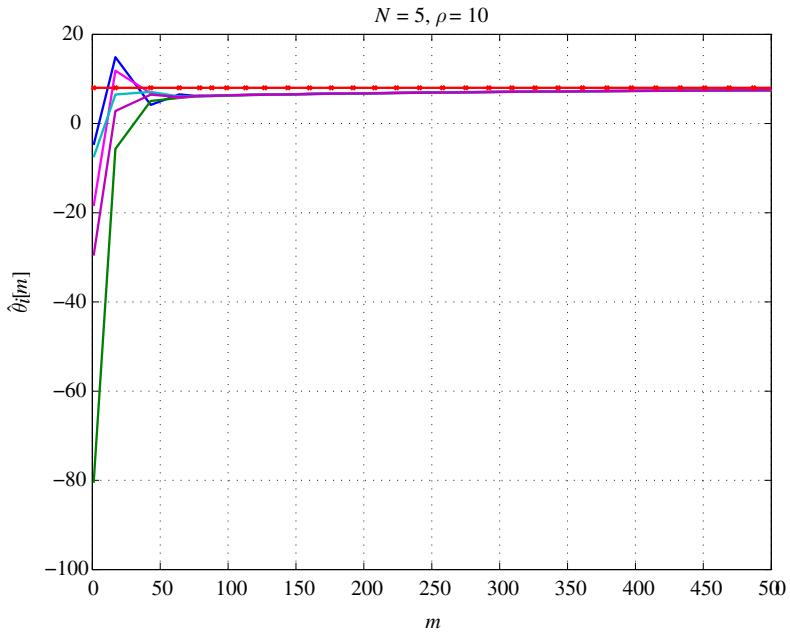
$$\theta_i[k+1] = \frac{1}{a_{ii} + \rho + 2 \sum_{j \in \mathcal{N}_i, j > i} a_{ij}} \left(\rho z[k] - \lambda_i[k] + a_{ii} x_i + \sum_{j \in \mathcal{N}_i, j > i} a_{ij}(x_i + x_j) \right) \quad (7.96)$$

Run consensus over $\theta_i[k+1]$ and $\lambda_i[k]$ to get $\overline{\theta}[k+1]$ and $\overline{\lambda}[k]$ until ϵ -convergence;

$$z[k+1] = \overline{\theta}[k+1] + \frac{1}{\rho} \overline{\lambda}[k] \quad (7.97)$$

$$\lambda_i[k+1] = \lambda_i[k] + \rho (\theta_i[k+1] - z[k+1]) \quad (7.98)$$

Set $k = k + 1$, if convergence criterion is satisfied stop, otherwise go to step 2.

**FIGURE 7.12**

Per node parameter estimation versus the iteration index m for spatially correlated observations using the ADMM approach.

2.07.4.2 Decentralized observations with centralized estimation

In many cases, the observations are gathered in distributed form, through sensors deployed over a certain area, but the decision (either estimation or detection) is carried out in a central fusion center. In this section, we review some of the problems related to distributed estimation, with centralized decision. In such a case, the measurements gathered by the sensors are sent to a fusion center through rate-constrained physical channels. The question is how to design the quantization step in each sensor in order to optimize some performance metric related to the estimation of the parameter of interest. Let us start with an example, to introduce the basic issues.

Let us consider a network of N sensors, each observing a value x_k containing a deterministic parameter θ , corrupted by additive noise v_k , i.e.,

$$x_k = \theta + v_k, \quad k = 1, \dots, N. \quad (7.99)$$

The noise variables v_k are supposed to be zero mean spatially uncorrelated random variables with variance σ_k^2 . Suppose that the sensors transmit their observations via some orthogonal multiple access scheme to a control center which wishes to estimate the unknown signal θ by minimizing the estimation mean square error (MSE) $E[(\hat{\theta} - \theta)^2]$. In the ideal case, where the observations are unquantized and received by the control center without distortion, the best linear unbiased estimator (BLUE) can be

performed by the control center and the estimate $\hat{\theta}$ is given by

$$\hat{\theta} = \left(\sum_{k=1}^N \frac{1}{\sigma_k^2} \right)^{-1} \sum_{k=1}^N \frac{x_k}{\sigma_k^2} \quad (7.100)$$

with MSE given by $E[(\hat{\theta} - \theta)^2] = \left(\sum_{k=1}^N \frac{1}{\sigma_k^2} \right)^{-1}$. This estimator coincides with the maximum likelihood estimator when the noise variables are jointly Gaussian and uncorrelated.

Let us consider now the realistic case, where each sensor quantizes the observation x_k to generate a discrete message m_k of n_k bits. Assuming an error-free transmission, the fusion center must then provide an estimate $\hat{\theta}$ of the true parameter, based on the messages m_k transmitted by all the nodes. More specifically, assuming a uniform quantizer which generates unbiased message functions, the estimator at the control center performs a linear combination of the received messages. Let us suppose that the unknown signal to be estimated belongs to the range $[-A, A]$ and each sensor uniformly divides the range $[-A, A]$ into 2^{n_k} intervals of length $W_k = 2A/2^{n_k}$ rounding x_k to the midpoint of these intervals. In this case, the quantized value m_k at the k th sensor can be written as $m_k = \theta + v_k + w_k$, where the quantization noise w_k is independent of v_k . It can be proved that m_k is an unbiased estimator of θ with

$$\text{Var}\{m_k\} \leq \delta_k^2 + \sigma_k^2, \quad (7.101)$$

where δ_k^2 denotes an upper bound on the quantization noise variance and is given by

$$\delta_k^2 = \frac{W_k^2}{12} = \frac{A^2}{3 \cdot 2^{2n_k}}. \quad (7.102)$$

A linear unbiased estimator of θ is [49]

$$\hat{\theta} = \left(\sum_{k=1}^N \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \sum_{k=1}^N \frac{m_k}{\sigma_k^2 + \delta_k^2}. \quad (7.103)$$

This estimate yields an MSE upper bound

$$E \left[(\hat{\theta} - \theta)^2 \right] \leq \left(\sum_{k=1}^N \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1}. \quad (7.104)$$

As mentioned before, the previous strategy assumes that there are no transmission errors. This property can be made as close as possible to reality by enforcing the transmission rate of sensor k to be strictly less than the channel capacity from sensor k to the fusion center. If we denote by p_k the transmit power of sensor k , h_k the channel coefficient between sensor k and control node and N_0 is the noise variance at the control node receiver, the bound on transmit rate guaranteeing an arbitrarily small error probability is

$$n_k \leq \frac{1}{2} \log \left(1 + \frac{p_k h_k^2}{N_0} \right). \quad (7.105)$$

The problem is then how to allocate power and bits over each channel in order to fulfill some optimality criterion dictated by the estimation problem. This problem was tackled in [49] where it was proposed the

minimization of the Euclidean norm of the transmit power vector under the constraint that the estimation variance is upper bounded by a given quantity and that the number of bits per symbol is less than the channel capacity. Here we formulate the problem as the minimization of the total transmit power under the constraint that the final MSE be upper bounded by a given quantity $\epsilon > 0$. From (7.105), defining $a_k = \frac{h_k^2}{N_0}$, we can derive the number of quantization level as a function of the transmit power,⁶

$$2^{2n_k} = (1 + p_k a_k). \quad (7.106)$$

Our aim is to minimize the sum of powers transmitted by all the sensors under the constraint

$$\left(\sum_{k=1}^N \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1} \leq \epsilon. \quad (7.107)$$

Denoting with $\mathbf{p} = [p_1, \dots, p_N]$ the power vector, the optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{p}} \quad & \sum_{k=1}^N p_k \\ \text{s.t.} \quad & \sum_{k=1}^N \frac{1}{\sigma_k^2 + \frac{A^2}{3 \cdot 2^{2n_k}}} \geq \frac{1}{\epsilon} \\ & \mathbf{p} \geq \mathbf{0}, \end{aligned} \quad (7.108)$$

where n_k is a function of p_k , as in (7.105). In practice, the values n_k are integer. However, searching for the optimal integer values n_k leads to an integer programming problem. To relax the problem, we assume that the variables n_k are real. Then, by using (7.106), the optimization problem in (7.108) can be formulated as

$$\begin{aligned} \min_{\mathbf{p}} \quad & \sum_{k=1}^N p_k \\ \text{s.t.} \quad & \sum_{k=1}^N \frac{1}{\sigma_k^2 + \frac{A^2}{3(1+p_k a_k)}} \geq \frac{1}{\epsilon} \quad (\mathcal{P}) \\ & \mathbf{p} \geq \mathbf{0} \end{aligned} \quad (7.109)$$

Problem (\mathcal{P}) is indeed a convex optimization problem and it is feasible if $\sum_{k=1}^N \frac{1}{\sigma_k^2} > \frac{1}{\epsilon}$.

The optimal solution of the convex problem (\mathcal{P}) can be found by imposing the KKT conditions of (\mathcal{P}) , i.e.,

$$\begin{aligned} 1 - \mu_k - \lambda \frac{3A^2 a_k}{[3\sigma_k^2(1 + p_k a_k) + A^2]^2} &= 0 \quad \forall k = 1, \dots, N \\ 0 \leq \lambda \perp \sum_{k=1}^N \frac{3(1 + p_k a_k)}{3\sigma_k^2(1 + p_k a_k) + A^2} - \frac{1}{\epsilon} &\geq 0 \end{aligned} \quad (7.110)$$

⁶We neglect here the discretization of n_k , to simplify the problem and arrive at closed form expressions.

$$0 \leq \mu_k \perp p_k \geq 0 \quad \forall k = 1, \dots, N$$

where λ and μ_k denote the Lagrangian multipliers associated to the $N + 1$ constraints. The solution for the optimal powers turns out to be

$$p_k^* = \left[\frac{1}{\sigma_k^2} \sqrt{\frac{\lambda A^2}{3a_k}} - \frac{1}{a_k} - \frac{A^2}{3a_k \sigma_k^2} \right]^+, \quad (7.111)$$

where $(x)^+ = \max(0, x)$ and $\lambda > 0$ is found by imposing the MSE constraint to be valid with equality.

It is now useful to present some numerical results. To guarantee the existence of a solution, we set the bound $\epsilon = \beta \epsilon_{\min}$ with $\beta > 1$ and $\epsilon_{\min} = \left(\sum_{k=1}^N \frac{1}{\sigma_k^2} \right)^{-1}$. In Figure 7.13 we report the sum of the optimal transmit powers vs. β , for different SNR values. The number of sensors is $N = 20$. We can notice that the minimum transmit power increases for smaller values of β , i.e., when we require the realistic system to perform closer and closer to the ideal communication case.

In the bottom subplot of Figure 7.14 we report an example of optimal power allocation obtained by solving the optimization problem (\mathcal{P}) , corresponding to the channel realization shown in the top subplot, assuming a constant observation noise variance $\sigma_k^2 = 0.01$. We can observe that the solution is that only

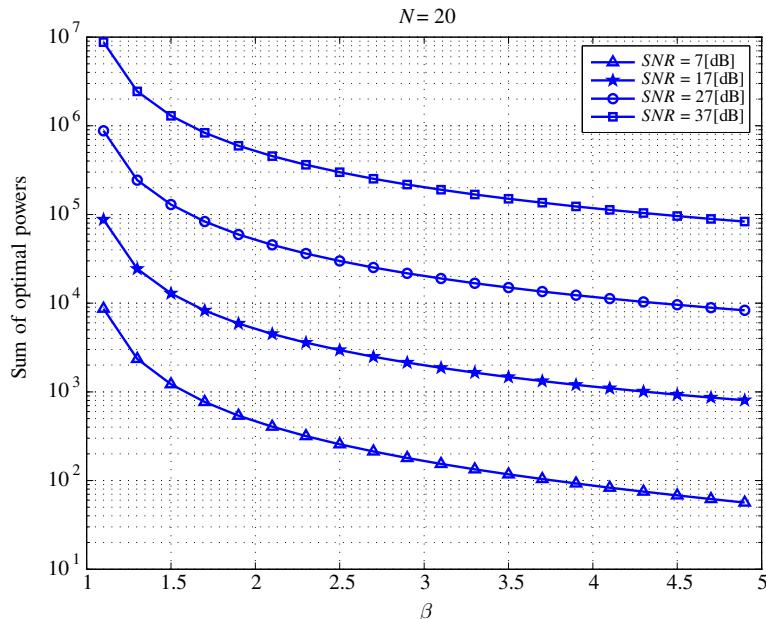
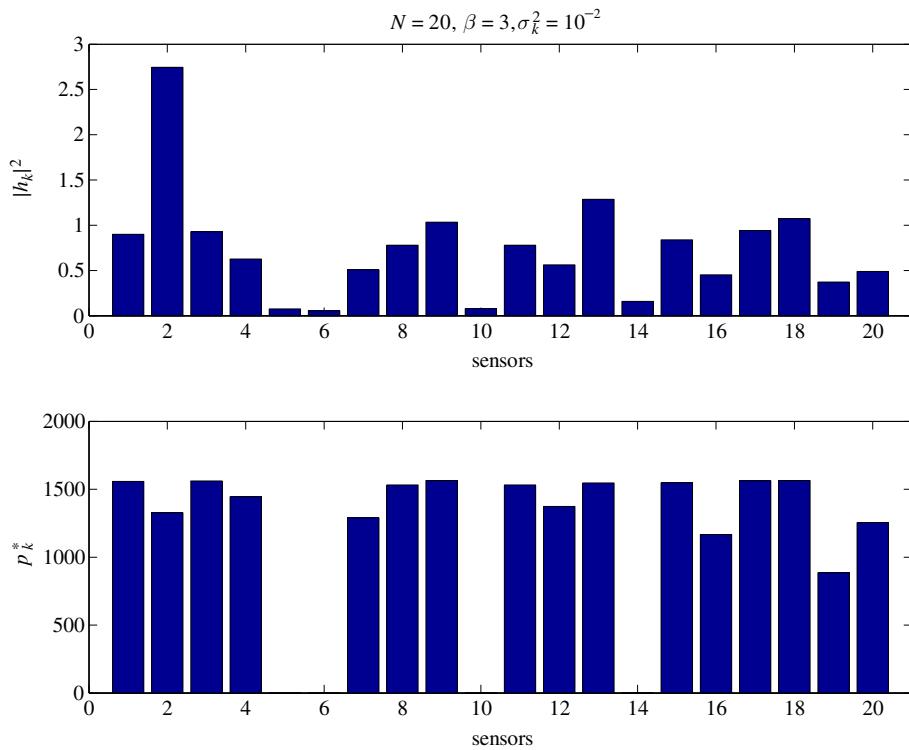


FIGURE 7.13

Sum of the optimal powers for problem (\mathcal{P}) versus β for several values of σ_k^2 .

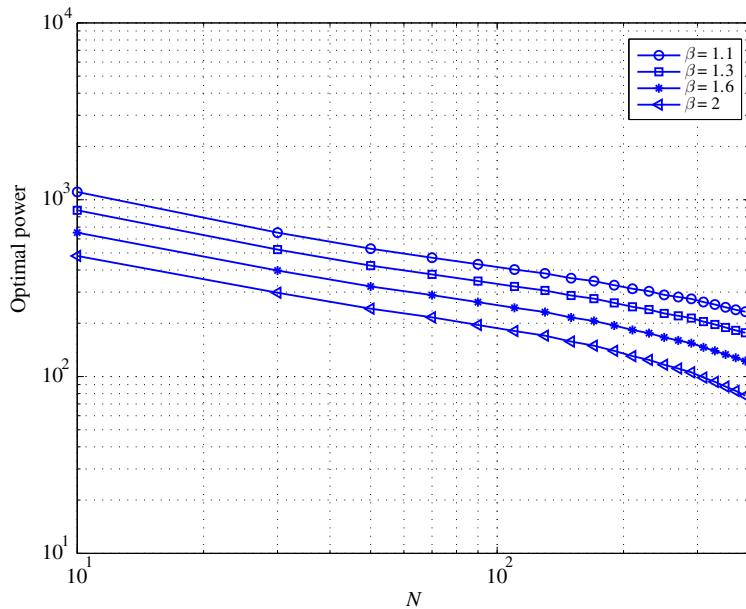
**FIGURE 7.14**

Optimal power allocation of the sensors for problem (\mathcal{P}) , fixing the per-node observation noise variance.

the nodes with the best channels coefficients are allowed to transmit. Finally, in Figure 7.15 we plot the sum of the optimal transmit powers versus the number of sensors N , for different values of β . We can see that, as N increases, a lower power is necessary to achieve the desired estimation variance, as expected.

2.07.5 Distributed detection

The distributed detection problem is in general more difficult to handle than the estimation problem. There is an extensive literature on distributed detection problem, but there is still a number of open problems. According to decision theory, an ideal centralized detector having error-free access to all the measurements collected by a set of nodes, should form the likelihood ratio and compare it with a suitable threshold [5]. Denoting with \mathcal{H}_0 and \mathcal{H}_1 the two alternative hypotheses, i.e., absence or presence of the event of interest, and with $p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_i)$ the joint probability density function of the whole set of observed data, under the hypothesis \mathcal{H}_i , many decision tests can be cast as threshold strategies where the likelihood ratio (LR) is compared with a threshold γ , which depends on the decision criterion. This is true, for example, for two important formulations leading to the Bayes approach and to the

**FIGURE 7.15**

Sum of optimal powers versus N for several values of β .

Neyman-Pearson criterion, the only difference between the two's being the values assumed by the threshold γ . The detection rule decides for \mathcal{H}_1 if the threshold is exceeded or for \mathcal{H}_0 , otherwise. In formulas,

$$\Lambda(\mathbf{x}) := \Lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_1)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_0)} \begin{cases} \geq \gamma & \mathcal{H}_1 \\ \leq \gamma & \mathcal{H}_0 \end{cases} \quad (7.112)$$

Ideally, with no communication constraints, every node should then send its observation vector \mathbf{x}_i , with $i = 1, \dots, N$ to the fusion center, which should then use all the received vectors to implement the LR test, as in (7.112). In reality, there are intrinsic limitations due to, namely: (a) the finite number of bits with which every sensor has to encode the measurements before transmission; (b) the maximum latency with which the decision has to be taken; (c) the finite capacity of the channel between sensors and fusion center. The challenging problem is then how to devise an optimum decentralized detection strategy taking into account the limitations imposed by the communication over realistic channels. The global problem, in the most general setting, is still an open problem, but there are many works in the literature addressing some specific cases. The interested reader may check the book [50] or the excellent tutorial reviews given in [51–53]. The situation becomes more complicated when we take explicitly into account the capacity bound imposed by the communication channel and we look for the number of bits to be used to quantize the local observations before transmitting to the fusion center. This problem was addressed in [54, 55], where it was shown that binary quantization is optimal for the problem of detecting deterministic signals in Gaussian noise and for detecting signals in Gaussian noise using a square-law detector. The interesting

indication, in these contexts, is that the gain offered by having more sensor nodes outperforms the benefits of getting more detailed (nonbinary) information from each sensor. A general framework to cast the problem of decentralized detection is the one where the topology describing the exchange of information among sensing nodes is not simply a tree, with all nodes sending data to a fusion center, but it is a graph. Each node is assumed to transmit a finite-alphabet symbol to its neighbors and the problem is how to find out the encoding (quantization) rule on each node. A class of problems admitting a message passing algorithm with provable convergence properties was proposed in [10]. The solution is a sort of distributed fusion protocol, taking explicitly into account the limits on the communication resources. An interesting and well motivated observation model is a correlated random field, as in many applications the observations concern physical quantities, like temperature or pressure, for example, which being subject to diffusion processes, are going to be spatially and temporally correlated. One of the first works addressing the detection of a known signal embedded in a correlated Gaussian noise was [56]. Using large deviations theory, the authors of [57] study the impact of node density, assuming that observations become increasingly correlated as sensors are in closer proximity of each other. More recently, the detection of a Gauss-Markov Random field (GMRF) with nearest-neighbor dependency was studied in [8]. Scaling laws for the energy consumption of optimal and sub-optimal fusion policies were then presented in [9]. The problem of energy-efficient routing of sensor observations from a Markov random field was analyzed in [58].

A classification of the various detection algorithms depends on the adopted criterion. A first important classification is the following:

1. Global decision is taken at the fusion center
 - a. Nodes send data to FC; FC takes global decision
 - b. Nodes send local decisions to FC; FC fuses local decisions
2. Every node is able to take a global decision
 - a. Nodes exchange data with their neighbors
 - b. Nodes exchange local decisions with their neighbors.

In the first case, the observation is distributed across the nodes, but the decision is centralized. This case has received most of the attention. The interested reader may check, for example, the book [50] or the tutorial reviews given in [51–53]. In the second case, also the decision is decentralized. This case has been considered only relatively recently. Some references are, for example, [59–65].

An alternative classification is between

1. Batch algorithms
2. Sequential algorithms.

In the first case, the network collects a given amount of data along the time and space domains and then it takes a decision. In the second case, the number of observations, either in time or in terms of number of involved sensors, is not decided a priori, but it is updated at every new measurement. The network stops collecting information only when some performance criterion is satisfied (typically, false alarm and detection probability) [66–68].

One of the major difficulties in distributed detection comes from establishing the optimal decision thresholds at local and global level. The main problem is how to optimize the local decisions, taking into

account that the final decisions will be only the result of the interaction among the nodes. Taking a local decision can be interpreted as a form of source coding. The simple (binary) hypothesis testing can be seen in fact as a form of binary coding. Whenever the observations are conditionally independent, given each hypothesis, the likelihood ratio test at the sensor nodes is indeed optimal [69]. However, finding the optimal quantization levels is a difficult task. Even when the observations are i.i.d., assuming identical decision rules is very common and apparently well justified. Nevertheless there are counterexamples showing that nonidentical decision rules are optimal [69]. Identical decision rules in the i.i.d. case turns out to be optimal only asymptotically, as the number of nodes tends to infinity [70].

A simple example may be useful to grasp some of the difficulties associated with distributed detection. For this purpose, we briefly recall the seminal work of Tenney and Sandell [71]. Let us consider two sensors, each measuring a real quantity x_i , with $i = 1, 2$. Based on its observation x_i , sensor i decides whether the phenomenon of interest is present or not. In the first case, it sets the decision variable $u_i = 1$, otherwise, it sets $u_i = 0$. The question is how to implement the decision strategy, according to some optimality criterion. The approach proposed in [71] is a Bayesian approach, where the goal of each sensor is to minimize the Bayes risk, which can be made explicit by introducing the cost coefficients and the observation probability model. Let us denote by C_{ijk} the cost of detector 1 deciding on \mathcal{H}_i , detector 2 deciding on \mathcal{H}_j , when the true hypothesis is \mathcal{H}_k . Denoting by P_k the prior probability of event \mathcal{H}_k and by $p(u_1, u_2, x_1, x_2, \mathcal{H}_k)$ the joint pdf of having \mathcal{H}_k , observing the pair (x_1, x_2) and deciding for the pair (u_1, u_2) , the average risk can be written as

$$\begin{aligned} \mathcal{R} &= \sum_{i,j,k} \int C_{ijk} p(u_1, u_2, x_1, x_2, \mathcal{H}_k) dx_1 dx_2 \\ &= \sum_{i,j,k} P_k \int C_{ijk} p(u_1, u_2, x_1, x_2 / \mathcal{H}_k) dx_1 dx_2 \\ &= \sum_{i,j,k} P_k \int C_{ijk} p(u_1, u_2 / x_1, x_2, \mathcal{H}_k) p(x_1, x_2 / \mathcal{H}_k) dx_1 dx_2. \end{aligned} \quad (7.113)$$

In this case, each node observes only its own variable and takes a decision independently of the other node. Hence, we can set

$$\mathcal{R} = \sum_{i,j,k} P_k \int C_{ijk} p(u_1 / x_1) p(u_2 / x_2) p(x_1, x_2 / \mathcal{H}_k) dx_1 dx_2. \quad (7.114)$$

Expanding the right hand side by explicitly summing over index i , we get

$$\mathcal{R} = \sum_{j,k} P_k \int p(u_2 / x_2) p(x_1, x_2 / \mathcal{H}_k) [C_{0jk} p(u_1 = 0 / x_1) + C_{1jk} p(u_1 = 1 / x_1)] dx_1 dx_2. \quad (7.115)$$

Considering that $p(u_1 = 1 / x_1) = 1 - p(u_1 = 0 / x_1)$ and ignoring all terms which do not contain u_1 , we get

$$\mathcal{R} = \int p(u_1 = 0 / x_1) \sum_{j,k} P_k \left\{ \int p(u_2 / x_2) p(x_1, x_2 / \mathcal{H}_k) [C_{0jk} - C_{1jk}] dx_2 \right\} dx_1 + \text{const.} \quad (7.116)$$

The average risk is minimized if $p(u_1 = 0/x_1)$ is chosen as follows

$$p(u_1 = 0/x_1) = \begin{cases} 0, & \text{if } \sum_{j,k} P_k \int p(u_2/x_2) p(x_1, x_2/\mathcal{H}_k) [C_{0jk} - C_{1jk}] dx_2 \geq 0 \\ 1, & \text{otherwise.} \end{cases} \quad (7.117)$$

This expression shows that the optimal local decision rule is a *deterministic* rule. After a few algebraic manipulations, (7.117) can be rewritten, equivalently, as [50]

$$\Lambda(x_1) := \frac{p(x_1/\mathcal{H}_1)}{p(x_1/\mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \frac{P_0 \sum_j \int p(u_2/x_2) p(x_2/x_1, \mathcal{H}_0) [C_{1j0} - C_{0j0}] dx_2}{P_1 \sum_j \int p(u_2/x_2) p(x_2/x_1, \mathcal{H}_1) [C_{0j1} - C_{1j1}] dx_2}, \quad (7.118)$$

where $\Lambda(x_1)$ is the LR at node 1. Eq. (7.118) has the structure of a LRT. However, note that the threshold on the right hand side of (7.118) depends on the observation x_1 , through the term $p(x_2/x_1/\mathcal{H}_1)$, which incorporates the statistical dependency between the observations x_1 and x_2 . Hence, Eq. (7.118) is not a proper LRT.

The situation simplifies if the observations are conditionally independent, i.e., $p(x_2/x_1/\mathcal{H}_k) = p(x_2/\mathcal{H}_k)$. In such a case, the threshold t_1 can be simplified into

$$t_1 = \frac{P_0 \int p(x_2/\mathcal{H}_0) \{p(u_2 = 0/x_2)[C_{100} - C_{000}] + p(u_2 = 1/x_2)[C_{110} - C_{010}]\} dx_2}{P_1 \int p(x_2/\mathcal{H}_1) \{p(u_2 = 0/x_2)[C_{001} - C_{101}] + p(u_2 = 1/x_2)[C_{011} - C_{111}]\} dx_2}. \quad (7.119)$$

Since $p(u_2 = 1/x_2) = 1 - p(u_2 = 0/x_2)$, (7.119) can be rewritten as

$$t_1 = \frac{P_0 \int p(x_2/\mathcal{H}_0) \{[C_{110} - C_{010}] + p(u_2 = 0/x_2)[C_{100} - C_{000} + C_{010} - C_{110}]\} dx_2}{P_1 \int p(x_2/\mathcal{H}_1) \{[C_{011} - C_{111}] + p(u_2 = 0/x_2)[C_{001} - C_{101} + C_{111} - C_{011}]\} dx_2}. \quad (7.120)$$

Hence, the threshold t_1 to be used at node 1 is a function of $p(u_2 = 0/x_2)$, i.e., on the decision taken by node 2. At the same time, the threshold t_2 to be used by node 2 will depend on the decision rule followed by node 1. This means that, even if the observations are conditionally independent and the decisions are taken autonomously by the two nodes, the decisions are still coupled through the thresholds. This simple example shows how the detection problem can be rather complicated, even under a very simple setting.

In the special case where $C_{000} = C_{111} = 0$, $C_{010} = C_{100} = C_{011} = C_{101} = 1$, and $C_{110} = C_{001} = 2$, i.e., there is no penalty if the decisions are correct, the penalty is 1, when there is one error, and the penalty is 2 when there are two errors, the threshold simplifies into

$$t_1 = \frac{P_0}{P_1}. \quad (7.121)$$

Hence, in this special case, the two thresholds are independent of each other and the two detectors become independent of each other.

After having pointed out through a simple example some of the problems related to distributed detection, it is now time to consider in more detail the cases where the nodes send their (possibly encoded) data to the FC or they take local decisions first and send them to the FC. In both situations, there are two extreme cases: (a) there is only one FC; (b) every node is a potential FC, as it is able to take a global decision.

2.07.5.1 Nodes send data to decision center

Let us consider for simplicity the simple (binary) hypothesis testing problem. Given a set of vector observations $\mathbf{x} := [\mathbf{x}_1, \dots, \mathbf{x}_N]$, where \mathbf{x}_i is the vector collected by node i , $i = 1, \dots, N$, the optimal decision rule for the simple hypothesis testing problem, under a variety of optimality criteria, amounts to compute the likelihood ratio (LR) $\Lambda(\mathbf{x})$ and compare it with a threshold. In formulas,

$$\Lambda(\mathbf{x}) := \Lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_1)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N; \mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \gamma. \quad (7.122)$$

In words, the detector decides for \mathcal{H}_1 if the LR exceeds the threshold, otherwise it decides for \mathcal{H}_0 . In general, what changes the distributed detection problem from the standard centralized detection is that the data are sent to the decision center after source encoding into a discrete alphabet. The simplest form of encoding is quantization. But also taking local decisions can be interpreted as a form of binary coding. Clearly, source coding is going to affect the detection performance. It is then useful to show, through a simple example, how local quantization affects the final detection performance and how we can benefit from the theoretical analysis to optimize the number of bits associated to the quantization step in order to optimize performance of the detection scheme.

2.07.5.1.1 Centralized detection of deterministic signal embedded in additive noise

Let us consider the detection of a deterministic (known) signal embedded in additive noise. In this section, we consider the case where the decision is taken at a FC, after having collected the data sent by the sensors. This case could refer for example to the detection of undesired resonance phenomena in buildings, bridges, etc. The form of the resonance is known. However, the measurements taken by the sensors are affected by noise and then it is of interest to check the performance as a function of the signal to noise ratio.

The measurement vector is $\mathbf{x} = (x_1, \dots, x_N)$, where x_i is the measurement taken by node i . Let us denote as \mathbf{s} the known deterministic signal. The observation can be modeled as

$$\mathbf{x} \sim \begin{cases} \mathbf{v} + \mathbf{w} & \text{under } \mathcal{H}_0 \\ \mathbf{s} + \mathbf{v} + \mathbf{w} & \text{under } \mathcal{H}_1 \end{cases}, \quad (7.123)$$

where \mathbf{v} is the background noise, whereas \mathbf{w} is the quantization noise. We assume the noise to be Gaussian with zero mean and (spatial) covariance matrix \mathbf{C}_n , i.e., $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$. To simplify the mathematical tractability, we consider a dithered quantization so that the quantization error can be modeled as a random process statistically independent of noise. We may certainly assume that, after dithering, the quantization noise variables over different sensors are statistically independent. Hence, we can state that the quantization noise vector \mathbf{w} has zero mean and a diagonal covariance matrix $\mathbf{C}_q = \text{diag}(\sigma_{q1}^2, \dots, \sigma_{qN}^2)$. If the amplitude of the useful signal spans the dynamic range $[-A, A]$ and the number of bits used by node i is n_i , the quantum range is $q_i = 2A/2^{n_i}$ so that the quantization noise variance at node i is

$$\sigma_{qi}^2 = \frac{(2A)^2}{12 \cdot 2^{2n_i}} = \frac{A^2}{3 \cdot 2^{2n_i}}. \quad (7.124)$$

The overall noise has then a zero mean and covariance matrix $\mathbf{C} = \mathbf{C}_n + \mathbf{C}_q$.

If the quantization noise is negligible, the Neyman-Pearson criterion applied to this case leads to the following linear detector

$$T(\mathbf{x}) = \mathcal{R}\{\mathbf{s}^H \mathbf{C}^{-1} \mathbf{x}\} \begin{cases} \geq \gamma, & \mathcal{H}_1 \\ \leq \gamma, & \mathcal{H}_0 \end{cases} \quad (7.125)$$

where $\mathcal{R}(x)$ denotes the real part of x , and the detection threshold γ is computed in order to guarantee the desired false alarm probability P_{fa} . Unfortunately, since the quantization noise is not Gaussian, the composite noise $\mathbf{v} + \mathbf{w}$ is not Gaussian and then the detection rule in (7.125) is no longer optimal. Nevertheless, the rule in (7.125) is still meaningful as it maximizes the signal to noise ratio (SNR). Hence, it is of interest to look at the performance of this detector in the presence of quantization noise. The exact computation of the detection probability is not easy, at least in closed form, because it requires the computation of the pdf of $T(\mathbf{x})$. Nevertheless, when the number of nodes is sufficiently high (an order of a few tens can be sufficient to get a good approximation), we can invoke the central limit theorem to state that $T(\mathbf{x})$ is approximately Gaussian. Using this approximation, the detection probability can be written in closed form for any fixed P_{fa} , following standard derivations (see, e.g., [5]), as

$$P_d = Q\left[Q^{-1}(P_{fa}) - \sqrt{\mathbf{s}^H \mathbf{C}^{-1} \mathbf{s}}\right] = Q\left[Q^{-1}(P_{fa}) - \sqrt{\mathbf{s}^H (\mathbf{C}_n + \mathbf{C}_q)^{-1} \mathbf{s}}\right]. \quad (7.126)$$

This formula is useful to assess the detection probability as a function of the bits allocated to each transmission. At the same time, we can also use (7.126) as a way to find out the bit allocation that maximizes the detection probability. This approach establishes an interesting link between the communication and detection aspects. In practice, in fact, encoded data are transmitted over a finite capacity channel. Hence, it is useful to relate the number of quantization bits used by each node and capacity of the channel between that node and the FC. For simplicity, we consider the optimization problem under the assumption of spatially uncorrelated noise, i.e., $\mathbf{C}_n = \text{diag}(\sigma_{n1}^2, \dots, \sigma_{nN}^2)$. The problem we wish to solve is the maximization of the detection probability, for a given false alarm rate and a maximum global transmit power. To guarantee an arbitrarily low transmission error rate, we need to respect Shannon's channel coding theorem, so that the number of bits per symbol must be less than channel capacity. Denoting with p_i the power transmitted by user i and assuming flat fading channel, with channel coefficient h_i^2 , the capacity is given by (7.105). From (7.126), maximizing P_d is equivalent to maximizing $\mathbf{s}^H (\mathbf{C}_n + \mathbf{C}_q)^{-1} \mathbf{s}$. Hence, using (7.124), the maximum P_d , for a given P_{fa} and a given global transmit power P_T , can be achieved by finding the power vector $\mathbf{p} = (p_1, \dots, p_N)$ that solves the following constrained problem

$$\max_{\mathbf{p}} \sum_{i=1}^N |s_i|^2 \left(\sigma_{ni}^2 + \frac{A^2}{3(1+a_i p_i)} \right)^{-1} \quad (7.127)$$

$$\text{s.t. } \sum_{i=1}^N p_i \leq P_T; p_i \geq 0, i = 1, \dots, N. \quad (7.128)$$

It is straightforward to check that this is a convex problem. Imposing the Karush-Kuhn-Tucker conditions, the optimal powers can be expressed in closed form as:

$$p_i = \left[\frac{1}{\sqrt{\lambda}} \sqrt{\frac{s_i^2 A^2}{3a_i \sigma_{ni}^4}} - \frac{A^2}{3a_i \sigma_{ni}^2} - \frac{1}{a_i} \right]^+, \quad (7.129)$$

where the Lagrange multiplier λ associated to the sum-power constraint can be determined as the value that makes $\sum_{i=1}^N p_i = P_T$.

A numerical example is useful to grasp some of the properties of the proposed algorithm. Let us consider a series of sensors placed along a bridge of length L . The purpose of the network is to detect one possible spatial resonance, which we represent as the signal $s(z) = A \cos(\pi z/L)$, where $z \in [-L/2, L/2]$ denotes the spatial coordinate. The sensors are uniformly spaced along the bridge, at positions $z_i = (i - 1)L/N$, with $i = 1, \dots, N$. Every sensor measures a shift $x_i = s(z_i) + v_i$, affected by the error v_i . To communicate its own measurement to the FC, every sensor has to quantize the measurement first. The optimal number of bits to be used by every sensor can be computed by using the previous theory. In this case, in Figure 7.16 we report the detection probability vs. the sum power P_T available to the whole set of sensors, for different numbers N of sensors. As expected, as the total transmit power increases, P_d increases because more bits per symbol can be transmitted and then the quantization errors become negligible. It is also important to notice how, increasing the number of sensors, the detection probability improves, for any given transmit power. Furthermore, in Figure 7.17 we can see the optimal per channel bit allocation (bottom), together with the channels profiles $|h_k|^2$ (top). Interestingly, we can see that the method allocates more bits in correspondence with the best channels and the central elements of the array, where the useful signal is expected to have the largest variations.

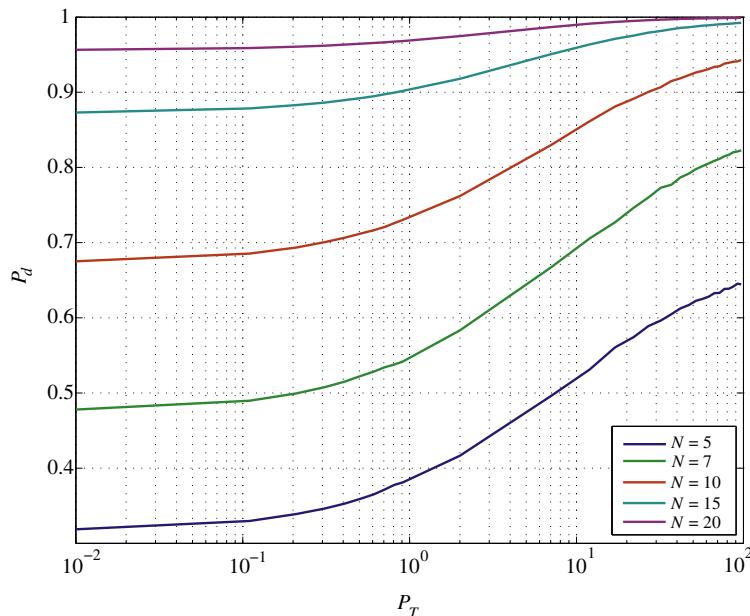
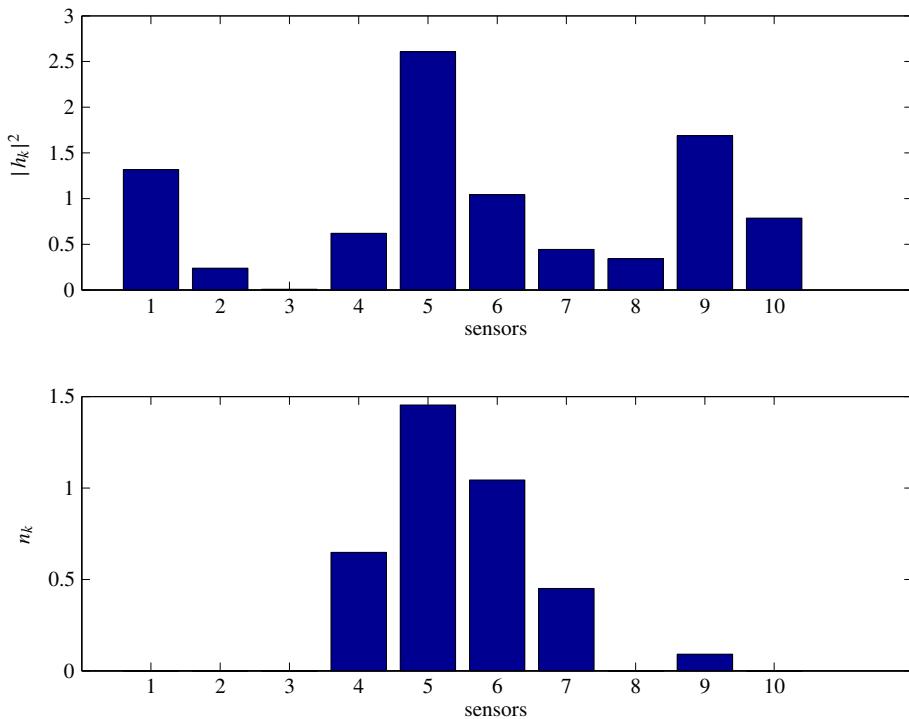


FIGURE 7.16

Detection probability vs. sum transmit power, for different number of sensors.

**FIGURE 7.17**

Optimal bit allocation.

2.07.5.1.2 Decentralized detection under conditionally independent observations

Let us consider now the case where the globally optimal decision can be taken, in principle, by any node. To enable this possibility, every node must be able to implement the statistical test (7.122). If the measurements collected by the sensors are conditionally independent, the logarithm of the likelihood ratio can be written as

$$\log \Lambda(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{i=1}^N \log \Lambda_i(\mathbf{x}_i) = \sum_{i=1}^N [\log p_{X_i}(\mathbf{x}_i, \mathcal{H}_1) - \log p_{X_i}(\mathbf{x}_i, \mathcal{H}_0)]. \quad (7.130)$$

This formula shows that, in the conditionally independent case, running a consensus algorithm is sufficient to enable every node to compute the global LR. It is only required that every sensor initializes its own state with the local log-LR $\log \Lambda_i(\mathbf{x}_i)$ and then runs the consensus iterations. If the network is connected, every node will end up with the average value of the local LR's. In practice, to send the local LR, every node must quantize it first. Then, we need to refer to the consensus algorithm in the presence of quantization errors. However, we have already seen in Section 2.07.3.2 that the consensus iterations may

be properly modified to make the algorithm robust against a series of drawbacks coming from communications through realistic channels, as, eg., random packet drops and quantization. Hence, a consensus algorithm, properly modified, can enable every node to compute the global LR with controllable error.

2.07.5.2 Nodes send local decisions to fusion center

Consider now the case where each node i takes a local decision, according to a locally optimal criterion, and encodes the decision into the binary variable u_i . Then, the node sends the variable u_i to the fusion center, which is asked to take a global decision on the basis of the vector $\mathbf{u} := (u_1, \dots, u_N)$ containing all local decisions. Let us consider for simplicity the binary hypothesis test. This problem was considered in [50] and we will now review the basic results. This problem is distinct from the case studied in the previous section because here the local decision thresholds are optimized according to a detection criterion, whereas in standard quantization the decision thresholds are not optimized.

Under both Bayesian and Neyman-Pearson (NP) formulations, the optimal test amounts to a likelihood ratio test, based on \mathbf{u} , i.e.,

$$\frac{p(u_1, \dots, u_N; \mathcal{H}_1)}{p(u_1, \dots, u_N; \mathcal{H}_0)} \geq \eta. \quad (7.131)$$

In the case of conditionally independent local decisions, the LRT converts into

$$\frac{\prod_{i=1}^N p(u_i; \mathcal{H}_1)}{\prod_{i=1}^N p(u_i; \mathcal{H}_0)} := \prod_{i=1}^N \Lambda_i(u_i) \geq \eta. \quad (7.132)$$

Since each variable u_i can only assume the values 0 or 1, we can group all the variables into two subsets: the subset S_0 containing all variables $u_i = 0$ and the subset S_1 containing all variables $u_i = 1$, thus yielding

$$\prod_{i \in S_0} \frac{p(u_i = 0; \mathcal{H}_1)}{p(u_i = 0; \mathcal{H}_0)} \prod_{i \in S_1} \frac{p(u_i = 1; \mathcal{H}_1)}{p(u_i = 1; \mathcal{H}_0)} \geq \eta. \quad (7.133)$$

Denoting with $P_{Mi} = p(u_i = 0; \mathcal{H}_1)$, and $P_{Fi} = p(u_i = 1; \mathcal{H}_0)$, the probabilities of miss and the probability of false alarm of node i , respectively, (7.133) can be rewritten as

$$\prod_{i \in S_0} \frac{P_{Mi}}{1 - P_{Fi}} \prod_{i \in S_1} \frac{1 - P_{Mi}}{P_{Fi}} \geq \eta. \quad (7.134)$$

Taking the logarithm of both sides and reintroducing the variables u_i , the fusion rule becomes

$$\sum_{i=1}^N \left[\log \left(\frac{1 - P_{Mi}}{P_{Fi}} \right) u_i + \log \left(\frac{P_{Mi}}{1 - P_{Fi}} \right) (1 - u_i) \right] \geq \log \eta \quad (7.135)$$

or, equivalently

$$\sum_{i=1}^N \log \left[\frac{(1 - P_{Mi})(1 - P_{Fi})}{P_{Mi} P_{Fi}} \right] u_i \geq \log \left[\eta \prod_{i=1}^N \frac{1 - P_{Fi}}{P_{Mi}} \right]. \quad (7.136)$$

The optimal fusion rule is then a simple weighted sum of the local decisions, where the weights depend on the reliabilities of the local decisions: Larger weights are assigned to the most reliable nodes.

If instead of having a single FC, we wish to enable every node to implement the decision fusion rule described above, we can see that, again, running a consensus algorithm suffices to reach the goal. In fact, if each local state variable is initialized with a value $x_i[0] = \log \left[\frac{(1-P_{Mi})(1-P_{Fi})}{P_{Mi}P_{Fi}} \right] u_i$, running a consensus algorithm allows every node to know the function in (7.136). The only constraint is, as always, network connectivity. The drawback of this simple approach is that running this sort of consensus algorithm requires the transmission of real variables, rather than the binary variables u_i . In fact, even if the local decision u_i is binary, the coefficient multiplying u_i is a real variable, which needs to be quantized before transmission over a realistic channel. Again, the consensus algorithm can be robustified against quantization errors by using dithered quantization and a decreasing step size, as shown in 2.07.3.2. However, it is important to clarify that we cannot make any claim of optimality of this kind of distributed decision. In principle, when the nodes exchange their decisions with the neighbors, the decision thresholds should be adjusted in order to accommodate some optimality criterion. This is indeed an interesting, yet still open, research topic.

2.07.6 Beyond consensus: distributed projection algorithms

In many applications, the field to be reconstructed by a sensor network is typically a smooth function of the spatial coordinates. This happens for example, in the reconstruction of the spatial distribution of the power radiated by a set of transmitters. The problem is that local measurements may be corrupted by local noise or fading effects. An important application of this scenario is given by cognitive networks. In such a case, a secondary node would need to know the channel occupation across space, to find out unoccupied channels, within the area of interest. This requires some sort of spectrum sensing, but in a localized area. The problem of sensing is that wireless propagation is typically affected by fading or shadowing effects, so that a sensor in a shadowed location might indicate that a channel is unoccupied, while this is not true. To avoid this kind of error, which would lead to undue channel occupation from opportunistic users, it is useful to resort to cooperative sensing. In such a case, nearby nodes exchange local measurements to counteract the effect of shadowing.

The problem with local averaging operations is that they should reduce the effect of fading, but without destroying valuable spatial variations. In the following, we recall a distributed algorithm proposed in [72] to recover a spatial map of a field, using local weighted averages where the weights are chosen so as to improve upon local noise or fading effects, but without destroying the spatial variation of the useful signal.

Let us consider a network composed of N sensors located at positions (x_i, y_i) , $i = 1, \dots, N$, and denote the measurement collected by the i th sensor by $g(x_i, y_i) = z(x_i, y_i) + v_i$, where $z(x_i, y_i)$ represents the useful field while v_i is the observation error. Let us also denote by $u_k(x, y)$, $k = 1, \dots, r$, a set of linearly independent spatial functions defining a basis for the useful signal. The useful signal can then be represented through the basis expansion model

$$z(x_i, y_i) = \sum_{k=1}^r s_k u_k(x_i, y_i). \quad (7.137)$$

In vector notation, introducing the N -size column vector $\mathbf{g} := [(g(x_1, y_1), g(x_2, y_2), \dots, g(x_N, y_N))^T]$ and similarly for the vector \mathbf{z} , we may write

$$\mathbf{g} = \mathbf{z} + \mathbf{v} = \mathbf{U}\mathbf{s} + \mathbf{v}, \quad (7.138)$$

where \mathbf{U} is the $N \times r$ matrix whose m th column is $\mathbf{u}_m = (u_m(x_1, y_1), \dots, u_m(x_N, y_N))$, $\mathbf{s} = (s_1, \dots, s_r)$ is an r -size vector of coefficients and $\mathbf{z} = \mathbf{U}\mathbf{s}$ is the useful signal. The spatial smoothness of the useful signal field may be captured by choosing the functions $u_k(x, y)$ to be the low frequency components of the Fourier basis or low-order 2 D polynomials. For instance, if the space under monitoring is a square of side L , we may choose the set

$$\{u_{nm}(x, y)\} = \left\{ 1, \cos\left(2\pi \frac{nx + my}{L}\right), \sin\left(2\pi \frac{nx + my}{L}\right) \right\}_{m=0, n=0; m+n \neq 0}^{m=\infty, n=\infty}. \quad (7.139)$$

In practice, the dimension r of the useful signal subspace is typically much smaller than the dimension N of the observation space, i.e., of the number of sensors. We can exploit this property to devise a distributed denoising algorithm.

If we use a Minimum Mean Square Error (MMSE) strategy, the goal is to find the useful signal vector $\hat{\mathbf{s}}$ that minimizes the mean square error

$$\mathcal{E} := E\{\|\mathbf{g} - \mathbf{U}\hat{\mathbf{s}}\|^2\}. \quad (7.140)$$

The solution is well known and is given by [44]:

$$\hat{\mathbf{s}} = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{g}. \quad (7.141)$$

Our goal is actually to recover the vector \mathbf{z} , rather than \mathbf{s} . In such a case, the estimate of \mathbf{z} is

$$\hat{\mathbf{z}} = \mathbf{U}(\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{g}. \quad (7.142)$$

The operation performed in (7.142) corresponds to projecting the observation vector onto the subspace spanned by the columns of \mathbf{U} . Assuming, without any loss of generality (w.l.o.g.), the columns of \mathbf{U} to be orthonormal, the projector simplifies into

$$\hat{\mathbf{z}} = \mathbf{U} \mathbf{U}^T \mathbf{g}. \quad (7.143)$$

The centralized solution to this problem is then very simple: The fusion center collects all the measurements $g(x_i, y_i)$, compute \mathbf{U} and then recovers $\hat{\mathbf{z}}$ from (7.143).

The previous approach is well known. The interesting point is that the MMSE solution can be achieved with a totally decentralized approach, where every sensor interacts only with its neighbors, with no need to send any data to a fusion center. The proposed approach is based on a very simple iterative procedure, where each node initializes a state variable with the local measurement, let us say $z_i[0] = g(x_i, y_i)$, and then it updates its own state by taking a linear combination of its neighbors' states, similarly with what happens with consensus algorithms, but with coefficients computed in order to solve the new problem.

More specifically, denoting by $\mathbf{z}[k]$, the N -size vector containing the states of all the nodes, at iteration k , and by \mathbf{g} the vector containing the initial measurements collected by all the nodes, the vector $\mathbf{z}[k]$ evolves according to the following linear state equation:

$$\mathbf{z}[k+1] = \mathbf{W}\mathbf{z}[k], \quad \mathbf{z}[0] = \mathbf{g} \in \mathbb{R}^N, \quad (7.144)$$

where $\mathbf{W} \in \mathbb{R}^{N \times N}$ is typically a *sparse* (not necessarily symmetric) matrix. The network topology is reflected into the sparsity of \mathbf{W} . In particular, the number of nonzero entries of, let us say, the i th row is equal to the number of neighbors of node i . In a WSN, the neighbors of a node are the nodes falling within the coverage area of that node, i.e., within a circle centered on the location of the node, with radius dictated by the transmit power of the node and by the power attenuation law. Our goal is to find the nonnull coefficients of \mathbf{W} that allow the convergence of $\mathbf{z}[k]$ to the vector $\hat{\mathbf{z}}$ given in (7.143). In general, not every network topology guarantees the existence of a solution of this problem. In the following, we will show that a solution exists only if each node has a number of neighbors greater than the dimension r of the useful signal subspace.

Let us denote by $\mathbf{P}_{\mathcal{R}(\mathbf{U})} \in \mathbb{R}^{N \times N}$ the orthogonal projector onto the r -dimensional subspace of \mathbb{R}^N spanned by the columns of (\mathbf{U}) , where $\mathcal{R}(\cdot)$ denotes the range space operator and $\mathbf{U} \in \mathbb{R}^{N \times r}$ is a full-column rank matrix, assumed, w.l.o.g., to be semi-unitary. System (7.144) converges to the desired orthogonal projection of the initial value vector $\mathbf{z}[0] = \mathbf{g}$ onto $\mathcal{R}(\mathbf{U})$, for any given $\mathbf{g} \in \mathbb{R}^N$, if and only if

$$\lim_{k \rightarrow +\infty} \mathbf{z}[k] = \lim_{k \rightarrow +\infty} \mathbf{W}^k \mathbf{g} = \mathbf{P}_{\mathcal{R}(\mathbf{U})} \mathbf{g}, \quad (7.145)$$

i.e.,

$$\lim_{k \rightarrow +\infty} \mathbf{W}^k = \mathbf{P}_{\mathcal{R}(\mathbf{U})}. \quad (7.146)$$

Resorting to basic algebraic properties of discrete-time systems, it is possible to derive immediately some basic properties of \mathbf{W} . In particular, denoting with OUD the Open Unit Disk, i.e., the set $\{x \in \mathbb{C} : |x| < 1\}$, a matrix \mathbf{W} is *semistable* if its spectrum $\text{spec}(\mathbf{W})$ satisfies $\text{spec}(\mathbf{W}) \subset \text{OUD} \cup \{1\}$ and, if $1 \in \text{spec}(\mathbf{W})$, then 1 is semisimple, i.e., its algebraic and geometric multiplicities coincide. If \mathbf{W} is semistable, then [73, p.447]

$$\lim_{k \rightarrow +\infty} \mathbf{W}^k = \mathbf{I} - (\mathbf{I} - \mathbf{W})^\sharp (\mathbf{I} - \mathbf{W}), \quad (7.147)$$

where $^\sharp$ denotes group generalized inverse [73, p.228]. Furthermore, setting, without loss of generality, the matrix \mathbf{W} in the form $\mathbf{W} = \mathbf{I} - \epsilon \mathbf{L}$, (7.147) can be rewritten as

$$\lim_{k \rightarrow +\infty} \mathbf{W}^k = \mathbf{I} - \mathbf{L}^\sharp \mathbf{L}. \quad (7.148)$$

But $\mathbf{I} - \mathbf{L}^\sharp \mathbf{L}$ is the projector onto the null-space of \mathbf{L} . Hence, we can state the following:

Proposition 1. *Given the dynamical system in (7.144) and the projection matrix $\mathbf{P}_{\mathcal{R}(\mathbf{U})}$, the vector $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ is globally asymptotically stable for any fixed $\mathbf{z}[0] \in \mathbb{R}^N$, if and only if the following conditions are satisfied:*

- i. \mathbf{L} has a nullspace of dimension r , spanned by the columns of \mathbf{U} ;
- ii. \mathbf{L} and ϵ must be chosen so that \mathbf{W} is semistable.

Alternatively, the previous conditions can be rewritten equivalently in the following form

Proposition 2. Given the dynamical system in (7.144) and the projection matrix $\mathbf{P}_{\mathcal{R}(\mathbf{U})}$, the vector $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ is globally asymptotically stable for any fixed $\mathbf{z}[0] \in \mathbb{R}^N$, if and only if the following conditions are satisfied:

$$\mathbf{W}\mathbf{P}_{\mathcal{R}(\mathbf{U})} = \mathbf{P}_{\mathcal{R}(\mathbf{U})}, \quad (\text{C.1})$$

$$\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{W} = \mathbf{P}_{\mathcal{R}(\mathbf{U})}, \quad (\text{C.2})$$

$$\rho(\mathbf{W} - \mathbf{P}_{\mathcal{R}(\mathbf{U})}) < 1, \quad (\text{C.3})$$

where $\rho(\cdot)$ denotes the spectral radius operator [23]. \square

Remark 1. Conditions C.1,C.2,C.3 have an intuitive interpretation. In particular, C.1 and C.2 state that, if system (7.144) asymptotically converges, then it is guaranteed to converge to the desired value. In fact, C.1 guarantees that the projection of vector $\mathbf{z}[k]$ onto $\mathcal{R}(\mathbf{U})$ is an invariant quantity for the dynamical system, implying that the system in (7.144), during its evolution, keeps the component $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ of $\mathbf{z}[0]$ unaltered. At the same time, C.2 makes $\mathbf{P}_{\mathcal{R}(\mathbf{U})}\mathbf{z}[0]$ a fixed point of matrix \mathbf{W} and thus a potential accumulation point for the sequence $\{\mathbf{z}[k]\}_k$. Both conditions C.1 and C.2 do not state anything about the convergence of the dynamical system. This is guaranteed by C.3, which imposes that all the modes associated to the eigenvectors orthogonal to $\mathcal{R}(\mathbf{U})$ are asymptotically vanishing.

Remark 2. The conditions C.1,C.2,C.3 contain, as a special case, the convergence conditions of average consensus algorithm. In fact, it is sufficient to set in (7.145), $r = 1$ and $\mathbf{U} = \mathbf{u} = \frac{1}{\sqrt{N}}\mathbf{1}_N$, where $\mathbf{1}_N$ is the N -length vector of all ones. In such a case, C.1,C.2,C.3 can be restated as following: the digraph associated to the network described by \mathbf{W} must be strongly connected and balanced.

The previous conditions do not make any explicit reference to the sparsity of matrix \mathbf{W} . However, when we consider a sparse matrix, reflecting the network topology, additional conditions are necessary to make sure that the previous conditions are satisfied. In other words, not every network topology is able to guarantee the asymptotic projection onto a prescribed signal subspace. One basic question is then what network topology is able to guarantee the convergence to a prescribed projector. We provide now the conditions on the sparsity of \mathbf{W} , or equivalently \mathbf{L} , guaranteeing the desired convergence.

From condition (i) of Proposition 1, given the matrix \mathbf{U} , \mathbf{L} must satisfy the equation $\mathbf{L}\mathbf{U} = \mathbf{0}$. Let us assume that every row of \mathbf{L} has K nonzero entries and let us indicate with $\{i_{j1}, \dots, i_{jK}\}$ the set of the column indices corresponding to the nonzero entries of the j th row of \mathbf{L} . Hence, every row of \mathbf{L} must satisfy the following equation

$$\begin{pmatrix} u_1(i_{11}), & u_1(i_{12}), & \dots & u_1(i_{1K}) \\ u_2(i_{21}), & u_2(i_{22}), & \dots & u_2(i_{2K}) \\ \vdots & \vdots & \ddots & \vdots \\ u_r(i_{r1}), & u_r(i_{r2}), & \dots & u_r(i_{rK}) \end{pmatrix} \begin{pmatrix} l_{j1} \\ \vdots \\ l_{jK} \end{pmatrix} = \mathbf{0}. \quad (7.149)$$

To guarantee the existence of a nontrivial solution to (7.149), the matrix on the left hand side must have a kernel of dimension at least one. This requires K to be strictly greater than r , the dimension of the signal subspace. Since the number of nonzero entries of, let us say the j th, row of \mathbf{L} is equal to the number of neighbors of node j plus one (the coefficient multiplying the state of node i itself), this implies that the minimum number K of neighbors of each node must be at least equal to the dimension

r of the signal subspace. Of course this condition is necessary but not sufficient. It is also necessary to check that the sparse matrix \mathbf{L} built with rows satisfying (7.149), with $j = 1, \dots, N$, had rank $N - r$. This depends on the location of the nodes and on the specific choice of the orthogonal basis.

An example can be useful to illustrate the benefits achievable with the proposed technique. We consider the case where the observation is corrupted by a multiplicative, spatially uncorrelated, noise, which models, for example a fading effect. Let us denote with $P(x_i, y_i) = A(x_i, y_i)S(x_i, y_i)$ the measurement carried out from node i , located in the point of coordinates (x_i, y_i) , where $S(x_i, y_i)$ denotes the useful field, whereas $A(x_i, y_i)$ represents fading. We consider, for instance, a useful signal composed by $N_s = 4$ transmitters and we assume a polynomial power attenuation, so that the useful signal measured at the point of coordinates (x, y) is

$$S(x, y) = \sum_{i=1}^{N_s} \frac{P_i}{1 + ((x - x_i)^2 + (y - y_i)^2)/\sigma^2}, \quad x \in \left[-\frac{L}{2}, \frac{L}{2}\right], y \in \left[-\frac{L}{2}, \frac{L}{2}\right] \quad (7.150)$$

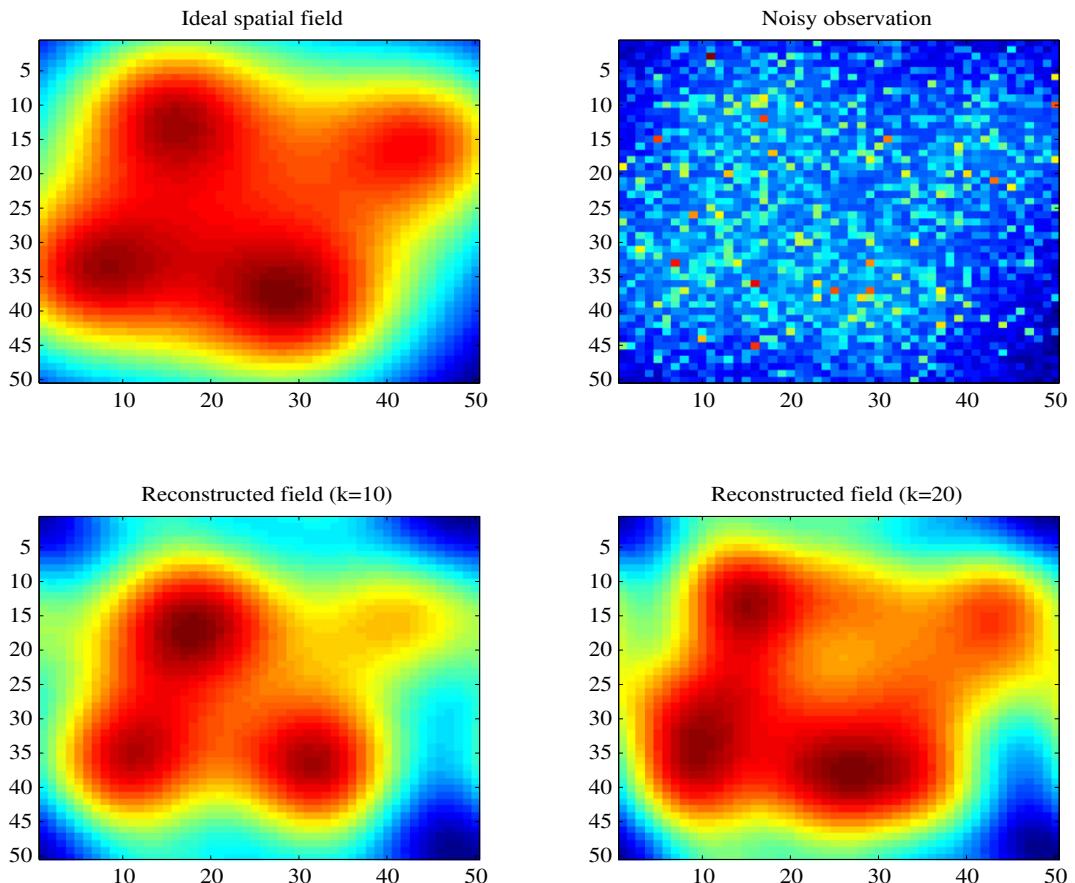
where P_i is the power emitted by source i , located at (x_i, y_i) , and σ specifies the power spatial spread. Furthermore, fading is modeled as a spatially uncorrelated multiplicative noise. The sensor network is composed of 2500 nodes uniformly distributed over a 2D grid. All the transmitters use the same power, i.e., $P_i = P$ in (7.150), and the noise has zero mean and variance $\sigma_n^2 = P$.

In this case, it is useful to apply a homomorphic filtering to the measured field. In particular, we take the log of the measurement, thus getting $\log(P(x_i, y_i)) = \log(S(x_i, y_i)) + \log(A(x_i, y_i))$. To smooth out the undesired effect of fading, we assume a signal model composed by the superposition of 2D sinusoids, so that the columns of the matrix \mathbf{U} in (7.138) are composed of signals of the form $\sin(2\pi(mx + ny)/L)$, and $\cos(2\pi(mx + ny)/L)$, with $m, n = 0, 1, \dots$. We set the initial value of the state of each node equal to $\log(P(x_i, y_i))$ and we run the distributed projection algorithm described above. After convergence, we simply take the exp of the result.

Figure 7.18 shows an example of application. In particular, the spatial behavior of the useful signal power is shown in the top left plot, while the observation corrupted by fading is reported in the top right figure. It is useful to consider that, in the example at hand, the useful signal would require a Fourier series expansion with an infinite number of terms to null the modeling error. Conversely, in our example, we used two different orders, $k = 10$ and $k = 20$. The corresponding reconstructions are shown in the bottom figures. From Figure 7.18 it is evident the capability of the proposed distributed approach to provide a significant attenuation of the fading phenomenon, without destroying valuable signal variations.

2.07.7 Minimum energy consensus

Although distributed algorithms to achieve consensus have received a lot of attention because of their capability of reaching optimal decisions without the need of a fusion center, the price paid for this simplicity is that consensus algorithms are inherently iterative. As a consequence the iterated exchange of data among the nodes might cause an excessive energy consumption. Hence, to make consensus algorithms really appealing in practical applications, it is necessary to minimize the energy consumption necessary to reach consensus. The network topology plays a fundamental role in determining the convergence rate [74]. As the network connectivity increases, so does the convergence rate. However, a highly connected network entails a high power consumption to guarantee reliable direct links between the nodes. On the other hand, if the network is minimally connected, with only neighbor nodes

**FIGURE 7.18**

Example of field reconstruction in the presence of fading: ideal spatial field (top left); measured field (top right); field reconstructed with order $k = 10$ (bottom left) and $k = 20$ (bottom right) [75].

connected to each other, a low power is spent to maintain the few short range links, but, at the same time, a large convergence time is required. Since what really matters in a WSN is the overall energy spent to achieve consensus, in [76,77] it was considered the problem of finding the optimal network topology that minimizes the overall energy consumption, taking into account convergence time and transmit powers *jointly*. More specifically, in [77] it is proposed a method for optimizing the network topology and the power allocation across every link in order to minimize the energy necessary to achieve consensus. Two different types of networks are considered: a) deterministic topologies, where node positions are arbitrary, but known; b) random geometries, where the unknown node locations are modeled as random variables. We will now review the methodology used in both cases.

2.07.7.1 Optimization criterion

By considering only the power spent to enable wireless communications, the overall energy consumption to reach consensus can be written as the product between the sum of the power P_{tot} necessary to establish the communication links among the nodes and the number of iterations N_{it} necessary to achieve consensus. The exchange of information among the nodes is supposed to take place in the presence of a slotted system, with a medium access control (MAC) mechanism that prevents packet collisions. The number of iterations can be approximated as $N_{\text{it}} = T_c/T_s$ where T_s denotes the duration of a time slot unit and

$$T_c = -\frac{\log(\gamma)}{\lambda_2(\mathbf{L})}$$

is the convergence time defined as the time necessary for the slowest mode of the dynamical system (7.23) to be reduced by a factor $\gamma \ll 1$. The total power spent by the network in each iteration is then $P_{\text{tot}} = \sum_{i,j} a_{ij} p_{ij}$ where the coefficient $p_{ij} = p_{ji}$, $i \neq j$ denotes the power transmitted by node i to node j , while the binary coefficients a_{ij} assess the presence ($a_{ij} = 1$) of a link between nodes i and j or not ($a_{ij} = 0$). Our goal is to minimize the energy consumption expressed by the following metric

$$\mathcal{E} = P_{\text{tot}} N_{\text{it}} = K \frac{\sum_{i=1}^N \sum_{j=1}^N a_{ij} p_{ij}}{\lambda_2(\mathbf{L}(\mathbf{a}))}, \quad (7.151)$$

where K incorporates all irrelevant constants, N is the number of sensors and $\mathbf{L}(\mathbf{a})$ is the Laplacian matrix depending on the vector $\mathbf{a} = \mathbf{A}(:)$ containing all the coefficients a_{ij} . More specifically, we aim to find the set of active links, i.e., the non-zero coefficients a_{ij} , and the powers p_{ij} that minimize the energy consumption (7.151), under the constraint of guaranteeing network connectivity, i.e., enforcing $\lambda_2(\mathbf{L}(\mathbf{a})) > 0$. The problem can be formulated as follows [77]:

$$\begin{aligned} \min_{\mathbf{a}, \mathbf{p}} \quad & \frac{\sum_{i=1}^N \sum_{j=1}^N a_{ij} p_{ij}}{\lambda_2(\mathbf{L}(\mathbf{a}))} \\ \text{s.t. } \epsilon \leq \lambda_2(\mathbf{L}(\mathbf{a})) \quad & [\mathbf{P.0}] \\ a_{ij} \in \{0, 1\} \quad & \\ p_{ij} \geq 0 \quad & \forall i, j = 1, \dots, N \end{aligned} \quad (7.152)$$

where ϵ is an arbitrarily small positive constant used to ensure network connectivity and \mathbf{p} is the vector with entries p_{ij} . Since the topology coefficients are binary variables, [P.0] is a combinatorial problem, with complexity increasing with the size N of the network as $2^{N(N-1)/2}$. In [77] we have modified [P.0] in order to convert it into a convex problem, with negligible performance losses. A first simplification comes from observing that the coefficients a_{ij} and p_{ij} are dependent of each other through the radio propagation model so that the set of unknowns can be reduced to the set of powers p_{ij} . More specifically, by assuming flat fading channel, we can assume that the power p_{Rj} received by node j when node i transmits is given by

$$p_{Rj} = \frac{p_{ij}}{1 + (r_{ij}/r_0)^\eta}, \quad (7.153)$$

where r_{ij} is the distance between nodes i and j , η is the path loss exponent, and the parameter r_0 corresponds to the so called Fraunhofer distance. We have included in the denominator the unitary term to avoid the unrealistic situation in which the received power could be greater than the transmitted one. Given the propagation model (7.153), the relation between the power coefficients p_{ij} and the topology

coefficients a_{ij} is then

$$a_{ij} = \begin{cases} 1 & \text{if } p_{ij} > p_{\min} \left[1 + \left(\frac{r_{ij}}{r_0} \right)^\eta \right], \\ 0 & \text{otherwise} \end{cases}, \quad (7.154)$$

where p_{\min} is the minimum power needed at the receiver side to establish a communication. In [77] we have shown how to relax this relation in order to simplify the solution of the optimal topology control problem considering both the deterministic and random topology.

2.07.7.2 Optimal topology and power allocation for arbitrary networks

In the case where the distances between the nodes are known, to find the optimal solution of problem [P.0] involves a combinatorial strategy that makes the problem numerically very hard to solve. In [77], we have relaxed problem [P.0] so that, instead of requiring a_{ij} to be binary, we assume a_{ij} to be a real variable belonging to the interval $[0, 1]$. This relaxation is the first step to transform the previous problem into a *convex* problem. More specifically, we have introduced the following relationship between the coefficients a_{ij} and the distances r_{ij} :

$$a_{ij} = \frac{1}{1 + (r_{ij}/r_{c_{ij}})^\alpha}, \quad (7.155)$$

where α is a positive coefficient and $r_{c_{ij}}$ is the coverage radius, which depends on the transmit power. According to (7.155), a_{ij} is close to one when node j is within the coverage radius of node i , i.e., $r_{ij} \ll r_{c_{ij}}$, whereas a_{ij} is close to zero, when $r_{ij} \gg r_{c_{ij}}$. The switching from zero to one can be made steeper by increasing the value of α . In [77] we have found the coefficients p_{ij} as a function of a_{ij}

$$p_{ij} = q(a_{ij}) = p_{\min} + k_1 \left(\frac{a_{ij}}{1 - a_{ij}} \right)^{\eta/\alpha}, \quad (7.156)$$

with $k_1 = p_{\min} \frac{r_{ij}^\eta}{r_0^\eta}$. Consequently, we can reduce the set of variables to the only power vector \mathbf{p} and problem [P.0] can be relaxed into the following problem:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \frac{\mathbf{p}^T \mathbf{1}}{\lambda_2(\mathbf{L}(\mathbf{p}))} \\ \text{s.t.} \quad & \epsilon \leq \lambda_2(\mathbf{L}(\mathbf{p})) \quad [\mathbf{P.1}], \\ & p_{\min} \mathbf{1} \leq \mathbf{p} \end{aligned} \quad (7.157)$$

The first important result proved in [77] is that the problem [P.1] is a convex-concave fractional problem if $\eta \geq \alpha$, so that we can use one of the methods that solve quasi-convex optimization problems, see e.g., [78, 79]. In [77] we have used the nonlinear parametric formulation proposed in [79]. Hence we have further converted the convex-concave fractional problem [P.1] into the following equivalent parametric problem in terms of vector \mathbf{a} , i.e.,

$$\begin{aligned} \min_{\mathbf{a}} \quad & \phi(\mathbf{a}) - \mu \lambda_2(\mathbf{L}(\mathbf{a})) \\ \text{s.t.} \quad & \epsilon \leq \lambda_2(\mathbf{L}(\mathbf{a})) \quad [\mathbf{P.2}], \\ & \mathbf{0} \leq \mathbf{a} < \mathbf{1}, \end{aligned} \quad (7.158)$$

where $\phi(\mathbf{a}) = \sum_{i=1}^N \sum_{j=1, i \neq j}^N q(a_{ij})$ and μ controls the trade-off between total transmit power and convergence time.

The optimization problem [P.2] is a convex parametric problem [77] and an optimal solution can be found via efficient numerical tools. Furthermore, using Dinkelbach's algorithm [79], we are also able to find the optimal parameter μ in [P.2].

2.07.7.3 Numerical examples

Since our optimization procedure is based on a relaxation technique, we have evaluated the impact of the relaxation on the final topology and performance.

More specifically, the topology coefficients a_{ij} obtained by solving [P.2] are real variables belonging to the interval $[0, 1]$, so that, to obtain the network topology, it is necessary a quantization step to convert them into binary values, 1 or 0, by comparing each a_{ij} with a threshold a_{th} . It has been shown that the loss in terms of optimal energy due to the relaxation of the original problem is negligible. To evaluate the impact of thresholding operation, in Figure 7.19 we show the topologies obtained by solving problem

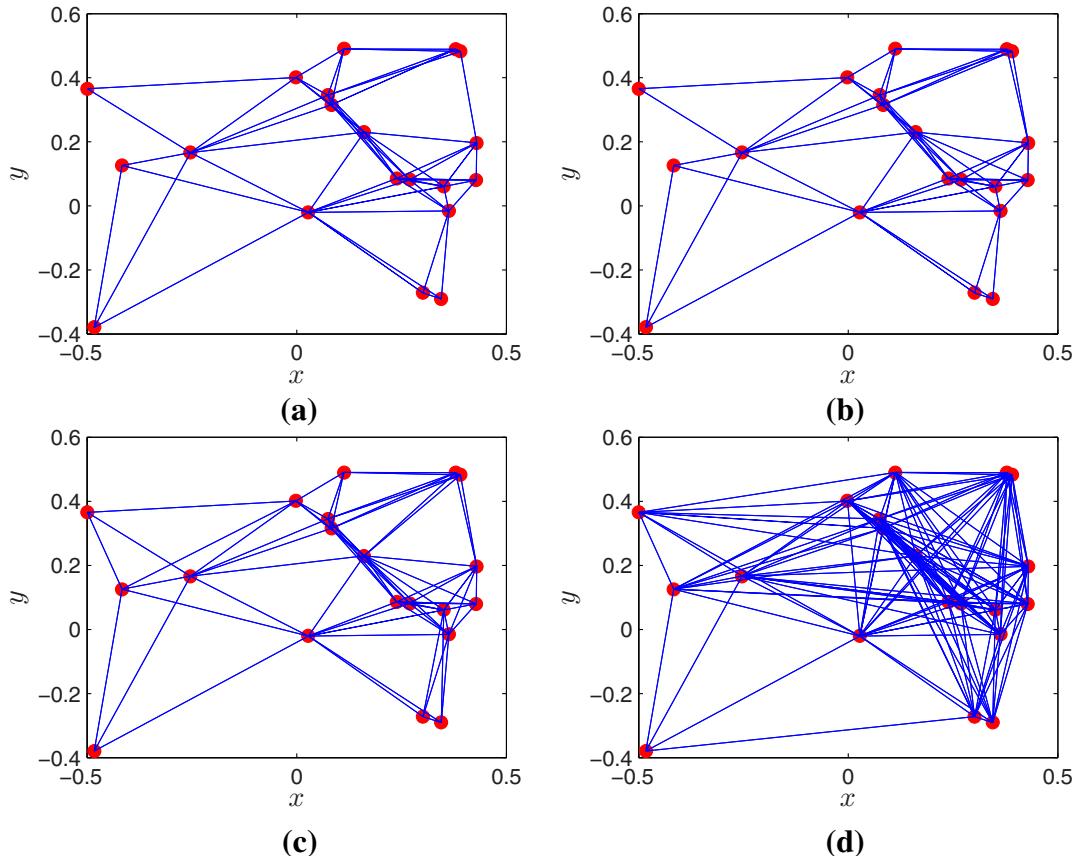
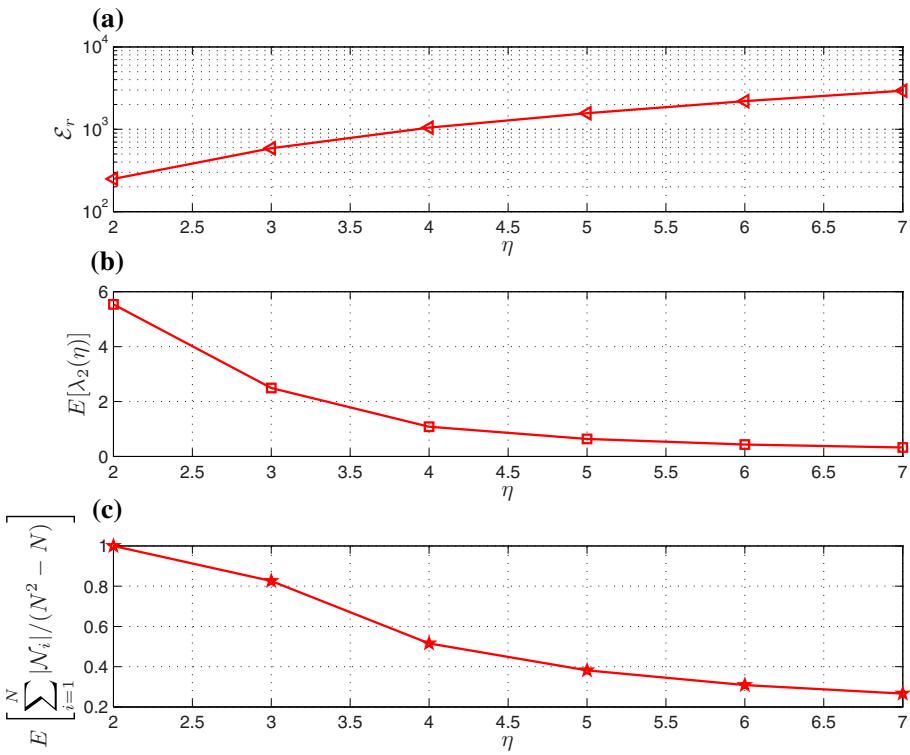


FIGURE 7.19

Optimal topologies, for different threshold values and $\eta = 6$: (a) $a_{th} = 0.09$; (b) $a_{th} = 0.05$; (c) $a_{th} = 10^{-4}$; (d) $a_{th} = 10^{-7}$.

**FIGURE 7.20**

Average value of (a) energy; (b) $\lambda_2(\mathbf{L})$; (c) fraction of active links vs. path loss η for $a_{th} = 0.09$.

[P.2], for a network composed of $N = 20$ nodes, using different values of a_{th} and assuming $\eta = 6$. Comparing the four cases reported in Figure 7.19, we can note that for a large range of values of a_{th} , the final topology is practically the same, while only for very low values of the threshold (i.e., case (d)), we can observe a sensitive change of topology. This means that the relaxation method is robust against the choice of the final threshold.

The previous results pertain to a specific realization of the node locations. To provide results of more general validity, in Figure 7.20, we report the average value of (a) the energy (E_r) (b) $\lambda_2(\mathbf{L})$, and (c) fraction of active links $\frac{\sum_{i=1}^N |\mathcal{N}_i|}{N(N-1)}$, as a function of the path loss exponent η , setting $a_{th} = 0.09$. From Figure 7.20, we observe that when the attenuation is high (i.e., η is large), reducing the number of links (making the topology sparser) is more important than reducing convergence time. Conversely, when the attenuation is low (i.e., η is small), increasing network connectivity is more important than reducing power consumption.

2.07.7.4 Minimization of the energy consumption over random geometric graphs

Let us consider now the problem of minimizing the energy consumption for a sensor network modeled as a random geometric graph. We will use the symbol $G(N, r)$ to indicate an RGG composed of N points, with coverage radius r .

In [74], it has been shown that the degree of an RGG $G(N, r)$ of points uniformly distributed over a two-dimensional unit torus⁷ is equal to

$$d(N) = \pi r^2 N \quad (7.159)$$

with high probability, i.e., with probability $1 - 1/N^2$, if the radius behaves as $r_0(N)$ in (7.2). This implies that if the coverage radius is chosen so as to guarantee connectivity with high probability, an RGG tends to behave, asymptotically, as a regular graph. In order to calculate the convergence rate we have to derive the second eigenvalue of the Laplacian, $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where \mathbf{D} is the degree matrix and \mathbf{A} is the adjacency matrix. From (7.159), $\mathbf{D} = \pi r^2 N \mathbf{I}$, so that we only need to calculate the second largest eigenvalue of \mathbf{A} . In Appendix A.2, we study the asymptotic behavior of the spectrum of \mathbf{A} and the result is that the second largest eigenvalue of \mathbf{L} tends asymptotically to

$$\lambda_2(\mathbf{L}) = \pi N r^2 - N r J_1(2\pi r), \quad (7.160)$$

where r is the coverage radius of each node.

2.07.7.4.1 An analytic approach for minimizing the energy consumption

In [77] we studied the energy minimization problem for RGG's, exploiting the previous analytic expressions. In the random topology case, since the distances are unknown, we cannot optimize the power associated with each link. However, we can seek the common transmit power that minimizes energy consumption. Thus, in the random setting we assume a broadcast communication model, where each node broadcasts the value to be shared with its neighbors. In the lack of any information about distances among the nodes, we assume that each node uses the same transmit power. In this case, the network topology can be modeled as a random graph model. In [80, 81] it has been shown that the dynamical system $\dot{\mathbf{x}}(t) = -\mathbf{L}\mathbf{x}(t)$ converges to consensus almost surely, i.e., $\Pr\{\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}^* \mathbf{1}\} = 1$ assuming that each node has a coverage radius so that the network is asymptotically connected with probability one. Then the rate of convergence to consensus is given [80, 81] by $E[e^{-2T_s \lambda_2(\mathbf{L})}]$. In [77] we proved that the convergence rate can be approximated as

$$E[e^{-2T_s \lambda_2}] \approx e^{-2T_s E[\lambda_2]} \quad (7.161)$$

so that the energy spent to achieve consensus can now be expressed as

$$\mathcal{E} = K \frac{N p}{2 E[\lambda_2(\mathbf{L}(p))]} \quad (7.162)$$

This is the performance metric we wish to minimize in the random scenario, with respect to the single unknown p .

⁷A torus geometry is typically used to get rid of border effects.

In particular, using the asymptotic expression (7.160) for the algebraic connectivity, we can introduce the following metric

$$\mathcal{E}(r) = \frac{N p_{\min} [1 + (r/r_0)^\eta]}{N \pi r^2 - r N J_1(2\pi r)} \quad (7.163)$$

that is a convex function of r , for $r_0(N) \leq r \leq 0.5$, where $r_0(N)$, behaves as in (7.2) to ensure connectivity.

Numerical examples. In Figure 7.21, we compare the value of $\mathcal{E}(r)$ obtained by our theoretical approach and by simulation, for various values of the path loss exponent η . The results are averaged over 100 independent realizations of random geometric graphs composed of $N = 1000$ nodes. For each η , we indicate the pair of radius and energy providing minimum energy consumption by a circle (simulation) or a star (theory). It can be noted that the theoretical derivations provide a very good prediction of the performance achieved by simulation and, for each η , there is a coverage radius value that minimizes energy consumption.

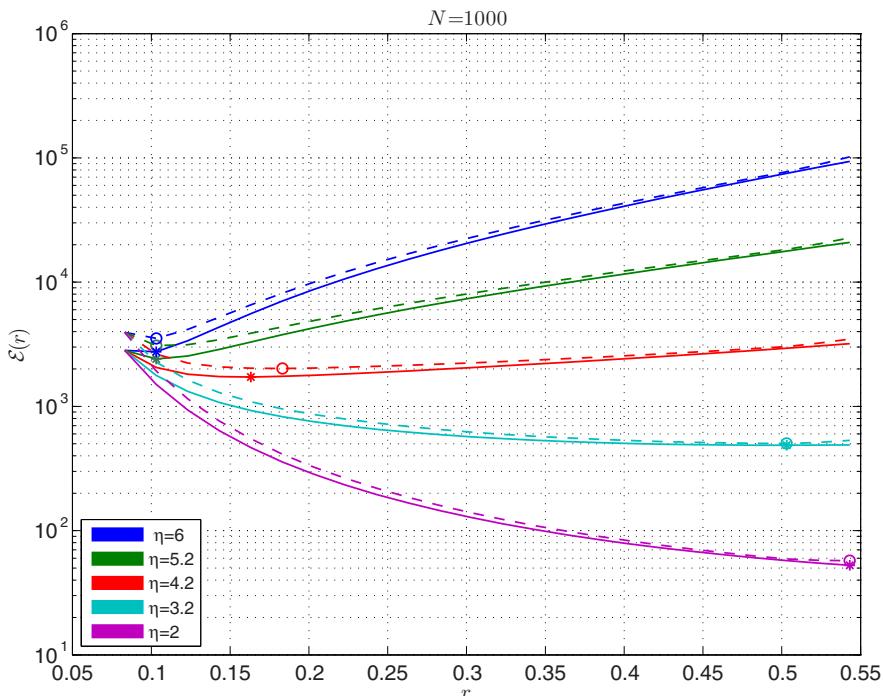


FIGURE 7.21

Global energy consumption versus transmission radius for an RGG; theoretical values (solid) and simulation results (dashed).

2.07.8 Matching communication network topology to statistical dependency graph

In Section 2.07.2.2, we saw that the topology of a sensor network observing a random field should depend on the structure of the graph describing the observed field. In this section, we recall a method proposed in [82, 83] to design the topology of a wireless sensor network observing a Markov random field in order to match the structure of the dependency graph of the observed field, under constraints on the power used to ensure the sensor network connectivity. As in [82, 83], our main task is to recover the sparsity of the dependency graph and to replicate it at the sensor network level, under the constraint of limiting the transmit power necessary to establish the link among the nodes. Also in this case, searching for an optimal topology is a combinatorial problem. To avoid the computational burden of solving the combinatorial problem, we propose an ad hoc relaxation technique that allows us to achieve the solution through efficient algorithms based on difference of convex problems.

Let us assume to have a network composed of N nodes, each one observing a spatial sample of a Gaussian Markov Random Field. We denote by $\mathbf{x} := (x_1, \dots, x_N)$ the vector of the observations collected by the N nodes and we assume that \mathbf{x} has zero mean and covariance matrix \mathbf{C} . The statistical dependency among the random variables x_i is well captured by the structure of the Markov graph whose vertices correspond to the random variables and whose links denote statistical dependencies among the variables. As discussed in 2.07.2.2 the main feature of a Markov graph is that it is sparse and there is no link between two nodes if and only if their observations are statistically independent. Moreover, if the random vector \mathbf{x} is also Gaussian, with covariance matrix \mathbf{C} , the sparsity of the Markov graph is completely specified by the sparsity of the precision matrix, which is the inverse of the covariance matrix, i.e., $\mathbf{B} := \mathbf{C}^{-1}$. On the other hand, the topology of the WSN can also be described by a graph, having adjacency matrix \mathbf{A} such that $a_{ij} \neq 0$ only if there is a physical link between nodes i and j . We use a simple propagation model such that there is a link between node i and j if the power received by node j exceeds a minimum power p_{\min} . The received power depends on the power $p_T(i, j)$ used by node i to transmit to node j and on the distance r_{ij} between nodes i and j through the equation

$$p_R(i, j) = \frac{p_T(i, j)}{1 + r_{ij}^\eta}, \quad (7.164)$$

where η is the path loss exponent.

As proposed in [82, 83], our goal is to design the topology of the WSN, and hence its adjacency matrix \mathbf{A} , in order to match as well as possible the topology of the dependency graph, compatibly with the power expenditure necessary to establish each link in the network. Without any power constraint, we would choose \mathbf{A} to be equal to $\mathbf{B} = \mathbf{C}^{-1}$, so as to reproduce the same sparsity of the dependency graph. Adding the power constraints, we will end up, in general, with a matrix \mathbf{A} different from \mathbf{B} . We measure the difference between the two matrices \mathbf{A} and \mathbf{B} using the so called Burg divergence, defined as

$$D_B(\mathbf{A}, \mathbf{B}) := \frac{1}{2} \text{trace} \left(\mathbf{AB}^{-1} - \mathbf{I} \right) - \frac{1}{2} \log (\det (\mathbf{AB}^{-1})). \quad (7.165)$$

Even though the Burg divergence does not respect all the prerequisites to be a distance, it holds true that $D_B(\mathbf{A}, \mathbf{B}) = 0$, if and only if $\mathbf{A} = \mathbf{B}$, otherwise, the divergence is strictly positive. If the matrices \mathbf{A} and \mathbf{B} are definite positive, the expression in (7.165) coincides with the Kullback-Leibler divergence

between the probability density function (pdf) of two Gaussian random vectors having zero mean and precision matrices \mathbf{A} and \mathbf{B} or, equivalently, covariance matrices \mathbf{A}^{-1} and \mathbf{C} .

2.07.8.1 Encouraging sparsity by preserving total transmit power

One of the most important tasks in wireless sensor networks is to minimize the energy consumption for reliable data transmission. This need can be accommodated by formulating the search for a sparse topology incorporating a penalization for the presence of links among distant nodes. The first strategy we propose is named Sparsity with Minimum Power (SMP) consumption. We consider both cases where the covariance matrix is perfectly known or estimated from the collected data.

Let us consider a wireless sensor network whose communication graph is a geometric graph, where each node communicates only with the nodes lying within its coverage area of radius r . We assume, initially, that the covariance matrix \mathbf{C} of the GMRF is perfectly known. Our goal is to find the optimal adjacency matrix \mathbf{A} that minimizes the divergence $D_B(\mathbf{A}, \mathbf{B})$ given in (7.165), under the constraint of limiting the transmit power necessary to maintain the links among the nodes of the WSN. This constraint can be incorporated in our optimization problem by introducing a penalty term given by the sum of the transmit powers over all active links, i.e.,

$$P_N(\mathbf{A}) = \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N p_T(i, j) \delta(a_{ij}), \quad (7.166)$$

where $p_T(i, j)$ denotes the power used by node i to transmit to node j and

$$\delta(a_{ij}) = \begin{cases} 0 & \text{if } a_{ij} = 0 \\ 1 & \text{otherwise,} \end{cases} \quad (7.167)$$

assuming that a_{ij} is different from zero only if the power $p_R(i, j)$ received by node j when node i transmits, as given in (7.164), exceeds a suitable minimum level p_{\min} , i.e., if

$$p_T(i, j) > (1 + r_{ij}^\eta) p_{\min}. \quad (7.168)$$

The optimization problem can then be formulated as

$$\min_{\mathbf{A} \in \mathbb{S}_{++}^N} D_B(\mathbf{A}, \mathbf{B}) + \rho P_N(\mathbf{A}), \quad (7.169)$$

where \mathbb{S}_{++}^N is the cone of definite positive symmetric $N \times N$ -dimensional matrices, while $\rho \geq 0$ is the penalty coefficient introduced to control sparsity. In fact, increasing the penalty coefficient, we assign a higher weight to power consumption so that sparse structures are more likely to occur.

Problem (7.169) is indeed quite hard to solve as the penalty function is a nonconvex discrete function. Optimization problems with a convex penalty have been largely considered in several signal processing applications, for example in compressed sensing [84] where these problems are often formulated as a penalized least-square problem in which sparsity is usually induced by adding a l_1 -norm penalty on the coefficients, as in Lasso algorithm [47].

Indeed, non-convex penalty functions such as l_q -norm, with $q < 1$, are even more effective to recover sparsity than l_1 -norm. Actually, using the so called l_0 norm would be even more effective to measure sparsity, even though the l_0 norm does not respect all requisites to be a norm.⁸ Here we adopt the so called Zhang penalty function analyzed in [85], i.e.,

$$z(a_{ij}) = \min\left(\frac{|a_{ij}|}{\epsilon}, 1\right) = \begin{cases} \frac{|a_{ij}|}{\epsilon} & \text{if } |a_{ij}| \leq \epsilon \\ 1 & \text{otherwise,} \end{cases} \quad (7.170)$$

where ϵ is an infinitesimal positive constant. Hence, by assuming $p_T(i, j) = (1 + r_{ij}^\eta)p_{\min}$, the second term in (7.169) can be written as

$$P_N(\mathbf{A}) = \sum_{i=1}^N \sum_{j=1}^N d_{ij} z(a_{ij}) = \text{trace}[z(\mathbf{A}), \mathbf{D}] \quad (7.171)$$

where \mathbf{D} is a $N \times N$ dimensional symmetric matrix with entries $d_{ij} = (1 + r_{ij}^\eta)p_{\min}$, $d_{ii} = 0$, $\forall i, j = 1, \dots, N$, while the matrix mapping $z(\mathbf{A})$ is defined applying the elementwise mapping $z(a_{ij}) : \mathbb{R} \rightarrow \mathbb{R}^+$ given in (7.170). The combinatorial problem in (7.169) can then be reformulated as

$$\min_{\mathbf{A} \in \mathbb{S}_{++}^N} D_B(\mathbf{A}, \mathbf{B}) + \rho \text{ trace}[z(\mathbf{A})\mathbf{D}]. \quad (7.172)$$

Unfortunately the second term in (7.172), is not convex so that the problem we have to solve is a nonconvex, nonsmooth optimization problem. Nevertheless, in [82, 83] we reformulated this problem as a difference of convex (DC) problem. Before proceeding, we simply illustrate how to extend our approach to the case where the covariance matrix of the observed vector is not known but estimated from the data. In such a case, the matrix \mathbf{C} in (7.172) is substituted by the estimated matrix $\widehat{\mathbf{C}}$, whose entry \widehat{C}_{ij} is

$$\widehat{C}_{ij} = \frac{1}{K} \sum_{k=1}^K x_i(k)x_j(k), \quad (7.173)$$

where $x_i(k)$ is the observation collected by node i , at time k , with $k = 1, \dots, K$. The practical, relevant, difference is that while the true precision matrix \mathbf{B} is sparse by hypothesis, the inverse of $\widehat{\mathbf{C}}$ in general is not sparse. Also in this case, encouraging sparsity in estimating the inverse of the covariance matrix can be beneficial to improve the quality of the estimation itself.⁹

The problem (7.172) can be reformulated as a Difference of Convex (DC) functions problem [86], by decomposing the function $z(a_{ij})$ as the difference of two convex functions $z(a_{ij}) = g_v(a_{ij}) - h(a_{ij})$ with $g_v(a_{ij}) = \frac{|a_{ij}|}{\epsilon}$ and

$$h(a_{ij}) = \begin{cases} 0 & \text{if } |a_{ij}| \leq \epsilon \\ \frac{|a_{ij}|}{\epsilon} - 1 & \text{otherwise.} \end{cases} \quad (7.174)$$

⁸The l_0 norm of a vector \mathbf{x} is defined as the number of nonzero entries of \mathbf{x} .

⁹Provided that the observed field is a Markov field.

Hence the optimization problem in (7.172) can be rewritten as

$$\min_{\mathbf{A} \in \mathbb{S}_{++}^N} D_B(\mathbf{A}, \mathbf{B}) + \rho \operatorname{trace}[(g_v(\mathbf{A}) - h(\mathbf{A})) \mathbf{D}]. \quad (7.175)$$

To solve this problem, we have used an iterative procedure, known as DC algorithm (DCA), based on the duality of DC programming. The usefulness of using DCA is that its convergence has been proved in [86] and it is simple to implement, as it iteratively solves a convex optimization problem. We refer the reader to [82, 83] for further analytical details.

2.07.8.2 Sparsification and estimation of the precision matrix

In this section we illustrate an alternative sparsification strategy that improves the estimate of the precision matrix with respect to the SMP strategy. In this alternative formulation, the penalty term is the sum of the absolute values of the entries of \mathbf{A} , weighted with the corresponding per-link transmit power consumption. In this way, although the power consumption should not be lower than the SMP method, we expect a sparse topology with a more accurate estimate of the precision matrix. We call this strategy Sparse Estimation Strategy (SES). The new problem is formulated as follows

$$\min_{\mathbf{A} \in \mathbb{S}_{++}^N} D_B(\mathbf{A}, \mathbf{B}) + \rho \operatorname{trace}[|\mathbf{A}| \mathbf{D}] \quad (7.176)$$

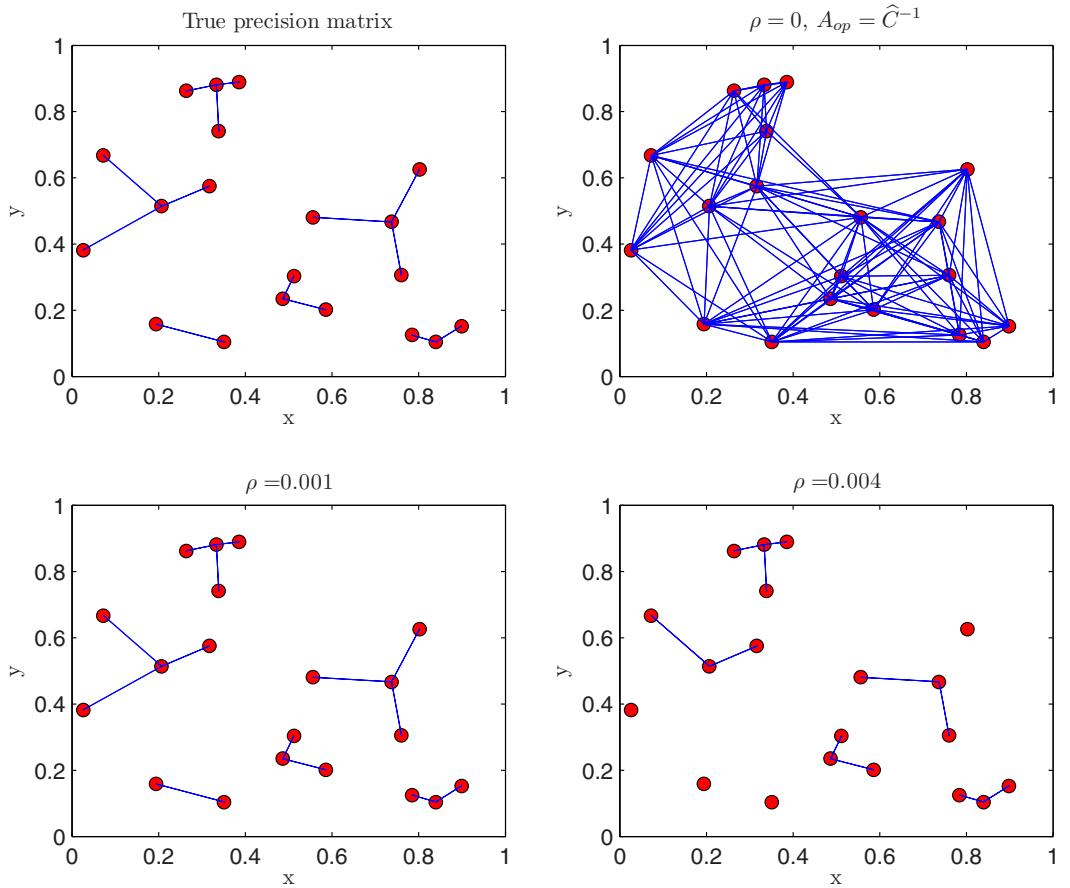
and it can be converted into a convex definite positive problem. In particular, splitting the matrix \mathbf{A} into the difference of two nonnegative matrices representing its positive and negative part, i.e., $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$, we can rewrite (7.176) as

$$\begin{aligned} \min_{\mathbf{A}^+, \mathbf{A}^- \in \mathbb{S}^N} \quad & D_B(\mathbf{A}^+ - \mathbf{A}^-, \mathbf{B}) + \rho \operatorname{trace}[(\mathbf{A}^+ + \mathbf{A}^-) \mathbf{D}] \\ \text{s.t.} \quad & \mathbf{A}^+ - \mathbf{A}^- \succ 0 \\ & \mathbf{A}^+ \geq \mathbf{0} \\ & \mathbf{A}^- \geq \mathbf{0}. \end{aligned} \quad (7.177)$$

This problem can be solved using standard numerical tools or by applying a projected gradient algorithm [87].

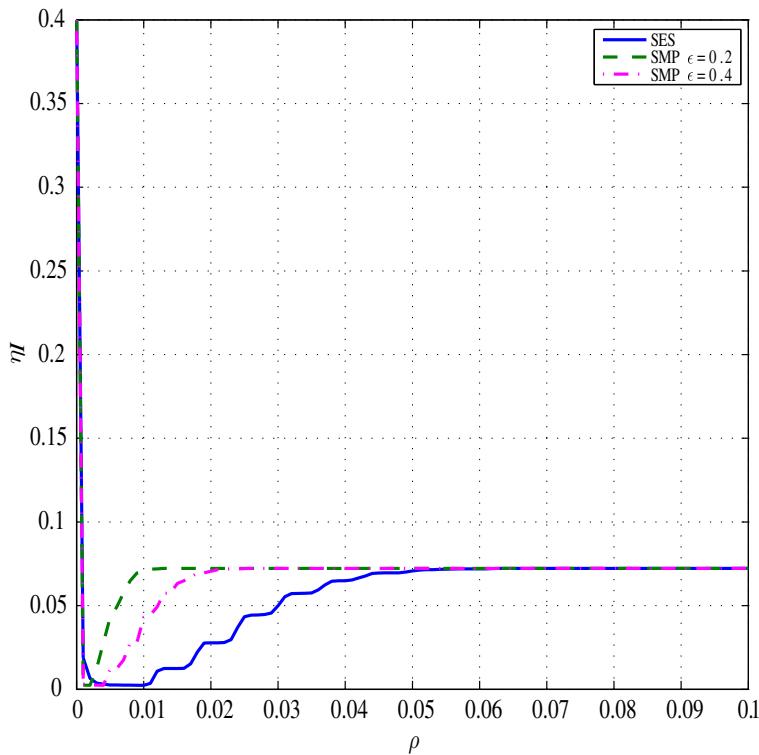
2.07.8.3 Numerical results

In this section we report some simulation results considering a sensors network composed of $N = 20$ nodes, uniformly deployed over a unit area square and observing correlated data from a GMRF. We adopt the Markov model proposed in [8], where the correlation between neighboring nodes is a decreasing function of their distance and the entries of the covariance matrix can be derived in closed form. In [82, 83] we have shown that even though the problem is not convex, the numerical results seem to indicate that the method always converges to the same value of the precision matrix entries, irrespective of the initializations. In Figure 7.22 we report the final optimal network topology referring to the case of a matrix estimated from the data. In particular, the top left plot of Figure 7.22 shows the true dependency graph and all other plots depict the network topologies obtained using the proposed SMP algorithm,

**FIGURE 7.22**

Optimal links configurations for the SMP strategy using the data estimated covariance matrix \hat{C} .

with different penalty coefficients. More specifically, the network topologies shown in Figure 7.22 are obtained by thresholding the values of the matrix A_{op} obtained through our SMP algorithm, i.e., the coefficients of matrix A_{op} are set to zero if $|a_{op}(i, j)| < 10^{-4}$. In the top right plot of Figure 7.22, it can be noted that the precision matrix achieved with a null penalty can be quite dense because of estimation errors. Nevertheless, it is interesting to observe that, as the penalty coefficient increases, the proposed method is not only able to recover the desired topology, but also to correct most of the errors due to estimation. We can say that the introduction of the penalty induces a robustness against estimation errors. Let us now compare the SMP method with the SES strategy, by considering a data estimated covariance matrix in the divergence term and averaging the simulation results over 100 independent realizations of the nodes deployment. Let us now evaluate the mismatch between the network topology and the

**FIGURE 7.23**

Fraction of incorrect links versus ρ for the SMP and SES strategies.

dependency graph, as a function of the penalty coefficient, obtained using the two proposed strategies. We assess the mismatch by counting the number of links appearing in the network topology, which do not appear in the dependency graph. To this end, Figure 7.23 shows the fraction of incorrect links η_I , normalized to the total number of links $N_T = N(N - 1)/2$, versus ρ . More specifically, considering the true and optimal precision matrices (A_t and A_{op} respectively), we can define $\eta_I = \frac{\sum_{i=1}^N \sum_{j=1, j>i}^N q_{ij}}{N_T}$ where $q_{ij} = 1$ if $\delta(a_t(i, j))$ and $\delta(a_{op}(i, j))$ are not equal, assuming $\delta(a_{ij}) = 1$ if $a_{ij} > 0$ and zero otherwise. From Figure 7.23, we can deduce that SES provides more correct links than SMP, as it achieves lower values of η_I .

2.07.9 Conclusions and further developments

In this article we have provided a general framework to show how an efficient design of a wireless sensor network requires a joint combination of in-network processing and communication. In particular, we have shown that inferring the structure of the graph describing the statistical dependencies among the observed data can provide important information on how to build the sensor network topology

and how to design the flow of information through the network. We have illustrated several possible network architectures where the global decisions, either estimation or hypothesis testing, are taken by a central node or in a totally decentralized way. In particular, various forms of consensus have been shown to be instrumental to achieve globally optimal performance through local interactions only. Consensus algorithms have then been generalized to more sophisticated signal processing techniques able to provide a cartography of the observed field. In a decentralized framework, the network topology plays an important role in terms of convergence time as well as structure of the final consensus value. Considering that most sensor networks exchange information through a wireless channel, we have addressed the problem of finding the network topology that minimizes the energy consumption required to reach consensus. Finally, we have showed how to match the network topology to the Markov graph describing the observed variables, under constraints imposed by the power consumption necessary to establish direct links among the sensor nodes.

Even though the field of distributed detection and estimation has accumulated an enormous amount of research works, there are still many open problems, both in the theoretical as well as in the application sides. In the following we make a short list of possible topics of future interest.

1. The general *multi-terminal source/channel coding* problem is still an open issue. The conventional paradigm established by the source/channel coding separation theorem does not hold for the multi-terminal case. This means that source coding should be studied jointly with channel coding.
2. Distributed decision establishes a *strict link between statistical signal processing and graph theory*. In particular, the network topology plays a fundamental role in the design of an efficient sensor network. In this chapter, we have shown some simple techniques aimed to matching the network topology to the statistical dependency graph of the observed variables, but significant improvements may be expected from cross-fertilization of methods from graph theory and statistical signal processing.
3. The design of fully decentralized detection algorithms has already received important contributions. Nevertheless, there are many open issues concerning the refinements of the local decision thresholds as a function of both local observations and the decisions taken from neighbors. In a more general setting, *social learning* is expected to play an important role in future sensor networks.
4. An efficient design of wireless sensor networks requires a *strict relation between radio resource allocation and decision aspects*, under physical constraints dictated by energy limitations or channel noise and interference. Some preliminary results have been achieved in the many-to-one setting, but the general many-to-many case needs to be thoroughly studied.
5. The application of wireless sensor networks to new fields may be easily expected. The important remark is that, to improve the efficiency of the network at various levels, it is necessary to take the application needs strictly into account in the network design. In other words, a *cross-layer design incorporating all layers* from the application down to the physical layer is especially required in sensor networks. Clearly, handling the complexity of the network will require some sort of layering, but this layering will not necessarily be the same as in telecommunication networks, because the requirements and constraints in the two fields are completely different.

Appendix A

In this appendix we briefly review some important notations and basic concepts of graph theory that have been adopted in the previous sections (for a more detailed introduction to this field see [13]).

A.1 Algebraic graph theory

Given N nodes let us define a directed graph or *digraph* $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ as a set of nodes $\mathcal{V} = \{v_i\}_{i=1}^N$ and a set of edges or links $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ where the links $e_{ij} \in \mathcal{E}$ connect the ordered pair of nodes (v_i, v_j) , with the convention that the information flows from v_j to v_i . In the case where a positive weight a_{ij} is associated to each edge, the digraph is called *weighted*. Let us assume that there are no loops, i.e., $a_{ii} = 0$.

The graph is called *undirected* if $e_{ij} \in \mathcal{E} \Leftrightarrow e_{ji} \in \mathcal{E}$. The in-degree and out-degree of node v_i are, respectively, defined as $\deg_{\text{in}} \triangleq \sum_{j=1}^N a_{ij}$ and $\deg_{\text{out}} \triangleq \sum_{j=1}^N a_{ji}$. In the case of undirected graphs $\deg_{\text{in}} = \deg_{\text{out}}$. Let \mathcal{N}_i denote the set of neighbors of node i , so that $|\mathcal{N}_i| \triangleq \deg_{\text{in}}(v_i)$.

The node v_i of a digraph is said to be *balanced* if and only if its in-degree and out-degree coincide, while a digraph is called *balanced* if and only if all its nodes are balanced.

We recall now the basic properties of the matrices associated to a digraph, as they play a fundamental role in the study of the connectivity of the network associated to the graph. Given a digraph \mathcal{G} , we introduce the following matrices associated with \mathcal{G} : (1) The $N \times N$ adjacency matrix \mathbf{A} whose entries a_{ij} are equal to the weight associated to the edge e_{ij} , or equal to zero, otherwise; (2) the degree matrix \mathbf{D} which is the diagonal matrix whose diagonal entries are $d_{ii} = \deg_{\text{in}}(v_i) = \sum_{j=1}^N a_{ij}$; (3) the weighted Laplacian matrix \mathbf{L} , defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ whose entries are

$$\ell_{ij} = \begin{cases} \deg_{\text{in}}(v_i) & \text{if } j = i, \\ -a_{ij} & \text{if } j \neq i. \end{cases} \quad (7.178)$$

According to this definition \mathbf{L} has the following properties: (a) its diagonal elements are positive; (b) it has zero row sum; (c) it is a diagonally row dominant matrix. It can be easily verified that $\mathbf{L}\mathbf{1} = \mathbf{0}$,¹⁰ i.e., zero is an eigenvalue of \mathbf{L} corresponding to a right eigenvector $\mathbf{1}$ in the $\text{Null}\{\mathbf{L}\} \supseteq \text{span}\{\mathbf{1}\}$, and all the other eigenvalues have positive real parts. Furthermore a digraph is balanced if and only if $\mathbf{1}$ is also a left eigenvector of \mathbf{L} associated with the zero eigenvalue or $\mathbf{1}^T \mathbf{L} = \mathbf{0}^T$. Note that for undirected graph the Laplacian matrix is a symmetric and then balanced matrix with non negative real eigenvalues.

The algebraic multiplicity of the zero eigenvalue of \mathbf{L} is equal to the number of connected components contained in \mathcal{G} . For undirected graphs \mathcal{G} is connected if and only if the algebraic multiplicity of the zero eigenvalue is 1, or, equivalently, $\text{rank}(\mathbf{L}) = N - 1$ if and only if \mathcal{G} is connected. Hence, if an undirected graph is connected, the eigenvector associated with the zero eigenvalue is $\mathbf{1}$, and the second smallest eigenvalue of \mathbf{L} , denoted as $\lambda_2(\mathbf{L})$ and called algebraic connectivity [22] of \mathcal{G} , is strictly positive.

A.1.1 Forms of connectivity for digraphs

Before to introduce several forms of graph connectivity [17] we have to define some useful concepts. A *strong path* of a digraph \mathcal{G} is a sequence of distinct nodes $v_1, v_2, \dots, v_p \in \mathcal{V}$ such that $(v_{j-1}, v_j) \in \mathcal{E}$, for $j = 2, \dots, p$. If $v_1 \equiv v_p$, the path is said to be closed. A *weak path* is a sequence of distinct nodes $v_1, v_2, \dots, v_p \in \mathcal{V}$ such that either $(v_{j-1}, v_j) \in \mathcal{E}$ or $(v_j, v_{j-1}) \in \mathcal{E}$, for $j = 2, \dots, p$. A closed strong path is said a *strong cycle*. A digraph with N nodes is a *directed tree* if it has $N - 1$ edges and there exists a node, called the root node, which can reach all the other nodes through an unique strong path. As a consequence a directed tree contains no cycles and every node, except the root, has one and only one incoming edge. A digraph is a *forest* if it consists of one or more directed trees. A subgraph

¹⁰We denote by $\mathbf{1}$ and $\mathbf{0}$ the vectors of all ones or zeros, respectively.

$\mathcal{G}_s = \{\mathcal{V}_s, \mathcal{E}_s\}$ of a digraph \mathcal{G} , with $\mathcal{V}_s \subseteq \mathcal{V}$ and $\mathcal{E}_s \subseteq \mathcal{E}$, is a directed spanning tree (or a spanning forest) if it is a directed tree (or a directed forest) and it has the same node set as \mathcal{G} .

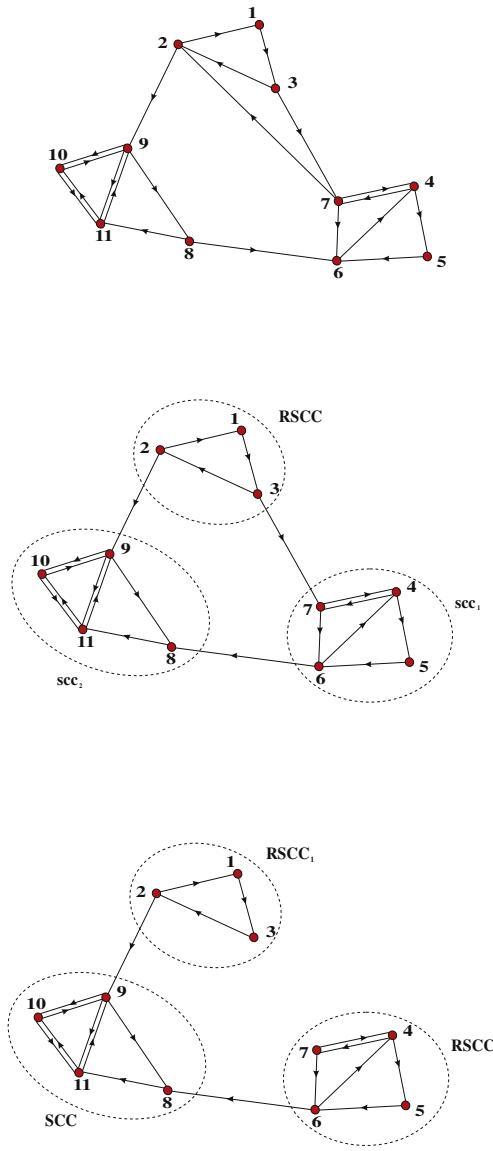
According to this definition we can define many forms of connectivity [17]: (a) a digraph is *strongly connected* (SC) if any ordered pair of distinct nodes can be joined by a strong path; (b) a digraph is *quasi strongly connected* (QSC) if, for every ordered pair of nodes v_i and v_j , there exists a node r that can reach both v_i and v_j via a strong path; (c) a digraph is *weakly connected* (WC) if any ordered pair of distinct nodes can be joined by a weak path; (d) a digraph is disconnected if it is not weakly connected. Note that for undirected graphs, the above notions of connectivity are equivalent. Moreover, it is easy to check that the quasi strong connectivity of a digraph is equivalent to the existence of a directed spanning tree in the graph.

A.1.2 Connectivity study from the condensation digraph

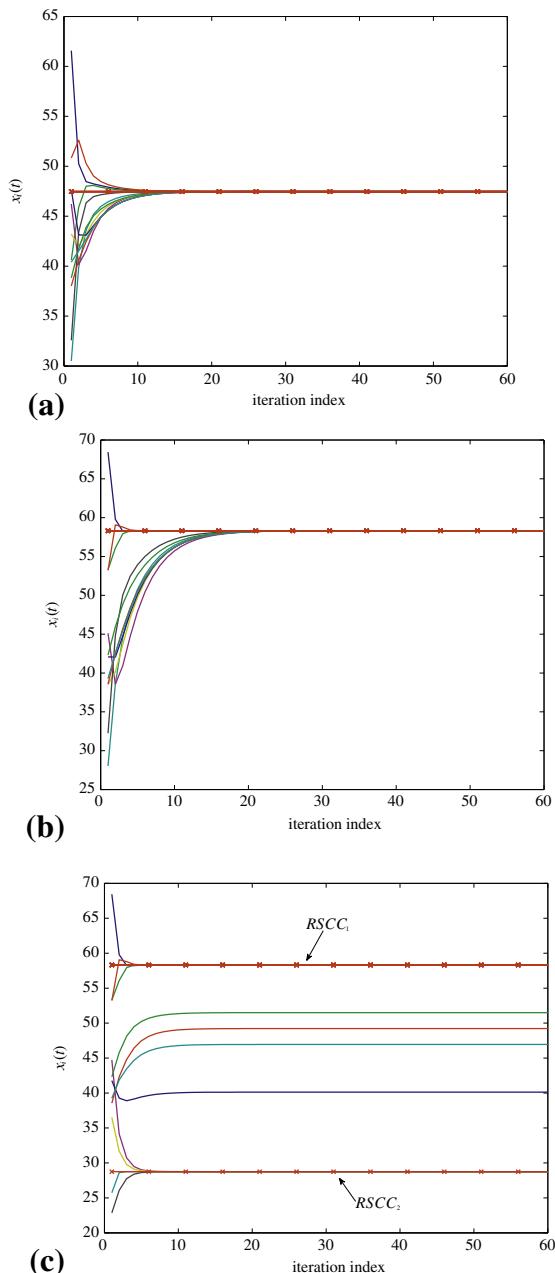
When a digraph \mathcal{G} is WC, it may still contain strongly connected subgraphs. A maximal subgraph of \mathcal{G} , which is also SC, is called a *strongly connected component* (SCC) of \mathcal{G} [17, 88]. Any digraph \mathcal{G} can be partitioned into SCCs, let us say $\mathcal{G}_k = \{\mathcal{V}_k, \mathcal{E}_k\}$ where $\mathcal{V}_k \subseteq \mathcal{V}$ and $\mathcal{E}_k \subseteq \mathcal{E}$ for $k = 1, \dots, r$. The connectivity properties of a digraph may be better studied by referring to its corresponding *condensation digraph*. We may reduce the original digraph \mathcal{G} to the condensation digraph $\mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\}$ by associating the node set \mathcal{V}_k of each SCC \mathcal{G}_k of \mathcal{G} to a single distinct node $v_k^* \in \mathcal{V}_k^*$ of \mathcal{G}^* and introducing an edge in \mathcal{G}^* from v_i^* to v_j^* , if and only if there exists some edges from the SCC \mathcal{G}_i and the SCC \mathcal{G}_j of the original graph. An SCC that is reduced to the root of a directed spanning tree of the condensation digraph is called the *root SCC* (RSCC). Looking at the condensation graph, we may identify the following topologies of the original graph: (1) \mathcal{G} is SC if and only if \mathcal{G}^* is composed by a single node; (2) \mathcal{G} is QSC if and only if \mathcal{G}^* contains a directed spanning tree; (3) if \mathcal{G} is WC, then \mathcal{G}^* contains either a spanning tree or a (weakly) connected forest.

The multiplicity of the zero eigenvalue of \mathbf{L} is equal to the minimum number of directed trees contained in a directed spanning forest of \mathcal{G} . Moreover, the zero eigenvalue of \mathbf{L} is simple if and only if \mathcal{G} contains a spanning directed tree or, equivalently, \mathcal{G} is QSC. If \mathcal{G} is SC then \mathbf{L} has a simple zero eigenvalue and positive left-eigenvector associated to the zero eigenvalue. If \mathcal{G} is QSC [17] with $Q \geq 1$ strongly connected components $\mathcal{G}_i \triangleq \{\mathcal{V}_i, \mathcal{E}_i\}$ with $\mathcal{V}_i \subseteq \mathcal{V}, \mathcal{E}_i \subseteq \mathcal{E}$ for $i = 1, \dots, Q$, $|\mathcal{V}_i| = r_i$ and $\sum_i r_i = N$, numbered w.l.o.g. so that \mathcal{G}_1 coincides with the root SCC of \mathcal{G} , then the left-eigenvector $\mathbf{y} = [\gamma_1, \dots, \gamma_N]^T$ of \mathbf{L} associated to the zero eigenvalue has entries $\gamma_i > 0$ iff $v_i \in \mathcal{V}_1$ and zero otherwise. If \mathcal{G}_1 is balanced then $\mathbf{y}_{r_1} = [\gamma_1, \dots, \gamma_{r_1}]^T \in \text{span}\{\mathbf{1}_{r_1}\}$ where $r_1 \triangleq |\mathcal{V}_1|$.

As a numerical example, in Figure 7.24 we report three network topologies: (a) a SC digraph; (b) a QSC digraph with three SCCs; (c) a WC digraph with a two-trees forest. We have also depicted for each digraph its decomposition into SCCs corresponding to the nodes of the associated condensation digraph; RSCC denotes the root SCC. For each network topology, we have also reported the dynamical evolution of the consensus algorithm in (7.23) versus time. It can be observed that the dynamical system in Figure 7.24a achieves a global consensus since the underlying digraph is SC. For the QSC digraph in Figure 7.24b, instead, there is a set of nodes in the RSCC component that is able to reach all other nodes so that the dynamical system can achieve a global consensus. Finally, in Figure 7.24c, the system cannot achieve a global consensus since there is no node that can reach all the others. Although we can observe two disjoint clusters corresponding to the two RSCC components, the nodes of the SCC

**FIGURE 7.24**

Consensus for different network topologies: (a) SC digraph; (b) QSC digraph with three SCCs; (c) WC digraph with a forest.



component (middle lines) are affected by the consensus in the two RSCC components but are not able to influence them.

A.2 RGG adjacency matrix

A random graph is obtained by distributing N points randomly over the d -dimensional space \mathbb{R}^d and connecting the nodes according to a given rule. The graph topology is captured by the adjacency matrix \mathbf{A} which, in this case, is a random matrix. An important class of random matrices, is the so called Euclidean Random Matrix (ERM) class, introduced in [89]. Given a set of N points located at positions $\mathbf{x}_i, i = 1, \dots, N$, an $N \times N$ adjacency matrix \mathbf{A} is an ERM if its generic (i, j) entry depends only on the difference $\mathbf{x}_i - \mathbf{x}_j$, i.e., $a_{ij} = F(\mathbf{x}_i - \mathbf{x}_j)$, where F is a measurable mapping from \mathbb{R}^d to \mathbb{R} . An important subclass of ERM is given by the adjacency matrices of the so called Random Geometric Graphs (RGG). In such a case, the entries a_{ij} of the adjacency matrix are either zero or one depending only on the distance between nodes i and j , i.e.,

$$a_{ij} = F(\mathbf{x}_i - \mathbf{x}_j) = \begin{cases} 1 & \text{if } \|\mathbf{x}_i - \mathbf{x}_j\| \leq r, \\ 0 & \text{otherwise,} \end{cases} \quad (7.179)$$

where r is the coverage radius. Next we discuss some important properties of the spectrum of the adjacency matrix of a random geometric graph.

A.2.1 Spectrum of a random geometric graph

Assuming that the RGG $G(N, r)$ is connected with high probability, we have derived in [77] an analytical expression for the algebraic connectivity of the graph, i.e., the second eigenvalue of the symmetric Laplacian, $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where \mathbf{D} is the degree matrix and \mathbf{A} is the adjacency matrix. From (7.159), $\mathbf{D} = \pi r^2 N \mathbf{I}$, so that we only need to investigate the second *largest* eigenvalue of \mathbf{A} . Hence, let us start by studying the spectrum of \mathbf{A} as discussed in [77]. In [90, 91], it is shown that the eigenvalues of the adjacency matrix tend to be concentrated, as the number of nodes tend to infinity. In particular, in [90] it is shown that the eigenvalues of the normalized adjacency matrix $\mathbf{A}_N = \mathbf{A}/N$ of an RGG $G(N, r)$, composed of points uniformly distributed over a unitary two-dimensional torus, tend to the Fourier series coefficients of the function F defined in (7.179),

$$\hat{F}(z) = \int_{\Omega_r} \exp(-2\pi j z^T \mathbf{x}) d\mathbf{x} \quad (7.180)$$

almost surely, for all $z = [z_1, z_2] \in \mathbb{Z}^2$, where $\Omega_r = \{\mathbf{x} = [x_1, x_2]^T \in \mathbb{R}^2 : \|\mathbf{x}\| \leq r\}$. Using polar coordinates, i.e., $x_1 = \rho \sin \theta$ and $x_2 = \rho \cos \theta$, with $0 \leq \rho \leq r$ and $0 \leq \theta \leq 2\pi$, we obtain

$$\hat{F}(z) = \int_0^r \int_0^{2\pi} \exp(-2\pi j \rho(z_1 \sin \theta + z_2 \cos \theta)) \rho d\rho d\theta.$$

This integral can be computed in closed form. Setting $z_1 = A \sin \phi$ and $z_2 = A \cos \phi$, we have

$$\hat{F}(A, \phi) = \int_0^r \int_{-\phi}^{2\pi-\phi} \exp(-2\pi j \rho A \cos(\xi)) \rho d\rho d\xi$$

with $\xi = \theta - \phi$. Furthermore, using the integral expression for the Bessel function of the first kind of order k , $J_k(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \sin(\xi) - jk\xi) d\xi$, we get

$$\hat{F}(A, \phi) = \hat{F}(A) = 2\pi \int_0^r J_0(2\pi\rho A) \rho d\rho.$$

Finally, using the identity $\int_0^u v J_0(v) dv = u J_1(u)$, we can make explicit the dependence of $\hat{F}(A)$ on the index pair $[z_1, z_2]$

$$\hat{F}(z_1, z_2) = \frac{r}{\sqrt{z_1^2 + z_2^2}} J_1\left(2\pi r \sqrt{z_1^2 + z_2^2}\right). \quad (7.181)$$

This formula allows us to rank the eigenvalues of $A_N = A/N$. In particular, we are interested in the second largest eigenvalue of A_N . Considering that the minimum coverage radius ensuring connectivity behaves as $r(N) \sim \sqrt{\frac{\log(N)}{N}}$, i.e., it is a vanishing function of N , we can use the Taylor series expansion of $\hat{F}(z_1, z_2)$, for small r . Recalling that, for small x , $J_1(x) = x/2 - x^3/16 + o(x^5)$, we can approximate the eigenvalues as

$$\hat{F}(z_1, z_2) = \pi r^2 - \frac{\pi^3(z_1^2 + z_2^2)r^4}{2} + o(r^6). \quad (7.182)$$

This expansion shows that, at least for small r , the largest eigenvalue equals πr^2 and occurs at $z_1 = z_2 = 0$, whereas the second largest eigenvalue corresponds to the cases $(z_1 = 1, z_2 = 0)$ and $(z_1 = 0, z_2 = 1)$. More generally, we can check numerically that, for $r \leq 1/2$ and $A \geq 1$, the following inequalities hold true:

$$\pi r^2 \geq r J_1(2\pi r) \geq \frac{r}{A} |J_1(2\pi r A)|. \quad (7.183)$$

In summary, denoting the spectral radius of A_N as $\zeta_1(A_N) = \max_{1 \leq i \leq N} \frac{|\lambda_i(N)|}{N}$, where $\{\lambda_i(N)\}_{i=1}^N$ is the set of eigenvalues of A , it follows that

$$\lim_{N \rightarrow \infty} \zeta_1(A_N) = \max_{z \in \mathbb{Z}^2} |\hat{F}(z)| = \hat{F}(0, 0) = \pi r^2, \quad (7.184)$$

while the second largest eigenvalue of A_N , $\zeta_2(A_N)$, converges to

$$\lim_{N \rightarrow \infty} \zeta_2(A_N) = \hat{F}(1, 0) = \hat{F}(0, 1) = r J_1(2\pi r). \quad (7.185)$$

We are now able to derive the asymptotic expression for the second largest eigenvalue of the normalized Laplacian $L_N = D_N - A_N$, where $D_N := D/N$ is the normalized degree matrix. Because of the asymptotic property of the degree of an RGG, shown in (7.159), the second largest eigenvalue of L_N tends asymptotically to

$$\lambda_2(L_N) = \pi r^2 - \zeta_2(A_N). \quad (7.186)$$

Thus, the algebraic connectivity of the graph can be approximated, asymptotically, as

$$\lambda_2(L) = \pi N r^2 - N r J_1(2\pi r). \quad (7.187)$$

References

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, *Comput. Networks* (2002) 393–422.
- [2] A. Giridhar, P.R. Kumar, Toward a theory of in-network computation in wireless sensor networks, *IEEE Commun. Mag.* (2006) 98–107.
- [3] P. Gupta, P. Kumar, Critical power for asymptotic connectivity in wireless networks, in: W. McEneaney, G. Yin, Q. Zhang (Eds.), *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, Boston, MA, Birkhauser, 1998.
- [4] A. Giridhar, P.R. Kumar, Computing and communicating functions over sensor networks, *IEEE J. Sel. Areas Commun.* (2005) 755–764.
- [5] S.M. Kay, *Fundamentals of Statistical Signal Processing, Detection Theory*, vol II, Prentice-Hall PTR, Englewood Cliffs, NJ, 1998.
- [6] J. Whittaker, *Graphical Models in Applied Multivariate Statistics*, John Wiley & Sons, 1990.
- [7] S.L. Lauritzen, *Graphical Models*, Oxford University Press, 1996.
- [8] A. Anandkumar, L. Tong, A. Swami, Detection of GaussMarkov random fields with nearest-neighbor dependency, *IEEE Trans. Inform. Theory* 55 (2009) 816–827.
- [9] A. Anandkumar, J.E. Yukich, L. Tong, A. Swami, Energy scaling laws for distributed inference in random fusion networks, *IEEE J. Sel. Areas Commun.* (2009) 1203–1217.
- [10] O.P. Kreidl, A.S. Willsky, An efficient message-passing algorithm for optimizing decentralized detection networks, *IEEE Trans. Autom. Contr.* (2010) 563–578.
- [11] M. Gastpar, Information-theoretic bounds on sensor network performance, in: A. Swami, Q. Zhao, Y.-W. Hong, L. Tong (Eds.), *Wireless Sensor Networks—Signal Processing and Communications Perspectives*, John Wiley & Sons, 2007 (Chapter 2).
- [12] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, NJ, 2006.
- [13] C. Godsil, G. Royle, *Algebraic Graph Theory*, Springer, New York, 2001.
- [14] W. Ren, R.W. Beard, E.M. Atkins, Information consensus in multivehicle cooperative control: collective group behavior through local interaction, *IEEE Control Syst. Mag.* 27 (2) (2007) 71–82.
- [15] R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems, *Proc. IEEE* 95 (2007) 215–233.
- [16] S. Barbarossa, G. Scutari, Decentralized maximum-likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems, *IEEE Trans. Signal Process.* 55 (7) (2007) 3456–3470.
- [17] G. Scutari, S. Barbarossa, L. Pescosolido, Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels, *IEEE Trans. Signal Process.* 56 (4) (2008) 1667–1684.
- [18] A. Jadbabaie, S. Morse, Coordination of groups of mobile autonomous agents using nearest neighbor rules, *IEEE Trans. Auto. Control* 48 (2003) 988–1001.
- [19] L. Xiao, S. Boyd, Fast linear iterations for distributed averaging, *Syst. Control Lett.* 53 (2004) 65–78.
- [20] A. Nedić, Asuman Ozdaglar, Cooperative distributed multi-agent optimization, in: D.P. Palomar, Y.C. Eldar, *Convex Optimization in Signal Processing and Communications*, Cambridge University Press, 2010.
- [21] S. Sardellitti, M. Giona, S. Barbarossa, Fast distributed average consensus algorithms based on advection-diffusion processes, *IEEE Trans. Signal Process.* 58 (2) (2010) 826–842.
- [22] M. Fiedler, Algebraic connectivity of graphs, *J. Czech. Math.* 23 (1973) 298–305.
- [23] R. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

- [24] Y. Hatano, A.K. Das, M. Mesbahi, Agreement in presence of noise: pseudogradients on random geometric networks, in: Proceedings of the IEEE Conference on Decision and Control (CDC), Seville, Spain, 2005, pp. 6382–6387.
- [25] A. Tahbaz Salehi, A. Jadbabaie, On consensus over random networks, in: Proceedings of the 44th Allerton Conference, UIUC, Illinois, USA, 2006, 27–29.
- [26] A. Tahbaz Salehi, A. Jadbabaie, Consensus over ergodic stationary graph processes, *IEEE Trans. Automat. Contr.* 55 (1) (2010).
- [27] T.C. Aysal, M. Coates, M. Rabbat, Distributed average consensus using probabilistic quantization, in: Proceedings of the IEEE/SP Workshop on Statistical Signal Processing Workshop (SSP), Maddison, Wisconsin, USA, 2007, pp. 640–644.
- [28] M. Huang, J. Manton, Stochastic approximation for consensus seeking: mean square and almost sure convergence, in: Proceedings of the IEEE Conference on Decision and Control (CDC), New Orleans, LA, USA, 2007, pp. 306–311.
- [29] S. Kar, J.M.F. Moura, Distributed consensus algorithms in sensor networks with imperfect communication: link failures and channel noise, *IEEE Trans. Signal Process.* 57 (5) (2009) 355–369.
- [30] S.P. Lipshitz, R.A. Wannamaker, J. Vanderkooy, Quantization and dither: a theoretical survey, *J. Audio Eng. Soc.* 40 (1992) 355–375.
- [31] R. Wannamaker, S. Lipshitz, J. Vanderkooy, J. Wright, A theory of nonsubtractive dither, *IEEE Trans. Signal Process.* 48 (2) (2000) 499–516.
- [32] L. Schuchman, Dither signals and their effect on quantization noise, *IEEE Trans. Commun. Technol. COMM*-12 (1964) 162–165.
- [33] B. Polyak, *Introduction to Optimization*, Optimization Software Inc., New York, 1987.
- [34] S. Barbarossa, G. Scutari, Bio-inspired sensor network design: distributed decision through self-synchronization, *IEEE Signal Process. Mag.* 24 (3) (2007) 26–35.
- [35] I.D. Schizas, A. Ribeiro, G.B. Giannakis, Consensus in Ad Hoc WSNs with noisy links: Part I—Distributed estimation of deterministic signals, *IEEE Trans. Signal Process.* 56 (2008) 350–364.
- [36] I.D. Schizas, G.B. Giannakis, S.I. Roumeliotis, A. Ribeiro, Consensus in Ad Hoc WSNs With noisy links: Part II—Distributed estimation and smoothing of random signals, *IEEE Trans. Signal Process.* 56 (2) (2008) 1650–1666.
- [37] F. Cattivelli, C.G. Lopes, A.H. Sayed, Diffusion recursive least-squares for distributed estimation over adaptive networks, *IEEE Trans. Signal Process.* 56 (5) (2008) 1865–1877.
- [38] P. Braca, S. Marano, V. Matta, Enforcing consensus while monitoring the environment in wireless sensor networks, *IEEE Trans. Signal Process.* 56 (7) (2008) 3375–3380.
- [39] A.G. Dimakis, S. Kar, J.M.F. Moura, M.G. Rabbat, A. Scaglione, Gossip algorithms for distributed signal processing, *Proc. IEEE* 98 (11) (2010) 1847–1864.
- [40] A. Bertrand, M. Moonen, Consensus-based distributed total least squares estimation in Ad Hoc wireless sensor networks, *IEEE Trans. Signal Process.* 59 (5) (2011) 2320–2330.
- [41] L. Li, A. Scaglione, J.H. Manton, Distributed principal subspace estimation in wireless sensor networks, *IEEE J. Sel. Top. Signal Process.* 5 (4) (2011) 725–738.
- [42] S. Kar, J.M.F. Moura, K. Ramanan, Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication, *IEEE Trans. Inform. Theory* 58 (2012) 3575–3605.
- [43] G. Mateos, I.D. Schizas, G.B. Giannakis, Distributed recursive least-squares for consensus-based in-network adaptive estimation, *IEEE Trans. Signal Process.* 57 (11) (2009) 4583–4588.
- [44] S.M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, vol I, Prentice-Hall PTR, Englewood Cliffs, NJ, 1993.
- [45] D.P. Bertsekas, J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, Belmont, MA, 1989.

- [46] S. Boyd, N. Parikh, E. Chu, *Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers*, Now Publishers, May 2011.
- [47] R. Tibshirani, Regression Shrinkage and Selection via the Lasso, *J. Roy. Stat. Soc.* 58 (1996) 267–288.
- [48] G. Mateos, J.A. Bazerque, G.B. Giannakis, Distributed sparse linear regression, *IEEE Trans. Signal Process.* 58 (2010) 5262–5276.
- [49] J.J. Xiao, S. Cui, Z.Q. Luo, A.J. Goldsmith, Power scheduling of universal decentralized estimation in sensors networks, *IEEE Trans. Signal Process.* 54 (2) (2006) 413–422.
- [50] P.K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, New York, 1997.
- [51] R. Viswanathan, P.K. Varshney, Distributed detection with multiple sensors: Part I—Fundamentals, *Proc. IEEE* 85 (1997) 54–63.
- [52] R.S. Blum, S.A. Kassam, H.V. Poor, Distributed detection with multiple sensors: Part II—Advanced topics, *Proc. IEEE* 85 (1997) 64–79.
- [53] J.F. Chamberland, V. Veeravalli, Wireless sensors in distributed detection applications, *IEEE Signal Process Mag.* (2007) 16–25.
- [54] J.F. Chamberland, V. Veeravalli, Decentralized detection in sensor networks, *IEEE Trans. Signal Process.* (2003) 407–416.
- [55] J.F. Chamberland, V. Veeravalli, Asymptotic results for decentralized detection in power constrained wireless sensor networks, *IEEE J. Sel. Areas Commun.* (2004) 1007–1015.
- [56] P. Willett, P.F. Swaszek, R.S. Blum, The good, bad, and ugly: distributed detection of a known signal in dependent gaussian noise, *IEEE Trans. Signal Process.* (2000) 3266–3278.
- [57] J.F. Chamberland, V. Veeravalli, How dense should a sensor network be for detection with correlated observations? *IEEE Trans. Inform. Theory* (2006) 5099–5106.
- [58] A. Anandkumar, L. Tong, A. Swami, Energy efficient routing for statistical inference of Markov random fields, in: *Proceedings of CISS 2007*, Baltimore, 2007, pp. 643–648.
- [59] M. Alanyali, V. Saligrama, O. Savas, S. Aeron, Distributed Bayesian hypothesis testing in sensor networks, in: *Proceedings of the American Control Conference*, Boston, MA, vol. 6, 2004, pp. 5369–5374.
- [60] V. Saligrama, M. Alanyali, O. Savas, Distributed detection in sensor networks with packet losses and finite capacity links, *IEEE Trans. Signal Process.* 54 (11) (2006) 4118–4132.
- [61] S.A. Aldosari, J.M.F. Moura, Topology of sensor networks in distributed detection, in: *Proceedings of the IEEE International Conference on Acoustic Speech Signal Processing (ICASSP)*, Toulouse, France, 2006, pp. V1061–V1064.
- [62] S. Kar, J.M.F. Moura, Consensus based detection in sensor networks: topology optimization under practical constraints, *Workshop Information Theory in Sensor Networks*, Santa Fe, NM, 2007.
- [63] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, J.M.F. Moura, Distributed detection via gaussian running consensus: large deviations asymptotic analysis, *IEEE Trans. Signal Process.* 59 (2011) 4381–4396.
- [64] F. Cattivelli, A.H. Sayed, Distributed detection over adaptive networks using diffusion adaptation, *IEEE Trans. Signal Process.* 59 (5) (2011) 1917–1932.
- [65] L. Xiao, S. Boyd, S. Lall, A space-time diffusion scheme for peer-to-peer least-squares estimation, in: *Proceedings of International Conference on Information Processing in Sensor Networks*, Nashville, TN, 2006, pp. 168–176.
- [66] S. Marano, V. Matta, P. Willett, L. Tong, Cross-layer design of sequential detectors in sensor networks, *IEEE Trans. Signal Process.* 54 (2006) 4105–4117.
- [67] Y. Mei, Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks, *IEEE Trans. Inform. Theory* 54 (2008) 2072–2089.
- [68] Sardellitti, S.; Barbarossa, S.; Pezzolo, L. Distributed double threshold spatial detection algorithms in wireless sensor networks, in: *Proceedings of the IEEE 10th Workshop on Signal Processing Advances in Wireless Communication (SPAWC '09)*, Perugia, Italy, 2009, pp. 51–55.

- [69] J.N. Tsitsiklis, Decentralized detection, *Adv. Stat. Signal Process.* 2 (1993) 297–344.
- [70] J.N. Tsitsiklis, Decentralized detection by a large number of sensors, *Math. Control Signal Syst.* 1 (2) (1988) 167–182.
- [71] R.R. Tenney, N.R. Sandell, Jr., Detection with distributed sensors, *IEEE Trans. Aerosp. Electron. Syst.* 17 (1981) 98–101.
- [72] S. Barbarossa, G. Scutari, T. Battisti, Distributed signal subspace projection algorithms with maximum convergence rate for sensor networks with topological constraints, in: ICASSP 2009, 2009.
- [73] D.S. Bernstein, *Matrix Mathematics*, Princeton University Press, 2005.
- [74] S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, Randomized gossip algorithms, *IEEE Trans. Inform. Theory* 52 (2006) 2508–2530.
- [75] S. Barbarossa, G. Scutari, T. Battisti, Cooperative sensing for cognitive radio using decentralized projection algorithms, in: IEEE 10th Workshop on Signal Processing in Wireless Communication, SPAWC’09, pp. 116–120.
- [76] S. Sardellitti, S. Barbarossa, A. Swami, Average consensus with minimum energy consumption: optimal topology and power allocation, in: Proceedings of the European Signal Processing Conference (EUSIPCO 2010), Aalborg, Denmark, 2010, pp. 189–193.
- [77] S. Sardellitti, S. Barbarossa, A. Swami, Optimal topology control and power allocation for minimum energy consumption in consensus networks, *IEEE Trans. Signal Process.* 60 (1) (2012) 383–399.
- [78] A. Jeflea, A parametric study for solving nonlinear fractional problems, *An. St. Univ. Ovidius Constanta* 11 (2003) 87–92.
- [79] W. Dinkelbach, On nonlinear fractional programming, *Manag. Sci.* 13 (1967) 492–498.
- [80] Y. Hatano, M. Mesbahi, Agreement over random networks, *IEEE Trans. Automat. Contr.* 50 (2005) 1867–1872.
- [81] T.C. Aysal, K.E. Barner, Convergence of consensus models with stochastic disturbances, *IEEE Trans. Inform. Theory* 56 (2010) 4101–4113.
- [82] S. Sardellitti, S. Barbarossa, Energy preserving matching of sensor network topology to the statistical graphical model of the observed field, in: 17th International Conference on Digital Signal Processing (DSP), Corfu, Greece, 2011.
- [83] S. Barbarossa, S. Sardellitti, Energy Preserving Matching of Sensor Network Topology to Dependency Graph of the Observed Field, *IEEE Trans. Signal Process.* (in press).
- [84] D. Donoho, Compressive sensing, *IEEE Trans. Inform. Theory* 52 (2006) 1289–1306.
- [85] G. Gasso, A. Rakotomamonjy, S. Canu, Recovering sparse signals with a certain family of non-convex penalties and DC programming, *IEEE Trans. Signal Process.* 57 (2009) 4686–4698.
- [86] P.D. Tao, L.T.H. An, A D.C. optimization algorithms for solving the trust-region subproblem, *SIAM J. Optim.* 8 (1998) 476–505.
- [87] R. Zdunek, A. Cichocki, Fast nonnegative matrix factorization algorithms using projected gradient approaches for large-scale problems, *Comput. Intell. Neurosci.* (2008).
- [88] R.A. Brualdi, H.J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, Cambridge, UK, 1991.
- [89] M. Mezard, G. Parisi, A. Zee, Spectra of Euclidean random matrices, *Nuclear Phys. B* 559 (1999) 689–701.
- [90] C. Bordenave, Eigenvalues of Euclidean Random Matrices, *Random Structures and Algorithms*, vol. 33, John Wiley & Sons, NY, USA, December 2008, pp. 515–532.
- [91] S. Rai, The Spectrum of a random geometric graph is concentrated, *J. Theor. Probab.* 20 (2007) 119–132.

Signal Processing and Optimal Resource Allocation for the Interference Channel

Mingyi Hong and Zhi-Quan Luo

Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA

2.08.1 Introduction

2.08.1.1 Resource allocation in communication networks

Resource allocation is a fundamental task in the design and management of communication networks. For example, in a wireless network, we must judiciously allocate transmission power, frequency bands, time slots, and transmission waveforms/codes across multiple interfering links in order to achieve high system performance while ensuring user fairness and quality of service (QoS). The same is true in wired networks such as the Digital Subscriber Lines (DSL).

The importance of resource allocation can be attributed to its key role in mitigating multiuser interference. The latter is the main performance limiting factor for heterogeneous wireless networks where the number of interfering macro/pico/femto base stations (BS) can be very large. In addition, resource allocation provides an efficient utilization of limited resources such as transmission power and communication spectrum. These resources are not only scarce but also expensive. In fact, wireless system operators typically spend billions of dollars to acquire licenses to operate certain frequency bands. Moreover, the rising cost of electricity for them to operate the infrastructure has already surpassed the salary cost to employees in some countries. Thus, from the system operator's perspective, efficient spectrum/power utilization directly leads to high investment return and low operating cost (see e.g., [1,2]). The transmission power of a mobile terminal is another scarce resource. In this case, careful and efficient power allocation is the key to effectively prolong the battery life of mobile terminals.

Current cellular networks allocate orthogonal resources to users. For example, in a time-division multiplex access (TDMA) or a frequency-division multiplex access (FDMA) network, users in the same cell transmit in different time slots/frequency bands, and users in the neighboring cells transmit using orthogonal frequency channels. Although the interference from neighboring cells is suppressed, the overall spectrum efficiency is reduced, as each BS only utilizes a fraction of the available spectrum. According to a number of recent studies [3,4] current spectrum allocation strategies are not efficient, as at any given time and location, much of the allocated spectrum appears idle and unused. Moreover, users in cell edges still suffer from significant interference from non-neighboring cells, or pico/femto cells. In addition, for cell edge users the signal power from their own cells are typically quite weak. All of these factors can adversely affect their service quality.

To improve the overall system performance as well as user fairness, future wireless standards [5] advocate the notion of a heterogeneous network, in which low-power BSs and relay nodes are densely

deployed to provide coverage for cell edge and indoor users. This new paradigm of network design brings the transmitters and receivers closer to each other, thus is able to provide high link quality with low transmission power [6,7]. Unfortunately, close proximity of many transmitters and receivers also introduces substantial in-network interference, which, if not properly managed, may significantly affect the system performance. Physical layer techniques such as multiple input multiple output (MIMO) antenna arrays and multiple cell coordination will be crucial for effective resource allocation and interference management in heterogeneous networks.

An effective resource allocation scheme should allow not only flexible coordination among BS nodes but also sufficiently distributed implementation. Coordination is very effective for interference mitigation among interfering nodes (e.g., Coordinated Multi-Point (CoMP)), but is also costly in terms of signaling overhead. For example, CoMP requires full BS coordination as well as the sharing of transmit data among all cooperating BSs. In contrast, a distributed resource allocation requires far less signaling overhead and no data sharing, albeit at the cost of possible performance loss. For in-depth discussions of various design issues in heterogeneous networks, we recommend the recent articles and books including [7–13].

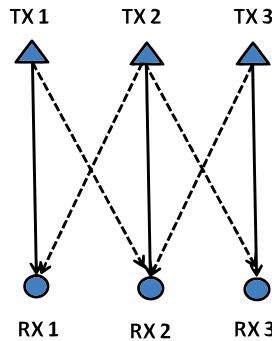
In this article, we examine several design and complexity aspects of the optimal physical layer resource allocation problem for a generic interference channel (IC). The latter is a natural model for multi-user communication networks. In particular, we characterize the computational complexity, the convexity as well as the duality of the optimal resource allocation problem. Moreover, we summarize various existing algorithms for resource allocation and discuss their complexity and performance tradeoff. We also mention various open research problems throughout the article.

2.08.1.2 Notation

Throughout, we use bold upper case letters to denote matrices, bold lower case letters to denote vectors, and regular lower case letters to denote scalars. For a symmetric (or Hermitian) matrix \mathbf{X} , the notation $\mathbf{X} \succeq \mathbf{0}$ (or $\mathbf{X} > \mathbf{0}$) signifies \mathbf{X} is positive semi-definite (or definite). User indices are denoted by subscripts, while frequency tone indices are denoted by superscripts. Zero mean normalized complex Gaussian distributions are denoted by $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.

2.08.1.3 Interference channels

An interference channel (IC) represents a communication network in which multiple transmitters simultaneously transmit to their intended receivers in a common channel. See Figure 8.1 for a graphical illustration of the IC. Due to the shared communication medium, each transmitter generates interference to all the other receivers. The IC model can be used to study many practical communication systems. The simplest example is a wireless ad hoc network in which transmitters and their intended receivers are randomly placed. When all these nodes are equipped with multiple antenna arrays, the channel becomes a MIMO IC. See Figure 8.2 for a graphical illustration of a 2-user MIMO IC. If each transmitter and receiver pair communicates over multiple parallel subchannels, the resulting overall channel model becomes a *parallel IC*. This parallel IC model can be used to describe communication networks employing Orthogonal Frequency Division Multiple Access (OFDMA) where the available spectrum is divided into multiple independent tones/channels. Networks of this kind include the DSL network or the IEEE 802.11 \times networks.

**FIGURE 8.1**

The Interference Channel model. The solid lines represent the direct channels, while the dotted lines represent the interfering channels.

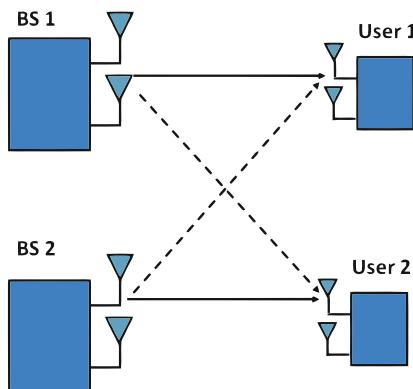
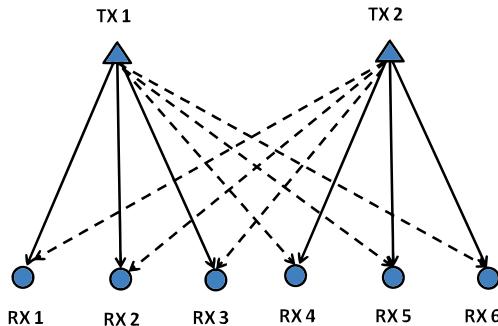
**FIGURE 8.2**

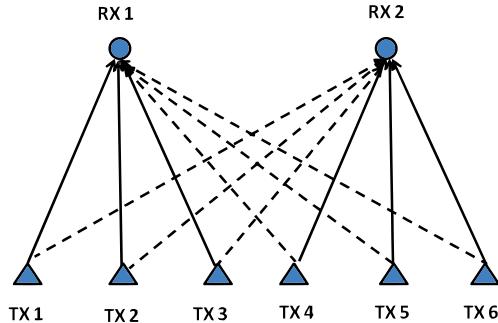
Illustration of a 2-user multiple input multiple output (MIMO) interference channel.

Another practical network is the multi-cell heterogenous wireless network. In the downlink of such network a set of interfering transmitters (BSs) simultaneously transmit to their respective groups of receivers. This channel is hitherto referred as an interfering broadcast channel (IBC). The uplink of this network can be modeled as an interfering multiple access channel (IMAC). See Figures 8.3 and 8.4 for graphical illustrations of these two channel models. Note that both IMAC and IBC reduce to an IC when there is only a single user in each cell.

Our ensuing discussions will be focussed on these channel models. We will illustrate key computational challenges associated with optimal resource allocation and suggest various practical resource allocation approaches to overcome them.

**FIGURE 8.3**

The Interfering Broadcast Channel model. The solid lines represent the direct channels, while the dotted lines represent the interfering channels.

**FIGURE 8.4**

The Interfering Multiple Access Channel model. The solid lines represent the direct channels, while the dotted lines represent the interfering channels.

2.08.1.4 System model

We now give mathematical description for three types of IC model—the scalar, parallel and MIMO IC models. Let us assume that there are K transmitter and receive pairs in the system, and we refer to each transceiver pair as a user. Let $\mathcal{K} = \{1, \dots, K\}$ denote the set of all the users.

2.08.1.4.1 Scalar IC model

In a scalar IC model each user transmits and receives a scalar signal. Let $x_k \in \mathcal{C}$ denote user k 's transmitted signal, and let $p_k = |x_k|^2$ denote its power. Let \bar{p}_k denote user k 's power constraint: $p_k \leq \bar{p}_k$. Let $z_k \sim \mathcal{CN}(0, 1)$ denote user k 's normalized complex Gaussian noise with unit variance. Note that we have normalized the power of the noise to unity. Let $H_{lk} \in \mathcal{C}$ denote the channel between

transmitter l and receiver k . Then user k 's received signal $y_k \in \mathcal{C}$ can be expressed as

$$y_k = \underbrace{H_{kk}x_k}_{\text{user } k\text{'s intended signal}} + \underbrace{\sum_{l \neq k} H_{lk}x_l}_{\text{multiuser interference}} + z_k. \quad (8.1)$$

The signal to interference plus noise ratio (SINR) for user k is defined as

$$\text{SINR}_k = \frac{|H_{kk}|^2 p_k}{1 + \sum_{l \neq k} |H_{lk}|^2 p_l}. \quad (8.2)$$

We denote the collection of all the users' transmit powers as $\mathbf{p} = [p_1, \dots, p_K]^T$.

2.08.1.4.2 Parallel IC model

In a parallel IC model, the spectrum is divided into N independent non-overlapping bands, each giving rise to a parallel subchannel. Let \mathcal{N} denote the set of all subchannels. Let $x_k^n \in \mathcal{C}$ denote the transmitted signal of user k on channel n , and let $p_k^n = |x_k^n|^2$ denote its power. We use \bar{p}_k to denote user k 's power budget so that $\sum_{n=1}^N p_k^n \leq \bar{p}_k$. Let $H_{lk}^n \in \mathcal{C}$ denote the channel coefficient between the transmitter of user l and the receiver of user k on channel n . Let $z_k^n \sim \mathcal{CN}(0, 1)$ denote the Gaussian channel noise. The received signal of user k on subchannel n , denoted as $y_k^n \in \mathcal{C}$, can be expressed as

$$y_k^n = \underbrace{H_{kk}^n x_k^n}_{\text{user } k\text{'s intended signal}} + \underbrace{\sum_{l \neq k} H_{lk}^n x_l^n}_{\text{multiuser interference}} + z_k^n. \quad (8.3)$$

We define the collection of user k 's transmit power as $\mathbf{p}_k = [p_k^1, \dots, p_k^N]^T$, and define all the users' transmit powers as $\mathbf{p} = [\mathbf{p}_1^T, \dots, \mathbf{p}_K^T]^T$.

2.08.1.4.3 MIMO IC model

In a MIMO IC model the receivers and transmitters are equipped with N_r and N_t antennas, respectively. Let $\mathbf{x}_k \in \mathcal{C}^{N_t}$ and $\mathbf{y}_k \in \mathcal{C}^{N_r}$ denote the transmitted and received signal of user k . Let $\mathbf{H}_{lk} \in \mathcal{C}^{N_r \times N_t}$ represent the channel gain coefficient matrix between transmitter l and receiver k .

Suppose each user k transmits/receives d_k data streams, and let $\mathbf{s}_k \in \mathcal{C}^{d_k \times 1}$ and $\hat{\mathbf{s}}_k \in \mathcal{C}^{d_k \times 1}$ denote the transmitted symbols and the received *estimated* symbols, respectively. Assume that the data vector \mathbf{s}_k is normalized so that $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] = \mathbf{I}$, and that the data signals for different users are independent from each other. Throughout this article, we will focus on *linear strategies* in which users use beamformers to transmit and receive data symbols. Let $\mathbf{V}_k \in \mathcal{C}^{N_t \times d_k}$ and $\mathbf{U}_k \in \mathcal{C}^{N_r \times d_k}$ denote the transmit and receive beamformers, respectively. Let $\mathbf{z}_k \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$ denote the normalized complex Gaussian noise vector at receiver k , where \mathbf{I}_{N_r} is the $N_r \times N_r$ identity matrix. Then the transmitted and received signal for user k can be expressed as

$$\mathbf{x}_k = \mathbf{V}_k \mathbf{s}_k, \quad (8.4)$$

$$\mathbf{y}_k = \underbrace{\mathbf{H}_{kk} \mathbf{x}_k}_{\text{user } k \text{ s intended signal}} + \underbrace{\sum_{l \neq k} \mathbf{H}_{lk} \mathbf{x}_l}_{\text{multiuser interference}} + \mathbf{z}_k, \quad (8.5)$$

$$\hat{\mathbf{s}}_k = \mathbf{U}_k^H \mathbf{y}_k. \quad (8.6)$$

Let $\mathbf{Q}_k = E[\mathbf{x}_k \mathbf{x}_k^H]$ denote the covariance matrix of the transmitted signal of user k . We assume that each transmitter has an averaged total power budget of the form

$$\text{Tr}(\mathbf{Q}_k) \leq \bar{p}_k, \quad k = 1, \dots, K. \quad (8.7)$$

When we have a single stream per user, \mathbf{V}_k and \mathbf{U}_k reduce to vectors $\mathbf{v}_k \in \mathcal{C}^{N_t \times 1}$ and $\mathbf{u}_k \in \mathcal{C}^{N_r \times 1}$. In this case the SINR for user k 's stream can be defined as

$$\text{SINR}_k = \frac{|\mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{v}_k|^2}{\|\mathbf{u}_k\|^2 + \sum_{l \neq k} |\mathbf{u}_k^H \mathbf{H}_{lk} \mathbf{v}_l|^2}. \quad (8.8)$$

Multiple Input Single Output (MISO) IC is a special case of MIMO IC in which the receivers only have a single antenna. In this case each user can only transmit a single stream ($d_k = 1, s_k \in \mathcal{C}$), and the beamforming matrix \mathbf{V}_k reduces to a beamforming vector $\mathbf{v}_k \in \mathcal{C}^{N_t \times 1}$. The channel coefficient matrix \mathbf{H}_{kl} becomes a *row vector* \mathbf{h}_{kl} , the received signal \mathbf{y}_k reduces to a scalar, which can be expressed as

$$y_k = \underbrace{\mathbf{h}_{kk} \mathbf{x}_k}_{\text{user } k \text{ s intended signal}} + \underbrace{\sum_{l \neq k} \mathbf{h}_{lk} \mathbf{x}_l}_{\text{multiuser interference}} + z_k. \quad (8.9)$$

The SINR for each user k can be expressed as

$$\text{SINR}_k = \frac{|\mathbf{h}_{kk} \mathbf{v}_k|^2}{1 + \sum_{l \neq k} |\mathbf{h}_{lk} \mathbf{v}_l|^2}. \quad (8.10)$$

The power budget constraint becomes

$$\|\mathbf{v}_k\|^2 \leq \bar{p}_k, \quad k = 1, \dots, K. \quad (8.11)$$

2.08.2 Information-theoretic results

2.08.2.1 Capacity results for IC model

In this subsection we briefly review some information theoretical results related to the capacity of the interference channel.

Consider a single user point to point additive white Gaussian noise (AWGN) scalar channel in the following form

$$y = Hx + z, \quad (8.12)$$

where x , y , H , z are the transmitted signal, the received signal, the channel coefficient, and the Gaussian noise, respectively. Assume that the noise is independently distributed as $z \sim \mathcal{CN}(0, 1)$, and that the

signal has a power constraint $|x|^2 \leq \bar{p}$. An achievable transmission rate R for this channel is defined as the rate that can be transmitted and decoded with diminishing error probability. The capacity of a channel C is the supremum of all achievable rates. Let us define the signal to noise ratio (SNR) of the channel as $\text{SNR} = \bar{p}|H|^2$, then the capacity of the Gaussian channel is given by

$$C(\text{SNR}) = \log_2(1 + \text{SNR}) \text{ bit per transmission.} \quad (8.13)$$

We refer the readers' to the classic books such as Cover Thomas [14] and the [online course](#) for an introductory treatment of information theory.

Now consider a 2-user interference channel

$$\begin{aligned} y_1 &= H_{11}x_1 + H_{21}x_2 + z_1, \\ y_2 &= H_{22}x_2 + H_{12}x_1 + z_2. \end{aligned} \quad (8.14)$$

The capacity region of this channel is the set of all achievable rate pairs of user 1 and user 2. Unlike the previous point to point channel, the complete characterization of the capacity region in this simplest 2-user IC case is an open problem in information theory. The largest achievable rate region for the interference channel is the Han-Kobayashi region [15], and it is achieved using superposition coding and interference subtraction. Recently, Etkin et al. [16] showed that this inner region is within one bit of the capacity region for scalar ICs. The capacity of the scalar interference channel under strong or very strong interference has been found in [15, 17, 18]. In particular, in the very strong interference case, i.e., $\frac{|H_{21}|^2}{|H_{11}|^2} \geq 1 + \bar{p}_1$ and $\frac{|H_{12}|^2}{|H_{22}|^2} \geq 1 + \bar{p}_2$, the capacity region is given as

$$R_k \leq \log_2(1 + |H_{kk}|^2 \bar{p}_k), \quad k = 1, 2, \quad (8.15)$$

where R_k is the transmission rate for user k . This result indicates that in very strong interference case the capacity is not reduced. The references [19, 20] include recent results that establish the capacity region for more general MIMO and parallel ICs in the strong interference case. However, for the general case where the interference is moderate, the capacity region remains unknown.

The capacity of a communication channel can be approximated by the notion of *degrees of freedom*. Recall that in the high SNR regime the capacity of a point to point link can be expressed as

$$C(\text{SNR}) = d \log_2(\text{SNR}) + o(\log_2(\text{SNR})). \quad (8.16)$$

In this case we say the channel has d degrees of freedom. In a 2-user interference channel, the *degrees of freedom region* can be characterized as follows. Let the sum transmit power across all the transmitters be ρ , and let $R_k(\rho)$ denote the transmission rate achievable for user k . Then the capacity region $C(\rho)$ of this 2-user channel is the set of all achievable rate tuples $(R_1(\rho), R_2(\rho))$. The degree of freedom region \mathcal{D} for this channel approximates the capacity region, and is defined as (see [21])

$$\begin{aligned} \mathcal{D} &= \left\{ (d_1, d_2) \in \mathbb{R}_+^2 : \forall (w_1, w_2) \in \mathbb{R}_+^2, w_1 d_1 + w_2 d_2 \right. \\ &\leq \left. \lim_{\rho \rightarrow \infty} \sup \left[\sup_{(R_1(\rho), R_2(\rho)) \in C(\rho)} [w_1 R_1(\rho) + w_2 R_2(\rho)] \frac{1}{\log_2(\rho)} \right] \right\}. \end{aligned} \quad (8.17)$$

The goal of resource allocation is to achieve the optimal performance established by information theory, subject to resource budget constraints. Unfortunately, optimal strategies for achieving the information theoretic limits are often unknown, too difficult to compute or too complicated to implement in practice. For practical considerations, we usually rely on simple transmit/receive strategies (such as linear beamformers) for resource allocation, with the goal of attaining an approximate information theoretic performance bound. The latter can be in terms of the degrees of freedom or some approximate capacity bounds which we describe next.

2.08.2.2 Achievable rate regions when treating interference as noise

Due to the difficulties in characterizing the capacity region and the optimal transmit/receive strategy for a general interference channel, many works in the literature study simplified transmit/receive strategies and the corresponding achievable rate regions. One such simplification, which is well motivated from practical considerations, is to assume that low-complexity single user receivers are used and that the multiuser interference is treated as additive noise. The authors of [22, 23] show that treating interference as noise in a Gaussian IC actually achieves the sum-rate channel capacity if the channel coefficients and power constraints satisfy certain conditions. These results serve as a theoretical justification for this simplification. In the rest of this article we will treat interference as noise at the receivers. Let us first review some achievable rate region results for different IC models with this simplified assumption.

2.08.2.2.1 Definition of rate region

Consider the 2-user scalar IC (8.14). The users' transmission powers are constrained by $0 \leq p_1 \leq \bar{p}_1$ and $0 \leq p_2 \leq \bar{p}_2$, respectively. The following rates are achievable when the users treat their respective interference as noise

$$\begin{aligned} R_1(p_1, p_2) &= \log_2(1 + \text{SINR}_1), \\ R_2(p_1, p_2) &= \log_2(1 + \text{SINR}_2), \end{aligned}$$

where the term SINR_k has been defined in (8.2). The *directly achievable* rate region $\overline{\mathcal{R}}$ is defined as the union of the achievable rate tuples $(R_1(p_1, p_2), R_2(p_1, p_2))$

$$\overline{\mathcal{R}} = \{(R_1(p_1, p_2), R_2(p_1, p_2)) : 0 \leq p_1 \leq \bar{p}_1, 0 \leq p_2 \leq \bar{p}_2\}. \quad (8.18)$$

The directly achievable rate region represents the set of achievable rates when the transmitters are not able to synchronize with each other [24]. If transmitter synchronization is possible, time-sharing among the extreme points of the directly achievable rate region can be performed. In this case, the achievable rate region becomes the convex hull of the directly achievable rate region (8.18). Sometimes for convenience, we will refer the directly achievable rate regions simply as *rate regions*. The exact meaning of the rate region should be clear from the corresponding context.

For a parallel IC model, user k 's achievable rate on channel n , R_k^n , can be expressed as

$$R_k^n(p_1^n, \dots, p_K^n) = \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right). \quad (8.19)$$

User k 's achievable sum rate is the sum of the rates achievable on all the channels

$$R_k = \sum_{n=1}^N R_k^n(p_1^n, \dots, p_K^n). \quad (8.20)$$

The directly achievable rate region $\bar{\mathcal{R}}$ in this case can be expressed as

$$\bar{\mathcal{R}} = \left\{ (R_1, \dots, R_K) : \sum_{n=1}^N p_k^n \leq \bar{p}_k, p_k \geq 0, \forall k \in \mathcal{K} \right\}. \quad (8.21)$$

For a MIMO IC model, user k 's achievable rate when treating all other users' interference as noise is

$$R_k(\mathbf{Q}_1, \dots, \mathbf{Q}_K) = \log_2 \det \left(\mathbf{I}_{N_r} + \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \left(\mathbf{I}_{N_r} + \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{Q}_l \mathbf{H}_{lk}^H \right)^{-1} \right). \quad (8.22)$$

The directly achievable rate region $\bar{\mathcal{R}}$ can be expressed as

$$\bar{\mathcal{R}} = \{(R_1, \dots, R_K) : \text{Tr}(\mathbf{Q}_k) \leq \bar{p}_k, \mathbf{Q}_k \succeq 0, \forall k \in \mathcal{K}\}. \quad (8.23)$$

2.08.2.2.2 Characterization of the directly achievable rate regions

Resource allocation requires a good understanding of the achievable rate regions. The (directly achievable) rate regions of the 2-user and the more general K -user scalar IC have been recently characterized in [25, 26]. We briefly elaborate the 2-user rate region and its properties. Let $\Phi(p_1, p_2)$ denote a point in the rate region with x, y coordinates representing $R_1(p_1, p_2)$ and $R_2(p_1, p_2)$, respectively. Let $\bar{p}_1 = \bar{p}_2 = \bar{p}$. Define two functions $\Phi_1(p_2) = \Phi(\bar{p}, p_2)$ and $\Phi_2(p_1) = \Phi(p_1, \bar{p})$. Then the boundary of the 2-user rate region consists of the union of two axis and the following two curves

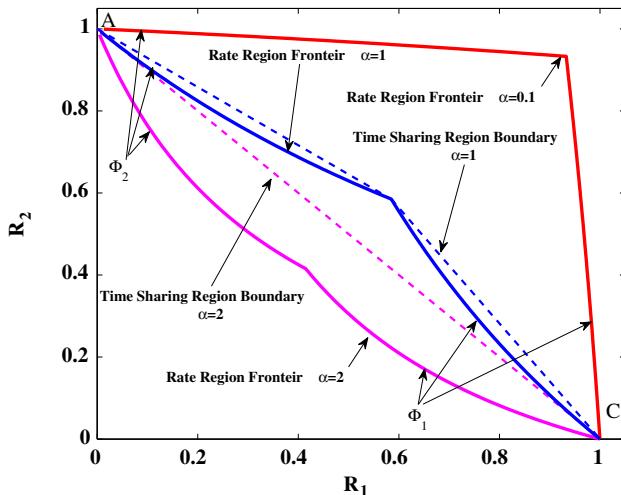
$$\Phi_1(p_2) = \log_2 \left(1 + \frac{\frac{|H_{22}|^2}{|H_{21}|^2} (|H_{11}|^2 \bar{p} - (2^{R_1} - 1))}{(2^{R_1} - 1)(1 + |H_{12}|^2 \bar{p})} \right), \quad 0 \leq p_2 \leq \bar{p}, \quad (8.24)$$

$$\Phi_2(p_1) = \log_2 \left(1 + \frac{|H_{22}|^2 \bar{p}}{1 + \frac{|H_{12}|^2}{|H_{11}|^2} (1 + |H_{21}|^2 \bar{p})(2^{R_1} - 1)} \right), \quad 0 \leq p_1 \leq \bar{p}. \quad (8.25)$$

Each of the above two curves consists of the set of rates achievable by one transmitter using its full power, while the other transmitter sweeping over its range of transmit powers. The convexity of this 2-user directly achievable rate region is studied in [25]. The following two conditions are sufficient to guarantee the convexity of the directly achievable rate region

$$\left. \frac{\partial^2 \Phi_1(p_2)}{\partial R_1^2} \right|_{p_2} < 0, \quad \forall 0 \leq p_2 \leq \bar{p}, \quad (8.26)$$

$$\left. \frac{\partial^2 \Phi_2(p_1)}{\partial R_1^2} \right|_{p_1} < 0, \quad \forall 0 \leq p_1 \leq \bar{p}. \quad (8.27)$$

**FIGURE 8.5**

The rate regions for a 2-user IC with different interference conditions. $\bar{p}_1 = \bar{p}_2 = 1$, $|H_{11}|^2 = |H_{22}|^2 = 1$, $|H_{12}|^2 = |H_{21}|^2 = \alpha$. At point A and C , a single user transmits using full power. The solid lines are the directly achievable rate frontier. The dotted lines represents the rate boundary that can be achieved by time sharing. Note that the time sharing boundary for $\alpha = 0.1$ is the same as the rate region frontier.

In particular, a necessary condition for (8.26) and (8.27) is

$$|H_{22}|^2 |H_{12}|^2 \bar{p}(1 + |H_{21}|^2 \bar{p}) - |H_{11}|^2 (1 + |H_{22}| \bar{p}) < 0 \quad (8.28)$$

which requires that the maximum possible interference to be sufficiently small. As the interference increases, the directly achievable rate regions become nonconvex. Figure 8.5 shows the transition of the directly achievable rate regions as well as the time-sharing regions when the interference levels change from strong to weak. Clearly, when the interference is strong ($\alpha = 2$ in this figure), orthogonal transmission such as TDMA or FDMA is optimal.

The same authors also characterize the achievable rate regions for the general K -user case. However the conditions for the convexity of the K -user regions are not available and deserve investigation. These conditions can be useful in solving resource allocation problems for an interference network.

More generally, it remains an open problem to derive a complete characterization of the (directly achievable) rate region for a parallel IC. The exact conditions for its convexity (or the lack of) are still unknown, although it is clear that the rate region will be convex if the interference coefficients are sufficiently small.

Several efforts have been devoted to characterizing certain interesting points (such as sum-rate optimal point) on the Pareto boundary of the rate region. Hayashi and Luo [27] have shown that in a

parallel IC model with channel gains satisfying the following strong interference conditions

$$\frac{|H_{lk}^n|^2}{|H_{kk}^n|^2} > \frac{1}{2}, \quad \text{and} \quad \frac{|H_{lk}^n|^2}{|H_{kk}^n|^2} \frac{|H_{kl}^n|^2}{|H_{ll}^n|^2} > \frac{1}{4} \left(1 + \frac{1}{C-1}\right)^2, \quad \forall n \in \mathcal{N}, (l, k) \in \mathcal{K} \times \mathcal{K} \quad (8.29)$$

where $C \geq 2$ is the minimum number of subchannels used by any user, then the sum rate maximization point can only be achieved using an FDMA strategy. In the special case of 2-user N channel model, the following condition is sufficient for the optimality of FDMA strategy

$$\frac{|H_{12}^n|^2}{|H_{22}^n|^2} \frac{|H_{21}^n|^2}{|H_{11}^n|^2} > \frac{1}{4} \left(1 + \frac{1}{C-1}\right)^2, \quad \forall n \in \mathcal{N}. \quad (8.30)$$

The MIMO IC model is even more general than the parallel IC, hence its achievable rate region is also difficult to characterize. To see this, assuming that $N_t = N_r = N$; let all the channel matrices be diagonal: $\mathbf{H}_{lk} = \text{diag}([H_{lk}^1, \dots, H_{lk}^N])$, $(l, k) \in \mathcal{K} \times \mathcal{K}$; let all the transmission covariances be diagonal as well: $\mathbf{Q}_k = \text{diag}([p_k^1, \dots, p_k^N])$, $k \in \mathcal{K}$. In this simplified model, user k 's transmission rate reduces to

$$\begin{aligned} R_k(\mathbf{Q}_1, \dots, \mathbf{Q}_K) &= \log_2 \det \left(\mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \left(\mathbf{I}_{N_r} + \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{Q}_l \mathbf{H}_{lk}^H \right)^{-1} + \mathbf{I}_{N_r} \right) \\ &= \sum_{n=1}^N \log_2 \left(1 + \frac{p_k^n |H_{kk}^n|^2}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right) \end{aligned} \quad (8.31)$$

which is exactly the rate expression for the N channel K user parallel IC as expressed in (8.19) and (8.20).

Larsson and Jorswieck [28, 29] have characterized the achievable rate region of a 2-user MISO IC. In this case, user k 's achievable transmission rate reduces to

$$R_k(\mathbf{v}_1, \dots, \mathbf{v}_k) = \log_2 (1 + \text{SINR}_k), \quad (8.32)$$

where the SINR for user k is defined in (8.10).

Define the maximum-ratio transmission (MRT) and the zero forcing (ZF) beamformers for both users as

$$\begin{aligned} \mathbf{v}_1^{\text{MRT}} &= \bar{p}_1 \frac{\mathbf{h}_{11}^H}{\|\mathbf{h}_{11}^H\|}, \quad \mathbf{v}_2^{\text{MRT}} = \bar{p}_2 \frac{\mathbf{h}_{22}^H}{\|\mathbf{h}_{22}^H\|}, \\ \mathbf{v}_1^{\text{ZF}} &= \bar{p}_1 \frac{\Pi_{\mathbf{h}_{12}^H}^\perp \mathbf{h}_{11}^H}{\|\Pi_{\mathbf{h}_{12}^H}^\perp \mathbf{h}_{11}^H\|}, \quad \mathbf{v}_2^{\text{ZF}} = \bar{p}_2 \frac{\Pi_{\mathbf{h}_{21}^H}^\perp \mathbf{h}_{22}^H}{\|\Pi_{\mathbf{h}_{21}^H}^\perp \mathbf{h}_{22}^H\|}, \end{aligned} \quad (8.33)$$

where $\Pi_{\mathbf{X}}^\perp$ represents the orthogonal projection onto the complement of the column space of \mathbf{X} . The authors show that any point on the Pareto boundary is achievable with the beamforming strategy

$$\begin{aligned} \mathbf{v}_1(\lambda_1) &= \bar{p}_1 \frac{\lambda_1 \mathbf{v}_1^{\text{ZF}} + (1 - \lambda_1) \mathbf{v}_1^{\text{MRT}}}{\|\lambda_1 \mathbf{v}_1^{\text{ZF}} + (1 - \lambda_1) \mathbf{v}_1^{\text{MRT}}\|}, \\ \mathbf{v}_2(\lambda_2) &= \bar{p}_2 \frac{\lambda_2 \mathbf{v}_2^{\text{ZF}} + (1 - \lambda_2) \mathbf{v}_2^{\text{MRT}}}{\|\lambda_2 \mathbf{v}_2^{\text{ZF}} + (1 - \lambda_2) \mathbf{v}_2^{\text{MRT}}\|}, \end{aligned} \quad (8.34)$$

where $0 \leq \lambda_1, \lambda_2 \leq 1$. Intuitively, it is clear that $\mathbf{v}_1(\lambda_1), \mathbf{v}_2(\lambda_2)$ should stay in the subspace spanned by the channel vectors $\mathbf{h}_{11}^H, \mathbf{h}_{22}^H$. Since this subspace is spanned by the MRT and ZF beamformers, it is no surprise that $\mathbf{v}_1(\lambda_1), \mathbf{v}_2(\lambda_2)$ can be written as linear combinations of the MRT and ZF beamformers. The novelty of (8.34) lies in the claim that the parameters λ_1, λ_2 are real numbers and lie in the interval $[0, 1]$. Similar to the characterization (8.24) and (8.25) for the rate region of a scalar IC, the characterization (8.34) of optimal beamforming strategy can be used to computationally determine the rate region for a 2-user MISO IC.

In [30], the authors extend their 2-user MISO channel work to a general K -user MISO IC. In particular, any point in the achievable rate region can be achieved using a set of beamformers $\{\mathbf{v}_k\}_{k=1}^N$ that is characterized by K^2 complex numbers $\{\epsilon_{kl}\}_{(k,l) \in \mathcal{K} \times \mathcal{K}}$ as

$$\begin{aligned}\mathbf{v}_k &= \sum_{l=1}^K \epsilon_{kl} \mathbf{h}_{kl}^H, \quad \forall k \in \mathcal{K}, \\ \|\mathbf{v}_k\|^2 &= \bar{p}_k, \quad \forall k \in \mathcal{K}.\end{aligned}$$

However, because of the large number of (complex) parameters involved, this characterization appears less useful computationally in the determination of the rate region. We refer the readers to the web pages of [Jorswieck](#) and [Larsson](#) for more details. We emphasize again that except for these limited results, the structure of a general MIMO IC rate region is still unknown when the interference is treated as noise.

2.08.3 Optimal resource allocation in interference channel

As is evident from the discussions in Section 2.08.2, the most interesting points on the boundaries of the rate regions can only be achieved by careful resource allocation. In this section we discuss optimal resource allocation schemes for the general IC models. Such optimality is closely related to the choice of a performance metric for the communication system under consideration.

2.08.3.1 Problem formulations

A communication system should provide users with QoS guarantees, and fairness through efficient resource utilization. Mathematically, the resource allocation problem can be formulated as the problem of optimizing a certain system level utility function subject to resource budget constraints.

A popular family of utility functions is the so called “ α -fair” utility functions, which can be expressed as

$$U\left(\{R_k\}_{k=1}^K\right) = \sum_{k=1}^K \frac{(R_k)^{1-\alpha}}{1-\alpha}, \quad (8.35)$$

where R_k denotes the transmission rate of user k . As pointed out in [31], different choices of the parameter α give different priorities to user fairness and overall system performance. We list four commonly used utility functions that belong to the family of α -fair utility functions:

- a. *The sum rate utility:* $U_1\left(\{R_k\}_{k=1}^K\right) = \sum_{k=1}^K R_k$, obtained by setting $\alpha = 0$;

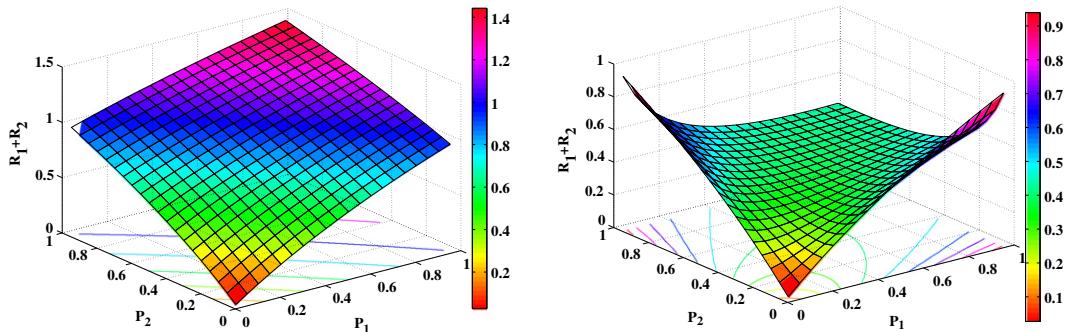


FIGURE 8.6

The sum rate utility for a 2-user scalar IC with different interference conditions. $\bar{p}_1 = \bar{p}_2 = 1$, $|H_{11}|^2 = |H_{22}|^2 = 1$, $|H_{12}|^2 = |H_{21}|^2 = \alpha$. In the left panel, $\alpha = 0.5$. In the right panel, $\alpha = 5$.

- b. *The proportional fair utility:* $U_2(\{R_k\}_{k=1}^K) = \sum_{k=1}^K \ln(R_k)$, obtained by letting $\alpha \rightarrow 1$;
- c. *The harmonic-rate utility:* $U_3(\{R_k\}_{k=1}^K) = \left(\sum_{k=1}^K R_k^{-1} \right)^{-1}$, obtained by setting $\alpha = 2$;
- d. *The min-rate utility ($\alpha \rightarrow \infty$):* $U_4(\{R_k\}_{k=1}^K) = \min_{1 \leq k \leq K} R_k$, obtained by letting $\alpha \rightarrow \infty$.

In terms of overall system performance, these utility functions can be ordered as

$$U_1 \geq U_2 \geq U_3 \geq U_4. \quad (8.36)$$

In terms of user fairness, the order is reversed. We note that except for the case in which the interference is weak, these utility functions are *nonconcave* in general. For example, in Figure 8.6 we plot the sum rate utility for a 2-user scalar IC in cases where the interference is either weak or strong. Moreover, in most cases, it is not possible to represent these utility functions as concave ones via a nonlinear transformation. See [32] for an impossibility result in scalar interference channel. This is consistent with the complexity status (NP-hard) of the utility maximization problems [33–35] (see discussions in Section 2.08.3.2).

If we wish to find a resource allocation scheme that maximizes the system level performance, then we need to determine the conditions under which the system level problem is easy to solve. Whenever such conditions are met, efficient system level resource allocation decision can be carried out by directly solving a convex optimization problem. Intuitively, when the crosstalk coefficients are zero or sufficiently small (low interference regime), the utility functions should be concave. It will be interesting to analytically determine how small the crosstalk coefficients need to be in order to preserve concavity.

From a practical perspective, the conditions for the concavity of the utility function (in terms of the crosstalk coefficients) are valuable because they can be used to find high quality approximately optimal resource allocation schemes. In particular, we can use these conditions to partition the users into small groups within which the interference is less and resource allocation is easy. Different groups can be put on orthogonal resource dimensions, because the groups cause too much interference to each other. Ultimately a good resource allocation scheme in an interference limited network will likely involve a hybrid scheme whereby some small groups of users share resources, while different groups are separated from competition.

The lack of concavity (or more generally, the lack of concave reformulation/transformation) has made it difficult to numerically maximize these utility functions for resource allocation. To circumvent the computational difficulties, and to reduce the amount of channel state information required for practical implementation, some researchers have proposed to use alternative utility functions for resource allocation. For example, both the mean squared error (MSE) and the leakage power cost functions have been proposed as potential substitutes for the rate-based utility functions listed above [36–39]. Recently, a number of studies [32, 40–42] have characterized a family of system utility functions that, under appropriate transformations, admit concave representations. Such transformations allow the associated utility maximization problems to be easily solvable. We refer the readers to [Holger Boche's web page](#) for details on this topic. Unfortunately these utility functions are not directly related to individual users' transmission rates, hence the solutions of the associated optimization problems tend to give suboptimal system performance (in terms of the users' achievable rates). We shall not further elaborate on these resource allocation approaches in this article. Instead, we will focus on the use of above listed rate-based utility functions for resource allocation.

Let us describe several utility maximization problems to be considered in this article.

1. Utility maximization for the scalar IC model:

$$\begin{aligned} \max_{\{p_k\}_k} \quad & U\left(\{R_k\}_{k=1}^K\right) \\ \text{s.t.} \quad & R_k = \log_2 \left(1 + \frac{|H_{kk}|^2 p_k}{1 + \sum_{l \neq k} |H_{lk}|^2 p_l} \right), \quad k = 1, \dots, K, \\ & 0 \leq p_k \leq \bar{p}_k, \quad k = 1, \dots, K. \end{aligned} \tag{8.37}$$

2. Utility maximization for the parallel IC model:

$$\begin{aligned} \max_{\{p_k^n\}_{k,n}} \quad & U\left(\{R_k\}_{k=1}^K\right) \\ \text{s.t.} \quad & R_k = \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right), \quad k = 1, \dots, K, \\ & \sum_{n=1}^N p_k^n \leq \bar{p}_k, \quad k = 1, \dots, K, \\ & p_k^n \geq 0, \quad (k, n) \in \mathcal{K} \times \mathcal{N}. \end{aligned} \tag{8.38}$$

3. Utility maximization for the MISO IC model:

$$\begin{aligned} \max_{\{\mathbf{v}_k\}_k} \quad & U\left(\{R_k\}_{k=1}^K\right) \\ \text{s.t.} \quad & R_k = \log_2 \left(1 + \frac{|\mathbf{h}_{kk} \mathbf{v}_k|^2}{1 + \sum_{l \neq k} |\mathbf{h}_{lk} \mathbf{v}_l|^2} \right), \quad k = 1, \dots, K, \\ & \|\mathbf{v}_k\|^2 \leq \bar{p}_k, \quad k = 1, \dots, K. \end{aligned} \tag{8.39}$$

4. Utility maximization for the MIMO IC model:

$$\begin{aligned} & \max_{\{\mathbf{Q}_k\}_k} U\left(\{R_k\}_{k=1}^K\right) \\ \text{s.t. } & R_k = \log_2 \det \left(\mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \left(\mathbf{I}_{N_r} + \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{Q}_l \mathbf{H}_{lk}^H \right)^{-1} + \mathbf{I}_{N_r} \right), \quad k = 1, \dots, K, \\ & \text{Tr}(\mathbf{Q}_k) \leq \bar{p}_k, \quad \mathbf{Q}_k \succeq 0, \quad k = 1, \dots, K. \end{aligned} \quad (8.40)$$

5. Utility maximization for the MIMO IC model (single stream per user):

$$\begin{aligned} & \max_{\{(\mathbf{u}_k, \mathbf{v}_k)\}_{k=1}^K} U\left(\{R_k\}_{k=1}^K\right) \\ \text{s.t. } & R_k = \log_2 \left(1 + \frac{|\mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{v}_k|^2}{\|\mathbf{u}_k\|^2 + \sum_{l \neq k} |\mathbf{u}_k^H \mathbf{H}_{lk} \mathbf{v}_l|^2} \right), \quad k = 1, \dots, K, \\ & \|\mathbf{v}_k\|^2 \leq \bar{p}_k, \quad k = 1, \dots, K. \end{aligned} \quad (8.41)$$

A “dual” paradigm for the design of the resource allocation algorithm is to provide QoS guarantees to all the users while minimizing the total power consumption. This formulation traditionally finds its application in voice communication networks where it is desirable to maintain a minimum communication rate (or SINR level) for each user in the system. Define $\{\gamma_k\}_{k=1}^K$ as the set of SINR targets. We list several QoS constrained min-power problem to be considered in this article.

6. Power minimization for the scalar IC model:

$$\begin{aligned} & \min \sum_{k=1}^K p_k \\ \text{s.t. } & \text{SINR}_k \geq \gamma_k, \quad k = 1, \dots, K, \\ & \text{SINR}_k = \frac{|\mathbf{H}_{kk}|^2 p_k}{1 + \sum_{l \neq k} |\mathbf{H}_{lk}|^2 p_l}, \quad k = 1, \dots, K. \end{aligned} \quad (8.42)$$

7. Power minimization for the MISO IC model:

$$\begin{aligned} & \min_{\{\mathbf{v}_k\}_{k=1}^K} \sum_{k=1}^K \|\mathbf{v}_k\|^2 \\ \text{s.t. } & \text{SINR}_k \geq \gamma_k, \quad k = 1, \dots, K, \\ & \text{SINR}_k = \frac{|\mathbf{h}_{kk} \mathbf{v}_k|^2}{1 + \sum_{l \neq k} |\mathbf{h}_{lk} \mathbf{v}_l|^2}, \quad k = 1, \dots, K. \end{aligned} \quad (8.43)$$

8. Power minimization for the MIMO IC model (single stream per user):

$$\min_{\{(\mathbf{u}_k, \mathbf{v}_k)\}_{k=1}^K} \sum_{k=1}^K \|\mathbf{v}_k\|^2 \quad (8.44)$$

$$\begin{aligned} \text{s.t. } & \text{SINR}_k \geq \gamma_k, \quad k = 1, \dots, K, \\ & \text{SINR}_k = \frac{|\mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{v}_k|^2}{\|\mathbf{u}_k\|^2 + \sum_{l \neq k} |\mathbf{u}_k^H \mathbf{H}_{lk} \mathbf{v}_l|^2}, \quad k = 1, \dots, K. \end{aligned}$$

A hybrid formulation combines the above two approaches. It aims to provide QoS guarantees while at the same time maximizing a system level utility function. This hybrid formulation is useful in data communication networks where besides the minimum rate constraints, it is preferable to deliver high system throughput. We list two such formulations to be considered later in this article.

9. Hybrid formulation for the scalar IC model:

$$\begin{aligned} \max_{\{p_k\}_{k=1}^K} & U(\{p_k\}_{k=1}^K) \\ \text{s.t. } & \text{SINR}_k \geq \gamma_k, \quad k = 1, \dots, K, \\ & 0 \leq p_k \leq \bar{p}_k, \quad k = 1, \dots, K. \end{aligned} \tag{8.45}$$

10. Hybrid formulation for the parallel IC model:

$$\begin{aligned} \max_{\{p_k^n\}_{k,n}} & U(\{R_k\}_{k=1}^K) \\ \text{s.t. } & R_k = \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right), \quad k = 1, \dots, K, \\ & R_k \geq \zeta_k, \quad k = 1, \dots, K, \\ & \sum_{n=1}^N p_k^n \leq \bar{p}_k, \quad k = 1, \dots, K, \\ & p_k^n \geq 0, \quad (k, n) \in \mathcal{K} \times \mathcal{N} \end{aligned} \tag{8.46}$$

where $\{\zeta_k\}_{k=1}^K$ is a set of rate targets.

We note that for the latter two formulations, the minimum rate/SINR requirements provide fairness to the users, while the optimization objectives are aimed at efficient utilization of system resource (e.g., spectrum or power). For both of these two problems, the *feasibility* of the set of rate/SINR targets needs to be carefully examined, as the rate/SINR requirements may not be simultaneously satisfiable.

2.08.3.2 Complexity of the optimal resource allocation problems

The aforementioned optimal resource allocation problems are nonconvex. However, the lack of convexity does not necessarily imply that the problem is difficult to solve. In some cases, it may be possible for a nonconvex problem to be appropriately transformed into an equivalent convex one and solved efficiently. A principled approach to characterize the intrinsic difficulty of an utility maximization problem is by way of the computational complexity theory [43].

In the following, we summarize a number of recent studies on the computational complexity status of these resource allocation problems. These complexity results suggest that in most cases solving the utility maximization problems to global optimality is computationally intractable as the number of users in the system increases.

Table 8.1 Complexity Status of Utility Maximization Problems for the Parallel and MISO IC Models [33, 34]

Problem Class	Utility Function			
	Sum Rate	Proportional Fair	Harmonic Mean	Min-Rate
Parallel IC, $K = 1$, N arbitrary	Convex opt	Convex opt	Convex opt	Convex opt
Parallel IC, $K \geq 2$ fixed, N arbitrary	NP-hard	NP-hard	NP-hard	NP-hard
Parallel IC, $N \geq 2$ fixed, K arbitrary	NP-hard	NP-hard	NP-hard	NP-hard
Parallel IC, $N = 1$, K arbitrary	NP-hard	Convex opt	Convex Opt	LP
MISO IC, $N_t \geq 2$, K arbitrary	NP-hard	NP-hard	NP-hard	Poly. time solvable

Table 8.2 Complexity Status of the Min-Rate Utility Maximization for the MIMO IC Model [35, 44, 45]

N_r	N_t		
	$N_t = 1$	$N_t = 2$	$N_t \geq 3$
$N_r = 1$	Poly. time solvable	Poly. time solvable	Poly. time solvable
$N_r = 2$	Poly. time solvable	NP-hard	NP-hard
$N_r \geq 3$	Poly. time solvable	NP-hard	NP-hard

Table 8.1 lists the complexity status for resource allocation problems with specific utility functions for the parallel and MISO IC models. Note that the scalar IC model is included as a special case.

Table 8.2 summarizes the complexity status for the minimum rate utility maximization problem and the sum power minimization problem with the QoS constraint in MIMO IC model (i.e., problem (8.41) with min-rate utility and problem (8.44)). Note that the results in Table 8.2 are based on the assumption that all transmitters and receivers use linear beamformers and that each mobile receives a single data stream.

Recall that the MIMO IC is a generalization of the Parallel IC (see Section 2.08.2.2). It follows that the complexity results in Table 8.1 hold true for the MIMO IC model with an arbitrary number of data streams per user. We refer the readers to the author's [web page](#) for recent developments in the complexity analysis as well as other resource allocation algorithms.

2.08.3.3 Algorithms for optimal resource allocation

We now describe various utility maximization based algorithms for resource allocation. These algorithms will be grouped and discussed according to their main algorithmic features. Since the min-rate utility function is non-differentiable, it requires a separate treatment that is different from the other utility functions. We begin our discussion with resource allocation algorithms based on the min-rate utility maximization.

2.08.3.3.1 Algorithms for min-rate maximization

Early works on resource allocation aimed to find optimal transmission powers that can maximize the min-SINR utility. In case of the scalar IC, this problem can be formulated as

$$\begin{aligned} \max_{\{p_k\}_{k \in \mathcal{K}}} \quad & \min_{k \in \mathcal{K}} \text{SINR}_k \\ \text{s.t.} \quad & p_k \geq 0, \quad \forall k \in \mathcal{K}, \\ & \text{SINR}_k \text{ defined in (8.2).} \end{aligned} \tag{8.47}$$

In [46–48], the authors studied the feasibility of this problem and proposed optimal power allocation strategies for it. For randomly generated scalar interference channels, they showed that with probability one, there exists a unique optimum value to the above problem. This optimal value, denoted as γ^* , can be expressed as

$$\gamma^* = \frac{1}{\rho(\mathbf{Z}) - 1} \tag{8.48}$$

where $\rho(\mathbf{Z})$ represents the maximum eigenvalue of the matrix \mathbf{Z} ; $\mathbf{Z} \in \mathbb{R}^{K \times K}$ is a matrix with its (k, l) th element defined as $[\mathbf{Z}]_{k,l} = \frac{|H_{lk}|^2}{|H_{kk}|^2}$. Distributed power allocation algorithms for this problem were also developed. For example, Foschini Miljanic [46] proposed an autonomous power control (APC) algorithm that iteratively adjusts the users' power levels as follows

$$\frac{|H_{kk}|^2 p_k^{(t+1)}}{1 + \sum_{l \neq k} |H_{lk}|^2 p_l^{(t)}} = \frac{|H_{kk}|^2 p_k^{(t)}}{1 + \sum_{l \neq k} |H_{lk}|^2 p_l^{(t)}} - \beta (\text{SINR}_k^{(t)} - \gamma^*), \tag{8.49}$$

where β is a small positive constant and t is the iteration index. We refer the readers to [49] and the [web page](#) of Hanly for further discussion of power control techniques for a scalar IC.

For a MISO IC model, the problem of finding optimal transmit beamformers for the maximization of the min-SINR utility has been considered by Bengtsson Ottersten [50] and Wiesel et al. [51]. The corresponding resource allocation problem can be equivalently formulated as

$$\begin{aligned} \max_{\{\mathbf{v}_k\}_{k \in \mathcal{K}}} \quad & \gamma \\ \text{s.t.} \quad & \text{SINR}_k \geq \gamma, \quad k = 1, \dots, K, \\ & \|\mathbf{v}_k\|^2 \leq \bar{p}_k, \quad k = 1, \dots, K, \\ & \text{SINR}_k \text{ defined in (8.10).} \end{aligned} \tag{8.50}$$

This optimization problem (8.50) is nonconvex, but can be relaxed to a semidefinite program (or SDP; see [52] for an introduction to the related concepts and algorithms). Surprisingly, [50] established that the SDP relaxation for (8.50) is tight; see the subsequent section “Algorithms for QoS Constrained Power Minimization” for more discussions. Later, Wiesel et al. [51] further showed that this nonconvex optimization problem can be solved via a sequence of second order cone programs (SOCP); see [52] for the definition of SOCP. The key observation is that for a fixed γ , checking the feasibility of (8.50) is an SOCP, which can be solved efficiently by the standard interior point methods. Let γ^* denote the optimal objective for problem (8.50), this max-min SINR problem can be solved by a bisection technique:

1. choose $\epsilon > 0$ (termination parameter), γ_l and γ_u such that γ^* lies in $[\gamma_l, \gamma_u]$;
2. let $\gamma_{\text{mid}} = (\gamma_l + \gamma_u)/2$;
3. check the feasibility of problem (8.50) with $\gamma = \gamma_{\text{mid}}$. If feasible, let $\gamma_l = \gamma_{\text{mid}}$, otherwise set $\gamma_u = \gamma_{\text{mid}}$;
4. terminate if $\gamma_u - \gamma_l \leq \epsilon$; else go to step 2. and repeat.

More recently, the max-min fairness resource allocation problem has been considered by Liu et al. [44] for the MIMO IC model. Unfortunately, the problem becomes NP-hard in this case (see Table 8.2).

The joint transceiver beamformer design for the min-SINR maximization problem in a MIMO IC (i.e., problem (8.41) with min-rate utility) has recently been considered in [44]. As shown in Section 2.08.3.2, this problem is in general NP-hard. Consequently, they proposed a low-complexity algorithm that converges to a stationary point of this problem. A key observation is that when the receive beamformers $\{\mathbf{u}_k\}_{k=1}^K$ are fixed, the considered problem can be written as

$$\begin{aligned} \max_{\{\mathbf{v}_k\}_{k=1}^K} \quad & \gamma \\ \text{s.t.} \quad & \frac{|\mathbf{u}_k^H \mathbf{H}_{kk} \mathbf{v}_k|^2}{\|\mathbf{u}_k\| + \sum_{l \neq k} |\mathbf{u}_k^H \mathbf{H}_{lk} \mathbf{v}_l|^2} \geq \gamma, \quad k = 1, \dots, K, \\ & \|\mathbf{v}_k\|^2 \leq \bar{p}_k, \quad k = 1, \dots, K \end{aligned} \quad (8.51)$$

which has the same form as the MISO min-SINR problem in (8.50), and thus can be solved using bisection and SOCP. As a result, the authors proposed to alternate between the following two steps to solve the min-rate maximization problem:

1. for fixed $\{\mathbf{u}_k^{(t)}\}_{k=1}^K$, solve (8.51) via SOCP to obtain $\{\mathbf{v}_k^{(t+1)}\}_{k=1}^K$;
2. for fixed $\{\mathbf{v}_k^{(t+1)}\}_{k=1}^K$, update $\{\mathbf{u}_k^{(t+1)}\}_{k=1}^K$ to the minimum mean square error (MMSE) receiver: $\mathbf{u}_k^{(t+1)} = \left(\sum_{l=1}^K \mathbf{H}_{lk} \mathbf{v}_l^{(t+1)} \left(\mathbf{H}_{lk} \mathbf{v}_l^{(t+1)} \right)^H + \mathbf{I} \right)^{-1} \mathbf{H}_{kk} \mathbf{v}_k^{(t+1)}$.

Unlike the MISO min-SINR case, only *local* optimal solutions can be found in the MIMO case. Extending the above algorithm to the MIMO IC/IBC/IMAC case with multiple data streams per user is not a trivial task. For a MIMO IC model, the feasibility problem becomes

$$R_k = \log_2 \det \left(\mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \left(\mathbf{I}_{N_r} + \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{Q}_l \mathbf{H}_{lk}^H \right)^{-1} + \mathbf{I}_{N_r} \right) \geq \zeta, \quad k = 1, \dots, K,$$

$$\text{Tr}(\mathbf{Q}_k) \geq 0, \quad \text{Tr}(\mathbf{Q}_k) \leq \bar{p}_k, \quad k = 1, \dots, K.$$

This problem is nonconvex and there is no known convex reformulation for it. Finding efficient and preferably distributed algorithms for these channel models is a challenging problem which deserves investigation.

2.08.3.3.2 Algorithms for weighted sum-utility maximization

In addition to the min-rate (min-SINR) utility, we can use other utility functions to allocate resources. For instance, let $\{\mu_k\}_{k=1}^K$ denote a set of positive weights that represent the relative priorities of the users in the system. Then the weighted sum-rate maximization (WSRM) problem for a parallel IC can be formulated as

$$\begin{aligned} \max_{\{p_k^n\}_{k,n}} \quad & \sum_{k=1}^K \mu_k R_k \\ \text{s.t.} \quad & R_k = \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right), \\ & \sum_{n=1}^N p_k^n \leq \bar{p}_k, \quad p_k \geq 0, \quad k = 1, \dots, K. \end{aligned} \tag{8.52}$$

This simply corresponds to the problem (8.38) with $\sum_{k=1}^K \mu_k R_k$ as the objective function. WSRM is a central problem for physical layer resource allocation. Many sum-utility maximization problems can be reduced to solving a sequence of WSRM problems for the single channel $N = 1$ case, see [53]. Unfortunately, the complexity results in Section 2.08.3.2 indicate that WSRM is in general a hard problem which can not be solved to global optimality by a polynomial time algorithm (unless NP = P). As a result, many works are devoted to finding high quality locally optimal solutions for the WSRM problem.

2.08.3.3.3 Algorithms based on Lagrangian dual decomposition

The linear additive structure of the power budget constraints in the weighted sum-utility maximization problem (8.38) can be exploited by Lagrangian dualization. In particular, [54] (see also [34, 53]) considered the Lagrangian dual relaxation of the utility maximization problem (8.38) for the parallel IC model. Let us define the *dual function* of the primal problem (8.38) as

$$d(\lambda) = \max_{\{p_k^n \geq 0\}_{k,n}} \left\{ U \left(\{R_k\}_{k=1}^K \right) - \sum_{k=1}^K \lambda_k \left(\sum_{n=1}^N p_k^n - \bar{p}_k \right) \right\}, \tag{8.53}$$

where $\lambda = \{\lambda_k \geq 0\}_{k=1}^K$ is the set of dual variables associated with the sum power constraints. Then the *dual problem* of the utility maximization problem can be expressed as follows

$$\begin{aligned} \min_{\lambda} \quad & d(\lambda) \\ \text{s.t.} \quad & \lambda_k \geq 0, \quad k = 1, \dots, K. \end{aligned} \tag{8.54}$$

Denote the optimal objective values of the primal problem (8.38) and the dual problem (8.54) with N channels as p_N^* and d_N^* , respectively. By the standard duality theory in optimization [55], we have that

the duality gap $d_N^* - p_N^*$ satisfies

$$d_N^* - p_N^* \geq 0. \quad (8.55)$$

When the primal problem is convex, strong duality holds and the inequality becomes equality. When restricted to the FDMA (Frequency Division Multiple Access) solutions, the Lagrangian dual problem decomposes across tones and is efficiently solvable [27, 53]. However, when the dual optimal solutions are not unique, it is difficult to construct a primal optimal solution for the problem (8.38). Luo and Zhang [53] proposed to use an additional randomized step to generate a primal feasible solution from the dual optimal solution.

When the primal problem is not restricted to the FDMA solutions, the Lagrangian dual function is difficult to compute, let alone optimize (see the complexity results in Section 2.08.3.2). Yu Liu [54] proposed an iterative spectrum balancing (ISB) algorithm that alternates between the following two steps to solve the WSRM problem (8.52):

1. given a $\lambda \geq \mathbf{0}$, use a coordinate ascent strategy to *approximately* evaluate the dual function until convergence;
2. update λ using the subgradient method or the ellipsoid method.

Due to the inexactness of step 1, this algorithm is not guaranteed to converge to a global optimal solution of the WSRM problem (8.52).

A surprising observation in [54] is that when N (the number of channels) goes to infinity, the duality gap vanishes. Luo and Zhang [34, 53] rigorously proved this result using Lyapunov theorem in functional analysis. In particular, Lyapunov's theorem implies that for the continuous formulation of the WSRM problem (infinite number of channels), the rate region is actually convex. With additional steps to estimate of the approximation of Lebesgue integrals, Luo and Zhang [53] showed that for some constant L , an estimate of the duality gap is bounded by

$$0 \leq d_N^* - p_N^* \leq \frac{L}{\sqrt{N}}. \quad (8.56)$$

Clearly the gap vanishes as N goes to infinity. Using this estimate, Luo Zhang [53] further developed a polynomial time approximation scheme to find an optimal FDMA solution for the continuous version of the WSRM problem (8.52).

Algorithms based on interference pricing

In a number of related works [56–59], the authors proposed a modified iterative water-filling (M-IWF) algorithm that iteratively solves K subproblems. The subproblem related to user k can be expressed as

$$\max_{\{p_k^n\}_{k,n}} R_k - \sum_{n=1}^N p_k^n T_k^n, \quad (8.57)$$

$$\text{s.t. } \sum_{n=1}^N p_k^n \leq \bar{p}_k, \\ p_k^n \geq 0, \quad n \in \mathcal{N}, \quad (8.58)$$

where T_k^n is defined as

$$T_k^n = \sum_{l \neq k} \left(\frac{|H_{kl}^n|^2}{\sum_{j \neq k} |H_{jl}^n|^2 p_j^n + 1} - \frac{|H_{kl}^n|^2}{\sum_{j=1}^K |H_{jl}^n|^2 p_j^n + 1} \right). \quad (8.59)$$

This term can be viewed as the *interference price* that user k needs to pay on channel n for the unit of interference it causes to all other users in the system. In other words, the price T_k^n corresponds to the marginal decrease in the sum-rate utility per unit increase in interference power p_k^n . If the interference price is set to zero, then we are led to the standard iterative water-filling algorithm [60]. The M-IWF algorithm works by iteratively performing the following steps:

1. fix $\{T_k^{n,(t)}\}_{n,k}$, each user iteratively computes the optimal solution $\{p_k^{n,(t+1)}\}_{n,k}$ to the convex subproblem (8.57) until convergence;
2. update $\{T_k^{n,(t+1)}\}_{n,k}$ according to (8.59) using $\{p_k^{n,(t+1)}\}_{n,k}$.

We note that the overall computational complexity of step 1 is $O(T K N \log N)$, where T is the total number of iterations needed for convergence. It was conjectured that this algorithm converges at least to a stationary point of the WSRM problem, but no formal proof was given. In [61], the authors successfully established the convergence (to the stationary point) of this type of pricing algorithm under the condition that the users act *sequentially*, i.e., in step 1 of M-IWF, only a *single user* solves its optimization problem (8.57). They interpreted this sequential M-IWF as a successive linear approximation of the WSRM problem, and showed that the term $-\sum_{n=1}^N \sum_{k=1}^K p_k^n T_k^n$ is the first order Taylor approximation (up to an additive constant term) of $\sum_{l \neq k} R_l$, the nonconcave part of the objective function. With this interpretation, the M-IWF algorithm can be seen as letting each user sequentially solve a partially linearized version of the WSRM problem. Since the first order Taylor approximation is a locally tight approximation of the weighted sum-rate objective function, the weighted sum-rates computed by the sequential M-IWF algorithm improve monotonically. Moreover, since the users update their power allocations locally, the M-IWF algorithm can be implemented in a distributed manner as long as the interference prices are exchanged among the users at each iteration. We shall refer to the sequential modification of the M-IWF as the multichannel distributed pricing (MDP) algorithm.

The interference pricing idea has been extended to the MISO IC in [62, 63], and to the MIMO IC with single stream per user in [64]. Ref. [65] considered interference pricing for the general MIMO IC without the single data stream per user restriction. Similar to the parallel IC situation, the convergence of the interference pricing algorithm for the MIMO IC has only been analyzed for the sequential user update case. It will be interesting to see how the pricing technique (and its convergence proof) can be extended to the MIMO IC/IBC/IMAC models with an arbitrary number of streams per user, while allowing simultaneous user updates. A step in this direction was taken by Venturino et al. [66] which extended the interference pricing technique to the MISO IBC model. Their algorithm (named Iterative Coordinated BeamForming (ICBF)) calculates proper pricing coefficients that enable each BS to update their respective beamformers. Convergence was always observed in the simulation, but no formal proof was given. A recent survey of various pricing techniques used in wireless networks can be found in [67]. We also refer the readers to the web pages of [Berry](#) and [Honig](#) for other related works on this topic.

Algorithms based on successive convex approximation

The MDP algorithm belongs to a class of algorithms called successive convex approximation (SCA). The idea is to construct and maximize a series of (concave) lower bounds of the original WSRM problem, so that a high quality solution can be obtained asymptotically. See Figure 8.7 for a graphical illustration of how this class of algorithms work. In [68], an algorithm called Successive Convex Approximation for Low complExity (SCALE) is proposed to improve the spectral efficiency of the DSL network. This algorithm transforms the nonconvex sum rate maximization problem into a series of convex problems by utilizing the following lower bound

$$\alpha \log_2(z) + \beta \leq \log_2(1+z), \quad (8.60)$$

$$\alpha = \frac{z_0}{1+z_0}, \quad \beta = \log_2(1+z_0) - \frac{z_0}{1+z_0} \log_2 z_0, \quad (8.61)$$

where the inequality (8.60) is tight at $z = z_0$. This lower bound allows the WSRM problem to be approximated by

$$\begin{aligned} \max_{\{p_k^n\}_{k,n}} \quad & \sum_{k=1}^K \mu_k \sum_{n=1}^N \alpha_k^n \log_2 \left(\frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right) + \beta_k^n \\ \text{s.t.} \quad & \sum_{n=1}^N p_k^n \leq \bar{p}_k, \quad k = 1, \dots, K \\ & p_k^n \geq 0, \quad (k, n) \in \mathcal{K} \times N. \end{aligned} \quad (8.62)$$

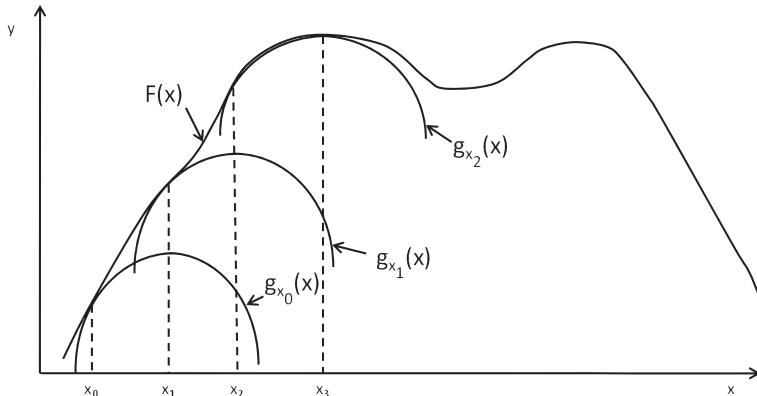


FIGURE 8.7

Graphical illustration of the family of algorithms using successive convex approximation. In this example, x_0 is the initial point. At x_0 , a (strictly) concave function $g_{x_0}(x)$ is used to approximate the original non-convex function $F(x)$. The optimal solution of $g_{x_0}(x)$, x_1 , is found by standard optimization technique. Then the (strictly) concave function $g_{x_1}(x)$ is constructed at the point x_1 . $g_{x_1}(x)$ is then optimized to obtain the point x_2 . Continue this process, a stationary solution of the original function $F(x)$ can be found.

After a log transformation $\tilde{p}_k^n = \log_2(p_k^n)$, $(k, n) \in \mathcal{K} \times N$, this relaxed problem turns out to be concave in \tilde{p}_n^k . The SCALE algorithm alternates between the following two steps:

1. fix $\{\alpha_k^{n,(t)}, \beta_k^{n,(t)}\}_{k,n}$, solve (8.62) and obtain $\{p_k^{n,(t+1)}\}_{n,k}$;
2. update the parameters $\{\alpha_k^{n,(t+1)}, \beta_k^{n,(t+1)}\}_{k,n}$ according to (8.61) using $\{p_k^{n,(t+1)}\}_{n,k}$.

Step 1 can be solved either in a centralized fashion using Geometric Programming (GP) technique, or by solving the dual problem of (8.62) in a distributed way. This algorithm is guaranteed to reach a stationary point of the original sum rate maximization problem. We briefly compare the major differences of the MDP and SCALE algorithms in Table 8.3.

In [69], a different lower bound is proposed for the WSRM problem. Specifically, the authors decompose the objective function as the difference of two concave functions of $\{p_k^n\}_{n,k}$ (referred to as the “dc” function)

$$\begin{aligned} & \sum_{k=1}^K \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right) \\ &= \sum_{k=1}^K \sum_{n=1}^N \log_2 \left(1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n + |H_{kk}^n|^2 p_k^n \right) \\ & \quad - \sum_{k=1}^K \sum_{n=1}^N \log_2 \left(1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n \right). \end{aligned} \quad (8.63)$$

Similar to the steps of SCA introduced earlier, in each iteration of the algorithm, the second sum is replaced with its linear lower bound, and the resulting concave maximization problem is solved. Compared to the MDP algorithm which linearizes $\sum_{l \neq k} R_l$, this algorithm linearizes all the interference terms in each iteration. As such, it linearizes more terms than the MDP algorithm per iteration.

A related algorithm has been proposed in the recent work [33] where the authors considered the general utility maximization problem in MISO IC (i.e., problem formulation (8.39)). Besides providing

Table 8.3 Comparison of SCALE and MDP Algorithms

	Properties			
	User update Schedule	Approximation Methods	Computation per Iteration	Dual Updates
SCALE	Simultaneously	Concave approximation	Iterative	Subgradient
MDP	Sequentially	Linear approximation	Closed form	Bisection

complexity results, the authors proposed an algorithm that is able to converge to local optimal solutions for problem (8.39) with any *smooth* (twice continuously differentiable) utility functions. The basic idea is to let the users cyclically update their beamformers using projected gradient ascent algorithm. In particular, at iteration t , user k takes a gradient projection step to compute the direction $\mathbf{d}_k^{(t+1)}$ by solving the following problem

$$\begin{aligned} \max_{\mathbf{d}_k} \quad & \langle \nabla_{\mathbf{v}_k} U \left(\left\{ \mathbf{v}_l^{(t)} \right\}_{l=1}^K \right), \mathbf{d}_k \rangle - \frac{1}{2} \|\mathbf{d}_k\|^2 \\ \text{s.t.} \quad & \|\mathbf{d}_k + \mathbf{v}_k^{(t)}\|^2 \leq \bar{p}_k. \end{aligned} \quad (8.64)$$

In contrast to the MDP and SCALE, the subproblem (8.64) linearizes the entire objective function of (8.39) at the current point $\left\{ \mathbf{v}_k^{(t)} \right\}_{k \in \mathcal{K}}$, and has an additional quadratic regularization term $\frac{1}{2} \|\mathbf{d}_k\|^2$. This subproblem is a convex quadratic minimization problem over a ball. As such, it is easier to solve than the corresponding subproblems of MDP and SCALE which are based on partial linearization of the original WSRM objective function. We list below the main steps of this cyclic coordinate ascent (CCA) algorithm:

1. select a user $k \in \mathcal{K}$ and compute its gradient projection direction $\mathbf{d}_k^{(t+1)}$ by solving (8.64);
2. determine stepsize $\alpha_k^{(t+1)}$ for user k using a line search strategy;
3. update beamformer: $\mathbf{v}_k^{(t+1)} = \mathbf{v}_k^{(t)} + \alpha_k^{(t+1)} \mathbf{d}_k^{(t+1)}$, and go to step 1.

The CCA algorithm only works in MISO IC case, and it is not clear how to extend it to the MIMO IC.

Algorithms based on weighted MMSE minimization

A different weighted sum-rate maximization approach was proposed in [70] for the MIMO broadcast downlink channel, where the WSRM problem is transformed to an equivalent weighted sum MSE minimization (WMMSE) problem with some specially chosen weight matrices. Since the optimal weight matrices are generally unknown, the authors of [70] proposed an iterative algorithm that adaptively chooses the weight matrices and updates the linear transmit/receive beamformers at each iteration. A nonconvex cost function was constructed in [70] and shown to monotonically decrease as the algorithm progresses. But the convergence of the iterates to a stationary point (or the global minimum) of the cost function is not known. Later, a similar algorithm was proposed in [71] for the interference channel where each user only transmits one data stream.

It turns out that this WMMSE based resource allocation approach can be extended significantly to handle the MIMO-IC and MIMO-IBC/IMAC models as well as general utility functions. In particular, the authors of [72, 73] established a general equivalence result between the global (and local) minimizers of the weighted sum-utility maximization problem (e.g., (8.40) and (8.52)) and a suitably defined weighed MMSE minimization problem. The latter can be effectively optimized by utilizing the block coordinate descent technique, resulting in independent, closed form iterative update across the transmitters and receivers. The resulting algorithm is named the WMMSE algorithm.

To gain some insight, let us consider the special case of a scalar IC system where the equivalence of the WSRM problem (8.52) and a weighted sum MSE minimization can be seen more directly. Let v_j, u_k

denote the complex gains used by the transmitter j and receiver k respectively. Consider the following weighted sum-MSE minimization problem

$$\begin{aligned} & \min_{\{w_k, u_k, v_k\}_{k=1}^K} \sum_{k=1}^K \mu_k (w_k e_k - \log w_k) \\ \text{s.t. } & |v_k|^2 \leq \bar{p}_k, \quad k = 1, 2, \dots, K, \end{aligned} \tag{8.65}$$

where w_k is a positive weight variable, and e_k is the mean square estimation error

$$e_k \triangleq |u_k H_{kk} v_k - 1|^2 + \sum_{j \neq k} |u_k H_{jk} v_j|^2 + |u_k|^2.$$

To see the equivalence, we can check the first order optimality condition to find the optimal w_k and u_k

$$u_k^{\text{opt}} = \frac{\overline{H_{kk} v_k}}{\sum_{j=1}^K |H_{jk}|^2 |v_j|^2 + 1}, \quad w_k^{\text{opt}} = e_k^{-1}, \quad \forall k = 1, 2, \dots, K. \tag{8.66}$$

Plugging these optimal values in (8.65) gives the following equivalent optimization problem

$$\begin{aligned} & \max_{\{v_k\}_{k=1}^K} \sum_{k=1}^K \mu_k \log \left(1 - \frac{|H_{kk}|^2 |v_k|^2}{\sum_{j=1}^K |H_{jk}|^2 |v_j|^2 + 1} \right)^{-1} \\ \text{s.t. } & |v_k|^2 \leq \bar{p}_k, \quad k = 1, 2, \dots, K, \end{aligned}$$

which, upon a change of variable $p_k = |v_k|^2$, is equivalent to

$$\begin{aligned} & \max_{\{p_k\}_{k=1}^K} \sum_{k=1}^K \mu_k \log \left(1 + \frac{|H_{kk}|^2 p_k}{\sum_{j \neq k} |H_{jk}|^2 p_j + 1} \right) \\ \text{s.t. } & p_k \leq \bar{p}_k, \quad k = 1, 2, \dots, K. \end{aligned}$$

This establishes the equivalence of the WMMSE problem (8.65) and the WSRM problem (8.52). More importantly, the equivalence goes one step further: there is a one-to-one correspondence between the local minimums of the two problems (see [72, 73]).

The equivalence relation implies that maximizing the weighted sum-rate can be accomplished via iterative weighted MSE minimization. The latter problem is in the space of (u, v, w) and is easier to handle since optimizing each variable while holding the others fixed is convex and easy (e.g., closed form). This property has been exploited in [72, 73] to design the WMMSE algorithm. In contrast, the original sum-rate maximization problem (8.52) is in the space of \mathbf{p} and is nonconvex, which makes the iterative optimization process difficult.

The general form of the WMMSE algorithm can handle any utility functions satisfying the following conditions

$$(\text{Separability}) \quad U(\{R_k\}_{k=1}^K) = \sum_{k=1}^K u_k(R_k), \tag{8.67}$$

$$(\text{Concavity}) \quad -u_k(-\log_2 \det(\mathbf{X})) \text{ strictly concave in } \mathbf{X} \succ \mathbf{0}, \quad \forall k \in \mathcal{K}, \tag{8.68}$$

$$(\text{Differentiability}) \quad u_k(x) \text{ increasing and twice continuously differentiable in } x, \quad \forall k \in \mathcal{K}. \tag{8.69}$$

In addition, it also handles a wide range of channel models, e.g., MIMO and parallel IC/IBC/IMAC. It is well known that $R_k = \max_{\mathbf{U}_k} \log_2 \det \left(\mathbf{E}_k^{-1} (\mathbf{U}_k, \{\mathbf{V}_k\}_{k=1}^K) \right)$, where \mathbf{E}_k is the mean square error (MSE) matrix for user k . Define a set of new functions: $c_k(\mathbf{E}_k) = -u_k(-\log \det(\mathbf{E}_k(\mathbf{U}_k, \{\mathbf{V}_k\}_{k=1}^K)))$, $k = 1, \dots, K$. Similar to the scalar IC case, the equivalence of the following two optimization problems can be established

$$\begin{aligned} & \min_{\{\mathbf{U}_k, \mathbf{V}\}_{k=1}^K} \quad \sum_{k=1}^K c_k(\mathbf{E}_k) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq \bar{p}_k, \quad k = 1, \dots, K, \end{aligned} \quad (8.70)$$

$$\begin{aligned} & \min_{\{\mathbf{U}_k, \mathbf{V}_k, \mathbf{W}_k\}_{k=1}^K} \quad \sum_{k=1}^K \text{Tr}(\mathbf{W}_k^H \mathbf{E}_k) + c_k(\phi_k(\mathbf{W}_k)) - \text{Tr}(\mathbf{W}_k^H \phi_k(\mathbf{W}_k)) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq \bar{p}_k, \quad k = 1, \dots, K, \end{aligned} \quad (8.71)$$

where $\phi_k(\mathbf{W}_k)$ is the inverse map of $\nabla c_k(\mathbf{E}_k)$. The WMMSE algorithm finds a stationary point of the alternative problem (8.71). In particular, it alternately updates the three sets of variables $\{\mathbf{U}_k\}_{k=1}^K$, $\{\mathbf{V}_k\}_{k=1}^K$, or $\{\mathbf{W}_k\}_{k=1}^K$ for problem (8.71), each time keeping two sets of variables fixed. The WMMSE algorithm for a MIMO IC is listed in the following table:

1. initialize $\{\mathbf{V}_k\}_{k \in \mathcal{K}}$ such that $\text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) = \bar{p}_k$;
2. **repeat**;
3. $\mathbf{W}'_k \leftarrow \mathbf{W}_k, \forall k \in \mathcal{K}$;
4. $\mathbf{U}_k \leftarrow (\sum_{l \in \mathcal{K}} \mathbf{H}_{lk} \mathbf{V}_l \mathbf{V}_l^H \mathbf{H}_{lk}^H + \mathbf{I})^{-1} \mathbf{H}_{kk} \mathbf{V}_k, \forall k \in \mathcal{K}$;
5. $\mathbf{W}_k \leftarrow (\mathbf{I} - \mathbf{U}_{kk}^H \mathbf{H}_{kk} \mathbf{V}_k)^{-1}, \forall k \in \mathcal{K}$;
6. $\mathbf{V}_k \leftarrow \mu_k (\sum_{l \in \mathcal{K}} \mu_l \mathbf{H}_{kl}^H \mathbf{U}_l \mathbf{W}_l \mathbf{U}_l^H \mathbf{H}_{kl} + \lambda_k^* \mathbf{I})^{-1} \mathbf{H}_{kk}^H \mathbf{U}_k \mathbf{W}_k, \forall k \in \mathcal{K}$;
7. **until** $|\sum_{l \in \mathcal{K}} \log \det(\mathbf{W}_l) - \sum_{l \in \mathcal{K}} \log \det(\mathbf{W}'_l)| \leq \epsilon$.

We note that in Step 6, $\lambda_k^* \geq 0$ is the Lagrangian multiplier for the constraint $\text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq \bar{P}_k$. This multiplier can be found easily by bi-section method. Also, notice that all updates are in closed form (except for λ_k^*) and can be performed simultaneously across users.

To compare the performance and the efficiency of various resource allocation methods, we consider a simple simulation experiment involving a parallel-IC and a MIMO-IC. We first specialize the WMMSE algorithm to the parallel-IC scenario and compare it with SCALE and MDP algorithms described earlier. To specialize the WMMSE algorithm for a parallel IC, let us restrict the transmit/receive matrices for each user to be diagonal. That is, the beamforming directions are fixed to be unit vectors and we only optimize power loading factors on the parallel channels. Let $\mathbf{v}_k \in \mathbb{C}^{N \times 1}$ denote the user k 's transmit filter vector, with v_k^n corresponding to the complex scaling coefficient to be used for the data stream

on channel n . Similarly, the receive filter vector and the weight vector are denoted by $\mathbf{u}_k, \mathbf{w}_k \in \mathcal{C}^{N \times 1}$ respectively. Then the WMMSE algorithm for the parallel IC channel can be described as:

1. Initialize $\{\mathbf{v}_k\}_{k \in \mathcal{K}}$ such that $\sum_{n \in \mathcal{N}} v_k^n = \bar{p}_k$;
2. **repeat**:
3. $(w_k^n)' \leftarrow w_k^n, \forall (n, k) \in \mathcal{N} \times \mathcal{K}$;
4. $u_k^n \leftarrow (\sum_{l \in \mathcal{K}} |H_{lk}^n|^2 |v_l^n|^2 + 1)^{-1} H_{kk}^n v_k^n, \forall (n, k) \in \mathcal{N} \times \mathcal{K}$;
5. $w_k^n \leftarrow (1 - \bar{u}_k^n H_{kk}^n v_k^n)^{-1}, \forall (n, k) \in \mathcal{N} \times \mathcal{K}$;
6. $v_k^n \leftarrow \mu_k \frac{\bar{H}_{kk}^n u_k^n w_k^n}{\sum_{l \in \mathcal{K}} \mu_l |H_{kl}^n|^2 |u_l^n|^2 w_l^n + \lambda_k^*}, \forall (n, k) \in \mathcal{N} \times \mathcal{K}$;
7. **until** $|\sum_{l \in \mathcal{K}} \sum_{n \in \mathcal{N}} \log(w_l^n) - \sum_{l \in \mathcal{K}} \sum_{n \in \mathcal{N}} \log((w_l^n)')| \leq \epsilon$.

In the simulation, we set the weights $\{\mu_k\}_{k \in \mathcal{K}}$ all equal to 1, and set the maximum power $\bar{p}_k = 10^{\text{SNR}/10}$ for all the users. We set the stopping criteria as $\epsilon = 0.01$ for all algorithms. The channel coefficients are generated from the complex Gaussian distribution $\mathcal{CN}(0, 1)$. For MIMO IC, all the transmitters and receivers are assumed to have the same number of antennas.

We first investigate the performance of SCALE, MDP and the parallel version of the WMMSE algorithm for a parallel IC. Figure 8.8 illustrates the sum rate performance of different algorithms when

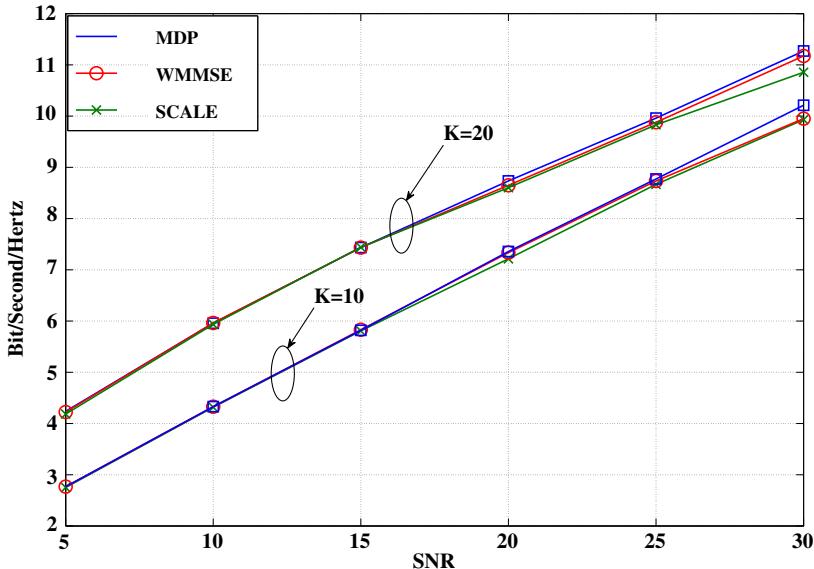
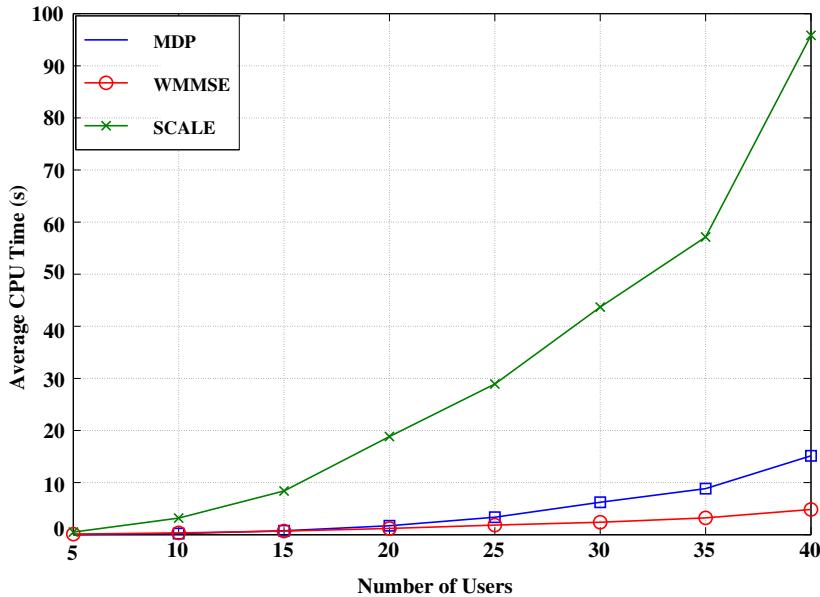


FIGURE 8.8

Comparison of the averaged sum rate performance versus SNR of different algorithms in parallel IC. $K = [10, 20]$, $N = 32$. Each curve in the figure is averaged over 100 random channel realizations.

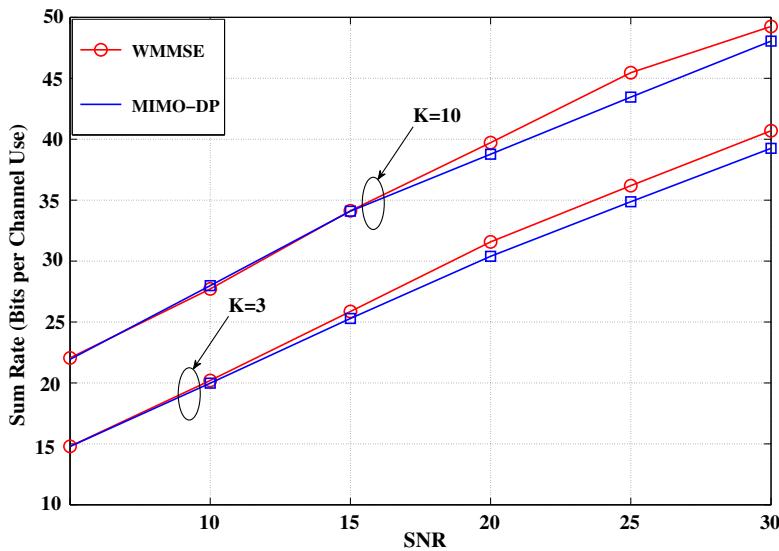
**FIGURE 8.9**

Comparison of the averaged CPU time versus the number of users of different algorithms in parallel IC. $N = 32$, $SNR = 10$. Each curve in the figure is averaged over 100 random channel realizations.

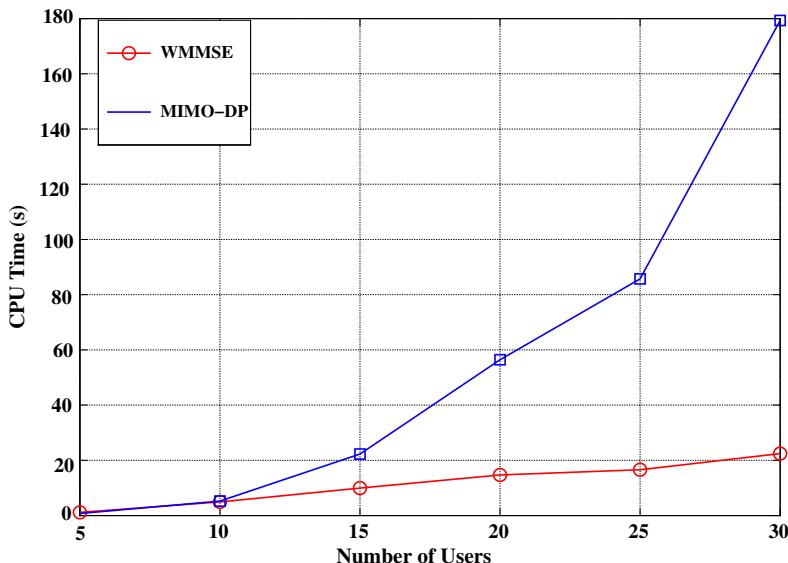
$K = [10, 20]$ and $N = 32$. We see that these algorithms all have similar performance across all the SNR values. Figure 8.9 shows the averaged CPU time comparison of these three algorithms under the same termination criteria and the same accuracy for the search of Lagrangian variables. We observe that the WMMSE requires much less computational time compared to the other two algorithms when the number of users becomes large. Note that the first step in the SCALE algorithm is implemented using the subgradient and the fixed point iterations suggested in [68, Section IV-A]. The stepsizes for the subgradient method as well as the number of the fixed point iterations need to be tuned appropriately to ensure fast convergence.

Next we examine the performance of the WMMSE and the MIMO distributed pricing (MIMO-DP) algorithm developed in [65] in the context of a MIMO-IC. Figure 8.10 illustrates the sum rate performance of the two algorithms when $K = [3, 10]$ and $N_r = N_t = 3$. Figure 8.11 shows the averaged CPU time comparison of the two algorithms. We again observe that the WMMSE requires much less computational time compared with the MIMO-DP algorithm when the number of users becomes large.

Different from many algorithms discussed earlier (e.g., CCA, MDP), the WMMSE algorithm allows all transmitters/receivers to update their beamformers simultaneously. This feature leads to simple implementation and fast convergence. It will be interesting to see how this algorithm can be further extended to include other utility functions such as the min-rate utility, and to other formulations like QoS constrained power minimization. Also, further research is needed to uncover the full algorithmic

**FIGURE 8.10**

Comparison of the averaged sum rate performance versus SNR of different algorithms in MIMO IC. $K = [3, 10]$, $N_r = N_t = 3$. Each curve in the figure is averaged over 100 random channel realizations.

**FIGURE 8.11**

Comparison of the averaged CPU time versus the number of users of different algorithms in MIMO IC. $N_r = N_t = 3$, $\text{SNR} = 10$. Each curve in the figure is averaged over 100 random channel realizations.

potential of WMMSE algorithm for a wide range of applications including joint base station assignment, power control, and beamforming.

Algorithms for cross layer resource allocation

We briefly mention a few cross-layer resource allocation algorithms which require solving a weighted sum-utility problem at each step. These algorithms jointly optimize physical layer as well as the media access (MAC) layer resources to improve the overall system performance.

Recently, Yu et al. [74] considered the joint MAC layer scheduling and physical layer beamforming and power control in a multicell OFDMA-MIMO network. The algorithm assigns the users to the BSs according to their individual priority and channel status. The beamformers are updated using a MSE duality results developed in [75, 76] for multicell network. The transmit powers of the BSs are updated by using the Newton's method. In [77–79], the authors considered the WSRM problem in a multicell downlink OFDMA wireless network. They proposed to let the BSs alternate between the following two tasks to achieve high system throughput: (1) optimally schedule the users on each channel; (2) jointly optimize their downlink transmit power using a physical layer resource allocation algorithm such as M-IWF or SCALE. Razaviyayn et al. [80] proposed to adaptively group users into non-interfering groups, and optimize the transceiver structure and the group membership jointly. Such grouping strategy results in fair resource allocation, as cell-edge users with weak channels are protected from the strong users. A generalized version of the WMMSE algorithm has been developed to perform such joint optimization. In all these works the resulting resource allocation schemes achieved a weighted sum-utility that is significantly higher than what is possible with only performing physical layer beamforming/power allocation.

Joint admission control and downlink beamforming is another example of cross layer resource allocation. For a single cell MISO network, this problem has been considered in [81, 82]. A related problem is the joint BS selection and power control/beamforming problem. This problem has been addressed in the traditional CDMA based network (see, e.g., [83–85]), in OFDMA networks (e.g., [86–88]) and in a more general MIMO-HetNet in which all BSs operate on the same frequency bands [89, 90]. An interesting research direction to pursue is to effectively incorporate these higher layer protocols to boost the system performance for a MIMO and parallel IBC/IMAC network.

2.08.3.3.4 Algorithms for QoS constrained power minimization

For the scalar IC model, the QoS constrained min-power problem as formulated in (8.42) has been considered in [83, 85, 91]. They derived conditions for the existence of a feasible power allocation given a set of SINR targets. Define a $K \times K$ matrix \mathbf{A} as follows

$$[\mathbf{A}]_{k,l} = \begin{cases} \frac{\gamma_k |H_{lk}|^2}{|H_{kk}|^2} & \text{if } l \neq k, \\ 0 & \text{otherwise.} \end{cases} \quad (8.72)$$

If $\rho(\mathbf{A}) < 1$, an optimal power allocation $\mathbf{p} = [p_1, \dots, p_K]^T$ can be found as follows

$$\mathbf{p} = (\mathbf{I}_K - \mathbf{A})^{-1} \mathbf{b}, \quad (8.73)$$

where $\mathbf{b} = \left[\frac{\gamma_1}{|H_{11}|^2}, \dots, \frac{\gamma_K}{|H_{KK}|^2} \right]^T$. The convexity of the feasible SINR region for this problem has been established in [92, 93].

Alternatively, Yates [94] has provided a framework that allows the users to compute the optimal solution of problem (8.42) *distributedly* by the following fixed point iteration

$$p_k^{(t+1)} = \gamma_k \left(\frac{1 + \sum_{l \neq k} |H_{lk}|^2 p_l^{(t)}}{|H_{kk}|^2} \right), \quad k = 1, \dots, K. \quad (8.74)$$

This algorithm is shown to have linear rate of convergence, that is,

$$\limsup_{t \rightarrow \infty} \frac{\|\mathbf{p}^{(t+1)} - \mathbf{p}^*\|}{\|\mathbf{p}^{(t)} - \mathbf{p}^*\|} = c < 1, \quad (8.75)$$

where \mathbf{p}^* is the optimal solution of the min-power problem, and $c < 1$ is some positive constant. Recently Boche and Schubert [95] has proposed a different algorithm based on a Newton-type update that exhibits even faster (super-linear) rate of convergence. This algorithm can be applied to more general scenarios when the receivers are equipped with multiple antennas.

When the set of SINR targets cannot be supported by the system (that is, the problem (8.42) is infeasible), the call admission control mechanism should be invoked. A couple of recent works [96, 97] have considered the problem of joint admission and power control arises in the QoS constrained power minimization problem. See [98] for a survey on the general topic of call admission control.

For a MISO IC, Bengtsson Ottersten [50, 99] considered the min-power transmit beamforming problem under QoS constraints (problem (8.43)). Define $\mathbf{V}_k = \mathbf{v}_k \mathbf{v}_k^H$ and $\mathbf{G}_{lk} = \mathbf{h}_{lk}^H \mathbf{h}_{lk}$, this problem can be equivalently formulated as

$$\begin{aligned} \min_{\{\mathbf{V}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \text{Tr}(\mathbf{V}_k) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{G}_{kk} \mathbf{V}_k) - \gamma_k \sum_{l \neq k} \text{Tr}(\mathbf{G}_{lk} \mathbf{V}_l) = \gamma_k, \quad k = 1, \dots, K \\ & \text{rank}(\mathbf{V}_k) = 1, \mathbf{V}_k \succeq 0, \quad k = 1, \dots, K. \end{aligned} \quad (8.76)$$

Relaxing the rank constraint, this problem is a convex semidefinite program and can be solved efficiently. Interestingly, the authors showed that a rank-one solutions must exist for the relaxed problem, revealing a certain hidden convexity in this problem. The following procedure can be used to construct a rank-1 solution from an optimal solution $\{\mathbf{V}_k^*\}_{k=1}^K$ of the relaxed problem:

1. Take $\mathbf{e}_k \in \text{span}(\mathbf{V}_k^*)$;
2. Define $\boldsymbol{\eta} = [\gamma_1, \dots, \gamma_K]^T$;
3. Define a matrix \mathbf{F} with its elements as

$$[\mathbf{F}]_{kl} = \begin{cases} -\gamma_k \mathbf{e}_l^H \mathbf{G}_{kk} \mathbf{e}_l & \text{if } l \neq k, \\ \mathbf{e}_k^H \mathbf{G}_{kk} \mathbf{e}_k & \text{otherwise} \end{cases}; \quad (8.77)$$

4. Find $\mathbf{p} = \mathbf{F}^{-1} \boldsymbol{\eta}$;
5. Obtain $\mathbf{v}_k^* = \sqrt{[\mathbf{p}]_k} \mathbf{e}_k$.

The approach works for the IBC model as well. It can also be extended to include other resource allocation options, such as admission control and base station assignment; see [84, 100, 101] for details.

2.08.3.3.5 Algorithms for hybrid formulations

For a scalar IC, Chiang et al. [102] proposed to use a technique called *geometric programming* (GP) to find an approximate solution to the WSRM problem with QoS constraint. They showed that after approximating the rate function $\log(1 + \text{SINR}_k)$ by $\log(\text{SINR}_k)$, the WSRM problem becomes a GP and can be solved efficiently. Moreover, with this approximation, the resource allocation problem falls into the family of problems considered in [32, 41], which can be transformed into equivalent convex optimization problems. For this family of problems, fast and distributed algorithms based on certain Newton-type iteration have been proposed in [103].

However, this approximation is not so useful in practice because (1) it is accurate only in high SINR region; (2) it always leads to a solution for which all links are active. The latter feature is undesirable because having all links active can be highly suboptimal when interference is strong. In fact, the main difficulty with WSRM is precisely how to identify which links should be shut off, an important option that is excluded by the GP approximation approach.

Recognizing such problems, the same authors further proposed in [102] a successive convex approximation (GP-SCA) method that aims at finding a stationary solution to the original WSRM problem. In particular, let $e_{lk}(\mathbf{p}) = |H_{lk}|^2 p_l$, $f_k(\mathbf{p}) = 1 + \sum_{l=1}^K e_{lk}(\mathbf{p})$ and $g_k(\mathbf{p}) = 1 + \sum_{l \neq k} |H_{lk}|^2 p_l$, for $k = 1, \dots, K$. Utilizing the arithmetic-geometric mean inequality, the users' rate functions can be lower-approximate as

$$\log_2 (1 + \text{SINR}_k) = \log_2 \left(\frac{f_k(\mathbf{p})}{g_k(\mathbf{p})} \right) \quad (8.78)$$

$$\geq \log_2 \left(\frac{\prod_{l=1}^K \left(\frac{e_{lk}(\mathbf{p})}{\alpha_l} \right)^{\alpha_l} \times \left(\frac{1}{\beta} \right)^{\beta}}{g_k(\mathbf{p})} \right), \quad (8.79)$$

$$\text{where, } \alpha_l = \frac{e_{lk}(\hat{\mathbf{p}})}{f_k(\hat{\mathbf{p}})}, \beta = \frac{1}{f_k(\hat{\mathbf{p}})}. \quad (8.80)$$

This lower bound is again concave (upon performing a log-transformation), and it is tight when $\mathbf{p} = \hat{\mathbf{p}}$. The QoS constrained WSRM problem with the approximated objective (8.79) can be again solved by a GP. A similar alternating procedure as the one we have introduced for the SCALE algorithm can be used to compute a stationary solution to the WSRM problem with QoS constraint.

Algorithms based on global optimization

There are a number of attempts to find globally optimal solution for the WSRM problem. However, these algorithms are all based on implicit enumeration (not surprising in light of the complexity results in Section 2.08.3.2). As a result, they can only solve small scale problems and are unlikely to be suitable for implementation in practical applications. However, this does not mean that global optimization algorithms for WSRM are useless. For one thing, they can be a valuable tool to benchmark various low-complexity suboptimal approaches for resource allocation (e.g., those described earlier in this section).

For a scalar IC, Qian et al. [104] proposed to use an existing algorithm [105, 106] for nonconvex fractional programming to find the *global* optimal solution for the hybrid problem (8.45). Specifically, introducing a set of auxiliary variable $\{z_k\}_{k=1}^K$, the scalar WSRM problem with SINR constraint can be formulated into the following equivalent form

$$\begin{aligned} & \max_{\mathbf{z}, \mathbf{p}} \prod_{k=1}^K (z_k)^{w_k} \\ \text{s.t. } & 0 \leq z_k \leq \frac{f_k(\mathbf{p})}{g_k(\mathbf{p})}, \\ & 0 \leq p_k \leq \bar{p}_k, \frac{f_k(\mathbf{p})}{g_k(\mathbf{p})} \geq \gamma_k, \quad \forall k \in \mathcal{K}. \end{aligned} \tag{8.81}$$

This reformulated problem has a concave objective (upon a log transformation), and a nonconvex feasible set \mathcal{G} . The global optimization algorithm of [105, 106] solves the reformulated problem via some convex optimization problems over a sequence *shrinking* convex sets $S_1 \supset S_2 \dots \supset \mathcal{G}$. The worst case complexity of this algorithm is exponential.

Several other global optimization methods have been proposed to solve the utility maximization problems for more general IC models. For example, [107–110] considered the parallel IC model, and [111] treated the two user-MISO IC model. In particular, the algorithm proposed in [110] utilized the dc structure (8.63) of the weighted sum rate function, and applied a branch-and-bound (BB) algorithm to find global optimal solution to the WSRM problem. Due to their exponentially increasing complexity, these algorithms are only suitable for benchmarking resource allocation algorithm for networks with relatively small number of links. For example, the work of [104, 107] compared their global algorithms for a small parallel IC with $N = 4$, $K = 2$, and a scalar IC with K up to 10.

An important open problem is how to develop efficient algorithms (suitable for large networks) that can find (provably) tight upper bounds for the system performance.

2.08.3.3.6 Algorithms for robust resource allocation

All of the aforementioned resource management schemes require perfect channel state information (CSI) at the transmitter side. However, in practice the CSI obtained at the transmitter is susceptible to various sources of uncertainties such as quantization error, channel estimation error or channel aging. These uncertainties may significantly degrade the performance of resource allocation schemes that are designed using perfect CSI. As a result, *robust* designs are needed for practical resource management.

Several recent contributions considered robust linear transmitter design in a MISO channel with a single transmitter and multiple receivers. Let $\mathbf{h}_k \in \mathcal{C}^{N_t}$ and $\hat{\mathbf{h}}_k \in \mathcal{C}^{N_t}$ denote the *true* channel and the *estimated* channel between the transmitter and the k -th receiver, respectively. Let \mathcal{U}_k denote the uncertainty set of channel \mathbf{h}_k , which is the set of possible values that \mathbf{h}_k may take after obtaining the estimated channel $\hat{\mathbf{h}}_k$. Consider the following specific form of uncertainty set

$$\mathcal{U}_k(\Delta_k) = \left\{ \mathbf{h}_k | \mathbf{h}_k = \hat{\mathbf{h}}_k + \Delta_k, \|\Delta_k\| \leq \delta_k \right\}, \tag{8.82}$$

where Δ_k is the vector of estimation error and δ_k is the uncertainty bound. One of the most popular formulation of the robust design is the following QoS constrained min power problem

$$\begin{aligned} \min_{\{\mathbf{v}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \|\mathbf{v}_k\|^2 \\ \text{s.t.} \quad & \frac{|\mathbf{h}_k \mathbf{v}_k|^2}{\sum_{l \neq k} |\mathbf{h}_k \mathbf{v}_l|^2 + 1} \geq \gamma_k, \quad \forall \mathbf{h}_k \in \mathcal{U}_k(\Delta_k), \quad k = 1, \dots, K. \end{aligned} \quad (8.83)$$

This formulation aims at minimizing the total transmission power while ensuring that the SINR constraints are satisfied *under all possible channel uncertainties*. Define

$$\begin{aligned} \mathbf{V} &= [\mathbf{v}_1, \dots, \mathbf{v}_K] \\ \underline{\mathbf{h}}_k &= [\operatorname{Re}\{\mathbf{h}_k\} \quad \operatorname{Im}\{\mathbf{h}_k\}] \\ \underline{\mathbf{V}} &= \begin{bmatrix} \operatorname{Re}\{\mathbf{V}\} & \operatorname{Im}\{\mathbf{V}\} \\ -\operatorname{Im}\{\mathbf{V}\} & \operatorname{Re}\{\mathbf{V}\} \end{bmatrix} \\ \underline{\mathbf{v}}_k &= \begin{bmatrix} \operatorname{Re}\{\mathbf{v}_k\} \\ -\operatorname{Im}\{\mathbf{v}_k\} \end{bmatrix}. \end{aligned}$$

In [112] problem (8.83) has been reformulated into the following semi-infinite SOCP

$$\min_{\underline{\mathbf{V}}, t} \quad t \quad (8.84)$$

$$\begin{aligned} \text{s.t.} \quad & \|\operatorname{vec}([\underline{\mathbf{v}}_1, \dots, \underline{\mathbf{v}}_K])\| \leq t \\ & \|[\underline{\mathbf{h}}_k \underline{\mathbf{V}}, 1]\| \leq \sqrt{1 + \gamma_k^{-1}} \underline{\mathbf{h}}_k \underline{\mathbf{v}}_k, \quad \forall \underline{\mathbf{h}}_k \in \mathcal{U}_k(\Delta_k), \quad k = 1, \dots, K. \end{aligned} \quad (8.85)$$

This reformulation is a *convex restriction* to the original problem (8.83) in that the complex magnitude $|\mathbf{h}_k \mathbf{v}_k|$ in the constraint is replaced by the lower bound equal to its real part $\operatorname{Re}(\mathbf{h}_k \mathbf{v}_k) = \underline{\mathbf{h}}_k \underline{\mathbf{v}}_k$. However, due to the presence of $\underline{\mathbf{h}}_k$ on both sides of the SOC constraint (8.85), this reformulated problem is still difficult to solve. A conservative design is then developed by assuming independent uncertainties for $\underline{\mathbf{h}}_k$ on the left and right hand sides of each constraint in (8.85). With such an assumption, problem (8.85) can be transformed to the following SDP problem and solved efficiently using standard interior point method.

$$\begin{aligned} \min_{\underline{\mathbf{V}}, \eta, \kappa, t} \quad & t \\ \text{s.t.} \quad & \|\operatorname{vec}([\underline{\mathbf{v}}_1, \dots, \underline{\mathbf{v}}_K])\| \leq t \\ & \begin{bmatrix} \kappa_k - \eta_k & \mathbf{0} & [\hat{\mathbf{h}}_k \underline{\mathbf{V}}, 1] \\ \mathbf{0} & \eta_k \mathbf{I}_{2N_t} & \delta_k [\underline{\mathbf{V}}, \mathbf{0}] \\ [\hat{\mathbf{h}}_k \underline{\mathbf{V}}, 1]^T & \delta_k [\underline{\mathbf{V}}, \mathbf{0}]^T & \kappa_k \mathbf{I}_{2K+1} \end{bmatrix} \succeq 0, \quad k = 1, \dots, K \\ & \begin{bmatrix} \sqrt{1 + \gamma_k^{-1}} \hat{\mathbf{h}}_k \underline{\mathbf{v}}_k - \kappa_k & \delta_k \sqrt{1 + \gamma_k^{-1}} \underline{\mathbf{v}}_k^T \\ \delta_k \sqrt{1 + \gamma_k^{-1}} \underline{\mathbf{v}}_k & \left(\sqrt{1 + \gamma_k^{-1}} \hat{\mathbf{h}}_k \underline{\mathbf{v}}_k - \kappa_k \right) \mathbf{I}_{2N_t} \end{bmatrix} \succeq 0, \quad k = 1, \dots, K. \end{aligned} \quad (8.86)$$

Instead of solving (8.84) by the SDP relaxation (8.86), Vucic and Boche [113] proposed to solve (8.84) by: (1) directly applying the ellipsoid method from convex optimization and (2) approximating (8.84) by a robust MSE constrained min power problem. Let $u_k \in \mathcal{C}$ denote the scalar receive filter used at receiver k . Let \mathbf{e}_k denote the unit vector with its k -th element being 1. Define the MSE of the k -th user as

$$\text{MSE}_k \left(\{\mathbf{v}_k, u_k\}_{k=1}^K \right) = |u_k|^2 \left(\left(\mathbf{h}_k \mathbf{V} - \frac{1}{u_k} \mathbf{e}_k \right) \left(\mathbf{h}_k \mathbf{V} - \frac{1}{u_k} \mathbf{e}_k \right)^H + 1 \right). \quad (8.87)$$

The robust MSE constrained min power problem is given as

$$\min_{\{\mathbf{v}_k, u_k\}_{k=1}^K} \sum_{k=1}^K \|\mathbf{v}_k\|^2 \quad (8.88)$$

$$\text{s.t.} \quad \text{MSE}_k \left(\{\mathbf{v}_k, u_k\}_{k=1}^K \right) \leq \frac{1}{1 + \gamma_k}, \quad \forall \mathbf{h}_k \in \mathcal{U}_k(\Delta_k), k = 1, \dots, K. \quad (8.89)$$

This problem is convex and can be equivalently formulated as an SDP problem and efficiently solved by interior point methods. It is shown in [113] that both the ellipsoid method approach and the robust MSE constrained reformulation approach achieve better performance than the SDP relaxation (8.86) in terms of various system level performance measures.

As noted earlier, the original min power SINR constrained problem (8.83) is not equivalent to the formulation (8.84), as the latter replaces the nonlinear term $|\mathbf{h}_k \mathbf{v}_k|$ by a linear lower bound $\text{Re}(\mathbf{h}_k \mathbf{v}_k) = \mathbf{h}_k \mathbf{v}_k$. Implicit in this reformulation is the additional requirement that $\text{Re}(\mathbf{h}_k \mathbf{v}_k)$ is positive for all the channels $\mathbf{h}_k \in \mathcal{U}_k(\delta_k)$. Recently, the authors of [114] showed that the direct SDP relaxation of the original problem (8.83) is actually tight as long as the size of the uncertainty set is sufficiently small. This implies that robust resource allocation for MISO channels can be solved to global optimality in polynomial time, provided the channel uncertainty is small. More precisely, define $\mathbf{V}_k = \mathbf{v}_k \mathbf{v}_k^H$, and $\mathbf{X}_k = \frac{1}{\gamma_k} \mathbf{V}_k - \sum_{l \neq k} \mathbf{V}_l$, the problem (8.83) can be equivalently reformulated as

$$\begin{aligned} & \min_{\{\mathbf{V}_k, \kappa_k\}_{k=1}^K} \sum_{k=1}^K \text{Tr}(\mathbf{V}_k) \\ & \text{s.t.} \quad \kappa_k \geq 0, \quad k = 1, \dots, K \\ & \quad \begin{bmatrix} \mathbf{X}_k + \kappa_k \mathbf{I} & \mathbf{X}_k \hat{\mathbf{h}}_k^H \\ \hat{\mathbf{h}}_k \mathbf{X}_k^H & \hat{\mathbf{h}}_k \mathbf{X}_k^H \hat{\mathbf{h}}_k^H - 1 - \kappa_k \delta_k^2 \end{bmatrix} \succeq 0, \quad k = 1, \dots, K. \\ & \quad \mathbf{V}_k \succeq 0, \quad \text{rank}(\mathbf{V}_k) = 1, \quad k = 1, \dots, K. \end{aligned} \quad (8.90)$$

When the rank constraints are dropped, this problem becomes the following SDP and can be efficiently solved

$$\begin{aligned} & \min_{\{\mathbf{V}_k, \kappa_k\}_{k=1}^K} \sum_{k=1}^K \text{Tr}(\mathbf{V}_k) \\ & \text{s.t.} \quad \kappa_k \geq 0, \quad k = 1, \dots, K \\ & \quad \begin{bmatrix} \mathbf{X}_k + \kappa_k \mathbf{I} & \mathbf{X}_k \hat{\mathbf{h}}_k^H \\ \hat{\mathbf{h}}_k \mathbf{X}_k^H & \hat{\mathbf{h}}_k \mathbf{X}_k^H \hat{\mathbf{h}}_k^H - 1 - \kappa_k \delta_k^2 \end{bmatrix} \succeq 0, \quad k = 1, \dots, K. \\ & \quad \mathbf{V}_k \succeq 0, \quad k = 1, \dots, K. \end{aligned} \quad (8.91)$$

Let $\delta = [\delta_1, \dots, \delta_K]$. Let P_δ denote the above SDP problem when the bounds on the uncertainty set is δ . Let $P^*(\delta)$ denote the optimal value of the the problem P_δ . Suppose that for some choice of uncertainty bounds $\bar{\delta} = [\bar{\delta}_1, \dots, \bar{\delta}_K]^T > 0$, the problem $P_{\bar{\delta}}$ is strictly feasible. Define the set

$$\Omega(\bar{\delta}) = \left\{ \delta \mid \delta_k \leq \bar{\delta}_k \text{ and } \delta_k < \sqrt{\frac{\gamma_k}{P^*(\bar{\delta})}}, \quad k = 1, \dots, K \right\}. \quad (8.92)$$

Then, according to Song et al. [114], for any vector of uncertainty bounds $\delta \in \Omega(\bar{\delta})$, the problem P_δ is feasible. Moreover, its optimal solution $\{V_k^*\}_{k=1}^K$ satisfies $\text{rank}(V_k^*) = 1$, $k = 1 \dots, K$, and it must be the optimal solution of the original problem (8.83).

Alternative system level objectives and constraints can be considered to result in different formulations of the robust resource allocation problem. For example, Ref. [115] considered robust design for both the averaged sum MSE minimization problem and the worst case sum MSE minimization problem. Ref. [116] considered the worst case weighted sum rate maximization problem and min-rate maximization problem. The authors of [117] considered the robust beamformer design in a cognitive radio network in which there are additional requirements that the transmitter's interference to the primary users should be kept under a prescribed level. However, most of the above cited works focus on robust design in a single cell network with a single transmitter. The extensions to the general MIMO IC/IBC/IMAC will be interesting.

To close this section, we summarize the properties of most of the algorithms discussed in this section in Table 8.4. These algorithms usually admit certain forms of decentralized implementation, in which the computational loads are distributed to different entities in the network. We emphasize that the per iteration computational complexity and the amount of message exchanges are important characteristics for practical implementation of these distributed algorithms. Efficient computation ensures real time implementation, while fewer number of message exchanges per iteration implies less signaling overhead. In Table 8.4, these characteristics are listed for each of the algorithms. We note that the computational complexity and the required message exchanges are calculated on a per iteration basis, where in one iteration each user $k \in \mathcal{K}$ completes one update. Also note that in Table 8.4, the variable T in ICBF, SCALE-Dual, and M-IWF represents the number of *inner* iterations needed; the variable ϵ in Bisection-SOCP and SCALE-GP represents the required precision for their respective inner solutions; the variable t in MAPEL represents the iteration index; the variable B in BB and ISB represents the maximum number of transmitted bits allowed for each subchannel; the variable C in the BB algorithm represents its computation overhead.

2.08.4 Distributed resource allocation in interference channel

Most of the algorithms introduced in the previous section are either centralized or require certain level of user coordination. Such coordination may be costly in infrastructure based networks, and is often infeasible for fully distributed networks. In this section we discuss fully distributed resource allocation algorithms that require no user coordination.

Table 8.4 Comparisons of Resource Allocation Algorithms

Algorithm	Optimality	Complexity per Iteration	Convergence Status	Coordination Level	Message Exchange	Channel Model	Update Schedule	Problem Formulation
APC [46]	Global	$O(K)$	Yes	Distributed	$O(K)$	Scalar IC	Sequential	Min SINR
BB [110]	Global	Lower bounded By $O((3 + B/2)K^3 + 3K^2/2 + C)$	Yes	Centralized	N/A	Parallel IC	N/A	Sum rate
Bisection-SOCP ([33])	Global	$O\left(K^{3.5}N_t^{3.5} \log\left(\frac{1}{\epsilon}\right)\right)$	Yes	Centralized	N/A	MISO IC	N/A	Min SINR
CCA [33]	Local	$O(N_t K^2)$	Yes	Distributed	$O(K^2)$	MISO IC	Sequential	Smooth utility
ICBF [66]	Unknown	$O(TK^2 N_t^3)$	Unknown	Distributed	$O(K^2)$	MISO IC	Sequential	Sum rate
						MISO IBC/IMAC		
ISB [54]	Unknown	$O(BNK)$	Unknown	Distributed	$O(K^2 N)$	Parallel IC	Sequential	Sum rate
GP [102]	Unknown	$O(K^3)$ (Scalar IC Case)	Yes	Centralized	N/A	Parallel IC	N/A	Mixed sum rate
NFP ([106])	Global	Upper bounded By K^t	Yes	Centralized	N/A	Parallel IC Scalar IC	N/A	Mixed sum rate
MDP [61]	Local	$O(KN \log N + K^2 N)$ (Parallel IC Case)	Yes	Distributed	$O(K^2 N)$	MISO IC Parallel IC	Sequential	Sum rate
MIMO-DP [65]	Unknown	$O(K^3(N_t N_r^2 + N_t^2 N_r) + K^2 N_r^3)$	Unknown	Distributed	$O(K^2 N_r^2)$	MIMO IC	Sequential	Sum rate
M-IWF [59]	Unknown	$O(TKN \log N + K^2 N)$	Unknown	Distributed	$O(K^2 N)$	Parallel IC	Simultaneous	Sum rate
SCALE-Dual [68]	Local	$O(TKN + NK^2)$	Yes	Distributed	$O(K^2 N)$	Parallel IC	Simultaneous	Sum rate and min power
SCALE-GP [68]	Local	$O(K^4 N^4 \log (KN/\epsilon))$	Yes	Centralized	N/A	Parallel IC	N/A	Sum rate and min power
WMMSE-MIMO [73]	Local	$O(K^2(N_t N_r^2 + N_t^2 N_r + N_t^3) + KN_r^3)$ (MIMO IC Case)	Yes	Distributed	$O(K^2 N_r^2)$	MIMO IC IBC/IMAC	Simultaneous	Utility satisfy 8.67–8.69
WMMSE-Parallel [73]	Local	$O(K^2 N^3)$ (MIMO IC Case)	Yes	Distributed	$O(K^2 N^2)$	Parallel IC Parallel IBC/IMAC	Simultaneous	Utility satisfy 8.67–8.69

2.08.4.1 Game theoretical formulations

In scenarios where users cannot exchange information explicitly, it is no longer possible to allocate resources using the maximizer of a system wide utility function. Instead, we need to rely on alternative solution concepts for distributed resource allocation. One such concept that is particularly useful in our context is the renowned notion of Nash equilibrium (NE) for a noncooperative game; see [118, 119], and the [Yale Open Course](#) online. In a noncooperative game, there are a number of players, each seeking to maximize its own utility function by choosing a strategy from an individual strategy set. However, the utility of one player depends on not only the strategy of its own, but also those of others in the system. As a result, when players have conflicting utility functions, there is usually no joint player strategy that will simultaneously maximize the utilities of all players. For such a noncooperative game, a NE solution is defined as a tuple of joint player strategies in which no single player can benefit by changing its own strategy unilaterally.

Mathematically, a K -person noncooperative game in the strategic form is a three tuple $(\mathcal{K}, \chi, \mathbf{U})$, in which $\mathcal{K} = \{1, \dots, K\}$ is the set of players of the game; $\chi = \prod_{k=1}^K \chi_k$ is the joint strategic space of all the players, with χ_k being player k 's individual strategy space; $\mathbf{U} = [U_1, \dots, U_K]$, where $U_k(\mathbf{x}_k, \mathbf{x}_{-k}) : \chi \mapsto \mathcal{R}$ is user k 's utility function. In the above definition we have used $\mathbf{x}_k \in \chi_k$ to denote player k 's strategy, $\mathbf{x}_{-k} = \{\mathbf{x}_l\}_{l \neq k}$ to denote the strategies of all remaining users. It is clear that player k 's strategy depends on its own strategy $\mathbf{x}_k \in \chi_k$ as well as those of others $\mathbf{x}_{-k} \in \chi_{-k}$. A NE of the game \mathcal{G} is defined as the set of joint strategies of all the players $\mathbf{x}^* \in \chi$ such that the following inequality is satisfied simultaneously for all players $k \in \mathcal{K}$

$$U_k(\mathbf{x}_k^*, \mathbf{x}_{-k}^*) \geq U_k(\mathbf{x}_k, \mathbf{x}_{-k}^*), \quad \forall \mathbf{x}_k \in \chi_k. \quad (8.93)$$

Clearly at a NE, the system is stable as none of the players has any intention to switch to a different strategy. We define a *best response function* for each player in the game, as its best strategy when all other players have their strategies fixed

$$BR_k(\mathbf{x}_{-k}) = \arg \max_{\mathbf{x}_k \in \chi_k} U_k(\mathbf{x}_k, \mathbf{x}_{-k}). \quad (8.94)$$

Using this definition, a NE of the game \mathcal{G} can be alternatively defined as

$$\mathbf{x}^* \in BR_k(\mathbf{x}_{-k}^*), \quad \forall k = 1, \dots, K. \quad (8.95)$$

Figure 8.12 is an illustration of the NE point of a game with 2-player and affine best response functions. This figure also shows how a sequence of best response may enable the players to approach the NE.

Let us illustrate the notion of NE in our 2-user scalar IC model (8.14). Suppose these two users are the players of a game, and their strategy spaces are $\chi_k = \{p_k \mid 0 \leq p_k \leq \bar{p}_k\}$, $k = 1, 2$. Assume that the users' utility functions are their maximum transmission rates defined in (8.18). Thus, for this example, user 1's best response function admits a particular simple expression

$$\begin{aligned} BR_1(p_2) &= \arg \max_{0 \leq p_1 \leq \bar{p}_1} \log_2 \left(1 + \frac{|H_{11}|^2 p_1}{1 + |H_{21}|^2 p_2} \right) \\ &= \bar{p}_1. \end{aligned} \quad (8.96)$$

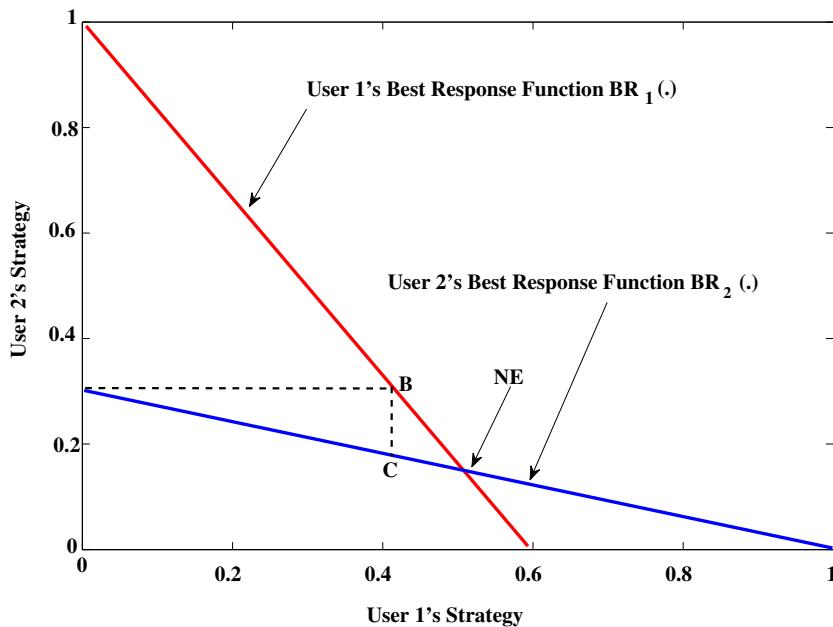
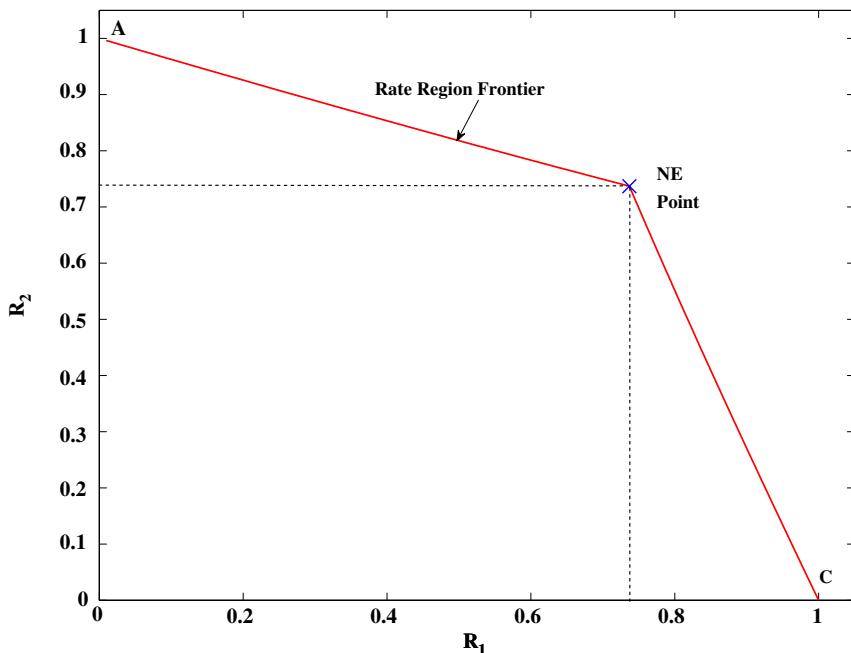
**FIGURE 8.12**

Illustration of the Nash Equilibrium (NE) of a 2-user scalar game. This figure also shows the process of a sequence of best responses that reach the NE. The function $BR_k(\cdot)$ represents user k 's best response function. Suppose both users initially choose 0. User 2 acts first and chooses point A which is its best response, User 1 acts next and chooses its best response point B. User 2 acts again and chooses point C. Continuing iteratively in this fashion, the NE will be reached in the limit.

This says that *regardless* of user 2's transmission strategy, user 1 will transmit with full power. The same can be said about user 2. Consequently, the only NE of this game is the transmit power tuple (\bar{p}_1, \bar{p}_2) . Obviously, assuming that each user is indeed selfish and they intend to maximize their own utility, the NE point (\bar{p}_1, \bar{p}_2) can be implemented *without* any explicit coordination between the users. Now let us assess the efficiency of such power allocation scheme in terms of system sum rate. In Figures 8.13 and 8.14, we plot the rate region boundary and the NE points for different interference levels. We see that when interference is low, the NE corresponds exactly to the maximum sum rate point. However, when interference is strong, the NE scheme is inferior to the time sharing scheme in which the users transmit with full power in an orthogonal and interference free fashion (e.g., TDMA or FDMA). Nonetheless, it should be pointed out that the NE point can be reached without user coordination, while the time sharing scheme requires the users to synchronize their transmissions.

We refer the readers to the September 2009 issue of IEEE Signal Processing Magazine for the applications of game theory to wireless communication and signal processing.

**FIGURE 8.13**

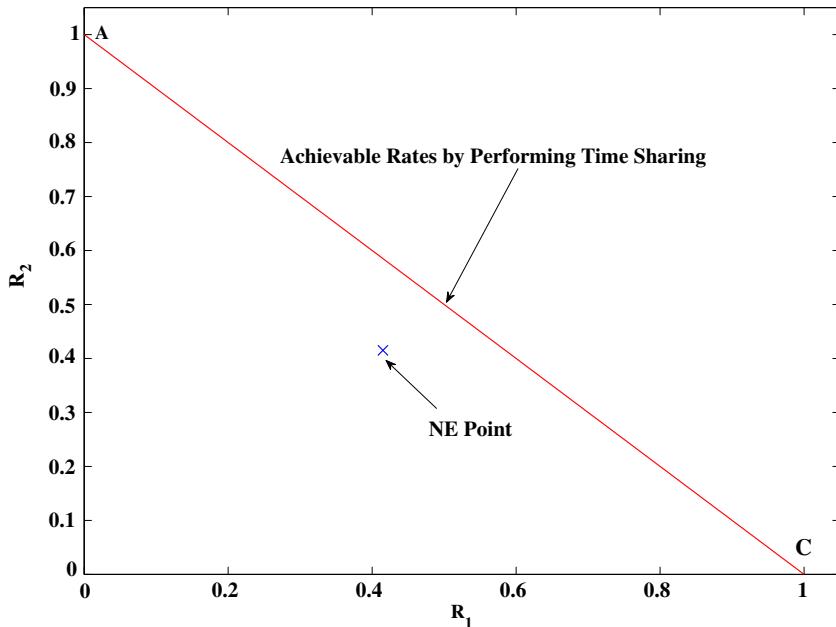
An illustration of the efficiency of the NE for 2-user IC when the interference is weak. $\bar{p}_1 = \bar{p}_2 = 1$, $|H_{11}|^2 = |H_{22}|^2 = 1$, $|H_{12}|^2 = |H_{21}|^2 = 0.5$, $\sigma_1^2 = \sigma_2^2 = 1$. At point A and C, a single user transmits using full power; at the NE point, both users transmit using full power.

2.08.4.2 Distributed resource allocation for interference channels

Early works on distributed physical layer resource allocation in wireless networks largely deal with the scalar IC models. Sarayda et al. [120, 121] and Goodman and Mandayam [122] are among the first to cast the general scalar power control problem in a game theoretic framework. They proposed to quantify the tradeoff between the users' QoS requirements and energy consumption by a utility function in the form:

$$U_k(p_k, \mathbf{p}_{-k}) = \frac{R_k f(\text{SINR}_k)}{p_k}, \quad (8.97)$$

where R_k is user k 's fixed transmission rate, and $f(\text{SINR}_k)$ is a function of user k 's SINR that characterizes its bit error rate (BER). They showed that any NE point is inefficient in the Pareto sense, i.e., it is possible to increase the utility of some of the terminals without hurting any other terminal. To improve the efficiency of the power control game, they proposed to charge the users with a price that is proportional to their transmit powers. Specifically, each users' utility function now becomes $U_k(p_k, \mathbf{p}_{-k}) = R_k f(\text{SINR}_k)/p_k - \alpha_k p_k$, where α_k is a positive scalar that can be appropriately chosen by the system operator. They showed that this modified game always admits a NE, and proposed an algorithm that allows the users to reach one of the NEs by adapting their transmit powers in the best

**FIGURE 8.14**

An illustration of the inefficiency of the NE for 2-user IC when the interference is strong. $\bar{p}_1 = \bar{p}_2 = 1$, $|H_{11}|^2 = |H_{22}|^2 = 1$, $|H_{12}|^2 = |H_{21}|^2 = 2$, $\sigma_1^2 = \sigma_2^2 = 1$. At point A and C, a single user transmits using full power; at the NE point, both users transmit using full power.

response fashion. Meshkati et al. [123, 124] extended the above works to the multi-carrier data network. They defined the utility function for each user as

$$U_k(\mathbf{p}_k, \mathbf{p}_{-k}) = \frac{\sum_{n=1}^N R_k^n f_k(\text{SINR}_k^n)}{\sum_{n=1}^N p_k^n}, \quad (8.98)$$

where the function $f_k(\cdot)$ represents the BER of user k , and it incorporates the underlying structure of different linear receivers. However, such multi-carrier power control game is more complicated than the scalar power control game introduced early, and in certain network configurations it is possible that no NE exists (see [123]).

An alternative approach in distributed power control is to directly optimize the individual users' transmission rates. Consider a K -user N -channel parallel IC model. Assume that each user $k \in \mathcal{K}$ is interested in maximizing its transmission rate, and again assume that its total transmission power budget is \bar{p}_k . Then the users' utility functions as well as their feasible spaces can be expressed as

$$U_k(\mathbf{p}_k, \mathbf{p}_{-k}) = \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n} \right), \quad k = 1, \dots, K, \quad (8.99)$$

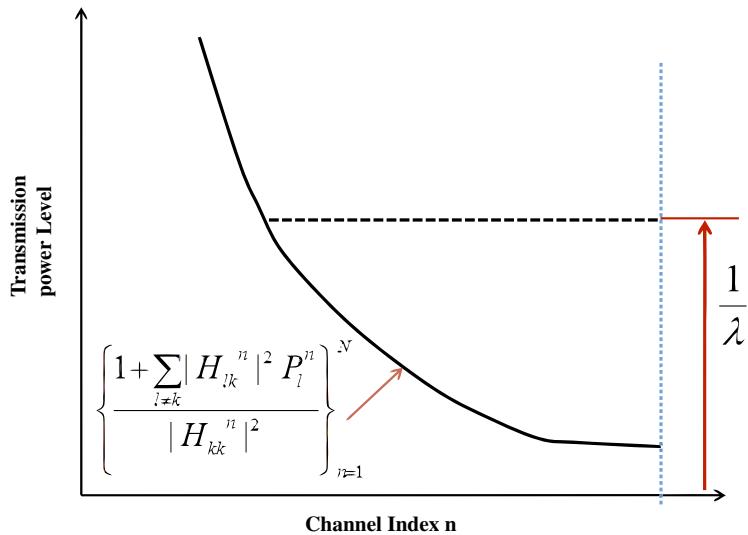

FIGURE 8.15

Illustration of the Water-Filling computation for user k .

$$\chi_k = \left\{ \mathbf{p}_k \left| \sum_{n=1}^N p_k^n \leq \bar{p}_k, \quad p_k^n \geq 0, \quad n = 1, \dots, N \right. \right\}, \quad k = 1, \dots, K. \quad (8.100)$$

Fixing \mathbf{p}_{-k} , user k 's best response solution \mathbf{p}_k^* is the classical *water-filling* (WF) solution

$$p_k^{n,*} = \left[\frac{1}{\lambda_k} - \frac{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n}{|H_{kk}^n|^2} \right]^+, \quad n = 1, \dots, N, \quad (8.101)$$

where $\lambda_k \geq 0$ is the dual variable ensuring the sum power constraint, and the operator $[x]^+ = \max\{x, 0\}$. Figure 8.15 illustrates the WF solution for user k . We note that in order to compute the WF solution, user k needs to know the terms $\{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n\}_{n=1}^N$, which is simply the set of noise-plus-interference (NPI) levels on all its channels. They can be measured *locally* at its receiver. We refer to this game as a WF game.

Yu et al. [60] is the first to formulate the distributed power control problem as a WF game. The authors proposed an iterative water-filling algorithm (IWFA) in which the following two steps are performed iteratively:

- 1. each user $k \in \mathcal{K}$ measures its NPI power $\{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^{n,(t)}\}_{n=1}^N$;
- 2. each user $k \in \mathcal{K}$ computes its power allocation $\mathbf{p}_k^{(t+1)}$ according to (8.101).

Variations of the IWFA algorithm allow the users to update using different schedules. The following three update schemes have been proposed: (a) *simultaneous update*, in which all the users update their power in each iteration, see [60, 125]; (b) *sequential update*, in which a single user updates in each iteration, see [125, 126]; (c) *asynchronous update*, in which a random fraction of users update in each iteration, and they are allowed to use *outdated* information in their computation, see [127]. Regardless of the specific update schedule used, the IWFA algorithm is a distributed algorithm because only local NPI measurements are needed for the users to perform their independent power update.

The properties of the WF game as well as the convergence conditions of the IWFA have been extensively studied. The original work has only provided sufficient conditions for the convergence of the IWFA in a 2-user network. Subsequent works such as [125, 126, 128–130] generalized this result to networks with arbitrary number of users. Luo and Pang [126] characterized the NE of the water-filling game as the solution to the following affine variational inequality (AVI)

$$(\mathbf{p}' - \mathbf{p})(1 + \mathbf{M}\mathbf{p}) \geq 0, \quad \forall \mathbf{p}' \in \chi, \quad (8.102)$$

where \mathbf{M} is a block partitioned matrix with its (ij) th block defined as $\mathbf{M}^{ij} = \text{diag}\left(\frac{|H_{ji}^1|^2}{|H_{ii}^1|^2}, \dots, \frac{|H_{ji}^N|^2}{|H_{ii}^N|^2}\right)$.

Using the AVI characterization (8.102), they also showed that the sequential version of the IWFA corresponds to the classical projection algorithm whose convergence to the unique NE of the water-filling game is guaranteed if the following contraction condition is satisfied

$$\rho((\mathbf{I} - \boldsymbol{\Upsilon}_{\text{low}})^{-1} \boldsymbol{\Upsilon}_{\text{upp}}) < 1, \quad (8.103)$$

where $\boldsymbol{\Upsilon}_{\text{low}}$ and $\boldsymbol{\Upsilon}_{\text{upp}}$ is the strictly lower and strictly upper triangular part of a $K \times K$ matrix $\boldsymbol{\Upsilon}$ given by

$$[\boldsymbol{\Upsilon}]_{q,r} \triangleq \begin{cases} \max_n \left\{ \frac{|H_{rq}^n|^2}{|H_{qq}^n|^2} \right\} & \text{if } r \neq q, \\ 0 & \text{otherwise.} \end{cases} \quad (8.104)$$

Also using the AVI characterization (8.102), the authors of [129, 125] further proved that the condition

$$\rho(\boldsymbol{\Upsilon}) < 1 \quad (8.105)$$

is sufficient for the convergence of the IWFA as well as the uniqueness of the NE. We refer the readers to [131] for a detailed comparison of various conditions for the convergence of the IWFA. It is worth noticing that all the sufficient conditions for the convergence of IWFA require that the interference among the users are weak. For example, a sufficient condition for (8.105) is

$$\sum_{l \neq k} |H_{kl}^n|^2 s_k < |H_{ll}^n|^2 s_l, \quad \forall k \in \mathcal{K}, n \in \mathcal{N}, \quad (8.106)$$

where $s_k > 0$, $k = 1, \dots, K$ is a set of constant scalars. Intuitively, this condition says that at the receiver of each user $k \in \mathcal{K}$, the power of the useful signal should be larger than the power of total interference. When the interference is strong, IWFA diverges. Leshem and Zehavi [132] provided an example in which *all* forms of IWFA diverge, regardless of their update schedules. We remark that

extending the IWFA so that it converges in less stringent conditions that do not require the interference to be weak is still an open problem. Without any algorithmic modifications, the standard IWFA is only known to converge [126] when the crosstalk coefficients are symmetric

$$\frac{|H_{rq}^n|^2}{|H_{qq}^n|^2} = \frac{|H_{qr}^n|^2}{|H_{rr}^n|^2}, \quad \forall r \neq q, \forall n,$$

regardless the interference levels.

The IWFA has been recently generalized to MIMO IC model. In the MIMO WF game, the strategy of each user $k \in \mathcal{K}$ is its transmission covariance matrix \mathbf{Q}_k . The rate utility function and strategy set for user k can be expressed as

$$U_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) = \log_2 \det \left(\mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \left(\mathbf{I}_{N_r} + \sum_{l \neq k} \mathbf{H}_{lk} \mathbf{Q}_l \mathbf{H}_{lk}^H \right)^{-1} + \mathbf{I}_{N_r} \right), \quad (8.107)$$

$$\chi_k = \{\mathbf{Q}_k : \text{Tr}(\mathbf{Q}_k) \leq \bar{p}_k, \mathbf{Q}_k \succeq 0\}. \quad (8.108)$$

In this case, each user k 's best response $BR_k(\mathbf{Q}_{-k})$ is again a water-filling solution, see [133]. Arslan et al. [134] suggested that in each iteration, the users' covariance can be updated as

$$\mathbf{Q}_k^{(t+1)} = \alpha_t \mathbf{Q}_k^{(t)} + (1 - \alpha_t) BR_k(\mathbf{Q}_{-k}^{(t)}), \quad (8.109)$$

where $\{\alpha_t\}_{t=1}^\infty$ is a set of constants that satisfy $\alpha_t \geq 0$, $\lim_{t \rightarrow \infty} \alpha_t = 0$ and $\lim_{t \rightarrow \infty} \sum_{t=1}^T \alpha_t < \infty$. They claimed that their algorithm converges when the interference is weak, but no specific conditions are given. This work has been generalized by Scutari et al. [131, 135], in which rigorous conditions for the convergence of the MIMO IWFA have been derived. In particular, consider a MIMO network in which $N_t = N_r$ and the direct link channel matrices $\{\mathbf{H}_{kk}\}_{k=1}^K$ are all nonsingular. Define a $K \times K$ matrix \mathbf{S} as

$$[\mathbf{S}]_{q,r} \triangleq \begin{cases} \rho \left(\mathbf{H}_{rq}^H \mathbf{H}_{qq}^{-H} \mathbf{H}_{qq}^{-1} \mathbf{H}_{rq} \right) & \text{if } r \neq q, \\ 0 & \text{otherwise.} \end{cases} \quad (8.110)$$

Then the condition $\rho(\mathbf{S}) < 1$ is sufficient for the convergence of the sequential/simultaneous/asynchronous MIMO IWFA. This condition is again a weak interference condition, and future work is needed to extend the MIMO IWFA to work in networks without this restriction. We refer the readers to web pages of [Yu](#), [Palomar](#), and [Pang](#) for other works related to the WF games and IWFA.

The above parallel and MIMO WF games have been extended in several directions. A series of recent works considered the robustness issue in a WF games. For instance, Gohary et al. [136] considered the WF game in the presence of a jammer. Let us denote user 0 as the jammer and denote its transmission power as $\mathbf{p}_0 = [p_0^1, \dots, p_0^N]^T$. The rate utility function of a normal user k ($k \neq 0$) becomes

$$U_k(\mathbf{p}_k, \mathbf{p}_{-k}, \mathbf{p}_0) = \sum_{n=1}^N \log_2 \left(1 + \frac{|\mathbf{H}_{kk}^n|^2 p_k^n}{1 + \sum_{l \neq k} |\mathbf{H}_{lk}^n|^2 p_l^n + |\mathbf{H}_{0k}^n|^2 p_0^n} \right). \quad (8.111)$$

Suppose the jammer's objective is to minimize the utility of the whole system. This can be reflected by its utility function and the strategy set

$$U_0(\mathbf{p}_0, \mathbf{p}) = - \sum_{k=1}^K U_k(\mathbf{p}_k, \mathbf{p}_{-k}, \mathbf{p}_0), \quad (8.112)$$

$$\chi_0 = \left\{ \mathbf{p}_0 : \sum_{n=1}^N p_0^n \leq \bar{p}_0, p_0^n \geq 0, \forall n \in \mathcal{N} \right\}. \quad (8.113)$$

Gohary et al. [136] proposed a generalized IWFA (GIWFA) algorithm in which the normal users and the jammer all selfishly maximize their respective utility functions. Notice that the selfish maximization problems are all convex. The convergence condition for the GIWFA is

$$\rho((\mathbf{I} - \boldsymbol{\Upsilon}_{\text{low}})^{-1} \boldsymbol{\Upsilon}_{\text{upp}}) \leq \frac{1}{1+a} - c < 1, \quad (8.114)$$

where the matrix $\boldsymbol{\Upsilon}$ is defined in (8.104), and $a > 0$ and $c > 0$ are constants related to the system parameters. Clearly this condition is more restrictive than those of the original IWFA, for example the condition (8.103). This is partly because the presence of the jammer introduces *uncertainty* to the NPI that each normal user experiences.

Uncertainty of the NPI is also caused by events such as sudden changes in the number of users in the system or errors of interference measurement at the receivers. Setoodeh and Haykin [137] seek a formulation that takes into consideration the worst case NPI errors. Let I_k^n denote the power of NPI that user k should have experienced on channel n if no measurement errors occur. Let $\tilde{I}_k^n = I_k^n - \Delta I_k^n$ be the measured NPI value, with ΔI_k^n representing the NPI uncertainty. Let $\Delta \mathbf{I}_k = [\Delta I_1^n, \dots, \Delta I_K^n]^T$ and suppose it is bounded, i.e., $\|\Delta \mathbf{I}_k\| \leq \epsilon_k$ for some $\epsilon_k > 0$. In this robust WF game, the objectives of the users are to maximize their *worst case* transmission rate. In other words, user k 's utility function can be expressed as

$$U_k(\mathbf{p}_k, \mathbf{p}_{-k}) = \min_{\|\Delta \mathbf{I}_k\| \leq \epsilon_k} \sum_{n=1}^N \log_2 \left(1 + \frac{|H_{kk}^n|^2 p_k^n}{\tilde{I}_k^n + \Delta I_k^n} \right). \quad (8.115)$$

This formulation trades performance in favor of robustness, thus the equilibrium solution obtained is generally less efficient than that of the original IWFA. In [138], an averaged version of IWFA was proposed which converges when the error of the NPI $\Delta \mathbf{I}_k$ satisfies certain conditions. In [139], the authors provided a probabilistically robust IWFA to deal with the quantization errors of the NPI at the receiver of each user. In this algorithm, users allocate their powers to maximize their total rate for a large fraction of the error realization.

Another thread of works such as [58, 140–143] generalized the original WF game and the IWFA to interfering cognitive radio networks (CRN). In a CRN, the secondary users (SUs) are allowed to use the spectrum that is assigned to the primary users (PUs) as long as the SUs do not create excessive interference to the primary network. Suppose the secondary network is a K -user N -channel parallel IC. Let $\mathcal{Q} = \{1, \dots, Q\}$ denote the set of PUs in the network. Let $|G_{kq}^n|^2$ denote the channel gain from SU k to PU q on channel n . The following aggregated interference constraints are imposed on the

secondary network (these constraints are also referred to as the interference temperature-constraints, see [144, 145])

$$\sum_{k=1}^K |G_{kq}^n|^2 p_k \leq \bar{I}_q^n, \quad \forall (q, n) \in \mathcal{Q} \times \mathcal{N}, \quad (8.116)$$

where \bar{I}_q^n represents the maximum aggregated interference allowed at the receiver of PU q on channel n . The original WF algorithm needs to be properly modified to strictly enforce these interference constraints in the equilibrium. Xie et al. [143] formulated the power allocation problem in this CRN as a competitive market model. In this model, each channel has a fictitious price per unit power, and the users must purchase the transmission power on each channel to maximize their data rates. Scutari et al. [135, 146] systematically studied the WF game with interference constraints. For each primary user q , they introduced a set of interference prices $\mathbf{v}_q = [v_q^1, \dots, v_q^N]^T$. Each SU is charged for their contribution of total interference at PUs' receiver. Specifically, a SU k 's utility function and feasible set is defined as

$$U_k(\mathbf{p}_k, \mathbf{p}_{-k}, \mathbf{v}) = R_k(\mathbf{p}_k, \mathbf{p}_{-k}) - \sum_{q=1}^Q \sum_{n=1}^N v_q^n |G_{kq}^n|^2 p_k^n, \quad (8.117)$$

$$\chi_k = \left\{ \mathbf{p}_k : \sum_{n=1}^N p_k^n \leq \bar{p}_k, p_k^n \geq 0, n = 1, \dots, N \right\}, \quad (8.118)$$

where $R_k(\mathbf{p}_k, \mathbf{p}_{-k})$ is user k 's transmission rate. The NE of this interference-constrained WF game is the tuple $(\mathbf{v}^*, \mathbf{p}^*)$ that satisfies the following conditions

$$\begin{aligned} \mathbf{p}_k^* &= \max_{\mathbf{p}_k \in \chi_k} U_k(\mathbf{p}_k, \mathbf{p}_{-k}^*, \mathbf{v}^*), \\ &\sum_{k=1}^K |G_{kq}^n|^2 p_k^{n,*} \leq \bar{I}_q^n, (q, n) \in \mathcal{Q} \times \mathcal{N}, \\ &v_q^{n,*} \left(\bar{I}_q^n - \sum_{k=1}^K |G_{kq}^n|^2 p_k^{n,*} \right) = 0, \quad v_q^{n,*} \geq 0, (q, n) \in \mathcal{Q} \times \mathcal{N}. \end{aligned}$$

Scutari et al. [135, 146] derived the conditions for the existence and uniqueness of the NE for this game. Introduce a $N \times N$ matrix $\widehat{\mathbf{\Upsilon}}$

$$[\widehat{\mathbf{\Upsilon}}]_{l,k} \triangleq \begin{cases} -\max_{n \in \mathcal{N}} \left\{ \frac{|H_{lk}^n|^2}{|H_{kk}^n|^2} \times \widehat{\text{innr}}_{lk}^n \right\} & \text{if } k \neq l, \\ 1 & \text{otherwise,} \end{cases} \quad (8.119)$$

where $\widehat{\text{innr}}_{lk}^n \triangleq 1 + \sum_{m \in \mathcal{K}} |H_{km}^n|^2 \bar{p}_k$. Then the condition $\widehat{\mathbf{\Upsilon}} \succeq 0$ guarantees the uniqueness of the NE. A set of distributed algorithms that alternately update the users' power allocation and the interference prices were proposed to reach the NE of this game. The MIMO generalization has been considered in [147], whereby both the SUs and PUs are equipped with multiple antennas. Another extension [142]

considered the possibility that the SU-PU channels may be uncertain, and formulated a robust WF game that ensures the SU-PU interference constraints are met even in the worst case channel conditions.

All the above mentioned WF games can be categorized as *rate adaptive* (RA) games, in which the users selfishly maximize their own data rates. One drawback of the RA formulation is that individual users have no QoS guarantees. Alternatively, a *fixed margin* (FM) formulation allows each user to minimize its transmission power while maintaining its QoS constraint. The FM formulation is more difficult to analyze due to the *coupling* of the users' strategy spaces resulted from the QoS constraint. For the parallel IC model, the utility function and the strategy set for user k in a FM formulation can be expressed as

$$U_k(\mathbf{p}_k) = - \sum_{n=1}^N p_k^n, \quad (8.120)$$

$$\chi_k(\mathbf{p}_{-k}) = \{\mathbf{p}_k : R_k(\mathbf{p}_k, \mathbf{p}_{-k}) \geq \zeta_k, p_k^n \geq 0, n \in \mathcal{N}\}, \quad (8.121)$$

where ζ_k is the rate target for user k . The solution to the individual users' utility maximization problems, assuming the feasibility of the rate targets, is again a water-filling solution

$$p_k^{n,*} = \left[\lambda_k - \frac{1 + \sum_{l \neq k} |H_{lk}^n|^2 p_l^n}{|H_{kk}^n|^2} \right]^+, \quad n = 1, \dots, N, \quad (8.122)$$

where λ_k is the water-level that is associated with user k 's rate constraint.

A NE of this FM game (which is usually referred to as the *generalized NE* due to the coupling of the users' strategy spaces) is defined as a power vector \mathbf{p}^* that satisfies

$$\mathbf{p}_k^* \in \arg \max_{\mathbf{p}_k \in \chi_k(\mathbf{p}_{-k}^*)} U_k(\mathbf{p}_k, \mathbf{p}_{-k}^*), \quad k = 1, \dots, K. \quad (8.123)$$

Similar to the min-power QoS constrained formulation discussed in the previous section, the first thing we need to characterize for this FM game is the feasibility of a given set of rate targets. The following condition is among many of those that have been derived in [148] which guarantee the existence of a bounded power allocation achieving the given set of rate targets

$$\sum_{l \neq k} \frac{|H_{kl}^n|^2}{|H_{kk}^n|^2} < \frac{1}{\exp\{\zeta_k\} - 1}, \quad (n, k) \in \mathcal{N} \times \mathcal{K}. \quad (8.124)$$

This condition is again a weak interference condition. The following condition is sufficient for the uniqueness of the (generalized) NE of the FM game

$$\sum_{l \neq k} \max_{n \in \mathcal{N}} \left\{ \frac{|H_{kl}^n|^2}{|H_{kk}^n|^2} \right\} < \frac{\beta}{\exp\{\zeta_l\} - 1}, \quad k \in \mathcal{K}, \quad (8.125)$$

where $\beta < 1$ is related to the set of given rate targets. This condition is also sufficient for the convergence of a FM-IWFA in which the users sequentially or simultaneously update their power using the WF solution. Algorithmic extension of this work to the CRN with interference constraints of the form

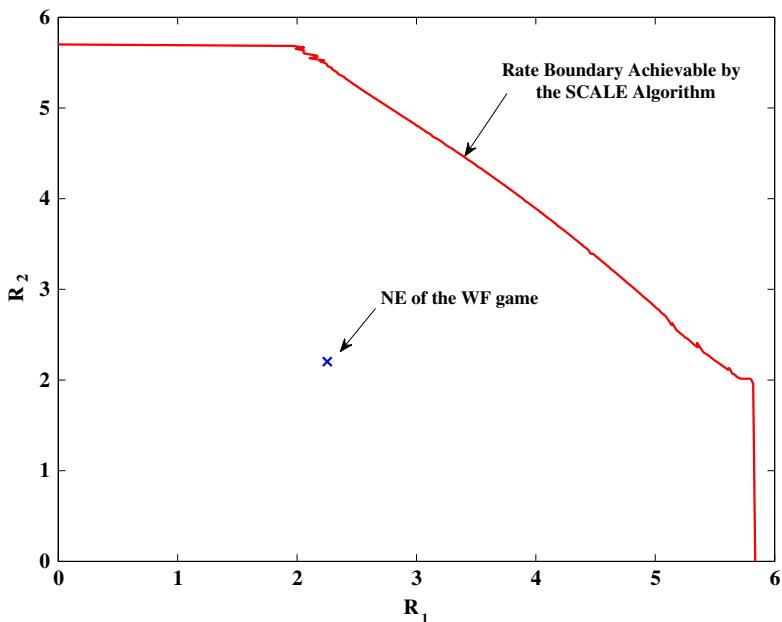
**FIGURE 8.16**

Illustration of the inefficiency of the NE point for a WF game with $K = 2$, $N = 32$, and $\bar{p}_1 = \bar{p}_2 = 1$.

(8.116) has been considered recently in [149]. It remains to see how the FM games and their theoretical properties (e.g., uniqueness of the NE, convergence of the FM-IWFA) can be extended to MIMO CRNs with interference constraints.

We remark that the NE points of the various RA based WF games introduced in this section is generally *inefficient*, in the sense that the sum rate of the users is often smaller compared with that of the socially optimal solutions.¹ In Figure 8.16, we illustrate such inefficiency of the NE in a parallel IC with $K = 2$, $N = 32$ and randomly generated channel coefficients. We plot the NE point of the WF game as well as the rate region boundary achieved by the SCALE algorithm. In order to improve the efficiency of the NE, user coordination must be incorporated into the original WF game. The pricing algorithms such as MDP, M-IWF or WMMSE introduced in the previous section are examples of such extensions. In those algorithms, system efficiency is improved due to explicit message exchange and cooperation among the users. Careful analysis is needed to identify the tradeoffs between the improvement of the system sum rate and the signaling overhead. Evidently, when the total number of users in the system is large, a complete cooperation of all users is too costly. An interesting problem is to decide how to partition the users into collaborative groups in a way that strike an optimal tradeoff between system performance and coordination overhead.

¹However, note that in a MAC channel, which is a special case of the IC, the NEs are indeed efficient. See [150]. In this case, the sequential version of the IWFA converge to a joint strategy that maximizes the system sum rate.

2.08.5 Resource allocation via interference alignment

Theoretically, the optimal resource allocation for a MIMO interference channel is related to the characterization of the capacity region of an interference channel, i.e., determining the set of rate tuples that can be achieved by the users simultaneously. In spite of intensive research on this subject over the past three decades, the capacity region of interference channels is still unknown (even for small number of users). The lack of progress to characterize the capacity region of the MIMO interference channel has motivated researchers to derive various approximations of the capacity region. For example, the maximum total degrees of freedom (DoF) corresponds to the first order approximation of sum-rate capacity in the high SNR regime. Specifically, in a K -user interference channel, we define the degrees of freedom region as the following [21]:

$$\mathcal{D} = \left\{ (d_1, d_2, \dots, d_K) \in \mathbb{R}_+^K \mid \forall (w_1, w_2, \dots, w_K) \in \mathbb{R}_+^K, \right. \\ \left. \sum_{k=1}^K w_k d_k \leq \limsup_{\text{SNR} \rightarrow \infty} \left[\sup_{\mathbf{R} \in \mathcal{C}} \frac{1}{\log \text{SNR}} \sum_{k=1}^K w_k R_k \right] \right\}, \quad (8.126)$$

where \mathcal{C} is the capacity region and R_k is the rate of user k . The total DoF in the system can be defined as the following:

$$\eta = \max_{(d_1, d_2, \dots, d_K) \in \mathcal{D}} d_1 + d_2 + \dots + d_K. \quad (8.127)$$

Roughly speaking, the total DoF is the number of independent data streams that can be communicated in the channel without interference.

For various channel models, the DoF region or the total DoF have been characterized recently. In particular, for a point-to-point MIMO channel with M antennas at the transmitter and N antennas at the receiver, the total DoF is $\eta = \min\{M, N\}$. Different approaches such as SVD precoder or V-BLAST can be used to achieve this DoF bound. For a 2-user MIMO fading interference channel with user k equipped with M_k transmit antennas and N_k receive antennas ($k = 1, 2$), Jafar and Fakhereddin [151] proves that the maximum total DoF is

$$\eta = \min\{M_1 + M_2, N_1 + N_2, \max\{M_1, N_2\}, \max\{M_2, N_1\}\}. \quad (8.128)$$

Therefore, for the case of $M_1 = M_2 = N_1 = N_2$, the total DoF in the system is the same as the single user case. In other words, we do not gain more DoF by increasing the number of users from one to two. Interestingly, if generic channel extensions (drawn from a continuous probability distribution) are allowed either across time or frequency, Cadambe and Jafar [21] showed that the total DoF is $\eta = K M / 2$ for a K -user MIMO interference channel, where M is the number of transmit/receive antennas per user. This surprising result implies that each user can effectively utilize half of the total system resources in an interference-free manner by aligning the interference at all receivers.² Moreover, this total DoF can be achieved by using a carefully designed linear beamforming strategy.

²The idea of interference alignment was introduced in [152–154] and the terminology “interference alignment” was first used in [155].

Mathematically, a linear beamforming strategy for a K -user MIMO IC can be described by the transmit beamforming matrices $\{\mathbf{V}_k\}_{k \in \mathcal{K}}$ and the receive beamforming matrices $\{\mathbf{U}_k\}_{k \in \mathcal{K}}$. The receiver k estimates the transmitted data vector \mathbf{s}_k as follows

$$\mathbf{x}_k = \mathbf{V}_k \mathbf{s}_k, \quad \hat{\mathbf{s}}_k = \mathbf{U}_k^H \mathbf{y}_k, \quad (8.129)$$

where the power of the data vector $\mathbf{s}_k \in \mathbb{R}^{d_k \times 1}$ is normalized such that $E[\mathbf{s}_k \mathbf{s}_k^H] = \mathbf{I}$, and $\hat{\mathbf{s}}_k$ is the estimate of \mathbf{s}_k at the k -th receiver. The matrices $\mathbf{V}_k \in \mathcal{C}^{M_k \times d_k}$ and $\mathbf{U}_k \in \mathcal{C}^{N_k \times d_k}$ are the beamforming matrices at the k -th transmitter and receiver respectively, where M_k (N_k) is the number of antennas at transmitter k (respectively receiver k). Without channel extension, the linear interference alignment conditions can be described by the following zero-forcing conditions [156, 157]

$$\mathbf{U}_k^H \mathbf{H}_{jk} \mathbf{V}_j = 0, \quad k = 1, \dots, K, \quad \forall j \neq k, \quad (8.130)$$

$$\text{rank}(\mathbf{U}_k^H \mathbf{H}_{kk} \mathbf{V}_k) = d_k, \quad k = 1, \dots, K. \quad (8.131)$$

The first equation guarantees that all the interfering signals at receiver k lie in the subspace orthogonal to \mathbf{U}_k , while the second one assures that the signal subspace $\mathbf{H}_{kk} \mathbf{V}_k$ has dimension d_k and is linearly independent of the interference subspace. Clearly, as the number of users K increases, the number of constraints on the beamformers $\{\mathbf{U}_k, \mathbf{V}_k\}$ increases quadratically in K , while the number of design variables in $\{\mathbf{U}_k, \mathbf{V}_k\}$ only increases linearly. This suggests the above interference alignment can not have a solution unless K or d_k is small.

If the interference alignment conditions (8.130) and (8.131) hold for some linear beamforming matrices $\{\mathbf{V}_k, \mathbf{U}_k\}_{k \in \mathcal{K}}$, then transmitter k can use \mathbf{V}_k to send d_k independent data streams to receiver k (per channel use) without any interference. Thus, d_k represents the DoF achieved by the k -th transmitter/receiver pair in the information theoretic sense of (8.126). In other words, the vector (d_1, d_2, \dots, d_K) in (8.130) and (8.131) represents the tuple of DoF achieved by linear interference alignment. Intuitively, the larger the values of d_1, d_2, \dots, d_K , the more difficult it is to satisfy the interference alignment conditions (8.130) and (8.131).

In principle, we can allocate resources by maximizing the total achievable DoF. In particular, for a specific channel realization $\{\mathbf{H}_{kj}\}_{k,j \in \mathcal{K}}$, we need to find the beamforming matrices $\{\mathbf{V}_k, \mathbf{U}_k\}$ to maximize the total DoF while satisfying (8.130) and (8.131).

$$\begin{aligned} & \max_{\{\mathbf{U}_k, \mathbf{V}_k\}_{k=1}^K} \sum_{k=1}^K d_k \\ & \text{subject to } \mathbf{U}_k^H \mathbf{H}_{kj} \mathbf{V}_j = 0, \quad k = 1, \dots, K, \quad j \neq k, \\ & \quad \text{rank}(\mathbf{U}_k^H \mathbf{H}_{kk} \mathbf{V}_k) = d_k, \quad k = 1, \dots, K. \end{aligned}$$

Unfortunately, according to Razaviyayn et al. [156], this problem is NP-hard. So we are led to find suboptimal solution for this problem. However, no efficient algorithms have been developed to approximately solve this problem at this point.

Instead of maximizing the total DoF, we can focus on a seemingly simpler problem: for a given channel realization $\{\mathbf{H}_{kj}\}_{k,j \in \mathcal{K}}$ and a fixed DoF tuple $\mathbf{d} = (d_1, \dots, d_K)$, check if there exist linear

beamformers $\{\mathbf{V}_k, \mathbf{U}_k\}_{k \in \mathcal{K}}$ satisfying the alignment conditions (8.130) and (8.131). Notice that the conditions (8.130) and (8.131) are quadratic polynomial equations, which are difficult to solve in general. However, if we fix either $\{\mathbf{U}_k\}_{k \in \mathcal{K}}$ or $\{\mathbf{V}_k\}_{k \in \mathcal{K}}$, the quadratic equations become linear and can be solved via the linear least squares. This suggests the following alternating directions method for solving (8.130) and (8.131) (for a fixed \mathbf{d}):

1. Fix the transmit beamformers $\{\mathbf{V}_k\}_{k \in \mathcal{K}}$. Each receiver k solves the following optimization problem

$$\begin{aligned} & \min \operatorname{Tr}(\mathbf{U}_k^H \mathbf{Q}_k \mathbf{U}_k) \\ \text{s.t. } & \mathbf{U}_k^H \mathbf{U}_k = \mathbf{I}_{d_k}, \end{aligned} \quad (8.132)$$

where $\mathbf{Q}_k = \sum_{j \neq k} \frac{p_j}{d_j} \mathbf{H}_{kj} \mathbf{V}_j \mathbf{V}_j^H \mathbf{H}_{kj}^H$, with $I_k = \operatorname{Tr}(\mathbf{U}_k^H \mathbf{Q}_k \mathbf{U}_k)$ being the total received interference power, and \bar{p}_j being the power budget of j th transmitter.

2. Fix $\{\mathbf{U}_k\}_{k \in \mathcal{K}}$ and update the transmit beamformers $\{\mathbf{V}_k\}_{k \in \mathcal{K}}$ in a symmetric fashion as in step 1 (by exchanging the roles of transmitter and receiver, and replacing the channel matrices $\{\mathbf{H}_{k,j}\}_{k,j \in \mathcal{K}}$ by $\{\mathbf{H}_{kj}^H\}_{kj \in \mathcal{K}}$).
3. Repeat steps 1 and 2 until convergence.

Notice that the optimal solution \mathbf{U}_k^* for (8.132) is given by the eigen-vectors of \mathbf{Q}_k corresponding to the d_k -smallest eigen-values. The above algorithm is proposed first in [158] and later in [159], albeit from a different perspective. Obviously, this algorithm cannot converge if the DoF vector \mathbf{d} is not achievable. However, even if \mathbf{d} is achievable, there has been no formal analysis that shows this alternating direction algorithm indeed will converge.

The lack of formal convergence proof may not be surprising. In fact, according to [156], even checking the feasibility of (8.130) and (8.131) is NP-hard when each transmitter/receiver is equipped with at least three antennas. Hence, for a given channel realization, assigning DoFs to the users in a manner that ensures feasibility is not easy. However, when the number of antennas at each transmitter/receiver is at most two, the problem of checking feasibility is polynomial time solvable ([156]).

Now let us turn our attention to the generic solvability of the interference alignment problem (8.130) and (8.131). In other words, we focus on the existence of a beamforming solution to the quadratic polynomial equations (8.130) and (8.131) when the channel matrices are randomly generated. To this end, it is natural to count the number of scalar equations and the number of scalar variables in the conditions (8.130) and (8.131). It is tempting to conjecture that there is an interference alignment solution if and only if the number of constraints is no larger than the number of variables (see [157]). Recently, Razaviyayn et al. [160] and Bresler et al. [161] have settled this conjecture completely in one direction, and partially in the other direction. They derive a general condition, described below, that must be satisfied by any DoF tuple (d_1, d_2, \dots, d_K) achievable through linear interference alignment.

Let us denote the polynomial equations in (8.131) by the index set

$$\mathcal{J} \triangleq \{(k, j) \mid 1 \leq k \neq j \leq K\}.$$

The following result ([160, 161]) provides an upper bound on the total achievable DoF when no channel extension is allowed. Consider a K -user flat fading MIMO interference channel where the channel matrices $\{\mathbf{H}_{ij}\}_{i,j=1}^K$ are generic (e.g., drawn from a continuous probability distribution). Assume no

channel extension is allowed. Then any tuple of degrees of freedom (d_1, d_2, \dots, d_K) that is achievable through linear interference alignment (8.130) and (8.131) must satisfy the following inequalities

$$\min\{M_k, N_k\} \geq d_k, \quad \forall k, \quad (8.133)$$

$$\max\{M_k, N_j\} \geq d_k + d_j, \quad \forall k, j, k \neq j, \quad (8.134)$$

$$\sum_{k:(k,j) \in \mathcal{I}} (M_k - d_k)d_k + \sum_{j:(k,j) \in \mathcal{I}} (N_j - d_j)d_j \geq \sum_{(k,j) \in \mathcal{I}} d_k d_j, \quad \forall \mathcal{I} \subseteq \mathcal{J}. \quad (8.135)$$

Roughly, the left hand side of (8.135) is equal to the number of independent scalar variables in (8.130) and (8.131) and the right hand side of (8.135) corresponds to the number of constraints in (8.130). Thus, the necessity of condition (8.135) for the existence of a feasible alignment scheme can be understood by counting the dimensions. However, a formal proof of this condition requires the use of field extension theory ([160]). We remark that condition (8.135) can be used to bound the total DoF achievable in a MIMO interference channel. In particular, the following upper bounds follow directly from condition (8.135).

- a. In the case of $d_k = d$ for all k , interference alignment is impossible unless

$$d \leq \frac{1}{K(K+1)} \sum_{k=1}^K (M_k + N_k).$$

- b. In the case of $M_k + N_k = M + N$, interference alignment requires

$$\left(\sum_{k=1}^K d_k \right)^2 + \sum_{k=1}^K d_k^2 \leq (M+N) \sum_{k=1}^K d_k,$$

which further implies

$$\sum_{k=1}^K d_k < (M+N).$$

The principal assumption enabling the surprising result of [21] is that the channel extensions are exponentially long in K^2 and are generic (e.g., drawn from a continuous probability distribution). If no channel extensions are allowed, part (b) above shows that the total achievable DoF in a MIMO interference channel is bounded by a constant $M + N - 1$, regardless of how many users are present in the system. While this bound is an improvement over the single user case which has a maximum DoF of $\min\{M, N\}$, it is significantly weaker than the maximum achievable total DoF of $K/2$ for a diagonal frequency selective (or time varying) interference channel with independent channel extensions. The latter grows linearly with the number of users in the system [21].

If channel extensions are restricted to have a polynomial length or are not generic, the total DoF for a MIMO interference channel is still largely unknown even for the Single-Input-Single-Output (SISO) interference channel. This is an interesting open problem. For the 3-user special case, Ref. [162] provided a characterization of the total achievable DoF as a function of the diversity.

Conversely, if all users have the same DoF d and the number of antennas M_k, N_k are divisible by d for each k , then condition (8.135) for each subsystem of (8.130) and (8.131) is also sufficient for the feasibility of interference alignment for generic choice of channel coefficients (e.g., drawn from a continuous probability distribution). If in addition, $M_k = M$ and $N_k = N$ for all k and M, N are divisible by d , then these results imply that interference alignment is achievable if and only if $(M+N) \geq d(K+1)$. Moreover, Bresler et al. [161] considered the symmetric case with $M_k = N_k = M, d_k = d$ for all k , and proved that the feasibility of interference alignment in this case is equivalent to $2M \geq d(K+1)$, regardless of the divisibility of M by d . When K is odd and $2M = d(K+1)$, then d divides M , so this result and Theorem 2 are in agreement. However, the case when K is even is not covered by Theorem 2.

To summarize, the initial work [21] is exciting and suggests that it may be possible to allocate resources in a MIMO IC based on DoF. However, the complexity and design of the interference alignment schemes have presented several challenges to the practicality of this approach for resource allocation.

- For a given channel realization, to determine whether a given set of DoF tuple is achievable is NP-hard (i.e., exponential effort is likely to be required for large number of users).
- Without channel extensions, the average DoF per user is shown to be at most $2M/(K+1)$, which is significantly smaller than $M/2$ when there are a large number of independent channel extensions (see [21]). Here M is the number of antennas at each transmitter and receive. Notice that the average per user DoF of $2M/(K+1)$ approximately doubles that of the orthogonal approaches (e.g. TDMA or FDMA).
- It requires too many channel extensions to reap the DoF benefit promised by Cadambe and Jafar [21].
- It requires full CSI, which can be difficult for large networks.
- It often requires selecting a set of feasible DoFs for the users a priori, which is difficult.

At this point, interference alignment appears most useful for a small system (e.g., 3–4 links) where a closed form interference alignment solution exists [21], and when using no or a small number of channel extensions. For a large network, direct maximization of the weighted sum-rate (or weighted sum utility maximization) seems to offer more potential for resource allocation and interference mitigation. For one thing, it requires the same amount of CSI, and yet can offer more sum-rate performance across all SNR regime than that of interference alignment. Moreover, it does not require selecting a DoF for each user in advance. As for future work, we suggest further investigation of the benefits of interference alignment for a small system with a few channel extensions.

Relevant Theory: Signal Processing Theory and Array Signal Processing

See Vol. 1, Chapter 3 Discrete-Time Signal and Systems

See Vol. 1, Chapter 12 Adaptive Filters

See Vol. 3, Chapter 19 Array Processing in the Face of Nonidealities

References

- [1] M. Cave, C. Doyle, W. Webb, *Essentials of Modern Spectrum Management*, Cambridge University Press, 2007.
- [2] A. Goldsmith, *Wireless Communications*, Combridge University Press, New York, 2005.

- [3] FCC, Report of the spectrum efficiency working group, 2002. <<http://www.fcc.gov/sptf/reports.html>>.
- [4] A. Sahai, M. Mishra, R. Tandra, K. Woyach, Cognitive radios for spectrum sharing, *IEEE Signal Process. Mag.* (2009) 140–146.
- [5] 3GPP, Evolved Universal Terrestrial Radio Access (EUTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); overall description, (2011) 3GPP TS 36.300, V8.9.0.
- [6] V. Chandrasekhar, J. Andrews, Femtocell networks: a survey. *IEEE Commun. Mag.* (2008) 59–67.
- [7] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, D. Malladi, A survey on 3GPP heterogeneous networks, *IEEE Wireless Commun.* 18 (2011) 10–21.
- [8] G. Foschini, K. Karakayali, R. Valenzuela, Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency, *IEE Proc. Commun.* 153 (2006) 548–555.
- [9] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, W. Yu, Multi-cell MIMO cooperative networks: a new look at interference, *IEEE J. Sel. Areas Commun.* 28 (2010) 1380–1408.
- [10] D. Gesbert, S. Kiani, A. Gjendemsj, ien, G., Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks, *Proc. IEEE* 95 (2007) 2393–2409.
- [11] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance*, Cambridge University Press, 2009.
- [12] D. Martín-Sacristán, J.F. Monserrat, J. Cabrejas-Peñauelas, D. Calabuig, S. Garrigas, N. Cardona, On the way towards fourth-generation mobile: 3GPP LTE and LTE-advanced, *EURASIP J. Wireless Commun. Networks* 4 (1–4) (2009) 10.
- [13] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, M. Tanno, Coordinated multipoint transmission/reception techniques for LTE-advanced, *IEEE Wireless Commun.* 17 (2010) 26–34.
- [14] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, second ed., Wiley, 2005.
- [15] T. Han, K. Kobayashi, A new achievable rate region for the interference channel, *IEEE Trans. Inform. Theory* 27 (1981) 49–60.
- [16] R. Etkin, D. Tse, H. Wang, Gaussian interference channel capacity to within one bit, *IEEE Trans. Inform. Theory* (2008).
- [17] A.B. Carleial, A case where interference does not reduce capacity, *IEEE Trans. Inform. Theory* 21 (1975) 569–570.
- [18] H. Sato, The capacity of the Gaussian interference channel under strong interference, *IEEE Trans. Inform. Theory* 27 (1981) 49–60.
- [19] S.T. Chung, J.M. Cioffi, The capacity region of frequency selective Gaussian interference channel under strong interference, *IEEE Trans. Commun.* 55 (2007) 1812–1821.
- [20] X. Shang, B. Chen, G. Kramer, H. Poor, Capacity regions and sum-rate capacities of vector Gaussian interference channels, *IEEE Trans. Inform. Theory* 56 (2010) 5030–5044.
- [21] V. Cadambe, S. Jafar, Interference alignment and degrees of freedom of the k -user interference channel, *IEEE Trans. Inform. Theory* 54 (2008) 3425–3441.
- [22] X. Shang, B. Chen, G. Kramer, H. Poor, Noisy-interference sum-rate capacity of parallel Gaussian interference channels, *IEEE Trans. Inform. Theory* 57 (2011) 210–226.
- [23] X. Shang, G. Kramer, B. Chen, A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels, *IEEE Trans. Inform. Theory* 55 (2009) 689–699.
- [24] J.Y.N. Hui, P.A. Humblet, The capacity region of the totally asynchronous multiple-access channel, *IEEE Trans. Inform. Theory* 31 (1985) 207–216.
- [25] M. Charafeddine, A. Paulraj, Maximum sum rates via analysis of 2-user interference channel achievable rates region, in: 43rd Annual Conference on Information Sciences and Systems, 2009, pp. 170–174.
- [26] M. Charafeddine, A. Sezgin, A. Paulraj, Rate region frontiers for n-user interference channel with interference as noise, in: Proceedings of Annual Allerton Conference on Communications, Control and Computing, 2007.

- [27] S. Hayashi, Z.-Q. Luo, Spectrum management for interference-limited multiuser communication systems, *IEEE Trans. Inform. Theory* 55 (2009) 1153–1175.
- [28] E.A. Jorswieck, E.G. Larsson, The MISO interference channel from a game-theoretic perspective: a combination of selfishness and altruism achieves pareto optimality, in: *IEEE ICASSP*, 2008, pp. 5364–5367.
- [29] E. Larsson, E. Jorswieck, Competition versus cooperation on the MISO interference channel, *IEEE J. Sel. Areas Commun.* 26 (2008) 1059–1069.
- [30] E.A. Jorswieck, E.G. Larsson, D. Danev, Complete characterization of the pareto boundary for the MISO interference channel, *IEEE Trans. Signal Process.* 56 (2008) 5292–5296.
- [31] J. Mo, J. Walrand, Fair end-to-end window-based congestion control, *IEEE/ACM Trans. Network.* 8 (2000) 556–567.
- [32] H. Boche, S. Naik, T. Alpcan, Characterization of convex and concave resource allocation problems in interference coupled wireless systems, *IEEE Trans. Signal Process.* 59 (2011) 2382–2394.
- [33] Y.-F. Liu, Y.-H. Dai, Z.-Q. Luo, Coordinated beamforming for MISO interference channel: complexity analysis and efficient algorithms, *IEEE Trans. Signal Process.* 59 (2011) 1142–1157.
- [34] Z.-Q. Luo, S. Zhang, Dynamic spectrum management: complexity and duality, *IEEE J. Sel. Top. Signal Process.* 2 (2008) 57–73.
- [35] M. Razaviyayn, M. Hong, Z.-Q. Luo, Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness, in: *2011 Asilomar Conference on Signals, Systems, and Computers*, 2011.
- [36] Z.K.M. Ho, D. Gesbert, Balancing egoism and altruism on interference channel: the MIMO case, in: *2010 IEEE International Conference on Communications (ICC)*, 2010, pp. 1–5.
- [37] M. Sadek, A. Tarighat, A. Sayed, A leakage-based precoding scheme for downlink multi-user mimo channels, *IEEE Trans. Wireless Commun.* 6 (2007) 1711–1721.
- [38] S. Shi, M. Schubert, H. Boche, Downlink mmse transceiver optimization for multiuser MIMO systems: duality and sum-mse minimization, *IEEE Trans. Signal Process.* 55 (2007) 5436–5446.
- [39] S. Ulukus, R. Yates, Iterative construction of optimum signature sequence sets in synchronous cdma systems, *IEEE Trans. Inform. Theory* 47 (2001) 1989–1998.
- [40] H. Boche, M. Schubert, A calculus for log-convex interference functions, *IEEE Trans. Inform. Theory* 54 (2008) 5469–5490.
- [41] H. Boche, M. Schubert, A unifying approach to interference modeling for wireless networks, *IEEE Trans. Signal Process.* 58 (2010) 3282–3297.
- [42] S. Stanczak, M. Wiczanowski, H. Boche, Distributed utility-based power control: objectives and algorithms, *IEEE Trans. Signal Process.* 55 (2007) 5058–5068.
- [43] M.R. Garey, D.S. Johnson, *Computers and Intractability: A guide to the Theory of NP-Completeness*, W.H. Freeman and Company, San Francisco, USA, 1979.
- [44] Y.-F. Liu, Y.-H. Dai, Z.-Q. Luo, Max-min fairness linear transceiver design for a multi-user MIMO interference channel, in: *Proceedings of the International Conference on Communicaitons 2011*, 2011.
- [45] M. Razaviyayn, M. Hong, Z.-Q. Luo, Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness, accepted by *Signal Processing*, special issue on sensor array processing, 2013.
- [46] G. Foschini, Z. Miljanic, A simple distributed autonomous power control algorithm and its convergence, *IEEE Trans. Veh. Technol.* 42 (1993) 641–646.
- [47] J. Zander, Distributed cochannel interference control in cellular radio systems, *IEEE Trans. Veh. Technol.* 41 (1992) 305–311.
- [48] J. Zander, Performance of optimal transmitter power control in cellular system, *IEEE Trans. Veh. Technol.* 41 (1992) 57–63.
- [49] S.V. Hanly, D. Tse, Power control and capacity of spread spectrum wireless networks, *Automatica* 35 (1999) 1987–2012.

- [50] M. Bengtsson, B. Ottersten, Optimal and Suboptimal Transmit Beamforming, *Handbook of Antennas in Wireless Communications*, CRC Press, 2001.
- [51] A. Wiesel, Y.C. Eldar, S. Shamai(Shitz), Linear precoding via conic optimization for fixed MIMO receivers, *IEEE Trans. Signal Process.* 54 (2006) 161–176.
- [52] Z.-Q. Luo, W. Yu, An introduction to convex optimization for communications and signal processing, *IEEE J. Sel. Areas Commun.* 24 (2006) 1426–1438.
- [53] Z.-Q. Luo, S. Zhang, Duality gap estimation and polynomial time approximation for optimal spectrum management, *IEEE Trans. Signal Process.* 57 (2009) 2675–2689.
- [54] W. Yu, R. Lui, Dual methods for nonconvex spectrum optimization of multicarrier systems, *IEEE Trans. Commun.* 54 (2006) 1310–1322.
- [55] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [56] J.W. Huang, R.A. Berry, M.L. Honig, Distributed interference compensation for wireless networks, *IEEE J. Sel. Area. Commun.* 24 (2006).
- [57] F. Wang, M. Krunz, S.G. Cui, Price-based spectrum management in cognitive radio networks, *IEEE J. Sel. Top. Signal Process.* 2 (2008).
- [58] Y. Wu, D.H.K. Tsang, 2008. Distributed multichannel power allocation algorithm for spectrum sharing cognitive radio networks, in: WCNC Proceedings, Las Vegas, USA, 2008.
- [59] W. Yu, Multiuser water-filling in the presence of crosstalk, in: Information Theory and Applications, Workshop 07, 2007.
- [60] W. Yu, G. Ginis, J.M. Cioffi, Distributed multiuser power control for digital subscriber lines, *IEEE J. Sel. Areas Commun.* 20 (2002) 1105–1115.
- [61] C. Shi, R.A. Berry, M.L. Honig, Monotonic convergence of distributed interference pricing in wireless networks, in: Proceedings of the 2009 IEEE international conference on Symposium on Information Theory, vol. 3, 2009, pp. 1619–1623.
- [62] D.A. Schmidt, A. Grundinger, W. Utschick, M.L. Honig, Distributed precoder optimization for interfering MISO channels, in: International ITG Workshop on Smart Antennas, 2008, pp. 285–290.
- [63] C. Shi, R.A. Berry, M.L. Honig, Distributed interference pricing with MISO channels, in: 46th Annual Allerton Conference on Communication, Control, and Computing, 2008, pp. 539–546.
- [64] C. Shi, D.A. Schmidt, R.A. Berry, M.L. Honig, W. Utschick, Distributed interference pricing for the MIMO interference channel, *IEEE Int. Conf. Commun.* (2009) 1–5.
- [65] S.-J. Kim, G. Giannakis, Optimal resource allocation for MIMO Ad Hoc cognitive radio networks, in: 46th Annual Allerton Conference on Communication, Control, and Computing, 2008, pp. 39–45.
- [66] L. Venturino, N. Prasad, X. Wang, Coordinated linear beamforming in downlink multicell wireless networks, *IEEE Trans. Wireless Commun.* 9 (2010) 1451–1461.
- [67] D. Schmidt, C. Shi, R. Berry, M. Honig, W. Utschick, Distributed resource allocation schemes, *IEEE Signal Process. Mag.* 26 (2009) 53–63.
- [68] J. Papandriopoulos, J.S. Evans, SCALE: a low-complexity distributed protocol for spectrum balancing in multiuser DSL networks, *IEEE Trans. Inform. Theory* 55 (2009).
- [69] P. Tsiaflakis, M. Diehl, M. Moonen, Distributed spectrum management algorithms for multiuser DSL networks, *IEEE Trans. Signal Process.* 56 (2008) 4825–4843.
- [70] S.S. Christensen, R. Agarwal, E.D. Carvalho, J.M. Cioffi, Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design, *IEEE Trans. Wireless Commun.* 7 (2008) 4792–4799.
- [71] D. Schmidt, C. Shi, R.A. Berry, M.L. Honig, W. Utschick, Minimum mean squared error interference alignment, in: Proceedings of Asilomar Conference, Pacific Grove, CA, 2009.
- [72] Q. Shi, M. Razaviyayn, Z.-Q. Luo, C. He, An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel, in: 2011 IEEE International Conference on Acoustics, Speech and, Signal Processing (ICASSP), 2011, pp. 3060–3063.

- [73] Q. Shi, M. Razaviyayn, Z.-Q. Luo, C. He, An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel, *IEEE Trans. Signal Process.* 59 (2011) 4331–4340.
- [74] W. Yu, T. Kwon, C. Shin, Multicell coordination via joint scheduling beamforming and power spectrum adaptation, in: Proceedings of IEEE International Conference on Computer Communication (INFOCOM), 2011.
- [75] H. Dahrouj, W. Yu, Coordinated beamforming for multicell multi-antenna wireless systems, *IEEE Trans. Wireless Commun.* 9 (2010) 1748–1759.
- [76] B. Song, R. Cruz, B. Rao, Network duality for multiuser MIMO beamforming networks and applications, *IEEE Trans. Commun.* 55 (2007) 618–630.
- [77] J. Kim, D. Cho, A joint power and subchannel allocation scheme maximizing system capacity in dense femtocell downlink systems, in: IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, 2009, pp. 1381–1385.
- [78] L. Venturino, N. Prasad, X. Wang, An improved iterative water-filling algorithm for multi-cell interference mitigation in downlink OFDMA networks, in: ACSSC, 2007, pp. 1718–1722.
- [79] L. Venturino, N. Prasad, X. Wang, Coordinated scheduling and power allocation in downlink multicell OFDMA networks, *IEEE Trans. Veh. Technol.* 58 (2009) 2835–2848.
- [80] M. Razaviyayn, H. Baligh, A. Callard, Z.-Q. Luo, Joint user grouping and transceiver design in a mimo interfering broadcast channel, submitted to IEEE Transactions on Signal Processing for publication.
- [81] E. Matskani, N. Sidiropoulos, Z.-Q. Luo, L. Tassiulas, Convex approximation techniques for joint multiuser downlink beamforming and admission control, *IEEE Trans. Wireless Commun.* 7 (2008) 2682–2693.
- [82] E. Matskani, N. Sidiropoulos, Z.-Q. Luo, L. Tassiulas, Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control, *IEEE Trans. Signal Process.* 57 (2009) 4882–4894.
- [83] S.V. Hanly, An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity, *IEEE J. Sel. Areas Commun.* 13 (1995) 1332–1340.
- [84] F. Rashid-Farrokhi, L. Tassiulas, K. Liu, Joint optimal power control and beamforming in wireless networks using antenna arrays, *IEEE Trans. Commun.* 46 (1998) 1313–1324.
- [85] R.D. Yates, C.Y. Huang, Integrated power control and base station assignment, *IEEE Trans. Veh. Technol.* 44 (1995) 1427–1432.
- [86] L. Gao, X. Wang, G. Sun, Y. Xu, A game approach for cell selection and resource allocation in heterogeneous wireless networks, in: The Proceeding of the SECON, 2011.
- [87] M. Hong, A. Garcia, J. Barrera, Joint distributed AP selection and power allocation in cognitive radio networks, in: The Proceedings of the IEEE INFOCOM, 2011.
- [88] S.M. Perlaza, E.V. Belmega, S. Lasaulce, M. Debbah, On the base station selection and base station sharing in self-configuring networks, in: Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools, 2009, pp. 71:1–71:10.
- [89] M. Hong, Z.-Q. Luo, Joint linear precoder optimization and base station selection for an uplink MIMO network: a game theoretic approach. in: Proceedings of the IEEE ICASSP, 2012.
- [90] M. Sanjabi, M. Razaviyayn, Z.-Q. Luo, Optimal joint base station assignment and downlink beamforming for heterogeneous networks, in: 2012 IEEE ICASSP, 2012.
- [91] S.V. Hanly, Information capacity of radio networks, Ph.D. Dissertation, Cambridge University, 1993.
- [92] H. Boche, S. Stanczak, Convexity of some feasible qos regions and asymptotic behavior of the minimum total power in cdma systems, *IEEE Trans. Commun.* 52 (2004) 2190–2197.
- [93] S. Stanczak, H. Boche, On the convexity of feasible qos regions, *IEEE Trans. Inform. Theory* 53 (2007) 779–783.
- [94] R.D. Yates, A framework for uplink power control in cellular radio systems, *IEEE J. Sel. Areas Commun.* 13 (1995) 1341–1347.

- [95] H. Boche, M. Schubert, A superlinearly and globally convergent algorithm for power control and resource allocation with general interference functions, *IEEE/ACM Trans. Network.* 16 (2008) 383–395.
- [96] Y.-F. Liu, Y.-H. Dai, Z.-Q. Luo, Joint power and admission control via linear programming deflation, in: *Proceedings of the ICASSP*, 2012.
- [97] I. Mitliagkas, N. Sidiropoulos, A. Swami, Joint power and admission control for ad-hoc and cognitive underlay networks: convex approximation and distributed implementation, *IEEE Trans. Wireless Commun.* 10 (2011) 4110–4121.
- [98] M. Ahmed, Call admission control in wireless networks: a comprehensive survey, *IEEE Commun. Survey Tut.* 7 (2005) 49–68.
- [99] M. Bengtsson, B. Ottersten, Optimal downlink beamforming using semidefinite optimization. in: *Proceedings of the 37th Annual Allerton Conference*, 1999.
- [100] M. Bengtsson, Jointly optimal downlink beamforming and base station assignment, in: *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*, vol. 5, 2001, pp. 2961–2964.
- [101] R. Stridh, M. Bengtsson, B. Ottersten, System evaluation of optimal downlink beamforming with congestion control in wireless communication, *IEEE Trans. Wireless Commun.* 5 (2006) 743–751.
- [102] M. Chiang, C.W. Tan, D.P. Palomar, D. O'Neill, D. Julian, Power control by geometric programming, *IEEE Trans. Wireless Commun.* 6 (2007) 2640–2651.
- [103] M. Wiczanowski, S. Stanczak, H. Boche, Providing quadratic convergence of decentralized power control in wireless networks—the method of min-max functions, *IEEE Trans. Signal Process.* 56 (2008) 4053–4068.
- [104] L. Qian, Y. Zhang, J. Huang, MAPEL: Achieving global optimality for a non-convex wireless power control problem, *IEEE Trans. Wireless Commun.* 8 (2009) 1553–1563.
- [105] J.B.G. Frenck, S. Schaible, Fractional Programming, *Handbook of Generalized Convexity and Generalized Monotonicity*, Springer, 2006, pp. 335–386.
- [106] N.T.H. Phuong, H. Tuy, A unified monotonic approach to generalized linear fractional programming, *J. Global Optim.* 26 (2003).
- [107] L. Qian, Y. Zhang, Monotonic optimization for non-concave power control in multiuser multicarrier network systems, in: *Proceedings of IEEE International Conference on Computer, Communication (INFOCOM)*, 2009, pp. 172–181.
- [108] C.W. Tan, M. Chiang, R. Srikant, Fast algorithms and performance bounds for sum rate maximization in wireless networks, in: *Proceedings of the IEEE International Conference on Computer Communication (INFOCOM)*, 2009, pp. 1350–1358.
- [109] C.W. Tan, S. Friedland, S.H. Low, Spectrum management in multiuser cognitive wireless networks: optimality and algorithm, *IEEE J. Sel. Areas Commun.* 29 (2011) 421–430.
- [110] Y. Xu, T. Le-Ngoc, S. Panigrahi, Global concave minimization for optimal spectrum balancing in multi-user DSL networks, *IEEE Trans. Signal Process.* 56 (2008) 2875–2885.
- [111] E.A. Jorswieck, E.G. Larsson, Monotonic optimization framework for the two-user MISO interference channel, *IEEE Trans. Commun.* 58 (2010) 2159–2168.
- [112] M.B. Shenouda, T. Davidson, Convex conic formulations of robust downlink precoder designs with quality of service constraints, *IEEE J. Sel. Top. Signal Process.* 1 (2007) 714–724.
- [113] N. Vucic, H. Boche, Robust QoS-Constrained optimization of downlink multiuser MISO systems, *IEEE Trans. Signal Process.* 57 (2009) 714–725.
- [114] E. Song, Q. Shi, M. Sanjabi, R. Sun, Z.-Q. Luo, Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight? in: *The Proceedings of IEEE International Conference on Acoustics, Speech, and, Signal Processing (ICASSP)*, 2011, pp. 3096–3099.
- [115] M.B. Shenouda, T. Davidson, On the design of linear transceivers for multiuser systems with channel uncertainty, *IEEE J. Sel. Areas Commun.* 26 (2008) 1015–1024.

- [116] A. Tajer, N. Prasad, X. Wang, Robust linear precoder design for multi-cell downlink transmission, *IEEE Trans. Signal Process.* 59 (2011) 235–251.
- [117] G. Zheng, K.-K. Wong, T.-S. Ng, Robust linear MIMO in the downlink: a worst-case optimization with ellipsoidal uncertainty regions, *EURASIP J. Adv. Signal Process.* 2008 (2008) 1–15.
- [118] T. Basar, G. Olsder, *Dynamic Noncooperative Game Theory*, SIAM, 1999.
- [119] M.J. Osborne, A. Rubinstein, *A Course in Game Theory*, MIT Press, 1994.
- [120] C.U. Sarayda, N.B. Mandayam, D.J. Goodman, Pricing and power control in a multicell wireless data network, *IEEE J. Sel. Areas Commun.* 19 (2001) 1883–1892.
- [121] C.U. Sarayda, N.B. Mandayam, D.J. Goodman, Efficient power control via pricing in wireless data network, *IEEE Trans. Commun.* 50 (2002) 291–303.
- [122] D. Goodman, N. Mandayam, Power control for wireless data, *IEEE Personal Commun.* 7 (2000) 48–54.
- [123] F. Meshkati, M. Chiang, H.V. Poor, S.C. Schwartz, A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems, *IEEE J. Sel. Areas Commun.* 24 (2006) 1115–1129.
- [124] F. Meshkati, H. Poor, S. Schwartz, Energy-efficient resource allocation in wireless networks, *IEEE Signal Process. Mag.* 24 (2007) 58–68.
- [125] G. Scutari, D.P. Palomar, S. Barbarossa, Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Part II: Algorithms, *IEEE Trans. Signal Process.* 56 (2008d).
- [126] Z.-Q. Luo, J.-S. Pang, Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines, *EURASIP J. Appl. Signal Process.* 2006 (2006) 1–10.
- [127] G. Scutari, D.P. Palomar, S. Barbarossa, Asynchronous iterative water-filling for Gaussian frequency-selective interference channels, *IEEE Trans. Inform. Theory* 54 (2008a).
- [128] R. Cendrillon, J. Huang, M. Chiang, M. Moonen, Autonomous spectrum balancing for digital subscriber lines, *IEEE Trans. Signal Process.* 55 (2007) 4241–4257.
- [129] G. Scutari, D.P. Palomar, S. Barbarossa, Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Part I: Nash equilibria, *IEEE Trans. Signal Process.* 56 (2008c).
- [130] K.W. Shum, K.K. Leung, C.W. Sung, Convergence of iterative waterfilling algorithm for Gaussian interference channels, *IEEE J. Sel. Areas Commun.* 25 (2007) 1091–1100.
- [131] G. Scutari, D.P. Palomar, S. Barbarossa, Competitive design of multiuser MIMO systems based on game theory: a unified view, *IEEE J. Sel. Areas Commun.* 26 (2008b).
- [132] A. Leshem, E. Zehavi, Game theory and the frequency selective interference channel, *IEEE Signal Process. Mag.* 26 (2009).
- [133] E. Telatar, Capacity of multi-antenna Gaussian channels, *Eur. Trans. Telecommun.* (1999) 585–596.
- [134] G. Arslan, M. Demirkol, Y. Song, Equilibrium efficiency improvement in MIMO interference systems: A decentralized stream control approach, *IEEE Trans. Wireless Commun.* 6 (2007) 2984–2993.
- [135] G. Scutari, D.P. Palomar, J.-S. Pang, F. Facchinei, Flexible design of cognitive radio wireless systems: from game theory to variational inequality theory, *IEEE Signal Process. Mag.* 26 (2009).
- [136] R.H. Gohary, Y. Huang, Z.-Q. Luo, J.-S. Pang, Generalized iterative water-filling algorithm for distributed power control in the presence of a jammer, *IEEE Trans. Signal Process.* 57 (2009) 2660–2674.
- [137] P. Setoodeh, S. Haykin, Robust transmit power control for cognitive radio, *Proc. IEEE* (2009) 915–939.
- [138] M. Hong, A. Garcia, Averaged iterative water-filling algorithm: robustness and convergence, *IEEE Trans. Signal Process.* 59 (2011) 2448–2454.
- [139] R.H. Gohary, T.J. Willink, Robust IWFA for open-spectrum communications, *IEEE Trans. Signal Process.* 57 (2009) 4964–4970.
- [140] M. Hong, A. Garcia, Equilibrium pricing of interference in cognitive radio networks, *IEEE Trans. Signal Process.* 59 (2011) 6058–6072.
- [141] J.-S. Pang, G. Scutari, D. Palomar, F. Facchinei, Design of cognitive radio systems under temperature-interference constraints: a variational inequality approach, *IEEE Trans. Signal Process.* 58 (2010) 3251–3271.

- [142] J. Wang, G. Scutari, D. Palomar, Robust MIMO cognitive radio via game theory, *IEEE Trans. Signal Process.* 59 (2011) 1183–1201.
- [143] Y. Xie, B. Armbruster, Y. Ye, Dynamic spectrum management with the competitive market model, *IEEE Trans. Signal Process.* 58 (2010) 2442–2446.
- [144] FCC, In the matter of establishment of an interference temperature metric to quantify and management interference and to expand available unlicensed operation in certain fixed, mobile and satellite frequency band, 2003, ET Docket No. 03-237.
- [145] R. Zhang, Y.-C. Liang, S. Cui, Dynamic resource allocation in cognitive radio network, *IEEE Signal Process. Mag.* (2010) 102–114.
- [146] G. Scutari, D.P. Palomar, F. Facchinei, J.-S. Pang, Convex optimization, game theory, and variational inequality theory, *IEEE Signal Process. Mag.* 27 (2010) 35–49.
- [147] G. Scutari, D. Palomar, MIMO cognitive radio: a game theoretical approach, *IEEE Trans. Signal Process.* 58 (2010) 761–780.
- [148] J.-S. Pang, G. Scutari, F. Facchinei, C. Wang, Distributed power allocation with rate constraints in Gaussian parallel interference channels, *IEEE Trans. Inform. Theory* 54 (2008) 3471–3489.
- [149] Y. Wu, D.H.K. Tsang, Distributed power allocation algorithm for spectrum sharing cognitive radio networks with QOS guarantee, in: Proceedings of IEEE International Conference on Computer Communication (INFOCOM), 2009.
- [150] W. Yu, W. Rhee, S. Boyd, J.M. Cioffi, Iterative water-filling for Gaussian vector multiple-access channels, *IEEE Trans. Inform. Theory* 50 (2004) 145–152.
- [151] S.A. Jafar, M. Fakhereddin, Degrees of freedom for the MIMO interference channel, *IEEE Trans. Inform. Theory* (2007).
- [152] Y. Birk, T. Kol, Informed-source coding-on-demand (ISCOD) over broadcast channels, in: Proceedings of the IEEE International Conference on Computer Communications, San Francisco, CA, 1998, pp. 1257–264.
- [153] S. Jafar, Degrees of freedom on the MIMO x channel-optimality of the MMK scheme. Available from: <[Arxiv:cs.IT/0607099v2, 2006](https://arxiv.org/abs/cs.IT/0607099v2)>.
- [154] M. Maddah-Ali, A. Motahari, A. Khandani, Communication over MIMO x channels: Interference alignment, decomposition, and performance analysis, *IEEE Trans. Inform. Theory* (2008).
- [155] S. Jafar, S. Shamai, Degrees of freedom region for the MIMO X channel, *IEEE Trans. Inform. Theory* (2008) 151–170.
- [156] M. Razaviyayn, M. Boroujeni, Z.-Q. Luo, Linear transceiver design for interference alignment: complexity and computation, *IEEE Trans. Inform. Theory* 58 (2012) 2896–2910.
- [157] C.M. Yetis, T. Gou, S.A. Jafar, A.H. Kayran, On feasibility of interference alignment in MIMO interference networks, *IEEE Trans. Signal Process.* 2010.
- [158] K. Gomadam, V.R. Cadambe, S.A. Jafar, Approaching the capacity of wireless networks through distributed interference alignment, in: Proceedings of the 2008 IEEE GLOBECOM, 2008.
- [159] S. Peters, R. Heath, Interference alignment via alternating minimization, in: 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001, Proceedings (ICASSP '09), 2009, pp. 2445–2448.
- [160] M. Razaviyayn, G. Lyubeznik, Z.-Q. Luo, On the degrees of freedom achievable through interference alignment in a MIMO interference channel, *IEEE Trans. Signal Process.* (in press).
- [161] G. Bresler, D. Cartwright, D. Tse, Settling the feasibility of interference alignment for the MIMO interference channel: the symmetric case, 2011. Available from: <[Arxiv:1104.0888v1](https://arxiv.org/abs/1104.0888v1)>.
- [162] G. Bresler, D. Tse, Degrees-of-freedom for the 3-user gaussian interference channel as a function of channel diversity, in: Proceedings of the Allerton Conference on Communication, Control, and Computing, Allerton, IL, 2009.

9

Advances in Spectrum Sensing and Cross-Layer Design for Cognitive Radio Networks¹

Seung-Jun Kim*, Emiliano Dall'Anese*, Juan Andrés Bazerque*,
Ketan Rajawat†, and Georgios B. Giannakis*

*Department of ECE, University of Minnesota, Minneapolis, MN, USA

†Department of Electrical Engineering, Indian Institute of Technology, Kanpur, India

Nomenclature

$\Phi(\mathbf{x}, f)$	power spectral density (PSD) at position \mathbf{x} and frequency f
$\varphi(\mathbf{x}, f)$	measured PSD at position \mathbf{x} and frequency f
$\hat{\Phi}(\mathbf{x}, f)$	estimated PSD at position \mathbf{x} and frequency f
$\Phi_s(f)$	transmit-PSD of source s at frequency f
$g_{\mathbf{x}_s \rightarrow \mathbf{x}_r}(t, f)$	gain of the channel from position \mathbf{x}_s to position \mathbf{x}_r at time t and frequency f
$\boldsymbol{\theta}$	vector of basis expansion coefficients for the PSD map
\mathbf{x}_r	position of the r th CR
N_r	number of CRs
N_g	number of grid points
N_b	number of known bases
$\boldsymbol{\theta}_g$	the g th group of basis expansion coefficients
N	number of samples
$s_{\mathbf{x} \rightarrow \mathbf{x}_s}$	shadow fading of the channel from position \mathbf{x} to position \mathbf{x}_s
$\ell(\mathbf{x}, t)$	spatial loss field at position \mathbf{x} and time t

2.09.1 Introduction

The cognitive radio (CR) paradigm endeavors to mitigate the scarcity of spectral resources for wireless communication through intelligent sensing and agile resource allocation techniques [1,2]. The motivating reason is that although most of the available spectrum has been licensed to primary users (PUs) for exclusive usage, it is often significantly underutilized depending on the time and the location that communication takes place [3]. The CRs aim to learn the RF landscape, and identify the unused spectral resources—often called “white space” or “spectrum holes”—in the time, frequency, and space domains

¹This work was supported by QNRF grant NPRP 09-341-2-128.

through spectrum sensing. Based on the information obtained, judicious resource management is then performed to communicate opportunistically without causing harmful interference to the licensed PU systems.

The sensing task can be as basic as detecting the presence of PU signals in a given band at a given time. It can become as sophisticated as estimating the channel gains, transmit-powers, modulation classes, and PU locations, as well as learning their traffic patterns. Deciding the PU presence is necessary for the *spectrum overlay* scenario [4], in which the CRs identify completely unoccupied bands to transmit on. Channel gains between CR transmitters and PU receivers are useful for interference control in the *spectrum underlay* scenario, in which the CRs are allowed to share bands with the PUs, provided that the interference experienced by the PUs is maintained below an acceptable level. The bottom line is that the richer the information collected on the PU systems and the surrounding RF environment, the more adaptable the CR operation can become via dynamic resource optimization.

Spectrum sensing is no easy task. The CRs often need to scan a huge swath of bandwidth in order to identify spectrum holes [5]. In the prevalent case of no cooperation occurring between PU and CR systems, the PU signals must be detected in a low signal-to-noise power ratio (SNR) regime. The lack of dedicated training signals may render it difficult for the CR systems to acquire the channels accurately. In fact, the CRs might not even have prior knowledge on the PU signal characteristics, often limiting the options to using simple energy detectors (radiometers) [6]. Thus, obtaining a decent detection performance can become quite challenging [7]. The difficulty is only aggravated with the hidden terminal issues and strenuous propagation environments which may include fading and shadowing [8].

These formidable challenges have invited intensive research in this area. At the physical layer, various *cooperative* sensing schemes have been developed to cope with the hidden terminal problems and combat fading through diversity combining of the samples taken by multiple CRs [9,10]. Cyclostationarity detectors have been developed for improved sensing performance [11,12]. Various signal processing and learning tools have been employed to effectively capture the RF environment in which the CR network is deployed [13–15]. To minimize sensing delay while meeting the detection performance targets, sequential detectors have been investigated [16,17]. Both parallel scanning of multiple bands as well as serial search have been considered [18,19]. In the case of serial search, selecting the bands to sense and coordinating the search along with the access among multiple CRs have attracted much research toward designing the MAC layer tailored for CR sensing [20–22].

It is largely part of ongoing research how to effectively permeate the benefits of enhanced cognition capability of the CRs to cross-layer network design and adaptation. Needless to say, this is of critical importance for efficient, reliable, quality-of-service-assuring operation of CR networks in the presence of dynamics and uncertainties in the CR deployment [23]. It has been recognized that the sensing algorithms must be designed with the cross-layer interaction in mind [24,25]. This tutorial paper aims to survey some of the advances made in CR spectrum sensing, and also highlight the intertwined cross-layer resource management issues, hopefully providing with fruitful directions for future research.

The organization of the rest of the paper is as follows. An overview of the physical layer sensing techniques are provided in Section 2.09.2 with emphasis on the recent RF cartography approaches. The MAC layer issues of sensing is briefly reviewed in Section 2.09.3. The implications of sensing to cross-layer design are illustrated in Section 2.09.4. Some conclusions are offered in Section 2.09.5.

2.09.2 Sensing at the physical layer

Signal detection is the core element of CR sensing, with Neyman-Pearson (NP) hypothesis testing offering the natural and most widely used framework. With the objective of deciding whether a PU is present or white space is available, the CR acquires samples of the ambient RF signal $x(t)$ and decides between two hypotheses: $H_0 : x(t) = n(t)$ versus $H_1 : x(t) = s(t) + n(t)$, where $n(t)$ denotes ambient noise and $s(t)$ a PU signal possibly affected by multipath and shadow fading effects. The *energy detection* is widely used because it is simple and does not require knowledge about the PU system parameters [6]. The test statistic (TS) $S_T = \sum_{t=1}^T x^2(t)$ is modeled as χ^2 -distributed assuming that the noise variance is known. In order to avoid self interference, multiple CRs competing for the same bands must schedule quiet periods, that is, time intervals where CRs suspend transmission and perform detection [26].

Modulated signals are typically cyclostationary processes, that is, their correlation function $r(t, \tau) := \mathbb{E}\{x(t)x^*(t - \tau)\}$ is periodic in t , which implies that its Fourier spectrum peaks at the cyclic frequency corresponding to its period. *Cyclostationarity detection* capitalizes on this property to detect a PU signal even in low SNR, since white noise yields zero correlation at nonzero lags regardless of its power level [12]. Furthermore, cyclostationarity detectors can separate signals with different cyclic frequencies, thus potentially not requiring quiet periods. These advantages come at the price of requiring larger data records to attain comparable performance relative to energy detection, since cyclostationarity detection entails sample estimates of the fourth-order moments.

Standard NP tests predefine the number of samples to acquire as a function of prescribed test performance, i.e., the probabilities of detection and false alarms. *Sequential alternatives* are available, in which the TS is updated sample by sample [27]. Depending on how informative this TS is, three actions are possible, namely, rejecting the null or the alternative hypotheses, or acquiring an additional sample. For a specified test performance, the technique is proven to reduce the number of required samples in average, although it may exceed the NP sample size on a bad realization.

In wideband sensing, the CRs often need to scan different bands in search for white space. These bands can be tested independently by applying single-band detectors separately. In *multi-band testing*, decision thresholds corresponding to different bands can be optimized jointly [19]. Accordingly, the test performance is not prescribed but designed to maximize the aggregate CR throughput across frequency bands. Increasing the thresholds increases the chance of CR transmissions, increasing the throughput. It also increases the probability of miss detection. This probability is controlled by setting a price for interfering with PUs, and prescribing an upper bound to the aggregate cost across bands that the CR can afford [19]. A compressive sampling approach has also been proposed for inspecting all bands together at sub-Nyquist sampling rates [28].

Collaboration among CRs adds spatial diversity to the sensing methods which is crucial for improved detection of white spaces, since shadowing effects may lead a single CR to miss detection [8]. In particular, the hidden terminal problem can be mitigated by collaboration. Such a situation arises when a CR is not in the range of a PU transmitter, but a PU receiver in-between falls in the range of the CR. In this setup, the CR will miss the PU transmitter, infer white space, and start communicating, thus causing harmful interference to the PU receiver. In addition to gaining spatial diversity, collaborative CRs can share sensing resources with the potential to reduce sensing time or to improve detection performance [29].

Cooperation protocols must be judiciously designed so that the overhead introduced for collaboration does not outweigh the increase in the throughput of the opportunistic access [10]. Centralized CR networks are often considered, with CRs communicating their TSs to a fusion center (FC), where these are combined to yield a fused decision. Combining unquantized (i.e., soft) TSs serves as a guideline for the design of optimal protocols [30,31]. However, quantization is important for striking a balance in the data transmission-sensing trade-off. To this end, schemes combining quantized TSs have been proposed, even with one-bit resolution [32].

Distributed processing offers an alternative to the FC-based cooperation. Information is shared among neighboring CRs only, a preferable architecture for large networks since long-range communication to an FC consumes excessive power and interferes with reports from all other CRs in the network. In addition, *decentralized* networks are more robust since their operation is not dependent on a single point of failure, and it is more flexible since an entering CR only needs to discover its neighbors to start operating without competing for access to the FC. The information shared in the neighborhood eventually percolates across the entire network after a number of local communication steps, and approaches the optimality of a centralized CR network [14].

Optimality here should be understood as achieving the prescribed test performance with minimal number of samples, or maximizing the throughput. These criteria are “socially optimal” for the CR network. *Game theoretic* approaches have also been proposed in which a single CR decides whether or not to sense and collaborate in order to maximize its own throughput [33]. A CR might decide not to sense and use its resources to transmit on available bands according to the sensing results of its peer CRs. This strategy is not optimal if followed by all CRs as no white spaces are revealed, and then individual CRs are compelled to sense the spectrum. A strategy is developed in [33] to achieve a stable equilibrium point in which CRs balance sensing and transmission times.

Censoring offers a complementary approach to quantization when it comes to reducing the overhead. Since communicating local TSs consumes energy and bandwidth, the CRs are allowed to do so only when S_T is sufficiently informative. A two-threshold approach was proposed in [34], where the CRs transmit S_T only when it is below the lower threshold, or above the higher one. The thresholds are designed to guarantee the prescribed test performance, considering the cost of sensing and TS transmission [34].

These multiple dimensions of the CR sensing problem are summarized in Figure 9.1, and overviewed more extensively in the tutorial reference [35]. Recently, *spectrum cartography* opened a new dimension for CR sensing by analyzing the spectral opportunities not only in frequency but also in space [13–15]. The underlying departure of these approaches relative to the conventional sensing algorithms is that the PU’s band occupancy is no longer deemed the same regardless of location. Rather, spectrum cartography sets the objective of revealing a map of the spectrum occupancy across space per frequency band. Estimating spectrum maps is particularly relevant for wide-area ad hoc networks, where PU transmissions typically reach only a small subset of CRs. Knowing the spectrum across space allows remote CRs to reuse idle bands. It also enables the CRs to adapt their transmit-powers or positions to minimally interfere with PUs [36]. Of course, more sophisticated sensing approaches demand higher implementation complexity. While a viable alternative may be to formally allow the CRs and the PUs collaborate, it is noted that the CR paradigm advocated here does not need to be confined to the CR networking per se, but may be instrumental to other complex wireless networking scenarios, where distributed and autonomous cognition of the environment is useful [37]. The cartography approach is described in more detail in the ensuing section.

Energy	Cyclostationarity
Single-band	Multi-band
Fixed-sample size	Sequential
FC-based	Decentralized
Socially optimal	Game-theoretic
Soft-combined	Hard-combined
Censored	Uncensored
Space-invariant	Cartography

FIGURE 9.1

Aspects of collaborative CR sensing at the physical layer.

2.09.2.1 Sparsity-aware power spectrum cartography

A parametric approach to cooperative RF cartography is developed in [14] with the goal of mapping out the ambient RF power levels in the geographical area $\mathcal{A} \subseteq \mathbb{R}^2$ of interest. The setup includes N_s sources (PU transmitters) located at position vectors $\{\mathbf{x}_s \in \mathcal{A}\}_{s=1}^{N_s}$, and N_r receivers (CRs) at locations $\{\mathbf{x}_r \in \mathcal{A}\}_{r=1}^{N_r}$. The transmit-power spectral density (PSD) of a source signal at position \mathbf{x}_s is represented by a basis expansion as $\Phi_s(f) = \sum_{v=1}^{N_b} \theta_{sv} b_v(f)$, where $\{b_v(f)\}_{v=1}^{N_b}$ is a collection of known bases, and $\{\theta_{sv}\}_{v=1}^{N_b}$ denotes the expansion coefficients to be estimated per source s . Assuming spatial uncorrelatedness of channels and sources, the PSD $\Phi(\mathbf{x}_r, f)$ measured at receiving CRs at $\{\mathbf{x}_r\}_{r=1}^{N_r}$ in the presence of white noise with respective variances $\{\sigma_r^2\}_{r=1}^{N_r}$ is expressed as

$$\Phi(\mathbf{x}_r, f) = \sum_{s=1}^{N_s} g_{\mathbf{x}_s \rightarrow \mathbf{x}_r} \Phi_s(f) + \sigma_r^2 = \sum_{s=1}^{N_s} g_{\mathbf{x}_s \rightarrow \mathbf{x}_r} \sum_{v=1}^{N_b} \theta_{sv} b_v(f) + \sigma_r^2 = \mathbf{b}_r^T(f) \boldsymbol{\theta} + \sigma_r^2, \quad (9.1)$$

where T denotes transposition; $g_{\mathbf{x}_s \rightarrow \mathbf{x}_r}$ the average channel gain of the $\mathbf{x}_s \rightarrow \mathbf{x}_r$ link; $\boldsymbol{\theta}$ the vector formed by stacking the columns of the matrix with entries θ_{sv} ; and $\mathbf{b}_r(f)$ the vector constructed by concatenating the columns of the matrix with (s, v) -entry $g_{\mathbf{x}_s \rightarrow \mathbf{x}_r} b_v(f)$. One possible choice for the channel gains is to adopt an inverse polynomial path loss model so that $g_{\mathbf{x}_s \rightarrow \mathbf{x}_r}$ is only a function of the distance between \mathbf{x}_r and \mathbf{x}_s . Alternatively, estimates can be obtained from channel gain cartography, which will be explained in Section 2.09.2.2.

Given PSD measurements $\{\varphi(\mathbf{x}_r, f_n)\}_{n=0}^{N-1}$ at N discrete frequencies per CR r , the goal is to estimate the *PSD maps* $\hat{\Phi}(\mathbf{x}, f) \forall \mathbf{x} \in \mathcal{A}$, one per frequency f . From the linear model (9.1), this is possible provided an estimate of $\boldsymbol{\theta}$ becomes available.

2.09.2.1.1 Compressed sensing approach

Let φ_r denote the $N \times 1$ vector with the n th entry $\varphi_{rn} := \varphi(\mathbf{x}_r, f_n)$ representing the PSD measurement obtained, e.g., by sample-average on periodograms. Defining the estimation error vector \mathbf{e}_r likewise, one arrives at the local vector-matrix model per CR r

$$\varphi_r = \mathbf{B}_r \boldsymbol{\theta} + \mathbf{e}_r, \quad r = 1, 2, \dots, N_r, \quad (9.2)$$

where matrix \mathbf{B}_r is formed to have rows $\{\mathbf{b}_r^T(f_n)\}_{n=1}^N$, and the noise variance is absorbed in φ_r or $\boldsymbol{\theta}$ without loss of generality.

Cooperative PSD sensing is possible because $\boldsymbol{\theta}$ is common to all N_r receiving CRs. This allows estimation of $\boldsymbol{\theta}$ in the linear regression model (9.2) using the nonnegative (NN), and thus non-linear, LS criterion

$$\min_{\boldsymbol{\theta} \geq \mathbf{0}} \sum_{r=1}^{N_r} \|\varphi_r - \mathbf{B}_r \boldsymbol{\theta}\|^2, \quad (9.3)$$

where the nonnegativity constraints are naturally imposed to prevent negative PSD estimates.

With position vectors \mathbf{x}_s (and/or \mathbf{x}_r) unknown, even the model in (9.1) is nonlinear and the NN-LS optimization in (9.3) is rendered non-convex with multiple local minima. To bypass this challenge, the idea in [14] relies on a virtual grid of N_g candidate source locations as the one depicted in Figure 9.2. Vectors \mathbf{x}_s are replaced by \mathbf{x}_g in Figure 9.2, which no longer describe the actual positions of the PUs, but rather the grid points with *known* spatial coordinates where the PUs could be present. This virtual grid model removes the model nonlinearity, while rendering (9.3) convex at the price of the increased number of unknowns. Aided by the virtual grid, one can in principle obtain the PU locations that best fit the measurements via exhaustive search, provided that the number of PUs is known. (The number of PUs may be determined from model complexity, e.g., using Akaike's information-theoretic criterion.)

The exhaustive search is clearly undesirable because it incurs combinatorial complexity that grows rapidly in the number of grid points. In fact, a dense grid may be preferred to achieve higher precision in

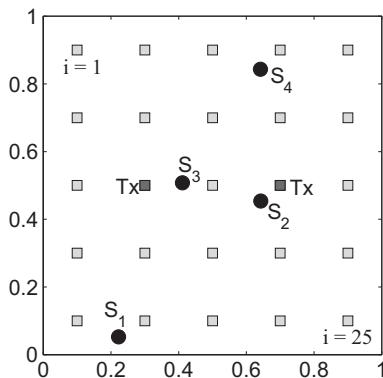


FIGURE 9.2

Virtual CR network grid.

localizing the PUs. Recent advances in compressive sampling can mitigate this hurdle by exploiting the sparsity present in θ [38–40]. Sparsity is manifested because individual transmissions typically occupy only a small fraction of the possibly huge system bandwidth. Moreover, active PUs are present only at a small fraction of candidate locations $\{\mathbf{x}_g\}$.

In particular, the least-absolute shrinkage and selection operator (Lasso) [40], also known as denoising basis pursuit [39], amounts to augmenting (9.3) with the ℓ_1 norm $\|\theta\|_1 := \sum_{s=1}^{N_g} \sum_{v=1}^{N_b} |\theta_{sv}|$ weighted by a sparsity-tuning parameter λ_1 . After incorporating PSD-imposed nonnegativity constraints, the Lasso for PSD map estimation amounts to solving the following convex optimization problem

$$\hat{\theta} = \arg \min_{\theta \geq 0} \sum_{r=1}^{N_r} \|\varphi_r - \mathbf{B}_r \theta\|_2^2 + \lambda_1 \|\theta\|_1. \quad (9.4)$$

Clearly, the larger λ_1 is chosen, the more entries of $\hat{\theta}$ will be shrunk to zero. The remaining non-zero entries of $\hat{\theta}$ yield the positions and power of active transmitters, the bands used for transmission, and eventually the entire PSD map $\hat{\Phi}(\mathbf{x}, f) \forall f$ and $\forall \mathbf{x} \in \mathcal{A}$ [cf. (9.1)].

2.09.2.1.2 Sparsity at group and coefficient levels

A refinement of the estimator in (9.4) is obtained by observing the hierarchical sparsity present in θ : when a point \mathbf{x}_g is unoccupied by a PU transmitter, the entire basis expansion coefficients $\theta_g := [\theta_{g1}, \dots, \theta_{gN_b}]^T$ corresponding to this point would become zero. On this account, the regularizer in (9.4) is augmented by adding the term $\lambda_2 \sum_{g=1}^{N_g} \|\theta_g\|_2$, which encourages sparsity at the group level [41, 42]. The ℓ_1 penalty is preserved in order to promote sparsity at the single coefficient level in the surviving θ_g , accounting for the sparsity in the frequency domain, yielding

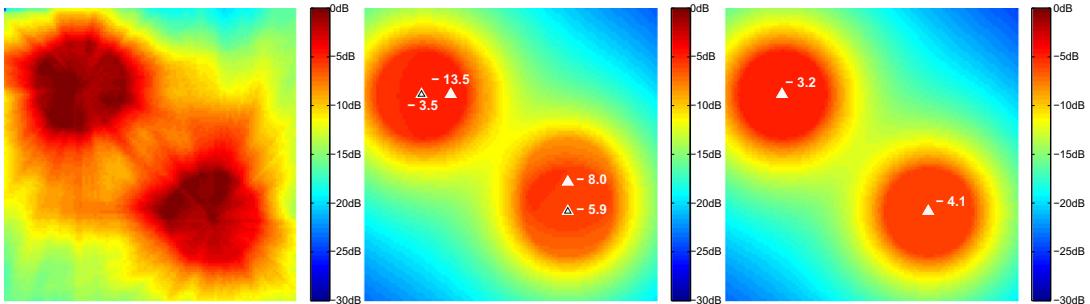
$$\hat{\theta} = \arg \min_{\theta \geq 0} \frac{1}{2} \sum_{r=1}^{N_r} \|\varphi_r - \mathbf{B}_r \theta\|_2^2 + \lambda_1 \|\theta\|_1 + \lambda_2 \sum_{g=1}^{N_g} \|\theta_g\|_2. \quad (9.5)$$

The group penalty encourages sparsity at the group level, either by shrinking to zero all variables within a group, or by retaining them altogether. As λ_2 is increased, more group estimates θ_g become zero.

2.09.2.1.3 Uncertainty on the propagation model

Further enhancement was introduced in [41] to achieve robustness against perturbations in matrices \mathbf{B}_r . Uncertainty in these matrices is due to: (i) errors in the estimates of $\{g_{\mathbf{x}_s \rightarrow \mathbf{x}_r}\}$; (ii) position offsets when PUs are located between grid points; and (iii) approximation errors in basis expansion. The resultant model mismatches are captured by an additive error matrix \mathbf{E} , yielding a perturbed model $\varphi = (\mathbf{B} + \mathbf{E})\theta$, where φ is obtained by stacking vectors φ_r , and \mathbf{B} by stacking matrices \mathbf{B}_r . Under this new model, it is pertinent to adopt a total LS formulation [41, 43]

$$\{\hat{\theta}, \hat{\mathbf{E}}\} = \arg \min_{\mathbf{E}, \theta \geq 0} \frac{1}{2} \|\varphi - (\mathbf{B} + \mathbf{E})\theta\|_2^2 + \frac{1}{2} \|\mathbf{E}\|_F^2 + \lambda_1 \|\theta\|_1 + \lambda_2 \sum_{g=1}^{N_g} \|\theta_g\|_2. \quad (9.6)$$

**FIGURE 9.3**

(Left) PSD map generated by two sources in 6 dB of log-normal shadowing aggregated across frequency; (Center) estimate obtained from (9.5) with $N_r = 50$ and $N_g = 100$; (Right) robust estimate (9.6).

Figure 9.3 shows how (9.5) is capable of recovering the PSD maps across space under log-normal shadowing, and how its robust version (9.6) is more effective on resolving sources located off the grid points. Estimator (9.6) can still be limited by fading, which is not accounted for in the model. A systematic method for identifying and rejecting such outliers can also be found in [41].

2.09.2.1.4 Nonparametric basis pursuit

A nonparametric basis pursuit method was applied to spectrum cartography in [44] and shown to offer a valuable augmentation of the parametric path loss model considered so far. The basis expansion model for the PSD maps is recast as

$$\Phi(\mathbf{x}, f) = \sum_{v=1}^{N_b} g_v(\mathbf{x}) b_v(f), \quad (9.7)$$

where $g_v(\mathbf{x})$ collects the aggregate power from all sources at receiving point \mathbf{x} . Without prescribing the functional form of $g_v(\mathbf{x})$ a priori, these functions are interpolated from the available data.

To capture the smooth variation of $\Phi(\mathbf{x}, f)$, the criterion for selecting $g_v(\mathbf{x})$ is regularized using the so-termed thin-plate penalty [45, p. 30]. Functions $\{g_v\}_{v=1}^{N_b}$ are estimated as

$$\{\hat{g}_v\}_{v=1}^{N_b} := \arg \min_{\{g_v \in \mathcal{S}\}} \frac{1}{N_r N} \sum_{r=1}^{N_r} \sum_{n=1}^N \left(\varphi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2 + \lambda_s \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} \|\nabla^2 g_v(\mathbf{x})\|_F^2 d\mathbf{x}, \quad (9.8)$$

where $\|\nabla^2 g_v\|_F$ denotes the Frobenius norm of the Hessian of g_v , and \mathcal{S} the space of Sobolev functions, for which the penalty is well defined [46]. The parameter $\lambda_s \geq 0$ controls the degree of smoothing. Specifically, for $\lambda_s = 0$ the estimates in (9.8) correspond to *rough* functions interpolating the data; while as $\lambda_s \rightarrow \infty$ the estimates yield linear functions (i.e., $\nabla^2 \hat{g}_v(\mathbf{x}) \equiv \mathbf{0}_{2 \times 2}$). A smoothing parameter in-between these limiting values is selected using, e.g., the leave-one-out cross-validation approach; see e.g., [47].

The optimization problem (9.8) is variational in nature, and in principle requires searching over the infinite-dimensional function space \mathcal{S} . Fortunately, it turns out that (9.8) admits closed-form, finite-dimensional minimizers

$$\hat{g}_v(\mathbf{x}) = \sum_{r=1}^{N_r} \beta_{vr} K(\|\mathbf{x} - \mathbf{x}_r\|_2) + \boldsymbol{\alpha}_{v1}^T \mathbf{x} + \alpha_{v0}, \quad v = 1, \dots, N_b, \quad (9.9)$$

where $K(\rho) := \rho^2 \log(\rho)$, and $\boldsymbol{\beta}_v := [\beta_{v1}, \dots, \beta_{vN_r}]^T$ satisfies $\sum_{r=1}^{N_r} \beta_{vr} = 0$, and $\sum_{r=1}^{N_r} \beta_{vr} \mathbf{x}_r = \mathbf{0}$. Optimal coefficients $\mathbf{c}_v^* := [\beta_{v1}^*, \dots, \beta_{vN_b}^*, \boldsymbol{\alpha}_{v1}^{*T}, \alpha_{v0}^*]$ can be found by substituting (9.9) back into (9.8) and solving it.

2.09.2.1.5 Group-Lasso on splines

An improved spline-based PSD estimator can be obtained by exploiting group sparsity to fit unknown spatial functions $\{g_v\}_{v=1}^{N_b}$ using (9.7) with $N_b \gg N_r, N$, possibly with an overcomplete set of bases $\{b_v\}_{v=1}^{N_b}$. The resultant model is particularly attractive when there is an inherent uncertainty on the PU transmission parameters, such as the center frequency or the pulse shape including the roll-off factor. Adaptive communication schemes frequently adjust such parameters [48, Ch. 9]. A sizable collection of bases can effectively accommodate most of the possible cases, providing model robustness. Thus, known bases are selected to describe frequency characteristics of the PSD map, while a variational approach is employed to capture spatial dependencies.

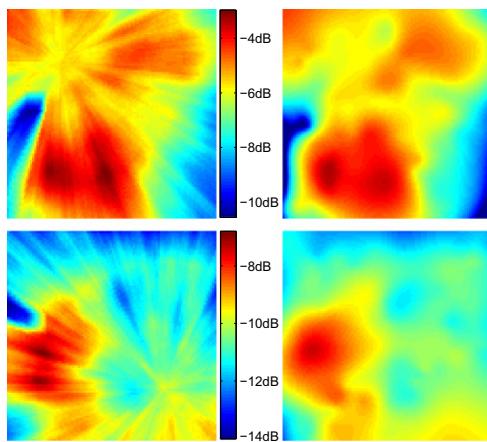
In this context, the envisioned estimation method provides the CRs with capability of selecting a few bases that better “explain” the actual transmitted signals. As a result, most functions g_v are expected to be identically zero; hence, there is an inherent form of sparsity present that can be exploited to improve estimation. A major departure from the conventional basis pursuit [39] is that (9.7) entails bases weighted by functions $\{g_v\}$ rather than scalars.

The proposed nonparametric basis pursuit method amounts to obtaining $\{\hat{g}_v\}_{v=1}^{N_b}$ from φ_{rn} as

$$\begin{aligned} \{\hat{g}_v\}_{v=1}^{N_b} &:= \arg \min_{\{g_v \in \mathcal{S}\}} \frac{1}{N_r N} \sum_{r=1}^{N_r} \sum_{n=1}^N \left(\varphi_{rn} - \sum_{v=1}^{N_b} g_v(\mathbf{x}_r) b_v(f_n) \right)^2 \\ &\quad + \lambda_s \sum_{v=1}^{N_b} \int_{\mathbb{R}^2} \left\| \nabla^2 g_v(\mathbf{x}) \right\|_F^2 d\mathbf{x} + \mu \sum_{v=1}^{N_b} \| [g_v(\mathbf{x}_1), \dots, g_v(\mathbf{x}_{N_r})] \|_2. \end{aligned} \quad (9.10)$$

Relative to (9.8), the cost in (9.10) is augmented with an additional regularization term weighted by a tuning parameter $\mu \geq 0$. Clearly, if $\mu = 0$, then (9.10) boils down to (9.8). To appreciate the role of the new penalty term, note that the minimization of $\| [g_v(\mathbf{x}_1), \dots, g_v(\mathbf{x}_{N_r})] \|_2$ intuitively shrinks all pointwise function values $\{g_v(\mathbf{x}_1), \dots, g_v(\mathbf{x}_{N_r})\}$ to zero for sufficiently large μ . Interestingly, it is shown in [44] that this suffices to guarantee that $\hat{g}_v(\mathbf{x}) \equiv 0 \forall \mathbf{x}$. This property is due to the fact that (9.9) still holds as a finite-dimensional solution to (9.10).

Figure 9.4 exhibits the capability of this approach for recovering the PSD maps on the 6th and 11th channel of the PUs abiding by the IEEE 802.11 wireless LAN standard. Compared to (9.6), the

**FIGURE 9.4**

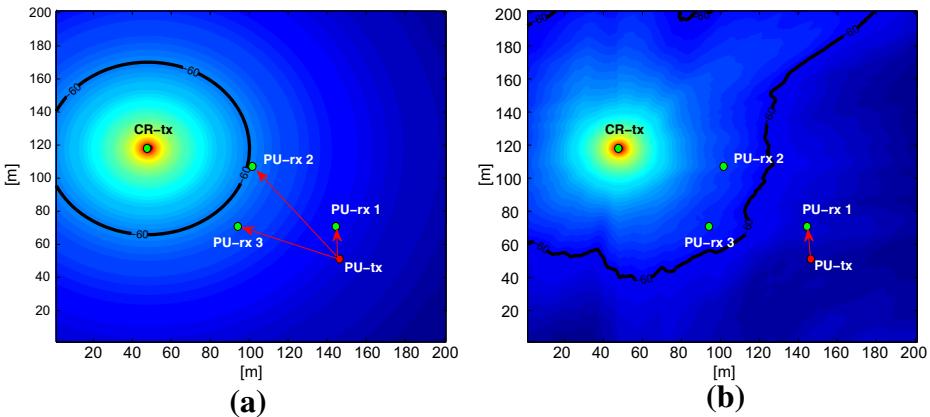
(Top) Power distribution across space with $g_6(\mathbf{x})$ in the band of 2437 MHz; (Top-left) actual distribution; (Top-right) estimated map from $N_r = 100$ CRs. (Bottom) Power distribution across space with $g_{11}(\mathbf{x})$ in the band of 2462 MHz; actual distribution and estimated map.

nonparametric approach in (9.10) can cope with shadowing effects at the price of increasing the number of CR sensors.

There are a number of practical issues in implementing the cartography algorithms. First, the PSD cartography algorithms must be run in a time scale commensurate with the coherence times of the RF emitter activities that the maps intend to capture. Typical implementation constraints would probably render it infeasible to track packetized transmission bursts or very high mobility. On the other hand, slowly varying RF landscape can be tracked based on the RF energy measurements accumulated over an appropriate duration. Online versions of the cartography algorithms have been developed for such purposes [14, 44]. The subsequent subsection provides an alternative mapping idea, which may be useful in tracing fast-varying PU activities. In addition to the robustness against channel uncertainties obtained via the total LS approach described in Section 2.09.2.1.3, robustness of the PSD maps to the grid granularity was considered [14]. Distributed synchronization as well as the position estimates of the CRs can be acquired via the GPS or other algorithms developed in the context of wireless sensor networking research.

2.09.2.2 Channel gain cartography

PSD cartography and PU localization algorithms are useful to identify regions that are “crowded” in terms of RF interference, and hence to be avoided by CR transmission. On the other hand, a complementary channel gain cartography is necessary to address the interference management issues in the spectrum underlay scenario. Channel coefficients and interference levels are generally acquired on a per-link basis by employing point-to-point training schemes. Although effective in conventional wireless networks, their application to the CR scenario is problematic due to the lack of cooperation between CR and PU nodes.

**FIGURE 9.5**

Coverage region of a transmitter. (a) Using a path loss-only model. (b) True.

As a motivating example, consider the setup in Figure 9.5a, where a CR transmitter aims to spatially reuse the frequency bands occupied by the PUs. As PU and CR systems do not generally cooperate, the CR transmitter relies upon a simple path loss model to ensure PU protection; i.e., to guarantee that the interference inflicted to PU receivers does not exceed a prescribed threshold [4,49], which is set to -60 dB in Figure 9.5a. However, due to random shadowing and small-scale propagation effects, employing a path loss-only model to calibrate the transmit-power may cause undesired disruption of PU communications [50], as depicted in Figure 9.5b. On the other hand, when significant signal attenuation exists due to shadowing, the CR links can benefit from it by raising the transmit-power levels. The key here is cognition of the spatio-temporal evolution of propagation channels.

Toward this end, a novel approach was recently put forth in [15,51], where the concept of *channel gain map* was introduced. For a CR node located at $\mathbf{x}_r \in \mathcal{A}$, the *local* channel gain (CG) map denoted by $g_{\mathbf{x} \rightarrow \mathbf{x}_r}(t, f)$ represents the CG of the link $\mathbf{x} \rightarrow \mathbf{x}_r$ for an arbitrary position $\mathbf{x} \in \mathcal{A}$, not necessarily occupied by a CR node. Similarly, for an *arbitrary* location $\mathbf{x}_s \in \mathcal{A}$, where none of the CRs resides, the *global* CG map collects the propagation coefficient of link $\mathbf{x} \rightarrow \mathbf{x}_s$ [51]; i.e., it contains CGs of links disjoint from any of the CR-to-CR links. Omitting the map's dependence on f for brevity (as separate maps can be constructed for each f), the CG $g_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ can be decomposed into path loss, shadowing and small-scale fading effects [52,53]. Averaging out small-scale fading [54], one obtains

$$G_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) := 10 \log_{10} g_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) = G_0 - 10\alpha \log_{10} (\|\mathbf{x} - \mathbf{x}_s\|_2) + s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t), \quad (9.11)$$

where G_0 denotes the path gain per unit distance, α the path loss exponent, and $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ the shadow fading in dB at time t . Once G_0 and α are known, the CG map $g_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ can be obtained provided the shadowing component can be predicted for every pair of points $\mathbf{x}_s, \mathbf{x} \in \mathcal{A}$.

The steps in constructing CG maps are: (1) characterization of the correlation among channel coefficients over different wireless links; and (2) development of an appropriate statistical inference algorithm leveraging the channel correlation to predict propagation gains of arbitrary links.

2.09.2.2.1 Spatio-temporal channel correlation

The shadowing, created by attenuation and diffraction of propagating signals owing to obstructions, can be accurately modeled as log-normal distributed [52,53]; thus, Gaussian-distributed when expressed in dB. However, characterization of its correlation is challenging, especially when samples are taken at different locations and time instants. Well-established correlation models for shadow fading are available for cellular networks, in which mobile terminals are assumed to move with constant velocity [55]. An extension involving one mobile and two base stations was proposed in [56], and multi-hop relay scenarios were studied in [57]. An experimentally validated parametric model for nomadic as well as mobile distributed channels was reported in [58]. The importance of shadowing in analyzing performance of wireless ad hoc networks was pointed out in [59], where a spatial correlation model was put forth to capture correlation of shadowing through a common “spatial loss” field $\ell(\mathbf{x}, t)$. Specifically, shadow fading is modeled as

$$s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) = \frac{1}{\|\mathbf{x} - \mathbf{x}_s\|_2^{1/2}} \int_{\mathbf{x}}^{\mathbf{x}_s} \ell(\mathbf{u}, t) d\mathbf{u}. \quad (9.12)$$

To allow spatio-temporal tracking of propagation gains, the spatial correlation model of [59] was judiciously extended to accommodate temporal variations in [51]. In particular, inspired by spatio-temporal modeling approaches in geostatistics and environmental science [60,61], the dynamics of the spatial loss field are captured as

$$\ell(\mathbf{x}, t) = \bar{\ell}(\mathbf{x}, t) + \tilde{\ell}(\mathbf{x}, t), \quad \text{and} \quad \bar{\ell}(\mathbf{x}, t) = \int_{\mathcal{A}} w(\mathbf{x}, \mathbf{u}) \bar{\ell}(\mathbf{u}, t-1) d\mathbf{u} + \eta(\mathbf{x}, t), \quad (9.13)$$

where $\bar{\ell}(\mathbf{x}, t)$ is spatially and temporally colored with $w(\mathbf{x}, \mathbf{u})$ modeling the interaction between $\ell(\mathbf{x}, t-1)$ at position \mathbf{x} at time t and $\ell(\mathbf{u}, t-1)$ at position \mathbf{u} at time $(t-1)$; and $\tilde{\ell}(\mathbf{x}, t)$ and $\eta(\mathbf{x}, t)$ are spatially colored yet temporally white zero-mean Gaussian stationary random fields. Plugging (9.13) into (9.12) yields $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) = \bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) + \tilde{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$, with the former spatially and temporally colored and the latter spatially colored but temporally white.

From a signal processing perspective, it would be desirable to reduce the dimensionality of the state-space model described by (9.13). One way to do it is through a basis expansion representation. Let $\{\psi_k(\cdot)\}_{k=1}^\infty$ be a set of complete prespecified orthonormal bases defined on \mathcal{A} . Then, $\bar{\ell}(\mathbf{x}, t)$ and $w(\mathbf{x}, \mathbf{u})$ can be approximated as $\bar{\ell}(\mathbf{x}, t) = \sum_{k=1}^K \chi_k(t) \psi_k(\mathbf{x})$ and $w(\mathbf{x}, \mathbf{u}) = \sum_{k=1}^K \beta_k(\mathbf{x}) \psi_k(\mathbf{u})$, where $\{\chi_k(t)\}$ and $\{\beta_k(\mathbf{x})\}$ are the basis expansion coefficients. Upon substituting these into (9.13), and sampling at locations $\{\mathbf{x}_r \in \mathcal{A}\}_{r=1}^{N_r}$, a finite-dimensional state equation for $\bar{\ell}(\mathbf{x}, t)$ is obtained as $\chi(t) = \mathbf{T}\chi(t-1) + \mathbf{J}\eta(t)$ with $\chi(t) := [\chi_1(t) \dots \chi_K(t)]^T$ denoting the state vector, and \mathbf{T} and \mathbf{J} determined from $\{\beta_k(\cdot)\}$ and $\{\psi(\cdot)\}$. This leads to a finite-dimensional representation of $\bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ as well. Specifically, upon defining $\phi_{\mathbf{x}_s \rightarrow \mathbf{x}, k} := (\|\mathbf{x}_s - \mathbf{x}\|)^{-1/2} \int_{\mathbf{x}_s}^{\mathbf{x}} \psi_k(\mathbf{u}) d\mathbf{u}$, $\bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ can be approximated as $\bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) \approx \boldsymbol{\phi}_{\mathbf{x} \rightarrow \mathbf{x}_s}^T \boldsymbol{\alpha}(t)$, where $\boldsymbol{\phi}_{\mathbf{x} \rightarrow \mathbf{x}_s} := [\phi_{\mathbf{x} \rightarrow \mathbf{x}_s, 1} \dots \phi_{\mathbf{x} \rightarrow \mathbf{x}_s, K}]^T$ depends only on the spatial coordinates \mathbf{x}_s and \mathbf{x} . Based on this spatio-temporal model for $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$, CG maps are constructed as follows.

2.09.2.2.2 CG map construction

Consider a network of N_r CRs $\{U_n\}_{n=1}^{N_r}$ at positions $\{\mathbf{x}_n\}_{n=1}^{N_r}$ known to one another, which exchange training signals in a time-division multiple-access (TDMA)-fashion to estimate their channel gains.

Suppose that each CR U_n can measure the received powers from the transmissions of the set \mathcal{M}_n of nodes, where $\mathcal{M}_n \subset \{U_1, \dots, U_{N_r}\} \setminus \{U_n\}$. With node $U_j \in \mathcal{M}_n$ transmitting its training sequence over a given TDMA slot at time t , receiver U_n can estimate $g_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)$ (and thus $G_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)$ after translating it to a dB scale) by measuring the received power. Subtracting the known deterministic path loss from the path gain, a measurement $\check{s}_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)$ of shadowing $s_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)$ is readily obtained. Let $\check{\mathbf{s}}_n(t)$ denote the vector collecting $\{\check{s}_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)\} \forall U_j \in \mathcal{M}_n$. Then, by pooling measurements from all CRs to a super-vector $\check{\mathbf{s}}(t) := [\check{\mathbf{s}}_1^T(t), \dots, \check{\mathbf{s}}_{N_r}^T(t)]^T$, one can write

$$\check{\mathbf{s}}(t) = \Phi \boldsymbol{\alpha}(t) + \tilde{\mathbf{s}}(t) + \boldsymbol{\epsilon}(t), \quad (9.14)$$

where Φ and $\tilde{\mathbf{s}}(t)$ are constructed with rows $\{\phi_{\mathbf{x}_j \rightarrow \mathbf{x}_n}^T\}$ and entries $\{\tilde{s}_{\mathbf{x}_j \rightarrow \mathbf{x}_n}(t)\}$, $n = 1, \dots, N_r$, respectively; and $\boldsymbol{\epsilon}(t)$ captures Gaussian-distributed measurement errors [54].

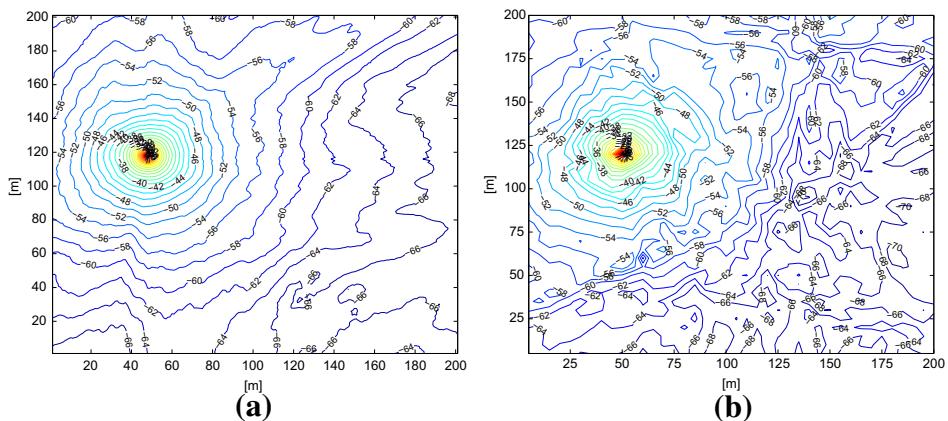
Based on the spatio-temporal model for $\bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$, and the measurement Eq. (9.14), an adaptation of the Kriged Kalman filtering (KKF) to track the time-varying shadowing field was proposed in [15, 51]. KKF is a universal Kriging approach [62], where the spatio-temporal evolution of $\bar{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}$ is tracked via Kalman filtering (KF) [60, 61]. Then, the shadow fading map $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ is obtained $\forall \mathbf{x}_s, \mathbf{x}, t$ by complementing the trend estimate with an estimate of $\tilde{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ obtained via Kriging interpolation [62]. Specifically, let $\hat{\boldsymbol{\alpha}}(t|t) := \mathbb{E}\{\boldsymbol{\alpha}(t)|\check{\mathbf{s}}_{1:t}\}$ be the minimum mean-square error (MMSE) estimate of $\boldsymbol{\alpha}(t)$ obtained via KF, given the accumulated data $\check{\mathbf{s}}_{1:t} := \{\check{\mathbf{s}}(\tau)\}_{\tau=1}^t$. Further, let $\mathbf{P}(t|t) := \text{cov}\{\boldsymbol{\alpha}(t)|\check{\mathbf{s}}_{1:t}\}$ be the KF estimation error covariance matrix. Then, conditioned on $\check{\mathbf{s}}_{1:t}$, the shadow fading process $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ for any $\mathbf{x}_s, \mathbf{x} \in \mathcal{A}$, is Gaussian distributed with mean and variance given, respectively, by

$$\hat{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) := \mathbb{E}\{s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)|\check{\mathbf{s}}_{1:t}\} = \phi_{\mathbf{x} \rightarrow \mathbf{x}_s}^T \hat{\boldsymbol{\alpha}}(t|t) + \mathbf{c}_{\tilde{s}}^T(\mathbf{x}_s, \mathbf{x}) \Sigma^{-1} [\check{\mathbf{s}}(t) - \Phi \hat{\boldsymbol{\alpha}}(t|t)], \quad (9.15a)$$

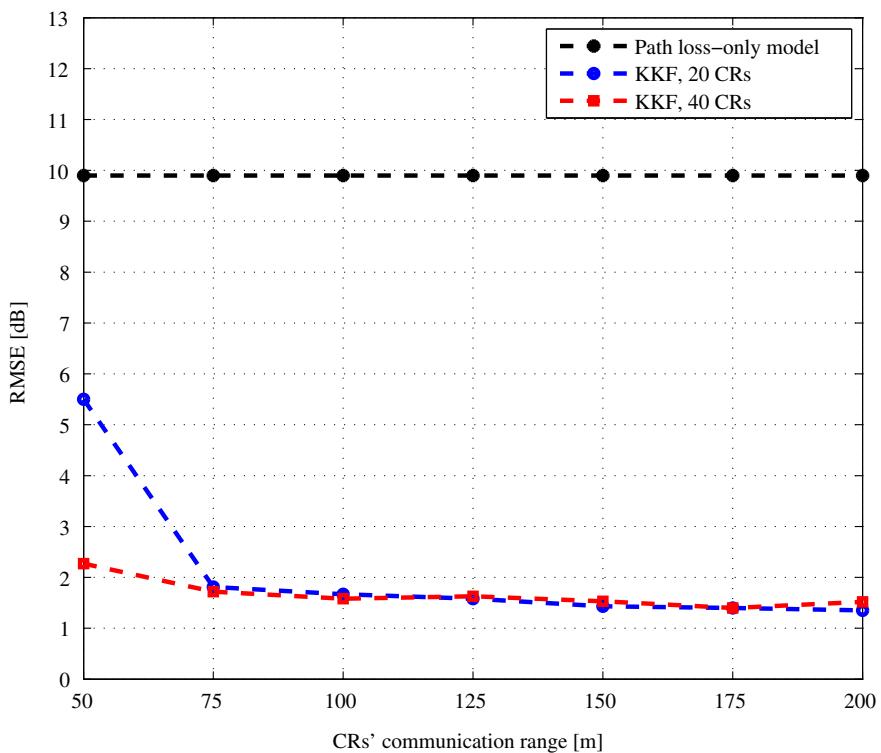
$$\begin{aligned} \sigma_{\mathbf{x} \rightarrow \mathbf{x}_s}^2(t) &:= \text{var}\{s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)|\check{\mathbf{s}}_{1:t}\} = \sigma_{\tilde{s}}^2 - \mathbf{c}_{\tilde{s}}^T(\mathbf{x}_s, \mathbf{x}) \Sigma^{-1} \mathbf{c}_{\tilde{s}}(\mathbf{x}_s, \mathbf{x}) \\ &\quad + \left[\phi_{\mathbf{x} \rightarrow \mathbf{x}_s}^T - \mathbf{c}_{\tilde{s}}^T(\mathbf{x}_s, \mathbf{x}) \Sigma^{-1} \Phi \right] \mathbf{P}(t|t) \left[\phi_{\mathbf{x} \rightarrow \mathbf{x}_s}^T - \Phi^T \Sigma^{-1} \mathbf{c}_{\tilde{s}}(\mathbf{x}_s, \mathbf{x}) \right], \end{aligned} \quad (9.15b)$$

where $\mathbf{c}_{\tilde{s}}(\mathbf{x}_s, \mathbf{x}) := \mathbb{E}\{\tilde{s}(t) \tilde{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)\}$, and $\Sigma := \text{cov}\{\tilde{s}(t)\} + \text{cov}\{\boldsymbol{\epsilon}(t)\}$. The CG map estimate $\hat{G}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ can now be constructed from $\hat{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ by adding back the deterministic path loss component; i.e., $\hat{G}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t) = G_0 - 10\alpha \log_{10}(\|\mathbf{x} - \mathbf{x}_s\|_2) + \hat{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$. Means of acquiring \mathbf{T} , $\mathbf{c}_{\tilde{s}}$ and Σ can be found in [15].

Figure 9.6a shows the (true) CG map corresponding to a PU located at $\mathbf{x}_s = (50, 120)$ m. Path loss parameters are set to $G_0 = 0$ and $\alpha = 3$. Clearly, the CG map exhibits a peak at location \mathbf{x}_s ; however, due to the spatially inhomogeneous shadowing component $s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$, whose standard deviation is approximately 10 dB, the overall CG map decays non-isotropically. Thus, estimating the shadowing field is essential for CR network operation and effective PU protection. To estimate the CG map, 20 CRs uniformly distributed over a square area of 200 m \times 200 m exchange signals to acquire propagation gains; the communication range was set to 125 m. An estimated version of the map in Figure 9.6a is depicted in Figure 9.6b. It can be seen that the KKF-based spatial interpolation can effectively predict the shadow fading process (and hence the channel gains) even in locations where no measurements were made. The error in reconstructing the map, evaluated over a grid of 36 evenly spaced locations was 1.5 dB, significantly lower than the standard deviation of the shadow fading. Figure 9.7 depicts the root-mean-square-errors (RMSEs) of the KKF for variable communication range and number of collaborating CRs. One can notice that the map estimation quality is maintained even when connectivity of the CR network is sparse due to shorter communication ranges.

**FIGURE 9.6**

Global CG map. (a) True map. (b) Estimated map.

**FIGURE 9.7**

Standard deviation of CG map estimation error.

2.09.2.2.3 Coverage region estimation

Although sensing schemes can locate active PUs, particularly challenging is to acquire the locations of “passive” PUs, which do not transmit but just listen. Nonetheless, those receivers still need to be protected from interference under the PU-CR hierarchy [4]. To this end, the coverage region of the PUs can be computed, where *potential* PU receivers can reside. This illustrates an application of the CG maps.

Let $\Pi_{\mathbf{x}}(t)$ denote the average power in dB received at location $\mathbf{x} \in \mathcal{A}$ due to the transmission of a PU located in \mathbf{x}_s signaling at power $P_s := 10 \log_{10} p_s$. Then, $\Pi_{\mathbf{x}}(t)$ can be expressed as $\Pi_{\mathbf{x}}(t) = P_s + G_0 - 10\alpha \log_{10} \|\mathbf{x}_s - \mathbf{x}\|_2 + s_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$. Based on the estimated CG map, $\Pi_{\mathbf{x}}(t)$ can be modeled as Gaussian distributed with mean $P_s + G_0 - 10\alpha \log_{10} \|\mathbf{x}_s - \mathbf{x}\|_2 + \hat{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)$ and variance $\sigma_{\mathbf{x} \rightarrow \mathbf{x}_s}^2(t)$.

Since a PU receiver can reliably decode the desired message only if the received power exceeds a certain threshold Π_{\min} (dB), one can compute the probability that a PU receiver at \mathbf{x} can decode as [48, Ch. 2]

$$P_{\mathbf{x}}^{\text{cov}}(t) := \Pr\{\Pi_{\mathbf{x}}(t) \geq \Pi_{\min}\} = Q\left(\frac{\Pi_{\min} - P_s - G_0 + 10\alpha \log_{10} \|\mathbf{x}_s - \mathbf{x}\|_2 - \hat{s}_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)}{\sigma_{\mathbf{x} \rightarrow \mathbf{x}_s}(t)}\right), \quad (9.16)$$

where $Q(\cdot)$ is the standard Gaussian tail function. The coverage region of the device located in \mathbf{x}_s is defined as the set of locations in \mathcal{A} , for which the coverage probability is no smaller than a threshold ν [48, Ch. 2]; i.e., $\mathcal{C}(t) := \{\mathbf{x} \in \mathcal{A} | P_{\mathbf{x}}^{\text{cov}}(t) \geq \nu\}$. In the absence of CG map knowledge, CRs would set $\hat{s}_{\mathbf{x}_s \rightarrow \mathbf{x}}(t) = 0$; consequently, \mathcal{C} would reduce to a time-invariant disc centered at \mathbf{x}_s ; see Figure 9.8a with $\Pi_{\min} = -60$ dB and $\nu = 0.4$. On the contrary, the CG map can portray the coverage region more

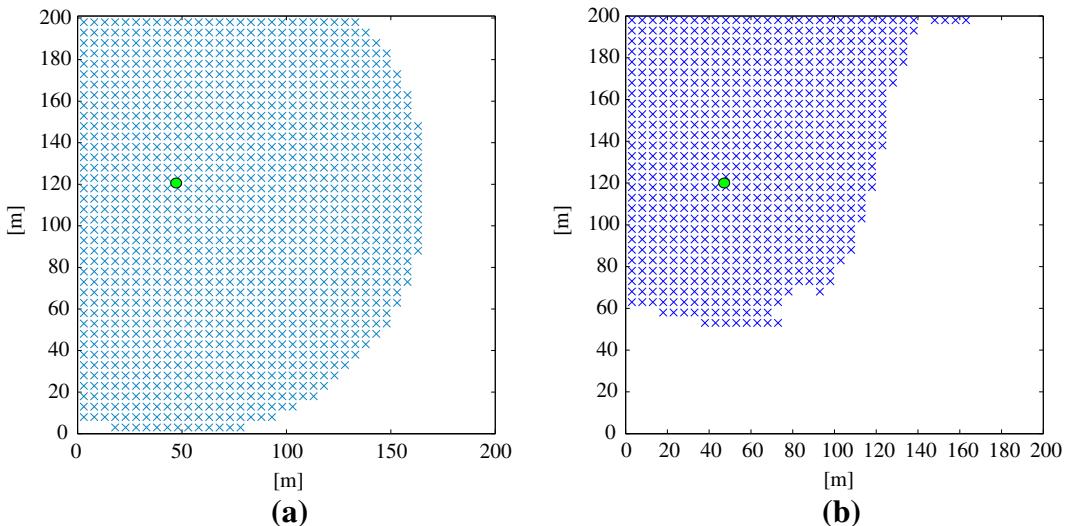


FIGURE 9.8

Coverage region of a PU transmitter. (a) Path loss-only model. (b) With CG map.

accurately as depicted in Figure 9.6b. The approach can also accommodate small-scale fading effects, as the composite log-normal and Nakagami fading turns out to be well approximated as log-normal [63, 64].

2.09.3 Sensing at the MAC layer

Spectrum sensing at the PHY layer is primarily concerned with detection of PU signals for the purpose of identifying transmission opportunities and creating spatio-temporal RF maps. In practice however, the PU spectrum is several times wider than the sensing bandwidth of a single CR transceiver. The *goal* of MAC layer sensing is to *schedule* the sensing operations of CRs across bands so as to locate maximal transmission opportunities. The MAC layer sensing problem is not trivial, as it involves balancing resources between sensing and communication. If multiple CRs are present, the MAC sensing algorithm must also take into account the possible contention among the CRs for medium access.

The design of MAC sensing algorithms is guided mainly by the rate at which the underlying spectrum occupancy changes. Slow temporal variations in the spectrum occupancy, as encountered in the TV bands, allow for longer sensing intervals. Consequently, MAC sensing algorithms in these bands simply search exhaustively through a wide range of frequencies before transmitting. On the other hand, fast spectral variations, as in the cellular bands, allow for only a few bands to be sensed before transmission. Therefore, the MAC sensing algorithms here need to utilize statistical inference to predict spectrum occupancy in order to better schedule sensing operations across bands. An alternative approach would be to design the analog front end of the CRs suitable for wideband spectrum sensing [65]. In this case, the trade-off between the complexity of the sensing hardware including the analog-to-digital converters (ADCs) and the sensing accuracy must be carefully examined [66]. In the sequel, the focus will be on the MAC sensing algorithms based on low-complexity narrowband sensing hardware.

2.09.3.1 Wireless regional area networks

The IEEE 802.22 standard for wireless regional area networks (WRAN) specifies a MAC sensing protocol for CRs operating on the spectrum assigned to TV services (between 47 and 910 MHz) [26]. The sensing in IEEE 802.22 operates in two stages: a *fast sensing* stage involving rapid probes on multiple bands, each lasting at most a millisecond; and a *fine sensing* stage lasting 25 ms, on a specific band determined by the outcome of the first stage. The MAC layer also schedules *in-band* sensing for determining if a PU starts transmitting before active CRs have completed their transmissions. Distributed sensing is enabled by scheduling periodic *quiet times* (during which all CRs must stop their transmissions in order to sense) and exchange of band occupancy reports among CRs.

Since the presence of periodic quiet times interrupts CR transmissions, an alternative dynamic frequency hopping mode has been proposed for CRs with multiple transceivers [67]. In this mode, a sensing transceiver is used to identify out-of-band opportunities *while* concurrent CR transmission, thus eliminating the need for quiet times. Related works have dealt with other practical issues pertaining to the coexistence of multiple CR networks, presence of hidden PU nodes [68], and co-channel interference in multi-cell environments [69].

2.09.3.2 Cellular networks: MAC sensing as an inference problem

As mentioned earlier, the two-stage sensing operation for WRANs is not feasible in cellular systems, where band occupancy varies far more quickly. MAC sensing algorithms in this case employ statistical inference methods by utilizing the observed history of spectrum occupancy. The first step for MAC sensing involves band occupancy prediction using time-series models, and is described in the next subsection.

2.09.3.2.1 Band occupancy prediction

A binary time-series prediction approach has been proposed in [70]. Consider a synchronous time-slotted PU network operating over N licensed bands. The PU network state at the MAC layer can be specified by an $N \times 1$ vector $\mathbf{s}(t)$ with entries $s_n(t) \in \{0(\text{occupied}), 1(\text{idle})\}$, $n = 1, \dots, N$. Given p samples $\mathbf{s}(t-p), \dots, \mathbf{s}(t-1)$, logistic regression is used to predict the probability of band occupancy at time t , i.e.,

$$\Pr(s_n(t) = 1) = \frac{1}{1 + \exp(- (a_{n0} + \sum_{j=1}^p a_{nj} s_n(t-j) + v_n(t)))} \quad 1 \leq n \leq N \quad (9.17)$$

with regression coefficients $\{a_{nj}\}_{j=0}^p$, intercepts (or offsets) a_{n0} , and noise term $v_n(t)$ for $n = 1, \dots, N$. The coefficients $\{a_{nj}\}_{j=0}^p$ may be estimated by taking M observations per band, and maximizing the log-likelihood function [71]

$$\hat{\mathbf{a}}_n = \arg \max_{\mathbf{a}_n} \sum_{m=p+1}^M s_{nm} \log q_{nm}(\mathbf{a}_n) + (1 - s_{nm}) \log (1 - q_{nm}(\mathbf{a}_n)) \quad 1 \leq n \leq N, \quad (9.18)$$

where \mathbf{a}_n stacks the terms $\{a_{nj}\}_{j=0}^p$, and $q_{nm}(\mathbf{a}_n) := 1/[1 + \exp(- (a_{n0} - \sum_{j=1}^p a_{nj} s_n(m-j)))]$.

An alternative approach in this context is described in [72], that utilizes nonstationary autoregressive time-series models. In this framework, PU packet arrivals (and subsequently the number of PU transmissions) for each band follow a non-homogeneous Poisson process $\{\mathcal{A}(t), t \geq 0\}$ with time-varying rate parameter $\lambda(t)$. Packet arrivals within a time slot $(t, t + T_s)$ may then be modeled by a Poisson process with constant rate λ_ℓ/T_s , where $\lambda_\ell := \int_{t=\ell T_s}^{(\ell+1)T_s} \lambda(t) dt$. A seasonal autoregressive integrated moving average (SARIMA) model is proposed in [72] to track the rates $\{\lambda_\ell\}$, which takes the form

$$\lambda_\ell = \lambda_{\ell-1} + \lambda_{\ell-24} - \lambda_{\ell-25} + z_\ell + \theta z_{\ell-1} + \Theta z_{\ell-24} + \theta \Theta z_{\ell-25}, \quad (9.19)$$

where $T_s = 1$ h, and $z_\ell \sim \mathcal{N}(0, \sigma^2)$ is an error term. The model parameters $(\theta, \Theta, \sigma^2)$ are estimated using techniques described in [73]. In contrast to the logistic regression approach of [70], SARIMA models allow nonparametric approaches for prediction of band occupancies, and handle trend and seasonal nonstationarities.

In a nutshell, high-order time-series models enable prediction of spectrum occupancies, and can be utilized for scheduling sensing operations across bands. However, since prediction must be carried out per band, these models become highly inefficient in terms of sensing overload and computational complexity if the number of available bands is large. The next section details joint prediction and scheduling MAC sensing algorithms that scale gracefully with the number of available bands.

2.09.3.2.2 Band occupancy scheduling

Since sensing is performed only on a fraction of available bands, the algorithms described in this section perform prediction based only on the observed spectrum occupancy history. One of the early MAC sensing algorithms in this context was proposed in [22], where evolution of the PU state $\mathbf{s}(t)$ is assumed to follow a Markov chain with known transition probabilities $p_{ij} := P(\mathbf{s}(t) = \mathbf{i} | \mathbf{s}(t-1) = \mathbf{j})$, for all $i, j \in \{0, 1\}^{2^N}$, that stay constant for at least T time slots, and can be estimated as described in [74]. Each CR seeks to access the spectrum opportunities that arise when one or more of the bands are idle. Owing to hardware constraints, a CR can only sense at most $L_1 \leq N$ bands, and access at most $L_2 \leq L_1$ bands. Consequently, at time slot t , the CR chooses a subset $\mathcal{B}_s(t)$ (with cardinality not exceeding L_1) of bands to sense, and a subset $\mathcal{B}_a(t) \subseteq \mathcal{B}_s(t)$ of bands to access, thus achieving a throughput of

$$R_{\{\mathcal{B}_s(t), \mathcal{B}_a(t)\}}(t) := \sum_{n \in \mathcal{B}_a(t)} s_n(t) \beta_n, \quad (9.20)$$

where β_n is the bandwidth of the n th band. The goal is to sequentially choose $\mathcal{B}_s(t)$ and $\mathcal{B}_a(t)$ to maximize the total throughput achieved in T time slots, averaged over all possible state vectors $\{s_n(t)\}$.

Since the network state is not directly observable by the CRs, beliefs $\lambda_i(t) := \Pr(\mathbf{s}(t) = \mathbf{i} | \mathcal{H}(t))$ are utilized to make the sensing and access decisions, using the observed spectrum availability history $\mathcal{H}(t) := \{s_n(\ell) \forall n \in \mathcal{B}_s(\ell), 1 \leq \ell \leq t-1\}$. Collecting the beliefs in $\boldsymbol{\lambda}(t)$, a policy is defined as the function $\pi(t) : \boldsymbol{\lambda}(t) \rightarrow \{\mathcal{B}_s(t), \mathcal{B}_a(t)\}$. The optimal policy is given by

$$\{\pi^*(t)\}_{t=1}^T = \arg \max_{\{\pi(t)\}_{t=1}^T} \mathbb{E} \left[\sum_{t=1}^T R_{\{\mathcal{B}_s(t), \mathcal{B}_a(t)\}}(t) | \mathcal{H}(1) \right]. \quad (9.21)$$

The problem of determining the optimal policy can now be solved using a partially observable Markov decision process (POMDP) framework, as detailed in [22]. Note however that the size of the state space of (9.21) is exponential in N , which makes this approach inefficient for a large number of bands. The state space and hence the problem size can be reduced to N if the bands are assumed to evolve independently. The independence assumption decouples the problem into N sub-problems, each with a two-state Markov chain as shown in [22, 75]. The case of independently evolving bands can also be related to the restless multi-armed bandit problem [76], which allows for sensing algorithms with even lower complexity. Further works have extended the approach to include energy constraints [77], and sensing imperfections [78]. To address the exponential complexity of the general POMDP approach, the impact of using myopic policies is examined in [79]. These policies only maximize the expected throughput over the next time slot, but are still optimal under certain conditions.

A similar problem is considered in [80], but a separate control channel is allowed for each CR. The control channel is shown to enable better synchronization, alleviation of the hidden node problem, and consequently improved spectrum utilization. A hardware constrained MAC design is pursued in [81], which optimizes the sensing time while respecting hardware constraints such as single transceiver, partial spectrum sensing and limited spectrum aggregation. The intuition is that sensing for longer periods uncovers more suitable bands but leaves little time for communication. The optimal stopping time is derived by employing a dynamic programming approach; a related example is described in detail in the context of cross-layer sensing in Section 2.09.4.1.1. Extending the approach to consider heterogeneous

bands with varying capacities, the problem of determining the optimal sensing order is investigated in [21].

In an attempt to relax the time-slotted assumption in [22] and related works, a continuous-time Markov chain approach has been proposed for modeling the band occupancy process [82]. In this case, the state is defined similarly as before, but a band stays in state i (busy/idle) for an exponentially distributed time period (with rate λ_i or μ_i). The sensing operation may then be viewed as sampling the state process $s(t)$ in a periodic fashion. Specifically, the N bands are sensed in a round-robin fashion, with exactly one sample per time slot (of duration T_s). After each sensing operation, the MAC layer schedules a transmission on a selected band. It is shown in [83] that periodic sampling yields a discrete time Markov chain of sampled state vectors, whose transition probabilities can be derived in closed form. This allows for the transmission scheduling problem to be cast within the POMDP framework, yielding an optimal medium access policy. The framework also allows for the inclusion of constraints on the collision probabilities in a straightforward way [84]. A more general semi-Markov model has also been proposed, where the busy/idle times have arbitrary distributions $f_i^1(x)$ and $f_i^0(x)$ [85]. Related problems include the choice of sensing period [85], sensing order [86], and more generally the sensing policy in time and frequency [87]. Finally, the sensing order and optimal sensing policy design for the multiuser scenario have been considered in [88,89].

2.09.4 CR Sensing and cross-layer design

As is the case with most PHY/MAC functions in wireless communication systems, spectrum sensing is not an isolated task that can be designed independently of other components in the CR system. It is important when designing sensing algorithms to also take into account the overall system objectives. Further, given the enhanced awareness of the RF environment in which the CRs operate, it is essential to adapt the higher-layer tasks to the environment to maximize efficiency and the overall CR network performance.

As was mentioned in Section 2.09.3, there is a trade-off between sensing accuracy and network-wide performance objectives such as throughput, delay, and reliability. Consider for example the trade-off between sensing duration and throughput. As was explained in Section 2.09.2, the basic element of the sensing device is the detector that discriminates the H_1 hypothesis that says there is an ongoing PU transmission, from the H_0 hypothesis that says the spectrum is unoccupied. Since the sensing is done typically in challenging situations such as at a low SNR and without explicit support from the PU systems, typically a large number of samples must be collected for reliable detection. This inevitably increases sensing time. On the other hand, only when the sensing is finished (and the medium is determined to be idle), can the CR proceed to actual data transmission. Therefore, given that the total idle duration is limited (as the PU can come back and transmit), the more time is devoted to sensing, the less time is left for useful data transmission. Such a trade-off was studied in detail in [24] for fixed sample size (FSS) test-based sensing. A similar trade-off was investigated in the context of MAC layer design of sensing algorithms in [81,90].

Even when the sensing duration is fixed, there is trade-off between the probability of miss detection and the probability of false alarm, which affects the system objectives. To see this, one needs to note that if the detector misses the presence of PUs, it is likely that the CR will proceed to data transmission, causing interference to the licensed users. Thus, a low miss detection probability is

desired, which in turn increases the false alarm probability. However, the false alarms result in wasted opportunities for CR transmission, thus reducing system efficiency. This problem becomes only more interesting when multiple bands must be sensed concurrently, as was explored in [18, 19, 91]; see also [17, 92–95] for sequential alternatives in this context. Some of these formulations will be reviewed in Section 2.09.4.1.

An equally important issue is how to effectively tap into the significantly improved awareness of the RF environment obtained through the sensing and the RF cartography, for design and operation of higher-layer networking protocols. Numerous challenges in this direction include distributed resource optimization, quality-of-service management, and maintaining network robustness under uncertainty [23, 96, 97]. In Section 2.09.4.2, a recently developed robust routing scheme exploiting the RF maps will be showcased.

2.09.4.1 Joint sensing and resource optimization

While the FSS test-based sensing algorithms must fix the number of samples used for detection before actually “seeing” those samples, the sequential alternatives can decide on the fly whether the samples collected so far are informative enough for reliable detection, and continue taking samples if not. For a binary hypothesis testing problem, sequential probability ratio test (SPRT) is a well-established algorithm that minimizes the sample size on the average, for given detection and false alarm probability specifications. It can be applied in a straightforward manner to the single-band sensing problem as outlined in Section 2.09.2. A low complexity alternative based on energy detection was analyzed in [95]. Next, the more challenging case of multi-band joint sensing is reviewed using the sequential detection approaches.

2.09.4.1.1 Throughput-aware sequential sensing

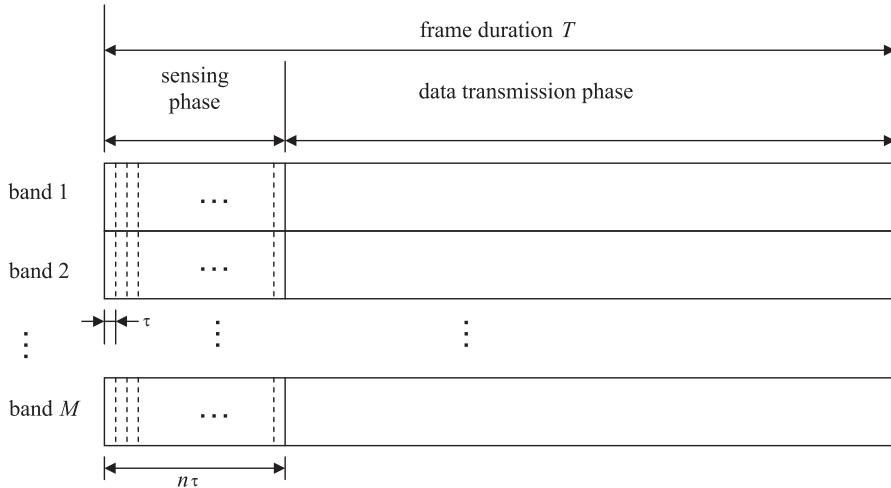
Consider a CR receiver that scans M bands in parallel for transmission opportunities. Due to self-interference issues, the radio is assumed to be half-duplex; i.e., it cannot sense on one band while transmitting on another. Thus, a CR frame of duration T is divided into a sensing phase of duration $n\tau$ (where τ is the sampling interval) and a data transmission phase of duration $(T - n\tau)$; see Figure 9.9.

Under the assumption that the spectrum occupancy of the PUs is independent across bands, binary hypothesis tests need to be performed on each band $m \in \{1, 2, \dots, M\}$. Denoting the samples obtained at time n as $\{r_n^{(m)}\}$, one aims to discriminate the following two hypotheses for band m

$$\begin{aligned} H_0^{(m)} : r_n^{(m)} &= z_n^{(m)}, \quad n \in \{1, \dots, N\}, \\ H_1^{(m)} : r_n^{(m)} &= h_n^{(m)} s_n^{(m)} + z_n^{(m)}, \quad n \in \{1, \dots, N\}, \end{aligned} \quad (9.22)$$

where $\{h_n^{(m)}\}$ denote the channel coefficients, and $\{z_n^{(m)}\}$ the noise. Adopting energy detection, the observation at time n is defined as $y_n^{(m)} := |r_n^{(m)}|^2$.

Let $\delta_n^{(m)} \in \{1, 0\}$ denote the permit-to-access decision for channel m , made after seeing up to the n th sample; if it is equal to 1, the channel is deemed idle (i.e., $H_0^{(m)}$ in effect), and 0 otherwise. Define $\delta_n := [\delta_n^{(1)} \ \delta_n^{(2)} \ \dots \ \delta_n^{(M)}]^T$. Denote also the PU occupancy over the M channels by \mathcal{H} , whose m th entry $\mathcal{H}^{(m)}$ takes values from $\{H_0^{(m)}, H_1^{(m)}\}$. We wish to characterize the effective throughput that the

**FIGURE 9.9**

CR frame structure.

CR can enjoy. Let $R^{(m)}$ denote the known rate that can be achieved when transmitting over channel m . Then, if the CR stops sensing after n sampling intervals and proceeds to data transmission, the overall throughput can be written as ($\mathbb{1}_{\{\cdot\}}$ denotes the indicator function)

$$f'_n(\mathcal{H}, \delta_n) = \frac{T - n\tau}{T} \sum_{m=1}^M R^{(m)} \mathbb{1}_{\{H_0^{(m)}\}} \delta_n^{(m)}, \quad n = 1, 2, \dots, N, \quad N\tau \leq T. \quad (9.23)$$

From (9.23), the throughput-sensing trade-off is apparent: as the number of observed samples increases, the factor $\frac{T - n\tau}{T}$ diminishes, but more available bands may be correctly identified to yield a higher value for the sum rate in (9.23).

Given the past observations $\mathbf{Y}_n \triangleq [\mathbf{y}_1 \mathbf{y}_2 \dots \mathbf{y}_n]$, the goal is to obtain the average throughput-optimal *stopping policy* $\Delta_n(\mathbf{Y}_n) \in \{S, \bar{S}\}$ that determines whether to stop (“ S ”) or not stop (“ \bar{S} ”) at each time n , and the *access policy* $\delta_n(\mathbf{Y}_n) \in \{1, 0\}^M$ indicating whether each band may be used for data transmission if the sensing stops at time n . In other words, the objective is to maximize the average throughput

$$E_{\mathbf{Y}_N, \mathcal{H}} \left\{ \sum_{n=1}^N \mathbb{1}_{\{\Delta_{n-1}=\bar{S}, \Delta_n=S\}} f'_n(\mathcal{H}, \delta_n) \right\} \quad (9.24)$$

over the control policies $\{\Delta_n(\cdot)\}_{n=1}^{N-1}$ and $\{\delta_n(\cdot)\}_{n=1}^N$. Here, $\Delta_{n-1} := [\Delta_0 \Delta_1 \dots \Delta_{n-1}]$ with $\Delta_0 \equiv \bar{S}$, and $\Delta_{n-1} = \bar{S}$ is a shorthand for $\Delta_0 = \Delta_1 = \dots = \Delta_{n-1} = \bar{S}$. Note also that $\Delta_N \equiv S$ by design, as we are dealing with a finite horizon problem. The indicator function in (9.24) ensures that the reward $f'_n(\cdot)$ is evaluated only at the smallest time slot n^* satisfying $\Delta_{n^*} = S$; for the rest of the time steps, $n < n^*$ and $n > n^*$, the summands are zero.

On the other hand, the CR access policy must ensure a low probability of “collision” with the ongoing PU transmissions due to miss detection. The “collision” probability $P_c^{(m)}$ on band m can be written as

$$P_c^{(m)} = \sum_{n=1}^N \Pr \left\{ \Delta_{n-1} = \bar{S}, \Delta_n = S, \delta_n^{(m)} = 1, |H_1^{(m)}| \right\} \quad (9.25)$$

which must be kept small; i.e., $P_c^{(m)} \leq \bar{\beta}$ for all m with $\bar{\beta}$ being a small positive threshold. Upon defining the so-called belief vector $\pi_n := [\pi_n^{(1)} \dots \pi_n^{(M)}]^T$ with entries $\pi_n^{(m)} := \Pr \left\{ H_0^{(m)} | y_1^{(m)}, \dots, y_n^{(m)} \right\}$, one can show that $P_c^{(m)}$ can be equivalently expressed as [17]

$$P_c^{(m)} = E_{\mathbf{Y}_N} \left\{ \sum_{n=1}^N \mathbb{1}_{\{\Delta_{n-1} = \bar{S}, \Delta_n = S\}} \delta_n^{(m)} \frac{1 - \pi_n^{(m)}}{1 - \pi_0^{(m)}} \right\}. \quad (9.26)$$

Furthermore, by first taking conditional expectation given \mathbf{Y}_N , and then taking unconditioned expectation, the average throughput in (9.24) can also be written as

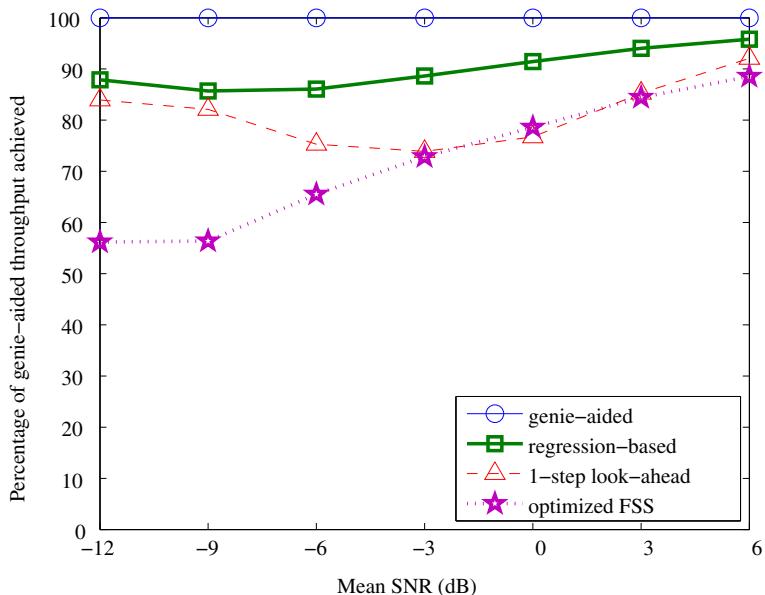
$$E_{\mathbf{Y}_N} \left\{ \sum_{n=1}^N \mathbb{1}_{\{\Delta_{n-1} = \bar{S}, \Delta_n = S\}} \frac{T - n\tau}{T} \sum_{m=1}^M R^{(m)} \pi_n^{(m)} \delta_n^{(m)} \right\}. \quad (9.27)$$

It can be verified that maximizing (9.27) over $\{\Delta_n(\cdot)\}_{n=1}^{N-1}$ and $\{\delta_n(\cdot)\}_{n=1}^N$ subject to $P_c^{(m)} \leq \bar{\beta}$ for all m adheres to the constrained dynamic programming (DP) formalism [98]. The problem can be tackled using Lagrange relaxation [99]. Interestingly, the optimal permit-to-access policy $\delta_n^{*(m)}$ for each band m is found to be a likelihood ratio test with the threshold depending on the Lagrange multiplier associated with the band [17]. A reduced-complexity, basis regression-based suboptimal stopping policy was also derived. It was seen that these policies could achieve a significant portion of the genie-aided policy (which non-causally possesses the information from the future observations), while significantly outperforming suboptimal 1-step look-ahead as well as the best FSS sensing schemes; see Figure 9.10.

2.09.4.1.2 Sequential sensing for real-time traffic

Another interesting challenge is to transport real-time traffic such as voice or video using CR networks. Real-time traffic often has stringent delay and minimum rate constraints. Thus, efficient design of the sensing algorithm is extremely important in this scenario. As a concrete paradigm, consider again the multi-band CR sensing problem, but with an outage constraint on the requested minimum rate. Thus, the goal is to minimize the sensing delay while finding enough number of idle bands that can support the given minimum rate with high probability.

Since the DP-based approach may be intractable with the chance constraint, a more structured approach was taken in [93]. Specifically, rather than formulating a rigorous DP problem to obtain the optimal access and stopping policies, a bank of SPRTs were used (one for each band) to sense the M bands in parallel. Then, an optimization problem was formulated to solve for the thresholds of the SPRTs to minimize the average sensing time while abiding by the minimum rate and the PU interference constraints. An important catch in this approach is that the stopping times of the individual SPRTs may be different. To mitigate this issue, a tractable objective of minimizing the largest average sensing time over the bands was employed. To much practical appeal, the associated optimization problem turns out to be convex [93].

**FIGURE 9.10**

Ratio of achieved average throughput relative to the genie-aided throughput.

2.09.4.2 Cartography-enabled route optimization

Volatile wireless connectivity in CR scenarios can be robustified via optimizing network operations leveraging spectrum sensing. Routing protocols for wireless networking hinge on the notion of network connectivity graph to find source-to-destination paths optimal in some sense [100]. The edge weights in CR network graphs should reflect spatio-temporal PU spectrum occupancy statistics, network-wide spectral opportunities, and propagation medium characteristics. Appropriate cross-layer design based on such information is instrumental for efficient resource allocation and for addressing end-to-end quality-of-service demands.

Based on the output of the sensing task, link weights can be used to indicate the amount of spectral resources available per CR-to-CR link [101]. Optimal source-to-destination paths can then be found via Dijkstra or Bellman Ford algorithms. A two-phase approach combining static mesh routing with per-packet dynamic routing is proposed in [102], where network nodes first compute an expected route cost as well as a set of candidate forwarding nodes, and then route packets via links with highest channel quality. In [103], a routing scheme is developed to avoid network zones with no guarantees of stable CR connectivity based on spatio-temporal statistics of the PU activities. The concept of coverage map is leveraged in [104] to devise routing strategies supporting multiple classes of routes, and hence different CR quality-of-service demands. The effects of random PU interference on CR links is accounted for in [105], where the predicted capacity of each CR-to-CR link is exploited to compute the path that is most likely to meet CR end-to-end requirements. An optimal cross-layer design problem was considered in

[106] to compute not only optimal routes, but also physical and medium access parameters that dictate the expected packet forwarding capabilities. In doing so, the statistics of propagation channels were exploited, along with PU state information provided by spectrum sensing.

Following [106], consider a CR wireless network sharing spectral resources with an incumbent PU system in an underlay setup [4] to route data packets to a sink node U_{N_r+1} . The dynamical and stochastic nature of the propagation medium naturally suggests stochastic routing strategies [107, 108], in which each CR node U_n decides per time slot whether to route packets toward a neighboring node U_i with probability $t_{\mathbf{x}_n \rightarrow \mathbf{x}_i} \in [0, 1]$. To capture channel- and interference-induced uncertainty, let $r_{\mathbf{x}_n \rightarrow \mathbf{x}_i} \in [0, 1]$ denote the probability that a packet transmitted from CR U_n is correctly decoded by U_i . As a result, the stochastic nature of data transport is captured by the pairwise packet delivery probabilities $\{t_{\mathbf{x}_n \rightarrow \mathbf{x}_i}, r_{\mathbf{x}_n \rightarrow \mathbf{x}_i}\}$.

A well-established criterion for successful packet reception is to require the signal-to-interference-plus-noise ratio (SINR) to stay above a certain threshold [109], which is determined by the receiver structure, transmit-power, modulation, and coding scheme. Recall that $g_{\mathbf{x}_n \rightarrow \mathbf{x}_i}$ denotes the log-normal-distributed propagation gain between U_n and U_i , which accounts for both shadowing and Nakagami fading [53, 64]. Then, the SINR of link $\mathbf{x}_n \rightarrow \mathbf{x}_i$ can be expressed as

$$\gamma_{\mathbf{x}_n \rightarrow \mathbf{x}_i} := \frac{p_n g_{\mathbf{x}_n \rightarrow \mathbf{x}_i}}{\sigma_i^2 + \sum_{S=1}^{N_S} \pi_S}, \quad (9.28)$$

where σ_i^2 stands for the receiver noise power at CR U_i , $p_n \in (0, p_n^{\max}]$ the transmission power of U_n , and π_S the received power from PU transmitter $S = 1, \dots, N_S$. As CR and PU nodes do not generally cooperate, interfering powers $\{\pi_S\}$ are not known. However, their statistics collected by sensing algorithms and CG cartography [cf. (15)] can be used instead. Exploiting the Fenton-Wilkinson method [110], the distribution of SINRs $\{\gamma_{\mathbf{x}_n \rightarrow \mathbf{x}_i}\}$ can be well approximated as log-normal, with mean and variance of the first- and the second-order moments of $\{g_{\mathbf{x}_n \rightarrow \mathbf{x}_i}\}$ and $\{\pi_S\}$ provided by the CG maps [63]. Let $\bar{\gamma}_{\mathbf{x}_n \rightarrow \mathbf{x}_i}$ denote the SINR threshold, and $\bar{\Gamma}_{\mathbf{x}_n \rightarrow \mathbf{x}_i} := 10 \log_{10} \bar{\gamma}_{\mathbf{x}_n \rightarrow \mathbf{x}_i}$. Assume that CRs adopt a random access strategy, and let μ_n and \mathcal{I}_{ni} denote the transmission probability of CR U_n and the set of nodes whose transmissions interfere with link $U_n \rightarrow U_j$, respectively. Then, the probability that a packet transmitted from the i th CR U_n is correctly received by U_i can be expressed as

$$\begin{aligned} r_{\mathbf{x}_n \rightarrow \mathbf{x}_i} &= \prod_{j \in \mathcal{I}_{ni}} (1 - \mu_j) \cdot \Pr\{\gamma_{\mathbf{x}_n \rightarrow \mathbf{x}_i} > \bar{\gamma}_{\mathbf{x}_n \rightarrow \mathbf{x}_i}\} \\ &\approx \prod_{j \in \mathcal{I}_{ni}} (1 - \mu_j) \cdot Q\left(\frac{\bar{\Gamma}_{\mathbf{x}_n \rightarrow \mathbf{x}_i} - P_n - m_{\mathbf{x}_n \rightarrow \mathbf{x}_i}}{\sigma_{\mathbf{x}_n \rightarrow \mathbf{x}_i}}\right), \end{aligned} \quad (9.29)$$

where $P_n := 10 \log_{10} p_n$; while mean and standard deviation of the dB-expressed SINR are denoted by $m_{\mathbf{x}_n \rightarrow \mathbf{x}_i}$ and $\sigma_{\mathbf{x}_n \rightarrow \mathbf{x}_i}$, respectively.

Assume that exogenous packet arrivals at node U_n are modeled as a stationary stochastic process with average rate $\rho_n \geq 0$. With λ_n denoting the average rate of packet departures from U_n , and assuming fully-backlogged queues per node [107, 111], the exogenous traffic rates $\{\rho_n\}$ and $\{\lambda_n\}$ abide by the flow conservation constraints

$$\rho_n = \lambda_n \sum_{i \in \mathcal{N}_{n \rightarrow}} t_{\mathbf{x}_n \rightarrow \mathbf{x}_i} r_{\mathbf{x}_n \rightarrow \mathbf{x}_i} - \sum_{j \in \mathcal{N}_{\rightarrow n}} \lambda_j t_{\mathbf{x}_j \rightarrow \mathbf{x}_n} r_{\mathbf{x}_j \rightarrow \mathbf{x}_n}, \quad (9.30)$$

where $\mathcal{N}_{n \rightarrow} := \{j | r_{\mathbf{x}_n \rightarrow \mathbf{x}_j} > 0, j = 1, \dots, N+1, j \neq n\}$ is the set of nodes that decode U_n 's transmissions with non-zero probability, and $\mathcal{N}_{\rightarrow n} := \{i | r_{\mathbf{x}_i \rightarrow \mathbf{x}_n} > 0, i = 1, \dots, N, i \neq n\}$ the set of nodes that route packets through U_n . For queue stability, it suffices to have $0 \leq \lambda_n \leq \mu_n$, for each CR U_n [112].

To complete the formulation, consider N_R PU receivers, whose locations $\{\mathbf{y}_R\}$ have been estimated via CG cartography [cf. (9.16)]. The interference caused to PU R is given by $i_R := \sum_n p_n g_{\mathbf{x}_n \rightarrow \mathbf{y}_R}$, where $g_{\mathbf{x}_n \rightarrow \mathbf{y}_R}$ denotes the channel gain between CR U_n and PU R . Approximate the channel gain $g_{\mathbf{x}_n \rightarrow \mathbf{y}_R}$ as log-normal [63], with mean and variance provided by the CG map of PU receiver R . Then, defining $\kappa := 0.1 \ln(10)$, the average interference experienced at the PU node R is given by

$$\mathbb{E}\{i_{\mathbf{y}_R}\} = \sum_{n=1}^N \mu_n e^{\kappa P_n + \kappa(G_0 - 10\alpha \log_{10} \|\mathbf{x}_n - \mathbf{y}_R\|_2 - s_{\mathbf{x}_n \rightarrow \mathbf{y}_R}) + \frac{\kappa^2}{2} \sigma_n^2}, \quad (9.31)$$

which must not exceed a predetermined threshold i_R^{\max} .

Based on these developments, the statistical routing task is formulated as the following optimization problem:

$$\max_{\substack{\{P_n \leq P_n^{\max}\}, \{\rho_n \geq 0\}, \{\mu_n \geq 0\}, \\ \{t_{\mathbf{x}_n \rightarrow \mathbf{x}_i} \geq 0\}, \{\lambda_n \geq 0\}}} \sum_{n=1}^N \mathcal{U}_n(\rho_n) - \sum_{n=1}^N \mathcal{C}_n(P_n) \quad (9.32.a)$$

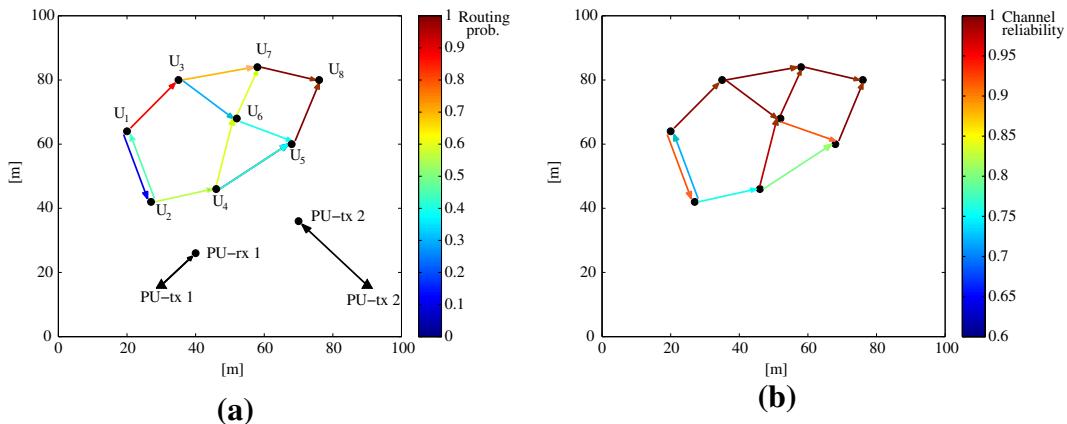
$$\text{subject to } \rho_n \leq \lambda_n \sum_{j \in \mathcal{N}_{n \rightarrow}} t_{\mathbf{x}_n \rightarrow \mathbf{x}_j} r_{\mathbf{x}_n \rightarrow \mathbf{x}_j} - \sum_{i \in \mathcal{N}_{\rightarrow n}} \lambda_i t_{\mathbf{x}_i \rightarrow \mathbf{x}_n} r_{\mathbf{x}_i \rightarrow \mathbf{x}_n} \quad (9.32.b)$$

$$\sum_{i \in \mathcal{N}_{n \rightarrow}} t_{\mathbf{x}_n \rightarrow \mathbf{x}_i} \leq 1, \quad \lambda_n \leq \mu_n - \epsilon, \quad \mu_n \leq 1 \quad (9.32.c)$$

$$\mathbb{E}\{i_{\mathbf{y}_R}\} \leq i_R^{\max} \quad (9.32.d)$$

with $\{r_{i \rightarrow n}\}$ given by (9.29), $\epsilon \ll 1$ ensuring queue stability, and $\mathcal{U}_n(\rho_n)$ and $\mathcal{C}_n(P_n)$ selected to be concave and convex functions, respectively, representing the reward of rate ρ_n and the cost of power P_n . Unfortunately, even without the interference constraint (9.32.d), problem (9.32) is nonconvex. However, a successive convex approximation approach [113] can be employed to efficiently find to a Karush-Kuhn-Tucker optimal solution [106].

To illustrate the attractive features of the resultant routing protocol, consider the simple scenario depicted in Figure 9.11, where $N_r = 7$ CR nodes route packets generated at U_1 and U_2 to the destination U_8 . Two PU sources also transmit with power 0 dBW. The path loss coefficients are set to $G_0 = 0$, and $\alpha = 3.5$, and $m = 1$ is used for Nakagami- m fading. Log-normal shadowing is generated with mean 0 and standard deviation 6 dB. The maximum transmit-power for the CR system is $P_n^{\max} = 0$ dBW, the noise power 10^{-8} , and the SINR threshold $\bar{\Gamma}_n = -10$ dB. The interference threshold is set to -80 dBW, and the sum of exogenous rates was maximized. Figure 9.11a depicts the optimal routing probabilities obtained by solving (9.32). It can be seen that due to the presence of PU links, there is a tendency to route packets generated by U_2 through links $U_4 \rightarrow U_6$ and $U_6 \rightarrow U_7$, rather than choosing the shortest path $U_2 \rightarrow U_4 \rightarrow U_5 \rightarrow U_8$. Conversely, packets generated by U_1 are routed through U_3 and U_7 with high probability, which in this case coincides with the shortest path. As can be noticed from Figure 9.11b, links to and from U_4 and U_5 manifest lower decoding capability compared to links that are farther from the PU system. This is not only due to the detrimental effect of PU interference on

**FIGURE 9.11**

Cartography-enabled optimal statistical routing. (a) Routing probabilities. (b) Link reliabilities.

the SINRs, but also due to the fact that U_4 and U_5 are confined to use a lower transmit-power in order to protect PU receivers from harmful interference.

2.09.5 Conclusion

Spectrum sensing at the PHY and MAC layers of CR networks, as well as its cross-layer design and application have been overviewed in this tutorial. At the PHY layer, the basic sensing task was to detect the presence of PU transmitters, for which collaboration of multiple CRs was seen effective to alleviate the challenges due to fading and shadowing effects. It has been also argued that a holistic characterization of the RF environment in which the CR network operates is of great importance for efficient and quality-assuring network design and adaptation. Various signal processing and learning techniques were employed to develop RF cartography algorithms to capture the spatio-temporal-spectral RF environment. At the MAC layer, the key issue was to schedule the per-band sensing operations over the wide bandwidth using limited sensing resources, based on estimated traffic patterns and accumulated sensing history. Cross-layer issues in sensing aimed to strike an optimal trade-off between sensing accuracy and system objectives. The rich cognition of the operating environment obtained via sensing holds a great potential for robust and efficient network operation, which can be realized through cross-layer design.

Glossary

Cognitive radio	an intelligent radio that can learn and adapt to the environment
Primary user	a radio system that possesses an exclusive license to use a given spectrum band
Neyman-Pearson test	a hypothesis test that maximizes the detection probability while ensuring a given false alarm probability

Physical (PHY) layer	a layer in the networking protocol stack concerned with electrical or optical interface to the communication medium
Medium access control (MAC) layer	a layer in the networking protocol stack concerned with sharing a physical connection among multiple communication entities
Cross-layer design	a design approach for networking protocols that takes into account interactions among different layers in the protocol stack

Relevant Theory: Statistical Signal Processing

See Vol. 3, Chapter 6 Quickest Change Detection

References

- [1] S. Haykin, Cognitive radio: brain-empowered wireless communications, *IEEE J. Sel. Areas Commun.* 23 (2) (2005) 201–220.
- [2] J. Mitola III, G.Q. Maguire Jr. Cognitive radio: making software radios more personal, *IEEE Pers. Commun.* 6 (4) (1999) 13–18.
- [3] FCC Spectrum Policy Task Force, Report of the spectrum efficiency working group, November 2002, <<http://www.fcc.gov/sptf/reports.html>>.
- [4] Q. Zhao, B.M. Sadler, A survey of dynamic spectrum access, *IEEE Signal Process. Mag.* 24 (3) (2007) 79–89.
- [5] Z. Tian, G.B. Giannakis, Compressed sensing for wideband cognitive radios, in: Proceedings of the IEEE ICASSP Conference, vol. 4, Honolulu, HI, April 2007, pp. 1357–1360.
- [6] P. Pawelczak, K. Nolan, L. Doyle, S.W. Oh, D. Cabric, Cognitive radio: ten years of experimentation and development, *IEEE Commun. Mag.* 49 (3) (2011) 90–100.
- [7] R. Tandra, A. Sahai, SNR walls for signal detection, *IEEE J. Sel. Topics in Signal Process.* 2(1) (2008) 4–17.
- [8] J. Unnikrishnan, V.V. Veeravalli, Cooperative sensing for primary detection in cognitive radio. *IEEE J. Sel. Topics Signal Process.* 2 (1) (2008) 18–27.
- [9] G. Ganesan, Y. Li, Cooperative spectrum sensing in cognitive radio networks, in: Proceedings of the IEEE DySPAN Conference, Baltimore, MD, November 2005, pp. 137–143.
- [10] A. Ghasemi, E.S. Sousa, Spectrum sensing in cognitive radio networks: the cooperation-processing tradeoff, *Wireless Commun. Mobile Comput.* 7 (9) (2007) 1049–1060.
- [11] D. Cabric, S.M. Mishra, R.W. Brodersen, Implementation issues in spectrum sensing for cognitive radios, in: Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers, vol. 1, Pacific Grove, CA, March 2005, pp. 772–776.
- [12] J. Lunden, V. Koivunen, A. Huttunen, H.V. Poor, Collaborative cyclostationary spectrum sensing for cognitive radio systems, *IEEE Trans. Sig. Proc.* 57 (11) (2009) 4182–4195.
- [13] A.B.H. Alaya-Feki, S.B. Jemaa, B. Sayrac, P. Houze, E. Moulines, Informed spectrum usage in cognitive radio networks: interference cartography, in: Proceedings of the PIMRC Conference, Cannes, France, September 2008, pp. 1–5.
- [14] J.A. Bazerque, G.B. Giannakis, Distributed spectrum sensing for cognitive radio networks by exploiting sparsity, *IEEE Trans. Signal Process.* 58 (3) (2010) 1847–1862.
- [15] S.-J. Kim, E. Dall’Anese, G.B. Giannakis, Cooperative spectrum sensing for cognitive radios using Kruged Kalman filtering, *IEEE J. Sel. Topics Signal Process.* 5 (2011) 24–36.

- [16] S. Chaudhari, V. Koivunen, H.V. Poor, Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios, *IEEE Trans. Sig. Proc.* 57 (7) (2009) 2690–2700.
- [17] S.-J. Kim, G.B. Giannakis, Sequential and cooperative sensing for multi-channel cognitive radios, *IEEE Trans. Signal Process.* 58 (8) (2010) 4239–4253.
- [18] R. Fan, H. Jiang, Optimal multi-channel cooperative sensing in cognitive radio networks, *IEEE Trans. Wireless Commun.* 9 (3) (2010) 1128–1138.
- [19] Z. Quan, S. Cui, A.H. Sayed, H.V. Poor, Optimal multiband joint detection for spectrum sensing in cognitive radio networks, *IEEE Trans. Sig. Proc.* 57 (3) (2009) 1128–1140.
- [20] S. Ahmad, M. Liu, T. Javidi, Q. Zhao, B. Krishnamachari, Optimality of myopic sensing in multichannel opportunistic access, *IEEE Trans. Info. Theory* 55 (9) (2009) 4040–4050.
- [21] H.T. Cheng, W. Zhuang, Simple channel sensing order in cognitive radio networks, *IEEE J. Sel. Areas Commun.* 29 (4) (2011) 676–688.
- [22] Q. Zhao, L. Tong, A. Swami, Y. Chen, Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework, *IEEE J. Sel. Areas Commun.* 25 (3) (2007) 589–600.
- [23] H.-P. Shiang, M. van der Schaar, Distributed resource management in multihop cognitive radio networks for delay-sensitive transmission, *IEEE Trans. Veh. Technol.* 58 (2) (2009) 941–953.
- [24] Y.-C. Liang, Y. Zeng, E.C.Y. Peh, T. Hoang, Sensing-throughput tradeoff for cognitive radio networks, *IEEE Trans. Wireless Commun.* 7 (4) (2008) 1326–1337.
- [25] R. Wang, V.K.N. Lau, L. Lv, B. Chen, Joint cross-layer scheduling and spectrum sensing for OFDMA cognitive radio systems, *IEEE Trans. Wireless Commun.* 8 (5) (2009) 2410–2416.
- [26] C. Cordeiro, K. Challapali, D. Birru, N. Sai Shankar, IEEE 802.22: the first worldwide wireless standard based on cognitive radios, in: Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access, Networks, November 2005, pp. 328–337.
- [27] K.W. Choi, W.S. Jeon, D.G. Jeong, Sequential detection of cyclostationary signal for cognitive radio systems, *IEEE Trans. Wireless Commun.* 8 (9) (2009) 4480–4485.
- [28] F. Zeng, C. Li, Z. Tian, Distributed compressive spectrum sensing in cooperative multihop cognitive networks, *IEEE J. Sel. Topics Signal Process.* 5 (1) (2011) 37–48.
- [29] G. Ganesan, Ye Li, B. Bing, Shaoqian Li, Spatiotemporal sensing in cognitive radio networks, *IEEE J. Sel. Areas Commun.* 26 (1) (2008) 5–12.
- [30] J. Ma, G. Zhao, Y. Li, Soft combination and detection for cooperative spectrum sensing in cognitive radio networks, *IEEE Trans. Wireless Commun.* 7 (11) (2008) 4502–4507.
- [31] Z. Quan, S. Cui, A.H. Sayed, Optimal linear cooperation for spectrum sensing in cognitive radio networks, *IEEE J. Sel. Topics Sig. Proc.* 2 (1) (2008) 28–40.
- [32] W. Zhang, R. Mallik, K. Letaief, Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks, *IEEE Trans. Wireless Commun.* 8 (12) (2009) 5761–5766.
- [33] B. Wang, K.J.R. Liu, T.C. Clancy, Evolutionary cooperative spectrum sensing game: how to collaborate? *IEEE Trans. Commun.* 58 (3) (2010) 890–900.
- [34] S. Maleki, A. Pandharipande, G. Leus, Energy-efficient distributed spectrum sensing for cognitive sensor networks, *IEEE Sensors Journal* 11 (3) (2011) 565–573.
- [35] I.F. Akyildiz, B.F. Lo, R. Balakrishnan, Cooperative spectrum sensing in cognitive radio networks: a survey, *Elsevier Phys. Commun.* 4 (1) (2011) 40–62.
- [36] J. Wang, P. Urriza, Y. Han, D. Cabric, Weighted centroid localization algorithm: theoretical analysis and distributed implementation, *IEEE Trans. Wireless Commun.* 10 (10) (2011) 3403–3413.
- [37] S. Geirhofer, L. Tong, B.M. Sadler, A sensing-based cognitive coexistence method for interfering infrastructure and ad hoc systems, *Wireless Commun. and Mobile Comput.* 10 (1) (2010) 16–30.
- [38] E.J. Candès, Y. Plan, Near-ideal model selection by ℓ_1 minimization, *Ann. Stat.* 37 (5A) (2009) 2145–2177.

- [39] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, *SIAM J. Sci. Comput.* 20 (1) (1998) 33–61.
- [40] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. Royal Stat. Soc. Ser. B*, 58 (1) (1996) 267–288.
- [41] E. Dall’Anese, J.A. Bazerque, G.B. Giannakis, Group sparse Lasso for cognitive network sensing robust to model uncertainties and outliers, *Elsevier Physical Commun.* 5 (1) (2012) 161–172.
- [42] M. Yuan, Y. Lin, Model selection and estimation in regression with grouped variables, *J. Royal Stat. Soc.* 68 (2006) 49–67.
- [43] H. Zhu, G. Leus, G.B. Giannakis, Sparsity-cognizant total least-squares for perturbed compressive sampling, *IEEE Trans. Signal Process.* 59 (2011) 2002–2016.
- [44] J.A. Bazerque, G. Mateos, G.B. Giannakis, Group-Lasso on splines for spectrum cartography, *IEEE Trans. Signal Process.* 59 (10) (2011) 4648–4663.
- [45] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.
- [46] J. Duchon, *Splines Minimizing Rotation-Invariant Semi-norms in Sobolev Spaces*, Springer-Verlag, Berlin, 1977.
- [47] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, second ed., Springer-Verlag, New York, NY, 2009.
- [48] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.
- [49] R. Zhang, On peak versus average interference power constraints for protecting primary users in cognitive radio networks, *IEEE Trans. Wireless Commun.* 8 (4) (2009) 2112–2120.
- [50] K. Kumaran, S.E. Golowich, S. Borst, Correlated shadow-fading in wireless networks and its effect on call dropping, *Wireless Networks* 8 (1) (2002) 61–71.
- [51] E. Dall’Anese, S.-J. Kim, G.B. Giannakis, Channel gain map tracking via distributed Kriging, *IEEE Trans. Veh. Technol.* 60 (3) (2011) 1205–1211.
- [52] T.S. Rappaport, *Wireless Communications: Principles & Practice*, Prentice Hall, 1996.
- [53] G.L. Stüber, *Principles of Mobile Communication*, second ed., Kluwer Academic Publishers, Boston, MA, 2001.
- [54] A. Goldsmith, L.J. Greenstain, G.J. Foschini, Error statistics of real-time power measurements in cellular channels with multipath and shadowing, *IEEE Trans. Veh. Technol.* 43 (3) (1994) 439–446.
- [55] M. Gudmundson, Correlation model for shadow fading in mobile radio systems, *Elec. Lett.* 27 (23) (1991) 2145–2146.
- [56] F. Graziosi, F. Santucci, A general correlation model for shadow fading in mobile radio systems, *IEEE Commun. Lett.* 6 (3) (2002) 102–104.
- [57] Z. Wang, E.K. Tameh, A. Nix, Simulating correlated shadowing in mobile multihop relay/ad hoc networks, Technical Report IEEE C802.16j-06/060, IEEE 802.16 Broadband Wireless Access Working Group, July 2006.
- [58] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, A.J. Paulraj, Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels, *IEEE Trans. Veh. Technol.* 59 (2010) 2253–2265.
- [59] P. Agrawal, N. Patwari, Correlated link shadow fading in multi-hop wireless network, *IEEE Trans. Wireless Commun.* 8 (8) (2009) 4024–4036.
- [60] K.V. Mardia, C. Goodall, E.J. Redfern, F.J. Alonso, The Kriged Kalman filter, *Test* 7 (2) (Dec. 1998) 217–285.
- [61] C.K. Wikle, N. Cressie, A dimension-reduced approach to space-time Kalman filtering, *Biometrika* 86 (4) (1999) 815–829.
- [62] B.D. Ripley, *Spatial Statistics*, John Wiley & Sons, 1981.
- [63] E. Dall’Anese, S.-J. Kim, G.B. Giannakis, S. Pupolin, Power control for cognitive radio networks under channel uncertainty, *IEEE Trans. Wireless Commun.* 10 (10) (2011) 3541–3551.

- [64] X. Hong, C.-X. Wang, J. Thompson, Interference modeling of cognitive radio networks, in: Proceedings of the IEEE Vehicle Technology Conference, Singapore, May 2008, pp. 1851–1855.
- [65] V. Blaschke, T. Renk, F.K. Jondral, A cognitive radio receiver supporting wide-band sensing, in: Proceedings of the International Conference on Communication, Beijing, China, May 2008, pp. 499–503.
- [66] B. Le, T.W. Rondeau, J.H. Reed, C.W. Bostian, Analog-to-digital converters, *IEEE Signal Process. Mag.* 22 (6) (2005) 69–77.
- [67] W. Hu, D. Willkomm, M. Abusubaih, J. Gross, G. Vlantis, M. Gerla, A. Wolisz, Dynamic frequency hopping communities for efficient IEEE 802.22 operation, *IEEE Commun. Mag.* 45 (5) (2007) 80–87.
- [68] S. Sengupta, S. Brahma, M. Chatterjee, N. Sai Shankar, Enhancements to cognitive radio based IEEE 802.22 air-interface, in: Proceedings of the IEEE ICC Conference, June 2007, pp. 5155–5160.
- [69] D. Willkomm, M. Bohge, D. Hollos, J. Gross, A. Wolisz, Double hopping: a new approach for dynamic frequency hopping in cognitive radio networks, in: Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC), September 2008, pp. 1–6.
- [70] S. Yarkan, H. Arslan, Binary time series approach to spectrum prediction for cognitive radio, in: Proceedings of the IEEE Vehicle Technology Conference, October 2007, pp. 1563–1567.
- [71] M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, New York, 2006.
- [72] X. Li, S.A. Zekavat, Traffic pattern prediction and performance investigation for cognitive radio systems, in: Proceedings of the IEEE Wireless Communication and Networks Conference, April 2008, pp. 894–899.
- [73] P.J. Brockwell, R.A. Davis, *Introduction to Time Series and Forecasting*, second ed., Springer-Verlag, New York, 2006.
- [74] X. Long, X. Gan, Y. Xu, J. Liu, M. Tao, An estimation algorithm of channel state transition probabilities for cognitive radio systems. In Proceedings of the IEEE International Conference on Cognitive Radio Oriented Wireless Network and Communication (CrownCom), May 2008, pp. 1–4.
- [75] Q. Zhao, A. Swami, A decision-theoretic framework for opportunistic spectrum access, *IEEE Wireless Commun.* 14 (4) (2007) 14–20.
- [76] K. Liu, Q. Zhao, Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access, *IEEE Trans. Info. Theory* 56 (11) (2010) 5547–5567.
- [77] A.T. Hoang, Y.-C. Liang, D. Wong, Y. Zeng, R. Zhang, Opportunistic spectrum access for energy-constrained cognitive radios, *IEEE Trans. Wireless Commun.* 8 (3) (2009) 1206–1211.
- [78] Y. Chen, Q. Zhao, A. Swami, Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors, *IEEE Trans. Info. Theory* 54 (5) (2008) 2053–2071.
- [79] Q. Zhao, B. Krishnamachari, K. Liu, On myopic sensing for multi-channel opportunistic access: structure, optimality, and performance, *IEEE Trans. Wireless Commun.* 7 (12) (2008) 5431–5440.
- [80] J. Unnikrishnan, V.V. Veeravalli, Algorithms for dynamic spectrum access with learning for cognitive radio, *IEEE Trans. Signal Process.* 58 (2) (2010) 750–760.
- [81] J. Jia, Q. Zhang, X. Shen, HC-MAC: a hardware-constrained cognitive MAC for efficient spectrum management, *IEEE J. Sel. Areas Commun.* 26 (1) (2008) 106–117.
- [82] S. Geirhofer, L. Tong, B.M. Sadler, Dynamic spectrum access in the time domain: modeling and exploiting white space, *IEEE Commun. Mag.* 45 (5) (2007) 66–72.
- [83] Q. Zhao, S. Geirhofer, L. Tong, B.M. Sadler, Opportunistic spectrum access via periodic channel sensing, *IEEE Trans. Signal Process.* 56 (2) (2008) 785–796.
- [84] X. Li, Q. Zhao, X. Guan, L. Tong, Optimal cognitive access of Markovian channels under tight collision constraints, *IEEE J. Sel. Areas Commun.* 29 (4) (2011) 746–756.
- [85] H. Kim, K.G. Shin, Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks, *IEEE Trans. Mob. Comput.* 7 (5) (2008) 533–545.
- [86] H. Kim, K.G. Shin, Fast discovery of spectrum opportunities in cognitive radio networks, in: Prof. of the IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, October 2008, pp. 1–12.

- [87] S. Huang, X. Liu, Z. Ding, Optimal transmission strategies for dynamic spectrum access in cognitive radio networks, *IEEE Trans. Mob. Comput.* 8 (12) (2009) 1636–1648.
- [88] R. Fan, H. Jiang, Channel sensing-order setting in cognitive radio networks: a two-user case, *IEEE Trans. Veh. Technol.* 58 (9) (2009) 4997–5008.
- [89] L. Lai, H. El Gamal, H. Jiang, H.V. Poor, Cognitive medium access: exploration, exploitation, and competition, *IEEE Trans. Mob. Comput.* 10 (2) (2011) 239–253.
- [90] H. Jiang, L. Lai, R. Fan, H.V. Poor, Optimal selection of channel sensing order in cognitive radio, *IEEE Trans. Wireless Commun.* 8 (1) (2009) 297–307.
- [91] Q. Zhao, J. Ye, When to quit for a new job: quickest detection of spectrum opportunities in multiple channels, in: Proceedings of the MILCOM Conference, San Diego, CA, November 2008.
- [92] S.-J. Kim, G.B. Giannakis, Rate-optimal and reduced-complexity sequential sensing algorithms for cognitive OFDM radios, *EURASIP J. Adv. Signal Process.* (2009) p. 11 (Article ID 421540).
- [93] S.-J. Kim, G. Li, G.B. Giannakis, Multi-band cognitive radio spectrum sensing for quality-of-service traffic, *IEEE Trans. Wireless Commun.* 10 (10) (2011) 3506–3515.
- [94] N. Kundargi, A. Tewfik, Hierarchical sequential detection in the context of dynamic spectrum access for cognitive radios, in: Proceedings of the International Conference on Electrical, Circuits and System, Marrakech, Morocco, December 2007, pp. 514–517.
- [95] Y. Xin, H. Zhang, S. Rangarajan, SSCT: a simple sequential spectrum sensing scheme for cognitive radio, in: Proceedings of the GLOBECOM Conference, Honolulu, HI, November–December 2009, pp. 1–6.
- [96] G. Cheng, W. Liu, Y. Li, W. Cheng, Spectrum aware on-demand routing in cognitive radio networks, in: Proceedings of the IEEE DySPAN Conference, Dublin, Ireland, April 2007, pp. 571–574.
- [97] Y. Shi, Y.T. Hou, A distributed optimization algorithm for multi-hop cognitive radio networks, in: Proceedings of the IEEE INFOCOM Conference, Phoenix, AZ, April 2008, pp. 1292–1300.
- [98] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, second ed., Athena Scientific, Belmont, MA, 2000.
- [99] D.A. Casta non, Approximate dynamic programming for sensor management, in: Proceedings of the 36th Conference on Decision and Control, vol. 2, San Diego, CA, December 1997, pp. 1202–1207.
- [100] D. De Couto, D. Aguayo, J. Bicket, R. Morris, A high-throughput path metric for multi-hop wireless routing, in: Proceedings of the International ACM Conference Mobile Computing, Networking, San Diego, CA, September 2003, pp. 134–156.
- [101] C. Xin, L. Ma, C.-C. Shen, A path-centric channel assignment framework for cognitive radio wireless networks, *Mobile Netw. Appl.* (Kluwer) 13 (5) (2008) 463–476.
- [102] I. Pefkianakis, S.H.Y. Wong, S. Lu, SAMER: spectrum aware mesh routing in cognitive radio networks, in: Proceedings of IEEE DySPAN, Chicago, IL, October 2008.
- [103] A. Abbagnale, F. Cuomo, Gymkhana: a connectivity-based routing scheme for cognitive radio ad hoc networks, in: Proceedings of the IEEE INFOCOM, San Diego, CA, March 2010.
- [104] K.R. Chowdhury, I.F. Akyildiz, CRP: a routing protocol for cognitive radio ad hoc networks, *IEEE J. Sel. Areas Commun.* 29 (4) (2011) 794–802.
- [105] H. Khalife, S. Ahuja, N. Malouch, M. M. Krantz, Probabilistic path selection in opportunistic cognitive radio networks, in: Proceedings of the IEEE Globe Telecom Conference, New Orleans, LA, December 2010.
- [106] E. Dall’Anese, G.B. Giannakis, Statistical routing for multihop wireless cognitive networks, *IEEE J. Sel. Areas Commun.* 30 (10) (2012) 1983–1993.
- [107] A. Ribeiro, Z.-Q. Luo, N. Sidiropoulos, G.B. Giannakis, Modelling and optimization of stochastic routing for wireless multihop networks, in: Proceedings of the IEEE International Conference on Computer Communication, Anchorage, AK, May 2007, pp. 1748–1756.
- [108] A. Ribeiro, N. Sidiropoulos, G.B. Giannakis, Optimal distributed stochastic routing algorithms for wireless multihop networks, *IEEE Trans. Wireless Commun.* 7 (11) (2008) 4261–4272.

- [109] M. Haenggi, On routing in random rayleigh fading networks, *IEEE Trans. Wireless Commun.* 4 (4) (2005) 1553–1562.
- [110] L.F. Fenton, The sum of lognormal probability distributions in scatter transmission systems, *IRE Trans. Commun. Syst.* 8 (1) (1960) 57–67.
- [111] R. Rao, A. Ephremides, On the stability of interacting queues in a multi-access system, *IEEE Trans. Info. Theory* 34 (1988) 918–930.
- [112] R. Loynes, The stability of a queue with non-independent interarrival and service times, *Math. Proc. Cambridge Philosophical Society* 58 (1962) 497–520.
- [113] B.R. Marks, G.P. Wright, A general inner approximation algorithm for nonconvex mathematical programs, *Oper. Res.* 26 (4) (1978) 681–683.

Introduction to the Radar Signal Processing Section

10

Fulvio Gini

Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy

Signal processing for radar systems is a vast and fascinating discipline that concerns many different techniques and fields of application. The complete Elsevier Reference includes this section along with that on Array Signal Processing, which is intimately related to radar techniques. This section focuses on the most recent and exciting findings on the applications of signal processing techniques in radar systems, described in a tutorial manner, such that newcomers in the area can learn what is going on in this always in progress field of research. The basics and the new findings on array signal processing techniques are covered in the Array Signal Processing section edited by Mats Viberg.

The history of radar started more than one hundred years ago, in 1904, when Christian Hülsmeyer demonstrated the first experimental radar in Cologne, Germany. Later, in 1920, Guglielmo Marconi also observed the radio detection of targets in his experiments, but it was not until World War II that dynamic development of radar began. It has since then evolved into an indispensable all-weather, long-range sensor. Military and security applications have always been the main drivers of radar developments. However, later on radar has become a key sensor for civil applications including air, maritime, and ground traffic control and the guidance of aircraft and vehicles on airport surfaces.

According to its acronym RAdio Detection And Ranging, the classical radar mission is to detect and locate objects. With the advent of coherent pulse radar, velocity measurements have become possible by exploiting the Doppler effect. Today specialized radars measure elevation, aid in weather monitoring, and help with target classification. To find an overview of basic radar concepts and radar signal processing algorithms the interested reader can refer to [1].

One of the main problems in radar technology has always been the mitigation of echoes reverberating from the environment surrounding the radar, named clutter. Since the clutter is one of the most limiting factor for the performance of a radar system, we start this section with a chapter on the statistical modeling and analysis of radar clutter data. The chapter titled “Radar Clutter Modeling and Analysis,” co-authored by Maria S. Greco and Simon Watts, describes how sea, ground, and rain clutter processes have been modeled to design optimal radar receivers. A number of families of statistical models are described in the first part of the chapter that can be used to fit the observed amplitude statistics over a wide range of conditions, including the log-normal, the Weibull, and, especially, the compound-Gaussian model. The second part of the chapter describes the statistical analysis performed on different experimental sea and land clutter data to test the fit of theoretical models with real data using different radar parameters and environmental conditions. The limitations of theoretical models in some conditions are also discussed. Concerning the clutter power spectral density (PSD), many models have been proposed in the literature

in the last 50 years [2]. Greco and Watts summarize the most successful results concerning clutter PSD modeling by describing the Autoregressive (AR), Gaussian, power law (PL), and exponential spectral models and their fit with real data. A definite conclusion on this complex matter cannot be reached, since the study of clutter phenomenon is an always in progress research area. As radar systems improve, new features of clutter and target scattering may be observed and new analyses are then necessary to develop more accurate models of these characteristics and to improve the design of target detectors (see e.g., [3]).

A major advance, leading to a new quality of radar, is represented by the invention of array antennas. One of the main advantages of array antennas is that they provide, in principle, spatial sampling of the received wavefront, i.e., multichannel data acquisition [4]. This enables the use of different spatial processing techniques according to the application, e.g., pattern shaping to suppress interference in certain directions. Spatial and temporal processing can be made data dependent resulting in space-time adaptive processing (STAP). STAP is one of the key topics in modern coherent radar. The practical relevance of STAP lies in its ability to suppress clutter from airborne and space-based radar systems, where, depending on angle and range, clutter is spread in Doppler frequency. Chapter 12 of this section, “Space-Time Adaptive Processing for Radar,” is entirely devoted to STAP. William Melvin describes a number of key aspects of STAP and a number of approaches to design optimal detectors. Several major algorithms are summarized, including post-Doppler STAP, pre-Doppler STAP, and parametric STAP. The author then characterizes the performance of various methods in homogeneous clutter using the signal to interference plus noise ratio (SINR) loss as the performance metric. Then, he discusses the challenges of STAP application in heterogeneous/nonstationary clutter. One section of this chapter is devoted to the recent concept of knowledge-aided (KA), or knowledge-based (KB), sensor signal processing. A KB radar exploits methods and techniques of artificial intelligence and uses a-priori information collected from the environment to improve the performance of adaptive sensor systems. Examples of a-priori information are geographic maps, digital elevation models, weather data, car traffic data, intelligence, e.m. data base, GIS, etc. More details on knowledge-based radar can be found in [5,6]. Finally, Melvin summarizes several STAP data collection programs, the characteristics of the various radar systems, and the type of data collected.

A recent trend in multichannel radar development is represented by the multiple input–multiple output (MIMO) radars. They are formed by several spatially separated transmitting, receiving, or transmitting-receiving systems. The distinguished feature is represented by the capability of joint processing all the received target information. Two kinds of MIMO radars have been considered, the MIMO radars with colocated antennas and the MIMO radars with widely separated antennas. Most of the achievements about the first kind of configuration are described in [7]. As concerning the second type, sometime called Multistatic Radar System (MSRS), the theoretical foundations can be found in [8], both for active and passive radars. In Chapter 13, “MIMO Radar with Widely Separated Antennas—From Concepts to Designs,” Qian He, Yang Yang, and Rick S. Blum summarize and discuss a large range of issues related to the MIMO radar with widely separated antennas. Both coherent and noncoherent MIMO radar performance in target localization and velocity estimation are investigated. Topics of realistic values are covered, which include, for example, phase synchronization algorithm design for coherent MIMO radar and MIMO radar waveform design for extended targets. The performance differences between coherent and noncoherent MIMO radars in the application of joint target location and velocity estimation are investigated. Then, the authors derive the diversity gain for a MIMO radar system adopting the Neyman-Pearson detection. Three phase synchronization approaches for coherent MIMO

radar, which includes the master-slave closed-loop method, the round-trip approach, and the broadcast consensus based algorithm, are described and compared, and some issues that may arise in practice are discussed. Finally, some waveform design schemes for MIMO radar with widely separated antennas are described.

The transmitted waveform has a crucial role in determining the performance of a radar system. In the well-known book by Levanon and Mozeson [9], classical (non-adaptive) techniques for waveform design are described in depth. The more recent topics of waveform diversity and adaptive waveform design are instead the subject of a recent book [10]. In [Chapter 14](#), “Optimal Radar Waveform Design,” Joseph R. Guerci describes how adaptivity on transmission can be achieved through waveform diversity and how this may increase the detection and identification performance of a radar system. The chapter develops the basic theory of optimal transmit/receive design using a multi-input, multi-output (MIMO) formulation that can account for all potential degrees of freedom (DOFs) such as waveform (fast-time), angle, and polarization. Various applications and examples are provided to further illustrate the potential impact of joint transmit/receive adaptivity.

The scope of radar post-processing involves estimating target parameters, such as range, Direction of Arrival (DoA), and velocity, and form the tracks of all the detected targets. These parameters usually represent the necessary information to automatically determine the identity (class) of the targets. The following steps are typically the situation assessment and the response selection, according to the mission of the system. Situation Awareness is the crucial step for a reliable decision making process, which results in an outcome leading to the selection of a suitable course of action among several alternatives (see [\[11\]](#) and references therein). [Chapter 15](#), “Multitarget Multisensor Tracking,” co-authored by X. Chen, R. Tharmarasa and T. Kirubarajan, introduces various multisensor-multitarget tracking architectures. The authors discuss in detail estimators for spatial clutter intensity, filters for linear and nonlinear systems, algorithms for data associations and multitarget-tracking, techniques used in centralized and distributed track-to-track fusion. In addition, their quantitative and qualitative merits are discussed. Various combinations of these algorithms provide a complete tracking framework for multisensor networks with application to civilian as well as military problems. For example, the tracking and fusion techniques discussed in this chapter are applicable to fields like air traffic control, air/ground/maritime surveillance, mobile communication, transportation, video monitoring, and biomedical imaging/signal processing. The tracker performance evaluation, including its guiding principle and several measures of performance, is also discussed in this chapter. A challenging scenario with many closely spaced targets is used to compare several multitarget tracking algorithms.

Radar systems where the transmitter and receiver are colocated, or the separation between the transmitter and the receiver is small with respect to the target range, are called monostatic. Instead, when the transmitter and receiver are largely separated, so that scattering phenomenology differs from the monostatic case, we talk about bistatic radar systems. In [Chapter 16](#), “Passive Bistatic Radar,” Hugh Griffiths introduces the reader to those bistatic radar systems that make use of emissions from broadcast, communications or radionavigation transmitters rather than a dedicated, cooperative radar transmitter. Such systems have a number of potential advantages. The receiver is passive and so potentially undetectable. There are many illumination sources that can be used, many of them of high power and favorably sited. The receiver of passive bistatic radar (PBR) can often be rather simple and low cost, and there is no need for any license for the transmitter. However, since the waveforms are not explicitly designed for radar use they may be far from optimum for radar purposes. It is therefore necessary to understand

the effect of the waveform on the performance of the passive bistatic radar, so as to be able to choose the most appropriate illuminator, and to process the waveform in the optimal way. This chapter describes the properties of such sources, the processing techniques that are used, and the performance that may be achieved. Griffiths finally concludes that applications to which PBR systems may be put need careful thought. It is important to understand very thoroughly the relationship between the desired application and the performance requirements that follow, and the properties of the illuminator source that might be used, for example in terms of coverage, power level, bandwidth, integration time, and so on.

The multichannel system concept described in [Chapter 12](#) is also applicable in cases where the spatial samples are taken sequentially in time, e.g., during the flight of an aircraft. This leads to synthetically generated antenna apertures of very large dimension, resulting in extremely high resolution. Imaging radars that use such a synthetic aperture (SAR) differ greatly from real aperture radars. They have been developed since the 1970s as a new all-weather sensor for ground observation [1]. The main goal of SAR is to achieve high resolution. Challenges to focus the synthetic array include mitigation of sensor position errors, and handling of large bandwidth, among others. Combining both types of arrays, i.e., using an antenna array as a sensor of the synthetic array, results in new techniques to solve the focusing problem. In addition, multichannel SAR offers new features, particularly the ability to detect moving targets (GMTI, ground moving target indication), and interferometric SAR, which can provide 3-D images of the terrain. Multichannel SAR for airborne or space-based remote sensing, both for military and civilian applications, is one of the most dynamic developing radar technologies. Recently, radar imaging techniques have also been developed for providing high quality imagery of enclosed structures. Each remote sensing application area has driven different sensing modalities and imaging algorithm development based upon propagation characteristics, sensor positioning, and safety issues. For example, the recent technique of through-the-wall radar imaging (TWRI) makes use of electromagnetic waves below the S-band to penetrate through building wall materials. The indoor scene can be illuminated from each antenna, and be reconstructed using the data from the receive antennas. Due to the “see”-through ability, TWRI has attracted much attention in the last decade and has found a variety of important civilian and military applications.

In [Chapter 17](#), “Through-the-Wall Radar Imaging: Theory and Applications,” Moeness Amin and Fauzia Ahmad describe the recent algorithmic advances in TWRI. Firstly, they focus on the problem of mitigating wall clutter. Front wall reflections are often stronger than target reflections, and they tend to persist over a long duration of time. Therefore, weak and close by targets behind walls become obscured and invisible in the image. Approaches based on both electromagnetic modeling and signal processing are advocated to significantly mitigate the front wall clutter. Then, the authors present an approach to exploit the rich indoor multipath environment for improved behind-the-wall target detection and localization. A change detection approach to moving target indication for through-the-wall applications is also described. Both coherent and noncoherent change detection techniques are examined and their performance compared using real data collected in a semi-controlled laboratory environment. The final part of the chapter is devoted to the application of the emerging compressive sensing techniques to circumvent possible logistic difficulties in collecting measurements in time and space and provide fast data acquisition and scene reconstruction for moving target indication.

One fundamental application of modern multichannel air- and spaceborne SAR sensor is ground moving target indication. The aim of SAR GMTI is to detect moving targets on ground and to estimate their positions, velocities, and moving directions. Although originated in the military field, nowadays

GMTI also has gained relevance for civilian road traffic monitoring to ensure the mobility and to increase the safety of the road users. In [Chapter 18](#), titled “Multi-channel SAR for Ground Moving Target Indication,” Stefan V. Baumgartner and Gerhard Krieger provide a comprehensive tutorial for GMTI with SAR systems. The SAR principle is explained and a single- and multi-channel moving point target signal model is derived. The effects on SAR imagery caused by moving targets are addressed in detail. The knowledge about these effects is essential for understanding the different parameter estimation methods. The classical dual-channel techniques Along-Track Interferometry (ATI) and Displaced Phase Center Antenna (DPCA) and their limitations are discussed in detail, since they are currently of special importance for state-of-the-art spaceborne SAR-GMTI systems. A general GMTI processing chain is also presented and basic Doppler parameter estimation methods introduced. Finally an outlook to future trends, applications, and open problems is given.

The imaging mode when the radar is stationary and the target is moving, or when both platforms are moving, is called inverse synthetic aperture radar (ISAR). Although the SAR and ISAR imaging modes have quite a lot in common, there are some significant differences that mark a line between them. In [Chapter 19](#), titled “Introduction to Inverse Synthetic Aperture Radar,” Marco Martorella introduces the readers to the main concepts and algorithms relative to ISAR imaging. Specifically, the concept of high resolution applied to radar is used to introduce ISAR imaging. A model-based approach is proposed as a method to derive the ISAR processor. ISAR geometry and received signal modeling have been defined. Polarimetric and bistatic ISAR imaging are also discussed as they represent more recent advances in ISAR imaging that are opening the doors to the use of ISAR when polarimetric radars are employed or when the radar configuration is multistatic. Examples are provided in a variety of scenarios.

The classical SAR imaging allows high resolution capability in 2-D, i.e., azimuth and range. The resulting image represents only a “projection” along the elevation direction of the illuminated 3-D scene backscattering properties over the azimuth slant-range plane. In [Chapter 20](#), titled “SAR Interferometry and Tomography: Theory and Applications,” Gianfranco Fornaro and Vito Pascazio describe the principles of SAR interferometry (InSAR) and SAR tomography (Tomo-SAR) for 3-D imaging. Interferometric SAR provides a 3-D imaging of the sensed scene relying only on the use of the phase difference between the signal acquired in at least two passes. However, InSAR implicitly assumes the presence of only a single scattering mechanism, i.e., does not comply with a possible integration of the scattering along the elevation direction. SAR Tomography is the extension of SAR Interferometry to allow a full 3-D imaging. It simply synthesizes, as in the case of the azimuth direction where an array is digitally formed, an array also in the elevation direction by exploiting the spatial multibaseline acquisition. In this chapter, the authors describe the most innovative processing techniques, like permanent scatterer interferometry, polarimetric-interferometry and tomography, and some examples of their many applications, such as the monitoring of volcanoes, earthquakes, landslides, ice sheet motion, and anthropogenic sources such as ground pumping of water and oil. The Differential SAR Interferometric (DInSAR) technique is also described in detail. In this case, multitemporal data are acquired over repeat passes and processed to generate very accurate deformation time series and therefore to achieve a regular monitoring of the deformation of the observed scene. DInSAR techniques also mitigate most of the limitations of the standard single-interferogram approaches, such as temporal and geometric decorrelation and the atmospheric phase delay and, unlike conventional InSAR, they allow increasing the measurement accuracy from centimeter up to millimeter. Finally, a mention to the new technique known as Polarimetric SAR Tomography is reported. Polarimetry offers a high potential for estimating

physical characteristics of objects and environments. A detailed description of radar Polarimetry can be found in [Chapter 21](#), authored by Eric Pottier and Laurent Ferro-Famil.

[Chapter 21](#), “Radar Polarimetry Basics and Selected Earth Remote Sensing Applications,” introduces the readers to some characteristic techniques and applications of modern SAR Polarimetry. The first part presents some basics of SAR Polarimetry. The polarization of a wave is first described in terms of vector representations. Both coherent and incoherent polarimetric operators are then defined and their properties are investigated for different kinds of scatterers. Various polarimetric decomposition techniques are presented. Their necessity and usefulness are discussed as well as their domain of application. Finally a simple polarimetric classification scheme, based on both physical interpretation and statistical segmentation is described. The second section presents advanced decomposition techniques and their applications to the estimation of physical quantities. The last part presents selected topics in multidimensional polarimetric SAR signal processing. The dimensionality of polarimetric signal being rather low, the use of other kinds of diversity is often required to solve an estimation problem. This part shows how three different applications, Time-Frequency Representation (TFR) of complex environments, Pol-InSAR characterization of forested areas, and polarimetric SAR Tomography, are all based on the modeling of the correlation properties of polarimetric signals acquired from different positions. Finally, some current trends in SAR Polarimetry are discussed.

The last chapter, “Integrated Sensor Systems and Data Fusion for Homeland Protection,” is motivated by the evidence that in the last few years technologies for identification and authentication, border security, and controlled access to critical infrastructures have become a very important concern to prevent unexpected attacks. Both future military operations and civil needs require innovative powerful sensors and sensor configurations that are able to change dynamically their behavior to overcome with the changing priorities and requirement of the mission. Alfonso Farina, Luciana Ortenzi, Branko Ristic, and Alex Skvortsov describe the most important issues related to an integrated system formed by a network of sensors, not necessarily all radar sensors, even if the radar is still the fundamental sensor in an integrated surveillance system. In particular, they address the application of data and information fusion to the design of integrated systems in the Homeland Protection (HP) domain. Data fusion is about combining, or fusing, data from different sources of information to provide knowledge that is not evident from individual sources. It aims to enhance situation awareness and decision making through the combination of information/data obtained by networks of homogeneous and/or heterogeneous sensors. The systems are integrated, to mean that it is generally not sufficient to study each subsystem in isolation, they are different in purpose and require a multidisciplinary approach for their design and analysis. In this chapter these aspects are investigated in depth for networks, respectively, of homogeneous and heterogeneous sensors with the description of real study cases applied to real world problems of homeland protection. In particular the possibility of netting different sensors operating with different characteristics of domain, coverage, frequency, and resolution allows a multiscale approach. This approach is particularly suitable for the surveillance of wide areas such as national borders or critical strategic regions.

What lies ahead? It comes as no surprise that we have yet to confront numerous fascinating technical challenges. These are potential R&D topics for the next editions of this encyclopedia. Some of them have only been “touched” in these chapters, e.g., advanced radar detection techniques for spiky non-Gaussian clutter, knowledge-based radar, cognitive radar, application of compressive sensing in radar, ground penetrating radar, automotive radar, etc. Until then, enjoy this outstanding collection of chapters.

With gratitude I want to acknowledge the extremely valuable work invested by the authors. They have provided highly professional and timely contributions. I also want to express my sincere gratitude to the referees who carefully read and reviewed the submitted manuscripts, thereby ensuring a high standard for this section. The referees for the chapters of this section are: Chris Baker, Alessio Benavoli, Paolo Braca, Victor Chen, Stefano Coraluppi, Antonio De Maio, Jean Dezert, Joachim Ender, Stefano Fortunati, Scott Hensley, Richard Klemm, Michael Leigsnering, Regis C.P. Marques, Ram M. Narayanan, Willie Nel, Jean-Philippe Ovarlez, Riccardo Paladini, Matteo Pardini, Fred Posner, Branko Ristic, Luke Rosenberg, Greg Showman, Francesco Soldovieri, Pietro Stinco, Keith Ward, Michael Wicks.

References

- [1] Principles of Modern Radar: Basic Principles, M.A. Richards, J.A. Scheer, W.A. Holm (Eds.), SciTech Publishing, Raleigh, NC, 2010.
- [2] J.B. Billingsley, Low-Angle Radar Land Clutter: Measurements and Empirical Models, William Andrew Publishing, Inc., Norwich, NY, 2002.
- [3] K.J. Sangston, F. Gini, M.V. Greco, A. Farina, Structures for optimal and suboptimal coherent radar detection in compound gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 35 (2) (1999) 445–458.
- [4] A. Farina, Antenna-Based Signal Processing Techniques for Radar Systems, Artech House Inc., Norwood, MA, 1992.
- [5] F. Gini, M. Rangaswamy (Eds.), Knowledge Based Radar Detection, Tracking and Classification, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.
- [6] J.R. Guerci, Cognitive Radar: The Knowledge-Aided Fully Adaptive Approach, Artech House, Norwood, MA, 2010.
- [7] J. Li, P. Stoica (Eds.), MIMO Signal Processing, John Wiley & Sons, Inc., 2009.
- [8] V. Chernyak, Fundamentals of Multisite Radar Systems, Multistatic Radars and Multiradar Systems, Gordon and Breach Science Publishers, 1998 (Revised and Extended Translation from Russian Edition, 1993).
- [9] N. Levanon, E. Mozeson, Radar Signals, John Wiley & Sons (Interscience Div.), New York, 2004.
- [10] F. Gini, A. De Maio, L.K. Patton (Eds.), Waveform Design and Diversity for Advanced Radar Systems, IET, Radar Sonar and Navigation Series, vol. 22, 2012.
- [11] S. Giompapa, F. Gini, A. Farina, A. Graziano, R. Croci, R. Di Stefano, Maritime border control multisensor system, *IEEE Aerosp. Electron. Syst. Mag.* 24 (8) (2009) 9–15.

Radar Clutter Modeling and Analysis

11

Maria S. Greco * and Simon Watts †

*Department of Ingegneria dell'Informazione, University of Pisa, Pisa, Italy

†Thales UK, Crawley, West Sussex, UK

2.11.1 Introduction

Radar systems operating in an open environment will receive returns from many sources. In addition to reflections from objects of interest, usually known as targets, the radar signal will include backscatter from the environment and other unwanted objects. These unwanted returns are usually called radar clutter.

Clutter signals that affect the radar performance are typically categorized in terms of backscatter from the land, the sea and the atmosphere (particularly from precipitation). Other objects may also provide unwanted reflections that can be called clutter. These include birds and insects, dust and man-made objects such as buildings, pylons, roads and so on. In the field of electronic warfare, the characteristics of chaff [1] may also be described in the same way as some types of clutter. In the early days of radar, unexplained backscatter signals were sometimes called “angels.” These are now mostly understood as atmospheric effects or returns from birds. For example, so called “ring angels,” which appeared like circular ripples on a pond when seen on a radar display, were found to be from starlings setting out from their roosts in waves [2].

A radar system is required to process the returns from targets in the presence of unwanted clutter (in addition to thermal noise that is always present in a radar receiver). The radar will usually need to detect the presence of a target and its position (typically at least its range and bearing from the radar) and perhaps to track it, if it is moving. The radar may also need to distinguish between different types of targets, including target classification and recognition [3]. In order to develop suitable processing algorithms for these tasks, the radar designer needs to be able to characterize clutter returns, in order to distinguish them from those of targets.

The characteristics of clutter are usually captured in terms of mathematical and numerical models. These models are designed to describe the various aspects of clutter that affect the design and use of radar systems. Models of clutter are used in all phases of the design cycle [4], including the following activities:

- Requirements definition.
- Predicting likely radar performance in different conditions.
- Radar system and signal processing algorithm design.
- Performance assessment and acceptance procedures for radar procurement.

- In-service tactics and training.
- In-service upgrades.

A mathematical model only has full value to a radar designer if it can be related to characteristics actually observed in a real radar. The fidelity required will depend on the application. High-level simulations may only require simple models, while models used for signal processing algorithm development may need to be very detailed. In particular, a model needs to be able to reflect the specific conditions in which the radar is being designed to operate. So, the parameter values of a mathematical model must be related to environmental conditions (wind speed, wave height, rainfall rate, etc.) and terrain type (land cultural features, rain or snow, etc.).

Throughout the following discussions of clutter modeling, the purposes and scope of the various models must be clearly understood. While many of the models described here can represent what is seen in a real radar with considerable fidelity, models are rarely precisely the same as real life. Models may be used to assist the design and assessment of radar systems, but the radar must still be able to deal with variations of characteristics that may be outside the average levels predicted by a model. For example, there are models that predict the average backscatter levels from the sea surface, as a function of the wind speed and direction, sea state (wave height), radar frequency, grazing angle and so on. These models predict very well the range of values that may be encountered over a wide range of conditions. However, they cannot reliably be used to predict the exact levels of backscatter on a given trial, not least because of the extreme difficulty of accurately measuring the prevailing conditions in the trial.

The study of clutter models is a continuing research topic. As radar systems improve, they are able to undertake more detailed analysis of the returns from targets and clutter. This in turn demands more detailed models of the returns, for example to characterize second order effects that had previously been ignored. The ability of radars to use very wide pulse bandwidths (providing finer range resolution), higher stability waveforms (allowing more precise analysis of Doppler shifts), polarization diversity and so on, all require improved clutter models to support their development. For example, the current research and gaps in our understanding of sea clutter are reviewed in [5], where it is noted that improved understanding is needed for the modeling of Doppler spectra. Previous radars have not routinely used coherent models for detecting small targets on the sea surface but as technology improves and operational requirements change, it is likely that pulse Doppler modes will be increasingly used. In addition, the use of very wide waveform bandwidths means that often the returns from sea clutter appear very spiky, with occasional large amplitude excursions against a lower clutter background. These sea clutter spikes need to be characterized to distinguish them from small targets.

This tutorial describes the various ways in which clutter is modeled and introduces some of the ways in which models are used in radar design. Section 2.11.2 introduces the characteristics of clutter as observed by radars and the methods used to characterize them. While some progress can be made in predicting clutter characteristics from the theory of electromagnetic scattering from rough surfaces, most of the models used by radar designers are empirical, with mathematical models designed to fit observed characteristics. Section 2.11.2 also introduces the empirical models widely used by radar designers, including the methods used to fit observed data to models. Section 2.11.3 provides some examples of the methods applied to develop models from the analysis of recorded data. One of the uses of these models is in the simulation of clutter returns for use in Monte Carlo computer simulation and for stimulating real hardware. Various methods for clutter simulation are therefore introduced in

Section 2.11.4. Section 2.11.5 then summarizes how the models are applied in radar design and analysis, including their use for performance prediction, the design of detection processing algorithms and in the specification and measurement of radar performance in the radar procurement process.

2.11.2 Clutter modeling

2.11.2.1 Generic clutter characteristics

There are various features of clutter that are of interest to a radar designer. These are usually characterized using the following types of model:

- The area reflectivity (normalized radar cross section) σ^0 for spatially distributed surface clutter.
- The volume reflectivity, η , for volume distributed clutter, such as precipitation and chaff.
- The amplitude distribution of clutter returns.
- The Doppler spectrum of returns.
- The spatial variation of clutter characteristics.
- Polarization characteristics (the polarization scattering matrix).
- Discrete features (sea clutter spikes, discrete land clutter features, etc.).

The mathematical models discussed here may include one or more of these features. The methodologies for these models are developed below. Specific empirical models for different types of clutter are described in Sections 2.11.2.2–2.11.2.4.

2.11.2.1.1 Normalized clutter reflectivity, σ^0

A perfectly smooth and flat conducting surface will act as a mirror, producing a coherent forward reflection, with the angle of incidence equal to the angle of reflection. However, if the surface has some roughness, the forward scatter component (called coherent or specular reflection) is reduced by diffuse, non-coherent scattering in other directions. For monostatic radar, clutter is the diffuse backscatter in the direction towards the radar. This is illustrated in Figure 11.1.

The magnitude of the backscattered signal is characterized by the normalized radar cross-section σ^0 . At some instance during the propagation of the pulse, a pulsed radar will illuminate a patch on the surface, defined (for low grazing angles) to a first order by the pulse length, the antenna azimuth beamwidth and the local grazing angle. The backscatter from land or sea is then modeled assuming multiple scatterers distributed spatially uniformly over this clutter patch. This is illustrated in Figure 11.2.

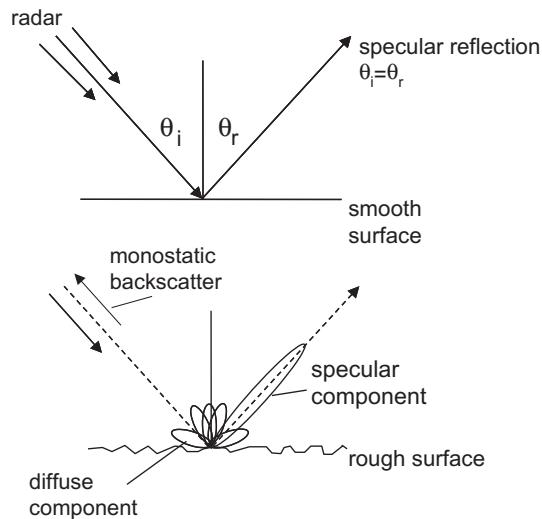
The normalized clutter reflectivity, σ^0 , is defined as the total RCS, σ , of the scatterers in the illuminated patch, normalized by the area, A_c , of the patch:

$$\sigma^0 = \sigma / A_c. \quad (11.1)$$

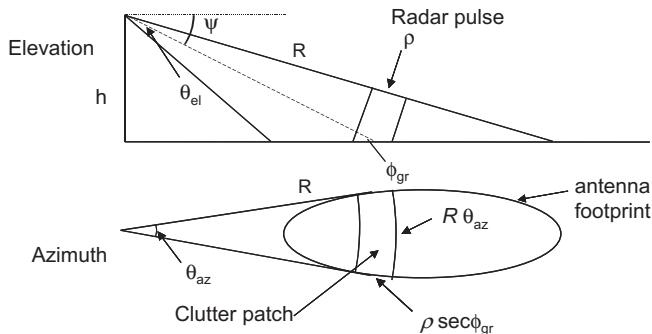
σ^0 is usually defined in units of dB m²/m² (dB relative to 1 m² radar cross-section, per m² of area). Referring to Figure 11.2, the area of the clutter patch is given by

$$A_c = \alpha \rho R \theta_{az} \sec(\phi_{gr}), \quad (11.2)$$

where θ_{az} is the antenna azimuth beamwidth and ϕ_{gr} is the local grazing angle. The range resolution, ρ , is related to the radar pulse bandwidth, B , by $\rho = \frac{c}{2B}$, where c is the velocity of light. The factor α

**FIGURE 11.1**

Reflection from smooth and rough surfaces.

**FIGURE 11.2**

Clutter illuminated patch size.

accounts for the actual compressed pulse shape and the azimuth beamshape, including the range and azimuth sidelobes. For a rectangular shaped pulse and beamshape, $\alpha = 1$, while for a Gaussian-shaped beam and rectangular pulse, $\alpha = 0.753$.

The grazing angle is defined in terms of the average local surface. For a nominally average flat surface, such as the sea, the grazing angle can be defined in terms of the radar height and propagation over a curved earth. In this case, the grazing angle, ϕ_{gr} , can be written as:

$$\phi_{gr} = \sin^{-1} \left(\frac{h}{R} + \frac{h^2}{2r_e R} - \frac{R}{2r_e} \right), \quad (11.3)$$

where h is the height (altitude) of the radar, r_e is the effective Earth's radius and R is the slant range. For routine calculations at sea level, $r_e \approx 4/3r$, where r is the true earth radius. This allows for the effects of atmospheric refraction for a typical refractive index profile. The actual effective grazing angle will depend on the local propagation and may be greatly changed under conditions of anomalous propagation, such as a surface ducts [2]. Over land, the local grazing angle will often be dominated by the terrain elevation and slope, which must be determined for specific cases.

The above expressions apply at low grazing angles, when the illuminated patch is defined by the azimuth beamwidth and pulse length. At high grazing angles, or for low bandwidth radars, the illuminated patch may be defined by the antenna azimuth and elevation beamwidths. Care should be taken if the grazing angle varies significantly over the illuminated patch, as σ^0 will not be constant.

2.11.2.1.2 Clutter volume reflectivity

A similar approach is taken for volume scattering. This is defined in terms of the volume reflectivity, η :

$$\eta = \sigma/V_c \quad (11.4)$$

with units of $\text{dB m}^2/\text{m}^3$.

The illuminated volume, V_c , is given approximately by:

$$V_c = \alpha\rho R^2 \theta_{az} \theta_{el}, \quad (11.5)$$

where θ_{el} is the one-way 3 dB elevation beamwidth. This is illustrated in Figure 11.3.

Clearly, this expression for the illuminated volume assumes that the volume scatterers fully fill the antenna beam and pulse length at a given range. If this is not the case, appropriate corrections must be made. For example, a volume search radar may have a narrow azimuth beam and a broad elevation beam. When illuminating rain, the rain ceiling may subtend a smaller angle than the upper edge of the beam.

2.11.2.1.3 Amplitude statistics

As illustrated in Figures 11.2 and 11.3, the return from clutter is usually assumed to comprise the backscatter from multiple scatterers, uniformly spatially distributed over an area or volume.

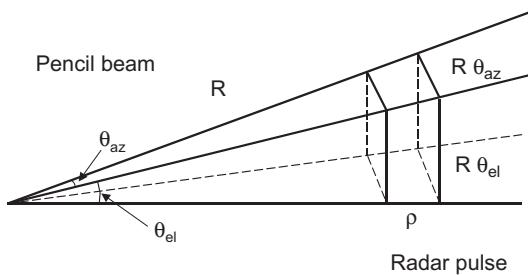


FIGURE 11.3

Volume clutter illuminated cell size.

The scattered field, y , can be written as the vector sum from N random scatterers:

$$y = \sum_N \sqrt{\sigma_i} \exp[j\phi_i], \quad (11.6)$$

where σ_i is the radar cross-section of a single scatterer, ϕ_i is a phase term (related to the reflection coefficient and the relative range of each scatterer from the radar). Provided $N \gg 1$, the probability density function of the real and imaginary parts of y will be Gaussian, through the application of the central limit theorem (CLT). The corresponding PDFs for the envelope and intensity of the signal will be:

$$\begin{aligned} p(r) &= \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right]; \quad 0 \leq r \leq \infty, \\ p(z) &= \frac{1}{2\sigma^2} \exp\left[-\frac{z}{2\sigma^2}\right]; \quad 0 \leq z \leq \infty, \end{aligned} \quad (11.7)$$

where r is $|y|$, $z = r^2$, $E[z] = 2\sigma^2$.

This representation of clutter is applicable to spatially uniform clutter observed with low-resolution radars (i.e., with a large clutter patch, so $N \gg 1$). This may apply to sea clutter, observed at high grazing angles and low resolution, or to, say, large flat areas of monoculture on land, such as woods or fields.

In many cases this representation is too simplistic. For example, the clutter mean intensity, $2\sigma^2$ above, may vary from one clutter cell to another, even though each cell still comprises multiple scatterers as in (11.6). Under these circumstances, the overall PDF will no longer have simple Gaussian statistics. Figure 11.4 illustrates the difference between returns with a Rayleigh PDF compared with clutter with strongly non-Gaussian statistics, modeled by a K distribution PDF, which is discussed below. A variation

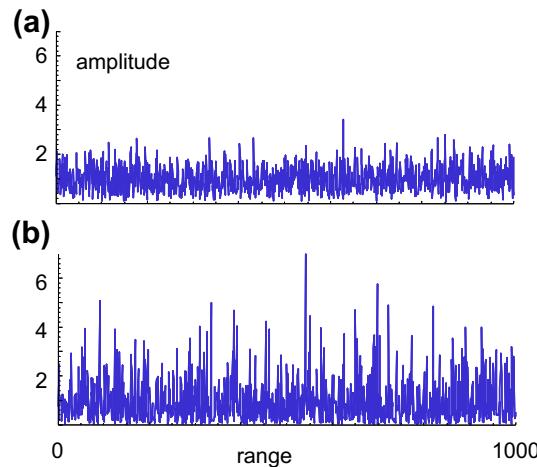


FIGURE 11.4

Envelope of uncorrelated signals versus time; (a) Rayleigh noise, mean level 1; (b) K distributed clutter, mean level 1, $v = 0.5$.

of the underlying mean intensity can be incorporated into the overall amplitude statistics by extending the simple model for PDF to include a dependence on the local mean intensity, which we shall call τ . Now:

$$p(r|\tau) = \frac{2r}{\tau} \exp\left[-\frac{r^2}{\tau}\right]; \quad 0 \leq r \leq \infty. \quad (11.8)$$

The mean level τ is itself a random variable with PDF $p(\tau)$, so that the overall PDF can be written

$$p(r) = \int_0^\infty p(r|\tau)p(\tau)d\tau; \quad 0 \leq r \leq \infty. \quad (11.9)$$

This way we obtain the so-called compound-Gaussian model. The use of this type of model to represent radar sea clutter was originally described in [6–8]. The compound K distribution form of the model, discussed below, was originally formulated by Ward et al. [9, 10]. There is further discussion in [9–21], and references therein. According to this model, each sample of the complex envelope of the sea clutter process is the product of two random variables: the texture and the speckle and can be represented as $y = \sqrt{\tau}x$. The term $x = x_I + jx_Q$ represents a stationary complex Gaussian process, called *speckle*, which accounts for local backscattering; x_I e x_Q are the in-phase and quadrature components of the speckle complex envelope x . They satisfy the property $E\{x_I\} = E\{x_Q\} = 0$ and $E\{x_I^2\} = E\{x_Q^2\} = 1/2$, so that $E\{|x|^2\} = 1$, i.e., the speckle complex samples have unit power. The factor τ is a non-negative real random process, called *texture* that, as said, models the local clutter power.

The compound-Gaussian model can be derived also as an extension of the CLT, allowing the number of scatterers N in Eq. (11.6) to be a random variable [16, 22]. In the particular case in which the number of scatterers is distributed following a negative binomial PDF, the texture can be shown to be a Gamma-distributed random variable (r.v.), and the amplitude a K-distributed r.v. (see Eq. (11.10)).

The compound-Gaussian model counts among its particular cases some families of distribution that are very popular in clutter modeling. The analytical expressions for these PDFs and their moments $m_R(n) = E\{r^n\}$ are reported below, where $r = |y|$ denotes the clutter amplitude.

K-model (K):

Replacing in Eq. (11.9) the generic $p(\tau)$ with the Gamma PDF $p(\tau) = \frac{1}{\Gamma(v)} \left(\frac{\nu}{\mu}\right)^v \tau^{v-1} \exp\left(-\frac{\nu}{\mu}\tau\right) u(\tau)$, we obtain

$$\text{PDF: } p_R(r) = \frac{\sqrt{2\nu/\mu}}{2^{v-1}\Gamma(v)} \left(\sqrt{\frac{2\nu}{\mu}}r\right)^v K_{v-1} \left(\sqrt{\frac{2\nu}{\mu}}r\right) u(r), \quad (11.10)$$

$$\text{moments: } m_R(n) = \left(\frac{2\mu}{\nu}\right)^{\frac{n}{2}} \frac{\Gamma\left(\nu + \frac{n}{2}\right)\Gamma\left(1 + \frac{n}{2}\right)}{\Gamma(v)}, \quad (11.11)$$

where $\Gamma(\cdot)$ is the gamma function, $K_{v-1}(\cdot)$ is the modified Bessel function of the second kind, of order $v - 1$, ν is the *shape parameter*, and μ is the *mean*.

Generalized K model with lognormal texture (LNT):

$$\text{PDF: } p_R(r) = \frac{r}{\sqrt{2\pi\sigma^2}} \int_0^\infty \frac{2}{\tau^2} \exp\left[-\frac{r^2}{\tau} - \frac{1}{2\sigma^2} \left[\ln\left(\frac{\tau}{2m}\right)\right]^2\right] d\tau, \quad (11.12)$$

$$\text{moments: } m_R(n) = (2m)^{\frac{n}{2}} \Gamma\left(1 + \frac{n}{2}\right) \exp\left(\frac{n^2 \sigma^2}{8}\right), \quad (11.13)$$

where σ is the *shape parameter*, and m is the *scale parameter*. This model can be obtained with the lognormal texture PDF $p(\tau) = \frac{1}{\tau \sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left[\left(\ln \tau - \ln 2m\right)^2\right]\right) u(\tau)$.

Generalized K model with generalized Gamma texture (GK):

Putting in Eq. (11.9) the generalized Gamma texture PDF $p(\tau) = \frac{\nu b}{\mu \Gamma(\nu)} \left(\frac{\tau \nu}{\mu}\right)^{\nu b-1} \exp\left[-\left(\frac{\nu}{\mu}\tau\right)^b\right] u(\tau)$ we obtain

$$\text{PDF: } p_R(r) = \frac{2br}{\Gamma(\nu)} \left(\frac{\nu}{\mu}\right)^{\nu b} \int_0^\infty \tau^{\nu b-2} \exp\left[-\frac{r^2}{\tau} - \left(\frac{\nu}{\mu}\tau\right)^b\right] d\tau, \quad (11.14)$$

$$\text{moments: } m_R(n) = \left(\frac{\mu}{\nu}\right)^{\frac{n}{2}} \frac{\Gamma\left(\nu + \frac{n}{2b}\right) \Gamma\left(1 + \frac{n}{2}\right)}{\Gamma(\nu)}, \quad (11.15)$$

Weibull model (W):

$$\text{PDF: } p_R(r) = \frac{c}{b^c} r^{c-1} \exp[-(r/b)^c] u(r), \quad (11.16)$$

$$\text{moments: } m_R(n) = b^n \Gamma\left(\frac{n}{c} + 1\right), \quad (11.17)$$

where c is the *shape parameter* and b is the *scale parameter*. The Rayleigh PDF is a particular case of the Weibull PDF for $c = 2$ [23]. Unfortunately, for the Weibull distribution, the PDF of the texture does not have a closed form and it is a compound-Gaussian model only for $c < 2$ [12].

Other sources of variability in clutter give rise to non-Gaussian statistics. While compound distributions can give some insight into the physical model underlying the non-Gaussian behavior, it is often sufficient to find an empirical fit of the overall amplitude statistics to a generalized PDF. Other popular distributions include the lognormal model, which does not belong to the compound-Gaussian family. The expressions of PDF and the moments are given below.

Lognormal model (LN):

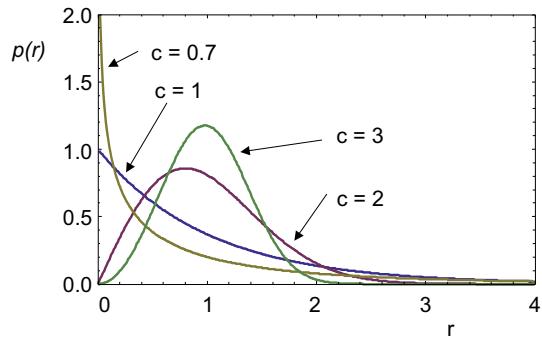
$$\text{PDF: } p_R(r) = \frac{1}{r \sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \left[\left(\ln r - \ln \delta\right)^2\right]\right) u(r), \quad (11.18)$$

$$\text{moments: } m_R(n) = \delta^n \exp\left[n^2 \sigma^2 / 2\right], \quad (11.19)$$

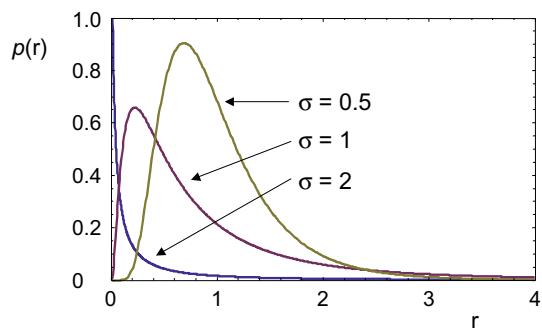
where σ is the *shape parameter*, and δ is the *scale parameter* and $u(r)$ the unit step function. Unfortunately, the LN model does not satisfy any of the compound-Gaussian properties [12].

The non-Gaussian PDFs K, W, and LN each have two parameters, a shape and a scale parameter, that can be adjusted to fit the observed data. The GK, conversely, has three parameters. Figures 11.5–11.7 illustrate examples of PDFs from different families.

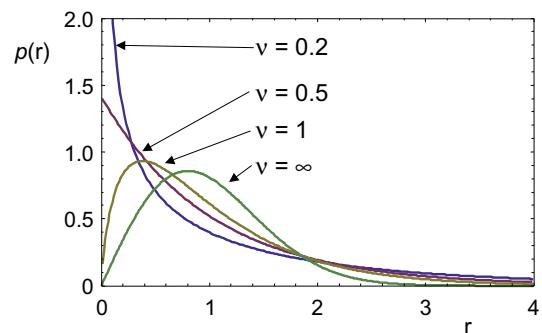
The K and Weibull distributions are very similar, both including the Rayleigh distribution as part of their family ($c = 2$ in the Weibull distribution and $\nu = \infty$ in the K distribution) and they are often used for sea clutter modeling. The lognormal distribution gives PDFs with much longer tails (i.e., a

**FIGURE 11.5**

Weibull distribution, mean power = 1.

**FIGURE 11.6**

Lognormal distribution, mean power = 1, m = 0.

**FIGURE 11.7**

K distribution, normalized to mean; mean power = 1.

higher probability of achieving larger amplitudes relative to the mean amplitude). This is often used for modeling land clutter, when the presence of large discrete scatterers can give rise to long tails if they are included in distributed clutter. A model that explicitly includes discrete spikes is the KA distribution, which is described in detail in [24,25].

2.11.2.1.4 Doppler spectrum

The simple model for clutter scattering given by (11.6), does not include any variation over time. If the individual clutter scatterers are moving radially with respect to the radar, the phase will vary with time, so that

$$y(t) = \sum_N \sqrt{\sigma_i} \exp[j\phi_i(t)]. \quad (11.20)$$

Assuming random motion of the scatterers, the temporal variations of the return must be described in terms of the autocorrelation function, ACF:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_{-T}^T y(t)y(t+\tau) dt \right|. \quad (11.21)$$

The power spectrum of the returns can be related to the ACF from its Fourier transform (Weiner Khintchine theorem):

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) \exp[-j\omega\tau] d\tau, \quad (11.22)$$

where ω is the Doppler radian frequency and $\omega = 2\pi f$, where f is the Doppler frequency.

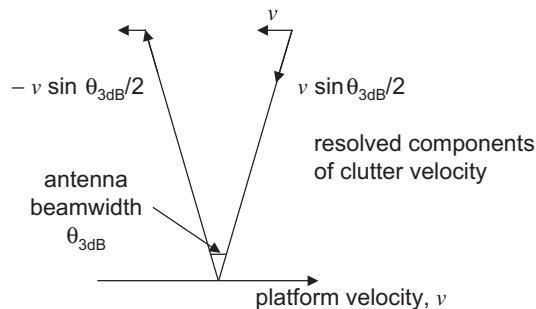
The Doppler power spectrum (power spectral density PSD) is often modeled as having a Gaussian shape:

$$S_G(f) = S_{0G} \exp \left[-\frac{(f - m_f)^2}{2\sigma_f^2} \right]. \quad (11.23)$$

This is usually a mathematical convenience rather than any attempt at realism. Often the Doppler spectrum will be strongly asymmetric and the mean Doppler shift, m_f , may not be zero. Clearly for land clutter m_f is usually zero, but for rain and sea clutter in general $m_f \neq 0$ and will be dependent on the wind speed and direction. Moreover, for sea clutter, as first reported by Pidgeon at C-band [26], and in X-band [27], the sea spectrum exhibits different peaks in the *HH* and *VV* polarizations [28,25,10]. Generally the spectral component corresponding to the lower frequency peak relative to the *VV* polarization, associated with the Bragg scattering component, is well described by the Gaussian function (11.23). On the contrary, the *HH* polarization is characterized by a higher frequency peak in the spectrum, maybe owing to the scattering from fast moving (faster than Bragg scatterers) and short-life scatterers. In this polarization the clutter PSD is well described by the Lorentzian function (autoregressive model of order 1)

$$S_L(f) = \frac{S_{0L}}{(f - m_f)^2 + k}, \quad (11.24)$$

where the constant k depends on the mean lifetime of scatterers.

**FIGURE 11.8**

Resolved component of clutter velocity across antenna beam, due to platform motion.

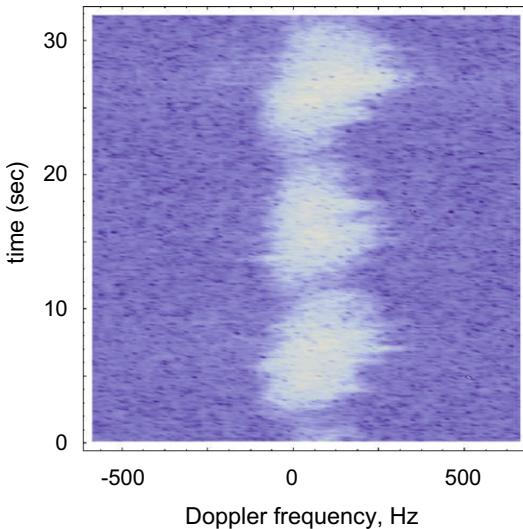
In some practical situations the above-mentioned two models do not suffice to fit the real spectra shape, especially for the *HH* polarization. In [28] the authors consider first the scattering from fast-to-intermediate scatterers (e.g., bound-Bragg waves, etc.) whose PSD is characterized by a convolution of the Gaussian and Lorentzian profiles, resulting in the Voigtian function, and then a linear combination of these three models to improve the fitting.

Another simple model often used for both sea and land clutter PSD is the autoregressive (AR) one. The rationale for adopting AR models for the radar echoes is to have a highly parameterized model with a minimum number of parameters that can be easily estimated. Some examples can be found in [29,30].

For moving platforms, the antenna motion with respect to the clutter will also modify the Doppler spectrum. This is illustrated in Figure 11.8 for a side looking antenna, with one-way 3 dB beamwidth θ_{3dB} and platform velocity v . If the antenna has a Gaussian shape, the combined Doppler spectrum due to platform motion and internal clutter motion will be approximately:

$$S(f) = S_0 \exp \left[-\frac{f^2}{2(0.6v\theta_{3dB}/\lambda)^2 + 2\sigma_f^2} \right]. \quad (11.25)$$

Figure 11.9 shows an example of the Doppler spectrum of sea clutter as function of time, derived from radar data collected by CSIR [31]. The radar was vertically polarized with a frequency of 9 GHz, a pulse repetition frequency, f_r , of 5 kHz and a range resolution of 15 m. The local grazing angle for the data collected was approximately 1° . The radar look direction was 165.5° , with wind of 15 kts from 247° and a wave direction of 135° , significant wave height 2.2 m. The raw data was then processed over bursts of $L = 512$ samples with an FFT, using a -55 dB Dolph-Chebyshev weighting function in the time domain. It can be seen that the clutter intensity varies in time in a periodic manner. The spectrum width also varies with time, with occasional extreme Doppler excursions, such as seen around 27 s into the time record, perhaps as the result of local wind gusting. Finally, the spectrum appears asymmetric in shape, with a non-zero mean Doppler shift. These features highlight the complexity of the relationship between the intensity modulation and the form of the spectrum, the former being dominated by the swell

**FIGURE 11.9**

Sea clutter Doppler spectrogram.

structure in the sea surface and the latter being additionally affected by the local gusting of the wind and the detailed scattering mechanism. This non-stationary behavior of sea clutter will be addressed with more detail in Section 2.11.3.1.7. However, despite this complexity it should be noted that the compound modulated Gaussian process is still applicable in the spectral domain and will affect the performance of both coherent and non-coherent radars.

2.11.2.1.5 *Polarization characteristics*

It is observed that most clutter characteristics are very dependent on the polarization of the radar signal and so an understanding of polarization is important.

A wave is said to be polarized if the direction of the electric, \mathbf{E} , and magnetic, \mathbf{H} , fields reside in a fixed plane. The plane in which the \mathbf{E} vector moves is called the plane of polarization. The polarization scattering matrix, \mathbf{S} , describes the amplitude and relative phase of returns from different combinations of polarizations on transmit and receive:

$$\mathbf{S} = \begin{vmatrix} \sqrt{\sigma_{HH}} e^{j\rho_{HH}} & \sqrt{\sigma_{HV}} e^{j\rho_{VH}} \\ \sqrt{\sigma_{VH}} e^{j\rho_{PH}} & \sqrt{\sigma_{VV}} e^{j\rho_{PV}} \end{vmatrix}, \quad (11.26)$$

$$\begin{bmatrix} E_H \\ E_V \end{bmatrix}_{\text{receive}} = A\mathbf{S} \begin{bmatrix} E_H \\ E_V \end{bmatrix}_{\text{transmit}},$$

where:

σ_{HH} , RCS for Tx on H and Rx on H polarization;

Table 11.1 Examples of Linear Polarization Scattering Matrices for Different Target Types

Sphere: $S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	Dihedral at ψ° : $S = \begin{bmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{bmatrix}$
Dihedral at 0° : $S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	Linear target at ψ° : $S = \begin{bmatrix} \cos^2 \psi & \frac{1}{2} \sin 2\psi \\ \frac{1}{2} \sin 2\psi & \sin^2 \psi \end{bmatrix}$
Horizontal linear target: $S = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	Left hand helix: $S = \begin{bmatrix} 1 & j \\ j & -1 \end{bmatrix}$
	Right hand helix: $S = \begin{bmatrix} 1 & -j \\ -j & -1 \end{bmatrix}$

σ_{VV} , RCS for Tx on V and Rx on V polarization;

σ_{VH} , RCS for Tx on V and Rx on H polarization;

ρ is the associated reflection phase.

For monostatic backscatter $\sigma_{HV} = \sigma_{VH}$ and $\rho_{HV} = \rho_{VH}$.

The discussion above is for linear polarization and some examples of polarization scattering matrices for different targets types are shown in Table 11.1. Similar matrices can be used to describe the relationship for circular polarized signals (right or left handed) or any orthogonal coordinate system.

If the V and H components of the electric field are in phase, linear polarization is obtained. In general, an arbitrary phase between the V and H fields produces an elliptical polarization. The special case of a $\pi/2$ phase shift gives circular polarization.

For circular polarization, left-hand polarization is defined when $E_V = jE_H$ on transmit; right-hand polarization is defined with $E_V = -jE_H$.

The circular-polarization scattering matrix is defined as

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix}_{\text{receive}} = \begin{bmatrix} c_{LL} & c_{RL} \\ c_{RL} & c_{RR} \end{bmatrix} \begin{bmatrix} E_L \\ E_R \end{bmatrix}_{\text{transmit}}, \quad (11.27)$$

while the linear-polarization has a scattering matrix given by

$$\begin{bmatrix} E_H \\ E_V \end{bmatrix}_{\text{receive}} = \begin{bmatrix} a_{HH} & a_{VH} \\ a_{HV} & a_{VV} \end{bmatrix} \begin{bmatrix} E_H \\ E_V \end{bmatrix}_{\text{transmit}}. \quad (11.28)$$

The circular polarization RCS terms can then be related to the linear polarization RCS terms by:

$$\begin{aligned} \sigma_{LL} &= k |c_{LL}|^2 = k \left| \frac{a_{HH} - a_{VV}}{2} + ja_{VH} \right|^2, \\ \sigma_{LR} &= k |c_{LR}|^2 = k \left| \frac{a_{HH} + a_{VV}}{2} \right|^2, \\ \sigma_{RR} &= k |c_{RR}|^2 = k \left| \frac{a_{HH} - a_{VV}}{2} - ja_{VH} \right|^2. \end{aligned} \quad (11.29)$$

Using the definitions above, it can be seen that for a sphere, where $a_{HH} = a_{VV} = 1$ and $a_{HV} = a_{VH} = 0$:

$$\begin{aligned}\sigma_{RR} &= \sigma_{LL} = 0, \\ \sigma_{LR} &= \sigma_{HH} = \sigma_{VV}.\end{aligned}\quad (11.30)$$

For this reason circular polarization is often used to reduce the return from rain clutter. Odd-bounce scatterers such as spheres or trihedrals will reverse the hand of polarization on reflection and a perfect sphere will have $\sigma_{RR} = \sigma_{LL} = 0$. Unfortunately, raindrops are not perfectly spherical but, even so, the reflectivity of rain may be reduced by about 15 dB to as much as 30 dB, dependent on conditions.

Target signatures, such as high-resolution range profiles, may be quite different according to polarization. For example using circular polarization, σ_{RL} will show odd bounce scatterers while σ_{RR} will show even bounce scatterers.

2.11.2.1.6 Spatial correlation

The returns from spatially uniform clutter will have Gaussian amplitude statistics, as described in Section 2.11.2.1.3. The magnitude of the return will change as the viewing geometry changes, such as when the antenna beam scans in azimuth or the range from the radar is changed. For a square beam and pulse shape (see Section 2.11.2.1.2), the returns from clutter patches spaced by more than one beamwidth or one pulse length will be independent and uncorrelated. However, if the successive clutter patches overlap spatially, then the returns will be correlated. A convenient measure of the spatial correlation of a sequence of intensity samples, $z(i)$ ($i = 1 - N$) is given by the estimation of the correlation coefficient:

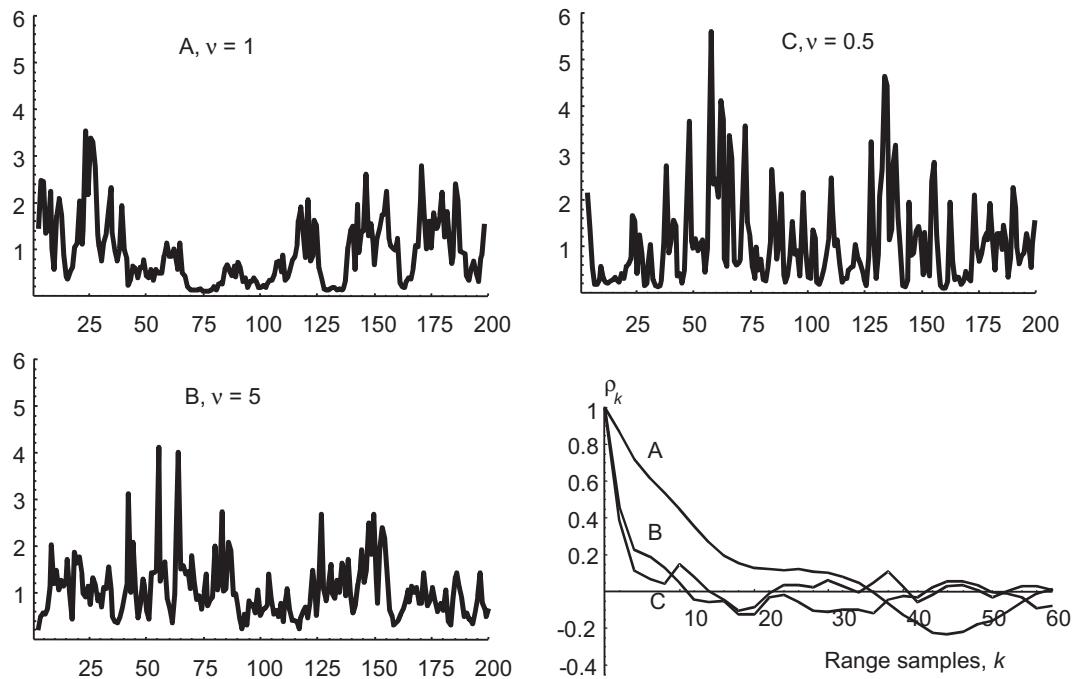
$$\rho_k = \frac{\sum_i^{N/2} (z(i) - \hat{m})(z(i + k) - \hat{m})}{\sum_i^{N/2} (z(i) - \hat{m})^2}, \quad (11.31)$$

where \hat{m} is an estimate of the mean intensity and k is the correlation lag ($0 \leq k \leq N/2$). For spatially uniform clutter with Gaussian statistics and a fractional overlap of the beam or pulses between successive samples of β ($0 \leq \beta \leq 1$), then the correlation coefficient of the clutter intensity is $\rho_k = \beta^2$.

Spatial correlation of the clutter returns may also be observed if the local normalized reflectivity is changing in a systematic way. One example of this is observed over the short term in sea clutter, when the spatial variations caused by the sea swell or waves causes a related variation in the local mean reflectivity. This is illustrated in Section 2.11.3.1. Some examples of the range profiles of the local mean intensity of sea clutter and their corresponding range correlation coefficients are shown in Figure 11.10.

2.11.2.1.7 Discrete scatterers

The models for spatially distributed clutter are very useful for representing the returns from rain and often from land and sea, especially at high grazing angles. However, under some conditions the underlying assumptions of spatially uniform scatterers is no longer valid. For example, at low grazing angles terrain scattering may become very patchy and spiky, and is dominated by local high structures [32]. At higher grazing angles a distributed clutter model for terrain becomes more useful, but there will often also be a number of very large discrete scatterers in any scene, due to natural and man-made features. Barton [33] has analyzed a number of results from the literature and suggests that discrete clutter echoes of 10^4 m^2 RCS might have a typical density of $0.2/\text{km}^2$; 10^3 m^2 RCS a density of $0.5/\text{km}^2$; and 10^2 m^2

**FIGURE 11.10**

Recorded data exhibiting different spatial correlations. Reproduced with permission from [25], © The Institution of Engineering and Technology.

RCS a density of $2/\text{km}^2$. Discrete scatterers as large as 10^6 m^2 RCS may be found. Long [34] suggest a density of about $0.004/\text{km}^2$ for 10^6 m^2 RCS and $0.04/\text{km}^2$ for 10^5 m^2 RCS.

Sea clutter may also include discrete spikes that have distinctly different properties from the surrounding clutter. In particular, specular scattering with *HH* polarization from the crests of incipient breaking waves may give rise to very localized returns having a large RCS [25]. The KA distribution [24,25] can be used to model discrete spikes that are added to the standard compound K distribution model. The amplitude distribution of clutter spikes has also been investigated in [35] who used the KK distribution to achieve a good fit to the tail of the distribution of clutter-plus-noise data recorded at medium grazing angles.

2.11.2.2 Sea clutter

Observations of radar sea clutter are usually associated with particular characteristics of the sea surface and environment, such as sea waves, sea swell and wind speed. Sea waves are the interaction between the wind and the sea surface. As the wind blows over the surface, waves are generated that increase in height and wavelength over time. Eventually an equilibrium is reached when the energy dissipated in the waves

Table 11.2 Douglas Sea State

Douglas Sea State	Description	Wave Height $h_{1/3}$ (ft)	Wind Speed Knots	Fetch (nmi)	Duration (h)
1	Smooth	0–1	0–6		
2	Slight	1–3	6–12	50	5
3	Moderate	3–5	12–15	120	20
4	Rough	5–8	15–20	150	23
5	Very rough	8–12	20–25	200	25
6	High	12–20	25–30	300	27
7	Very high	20–40	30–50	500	30
8	Precipitous	>40	>50	700	35

Table 11.3 Beaufort Wind Scale

Beaufort Number	Description	Wind Speed, Knots	Sea Conditions
0	Calm	<1	Calm (glassy)
1	Light air	1–3	Calm (rippled)
2	Light breeze	4–6	Smooth (wavelets)
3	Gentle breeze	7–10	Slight
4	Moderate breeze	11–16	Slight-moderate
5	Fresh breeze	17–21	Moderate
6	Strong breeze	22–27	Rough
7	Near gale	27–33	Rough-very rough
8	Gale	34–40	Very rough-high
9	Severe gale	41–47	High
10	Storm	48–55	Very high
11	Violent storm	56–63	Very high
12	Hurricane	>64	Phenomenal

matches the energy input by the wind. This allows an average wave height to be associated with a specific wind speed, provided that the duration (the length of time that the wind has been blowing at a given speed) and the fetch (the range extent over which the wind has been blowing) are known. Wave heights are usually measured in terms of $h_{1/3}$, the specific wave height, defined at the average peak to trough wave height of the highest one third of the waves. Ranges of significant wave heights are associated with different sea states, which can be associated with wind speed, as discussed above. Table 11.2 shows the relationships for the Douglas sea state, which is usually used for radar sea clutter modeling. Wind speed is usually characterized in terms of the Beaufort wind scale, which is shown in Table 11.3.

It should be noted that assessing the environment in a particular trial is notoriously difficult and the sea state reported by observers can show a wide variation. Wave-rider buoys can be used to estimate local wave heights but these can usually only give a rough guide to likely clutter characteristics, which depend on things such as the “wind friction velocity,” “wave age” and so on. These issues are discussed in detail in [25].

2.11.2.2.1 Theoretical and empirical models for sea clutter reflectivity

The radar backscatter from the sea is derived from a complex interaction between the incident electromagnetic waves and the sea surface. There are many theoretical models for backscatter, based on the physics of scattering from rough surfaces and approximations to scattering mechanisms. The simplest models attempt to represent the surface as many small segments, called facets, with orientations modulated by the waves. Scattering from wind-driven ripples may be approximated by Bragg scattering. The tilting of the ripples by longer sea waves changes the scattered power. This type of model, introduced by Wright [36] and Bass et al. [37], is discussed in detail in [25] and can give good results at medium to high grazing angles. However, at low grazing angles and high sea states the electromagnetic scattering becomes much more complex, with multiple reflection paths and shadowing from adjacent waves. There will also be breaking waves that can contribute considerably to the backscatter and are not modeled by simple modulated Bragg scattering. Some progress has been made recently in the understanding of electromagnetic scattering at low grazing angles [25]. However, the practical development of sea clutter models still mainly relies on empirical measurements.

Figure 11.11 shows a typical plot of normalized clutter reflectivity, σ^0 , for sea clutter as a function of grazing angle, for VV and HH polarizations. At near vertical incidence, the backscatter is quasi-specular. In this region, the backscatter varies inversely with surface roughness with maximum backscatter at vertical incidence for a perfectly smooth surface. At medium grazing angles the reflectivity shows a lower dependence on grazing angle. This is often called the plateau region. Here the reflectivity is well modeled

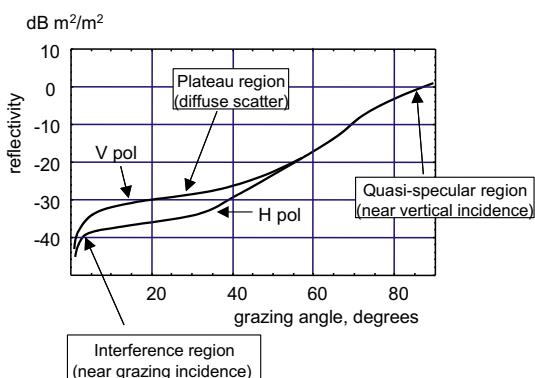


FIGURE 11.11

Typical variation with grazing angle and polarization of sea clutter reflectivity at X-band (for a wind speed of about 15 km). Reproduced with permission from [25], © The Institution of Engineering and Technology.

by the composite model. Below some critical angle (typically around 10° grazing angle, dependent on the roughness) it is found that the reflectivity reduces much more rapidly with smaller grazing angles. This is known as the interference region, where propagation is strongly affected by multipath scattering and shadowing. Also shown in Figure 11.11 is the dependence of the reflectivity on radar polarization. In the plateau region, the backscatter for *HH* polarization is significantly lower than for *VV* polarization. This is evident of the significantly different scattering mechanisms for *VV* and *HH* polarizations.

In addition to a dependence on polarization and grazing angle, it is found that the reflectivity is strongly dependent on wind speed, which creates local surface roughness. This is often associated with the sea state but it should be noted that a strong sea swell in the absence of local wind may have a low reflectivity, while a strong wind may create a high reflectivity from a comparatively flat sea. The reflectivity will also depend to some extent on radar frequency. Another important consideration is propagation effects, such as ducting, which can change the local grazing angle of the signal incident on the sea surface. Indeed, observation of variation of surface reflectivity may be used to infer the presence of ducts [38].

There are various empirical models for the normalized clutter reflectivity that are used by radar designers. Tables of σ^0 for different radar frequencies, grazing angles and sea states and for *V* and *H* polarizations are given in [1]. These values are the result of averaging measurements from many experiments. It may be noted that they do not model the variation of reflectivity with wind direction.

Another useful model, known as the GIT model [39], was developed by workers at the Georgia Institute of Technology in the 1970s. This model covers radar frequencies from 1 to 100 GHz and grazing angles from 0.1 to 10°. It is based on an underlying multipath model as well as more general trends observed in experimental data sets. The normalized reflectivities modeled for *H* and *V* polarizations, $\sigma^0(H)$ and $\sigma^0(V)$, respectively, are given by the following expressions:

Radar frequency 1–10 GHz

$$\begin{aligned}\sigma^0(H) &= 10 \log \left[3.9 \times 10^{-6} \lambda \psi^{0.4} G_a G_u G_w \right], \\ \sigma^0(V) &= \begin{cases} \sigma^0(H) - 1.05 \ln(h_a + 0.015) + 1.09 \ln(\lambda) \\ \quad + 1.27 \ln(\psi + .0001) + 9.70 & (3-10 \text{ GHz}), \\ \sigma^0(H) - 1.73 \ln(h_a + 0.015) + 3.76 \ln(\lambda) \\ \quad + 2.46 \ln(\psi + .0001) + 22.2 & (\text{below } 3 \text{ GHz}). \end{cases} \end{aligned} \quad (11.32)$$

The adjustment factors are

$$G_a = \frac{a^4}{1+a^4}; \quad G_u = \exp \left[0.2 \cos \phi (1 - 2.8\psi)(\lambda + 0.015)^{-0.4} \right]; \quad G_w = \left[\frac{1.94 U_w}{(1 + U_w/15.4)} \right]^q$$

with $q = 1.1/(\lambda + 0.015)^{0.4}$ and $a = (14.4\lambda + 5.5)\psi h_a/\lambda$.

Radar frequency 10–100 GHz

$$\begin{aligned}\sigma^0(H) &= 10 \log \left[5.78 \times 10^{-6} \psi^{0.547} G_a G_u G_w \right], \\ \sigma^0(V) &= \sigma^0(H) - 1.38 \ln(h_a) + 3.43 \ln(\lambda) + 1.31 \ln(\psi) + 18.55. \end{aligned} \quad (11.33)$$

and the adjustment factors are

$$G_a = \frac{a^4}{1+a^4}; \quad G_u = \exp \left[0.25 \cos \phi (1 - 2.8\psi) \lambda^{-0.33} \right]; \quad G_w = \left[\frac{1.94 U_w}{(1 + U_w/15.4)} \right]^q$$

with $q = 1.93 \lambda^{-0.04}$ and $a = (14.4\lambda + 5.5)\psi h_a/\lambda$.

The units and symbols used here are:

$\sigma^0(H), \sigma^0(V)$	reflectivity for H and V polarizations, dBm^2/m^2
h_a	average wave height, m ($h_a \cong 4.52 \times 10^{-3} U_w^{2.5}$)
λ	radar wavelength, m
U_w	wind velocity, m s^{-1}
ψ	grazing angle, rad ($0.1 \leq \psi \leq 10^\circ$)
ϕ	look direction relative to wind direction, rad

Radar performance is often specified in terms of sea state and a useful relationship between sea state, s , and wind speed for a fully developed sea is:

$$U_w = 3.16 s^{0.8}. \quad (11.34)$$

Figures 11.12 and 11.13 show examples, using the GIT model, of $\sigma^0(V)$ and $\sigma^0(H)$ as a function of sea state and grazing angle, for the radar looking cross-wind and radar frequency 10 GHz (using Eq. (11.32)).

It may be noted that the values of normalized reflectivity given by the GIT model are significantly different from some of the values given in [1], especially at low grazing angles. This is not surprising, given that the models were derived from different data sets, and reflects the wide variation of values that may be encountered for nominally similar conditions.

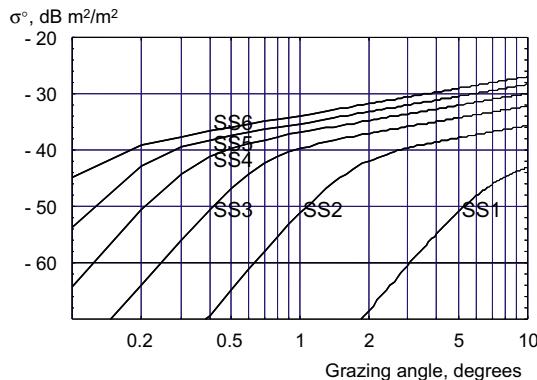
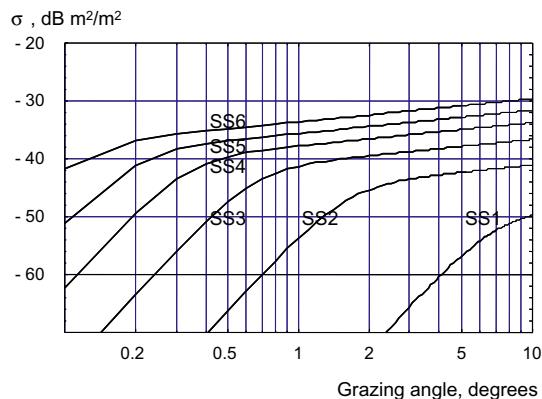


FIGURE 11.12

GIT model: $\sigma^0(V)$ as a function of grazing angle for sea states 1–6, for the radar looking cross-wind and radar frequency 10 GHz.

**FIGURE 11.13**

GIT model: $\sigma^0(H)$ as a function of grazing angle for sea states 1–6, for the radar looking cross-wind and radar frequency 10 GHz.

2.11.2.2 Sea clutter amplitude statistics

Observation of the returns from the sea surface has identified two distinct components of the amplitude fluctuations. The first is a spatial variation, the texture, often associated with the sea swell. This represents a spatial and longer-term temporal variation of the local mean clutter level. This component de-correlates only slowly with time and is not affected by radar frequency agility. The second is a speckle component, associated with scattering from multiple scatterers in a given range cell. The speckle component de-correlates with time due to relative motion of the scatterers or due to changes in radar frequency.

These characteristics are shown in Figures 11.14 and 11.15 [25], which show recordings of radar sea clutter data, with a radar operating in X band (9 GHz), with a 1.2° antenna beamwidth and 4.2 m range resolution.

Figure 11.14 shows range-time intensity plots of returns for a range interval of 800 m and a period of 125 ms, with a PRF of 1 kHz. The radar range was 5 km and the grazing angle was 1.5° . The upper part of the figure shows fixed frequency returns. At a given range the returns exhibit a correlation time of ~ 10 ms. The underlying swell pattern is clearly visible. The lower figure shows frequency agile returns. Now returns at each range are decorrelated from pulse to pulse but the swell pattern is not affected.

In Figure 11.15, the fluctuating component (speckle) has been reduced by adding successive pulses. After 60 s the polarization changes from VV to HH. The VV POL returns show a clear swell-like component while the HH POL returns show short lived (~ 1 s) clutter “spikes” which still appear to be associated with the swell peaks. It may also be noted that the overall mean level of the HH returns appears to be lower than that of the VV returns, evidence of a lower reflectivity, σ^0 .

Using data of the type shown in Figure 11.15, it has been shown [10] that the distribution of the local mean power fits well to the Gamma distribution. The local scattering (at a given range in Figure 11.14) has Gaussian statistics, resulting in the envelope of the overall return having the compound form of the K distribution (11.10). An empirical model for the dependence of the shape parameter v on radar,

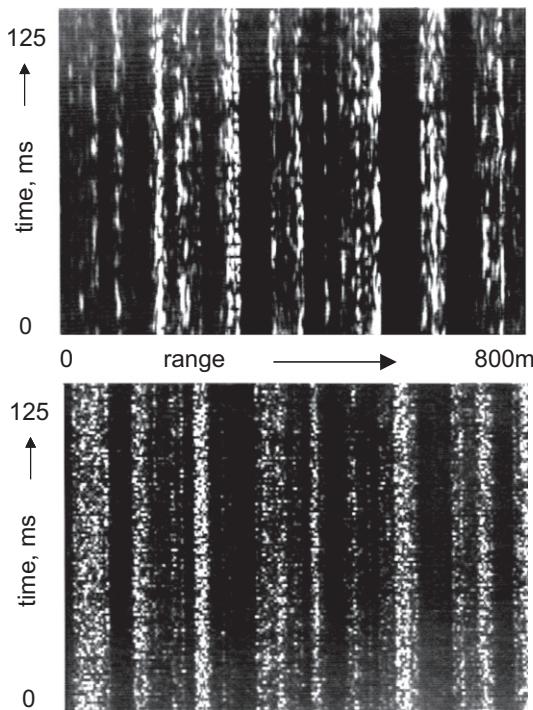


FIGURE 11.14

Range-time intensity plots of sea clutter; upper plot shows returns from a fixed frequency radar and lower plot shows returns with pulse-to-pulse frequency agility. Reproduced with permission from [25], © The Institution of Engineering and Technology.

environmental and geometric parameters has been developed through the analysis of experimental data at X-band (9–10 GHz). The model is [25]:

$$\log_{10}(\nu) = \frac{2}{3} \log_{10}(\phi_{gr}^o) + \frac{5}{8} \log_{10}(A_c) - k_{\text{pol}} - \frac{\cos(2\theta_{\text{sw}})}{3}, \quad (11.35)$$

where

ϕ_{gr}^o is the grazing angle in degrees,

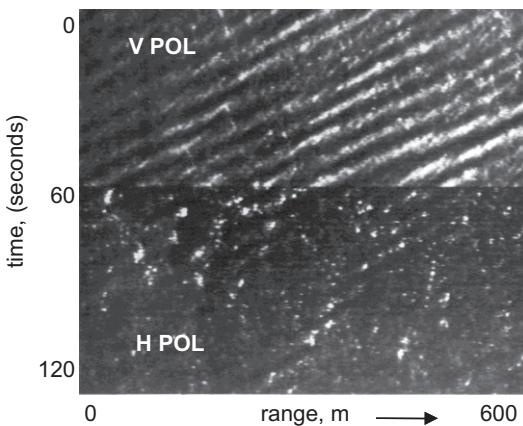
A_c is the radar resolved area, m^2 ,

k_{pol} is a polarization dependent parameter (1.39 for VV and 2.09 for HH), and

θ_{sw} is the aspect angle with respect to the swell direction.

(The last term is omitted if there is no swell).

There are no known comparable models for the shape parameters of the Weibull and Lognormal distribution models.

**FIGURE 11.15**

Range-time intensity plot of sea clutter averaged over 250 successive pulses to reduce the speckle component, revealing the underlying mean level. After 60 s, the radar was switched from vertical to horizontal polarization. Reproduced with permission from [25], © The Institution of Engineering and Technology.

2.11.2.2.3 Sea clutter Doppler spectrum

The Doppler spectrum of sea clutter varies with environmental condition, just as do the values of σ^0 and the amplitude statistics. A useful simple low grazing angle model [40] for the shape of the average clutter velocity spectrum upwind at a given wind speed U_w is a Gaussian shape with a mean velocity of m_v and standard deviation of σ_v given by (in m s^{-1})

$$\begin{aligned} m_v^{(VV)} &= 0.25 + 0.13U_w, \\ m_v^{(HH)} &= m_v^{(VV)} + 0.185, \\ \sigma_v &= m_v^{(VV)}, \end{aligned} \quad (11.36)$$

where (VV) and (HH) denote the values for vertical and horizontal polarizations, respectively. This model can be applied to Eq. (11.23) with $m_f = \frac{2m_v}{\lambda}$ and $\sigma_f = \frac{2\sigma_v}{\lambda}$.

There are many observations from data that can be used to improve this simple model. As the radar scans away from upwind, the mean velocity changes approximately proportionately with the view direction component of the wind vector, while the standard deviation remains approximately unchanged. The shape of the spectrum is skewed somewhat in the direction of the wind, and the shape varies in a complicated manner as the local mean power in the compound K distribution changes [41,42]. This is dependent on polarization and sea conditions. It causes the shape parameter of the magnitude distributions in individual frequency filters (after Doppler radar processing) to vary with frequency [10,42,43], thus causing difficulties for false alarm rate control in Pulsed Doppler and MTI radar systems. More research is needed to develop quantitative models to characterize these effects.

2.11.2.3 Ground clutter

Land (or ground) clutter is the most difficult to characterize of the common clutter categories. It can rarely be described as spatially uniform except perhaps for local regions of woods, open fields or desert.

Almost invariably there are abrupt changes in clutter due to natural or man-made boundaries (river banks, hedges, edges of woods, etc.), significant local variation in ground slope and many isolated discrete scatterers (rocks, isolated trees, pylons, buildings). Urban environments are particularly complex, as might be expected.

For modeling convenience and mathematical tractability ground clutter is often modeled quite simply as uniformly distributed scatters over a “flat” earth. Of course, amplitude statistics over large areas of land are most unlikely to be described by single Gaussian statistics with a steady backscatter power. If the backscatter coefficient is taken as representing the global mean backscatter, amplitude distributions such as the K-distribution or Log-Normal distribution are often used to fit to measured data. More detailed modeling of backscatter from land usually requires modeling of specific sites. An excellent reference for a very detailed exposition of the nature and statistics of land clutter has been written in [32]. Further useful sources of data are to be found in [2, 33, 44–47].

2.11.2.3.1 Empirical models of σ^0 for land clutter

Normalized clutter RCS, σ^0 , at low grazing angles

At low grazing angles the returns from land clutter become very spiky. Shadowing due to terrain height variations and cultural features become very marked. Under these conditions, it becomes very difficult to distinguish between spiky distributed clutter and discrete scatterers, whether man-made or natural. An excellent source of empirical measurements is the work at the MIT Lincoln Laboratory by Billingsley and others [32]. This has resulted in a very large database of land clutter data for a wide variety of terrain at low grazing angles. Measurements were made at 42 sites across North America at frequencies ranging from VHF (167 MHz) to X-band (9.2. GHz). Range resolutions of 150 m and either 36 m or 15 m were used, with both vertical and horizontal polarizations. However, Billingsley reported that variations in reflectivity due to polarization and resolution were small (1–2 dB). Because of the difficulty of defining the local grazing angle in uneven terrain, the depression angle from the radar was recorded, taking into account the earth curvature but not the effect of local terrain slope. Billingsley reported that for very low grazing angles, where masking occurs, the amplitude statistics can be represented by the Weibull PDF. At higher grazing angles, with less masking, the clutter backscatter increases and for depression angles above about 6° the clutter can be represented by a Rayleigh PDF.

Another useful empirical model was developed at the Georgia Institute of Technology [48]. This model provides an empirical fit to clutter reflectivity for a range of different terrains and radar frequencies:

$$\sigma^0 = 10 \log \left[a(\psi + C)^B \right] \exp \left[-\frac{D}{(1 + 0.1 \frac{\sigma_h}{\lambda})} \right], \quad (11.37)$$

where

- ψ is the depression angle in degrees,
- σ_h is the standard deviation of the surface roughness in cm,
- λ is the radar wavelength,
- A , B , C , and D are empirically derived constants.

Table 11.4 shows values of the parameters A , B , C , and D given in [48].

Table 11.4 Parameter Values for Land Clutter Mode [48]

Terrain Type	A				B	C	D
Frequency (GHz)	15	9.5	5	3	All	All	All
Soil, sand, and rocks	0.05	0.025	0.0096	0.0045	0.83	0.0013	2.3
Grass and crops	0.079	0.039	0.015	0.0071	1.5	0.012	0.0
Trees	0.019	0.003	0.0012	0.00054	0.64	0.002	0.0
Urban	2.0	2.0	0.779	0.362	1.8	0.015	0.0

Normalized clutter RCS—medium grazing angles

At medium grazing angles, more measurements of clutter are available, although ground truth is often difficult to obtain. It has been found that in the “plateau region” of backscatter (see Figure 11.11), the clutter reflectivity is approximately proportional to the sine of the grazing angle, ϕ_{gr} (see Figure 11.2) leading to clutter normalized RCS being defined in terms of a parameter γ

$$\gamma = \frac{\sigma^\circ}{\sin(\phi_{gr})}. \quad (11.38)$$

Typical values of γ have been reported by Barton [33], as summarized in Table 11.5.

Other empirical models for land clutter can be found in [46–49].

2.11.2.3.2 Empirical models of Doppler spectra for land clutter

Billingsley [32] has developed an empirical model for the Doppler spectrum of ground clutter. This is discussed in detail in Section 2.11.3.2.

2.11.2.4 Rain clutter

Backscatter from rain and other precipitation, such as hail and snow, can have a significant effect on radar performance. In addition, for frequencies significantly above 9 GHz the attenuation of the radar signal can be considerable. A useful summary of the effects of precipitation and weather on radar is given by Nathanson [1].

Table 11.5 Typical Values of γ for Different Types of Terrain

Terrain	Mean γ ($\text{dB m}^2/\text{m}^2$)
Mountains, urban	-5
Wooded hills	-10
Rolling hills	-12
Farmland, desert	-15
Flatland	-20

As with other types of clutter, some progress has been made with theoretical modeling of electromagnetic scattering but in general radar designers resort to empirical models of precipitation clutter.

2.11.2.4.1 Atmospheric attenuation

A detailed description of the attenuation of radar signals by the atmosphere has been provided by Blake [50]. Atmospheric attenuation varies with altitude, radar frequency and humidity. The references give typical figures for attenuation through the whole troposphere, attenuation over paths at sea level and so on. A simple rule of thumb for two-way atmospheric attenuation, α , at sea level for a frequency f GHz is given by:

$$\alpha = 10^{-2} f^{0.3} \text{ dB/km two-way.} \quad (11.39)$$

2.11.2.4.2 Rain attenuation

Rainfall is usually specified in terms of a rainfall rate, r , with the values for different levels of precipitation are given by Nathanson [1]:

Drizzle	$r = 0.25 \text{ mm/h}$
Light Rain	$r = 1.0 \text{ mm/h}$
Moderate Rain	$r = 4.0 \text{ mm/h}$
Heavy Rain	$r = 16.0 \text{ mm/h}$
Excessive rain	$r = 40.0 \text{ mm/h}$

The attenuation through rain is difficult to calculate theoretically, not least because conditions (humidity, drop size distribution and so on) can vary considerably for a nominal rainfall rate. Rainfall rates also vary with altitude and spatially across a given rainstorm. Again Nathanson [1] gives useful rules of thumb. The diameter, D , of a rainstorm can be represented by

$$D = 41.60 - 23.62 \log_{10}(r) \text{ km} \quad (11.40)$$

and if the rainfall rate at ground level is r_s , the rate r_h at height h km, is given by:

$$r_h/r_s = \exp[-0.2 h^2]. \quad (11.41)$$

The attenuation of two-way radar signals through rain is approximated by:

$$\alpha = 3.7 \times 10^{-4} f^{1.85} \text{ (dB/km)/(mm/h).} \quad (11.42)$$

2.11.2.4.3 Theoretical and empirical model for rain reflectivity

Backscatter from rain can be modeled as scattering from multiple spheres. This is generally valid for small raindrops and theoretical models can be used to predict reflectivity, within their limits of validity. For higher rainfall rates (larger drops) these assumptions may no longer be valid and again the radar designer will resort to empirical models. For Rayleigh scattering from spherical drops ($\pi D/\lambda < 1$), the volume reflectivity can be expressed as [2]:

$$\eta = \sum_N \frac{\pi^5 |K|^2 D^6}{\lambda^4}, \quad (11.43)$$

Table 11.6 Rainfall Reflectivity

Rainfall Rate, (mm/h)	Probability of Occurrence in UK (%)	Reflectivity, η , (dB m ² /m ³)					
		L-band, 1.25 GHz	S-band, 3.0 GHz	C band, 5.6 GHz	X-band, 9.3 GHz	K _u -band, 24 GHz	K _a -band, 35 GHz
0.25	5		−102	−91	−82	−64	−57
1	2.5	−107	−92	−81.5	−72	−54	−47
4	1	−97	−83	−72	−62	−46	−39
16	0.1		−73	−62	−53		−32

where D is the drop diameter, N is the number of drops, λ is the radar wavelength and $K \approx 0.95$, dependent on the dielectric constant and λ . It is found that

$$Z = \sum_N D^6 \approx 200r^{1.6} \text{ mm}^6/\text{m}^3, \quad (11.44)$$

where r is the rainfall rate in mm/h.

Empirical values for rain reflectivity are given by Nathanson [1], Currie [48], and Skolnik [2]. As an example, Table 11.6 shows some typical values of reflectivity given in [1]. Also shown in this table are values for the probability of different rainfall rates occurring, which is useful information when designing a radar for a particular application. The probabilities given in Table 11.6 are for the UK and values for other areas of the world can be found in [1].

2.11.2.4.4 Rain Doppler spectrum

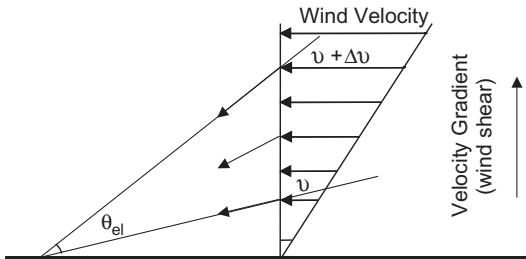
The spectrum of wind-driven rain is of particular interest to radar designers. For radar meteorologists it provides detailed information on weather patterns. For weather avoidance radar on aircraft, Doppler provides important information on wind shear and other dangerous conditions.

A good discussion of the Doppler spectrum of rain is given in [1] and this is summarized here. The velocity spectrum and equivalent Doppler spectrum of rain clutter are conveniently modeled as having a Gaussian shape, with a standard deviations $\sigma_v \text{ m s}^{-1}$ and $\sigma_f \text{ Hz}$ ($\sigma_f = 2\sigma_v/\lambda$), respectively. The spectrum of rain clutter is found to derive from various physical mechanisms, of which the principle ones are:

- Wind shear, σ_{shear} , caused by the variation of wind speed with height.
- Beam broadening, σ_{beam} , caused by the variation in the radial velocity component of velocity across the radar beam.
- Turbulence, σ_{turb} , caused by turbulent atmospheric effects.
- Fall velocity distribution, σ_{fall} , caused by variations in raindrop fall velocities.

On the assumption that these spectral components are independent, the total velocity spectrum width can be given as:

$$\sigma_v^2 = \sigma_{\text{beam}}^2 + \sigma_{\text{turb}}^2 + \sigma_{\text{fall}}^2 + \sigma_{\text{shear}}^2. \quad (11.45)$$

**FIGURE 11.16**

Wind shear.

Figure 11.16 illustrates how wind shear contributes to variations in Doppler shift across the radar beam. Wind shear causes a change in wind velocity with height, which can be approximated as a constant velocity gradient with zero velocity at ground level. Across the elevation beam of a ground-based antenna there will be a linear change of radial velocity, Δv , which for low elevation angles will equal the difference in horizontal wind speeds. For a Gaussian shaped beam:

$$\sigma_{\text{shear}} = 0.6 k R \theta_{el}, \quad (11.46)$$

where k is the shear gradient (typical value $4 \text{ m s}^{-1} \text{ km}^{-1}$), R is the range to the radar and θ_{el} is the one-way 3 dB elevation beamwidth.

For altitudes up to about 3 km, $\sigma_{\text{turb}} \approx 1 \text{ m s}^{-1}$. Beam broadening is due to tangential velocity variation across the antenna in azimuth and

$$\sigma_{\text{beam}} = 0.6 v_0 \theta_{az} \sin \theta_r, \quad (11.47)$$

where v_0 is the tangential wind velocity at the beam centre, θ_{az} is the one-way antenna azimuth 3 dB beamwidth, and θ_r is the azimuth angle relative to the wind direction at the beam centre. σ_{beam} is usually only a small component of the spectrum.

Finally, the distribution of vertical velocity amongst the raindrops will cause a spread in the velocity spectrum. A typical spread of vertical velocities is about 1 m s^{-1} so that at an elevation angle ϕ we have

$$\sigma_{\text{fall}} = 1.0 \sin \phi. \quad (11.48)$$

2.11.3 Radar clutter analysis

For the effective application of theoretical models as presented in the previous Sections it is necessary to test their fit with real data using different radar parameters and environmental conditions. We have seen that a number of families of distributions can be used to fit the observed amplitude statistics over a wide range of conditions, including the log-normal, the Weibull and, especially, the compound-Gaussian model. The main goal of this section is to describe the statistical analysis performed on different experimental data of sea and land clutter, to comment on possible results and on the limitations of theoretical models in some conditions.

2.11.3.1 Sea clutter

Most of the results shown in this paragraph on sea clutter relate to the statistical analysis performed on the data recorded by IPIX radar during two campaigns, located in Dartmouth in 1993 and in Grimsby in 1998 [15, 41, 51]. Further extensive work on the modeling of sea clutter and the associated methods used for the statistical analysis of radar data are given in [25].

IPIX is an X-band (9.4 GHz) experimental instrumentation class radar, capable of dual polarized and frequency agile operation. During the first campaign, the radar site was located on a cliff facing the Atlantic Ocean, at a height of 100 ft above mean sea level and had an open view of about 130° . During the second campaign it was on the shore of Lake Ontario, East of the “Place Polonaise” at Grimsby, between Toronto and Niagara Falls Ontario, looking at Lake Ontario from a height of 20 m.

The data of Dartmouth are stored as 8 bits integers. For the second campaign, the radar was upgraded to a dynamic range of 10 bits (instead of 8 bits), so that strong target and weak clutter signals could be observed simultaneously without clipping or large quantization error. There are always like-polarizations, HH and VV (Lpol) and cross-polarizations, HV and VH (Xpol), coherent reception, leading to a quadruplet of in-phase and quadrature values for Lpol and Xpol (see Table 11.7).

For the results shown in the following, the clutter data were collected during the Grimsby campaign. Many files with different range resolutions and recorded in different days have been analyzed [51], but here we summarize the results relating to only five files recorded on February 4, 1998 at about 22.30 h (local time) as representative of most of the results obtained from all the processed data. Unfortunately, there is no available information about the wind and wave observations for these datasets. The relevant parameters are summarized in Tables 11.8a and 11.8b.

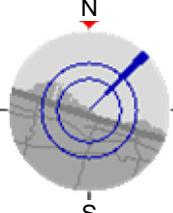
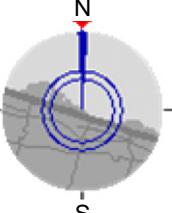
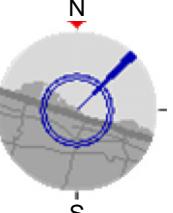
It is important to observe that in each file the range resolution is different, with data collected at 60 m, 30 m, 15 m, 9 m, and 3 m range resolution. The analysis here is aimed at highlighting the differences in clutter characteristics due to the change in the resolution.

The IPIX receiver has two operational modes depending upon the selected RF pulse width (PW). When the system was operating with $PW \geq 200$ ns, a 5 MHz filter was used to limit the receiver bandwidth to approximately 5 MHz. When $PW < 200$ ns this filter was by-passed, so the bandwidth of

Table 11.7 Characteristics of the IPIX Radar During the Grimsby Campaign

Transmitter	Receiver	Parabolic Dish Antenna
TWT peak power: 8 kW	Two receivers	Diameter: 2.4 m
Dual frequency simultaneous transmission: 8.9–9.4 GHz	Outputs: Linear, I and Q	Pencil beam width (azimuth resolution): 1.1°
H-V polarization, agile	Receiving polarizations: $H-V$	Antenna gain: 45.7 dB
Pulse width (PW): 20–5000 ns (real) 5000 ns (expanded) 32 ns (compressed)	<i>Dataacquisition</i> : Sample rate from 0 to 50 MHz Outputs: Linear, I and Q Quantization: 10 bit	Cross polarization isolation: 30 dB
PRF: from 0 to 20 kHz		Double polarization with central feeder

Table 11.8a Characteristics of the Analyzed Files, Grimsby Campaign

Name of the Data Set	19980204_223753_ANTSTEP	19980204_220849_ANTSTEP	19980204_223220_ANTSTEP
Date, time of acquisition	02/04/1998 22:37:53	02/04/1998 22:08:49	02/04/1998 22:32:20
# Range cells	28	28	28
Start range	3201 m	3201 m	3201 m
Range resolution	60 m	30 m	15 m
Pulse width	400 ns	200 ns	100 ns
Total # sweep	60,000	60,000	60,000
Sample for cell	60,000	60,000	60,000
Radar and wave geometry	N 	N 	N 

the receiver was about 50 MHz to match the minimum 20 ns pulse width. Therefore, for data collected with $\text{PW} < 200 \text{ ns}$, the receiver thermal noise level is about 10 dBs higher than for data collected with $\text{PW} \geq 200 \text{ ns}$ [52]. In the following figures the amplitude of the clutter is expressed in Volt (V).

2.11.3.1.1 Statistical models of clutter amplitude

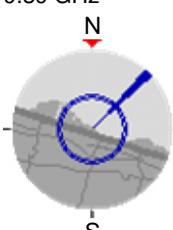
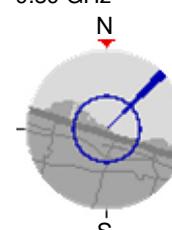
As already said, many distributions are described in the literature to model the amplitude probability density function (PDF) of high-resolution non-Gaussian clutter. Here we compare the empirical PDF with lognormal (LN), Weibull (W), K, and Generalized K (GK) and Generalized K with lognormal texture (LNT) PDFs whose expressions are given in Section 2.11.2.1.3.

The characteristic parameters of the theoretical PDFs can be estimated by the classical *method of moments* (MoM) [53], which consists of equating experimental moments with the corresponding theoretical moments. The estimated moments are given by:

$$\hat{m}_R(n) = \frac{1}{N_S} \sum_{i=1}^{N_S} |z(i)|^n. \quad (11.49)$$

For the data at hand, $N_s = 60,000$ samples have been processed for each range cell.

Table 11.8b Characteristics of the Analyzed Files, Grimsby Campaign

Name of the Data Set	19980204_224024_ANTSTEP	19980204_223506_ANTSTEP
Date, time of acquisition	02/04/1998 22:40:24	02/04/1998 22:35:06
#Range cells	28	27
Start range	3201 m	3321 m
Range resolution	9 m	3 m
Pulse width	60 ns	20 ns
Total # sweep	60,000	60,000
Sample for cell	60,000 Sampled at 9 m	60,000 Sampled at 3 m
PRF	1 kHz	1 kHz
Frequency RF	9.39 GHz	9.39 GHz
Radar and wave geometry		

Range resolutions of 60 m, 30 m, and 15 m

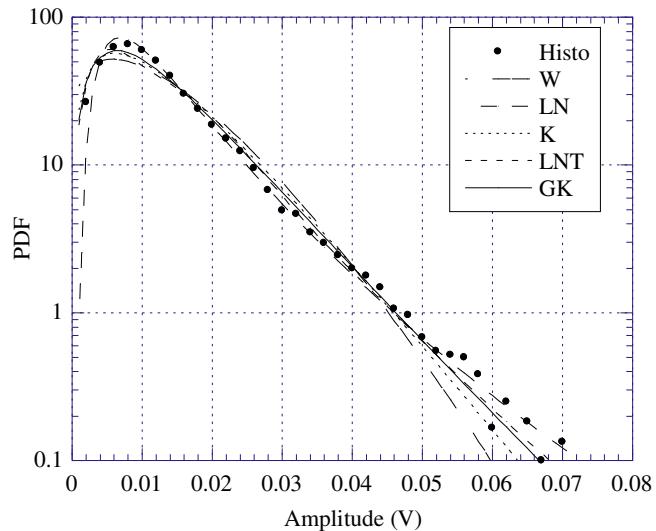
The results of the statistical analysis by means of histograms and estimated moments reveal that the GK-PDF yields a good fit for both co- and cross-polarized data and for all the three resolutions. Therefore, the analyzed clutter process can be accurately modeled by a compound-Gaussian process with Generalized K-PDF, provided that the size of the range resolution cell is greater than or equal to 15 m (note that the Gaussian model is a particular case of the Generalized K model).

In Figures 11.17 and 11.18, we report the histogram and the moments for the 15th range cell, VV data, and 60 m range resolution. The numerical results for the other range cells and the two other range resolutions are very similar [51].

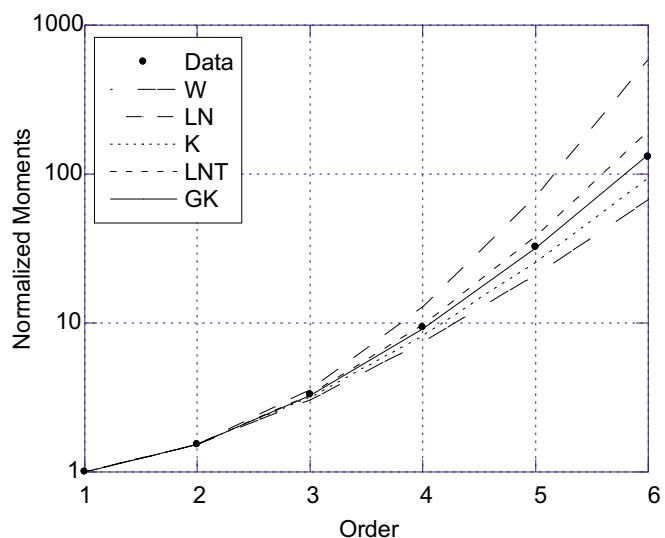
In Table 11.9 we report the mean values of the parameters estimated for each theoretical PDF. The results show that, for a resolution of 60 m, on the average, the *HH* component is spikier ($\bar{c} = 1.226$) than both *VV* ($\bar{c} = 1.293$) and *VH* ($\bar{c} = 1.292$) components.¹ Moreover, the parameters estimated when the range resolution is 30 m (Table 11.10) show that the data are spikier at 30 m range resolution than at 60 m; this was found for all polarizations. The results show also that *VV* data ($\bar{c} = 1.094$) and *VH* data ($\bar{c} = 1.093$) are spikier than *HH* data ($\bar{c} = 1.218$).

For the range resolution of 15 m and for co-polarizations generally the GK-model provides a good fit to the data. The values of the estimated parameters confirm that the clutter gets spikier when range

¹ Results for different polarizations can be compared with respect to the mean value (\bar{c}) of the estimates \hat{c} of parameter c of the Weibull distribution, because the meaning of this parameter is quite easy to understand.

**FIGURE 11.17**

Clutter amplitude PDF, VV polarization, 15th range cell, 60 m.

**FIGURE 11.18**

Normalized clutter moments, VV polarization, 15th range cell, 60 m

Table 11.9 Estimated Parameters, 60 m

Cell	W			LN		K		LNT		GK		
	\hat{c}	\hat{b}	$\hat{\sigma}$	$\hat{\delta}$	\hat{v}	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	\hat{v}	$\hat{\mu}$	\hat{b}	
VV-15th	1.391	0.014	0.652	0.010	1.331	1.268e-4	0.736	8.760e-5	5.500	2.798e-5	0.461	
mean	1.293	0.021	0.695	0.015	1.119	4.718e-4	1.020	2.694e-4	20.57	2.604e-4	0.427	
HH-15th	1.324	0.016	0.677	0.012	1.122	1.671e-4	0.867	1.083e-4	6.010	1.888e-5	0.408	
mean	1.226	0.022	0.722	0.016	0.927	4.162e-4	1.133	2.210e-4	15.55	2.758e-4	0.428	
HV-15th	1.386	0.014	0.654	0.011	1.316	1.321e-4	0.744	9.106e-5	5.341	3.123e-5	0.465	
mean	1.292	0.021	0.695	0.016	1.117	4.179e-4	0.980	2.330e-4	19.70	2.734e-4	0.430	

Table 11.10 Estimated Parameters, 30 m

Mean	W		LN		K		LNT		GK		
	\hat{c}	\hat{b}	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	$\hat{\nu}$	$\hat{\mu}$	\hat{b}
VV	1.094	0.015	0.791	0.011	0.681	2.331e-4	1.573	1.033e-4	12.80	1.946e-4	0.436
HH	1.218	0.005	0.729	0.003	0.944	1.701e-5	1.182	8.966e-6	29.77	8.766e-6	0.335
HV	1.093	0.016	0.792	0.011	0.678	2.468e-4	1.577	1.094e-4	12.28	2.061e-4	0.437

Table 11.11 Estimated Parameters, 15 m

Mean	W			LN			K			LNT			GK		
	\hat{c}	\hat{b}	$\hat{\sigma}$	$\hat{\delta}$	\hat{v}	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{b}			
VV	0.933	0.029	0.895	0.021	0.466	0.001	2.309	3.450e-4	3.602	0.001	0.459				
HH	0.874	0.024	0.935	0.017	0.383	9.364e-4	2.592	2.305e-4	3.281	9.979e-4	0.474				
HV	1.018	0.011	0.840	0.008	0.592	1.424e-4	1.905	5.198e-5	3.845	1.078e-4	0.467				

resolution increases (i.e., the size of the resolution cell decreases). We can also notice that on the average HH data are spikier ($\bar{c} = 0.874$) than VV data ($\bar{c} = 0.933$) and VH data ($\bar{c} = 1.018$); the same happens for the 15 m resolution data (see Table 11.11).

Range resolutions of 9 m and 3 m

Examining the histograms obtained by analyzing the file at a resolution of 9 m, it is possible to see that many cells of co-polarized data exhibit heavy-tails and none of the proposed models yields a good fit to the data. One of these cells is the fifth, plotted in Figure 11.19 for VV and HH data. This problem could be due to a non-Gaussian distribution of the speckle because of the very high range resolution. On the contrary, in [51] it was observed that for cross-polarizations, the clutter process can still be accurately modeled by a compound-Gaussian process with Generalized K-PDF. Again, the values of the estimated parameters show that HH data have the spikiest behavior ($HH: \bar{c} = 0.991$, $VV: \bar{c} = 1.099$, $VH: \bar{c} = 1.175$, see Table 11.12).

The results obtained for co-polarizations at a range resolution of 3 m do not show significant differences with respect to the results obtained at 9 m. Conversely, the analysis for VH polarization presents some difference. There are cells showing histograms with tails longer than the average length recorded at lower resolutions; in this case the compound model cannot be used to model clutter data.

The estimates of the parameters are reported in Table 11.13. The results show again that HH data ($\bar{c} = 1.307$) are spikier than VV ($\bar{c} = 1.417$) and VH ($\bar{c} = 1.542$) data. With respect to the other resolutions estimated values of \bar{c} are slightly higher, probably because of the effect of added thermal noise.

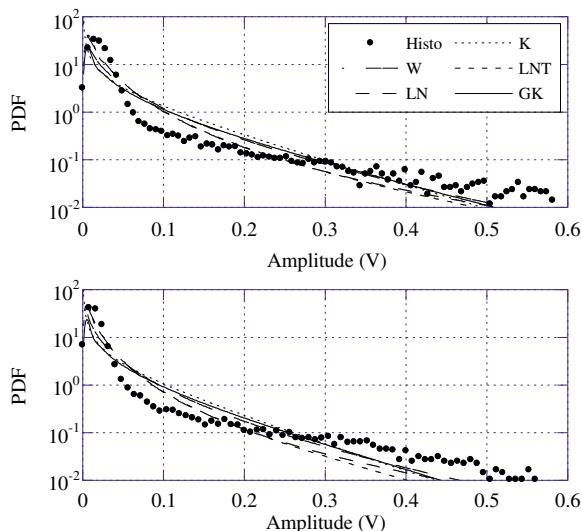


FIGURE 11.19

Clutter amplitude PDF, VV (upper) and HH (lower) polarizations, fifth range cell, 9 m.

Table 11.12 Estimated Parameters, 9 m

Cell	W			LN			K			LNT			GK		
	\hat{c}	\hat{b}	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	$\hat{\nu}$	$\hat{\mu}$	\hat{b}
VV-5th	0.585	0.021	1.207	0.016	0.129	0.002	4.861	2.032e-04	0.216	0.005	0.638				
mean	1.099	0.028	0.806	0.020	0.781	8.523e-4	1.732	3.261e-4	35.69	4.025e-4	0.227				
HH-5th	0.548	0.015	1.261	0.011	0.108	0.002	5.395	1.041e-4	0.147	0.003	0.727				
mean	0.991	0.021	0.868	0.015	0.564	5.852e-4	2.150	1.829e-4	22.59	3.777e-4	0.278				
HV-5th	0.626	0.007	1.153	0.005	0.154	2.085e-4	4.352	2.367e-5	0.459	2.680e-4	0.478				
mean	1.175	0.009	0.765	0.007	0.957	8.157e-5	1.461	3.578e-5	33.01	3.832e-5	0.259				

Table 11.13 Estimated Parameters, 3 m

Mean	W		LN		K		LNT		GK	
	\hat{c}	\hat{b}	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\nu}$	$\hat{\mu}$	$\hat{\sigma}^2$	\hat{m}	$\hat{\nu}$	$\hat{\mu}$
VV	1.417	0.038	0.668	0.028	2.621	0.001	0.889	6.470e-4	59.17	1.997e-6
HH	1.307	0.027	0.709	0.019	1.588	5.580e-4	1.118	3.071e-4	54.83	2.672e-6
HV	1.542	0.009	0.619	0.006	3.960	4.550e-5	0.608	3.284e-5	63.98	2.01e-10

Table 11.14 Probability of Type I Error ($1 - \alpha$), KS Test, Seventh Range Cell, VV Data, 9 m

W	LN	K	GK	LNT
52%	0%	51%	33%	0%

Kolmogorov-Smirnoff (KS) goodness-of-fit test

The statistical analysis of clutter amplitude can be completed by applying the Kolmogorov-Smirnoff (KS) goodness-of-fit test. This test is largely used to determine which distribution provides the best fit to the data. Unfortunately in some cases, it is not useful in distinguishing between different long-tailed models, because it places an equal importance on all regions in the probability space. Therefore, in the heaviest part of the PDF, i.e., the most affecting the results of the test, the “bell” area or the body of the PDF, many of the tested PDF are very similar.

The test is characterized by the probability of Type I error α . The type-I error is the probability of observing under H_0 a sample outcome at least as extreme as the one observed and hence provides the smallest level at which the observed sample statistic is significant. [54]. Roughly speaking, α represents the probability of having an error if we reject the null hypothesis (empirical distribution equal to theoretical distribution). If this probability α is very low, say less than 1%, the hypothesis H_0 can be rejected.

For resolutions of 60 m and 30 m, for all the tested models, the probability α is in the range of 95–99%. Some differences were found at higher resolutions. The results of the KS test for this cell are summarized in Table 11.14. For resolution higher than 30 m (i.e., values <30 m), the KS test provides the lowest probability of Type I error α generally for the LN distribution.

Summarizing, for low grazing angle and range resolution values >15 m, moments and histogram analysis generally confirm that the K and GK models provide a good fit to the data for both like and cross-polarizations (see also [55]). The HH data are spikier than VV for almost all the resolutions. For range resolution values ≤ 15 m, the compound-Gaussian model starts failing and in some range cells the data histogram shows long tails that are not well modeled with any of the tested PDFs.

2.11.3.1.2 Cumulant domain analysis

To perform additional analysis of the compound-Gaussian model and to investigate whether the deviation from the theoretical models in the highest two resolutions, i.e., 9 m and 3 m, may be due to the presence of non-negligible thermal noise, we can apply the theory of cumulants² [56]. It is widely known in the literature that cumulants of order greater than two for a Gaussian process are identically zero [57, 56]. Thus, if we consider the clutter process $z(i) = y(i) + v(i)$, where $v(i)$ is a Gaussian process and $y(i)$ is a non-Gaussian process, independent of $v(i)$, we have:

$$c_k^z(l_1, \dots, l_{k-1}) = c_k^y(l_1, \dots, l_{k-1}) + c_k^v(l_1, \dots, l_{k-1}) = c_k^y(l_1, \dots, l_{k-1}), \quad \text{for } k > 2, \quad (11.50)$$

so the cumulants of $y(i)$ can be derived from the cumulants of $z(i)$, that is, the only contribution in the cumulants of the overall process is that of the non-Gaussian component. In our case the in-phase

²For the definition of cumulants and their relation with the moments see [56]. Particularly, for real processes $x(n)c_k^x(l_1, \dots, l_{k-1}) = E\{x(n)x(n+l_1), \dots, x(n+l_{k-1})\}$.

(I) and quadrature (Q) components of the thermal noise are zero-mean Gaussian processes, then only non-Gaussian clutter contributes to the cumulants of order $k > 2$ of the observed complex data.

In [51] the authors estimated from the data the second, third, and fifth order normalized cumulants at zero-lags, i.e., for $l_1 = l_2 = \dots = l_{k-1} = 0$ whose definition is:

$$\mu_k = \frac{c_k^I(0, 0, \dots, 0)}{(c_2^I(0))^{k/2}} = \frac{c_k^Q(0, 0, \dots, 0)}{(c_2^Q(0))^{k/2}}, \quad (11.51)$$

where superscripts I and Q refer to the in-phase and quadrature components, i.e., the real and imaginary parts of the complex data. Then the estimates have been compared with the (normalized) theoretical cumulants of the compound-Gaussian model calculated at zero-lags. All the theoretical cumulants of the compound-Gaussian model of odd order calculated at the origin are equal to zero since the PDF of a complex compound-Gaussian process is symmetric with respect to the mean value (zero, in our case).

In Figures 11.20 and 11.21 we show the normalized cumulants μ_3 and μ_5 versus the second order cumulant for the resolutions of 9 m and 3 m, respectively. The results show that at a range resolution of 9 m the compound-Gaussian model is still accurate for VV and VH data because μ_3 and μ_5 are close to zero. Conversely, the fifth order cumulant for HH data shows a large deviation from zero in most cells. At a range resolution of 3 m, for most cells and all polarizations, the estimated cumulants deviate significantly from zero. This is an indication that the thermal noise, whose contribution is null in the cumulants of order higher than two, in these cases is not the cause of the deviation from the compound-Gaussian family.

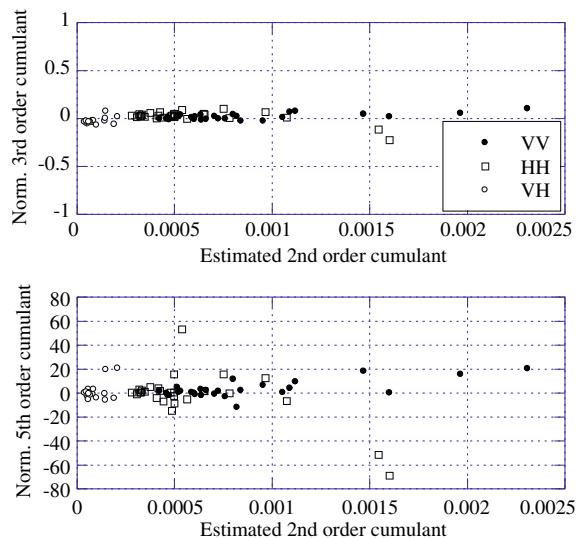
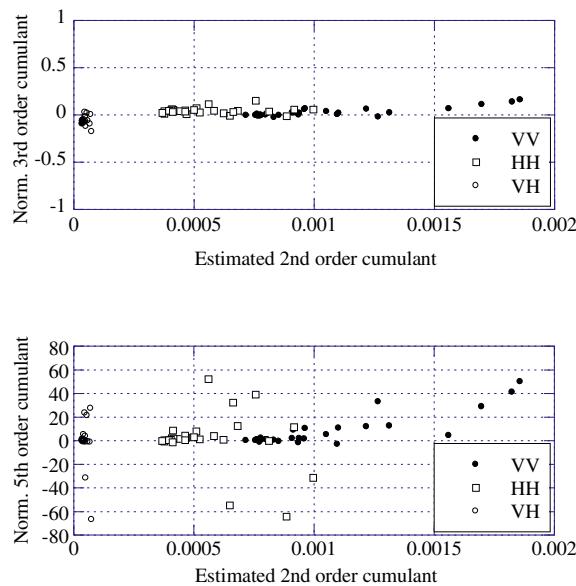


FIGURE 11.20

Normalized third and fifth order cumulants versus second order cumulant, 9 m.

**FIGURE 11.21**

Normalized third and fifth order cumulants versus second order cumulant, 3 m.

2.11.3.1.3 Correlation analysis and power spectrum estimation

As said, the compound-Gaussian clutter model assumes the presence of two components, speckle and texture, with very different correlation times (some milliseconds for the first component and some seconds for the second one, in X-band). If the two components are statistically independent the overall autocorrelation function is the product of the autocorrelation functions of the two components [15,58]:

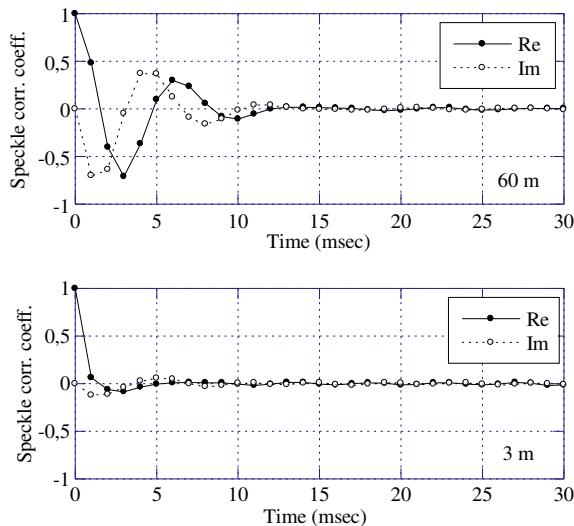
$$R_Y(m) = E \{y(i)y^*(i+m)\} = R_{\sqrt{\tau}}(m)R_X(m) = 2R_{\sqrt{\tau}}(m) (R_{X_I}(m) + jR_{X_I X_Q}(m)), \quad (11.52)$$

(if the speckle is a complex-valued stationary circular process, then $R_{X_I}(m) = R_{X_Q}(m)$). In practice, the decorrelation time of the coherent signal $y(i)$ is equal to that of the faster component [15].

2.11.3.1.4 Estimation of the speckle autocorrelation and cross-correlation sequences

Since the texture can be considered constant over short time intervals, we can estimate the speckle autocorrelation functions $R_{X_I}(m)$ and $R_{X_I X_Q}(m)$ by using coherent signal samples from such short intervals with or without overlapping.

$$\hat{R}_{X_I}(m) = \frac{1}{N_b} \left[\sum_{k=1}^{N_b} \frac{1}{2N\hat{\tau}_k} \operatorname{Re} \left\{ \sum_{i=0}^{N-1-m} y_k(i)y_k^*(i+m) \right\} \right], \quad (11.53)$$

**FIGURE 11.22**

Speckle correlation coefficients, HH polarization first range cell.

$$\hat{R}_{X_I X_Q}(m) = \frac{1}{N_b} \left[\sum_{k=1}^{N_b} \frac{1}{2N\hat{\tau}_k} \text{Im} \left\{ \sum_{i=0}^{N-1-m} y_k(i)y_k^*(i+m) \right\} \right], \quad (11.54)$$

where N_b is the number of data bursts and $\hat{\tau}_k$ is the estimated value of the texture in the k th burst $\hat{\tau}_k = \sum_{i=0}^{N-1} |y_k(i)|^2/N$ where $y_k(i) = y((k-1)N + i)$.

Figure 11.22 shows two plots of the real and the imaginary parts of the speckle autocorrelation function for 60 m and 3 m estimated with $N = 128$. This result provides a clear indication that for all resolutions and all polarizations, the speckle correlation time is about 10 ms and the behavior is oscillatory.

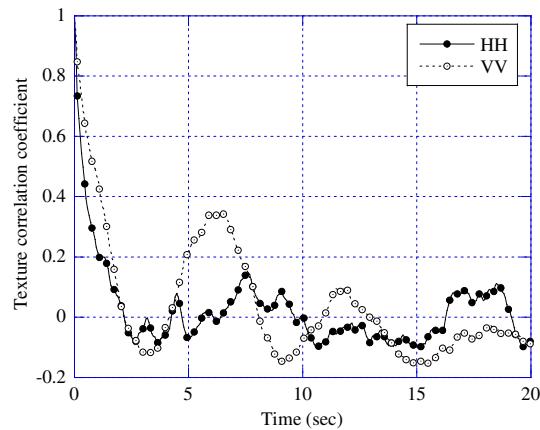
2.11.3.1.5 Estimation of the texture autocorrelation sequence

To check the validity of the hypothesis made on the correlation times of the two components, we can estimate the texture autocorrelation sequence with the formula:

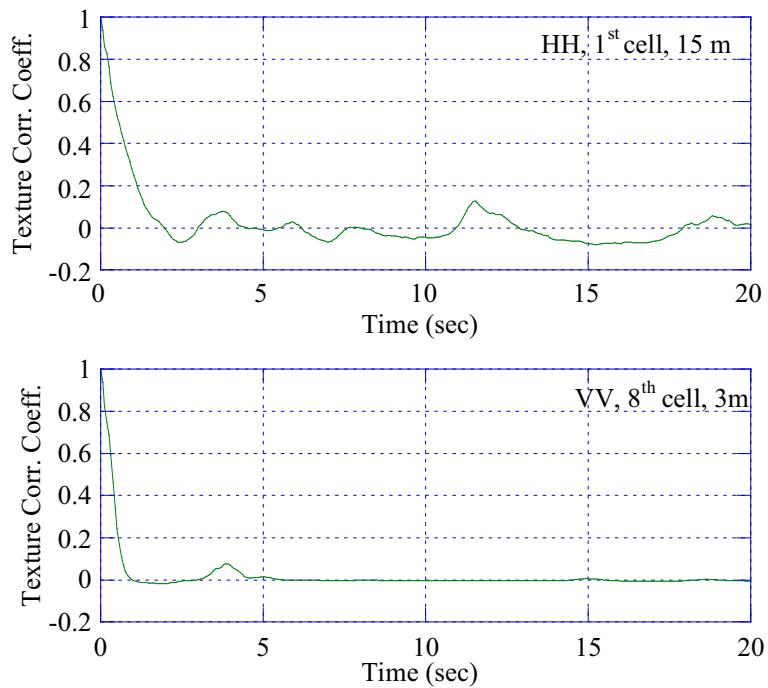
$$\hat{R}_\tau(Nm/2) = \frac{1}{N_b} \sum_{k=1}^{N_b-|m|} \hat{\tau}_k \hat{\tau}_{k+m}, \quad (11.55)$$

and the texture covariance as

$$\hat{C}_\tau(Nm/2) = \frac{1}{N_b} \sum_{k=1}^{N_b-|m|} \hat{\tau}_k \hat{\tau}_{k+m} - \left[\frac{1}{N_b} \sum_{k=1}^{N_b} \hat{\tau}_k \right]^2. \quad (11.56)$$

**FIGURE 11.23**

Texture autocovariance function, first cell, $N = 128$, 60 m.

**FIGURE 11.24**

Texture autocovariance function.

It is useful to observe that with a 50% of overlap, we can estimate the texture correlation and covariance every $N/2$ lags. In the figures we plot the texture correlation coefficient, that is $c_\tau(Nm/2) = \hat{C}_\tau(Nm/2)/\hat{C}_\tau(0)$, with $N = 128$.

Figures 11.23 and 11.24 show the texture correlation coefficient. At the same resolutions and without differences in the polarizations, the texture correlation time is on the order of seconds. Furthermore, the texture presents periodicities with a period of 8 s at a range resolution of 60 m and of 3 s at a range resolution of 30 m. The periodicity is particularly evident in the VV polarized data for the resolution of 60 m in all the analyzed files. These results show as well that, with increasing resolution, the texture correlation time decreases, but still in the order of few seconds and the periodicities tend to disappear, due to the strong contribution of the thermal noise (see Figure 11.24).

Figure 11.25 reports two examples of the average spectrogram in semi-logarithm scale calculated as

$$P(k) = \frac{1}{N_{\text{seq}}} \sum_{r=1}^{N_{\text{seq}}} P_r(k) = \frac{1}{N_{\text{seq}}} \sum_{r=1}^{N_{\text{seq}}} \frac{1}{N_c} \left| \sum_{i=(r-1)N_c+1}^{rN_c} x(i)e^{-jkn} \right|^2, \quad k = 1, 2, \dots, N_c, \quad (11.57)$$

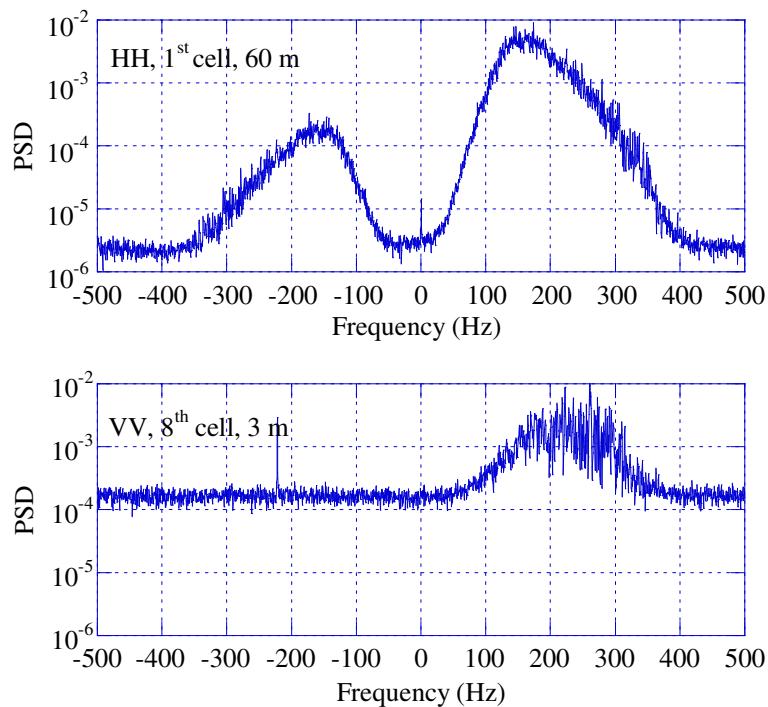


FIGURE 11.25

Average power spectral density, PSD, $CNR \approx 24 \text{ dB}$ and $CNR \approx 4 \text{ dB}$ respectively.

where N_{seq} is the number of sequences in which the received vector for each cell has been divided, N_c is the number of samples per sequence, k is the normalized frequency and $P_r(k)$ is the k th sample of the periodogram of the r th sequence.

In these results, the periodogram shows, for all the polarizations, for all the range resolutions, a peak located around 150 Hz. Moreover, with a resolution of 60 m and 30 m, most of the analyzed cells show a bimodal spectrum, particularly evident in *HH* polarization, and then a second peak near -150 Hz; the power of the second peak is much lower than the power of the main one (see upper figure). From the resolution of 15 m, the IPIX radar seems to add a frequency interfering line in the spectrum at about -220 Hz (see lower figure). The line at 0 Hz is due to a residual of the continuous component. It is evident that, as resolution increases, the thermal noise effect becomes very important. In fact, the clutter-to noise ratio (*CNR*) decreases from $\cong 24$ dB for a resolution of 60 m and *VV* data to less than -5 dB for a resolution of 3 m and *VH* data. The *CNR* has been roughly estimated from the spectrum figures reading the value of the noise floor from each figure and calculating the clutter power as the difference between the overall disturbance power and the noise power. The values calculated for each cell is reported in the figure captions.

2.11.3.1.6 Mean range texture autocovariance sequence

To conclude the correlation analysis, in order to highlight further differences due to the resolution, we can also calculate the average range autocovariance function of the texture given by

$$\hat{R}_\tau(n) = \frac{1}{N_b} \sum_{m=1}^{N_b} \hat{R}_{\tau_m}(n) = \frac{1}{N_b N_c} \sum_{m=1}^{N_b} \sum_{i=1}^{N_c-n} [\hat{\tau}_m(i) - \bar{\tau}_m] [\hat{\tau}_m(i+n) - \bar{\tau}_m], \quad (11.58)$$

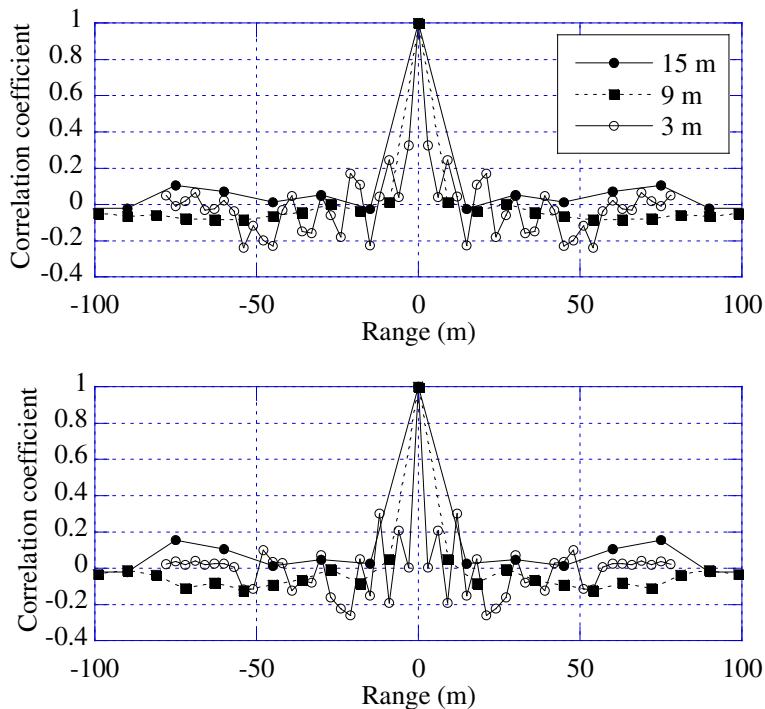
where $\hat{\tau}_m(i)$ is the estimate of the texture on the m th burst of the i th cell, N_c is the number of illuminated cells, N_b is the number of bursts and $\bar{\tau}_m$ is the texture average value in the m th burst $\bar{\tau}_m = \sum_{i=1}^{N_c} \hat{\tau}_m(i)/N_c$.

Since at 30 m range resolution the illuminated zone is different, we compare here only the results found at 15 m, 9 m, and 3 m range resolutions.

Figure 11.26 shows the results obtained for *VV* and *HH* polarizations. From the figure it is evident that with a resolution of 3 m, it is possible to highlight and resolve shorter-range periodicities that are not visible in the other resolutions in all the polarizations.

2.11.3.1.7 Sea clutter non-stationarity: Bragg scattering and long waves

So far we have considered the sea clutter process as a stationary stochastic process. Actually this is not true, and the behavior we have proved in the previous paragraphs holds true only "in average." The non-stationarity of the sea clutter is due to its nature and some explanation is needed. Sea clutter is backscattered by a moving rough surface [59,60] characterized in terms of two fundamental types of waves. The first type is represented by capillary waves with wavelengths (λ) on the order of centimeters or less, the second by the longer gravity waves (sea or swell) with wavelengths ranging from a few hundred meters to less than a meter [59]. In deep water, for the capillary waves we have $\lambda < 1.73$ cm, whereas for the gravity waves the wavelength is $\lambda > 1.73$ cm. Capillary waves are usually generated by turbulent gusts of near surface wind and their restoring force is the surface tension. On the contrary,

**FIGURE 11.26**

Mean range texture auto-covariance, *VV* polarization and *HH* polarization.

swells are produced by stable winds and their restoring force is the gravity. Then, at any point on the surface the waves are complex summations of the locally generated wind waves and waves that have propagated in from other areas and different directions, resulting in a complex interaction [61].

To take into account the presence of different scales of roughness in the sea surface, Wright [36] and Bass et al. [37] developed a two-scale model of the sea surface scattering in which the surface height is partitioned into a large-scale displacement and a small-scale displacement. For this model, it is assumed that over any patch of the surface that is large compared with small-scale lengths, but small compared with large-scale lengths, the scattering can be modeled as first-order Bragg scattering from the small-scale structure. Thus, the effect of the large-scale structure is to change the distance between the antenna and each point of the considered patch, by tilting the surface and advecting the small-scale structure both vertically and horizontally. The effect of large-scale surface tilt is to introduce an effective amplitude modulation of the small-scale scattering [62]. Conversely, the effect of the advection is to influence the frequency content of the overall scattering.

The Bragg scattering is based on the principle that the return signals from scatterers that are half a radar wavelength apart, measured along the line of sight from the radar, reinforce each other since they are in phase [59]. The Bragg resonant length is $\lambda_B = \lambda_0 / (2 \cos \theta_0)$ where λ_0 is the wavelength of the radar signal and θ_0 is the grazing angle. At microwave frequencies, the Bragg scattering is from capillary

waves and $f_D \cong \sqrt{g/(2\pi\lambda_B)}$, where g is the acceleration of free fall. As a consequence, the capillary waves approaching and receding in the radar line-of-sight direction give rise to two Bragg spectral lines located at $\pm f_D$, at least in absence of other scattering phenomena and of the long waves. The magnitude of these lines depends on the azimuth look direction of the radar relative to the wind direction [41]. In a real scenario, the Bragg scatterers are advected by the orbital velocity of the intermediate waves (waves with wavelengths longer than the Bragg wavelength but shorter than the radar resolution cell) and of the long waves (waves with wavelengths longer than the radar resolution cell). The sum of the orbital velocities of the unresolved intermediate scale waves causes a spectral broadening around the Bragg lines. For many ocean conditions, these orbital motions broaden the Bragg lines by more than their separation, causing the lines to be unresolved, and generating only one Doppler peak. At X-band frequencies, this is generally the case [61, 63]. Therefore, according to the Bragg theory, long waves that are resolved by high-resolution radars may be assumed to be constant over each illuminated cell. Consequently, their effect on the Doppler spectrum is to shift the Doppler peak according to the long wave orbital velocity. The orbital velocities are given by the simple harmonic motion $V_0 = \pi f H = \pi H/T$, where f is the frequency of the long gravity wave of period T and H is its height from crest to trough. The contribution of the orbital velocity to the motion of the Bragg scatterers is the horizontal component, that is $V_{OR} = V_0 \cos(2\pi ft - Kx)$, where K is the wave number and x the spatial position [64]. The velocities of the scatterers in the nodes, crests and troughs of the waves are very different. Finally, an additional Doppler shift results from any surface currents present, including wind drift. A formula often used to represent this drift is $D_w = 0.03U_w$, where U_w is the wind speed. Therefore, by considering the contribution of orbital velocity, current velocity V_c and wind drift, we can calculate the time-varying instantaneous Doppler shift as (see the picture in Figure 11.27) $f_D = 2 \cos \theta_0 (\pm C_0 + V_{OR} + D_w + V_c)/\lambda_0$ where C_0 is the intrinsic phase speed of the Bragg wave given by the wave dispersion relation. Due to the periodicity of V_{OR} , f_D should be periodic as well. However, the Bragg scattering is not the only phenomenon determining the clutter return, particularly if breaking waves are present on the sea surface as clearly showed in [65, 66] but it is generally dominant in down-wind VV data. Therefore, to see clearly the effect of the long waves on the sea return, we show here in detail some results of the analyses carried out on down-wind VV data [41].

It is worth observing that in the modern statistical sea clutter literature, as already written, the small-scale structure scattering of the two-scale model corresponds to the *speckle* ([11, 12, 15, 55]). The

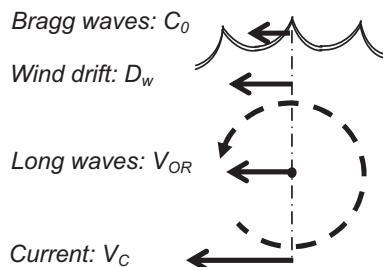


FIGURE 11.27

Different contributions to the surface velocity.

variations of the local power, due to the amplitude modulation of the speckle introduced by the tilting of the small-scale structure correspond to the *texture*. According to the compound-Gaussian model the texture and the speckle are two independent processes, the speckle is stationary and the texture amplitude modulates the speckle. However, as suggested by Haykin et al. [59], and as provided for by the two-scale model, the relationship between slowly and rapidly time-varying processes is much more complicated: the slowly varying swell motion does not only modulate the amplitude of the speckle, but also its mean frequency (Doppler centroid) and its bandwidth [41].

In this section of this tutorial, evidence of the modulating effect of long waves on speckle backscattering is verified through the analysis of experimental sea clutter data, collected in Dartmouth, Nova Scotia, at Osborne Head Gunnery Range (OHGR), with the IPIX radar.

The effect of the long waves on the overall scattering has been already considered in literature, but generally the analyses are theoretical or, when experimental, they consider only the average spectrum, that is, the spectrum calculated on the entire data set [27, 28, 61, 67]. This kind of analysis can be exhaustive only when the radar resolution is low. In this case in fact, as explained in [67], the low-resolution radar performs a spatial averaging over many waves and the radar cannot see the different features of the waves passing through the resolution cell during the recording. Some evidence of the long wave and breaking wave amplitude and frequency modulation on the time-varying spectrum of real sea clutter is presented, for instance, in [59, 64, 68].

The sea clutter data shown here were collected at Osborne Head Gunnery Range (OHGR) with IPIX radar during the Dartmouth campaign. The characteristics and acquisition conditions of the analyzed file are summarized in Table 11.15. Data from all polarizations were processed, however, in this section we describe in more details the experimental results relative to the VV polarized data of the file Starea4 recorded on the 7 November 1993 at 11:23 pm. This dataset was recorded in conditions particularly apt to reveal the long wave modulating effects (that is our main interest here), both for the fully developed sea state and for the radar downwind direction. During the time the data was collected, the Canadian Forecast Service reported that the significant wave height was of 2.23 m with an average period of 8.3 s. The wind was blowing with a speed of 4–15 km/h since 8 pm of 6 November, coming to a relative calm state (1–3 in the Beaufort scale) after it has been blowing with strong velocity (35–45 km/h) for the previous 24 h. The wind direction was 280° from the North since 10 am, and the azimuth angle was fixed at 134° from the North (146° of difference, approximately down-wind). Due to these conditions in the VV data the Bragg scattering was dominant.

A statistical analysis performed on these data as described in the previous paragraphs, showed that this clutter is GK distributed [41].

As already stated, the selection of the vertical polarization is justified because Bragg scattering was dominant (in the recording conditions of file Starea4) and the effect of long-waves should be clear and evident. Figure 11.28 shows the time evolution of the texture, estimated as $\hat{\tau}(l) = \sum_{n=1+(l-1)L_b/2}^{(l+1)L_b/2} |y(n)|^2 / L_b$, where $l = 1, 2, \dots, N_B$, $y(n)$, is the n th sample of the complex envelope of the data, $N_B = 2047$ is the number of bursts in which the entire sequence of the range cell data has been divided, $L_b = 128$ is the number of samples in the l th burst (adjacent bursts have an overlap of 50%). It has been assumed (and verified) that the texture is constant within each short burst of 0.128 s. The almost-periodic behavior of the texture is evident for both polarizations.

In the frequency domain, the presence of periodicity in the power, and then of an amplitude modulating effect of the swells, is evidenced by the time-varying spectrogram in Figure 11.29a, where the

Table 11.15 Operative Data from OHGR Database [69]. Weather Data from the Canadian Forecast Service

Data Set Name	Starea4
Date, time of the acquisition	November 7, 1993, 11:23
# range cells	7
Start range	2574 m
Range resolution	30 m
Range acquisition window	210 m
Pulse width	200 ns
Range sample rate	10 MHz
Total # sweeps	262,144
Samples for cell	131,072
PRF (polarization agility)	2 kHz
RF frequency	9.39 GHz
Grazing angle	0.305°
Azimuth angle (from N)	134°
Wind direction (from N)	280°
Approximate look direction	Down-wind
Wind speed	7 km/h (1.94 m s)
Significant wave height	2.23 m
Significant wave period	8.3 s
Sea state	Fully developed

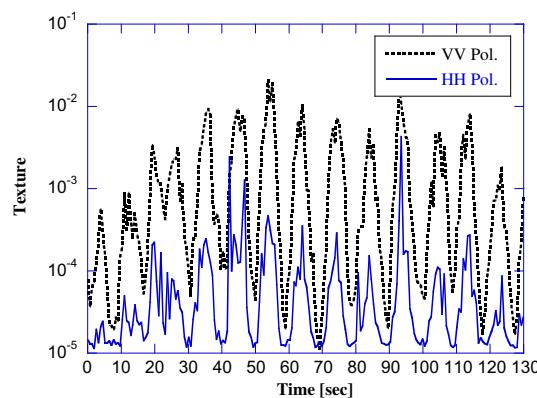
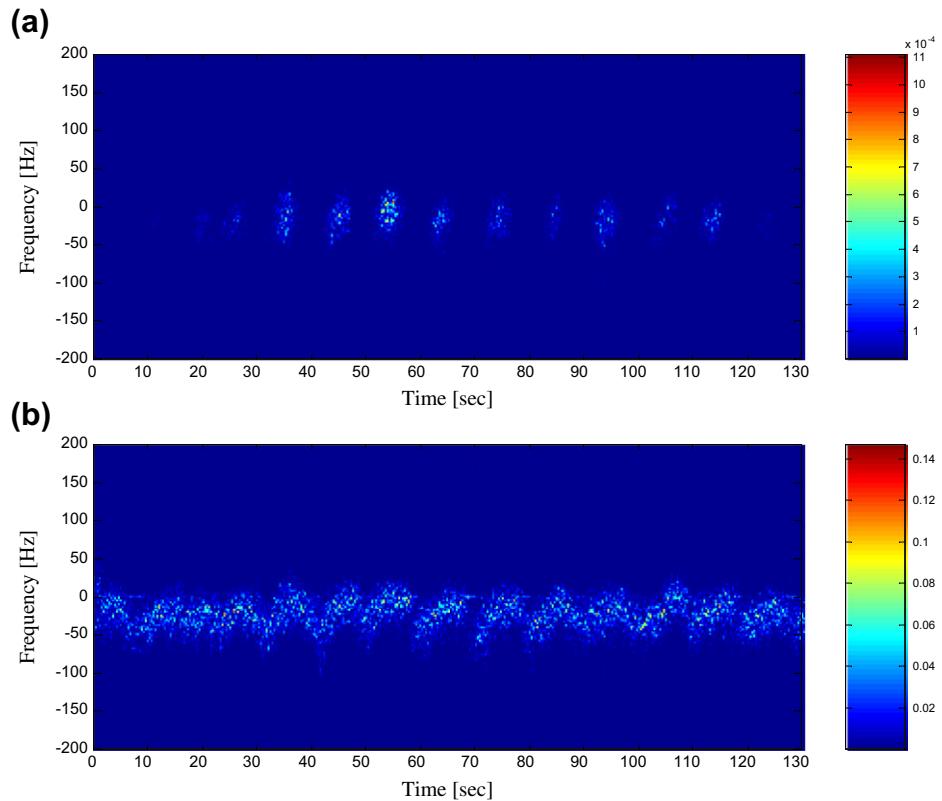


FIGURE 11.28

HH and *VV* polarized backscattering mean texture, versus the observation time interval. Dataset: Starea4 of November 7, range cell 3.

**FIGURE 11.29**

(a, b) VV spectrogram and normalized spectrogram. Starea4 of November 7, range cell 3.

spectrogram is plotted in a linear colormap from blue to red in the frequency range -200 to 200 Hz, where we have most of the sea clutter contribution (the variability of the sea spectrum was also visible in Figure 11.9). The plots in this figure were obtained by computing the FFT over a sliding window of 512 samples (0.512 s), weighed by a Hanning window, and with an overlap of 50%. The 512-points FFT ($N_F = 512$) provides the spectrogram in the frequency range -500 to 500 Hz. It is apparent that the mean texture and spectrogram illustrate a common periodic trend with a period of about 10 s that is comparable with those of the swells after projection on the radar line of sight. However, the effect of long waves is not limited to an amplitude modulation. It involves also a periodic change in the shape, bandwidth and maximum of the time-varying spectrum. This aspect is better revealed by the normalized spectrogram reported in Figure 11.29b (the spectrogram in Figure 11.29a was not normalized). The entire sequence of data has been again divided into $N_B = 2N_s/L_b - 1 = 511$ bursts of $L_b = 512$ samples each, with an overlap between adjacent bursts of 50%. Individual spectra have been calculated using a 512-point FFT and Hanning windowing, and then normalized with respect to the local clutter power. The non-stationarity of the sea clutter related to the periodic spectral variations

is evident. These variations correspond to the changes of the frequency extent and of the Doppler peak due to the frequency modulation induced by the long waves on the Bragg-wave scattering.

In order to measure the temporal variation of the backscattered Doppler spectra, we can measure the Doppler centroid of the clutter spectrum and the rms (root mean square) bandwidth, defined respectively as

$$f_C = \frac{\int_{-\infty}^{+\infty} f S(f) df}{\int_{-\infty}^{+\infty} S(f) df}, \text{ and } B_W = \sqrt{\frac{\int_{-\infty}^{+\infty} (f - f_C)^2 S(f) df}{\int_{-\infty}^{+\infty} S(f) df}}, \quad (11.59)$$

where f is the frequency and $S(f)$ is the clutter PSD. The behavior of the Doppler centroid is investigated instead of that of the PSD peak because it is generally more representative of the short-time spectrum change, especially in the case of non-symmetric spectral shape. To estimate the time-varying centroid and bandwidth we can calculate:

$$\hat{f}_C(l) = \frac{1}{Q} \sum_{n=-N_F/2}^{N_F/2-1} f(n) P_l(n), \quad \text{and}$$

$$\hat{B}_W(l) = \sqrt{\frac{1}{Q} \sum_{n=-N_F/2}^{N_F/2-1} [f(n) - \hat{f}_C(l)]^2 P_l(n)}, \quad l = 1, 2, \dots, N_B \quad (11.60)$$

where $f(n) = 1000n/N_F \text{ Hz}$ is the digital frequency, $P_l(n)$ is the periodogram of the l th data burst (composed of 512 samples) corresponding to $f(n)$, computed on $N_F = 512$ points and $Q = \sum_{n=-N_F/2}^{N_F/2-1} P_l(n)$.

Figure 11.30 shows the time evolution of the texture (on arbitrary scale and units, for ease of representation), Doppler centroid, and bandwidth, obtained by processing the data from the third range cell. They exhibit a common periodic trend with their own reciprocal delay. The Doppler centroid varies

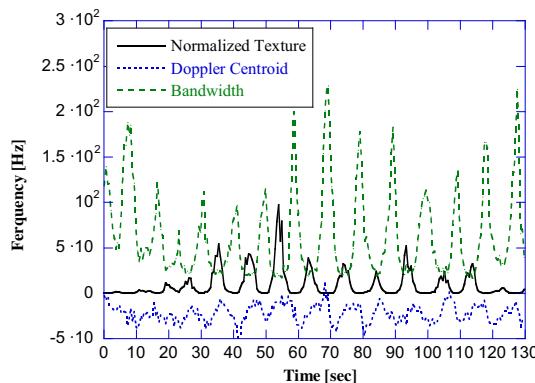


FIGURE 11.30

Time evolution of VV polarized backscattering mean texture, Doppler centroid and bandwidth. Dataset: Starea4 of November 7, range cell 3, VV polarization.

between -50 Hz and 5 Hz, the bandwidth between 25 Hz and 225 Hz. We observe that the maxima in the texture values are coincident with the minima of the bandwidth and vice versa. The spectral enlargement could be due to the modulation induced by the long waves or the presence of thermal noise. The effect of this noise in the spectral and temporal features of the received echo may be quite different depending on the clutter-to-noise power ratio (*CNR*). In the analyzed file, the *CNR* ranges from about 80 dB when the texture is maximum to about -5 dB when the texture is minimum [41]. In the latter case, the contribution of the noise to the echo is strong. The Doppler centroid and the bandwidth follow the texture behavior but with small delays of $2\text{--}3$ s and 5 s respectively, i.e., about $1/4$ and $1/2$ of the swell period.

In order to quantify the relationship between texture, centroid and bandwidth, we can use their cross-covariance functions (CCF). The CCF between the time-varying texture and the Doppler centroid signals can be calculated as

$$C_{\tau_C}(m) = \frac{\sum_{l=1}^{N_B-|m|} [\hat{\tau}(l) - \hat{\eta}_\tau] \cdot [\hat{f}_C(l+m) - \hat{\eta}_C]}{\sqrt{\sum_{l=1}^{N_B} [\hat{\tau}(l) - \hat{\eta}_\tau]^2 \sum_{l=1}^{N_B} [\hat{f}_C(l) - \hat{\eta}_C]^2}}, \quad m = -(N_B - 1), \dots, N_B - 1, \quad (11.61)$$

and for texture and bandwidth:

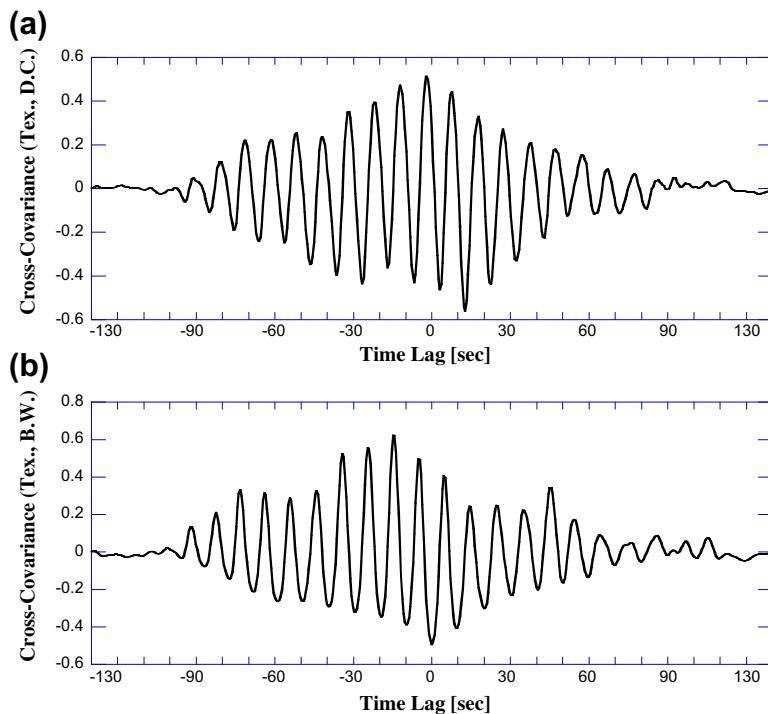
$$C_{\tau_W}(m) = \frac{\sum_{l=1}^{N_B-|m|} [\hat{\tau}(l) - \hat{\eta}_\tau] \cdot [\hat{B}_W(l+m) - \hat{\eta}_w]}{\sqrt{\sum_{l=1}^{N_B} [\hat{\tau}(l) - \hat{\eta}_\tau]^2 \sum_{l=1}^{N_B} [\hat{B}_W(l) - \hat{\eta}_w]^2}}, \quad m = -(N_B - 1), \dots, N_B - 1, \quad (11.62)$$

where, $\hat{\tau}(l)$, $\hat{f}_C(l)$ and $\hat{B}_W(l)$ are the estimated texture, Doppler centroid and bandwidth, and

$$\hat{\eta}_\tau = \frac{1}{N_B} \sum_{l=1}^{N_B} \hat{\tau}(l), \quad \hat{\eta}_C = \frac{1}{N_B} \sum_{l=1}^{N_B} \hat{f}_C(l), \quad \hat{\eta}_w = \frac{1}{N_B} \sum_{l=1}^{N_B} \hat{B}_W(l), \quad (11.63)$$

are their respective sample means. Figure 11.31a and b shows the plots of $C_{\tau_C}(m)$ and $C_{\tau_W}(m)$, respectively. They show a behavior very similar to the cross-covariance functions between two truncated sinusoids with a common period of about 10 s separated by a lag of about 2.5 s and 5 s, respectively.

In order to retrieve the actual frequency values of the common sinusoidal components, it is useful to calculate the absolute value of the mutual PSDs, i.e., the absolute value of the Fourier transform of $C_{\tau_C}(m)$ and $C_{\tau_W}(m)$. The results of this calculation are shown in Figure 11.32a and b. The information on the frequency “similarity” of texture and centroid provided by these mutual PSDs is analogous to that contained in the modulation transfer function (MTF) known to the geophysics community (see, for instance, [62] and references therein). The peaks in the mutual PSD correspond to the frequencies common to the two analyzed signals. This means, for instance, that the frequency component at $f_{LW} = 0.1$ Hz is present in both texture and Doppler centroid signals and in both texture and bandwidth signals. This fact is confirmed by the normalized Fourier spectrum of texture, Doppler centroid and bandwidth (calculated after sample mean removal). The curves are in Figure 11.33a–c. They confirm the hypothesis of a near line-shape Doppler spectrum with components located in correspondence with the long wave frequency $f_{LW} = 0.1$ Hz. Moreover, secondary components are present close to $f \approx 0$ and, especially

**FIGURE 11.31**

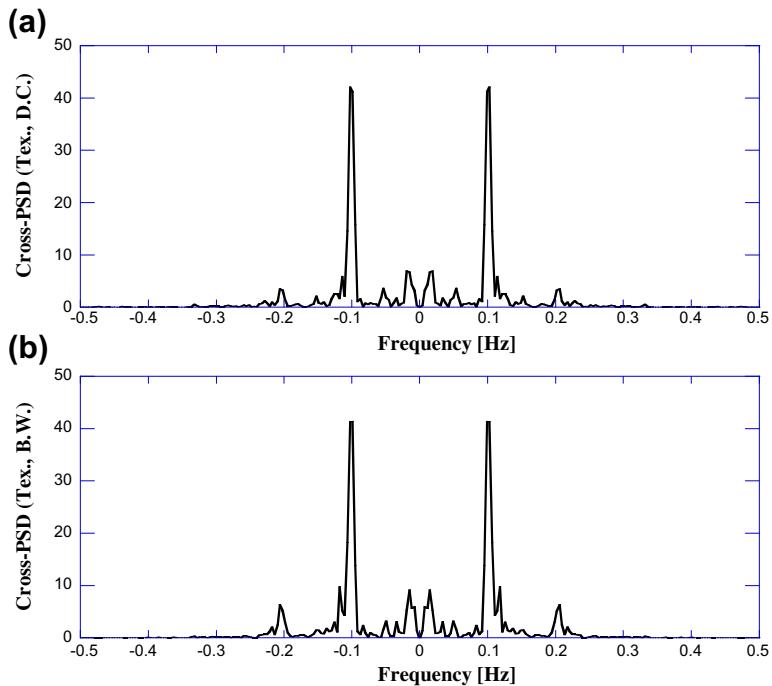
(a, b) Cross-covariance between mean texture and Doppler centroid, and between mean texture and bandwidth. Dataset: Starea4 of November 7, range cell 3, VV polarization.

for texture and bandwidth spectra, at $f = 2f_{\text{LW}}$. These results suggest that, at least when Bragg scattering is dominant, the long waves modulate the speckle scattering in amplitude, causing the variation of the local power, and in frequency, causing a periodic variation of the spectrum centroid and bandwidth.

2.11.3.1.8 AR model

As apparent, there are two goals related to clutter modeling. The first is to provide further insight into the physical and electromagnetic factors that play a role in forming the clutter signal. The second is to produce a mathematical model, physically grounded, with which the clutter signal can be generated and processed to test detection algorithms. In this section, we suggest the use of AR modeling to describe the physical phenomenon of the long wave modulation analyzed in the previous section. As known, the autoregressive models can fit both Gaussian and non-Gaussian (as in our case) processes [70, Chapter 4].

The model's ability to describe a physical phenomenon is not the main constraint, its mathematical tractability is also very important. In [41] the authors found that an AR model of the third order, AR(3), allows these constrictions. As regards the tractability, the AR(3) model is defined by seven real

**FIGURE 11.32**

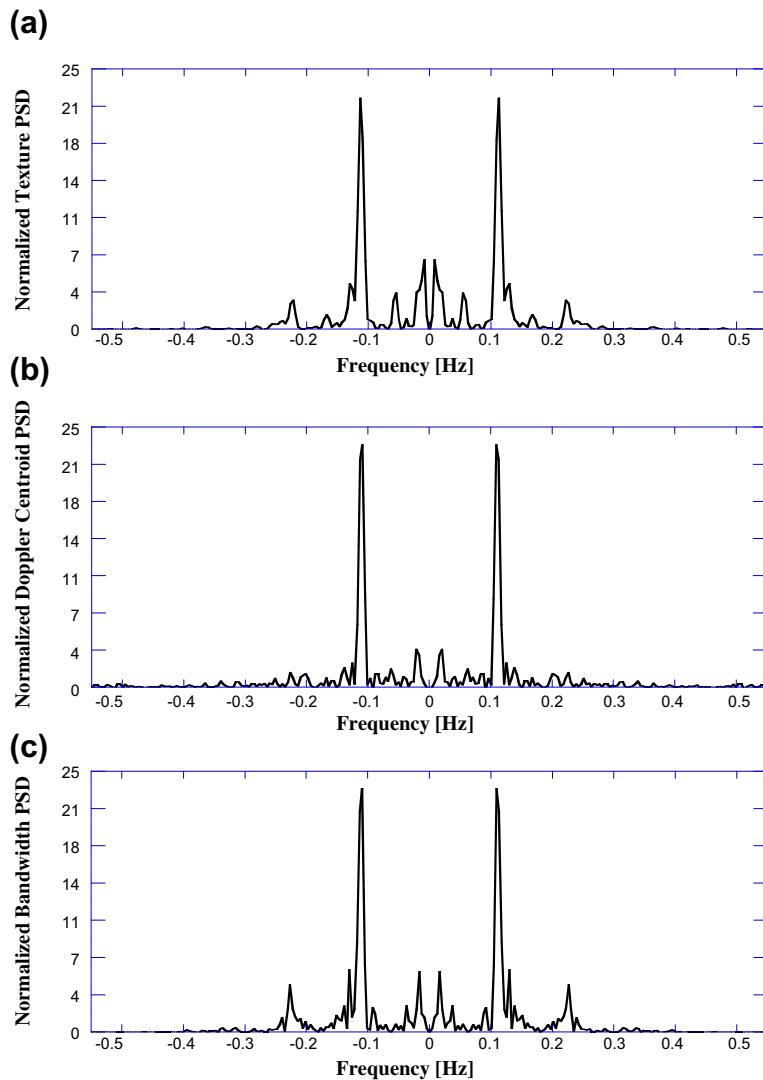
(a, b) Absolute value of the normalized cross-spectrum between mean texture and Doppler centroid, and between mean texture and bandwidth, respectively. Dataset: Starea4 of November 7, range cell 3, *VV* polarization.

parameters, the complex value of its three poles and the final prediction error. Moreover, an AR process can be simply generated by properly filtering white noise.

The normalized periodogram of the data was obtained by averaging the individual normalized Fourier spectra related to single sliding bursts. The entire sequence of data has been divided into $N_B = 2N_s/L_b - 1 = 511$ bursts of $L_b = 512$ samples each, with an overlap between adjacent bursts of 50%. Next, individual Fourier spectra were calculated using a 512-point FFT and a Hanning window and then normalized with respect to the local clutter power. As regards the AR spectrum, it was obtained by processing all the range cell samples via Yule-Walker's method [71]. In Figure 11.34 we report some results for an AR(3) process. It seems that the AR(3) can describe the basic shape features such as Doppler centroid and bandwidth.

This is confirmed by Figure 11.35a and b that represent the time evolution of these quantities, estimated both by the spectrogram and by the AR method. In particular, the estimated AR centroid is given by

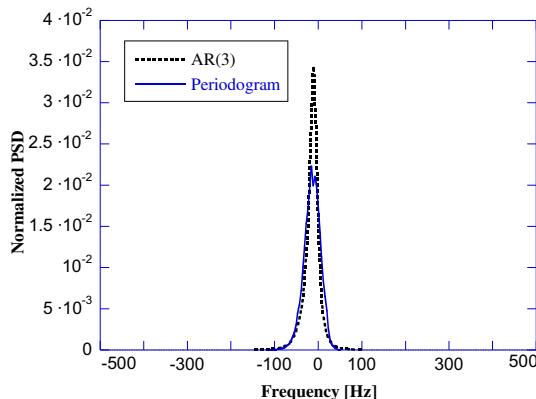
$$\hat{f}_{C, \text{AR}}(l) = \frac{1}{Q_{\text{AR}}} \sum_{n=-N_F/2}^{N_F/2-1} f(n) P_{\text{AR}, l}(n), \quad l = 1, 2, \dots, N_B, \quad (11.64)$$

**FIGURE 11.33**

(a–c) Normalized Power Spectral Density of the mean texture, Doppler centroid, and bandwidth, respectively.
Dataset: Starea4 of November 7, range cell 3, VV polarization.

while the estimated AR bandwidth is

$$\hat{B}_W^{\text{AR}}(l) = \sqrt{\frac{1}{Q_{\text{AR}}} \sum_{n=-N_F/2}^{N_F/2-1} \left(f(n) - \hat{f}_C^{\text{AR}} \right)^2 P_{\text{AR},l}(n)}, \quad l = 1, 2, \dots, N_B, \quad (11.65)$$

**FIGURE 11.34**

Normalized periodogram and AR(3) PSD. Dataset: Starea4 of November 7, range cell 3.

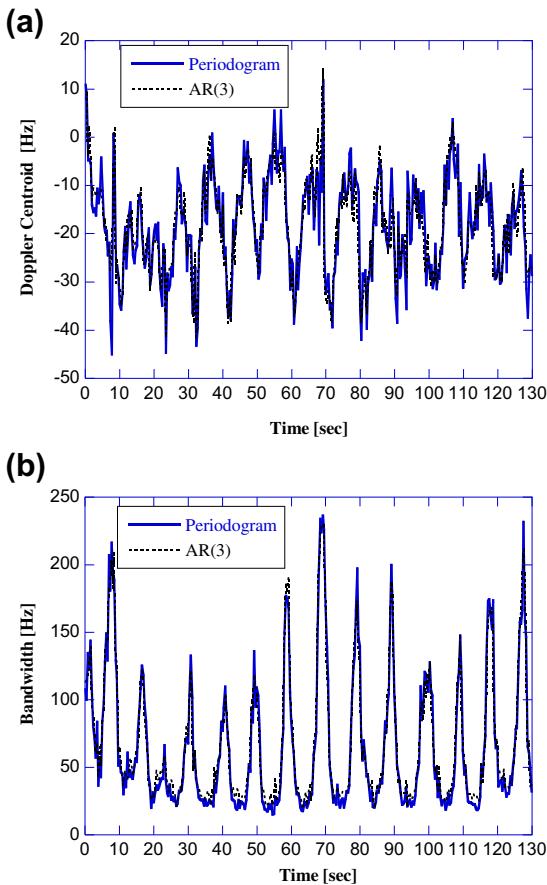
where $f(n) = 1000n/N_F \text{Hz}$ is the digital frequency, $P_{\text{AR},l}(n)$ is the AR spectrum computed by processing the 512 samples of the l th data burst, via Yule-Walker's method on $N_F = 512$ points, and $Q_{\text{AR}} = \sum_{n=-N_F/2}^{N_F/2-1} P_{\text{AR},l}(n)$. Figure 11.36a and b shows the Fourier spectrum of the AR(3)-centroid and AR(3)-bandwidth. AR centroid and AR bandwidth still preserve the spectral components of the corresponding quantities obtained by the Fourier analysis. Also, the time delay with respect to texture is preserved, as shown by the cross-covariance (Figure 11.37a and b). Therefore, this analysis suggests that AR(3) model is able to model the basic effects of long waves modulation on the capillary backscattering.

It is also of interest to investigate how its parameters vary. Figures 11.36a–c show the time evolution of the two dominant AR poles and final prediction error, together with texture. The elements of correlation are evident. The dominant pole, with almost constant unitary modulus, has an instantaneous frequency that varies with the same period of the texture. Also, the second pole follows the same periodic trend, being almost constant when the signal is strongly reflected by the crest of a wave and noisy otherwise. The third pole is dominated by noise, while the final prediction error follows the texture trend.

From the results shown in the previous pages, it is apparent that there is not a unique optimum amplitude and spectral model for the sea clutter in all the resolutions. The statistical parameters change with time, range resolution, polarization, sea conditions. Generally the compound-Gaussian model provides a good fitting to the data but it shows some limits at the very high resolutions particularly for *HH* data, which are almost always spikier than *VV* data. Moreover, the sea clutter cannot be considered stationary on long periods of time. As particularly evident in X-band and *VV* pol, the long waves modulate the sea scattering in amplitude and frequency, determining the temporal and spatial periodical changes of the texture and of the bandwidth and Doppler centroid of the speckle spectrum.

2.11.3.2 Land clutter analysis: farmland area

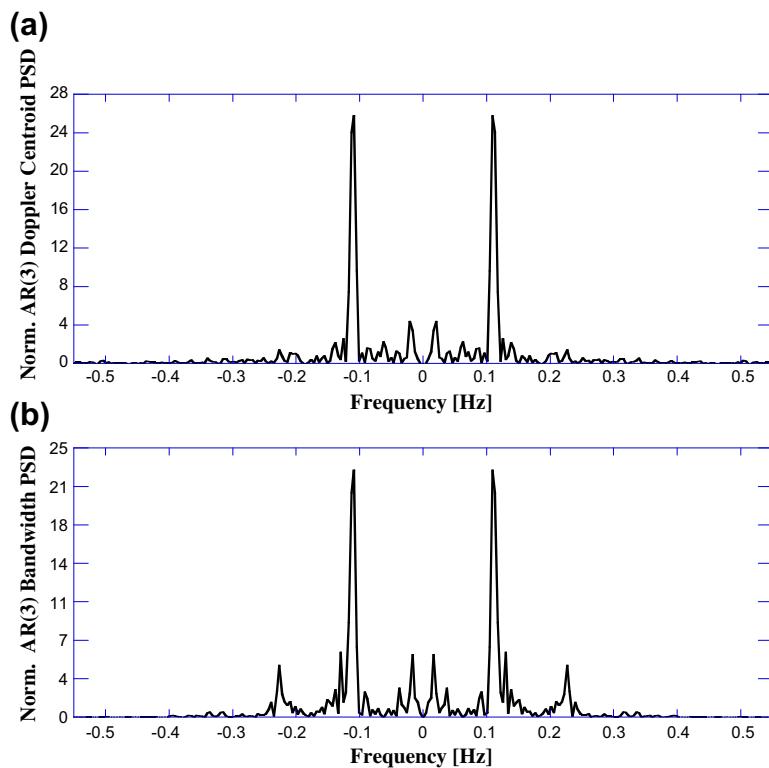
The results of the analysis of ground clutter, reported in this section, mainly relate to the data recorded with the MIT Lincoln Laboratory Phase One radar [29, 32].

**FIGURE 11.35**

(a,b) Time evolution of the Doppler centroid and bandwidth, obtained from the Fourier and AR(3) estimated spectra. Dataset: Starea4 of November 7, range cell 3, VV polarization.

The Phase One radar could operate in any one of five different radar bands (VHF, UHF, L-, S-, and X-bands). The results of this section are for ground clutter data recorded at X-band. Results in all bands are given in [32]. The overall radar system block diagram description can be found in [29].

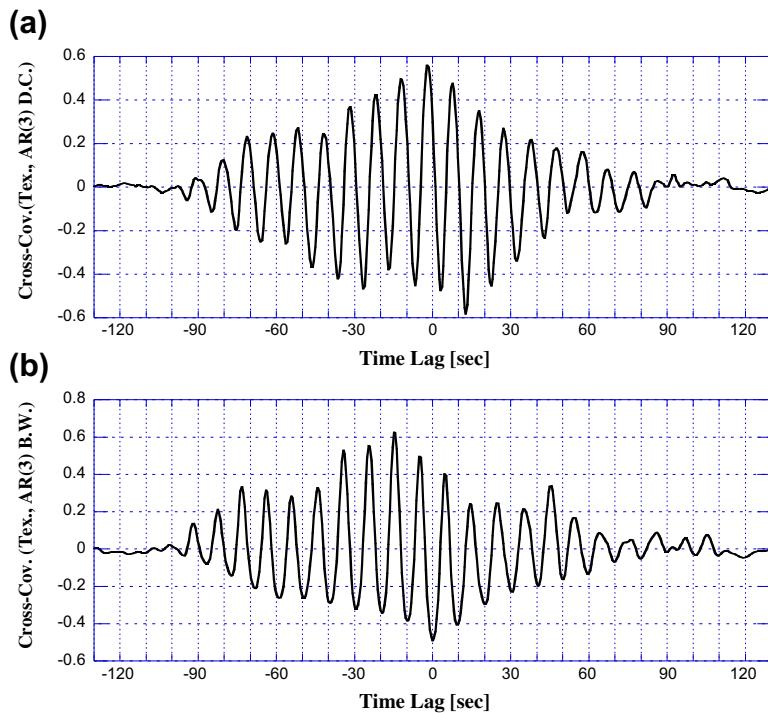
The radar maintained coherence and stability sufficient for 60 dB, two-pulse-canceller clutter attenuation in post-processing and also had uncoded pulsed waveforms, with two pulse lengths available in each band to provide high and low range resolutions. Polarization was selectable as vertical or horizontal, with transmit and receive antennas always co-polarized, i.e., the cross-polarized component in the radar return signal could not be received. Frequency, polarization, and pulse length as well as spatial extent in range and azimuth of the recording window, number of pulses, and pulse-repetition rate were selectable by computer console for each recorded clutter experiment.

**FIGURE 11.36**

(a,b) Normalized Fourier spectrum of the AR(3) Doppler centroid, and AR(3) bandwidth. Dataset: Starea4 of November 7, range cell 3.

The analyzed clutter data files described here were recorded at X-band at Wolseley, Saskatchewan, located on the Canadian prairie at a latitude of 50.36°N and a longitude of 103.150W. The illuminated area was covered by agricultural crops (83%), deciduous trees (11%), lakes (4%), and rural farm buildings (2%). The terrain was a sequence of gentle slopes (<10°) with a relief of 25–150 ft.

The analyzed X-band (9.2 GHz) data are divided in two different sets. In the first one (N007001.35) the polarization is *HH*, in the second one (N007001.34) it is *VV*. Each set contains four range intervals recorded in scan mode, and each range interval contains 316 range cells. The scan velocity was 2°/s, i.e., about 2 beams/s. The emitted pulse repetition frequency (PRF) was 500 Hz, but only 1 out of 2 pulse returns was recorded. The pulses were further coherently integrated in groups of 16, so the effective PRF of the data from this experiment is 15.625 Hz. The data were stored in a 316×703 matrix, each row for a fixed range, each column for a fixed azimuth. For each integrated pulse, 316 range samples are provided at 10 MHz sampling rate. The data were collected one range interval after another in “windshield-wiper” mode (see [29] for more details).

**FIGURE 11.37**

(a,b) Cross-covariance between mean texture and the AR(3) Doppler centroid, and between mean texture and the AR(3) bandwidth. Dataset: Starea4 of November 7, range cell 3.

The radar depression angle decreased from about 1° to 0.7° across the first range interval, and further decreased to 0.5° in the second interval, to 0.3° in the third interval and to 0.2° in the fourth range interval. The azimuth beamwidth of the antenna was 0.018 rad at HH polarization and 0.019 rad about 10 (i.e., at VV polarization). The nominal azimuth sampling interval was $0.128^\circ/\text{pulse}$; thus for each range cell the radar recorded seven to eight azimuth samples. The range resolution was 15 m (i.e., pulse length = 100 ns, which matches the 10 MHz sampling rate). The return from each pulse was provided in I and Q format calibrated in units of radar cross section [32]. The analysis was performed on the data after normalizing them with respect to the square root of the cell area. This normalization is helpful in comparing results for statistical populations extended in range, as these Wolseley data. The normalization makes the amplitude values of clutter returns independent of the distance.

In Figure 11.39a the 2D clutter map relative to the first range interval of the VV polarized data is reported. The data are plotted in logarithmic scale to span the grey-scale code (0–255). The black areas of the image indicate regions of high radar reflectivity, usually from discrete vertical clutter sources, such as buildings, fence lines, trees, and bushes, aligned along roads and field borders; the white areas indicate regions of relatively low reflectivity, such as field surfaces. On the x -axis the 316 range samples are reported from 1 km to 5.7 km, whereas on the y -axis the 703 azimuth samples are from 270° to 360° .

Figure 11.39b shows the same data as Figure 11.39a, but in a 3D format to highlight the presence of spikes, well evident in this range interval. Wolseley is a generally open farmland site of low relief. At open farmland sites like Wolseley, spatial clutter statistics are dominated by spatially localized, fixed, discrete scatterers that comprise all the vertical objects that occur on the landscape. These include the isolated trees and small clusters of trees; agricultural field boundaries and the vertical objects along them such as fences and higher uncultivated vegetation (tall grass, shrubs); roads and the vertical objects along them such as utility poles and wires; banks of streams and creeks; complexes of farm buildings and farm machinery; other cultural artifacts such as water towers and radio towers; and, also, locally high points in the microtopography itself. Such discrete objects cause strong spikes of clutter distributed randomly over the agricultural field surfaces, which themselves constitute a relatively weak, area-extensive, backscattering medium. These spikes of clutter are observed in Figure 11.38b, where the curvilinear patterns indicate field and road boundaries on a predominantly north-south, east-west grid. Such spikes are of extremely wide variation in amplitude and result in long tails in empirical clutter spatial amplitude distribution applicable to such terrain.

The first step of the analysis done on these data was the check of the Gaussianity of the I and Q components. To this purpose the histograms of the I and Q components have been drawn, for each range interval and for each polarization (VV and HH). The histogram has been compared with the Gaussian PDF having the same mean and variance of the data. The dc offset of each channel has been estimated from each of the four range intervals of $316 \times 703 = 222,148$ samples, and then subtracted from the data. This analysis, performed on each range interval, has shown that I and Q PDFs deviate considerably from Gaussianity; the clutter amplitude is therefore not Rayleigh distributed. This deviation is well evident in Figure 11.40, where the histogram of the I component for the fourth range interval (VV polarization) is compared with the Gaussian PDF having the same variance and zero-mean.

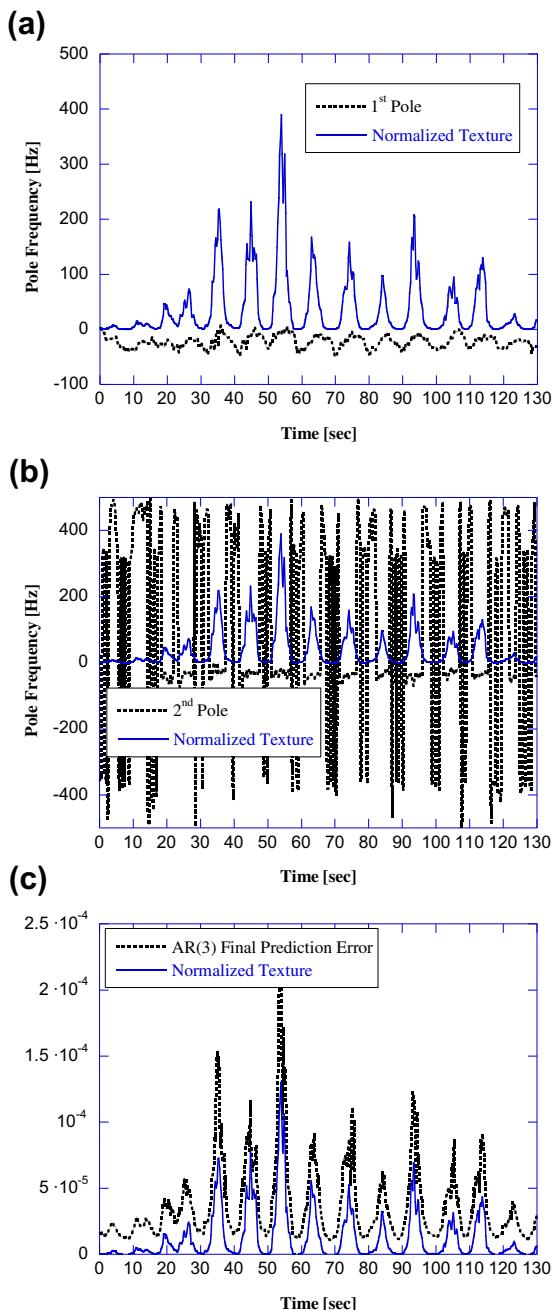
These results were confirmed by estimating the *skewness* and *kurtosis*, defined, respectively, as

$$\gamma_3(Z) \triangleq \frac{E\{(Z - E\{Z\})^3\}}{E^{3/2}\{(Z - E\{Z\})^2\}} \quad \text{and} \quad \gamma_4(Z) \triangleq \frac{E\{(Z - E\{Z\})^4\}}{E^2\{(Z - E\{Z\})^2\}} - 3, \quad (11.66)$$

where $E\{Z\}$ is the mean value of the random variable Z . The skewness characterizes the degree of asymmetry of a distribution around its mean value. A positive value of skewness corresponds to a distribution with an asymmetric tail extending on the right of the mean. A negative value of the skewness corresponds to a distribution with an asymmetric tail extending on the left. The kurtosis measures the relative peakedness or flatness of a distribution. For a Gaussian PDF these two parameters are identically zero, so they are a measure of the deviation from Gaussianity. The estimates of the skewness and kurtosis from two range intervals are reported in Table 11.16.

The study of the data has been completed performing a statistical analysis similar to that performed on the sea clutter data, comparing the histogram with some known distributions. The results show that the data seem to fit best the Weibull distribution for the first and second range intervals, while for the third and fourth range intervals the data show a behavior that is intermediate between Weibull and log-normal [29]. The results of histogram analysis for the fourth range cell are reported in Figure 11.41 on a log-scale.

The same results were obtained for the VV data of the same site.

**FIGURE 11.38**

(a-c) AR(3) pole frequencies and final prediction error, compared with the normalized texture.

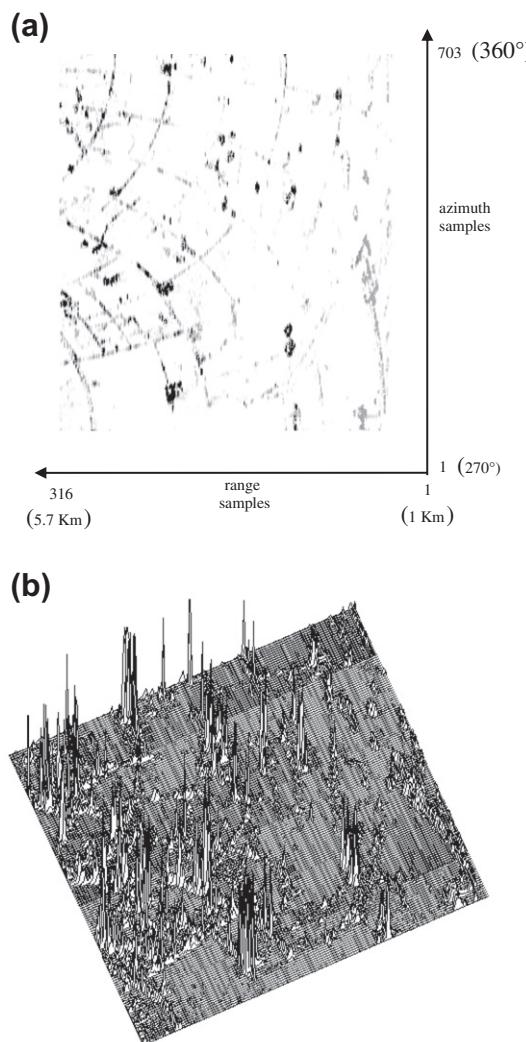
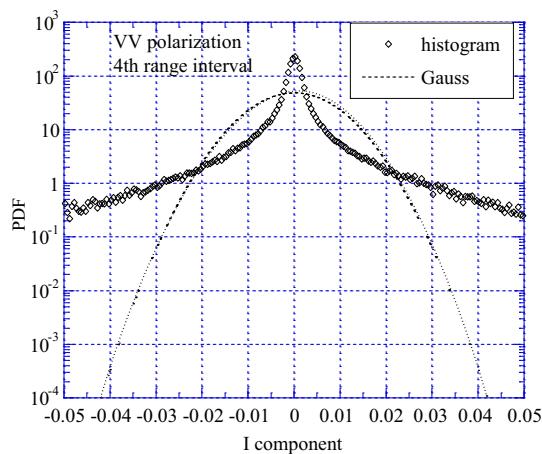


FIGURE 11.39

(a,b)2D clutter map, VV data, Wolseley site, 3D clutter map, VV data, Wolseley site.

2.11.3.2.1 **Modified Kolmogorov-Smirnoff statistical test**

The statistical analysis of data amplitude was concluded by applying a statistical hypothesis test as for the sea clutter data. The KS goodness-of-fit test has been largely used to determine which distribution (Rayleigh, log-normal, Weibull, or K in our case) best fits the data. Unfortunately, as already said for sea clutter, in some cases it is not useful, because it places an equal importance on all regions in the probability space. In practical radar applications a good fit is important in the tail regions of the PDFs.

**FIGURE 11.40**

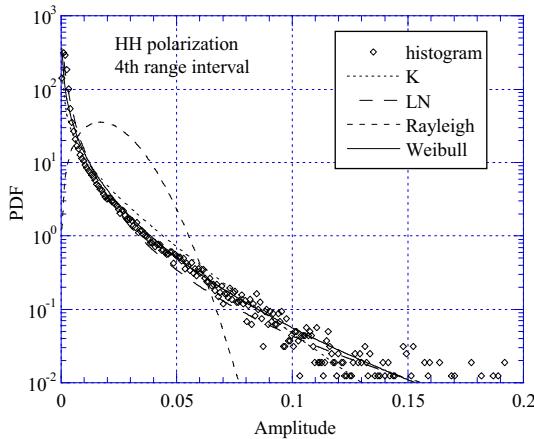
Histogram of I component, Wolseley site.

Table 11.16 Standard Deviation, Skewness and Kurtosis, Wolseley site

	HH I Comp	HH Q Comp	VV I Comp	VV Q Comp
<i>First range interval</i>				
Std dev	0.00931	0.00924	0.0110	0.0107
Skewness	-0.22284	0.18859	-0.1199	0.2341
Kurtosis	260.929	257.923	277.728	271.549
<i>Fourth range interval</i>				
Std dev	0.0081	0.0086	0.0111	0.0111
Skewness	0.0918	0.0522	0.078	-0.173
Kurtosis	115.346	111.309	131.982	133.314

The tails, in fact, contain the strong values (i.e., the spikes) that, considered as target returns by a detector, can increase the FAR. Since good fitting in the tails is mandatory for correct design of CFAR processors, especially when low P_{FA} values are required, the KS test is of limited use for clutter data (as recognized also in [32]). To overcome this problem we can use a *modified* Kolmogorov-Smirnov (MKS) goodness-of-fit test [29].

The idea is simple: apply the standard KS test by taking into account only the tail regions, i.e., by considering only the data above a given threshold AMKS and the *modified* theoretical PDFs $p_{Z,MKS}(z) = p_Z(z)u(z - \lambda_{MKS})$ where $u(\cdot)$ is the unit step function. The standard two-sample KS test verifies whether the recorded data are distributed in accordance with a hypothetical PDF, as already said for the sea clutter data. If the probability of type I is very low, for instance <1%, then this hypothesis

**FIGURE 11.41**

Histogram analysis.

should be rejected [72]. By applying the standard two-sample test, which considers the entire definition range of the random variable Z under investigation, the obtained probability of Type I error (α) is always $<1\%$, for all the distributions (Rayleigh, log-normal, Weibull, and K); thus, in the classical formulation, the KS test is not useful with these data. This is due to the differences exhibited by the distributions in the region of low values of Z . On the contrary, if we apply the MKS test setting the threshold AM_{KS} in the region $(0.01, 0.03)$, where the distributions are very similar (except for the Rayleigh), the result for the probability of type I is always 100% . This means that, in this central region, the KS test cannot distinguish between the different proposed models. But, as written above, we are mainly interested in the tails of the distribution, so we can applied the MKS test with, for instance, $\lambda_{MKS} = 0.03$. In this way we leave out of the analysis the central region of the PDF, corresponding to high values of P_{FA} (say, $P_{FA} = 10^{-2}$). The results are reported in Table 11.17.

Generally, the value of probability of error of type I, α , for Weibull PDF is higher than for the other PDFs, so the good fit of these ground clutter data to the Weibull model is confirmed.

2.11.3.2.2 Windblown trees data

The same analysis of farmland data was performed in [73] on a different Phase One X-band file, namely H067032.2, for which the data were measured from range cells containing windblown trees in contrast to open farmland. The clutter data shown here were recorded at Katahdin Hill site by Lincoln Laboratory personnel on 17 April 1985 [32]. The analyzed data set contains 30,720 samples per range cell, of *HH*-polarization, with pulse repetition frequency (PRF) equal to 500 Hz, and therefore pulse repetition interval (PRI) of $T = 1/PRF = 2$ ms. The return from each pulse is provided in in-phase and quadrature format. Data were recorded from 76 contiguous range gates utilizing the Phase One X-band stationary antenna in a fixed azimuth position (235°). These 76 range cells were located from 2.0 km to 3.1 km. The radar depression angle was about 0.65° and the azimuth beamwidth of the antenna was 0.018 rad.

Table 11.17 Results of MKS Test

Range Interval	α (Type I Error)		
	LN (%)	Weibull (%)	K (%)
1stHH	99	99	19
1stVV	6.4	96	4.2
2ndHH	1.4	96	31
2ndVV	14.6	99	3.3
3rdHH	11	87	1.3
3rdVV	62	55	<1
4thHH	84	97	1.5
4thVV	94	79	<1

The radar range resolution was 15 m. The illuminated area was tree covered, primarily with mixed deciduous trees, but also with occasional pine and cedar. At the time of the experiments, the deciduous trees did not yet have their leaves. During this experiment the wind was quite strong: 15–20 knots. Details about how the data were collected can be found in [32]. The 3D power map of the clutter is shown in Figure 11.42.

In this figure the presence of two regions, with different power values, is evident. The curves representing the received power versus range exhibit a stepwise behavior. This is due to the transition between agricultural fields/wetland (first region) and windblown trees (second region). In our analysis here we consider only single range cells from the second region. In Figure 11.43 the histogram of the data relative only to the range cells 34–36 is compared with the Rayleigh PDF with the same variance.

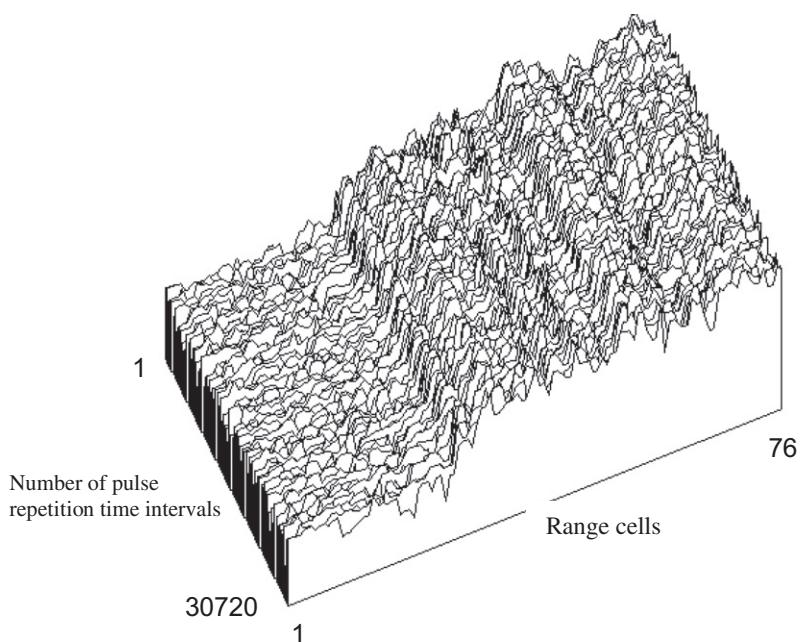
The results obtained for the open farmland data (two data files N007001.34 and N007001.35) are very different from these windblown trees data. This is partly due to the different land covers of the illuminated areas, but also importantly, partly due to the forest data embodying temporal variations, not spatial. The area relative to the H067032.3 data file was homogeneously tree covered, primarily with mixed deciduous trees and with occasional pine and cedar. In the other two analyzed data files, the returns came from a large spatial population of fixed discrete sources on open farmland. This heterogeneity introduces a considerable spread in the distributions as already noted in [32]. The differences are also due to the way of recording the data: the windblown tree data are temporal statistics (variations in time on a given range cell, or on few cells) recorded with fixed antenna; the Wolseley farmland data are spatial statistics, recorded with a scanning antenna on many range cells.

2.11.3.2.3 The experimentally-measured Doppler spectrum of ground clutter

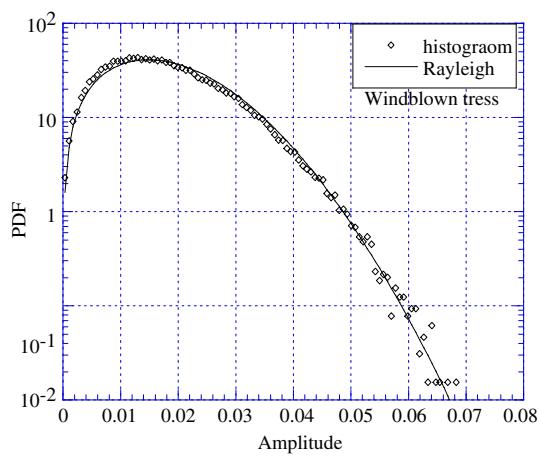
The Katahdin Hill data are particularly interesting for the shape of the PSD.

A suitable analytic expression of the PSD of ground clutter, $P_{\text{tot}}(f)$, back-scattered from regions containing windblown vegetation [32], is

$$P_{\text{tot}}(f) = \frac{r}{1+r} \delta(f) + \frac{1}{1+r} P_{\text{ac}}(f), \quad -\infty < f < \infty \quad (11.67)$$

**FIGURE 11.42**

3D power map of the analyzed clutter data.

**FIGURE 11.43**

Data histogram, windblown trees.

where f is the Doppler frequency in Hz, r is the ratio of dc power to ac power in the spectrum, $\delta(f)$ is the Dirac delta function which represents the shape of the dc component in the spectrum and $P_{\text{ac}}(f)$ represents the shape of the ac component of the spectrum, normalized such that $\int_{-\infty}^{\infty} P_{\text{ac}}(f) df = 1$. A phenomenological interpretation of this spectral shape was proposed in [74], where the exponential spectrum is obtained as the limit power spectrum in the presence of a random number of fluctuation scales in the Doppler components of windblown vegetation. From the detailed analyses of the MIT-LL measurements, it is apparent that the value of the dc/ac ratio r in Eq. (11.67) is strongly dependent on both wind speed and radar frequency and generally follows the empirically-derived analytical expression $10 \log_{10}(r) = -15.5 \log_{10}(w) - 12.1 \log_{10}(f_o) + 63.2$ where w is wind speed in miles/hour, and f_o is the radar carrier frequency in MHz. As described in [32], the MIT-LL measurements show a good fit to the model of Eq. (11.67) when the ac component of the spectrum is characterized by the two-sided exponential spectral shape

$$P_{\text{ac}}(f) = \frac{\lambda\beta}{4} e^{-\frac{\lambda\beta}{2}|f|}, \quad -\infty < f < \infty, \quad (11.68)$$

where λ is the radar transmission wavelength and β is the exponential shape parameter. The value of β is a function of the wind conditions such that the spectral width increases with increasing wind speed (see Table 11.18); however, β is largely independent of radar carrier frequency over the range from VHF to X-Band. An algebraic expression for β that incorporates the linear dependency of spectral width on the logarithm of the wind speed w (mph) as observed in the data is $\beta^{-1} = [\log_{10} w - \log_{10}(2/3^{1.5})] \cdot (20 \log_{10} 3)^{-1}$.

The described equations and Table 11.18 constitute a simple but complete model for characterizing the complex physical phenomenon of windblown clutter spectra over spectral dynamic ranges reaching 60–80 dB below zero-Doppler peaks, applicable over the range from VHF to X-band [32].

The Gaussian model for the windblown clutter spectrum, which was supported by early measurements with very limited dynamic range, is given by

$$P_{\text{ac}}(f) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{f^2}{2\sigma^2}}, \quad -\infty < f < \infty, \quad (11.69)$$

where σ is the standard deviation of the Gaussian spectrum. The exponential model has wider tails than the Gaussian model. For use with somewhat increased dynamic range, the other alternative popular spectral shape function $P_{\text{ac}}(f)$ has been the power-law spectrum, which is characterized by two

Table 11.18 Exponential AC Shape Parameter β versus Wind Speed

Wind Conditions	Wind Speed (mph)	β (s/m) (Typical)	β (s/m) (Worst Case)
Light air	1–7	12	–
Breezy	7–15	8	–
Windy	15–30	5.7	5.2
Gale force	30–50	4.3	3.8

parameters, the shape parameter n , and the break-point Doppler frequency f_c that determines where the shape function is 3 dB below its peak zero-Doppler level:

$$P_{ac}(f) = \frac{n \sin(\pi/n)}{2\pi f_c} \frac{1}{1 + (f/f_c)^n}, \quad -\infty < f < \infty. \quad (11.70)$$

The MIT-LL measurements of much increased dynamic range clearly indicate that clutter spectral shapes are indeed wider than Gaussian; but that observed rates of decay modeled as power-law at upper levels of spectral power do not continue as power-law to lower levels of spectral power, but fall off much faster at the lower levels. The tails of the exponential model show an intermediate rate of decay—between the fast decay of the Gaussian model and the slow decay of the power-law model—that yields a good fit to the measured data.

For analyzing the PSD of windblown tree clutter, many data from various windblown-tree range cells were processed. Here we report only the results from range cell #35 (second region) as representative of results achieved by processing other range cells in the same region. The PSD has been first estimated non-parametrically, by using the modified periodogram method (Welch method) with the four-sample Blackman-Harris window [73]:

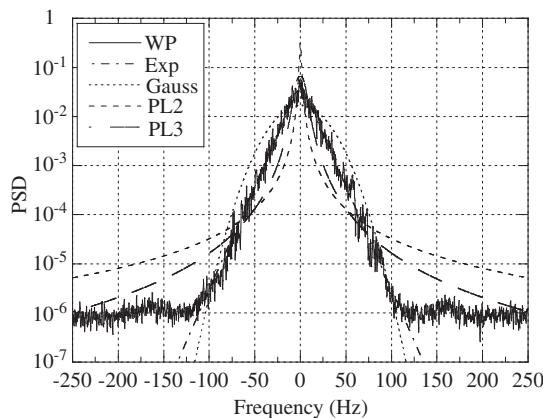
$$w[n] = \sum_{i=0}^3 a_i (-1)^i \cos\left(\frac{2\pi n}{N}\right), \quad n = 0, 1, \dots, N-1, \quad (11.71)$$

with coefficients $a_0 = 0.40217$, $a_1 = 0.49703$, $a_2 = 0.09392$, and $a_3 = 0.00183$. The set of 30,720 data samples has been divided in $L = 15$ subsets $\{z^{(k)}[n]\}_{n=0}^{N-1}$, with $k = 1, 2, \dots, L$, of $N = 2048$ samples each and then the windowed periodogram calculated from each subset as follows

$$P_k(f_i) = \left| \sum_{n=0}^{N-1} w[n] z^{(k)}[n] e^{-j2\pi f_i n} \right|^2, \quad (11.72)$$

where $f_i = i/(NT)$, for $-N/2 \leq i \leq N/2 - 1$, and $z^{(k)}[n] = z_I^{(k)}[n] + jz_Q^{(k)}[n]$ is the n th complex sample relative to the k th subset. To obtain the final spectrum estimate, the K estimates are averaged as $s(f_i) = (1/K) \sum_{K=1}^K P_K(f_i)$. Finally, the estimate is normalized such that the area is unity, i.e., divided by the estimate of the total disturbance power $P_{C+N} = P_C + P_N$. The disturbance power is estimated to be $\hat{P}_{C+N} = 5.9276 \cdot 10^{-4}$. The resulting curve is shown in Figure 11.44, labeled as “WP,” which stands for weighted periodogram. Note that the PSD is plotted from -250 Hz to 250 Hz because the PRF is 500 Hz. It is quite evident in these figures that, for $|f| > 100$ Hz, thermal noise predominates over clutter; thus we fit the clutter PSD only over the frequency range $[-100, 100]$ Hz. In [73] the authors used the frequency bins in the range $[-250, -100] \cup [100, 250]$ Hz to estimate the clutter-to-noise power ratio (CNR), which is obtained by solving the equation $(5/3) \sum_i S(f_i) = P_N / P_{C+N} = 1 / (1 + CNR)$ with respect to CNR , where P_N is the thermal noise power, $S(f_i)$ denotes the normalized PSD estimate and the sum is over all i 's such that $|f_i| \in [100, 250]$ Hz. For the 35th range cell it results that $CNR \cong 33$ dB.

To estimate the clutter PSD parameters, nonlinear least squares (NLLS) method [15] has been applied. The NLLS method can be applied directly to either $P(f_k)$ or to its logarithm. In the second case, the

**FIGURE 11.44**

Power spectral density, Log-NLLS method, 35th range cell.

estimates of the model parameters are derived by solving the following minimization problem:

$$\boldsymbol{\theta}_{\text{Log-NLLS}} = \arg \min_{\boldsymbol{\theta}} \sum_i |\log_{10} P_{\text{ac}}(f_i, \boldsymbol{\theta}) - \log_{10} S(f_i)|^2, \quad |f_i| < 100 \text{ Hz}. \quad (11.73)$$

where $\boldsymbol{\theta}$ is the vector of model parameters that have to be estimated. We refer to the method in (64) as the NLLS method. For the Power Law model in Eq. (11.61), we consider the two cases $n = 2$ (PL2) and $n = 3$ (PL3), and then we use the NLLS method to estimate the PSD parameters. In Figure 11.44 the various models obtained by the NLLS method are compared with the weighted periodogram. In Table 11.19 we report the estimates of β , σ , and f_c obtained by the NLLS method. The value of r is always lower than 10^{-4} , so the dc component in cell #35 is negligible. The best fit is provided by the exponential model.

The analysis of the land clutter shows that the clutter statistics are heavily site-dependent. As a matter of fact the amplitude of the clutter at X- and L-band is Gaussian distributed for homogeneous clutter as that of windblown trees, non-Gaussian distributed if backscattered by discrete sources as pylons, fences, etc. Even the PSD depends on the sites, and, often, on the dynamic range of the radar as well.

Table 11.19 Estimates of the PSD Model Parameters, Log-NLLS Method

Cell #35	Exp	Gauss	PL2	PL3
$\beta/\sigma/f_c/f_c$	$5.95 (2.3. \text{ Hz m})^{-1}$	23.63 Hz	1.02 Hz	6.33 Hz

2.11.4 Simulation methods

In circumstances where mathematical analysis is not possible because it is too complex, it is often necessary to resort to numerical simulation of clutter returns. This data may be used in computer simulation of signal processing. In addition, the testing and assessment of real hardware often requires the processing of data in real time in the laboratory. This can sometimes be achieved with recorded data but often it is necessary to use numerically simulated data.

When simulating clutter returns, the problem addressed here is that of the generation of random variates with prescribed one and two point statistics (i.e., the probability density function, PDF, and correlation function). The methods for generating this data will reflect the statistical models described in Section 2.11.2.1.3.

2.11.4.1 Generating uncorrelated random numbers with a prescribed PDF

There are many well-proven techniques for generating very long sequences of random numbers uniformly distributed between 0 and 1. This data will form the basis of techniques for generating random numbers with other PDFs.

The general solution for transforming data, x , with one PDF, $P_1(x)$, to data y with another PDF, $P_2(y)$, is the memoryless non-linear transform (MNLT) is given by Tough and Ward [75]:

$$\int_x^{\infty} P_1(x')dx' = \int_y^{\infty} P_2(y')dy'. \quad (11.74)$$

Given a value for x , this equation is solved for y . A simple example is the generation of data, y , with an exponential PDF, starting with random numbers, x , uniformly distributed between 0 and 1. Now:

$$\begin{aligned} \text{Now } P(y) &= \exp(-y); \quad y \geq 0 \\ \text{and } \int_0^x dx &= \int_0^y \exp(-y')dy' = 1 - \exp(-y), \\ y &= -\log_e(1-x). \end{aligned} \quad (11.75)$$

As x and $(1-x)$ have the same distribution, we can also write $y = -\log_e(x)$. It follows that Rayleigh variates can be generated as

$$y = \sqrt{-\log_e(x)}. \quad (11.76)$$

This result then produces a simple method for generating Gaussian variates. Gaussian variates could be obtained by solving Eq. (11.75) but this would be numerically difficult. An easier method is to recall that independent in-phase and quadrature Gaussian variates will have a Rayleigh distributed amplitude and random phase, uniformly distributed over 0 to 2π . Then using two uniformly distributed random numbers, x_1 and x_2 , we can produce two independent Gaussian random numbers, y_1 and y_2 , from:

$$y_1 = \sqrt{-\log_e(x_1)}\cos(2\pi x_2); \quad y_2 = \sqrt{-\log_e(x_1)}\sin(2\pi x_2). \quad (11.77)$$

These Gaussian variables will have a mean of 0 and variance of 0.5. The more general case of variables with mean μ and variance σ^2 are obtained by using:

$$g = \mu + \sqrt{2\sigma^2}y. \quad (11.78)$$

2.11.4.2 Generating correlated Gaussian random numbers

Gaussian random numbers with arbitrary autocorrelation functions can be generated by appropriate filtering. A simple filter example is the recurrence relation (AR process of order 1):

$$x_n = \eta x_{n-1} + \beta g_n, \quad (11.79)$$

where the g_n are zero mean, unit variance independent Gaussian numbers and η , β are constants ($\eta < 1$). It can be easily shown [25] that:

$$\begin{aligned} E \{x_n\} &= \eta^n x_0 \\ \text{and } E \{x_n^2\} &= \eta^{2n} x_0^2 + \frac{\beta^2 (1 - \eta^{2n})}{1 - \eta^2}. \end{aligned} \quad (11.80)$$

The terms in η^n represent transient terms in the filter output and go to zero as $n \rightarrow \infty$. The resultant random numbers x_n have a Gaussian PDF with zero mean and variance $\frac{\beta^2}{1 - \eta^2}$. More importantly, they will have an autocorrelation function:

$$E \{x_n x_{n+m}\} = \eta^{|m|} E \{x_n^2\}. \quad (11.81)$$

This method can be extended to more general filter functions. The filtering may also be done using an FFT. If we wish to generate random numbers with a power spectral density ($S\omega$), we can generate a batch of random Gaussian random numbers, $g_I + jg_Q$, representing complex data in the frequency domain with a uniform power spectrum. This data is then weighted by the amplitude of the required spectrum $\sqrt{S(\omega)}$ and the appropriate complex time domain sequence obtained by Fourier inversion.

2.11.4.3 Correlated non-Gaussian random numbers

Generating sequences of correlated random numbers with non-Gaussian PDFs is more complex. A sequence of random variates with a given PDF can be correlated by filtering, as above, but the resulting data will in general have a different PDF. For a given filter function and PDF, empirical relationships can sometimes be developed to approximately simulate the required characteristics [76, 77]. An example of a method for generating correlated Gamma distributed variates is given in [25]. Random Gamma variates can be generated from random Gaussian variates by the MNLT (Memoryless Non-linear Transform) method. If the Gaussian variates are correlated, the resulting PDF after the MNLT will be Gamma distributed as required, but with a changed autocorrelation function, ACF. It has been found that if the Gaussian process has an exponentially decaying ACF with decay time τ_G , then the resulting Gamma variates, with shape parameter v , will have a decay time τ_γ given by:

$$\frac{\tau_G}{\tau_\gamma} = 1 + \frac{0.15}{v^{0.7}}. \quad (11.82)$$

A more general process has been developed by Tough and Ward [75] to produce correlated variates with known PDF and ACF. This is illustrated here by the problem of generating correlated K distributed random numbers. Starting with random Gamma distributed variates, g , random Gamma variates can be produced by a MNLT so that

$$\begin{aligned} \frac{\Gamma(\nu, x)}{\Gamma(\nu)} &= 0.5 \operatorname{erfc}\left(g/\sqrt{2}\right), \\ x(g) &= Q_{\text{dist}}\left(0.5 \operatorname{erfc}\left(g/\sqrt{2}\right)\right), \\ \text{with } E\{x\} &= \nu; \quad E\{x\} = \nu(\nu + 1), \end{aligned} \quad (11.83)$$

where $\Gamma(\nu)$ is a Gamma function, $\Gamma(\nu, x)$ is an incomplete Gamma function and $\operatorname{erfc}()$ is the error function complement.

Independent K distributed random numbers representing clutter intensity can be generated by modulating exponentially distributed variates by the Gamma variates. If correlated Gaussian variates are used, the gamma variates will also be correlated. If the Gaussian variates have a normalized ACF (the correlation coefficient) given by

$$R_G(m) = \frac{E\{g_n g_{n+m}\}}{E\{g_n^2\}} \quad (11.84)$$

then the normalized ACF of the Gamma variates is given by:

$$E\{x_n x_{n+m}\} = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{R_G(m)^n}{2^n n!} \left(\int_{-\infty}^{\infty} e^{-y^2/2} H_n\left(y/\sqrt{2}\right) Q_{\text{dist}}\left(0.5 \operatorname{erfc}\left(y/\sqrt{2}\right)\right) dy \right)^2, \quad (11.85)$$

where $H_n(\cdot)$ is a Hermite polynomial, defined as

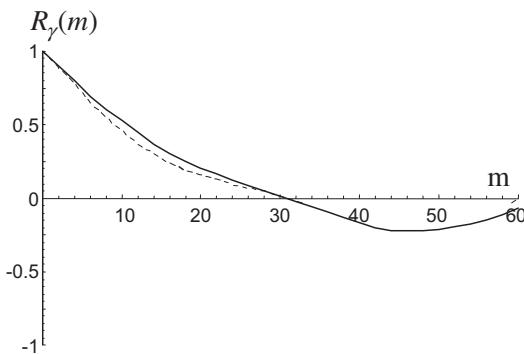
$$H_n(u) = (-1)^n \exp\left(u^2\right) \left(\frac{d}{du}\right)^n \exp\left(-u^2\right). \quad (11.86)$$

As an example, for $\nu = 0.5$,

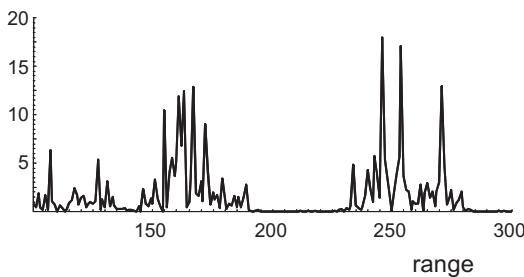
$$\begin{aligned} E\{x_n x_{n+m}\} &\approx 0.253 + 0.348 R_G(m) + 0.140 R_G(m)^2 + 0.013 R_G(m)^3 + 0.000054 R_G(m)^4 \\ &\quad + 0.000174 R_G(m)^5 + 9.36 \times 10^{-6} R_G(m)^6 + 8.87 \times 10^{-6} R_G(m)^7. \end{aligned} \quad (11.87)$$

Here $E\{x_n x_{n+m}\} = E^2\{x\} + (E\{x^2\} - E^2\{x\}) R_\gamma(m) = \nu^2 + R_\gamma(m)\nu$, where $R_\gamma(m)$ is the normalized ACF of the gamma variates. For $\nu \gg 10$, $R_G(m) \approx R_\gamma(m)$. As ν reduces, the relationship becomes non-linear and this produces a limitation for negative values of $R_\gamma(m)$. For example, for $\nu = 0.5$ and $R_G(m) = -1$ it is found that $R_\gamma(m) = -0.437$, which means that $R_\gamma(m) < -0.437$ cannot be simulated. This is not usually a significant limitation and this method is extremely powerful for generating random numbers with a wide range of PDFs and ACFs.

Figure 11.45 shows a normalized ACF for data with $\nu = 0.5$, showing the desired values and the values achieved with a sample of simulated data. Figure 11.46 shows an example of a time series of simulated correlated K distributed variates with this ACF.

**FIGURE 11.45**

Normalized ACF of correlated Gamma variates. Dashed line: original data; solid line: simulated data.

**FIGURE 11.46**

Correlated Gamma variates, with normalized ACF of Figure 11.45.

This method can readily be extended to two dimensions, as described in [25, 75].

In some situations, particularly for detection purposes, we may be interested in generating many independent vectors of correlated compound-Gaussian clutter. If the Time-on-Target (ToT) is short (tens of ms) the texture can be considered a constant in each vector but randomly changing from one vector to the other. In this case, the compound-Gaussian model degenerates into the Spherically Invariant Random Process (SIRP), proposed by Conte and Longo [12] for modeling the radar sea clutter. The samples of the SIRV constitutes a spherically invariant random vector (SIRV) whose PDF is given by

$$p_Z(\mathbf{z}) = \int_0^\infty \frac{1}{(2\pi\tau)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{z}^H \mathbf{M}^{-1} \mathbf{z}}{\tau}\right) p_\tau(\tau) d\tau, \quad (11.88)$$

where the vector $\mathbf{z} = [z(1) z(2) \dots z(m)]^T$ is the vector we want to generate. Then the vector \mathbf{z} we need is the product between the texture real random variable τ and a speckle complex Gaussian vector \mathbf{x} . Conditionally to the texture, the covariance matrix of the vector \mathbf{z} is the covariance matrix \mathbf{M} of the

speckle. The problem is then reduced to the generation of a Gaussian vector with a certain covariance matrix \mathbf{M} . To this purpose, we can use a linear filter as described at the beginning of this section (sequential method), or a batch method based on the invariance property of the Gaussian processes with respect to linear operations. Consider the vector $\mathbf{x} = \mathbf{Aw}$ where \mathbf{w} is a vector of independent identically distributed complex Gaussian random variables with zero-mean and unit variance. The covariance matrix of \mathbf{x} is $\mathbf{M} = E\{\mathbf{xx}^H\} = \mathbf{AA}^H$. There are many ways to find a matrix \mathbf{A} satisfying this relation. One of them is the Cholesky decomposition that results in a lower triangular matrix $\mathbf{A}=\mathbf{L}$ [78]. Summarizing, to generate the vector \mathbf{z} with must generate first \mathbf{w} , multiply it by the matrix \mathbf{A} and then by the random variable τ . The advantage of this method with respect to the sequential one is that it does not suffer the problem of the transient term.

2.11.5 Use of clutter models in radar design and analysis

2.11.5.1 Performance prediction

One of the purposes of clutter models is to predict the performance of radar systems. This is typically achieved by initial application of the radar equation to establish the relative power levels for a given situation of returns from targets and clutter and thermal noise.

The basic radar equation for noise-limited target detection performance can be written as

$$\frac{S}{N} = \frac{P_t G_t}{4\pi R^2} \frac{\sigma}{4\pi R^2} A_r \frac{1}{L} \frac{C}{p_n}, \quad (11.89)$$

where

- P_t peak transmitter power, W
- G_t antenna gain on transmit
- G_r antenna gain on receive
- σ target RCS, m^2
- A_r receive aperture, m^2 ; $A_r \approx \frac{G_r \lambda^2}{4\pi}$
- p_n noise power, W; $p_n = kT_0 F_n B$
- L system losses
- C pulse compression gain; $C \approx B\tau$
- k :Boltzmann's constant; $k = 1.38 \times 10^{-23} \text{ JK}^{-1}$
- T_0 Standard noise temperature, 290K.
- B receiver bandwidth, Hz
- F_n receiver noise figure
- τ radar pulse length, s

The basic radar hardware and waveform parameters can be measured accurately in the factory. The system losses will usually include the microwave losses in the radar hardware, which can be measured, and the propagation losses, which usually cannot be measured but must be estimated from the prevailing conditions and models. It is sometimes convenient to include signal processing "losses" in this term to

account for the non-ideal nature of subsequent signal processing compared with some idealized form used to calculate performance analytically.

For a single pulse return, detection performance can be described in terms of the probability of false alarm, P_{fa} , and probability of detection, P_d , following a detection threshold. For a square law detector, Gaussian clutter and an idealized fixed threshold detection threshold, t :

$$P_{fa} = \int_t^{\infty} \frac{1}{p_n} \exp\left(-\frac{x}{p_n}\right) dx = \exp\left(-\frac{t}{p_n}\right). \quad (11.90)$$

For a Swerling Case 1 or Case 2 target ([79], exponential PDF), with mean power s

$$P_d = \int_t^{\infty} \frac{1}{p_n + s} \exp\left(-\frac{x}{p_n + s}\right) dx = \exp\left(-\frac{t/p_n}{1 + SNR}\right) = \exp\left(\frac{\log_e(P_{fa})}{1 + SNR}\right), \quad (11.91)$$

where $SNR = s/p_n$, the signal-to-noise ratio.

For other target models, equivalent PDFs of target-plus-noise must be derived. For example, a non-fading target in noise results in a Rice PDF [33].

If the returns include clutter, then the clutter to noise ratio can be determined by replacing the target RCS by that of the clutter return, σ_c . As described in Section 2.11.2.1.1, for surface distributed clutter, $\sigma_c = \sigma^0 A_c$. Appropriate values of σ^0 are usually derived from empirical models, as described in Section 2.11.2.2.1 and 2.11.2.3.1. The values of P_{fa} and P_d must now be calculated for a return that includes clutter-plus-noise. This requires a knowledge of the PDFs of the envelope of clutter-plus-noise and target-plus-clutter-plus-noise, which can sometimes be difficult to derive. For a known clutter PDF, the PDF of clutter-plus-noise will generally have to be calculated numerically. In the particular case of the compound K distribution model, the PDF of clutter-plus-noise can be defined exactly [80]. Similarly, the PDF of the envelope of target-plus-clutter-plus-noise is required to calculate P_d . Again, this can be calculated for the compound K distribution model for a wide range of target fading models [25].

2.11.5.1.1 Signal processing and design of detectors

The radar equation provides a means of calculating the signal-to-clutter-plus-noise ratio, $SCNR$, and the clutter-to-noise ratio, CNR , for a single pulse return. For many radar systems, the returns from several successive pulses are integrated, to benefit from more radar power on the target. These pulses may sometimes be at different frequencies (frequency agility). In a typical example, during a dwell on a target, a pulse Doppler radar may transmit a burst of pulses with a fixed PRF and frequency, followed by successive bursts at different PRFs and frequencies. Returns from trains of pulses are used in subsequent signal processing to improve target detection and discrimination against clutter and noise.

To calculate P_{fa} and P_d following detection processing in the radar requires knowledge of the effect of the processing on the amplitude statistics immediately prior to the final detection threshold or decision. For coherent processing with Doppler filter banks (typically achieved with an FFT over a fixed frequency burst of pulses), the Doppler spectra of target and clutter must be known to calculate the subsequent $SCNR$ and CNR in each Doppler filter. The outputs from each Doppler filter may then be detected and then subsequent detections from successive bursts may be non-coherently integrated. Again the PDF of the integrated signals must be calculated. In the case of clutter, it is very important to understand any

correlation between the samples that are being integrated. Thermal noise will always be independent from pulse to pulse. For clutter, the situation is more complex, with different components fluctuating at different rates. For example, pulse-to-pulse frequency agility can decorrelate the speckle component of sea clutter [10], but the underlying mean level component is not affected and may effectively be constant over a dwell. The target returns may also fluctuate at different rates over a dwell. This type of behavior is usually approximated by estimating the effective number of independent samples observed over a dwell period. Again, for the compound K distribution model it is possible to model the effects of pulse-to-pulse integration taking into account the effective number of independent samples of target, thermal noise and the two clutter components, each of which may be different [25].

The design of optimum and suboptimum, adaptive and non-adaptive, coherent and non-coherent detectors for detecting point-targets embedded in Gaussian disturbance is a mature and well understood topic. There are many excellent books and papers addressing different aspect of the problem. Some references are listed in [81]. Conversely, the detection of targets embedded in non-Gaussian clutter is a less mature topic, even though there are many references published in the last 20 years (see for instance [19, 22, 25, 82–92] and references therein). The main problems that must be taken into account are the estimation of the disturbance covariance matrix, the estimation of the parameters of the clutter distribution [17, 25, 93–95], the non-stationarity of the clutter in the case of sea scattering [96] and the site-dependence of the clutter in the case of ground scattering [97, 98]. New radar systems must be able to rapidly adapt to changing conditions. Knowledge or estimation of clutter conditions can assist with the selection of both transmitted waveforms and detection algorithms.

2.11.5.2 The specification and measurement of radar performance

Another important role for clutter models relates to the specification and measurement of real radar performance. This is discussed in detail in [4, 25, 99–101].

Modern radars, especially those that use adaptive processing, are extremely complex and their performance is difficult to specify. A customer needs to ensure that those features of the radar that will define its ability to achieve the required operational capability are adequately specified. The ability to do this is in turn dependent upon a detailed understanding of the likely radar performance in different environmental and operational conditions. The supplier, however, needs to have a simple set of performance criteria upon which to base the radar design. One way to bridge this gap is to use a model that relates the basic performance of the radar to its behavior in complex operational environments. An important component of such models is likely to be the behavior of radar clutter.

A typical radar life cycle involves the use of models by the customer, the supplier and the user at various times. Table 11.20 shows a typical range of applications for models [4] during a radar's life.

2.11.5.2.1 Clutter models for performance acceptance and trials analysis

An increasingly important role for models is for the analysis of radar trials and the acceptance by a customer of a new radar into service.

Acceptance trials are an integral part of any radar development program. These are intended to demonstrate to the customer that the radar provides the required functionality and is “fit for purpose.” It is also usually desired that trials be used to quantify the key performance parameters of the radar (e.g., detection range against a given target type). However, in many cases the reliable measurement

Table 11.20 Applications of Clutter Models during the Life-Cycle of a Radar

Modelers	Life-Cycle Stages	Requirements for Clutter Models
Customer	Requirements definition	Generic performance modeling; High level synthetic environment simulations
Customer and potential suppliers	Modeling of potential performance	Performance predictions of different designs Comparison of competing designs;
Supplier	System and algorithm design	Performance prediction; Algorithm design (insight from models); Algorithm simulation Hardware and software stimulation.
Customer and supplier	Performance assessment and acceptance trials	Analysis of trials data; comparison of trials results with performance predictions
Radar user	In-service tactics and training	Simulation and hardware stimulation
Customer, supplier and user	In-service upgrades	Performance predictions; Algorithm simulation; synthetic environment simulation

of performance on trials is very difficult and may at the very least require many hours of trials. This is particularly true for radars required to detect very small targets in difficult clutter conditions or for radars that are continuously adapting their parameters (such as signal processing, waveforms, antenna beamshape, etc.) in response to changing environmental conditions. The customer may specify that a target of given RCS must be detected at a certain range, in particular environmental conditions (rainfall rate, sea state, etc.). Given an agreed set of models, such a situation could be modeled and provide reasonably accurate estimates of performance, within the limitations of the models used. However, measurement of these parameters in a trial can be very difficult. Even the apparently simple task of measuring the noise-limited detection range of a moving target is not straightforward [102]. If the original requirement was, say, to detect a target in sea state 3, in 1mm/h of rain, the radar designer would have to put a very large margin on expected trials results to allow for uncertainty in assessing whether sea state three is present, what the rainfall rate was (at the target and over the intervening path), combined with the large range of clutter characteristics, such as σ^0 associated with a given sea state. Although the radar designer would be able to model performance for given clutter model parameters, it would not usually be possible to relate these to the conditions actually prevailing on a trial.

One solution is to combine trials measurements with modeling and factory measurements. Factory measurements can determine the basic radar parameters that determine noise-limited performance in, say, clear air. Modeling can predict the expected mean performance for a wide range of target sizes and clutter characteristics, which could never be tested in practical trials of reasonable duration. The empirical models that are used in radar design have themselves been developed from many hours of trials data, representing an average expected conditions that do not then have to be reproduced in acceptance

trials. Then on a trial, detection performance can be measured (detection range, for example) for the target and environmental conditions applying on the trial. At the same time, the raw radar returns can be recorded for subsequent analysis. This analysis can yield the actual target and clutter characteristics that were present, provided that the radar noise-limited performance is suitably calibrated. These observed characteristics can then be input to the performance prediction model with the results compared with those actually achieved in the trial.

If raw radar data is recorded on a trial, it should also be possible to replay it through the radar signal processing, to assess the effects of different detection algorithms. In addition, it may be possible to inject synthetic targets, to further investigate detection performance.

This type of approach, combining trials with performance prediction modeling, can allow quantitative assessment of performance that would be effectively impossible from trials alone (due to uncertainty over the prevailing conditions and the time and cost of extensive trials). Of course, it is far from straightforward to achieve. It implies the measurement of detection ranges and false alarms rates from trials observations, and the estimation of clutter and target characteristics from recorded data. Each of these measurements is subject to its own measurement errors and the combined results would have to be analyzed in detail to assess the buyer and seller risks [25] associated with eventual equipment acceptance or rejection. The analysis methods and also the underlying performance prediction model must be agreed in detail between customer and supplier (preferably well in advance of any acceptance trials actually take place).

2.11.6 Conclusions

Modeling and analysis of clutter, as described in this tutorial, is a very important and complex matter. The characteristics of clutter depend on the environmental conditions, the viewing geometry and the radar characteristics. In general, there are not unique empirical or theoretical models that exhibit a satisfactory fit to real data under all possible conditions. Nevertheless, we have shown some general trends focusing particularly on sea and ground clutter recorded by high resolution pulsed radars. For low grazing angles and high spatial resolution, sea clutter amplitude statistics are not Gaussian and the returns exhibit some distinct amplitude peaks, called spikes, that can be confused by a radar system for targets. The spikiness of sea clutter increases with resolution and is more evident in *HH* polarized data. A good and flexible statistical model for sea clutter is the compound-Gaussian model, which comprises among its particular cases the popular Rayleigh, Weibull and K models. One of the advantages of the compound-Gaussian model is its mathematical tractability that helps in designing optimum and suboptimum detectors and the possibility of describing separately the amplitude PDF and the correlation characteristics. For sea clutter, this family of distributions has been found to give a good fit over a wide range of conditions, although this model may sometimes fail at very high spatial resolutions. These models can also be used successfully for ground clutter, especially when the ground clutter is scattered by discrete objects as pylons, fences, tree trunks. In some scenarios, when the ground surface is covered by wind-blown trees and vegetation, the clutter still exhibits Gaussian statistics, even at high resolutions.

Concerning the clutter spectrum, many models have been proposed in the literature in the last 50 years. We have presented some results with AR models, Gaussian, power law and exponential PSDs. Again, there is not a unique optimum model for all the conditions. The AR model is very flexible and shows a

good fitting with sea clutter and ground clutter in some cases, but the exponential one seems to be the best for windblown tree clutter recorded by high dynamic-range radars.

Finally, we cannot make a “final point” or a definite conclusion on this complex matter because the study of clutter models is a continuing research topic. As radar systems improve, new features of clutter and target scattering may be observed and new analyses are then necessary to develop more accurate models of these characteristics and to improve the design of target detectors.

Relevant Theory: Signal Processing Theory

See Vol. 1, Chapter 4 Random Signals and Stochastic Processes

References

- [1] F.E. Nathanson, Radar Design Principles, McGraw Hill, 1969.
- [2] M.I. Skolnik, Introduction to Radar Systems, third ed., McGraw Hill, 2001.
- [3] P. Tait, Introduction to Radar Target Recognition, IET, 2006.
- [4] K.D. Ward, S.Watts, Use of sea clutter models in radar design and development, *IET Radar Sonar Navig.* 4 (2) (2010) 146–157 (Special Issue on Radar Clutter).
- [5] S. Watts, Radar sea clutter modelling—recent progress and future challenges, in: International Radar Conference 2008, Adelaide, September 2008 (Invited Paper).
- [6] J.M. Blythe, D.E. Rice, W.L. Attwood, The CFE clutter model, with application to automatic detection, *Marconi Rev.* 32 (1969) 185–206.
- [7] G.R. Valenzuela, M.B. Lang, On the statistics of sea clutter, *NRL Report 7349*, 1971.
- [8] G.V. Trunk, Radar properties of non-Rayleigh sea clutter, *IEEE Trans. AES* 8 (1972) 196–204.
- [9] K.D. Ward, Compound representation of high resolution sea clutter, *Electron. Lett.* 17 (1981) 561–563.
- [10] K.D. Ward, C.J. Baker, S. Watts, Maritime surveillance radar Part 1: radar scattering from the ocean surface, *Proc. IEE* 137 (2) (1990) 51–62 (part F).
- [11] T.J. Barnard, D.D. Weiner, Non-Gaussian clutter modeling with generalized spherically invariant random vectors, *IEEE Trans. Signal Process.* 44 (10) (1996).
- [12] E. Conte, M. Longo, Characterisation of radar clutter as a spherically invariant random process, *IEE Proc.-F* 134 (2) (1987) 191–197.
- [13] E. Conte, A. De Maio, C. Galdi, Statistical analysis of real clutter at different range resolutions, *IEEE Trans. Aerosp. Electron. Syst.* 40 (3) (2004) 903–918.
- [14] F.A. Fay, J. Clarke, R.S. Peters, Weibull distribution applied to sea-clutter, in: Proceedings of the IEE Conference Radar '77, London, UK, 1977.
- [15] A. Farina, F. Gini, M. Greco, L. Verrazzani, High resolution sea clutter data: a statistical analysis of recorded live data, *IEE Proc.-F* 144 (3) (1997) 121–130.
- [16] E. Jakeman, P.N. Pusey, A model for non-Rayleigh sea echo, *IEEE Trans. Antenna Propag.* AP-24 (1976) 806–814.
- [17] M. Rangaswamy, D.D. Weiner, A. Ozturk, Non-Gaussian vector identification using spherically invariant random processes, *IEEE Trans. Aerosp. Electron. Syst.* 29 (1) (1993) 111–124.
- [18] M. Rangaswamy, D.D. Weiner, A. Ozturk, Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. Aerosp. Electron. Sys.* 31 (1) (1995) 106–115.
- [19] M. Rangaswamy, J.H. Michels, D.D. Weiner, Multichannel detection algorithm for correlated non-Gaussian random processes based on innovations, *IEEE Trans. Signal Process.* 43 (8) (1995) 1915–1922.

- [20] M. Rangaswamy, J.H. Michels, A parametric multichannel detection for correlated non-Gaussian random processes, in: Proceedings of the National Radar Conference, Syracuse, NY, USA, May 1997, pp. 349–354.
- [21] M. Rangaswamy, J.H. Michels, Adaptive signal processing in non-Gaussian noise backgrounds, in: Proceedings of the Ninth IEEE-SSAP Workshop, Portland, OR, September 1998.
- [22] K.J. Sangston, K.R. Gerlach, Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. Aerosp. Electron. Syst.* 30 (2) (1994) 330–340.
- [23] M. Sekine, Y. Mao, Weibull Radar Clutter, Peter Peregrinus Ltd., London 1990.
- [24] D. Middleton, New physical-statistical methods and models for clutter and reverberation: the KA distribution and related probability structures, *IEEE J. Oceanic Eng.* 24 (3) (1999) 261–284.
- [25] K.D. Ward, R.J.A. Tough, S. Watts, Sea Clutter: Scattering, the K Distribution and Radar Performance, Institution of Engineering and Technology, 2006.
- [26] V.W. Pidgeon, Doppler dependence of radar sea return, *J. Geophys. Res.* 73 (1968) 1333–1341.
- [27] P.H.Y. Lee, J.D. Barter, K.L. Beach, C.L. Hindman, B.M. Lake, H. Rungaldier, J.C. Shelton, A.B. Williams, R. Yee, H.C. Yuen, X-band microwave backscattering from ocean waves, *J. Geophys. Res.* 100 (C2) (1995) 2591–2611.
- [28] P.H.Y. Lee, J.D. Barter, K.L. Beach, E. Caponi, C.L. Hindman, B.M. Lake, H. Rungaldier, J.C. Shelton, Power spectral lineshapes of microwave radiation backscattered from sea surfaces at small grazing angles, *IEE Proc. F* 142 (5) (1995) 252–258.
- [29] J.B. Billingsley, A. Farina, F. Gini, M. Greco, L. Verrazzani, Statistical analyses of measured radar ground clutter data, *IEEE Trans. Aerosp. Electron. Syst.* 35 (2) (1999) 579–593.
- [30] S. Haykin, A. Steinhardt, Adaptive Radar Detection and Estimation, Wiley, New York, 1992.
- [31] CSIR, small boat detection research, <<http://www.csir.co.za/smallboatdetection/>>.
- [32] J.B. Billingsley, Low-Angle Radar Land Clutter—Measurements and Empirical Models, William Andrew Publishing, Norwich, NY, 2002.
- [33] D.K. Barton, Radar System Analysis and Modeling, Artech House, 2005.
- [34] W.H.D. Long, D. Mooney, W.A. Skillman, Pulse Doppler radar, in: M. Skolnik (Ed.), Radar Handbook, second ed., McGraw Hill, New York 1990, pp. 17.11–17.16 (Chapter 17).
- [35] L. Rosenberg, D.J. Crisp, N.J. Stacy, Analysis of the kk-distribution with medium grazing angle sea-clutter, *IET Proc. Radar Sonar Navig.* 4 (2010) 209–222.
- [36] J.W. Wright, A new model for sea clutter, *IEEE Trans. AP* 16 (1968) 217–223.
- [37] F.G. Bass, I.M. Fuks, A.E. Kalmykov, I.E. Ostrovsky, A.D. Rosenberg, Very high frequency radiowave scattering by a disturbed sea surface, *IEEE Trans. Antenna Propag. AP-16* (1968) 554–568.
- [38] P. Gerstoft, L.T. Rogers, L.J. Wagner, W.S. Hodgkiss, Estimation of radio refractivity structure using radar clutter, in: 30th International Conference on Radar Meteorology, München Germany, 19–24 July 2001.
- [39] M.M. Horst, F.B. Dyer, M.T. Tuley, Radar sea clutter model, in: IEEE International Conference Antennas and Propagation, November 1978, pp. 6–10.
- [40] L.B. Wetzel, Sea Clutter, in: M. Skolnik (Ed.), Radar Handbook, second ed., McGraw Hill, New York, 1990 (Chapter 13).
- [41] M. Greco, F. Bordoni, F. Gini, X-band sea clutter non-stationarity: the influence of long waves, *IEEE J. Ocean Eng.* 29 (2) (2004) 269–283 (Special Issue on Non-Rayleigh Reverberation and Clutter).
- [42] S. Watts, A new method for the simulation of coherent sea clutter, in: IEEE International Conference, RadarCon 2011, May 2011.
- [43] M.A. Ritchie, K. Woodbridge, A.G. Stove, Analysis of sea clutter distribution variation with Doppler using the compound K-distribution, in: International Conference on IEEE Radar 2010, May 2010, pp. 495–499.
- [44] R.R. Boothe, The Weibull distribution applied to the ground clutter backscattering coefficient, US Army Missile Command, RE-TR-69-15, ADA691109, 1969.
- [45] M.W. Long, Radar reflectivity of Land and Sea, third ed., Artech House, 2001.

- [46] R.K. Moore, Ground echo, in: M. Skolnik (Ed.), *Radar Handbook*, third ed., McGraw Hill, New York, 2008 (Chapter 16).
- [47] F.T. Ulaby, M.C. Dobson, *Handbook of Radar Scattering Statistics for Terrain*, Artech House, Norwood MA, 1989.
- [48] N.C. Currie, Clutter characteristics and effects, in: J.L. Eaves, E.K. Ready (Eds.), *Principles of Modern Radar*, Van Nostrand Reinhold, New York, 1987, pp. 281–340.
- [49] W.C. Morchin, *Airborne Early Warning Radar*, Artech House, 1990.
- [50] L.V. Blake, Prediction of radar range, in: M. Skolnik (Ed.), *Radar Handbook*, first ed., McGraw Hill, 1970.
- [51] M. Greco, F. Gini, M. Rangaswamy, Statistical analysis of real polarimetric clutter data at different range resolutions, *IEE Proc. Radar Sonar Navig.* (2006).
- [52] B. Currie, Private Communications, October 2003.
- [53] S.M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, vol. I, Prentice Hall, 1993.
- [54] E.L. Lehmann, *Testing Statistical Hypotheses*, second ed., Springer texts in statistics, New York, 2010.
- [55] T. Nohara, S. Haykin, Canadian East coast radar trials and the K-distribution, *IEE Proc.-F* 138 (2) (1991) 80–88.
- [56] B. Sadler, G.B. Giannakis, K.S. Lii, Estimation and detection in the presence of non-Gaussian noise, *IEEE Signal Process.* 42 (10) (1994) 2729–2741.
- [57] F. Gini, A Cumulant-based adaptive technique for coherent radar detection in a mixture of K-distributed clutter and Gaussian disturbance, *IEEE Trans. Signal Process.* 45 (6) (1997) 1507–1519.
- [58] F. Gini, M. Greco, Texture modelling, estimation, and validation using measured sea clutter data, *IEE Proc. Radar Sonar Navig.* 149 (3) (2002).
- [59] S. Haykin, R. Bakker, B.W. Currie, Uncovering nonlinear dynamics—the case study of sea clutter, in: *Proceedings of the IEEE* 90 (5) (2002) 860–881.
- [60] G.R. Valenzula, Theories for the interaction of electromagnetic waves and oceanic waves—a review, *bound. Layer Meteorol.* 13 (1–4) (1978) 61–65.
- [61] W.J. Plant, W.C. Keller, Evidence of Bragg scattering in microwave Doppler spectra of sea return, *J. Geophys. Res.* 95 (C9) (1990) 16299–16310.
- [62] W.J. Plant, W.C. Keller, The two-scale radar wave probe and SAR imagery of the ocean, *J. Geophys. Res.* 88 (C14) (1983) 9776–9784.
- [63] C.L. Rino, E. Eckert, A. Siegel, T. Webster, A. Ochadlick, M. Rankin, J. Davis, X-band low-grazing-angle ocean backscatter obtained during LOGAN 1993, *IEEE J. Oceanic Eng.* 22 (1) (1997) 18–26.
- [64] A.D. Rozenberg, D.C. Quigley, W.K. Melville, Laboratory study of polarized microwave scattering by surface waves at grazing incidence: the influence of long waves, *IEEE Trans. Geosci. Remote Sens.* 34 (6) (1996) 1331–1342.
- [65] D. Walker, Experimentally motivated model for low grazing angle radar Doppler spectra of the sea surface, *IEE Proc. Radar Sonar Navig.* 147 (3) (2000) 114–120.
- [66] D. Walker, Doppler modelling of radar sea clutter, *IEE Proc. Radar Sonar Navig.* 148 (2) (2001).
- [67] D.B. Trizna, A model for Doppler peak spectral shift for low grazing angle sea scatter, *IEEE J. Oceanic Eng.* 10 (4) (1985) 368–375.
- [68] A.T. Jessup, W.K. Melville, W.C. Keller, Breaking waves affecting microwave backscatter. 1. Detection and verification, *J. Geophys. Res.* 96 (C11) (1991) 20547–20559.
- [69] A. Drosopoulos, Description of the OHGR database, Technical Note No. 94–14, Defence Research Establishment Ottawa, December 1994.
- [70] S. Haykin, in: Simon Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. I, Prentice Hall, Englewood Cliffs, NJ 07632, 1991.
- [71] C.W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, Englewood Cliffs, NJ 07632, 1992.

- [72] P.J. Kim, R.I. Jennrich, Tables of the exact sampling distribution of the two sample Kolmogorov-Smirnov criterion ($D_{mn} < m < n$), in: H.L. Harter, D.B. Owen (Eds.), Selected Tables in Mathematical Statistics, vol. I, Providence, RI: American Mathematical Society, 1973.
- [73] M. Greco, F. Gini, A. Farina, J.B. Billingsley, Validation of windblown ground clutter spectral shape by analysis of clutter cancellation in the presence of measured clutter data, *IEEE Trans. Aerosp. Electron. Syst.* 37 (2) (2001) 538–548.
- [74] K.S. Chen, A.K. Fung, Frequency dependence of backscattered signals from forest components, *IEE Proceedings, Pt. F* 142 (6) (1995) 310–315.
- [75] R.J.A. Tough, K.D. Ward, The correlation properties of gamma and other non-Gaussian processes generated by memoryless nonlinear transformation, *J. Phys. D* 32 (1999) 3075–3084.
- [76] A. Farina, A. Russo, F.R. Studer, Coherent radar detection in log-normal clutter, *IEE Proc.-F* 133 (1) (1986) 39–54.
- [77] A. Farina, A. Russo, F. Scannapieco, S. Barbarossa, Theory of radar detection in coherent Weibull clutter, *IEE Proc.-F* 134 (2) (1987) 174–190.
- [78] Gene H. Golub, Charles F. Van Loan, Matrix computations, third ed., Johns Hopkins University Press, 1996 (Section 4.2, ISBN: 0-8018-5414-8).
- [79] P. Swerling, Probability of detection for fluctuating targets, *IRE Trans. IT-6* (1960) 269–308.
- [80] S. Watts, Radar detection prediction for targets in both K-distributed sea clutter and thermal noise, *IEEE Trans. AES-23* (2) (1987) 40–45.
- [81] F. Gini, A. Farina, M. Greco, Selected list of references on non-Gaussian radar detection, *IEEE Trans. Aerosp. Electron. Syst.* 37 (1) (2001) 329–359.
- [82] E. Conte, G. Ricci, Performance prediction in compound-Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 30 (2) (1994) 611–616.
- [83] E. Conte, M. Lops, G. Ricci, Asymptotically optimum radar detection in compound-Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 31 (2) (1995) 617–625.
- [84] E. Conte, A. De Maio, C. Galati, Signal detection in compound-Gaussian noise: Neyman-Pearson and CFAR detectors, *IEEE Trans. Signal Process.* 48 (2) (2000) 419–428.
- [85] F. Gini, M.V. Greco, A. Farina, P. Lombardo, Optimum and mismatched detection against K-distributed plus Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 34 (3) (1998) 860–876.
- [86] F. Gini, M. Greco, A suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 35 (3) (1999) 1095–1104.
- [87] F. Gini, M. Greco, A. Farina, Clairvoyant and adaptive signal detection in non-Gaussian clutter: a data-dependent threshold interpretation, *IEEE Trans. Signal Process.* 47 (6) (1999) 1522–1531.
- [88] S. Haykin, D.J. Thomson, Signal detection in nonstationary environment reformulated as an adaptive pattern classification problem, *Proceedings of the IEEE* 86 (11) (1998) 2325–2344.
- [89] C.D. Richmond, Performance of a class of adaptive detection algorithms in non-homogeneous environment, *IEEE Trans. Signal Process.* 48 (5) (2000) 1248–1262.
- [90] K.J. Sangston, F. Gini, M. Greco, A. Farina, Structures for optimum and suboptimal coherent radar detection in compound-Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 35 (2) (1999) 445–458.
- [91] K.J. Sangston, F. Gini, M. Greco, Coherent radar detection in heavy-tailed compound-Gaussian clutter, *IEEE Trans. Aerosp. Electron. Syst.* 42 (1) (2012) 64–77.
- [92] S. Watts, C.J. Baker, K.D. Ward, Maritime surveillance radar. Part 2: Detection performance prediction in sea clutter, in: *IEE Proceedings-F*, vol. 137, April 1990, pp. 63–72.
- [93] D.R. Iskander, A.M. Zoubir, Estimation of the parameters of the K-distribution using higher order and fractional moments, *IEEE Trans. Aerosp. Electron. Syst.* 35 (4) (1999) 1453–1457.
- [94] P. Lombardo, C.J. Oliver, Estimation of texture parameters in K-distributed clutter, *IEE Proc. F* 141 (4) (1994) 196–204.

- [95] R.S. Raghavan, A method for estimating parameters of K-distributed clutter, *IEEE Trans. Aerosp. Electron. Syst.* 27 (2) (1991) 238–246.
- [96] M. Greco, P. Stinco, F. Gini, M. Rangaswamy, Impact of sea clutter non-stationarity on disturbance covariance matrix estimation and CFAR detector performance, *IEEE Trans. Aerosp. Electron. Syst.* 46 (3) (2010) 1502–1513.
- [97] H.D. Griffiths, Knowledge-based solutions as they apply to the general radar problem, in: Proceedings of the RTO NATO Lecture Series 233 on Knowledge-Based Radar Signal and Data Processing, Rome, Italy, 6–7 November, 2003.
- [98] M. Lops, Hybrid clutter-map/L-CFAR procedure for clutter rejection in nonhomogeneous environment, *IEE Proc.-F* 143 (1996) 239–245.
- [99] S. Watts, Specification and measurement of performance for airborne maritime surveillance radars, in: International Radar Conference 1999, Brest, 17–21 May, 1999 (Session 2.6).
- [100] S. Watts, H.D. Griffiths, J.R. Holloway, A.M. Kinghorn, D.G. Money, D.J. Price, A.M. Whitehead, A.R. Moore, M.A. Wood, D.J. Bannister, The specification and measurement of radar performance, in: IEE International Radar Conference 2002, 14–17 October 2002, IEE Publication No.490, pp. 542–546.
- [101] S. Watts, H.D. Griffiths, J.R. Holloway, A.M. Kinghorn, D.G. Money, D.J. Price, A.M. Whitehead, A.R. Moore, M.A. Wood, D.J. Bannister, The specification and measurement of radar performance—future research challenges, *J. Defence Sci.* 8 (2) (2003) 83–91.
- [102] A.G. Stove, D.L. Hurd, Performance evaluation for modern radars, in: Proceedings of IEEE International Radar Conference, Adelaide, September 2003, pp. 547–53 (paper 44).

Space-Time Adaptive Processing for Radar

12

William L. Melvin

Georgia Institute of Technology, Atlanta, GA, USA

2.12.1 Introduction

Space-time adaptive processing (STAP) is a two-dimensional, adaptive filtering technique foundational to modern radar system design and implementation. Generally, spatial sampling is given by the outputs of an array of antenna elements (sometimes called “channels”); the Fourier transform of space yields angle. There are essentially two time measurements in radar: fast-time and slow-time. Fast-time samples correspond to outputs of the analog-to-digital converter abiding with the Nyquist rate for the reflected radar waveform; the Fourier transform of fast-time is radio-frequency (RF), which relates to radar range. The processor collects slow-time voltage samples at a given range realization from distinct, reflected, pulsed transmissions; the Fourier transform of slow-time yields Doppler frequency. Depending on the statistical properties of the interference source, either slow-time or fast-time is applicable. For example, the adaptive suppression of ground clutter received by airborne or spaceborne radar requires joint spatial and slow-time adaptivity since clutter returns appear correlated in the angle-Doppler domain, but are uncorrelated in fast-time. Fast-time adaptive degrees of freedom play in role in wideband and terrain-bounce interference mitigation in both aerospace and surface-based radar systems.

Suppose we consider adaptive mitigation of Doppler-spread ground clutter returns in further detail. This problem arises in airborne and spaceborne radar systems searching for moving targets. The clutter returns are generally much stronger than the target signal, in many cases by orders of magnitude, thereby rendering detection in certain regions of the range-Doppler map virtually impossible using conventional antenna beamforming and Doppler processing techniques. Spatial and Doppler beamforming represent one-dimensional matched filtering operations. The matched filter maximizes signal-to-noise ratio (SNR), thereby maximizing probability of detection for a fixed false alarm rate for uncorrelated, circular Gaussian disturbance [1,2]. STAP is needed when clutter and radio frequency interference (RFI)—colored noise—are present; the STAP applies a space-time weighting incorporating estimated characteristics of the interference environment to asymptotically maximize signal-to-interference-plus-noise ratio (SINR), a sufficient condition to maximize the probability of detection for a fixed false alarm rate under the assumption the disturbance is circular Gaussian [3].

Fixed points on the Earth’s surface exhibit a Doppler frequency shift consistent with their angular displacement from the platform velocity vector, \mathbf{v}_p . In turn, this angular displacement corresponds to a specific angle of arrival at the antenna array. Thus, given a particular azimuth and elevation angle in the antenna coordinate system, the clutter Doppler frequency is precisely specified, and vice versa.

The clutter angle-Doppler region of support is described as a ridge in airborne and spaceborne radar. This ridge is a line in sidelooking radar, opens up into an ellipse for varying degrees of platform yaw, and becomes a circle for the forward-looking array. Figure 12.1 provides example power spectral density (PSD) and minimum variance distortionless response (MVDR) super-resolution plots of the clutter angle-Doppler region of support for the side-looking array radar (SLAR), a forward-looking array radar (FLAR), and an array with forty-five (45°) of yaw. Typical radar parameters are used in this example and will be described in further detail in subsequent sections.

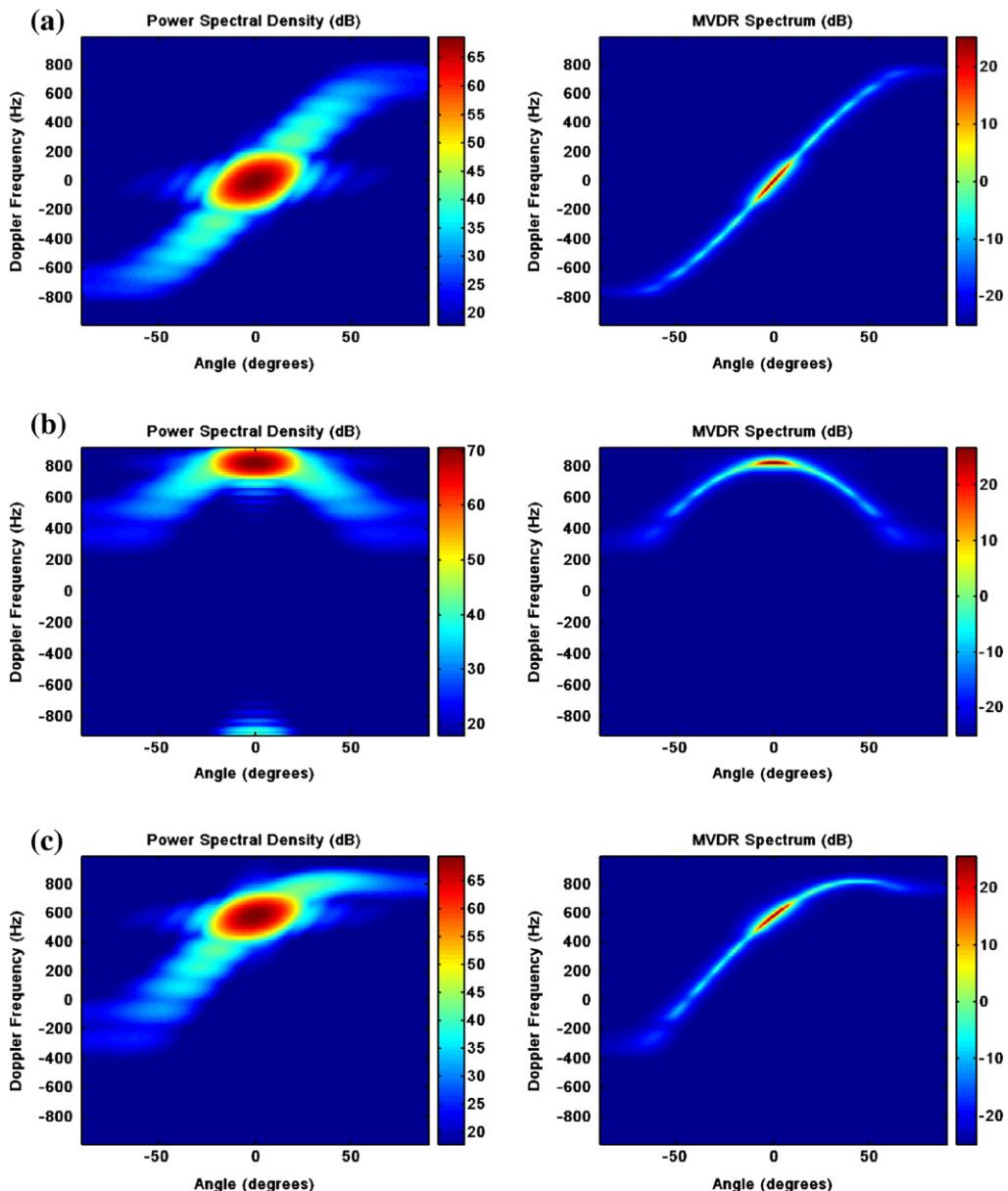
The PSD plots are a result of classical, two-dimensional, Fourier analysis. Resolution is limited in Fourier analysis by roughly the wavelength of the signal divided by the aperture length; to resolve more closely spaced signals, either the wavelength must be decreased for a fixed aperture size, or the physical aperture must be increased for a fixed wavelength. Super-resolution spectral estimators overcome the diffraction limitations of Fourier-based methods, oftentimes providing resolution enhancement by a factor of two to five, or even more. It turns out that STAP is intimately related to the MVDR spectrum, as discussed in Section 2.12.2.2. Hence, STAP likewise exhibits super-resolution properties, enabling it to detect slow moving targets buried in mainbeam clutter or targets closely spaced to other interference sources.

As mentioned, STAP coherently combines spatial and temporal voltage samples from a particular range realization using a data-dependent weighting to maximize SINR. The corresponding adaptive filter frequency response exhibits sharp nulls along the clutter region of support while maximizing the gain at another, specified angle-Doppler location where a potential target might exist. A target with a Doppler frequency outside of the mainlobe clutter Doppler spread (given as the highest power, circular region in the PSD) is called *exoclutter*, whilst targets within the mainlobe clutter spread are *endoclutter*. The STAP frequency response null is the inverse of the super-resolution contour shown in Figure 12.1. Hence, STAP itself exhibits super-resolution properties, providing detection performance within the diffraction limits of the traditional space-time aperture. This important characteristic makes STAP invaluable on airborne and spaceborne platforms, enabling drastic reduction in antenna aperture length for a desired minimum detectable velocity (MDV). The asymptotic STAP frequency responses for each of the PSDs in Figure 12.1 are given in Figure 12.2. (The asymptotic STAP response is optimal, since this is the weighting that precisely maximizes SINR.) The filter is tuned to broadside and 400 Hz Doppler.

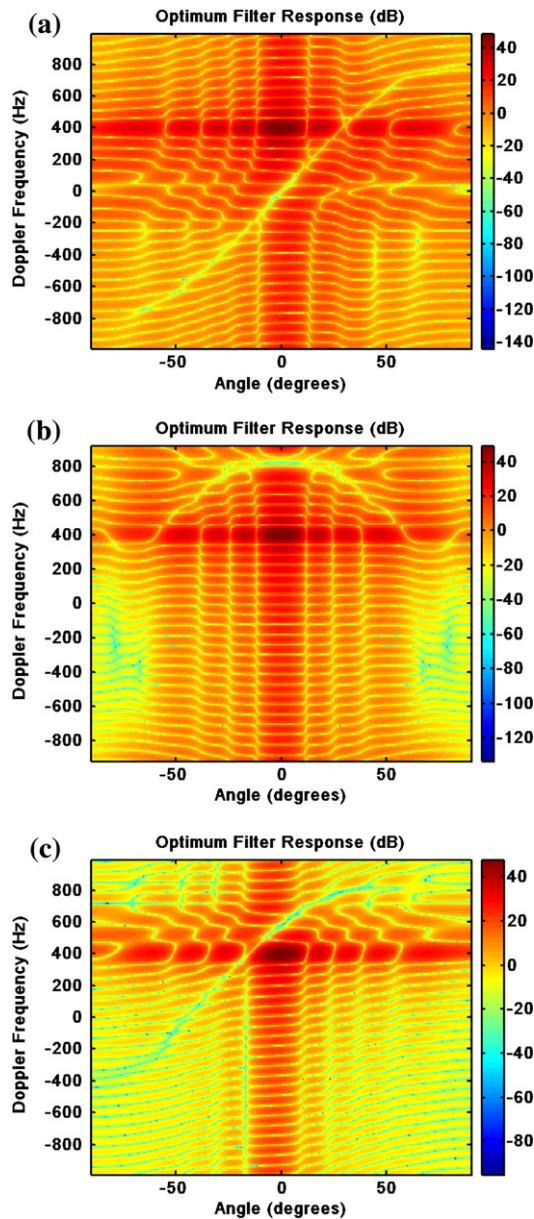
The MDV is the lowest target range-rate where sufficient SINR exists to permit a specified probability of detection for an acceptable probability of false alarm. Conversely, the MDV occurs at those points where the loss in SINR due to clutter is maximally acceptable. Figure 12.3 provides an example of the loss in SINR due to clutter for the radar array configuration of Figures 12.1 and 12.2. SINR loss will be described further in Section 2.12.2. Noise-limited performance is at 0 dB. The depth of the loss null is related to the clutter-to-noise ratio (CNR).

2.12.1.1 Historical overview

Historical overviews of STAP are given in [4,5]. Reed [4] describes the earlier history of adaptive arrays, whilst Klemm [5] covers more contemporary developments in STAP. Additionally, Entzminger et al. [6] discusses the history and future of Joint STARS and ground moving target indication (GMTI), identifying the central role of STAP in the success of both.

**FIGURE 12.1**

Clutter angle-Doppler region of support for three array configurations. (a) PSD (left) and super-resolution clutter image (right) for side-looking array, (b) PSD (left) and super-resolution clutter image (right) for forward-looking array, and (c) PSD (left) and super-resolution clutter image (right) for array with 45° of yaw.

**FIGURE 12.2**

Comparison of STAP frequency response for the three scenarios in Figure 12.1. (a) Side-looking array, (b) forward-looking array, and (c) array with 45° of yaw.

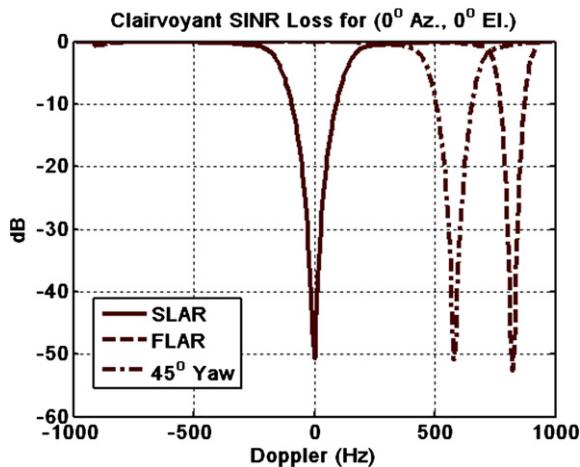


FIGURE 12.3

Clairvoyant SINR loss for example radar scenarios.

Table 12.1 is an historical accounting of STAP from the author’s perspective, as well as information derived from [4–6]. As such, it attempts to convey some key developments and general trends, but is necessarily incomplete. (Note: references to specific works are provided throughout the remainder of this text. Additionally, Klemm [5] provides an excellent, historical accounting and perspective, and is highly recommended to the interested reader.)

2.12.1.2 Organization

We organize the remaining discussion as follows:

- Section 2.12.2 describes basic concepts, include target detection; different space-time optimal filter formulations, including the maximum SINR filter, the minimum variance beamformer, and the generalized sidelobe canceler; the sample matrix inversion (SMI) approach to adaptive filtering; and, key performance metrics, such as SINR loss.
- Section 2.12.3 discusses clutter, RFI, target, and receiver space-time signal models.
- Section 2.12.4 discusses different adaptive filter topologies, including several reduced-rank methods, reduced-dimension STAP, and parametric adaptive matched filtering.
- Section 2.12.5 describes the utility of STAP to clutter suppression in airborne radar, as an example. This section also includes benchmark results for some of the algorithms discussed in Section 2.12.4.
- Much of current STAP research focuses on extending textbook discussion to real-world application. Section 2.12.6 describes various STAP challenges and research areas, including the impact of heterogeneous or nonstationary clutter on detection performance and mitigating techniques, application to airborne bistatic and conformal array radar, and knowledge-aided STAP. Each of these challenge areas predominantly focuses on issues surrounding effective estimation of the interference covariance matrix.

Table 12.1 Some Significant Events in the History of STAP

Year	Topic
1957	Paul Howells of General Electric, Syracuse (now Lockheed-Martin), develops technique to electronically scan antenna null in direction of a jammer
1959	Howells receives US Patent, "Intermediate frequency side-lobe canceller"
1962	Howells and Sidney Applebaum successfully test five-loop side-lobe canceller
1963	Howells and Applebaum, at Syracuse University Research Corp. (then Syracuse Research Corp., now SRC) investigate application of adaptive techniques to radar, including OTHR and BMD radar
1965 (1976)	Applebaum publishes "Adaptive Arrays," SURC TR-66-001, later available in IEEE Transactions on Antennas and Propagation in 1976
1969	Lloyd Griffiths publishes an adaptive algorithm for wideband antennas in Proceedings of the IEEE
1969	V. Anderson, H. Cox and N. Owsley separately publish works on adaptive arrays for sonar
1971	R.T. Compton describes the application of adaptive arrays for communications systems in Ohio State University Quarterly Rept. 3234-1, December 1971
1972	O.L. Frost publishes an adaptive algorithm for antenna arrays incorporating constraints
1972	Howells and Applebaum investigate adaptive radar for Airborne Early Warning (AEW) radar
1973	Larry Brennan and Irving Reed publish the seminal paper, "Theory of adaptive radar," in IEEE Trans. AES
1974	Reed, John Mallett and Brennan publish a paper describing the Sample Matrix Inverse (SMI) method for adaptive arrays in IEEE Transactions on Aerospace and Electronic Systems
1976	"Adaptive antenna systems," published by Bernard Widrow et al. in IEEE Proceedings
1980	R.A. Monzingo and T.W. Miller publish the book <i>Adaptive Arrays</i>
1983	Rule for calculating clutter subspace dimension proposed by Richard Klemm in IEE Proceedings Part F, Vol 130, February 1983
1991	Joint STARS prototypes deploy to Gulf War using adaptive clutter suppression methods (1978, Pave Mover was pre-cursor)
1992	Klemm proposes spatial transform techniques for dimensionality reduction in paper on antenna design for STAP
1992	Real-time STAP implementation by Alfonso Farina et al.
1992	Bob DiPietro presents post-Doppler STAP algorithm at Asilomar Conference on Signals, Systems and Computers
1994	Post-Doppler, beamspace method proposed by Wang and Cai in IEEE Transactions on Aerospace and Electronic Systems
1994	Jim Ward of MIT Lincoln Lab describes STAP techniques in ESC-TR-94-109, <i>Space-Time Adaptive Processing for Airborne Radar</i>
1995	Air Force Rome Laboratory (now Air Force Research Laboratory) and Westinghouse (now Northrop Grumman) collect twenty-four channel airborne radar data for STAP research and development under the Multichannel Airborne Radar Measurements (MCARM) program

Continued

Table 12.1 *Continued*

Year	Topic
1998	Richard Klemm publishes first STAP text book, <i>Space-Time Adaptive Processing: Principles and Applications</i>
1999	IEE Electronics and Communications Engineering Journal (ECEJ) Special Issue on STAP (Klemm, Ed.)
1999	STAP techniques for space-based radar, by Rabideau and Kogon, presented at IEEE Radar Conference
1999	3-D STAP for hot and cold clutter mitigation appears in IEEE Trans. AES by Techau, Guerci, Slocumb and Griffiths
2000	Bistatic STAP techniques appear in literature (numerous authors)
2000	IEEE Transactions on Aerospace and Electronic Systems Special Section on STAP (William Melvin, Ed.)
2002	DARPA's Knowledge-Aided Sensor Signal Processing & Expert Reasoning (KASSPER) commences under leadership of Joseph Guerci
2003	Guerci publishes the book <i>Space-Time Adaptive Processing for Radar</i>
2004	IEE publishers release the book <i>The Applications of Space-Time Processing</i> (Klemm, Ed.)
c. 2005	Multi-Input, Multi-Output (MIMO) adaptive systems (numerous authors)
2006	IEEE Transactions on Aerospace and Electronic Systems, Special Section on Knowledge-Aided Signal and Data Processing (Melvin, Guerci, Ed.)
2009	Cognitive Radar (Guerci)

- A description of some additional, practical issues in STAP implementation is given in Section 2.12.7, including maximum likelihood angle estimation and adaptive matched filter normalization.
- In Section 2.12.8 we briefly discuss some multichannel radar data collection efforts.
- Summary comments are given in Section 2.12.9, along with notation and acronym lists in the appendices.

2.12.1.3 Key points

STAP is a key component of modern radar. Some important points are given below:

- STAP is really a catch-all phrase for a variety of weight calculation schemes and adaptive architectures.
- Radar terminology usually includes two time measurements: fast-time, corresponding to a sample rate consistent with the analog-to-digital converter clock speed and effectively measuring range; and, slow-time, where the pulse repetition interval (PRI) defines the sample rate and separating targets and clutter in Doppler frequency is the primary motivation. Space and slow-time are used to mitigate ground clutter returns in aerospace radar, whereas space and fast-time are used to cope with wideband RFI or multipath signals in aerospace or surface radar systems.
- Herein, we focus on space and slow-time adaptive processing for clutter suppression. The ideas discussed are easily extended to other radar degrees of freedom (DoFs), include fast-time, polarization, and multiple passes.

- Implementation of the optimal processor requires precise knowledge of the null-hypothesis covariance matrix and the target space-time steering vector. The adaptive filter is an approximation to the optimal processor, since neither covariance matrix nor steering vector are known in practice. There are two primary factors leading to differences between adaptive and optimal filter outputs: covariance matrix estimation error and unknown array manifold due to amplitude and phase error as a function of frequency. Adaptive SINR loss describes the impact of both of these errors on detection performance potential.
- Current STAP research focuses on coupling basic STAP theory, physics-based models, and the realities of the real-world operating environment to construct practical detection architectures.
- STAP is applied at the coherent processing interval (CPI) level, often typically prior to noncoherent—or post-detection—integration. Maximizing SINR is always effective at improving detection performance, since the separation between the null and alternative hypothesis output probability density functions increases.
- In many respects, STAP development is in its early stages, at least from an implementation perspective. It has been only recently that embedded computing power has improved to the point where more flexible, real-time STAP implementation is feasible.

2.12.2 Basic concepts

In this section we discuss radar space-time detection and several approaches to calculate the optimal space-time weight vector: the maximum SINR filter, the minimum variance beamformer, the generalized sidelobe canceler, and use of the Rayleigh quotient and generalized eigen-analysis. We introduce the sample matrix inverse adaptive filter as the most common of the various adaptive implementations used in radar application. We also discuss several important STAP metrics, such as clairvoyant and adaptive SINR loss and improvement factor.

2.12.2.1 Detection

Suppose we collect M spatial samples across the antenna array for range realization k and each of the N received pulses. These voltages are organized in a vector as

$$\mathbf{x}_k = \left[\mathbf{x}_{s/k}^T(1) \quad \mathbf{x}_{s/k}^T(2) \quad \cdots \quad \mathbf{x}_{s/k}^T(N) \right]^T, \quad (12.1)$$

where $\mathbf{x}_k \in C^{NM \times 1}$ is the space-time snapshot and $\mathbf{x}_{s/k}(n) \in C^{M \times 1}$ is the spatial snapshot for the n th pulse. Two hypotheses correspond to (12.1): the null hypothesis, H_0 , where only additive interference and noise are present; and, the alternative hypothesis, H_1 , which includes target presence in addition to the components of the null hypothesis. Defining $\mathbf{c}_k \in C^{NM \times 1}$ as the clutter space-time snapshot, $\mathbf{j}_k \in C^{NM \times 1}$ as the RFI space-time snapshot, $\mathbf{n}_k \in C^{NM \times 1}$ as the uncorrelated noise space-time snapshot, and $\mathbf{t}_k \in C^{NM \times 1}$ the target space-time snapshot, we have

$$\begin{aligned} H_0 : \mathbf{x}_k &= \mathbf{c}_k + \mathbf{j}_k + \mathbf{n}_k, \\ H_1 : \mathbf{x}_k &= \mathbf{c}_k + \mathbf{j}_k + \mathbf{n}_k + \mathbf{t}_k = \mathbf{x}_k/H_0 + \mathbf{t}_k. \end{aligned} \quad (12.2)$$

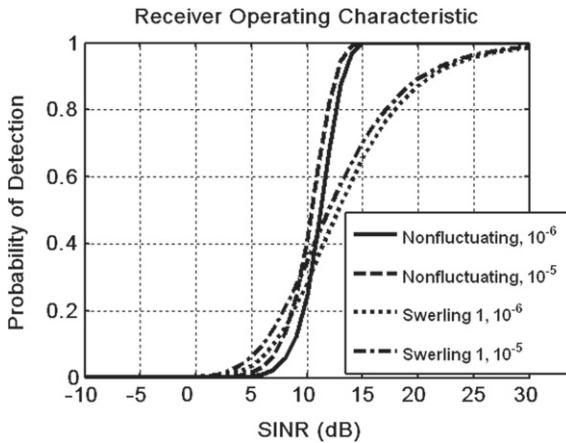


FIGURE 12.4

Receiver operating characteristic (ROC) curves for nonfluctuating and Swerling 1 targets (probability of false alarm given in legend).

The radar detection problem is to decide, given \mathbf{x}_k , which of the two hypotheses is most probable. It is shown in [3] that maximizing SINR is equivalent to maximizing probability of detection for a fixed false alarm rate when $\mathbf{x}_k/\mathbf{H}_0$ is jointly complex normal with zero mean and covariance matrix, \mathbf{R}_k , viz. $\mathbf{x}_k/\mathbf{H}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$. We wish to combine the elements of \mathbf{x}_k using a linear, finite impulse response (FIR) filter, and compare the scalar to a threshold, η_T , to decide which hypothesis is a best fit,

$$y_k = \mathbf{w}_k^H \mathbf{x}_k = \sum_{m=1}^{NM} [\mathbf{w}_k]_m^* [\mathbf{x}_k]_m^* \begin{matrix} > \\ H_1 \\ < \\ H_0 \end{matrix} \eta_T, \quad (12.3)$$

where \mathbf{w}_k is the space-time weight vector.

Figure 12.4 shows probability of detection, P_d , as a function of output SINR for varying probability of false alarm, P_{fa} . The curves correspond to nonfluctuating and Swerling 1 target types for $P_{fa} = 10^{-5}$ and $P_{fa} = 10^{-6}$. As seen, P_d increases monotonically with SINR. Also, it is seen that for reasonably high P_d we generally need higher SINR to detect a Swerling 1 target at the same rate as the nonfluctuating target due to fading; the SINR difference between Swerling 1 and nonfluctuating detection curves is called fluctuation loss.

2.12.2.2 Space-time filter formulations

2.12.2.2.1 Maximum SINR filter

A primary objective in STAP is to approximate as closely as possible the optimal weighting that maximizes SINR. By attempting to maximize output SINR, the STAP maximizes the probability of

detection for a fixed false alarm rate, as discussed in the prior section. We now formulate the optimal weight vector.

The target snapshot is of the form, $\mathbf{t}_k = \alpha_T \mathbf{s}_{s-t}(\phi, \theta, f_d)$, where α_T is a complex gain term proportional to the square root of the target cross section, $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the space-time steering vector of dimension NM -by-one, ϕ is azimuth, θ is elevation, and f_d is Doppler frequency. In the nonfluctuating target case, α_T is a constant, whereas in the Swerling 1 case, $\alpha_T \sim \mathcal{CN}(0, \sigma_T^2)$, where σ_T^2 is the target signal variance [1, 7, 8]. The space-time steering vector is the Kronecker product of the spatial steering vector, $\mathbf{s}_s(\phi, \theta) \in \mathbb{C}^{M \times 1}$, and the temporal steering vector, $\mathbf{s}_t(f_d) \in \mathbb{C}^{N \times 1}$. Generally, the spatial steering vector describes the phase variation among the spatial antenna channels (of common size and gain) to a signal with a direction of arrival of (ϕ, θ) , whereas the temporal steering vector describes the phase variation over the slow-time aperture to a signal with a particular range-rate [9–11]. If the aperture in either space or time is uniformly sampled, the corresponding steering vector will be Vandermonde.

The output SINR of the FIR filter with weight vector, \mathbf{w}_k , is the ratio of target power, P_T , to interference-plus-noise power, P_{I+N} , given as

$$\text{SINR} = \frac{P_T}{P_{I+N}} = \frac{E[y_T y_T^*]}{E[y_{k/H_0} y_{k/H_0}^*]} = \frac{E[\mathbf{w}_k^H \mathbf{t}_k \mathbf{t}_k^H \mathbf{w}_k]}{E[\mathbf{w}_k^H \mathbf{x}_k/\mathbf{H}_0 \mathbf{x}_k^H/\mathbf{H}_0 \mathbf{w}_k]} = \frac{\mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{R}_{k/H_0} \mathbf{w}_k}. \quad (12.4)$$

Using the definition for \mathbf{t}_k gives, $\mathbf{R}_T = E[\mathbf{t}_k \mathbf{t}_k^H] = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, and the Schwarz Inequality lets us write (12.4) as

$$\text{SINR} = \sigma_T^2 \frac{|\tilde{\mathbf{w}}^H \tilde{\mathbf{s}}|^2}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}} \leq \sigma_T^2 \frac{(\tilde{\mathbf{w}}^H \tilde{\mathbf{w}})(\tilde{\mathbf{s}}^H \tilde{\mathbf{s}})}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}}. \quad (12.5)$$

In this case, $\tilde{\mathbf{w}} = \mathbf{R}_{k/H_0}^{-1/2} \mathbf{w}_k$ and $\tilde{\mathbf{s}} = \mathbf{R}_{k/H_0}^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$. By inspection, we find that the right side of the inequality achieves the upper bound when $\tilde{\mathbf{w}} = \mu \tilde{\mathbf{s}}$, thereby providing the solution for the optimal weight vector. Through substitution, it follows that

$$\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.6)$$

for arbitrary scalar, μ , is the weighting that yields maximal output SINR. When the disturbance is uncorrelated noise, $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$, where σ_n^2 is the noise power; then, (12.6) becomes the space-time matched filter, viz. $\mathbf{w}_k = \mu \mathbf{s}_{s-t}(\phi, \theta, f_d)$.

Inserting the optimal weight vector, $\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, into (12.4) yields the maximum output SINR,

$$\text{SINR}_{\text{Max}} = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.7)$$

For simplicity, and without loss in meaning, in subsequent sections we use $\mathbf{R}_k = \mathbf{R}_{k/H_0}$.

The maximum SINR weight vector has the following interpretation. If $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, then $\tilde{\mathbf{x}}_k = \mathbf{R}_k^{-1/2} \mathbf{x}_k$ is the whitened data vector, where $\tilde{\mathbf{x}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{NM})$ and \mathbf{I}_{NM} is the NM -by- NM identity matrix. This results because

$$E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^H] = \mathbf{R}_k^{-1/2} E[\mathbf{x}_k \mathbf{x}_k^H] \mathbf{R}_k^{-1/2} = \mathbf{R}_k^{-1/2} \mathbf{R}_k \mathbf{R}_k^{-1/2} = \mathbf{I}_{NM}. \quad (12.8)$$

Then, the filter employing the maximum SINR weight vector is seen as a cascade of whitening and warped matched filtering operations,

$$y_k = \mathbf{w}_k^H \mathbf{x}_k = \left(\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \right) \left(\mathbf{R}_k^{-1/2} \mathbf{x}_k \right) = \tilde{\mathbf{s}}_{s-t}^H \tilde{\mathbf{x}}_k, \quad (12.9)$$

where $\tilde{\mathbf{s}}_{s-t} = \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the warped matched filter accounting for the linear transformation of the target signal through the whitening process and $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the original matched filter weight vector.

2.12.2.2 Minimum variance beamformer

The minimum variance (MV) space-time beam former is another common formulation leading to the adaptive processor [12,13]. The MV beamformer employs a weight vector yielding minimal output power subject to a linear constraint on the desired target spatial and temporal response:

$$\min_{\mathbf{w}_k} E[y_k y_k^*] \text{ subject to } \mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = g, \quad (12.10)$$

where g is a complex scalar. We can determine the weight vector satisfying the problem statement in (12.10) using the method of Lagrange. Thus, define the following cost function

$$J = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k + \operatorname{Re} \left\{ \tilde{\lambda}^* \left(\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) - g \right) \right\}, \quad (12.11)$$

where $\tilde{\lambda}^*$ is a complex Lagrangian multiplier and $E[y_k y_k^*] = E[\mathbf{w}_k^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{w}_k] = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k$. Taking the gradient of (12.11) with respect to the conjugate of the weights and setting to zero gives

$$\nabla J = \mathbf{R}_k \mathbf{w}_k + \tilde{\lambda}^* \mathbf{s}_{s-t}(\phi, \theta, f_d) = 0. \quad (12.12)$$

Solving (12.12) for the weight vector yields

$$\mathbf{w}_k = -\tilde{\lambda}^* \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.13)$$

Next, we find by applying the constraint in (12.10) that

$$\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = -\tilde{\lambda} \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = g. \quad (12.14)$$

Solving (12.14) for $\tilde{\lambda}$ then gives

$$\tilde{\lambda} = \frac{-g}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.15)$$

The minimum variance weight vector follows from (12.13) and (12.15) as

$$\mathbf{w}_k = \frac{g^* \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.16)$$

Equation (12.16) abides by the form $\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, where

$$\mu = \frac{g^*}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.17)$$

Hence, (12.16) likewise yields maximal SINR.

Setting $g = 1$ is known as the *distortionless response*, and the corresponding weight vector yields the minimum variance distortionless response (MVDR) beamformer. The output power of the MVDR beamformer is given as

$$P_{\text{MVDR}} = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k = \frac{1}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.18)$$

The MVDR spectrum, as given in Figure 12.1, follows from (12.18) by scanning the space-time steering vector in the denominator over the angle and Doppler locations of interest.

2.12.2.2.3 Generalized sidelobe canceler

The GSLC is a formulation that conveniently converts the minimum variance constrained beamformer described in the preceding section into an unconstrained form [12, 14]. Many prefer the GSLC interpretation of STAP over the whitening filter-warped matched filter interpretation of Section 2.12.2.2.1; since the GSLC implements the MV beamformer for the single linear constraint [14], this structure likewise can be shown to maximize output SINR.

Figure 12.5 provides a block diagram of the GSLC. The top leg of the GSLC generates a quiescent response by forming a beam at the angle and Doppler of interest. A blocking matrix, $\mathbf{B}_G = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$, prevents the target signal from entering the lower leg; essentially, the blocking matrix forms a notch filter tuned to (ϕ, θ, f_d) . With the desired signal absent from the lower leg, the processor tunes the weight vector to provide a minimal mean square error (MMSE) estimate of the interference in the top leg. In the final step, the processor differences the desired signal, $d_{o/k}$, with the estimate from the lower leg, $\hat{d}_{o/k}$, to form the filter output, $y_k = d_{o/k} - \hat{d}_{o/k}$. Ideally, any large residual at the output corresponds to an actual target.

The desired signal is given as

$$d_{o/k} = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{x}_k. \quad (12.19)$$

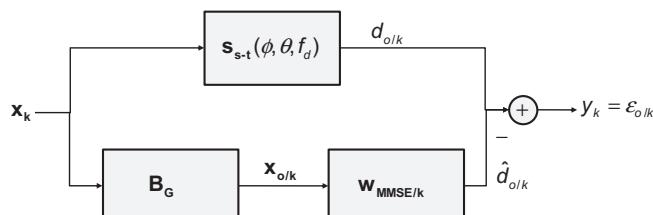


FIGURE 12.5

GSLC block diagram.

The signal vector in the lower leg is

$$\mathbf{x}_{0/k} = \mathbf{B}_G \mathbf{x}_k; \quad \mathbf{B}_G \in C^{(NM-1) \times NM}; \quad \mathbf{x}_{0/k} \in C^{(NM-1) \times 1}. \quad (12.20)$$

Forming the quiescent response of (12.19) uses a single degree of freedom (DoF), resulting in the reduced dimensionality of $\mathbf{x}_{0/k}$. By weighting the signal vector in the lower leg, the GSLC forms a MMSE estimate of $d_{o/k}$ as

$$\hat{d}_{o/k} = \mathbf{w}_{\text{MMSE}/k}^H \mathbf{x}_{0/k}, \quad (12.21)$$

where the MMSE weight vector follows from the well-known Wiener-Hopf equation [12, 13],

$$\mathbf{w}_{\text{MMSE}/k} = \mathbf{R}_{\mathbf{x}_{0/k}}^{-1} \mathbf{r}_{\mathbf{x}_{0/k} \mathbf{d}}. \quad (12.22)$$

The lower leg covariance matrix is

$$\mathbf{R}_{\mathbf{x}_{0/k}} = E \left[\mathbf{x}_{0/k} \mathbf{x}_{0/k}^H \right] = \mathbf{B}_G \mathbf{R}_k \mathbf{B}_G^H \quad (12.23)$$

whilst the cross-correlation between lower and upper legs is

$$\mathbf{r}_{\mathbf{x}_{0/k} \mathbf{d}} = E \left[\mathbf{x}_{0/k} d_{o/k}^* \right] = \mathbf{B}_G \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.24)$$

The GSLC filter output is then

$$\begin{aligned} y_k &= \varepsilon_{o/k} = d_{o/k} - \hat{d}_{o/k} = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{x}_k - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{x}_{0/k} \\ &= \left(\mathbf{s}_{s-t}^H(\phi, \theta, f_d) - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \right) \mathbf{x}_k. \end{aligned} \quad (12.25)$$

Comparing (12.25) and (12.3), we identify

$$\mathbf{w}_{\text{GSLC}/k} = \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k}. \quad (12.26)$$

We compute the output SINR as the ratio of output signal power to interference-plus-noise power, as in (12.4). The signal-only output power of the GSLC is

$$\begin{aligned} P_T &= E \left[\mathbf{w}_{\text{GSLC}/k}^H \mathbf{t}_k \mathbf{t}_k^H \mathbf{w}_{\text{GSLC}/k} \right] = E \left[\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{t}_k \mathbf{t}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right] \\ &= \sigma_T^2 \left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2 = \sigma_T^2 (NM)^2, \end{aligned} \quad (12.27)$$

and the output interference-plus-noise power is

$$P_{I+N} = E \left[\mathbf{w}_{\text{GSLC}/k}^H \mathbf{x}_k / \mathbf{H}_0 \mathbf{x}_{k/\mathbf{H}_0}^H \mathbf{w}_{\text{GSLC}/k} \right] = \mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k}. \quad (12.28)$$

The ratio of (12.27) and (12.28) is

$$\text{SINR}_{\text{GSLC}} = \frac{\sigma_T^2 N^2 M^2}{\mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k}} = \frac{\sigma_T^2 N^2 M^2}{\sigma_d^2 - \mathbf{r}_{\mathbf{x}_{0/k} \mathbf{d}}^H \mathbf{R}_{\mathbf{x}_{0/k}}^{-1} \mathbf{r}_{\mathbf{x}_{0/k} \mathbf{d}}}, \quad (12.29)$$

where $\sigma_d^2 = E[|d_{o/k}|^2] = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)$. We arrive at the denominator in (12.29) by using (12.22)–(12.24),

$$\begin{aligned} \mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k} &= \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k} \\ &\quad - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d) + \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \mathbf{R}_k \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k} \\ &= \sigma_d^2 - \mathbf{r}_{x_0/k}^H \mathbf{w}_{\text{MMSE}/k} - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{r}_{x_0/k} + \mathbf{w}_{\text{MMSE}/k}^H \mathbf{R}_{x_0/k} \mathbf{w}_{\text{MMSE}/k} \\ &= \sigma_d^2 - \mathbf{r}_{x_0/k}^H \mathbf{R}_{x_0/k}^{-1} \mathbf{r}_{x_0/k}. \end{aligned} \quad (12.30)$$

2.12.2.2.4 Rayleigh quotient

The optimal weight vector in Section 2.12.2.2.1 maximizes SINR. As seen from (12.4), for weight vector \mathbf{w}_k , SINR is given as the ratio of quadratics, $\text{SINR} = \mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k / \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k$. Thus, our problem is to find the weight vector that maximizes this function. In accord with [15], and similar to our discussion in Section 2.12.2.2.1, SINR is written as

$$\text{SINR} = \frac{\tilde{\mathbf{w}}^H \left(\left(\mathbf{R}_k^{-1/2} \right)^H \mathbf{R}_T \mathbf{R}_k^{-1/2} \right) \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}}, \quad (12.31)$$

where $\tilde{\mathbf{w}} = \mathbf{R}_k^{1/2} \mathbf{w}_k$. The weight vector, \mathbf{w}^0 , that maximizes this Rayleigh quotient form is known to be proportional to the eigenvector \mathbf{v}_{\max} associated with the largest eigenvalue λ_{\max} of the matrix in the numerator [13],

$$\tilde{\mathbf{w}} \propto \mathbf{v}_{\max} = \max_{\text{Eigenvector}} \left(\left(\mathbf{R}_k^{-1/2} \right)^H \mathbf{R}_T \mathbf{R}_k^{-1/2} \right) = \max_{\text{Eigenvector}} \left(\mathbf{R}_k^{-1/2} \mathbf{R}_T \mathbf{R}_k^{-1/2} \right), \quad (12.32)$$

and so

$$\mathbf{w}^0 \propto \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \max_{\text{Eigenvector}} \left(\mathbf{R}_k^{-1/2} \mathbf{R}_T \mathbf{R}_k^{-1/2} \right). \quad (12.33)$$

Take the case where $\mathbf{R}_T = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, as previously introduced in Section 2.12.2.2.1. The maximum eigenvector in (12.32) must abide by

$$\sigma_T^2 \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \lambda_{\max} \mathbf{v}_{\max}. \quad (12.34)$$

The solution for \mathbf{v}_{\max} in (12.34) is thus seen to be

$$\mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.35)$$

and so, as expected,

$$\mathbf{w}^0 \propto \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.36)$$

(Note: inserting (12.35) in (12.34) gives $\mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$ times a scalar on each side of the equation, thereby satisfying the eigen-relation.)

By plugging $\mathbf{w}_k = \mathbf{w}^0$ found in (12.33) into (12.4), it is straightforward to show that the maximum SINR is equal to λ_{\max} . In this vein, using (12.35) in (12.34) gives

$$\lambda_{\max} = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.37)$$

which is the desired result and matches the expression given in (12.7).

2.12.2.2.5 Generalized eigen-analysis

The second approach, described in [13], begins by noting that the solution to maximizing SINR is equivalent to finding the maximum eigenvalue and associated eigenvector for the generalized Eigen-problem

$$\mathbf{R}_T \mathbf{w}_k = \lambda \mathbf{R}_k \mathbf{w}_k. \quad (12.38)$$

This can be solved as an ordinary eigen-equation by writing it as

$$\left(\mathbf{R}_k^{-1} \mathbf{R}_T \right) \mathbf{w}_k = \lambda \mathbf{w}_k. \quad (12.39)$$

Then, the weight vector that maximizes (12.4) is proportional to the eigenvector \mathbf{v}_{\max} associated with the largest eigenvalue λ_{\max} of $\mathbf{R}_k^{-1} \mathbf{R}_T$,

$$\mathbf{w}^0 \propto \mathbf{v}_{\max} = \max_{\text{Eigenvector}} \left(\mathbf{R}_k^{-1} \mathbf{R}_T \right). \quad (12.40)$$

The maximum eigenvalue of $\mathbf{R}_k^{-1} \mathbf{R}_T$ is the maximum achievable SINR in (12.4).

If we again consider the case where $\mathbf{R}_T = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, the generalized eigen-analysis equation of (12.39) becomes

$$\left(\sigma_T^2 \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \right) \mathbf{w}_k = \lambda \mathbf{w}_k, \quad (12.41)$$

which, we find, is satisfied when $\mathbf{w}_k = \mathbf{w}^0 = \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$. Substituting this eigenvector, \mathbf{w}_k , into (12.41) gives,

$$\left(\sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \right) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \lambda_{\max} \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.42)$$

from which it is seen the maximum eigenvalue, λ_{\max} , is identical to the result in (12.37) and also the same as the maximum SINR expression found in Section 2.12.2.2.1.

2.12.2.6 Summary

Table 12.2 summarizes the weight vector formulation given in the prior sections.

2.12.2.3 Sample matrix inversion

All of the previously described approaches to filter design presume the availability of the interference-plus-noise covariance matrix for the k th realization, \mathbf{R}_k , and perfectly known target space-time steering

Table 12.2 Summary of Space-Time Filter Formulations

Formulation	Objective	Weight Vector
Maximum SINR filter	Maximize output SINR	$\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$
Minimum variance (MV) beamformer	Minimize output power subject to linear constraint maintaining gain in the look direction	$\mathbf{w}_k = \frac{\mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}$
Generalized sidelobe canceler	Convert constrained beamformer formulation to unconstrained form	$\mathbf{w}_k = \mathbf{w}_{GSLC/k} = \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{B}_G^H \mathbf{w}_{MMSE/k}$
Rayleigh quotient	Maximize output SINR	$\mathbf{w}_k \propto \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$
Generalized eigen-analysis	Maximize output SINR	$\mathbf{w}_k \propto \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$

vector, $\mathbf{s}_{s-t}(\phi, \theta, f_d)$. In practice, neither \mathbf{R}_k nor $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is known. The disturbance covariance matrix, \mathbf{R}_k , changes as look-angle, platform location, and platform attitude change. Errors in $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ are generally a result of straddling the precise target spatial or temporal frequency when searching over angle and Doppler, as well as system hardware errors that most greatly affect the spatial component, $\mathbf{s}_s(\phi, \theta)$. These system errors generally manifest as spatially varying, complex gain errors due to factors such as spatial variation in element gain patterns, varying line lengths, radome and airframe interactions, etc.

A number of techniques are available to adapt the weight vector, including least means square (LMS) and recursive least squares (RLS) formulations [12]. However, the sample matrix inversion (SMI) method is by far the most popular choice for radar application for two primary reasons: (1) convergence of the SMI approach depends only on the number of samples used to estimate the unknown interference-plus-noise covariance matrix, regardless of environmental conditions as long as the data are independent and identically distributed (IID), for a system of specified DoFs; and, (2) prior to processing, radar systems generally collect their data in blocks, called coherent processing intervals (CPIs), and so a batch processing strategy using SMI is fully acceptable. In the SMI approach, the optimal weight vector,

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.43)$$

is simply replaced with the adaptive weight vector,

$$\hat{\mathbf{w}}_k = \hat{\mu} \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d), \quad (12.44)$$

where $\hat{\mu}$ is a scalar that often depends on estimated quantities, $\hat{\mathbf{R}}_k$ is the interference-plus-noise covariance matrix estimate, and $\mathbf{v}_{s-t}(\phi, \theta, f_d)$ is the hypothesized space-time steering vector. In the absence of straddle error, system errors dominate the error between $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{v}_{s-t}(\phi, \theta, f_d)$. Generally, array errors are explicitly modeled as

$$\mathbf{s}_{s-t}(\phi, \theta, f_d) = \mathbf{s}_t(f_d) \otimes (\mathbf{e}_s(\phi, \theta) \odot \mathbf{s}_s(\phi, \theta)), \quad (12.45)$$

where $\boldsymbol{\epsilon}_s(\phi, \theta)$ is the spatially-varying error between the ideal and actual array manifold, and \otimes and \odot are Kronecker (tensor) and Hadamard (element-wise) products, respectively.

In practice, $\boldsymbol{\epsilon}_s(\phi, \theta)$ is “calibrated out” of the system using a variety of techniques, including array tuning on an antenna range or *in situ* using the clutter background and navigation data to estimate differences between ideal and actual multi-channel antenna responses [16]. Correspondingly, the hypothesized space-time steering vector is then

$$\mathbf{v}_{s-t}(\phi, \theta, f_d) = \mathbf{v}_t(f_d) \otimes (\hat{\boldsymbol{\epsilon}}_s(\phi, \theta) \odot \mathbf{v}_s(\phi, \theta)), \quad (12.46)$$

where $\hat{\boldsymbol{\epsilon}}_s(\phi, \theta) \in C^{M \times 1}$ is an estimate of the array error vector, and $\mathbf{v}_t(f_d)$ and $\mathbf{v}_s(\phi, \theta, f_d)$ are otherwise the hypothesized temporal and spatial steering vectors accounting for potential straddle. Temporal errors due to system non-ideality are commonly very small for the typical STAP CPI, and thus usually are not considered further.

The complex baseband, pulse compressed, space-time snapshots comprise the voltage vectors for the p th CPI,

$$\mathbf{X}_{\text{CPI}}(p) = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_L] \in C^{NM \times L}. \quad (12.47)$$

A particular realization, \mathbf{x}_k , chosen from amongst the space-time vectors in (12.47) for filtering and detection thresholding is called the cell-under-test (CUT); multiple CUTs form the primary data set. The remaining vectors in (12.47) are available to estimate the unknown interference-plus-noise covariance matrix and are called training or secondary data. The primary data set can be as small as a single CUT with several adjacent realizations serving as guard cells; the purpose of the guard cell region is to prevent any target energy from leaking into the training interval. Let L_g be the number of guard cells on either side of the primary data region. Then, define the training set as

$$\mathbf{X}_{\text{Train}}(p) = \{\mathbf{x}_m\}_{m \in 1:\kappa; m \neq k-L_g:k+L_g}. \quad (12.48)$$

It is shown in [17] that if the data vectors comprising the training interval—the columns of (12.48)—are IID with respect to the null-hypothesis of the CUT, a maximum likelihood estimate for \mathbf{R}_k is

$$\hat{\mathbf{R}}_k = \frac{1}{\kappa} \mathbf{X}_{\text{Train}}(p) \mathbf{X}_{\text{Train}}^H(p) = \frac{1}{\kappa} \sum_{m=1}^{\kappa} \mathbf{x}_m \mathbf{x}_m^H. \quad (12.49)$$

A fundamental question centers on how many training samples, κ , lead to an acceptable covariance matrix estimate. This matter was addressed in [17] and considered in further detail in Section 2.12.2.4.2.

2.12.2.4 Metrics

In radar, performance is generally measured according to the specific goals of the collection and processing mode. In moving target indication (MTI) radar, probability of detection, P_d , and false alarm rate (FAR)—or probability of false alarm, P_{fa} —are the primary measures of performance. As discussed in Section 2.12.2.1, the probability of detection is a monotonic function of output SINR for a fixed false alarm rate. Thus, measures of SINR are key to understanding achievable system performance. Among these measures, SINR loss metrics are preferable, since they characterize performance over certain independent variables—usually range rate or Doppler frequency, as shown in Figure 12.3—and are not

tied to a specific radar cross section, thus folding into traditional link budget analysis. Specifically, given SNR calculated from the radar range equation [1, 7, 8] or measured in the field, the output SINR is

$$\text{SINR}(\phi, \theta, f_d) = \text{SNR}(\phi, \theta) \cdot \prod_{m=1}^Q L_{s,m}(\phi, \theta, f_d), \quad (12.50)$$

where $L_{s,m}(\phi, \theta, f_d)$ is the m th SINR loss term and $0 \leq L_{s,m}(\phi, \theta, f_d) \leq 1$. (N.b., the term “SINR loss” is widely used, even though such losses are negative valued on a decibel scale. This oftentimes contradicts typical radar systems engineering usage. However, the losses are numerator terms in the full link calculation.) In general, SINR losses vary with angle and Doppler, since the interference PSD likewise varies.

In imaging radar, terrain-to-noise ratio (TNR) is a primary metric and is a function of multiplicative noise ratio and SINR loss. We focus our attention specifically on MTI radar, but the basic metrics are adaptable.

2.12.2.4.1 *Clairvoyant SINR loss*

Clairvoyant SINR loss is the ratio of output SINR for a filter implementation, \mathbf{w}_k , where all parameters are known precisely, to the maximum SNR. Clairvoyant SINR is given as

$$L_{s,1} = \frac{\text{SINR}}{\text{SNR}} = \frac{\left(\sigma_T^2 \frac{|\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right)}{\left(\frac{\sigma_T^2}{\sigma_n^2} NM \right)} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}. \quad (12.51)$$

A few cases are worth considering:

Case 1—Noise-Limited Condition: In this case, $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$, where σ_n^2 is the noise power. Then, the maximum SINR weight vector,

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \rightarrow \mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.52)$$

and so

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\sigma_n^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)} = 1. \quad (12.53)$$

As expected, the optimal weighting defaults to the matched filter and there is no loss relative to the bound defined by the uncorrelated noise.

Case 2—Color-Limited Condition, Matched Filter Weights: Clairvoyant loss characterizes the impact of interference on detection performance. Consider the case where the weight vector is set to the space-time matched filter, $\mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d)$. The clairvoyant SINR loss follows from (12.51) as

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)} = \frac{\sigma_n^2 NM}{\text{PSD}(\phi, \theta, f_d)} \leq 1, \quad (12.54)$$

where $\text{PSD}(\phi, \theta, f_d) = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the power spectral density (PSD), equal to the two-dimensional Fourier transform of the space-time covariance matrix [12]; Figure 12.1 provides

example clutter-plus-noise PSD plots. Equation (12.54) shows the impact of interference on performance relative to the noise-limited case. Since the PSD is diffraction-limited (the mainlobe is determined by the size of the space-time aperture), (12.54) characterizes the performance impacts of using the deterministic matched filter. At those angles and Doppler frequencies away from the clutter angle-Doppler region of support, the clairvoyant SINR loss approaches the noise-limited case.

Case 3—Color-Limited Condition, Optimal Weights: Using the maximum SINR weight vector, $\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, in (12.51), gives

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{\left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)} = \frac{\sigma_n^2}{NM} P_{MVDR}^{-1}(\phi, \theta, f_d), \quad (12.55)$$

where $P_{MVDR}(\phi, \theta, f_d)$ is a sample of the MVDR spectrum given in (12.18). Considering the MVDR plots in Figure 12.1, (12.55) suggests regions of loss confined to the sharp, super-resolution contours of the clutter MVDR response (note the inverse in (12.55)). For this reason, STAP is able to detect targets in close proximity to the center of the clutter angle-Doppler region of support. In contrast, (12.54) suggests that SINR loss extends to the full width of the diffraction-limited spectrum when using nonadaptive weights. Figure 12.25, given in Section 2.12.5.3, compares SINR loss for adaptive and nonadaptive solutions, confirming these observations.

2.12.2.4.2 Adaptive SINR loss

Adaptive SINR loss is the ratio of the output SINR for the filter using estimated quantities, $\hat{\mathbf{w}}_k$, to the optimal filter output, \mathbf{w}_k , where all parameters are clairvoyantly known, viz.

$$L_{s,2} = \left(\frac{\hat{\mathbf{w}}_k^H \mathbf{R}_T \hat{\mathbf{w}}_k}{\hat{\mathbf{w}}_k^H \mathbf{R}_k \hat{\mathbf{w}}_k} \right) \Bigg/ \left(\frac{\mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right). \quad (12.56)$$

Observe that when $\hat{\mathbf{w}}_k = \mathbf{w}_k$, $L_{s,2} = 1$. Moreover, $0 \leq L_{s,2} \leq 1$, since \mathbf{w}_k yields the maximum output SINR and so the denominator in (12.56) will always be greater than or equal to the numerator.

Substituting (12.43) and (12.44) into (12.56) yields

$$L_{s,2}(\hat{\mathbf{R}}_k) = (\text{SINR}|\hat{\mathbf{w}}_k(\hat{\mathbf{R}}_k)) / (\text{SINR}_{\text{Max}}), \quad (12.57)$$

where SINR_{Max} is given by (12.7). Assuming $\hat{\mathbf{w}}_k$ uses (12.49) in its calculation, as (12.57) suggests, then determining κ is critical. This important matter was addressed in [17], where it is shown that (12.57) is Beta distributed, with mean value

$$E[L_{s,2}(\hat{\mathbf{R}}_k)] = (\kappa + 2 - NM) / (\kappa + 1). \quad (12.58)$$

Setting (12.58) equal to 0.5 (or, 3 dB of loss) and solving for κ yields

$$\kappa = 2NM - 3. \quad (12.59)$$

It is popular to refer to the result in (12.59), where setting κ roughly equal to twice the processor's DoFs yields 3 dB of loss, as the Reed-Mallett-Brennan (RMB) rule after its originators. In practice, 3 dB of loss is substantial, and so choosing κ to be at least five times the processor's DoF is desirable.

The IID assumption inherent in the calculation of (12.57) sets the bound on performance for the adaptive processor given κ homogeneous training samples. We address the impact of non-IID clutter conditions in further detail in Section 2.12.5.

2.12.2.4.3 Improvement factor

Improvement factor (IF) is given as the ratio of the output SINR to the input (element-level) SINR [18, 19]. IF is given as

$$\text{IF} = \frac{\left(\sigma_I^2 \frac{|\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right)}{\left(\frac{\sigma_I^2}{\sigma_I^2 + \sigma_n^2} \right)} = \left| \mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2 \frac{(\sigma_I^2 + \sigma_n^2)}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}, \quad (12.60)$$

where σ_I^2 is the input (element-level) interference power.

Case 1—Noise-Limited Condition: In the noise-limited case, $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$ and $\sigma_I^2 = 0$, and considering the matched filter, $\mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d)$, then

$$\text{IF} = \frac{\sigma_n^2 |\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\sigma_n^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)} = NM. \quad (12.61)$$

As expected, the improvement factor equals the space-time integration gain in this instance.

Case 2—Clutter-Limited Condition, Matched Filter Weights: In the clutter-limited case, with $\mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d)$, the improvement factor is

$$\text{IF} = \left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2 \frac{\sigma_I^2 + \sigma_n^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)} = (NM)^2 \frac{\sigma_I^2 + \sigma_n^2}{\text{PSD}(\phi, \theta, f_d)} \leq NM. \quad (12.62)$$

The presence of interference in the PSD degrades the improvement factor. If we let $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$ and $\sigma_I^2 = 0$ in (12.62), we arrive at (12.61).

Case 3—Clutter-Limited Condition, Optimal Weights: Using the maximum SINR weight vector, $\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, in (12.60), gives

$$\text{IF} = (\sigma_I^2 + \sigma_n^2) \frac{\left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)} = (\sigma_I^2 + \sigma_n^2) P_{\text{MVDR}}^{-1}(\phi, \theta, f_d). \quad (12.63)$$

This expression indicates reduced improvement for those angle-Doppler locations aligning with the clutter, and otherwise good performance in proximity to the clutter response owing to the super-resolution characteristic of $P_{\text{MVDR}}(\phi, \theta, f_d)$ (see Figure 12.1).

2.12.2.4.4 Optimal and adaptive filter patterns

The filter gain pattern follows directly by evaluating the filter response over the angles (or spatial frequencies) and Doppler frequencies of interest. Define the space-time steering matrix,

$$\mathbf{S} = [\mathbf{s}_{s-t}(\phi(1), \theta(1), f_d(1)) \ \mathbf{s}_{s-t}(\phi(2), \theta(1), f_d(1)) \ \cdots \ \mathbf{s}_{s-t}(\phi(P), \theta(U), f_d(V))], \quad (12.64)$$

for all $\{\phi(1), \phi(2), \dots, \phi(P)\}$, $\{\theta(1), \theta(2), \dots, \theta(U)\}$, and $\{f_d(1), f_d(2), \dots, f_d(V)\}$. The optimal gain pattern is

$$\mathbf{g}_o = 20 \log_{10} \left| \mathbf{w}_k^H \mathbf{S} \right|, \quad (12.65)$$

where \mathbf{w}_k is the optimal weighting given by (12.43). The adaptive gain pattern follows similarly as

$$\mathbf{g}_a = 20 \log_{10} \left| \hat{\mathbf{w}}_k^H \mathbf{S} \right| \quad (12.66)$$

with (12.44) yielding $\hat{\mathbf{w}}_k$.

Example gain patterns corresponding to the clutter environments shown in Figure 12.1 are given in Figure 12.2. This gain pattern corresponds to (12.65).

2.12.3 Signal models

In this section we consider basic space-time signal and covariance models.

2.12.3.1 Clutter

A plethora of reflected signals from the Earth's surface comprise what is known as radar ground clutter. Ground clutter is a primary impediment to the detection of moving targets. The clutter signal is predominantly made up of signal reflections from distributed objects, such as returns from soil, forests, fields, etc. Discrete clutter sources are less frequently occurring and spatially distributed with a particular density as a function of RCS, leading to strong, point-like returns. Since clutter discrete returns occur relatively infrequently, their suppression is more challenging and they tend to drive false alarm rate. Many discrete returns result from manmade objects, like buildings, water towers, utility poles, etc. Fence lines, train tracks, and power lines are examples of extended sources of discrete clutter.

We discuss basic distributed and discrete clutter models.

2.12.3.1.1 Distributed clutter

Distributed clutter returns result from the integrated response of all scatterers within the range resolution cell. A model for distributed clutter involves discretizing the range resolution cell into fine angle bins called clutter patches. The location of the angle bin relative to the platform velocity vector determines the clutter patch Doppler frequency. The clutter snapshot then follows as the sum of the returns from each of the individual patches.

Figure 12.6 depicts the discretized clutter patch model for distributed clutter. The antenna gain varies around the range resolution annulus on the Earth's surface in accordance with the array normal and steering direction. It is common to model the clutter complex voltage as a Gaussian random variable due to the constructive and destructive sum of the returns from the sub-scatterers comprising each resolution cell. The patch clutter-to-noise ratio (CNR) at the channel-level (assuming matched channels) is

$$\text{CNR}_k(p) = \left(\frac{P_t G_t(\phi_p, \theta_p)}{4\pi r^2} \right) \left(\frac{\sigma_{k,p}}{4\pi r^2} \right) \left(\frac{g_c(\phi_p, \theta_p) \lambda^2}{4\pi k_B T_o B_r F_n L_{rf}} \right) G_{sp}, \quad (12.67)$$

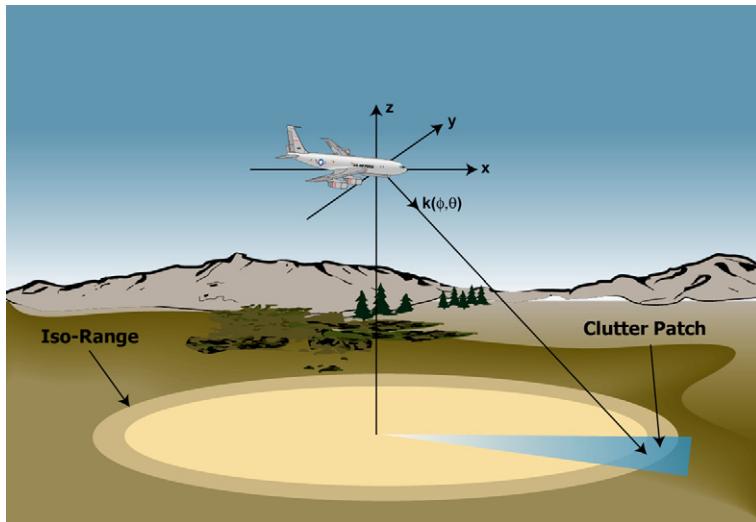
**FIGURE 12.6**

Illustration of space-time clutter patch calculation (after [9], © 2003 IEEE).

where P_t is the transmit power, $G_t(\phi_p, \theta_p)$ is the transmit gain in the p th direction of interest, r is the slant range, $\sigma_{k,p}$ is the p th clutter patch radar cross section, $g_c(\phi_p, \theta_p)$ is the receive channel gain, λ is the center wavelength, k_B is Boltzman's constant, T_o is standard operating temperature, B_r is the receiver bandwidth, F_n is noise figure, L_{rf} is RF system loss, and G_{sp} is signal processing gain (the pulse compression ratio, in this case) [1, 7, 8]. The clutter RCS is a function of the patch reflectivity, $\sigma_{k,p}^0$, and area, A_p , where the area is determined to be a fraction of the resolution cell,

$$\sigma_{k,p} = \sigma_{k,p}^0 A_p. \quad (12.68)$$

The constant gamma model is a common choice used to model reflectivity and is given as $\sigma_{k,p}^0 = \gamma_{k,p} \sin(\psi_p)$, where $\gamma_{k,p}$ is a normalized reflectivity term dependent on the terrain type and ψ_p is the grazing angle to the clutter patch [20]. The clutter patch power then follows by multiplying (12.67) by the receiver noise, σ_n^2 . The clutter voltage for the k th range bin and p th patch voltage is then

$$v_k(p) = \varsigma_{k,p} \sigma_n \sqrt{\text{CNR}_k(p)}, \quad (12.69)$$

where $\varsigma \sim \text{CN}(0, 1)$ [9–11, 18, 19]. (If the response is non-fluctuating, as in the case of a clutter discrete subsequently discussed, then ς is a complex scalar with unity magnitude and uniformly distributed phase.)

The patch voltage in (12.69) varies over the space-time aperture: there is a phase change from channel-to-channel due to direction of arrival, described by the spatial steering vector; and, a phase change from pulse-to-pulse due to the change in range-rate between the clutter patch and the radar platform, described by the temporal steering vector. The clutter space-time snapshot then follows by summing the voltages over the P patches for each of Q range ambiguities, where $q = 0$ is the first

(unambiguous) range return

$$\mathbf{c}_k = \sum_{q=0}^Q \sum_{p=1}^P v_k(p, q) \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)). \quad (12.70)$$

It is common to assume each patch is statistically independent, in which case the clutter covariance matrix corresponding to (12.70) is

$$\mathbf{R}_{c/k} = \sum_{q=0}^Q \sum_{p=1}^P \sigma_{c,k}^2(p, q) \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)) \mathbf{s}_{s-t}^H(\phi_p(q), \theta_p(q), f_{d,p}(q)), \quad (12.71)$$

where $\sigma_{c,k}^2(p, q) = E[|v_k(p, q)|^2]$ is the clutter patch power and follows from (12.69). Equations (12.70) and (12.71) form the basic ground clutter models.

The clutter voltage decorrelates, mainly in time due to intrinsic clutter motion (ICM). Windswept vegetation and moving water are two cases where ICM results. The two basic models describing clutter temporal decorrelation include the Billingsley model [21] and the Gaussian model [7, 10]. The Billingsley model is most common for overland surveillance; it allows a certain amount of the clutter power to decorrelate, say due to leaves fluctuating in the breeze, whilst some of the clutter power is persistent, thus modeling the tree trunk, for example. The Billingsley model leads to an exponential autocorrelation function. In contrast, the Gaussian power spectrum of [7] leads to a Gaussian autocorrelation; the Gaussian model leads to complete decorrelation over a specified time interval and is most appropriate in marine or riverine environments. Dispersion (group delay) across the array is the common source of spatial decorrelation; a sinc autocorrelation model is commonly chosen to model this effect, since it is the inverse transform of a rectangular function in the frequency domain [22, 23].

Given the aforementioned discussion, define the length- NM space-time correlation taper as $\mathbf{a}_{s-t} = \mathbf{a}_t \otimes \mathbf{a}_s$, with correlation matrix $\mathbf{A}_{s-t} = E[\mathbf{a}_{s-t} \mathbf{a}_{s-t}^H] = \mathbf{A}_t \otimes \mathbf{A}_s$. \mathbf{A}_t is the temporal correlation matrix and \mathbf{A}_s is the spatial correlation matrix. The resulting clutter covariance matrix is

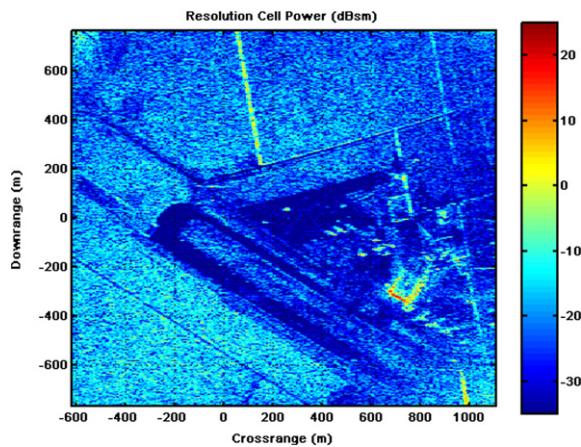
$$\mathbf{R}_{c/k} = \sum_{q=0}^Q \sum_{p=1}^P \sigma_{c,k}^2(p, q) \mathbf{A}_{s-t,k}(p, q) \odot \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)) \mathbf{s}_{s-t}^H(\phi_p(q), \theta_p(q), f_{d,p}(q)). \quad (12.72)$$

Generating space-time clutter snapshots exhibiting the correlation described by \mathbf{A}_{s-t} typically involves shaping white noise by multiplying the matrix square root of \mathbf{A}_t and \mathbf{A}_s by random vectors of length N or M , where the elements of each vector are uncorrelated, zero mean, unity variance Gaussian variates.

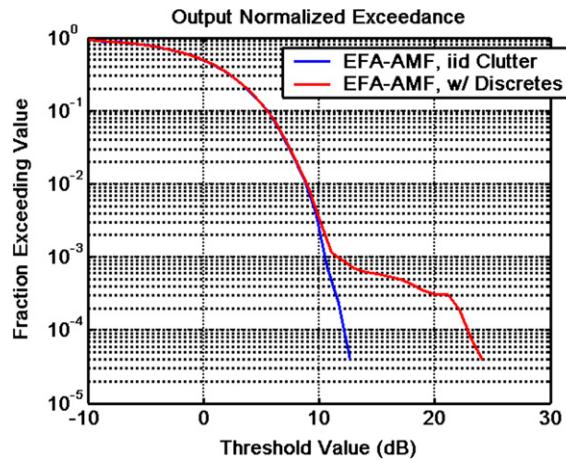
2.12.3.1.2 Discrete clutter

Clutter discretes are strong returns that occur relatively infrequently and correspond to stationary, point-like objects, such as parked vehicles, utility poles, water towers, etc. Generally, the RCS of the clutter discrete is large relative to the RCS of a typical, resolved clutter patch. In some cases, the discrete clutter can appear extended, such as in the case of a fence line or train track.

Figure 12.7 shows an example of discrete-like returns in a spotlight synthetic aperture radar (SAR) image of the Mojave Desert Airport. The hangar edges and fence lines, as well as some aircraft on the tarmac, are evident in the figure. Figure 12.8 shows an exceedance plot characterizing the typical

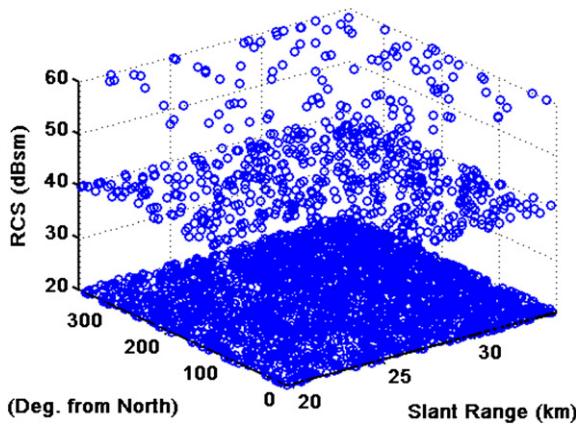
**FIGURE 12.7**

Spotlight SAR image of Mojave Desert Airport at 1 m resolution ([24], © 2004 IEEE).

**FIGURE 12.8**

Example exceedance plot showing the impact of discrete clutter on STAP performance.

impact of clutter discretes on detection performance. (EFA is a post-Doppler STAP method described in Section 2.12.4 and AMF refers to a STAP normalization discussed in Section 2.12.7.) In a homogeneous environment, the output of the STAP is well-behaved and the cumulative distribution follows the expected exponentially-shaped exceedance; in contrast, with discretes present, the tails of the exceedance plot are extended, indicating a requirement to raise the detection threshold to maintain a constant false alarm rate.

**FIGURE 12.9**

Example of clutter discrete seeding.

A plausible discrete model involves laying out the spatial locations of discrete clutter of varying densities. One way to do this is to specify the density per km² for a particular discrete RCS and employ a Poisson distribution to identify the random clutter discrete locations. An example of a range-angle seeding of discrete clutter of varying RCS from 20 dBsm to 60 dBsm is show in Figure 12.9.

Since discrete clutter is stationary, the angular position uniquely specifies Doppler frequency. In other words, clutter discretes reside along the angle-Doppler region of support corresponding to stationary objects. As in the case of distributed clutter, the discrete-to-noise ratio (DNR) follows in a form similar to (12.67) for the l th discrete at range, r_l , and angle (ϕ_l, θ_l) , as

$$\text{DNR}(l) = \left(\frac{P_t G_t(\phi_l, \theta_l)}{4\pi r_l^2} \right) \left(\frac{\sigma_{d,l}}{4\pi r_l^2} \right) \left(\frac{g_c(\phi_l, \theta_l)\lambda^2}{4\pi k_B T_o B_r F_n L_{rf}} \right) G_{sp}. \quad (12.73)$$

As the discrete is considered point-like, the RCS is taken as nonfluctuating, whereas the phase is uniformly distributed. The corresponding discrete space-time snapshot is

$$\mathbf{c}_d(l; k(l)) = e^{j2\pi u} \sigma_n \sqrt{\text{DNR}(l)} \mathbf{s}_{s-t}(\phi_l, \theta_l, f_{d,l}), \quad (12.74)$$

where u is a uniformly-distributed random variable between [0, 1] and $k(l)$ is the range realization for the l th clutter discrete. The discrete clutter is additive. The corresponding covariance matrix for each discrete is

$$\mathbf{R}_{cd}(l; k(l)) = \sigma_n^2 \text{DNR}(l) \mathbf{s}_{s-t}(\phi_l, \theta_l, f_{d,l}) \mathbf{s}_{s-t}^H(\phi_l, \theta_l, f_{d,l}). \quad (12.75)$$

2.12.3.2 Radio frequency interference

Typically, RFI appears as a noise-like, in-band signal source. The RFI can be intentional or not. In the narrowband case, the k th spatial snapshot for the i th RFI source and n th pulse is

$$\mathbf{j}_{s/k}(n; i) = \omega_k(nT) \mathbf{s}_s(\phi_i, \theta_i), \quad (12.76)$$

where $\omega_k(nT)$ is an uncorrelated source and T is the PRI. The waveform, $\omega_k(nT)$, is generally uncorrelated at nonzero lags, viz.,

$$E[\omega_k(nT)\omega_k^*(mT)] = \sigma_{J,i}^2 \delta((n-m)T) \quad (12.77)$$

but otherwise fully correlated for very short time lags corresponding to propagation across the multi-channel array, as (12.76) indicates (thus, $E[\omega_k(nT)\omega_j^*(nT)] = 0$, for $k \neq j$, in the narrowband case). The corresponding spatial covariance matrix is

$$\mathbf{R}_{\text{s/RFI}}(i) = \sigma_{J,i}^2 \mathbf{s}_s(\phi_i, \theta_i) \mathbf{s}_s^H(\phi_i, \theta_i) \in C^{M \times M}. \quad (12.78)$$

As a result of (12.77), the space-time covariance matrix is

$$\mathbf{R}_{\text{RFI}}(i) = E \left[\begin{bmatrix} \mathbf{j}_{s/\mathbf{k}}(0; i) \\ \vdots \\ \mathbf{j}_{s/\mathbf{k}}((N-1); i) \end{bmatrix} \begin{bmatrix} \mathbf{j}_{s/\mathbf{k}}^H(0; i) & \cdots & \mathbf{j}_{s/\mathbf{k}}^H((N-1); i) \end{bmatrix} \right] = \mathbf{I}_N \otimes \mathbf{R}_{\text{s/RFI}}(i) \in C^{NM \times NM}. \quad (12.79)$$

Thus, a spatial null is sufficient to remove the narrow-band RFI, as no temporal correlation is present.

Each RFI source is considered statistically independent, so that for J sources, $\mathbf{j}_{\mathbf{k}} = \sum_{i=1}^J \mathbf{j}_{\mathbf{k}}(i)$, where $\mathbf{j}_{\mathbf{k}} \in C^{NM \times 1}$ is the space-time snapshot and $\mathbf{R}_{\text{RFI}} = \sum_{i=1}^J \mathbf{R}_{\text{RFI}}(i)$.

As the fractional bandwidth—the ratio of waveform bandwidth to center frequency—of the receive signal increases, dispersion occurs. Dispersion leads to decorrelation of the RFI over the receive array; in this case, a single RFI signal appears as multiple, closely spaced sources in angle [23]. In general, the wideband RFI suppression case is handled similar to the narrowband case through the use of subband filtering or the use of fast-time taps. In the former case, the subbanding is commonly implemented using polyphase filtering, allowing the processor to implement the narrowband canceler in each subband prior to recombining [25].

2.12.3.3 Receiver noise

Uncorrelated noise sources—such as receiver noise or sky noise—are modeled as a white Gaussian noise (WGN) disturbance, $\mathbf{n}_{\mathbf{k}} = [\beta_k(m, nT)]_{m=1:M, n=0:N}$, where m and n are channel and pulse indices, respectively, and $\beta_k(m, nT) \sim \text{CN}(0, \sigma_n^2)$. The waveform samples are assumed uncorrelated such that

$$\mathbf{R}_{\mathbf{n}} = \sigma_n^2 \mathbf{I}_{NM}, \quad (12.80)$$

where σ_n^2 is the channel noise power. This noise source is also uncorrelated, independent, and identically distributed over the range dimension.

2.12.3.4 Target

The target snapshot is generally modeled using both Swerling 1 and Swerling 2 statistics [1, 7, 8]. Swerling 1 and Swerling 2 targets each exhibit voltages corresponding to a circular Gaussian distribution; the Gaussian nature of the voltage distribution models target fading effects. The target voltage is assumed

perfectly correlated within the CPI, in accord with the Swerling 1 model, and uncorrelated from CPI-to-CPI according to the Swerling 2 target model. Frequency hopping from CPI-to-CPI is a common reason for target voltage decorrelation and is used to minimize the impact of target fading on detection performance. STAP is applied on a CPI basis, as it is a coherent signal processing technique, with noncoherent addition (NCA) applied to the STAP output—from CPI-to-CPI—to mitigate target fading effects.

Using the Swerling 1 target model, the target snapshot at the k th range, Doppler frequency, $f_{d,T}$, and angle, (ϕ_T, θ_T) , is

$$\mathbf{t}_k = \varsigma_T \sigma_n \sqrt{\text{SNR}} s_{\mathbf{s}, \mathbf{t}}(f_{d,T}, \phi_T, \theta_T), \quad (12.81)$$

where $\varsigma_T \sim \text{CN}(0, 1)$ and SNR is the single channel, single pulse signal-to-noise ratio. The random variable, ς_T , models targeting fading resulting from subscatterers adding in and out of phase. As indicated, to handle fading, it is common to frequency hop from CPI-to-CPI, in which case target voltages appear decorrelated (Swerling 2).

Nonfluctuating, point-target analysis employs (12.81) with $\varsigma_T \rightarrow e^{j2\pi u}$, where u is uniformly distributed between $[0, 1]$. Sometimes the nonfluctuating target model is used in analysis. However, the Swerling 2 model is the preferred choice, with the Swerling 1 model applicable at the fixed frequency, CPI-level.

2.12.3.5 The space-time snapshot

The space-time snapshot for realization, k , is given by (12.2) with the addition of clutter discrete, where the possibility of the single target case is

$$\begin{aligned} H_0 : \mathbf{x}_k &= \mathbf{c}_k + \sum_{l=1}^D \mathbf{c}_{\mathbf{d}}(l; k(l)) + \mathbf{j}_k + \mathbf{n}_k, \\ H_1 : \mathbf{x}_k &= \mathbf{x}_k / \mathbf{H}_0 + \mathbf{t}_k, \end{aligned} \quad (12.82)$$

where D is the total number of discrete scatterers and the snapshot terms, $\mathbf{c}_{\mathbf{d}}(l; k(l))$, are only added when $k = k(l)$. The null-hypothesis covariance matrix corresponding to (12.82) is

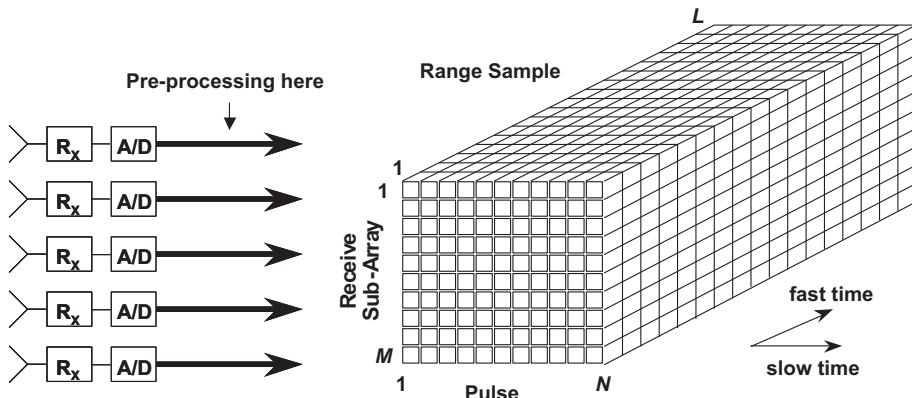
$$\mathbf{R}_k = \mathbf{R}_{\mathbf{c}, k} + \sum_{l=1}^D \mathbf{R}_{\mathbf{cd}}(l; k(l)) + \sum_{i=1}^J \mathbf{R}_j(i) + \mathbf{R}_n, \quad (12.83)$$

where, as expected, $\mathbf{R}_{\mathbf{cd}}(l; k(l))$ is included only for those terms where $k = k(l)$.

It is common to envision the collection of space-time snapshots organized as space-time data matrices,

$$\mathbf{X}_k = [\mathbf{x}_{s,k}(1) \ \mathbf{x}_{s,k}(2) \ \cdots \ \mathbf{x}_{s,k}(N)] \in C^{M \times N}. \quad (12.84)$$

The spatial snapshot for pulse, n , is denoted as $\mathbf{x}_{s,k}(n) \in C^{M \times 1}$; it results by removing N length- M segments from \mathbf{x}_k and stacking them side-by-side. The pictorial of (12.84) is given in Figure 12.10 and is known as the radar datacube. Generally, the STAP operates on each space-time data matrix, or slice, of the cube in Figure 12.10, while using slices at other ranges (or realizations) for training. We subsequently describe various adaptive implementations. (Note: space/fast-time processing uses slices along the pulse/fast-time domain and can train over the pulse dimension. As previously mentioned, we focus on space/slow-time adaptivity, but the basic formulations of the next section generally apply to any appropriately formatted data with the corresponding restrictions of each method.)

**FIGURE 12.10**

Radar datacube (after [9], © 2003 IEEE).

2.12.4 Adaptive filter implementations

Section 2.12.2.2 describes space-time filter formulations. As seen from the discussion, all solutions are similar, involving space-time weightings of the form of a covariance inverse and a steering vector to implement the matched filter.

We now discuss three basic paradigms to implement the weighting strategies discussed in Section 2.12.2.2: reduced-rank STAP (RR-STAP), reduced-dimension STAP (RD-STAP), and parametric STAP.

2.12.4.1 Reduced-rank STAP

As seen from (12.78), each narrowband RFI source is rank-one. Distributed clutter is oftentimes of rank significantly less than the dimensionality of \mathbf{x}_k , viz. $r_\lambda \ll \dim(\mathbf{x}_k)$. Through empirical analysis, the clutter rank for a sidelooking radar with minimal yaw is approximated as

$$r_\lambda \approx \left\lfloor N + (M - 1) \left(\frac{v_p T}{d/2} \right) \right\rfloor, \quad (12.85)$$

where v_p is the platform along-track velocity and d is the separation between channels of a uniform linear array (ULA) [10]. The expression in (12.85) is known as Brennan's Rule; a related rule is given by Klemm [26]. Brennan's rule is closely related to the number of independent antenna channel positions during the collection interval; redundancy in the measurements lowers the rank, and in fact leads to coloration of the clutter return. The idea behind RR-STAP is to essentially project the interference subspace—those eigenspaces corresponding to larger eigenvalues above the noise floor—out of the space-time snapshot.

The eigendecomposition of the space-time covariance matrix yields

$$\mathbf{R}_k = \sum_{m=1}^{NM} \lambda_{k/m} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H = \mathbf{Q}_{k/I} \Lambda_{k/I} \mathbf{Q}_{k/I}^H + \mathbf{Q}_{k/N} \Lambda_{k/N} \mathbf{Q}_{k/N}^H, \quad (12.86)$$

where $\lambda_{k/m}$ is the m th eigenvalue corresponding to eigenvector $\mathbf{q}_{\mathbf{k}/\mathbf{m}} \in C^{NM \times 1}$, $\mathbf{\Lambda}_{\mathbf{k}/\mathbf{I}} = \text{diag}(\lambda_{k/m})_{m=1}^P$ represents the interference subspace eigenvalues, and $\mathbf{\Lambda}_{\mathbf{k}/\mathbf{N}} = \text{diag}(\lambda_{k/m})_{m=P+1}^{NM}$ characterizes the noise subspace eigenvalues. The interference eigenvectors have the special property that each span the collection of interference steering vectors. Moreover, $\mathbf{Q} = [\mathbf{Q}_{\mathbf{k}/\mathbf{I}} \ \mathbf{Q}_{\mathbf{k}/\mathbf{N}}]$ is unitary, so

$$\mathbf{Q}_{\mathbf{k}/\mathbf{I}} \mathbf{Q}_{\mathbf{k}/\mathbf{I}}^H + \mathbf{Q}_{\mathbf{k}/\mathbf{N}} \mathbf{Q}_{\mathbf{k}/\mathbf{N}}^H = \mathbf{I}_{NM}. \quad (12.87)$$

We can write the distributed clutter-plus-noise part of (12.83) in a generic, simplified form

$$\mathbf{R}_{\mathbf{k}} = \sum_{p=1}^{\tilde{N}} \sigma_{s/p}^2 \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi_p, \theta_p, f_{d/p}) + \sigma_n^2 \mathbf{I}_{NM}, \quad (12.88)$$

where \tilde{N} is the number of signal sources and $\sigma_{s/p}^2$ is the power for the p th signal source. While this simplified covariance form suggests a model for clutter-plus-noise, the subsequent results are extensible to other correlated sources. Then, from (12.88),

$$\mathbf{R}_{\mathbf{k}} \mathbf{q}_{\mathbf{k}/\mathbf{m}} = \lambda_{k/m} \mathbf{q}_{\mathbf{k}/\mathbf{m}} = \sum_{p=1}^{\tilde{N}} \sigma_{s/p}^2 \gamma_{k/m,p} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) + \sigma_n^2 \mathbf{q}_{\mathbf{k}/\mathbf{m}}, \quad (12.89)$$

where

$$\gamma_{k/m,p} = \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi_p, \theta_p, f_{d/p}) \mathbf{q}_{\mathbf{k}/\mathbf{m}}. \quad (12.90)$$

Solving (12.89) for the m th eigenvector, $\mathbf{q}_{\mathbf{k}/\mathbf{m}}$,

$$\mathbf{q}_{\mathbf{k}/\mathbf{m}} = \sum_{p=1}^{\tilde{N}} \left(\frac{\sigma_{s/p}^2 \gamma_{k/m,p}}{\lambda_{k/m} - \sigma_n^2} \right) \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) = \sum_{p=1}^{\tilde{N}} \alpha_{k/m,p} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}), \quad (12.91)$$

where in this case

$$\alpha_{k/m,p} = \frac{\sigma_{s/p}^2 \gamma_{k/m,p}}{\lambda_{k/m} - \sigma_n^2}. \quad (12.92)$$

Naturally, these equations are only valid for the dominant subspace, $\lambda_{k/m} - \sigma_n^2 > 0$. Equation (12.91) shows $\mathbf{q}_{\mathbf{k}/\mathbf{m}} \in \text{span} \left(\{\mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p})\}_{p=1}^{\tilde{N}} \right)$.

Also, it is known that any wide-sense stationary (WSS) process with zero mean and covariance matrix, $\mathbf{R}_{\mathbf{k}}$, can be written as a linear combination of the eigenvectors of $\mathbf{R}_{\mathbf{k}}$, via what is known as the Karhunen-Loëve Transform (KLT),

$$\mathbf{x}_{\mathbf{k}} = \sum_{m=1}^{NM} a_{k/m} \mathbf{q}_{\mathbf{k}/\mathbf{m}}, \quad (12.93)$$

where $a_{k/m} = \mathbf{q}_{\mathbf{k}/\mathbf{m}}^H \mathbf{x}_{\mathbf{k}}$ are the Karhunen-Loëve (KL) coefficients [12].

The prior expressions, (12.86)–(12.93), provide necessary mathematical background for our subsequent discussion on RR-STAP methods.

Motivation for RR-STAP includes the following:

- Clutter and jamming tend to be of low numerical rank and the processor explicitly removes only those signal subspaces corresponding to interference.
- The eigendecomposition maximally compresses the interference into the fewest basis vectors [12].
- The RMB rule applies to the RR-STAP case, with rank replacing DoFs in the formulation, viz. training over twice the rank yields roughly 3 dB loss on average [27].

In effect, RR-STAP is a weight calculation strategy. There are challenges implementing RR-STAP, including high computational burden and difficulties in rank determination. Nevertheless, RR-STAP methods hold meaningful insight and, in some cases, are useful in weight vector determination.

2.12.4.1.1 Adaptive beamformer pattern synthesis

It is shown in [28] that (12.43) can be written

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \frac{\mu}{\lambda_0} \left[\mathbf{s}_{s-t}(\phi, \theta, f_d) - \sum_{m=1}^{NM} \frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \vartheta_{k/m} \mathbf{q}_{k/m} \right], \quad (12.94)$$

where $\lambda_0 = \min(\{\lambda_{k/m}\}_{\forall m})$ is the noise-floor eigenvalue level and $\vartheta_{k/m} = \mathbf{q}_{k/m}^H \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the projection of the m th eigenvector onto the quiescent pattern. As seen from (12.94), the STAP response appears as a notching of the space-time beampattern given by $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ by the weighted, interference eigenvectors. When $\lambda_{k/m} = \lambda_0$, no subtraction occurs, since the corresponding eigenvector lies in the noise subspace.

Naturally, (12.94) applies to the SMI formulation of (12.44), and can provide robustness in the presence of training sample support limitations if the interference rank is known (which is usually only practical for strong interferers). Specifically, running the sum in (12.94) only over $r_\lambda \ll NM$ subspaces leads to adaptive pattern robustness, since low sample support tends to predominantly perturb the noise subspace; the perturbed noise subspace estimate leads to poor adaptive sidelobe performance when the sum in (12.94) is run over all values of the index, m .

2.12.4.1.2 Principle components inverse

The idea behind principle components inverse (PCI) is to apply an orthogonal projection to \mathbf{x}_k , and then apply a matched filter to the remaining transformed data [29]. The PCI formulation starts by writing (12.94) applied to the data as

$$y_k = \frac{\mu \mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \left[\mathbf{I}_{NM} - \sum_{m=1}^{NM} \frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k. \quad (12.95)$$

Then, for the $P \ll NM$ eigenvalues where $\lambda_{m/k} \gg \lambda_0$,

$$y_k \approx \frac{\mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \left[\mathbf{I}_{NM} - \sum_{m=1}^{P \ll NM} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k = \frac{\mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \tilde{\mathbf{x}}_k. \quad (12.96)$$

We see that

$$\tilde{\mathbf{x}}_k = \left[\mathbf{I}_{NM} - \sum_{m=1}^{P \ll NM} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k \stackrel{\text{e.g.}}{=} \mathbf{x}_k - a_{k/1} \mathbf{q}_{k/1} - a_{k/2} \mathbf{q}_{k/2}, \quad (12.97)$$

where, for the case of $P = 2$ as an example, (12.97) shows the coherent removal of $\mathbf{q}_{k/1}$ and $\mathbf{q}_{k/2}$. The term in brackets in (12.97) is called an orthogonal projection for this reason.

2.12.4.1.3 Eigencanceler

Haimovich [30] gives two different developments; one of these essentially leads to (12.96) via an alternate route.

The minimum power eigencanceler (MPE) weighting results from

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k \text{ such that } \mathbf{Q}_{k/I}^H \mathbf{w}_k = \mathbf{0} \text{ and } \mathbf{C}^H \mathbf{w}_k = \mathbf{f}, \quad (12.98)$$

where \mathbf{C} is a constraint matrix and \mathbf{f} is a desired response vector. The solution to (12.98) is shown in [30] to be

$$\mathbf{w}_{k/\text{EC-MPE}} = \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{C} \left[\mathbf{C}^H \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{C} \right]^{-1} \mathbf{f}. \quad (12.99)$$

For the case where $\mathbf{C} = \mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{f} = g$, (12.99) takes the form

$$\mathbf{w}_{k/\text{EC-MPE}} = \mu_0 \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.100)$$

with $\mu_0 = \left[\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right]^{-1} g$ a scalar. The weight vector in (12.100) lies entirely in the noise subspace—as required by (12.98)—and will, by virtue of the orthonormal property of the set of eigenvectors, annihilate the clutter subspace.

The minimum norm eigencanceler (MNE) is formulated as

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{w}_k \text{ subject to } \mathbf{Q}_{k/I}^H \mathbf{w}_k = \mathbf{0} \text{ and } \mathbf{C}^H \mathbf{w}_k = \mathbf{f}. \quad (12.101)$$

The MNE weight vector is then shown in [30] to be

$$\mathbf{w}_{k/\text{EC-MNE}} = \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{C} \left[\mathbf{C}^H \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{C} \right]^{-1} \mathbf{f}. \quad (12.102)$$

Again, for the case where $\mathbf{C} = \mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{f} = g$, and using (12.87), we see (12.102) can be written in the same form as PCI,

$$\mathbf{w}_{k/\text{EC-MNE}} = \mu_1 \left[\mathbf{I}_{NM} - \mathbf{Q}_{k/I} \mathbf{Q}_{k/I}^H \right] \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.103)$$

with $\mu_1 = \left[\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right]^{-1} g$ a scaling that replaces $1/\lambda_0$ in (12.96). Using (12.87), the expression in (12.103) can be written

$$\mathbf{w}_{k/\text{EC-MNE}} = \mu_1 \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.104)$$

which closely relates to the MPE solution in (12.100) with the inverse eigenvalue weighting given by $\Lambda_{k/N}^{-1}$ absent. It is known that limited training sample support leads to perturbed estimates of the noise eigenvalues appearing along the diagonal of $\Lambda_{k/N}$. For this reason, the MNE provides robustness relative to MPE when training data samples are lacking.

2.12.4.1.4 Hung turner projection

The Hung Turner Projection (HTP) is applicable to RFI suppression [31]. While clutter mitigation is paramount to our discussion, it is worth taking a moment to describe the HTP as a general tool in adaptive filter implementation. Reducing computational burden associated with the eigendecomposition leading to the eigenvalues and eigenvectors of the null-hypothesis covariance matrix is a primary goal.

The basic idea behind the HTP is to use the Gram-Schmidt method to characterize the interference subspace. Specifically, suppose we identify J snapshots where J RFI sources are present,

$$\begin{aligned} \mathbf{x}_k &= \mathbf{j}_k(1) + \mathbf{j}_k(2) + \cdots + \mathbf{j}_k(J) + \mathbf{n}_k, \\ \mathbf{x}_{k+1} &= \mathbf{j}_{k+1}(1) + \mathbf{j}_{k+1}(2) + \cdots + \mathbf{j}_{k+1}(J) + \mathbf{n}_{k+1}, \\ &\vdots \\ \mathbf{x}_{k+J-1} &= \mathbf{j}_{k+J-1}(1) + \mathbf{j}_{k+J-1}(2) + \cdots + \mathbf{j}_{k+J-1}(J) + \mathbf{n}_{k+J-1}. \end{aligned} \quad (12.105)$$

Next, the modified Gram-Schmidt technique is applied to (12.105) to generate the orthonormal basis, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_J\}$, where $\mathbf{v}_m \in \text{span}\{\mathbf{j}_1, \mathbf{j}_2, \dots, \mathbf{j}_J\}$. Then, in accord with [31], the HTP weight vector is chosen as $\mathbf{w}_{\text{HTP}} = \bar{\mathbf{w}}_Q^0 / \|\bar{\mathbf{w}}_Q^0\|_2$, with

$$\bar{\mathbf{w}}_Q^0 = \bar{\mathbf{w}}_Q - \sum_{m=1}^J \left(\mathbf{v}_m^H \bar{\mathbf{w}}_Q \right) \mathbf{v}_m, \quad (12.106)$$

and where $\bar{\mathbf{w}}_Q$ is the quiescent weight vector.

2.12.4.1.5 Diagonal loading

Diagonal loading involves adding a scaled diagonal matrix to the covariance matrix estimate of (12.49) [32],

$$\hat{\mathbf{R}}_k = \frac{1}{K} \sum_{m=1}^K \mathbf{x}_m \mathbf{x}_m^H + \sigma_{\text{DL}}^2 \mathbf{I}_{NM} = \hat{\mathbf{Q}}_k \hat{\Lambda}_k \hat{\mathbf{Q}}_k^H + \sigma_{\text{DL}}^2 \hat{\mathbf{Q}}_k \hat{\mathbf{Q}}_k^H = \hat{\mathbf{Q}}_k \left(\hat{\Lambda}_k + \sigma_{\text{DL}}^2 \mathbf{I}_{NM} \right) \hat{\mathbf{Q}}_k^H, \quad (12.107)$$

where σ_{DL}^2 is the diagonal loading level and $\hat{\mathbf{Q}}_k$ and $\hat{\Lambda}_k$ follow from the decomposition of the covariance matrix estimate. It is common to set $0.1\sigma_n^2 < \sigma_{\text{DL}}^2 < 10\sigma_n^2$, with $\sigma_{\text{DL}}^2 = \sigma_n^2$ most typical. Primary benefits of diagonal loading include improved conditioning of the covariance matrix estimate (which improves the numerical stability of inversion routines) and compression of the range corresponding to the smallest eigenvalues. This latter benefit leads to a reduction in spurious sidelobes of the adapted pattern at the expense of null depth. Effectively, diagonal loading removes some of the adaptivity of the system to better condition the filter's sidelobe response.

While often considered ad hoc, it turns out diagonal loading is a component of the optimal solution to the constrained optimization,

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k \text{ such that } \mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = 1 \text{ and } \mathbf{w}_k^H \mathbf{w}_k \leq \beta^2. \quad (12.108)$$

The Lagrangian for (12.108) is readily found to be

$$L_c = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k + \lambda_1 \left(\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) - 1 \right) + \lambda_2 \left(\mathbf{w}_k^H \mathbf{w}_k - \beta^2 \right), \quad (12.109)$$

where λ_1 and λ_2 are Lagrangian multipliers. Taking the gradient of (12.109) with respect to the conjugate of the weight vector and setting the result to zero yields [13]

$$\nabla_{\mathbf{w}^H} L_c = \mathbf{R}_k \mathbf{w}_k + \lambda_1 \mathbf{s}_{s-t}(\phi, \theta, f_d) + \lambda_2 \mathbf{w}_k = 0. \quad (12.110)$$

Solving for the weight vector gives

$$\mathbf{w}_k = -\lambda_1 (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.111)$$

Using the linear constraint, it is seen that

$$\lambda_1 = \frac{-1}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.112)$$

The resulting weight vector is

$$\mathbf{w}_k = \frac{(\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.113)$$

From (12.113) we see that diagonal loading plays a key role in the solution to the constrained optimization. In this case, λ_2 is the diagonal loading level, σ_{DL}^2 , mentioned previously. If the loading level is set to zero, (12.113) defaults to the distortionless MV beamformer solution.

2.12.4.1.6 Cross spectral metric

Consider the GSLC structure of Figure 12.5. A principal components approximation to $\mathbf{R}_{x_{0/k}}$ is

$$\mathbf{R}_{x_{0/k}} \approx \sum_{m=1}^P \lambda_{x_{0/k};m} \mathbf{q}_{x_{0/k};m} \mathbf{q}_{x_{0/k};m}^H \quad (12.114)$$

with, ideally, $P \ll NM$, and where $\lambda_{x_{0/k};m}$ and $\mathbf{q}_{x_{0/k};m}$ represent the m th eigenvalue and eigenvector of the GSLC lower leg covariance matrix. Then,

$$\mathbf{w}_{MMSE-PC/k} = \sum_{m=1}^P \left(\frac{\mathbf{q}_{x_{0/k};m}^H \mathbf{r}_{x_{0/k}} \mathbf{d}_{0/H0}}{\lambda_{x_{0/k};m}} \right) \mathbf{q}_{x_{0/k};m}. \quad (12.115)$$

The computational burden of the GSLC does not generally justify its use in practice. Improved statistical convergence is a potential advantage of the principal components decomposition of the GSLC (PC-GSLC).

The cross spectral metric operates in a fashion similar to the PC-GSLC method, except now the selection of the basis defining the rank-reducing transformation is target signal-dependent [33]. The error signal at the output of the GSLC follows as the difference between (12.19) and (12.21), viz.

$$\varepsilon_{o/k} = d_{o/k} - \hat{d}_{o/k}. \quad (12.116)$$

The corresponding MMSE is given as the expected value of the magnitude squared of (12.116),

$$\begin{aligned} E \left[|\varepsilon_{o/k}|^2 \right] &= E \left[\varepsilon_{o/k} \varepsilon_{o/k}^* \right] = \sigma_{d_{o/H0}}^2 - \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}^H \mathbf{R}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}^{-1} \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}} \\ &= \sigma_{d_{o/H0}}^2 - \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}^H \left(\sum_{m=1}^{NM} \lambda_{\mathbf{x}_{o/\mathbf{k};m}}^{-1} \mathbf{q}_{\mathbf{x}_{o/\mathbf{k};m}} \mathbf{q}_{\mathbf{x}_{o/\mathbf{k};m}}^H \right) \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}} \\ &= \sigma_{d_{o/H0}}^2 - \sum_{m=1}^{NM} \lambda_{\mathbf{x}_{o/\mathbf{k};m}}^{-1} \left| \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}^H \mathbf{q}_{\mathbf{x}_{o/\mathbf{k};m}} \right|^2. \end{aligned} \quad (12.117)$$

In this case, $\sigma_{d_{o/H0}}^2 = E \left[|d_{o/k}|^2 \right]$, where no target signal energy is present in the upper leg (an assumption that carries through in the calculation of the cross correlation vector, $\mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}$). When choosing only P terms in the summation of (12.117), where $P \ll NM$, the processor attains the lowest possible mean square error (MSE) by choosing subspaces with the largest cross-spectral metric terms, given as

$$\text{CSM}_m = \frac{\left| \mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}^H \mathbf{q}_{\mathbf{x}_{o/\mathbf{k};m}} \right|^2}{\lambda_{\mathbf{x}_{o/\mathbf{k};m}}}. \quad (12.118)$$

The target signal influences the selection via $\mathbf{r}_{\mathbf{x}_{o/\mathbf{k}} \mathbf{d}_{o/\mathbf{H0}}}$. Thus, the idea behind the CSM approach is to select only those subspaces from $\mathbf{R}_{\mathbf{x}_{o/\mathbf{k}}}$ that most greatly enhance performance.

A direct-form implementation of the CSM method is given in [34]. In this case, the objective involves maximizing SINR by choosing the appropriate, signal-dependent, reduced-rank subspace (hence, the analogous metric is called the SINR metric in [34]). Thus, decomposing output SINR in (12.7) using the inverse of the eigendecomposition in (12.86) gives

$$\text{SINR}_{\text{Max}} = \sigma_T^2 \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \left(\sum_{m=1}^{NM} \lambda_{k/m}^{-1} \mathbf{q}_{\mathbf{k}/\mathbf{m}} \mathbf{q}_{\mathbf{k}/\mathbf{m}}^H \right) \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi, \theta, f_d). \quad (12.119)$$

Maximizing SINR requires choosing the largest terms in (12.119) of the form,

$$\text{CSM}_{\text{DFP};m} = \frac{\left| \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \mathbf{q}_{\mathbf{k}/\mathbf{m}} \right|^2}{\lambda_{k/m}}, \quad (12.120)$$

where DFP signifies “direct form processor” (i.e., the approach using a weight vector of the form given in (12.43) or (12.44)). Thus, with limited sample support and uncertainty over the interference rank, selecting the $P \ll NM$ largest CSM terms of (12.120) leads to maximal output SINR. Commonly, then, the CSM-DFP chooses the components in the noise subspace, but this is an artifact of the development and not useful: terms with the smallest eigenvalues are generally chosen, with any signal-dependent correlation meant to aid in selecting the most significant terms lost in the formulation. To avoid this scenario, [34] suggests choosing those terms maximally impacting the weight vector. In other words, considering (12.94), selecting the P largest terms of

$$\frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \vartheta_{k/m} \mathbf{q}_{\mathbf{k}/\mathbf{m}} \quad (12.121)$$

most greatly impacts the adaptive process. Equation (12.121) is signal-dependent and forces the selection to lie in the dominant subspace.

2.12.4.1.7 Multi-Stage Wiener Filter

The Multi-Stage Wiener Filter (MWF) is a truncated decomposition of the GSLC of Figure 12.5 [35]. Its operation is best understood from the diagram in Figure 12.11, which shows a two-stage decomposition. From this figure we see that the estimation stage is broken into a series of smaller problems, where the weight vector is calculated as a scalar using the Wiener-Hopf equation [12, 13] in each stage. The filter, \mathbf{h}_i , is a normalized cross-correlation vector between its input, $\mathbf{x}_{i-1/k}$, and the output of the filter in the preceding leg, $d_{i-1/k}$; the intent of choosing \mathbf{h}_i in this manner is to maximize the correlation between the interference signal in each leg. Specifically,

$$\mathbf{h}_i = \frac{\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}}}{\|\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}}\|_2}, \quad (12.122)$$

where $\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}} = E[\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}^*]$. The blocking matrix, \mathbf{B}_i , lies in the null space of \mathbf{h}_i (except for $i = 0$, where $\mathbf{B}_0 = \mathbf{B}_0 = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$). The weights, w_i , are scalars equal to the Wiener weight minimizing the mean square error between $d_{i-1/k}$ and $\hat{d}_{i-1/k} = w_i^* \varepsilon_i$. In this figure, for the MWF output to equal the GSLC output, the processor must calculate all quantities, including the vector weight, $\mathbf{w}_{MWF,3} = (E[\mathbf{x}_{2/k} \mathbf{x}_{2/k}^H])^{-1} \mathbf{r}_{\mathbf{x}_{2/k} \mathbf{d}_{2/k}}$. The MWF truncates the number of stages by simply dropping the lower leg of the last stage, setting the vector weight to all zeros (i.e., $\mathbf{w}_{MWF,3} = \mathbf{0}$) so that $\varepsilon_{N_{MWF}} = d_{N_{MWF}/k}$, where N_{MWF} is the number of MWF stages. So, for $N_{MWF} = 2$, $\varepsilon_2 = d_{2/k} = \mathbf{h}_2^T \mathbf{x}_{1/k}$.

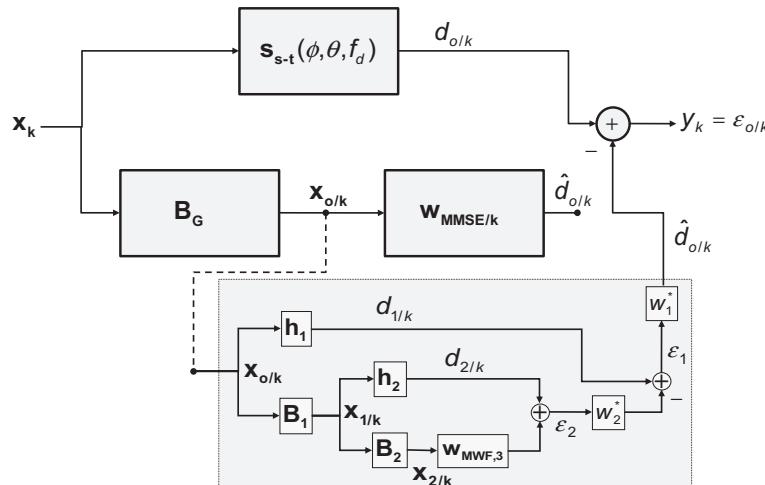


FIGURE 12.11

Multi-Stage Wiener Filter flow diagram.

In the adaptive version of the MWF, the processor replaces the known covariance matrices and steering vectors with estimates. The quantities, \mathbf{h}_i , \mathbf{B}_i , w_i , and, naturally, ε_i , are all expressible as linear transformations applied to \mathbf{R}_k and $\mathbf{s}_{s-t}(\phi, \theta, f_d)$; in the adaptive filter, $\mathbf{R}_k \rightarrow \hat{\mathbf{R}}_k$ and $\mathbf{s}_{s-t}(\phi, \theta, f_d) \rightarrow \mathbf{v}_{s-t}(\phi, \theta, f_d)$. The implementation requires first calculating and applying $\mathbf{B}_G = \mathbf{B}_0 = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$; next, the processor implements the first stage, calculating \mathbf{h}_1 , \mathbf{B}_1 , and $d_{1/k}$; with these first stage quantities in hand, the processor calculates the second stage quantities, \mathbf{h}_2 , \mathbf{B}_2 , and $d_{2/k}$, then third stage quantities, and so forth, working out to the last stage; then, as noted above, the error signal in the last stage is set to $\varepsilon_L = d_{L/k} = \mathbf{h}_L^H \mathbf{x}_{L-1/k}$, where $i = L$ is the last stage; and, finally, the scalar weights, w_i , are solved from the outer (last) stage into the first stage.

While the computational loading of the MWF is generally high, the target signal-dependent nature of the stage decomposition identifies the interference subspace most greatly impacting the MMSE through the cross-correlation process, estimates it, and subtracts it out. This allows the MWF to converge towards the optimal solution with minimal use of available training data.

2.12.4.2 Reduced-dimension STAP

Reduced-dimension STAP (RD-STAP) methods take advantage of the structure of the clutter angle-Doppler region of support to minimize filter length, thus reducing computational burden and training sample requirements. A taxonomy of RD-STAP methods, organized similarly to the discussion in [10], is given in Figure 12.12.

RD-STAP methods apply deterministic, linear filtering operations and dimensionality reduction prior to adaptive filtering. A linear transformation, $\mathbf{T} \in C^{NM \times \tilde{P}}$, where \tilde{P} is the length of the reduced

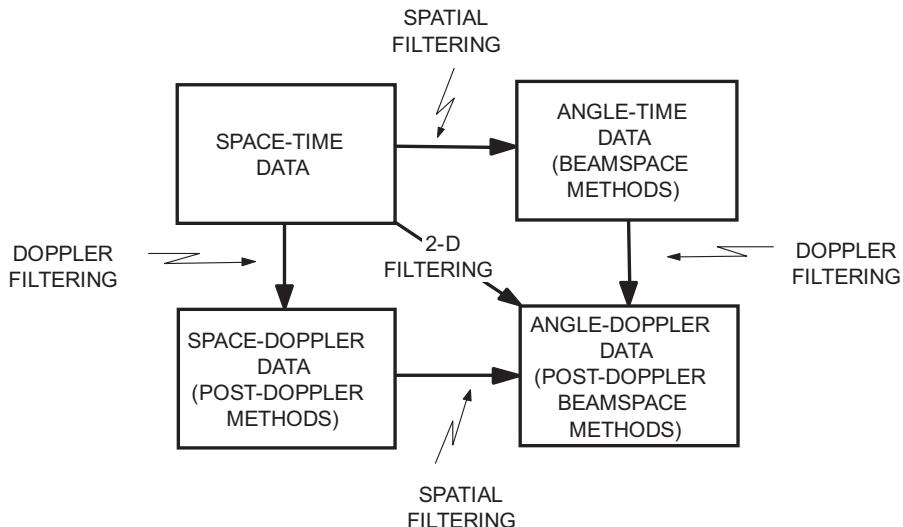


FIGURE 12.12

RD-STAP taxonomy.

dimension, describes these deterministic operations. Starting with the space-time snapshot, \mathbf{x}_k , the reduced-dimension snapshot is

$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \in C^{\tilde{P} \times 1}. \quad (12.123)$$

Via this transformation, the null-hypothesis covariance matrix is

$$\tilde{\mathbf{R}}_k = \mathbf{T}^H \mathbf{R}_k \mathbf{T} \in C^{\tilde{P} \times \tilde{P}} \quad (12.124)$$

and the target steering vector becomes

$$\tilde{\mathbf{s}}_{RD} = \mathbf{T}^H \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.125)$$

The corresponding optimal weighting follows from (12.43) as

$$\tilde{\mathbf{w}}_{RD/k} = \tilde{\mu} \left(\mathbf{T}^H \mathbf{R}_k \mathbf{T} \right)^{-1} \left(\mathbf{T}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right). \quad (12.126)$$

The adaptive weighting usually follows by forming the sample covariance matrix of (12.49) from the reduced-dimension data,

$$\hat{\mathbf{R}}_k = \frac{1}{\tilde{k}} \sum_{m=1}^{\tilde{k}} \mathbf{T}^H \mathbf{x}_m \mathbf{x}_m^H \mathbf{T}, \quad (12.127)$$

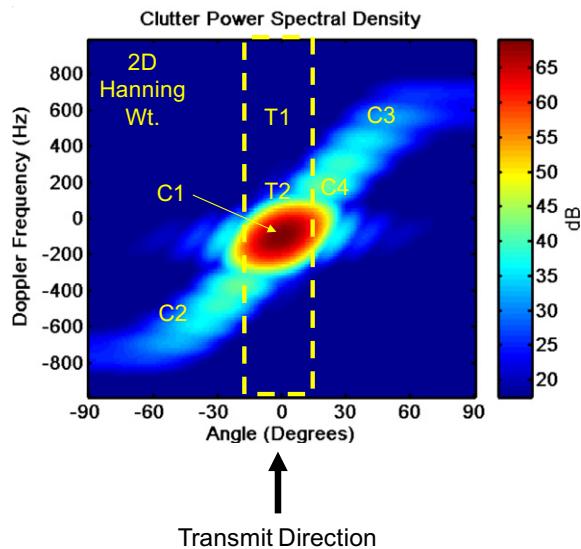
forming its inverse, and using $\tilde{\mathbf{v}}_{RD} = \mathbf{T}^H \mathbf{v}_{s-t}(\phi, \theta, f_d)$ in place of (12.125), viz.

$$\hat{\mathbf{w}}_{RD/k} = \hat{\mu} \hat{\mathbf{R}}_k^{-1} \tilde{\mathbf{v}}_{RD}. \quad (12.128)$$

In this case, $\hat{\mu}$ is a scalar that usually depends on estimated quantities.

Motivations for RD-STAP include the following:

- The clutter exhibits a highly structured angle-Doppler region of support, as shown in Figure 12.13. In this figure, T1 and T2 are target locations and C1–C4 are clutter regions. Target T1 predominantly competes with clutter located at C3, and a spatial null within the corresponding Doppler filter suffices to mitigate clutter. Clutter at other locations—such as at C1, C2, and C4—has no significant bearing on target detection. Alternately, target T2 is impacted by clutter in the region of C1 and C4; a two-dimensional null, positioned in angle and Doppler, is needed to suppress mainlobe and near-in sidelobe clutter.
- The RMB rule of Section 2.12.2.4.2 applies to RD-STAP, with $NM \rightarrow \tilde{P}$. It is not uncommon to have $9 \leq \tilde{P} \leq 50$, whereas $500 \leq NM \leq 5000$ is reasonable. Thus, limited and potentially heterogeneous or nonstationary training sample support is reduced by a factor of 10–100 or more.
- The implicit inverse of the sample covariance matrix is $O(N^3 M^3)$, being reduced to $O(\tilde{P}^3)$ in the RD-STAP case. For each halving of the processor’s DoFs, computational burden decreases by a factor of eight. Real-time operation requires RD-STAP methods.

**FIGURE 12.13**

Description of RD-STAP benefits.

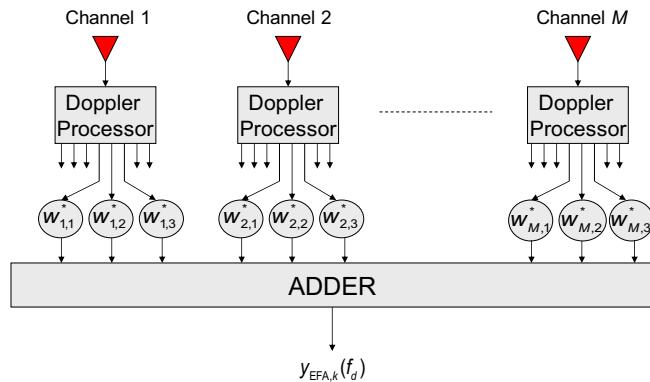
2.12.4.2.1 Post-Doppler STAP

Of the various RD-STAP methods, post-Doppler STAP techniques are most popular. Post-Doppler STAP is shown in the lower left and right of Figure 12.12, where the lower left box corresponds to the so-called “channel space,” post-Doppler methods that operate on space-Doppler data, and the lower right box characterizes the post-Doppler, (spatial) beamspace methods. The Extended Factored Algorithm (EFA) is an example of a “channel space,” post-Doppler method [36]. In EFA, each channel is Doppler filtered, and then the adaptive weighting of (12.128) is applied to the transformed data vector,

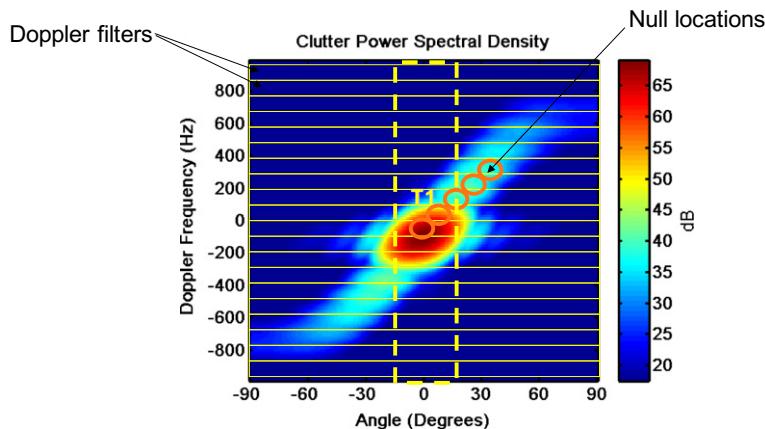
$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \stackrel{\text{e.g.}}{=} [\mathbf{f}_{s,k}^T(n-2) \quad \mathbf{f}_{s,k}^T(n-1) \quad \mathbf{f}_{s,k}^T(n) \quad \mathbf{f}_{s,k}^T(n+1) \quad \mathbf{f}_{s,k}^T(n+2)]^T, \quad (12.129)$$

where $\mathbf{f}_{s,k}(q) \in C^{M \times 1}$ is the spatial snapshot corresponding to the q th Doppler bin output (i.e., the q th Doppler bin output from channel 1 through M stacked in a vector). Equation (12.129) shows, as an example, the case of five adjacent Doppler bins. The EFA output for the k th range bin and n th Doppler bin is then given by $y_{\text{EFA}/k}(f_d) = \hat{\mathbf{w}}_{\text{RD}/k}^H \tilde{\mathbf{x}}_k$, where f_d corresponds to the n th Doppler bin output. Figure 12.14 depicts the EFA processing architecture, where $w_{m,q}$ is the EFA weight applied to the m th channel and q th Doppler bin. This figure shows the case of processing three adjacent Doppler bins; EFA typically employs three or five adjacent bins. The special case of one Doppler bin—where the processor only generates a spatial null—is known as Factored Time-Space (FTS) [10].

Figure 12.15 describes EFA operation in the angle-Doppler domain. Each horizontal, rectangular box corresponds to a Doppler filter extent. The circles correspond to null locations EFA is able to generate. EFA can generate up to $N_r M - 1$ nulls, where N_r is the number of temporal DoFs (the number of Doppler

**FIGURE 12.14**

Extended Factored Algorithm (EFA) implementation.

**FIGURE 12.15**

EFA operation in the angle-Doppler domain.

bins used in the adaptive combiner). The target location, T1, is shown close to mainlobe clutter. Nulling along the section of the clutter ridge in proximity to T1 is required for target detection. EFA is able to effectively suppress the clutter local to the target position. Its performance benchmarks very close to that of the space-time optimal solution, as will be shown in Section 2.12.5. FTS can generate up to $M - 1$ spatial nulls within the Doppler bin; as seen in this case, given the closeness of T1 to mainlobe clutter, only nulling sidelobe clutter is insufficient.

References [37,38] describe post-Doppler, (spatial) beamspace methods. The Joint Domain Localized (JDL) algorithm generalizes EFA by applying a beamspace transformation prior to Doppler filtering in Figure 12.14. JDL then combines M_s spatial beams and N_t Doppler bins to suppress clutter, in a

manner similar to a two-dimensional sidelobe canceler. For example, a common configuration is $M_s = 3$ and $N_t = 3$; in this three beam configuration, the processor uses the eight angle-Doppler beams surrounding the center beam to estimate and coherently remove the clutter signal in this ninth angle-Doppler direction of interest, thereby exploiting the structure of the clutter local to the target position. JDL performance tends to benchmark well relative to the space-time optimal solution. A special case of JDL is given in [38], where the Doppler filtering is applied to sum and difference beams, and then “auxiliary” beams about the sum beam and Doppler bin of interest are adapted to coherently remove the clutter signal. This latter method has been called Sigma-Delta STAP. Reference [39] also includes germane discussion.

Additional post-Doppler STAP methods based on PRI designs are given in [10].

2.12.4.2.2 Pre-Doppler STAP

Pre-Doppler STAP—sometimes called adaptive displaced phase center antenna (ADPCA)—involves processing overlapped sub-CPIs to mitigate clutter, then applying Doppler processing to the aggregated output [10,40]. It has particular application in those situations where the clutter response decorrelates during the course of the CPI, such as when the antenna array rotates. Figure 12.16 shows the pre-Doppler STAP, or ADPCA, architecture for the three pulse case.

The corresponding ADPCA data snapshot for three pulses is simply

$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \stackrel{\text{e.g.}}{=} [\mathbf{x}_{s,k}^T(n-1) \quad \mathbf{x}_{s,k}^T(n) \quad \mathbf{x}_{s,k}^T(n+1)]^T. \quad (12.130)$$

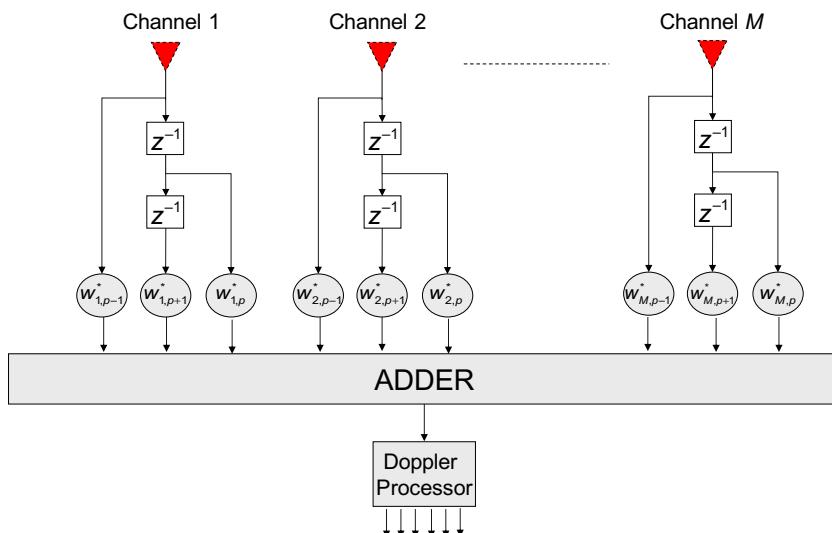


FIGURE 12.16

Pre-Doppler STAP.

The ADPCA weight vector is

$$\hat{\mathbf{w}}_{\text{adpca}/k} = \hat{\mathbf{R}}_{\text{adpca}/k}^{-1} \left(\left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \odot \begin{bmatrix} 1 \\ \exp(j2\pi f_1) \\ \exp(j2\pi 2f_1) \end{bmatrix} \right) \otimes \mathbf{v}_s(\phi, \theta) \right) \quad (12.131)$$

with f_1 the peak clutter Doppler times the PRI. The $3M$ -by- $3M$ ADPCA covariance matrix estimate, $\hat{\mathbf{R}}_{\text{adpca}/k}$, is an approximation to

$$\mathbf{R}_{\text{adpca}/k} = E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^H], \quad (12.132)$$

where $\tilde{\mathbf{x}}_k$ is given in (12.130). The covariance inverse in the ADPCA weight vector provides a dynamic response to whiten ground clutter returns. The steering vector term in parentheses suppresses mainlobe clutter with the binomial weights, identified as $[1 \ -2 \ 1]^T$, while forming a beam in a specified angular direction; the steering vector incorporates an additional linear phase variation over the temporal pulses to steer the Doppler null in cases where clutter is not centered at 0 Hz.

The performance of ADPCA is generally not as good as post-Doppler STAP techniques, thus limiting its use in practice. Performance benchmark results are given in Section 2.12.5.

2.12.4.3 Parametric adaptive matched filter

The parametric adaptive matched filtering (PAMF) employs multichannel, linear prediction to estimate the covariance matrix inverse [41]. We now discuss the PAMF based on [41, 42].

Consider the LDU decomposition of the covariance matrix inverse,

$$\mathbf{R}_k^{-1} = (\mathbf{A}_k^{-1})^H \mathbf{D}_k^{-1} \mathbf{A}_k^{-1}, \quad (12.133)$$

where $\mathbf{A}_k = \mathbf{L}$, $\mathbf{A}_k^H = \mathbf{U}$, and \mathbf{D}_k is diagonal. Since we are considering space-time data, a block LDU decomposition is required. In this case, \mathbf{A}_k^{-1} takes the form, for the example of $M = 4$ and $N = 8$,

$$\mathbf{A}_k^{-1} = \begin{bmatrix} \mathbf{I}_4 & \mathbf{0} \\ \mathbf{A}_1^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_3^H(3) & \mathbf{A}_3^H(2) & \mathbf{A}_3^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_4^H(4) & \mathbf{A}_4^H(3) & \mathbf{A}_4^H(2) & \mathbf{A}_4^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_5^H(5) & \mathbf{A}_5^H(4) & \mathbf{A}_5^H(3) & \mathbf{A}_5^H(2) & \mathbf{A}_5^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_6^H(6) & \mathbf{A}_6^H(5) & \mathbf{A}_6^H(4) & \mathbf{A}_6^H(3) & \mathbf{A}_6^H(2) & \mathbf{A}_6^H(1) & \mathbf{I}_4 & \mathbf{0} \\ \mathbf{A}_7^H(7) & \mathbf{A}_7^H(6) & \mathbf{A}_7^H(5) & \mathbf{A}_7^H(4) & \mathbf{A}_7^H(3) & \mathbf{A}_7^H(2) & \mathbf{A}_7^H(1) & \mathbf{I}_4 \end{bmatrix}, \quad (12.134)$$

where $\mathbf{A}_p^H(n)$ are the 4-by-4 multi-channel linear prediction (MCLP) coefficients for the p -order filter and the n th lag, $\mathbf{0}$ is the 4-by-4 zero matrix, and the realization index, k , is dropped for notational

convenience [42]. The matrix \mathbf{D}_k^{-1} is

$$\mathbf{D}_k^{-1} = \begin{bmatrix} \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_5 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_6 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_7 \end{bmatrix}, \quad (12.135)$$

where \mathbf{D}_p is the 4-by-4 inverse of the MCLP error for the p -order filter. As noted in [42], the MCLP development, while not necessary, is useful and provides a framework to calculate $\mathbf{A}_p^H(n)$ and \mathbf{D}_p .

The PAMF uses the LDU decomposition to approximate \mathbf{R}_k^{-1} . For example, let the maximum MCLP filter order $P = 2$, then

$$\mathbf{A}_k^{-1} \approx \begin{bmatrix} \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 \end{bmatrix}. \quad (12.136)$$

Extending the Wiener-Hopf equation to calculate the linear prediction coefficients, $\mathbf{A}_p^H(n)$, and augmenting with the identity matrices shown in (12.136), gives an approximation to \mathbf{A}_k^{-1} . The linear prediction error terms are given as

$$\mathbf{D}_k^{-1} \approx \begin{bmatrix} \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 \end{bmatrix}. \quad (12.137)$$

Generally, each of the diagonal blocks is inverted after solving for $\mathbf{D}_k = \mathbf{A}_k^{-1} \mathbf{R}_k (\mathbf{A}_k^{-1})^H$. The approximation to the covariance matrix inverse follows from (12.133).

2.12.4.4 Summary of adaptive filter implementations

Table 12.3 summarizes the adaptive filter implementations discussed in this section. Performance assessment of a number of these methods is given in Section 2.12.5.2.

Table 12.3 Summary of Adaptive Filter Implementations

Type	Method	Section	Comments
RR-STAP	Adaptive beam pattern synthesis (principal components)	2.12.4.1.1	Weight calculation approach, interprets STAP as pattern synthesis, where notches are incorporated into the quiescent beam pattern based on information in the dominant subspace. Insightful view of STAP, not as practical as other methods, especially due to high computational burden associated with the eigendecomposition
RR-STAP	Principal Components Inverse (PCI)	2.12.4.1.2	Weight calculation approach, involves forming a projection operation orthogonal to the clutter and interference subspaces. May be used with RD-STAP methods. Computationally burdensome due to eigendecomposition
RR-STAP	Eigencanceler	2.12.4.1.3	Two formulations, one leading to Principal Components Inverse weight vector solution
RR-STAP	Hung-Turner Projection (HTP)	2.12.4.1.4	Uses Gram-Schmidt to approximate PCI solution at much lower computational cost. Developed for use with large adaptive DoF systems. Best performance against stronger interference sources
Other	Diagonal loading	2.12.4.1.5	Adds scaled diagonal matrix to covariance matrix estimate, enhancing numerical stability and mitigating spurious sidelobe effects due to perturbed noise subspace
RR-STAP	Cross Spectral Metric (CSM)	2.12.4.1.6	Signal-dependent approach to select interference subspaces in weight calculation most greatly enhancing output SINR. Generally limited to the GSLC architecture, which tends to greatly increase computational burden and limits practical application
Other	Multi-Stage Wiener Filter (MWF)	2.12.4.1.7	Truncated decomposition of the GSLC architecture, involves calculating a series of scalar weights for each stage, with each subsequent stage maximally correlated with the “desired signal” (residual interference) of the prior stage. Very high computational burden relative to competing methods

Continued

Table 12.3 *Continued*

Type	Method	Section	Comments
RD-STAP	Extended Factored Algorithm (EFA)	2.12.4.2.1	Post-Doppler, “element” space STAP. Typically involves adaptive weighting three to five adjacent Doppler bins and all spatial channels to suppress clutter. Efficient, effective technique: reasonable computational burden and low training sample support requirements
RD-STAP	Factored Time-Space (FTS)	2.12.4.2.1	Special case of EFA when using only one Doppler bin. Places a spatial null within a given Doppler bin. Generally performs poorly at low range rates, since it only has spatially adaptive DoFs
RD-STAP	Joint-Domain Localized (JDL)	2.12.4.2.1	Post-Doppler, beamspace STAP. Essentially a two-dimensional, linear prediction technique: employ angle-Doppler beams surrounding the target beam to estimate and coherently remove competing interference. Low computational burden and minimal training sample support requirements
RD-STAP	Adaptive Displaced Center Antenna (ADPCA)	2.12.4.2.2	Pre-Doppler STAP method, involves adaptively combining all spatial channels and a small collection of pulses from within the current CPI. Adaptive canceler is applied to overlapped segments of pulses from the CPI. Collection of clutter canceled outputs then fed through a Doppler filter bank to achieve temporal integration gain. Useful for situations where interference may appear slightly nonstationary over the CPI, such as in the case with a rotating array. Computational burden and training sample requirements similar to EFA
Other	Parametric Adaptive Matched Filter (PAMF)	2.12.4.3	Low order, multi-channel linear prediction approach to estimate the covariance matrix inverse. Computationally efficient. Model order selection is generally ad hoc

Table 12.4 Radar Simulation Parameters

Parameter	Value	Parameter	Value
Transmit frequency	L-Band (1240 MHz)	Aperture size	~1.2 m width × 1 m height
Receiver bandwidth	0.8 MHz	Receive aperture configuration	11 channels configured as a uniform linear array
Peak power	15 kW	Array attitude	Sidelooking, canted 5° down, 6.78° yaw, 0° roll
PRF	1984 Hz	Platform altitude	3000 m
CPI length	32 pulses	Platform velocity	100 m/s
Clutter model	–20 dB reflectivity, Constant Gamma Model	ICM	Gaussian, 0.25 m ² /s ²

2.12.5 Application

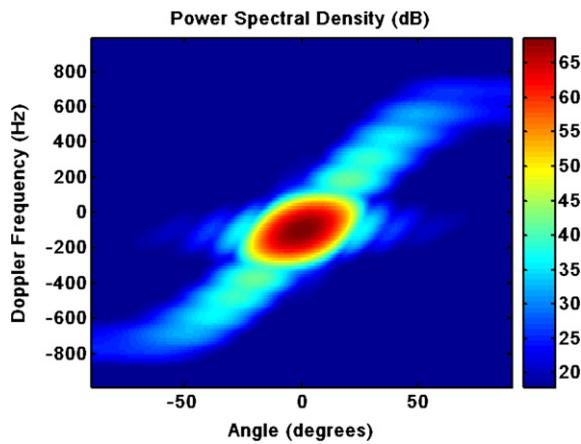
In this section we consider the application of STAP to the important problem of radar detection in clutter. We employ the simulation models of Section 2.12.3 and some of the metrics from Section 2.12.2 to characterize the interference, algorithm performance, and STAP improvement over the non-adaptive processor.

Table 12.4 provides some of the parameters of the simulated, multi-channel radar system. Chosen parameters are similar to those of the Multi-Channel Airborne Radar Measurements (MCARM) system described in Section 2.12.8. Also, these are the same parameters used to generate Figures 12.1–12.3 in Section 2.12.1.

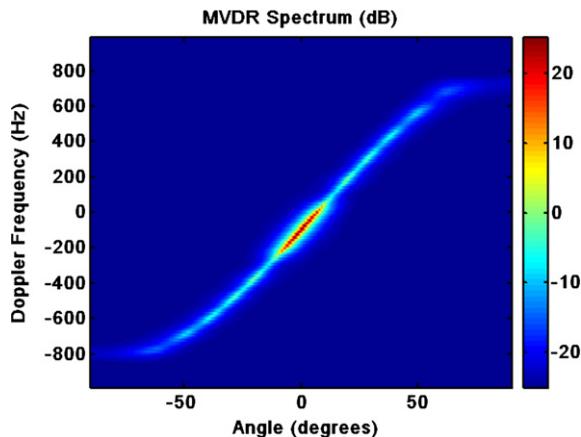
2.12.5.1 Interference characteristics

Using the models described in Section 2.12.3, we simulate clutter and noise for an eleven channel, thirty-two pulse airborne radar scenario. The clutter environment is homogeneous. Additional parameters are given in Table 12.4 and are similar in nature to a MCARM data collection referenced in Section 2.12.8. In this case, we simulated one CPI, where the aircraft experiences 6.78° of yaw. We focus our attention somewhat arbitrarily at a 32 km slant range for the analysis. Additionally, we simulated the eleven channel uniform linear array without array amplitude and phase errors. The antenna is steered to broad side: 0° azimuth and 0° elevation angle.

Figure 12.17 shows the PSD for this scenario. The PSD uses a Hanning weight in both space and time. The slight offset in peak clutter Doppler frequency in the look-direction from 0 Hz is due to platform yaw. The overall CNR is in the range of slightly greater than 50 dB. For comparison, Figure 12.18 shows the MVDR spectrum, clearly outlining the clutter angle-Doppler region of support. (Note: while MVDR provides units of power, it is widely known to be a poor estimator of the signal strength.) Figure 12.19 shows the eigenspectrum calculated using the known covariance matrix; the eigenspectrum displays the

**FIGURE 12.17**

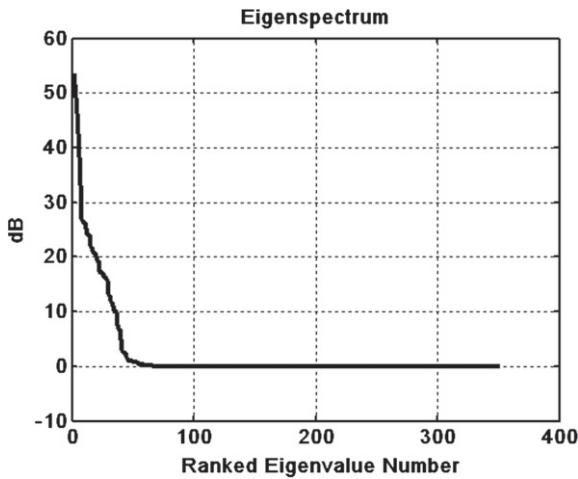
Power spectral density for sidelooking array radar example.

**FIGURE 12.18**

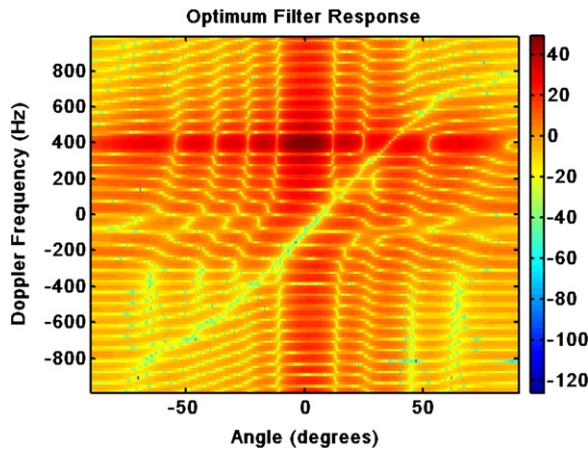
MVDR spectrum for sidelooking array radar example.

covariance matrix eigenvalues, sorted from largest to smallest. The noise eigenvalues are set to 0 dB, and those values greater than the noise floor correspond to the clutter signal. The largest eigenvalue is indicative of the CNR to within a few decibels, suggesting CNR in the range of 53–55 dB.

Considering the MVDR spectrum in Figure 12.18, we anticipate an optimal filter response that drives a deep null along the clutter angle-Doppler region of support. Figure 12.20 shows the optimal space-time filter response, steered to broadside and 400 Hz Doppler; the clutter null and peak gain is evident in the figure.

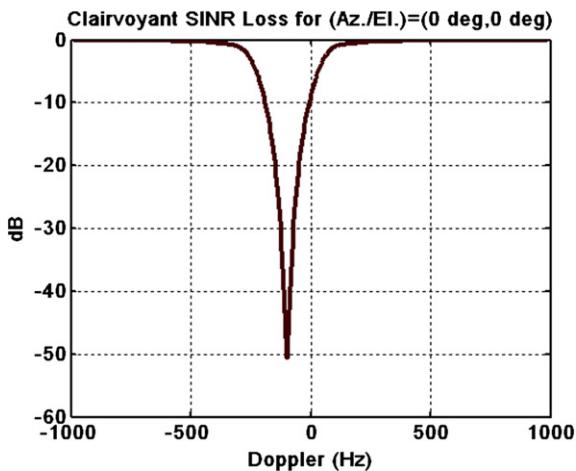
**FIGURE 12.19**

Eigenspectrum for example scenario.

**FIGURE 12.20**

Optimal space-time frequency response, steered to array broadside and 400 Hz Doppler frequency.

Finally, Figure 12.21 shows the clairvoyant (known covariance) SINR loss, $L_{s,1}$, over unambiguous Doppler ($-\text{PRF}/2 \leq f_d \leq \text{PRF}/2$) in the broadside direction. The loss cut shows the impact of clutter on detection performance, with the null offset from 0 Hz due to the platform yaw. As seen from the figure, the null depth is slightly greater than 50 dB, consistent with the mainlobe CNR and the eigenspectrum of Figure 12.19. We will consider the performance of various STAP techniques relative to the clairvoyant SINR loss curve momentarily.

**FIGURE 12.21**

Clairvoyant SINR loss.

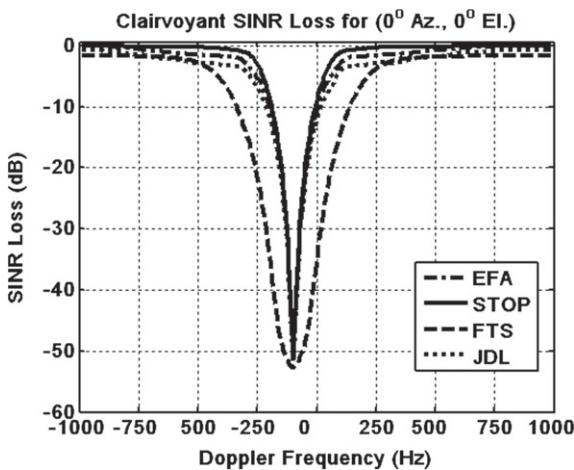
2.12.5.2 STAP algorithm performance

Based on the discussion in Section 2.12.4, we benchmark the performance of a number of STAP techniques using the scenario from the prior section. Again, we focus on a slant range of 32 km. Also, we consider clairvoyant loss, $L_{s,1}$. For the homogeneous clutter scenario considered herein, the adaptive loss, $L_{s,2}$, precisely follows the RMB rule: using sample support equal to twice the processor's DoF yields, on average, 3 dB loss. The different techniques have different sample support requirements, based on their adaptive DoFs. An exception in this case involves the PAMF; the PAMF implementation averages over the pulse domain, and so can rely on less training data over range. However, this advantage is not unique to the PAMF, and other methods can also average over pulses (like the ADPCA technique) or employ smoothing techniques to effectively increase the sample support [43].

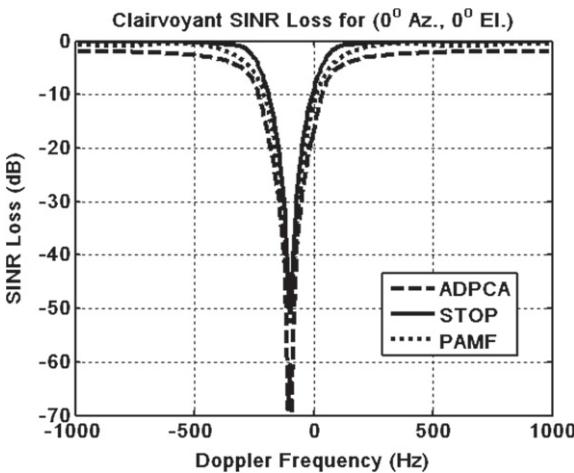
Figure 12.22 compares the clairvoyant SINR loss of EFA, FTS, and JDL to the bound provided by the space-time optimal processor (STOP; the optimal space-time filter using the clairvoyant covariance matrix). In this case, the processors are configured as follows:

- *FTS*—uses all eleven channels and a Hanning weighting on the Doppler filters. Eleven adaptive spatial DoFs.
- *EFA*—uses all eleven spatial channels and three adjacent, Hanning-weighted Doppler filters. Thirty-three adaptive space-Doppler DoFs.
- *JDL*—uses three adjacent, uniform weighted spatial beams and three adjacent, Hanning-weighted Doppler filters. Nine total adaptive angle-Doppler DoFs. The beam spacing is three degrees (a little less than one-third of the full aperture 3 dB beamwidth).

From Figure 12.22, we find that EFA and JDL provide excellent performance potential relative to the achievable bound set by STOP. The FTS performance is disappointing, but not a surprise since this algorithm does not adaptively combine any temporal DoFs. One then might conclude JDL to be a better selection, since it only requires nine DoFs. However, the additional spatial DoFs afforded by EFA may

**FIGURE 12.22**

Clairvoyant SINR loss for post-Doppler STAP techniques.

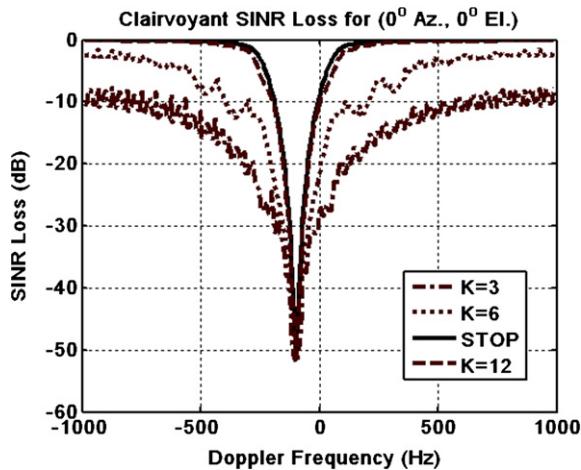
**FIGURE 12.23**

Clairvoyant SINR loss for pre-Doppler STAP and PAMF.

prove beneficial when RFI or Doppler ambiguities are present. Thus, a number of analyses should take place before final algorithm selection.

Figure 12.23 duplicates the aforementioned analysis for pre-Doppler STAP (ADPCA) and the PAMF. The processors are configured as follows:

- *ADPCA*—uses three adjacent pulses and all spatial channels for cancellation, then Doppler filters the resulting outputs using a Hanning weight. Thirty-three adaptive space-time DoFs. The

**FIGURE 12.24**

Clairvoyant SINR loss for the MWF.

implementation steers the temporal gain response to center the temporal steering vector null at -100 Hz, the center of mainlobe clutter.

- PAMF—uses a fourth-order filter model. No weighting is used on the Doppler steering vector.

The clairvoyant SINR loss curves in Figure 12.23 indicate that both ADPCA and PAMF provide very good performance, rivaling that of EFA and JDL in Figure 12.22.

As a last example, Figure 12.24 shows the benchmark performance of the MWF using three, six, and twelve stages ($K = 3, 6, 12$). As seen from the figure, the performance for $K = 3$ or $K = 6$ is poor, but very good capability is observed when $K = 12$. The computational loading of MWF, for the implementation used by the author, is significantly higher than the other methods examined.

2.12.5.3 STAP Comparison with nonadaptive solution

To conclude this section, we simply compare the performance of the space-time optimal filter to a nonadaptive processing scheme involving beamforming and Doppler processing. The nonadaptive filter implementation uses a Hanning temporal weight and uniform illumination spatially on receive. (Note: the transmit illumination uses a 30 dB Taylor weighting.) As seen from Figure 12.25, the space-time optimal filter provides a very significant performance advantage over the nonadaptive processor, thereby indicating the tremendous benefits of STAP.

2.12.5.4 Application summary

In this section we considered a typical airborne radar example. Clutter and noise models from Section 2.12.3 were used to simulate data snapshots and covariance matrices. The covariance matrices were used to assess the performance potential of several STAP techniques discussed in Section 2.12.4. From this analysis, we find that a number of the methods perform similarly well relative to the achievable

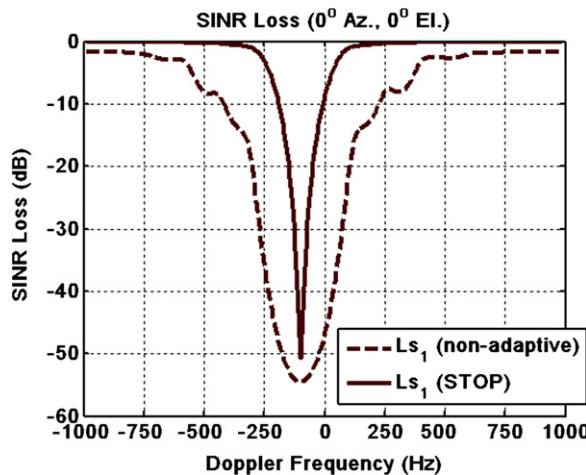


FIGURE 12.25

Comparison of the optimal and nonadaptive filter performance potential.

performance bound. We caution the reader, however, that many other practical matters drive algorithm selection. Of the methods discussed, EFA often provides the best performance and greatest flexibility when considering a number of issues.

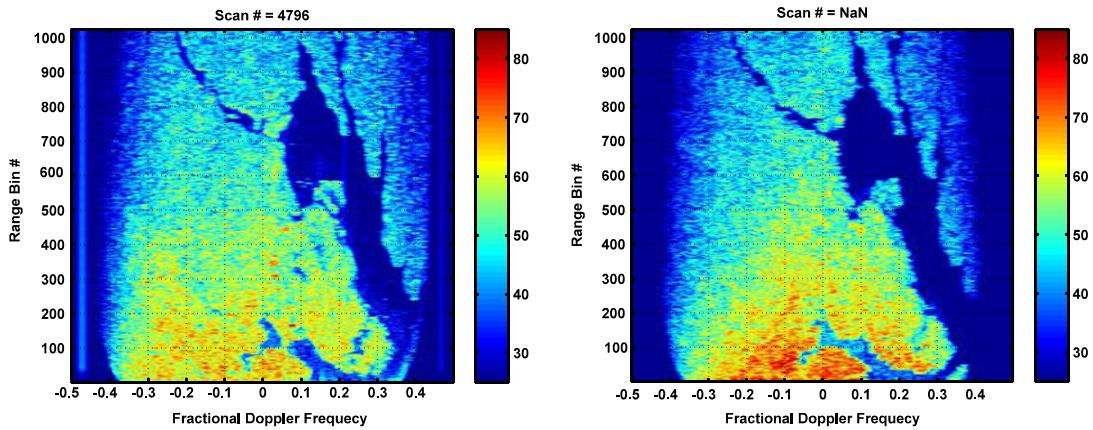
2.12.6 Challenges

Contemporary STAP research topics generally focus on challenges associated with covariance matrix estimation or mitigating computational burden. In this section we provide commentary on the former.

2.12.6.1 Heterogeneous clutter

Real world clutter environments are heterogeneous, thus tacitly undermining the IID assumption central to the development of (12.49) [44–46]. Heterogeneous clutter is a result of culturally-varying terrain type. Examples include: variation in clutter amplitude or spectral spread due to a mixture of clutter types; abrupt edges characteristic of clutter interfaces (e.g., between urban and rural regions); the presence of target-like signals in the training data (TSD) [45]; and, stationary, manmade objects resulting in strong, discrete responses. Each of the aforementioned effects is localized within the training set. Hence, the estimate of (12.49) captures the average behavior of the training data, thus potentially appearing mismatched to any particular cell under test. This mismatch translates to an erroneous, adaptive response relative to the optimal condition. SINR loss or threshold bias results in such instances. References [44–46] characterize in detail the nature of such system degradation.

As an example, distributed clutter is typically site-specific: the clutter reflectivity varies as a function of range and angle due to changes in the terrain features. Equation (12.70) accommodates site-specific

**FIGURE 12.26**

Comparison of measured (left) and simulated (right) multichannel UHF radar data (colorbar in decibels) (after [47], © 2006 IEEE).

simulation by modifying the term, $v_k(p, q)$, through access of a database to determine the clutter RCS, $\sigma_{k,p,q}$, in (12.67)–(12.69) (where the reference to the q th range ambiguity is added). Figure 12.26 compares actual and site-specific simulation of UHF radar data taken in the Delaware–Maryland–Virginia (Delmarva) Peninsula region; this region is dominated by rural clutter, bodies of water and rivers, and also covered by a number of roadways. As seen from Figure 12.26, the range and cross-range (Doppler, or angle) variation of the clutter response is evident and predictable. The clutter variation shown is one source of heterogeneity leading to covariance matrix estimation error.

Table 12.5 provides a summary of various sources of heterogeneous clutter and their impact on STAP performance. Of the effects listed, it is observed that target-like signals corrupting the secondary (training) data (TSD) and clutter discretes tend to lead to the greatest performance loss.

A variety of techniques have been developed to enhance STAP detection performance given the challenges of heterogeneous clutter. We summarize some of the available methods in Table 12.6. As discussed in [47], the different approaches fall in one of two categories: indirect or direct. The indirect methods attempt to manipulate the training set to improve the covariance matrix estimate, oftentimes using knowledge of the platform motion or surrounding terrain or operating environment, whereas the direct methods attempt to modify the filter response using models and ancillary information. Both indirect and direct methods attempt to improve the instantaneous adaptive filter response to minimize heterogeneous clutter residue at the filter output; in this manner, STAP performance in complex, heterogeneous environments can approach that attainable in a homogeneous setting.

As an example of the degradation of heterogeneous interference effects, and the improvement potential of available STAP techniques, Figure 12.27 compares detection rates using MCARM Flight 5, Acquisition 575 data using block training (bins 200–320) versus intelligent training and filter selection (ITFS, [47]), where maps facilitate training data excision in regions overlaying certain roadways. As seen from the figure, TSD leads to excessive, additional loss in the lobed regions near ± 20 m/s;

Table 12.5 Summary of Heterogeneous Clutter and STAP Impact

Source	Description	Impact on STAP
Spatially-varying, distributed clutter RCS	Range-angle variation in clutter RCS leads to varying power across clutter angle-Doppler region of support	Over- or undernull clutter. Overnull clutter can lead to signal cancellation, whereas undernulling leads to clutter residue and degraded SINR
Spatially-varying, distributed clutter spectral spread	Range-angle variation of intrinsic clutter motion varies spectral spread across clutter angle-Doppler region of support	Training on regions where ICM is less than that of the application region leads to insufficient null width and an increase in clutter residual, whereas regions with ICM less than that present in the training set experience suppression of low speed targets
Clutter edges and shadowing	Gross variation in clutter type (e.g., land-sea interface) or regions of extended obscuration, predominantly in the mainlobe direction	Over- or undernull clutter
Clutter discretes	Stationary objects with relatively large RCS, often manmade objects such as cars, utility poles, etc., also includes extended discretes, such as fence lines and train tracks	Increased false alarm rate, upward threshold bias leading to increase in missed detections
Target-like signals	Vehicles on roadways and within airspace, predominantly through the mainbeam	Signal cancellation due to nulling off the clutter ridge in angle-Doppler locations consistent with targets of interest

removing a specific highway using ITFS—as noted in the figure’s legend—fully mitigates this loss, raising the detection rate from roughly 55% to 98%.

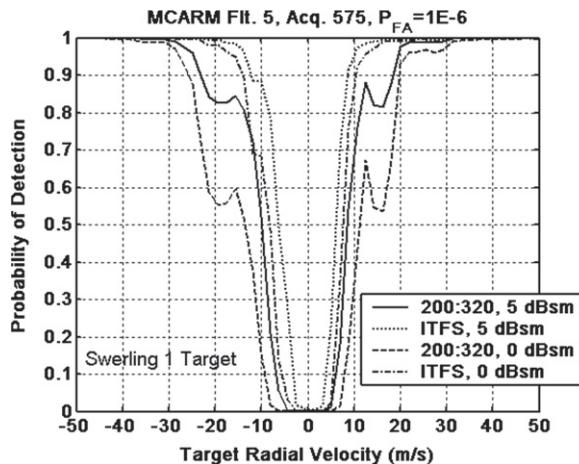
2.12.6.2 Nonstationary clutter

STAP maximizes SINR by filtering ground clutter in the angle-Doppler domain. Radar geometry determines the filter null location in this higher-dimensional space. When the null location varies over range, the adaptive filter produces an incorrect frequency response. In the monostatic radar case, the angle-Doppler region of support can exhibit range variation when the velocity vector and array normal are non-orthogonal. This occurs for forward-looking arrays, or under conditions when the platform is yawed [48, 49].

In sidelooking array radar, the spatial frequency measured by a uniform linear array is proportional to the corresponding Doppler frequency of a stationary clutter patch on the Earth’s surface. Specifically,

Table 12.6 STAP Techniques to Mitigate the Impact of Heterogeneous Clutter on Detection Performance

Technique	Description	Reference
Nonhomogeneity Detector (NHD)	Data-dependent screening measure, tests training data for similarity to the average, selects homogeneous training samples	[50–52]
Power Selected Training (PST)	Data-dependent screening measure, selects training samples strongest in power to drive null as deeply as possible	[53, 54]
Power Comparable Training (PCT)	Data-dependent screening measure, sorts data into tiles of similar power levels to match training data to cell under test	[55]
Power Variable Training (PVT)	Scale the power of the estimated covariance matrix's dominant (clutter) subspace to match the power level in the cell under test, requires eigendecomposition	[56]
Map-aided training	Uses mapping data and distance measures to screen training samples	[24, 57]
Covariance matrix taper (CMT)	Purposely spreads the estimated clutter response to increase null width	[58]
Subaperture smoothing	Exploits similar covariance structure among sub-apertures of a uniformly sampled spatial or temporal aperture to enhance the quantity of homogeneous training samples	[43, 59]
Nonlinear, nonadaptive STAP	Employs a model of the clutter covariance matrix to suppress clutter	[60]
Color loading	Adds a scaled clutter covariance model to the covariance matrix estimate to emphasize the anticipated clutter response and enable localized training	[61, 62]
Discrete Matched Filter (DMF)	Coherently removes clutter discrete signals using a modified version of the CLEAN algorithm	[47]
Signal and Clutter as Highly Independent Structured Modes (SCHISM)	Coherently removes distributed clutter returns using a modified version of the CLEAN algorithm	[63]
Adaptive Coherence Estimator (ACE)	Incorporates a statistical measure of the “whiteness” of the clutter residue in a particular cell under test to suppress clutter discretes?	[64]
Knowledge-Aided Parametric Covariance Estimation (KAPE)	Estimates parameters of a validated covariance matrix model, including clutter amplitude, spread, mainlobe centroid, and error components of the array manifold	[16, 65]

**FIGURE 12.27**

Estimated detection probability improvement over Doppler frequency using KA training on MCARM data (after [24], © 2004 IEEE).

normalized Doppler, $\tilde{f}_d = f_d T$, can be written

$$\tilde{f}_d = \frac{2v_p T}{\lambda} \cos \phi_{cone} = \left(\frac{v_p T}{d/2} \right) f_s, \quad (12.138)$$

where v_p is the platform velocity in the direction orthogonal to the array normal, T is PRI, λ is wavelength, ϕ_{cone} is the cone angle measurement from the platform centerline to the clutter patch location, d is channel spacing, and f_s is spatial frequency. Spatial frequency is given by

$$f_s = \frac{d}{\lambda} \cos \phi_{cone}. \quad (12.139)$$

Physically, the clutter iso-Doppler contours and the array beam traces align, translating to angle-Doppler behavior that is range invariant.

In the non-sidelooking radar case, the beam traces and iso-Dopplers of the ground clutter returns misalign at some ranges, mainly those where the range divided by the altitude is less than five [18, 19]. The variation of the clutter response through the training set leads to an adaptive filter response with incorrect null placement, leading to increased clutter residue and degraded detection performance. References [48, 49] discuss effective compensation methods.

Range varying angle-Doppler loci further exacerbates culturally induced heterogeneous clutter effects. For example, residue from clutter discretes further increase in the presence of adaptive filter null migration associated with the non-stationary clutter mechanism described in this section.

2.12.6.3 Bistatic STAP

Bistatic STAP development received considerable attention in the early-to-mid 2000s. Airborne bistatic radar involves a moving transmitter and receiver separated a considerable distance. The bistatic geometry

Table 12.7 Summary of Bistatic STAP Methods

Bistatic STAP approach	Description	Reference
Localized processing	Simple strategy, attempts to choose training data in the vicinity of the cell under test to minimize nonstationary impacts	[66, 67]
Time-varying weights	Employs truncated Taylor series expansion of the weight vector, presumes linear evolution in the weight vector elements over range	[68, 69]
Doppler warping	Aligns a point of the clutter ridge for each bistatic range to a designated reference point (e.g., 0 Hz Doppler) using a range-varying, complex modulation prior to STAP application	[49]
Higher-order Doppler warping	Aligns sections of the clutter ridge to a reference ridge using a range varying modulation, a multi-point extension of the Doppler warping method prior to STAP application	[70]
Angle-Doppler compensation	Aligns “spectral centers,” or regions of maximum angle-Doppler return, over range using a complex, range-varying, space-time modulation prior to STAP application	[71]
Adaptive angle-Doppler compensation	Derives key information on clutter range variation directly from the data, applies a complex, space-time, range-varying modulation to the data to align dominant clutter subspaces prior to STAP application	[72, 73]
Registration of “direction-Doppler curves”	Uses curve fitting methods to warp the power spectral density of a given range sum to a reference prior to STAP application	[74]

and independent motion between transmitter and receiver leads to significant angle-Doppler variation over range [75, 72].

Table 12.7 summarizes a number of recently developed bistatic STAP techniques, including a description of the approach and germane references. As seen from the table, the methods generally fall in one of three categories: localized processing (or training); data pre-warping, prior to STAP application; and, time-varying weights. Several of the bistatic STAP methods in Table 12.7 have proven extraordinarily effective, restoring performance to levels similar to that achievable in homogeneous clutter environments.

2.12.6.4 Conformal array STAP

A conformal array’s shape is compliant with the contours of the radar bearing platform. For example, Figure 12.28 depicts conformal antenna elements mounted to a chined nose cone; the short lines emanating from each dot representing an antenna element indicate the surface normal.

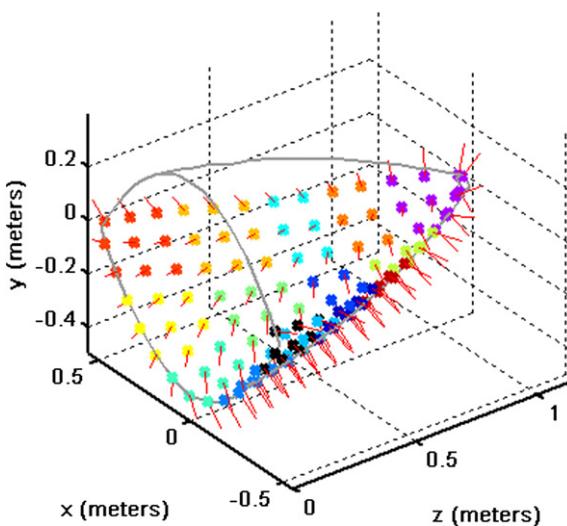


FIGURE 12.28

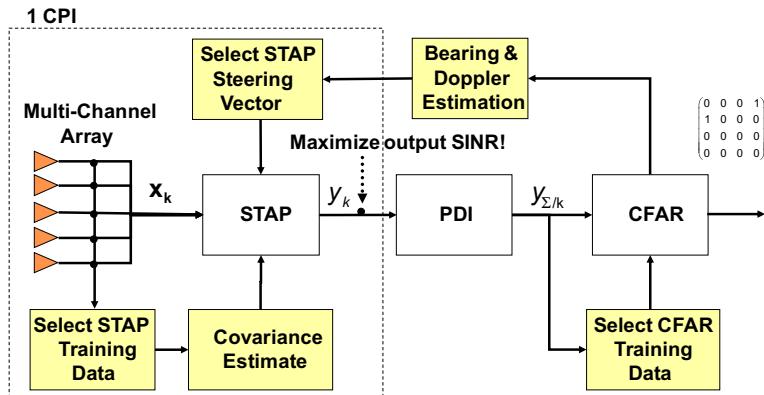
Example of conformal antenna elements mounted to a chined nose cone.

The curvature of the conformal array, coupled with the look-direction relative to the velocity vector, results in a nonstationary clutter angle-Doppler response [76–78]. SINR loss resulting from the nonstationary clutter response rivals that seen in bistatic radar examples. Hersey et al. [76] discusses several solutions to mitigate nonstationary clutter behavior based on adaptations of solutions developed for bistatic STAP implementation, including localized processing, localized processing with time-varying weights, angle-Doppler warping, and higher-order angle-Doppler warping. Additionally, Hersey et al. [76] develops a method called equivalent uniform linear array transformation which attempts to resample the data to a linear configuration. The conformal array STAP methods in [76] lead to significant performance enhancement.

2.12.7 Implementation

We discuss a generic, STAP-based, airborne radar detection architecture in this section. A flow diagram of the detection process is given in Figure 12.29.

STAP is applied at the CPI level. As a result of target fading, the MTI radar typically transmits bursts of pulses at several different frequencies, where each burst comprises the CPI. Being a coherent signal processing technique, STAP operates on each CPI as indicated in the figure, essentially generating a range-Doppler map (RDM) at a receive angle consistent with the transmit direction: a space-time snapshot for each range bin, \mathbf{x}_k , is processed using a space-time weighting steered to a given angle and Doppler to yield a single pixel in the RDM, $|y_k|^2$. Each CPI is processed, generating aligned RDMs; the alignment may require additional processing steps, such as motion compensation. Then, the RDMs

**FIGURE 12.29**

Detection processing flow.

at the different frequencies are noncoherently summed in a step called post-detection integration (PDI; also called noncoherent addition, NCA). PDI boosts detection performance by taking advantage of the purposeful decorrelation created by frequency hopping to avoid target fading that adversely affects operation using a single coherent dwell. Frequency hopping forces the target to take on a Swerling 2 characteristic [1, 7].

After PDI, a detection threshold is set using a constant false alarm rate (CFAR) algorithm; the CFAR algorithm multiplies an estimate of the residual interference power by a threshold multiplier in an attempt at achieving a constant false alarm rate [1]. Pixel values crossing the detection threshold are declared targets. The signal processor then estimates the bearing and Doppler frequency of the threshold crossing generally using a maximum likelihood estimator. The corresponding SINR, bearing, Doppler, and range—among other potential characteristics—for each detection is then passed to an analyst or an automatic tracker. Figure 12.29 is meant to show the basic processing steps; it is common to incorporate additional functionality to cope with clutter heterogeneity or target motion through resolution cells, for instance.

The selection of the scalar, $\hat{\mu}$, in (12.44) is an important, practical consideration. It is common to use the adaptive matched filter (AMF) normalization [79],

$$\hat{\mu} = \frac{1}{\sqrt{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)}}. \quad (12.140)$$

The AMF normalization sets the residual interference-plus-noise floor to unity as $\hat{\mathbf{R}}_k \rightarrow \mathbf{R}_k$, viz.

$$P_{I+N} = \hat{\mathbf{w}}_k^H \mathbf{R}_k \hat{\mathbf{w}}_k = \frac{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{R}_k \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)}{\left(\sqrt{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)} \right)^2} \rightarrow 1. \quad (12.141)$$

The AMF normalization accommodates variable training intervals and other segmentation of the weight estimation and application over range, minimizing shifts in the mean, residual interference-plus-noise

power that otherwise leads to threshold bias and increased false alarm rate. More specifically, assuming the segmented interference-plus-noise environment is IID, the AMF normalization possesses CFAR properties: as power fluctuates from one weight application region to the next, the filter output is scaled by the inverse of the residual interference-plus-noise power so that a fixed detection threshold maintains a constant false alarm rate. This is seen by examining the AMF decision statistic,

$$\eta_{\text{AMF}} = \frac{\left| \mathbf{v}_{\text{s-t}}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}} \right|^2}{\mathbf{v}_{\text{s-t}}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_{\mathbf{k}}^{-1} \mathbf{v}_{\text{s-t}}(\phi, \theta, f_d)} \stackrel{H_1}{>} \stackrel{H_0}{<} T_1, \quad (12.142)$$

where T_1 is a fixed decision threshold. Moving the denominator of the decision statistic to the right, (12.142) is interpreted as the square of the filter output (the numerator term) compared to a threshold multiplier, T_1 , scaled by the residual interference-plus-noise output power. This is the conventional view of CFAR in radar signal processing.

A virtually identical formulation of the AMF result is given in [80] and referred to as Modified Sample Matrix Inversion (MSMI).

A maximum likelihood estimate for target bearing and Doppler integrates seamlessly with the architecture in Figure 12.29 and use of the AMF normalization. This can be seen by expressing the alternative hypothesis of (12.2) as

$$\mathbf{x}_{\mathbf{k}} = \alpha \mathbf{s}_{\text{s-t}}(\mathbf{p}) + \mathbf{n}_{\text{tot}}; \mathbf{p} = \begin{bmatrix} \phi \\ \theta \\ f_d \end{bmatrix}, \quad (12.143)$$

where α is a complex constant, and $\mathbf{n}_{\text{tot}} = \mathbf{c}_{\mathbf{k}} + \mathbf{j}_{\mathbf{k}} + \mathbf{n}_{\mathbf{k}}$ is the interference-plus-noise (total noise) vector. The objective is to estimate \mathbf{p} with $\hat{\mathbf{p}}$. Since $\mathbf{n}_{\text{tot}} \sim \text{CN}(\mathbf{0}, \mathbf{R}_{\mathbf{k}})$, the joint probability density function (pdf) is

$$p(\mathbf{x}_{\mathbf{k}} | \mathbf{p}) = \frac{1}{\pi^M |\mathbf{R}_{\mathbf{k}}|} \exp \left(-(\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\text{s-t}}(\mathbf{p}))^H \mathbf{R}_{\mathbf{k}}^{-1} (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\text{s-t}}(\mathbf{p})) \right). \quad (12.144)$$

Given the likelihood function defined by (12.144), we first require an estimate for the complex constant. Equation (12.144) is maximal when

$$Q(\alpha, \mathbf{p}) = (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\text{s-t}}(\mathbf{p}))^H \mathbf{R}_{\mathbf{k}}^{-1} (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\text{s-t}}(\mathbf{p})) \quad (12.145)$$

is minimal. Differentiating (12.145) and setting the result to zero yields

$$\hat{\alpha} = \frac{\mathbf{s}_{\text{s-t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}}}{\mathbf{s}_{\text{s-t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\text{s-t}}(\mathbf{p})}. \quad (12.146)$$

Substituting (12.146) into (12.145) leads to

$$Q(\hat{\alpha}, \mathbf{p}) = \mathbf{x}_{\mathbf{k}}^H \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}} - \frac{\left| \mathbf{x}_{\mathbf{k}}^H \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\text{s-t}}(\mathbf{p}) \right|^2}{\mathbf{s}_{\text{s-t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\text{s-t}}(\mathbf{p})}. \quad (12.147)$$

Differentiating (12.147) with respect to \mathbf{p} and setting the result to zero yields the maximum likelihood estimate (MLE), $\hat{\mathbf{p}}$. Consequently, the estimator is [81]

$$\hat{\mathbf{p}} = \begin{bmatrix} \hat{\phi} \\ \hat{\theta} \\ \hat{f}_d \end{bmatrix} = \arg \max_{\mathbf{p}=[\phi \ \theta \ f_d]^T} \left(\frac{|\mathbf{s}_{s-t}^H(\mathbf{p}) \mathbf{R}_k^{-1} \mathbf{x}_k|^2}{\mathbf{s}_{s-t}^H(\mathbf{p}) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\mathbf{p})} \right). \quad (12.148)$$

The MLE takes the form of an AMF normalized STAP cost surface and in practice \mathbf{v}_{s-t} replaces \mathbf{s}_{s-t} . Implementing the estimator requires a very fine grid search to find the peak of the likelihood function (via the MLE approach). The grid search—which essentially amounts to stepping the space-time steering vector over potential target angles and Dopplers—is computationally burdensome and can be sub-optimum. Specifically, if the step size is too large, the estimator can exhibit bias and increased variance. For this reason, numerical approximations to the MLE, which exhibit good practical performance, are commonly used.

When processing multiple CPIs, the MLE appears as the sum of the individual cost surfaces at each frequency, where each cost surface takes the form of the term in parentheses in (12.148). An equivalent development for the multistatic radar case is given in [82].

STAP remains computationally challenging to implement. The weight calculation generally takes place in either the voltage domain using QR decomposition (QRD) or in the power domain using Cholesky factorization. Calculating the weight vector using QRD in the voltage domain seems favorable, since it essentially doubles the dynamic range in decibels over the power domain solution and avoids explicitly forming outer products in the covariance estimation step. Cholesky factorization in the power domain is also a common approach when considering all the details of real-time implementation on specific computing devices.

2.12.8 STAP data collection programs

In this section we highlight some measured data collection programs.

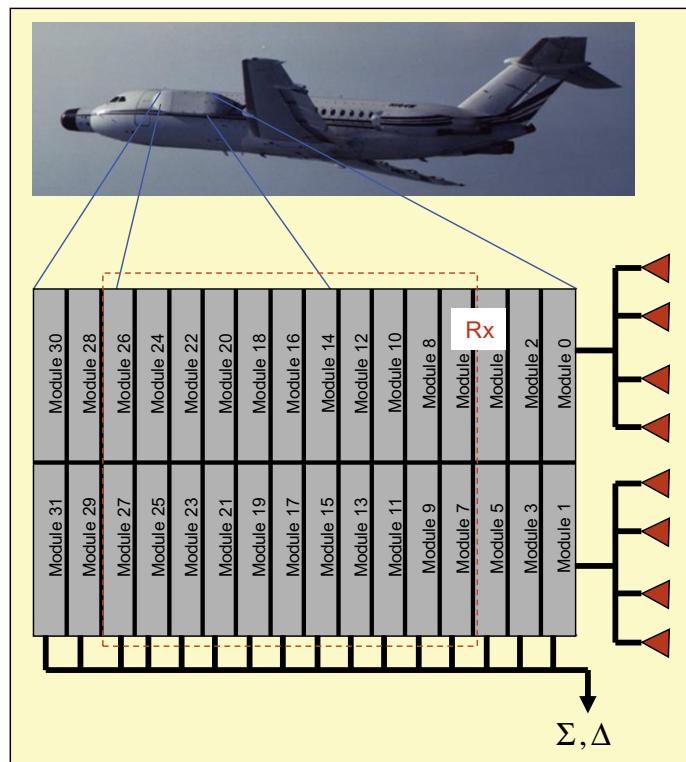
2.12.8.1 Multichannel Airborne Radar Measurements (MCARM)

In the mid-1990s, the US Air Force Rome Laboratory (now the Air Force Research Laboratory) contracted with Westinghouse Corporation (now Northrop Grumman) to build and fly a multi-channel, L-band radar under the Multi-Channel Airborne Radar Measurements (MCARM) Program [83]. The MCARM program produced a meaningful collection of multi-channel radar data over a range of clutter conditions to evaluate the performance potential of STAP and support continued technology development.

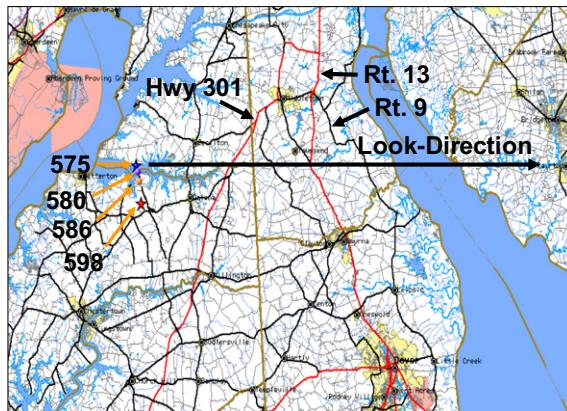
The salient parameters of the MCARM system are given in Table 12.8. As seen from the table, the system has relatively high peak power, an electrically modest aperture at L-band, relatively low range resolution, and built-in features to calibrate the antenna and radio frequency front-end. Additionally, as Figure 12.30 shows, MCARM has a highly flexible array configuration, supporting a range of study options, including planar array configuration, uniform linear array, and sum and difference channel analyses. Figure 12.31 shows the region of the Eastern United States—the Delmarva Peninsula—where

Table 12.8 MCARM Nominal System Parameters

Parameter	Value	Parameter	Value
Transmit frequency	<i>L</i> -band (1240 MHz)	Beamwidth	7.5°, and 3 × “blob” pattern
Bandwidth	1 MHz Transmit, 0.8 MHz Receive	Array configuration	Most collections use 11-over-11 receive channels, plus sum and difference (see figure)
Peak power	15 kW	Sample rate	5 MHz
Waveform	Linear frequency modulated or gated RF	CPI length	Nominally 128 pulses
PRF	0.5 kHz, 2 kHz, 7 kHz	Other	Test manifold for channel balance, range measured steering vectors

**FIGURE 12.30**

MCARM data collection system.

**FIGURE 12.31**

CPI locations for several acquisitions during MCARM Flight 5.

several acquisitions were collected during MCARM Flight 5. As previously mentioned, this region is relatively flat, dominated by rural clutter and water, and covered by a number of roadways, including several highways. Results from Acquisition 575, in particular, have been presented in a number of papers, starting with [50].

It is known that some of the MCARM acquisitions were previously made publicly releasable, as described in [84]. The Air Force Research Laboratory, Sensors Directorate, maintains the MCARM database.

2.12.8.2 Naval Research Laboratory (NRL) database

The Naval Research Laboratory (NRL) Adaptive Array Flight Test Database is described in [85]. Table 12.9 provides some flight test parameters. According to [85], the effort remarkably consisted of over thirty flights, collecting in excess of 2500 data files. Data were collected predominantly in the mountainous regions of Virginia and West Virginia, with urban clutter data collected in the vicinity of Washington, DC, and rural clutter data collected in North Carolina, Georgia, Ohio, and Indiana. Sea clutter data were collected over the Atlantic Ocean, east of Virginia, with some littoral clutter data also recorded.

Lee and Staudaher [85] indicates a database was constructed, with data formatted for MATLAB[®] analysis, with the intent “that the database will provide the adaptive signal processing community with a valuable tool for validating both existing and future adaptive algorithms.”

2.12.8.3 Mountaintop database

According to [86,87], the Mountaintop Program goal sought to advance airborne early warning radar technology. Collection of surveillance radar clutter and jamming data from mountaintops at White Sands Missile Range, New Mexico, and Pacific Missile Range Facility, Hawaii, served as central objectives in this effort. The Mountaintop radar, known as the Radar Surveillance Technology Experimental Radar (RSTER), employed an inverse displaced phase center antenna (IDPCA) technique to give the

Table 12.9 NRL Adaptive Array Flight Test System Parameters (from [85])

Parameter	Value	Parameter	Value
Transmit frequency	UHF	Beamwidth	15.7° azimuth, 80° elevation (16.5 dBi gain)
Bandwidth	2.6 MHz	Array configuration	Eight channel, uniform linear array
Peak power	In excess of 1 MW	Sample rate	5 MHz (separate in-phase and quadrature channels)
Waveform	Unstated, presumably linear frequency modulated or gated RF	CPI length	3–18 pulses at low PRF, 34–66 pulses for high PRF
PRF	Typically 300 Hz or 750 Hz	Other	Test manifold for channel balance

Table 12.10 Mountaintop RSTER Parameters (from [87, 88])

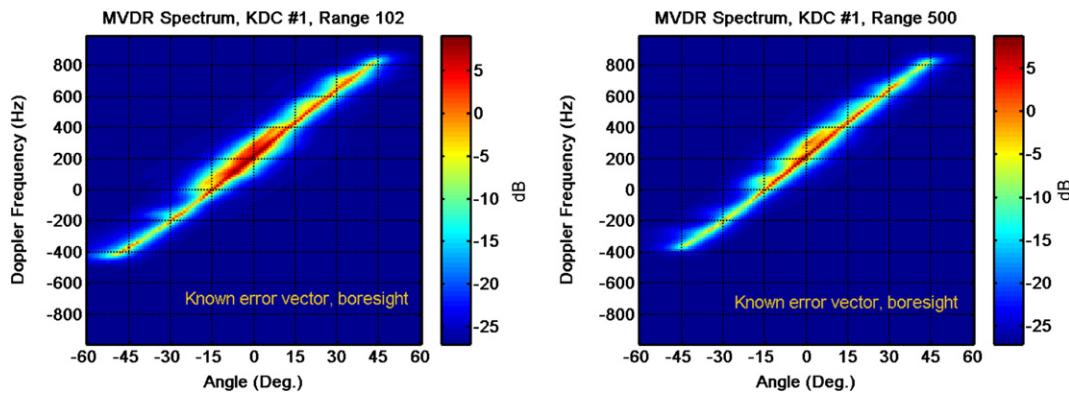
Parameter	Value	Parameter	Value
Transmit frequency	UHF (400–500 MHz)	Beamwidth	9° azimuth, 6° elevation
Bandwidth	200 kHz	Array configuration	14 channel uniform linear array
Peak power	100 kW	Sample Rate	1 MHz [88]
Waveform	Linear frequency modulated [yy]	CPI length	16 pulses
PRF	250–1500 Hz	Other	Horizontally polarized, relocatable

appearance of platform motion. RSTER parameters are given in Table 12.10. In IDPCA, the system steps the transmit location to one (or three) of eighteen total transmit antenna columns.

According to [87], Mountaintop RSTER data are archived using the Common Research Environment for STAP (CREST) database. Additionally, some data is accessible at the IEEE Signal Processing Information Base at Rice University [88].

2.12.8.4 Knowledge-aided sensor signal processing and expert reasoning (KASSPER) data

Improving the performance of STAP in complex clutter environments served as a primary aim of the DARPA KASSPER Program [89]. Under this program, DARPA funded the creation of several sets of synthetic, multichannel, multi-pulse, multi-range datacubes at *X*-band and *L*-band. The *L*-band datacube—KASSPER datacube #1, or KDC #1—is publicly releasable and was provided to attendees at the 2002 DARPA KASSPER Workshop [90]. Many of the *L*-band data parameters resemble those of

**FIGURE 12.32**

MVDR spectra for KDC #1, range bins 102 (left) and 500 (right).

the MCARM system described in Table 12.8, the exception mainly being the use of an eleven channel uniform linear array and typically fewer pulses in the CPI. Each of the KASSPER datacubes is a result of high-fidelity, site-specific clutter simulation. Dense moving targets are present in the data. A unique aspect of the simulation is the availability of truth information, including known covariance matrices on a range cell basis and target location data for some of the datasets (including the *L*-band datacube and some of the *X*-band cases).

Figure 12.32 shows MVDR spectra for KDC #1 range bins 102 and 500, with pre-steering applied to the antenna boresight. Additionally, the KASSPER datacubes included simulated antenna errors, with the true error vector provided with the data sets; the two-dimensional MVDR spectra in Figure 12.32 employ knowledge of the precise array manifold.

2.12.9 Summary

This tutorial describes a number of key aspects of STAP. Naturally, it is challenging to fully cover all the details of a topic so vast. Our goals for this effort are simpler: provide a concise, single reference to augment the available text books, reports, and papers on this important subject and highlight some of the more significant developments within the research community.

In this tutorial, we describe a number of approaches to design optimal filters. Of the methods described, all end up being closely related and ultimately maximize the probability of detection for a fixed probability of false alarm. We then discuss signal models for clutter, noise, radio frequency interference and target. These models form the basis for STAP algorithm development and analysis. STAP, as it turns out, refers to a collection of a number of practical techniques; we discuss several major algorithm themes, include post-Doppler STAP, pre-Doppler STAP, and parametric STAP. We further discuss a number of eigen- or subspace-based cancellation architectures, including Principle Components Inverse, the cross spectral metric, and the Multi-Stage Wiener Filter.

Using the simulation models and algorithm descriptions, we then characterize the performance of the various methods using SINR loss, a key performance metric in the STAP community, using a homogeneous clutter scenario. We find that, for the approach taken, all but one of the methods provides similarly good performance relative to the bound set by the optimal space-time filter.

Key challenges in the STAP research community center on modifying textbook principles to support real-world application. Invariably, the accurate estimation of the unknown clutter covariance matrix and implementing the adaptive filter in a real-time computing environment are driving considerations. In this vein, we briefly discuss the challenges of STAP application in heterogeneous clutter environments, for bistatic or conformal array configurations, and in the presence of otherwise nonstationary clutter.

A number of concepts are concisely summarized in Section 2.12.7, where we describe the end-to-end implementation of a basic MTI radar processor. STAP normalization and target parameter estimation is considered in this section.

We conclude the discussion by summarizing several STAP data collection programs, the characteristics of the various radar systems, and the type of data collected or generated.

Symbols and notation

We provide a key to some common notation used in this paper.

f_s	spatial frequency
f_d	Doppler frequency (Hz)
ϕ, θ	azimuth and elevation (rads)
ϕ_{cone}	cone angle (rads)
ψ	grazing angle (rads)
λ	wavelength (m)
M	number of channels
N	number of pulses
L	number of available range bins
T	pulse repetition interval (s)
σ_n^2	noise variance (w)
σ_T^2	target signal variance (w)
\mathbf{v}_p	platform velocity vector (m/s, m/s, m/s)
\mathbf{s}_s	spatial steering vector
\mathbf{v}_s	hypothesized spatial steering vector
\mathbf{s}_t	temporal steering vector
\mathbf{v}_t	hypothesized temporal steering vector
\mathbf{s}_{s-t}	space-time steering vector
\mathbf{v}_{s-t}	hypothesized space-time steering vector
$\mathbf{x}_{s/k}(n)$	spatial data snapshot, k th range cell, n th pulse
\mathbf{x}_k	space-time data snapshot, k th range cell
\mathbf{c}_k	clutter space-time snapshot, k th range cell
\mathbf{j}_k	RFI space-time snapshot, k th range cell

n_k	uncorrelated noise space-time snapshot, k th range cell
t_k	target space-time snapshot, k th range cell
R_k	null-hypothesis covariance matrix (k th range cell)
$R_{c/k}$	clutter covariance matrix
R_{RFI}	RFI covariance matrix
R_n	receive noise covariance matrix
\hat{R}_k	null-hypothesis covariance estimate, k th range cell
w_k	space-time weight vector
\hat{w}_k	adaptive space-time weight vector
P_d	probability of detection
P_{fa}	probability of false alarm.

Acronyms

Following is a list of common acronyms:

ADPCA	adaptive displaced phase center antenna
AMF	adaptive matched filter
CFAR	constant false alarm rate
CNR	clutter-to-noise ratio
CPI	coherent processing interval
CREST	Common Research Environment for STAP
CUT	cell-under-test
DNR	discrete-to-noise ratio
DoFs	degrees of freedom
FIR	finite impulse response
FLAR	forward-looking array radar
GLSC	generalized sidelobe canceler
GMTI	ground moving target indication
HTP	Hung-Turner projection
ICM	intrinsic clutter motion
IDPCA	inverse displaced phase center antenna processing
IF	improvement factor
IID	independent and identically distributed
JDL	joint domain localized
KASSPER	Knowledge-Aided Sensor Signal Processing and Expert Reasoning
MCARM	Multi-Channel Airborne Radar Measurements
MCLP	multichannel linear prediction
MDV	minimum detectable velocity
MLE	maximum likelihood estimate
MMSE	minimum mean square error
MSE	mean square error

MSMI	modified sample matrix inversion
MPE	minimum power eigencanceler
MNE	minimum norm eigencanceler
MTI	moving target indication
MV	minimum variance
MVDR	minimum variance distortionless response
MWF	multi-stage Wiener filter
NCA	noncoherent addition
pdf	probability density function
PDI	post-detection integration
PRI	pulse repetition interval
PSD	power spectral density
RDM	range-Doppler map
RFI	radio frequency interference
RD-STAP	reduced-dimension STAP
RR-STAP	reduced-rank STAP
RSTER	Radar Surveillance Technology Experimental Radar
SAR	synthetic aperture radar
SLAR	sidelooking array radar
SMI	sample matrix inversion
SNR	signal-to-noise ratio
SINR	signal-to-interference-plus-noise ratio
STAP	space-time adaptive processing
STOP	space-time optimal processing
TNR	target-to-noise ratio
TSD	targets in the secondary data
WGN	white Gaussian noise
WSS	wide-sense stationary

Relevant Theory: Statistical Signal Processing and Array Signal Processing

See [Vol. 3, Chapter 5](#) Distributed Signal Detection

See [Vol. 3, Chapter 19](#) Array Processing in the Face of Nonidealities

References

- [1] M. Richards, et al. (Eds.), *Principles of Modern Radar: Basic Principles*, SciTech Publishing, Inc., North Carolina, 2010.
- [2] J.V. DiFranco, W.L. Rubin, *Radar Detection*, Artech-House, Dedham, MA, 1980.
- [3] L.E. Brennan, I.S. Reed, *Theory of adaptive radar*, IEEE Trans. AES 9 (2) (1973) 237–252.
- [4] I.S. Reed, *A brief history of adaptive arrays*, Sudbury/Wayland Lecture Series (Raytheon Div. Education), Notes, 23 October 1985.

- [5] R. Klemm, Doppler properties of airborne clutter, in: Proceedings of the Research and Technology Organization, North Atlantic Treaty Organization (RTO-NATO) Lecture Series 228—Military Applications of Space-Time Adaptive Processing, RTO-ENP-027, September 2002, pp. 2-1–2-24.
- [6] J.N. Entzminger, C.A. Fowler, W.J. Kenneally, JointSTARS and GMTI: past, present and future, IEEE Trans. AES 35 (2) (1999) 748–761.
- [7] M.I. Skolnik, Introduction to Radar Systems, second ed., McGraw Hill, New York, NY, 1980.
- [8] N. Levanon, Radar Principles, John Wiley & Sons, New York, 1988.
- [9] W.L. Melvin, A STAP overview, in: Peter Willett (Ed.), IEEE AES Systems Magazine—Special Tutorials Issue, 19 (1) (2004) 19–35.
- [10] J. Ward, Space-Time Adaptive Processing for Airborne Radar, Lincoln Laboratory Technical Report, ESC-TR-94-109, December 1994.
- [11] J.R. Guerci, Space-Time Adaptive Processing for Radar, Artech House, Norwood, MA, 2003.
- [12] S. Haykin, Adaptive Filter Theory, third ed., Prentice-Hall, Upper Saddle River, NJ, 1996.
- [13] D.H. Johnson, D.E. Dudgeon, Array Signal Processing: Concepts and Techniques, Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [14] L.J. Griffiths, C.W. Jim, An alternative approach to linearly constrained adaptive beamforming, IEEE Trans. Antenn. Propag. 30 (1) (1982) 27–34.
- [15] L.M. Novak, M.B. Sechtin, M.J. Cardullo, Studies of target detection algorithms that use polarimetric radar data, IEEE Trans. AES 25 (2) (1989) 150–165.
- [16] W.L. Melvin, G.A. Showman, Knowledge-aided parametric covariance estimation, IEEE Trans. AES (2006) 1021–1042.
- [17] I.S. Reed, J.D. Mallett, L.E. Brennan, Rapid convergence rate in adaptive arrays, IEEE Trans. AES 10 (6) (1974) 853–863.
- [18] R. Klemm, Space-time adaptive processing: principles and applications, IEE Radar, Sonar, Navigation and Avionics 9, IEE Press, 1998.
- [19] R. Klemm, Principles of space-time adaptive processing, second ed., IEE Radar, Sonar, Navigation and Avionics 12, IEE Press, UK, 2002. (Note: a 3rd edition of this text, published in 2006, is also available.)
- [20] D. Barton, Land clutter models for radar design and analysis, Proc. IEEE 73 (2) (1985) 198–204.
- [21] J.B. Billingsley, Low-Angle Radar Land Clutter: Measurements and Empirical Models, William Andrew Publishing, Inc., 2002.
- [22] R.J. Mailloux, Phased Array Antenna Handbook, Artech House, Boston, MA, 1994.
- [23] M. Zatman, How narrow is narrowband? IEE Proc. Radar Sonar Navig. 145 (2) (1998) 85–91.
- [24] W.L. Melvin, G.A. Showman, J.R. Guerci, A knowledge-aided GMTI detection architecture, in: Proceedings of the 2004 IEEE Radar Conference, Philadelphia, PA, 26–29 April, 2004, ISBN: 0-7803-8235-8.
- [25] D.V. Rabinkin, N.B. Pulsone, Subband-domain signal processing for radar array systems, in: Proceedings of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations IX, Denver, CO., July 1999, pp. 174–187.
- [26] R. Klemm, Adaptive clutter suppression for airborne phased array radar, Proc. IEE 130 (1) (1983) 125–132.
- [27] C.H. Gierull, B. Balaji, Minimal sample support space-time adaptive processing with fast subspace techniques, IEE Proc. Radar Sonar Navig. 149 (5) (2002) 209–220.
- [28] W.F. Gabriel, Using spectral estimation techniques in adaptive processing antenna systems, IEEE Trans. Antenn. Propag. 34 (3) (1986) 291–300.
- [29] D.W. Tufts, I. Kirsteins, R. Kumaresan, Data-adaptive detection of a weak signal, IEEE Trans. AES 19 (2) (1983) 313–316.
- [30] A.M. Haimovich, The Eigencanceler: adaptive radar by eigenanalysis methods, IEEE Trans. AES 32 (2) (1996) 532–542.

- [31] E.K.L. Hung, R.M. Turner, A fast beamforming algorithm for large arrays, *IEEE Trans. AES* 19 (4) (1983) 598–607.
- [32] B.D. Carlson, Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans. AES* 24 (4) (1988) 397–401.
- [33] J.R. Guerci, J.S. Goldstein, I.S. Reed, Optimal and adaptive reduced-rank STAP, *IEEE Trans. AES* 36 (2) (2000) 647–661.
- [34] S.D. Berger, B.M. Welsh, Selecting a reduced-rank transformation for STAP—a direct form perspective, *IEEE Trans. AES* 35 (2) (1999) 722–729.
- [35] J.S. Goldstein, I.S. Reed, P.A. Zulch, Multistage partially adaptive STAP CFAR detection algorithm, *IEEE Trans. AES* 35 (2) (1999) 645–661.
- [36] R.C. DiPietro, Extended factored space-time processing for airborne radar, in: Proceedings of the 26th Asilomar Conference, Pacific Grove, CA, October 1992, pp. 425–430.
- [37] H. Wang, L. Cai, On adaptive spatial-temporal processing for airborne surveillance radar systems, *IEEE Trans. AES* 30 (3) (1994) 660–670.
- [38] R. Brown, M. Wicks, Y. Zhang, Q. Zhang, H. Wang, A space-time adaptive processing approach for improved performance and affordability, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 321–326.
- [39] R. Klemm, Antenna design for airborne MTI, in: Proceedings of the Radar 92, October 1992, Brighton, UK, pp. 296–299.
- [40] R. Blum, W. Melvin, M. Wicks, An analysis of adaptive DPCA, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 303–308.
- [41] J.R. Roman, M. Rangaswamy, D.W. Davis, Q. Zhang, B. Himed, J.H. Michels, Parametric adaptive matched filter for airborne radar applications, *IEEE Trans. AES* 36 (2) (2000) 677–692.
- [42] G.A. Showman, Personal communication, 13 July 2003.
- [43] R.L. Fante, E.C. Barile, T.P. Guella, Clutter covariance smoothing by subaperture averaging, *IEEE Trans. AES* 30 (3) (1994) 941–945.
- [44] W.L. Melvin, Space-time adaptive radar performance in heterogeneous clutter, *IEEE Trans. AES* 36 (2) (2000) 621–633.
- [45] W.L. Melvin, J.R. Guerci, Adaptive detection in dense target environments, in: Proceedings of the 2001 IEEE Radar Conference, Atlanta, GA, 1–3 May 2001, pp. 187–192.
- [46] W.L. Melvin, STAP in heterogeneous clutter environments, in: R. Klemm (Ed.), *The Applications of Space-Time Processing*, IEE Radar, Sonar, Navigation and Avionics 9, IEE Press, 2004.
- [47] W.L. Melvin, J.R. Guerci, Knowledge-aided sensor signal processing: a new paradigm for radar and other sensors, *IEEE Trans. AES* (2006) 983–996.
- [48] O. Kreyenkamp, R. Klemm, Doppler compensation in forward-looking STAP radar, *IEE Proc. Radar Sonar Navig.* 148 (5) (2001) 252–258.
- [49] G. Borsari, Mitigating effects on STAP processing caused by an inclined array, in: Proceedings of the 1998 IEEE Radar Conference, Dallas, Tx, May 1998, pp. 135–140.
- [50] W.L. Melvin, M.C. Wicks, R.D. Brown, Assessment of multichannel airborne radar measurements for analysis and design of space-time processing architectures and algorithms, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 130–135.
- [51] W.L. Melvin, M.C. Wicks, Improving practical space-time adaptive radar, in: Proceedings of the 1997 IEEE National Radar Conference, Syracuse, New York, May 13–15, 1997, pp. 48–53.
- [52] P. Chen, W.L. Melvin, M.C. Wicks, Screening among multivariate normal data, *J. Multivariate Anal.* 69 (1999) 10–29.

- [53] D.J. Rabideau, A.O. Steinhardt, Improving the performance of adaptive arrays in non-stationary environments through data-adaptive training, in: Proceedings of the 30th Asilomar Conference, Pacific Grove, CA, 3–6 November 1996, pp. 75–79.
- [54] D.J. Rabideau, A.O. Steinhardt, Improved adaptive clutter cancellation through data-adaptive training, *IEEE Trans. Aerosp. Electron. Syst.* 35 (3) (1999) 879–891.
- [55] G.A. Showman, W.L. Melvin, Knowledge-aided discrete removal techniques, in: Proceedings of the 2005 DARPA ISIS/KASSPER Workshop, Las Vegas, NV, 22–24 February 2005, CD ROM.
- [56] N. Pulsone, Improving ground moving target indication performance, in: Proceedings of the DARPA/AFRL KASSPER Workshop, 14–16 April 2003, Las Vegas, NV, 2003.
- [57] S.D. Berger, W.L. Melvin, G.A. Showman, Map-aided secondary data selection, in: Proceedings of the 2007 IEEE Radar Conference, Boston, MA, April 2007.
- [58] J.R. Guerci, Theory and application of covariance matrix tapers for robust adaptive beamforming, *IEEE Trans. Signal Process.* 47 (4) (1999) 977–985.
- [59] S.U. Pillai, Y.L. Kim, J.R. Guerci, Generalized forward/backward subaperture smoothing techniques for sample starved STAP, *IEEE Trans. Signal Process.* 48 (12) (2000) 3569–3574.
- [60] A. Farina, P. Lombardo, M. Pirri, Nonlinear nonadaptive space-time processing for airborne early warning radar, *IEE Proc. Radar Sonar Navig.* 145 (1) (1998) 9–18.
- [61] J.S. Bergin, C.M. Teixeira, P.M. Techau, J.R. Guerci, STAP with knowledge-aided data pre-whitening, in: Proceedings of the 2004 IEEE Radar Conference, April 2004, pp. 289–294.
- [62] J.S. Bergin, C.M. Teixeria, P.M. Techau, J.R. Guerci, Space-time beamforming with knowledge-aided constraints, in: Proceedings of the 2003 Adaptive Sensor Array Processing Workshop, MIT Lincoln Laboratory, Lexington, MA, March, 2003.
- [63] G.R. Letgers, J.R. Guerci, Physics-based airborne GMTI radar signal processing, in: Proceedings of the 2004 IEEE Radar Conference, Philadelphia, PA, April 2004, pp. 283–288, ISBN: 0-7803-8234-X.
- [64] C.D. Richmond, Statistical performance analysis of the adaptive sidelobe blanker detection algorithm, in: Proceedings of the 31st Annual Asilomar Conference—Signal, Systems & Computers, Pacific Grove, CA, 2–5 November 1997, pp. 872–876.
- [65] W.L. Melvin, G.A. Showman, Knowledge-aided, physics-based signal processing for next-generation radar, in: Proceedings of the Asilomar Conference on Signals, Systems, Computers, Pacific Grove, CA, November 2007.
- [66] W.L. Melvin, M.J. Callahan, M.C. Wicks, Adaptive clutter cancellation in bistatic radar, in: Proceedings of the 34th Asilomar Conference, Pacific Grove, CA, October 2000, pp. 1125–1130.
- [67] B. Himed, J.H. Michels, Y. Zhang, Bistatic STAP performance analysis in radar applications, in: Proceedings of the 2001 IEEE Radar Conference, Atlanta, GA, May 2001, pp. 198–203.
- [68] S.D. Hayward, Adaptive beamforming for rapidly moving arrays, in: Proceedings of the CIE International Conference on Radar, IEEE Press, Beijing, CN, 8–10 October 1996, pp. 480–483.
- [69] S.M. Kogon, M.A. Zatman, Bistatic STAP for airborne radar systems, in: Proceedings of the IEEE SAM, Lexington, MA, March 2000.
- [70] F. Pearson, G. Borsari, Simulation and analysis of adaptive interference suppression for bistatic surveillance radars, in: Proceedings of the 2001 ASAP Symposium, Lexington, MA, 13 March 2001.
- [71] B. Himed, Y. Zhang, A. Hajjari, STAP with angle-Doppler compensation for bistatic airborne radars, in: Proceedings of the 2002 IEEE Radar Conference, Long Beach, CA, 22–25 April 2002, ISBN: 0-7803-7358-8.
- [72] W.L. Melvin, M.E. Davis, Adaptive cancellation method for geometry-induced non-stationary bistatic clutter environments, *IEEE Trans. AES* (2007) 651–672.
- [73] W.L. Melvin, B. Himed, M.E. Davis, Doubly-adaptive bistatic clutter filtering, in: Proceedings of the 2003 IEEE Radar Conference, Huntsville, AL, 5–8 May 2003, pp. 171–178.

- [74] F.D. Lapierre, J.G. Verly, M. Van Droogenbroeck, New solutions to the problem of range dependence in bistatic STAP radars, in: Proceedings of the 2003 IEEE Radar Conference, Huntsville, AL, 5–8 May 2003, pp. 452–459.
- [75] W.L. Melvin, Adaptive moving target indication, in: N. Willis, H. Griffiths (Eds.), *Advances in Bistatic Radar*, Sci-Tech Publishing, 2007 (Chapter 11).
- [76] R.K. Hersey, W.L. Melvin, J.H. McClellan, E. Culpepper, Adaptive ground clutter suppression for conformal array radar systems, *IET Radar Sonar Navig.* 3 (4) (2009) 357–372.
- [77] R.K. Hersey, W.L. Melvin, E. Culpepper, Adaptive filtering for conformal array radar, in: Proceedings of the IEEE Radar Conference, May 2008, Rome, Italy.
- [78] R.K. Hersey, W.L. Melvin, J.H. McClellan, Clutter-limited detection performance of multi-channel conformal arrays, *Signal Process.* 84 (2004) 1481–1500 (special issue on New Trends and Findings in Antenna Array Processing for Radar).
- [79] F.C. Robey, D.R. Fuhrman, E.J. Kelly, R. Nitzberg, A CFAR adaptive matched filter detector, *IEEE Trans. AES* 28 (1) (1992) 208–216.
- [80] W.S. Chen, I.S. Reed, A new CFAR detection test for radar, *Digital Signal Processing*, vol. 1, Academic Press, 1991, pp. 198–214.
- [81] R.C. Davis, L.E. Brennan, I.S. Reed, Angle estimation with adaptive arrays in external noise fields, *IEEE Trans. AES* AES-12 (2) (1976) 179–186.
- [82] W.L. Melvin, R. Hancock, M. Rangaswamy, J. Parker, Adaptive distributed radar, in: Proceedings of the 2009 International Radar Conference, Bordeaux, France, October 2009.
- [83] D.K. Fenner, W.F. Hoover, Test results of a space-time adaptive processing system for airborne early warning radar, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, MI, 13–16 May 1996, pp. 88–93.
- [84] V. Cavo, MCARM/STAP Data Analysis, Final Technical Report, Air Force Research Laboratory, AFRL-SN-RS-TR-1999-48, May, Volume One (of Two), 1999.
- [85] F. Lee, F. Staudaher, The NRL adaptive array flight test database, in: Proceedings of the IEEE Adaptive Antenna Systems Symposium, 1992, pp. 101–104.
- [86] G.W. Titi, An overview of the ARPA/Navy mountaintop program, in: Proceedings of the IEEE Adaptive Antenna Systems Symposium, 1994.
- [87] G.W. Titi and D.F. Marshall, The ARPA/Navy Mountaintop Program: adaptive signal processing for airborne early warning radar, in: Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 2, pp. 1165–1168.
- [88] <http://spib.rice.edu/spib/mtn_top.html>.
- [89] J.R. Guerci, Knowledge-aided sensor signal processing and expert reasoning, in: Proceedings of the 2002 Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER) Workshop, Washington, DC, 3 April, 2002, CD ROM.
- [90] J.S. Bergin, P.M. Techau, (Electronic data) Workshop datacube, in: Proceedings of the 2002 KASSPER Conference, Washington, DC, 3 April 2002, CD ROM.

MIMO Radar with Widely Separated Antennas—From Concepts to Designs

13

Qian He^{*}, Yang Yang[†], and Rick S. Blum[†]

^{*}*Department of Electronic Engineering, University of Electronic Science and Technology of China,
Chengdu, Sichuan, China*

[†]*Department of Electrical and Computer Engineering, Lehigh University,
Bethlehem, PA, USA*

2.13.1 Introduction

Recent advances in wireless communications featuring the innovative multiple-input multiple-output (MIMO) technology [1] have catalyzed a wave of interest in understanding and exploiting the concept of MIMO radar, e.g., [2–7]. The similarity between MIMO communications and MIMO radar systems that employ widely separated antennas is rather intriguing: in communications, MIMO systems combat the fading effects of the multipath channel through its spatial diversity advantage; in radar, the complex targets consisting of several scatterers resemble very much the multipath channel in wireless communications, and likewise, MIMO radar with widely separated antennas also offers the diversity gain. To be more specific, a target's radar cross section (RCS), which determines the amount of returned power, varies greatly with respect to the considered aspect angle. Those variations can significantly degrade the ability of a conventional radar in detecting and estimating the target. MIMO radar with widely separated antennas, whereas, through observing a target simultaneously from different (uncorrelated) aspect angles, provide spatial diversity which can substantially countervail the fluctuations in received power.

Since the inception of MIMO radar in the early 2000s, a great deal of efforts have been devoted to studying its performance potentials or developing a variety of application paradigms for it, e.g., [8–13]. However, note that MIMO radar systems studied in most of the literature can be roughly classified into two categories. The first category features the use of closely spaced antennas [3], i.e., the array configuration of these radars is close to that of the conventional phased array radar. But the utilization of specific (e.g., orthogonal) transmit waveforms in this type of MIMO radar systems can render many benefits which are, otherwise, not achievable with the conventional phased array radars. The second category, as surveyed in [4], takes advantage of multiple transmit signals as well. But it employs widely separated antennas at both the transmit and receive ends, and thus enjoys the spatial diversity gain.

Among all these widely-separated-antenna cases are two cases that have received special attention: one called coherent MIMO radar and the other called noncoherent MIMO radar [4]. The distinguishing features are whether the target reflection model is coherent and whether the processing is coherent.

For the case of coherent MIMO radar, the antennas are within a given target beamwidth¹ which leads to identical (coherent) reflection coefficients, and a coherent processing approach is adopted. While for the case of noncoherent MIMO radar, the antennas lie in different target beamwidths which leads to distinct (noncoherent) reflection coefficients, and a noncoherent processing approach is adopted. The coherent processing requires phase synchronization of the oscillators employed at widely separated antennas, while the noncoherent processing does not. In fact, phase synchronization embodies a major difference between the operations of noncoherent MIMO radar and coherent MIMO radar, as will be mathematically demonstrated later in this chapter.

In this document, we will focus on the MIMO radar with widely separated antennas. Although far from being exhaustive, this document strives to summarize and discuss a large range of issues related to this particular type of MIMO radar—both coherent and noncoherent processing approaches included. Topics of interest encompass those with strong theoretical significance, for example, performance evaluation of both coherent and noncoherent MIMO radar in target localization and velocity estimation, study of diversity gain for MIMO radar under the Neyman-Pearson (NP) criteria, etc. Topics of realistic values are also covered, which include, for example, phase synchronization algorithm design for coherent MIMO radar, and MIMO radar waveform design for extended targets. Through reviewing some state of the art in this area, ranging from theoretical concepts to practical designs, this document is intended to serve as an easy beginning as well as a handy reference for researchers who have interest in delving into this field. It is also expected to foster further discussions on those related research topics, and to spur further research interest within this field.

The remainder of this document is organized as follows. In Section 2.13.2, after a brief introduction to the coherent MIMO radar, we derive the mean square error (MSE) of the maximum likelihood (ML) estimate and the Cramer-Rao bound (CRB) for joint target location and velocity estimation, and investigate the impact of static phase errors at the transmitters and receivers on the performance of target localization. In Section 2.13.3, we turn our attention to noncoherent MIMO radar. Parallel to the study for the coherent case, the joint estimation for target location and velocity using noncoherent MIMO radar is presented, the MSE of the ML estimate is analyzed, and the CRB is calculated. Then, the noncoherent MIMO radar ambiguity function (AF) is introduced. In Section 2.13.4, we discuss the MSE performance differences between coherent and noncoherent MIMO radars in the application of joint target location and velocity estimation. We demonstrate that the magnitude of these differences decreases with an increase in the product of the number of transmit and receive antennas, when the antennas for both coherent and noncoherent systems are properly placed. In Section 2.13.5, we derive the diversity gain for a MIMO radar system adopting the Neyman-Pearson detection. The relationship between the cumulative distribution function (cdf) of the reflection coefficients, the cdf of the clutter-plus-noise, and the signal space dimension of the transmitted waveforms is described. In Section 2.13.6, we present three phase synchronization approaches for coherent MIMO radar, which include the master-slave closed-loop method, the round-trip approach, and the broadcast consensus based algorithm. We compare these three phase synchronization approaches and discuss some issues that may arise in practice. In Section 2.13.7, we introduce some waveform design schemes for MIMO radar with widely separated antennas. Finally we conclude this document with a summary in Section 2.13.8.

¹If we model the target as an antenna, we can define a target beamwidth based on its size using standard antenna array calculations.

2.13.2 Coherent MIMO radar

Before exposing readers to a variety of interesting issues related to the coherent MIMO radar, we firstly introduce some settings that hold true for both the coherent and noncoherent MIMO radars, which lays a common ground for the subsequent analysis. Let us consider a MIMO radar which is equipped with M transmitters and N receivers. The positions of the k th, $k = 1, \dots, M$ transmitter and the l th, $l = 1, \dots, N$ receiver are (x_k^t, y_k^t) and (x_l^r, y_l^r) respectively, in a two-dimensional Cartesian coordinate system. The lowpass equivalent of the signal transmitted from the k th transmitter is $\sqrt{E/M}s_k(t)$, where E denotes the total transmitted energy, and the waveform is normalized such that

$$\int_{-\infty}^{\infty} |s_k(t)|^2 dt = 1. \quad (13.1)$$

Assume a target, if present, is located at (x, y) and moving with velocity (v_x, v_y) . The time delay τ_{lk} and Doppler shift f_{lk} involved in the path from transmitter k to receiver l , via the target reflection, are

$$\tau_{lk} = \frac{\sqrt{(x_k^t - x)^2 + (y_k^t - y)^2} + \sqrt{(x_l^r - x)^2 + (y_l^r - y)^2}}{c} = \frac{d_k^t + d_l^r}{c} \quad (13.2)$$

and

$$f_{lk} = \frac{v_x(x_k^t - x) + v_y(y_k^t - y)}{\lambda d_k^t} + \frac{v_x(x_l^r - x) + v_y(y_l^r - y)}{\lambda d_l^r}, \quad (13.3)$$

where c is the speed of light, d_k^t denotes the distance between the target and the k th transmitter, d_l^r denotes the distance between the target and the l th receiver, and $\lambda = c/f_c$ represents the wavelength of the carrier with frequency f_c . In both the coherent and noncoherent MIMO radars we assume all transmitter and receiver nodes have oscillators which are locked in frequency, possibly due to the use of a beacon. For the coherent MIMO radar, we also assume these oscillators are locked in phase.

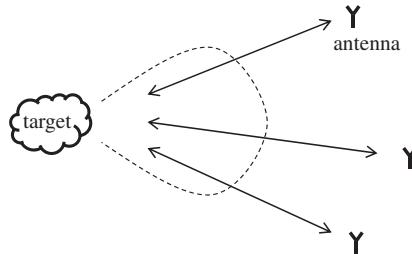
Throughout this chapter, we use the term *coherent MIMO radar* to refer to MIMO radar system employing coherent processing of received signals obeying a coherent target reflection model. If we model the target as an antenna then we can define an equivalent target beamwidth based on the target size [2]. In a coherent MIMO radar system, the antennas are assumed to be all within the same target beamwidth, as illustrated in Figure 13.1. In this case, the effective target reflection coefficient is assumed to be *identical* for each transmitter-target-receiver path (which gives a coherent target reflection model) and is denoted by $\zeta = \zeta_R + j\zeta_I$. We assume ζ is unknown but deterministic. Thus, the received signal at receiver l can be modeled as

$$r_l(t) = \sqrt{\frac{E}{M}} \zeta \sum_{k=1}^M s_k(t - \tau_{lk}) e^{-j2\pi f_c \tau_{lk}} e^{j2\pi f_{lk} t} + w_l(t) \quad (13.4)$$

$$= \sqrt{\frac{E}{M}} \zeta \mathbf{u}_l^T \mathbf{1}_{M \times 1} + w_l(t), \quad (13.5)$$

where $w_l(t)$ represents the clutter-plus-noise component at the l th receiver for any given time t , and

$$\mathbf{u}_l = [u_{l1}, u_{l2}, \dots, u_{lM}]^T \quad (13.6)$$

**FIGURE 13.1**

A coherent MIMO radar system with all three antennas in the same target beamwidth.

in which the k th element

$$u_{lk} = s_k(t - \tau_{lk}) e^{-j2\pi f_c \tau_{lk}} e^{j2\pi f_{lk} t}$$

represents the signal transmitted from transmitter k and received by receiver l . Collecting the time delayed and Doppler shifted signals from all paths in an $N \times MN$ block diagonal matrix, we arrive at

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1^T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{u}_2^T & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{u}_N^T \end{bmatrix}. \quad (13.7)$$

Then, the signals received at all N antennas can be written as

$$\mathbf{r}(t) = [r_1(t), r_2(t), \dots, r_N(t)]^T = \sqrt{\frac{E}{M}} \xi \mathbf{U} \mathbf{1}_{NM \times 1} + \mathbf{w}(t), \quad (13.8)$$

where

$$\mathbf{w}(t) = [w_1(t), w_2(t), \dots, w_N(t)]^T \quad (13.9)$$

represents the clutter-plus-noise vector, which is assumed to be a zero mean Gaussian random vector that satisfies

$$\mathbb{E}\{\mathbf{w}(t)\mathbf{w}^H(t - \tau)\} = \mathbf{Q}\delta(\tau) \quad (13.10)$$

with \mathbf{Q} assumed to be a known² time invariant constant matrix that determines the clutter-plus-noise covariance matrix at the output of the matched filters.

2.13.2.1 Joint location and velocity estimation

Let the signals observed by a MIMO radar system be

$$\tilde{\mathbf{r}}(t) = [\tilde{r}_1(t), \tilde{r}_2(t), \dots, \tilde{r}_N(t)]^T, \quad (13.11)$$

²If the covariance matrix \mathbf{Q} is not known, one may need to estimate it using secondary data. The readers are referred to literature on space-time adaptive processing (STAP) for more details [14].

which is a realization of the random vector $\mathbf{r}(t)$ in (13.8), we hope to jointly estimate the target location and velocity in the maximum likelihood (ML) sense. As discussed in [15], the ML estimates of the unknown parameters can be found by examining the corresponding log-likelihood ratio. Stack the parameters of interest into a vector as follows:

$$\boldsymbol{\theta}_J = [x, y, v_x, v_y]^T \quad (13.12)$$

and define a bigger vector

$$\boldsymbol{\Theta}_J = [x, y, v_x, v_y, \zeta_R, \zeta_I]^T \quad (13.13)$$

to include all the unknown parameters involved. Using the signal model in (13.8), it can be derived that the log-likelihood ratio with respect to $\boldsymbol{\Theta}_J$ is given by [16]

$$L_J(\boldsymbol{\Theta}_J; \tilde{\mathbf{r}}(t)) = 2\Re \left\{ \sqrt{\frac{E}{M}} \zeta \mathbf{z}^H \mathbf{1}_{NM \times 1} \right\} - \frac{E}{M} |\zeta|^2 \mathbf{1}_{1 \times NM} \mathbf{V} \mathbf{1}_{NM \times 1} + C_1, \quad (13.14)$$

where C_1 is a constant independent of the parameters to be estimated,

$$\mathbf{V} = \int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{Q}^{-1} \mathbf{U} dt \quad (13.15)$$

and³

$$\mathbf{z} = \int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{Q}^{-1} \tilde{\mathbf{r}}(t) dt. \quad (13.16)$$

The vector \mathbf{z} can be regarded as the output of a matched filter which considers the correlations allowed in the given analysis. It can be shown that, for any value of $\boldsymbol{\theta}_J$, the ML estimate of ζ (i.e., ζ_R and ζ_I) is [16]

$$\hat{\zeta}_{ML} = \hat{\zeta}_{R_{ML}} + j \hat{\zeta}_{I_{ML}} = \frac{\mathbf{1}_{1 \times NM} \mathbf{z}}{\sqrt{\frac{E}{M}} \mathbf{1}_{1 \times NM} \mathbf{V} \mathbf{1}_{NM \times 1}}. \quad (13.17)$$

Then, substituting $\zeta = \hat{\zeta}_{ML}$ in (13.14), it yields

$$L_J(\boldsymbol{\theta}_J; \tilde{\mathbf{r}}(t), \hat{\zeta}_{ML}) = \frac{|\mathbf{1}_{1 \times NM} \mathbf{z}|^2}{\mathbf{1}_{1 \times NM} \mathbf{V} \mathbf{1}_{NM \times 1}} + C_1. \quad (13.18)$$

Note that in (13.18) we changed the notation in the parenthesis to emphasize that, after we have the ML estimate for ζ , the parameters which need to be estimated are the elements of $\boldsymbol{\theta}_J$. Thus, the ML estimate of the unknown parameter vector $\boldsymbol{\theta}_J$ can be expressed as

$$\hat{\boldsymbol{\theta}}_{J,ML} = \arg \max_{\boldsymbol{\theta}_J} L_J(\boldsymbol{\theta}_J; \tilde{\mathbf{r}}(t), \hat{\zeta}_{ML}) = \arg \max_{\boldsymbol{\theta}_J} \frac{|\mathbf{1}_{1 \times NM} \mathbf{z}|^2}{\mathbf{1}_{1 \times NM} \mathbf{V} \mathbf{1}_{NM \times 1}}. \quad (13.19)$$

³Note that in this chapter, the integral over time in equations, such as (13.15) and (13.16) operates on each element of the corresponding matrix.

It is worth noting that the estimator (13.19) combines signals from different antennas in a coherent way, but we skip the discussion for now and postpone the explanation after (13.22) until some simplifying assumptions are introduced.

In the previous discussions, the clutter-plus-noise was assumed to be temporally white but possibly spatially colored. In order to simplify the analysis, in the following, we introduce two additional assumptions: orthogonal transmitted signals and spatially white clutter-plus-noise. Leveraging these assumptions, we are able to simplify the problem to be tackled and provide a handful of analytical results, which can not only characterize some typical behaviors of the system when these assumptions are satisfied, but can as well render insight into the system behaviors for cases where these assumptions do not hold. Since these analytical results are relatively easy to obtain and explain, we are also able to shed light on the relationship between the system performance and a few important system parameters. Due to the convenience afforded by these assumptions, we will use them repeatedly in the rest of this chapter.

Assumption 1. Assume the transmitted signals are approximately orthogonal

$$\int_{-\infty}^{\infty} s_k(t)s_{k'}^*(t)dt \approx \begin{cases} 1, & \text{if } k = k', \\ 0, & \text{if } k \neq k', \end{cases}$$

and maintain approximate orthogonality for time delays τ_k , $\tau_{k'}$ and Doppler shifts f_{dk} , $f_{dk'}$ of interest so that

$$\int_{-\infty}^{\infty} s_k(t - \tau_k)s_{k'}^*(t - \tau_{k'})e^{j2\pi(f_{dk} - f_{dk'})t} dt \approx \begin{cases} 1, & k = k', \\ 0, & k \neq k', \end{cases}$$

such that the signals contributed from different transmitters can be separated at each receiver. See Appendix A.1 for further discussion.

Assumption 2. The clutter-plus-noise corresponding to the lk th path $w_{lk}(t)$ is a temporally white zero-mean complex Gaussian random process with $\mathbb{E}\{w_{lk}(t)w_{lk}^*(u)\} = \sigma_w^2\delta(t - u)$, where σ_w is a constant and $\delta(t)$ is a unit impulse function. The clutter-plus-noise components are spatially white, such that $\mathbb{E}\{w_{lk}(t)w_{l'k'}^*(u)\} = 0$ if $l \neq l'$ or $k \neq k'$.

Under Assumption 1, it can be obtained that $\int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{U} dt = \mathbf{I}$. Assumption 2 leads to $\mathbf{Q} = \sigma_w^2 \mathbf{I}$, which can be regarded as a result of perfect whitening if the covariance matrix of $\mathbf{w}(t)$, called \mathbf{Q} previously, can be accurately estimated. Then, we have $\mathbf{z} = \int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{r}(t)dt$ and $\mathbf{V} = (1/\sigma_w^2)\mathbf{I}$.

Applying these results into (13.14), we get a simplified log-likelihood ratio

$$\begin{aligned} L_J(\Theta_J; \tilde{\mathbf{r}}(t)) &= \frac{2}{\sigma_w^2} \sqrt{\frac{E}{M}} \Re \left\{ \zeta \sum_{l=1}^N \sum_{k=1}^M \int_{-\infty}^{\infty} r_{lk}^*(t) s_k(t - \tau_{lk}) e^{-j2\pi f_c \tau_{lk}} e^{j2\pi f_{lk} t} dt \right\} \\ &\quad - \frac{1}{\sigma_w^2} \frac{E}{M} |\zeta|^2 NM + C_1, \end{aligned} \tag{13.20}$$

where $\tilde{r}_{lk}(t)$ represents the observed signal corresponding to the received signal model for the lk th path

$$r_{lk}(t) = \sqrt{\frac{E}{M}} \zeta s_k(t - \tau_{lk}) e^{-j2\pi f_c \tau_{lk}} e^{j2\pi f_{lk} t} + w_l(t). \tag{13.21}$$

Likewise, applying the simplification assumption to (13.19) gives the simplified ML estimate

$$\hat{\boldsymbol{\theta}}_{J,ML}^{coh} = \arg \max_{\boldsymbol{\theta}_J} \left| \sum_{l=1}^N \sum_{k=1}^M e^{j2\pi f_c \tau_{lk}} \int_{-\infty}^{\infty} \tilde{r}_{lk}(t) s_k^*(t - \tau_{lk}) e^{-j2\pi f_{lk} t} dt \right|^2. \quad (13.22)$$

In the estimator in (13.22), phase shifts imposed on various paths have a measurable impact on the estimation output via the term $e^{j2\pi f_c \tau_{lk}}$. Intuitively, we want the terms inside the $|\cdot|^2$ to add together in phase to maximize (13.22), so as to achieve high estimation performance. This is the motivation for the *coherent processing* (see Appendix A.2 for a distinction between coherent and noncoherent processing). To make the best use of the phase information, phase synchronization is required, so that all transmitters and receivers employ a common phase reference. Some approaches for achieving phase synchronization are described in Section 2.13.6. Note that under *Assumptions 1 and 2*, the coherent processing is optimal for the coherent target reflection model in Figure 13.1.

The mean square error (MSE), the average squared difference between an estimate and the true value of the parameter being estimated, is a useful metric for forecasting the performance of an estimator. The MSE of any unbiased estimator is lower bounded by the Cramer-Rao bound (CRB). Since attaining the MSE is often computationally expensive, the CRB that indicates the best MSE an estimator can provide, serves as an important tool for evaluating the estimation performance and system configuration. In (13.22), we have derived the ML estimator for the coherent MIMO radar joint location and velocity estimation problem. It is known that the ML estimator is asymptotically unbiased, and the MSE of the ML estimate asymptotically approaches the CRB. Thus, we can derive the CRBs for the parameters of interest to provide an approximate MSE of the corresponding ML estimates in the asymptotic region. The Fisher information matrix (FIM) with respect to $\boldsymbol{\Theta}_J$ can be derived from the log-likelihood function in (13.20) as follows [15, 17]

$$\begin{aligned} \mathbf{J}_{coh}(\boldsymbol{\Theta}_J) &= \mathbb{E} \left\{ \nabla_{\boldsymbol{\Theta}_J} L_J(\boldsymbol{\Theta}_J; \tilde{\mathbf{r}}(t)) [\nabla_{\boldsymbol{\Theta}_J} L_J(\boldsymbol{\Theta}_J; \tilde{\mathbf{r}}(t))]^T \right\} \\ &= -\mathbb{E} \left\{ \nabla_{\boldsymbol{\Theta}_J} [\nabla_{\boldsymbol{\Theta}_J} L_J(\boldsymbol{\Theta}_J; \tilde{\mathbf{r}}(t))]^T \right\}. \end{aligned}$$

Under *Assumptions 1 and 2*, according to the derivations provided in Appendix A.3, we obtain the FIM as shown below:

$$\mathbf{J}_{coh}(\boldsymbol{\Theta}_J) = \frac{8\pi^2 |\zeta|^2 E}{\sigma_w^2 M} \sum_{k=1}^M \sum_{l=1}^N \begin{bmatrix} \mathbf{J}_{coh}^{UL} & \mathbf{J}_{coh}^{UR} \\ \mathbf{J}_{coh}^{LL} & \mathbf{J}_{coh}^{LR} \end{bmatrix}, \quad (13.23)$$

where

$$\mathbf{J}_{coh}^{UL} = \begin{bmatrix} \varepsilon_k^c a_{lk}^2 + 2\gamma_{lk}^c a_{lk} e_{lk} + \eta_{lk}^c e_{lk}^2 & \begin{pmatrix} (\varepsilon_k^c a_{lk} + \gamma_{lk}^c e_{lk}) b_{lk} + \\ (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) g_{lk} \end{pmatrix} & (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) \beta_{lk} & (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) q_{lk} \\ \begin{pmatrix} (\varepsilon_k^c a_{lk} + \gamma_{lk}^c e_{lk}) b_{lk} + \\ (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) g_{lk} \end{pmatrix} & \varepsilon_k^c b_{lk}^2 + 2\gamma_{lk}^c b_{lk} g_{lk} + \eta_{lk}^c g_{lk}^2 & (\gamma_{lk}^c b_{lk} + \eta_{lk}^c g_{lk}) \beta_{lk} & (\gamma_{lk}^c b_{lk} + \eta_{lk}^c g_{lk}) q_{lk} \\ (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) \beta_{lk} & (\gamma_{lk}^c b_{lk} + \eta_{lk}^c g_{lk}) \beta_{lk} & \eta_{lk}^c \beta_{lk}^2 & \eta_{lk}^c \beta_{lk} q_{lk} \\ (\gamma_{lk}^c a_{lk} + \eta_{lk}^c e_{lk}) q_{lk} & (\gamma_{lk}^c b_{lk} + \eta_{lk}^c g_{lk}) q_{lk} & \eta_{lk}^c \beta_{lk} q_{lk} & \eta_{lk}^c q_{lk}^2 \end{bmatrix},$$

$$\mathbf{J}_{\text{coh}}^{\text{UR}} = \begin{bmatrix} \frac{(a_{lk}\alpha_k - e_{lk}\Xi_{lk})\zeta_I}{2\pi|\zeta|^2} & \frac{(e_{lk}\Xi_{lk} - a_{lk}\alpha_k)\zeta_R}{2\pi|\zeta|^2} \\ \frac{(b_{lk}\alpha_k - g_{lk}\Xi_{lk})\zeta_I}{2\pi|\zeta|^2} & \frac{(g_{lk}\Xi_{lk} - b_{lk}\alpha_k)\zeta_R}{2\pi|\zeta|^2} \\ \frac{-\beta_{lk}\Xi_{lk}\zeta_I}{2\pi|\zeta|^2} & \frac{\beta_{lk}\Xi_{lk}\zeta_R}{2\pi|\zeta|^2} \\ \frac{-q_{lk}\Xi_{lk}\zeta_I}{2\pi|\zeta|^2} & \frac{q_{lk}\Xi_{lk}\zeta_R}{2\pi|\zeta|^2} \end{bmatrix},$$

$$\mathbf{J}_{\text{coh}}^{\text{LL}} = \begin{bmatrix} \frac{(a_{lk}\alpha_k - e_{lk}\Xi_{lk})\zeta_I}{2\pi|\zeta|^2} & \frac{(b_{lk}\alpha_k - g_{lk}\Xi_{lk})\zeta_I}{2\pi|\zeta|^2} & \frac{-\beta_{lk}\Xi_{lk}\zeta_I}{2\pi|\zeta|^2} & \frac{-q_{lk}\Xi_{lk}\zeta_I}{2\pi|\zeta|^2} \\ \frac{(e_{lk}\Xi_{lk} - a_{lk}\alpha_k)\zeta_R}{2\pi|\zeta|^2} & \frac{(g_{lk}\Xi_{lk} - b_{lk}\alpha_k)\zeta_R}{2\pi|\zeta|^2} & \frac{\beta_{lk}\Xi_{lk}\zeta_R}{2\pi|\zeta|^2} & \frac{q_{lk}\Xi_{lk}\zeta_R}{2\pi|\zeta|^2} \end{bmatrix},$$

$$\mathbf{J}_{\text{coh}}^{\text{LR}} = \begin{bmatrix} \frac{1}{4\pi^2|\zeta|^2} & 0 \\ 0 & \frac{1}{4\pi^2|\zeta|^2} \end{bmatrix},$$

with the terms $a_{lk} = \partial\tau_{lk}/\partial x$, $b_{lk} = \partial\tau_{lk}/\partial y$, $e_{lk} = \partial f_{lk}/\partial x$, $g_{lk} = \partial f_{lk}/\partial y$, $\beta_{lk} = \partial f_{lk}/\partial v_x$, $q_{lk} = \partial f_{lk}/\partial v_y$ determined by the target position and velocity, and the antenna positions. The terms ε_k^c , γ_{lk}^c , η_{lk}^c , α_k , and Ξ_{lk} in $\mathbf{J}_{\text{coh}}(\Theta_J)$ are dependent on the characteristics of the received waveforms

$$\varepsilon_k^c = \int_{-\infty}^{\infty} f^2 |S_k(f)|^2 df - 2f_c \int_{-\infty}^{\infty} f |S_k(f)|^2 df + f_c^2, \quad (13.24)$$

$$\gamma_{lk}^c = \frac{1}{2\pi} \Im \left\{ \int_{-\infty}^{\infty} ts_k^*(t - \tau_{lk}) \frac{\partial s_k(t - \tau_{lk})}{\partial \tau_{lk}} dt \right\} - f_c \int_{-\infty}^{\infty} t |s_k(t - \tau_{lk})|^2 dt, \quad (13.25)$$

$$\eta_{lk}^c = \int_{-\infty}^{\infty} t^2 |s_k(t - \tau_{lk})|^2 dt, \quad (13.26)$$

$$\alpha_k = f_c - \int_{-\infty}^{\infty} f |S_k(f)|^2 df, \quad (13.27)$$

and

$$\Xi_{lk} = \int_{-\infty}^{\infty} t |s_k(t - \tau_{lk})|^2 dt, \quad (13.28)$$

where $S_k(f)$ represents the Fourier transform of $s_k(t)$. The CRBs for the estimates of the unknown target locations and velocities are the first four diagonal elements of the inverse of the FIM

$$\text{CRB}_x^{\text{coh}} = \left[\mathbf{J}_{\text{coh}}^{-1}(\Theta_J) \right]_{1,1}, \quad \text{CRB}_y^{\text{coh}} = \left[\mathbf{J}_{\text{coh}}^{-1}(\Theta_J) \right]_{2,2},$$

$$\text{CRB}_{v_x}^{\text{coh}} = \left[\mathbf{J}_{\text{coh}}^{-1}(\Theta_J) \right]_{3,3}, \quad \text{CRB}_{v_y}^{\text{coh}} = \left[\mathbf{J}_{\text{coh}}^{-1}(\Theta_J) \right]_{4,4}. \quad (13.29)$$

The CRBs can be used to optimize system configurations with respect to certain parameters. Collect the parameters of interest (β_1, β_2, \dots) in a vector as $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots]^T$. These parameters could be, for example, the number of antennas, the antenna placement, the waveform parameters, and so forth. Suppose we require the optimum system configuration design to minimize a metric formed by the weighted sum of the CRBs of the estimates of x , y , v_x , and v_y . Then, the value of $\boldsymbol{\beta}$ that optimizes this metric can be expressed as

$$\tilde{\boldsymbol{\beta}} = \arg \left\{ \min_{\boldsymbol{\beta} \in \Phi_{\boldsymbol{\beta}}} \left(W_1 \text{CRB}_x^{\text{coh}} + W_2 \text{CRB}_y^{\text{coh}} + W_3 \text{CRB}_{v_x}^{\text{coh}} + W_4 \text{CRB}_{v_y}^{\text{coh}} \right) \right\}, \quad (13.30)$$

where $W_i, i = 1, \dots, 4$ are weighting factors and $\Phi_{\boldsymbol{\beta}}$ represents the feasible set of $\boldsymbol{\beta}$. Numerical methods may be needed to solve this problem. A concrete example of finding the optimum antenna placement for the MIMO radar velocity estimation is provided in [18], where only $\text{CRB}_{v_x}^{\text{coh}}$ and $\text{CRB}_{v_y}^{\text{coh}}$ are involved and equal weighting is considered, and it is shown analytically that under certain conditions only symmetrical placement can be optimum. Interested readers are referred to [18] for more details.

2.13.2.2 Coherent ambiguity function

Besides the CRB, the ambiguity function (AF) is another useful tool for evaluating the estimation performance for radar systems. The performance of the ML estimate can be reflected by the shape of the AF. Actually, the CRB describes the shape of the AF around its maximum and this information influences the MSE in the high SCNR region or when MN is large. Other aspects of the AF, including the existence of sidelobes which can not be captured by the CRB, also influence the MSE performance in the low SCNR region when MN is not large.

Under Assumptions 1 and 2, we develop the AF for coherent MIMO radar system. Assume a stationary target is present at the origin of the Cartesian coordinate system, causing time delay τ_{0lk} and zero Doppler shift to the signal transmitted over the lk th path. Thus, the clutter-plus-noise free received signal can be expressed as

$$r_{0lk}(t) = \sqrt{\frac{E}{M}} \zeta s_k(t - \tau_{0lk}) e^{-j2\pi f_c \tau_{0lk}}, \quad (13.31)$$

where τ_{0lk} can be obtained by substituting $x = 0$ and $y = 0$ into (13.2). Consider the log-likelihood ratio in (13.20), which assumes the expected signals have arbitrary delay τ_{lk} and Doppler shift f_{lk} for the lk th path, corresponding to a target at location (x, y) with velocity (v_x, v_y) . Substituting the just described clutter-plus-noise free received signals $r_{0lk}(t)$ for $\tilde{r}_{lk}(t)$ in (13.20), we obtain

$$L_J(\Theta_J; \tilde{\mathbf{r}}(t)) = \frac{2E|\zeta|^2}{\sigma_w^2 M} \Re \left\{ \sum_{l=1}^N \sum_{k=1}^M \int_{-\infty}^{\infty} s_k^*(t) s_k(t - \tau_{\Delta lk}) e^{-j2\pi f_c \tau_{\Delta lk}} e^{j2\pi f_{lk}(t + \tau_{0lk})} dt \right\} + \tilde{C}_1, \quad (13.32)$$

where $\tau_{\Delta lk} = \tau_{lk} - \tau_{0lk}$ denotes the time difference for the lk th path. The quantity $L_J(\Theta_J; \tilde{\mathbf{r}}(t))$ in (13.32) measures the likelihood that a ML processor believes the target corresponds to path delays τ_{lk} and Doppler shifts f_{lk} for $k = 1, \dots, M$ and $l = 1, \dots, N$, when the actual path delays are τ_{0lk} and the actual Doppler shifts are zero. Of course, the parts of (13.32) that do not depend on the delays and Doppler shifts are not important in this consideration. Here we lump them into the constant \tilde{C}_1 . Further,

one can define any suitable normalized version of (13.32) to be the AF. In this way, it is clear that the AF is closely related to the performance of a processor that computes ML estimates.⁴ Thus, let us define the coherent AF as

$$\chi^{\text{coh}}(x, y, v_x, v_y) = \frac{1}{(MN)^2} \left| \sum_{l=1}^N \sum_{k=1}^M e^{-j2\pi f_c \tau_{\Delta lk}} \int_{-\infty}^{\infty} s_k^*(t) s_k(t - \tau_{\Delta lk}) e^{j2\pi f_{lk}(t+\tau_{0lk})} dt \right|^2, \quad (13.33)$$

where we have replaced the $\Re\{\cdot\}$ with $|\cdot|^2$ to define the coherent MIMO radar AF following the lead of Woodward [19], $1/(MN)^2$ is introduced as a normalization factor, and the $\tau_{\Delta lk}$ and f_{lk} in (13.33) are functions of (x, y) and (x, y, v_x, v_y) , respectively. An ideal AF has a single peak at $(0, 0, 0, 0)$ and is zero elsewhere, which is however, impossible to realize. In the real world, the AF always comes with a non-zero width mainlobe and several sidelobes, and researchers endeavor to design a better waveform/system with a narrow mainlobe and lower sidelobe peaks.

A coherent MIMO radar AF is studied in [20] and excellent properties are found through numerical investigations.⁵ For some cases with a small number of antennas, advantages are discussed over noncoherent MIMO radar for target location estimation. The coherent AF provided in [20] can be considered as a special case of (13.33) by letting $\tau_{\Delta lk} = \tau_{lk}$ and $f_{lk} = 0$, such that the resulting coherent AF is reduced to a function of (x, y) . We refer interested readers to [20] for more details.

2.13.2.3 Localization with phase errors

The previous section talks about location and velocity estimation for MIMO radar with ideal coherent processing, where the target reflections follow the coherent model shown in Figure 13.1 and the phase is assumed to be perfectly aligned across sensors. However, the difficulty in realizing perfect phase synchronization may bring problems for coherent MIMO radar. In this section, we investigate the impact of phase errors (e.g., due to imperfect phase synchronization) on the estimation performance. The focus will be on localization, so here we do not discuss the velocity estimation for simplicity. Assuming frequency synchronization, possibly through reception of a beacon, and white clutter-plus-noise, possibly due to estimating the covariance matrix and whitening the observations, we study the impact of phase errors on the target localization performance for coherent MIMO radar with widely dispersed antennas. We consider cases with sufficiently high SCNR such that the CRB provides accurate performance estimates. The CRBs with phase errors are computed in a few example cases and compared with the CRBs without phase errors. For these examples, using numerical results, we will show that at sufficiently high SCNR, phase errors degrade performance by only a relatively small amount.

Let Ω_k^t and Ω_l^r denote the phase errors induced by the k th transmitter and the l th receiver, respectively. Assume the phase errors are static (during the entire CPI) i.i.d. random variables with uniform distribution $\Omega_k^t \sim U(-\Delta\pi, \Delta\pi)$ and $\Omega_l^r \sim U(-\Delta\pi, \Delta\pi)$, where $0 \leq \Delta \leq 1$. Using (13.4) and further

⁴To be more precise, the relationship is built upon the facts that the CRB depends on the mainlobe of the AF, and that the performance of the ML estimator tends to the CRB.

⁵Note that some of the open literature [21, 22] may have different definitions for the MIMO radar AF.

taking account of the phase errors, the received signal model at the l th receiver can be expressed as

$$r_l(t) = \sqrt{\frac{E}{M}} \zeta e^{j\Omega_l^t} \sum_{k=1}^M e^{j\Omega_k^t} s_k(t - \tau_{lk}) e^{-j2\pi f_c \tau_{lk}} + w_l(t), \quad (13.34)$$

where $\zeta = \zeta_R + j\zeta_I$ denotes the deterministic, unknown complex reflection coefficient, $s_k(t)$ is the transmitted signal at the k th transmitter, E is the total transmitted energy, and $\tau_{lk} = (d_k^t + d_l^r)/c$ denotes the time delay between transmitter k and receiver l as per (13.2). The clutter-plus-noise $w_l(t)$ in (13.34) is assumed to be white complex Gaussian with power spectral density (PSD) σ_w^2 , which is assumed to be independent for different l .

Collect parameters of interest in a vector as follows

$$\boldsymbol{\theta}_P = [x, y, \zeta_R, \zeta_I]^T, \quad (13.35)$$

which is assumed to be deterministic and unknown. The random phase errors, regarded as nuisance parameters, are collected in an $(N + M)$ vector

$$\boldsymbol{\Omega} = [\Omega_1^t, \dots, \Omega_M^t, \Omega_1^r, \dots, \Omega_N^r]^T. \quad (13.36)$$

Based on the results in [23], it is easy to show that the likelihood ratio conditioned on the phase errors is

$$\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t) | \boldsymbol{\Omega}) = C_2 \exp \{g\} \quad (13.37)$$

with g defined as

$$g = \frac{-1}{\sigma_w^2} \sum_{l=1}^N \int_{\mathcal{T}} \left| \tilde{r}_l(t) - \zeta \sqrt{\frac{E}{M}} \sum_{k=1}^M e^{j\Omega_k^t} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) \right|^2 dt, \quad (13.38)$$

where $\tilde{\mathbf{r}}(t) = [\tilde{r}_1(t), \tilde{r}_2(t), \dots, \tilde{r}_N(t)]^T$ represents the signal observed over the observation interval \mathcal{T} , C_2 is a constant not dependent on $\boldsymbol{\theta}_P$, and $\Omega_{lk} = \Omega_l^r + \Omega_k^t$. If we remove the conditioning on the phase errors by averaging them out, the likelihood ratio can be written as

$$\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t)) = C_2 \int_{-\Delta\pi}^{\Delta\pi} \cdots \int_{-\Delta\pi}^{\Delta\pi} \exp \{g\} \left(\frac{1}{2\Delta\pi} \right)^{N+M} d\Omega_1^t \cdots d\Omega_N^r \quad (13.39)$$

which requires an $(N + M)$ dimensional multiple integration.

Assume an estimator is employed to produce an unbiased estimate, $\hat{\boldsymbol{\theta}}_P$, of the parameters of interest from the observed signal vector $\tilde{\mathbf{r}}(t)$. Then the variances of these estimates are bounded from below by the CRBs which are the diagonal elements of the inverse of the FIM. Following the derivation given in [23] for a similar single antenna system without phase errors, the FIM is given by

$$\mathbf{J}(\boldsymbol{\theta}_P) = -\mathbb{E} \left\{ \nabla_{\boldsymbol{\theta}_P} [\nabla_{\boldsymbol{\theta}_P} \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))]^T \right\}. \quad (13.40)$$

Since $\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))$ in (13.39) is explicitly a function of the time delays τ_{lk} , we introduce an alternative parameter vector

$$\boldsymbol{\vartheta}_P = [\tau_{11}, \tau_{12}, \dots, \tau_{NM}, \zeta_R, \zeta_I]^T, \quad (13.41)$$

which is a column vector with $(NM + 2)$ elements containing the time delays for every transmit-receive path and the real and imaginary parts of the target complex reflectivity. Using the chain rule, we have

$$\mathbf{J}(\boldsymbol{\theta}_P) = \left(\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T \right) \mathbf{J}(\boldsymbol{\vartheta}_P) \left(\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T \right)^T. \quad (13.42)$$

The computations of $\mathbf{J}(\boldsymbol{\vartheta}_P)$ and $\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T$ are provided in Appendix A.4. Substituting $\mathbf{J}(\boldsymbol{\vartheta}_P)$ and $\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T$ into (13.42) yields the expression for $\mathbf{J}(\boldsymbol{\theta}_P)$. Thus, for a MIMO radar with M transmit and N receive antennas, assuming white complex Gaussian clutter-plus-noise and that the phase errors at the transmitters and receivers are static (during the entire CPI) i.i.d. random variables with uniform distribution in $[-\Delta\pi, \Delta\pi]$, $0 \leq \Delta \leq 1$, the CRB for the estimates of the target location $X = (x, y)$ are the first two diagonal elements of the inverse of the FIM

$$\text{CRB}_x = \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_P) \right]_{1,1}, \quad \text{CRB}_y = \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_P) \right]_{2,2}.$$

Example

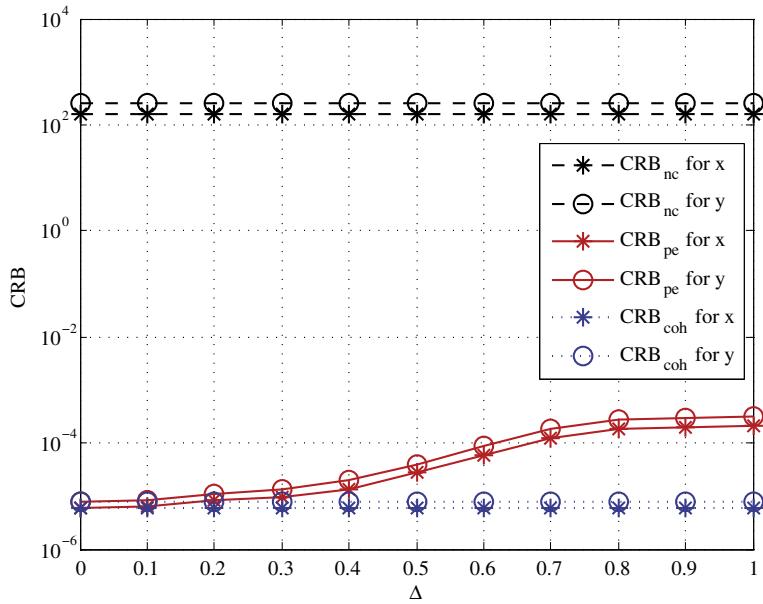
Assume a target is present at $(15, 17.5)$ m. Consider a MIMO radar with $M = 2$ transmitters located at $(-137.87, -24.31)$ km and $(129.33, 129.33)$ km, and $N = 2$ receivers located at $(-31.26, -177.27)$ km and $(149.82, 86.50)$ km. The transmitted signals are orthogonal frequency spread signals [18]

$$s_k(t) = \begin{cases} \frac{1}{\sqrt{T}} \exp(j2\pi kf_\Delta t) & 0 < t < T, \\ 0 & \text{otherwise,} \end{cases} \quad (13.43)$$

where $k = 1, 2$, the term T denotes the pulse duration, and $f_\Delta = f_{k+1} - f_k \geq 0$ is the frequency increment between $s_k(t)$ and $s_{k+1}(t)$. A single pulse is employed in the simulation. Assume the pulse duration is 1 ms, the frequency increment is 0.1 MHz, and the carrier frequency is 1 GHz. The SCNR, defined as $E|\zeta|^2/(M\sigma_w^2)$, is fixed at 20 dB. Letting Δ go from 0 to 1, we plot the CRBs with phase errors (CRB_{pe}) for the estimates of x and y in Figure 13.2. Note that the value of Δ describes the severity of the phase errors, which are distributed uniformly in $[-\Delta\pi, \Delta\pi]$. The corresponding coherent (CRB_{coh}) and noncoherent (CRB_{nc}) CRBs without phase errors were plotted in the same figure using the results in [20].

The figure indicates that a certain amount of phase error is tolerable if we are willing to accept a certain amount of loss. For a given acceptable loss, Figure 13.2 gives the tolerable phase errors. It is observed that CRB_{pe} approaches CRB_{coh} as Δ goes to zero, and deviates away from CRB_{coh} as Δ increases. In the worst case ($\Delta = 1$), CRB_{pe} is approximately 40 times worse than CRB_{coh} , but is still approximately 10^6 times better than CRB_{nc} . The degradation due to the loss of knowing phase is much smaller than the degradation caused by using noncoherent processing. These results document that having noisy phase measurements is different from, and better than, not having any phase measurements.

This is a representative example (we have tested other examples and obtained similar results) for cases with small M and N . We noted that cases with large M and N will also exhibit very small degradations due to phase errors since, as shown in [24], the noncoherent estimates must be perfect as $MN \rightarrow \infty$.

**FIGURE 13.2**

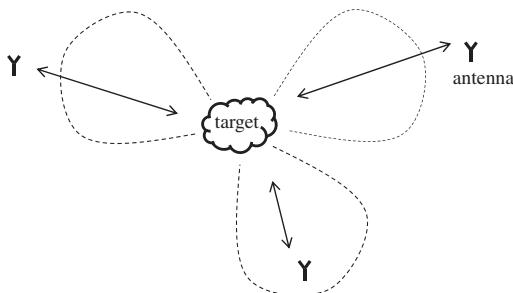
CRB for target localization versus Δ , where the phase errors are uniformly distributed in $[-\Delta\pi, \Delta\pi]$.

2.13.3 Noncoherent MIMO radar

In the previous section, we discussed coherent MIMO radar with widely separated antennas, where the target reflections follow a coherent model and coherent processing is employed. Now we define a different system, called *noncoherent MIMO radar*, employing a noncoherent processing of received signals obeying a noncoherent target reflection model. For noncoherent MIMO radar, the antennas are separated widely enough such that they fall in different target beamwidths and the effective target reflection coefficients for different paths are *distinct* (which implies a noncoherent target reflection model). An example of such a case is illustrated in Figure 13.3. Denote the complex Gaussian reflection coefficient for the lk th path by ζ_{lk} . The received signal at receive antenna l is modeled as

$$\begin{aligned} r_l(t) &= \sqrt{\frac{E}{M}} \sum_{k=1}^M \zeta_{lk} e^{j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) e^{j2\pi f_{lk} t} + w_l(t) \\ &= \sqrt{\frac{E}{M}} \sum_{k=1}^M \zeta_{lk} s_k(t - \tau_{lk}) e^{j2\pi f_{lk} t} + w_l(t), \end{aligned} \quad (13.44)$$

where $w_l(t)$ denotes the noise at receiver l . The simplification in going from the first to the second line in (13.44) follows since the distributions of ζ_{lk} and $\zeta_{lk} e^{j2\pi f_c \tau_{lk}}$ are identical. The vector that contains

**FIGURE 13.3**

A noncoherent MIMO radar system with three antennas each lying in a different target beamwidth.

all the received signals can be expressed as

$$\mathbf{r}(t) = [r_1(t), r_2(t), \dots, r_N(t)]^T = \sqrt{\frac{E}{M}} \mathbf{U} \boldsymbol{\xi} + \mathbf{w}(t), \quad (13.45)$$

where the clutter-plus-noise vector $\mathbf{w}(t) = [w_1(t), w_2(t), \dots, w_N(t)]^T$ is assumed to be zero mean Gaussian and satisfy $\mathbb{E}\{\mathbf{w}(t)\mathbf{w}^H(t-\tau)\} = \mathbf{Q}\delta(\tau)$. The reflection coefficient vector

$$\boldsymbol{\xi} = [\xi_{11}, \xi_{12}, \dots, \xi_{NM}]^T \quad (13.46)$$

is a complex Gaussian random vector, where $\mathbb{E}\{\boldsymbol{\xi}\boldsymbol{\xi}^H\} = \mathbf{R}$. Assume the matrices \mathbf{Q} and \mathbf{R} are both known, possibly from the pre-processing stage (e.g., the detection stage) or an adaptive procedure.⁶ For notational simplicity, we further assume that $\boldsymbol{\xi}$ has zero mean. The matrix \mathbf{U} in (13.45) has the same form as (13.7) with $\mathbf{u}_l = [u_{l1}, u_{l2}, \dots, u_{lM}]^T$. However, in this case, the u_{lk} has a different definition, which is given by

$$u_{lk} = s_k(t - \tau_{lk}) e^{j2\pi f_{lk} t}.$$

2.13.3.1 Joint location and velocity estimation

Similar to the discussion in Section 2.13.2.1, suppose the observed signal vector $\tilde{\mathbf{r}}(t)$, a realization of the $\mathbf{r}(t)$ in (13.45), is available. Our goal is to obtain the ML estimates of the target location and velocity based on the knowledge of $\tilde{\mathbf{r}}(t)$. Using the signal model in (13.45), the likelihood ratio with respect to $\boldsymbol{\theta}_J = [x, y, v_x, v_y]^T$ conditioned on the reflection coefficient $\boldsymbol{\xi}$ can be derived

$$\Lambda_J(\boldsymbol{\theta}_J; \tilde{\mathbf{r}}(t) | \boldsymbol{\xi}) = C'_1 \exp \left\{ \sqrt{\frac{E}{M}} \mathbf{z}^H \boldsymbol{\xi} + \sqrt{\frac{E}{M}} \boldsymbol{\xi}^H \mathbf{z} - \frac{E}{M} \boldsymbol{\xi}^H \mathbf{V} \boldsymbol{\xi} \right\}, \quad (13.47)$$

⁶ Assume the target has already been detected using a radar with a fairly coarse grid so we know the distances approximately, and one can develop specific relative values for variances for the reflection coefficients for different paths from transmit to receive antennas based on this.

where C'_1 is a constant not dependent on θ_J . In (13.47), $\mathbf{V} = \int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{Q}^{-1} \mathbf{U} dt$ and $\mathbf{z} = \int_{-\infty}^{\infty} \mathbf{U}^H \mathbf{Q}^{-1} \tilde{\mathbf{r}}(t) dt$, where the integral over time operates on each element of the corresponding matrix. Employing the probability density function (pdf) of the reflection coefficient

$$p_\zeta(\zeta) = \frac{1}{\pi^{NM} \det \mathbf{R}} \exp \left\{ -\zeta^H \mathbf{R}^{-1} \zeta \right\} \quad (13.48)$$

to average ζ out, we find the likelihood ratio as a function of the unknown parameter θ_J

$$\begin{aligned} \Lambda_J(\theta_J; \tilde{\mathbf{r}}(t)) &= \int_{\zeta} \Lambda_J(\theta_J; \tilde{\mathbf{r}}(t)|\zeta) p_\zeta(\zeta) d\zeta \\ &= C'_2 \det(\mathbf{X}) \exp \left\{ \frac{E}{M} \mathbf{z}^H \mathbf{X}^{-1} \mathbf{z} \right\}, \end{aligned} \quad (13.49)$$

where C'_2 is a constant not dependent on θ_J , \int_{ζ} implies the multidimensional integral over all possible values of ζ , and

$$\mathbf{X} = \frac{E}{M} \mathbf{V} + \mathbf{R}^{-1}. \quad (13.50)$$

In the calculations, we have assumed that the matrices \mathbf{Q} , \mathbf{R} , and \mathbf{X} have full rank, and are hence invertible. Note that the invertibility of \mathbf{R} is reasonable for MIMO radar with widely spread antennas and a target composed of a large number of scatters [5]. Further, due to thermal noise, \mathbf{Q} is also likely a full-rank matrix. From (13.49), the log-likelihood ratio can be expressed as

$$\begin{aligned} L_J(\theta_J; \tilde{\mathbf{r}}(t)) &= \ln \Lambda_J(\theta_J; \tilde{\mathbf{r}}(t)) \\ &= \frac{E}{M} \mathbf{z}^H \mathbf{X}^{-1} \mathbf{z} + \ln \det(\mathbf{X}) + \ln C'_2. \end{aligned} \quad (13.51)$$

Thus, the ML estimate of the unknown parameter vector θ_J is

$$\begin{aligned} \hat{\theta}_{J,ML} &= \arg \max_{\theta_J} L_J(\theta_J; \tilde{\mathbf{r}}(t)) \\ &= \arg \max_{\theta_J} \left\{ \frac{E}{M} \mathbf{z}^H \mathbf{X}^{-1} \mathbf{z} + \ln \det(\mathbf{X}) \right\}. \end{aligned} \quad (13.52)$$

Next we employ some simplifying assumptions,⁷ which are quite reasonable for cases with widely spaced antennas, which simplify matters.

Assumption 3. Assume that the reflection coefficients for different paths ζ_{lk} , $l = 1, \dots, N$, $k = 1, \dots, M$ are independent, such that the covariance matrix of $\zeta = [\zeta_{11}, \zeta_{12}, \dots, \zeta_{NM}]^T$ is diagonal, i.e.,

$$\mathbf{R} = \mathbb{E} \left\{ \zeta \zeta^H \right\} = \text{diag} \left\{ \sigma_{11}^2, \sigma_{12}^2, \dots, \sigma_{NM}^2 \right\}, \quad (13.53)$$

where σ_{lk}^2 represents the variance of ζ_{lk} .

⁷In [24] the authors have simulated a few cases where such assumptions do not exist and found similar results.

Consider the case where *Assumptions 1–3* hold true. Thus, $\mathbf{Q} = \sigma_w^2 \mathbf{I}$, $\mathbf{V} = (1/\sigma_w^2) \mathbf{I}$, and the log-likelihood ratio in (13.51) is reduced to

$$L_J(\boldsymbol{\theta}_J; \tilde{\mathbf{r}}(t)) = \sum_{k=1}^M \sum_{l=1}^N L_{lk}(\boldsymbol{\theta}_J; \tilde{r}_{lk}(t)), \quad (13.54)$$

where

$$L_{lk}(\boldsymbol{\theta}_J; \tilde{r}_{lk}(t)) = \frac{\sigma_{lk}^2 E}{(\sigma_{lk}^2 E + \sigma_w^2 M) \sigma_w^2} \left| \int_{-\infty}^{\infty} \tilde{r}_{lk}(t) s_k^*(t - \tau_{lk}) e^{-j2\pi f_{lk} t} dt \right|^2 + C'_3. \quad (13.55)$$

In (13.55), C'_3 is a constant not dependent on $\boldsymbol{\theta}_J$ and $\tilde{r}_{lk}(t)$ represents the observed signal corresponding to the lk th path, which is modeled as

$$r_{lk}(t) = \sqrt{\frac{E}{M}} \zeta_{lk} s_k(t - \tau_{lk}) e^{j2\pi f_{lk} t} + w_l(t). \quad (13.56)$$

Accordingly, the ML estimate can be reduced to

$$\hat{\boldsymbol{\theta}}_{J,ML}^{\text{non}} = \arg \max_{\boldsymbol{\theta}_J} \left\{ \sum_{l=1}^M \sum_{k=1}^N \frac{\sigma_{lk}^2 E}{(\sigma_{lk}^2 E + \sigma_w^2 M) \sigma_w^2} \left| \int_{-\infty}^{\infty} \tilde{r}_{lk}(t) s_k^*(t - \tau_{lk}) e^{-j2\pi f_{lk} t} dt \right|^2 \right\}. \quad (13.57)$$

Unlike (13.22), in the estimator from (13.57), the squared magnitude is taken before the summation. Thus, these terms will always add in phase and so we do not need all transmitters and receivers to be synchronized in phase. Therefore, this is a *noncoherent processing* (across sensors). Note that under *Assumptions 1–3*, the noncoherent processing is optimal for the noncoherent target reflection model in Figure 13.3.

Having obtained the ML estimate for the joint location and velocity estimation, let us introduce a theorem that shows how the number of antennas affects the estimation performance.

Theorem 1. *For a MIMO radar equipped with M transmit and N receive antennas, under Assumptions 1–3 and assume that the observations follow the assumed model in (13.56), the ML estimate as described in (13.57) converges almost surely to the true parameter value when MN is sufficiently large*

$$\hat{\boldsymbol{\theta}}_{J,ML}^{\text{non}} \xrightarrow{a.s.} \boldsymbol{\theta}_a, \text{ as } MN \rightarrow \infty. \quad (13.58)$$

It follows that the ML estimate is asymptotically ($MN \rightarrow \infty$) unbiased and its variance must asymptotically approach the CRB as $MN \rightarrow \infty$.

This theorem indicates that increasing the number of antennas properly can improve the estimation performance of noncoherent MIMO radar. Interested readers are referred to [24] for a proof of this theorem.

To evaluate the estimation performance of noncoherent MIMO radar, in the sequel, we develop the CRBs for the joint location and velocity estimation assuming that *Assumptions 1–3* hold true. The first

step in obtaining the Cramer-Rao bound is to compute the FIM, which is a four-dimensional matrix related to the second order derivatives of the log-likelihood ratio in (13.54) [15]

$$\mathbf{J}(\boldsymbol{\theta}_J) = -\mathbb{E} \left\{ \nabla_{\boldsymbol{\theta}_J} \left[\nabla_{\boldsymbol{\theta}_J} L_J(\boldsymbol{\theta}_J; \tilde{\mathbf{r}}(t)) \right]^T \right\}.$$

After lengthy algebraic manipulations, following the steps similar to those in Section 2.13.2, the expression of the FIM can be obtained as below [24]:

$$\begin{aligned} \mathbf{J}(\boldsymbol{\theta}_J) &= \sum_{k=1}^M \sum_{l=1}^N \frac{8\pi^2 \sigma_{lk}^4 E^2}{(\sigma_{lk}^2 E + \sigma_w^2 M) \sigma_w^2 M} \\ &\quad \begin{bmatrix} \varepsilon_k a_{lk}^2 + 2\gamma_{lk} a_{lk} e_{lk} + \eta_{lk} e_{lk}^2 & (\varepsilon_k a_{lk} + \eta_{lk} e_{lk}) b_{lk} + (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) g_{lk} & (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) \beta_{lk} & (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) q_{lk} \\ (\varepsilon_k a_{lk} + \eta_{lk} e_{lk}) b_{lk} + (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) g_{lk} & \varepsilon_k b_{lk}^2 + 2\gamma_{lk} b_{lk} g_{lk} + \eta_{lk} g_{lk}^2 & (\gamma_{lk} b_{lk} + \eta_{lk} g_{lk}) \beta_{lk} & (\gamma_{lk} b_{lk} + \eta_{lk} g_{lk}) q_{lk} \\ (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) \beta_{lk} & (\gamma_{lk} b_{lk} + \eta_{lk} g_{lk}) \beta_{lk} & \eta_{lk} \beta_{lk}^2 & \eta_{lk} \beta_{lk} q_{lk} \\ (\gamma_{lk} a_{lk} + \eta_{lk} e_{lk}) q_{lk} & (\gamma_{lk} b_{lk} + \eta_{lk} g_{lk}) q_{lk} & \eta_{lk} \beta_{lk} q_{lk} & \eta_{lk} q_{lk}^2 \end{bmatrix}, \end{aligned} \quad (13.59)$$

where $a_{lk} \equiv \partial \tau_{lk} / \partial x$, $b_{lk} \equiv \partial \tau_{lk} / \partial y$, $e_{lk} \equiv \partial f_{lk} / \partial x$, $g_{lk} \equiv \partial f_{lk} / \partial y$, $\beta_{lk} \equiv \partial f_{lk} / \partial v_x$, $q_{lk} \equiv \partial f_{lk} / \partial v_y$,

$$\varepsilon_k = \int_{-\infty}^{\infty} f^2 |S_k(f)|^2 df - \left| \int_{-\infty}^{\infty} f |S_k(f)|^2 df \right|^2, \quad (13.60)$$

$$\gamma_{lk} = \frac{1}{2\pi} \Im \left\{ \int_{-\infty}^{\infty} t s_k(t - \tau_{lk}) \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} dt \right\} - \int_{-\infty}^{\infty} f |S_k(f)|^2 df \int_{-\infty}^{\infty} t |s_k(t - \tau_{lk})|^2 dt, \quad (13.61)$$

and

$$\eta_{lk} = \int_{-\infty}^{\infty} t^2 |s_k(t - \tau_{lk})|^2 dt - \left| \int_{-\infty}^{\infty} t |s_k(t - \tau_{lk})|^2 dt \right|^2, \quad (13.62)$$

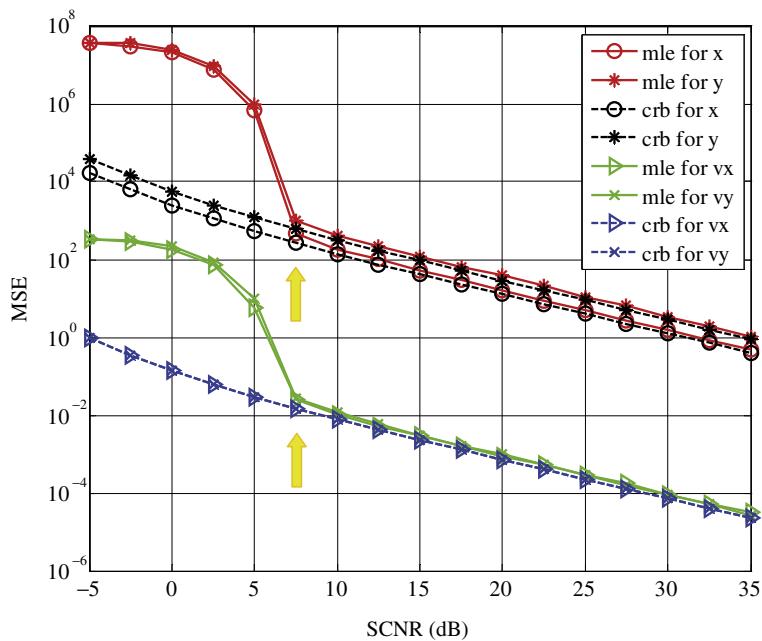
with $S_k(f)$ representing the Fourier transform of $s_k(t)$. Thus, the Cramer-Rao bounds for the estimates of the unknown parameters can be determined by the diagonal elements of the inverse of the FIM as follows:

$$\begin{aligned} \text{CRB}_x^{\text{non}} &= \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_J) \right]_{1,1}, & \text{CRB}_y^{\text{non}} &= \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_J) \right]_{2,2}, \\ \text{CRB}_{v_x}^{\text{non}} &= \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_J) \right]_{3,3}, & \text{CRB}_{v_y}^{\text{non}} &= \left[\mathbf{J}^{-1}(\boldsymbol{\theta}_J) \right]_{4,4}. \end{aligned}$$

Note that for any non-singular FIM, a closed-form expression for the Cramer-Rao bound can be easily obtained, since the analytical form of $\mathbf{J}^{-1}(\boldsymbol{\theta}_J)$ can be derived from (13.59) using Cramer's rule.

Example

Consider a noncoherent MIMO radar that has M transmitters and N receivers. The distance between each antenna and the origin is 7000 m. If the angles are measured with respect to the horizontal axis, then the transmitters are assumed to be uniformly distributed in $[0, 2\pi)$, where the angle of the k th transmitter

**FIGURE 13.4**

MSE versus SCNR for a noncoherent MIMO radar with 5×4 antennas uniformly placed in $[0, 2\pi)$.

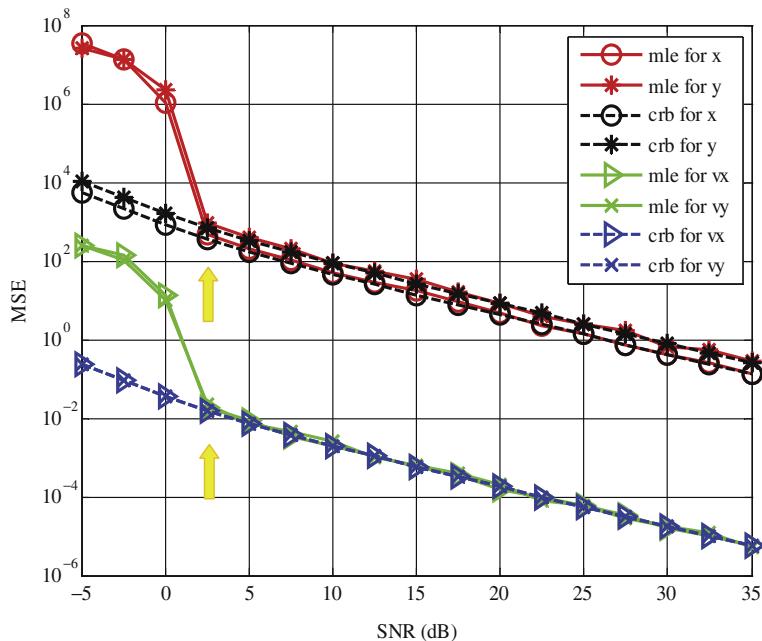
is $\phi_k^t = 2\pi(k-1)/M$, $k = 1, \dots, M$. The receivers are also assumed to be uniformly distributed in $[0, 2\pi)$, where the angle of the l th receiver is $\phi_l^r = 2\pi(l-1)/N$, $l = 1, \dots, N$. Assume the lowpass equivalents of the transmitted waveforms are frequency spread single Gaussian pulse signals

$$s_k(t) = \left(\frac{2}{T^2} \right)^{1/4} \exp \left(\frac{-\pi t^2}{T^2} + j2\pi k f_\Delta t \right), \quad (13.63)$$

where T is proportional to the pulse width and $f_\Delta = f_{k+1} - f_k \geq 0$ is the frequency increment between $s_k(t)$ and $s_{k+1}(t)$. We choose $T = 0.1$, $f_\Delta = 500$ kHz, and set the carrier frequency to $f_c = 1$ GHz. Suppose a target moving with velocity (50, 30) m/s is present at (150, 127.5) m. Assume the variance of each target reflection is the same for every path so that $\sigma_{lk} = \sigma$. The SCNR is defined as $E\sigma^2/(M\sigma_w^2)$.

Under *Assumptions 1–3*, the CRB and the MSE curves of the joint ML estimates are plotted versus SCNR in Figure 13.4 for the MIMO radar system with $M = 5$ transmitters and $N = 4$ receivers. It is observed from the simulation results that the MSEs of the ML estimates approach the corresponding CRBs as the SCNR becomes large. This is consistent with the theoretical asymptotic efficiency of ML estimates and also corroborates the correctness of the derived CRBs presented earlier.

The curves for a noncoherent MIMO radar with $M = 9$ transmitters and $N = 9$ receivers are plotted in Figure 13.5. Compared with Figure 13.4, it is seen that increasing the number of antennas decreases the CRB uniformly and lowers the threshold (i.e., the SCNR at which the MSE curve changes slope drastically, see arrows in the figure) to a smaller value. We also find that these MSE curves get more

**FIGURE 13.5**

MSE versus SCNR for the 9×9 noncoherent MIMO radar with transmit antennas placed uniformly in $[0, 2\pi)$ and receive antennas also placed uniformly in $[0, 2\pi)$.

favorable (closer to CRB in the asymptotic region) with more antennas. These results show that more antennas means better performance in both asymptotic (large MN) and non-asymptotic cases.

More explorations on the noncoherent MIMO radar joint location and velocity estimation are presented in [24], which include the analyses on the threshold phenomenon, the impact of finite system resources, the effects of changing antenna placement, the effects of employing different waveforms, and the noncoherent ambiguity function. Extensions to several general cases, such as for dependent reflection coefficients, nonorthogonal signals, and for spatially colored clutter-plus-noise, are also provided in [24].

2.13.3.2 Noncoherent ambiguity function

Having studied the CRB, we now consider the ambiguity function (AF) for noncoherent MIMO radar. Under Assumptions 1–3, the noncoherent AF has been developed in [24] in two different ways yielding

$$\chi^{\text{non}}(x, y, v_x, v_y) = \frac{1}{MN} \sum_{k=1}^M \sum_{l=1}^N \left| \int_{-\infty}^{\infty} s_k(t) s_k^*(t - \tau_{\Delta lk}) e^{-j2\pi f_{lk} t} dt \right|^2, \quad (13.64)$$

where $\tau_{\Delta lk} = \tau_{lk} - \tau_{0lk}$, with τ_{0lk} determined by the antenna positions which can be obtained by substituting $(x, y) = (0, 0)$ into (13.2). Note that since the $\tau_{\Delta lk}$ and f_{lk} in (13.64) are functions of

(x, y) and (x, y, v_x, v_y) , respectively, the AF is essentially a four-dimensional function with respect to variables (x, y, v_x, v_y) . Interested readers are referred to [24] for more detailed derivations.

Example

Consider a noncoherent MIMO radar system with M transmit and N receive antennas uniformly placed in direction $[0, 2\pi]$, where the look angles are $\varphi_k^t = 2\pi(k-1)/M$, $k = 1, \dots, M$ and $\varphi_l^r = 2\pi(l-1)/N$, $l = 1, \dots, N$. To make the AF a simple two-dimensional function that we can plot and easily interpret, we assume the target only move along the y -axis such that $(x, v_x) = (0, 0)$. Further assume that the monitored area is relatively small and all antennas are located sufficiently far away, so that the $\tau_{\Delta lk}$ and f_{lk} can be approximated to be approximately linearly related to y and v_y through

$$\tau_{\Delta lk} \approx \frac{y}{c} (\sin \varphi_k^t + \sin \varphi_l^r), \quad (13.65)$$

$$f_{lk} \approx \frac{v_y}{\lambda} (\sin \varphi_k^t + \sin \varphi_l^r). \quad (13.66)$$

Thus, the noncoherent AF from (13.64) can be simplified to

$$\chi^{\text{non}}(y, v_y) = \frac{1}{MN} \sum_{k=1}^M \sum_{l=1}^N \left| \int_{-\infty}^{\infty} s_k(t) s_k^*(t - y(\sin \varphi_k^t + \sin \varphi_l^r)/c) e^{-j2\pi v_y(\sin \varphi_k^t + \sin \varphi_l^r)t/\lambda} dt \right|^2. \quad (13.67)$$

Assume frequency spread single Gaussian pulse signals are used for transmission, whose complex envelopes $s_k(t)$ are given in (13.63). Plugging $s_k(t)$ into (13.67), the simplified noncoherent AF for the single Gaussian pulse can be obtained

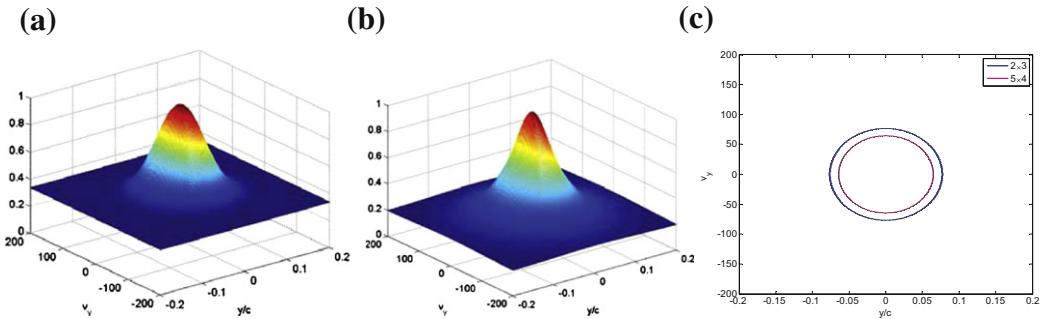
$$\chi^{\text{non}}(y, v_y) = \frac{1}{MN} \sum_{k=1}^M \sum_{l=1}^N e^{-\pi(\sin \varphi_k^t + \sin \varphi_l^r)^2 (y^2/(T^2 c^2) + T^2 v_y^2/\lambda^2)}.$$

The AFs of noncoherent MIMO radar systems with different configurations are plotted in Figure 13.6. The system considered in Figure 13.6a has 2 transmitters and 3 receivers, while the system considered in Figure 13.6b has 5 transmitters and 4 receivers. Using the previously defined simplified AF for the specific examples under study, it can be seen that the mainlobe of AF becomes narrower as more antennas are added. This is easier to read in the 0.5-level contour plots shown in Figure 13.6c. Narrower AF means better resolution, thus the performance is improved by using more antennas in this example.

Assume frequency spread Gaussian pulse trains are used for transmission, thus the signal transmitted from the k th transmitter is

$$s_k(t) = \sum_{n=-\infty}^{\infty} z(t - nT_r) g(t) e^{j2\pi kf_\Delta t}, \quad (13.68)$$

where T_r is the pulse repetition interval, $z(t)$ the narrow Gaussian pulse with pulsedwidth T , $g(t)$ the broad Gaussian envelope with pulsedwidth T_b , and f_Δ the frequency increment. Plugging $s_k(t)$ into

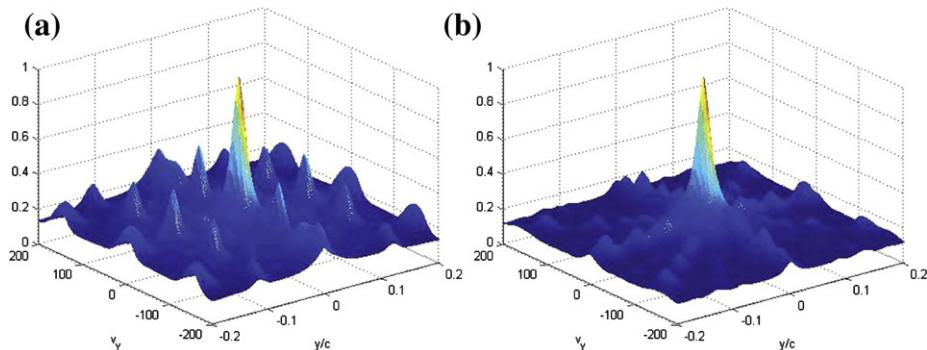
**FIGURE 13.6**

AF of noncoherent MIMO radar systems transmitting frequency spread single Gaussian pulse signals with (a) 2×3 , (b) 5×4 antennas, and their (c) 0.5-level contour, assuming that $(x, v_x) = (0, 0)$.

(13.67) gives the simplified noncoherent AF for the Gaussian pulse train

$$\chi^{\text{non}}(y, v_y) = \frac{1}{MN} \sum_{k=1}^M \sum_{l=1}^N \left\{ \frac{1}{T_r^2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \exp \left[-\pi \left(\frac{y}{c} (\sin \varphi_k^t + \sin \varphi_l^r) - n T_r \right)^2 \frac{1}{T_r^2} - \pi \left(\frac{m}{T_r} \right)^2 T_b^2 - \frac{y^2 \pi}{T_b^2 c^2} (\sin \varphi_k^t + \sin \varphi_l^r)^2 - \pi \left(\frac{v_y}{\lambda} (\sin \varphi_k^t + \sin \varphi_l^r) - \frac{m}{T_r} \right)^2 T_b^2 \right] \right\}.$$

The AF for the noncoherent MIMO radar system with 5 transmitters and 4 receivers is plotted in Figure 13.7a, while the AF for the noncoherent MIMO radar system with 9 transmitters and 9 receivers is plotted in Figure 13.7b. We see that ambiguity sidelobes are created by the use of the Gaussian pulse

**FIGURE 13.7**

AF of noncoherent MIMO radar systems transmitting frequency spread Gaussian pulse trains with (a) 5×4 and (b) 9×9 antennas, assuming that $(x, v_x) = (0, 0)$.

train. Comparing Figure 13.7b to a, it is seen that when the number of antennas increases, the sidelobes decease for the example considered here and our simplified AF. Again, these examples illustrate that using more antennas can give a better AF for noncoherent MIMO radar.

2.13.4 Performance and complexity analysis for coherent and noncoherent MIMO radar

In the previous sections, we have introduced the coherent and noncoherent MIMO radar systems and investigated their joint target location and velocity estimation abilities. The characteristics of the coherent and noncoherent MIMO radar merit more discussions. As opposed to the noncoherent counterpart, the coherent MIMO radar requires extra phase synchronization, which may not be very amenable to practical implementation. Aiming to identify possible scenarios that enable us to replace the coherent MIMO radar with the easier-to-implement noncoherent one but without inducing much loss in performance, we devote this section to studying the MSE performance difference between these two approaches.

It was indicated in [11] that when target localization is the application of interest, significant gains can be obtained through the use of coherent MIMO radar, over noncoherent radar. Here we will demonstrate that the magnitude of these gains decreases with an increase in the product of the number of transmit and receive antennas. In particular, the performance of the noncoherent MIMO radar approaches that of the coherent MIMO radar when the product of the number of transmit and receive antennas is sufficiently large.

2.13.4.1 Normalized mean square error difference

Let $\boldsymbol{\theta}_a = [\theta_{a,1}, \theta_{a,2}, \dots]^T$ and $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \hat{\theta}_2, \dots]^T$ denote the actual and estimated value of the parameter vector to be estimated $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots]^T$. Define the *total MSE* of the joint estimation as

$$\text{MSE} = \mathbb{E} \left\{ \sum_i W_i (\hat{\theta}_i - \theta_{a,i})^2 \right\} = \sum_i W_i \mathbb{E} \left\{ (\hat{\theta}_i - \theta_{a,i})^2 \right\} = \sum_i W_i \text{MSE}_i, \quad (13.69)$$

where $\mathbf{W} = [W_1, W_2, \dots]^T$ defines a constant weighting vector with the same length as $\boldsymbol{\theta}$ and $\text{MSE}_i = \mathbb{E}\{(\hat{\theta}_i - \theta_{a,i})^2\}$ represents the MSE corresponding to the i th component of $\boldsymbol{\theta}$ (i.e., θ_i). Note here we consider a weighted MSE that allows one to give extra priority to some components over others. Of course, taking the weight as a constant vector is also possible, which gives equal weight to all the components.

Theorem 1. Consider a MIMO radar equipped with M separated transmit antennas and N separated receive antennas for joint location and velocity estimation. Assume the coherent MIMO radar (13.22) performs better (in the sense of leading to a smaller MSE) than the noncoherent MIMO radar (13.57) for all M and N . Under Assumptions 1 and 2, the MSE of the ML estimate for the noncoherent MIMO radar (MSE^{non}) gets as close as desired to that of the coherent MIMO radar (MSE^{coh}) for sufficiently large MN .

For a detailed proof of the theorem, please refer to [16].

Remark 1. *Theorem 1* implies that under the given assumptions, the noncoherent MIMO radar, which has less stringent synchronization requirements, is generally more preferable than its coherent counterpart for sufficiently large MN . Stated a different way: (1) If the noncoherent MIMO radar leads to smaller MSE than the coherent MIMO radar, the noncoherent MIMO radar is preferred without a doubt, since it performs better and has less stringent synchronization requirements; (2) otherwise, if the coherent MIMO radar leads to smaller MSE than the noncoherent MIMO radar, as shown by *Theorem 1*, one can add more antennas to compensate, so that at some large MN the MSE performance of the noncoherent MIMO radar is sufficiently close to that of the coherent MIMO radar and thus the noncoherent MIMO radar is good enough.

Based on the total MSEs of the joint estimation for the coherent and noncoherent MIMO radars, assuming a known constant θ_a , we define the *normalized difference of the root mean square errors* (NDRMSE) to evaluate the overall difference in the MSE of the coherent and noncoherent MIMO radars as below

$$\text{NDRMSE} = \frac{(\text{MSE}^{\text{non}})^{1/2} - (\text{MSE}^{\text{coh}})^{1/2}}{(\sum_i W_i \theta_{a,i}^2)^{1/2}} \times 100\%, \quad (13.70)$$

where it is assumed that $\theta_a \neq \mathbf{0}$. If $\theta_a = \mathbf{0}$, then we would use $[(\text{MSE}^{\text{non}})^{1/2} - (\text{MSE}^{\text{coh}})^{1/2}]$ directly. The smaller the NDRMSE, the closer the MSE performance between the two estimators. For a predetermined value of ϵ , which is chosen to satisfy particular requirements, the MSE performance of the noncoherent MIMO radar is considered to be good enough if $\text{NDRMSE} \leq \epsilon\%$, and in such cases the noncoherent MIMO radar is preferred for its less stringent synchronization requirements.

Particularly, in the joint location and velocity estimation problem discussed previously, $\theta = [x, y, v_x, v_y]^T$, so we have

$$\begin{aligned} \text{MSE}^{\text{non}} &= W_1 \mathbb{E} \left\{ \left(\hat{\theta}_x^{\text{non}} - \theta_{a,x} \right)^2 \right\} + W_2 \mathbb{E} \left\{ \left(\hat{\theta}_y^{\text{non}} - \theta_{a,y} \right)^2 \right\} + W_3 \mathbb{E} \left\{ \left(\hat{\theta}_{v_x}^{\text{non}} - \theta_{a,v_x} \right)^2 \right\} \\ &\quad + W_4 \mathbb{E} \left\{ \left(\hat{\theta}_{v_y}^{\text{non}} - \theta_{a,v_y} \right)^2 \right\}, \end{aligned} \quad (13.71)$$

$$\begin{aligned} \text{MSE}^{\text{coh}} &= W_1 \mathbb{E} \left\{ \left(\hat{\theta}_x^{\text{coh}} - \theta_{a,x} \right)^2 \right\} + W_2 \mathbb{E} \left\{ \left(\hat{\theta}_y^{\text{coh}} - \theta_{a,y} \right)^2 \right\} + W_3 \mathbb{E} \left\{ \left(\hat{\theta}_{v_x}^{\text{coh}} - \theta_{a,v_x} \right)^2 \right\} \\ &\quad + W_4 \mathbb{E} \left\{ \left(\hat{\theta}_{v_y}^{\text{coh}} - \theta_{a,v_y} \right)^2 \right\}, \end{aligned} \quad (13.72)$$

and

$$\text{NDRMSE} = \frac{(\text{MSE}^{\text{non}})^{1/2} - (\text{MSE}^{\text{coh}})^{1/2}}{(W_1 \theta_{a,x}^2 + W_2 \theta_{a,y}^2 + W_3 \theta_{a,v_x}^2 + W_4 \theta_{a,v_y}^2)^{1/2}} \times 100\%, \quad (13.73)$$

where it is assumed that $\mathbf{W} = [W_1, W_2, W_3, W_4]^T$. Note that one can easily adjust for units by proper selection of W_i . Later in the numerical examples, we employ $W_i = 1$ for all i .

2.13.4.2 Numerical examples

Example 1

Consider a MIMO radar that has M transmitters and N receivers. When employing the coherent MIMO radar, the antennas are uniformly placed in $[\pi/6, \pi/3]$, where $\varphi_k^t = \pi(k-1)/[6(M-1)] + \pi/6$ and $\varphi_l^r = \pi(l-1)/[6(N-1)] + \pi/6$. In the example, Assumptions 1 and 2 are adopted. When employing the noncoherent MIMO radar, the antennas are uniformly placed in direction $[0, 2\pi)$, where the angle of transmitter k with respect to the horizontal axis is $\varphi_k^t = 2\pi(k-1)/M + \pi/8$, $k = 1, \dots, M$ and the angle of receiver l is $\varphi_l^r = 2\pi(l-1)/N$, $l = 1, \dots, N$. Let $\Delta f = 500$ kHz, such that $\Delta f > 5/T$, and thus the transmitted signals are approximately orthogonal ([24]). Assume the SCNR is 20 dB.

Suppose the system resources employed were not considered, the energy assigned to each transmit antenna was fixed, so adding a transmit antenna implies an extra energy source and increased total transmitted energy. In this case, the NDRMSEs are plotted versus MN with N (or M) going from 1 to 20 for a fixed $M = 4$ (or $N = 4$) in Figure 13.8a. It is observed that increasing either M or N decreases the NDRMSE, which makes the MSE performance of the noncoherent MIMO radar get closer to that of the coherent MIMO radar. When the system resources employed are taken into account, the situation becomes complicated. In such cases, the total transmitted energy is fixed, so that adding transmit antennas means splitting the energy over more antennas. Figure 13.8b shows such an example. Under the energy constraint, the NDRMSEs are plotted versus MN with the other parameters set the same as Figure 13.8a. It is observed that, when the total energy is fixed, increasing the number of receive antennas N always decreases NDRMSE, but this is not true when we increase the number of transmit antennas M for a fixed N . Increasing M first decreases the NDRMSE (when M is small, e.g., $M \leq 4$ in this example), but later increases the NDRMSE when M is large (e.g., $M > 4$ in this example). Here

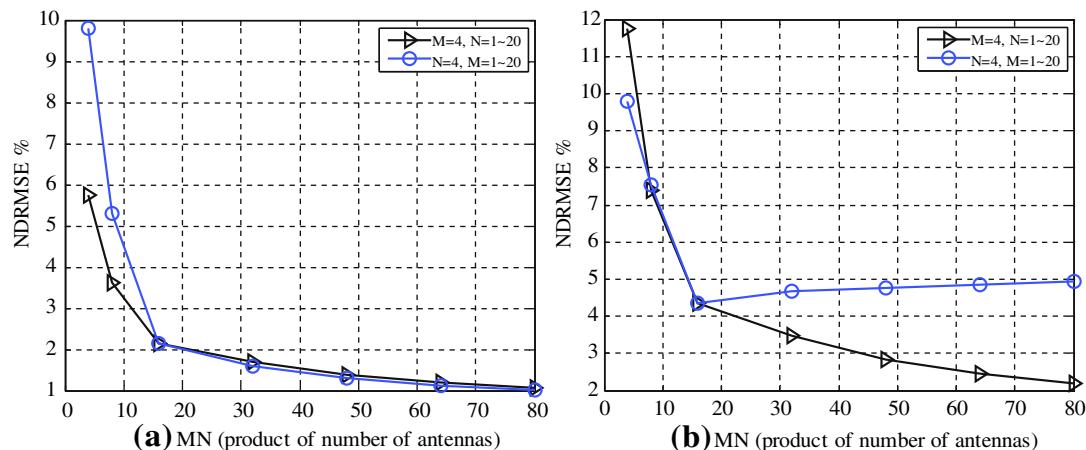


FIGURE 13.8

NDRMSEs versus MN with an increasing N (or M) for a fixed M (or N) for the joint estimation with (a) unconstrained energy or (b) constrained energy. SCNR = 20 dB.

we see the effects of spreading the energy too thinly between the transmit antennas. This occurs when M is too large.

2.13.4.2.1 Example 2

Consider the same scenario in Figure 13.8, except that the antenna positions for the noncoherent MIMO radar are different. Here the transmit antennas are assumed to be at $(x_k^t, 7000)$ m, $x_k^t = 10,000(k - 1)/(M - 1) - 3000$, $k = 1, \dots, M$ and the receive antennas are assumed to be at $(7000, y_l^r)$ m, $y_l^r = 10,000(l - 1)/(N - 1) - 3000$, $l = 1, \dots, N$. In this example, we fix the total number of antennas to $M + N = 50$. This is an approximate way to fix the total system complexity, since each added antenna implies the addition of several accompanying hardware components. Further, counting system complexity through the total number of antennas used can obviate the need for getting involved with details related to the hardware implementation. This is very useful in a high level study like the one we undertake here.

When the energy is unconstrained, i.e., the transmit energy per transmit antenna is fixed so that adding a transmit antenna increases the total transmit energy, the RMSE_i 's are plotted in Figure 13.9a. It is seen that there is a symmetry between M and N , and the minima for both the coherent and noncoherent MIMO radars occur at $M = 25 = (M + N)/2$, as expected. When the total energy is constrained, i.e., adding a transmit antenna decreases the transmit energy per transmit antenna, the RMSE_i 's are

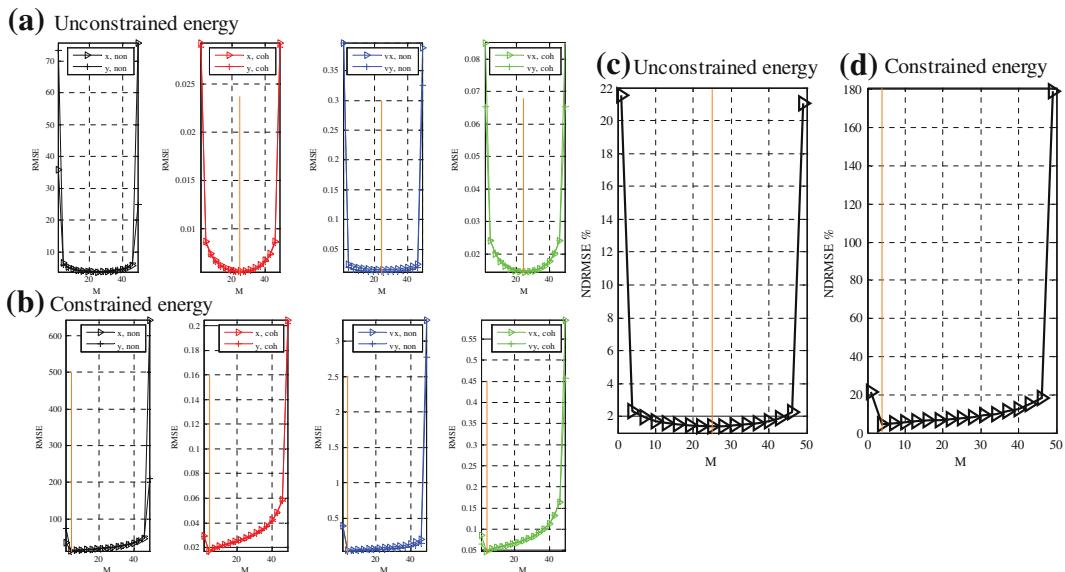


FIGURE 13.9

RMSE_i 's (in m for location and m/s for velocity) for the joint estimation of location and velocity with (a) unconstrained energy or (b) constrained energy, and the corresponding NDRMSEs (c)–(d) versus M . $M + N = 50$. SCNR = 20 dB.

plotted in Figure 13.9b. It is seen that the effects of increasing M and N are asymmetric. For both the coherent and noncoherent MIMO radars, $M = 4$ is the best, and further increasing M will degrade the performance. The NDRMSEs for the unconstrained and constrained energy cases are plotted versus M in Figure 13.9c and d, respectively. The best MSE performance occurs at $M = (M + N)/2$ for the unconstrained case and for an M such that $1 < M < (M + N)/2$ for the constrained case. In fact we have tested other cases (different placements and fixed $M + N$ value) and obtained similar results.

The main point is that if we fix the total energy and if we add too many transmit antennas then we will spread the energy too thinly among the transmit antennas and this will lead to bad performance. On the other hand, there are cases where we get gains from adding transmit antennas with fixed total energy, for example, when we do not spread the energy too thinly. Usually, it is better to use more than one transmit antenna to obtain good performance. Further numerical investigations and discussions can be found in [16].

2.13.5 Diversity gain for MIMO radar Neyman-Pearson signal detection

So far, the chapter has been about MIMO radar for estimation. In this section, we discuss the diversity gain that a MIMO radar system can render in the context of target detection. Diversity gain is one significant advantage that MIMO radar systems employing separated antennas can bring forth when compared with more traditional radar systems. When the Neyman-Pearson criterion is employed, for a fixed false alarm probability, *diversity gain* is defined as the negative of the slope of the miss probability versus SCNR for the high SCNR region when a logarithmic scale is employed for both axes. Assuming linear decay of the miss probability for sufficiently large SCNR when such scales are employed, large diversity gain implies good target detection performance for sufficiently high SCNR and fixed probability of false alarm. If we have two detectors and one has larger diversity gain, then at some sufficiently large SCNR, the detector with larger diversity gain must have smaller miss probability. Intuitively, diversity gain tells us about the value of the information we get from multiple looks (from several antennas, frequencies, or retransmissions, etc.).

2.13.5.1 Signal model

Consider a radar system that has M transmit and N receive antennas. The positions of the transmit and receive antennas of the radar system are (x_k^t, y_k^t) , $k = 1, \dots, M$ and (x_l^r, y_l^r) , $l = 1, \dots, N$, respectively. Denote the waveform transmitted by the k th antenna by $s_k(t)$, where $s_k(t)$ and $s_{k'}(t)$ can be non-orthogonal for $k \neq k'$, in the sense of [4]. Assume each waveform is normalized to give transmitted power E . The radar will break up the monitored space into cells and sequentially probe each cell for a target. Thus, our goal is to find out whether a target is at a particular position or not. Assume a target, if present, is composed of Q point scatterers, located at (x^q, y^q) , $q = 1, \dots, Q$. The reflection coefficient, ζ_q for the q th scatterer, is assumed to be constant over the observation interval.

The clutter-plus-noise-free received signal at receive antenna l due to the transmission from transmit antenna k and the reflection from the q th scatterer is modeled as [4]

$$r_{lk}^q(t) = \sqrt{\frac{E}{M}} \zeta_q s_k(t - \tau_{lk}^q) e^{-j2\pi f_c \tau_{lk}^q}, \quad (13.74)$$

where $\tau_{lk}^q = (d_{tk}^q + d_{rl}^q)/c$ denotes the time delay from transmit antenna k to receive antenna l due to the reflection from the q th scatterer, $d_{tk}^q = [(x_k^t - x^q)^2 + (y_k^t - y^q)^2]^{1/2}$, $d_{rl}^q = [(x_l^r - x^q)^2 + (y_l^r - y^q)^2]^{1/2}$, and c is the speed of light. Assume the transmitted waveforms are relatively narrow band, where the bandwidth of the waveforms is given such that they are not capable of resolving individual scatterers [5]. Therefore,

$$s_k(t - \tau_{lk}^q) \approx s_k(t - \tau_{lk}^0), \quad q = 1, \dots, Q, \quad (13.75)$$

where τ_{lk}^0 denotes the time delay from a reflection off a scatterer located at the gravity center of the scatterers (x^0, y^0) . Then (13.74) can be written as

$$r_{lk}^q(t) = \sqrt{\frac{E}{M}} \zeta_q s_k(t - \tau_{lk}^0) e^{-j2\pi f_c \tau_{lk}^q}. \quad (13.76)$$

Thus, the received signal at receive antenna l due to the propagation from all M transmit antennas and all Q scatterers is given by

$$r_l(t) = \sqrt{\frac{E}{M}} \sum_{q=1}^Q \sum_{k=1}^M \zeta_q s_k(t - \tau_{lk}^0) e^{-j2\pi f_c \tau_{lk}^q} + w_l(t), \quad (13.77)$$

where $w_l(t)$ denotes the clutter-plus-noise observed at the l th receiver.

2.13.5.2 Effect of signal space dimension on diversity gain

Suppose we attempt to implement the optimum receiver for the Gaussian reflection coefficients and Gaussian clutter-plus-noise case following (13.77), by first projecting the received continuous-time observations onto a basis that spans the space spanned by the first term, the signal term, in (13.77). Let the set of M_l orthonormal basis functions $\{\phi_{l1}(t), \phi_{l2}(t), \dots, \phi_{lM_l}(t)\}$ span the M_l dimensional space spanned by the first term, the signal term, of (13.77). The expansion, in terms of this basis, of any delayed signal term appearing in (13.77) can be expressed as

$$s_k(t - \tau_{lk}^0) = \sum_{i=1}^{M_l} s_{lki} \phi_{li}(t) \quad (13.78)$$

with the coefficients defined by

$$s_{lki} = \int_{-\infty}^{\infty} s_k(t - \tau_{lk}^0) \phi_{li}(t) dt, \quad i = 1, \dots, M_l, \quad (13.79)$$

which are obtained by using (orthonormal basis)

$$\int_{-\infty}^{\infty} \phi_{li}(t) \phi_{li'}(t) dt = \begin{cases} 1, & i = i', \\ 0, & \text{otherwise,} \end{cases} \quad (13.80)$$

after taking the inner product with $\phi_{li}(t)$ (right hand side of (13.79)) on both sides of (13.78). We note that the largest possible value of M_l is M and this occurs when each of the delayed signal terms in (13.77), the terms expanded in (13.78), are linearly independent. Now we can project the received signal $r_l(t)$ in (13.77) onto this basis to obtain⁸

$$\begin{aligned} r_{li} &= \int_{-\infty}^{\infty} r_l(t) \phi_{li}(t) dt \\ &= \sqrt{\frac{E}{M}} \sum_{k=1}^M s_{lk} \sum_{q=1}^Q \zeta_q e^{-j2\pi f_c \tau_{lk}^q} + w_{li} \\ &= \sqrt{\frac{E}{M}} \sum_{k=1}^M s_{lk} \mathbf{A}_{lk} \boldsymbol{\zeta} + w_{li}, \quad i = 1, \dots, M_l, \end{aligned} \quad (13.81)$$

where $\mathbf{A}_{lk} = [e^{-j2\pi f_c \tau_{lk}^1}, \dots, e^{-j2\pi f_c \tau_{lk}^Q}]$, and $\boldsymbol{\zeta} = [\zeta_1, \dots, \zeta_Q]^T$ is a vector of reflection coefficients for those Q scatterers. Collecting the M_l outputs and letting $\mathbf{r}_l = [r_{l1}, \dots, r_{lM_l}]^T$, we have

$$\mathbf{r}_l = \sqrt{\frac{E}{M}} \sum_{k=1}^M \mathbf{s}_{lk} \mathbf{A}_{lk} \boldsymbol{\zeta} + \mathbf{w}_l, \quad (13.82)$$

where

$$\mathbf{s}_{lk} = [s_{l1}, \dots, s_{lM_l}]^T \quad (13.83)$$

and $\mathbf{w}_l = [w_{l1}, \dots, w_{lM_l}]^T$. Further, stacking the outputs across all the N receivers into a single vector gives

$$\begin{aligned} \mathbf{r} &= [\mathbf{r}_1^T, \dots, \mathbf{r}_N^T]^T = [r_{11}, r_{12}, \dots, r_{NM_N}]^T \\ &= \sqrt{\frac{E}{M}} \text{Diag}\{\mathbf{S}_1, \dots, \mathbf{S}_N\} \mathbf{A} \boldsymbol{\zeta} + \mathbf{w} \\ &= \sqrt{\frac{E}{M}} \mathbf{C} \mathbf{A} \boldsymbol{\zeta} + \mathbf{w}, \end{aligned} \quad (13.84)$$

where

$$\mathbf{C} = \text{Diag}\{\mathbf{S}_1, \dots, \mathbf{S}_N\} \quad (13.85)$$

denotes a block diagonal matrix with submatrices $\mathbf{S}_1, \dots, \mathbf{S}_N$ on its diagonal, and the size of each submatrix

$$\mathbf{S}_l = [\mathbf{s}_{l1}, \dots, \mathbf{s}_{lM_l}] \quad (13.86)$$

⁸Choosing an orthogonal set in a particular manner, say using the Karhunen-Loeve expansion [23], allows one to represent the received continuous signals by a set of random variables that compose a sufficient statistic for the target-absent versus target-present hypothesis testing problem.

is $M_l \times M$; $\mathbf{A} = [\mathbf{A}_{11}^T, \mathbf{A}_{12}^T, \dots, \mathbf{A}_{NM}^T]^T$ is an $NM \times Q$ matrix. Let

$$\tilde{M} = M_1 + \dots + M_N, \quad (13.87)$$

then the dimension of matrix \mathbf{C} is given by $\tilde{M} \times NM$. The $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_N^T]^T$ in (13.84) is an $\tilde{M} \times 1$ clutter-plus-noise vector with covariance matrix \mathbf{R}_w . Based on (13.84), we have the following results which elaborate the diversity gain offered by MIMO radar systems employing non-orthogonal waveforms.

Theorem 1. Consider a MIMO radar system with M transmit and N receive antennas and observations following (13.77) with possibly nonGaussian clutter-plus-noise and reflection coefficients. After expanding the received signals using basis functions as per (13.81) and collecting the coefficients as per (13.84), the detection system can be described by a hypothesis testing problem as follows:

$$\begin{aligned} H_1 : \mathbf{r} &= \sqrt{\frac{E}{M}} \mathbf{CA}\boldsymbol{\zeta} + \mathbf{w}, \\ H_0 : \mathbf{r} &= \mathbf{w}. \end{aligned} \quad (13.88)$$

Assume that a detector of the class of

$$T = \mathbf{r}^H \mathbf{D} \mathbf{D}^H \mathbf{r} = \|\mathbf{D}^H \mathbf{r}\|^2 \geq_{H_0}^{H_1} \delta. \quad (13.89)$$

is employed with an arbitrary \mathbf{D} that meets $\|\mathbf{D}\| < \infty$ and a positive finite threshold δ determined by the required false alarm probability. Plug (13.88) under H_1 into (13.89) and write the test statistic in the form of

$$T = \frac{E}{M} \tilde{u} + w, \quad (13.90)$$

to define \tilde{u} and w . If the cumulative distribution function (cdf) of \tilde{u} , denoted as $F_{\tilde{u}}(\tilde{u})$, satisfies

$$F_{\tilde{u}}(\tilde{u}) = \tilde{u}^p G + o(\tilde{u}^p), \quad \text{as } \tilde{u} \rightarrow 0^+ \quad (13.91)$$

for some constant⁹ $0 \leq p < \infty$ and nonzero constant $G < \infty$, and

$$\mathbb{E} \left\{ \left| \sqrt{\delta} \pm \|\mathbf{D}^H \mathbf{w}\| \right|^{2p} \right\} < \infty \quad (13.92)$$

holds true, then the diversity gain achieved by this system is $d = p$.

Remark 2. For the hypothesis testing problem in (13.88), if the clutter-plus-noise \mathbf{w} is zero-mean Gaussian with covariance matrix \mathbf{R}_w and the reflection coefficient $\boldsymbol{\zeta}$ is zero-mean Gaussian with covariance matrix \mathbf{R} , then the detector in (13.89) is optimum for

$$\mathbf{D} = \mathbf{R}_w^{-1} \mathbf{C} \mathbf{A} \mathbf{X}_1^H, \quad (13.93)$$

where \mathbf{X}_1 satisfies $\mathbf{X}_1^H \mathbf{X}_1 = [\gamma'(\mathbf{C} \mathbf{A})^H \mathbf{R}_w^{-1} \mathbf{C} \mathbf{A} + \mathbf{R}^{-1}]^{-1}$ and γ' denotes the SCNR the detector is designed for.

⁹The notation $\tilde{u} \rightarrow 0^+$ means \tilde{u} approaches 0 from right. Note that in the cases where $p = 0$ we will get no diversity gain.

Theorem 2. Consider a MIMO radar system with M transmit and N receive antennas attempting to detect the presence of a target with Q scatterers. Assume observations following (13.77) with possibly nonGaussian clutter-plus-noise and reflection coefficients. After expanding the M_l -dimensional signal at the l th receive antenna, $l = 1, \dots, N$ using basis functions as per (13.81) and collecting the coefficients as per (13.84), assume that a detector of the class of (13.89) is employed, which includes the optimum detector for Gaussian clutter-plus-noise and reflections. Plug (13.88) under H_1 into (13.89) and write the test statistic in the form of

$$T = \frac{E}{M} \sum_{\ell=1}^L \tilde{u}_\ell + w = \frac{E}{M} \tilde{u} + w, \quad (13.94)$$

where $\tilde{u} = \sum_{\ell=1}^L \tilde{u}_\ell$, to define \tilde{u}_ℓ and w . If each \tilde{u}_ℓ for $\ell = 1, \dots, L$ has a cdf $F_{\tilde{u}_\ell}(\tilde{u}_\ell)$ that satisfies

$$F_{\tilde{u}_\ell}(\tilde{u}_\ell) = \tilde{u}_\ell^{p_\ell} G_\ell + o(\tilde{u}_\ell^{p_\ell}), \text{ as } \tilde{u}_\ell \rightarrow 0^+, \quad (13.95)$$

for some constant $0 \leq p_\ell < \infty$ and nonzero $G_\ell < \infty$, and (13.92) holds true, then the maximum diversity gain is

$$d = \sum_{\ell=1}^{\min(M_1+\dots+M_N, Q, \text{rank}(\mathbf{D}))} p_\ell. \quad (13.96)$$

This maximum diversity gain can be achieved under some conditions, for example statistically independent $\tilde{u}_1, \dots, \tilde{u}_L$, $\text{rank}(\mathbf{CA}) = \text{rank}(\mathbf{D})$, and \mathbf{CA} has full row or column rank.

For the special case of Gaussian reflection coefficients (Gaussian signals), it is known from [25] that $p_\ell = 1$, so the diversity gain is given by $d \leq \min(M_1 + \dots + M_N, Q, \text{rank}(\mathbf{D}))$. When the Gaussian optimum detector is employed, it can be shown that $\text{rank}(\mathbf{D}) = \text{rank}(\mathbf{CA}) \leq \min(M_1 + \dots + M_N, Q)$, which leads to a diversity gain $d \leq \min(M_1 + \dots + M_N, Q)$. Further, suppose the dimension of the space spanned by the target-reflected noise-free received waveforms at different receivers are the same for all l , such that $M_l = M'$. Then the largest possible diversity gain described in Theorem 2 becomes $d = \min\{NM', Q\}$.

Assuming the number of scatterers is very large such that $Q \gg NM$, Theorem 2 implies that to achieve the largest possible diversity gain, $\sum_{\ell=1}^{\min(NM, \text{rank}(\mathbf{D}))} p_\ell$, the noise-free received waveforms, due to target reflection, at the l th receiver must span the entire M dimensional space. If the clutter-plus-noise free received waveforms span a lower dimension, say $M_l < M$, the diversity gain will be equal to or smaller than $\sum_{\ell=1}^{\min(M_1+\dots+M_N, \text{rank}(\mathbf{D}))} p_\ell$, which is less than $\sum_{\ell=1}^{\min(NM, \text{rank}(\mathbf{D}))} p_\ell$ in this case. In other words, MIMO radars that employ either orthogonal waveforms or non-orthogonal waveforms can provide the maximum achievable diversity gain, as long as the noise-free received waveforms, due to target reflection, at each receiver span an M dimensional space.

Example 1

Consider a MIMO radar system that has $M = 2$ transmit antennas located at $(x_1^t, y_1^t) = (2, -2)$ km and $(x_2^t, y_2^t) = (6, -4)$ km, and $N = 2$ receive antennas located at $(x_1^r, y_1^r) = (8, 2)$ km and $(x_2^r, y_2^r) = (4, 0)$ km. The waveforms emitted from these two transmitters are $s_1(t)$ and $s_2(t)$ respectively. Gaussian optimum detector is employed. Assume the Q scatterers constituting the target are uniformly distributed

over $[0.3, 1] \times [9.4, 10.5] \text{ km}^2$. Each scatterer has a statistically independent scattering coefficient. In each example, the probability of miss versus SCNR curve is obtained from 100,000 Monte Carlo simulations per SCNR, and the resulting diversity gains are compared with those calculated using the corresponding theorem from the previous sections.

Assume the reflection coefficients and the clutter-plus-noise are statistically independent complex Gaussian with zero mean. Assume the noise-free received waveforms (due to target reflection) at each receiver, $s_1(t - \tau_{l1}^0)$ and $s_2(t - \tau_{l2}^0)$, $l = 1, 2$, are normalized and orthogonal [4]. In this case, these waveforms themselves are the orthonormal basis functions, $\phi_{l1}(t) = s_1(t - \tau_{l1}^0)$ and $\phi_{l2}(t) = s_2(t - \tau_{l2}^0)$, such that $M_l = 2$, $\mathbf{s}_{l1} = [1, 0]^T$, and $\mathbf{s}_{l2} = [0, 1]^T$ as per (13.83), and the \mathbf{S}_l in (13.86) is a 2×2 identity matrix, for $l = 1, 2$. Therefore the matrix \mathbf{C} in (13.85) is a 4×4 identity matrix and $M_1 + M_2 = 4$. Further, assume the number of scatterers is $Q = 1500$. When the probability of false alarm is fixed at $P_{\text{FA}} = 10^{-3}$, the probability of miss detection P_M is plotted versus SCNR in the solid curve with points marked with squares in Figure 13.10. It is seen that increasing SCNR by 10 dB decreases the P_M by four decades approximately. Thus, the slope of the miss probability versus SCNR, if plotted on a log-log scale, is approximately $d = 4$. This agrees well with *Theorem 2*, where the maximum diversity gain is $d = \min(M_1 + M_2, Q) = \min(4, 1500) = 4$, recalling that $p_\ell = 1$ for Gaussian reflections. Note that in this example $NM = M_1 + M_2 = 4$, the maximum diversity gain is achieved as expected, which is guaranteed by the statistically independent reflections of the Q scatterers, according to the discussion after *Theorem 2*.

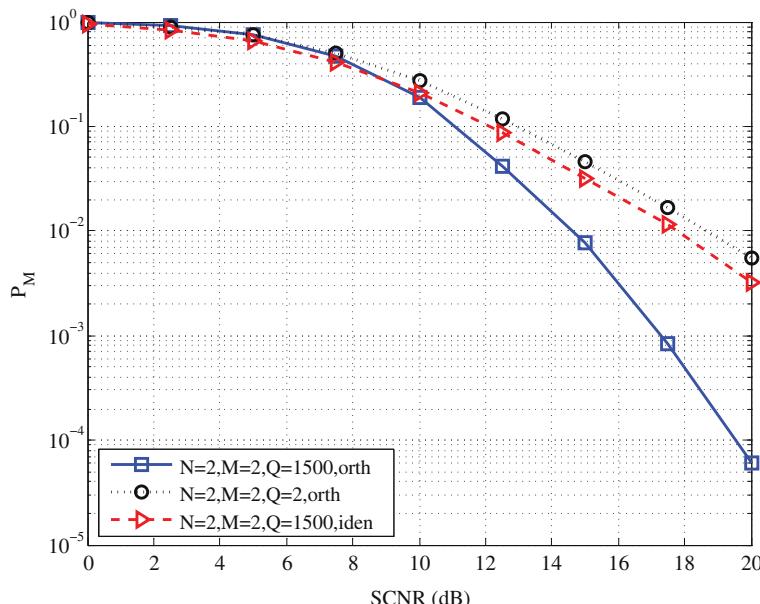


FIGURE 13.10

Probability of miss versus SCNR for three difference cases under the condition of Gaussian reflections and Gaussian clutter-plus-noise.

In the second example, we assume the number of scatterers constituting the target is reduced to $Q = 2$. The other parameters are kept unchanged. The P_M versus SCNR is shown in the dotted curve with points marked with circles in Figure 13.10. It is observed that increasing SCNR by 10 dB decreases the P_M by two decades approximately. The slope of the decrease of the miss probability versus SCNR on a log-log scale is thereby $d = 2$. This also justifies *Theorem 2*, where the maximum diversity gain is $d = \min(M_1 + M_2, Q) = \min(4, 2) = 2$.

The third example considers the same scenario as the first example, except that we assume the two noise-free received waveforms at each receiver (due to target reflection) are identical, $s_1(t - \tau_{l1}^0) = s_2(t - \tau_{l2}^0)$, where $l = 1, 2$. Thus, $M_l = 1$, $\mathbf{s}_{l1} = 1$ and $\mathbf{s}_{l2} = 1$, and $\mathbf{S}_l = [1, 1]$. Hence, the matrix \mathbf{C} in (13.85) is a 2×4 matrix, and $M_1 + M_2 = 2$. In this case, P_M versus SCNR is plotted in the dashed curve with points marked with triangles in Figure 13.10. It is seen that the diversity gain is approximately $d = 2$. Again, *Theorem 2* is justified, since the maximum diversity gain $d = \min(M_1 + M_2, Q) = \min(2, 1500) = 2$. This example illustrates the importance of the signal space spanned by the received waveforms.

Example 2

This numerical example illustrates the maximum diversity gains that can be achieved by the MIMO radar system employing orthogonal waveforms and the traditional phased array radar system. We focus on Gaussian signals and clutter-plus-noise. Employ the test in (13.89) with $\mathbf{D} = \mathbf{I}_{N \times M} = \mathbf{I}_{2 \times 2}$ which we

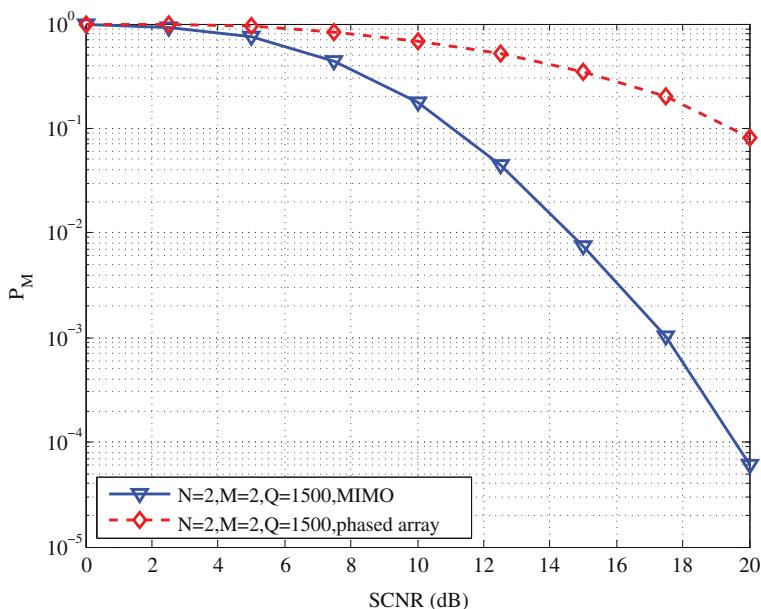


FIGURE 13.11

Probability of miss versus SCNR for MIMO and phased array (SISO) systems.

call noncoherent MIMO radar employing orthogonal waveforms (which maintain orthogonality after each experiencing different delays) [2], and $\mathbf{D} = \mathbf{a}^H(\theta) = [1, 1]$ which we call coherent processing. The resulting P_M versus SCNR curves are shown in Figure 13.11. It is found that, for the noncoherent processing (solid curve with triangles), increasing SCNR by 10 dB approximately decreases the P_M by four decades, so that $d = 4$, which justifies $d = \sum_{\ell=1}^{NM} p_\ell = NM$ for the noncoherent MIMO radar employing orthogonal waveforms. On the other hand, for the coherent processing (dashed curve with diamonds) used in the phased array radar, increasing SCNR by 10 dB approximately decreases the P_M by one decade so that $d = 1$, which justifies $d = p_1 = 1$ for the phased array radar performing beamforming.

For more theoretical analysis and numerical examples on non-Gaussian signals and non-Gaussian clutter-plus-noise, please refer to [25].

2.13.6 Phase synchronization for coherent MIMO radar

In the previous sections, we have elucidated the characteristics of both coherent MIMO radar and non-coherent MIMO radar. Regardless of which mode is employed, all the elements¹⁰ in MIMO radar need to synchronize firstly in terms of, for example, time, frequency, or phase. In fact, phase synchronization embodies a major difference between the operations of noncoherent MIMO radar and coherent MIMO radar: the coherent approach requires phase synchronization between the separated antenna elements, while the noncoherent counterpart does not. Note that in reality, MIMO radar elements are usually operated with physically different local oscillators, and each of them suffers from statistically independent phase offset, indicating that the phase of the carrier signal transmitted gets rotated by an unknown amount. Such imperfection in phase coherence can have an unfavorable effect on the performance of coherent MIMO radar, e.g., [26–29]. As a result, practical implementation of coherent MIMO radar systems needs to cope with the fundamental challenge of ensuring a common notion of phase among all the antennas, and calls for efficient algorithms which can lead to good performance in terms of phase coherence. Although the synchronization problem has been intensely studied in several closely related contexts, for example, in traditional radar systems (see, e.g., [30–36]) and in wireless networks including sensor networks (see e.g., [37–44]), work that is specifically targeted at attaining phase coherence, or phase synchronization in coherent MIMO radar seems still fragmented.

In the sequel, we will briefly discuss three different approaches that have been recently developed by us in [45] to achieve phase coherence in coherent MIMO radar, namely, the master-slave closed-loop algorithm, the round-trip algorithm, and the broadcast consensus based algorithm. These phase synchronization algorithms range from centralized to distributed types, and include both non-iterative and iterative methods. They are all time asynchronous algorithms, and do not require the establishment of time synchronization in advance. In particular, the master-slave closed-loop algorithm and the round-trip algorithm also require no *a priori* frequency synchronization. These algorithms are versatile and practically implementable, and should cater to different needs arising as a result of different number of MIMO radar elements, different sensor placements, as well as different operational preferences.

¹⁰In this section, we use the terms *sensor*, *node*, and *element* interchangeably.

2.13.6.1 Master-slave closed-loop phase synchronization

The master-slave closed-loop approach is a very simple scheme. It employs a master-slave architecture as illustrated in Figure 13.12, where sensor S_1 is denoted as the master node, and S_i , $i = 2, \dots, K$, as slave nodes. All the slave sensors will then synchronize their local phase with that of the master node. We assume no *a priori* establishment of time synchronization across all the radar elements, at least not one with sufficient accuracy to enable phase synchronization. Each radar element maintains its own time using its own independent local oscillator, and is able to track the elapsed time using its local clock with relative precision. As a result of the absence of a common time reference, the instantaneous phase of each radar element can no longer be precisely determined by the local time at that element. For simplicity, we assume all the radar elements have a similar notion of the length of a specified time interval, e.g., T for a time slot. In practice, this can be achieved by using identical stable clocks across all the radar elements. Moreover, we also assume no *a priori* frequency synchronization among these nodes. But we assume all the radar elements are equipped with identical frequency and phase estimators.

For this master-slave architecture, there exist two basic variations, depending on how the master element interacts with each of these slave elements. The first one is the open-loop approach, in which the master node broadcasts an unmodulated sinusoidal reference signal to all the slave nodes, and each of these slave nodes will then use such a reference signal to estimate and correct its phase offset. In this way, interaction between the master node and slave nodes is minimal. However, the absence of a common time reference, as we pointed out earlier, makes it difficult for each sensor to generate absolute phase with respect to the true time. As a result, phase shift induced by the channel between the master node and each slave node can not be disambiguated at each slave node, and hence cannot be compensated either. So, this open-loop method may lead to inferior accuracy in phase synchronization. On the other hand, suppose the channel phase information can be measured in a certain way and then fed back explicitly to each slave node for appropriate phase compensation, the accuracy of phase synchronization can be much improved. This is the basic idea behind the closed-loop approach, another type of master-slave synchronization. However, note that a fundamental premise that lies underneath this approach is the channel reciprocity, which is assumed in this work, meaning that the channel between a pair of sensors remains the same for both forward and reverse directions.

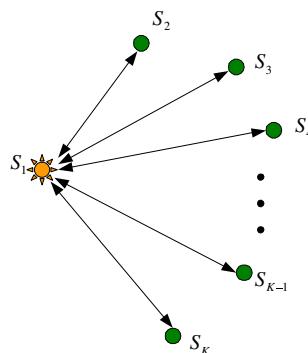


FIGURE 13.12

Master-slave architecture.

In the following, we describe a specific time-slotted closed-loop phase synchronization algorithm. In particular, we assume each slave element is assigned a specific time slot, during which they can exchange information with the master element S_1 . We further assume each slave element knows the number (or order) of the time slot assigned to it. This time-slotted closed-loop approach is then summarized as follows:

1. S_1 broadcasts an unmodulated beacon signal to all the slave elements.
2. Upon receiving the beacon signal, each slave sensor performs the following steps concurrently:
 - a. It estimates the frequency and phase of the received signal.
 - b. It begins to count the elapsed time using its own local clock, and waits for its time slot to come.
3. When the time slot for slave element S_i starts according to its local clock, the following steps will take place in a sequential order:
 - a. S_i passes the received beacon signal back to S_1 , using the local phase and frequency estimates generated from its observation.
 - b. Upon receiving the returned beacon signal from S_i , S_1 estimates the received phase, and computes the phase difference relative to the originally transmitted reference signal. It divides the phase difference by two, pre-compensates the phase offset, and then transmits the beacon signal back to S_i .
 - c. After receiving the reference signal, S_i estimates the phase information and adjusts its carrier phase accordingly.

As is evident from the above procedures, it is essential to estimate the pre-compensation phase offset for each slave node when performing this closed-loop phase synchronization scheme. But since the channel between the master node and each slave node may vary from time to time, though at a slow rate, the value of the pre-compensation offset has to be estimated and updated in a periodic manner. A detailed timing analysis and phase synchronization error analysis of this algorithm can be found in [45].

2.13.6.2 Round-trip phase synchronization

The round-trip phase synchronization algorithm, as detailed in [45], was inspired by the all-node-based method proposed in [41] to globally synchronize the clocks in a sensor network, and the time-slotted round-trip approach that was developed in [42] and further improved or modified in [43,44] to attain carrier synchronization for distributed beamforming in multiuser wireless communication systems. The above mentioned approaches, nevertheless, do not appear to be applicable to this MIMO radar phase synchronization problem directly, even though a common feature shared by these approaches, that when in operation, a reference signal or message traverses in loops through all the sensors, remains essential.

The round-trip phase synchronization scheme for coherent MIMO radar [45] employs an unmodulated beacon signal to travel through all these radar elements in a round-trip manner. Each node will record the phase information twice in order to compute the final phase. The equivalence of the cumulative phase shift at each sensor, a quantity that is directly related to the total propagation delay for the round-trip circuit, is the fundamental idea behind this technique. This technique employs pretty much the same assumptions as those in the master-slave closed-loop method. For example, this scheme

also assumes no *a priori* establishment of both time and frequency synchronization across all the radar elements, and assumes the channel reciprocity. However, specific to this round-trip method, we further assume that when one sensor transmits, the signal is only properly received by its intended receiver, i.e., its downstream neighbor. This can be attained in practice, for example, by using directional transmissions or appropriate power control schemes at the transmitter (assuming each radar element loosely knows its and its immediate neighbors' locations). In what follows, we summarize this round-trip phase synchronization algorithm. Its mathematical characterization, timing analysis, and error analysis can be found in [45]:

1. Find a cycle which passes through all the radar elements. The radar elements are then labeled in order as S_1, S_2, \dots, S_K . Let S_1 be the initiating element.
2. In time slot TS_1 , an unmodulated sinusoidal reference signal is transmitted from S_1 to its downstream neighbor S_2 . This reference signal is further passed along the cycle in a clockwise manner, as illustrated by solid lines in Figure 13.13.
3. Upon receiving the reference signal, each radar element S_i generates its local phase and frequency estimates from the noisy observation. In time slot TS_i , $i = 2, \dots, K$, which starts immediately after the conclusion of the received signal, S_i relays the received reference signal to the next node in the cycle. S_1 will be the downstream neighbor of S_K .
4. After the initiating node S_1 receives the signal from S_K , it generates a phase estimate from its noisy observation. Time slot TS_{K+1} starts right after the conclusion of the received signal, wherein S_1 transmits the original unmodulated sinusoidal reference signal as in step (1) to sensor S_K . This signal is passed along the cycle in the counterclockwise direction, as shown by dashed lines in Figure 13.13.
5. Upon receiving the signal transmitted from its upstream neighbor, each sensor S_i generates its local phase and frequency estimates from its observation, and calculates its final phase based on its two-time phase estimates. In time slot TS_{2K-i+2} , S_i relays the received signal to the next node in the cycle, i.e., S_{i-1} .

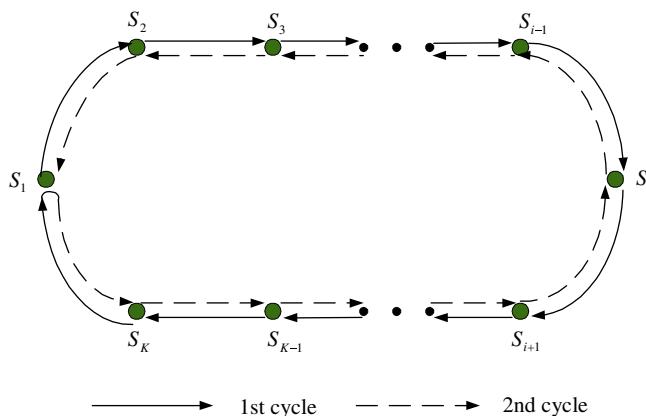


FIGURE 13.13

Illustration of the round-trip phase synchronization method.

6. Relay of the beacon signal ends after S_1 finishes its receipt of the signal from S_2 . S_1 generates a phase estimate from its noisy observation, and then calculates its final phase.

2.13.6.3 Broadcast consensus based phase synchronization

It can be easily seen that the aforescribed master-slave closed-loop and round-trip phase synchronization algorithms are not distributed in nature. Thus, both methods are not very robust against node failures. As far as robustness and scalability is concerned, distributed algorithms have enormous advantages, and may be more desirable for phase synchronization in coherent MIMO radar systems. The essence of these algorithms is that one or more sensors can communicate information with its immediate neighbors in a round, and the computation is distributed over all the sensors involved. In particular, when information averaging is concerned, these distributed algorithms are usually referred to as distributed consensus algorithms or gossip algorithms, e.g., [46–54]. Distributed consensus algorithms have been applied to a variety of network-related scenarios. For example, they have been considered in [53,54] for clock or time synchronization in wireless sensor networks.

Among the various distributed consensus algorithms, of particular note is the broadcast-based consensus algorithm, which has been studied in [50,51] without considering the stochastic disturbances, and more recently in [52] by taking non-zero-mean stochastic perturbations into account. Here we describe a broadcast consensus based phase synchronization algorithm for coherent MIMO radar systems, which was developed in [45], and is a direct extension of the work given in [52]. Similar to the aforescribed two phase synchronization approaches, this algorithm assumes no *a priori* time synchronization; thus, there is no common time scale established among all these sensors. Each radar element is assumed to operate according to its own local clock which ticks independently from others. Also, for the ease of exposition, we assume the carrier frequency synchronization is already established, while noting that this assumption can be also relaxed. Thus, when a sensor receives beacon signals from its neighbors, it only needs to estimate the received phase.

In this algorithm, each element broadcasts a reference signal containing its local phase information to all its immediate neighbors. Its neighbors then average their local phase with the received phase. Thus, no global operations or information exchanges are required in the process. As an inherent characteristic, this algorithm exploits the broadcast nature of the wireless communication environment, and obviates the need for sophisticated underlying media access control mechanisms. To this end, we briefly outline this phase synchronization algorithm. Its mathematical characterization and analysis is furnished in [45], which examines in detail how the propagation delays among neighboring sensors and the phase errors induced by inaccurate local estimation affect the phase synchronization accuracy:

1. Sensor S_i elects to broadcast an unmodulated reference signal.
2. Each neighbor of S_i , upon receiving the broadcast reference signal, generates an estimate of the received phase. If it is the first time for S_i to receive a broadcast reference signal, it updates its phase with a weighted value of the phase estimate; otherwise, it updates its phase with a weighted average of its current phase and the phase estimate.
3. The phase of each of the remaining sensors, including S_i , remains unchanged.
4. This procedure continues, with each radar element having the same probability to broadcast in every round.

Note that the absence of a common time reference implies that none of these sensors know the true time, and as a result, they do not know the “true” initial phase of their local oscillator. Thus, when a sensor receives a broadcast reference signal for the first time, we choose to update its phase using only the estimate of the received phase.

2.13.6.4 Discussion

As can be clearly seen, the master-slave closed-loop method is the simplest one among all these three approaches. It, however, has the potential to lead to the best phase synchronization accuracy. This is because there are only two (the least number) nodes which participate in each round of the phase synchronization process, i.e., the master node and one slave node. Thus, the propagating accumulation of each individual’s phase and frequency estimation errors is very limited, and is less of a concern. This, in fact, is corroborated by a rigorous phase synchronization error analysis included in [45]. Moreover, in this method, the phase offset induced by the propagation channel is explicitly compensated via the closed-loop operation, which further improves the phase synchronization accuracy. Therefore, this approach is particularly useful when a very precise phase synchronization is required. It is generally applicable to situations where a relatively powerful centralized entity exists or can be identified, and can be repeatedly used in systems having tree-like topologies or hierarchical structures. To complete the whole phase synchronization process, a total of $2K - 1$ transmissions are needed, including the broadcast transmission made by the master node.

Unlike the master-slave closed-loop method, both the round-trip and the broadcast consensus based algorithms do not require a powerful master entity. However, the round-trip scheme does need a global initiator to start the phase synchronization process. In theory, the global initiator can be any radar sensor in the system. This round-trip method is able to synchronize all the radar sensors to the same mean carrier phase, regardless of the topology of these sensors as long as a cycle traversing all these sensors can be identified. But since sensors in the system participate in the phase synchronization process all together in a cooperative fashion, this round-trip method may suffer from the propagating accumulation of each individual’s phase and frequency estimation errors, as mathematically demonstrated in [45]. As a result, this method may only be suitable to coherent MIMO radar systems with a limited number of sensors. In other words, it does not scale very well. This round-trip method needs a total of $2K$ transmissions to complete the phase synchronization process.

The broadcast consensus based method is distributed and localized in nature. It needs neither a centralized entity nor a particular global initiator. It is scalable and applicable to coherent MIMO radar systems composed of a large number of densely scattered sensors. It is also robust to sensor failures. But unfortunately, as shown in [45], the functionality of this approach is highly dependent on, and varies from topology to topology of the radar sensors. For arbitrary topologies, this approach may be only useful when a coarse phase synchronization is required. Since this algorithm runs in an iterative manner, it generally requires more transmissions than the other two approaches to complete the whole synchronization process. Thus, it has to be implemented at a relaxed time constraint as opposed to a stringent one.

Of further note is that in each of the presented phase synchronization algorithms, the propagation delay for the channel between any pair of MIMO radar elements that are involved in the reference signal exchange, can be critical in determining the resulting phase synchronization accuracy. A small

uncertainty in the channel propagation delay, which may result from, for example, the ambiguity in the location of certain radar elements or the mobile scatters and other propagation anomalies, can lead to intolerable phase variations across some or all the radar elements. This is particularly true when high frequency carriers are chosen as the reference signals. One way to alleviate this unfavorable effect is to use relatively low carrier frequency. If the carrier frequency is low enough, the location uncertainty of the MIMO radar elements will not play an important role. Another way is to recalibrate these sensors periodically to track the channel changes. This may induce some extra overhead, but the accompanying benefits are in general well worth it. Besides, the variations of the propagation channels among all these MIMO radar elements, though dependent on the specific operating environment, are usually very slow; it takes a relatively long time to reach the point where the physical changes in the medium are so pronounced that the resulting phase mismatch seriously deteriorates the performance of the coherent processing, especially for applications such as target localization [11]. Therefore, very frequent phase resynchronization may not be needed.

Another practical issue that merits attention is the phase dynamics or phase instability [55] in the coherent MIMO radar system. Note that the oscillator phase noise [56] is always present, and is an irreducible error which can affect the phase synchronization performance. To counter the effect of phase noise in oscillators, periodic phase synchronization remains indispensable. How often the phase synchronization procedure needs to be repeated, however, is dependent on, among other factors, the aggregate impact of the phase noise across all the local oscillators on the phase synchronization accuracy, as well as the phase accuracy or stability requirements of the coherent MIMO radar system and the specific application. It is indicated in [34, p. 261] that the phase accuracy or stability requirements for coherent processing by a monostatic or bistatic radar can range from less than one degree to many tens of degrees of RF phase over a coherent processing interval, depending on the type and duration of coherent processing. We conjecture that similar phase accuracy or stability requirements may hold as well for coherent processing by a MIMO radar system.

2.13.7 Waveform design for MIMO radar

Most radar systems operate by radiating a specific electromagnetic signal into a region and detecting the echo returned from the reflecting targets. The nature of the echo signal provides information about the target, such as range, radial velocity, angular direction, size, shape and so on [57]. This signal is usually referred to as the radar waveform, and plays a key role in the accuracy, resolution, and ambiguity for radar in performing the above mentioned tasks [58]. For MIMO radar, it also relies on waveforms to realize its excellent performance advantages. Thus, the design of appropriate waveforms for MIMO radar has become a crucial task, which has attracted a lot of attention over the years. Recent work on waveform design for MIMO radar (with both widely separated antennas and colocated antennas) includes, among others, [9, 12, 59–74].

In this section, we will briefly describe a specific waveform design scheme for MIMO radar with widely separated antennas. This scheme was originally developed by us in [9], and further improved or extended later in, for example [59, 60]. It is primarily focused on optimizing MIMO radar's performance in information extraction and estimation of extended targets which are modeled using an impulse response as in [75] (point targets are a special case). Assuming that the radar transmitter has some

knowledge (full or partial) of the target's second-order statistics and that the transmitted power is constrained, Yang and Blum [9] detail methods for finding waveform design strategies based on the following two criteria: (1) maximizing the conditional mutual information (MI) between the random target impulse response and the reflected radar waveforms; (2) minimizing the value of minimum mean-square error (MMSE) in estimating the target impulse response. In this part, we will concisely summarize some major results of [9], and shed some light on the waveform design problem for MIMO radar.

2.13.7.1 Signal model and design criteria

Figure 13.14 illustrates a bistatic MIMO radar scenario, where the radar is equipped with P transmit elements and Q receive elements. This model, appropriate for both bistatic and monostatic scenarios, is a slight generalization of that in [2] for the case of extended targets. For simplicity we consider this model in discrete-time and in baseband. For an extended target, the reflection from the signal sent from the p th transmit element and captured at the q th receive element can be modeled using a finite impulse response (FIR) linear system with order ν , whose impulse response is $g^{(p,q)}(l)$, $l \in [0, \nu]$. Then the received waveform at the q th receive element and discrete-time k is given by

$$y_q(k) = \sum_{p=1}^P \sum_{l=0}^{\nu} g^{(p,q)}(l)x_p(k-l) + n_q(k), \quad (13.97)$$

where $x_p(k)$ is the waveform transmitted from the p th transmit element and $n_q(k)$ is the additive complex Gaussian noise measured at the q th receive element.

Let L denote the length of the observed signal vector starting at an arbitrary discrete time k , and assume that $L > PM$ (typically $L \gg PM$). Denote $\mathbf{g}^{(p,q)} = [g^{(p,q)}(0), \dots, g^{(p,q)}(\nu)]^T$,

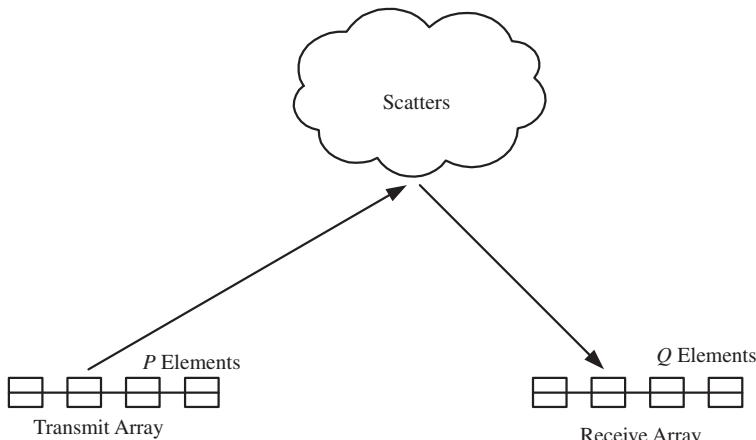


FIGURE 13.14

Illustration of MIMO radar for a bistatic radar scenario.

$\mathbf{y}_q = [y_q(k), \dots, y_q(k+L-1)]^T$ and $\mathbf{n}_q = [n_q(k), \dots, n_q(k+L-1)]^T$, then using the matrix-vector notation (13.97) can be rewritten into

$$\mathbf{y}_q = \sum_{p=1}^P \mathbf{X}_p \mathbf{g}^{(p,q)} + \mathbf{n}_q, \quad (13.98)$$

where \mathbf{X}_p is an $L \times M (M = v + 1)$ Toeplitz matrix which contains the waveforms transmitted from the p th transmit element, i.e.,

$$\mathbf{X}_p = \begin{pmatrix} x_p(k) & \cdots & x_p(k-v) \\ \vdots & \ddots & \vdots \\ x_p(k+L-1) & \cdots & x_p(k+L-1-v) \end{pmatrix}. \quad (13.99)$$

We further define $\bar{\mathbf{X}} = [\mathbf{X}_1, \dots, \mathbf{X}_P]$ and $\bar{\mathbf{g}}_q = [(\mathbf{g}^{(1,q)})^T, \dots, (\mathbf{g}^{(P,q)})^T]^T$, so that (13.98) could be reformulated into

$$\mathbf{y}_q = \bar{\mathbf{X}} \bar{\mathbf{g}}_q + \mathbf{n}_q. \quad (13.100)$$

By stacking the received waveforms across all the Q receive elements, we create $\bar{\mathbf{y}} = [\mathbf{y}_1^T, \dots, \mathbf{y}_Q^T]^T$. Defining $\mathcal{X} = \mathbf{I}_Q \otimes \bar{\mathbf{X}}$, we finally obtain

$$\bar{\mathbf{y}} = \mathcal{X} \bar{\mathbf{g}} + \bar{\mathbf{n}}, \quad (13.101)$$

where $\bar{\mathbf{g}} = [\bar{\mathbf{g}}_1^T, \dots, \bar{\mathbf{g}}_Q^T]^T$ and $\bar{\mathbf{n}} = [\mathbf{n}_1^T, \dots, \mathbf{n}_Q^T]^T$.

To facilitate the ensuing analysis on the waveform design of MIMO radar for extended targets, the model in (13.101) is assumed to have the following properties [9]:

Assumption 4. *The target impulse response vector $\bar{\mathbf{g}}$ is a Gaussian random vector with zero mean and full rank covariance matrix $\Sigma_{\bar{\mathbf{g}}}$.*

Assumption 5. *Components of the noise vector $\bar{\mathbf{n}}$ are assumed to be independently and identically distributed (i.i.d) and complex Gaussian, with zero mean and variance σ_n^2 .*

Note that $\Sigma_{\bar{\mathbf{g}}}$ can be diagonalized through its eigenvalue decomposition, i.e.,

$$\Sigma_{\bar{\mathbf{g}}} = \mathbf{U} \Lambda \mathbf{U}^H, \quad (13.102)$$

where \mathbf{U} is a unitary matrix whose columns are eigenvectors and $\Lambda = \text{diag}\{\Lambda_{11}, \dots, \Lambda_{PQM,PQM}\}$ is a diagonal matrix with each diagonal entry given by a real and nonnegative eigenvalue.

Now let us introduce the design criteria. The first one is the conditional mutual information (MI) between $\bar{\mathbf{y}}$ and $\bar{\mathbf{g}}$ given the knowledge of $\bar{\mathbf{X}}$, which is hereafter referred to as MI. We have [76] ($h(\cdot)$ denotes differential entropy)

$$I = I(\bar{\mathbf{y}}; \bar{\mathbf{g}} | \bar{\mathbf{X}}) = h(\bar{\mathbf{y}} | \bar{\mathbf{X}}) - h(\bar{\mathbf{y}} | \bar{\mathbf{g}}, \bar{\mathbf{X}}). \quad (13.103)$$

Conditioned on $\bar{\mathbf{X}}$ (equivalent to conditioning on \mathcal{X}), we can easily find that $\bar{\mathbf{y}}$ is Gaussian distributed with zero mean and covariance $(\mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \sigma_n^2\mathbf{I}_{QL})$. Using (13.103), the MI is [76]

$$\begin{aligned}\mathcal{I} &= \log \left[\det (\mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \sigma_n^2\mathbf{I}_{QL}) \right] - \log \left[\det (\sigma_n^2\mathbf{I}_{QL}) \right] \\ &= \log \left[\det (\sigma_n^{-2}\mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \mathbf{I}_{QL}) \right] \\ &= \log \left[\det (\sigma_n^{-2}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H\mathcal{X} + \mathbf{I}_{PQM}) \right],\end{aligned}\quad (13.104)$$

where (13.104) follows from

$$\det (\mathbf{I}_r + \mathbf{AB}) = \det (\mathbf{I}_t + \mathbf{BA}). \quad (13.105)$$

The design objective is to find those \mathcal{X} (transmitted waveform) that maximize the MI (\mathcal{I}) between the random target impulse response $\bar{\mathbf{g}}$ and the received (reflected) radar waveform $\bar{\mathbf{y}}$ under the constraint $\text{tr}\{\mathcal{X}^H\mathcal{X}\} \leq LQ\mathcal{P}_0$ which effectively limits the total transmit power. Therefore, the problem of waveform design based on MI can be cast as:

$$\begin{aligned}\max_{\mathcal{X}} \quad &\det (\sigma_n^{-2}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H\mathcal{X} + \mathbf{I}_{PQM}) \\ \text{s.t.} \quad &\text{tr}\{\mathcal{X}^H\mathcal{X}\} \leq LQ\mathcal{P}_0.\end{aligned}$$

Next, let us consider the problem of radar waveform design from the viewpoint of estimation. It is easy to verify that, conditioned on \mathcal{X} , $\bar{\mathbf{g}}$, and $\bar{\mathbf{y}}$ are jointly Gaussian distributed as

$$\begin{bmatrix} \bar{\mathbf{g}} \\ \bar{\mathbf{y}} \end{bmatrix} \sim \mathcal{CN} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_{\bar{\mathbf{g}}} & \Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H \\ \mathcal{X}\Sigma_{\bar{\mathbf{g}}} & \mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \sigma_n^2\mathbf{I}_{QL} \end{bmatrix} \right).$$

The conditional distribution of $\bar{\mathbf{g}}$ given $\bar{\mathbf{y}}$ and \mathcal{X} is also Gaussian, with conditional mean $\hat{\mu}_{\bar{\mathbf{g}}}$ given by (using Eqs. (IV.B.53) and (IV.B.55) on p. 156 of [17])

$$\begin{aligned}\hat{\mu}_{\bar{\mathbf{g}}} &= \Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H(\mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \sigma_n^2\mathbf{I}_{QL})^{-1}\bar{\mathbf{y}} \\ &= (\mathcal{X}^H\mathcal{X} + \sigma_n^2\Sigma_{\bar{\mathbf{g}}}^{-1})^{-1}\mathcal{X}^H\bar{\mathbf{y}}\end{aligned}\quad (13.106)$$

and conditional covariance matrix $\hat{\Sigma}_{\bar{\mathbf{g}}}$ given by (using Eq. (IV.B.54) on pp. 156 of [17])

$$\begin{aligned}\hat{\Sigma}_{\bar{\mathbf{g}}} &= \mathbb{E} \left\{ (\bar{\mathbf{g}} - \hat{\mu}_{\bar{\mathbf{g}}})(\bar{\mathbf{g}} - \hat{\mu}_{\bar{\mathbf{g}}})^H \right\} \\ &= \Sigma_{\bar{\mathbf{g}}} - \Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H \left(\mathcal{X}\Sigma_{\bar{\mathbf{g}}}\mathcal{X}^H + \sigma_n^2\mathbf{I}_{QL} \right)^{-1} \mathcal{X}\Sigma_{\bar{\mathbf{g}}} \\ &= \left(\sigma_n^{-2}\mathcal{X}^H\mathcal{X} + \Sigma_{\bar{\mathbf{g}}}^{-1} \right)^{-1},\end{aligned}\quad (13.107)$$

where the *matrix inversion lemma*¹¹ was employed in obtaining both (13.106) and (13.107).

¹¹ $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1}$.

Let $\hat{\bar{\mathbf{g}}}$ denote the Bayes estimate of $\bar{\mathbf{g}}$. When the cost function is defined as $\|\bar{\mathbf{g}} - \hat{\bar{\mathbf{g}}}\|^2$, the Bayes estimate $\hat{\bar{\mathbf{g}}}$ will be a linear MMSE estimator [17] which is given by the conditional mean of $\bar{\mathbf{g}}$ given $\bar{\mathbf{y}}$ (for a given value of \mathcal{X}), i.e.,

$$\hat{\bar{\mathbf{g}}} = \hat{\mu}_{\bar{\mathbf{g}}} = \left(\mathcal{X}^H \mathcal{X} + \sigma_n^2 \Sigma_{\bar{\mathbf{g}}}^{-1} \right)^{-1} \mathcal{X}^H \bar{\mathbf{y}}. \quad (13.108)$$

In this case, the Bayes risk for a given \mathcal{X} will be

$$\begin{aligned} \xi &= \mathbb{E}\{\|\bar{\mathbf{g}} - \hat{\bar{\mathbf{g}}}\|^2\} = \text{tr}\{\hat{\Sigma}_{\bar{\mathbf{g}}}\} \\ &= \text{tr}\left\{\left(\sigma_n^{-2} \mathcal{X}^H \mathcal{X} + \Sigma_{\bar{\mathbf{g}}}^{-1}\right)^{-1}\right\}. \end{aligned} \quad (13.109)$$

The goal is then to find those \mathcal{X} that minimize the value of MMSE under a constraint on the total transmit power of $\text{tr}\{\mathcal{X}^H \mathcal{X}\} \leq LQ\mathcal{P}_0$. Or equivalently, the problem of waveform design based on MMSE estimation can be expressed as:

$$\begin{aligned} \min_{\mathcal{X}} \quad &\text{tr}\{(\sigma_n^{-2} \mathcal{X}^H \mathcal{X} + \Sigma_{\bar{\mathbf{g}}}^{-1})^{-1}\} \\ \text{s.t.} \quad &\text{tr}\{\mathcal{X}^H \mathcal{X}\} \leq LQ\mathcal{P}_0. \end{aligned}$$

2.13.7.2 Optimal waveform design for MIMO radar

As mathematically demonstrated in [9], the maximum value of \mathcal{I} will be achieved when the matrix $(\sigma_n^{-2} \Sigma_{\bar{\mathbf{g}}} \mathcal{X}^H \mathcal{X} + \mathbf{I}_{PQM})$ is diagonal, and the minimum value of ξ will be attained when $(\sigma_n^{-2} \mathcal{X}^H \mathcal{X} + \Sigma_{\bar{\mathbf{g}}}^{-1})$ is diagonal, also. The following theorem summarizes the result of optimum \mathcal{X} . A detailed proof of this result can be found in [9].

Theorem 3. *Let Assumptions 4 and 5 hold true. Then, under the constraint $\text{tr}\{\mathcal{X}^H \mathcal{X}\} \leq LQ\mathcal{P}_0$, only \mathcal{X} that satisfy*

$$\mathcal{X} = \Psi \left(\text{diag} \left[\left(\eta - \frac{\sigma_n^2}{\Lambda_{11}} \right)^+, \dots, \left(\eta - \frac{\sigma_n^2}{\Lambda_{PQM,PQM}} \right)^+ \right] \right)^{1/2} \mathbf{U}^H, \quad (13.110)$$

can be optimum in the sense of either maximizing the MI or minimizing the MMSE. In (13.110), Λ and \mathbf{U} are defined in (13.102); Ψ is an $LQ \times PQM$ matrix with orthonormal columns; the scalar constant η is chosen to satisfy

$$\sum_{i=1}^{PQM} \left(\eta - \frac{\sigma_n^2}{\Lambda_{ii}} \right)^+ = LQ\mathcal{P}_0.$$

The resulting maximum value of MI is

$$\mathcal{I}_{max} = \sum_{i=1}^{PQM} \left(\log (\sigma_n^{-2} \Lambda_{ii} \eta) \right)^+,$$

and the corresponding minimum value of MMSE is

$$\xi_{min} = \sum_{i=1}^{PQM} \frac{\Lambda_{ii}}{(\Lambda_{ii}\sigma_n^{-2}\eta - 1)^+ + 1}.$$

Further, provided $\mathbf{X} = \mathbf{I}_Q \otimes \bar{\mathbf{X}}$ for some $\bar{\mathbf{X}}$, then any \mathbf{X} satisfying (13.110) is optimum for both the MI and MMSE criteria.

A review of *Theorem 3* indicates that knowledge of the full matrix $\Sigma_{\bar{\mathbf{g}}}$ is needed to perform an eigenvalue decomposition. An asymptotic formulation has been developed in [9] which lessens the required information to just a few samples of the PSD and would be much more suitable in practice. The asymptotic approach employs one important assumption as shown below:

Assumption 6. *The covariance matrix $\Sigma_{\bar{\mathbf{g}}}$ of $\bar{\mathbf{g}}$ has a Toeplitz structure.*

Clearly this assumption is not very restrictive for single transmit and receive antenna cases, implying that the target impulse response samples form a wide-sense stationary random process asymptotically. For multiple transmit and receive antenna cases, this assumption implies that the statistics over space are indistinguishable from the statistics over time. For the extreme cases of complete independence or dependence over both space and time, the assumption is very reasonable, for example. However, it is also straightforward to relax this assumption for multiple-antenna cases by employing a similar asymptotic approximation as that described below. For a further discussion on this assumption, readers are referred to [59].

It is shown in [77] that Toeplitz matrices can be approximated by their associated circulant matrices, and asymptotically (in the dimension of these matrices), they are equivalent. The asymptotic equality of two matrices implies that their eigenvalues and inverses (and certain products) behave similarly [77]. Now, let us denote the (i, j) th entry of the covariance matrix $\Sigma_{\bar{\mathbf{g}}}$ as $\Sigma_{\bar{\mathbf{g}}}(i, j)$, then we have $\Sigma_{\bar{\mathbf{g}}}(i, j) = r_{i-j}$ and $r_{-k} = r_k^*$, where the sequence $\{r_k; k = 0, \pm 1, \dots, \pm(PQM - 1)\}$ is absolutely summable, and its 2π -periodic truncated Fourier spectrum

$$f(\omega) = \sum_{k=1-PQM}^{PQM-1} r_k e^{-jk\omega}$$

is real valued, finite, and nonnegative, and represents the PSD of the random process $\bar{\mathbf{g}}$ when PQM approaches infinity. Then a circulant matrix $\tilde{\Sigma}_{\bar{\mathbf{g}}}$ that is asymptotically equivalent to $\Sigma_{\bar{\mathbf{g}}}$ is [78]

$$\tilde{\Sigma}_{\bar{\mathbf{g}}} = \mathbf{F}_{PQM} \mathbf{V} \mathbf{F}_{PQM}^H, \quad (13.111)$$

where \mathbf{F}_{PQM} is the $PQM \times PQM$ unitary discrete Fourier transform (DFT) matrix with its (k, l) th entry given by

$$\frac{1}{\sqrt{PQM}} \exp \left\{ \frac{-j2\pi(k-1)(l-1)}{PQM} \right\}, \quad \forall k, l \in [1, PQM];$$

and the diagonal matrix \mathbf{V} contains samples of the PSD of $\bar{\mathbf{g}}$ along its diagonal, or equivalently

$$\begin{aligned}\mathbf{V}_{ii} &= f(2\pi(i-1)/(PQM)) \\ &= r_0 + 2 \Re \left\{ \sum_{k=1}^{PQM-1} r_k e^{-j2\pi(i-1)/(PQM)} \right\},\end{aligned}$$

where $i = 1, \dots, PQM$.

Given the above asymptotic approximation, the waveform design problems posed earlier can be resolved using standard optimization theory. For brevity, we summarize the results in the following theorem. The proof can be found in [9].

Theorem 4. Suppose Assumptions 4–6 hold true and perfect knowledge is available on the samples of the target PSD (i.e., entries in the diagonal matrix \mathbf{V}). Then under a constraint $\text{tr}\{\mathbf{X}^H \mathbf{X}\} \leq LQ\mathcal{P}_0$, only a signaling matrix \mathbf{X} that satisfies

$$\mathbf{T} = \mathbf{X}^H \mathbf{X} = \mathbf{F}_{PQM} \mathbf{D} \mathbf{F}_{PQM}^H \quad (13.112)$$

can be asymptotically optimum in the sense of either maximizing the MI or minimizing the MMSE, where we have

$$\mathbf{D} = \text{diag} \left[\left(\eta - \frac{\sigma_n^2}{\mathbf{V}_{11}} \right)^+, \dots, \left(\eta - \frac{\sigma_n^2}{\mathbf{V}_{PQM,PQM}} \right)^+ \right], \quad (13.113)$$

and the constant η is chosen to satisfy

$$\sum_{i=1}^{PQM} \left(\eta - \frac{\sigma_n^2}{\mathbf{V}_{ii}} \right)^+ = LQ\mathcal{P}_0. \quad (13.114)$$

Further, any \mathbf{X} from (13.112) is asymptotically optimum for both MI and MMSE.

In fact, Theorem 4 is simply a modification of Theorem 3 for the asymptotic case. The essential change is that \mathbf{F}_{PQM} replaces \mathbf{U} and \mathbf{V} replaces $\mathbf{\Lambda}$. Under such asymptotic approximation, the solution for \mathbf{X} can be explicitly written as

$$\mathbf{X} = \mathbf{\Psi} \left(\text{diag} \left[\left(\eta - \frac{\sigma_n^2}{\mathbf{V}_{11}} \right)^+, \dots, \left(\eta - \frac{\sigma_n^2}{\mathbf{V}_{PQM,PQM}} \right)^+ \right] \right)^{1/2} \mathbf{F}_{PQM}^H. \quad (13.115)$$

It is obvious that to produce \mathbf{X} in (13.115), the full covariance matrix $\Sigma_{\bar{g}}$ is not needed, only samples of the PSD for \mathbf{V} are necessary. This is a more reasonable assumption in practice.

2.13.7.3 Discussion

Based on Theorem 3, it can be concluded that any $\mathbf{X}^H \mathbf{X}$ meeting the given power constraint which maximizes MI will also minimize MMSE, and this $\mathbf{X}^H \mathbf{X}$ is unique for a given $\Sigma_{\bar{g}}$. In fact, $\mathbf{X}^H \mathbf{X}$ can be deemed as the covariance matrix of the transmitted waveforms. Meanwhile, it should be noted that the

solution expressed by (13.110) is a closed-form and general solution which does not require any level of approximation. Its optimality is manifested in two ways. On the one hand, only \mathcal{X} satisfying (13.110) can be optimum in the sense of either maximizing the MI or minimizing the MMSE. On the other hand, any \mathcal{X} from (13.110) is optimum for both MI and MMSE provided such \mathcal{X} satisfies $\mathcal{X} = \mathbf{I}_Q \otimes \bar{\mathbf{X}}$. In other words, the condition stipulated by (13.110) is only a necessary condition for \mathcal{X} to attain the optimality, but not a sufficient one. This is because there may exist some cases where a solution \mathcal{X} satisfies (13.110) but does not have the desired Kronecker structure. Note for these cases, it can be possible that the equivalence between the MI and the MMSE does not hold any more. However, in a recently published paper [60], an iterative algorithm has been developed, which can identify waveform solutions that not only optimize the performance criteria of interest (i.e., MI and MMSE), but have a specific Kronecker structure. It is also found in [60] that waveform solutions generated through this algorithm can lead to performance which is very close to, and almost indistinguishable from that predicted by *Theorem 3*. Detailed information about this algorithm can be found in [60].

A close review of *Theorem 4* reveals that for the asymptotic simplification approach, exact knowledge of the target PSD samples is required in the waveform design. This requirement appears rather demanding, particularly in practical applications. In order to bring the waveform design even closer to practice and deliver more applicable waveform solutions for MIMO radar, it is necessary to accommodate some issues that may arise in reality, such as the uncertainty in the target's statistics. In such a circumstance, robust procedures, which can overcome those problems by incorporating modeling uncertainty into the design from the outset [79], seem quite attractive. As a result, a minimax robust waveform design problem for MIMO radar has been formulated and solved in [59], where a band model was adopted by assuming the PSD lies in an uncertainty class of spectra bounded by known upper and lower bounds. This band model might arise in practice, for example, if a confidence band for the spectrum could be determined via spectrum estimation, or the upper and lower bounds can be obtained simply by field measurement and modeling. It is shown in [59] that the resulting minimax robust waveforms can achieve good performance for any PSD in the uncertainty class, and in particular, they can bound the worst-case performance at an acceptable limit. One very interesting finding from the results in [59] is that the minimax robust waveforms for the MMSE criterion are generally different from those for the MI criterion, which is in stark contrast to the results for the completely known PSD case. We direct interested readers to [59] for more details.

2.13.8 Summary

In this document, we provided a review of some key topics related to MIMO radar with widely separated antennas. This type of MIMO radar features two major configurations, one with all antennas within a same target beamwidth and one with antennas placed in distinct target beamwidths, which, in turn, lead to two different operation modes called coherent and noncoherent MIMO radars, respectively. In Sections 2.13.2 and 2.13.3, signal models for coherent and noncoherent MIMO radars were introduced for the problem of joint target location and velocity estimation. Equations specifying the ML estimates were derived for some general cases, which accommodate scenarios featuring, for example, nonorthogonal transmitted waveforms, spatially colored clutter-plus-noise, and correlated reflection coefficients. However, to provide further insight, we also considered some simplified cases, where orthogonal waveforms,

temporally and spatially white clutter-plus-noise, and independent reflections are assumed. The MSEs of the ML estimates, the corresponding CRBs, and the coherent and noncoherent AFs were presented.

In Section 2.13.4, the performance and complexity issues for the coherent and noncoherent MIMO radars were investigated. Although the coherent MIMO radar outperforms the noncoherent one in general, it has more stringent operational requirement and incurs more implementation complexity. Fortunately, our results indicate that for some cases, e.g., when MN is sufficiently large, we can replace the coherent approach by its noncoherent counterpart, which, while reducing a great deal the implementation complexity, induces very little loss of performance. As a result, if large MN can be afforded in practice, the noncoherent MIMO radar is naturally preferred to the coherent one. The NDRMSE was defined in this section, and was used to compare the overall performance difference between the coherent and noncoherent MIMO radars for joint parameter estimation. A sufficiently small value of the parameter NDRMSE can serve as an indicator, suggesting the noncoherent MIMO radar performs very well relative to the coherent one.

The diversity gains were analyzed for MIMO radar target detection under the Neyman-Pearson criterion in Section 2.13.5. We showed that the maximum achievable diversity gain for a MIMO radar system is determined by the number of transmit antennas, the number of scatterers that compose the target, the dimension spanned by the clutter-plus-noise free received waveforms at each receiver, and the cdf of the reflection coefficients, while it is invariant to the clutter-plus-noise cdf, provided the magnitude of the processed clutter-plus-noise has bounded moments up to a required order. It was shown that for some cases of interest, the maximum diversity gain can be achieved. Further, properly chosen nonorthogonal waveforms can achieve the same diversity gain as orthogonal waveforms.

In Section 2.13.6, we discussed three effective approaches to achieve phase synchronization in coherent MIMO radar systems. The first one is the master-slave closed-loop approach, which employs a master-slave architecture. This method is very simple, and is suitable for fine phase synchronization. The second one is the round-trip algorithm, which employs an unmodulated beacon signal to travel through all the radar elements in a round-trip manner. This method is applicable to arbitrary topologies of the radar sensors as long as a cycle traversing all these sensors can be identified. The third one is the broadcast consensus based algorithm, which is scalable, and is distributed and localized in nature. Distinct as they are, these algorithms are all time asynchronous, and in particular, the master-slave closed-loop algorithm and the round-trip algorithm also require no *a priori* frequency synchronization. These algorithms should cater to different needs arising as a result of different number of MIMO radar elements, different sensor placements, as well as different operational preferences.

Several MIMO radar waveform design methods were presented in Section 2.13.7. The discussed methods capitalize on the knowledge of the covariance matrix of the extended target impulse response, and optimize two criteria: maximization of the MI and minimization of the MMSE. For both design criteria, a constraint on the transmitted power is imposed. The findings indicate that these two criteria are equivalent in the sense that they lead to the same covariance matrix of the transmit waveforms, which is a fundamental quantity specifying the waveform design. The optimum solution employs water-filling, which allocates the transmitted power in proportion to the quality of the particular mode in question. In particular, more transmitted waveform power is allocated to modes that have higher power, indicating the presence of significant target scattering, and modes with low power deserve excitation with lower transmitted waveform power. In addition, an asymptotic formulation was also provided in Section 2.13.7, which lessens the required knowledge about the target statistical model to just a few samples of the PSD, and would be much more suitable in practice.

Notation

Throughout this chapter, we use bold upper case letters to denote matrices, and bold lower case letters to signify column vectors. The definition of symbols used in this chapter are listed below:

$\mathbb{E}\{\cdot\}$	expectation with respect to all the random variables within the brackets
$\{\cdot\}^H$	conjugate transpose of the argument
$\{\cdot\}^*$	conjugate of the argument
$\{\cdot\}^T$	transpose of the argument
$(a)^+$	positive part of a , i.e., $(a)^+ = \max[0, a]$
$ \cdot $	absolute value of a complex number
$\ \cdot\ $	Euclidean norm of the argument
$\Re\{\cdot\}$	real part of a complex number
$\Im\{\cdot\}$	imaginary part of a complex number
$o(\cdot)$	little-o notation, i.e., Suppose $x(u)$ and $y(u)$ are both functions of a variable u , then $x(u) = o(y(u))$ as $u \rightarrow a$ implies $\lim_{u \rightarrow a} x(u)/y(u) = 0$
$\delta(\tau)$	unit impulse function, i.e., $\delta(\tau) = 1$ if $\tau = 0$ or 0 if $\tau \neq 0$
$\mathbf{1}_{a \times b}$	all one matrix of size $a \times b$
$\mathbf{0}_{a \times b}$	zero matrix of size $a \times b$
\mathbf{I}_K	identity matrix of size $K \times K$
$\det\{\cdot\}$	determinant of a matrix
$\text{tr}\{\cdot\}$	trace of a matrix
$\text{diag}\{a_1, a_2, \dots\}$	diagonal matrix with elements a_1, a_2, \dots on its diagonal
$\text{Diag}\{\mathbf{S}_1, \mathbf{S}_2, \dots\}$	block diagonal matrix with submatrices $\mathbf{S}_1, \mathbf{S}_2, \dots$ on its diagonal
$\nabla_{\theta}\{\cdot\}$	gradient of the argument with respect to θ
\otimes	Kronecker product operator

Appendix

A.1 Frequency spread signals

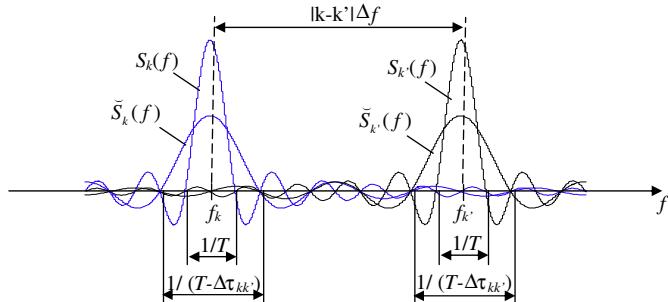
In this appendix, we describe one class of signals which could be employed. Under appropriate assumptions, these signals approximately satisfy the conditions of *Assumption 1* for orthogonal signals. We discuss these assumptions. These assumptions essentially ignore small sidelobes in the Fourier transform of these signals to approximate them with truncated (in frequency domain) signals.

Consider a set of pulsed sinusoidal signals¹²

$$s_k(t) = \begin{cases} \frac{1}{\sqrt{T}} \exp(j2\pi k \Delta f t) & 0 < t < T, \\ 0 & \text{otherwise,} \end{cases} \quad (\text{A.1})$$

where $k = 1, \dots, M$, the term T denotes the signal time duration, and $\Delta f = f_{k+1} - f_k \geq 0$ is the frequency increment between $s_k(t)$ and $s_{k+1}(t)$. The approximate bandwidth of each signal is $F = 1/T$.

¹²Windowing can be applied in practice but we omit discussion of this well studied topic.

**FIGURE A.1**

Frequency spectrum for transmitted signals (solid curves) and the effective parts of their time delayed versions (dotted curves).

The spectrum of $s_k(t)$ and $s_{k'}(t)$, $S_k(f)$ and $S_{k'}(f)$, are shown in Figure A.1. Note that throughout this appendix it is always assumed that $k \neq k'$. According to Parseval's theorem

$$\int_T s_k(t)s_{k'}^*(t)dt = \int_{-\infty}^{\infty} S_k(f)S_{k'}^*(f)df. \quad (\text{A.2})$$

Observing the spectrum, we see that when $|k - k'| \Delta f \gg \frac{1}{T}$, the main lobes of $S_k(f)$ and $S_{k'}(f)$ are sufficiently separated so that $S_k(f)S_{k'}^*(f) \approx 0$ and (A.2) approximates 0, which makes $s_k(t)$ and $s_{k'}(t)$ approximately orthogonal. Here we are ignoring contributions from small sidelobes which is similar to truncating the signals in the frequency domain. As $|k - k'| \geq 1$, if the frequency increment satisfies

$$\Delta f \gg \frac{1}{T}, \quad (\text{A.3})$$

then any two signals from the signal set $s_k(t)$ ($k = 1, \dots, M$) are approximately orthogonal.

Assume that $s_k(t)$ and $s_{k'}(t)$ are time delayed by τ_k and $\tau_{k'}$. The non-zero part of the delayed signal $s_k(t - \tau_k)$ occupies $\tau_k < t < \tau_k + T$, and the non-zero part of $s_{k'}(t - \tau_{k'})$ occupies $\tau_{k'} < t < \tau_{k'} + T$. Hence, the time duration of the overlap between them is $T - \Delta\tau_{kk'}$, where $\Delta\tau_{kk'} = |\tau_k - \tau_{k'}|$ denotes the time delay difference. Note that when $\Delta\tau_{kk'}$ is greater than or approximates T , the overlap part of these two signals vanishes, which implies (A.2) is close to 0 and the signals can be considered as orthogonal. Therefore, it is possible for the two delayed signals to be non-orthogonal only if $\Delta\tau_{kk'} \ll T$, which makes the following discussion necessary.

We denote the overlap interval as T_Δ . Thus

$$\begin{aligned} \int_T s_k(t - \tau_k)s_{k'}^*(t - \tau_{k'})dt &= \int_{T_\Delta} s_k(t - \tau_k)s_{k'}^*(t - \tau_{k'})dt \\ &= \int_{T_\Delta} \check{s}_k(t)\check{s}_{k'}^*(t)dt, \end{aligned}$$

where $\check{s}_k(t)$ and $\check{s}_{k'}(t)$ represent the overlapping parts of $s_k(t - \tau_k)$ and $s_{k'}(t - \tau_{k'})$, the time duration of which are both $T - \Delta\tau_{kk'}$. More specifically, suppose $\tau_{k'} < \tau_k < \tau_{k'} + T$, then we have

$$\begin{aligned}\check{s}_k(t) &= s_k(t - \tau_k), \quad \tau_k < t < \tau_{k'} + T, \\ \check{s}_{k'}(t) &= s_{k'}(t - \tau_{k'}), \quad \tau_k < t < \tau_{k'} + T.\end{aligned}$$

The spectrum of $\check{s}_k(t)$ and $\check{s}_{k'}(t)$, i.e., $\check{S}_k(f)$ and $\check{S}_{k'}(f)$, are plotted by dotted curves in Figure A.1. In a similar way, for the time delayed signals, the orthogonality is approximately maintained if

$$|k - k'| \Delta f \gg \frac{1}{T - \Delta\tau_{kk'}}. \quad (\text{A.4})$$

Since $|k - k'| \geq 1$, the inequality is maintained by requiring $\Delta f \gg \frac{1}{T - \Delta\tau_{kk'}}$. As $\Delta\tau_{kk'} \ll T$, the condition becomes $\Delta f \gg \frac{1}{T}$, which is the same as in (A.3). Hence, the requirement for (A.3) guarantees (A.4). That is, if (A.3) is met, the time delayed signals are orthogonal.

Next, we further discuss the Doppler shift restriction for orthogonality. The effective time delayed and Doppler shifted signals are given by $\tilde{s}_k(t) = \check{s}_k(t)e^{j2\pi f_{dk}t}$ and $\tilde{s}_{k'}(t) = \check{s}_{k'}(t)e^{j2\pi f_{dk'}t}$. The frequency spectrum of the effective time delayed signals $\tilde{s}_k(t)$ and $\tilde{s}_{k'}(t)$, i.e., $\tilde{S}_k(f)$ and $\tilde{S}_{k'}(f)$, are replotted in Figure A.2 using solid curves, the center frequency of which are f_k and $f_{k'}$. After Doppler frequency shift, the spectrum of the effective time delayed and Doppler shifted signals are $\tilde{S}_k(f) = S_k(f - f_{dk})$ and $\tilde{S}_{k'}(f) = S_{k'}(f - f_{dk'})$. They are shown in Figure A.2 using dashed curves. Thus, the center frequency of the effective signal $\tilde{s}_k(t)$ is $f_k + f_{dk}$, while the center frequency of $\tilde{s}_{k'}(t)$ becomes $f_{k'} + f_{dk'}$. Clearly, provided $|(k - k')\Delta f + f_{dk} - f_{dk'}| \gg \frac{1}{T - \Delta\tau_{kk'}}$, that is

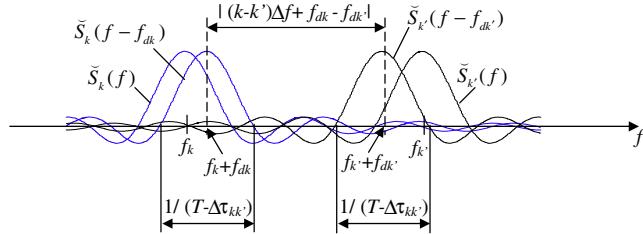
$$\begin{aligned}f_{dk} - f_{dk'} &\ll (k' - k)\Delta f - \frac{1}{T - \Delta\tau_{kk'}} \\ \text{or} \\ f_{dk} - f_{dk'} &\gg (k' - k)\Delta f + \frac{1}{T - \Delta\tau_{kk'}},\end{aligned} \quad (\text{A.5})$$

the overlapping part between $S_k(f - f_{dk}(\mathbf{v}))$ and $S_{k'}^*(f - f_{dk'}(\mathbf{v}))$ ($k' \neq k$) can be ignored. Thus the orthogonality is well approximated. Note that since (A.4) is satisfied, we have $|(k' - k)\Delta f| \gg |\frac{1}{T - \Delta\tau_{kk'}}|$. Hence the term $\frac{1}{T - \Delta\tau_{kk'}}$ in (A.5) can be ignored, resulting in $f_{dk} - f_{dk'} \ll (k' - k)\Delta f$ or $f_{dk} - f_{dk'} \gg (k' - k)\Delta f$.

It has been shown in [24] that, in the example case of target velocity estimation, the CRBs obtained using untruncated signals under certain conditions that are of practical interest are extremely close to the CRBs calculate by assuming ideal orthogonal signals. Interested readers are referred to [24] for detail.

A.2 Distinguishing coherent and noncoherent processing

In the MIMO radar context, the term *coherent processing* implies that the relative phase information embedded in the signals traveling over the paths between different transmit and receive antennas will be exploited (e.g., see (13.22)). Ultimately, this requires the oscillators in the up and down converters at different transmit and receive antenna nodes to be locked in phase. This is not the case for noncoherent processing where the relative phase information is not exploited (e.g., see (13.57)).

**FIGURE A.2**

Frequency spectrum for the time delayed equivalent signals (solid curves) and their Doppler shifted versions (dotted curves).

To distinguish between the coherent and noncoherent processing, one can examine the output of the processor (e.g., (13.22) or (13.57)), say denoted by P , as a function of a set of observations coming from different spatial paths, say denoted by $\{\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_{\mathcal{I}}\} = \{\tilde{r}_i\}_{i=1}^{\mathcal{I}}$, where \mathcal{I} represents the number of available observations coming from different spatial paths. That is $P = \text{func}(\{\tilde{r}_i\}_{i=1}^{\mathcal{I}})$. If an arbitrary constant¹³ phase set $\{\Omega_i\}_{i=1}^{\mathcal{I}}$ is imposed on the observations coming from different spatial paths, but the processing is insensitive to the phase shifts over the different paths, i.e., the output of the processor P is not affected, such that

$$P^{\text{non}} = \text{func}(\{\tilde{r}_i\}_{i=1}^{\mathcal{I}}) = \text{func}\left(\left\{\tilde{r}_i e^{j\Omega_i}\right\}_{i=1}^{\mathcal{I}}\right) \quad (\text{A.6})$$

then the processor is regarded as being noncoherent. Otherwise, if the output of the processor is affected by the arbitrary phase set, such that

$$P^{\text{coh}} = \text{func}(\{\tilde{r}_i\}_{i=1}^{\mathcal{I}}) \neq \text{func}\left(\left\{\tilde{r}_i e^{j\Omega_i}\right\}_{i=1}^{\mathcal{I}}\right), \quad (\text{A.7})$$

then the processor is regarded as being coherent.

The estimators $\hat{\theta}_{J,ML}^{\text{coh}}$ in (13.22) and $\hat{\theta}_{J,ML}^{\text{non}}$ in (13.57), which satisfy (A.7) and (A.6), are examples of noncoherent and coherent processors respectively, where

$$\hat{\theta}_{J,ML}^{\text{coh}} = \text{func}\left(\{\tilde{r}_{lk}(t)\}_{(l,k)=(1,1)}^{(N,M)}\right) \neq \text{func}\left(\left\{\tilde{r}_{lk}(t)e^{j\Omega_{lk}}\right\}_{(l,k)=(1,1)}^{(N,M)}\right)$$

and

$$\hat{\theta}_{J,ML}^{\text{non}} = \text{func}\left(\{\tilde{r}_{lk}(t)\}_{(l,k)=(1,1)}^{(N,M)}\right) = \text{func}\left(\left\{\tilde{r}_{lk}(t)e^{j\Omega_{lk}}\right\}_{(l,k)=(1,1)}^{(N,M)}\right).$$

¹³Here we model the oscillator phase drift as varying slowly enough so as to be considered constant over any pulse train interval. We define noncoherent processing approaches as those approaches that do not require phase locking between the oscillators at all the different transmitter and receiver nodes.

A.3 FIM for coherent MIMO radar joint estimation

Considering that the log-likelihood ratio in (13.20) is explicitly a function of τ_{lk} and f_{lk} , but implicitly as a function of Θ_J , we defined a new parameter vector

$$\vartheta_J = [\tau_{11}, \tau_{12}, \dots, \tau_{NM}, f_{11}, f_{12}, \dots, f_{NM}, \zeta_R, \zeta_I]^T. \quad (\text{A.8})$$

Then, the FIM can be derived using the chain rule as

$$\mathbf{J}_{\text{coh}}(\Theta_J) = \left(\nabla_{\Theta_J} \vartheta_J^T \right) \mathbf{J}(\vartheta_J) \left(\nabla_{\Theta_J} \vartheta_J^T \right)^T. \quad (\text{A.9})$$

We first compute $\nabla_{\Theta_J} \vartheta_J^T$ and obtain

$$\nabla_{\Theta_J} \vartheta_J^T = \begin{bmatrix} \frac{\partial \tau_{11}}{\partial x} & \frac{\partial \tau_{12}}{\partial x} & \dots & \frac{\partial \tau_{NM}}{\partial x} & \frac{\partial f_{11}}{\partial x} & \frac{\partial f_{12}}{\partial x} & \dots & \frac{\partial f_{NM}}{\partial x} & 0 & 0 \\ \frac{\partial \tau_{11}}{\partial y} & \frac{\partial \tau_{12}}{\partial y} & \dots & \frac{\partial \tau_{NM}}{\partial y} & \frac{\partial f_{11}}{\partial y} & \frac{\partial f_{12}}{\partial y} & \dots & \frac{\partial f_{NM}}{\partial y} & 0 & 0 \\ 0 & 0 & \dots & 0 & \frac{\partial f_{11}}{\partial v_x} & \frac{\partial f_{12}}{\partial v_x} & \dots & \frac{\partial f_{NM}}{\partial v_x} & 0 & 0 \\ 0 & 0 & \dots & 0 & \frac{\partial f_{11}}{\partial v_y} & \frac{\partial f_{12}}{\partial v_y} & \dots & \frac{\partial f_{NM}}{\partial v_y} & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}.$$

Partitioning $\nabla_{\Theta_J} \vartheta_J^T$ into a block matrix, we have

$$\nabla_{\Theta_J} \vartheta_J^T = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \end{bmatrix},$$

where $\mathbf{A}, \mathbf{B}, \mathbf{D}$ are $2 \times NM$ submatrices. The $\mathbf{J}(\vartheta_J)$ in (A.9) can be derived from (13.20) as

$$\mathbf{J}(\vartheta_J) = -\mathbb{E} \left\{ \nabla_{\vartheta_J} \left[\nabla_{\vartheta_J} L_J(\Theta_J; \tilde{\mathbf{r}}(t)) \right]^T \right\},$$

which is a $2(NM + 1) \times 2(NM + 1)$ matrix. Write $\mathbf{J}(\vartheta_J)$ in the form of a block matrix

$$\mathbf{J}(\vartheta_J) = \begin{bmatrix} \mathbf{X} & \mathbf{F} & \mathbf{P} \\ \mathbf{F}^H & \mathbf{Y} & \mathbf{K} \\ \mathbf{P}^H & \mathbf{K}^H & \mathbf{Z} \end{bmatrix},$$

where $\mathbf{X}, \mathbf{F}, \mathbf{Y}$ are $NM \times NM$ matrices. Carrying out the computation, we get

$$\mathbf{X} = \frac{8\pi^2 |\zeta|^2 E}{\sigma_w^2 M} \mathbf{I}_N \otimes \text{diag} \{ \varepsilon_1^c, \varepsilon_2^c, \dots, \varepsilon_M^c \}, \quad (\text{A.10})$$

$$\mathbf{F} = \frac{8\pi^2 |\zeta|^2 E}{\sigma_w^2 M} \text{diag} \{ \gamma_{11}^c, \gamma_{12}^c, \dots, \gamma_{NM}^c \}, \quad (\text{A.11})$$

$$\mathbf{Y} = \frac{8\pi^2 |\zeta|^2 E}{\sigma_w^2 M} \text{diag} \left\{ \eta_{11}^c, \eta_{12}^c, \dots, \eta_{NM}^c \right\}, \quad (\text{A.12})$$

$$\mathbf{P} = \frac{4\pi E}{\sigma_w^2 M} [\zeta_I \mathbf{1}_{N \times 1}, -\zeta_R \mathbf{1}_{N \times 1}] \otimes [\alpha_1, \alpha_2, \dots, \alpha_M]^T, \quad (\text{A.13})$$

$$\mathbf{K} = \frac{4\pi E}{\sigma_w^2 M} [-\zeta_I, \zeta_R] \otimes [\Xi_{11}, \Xi_{12}, \dots, \Xi_{NM}]^T, \quad (\text{A.14})$$

and

$$\mathbf{Z} = \frac{2EN}{\sigma_w^2} \mathbf{I}_2, \quad (\text{A.15})$$

where the terms $\varepsilon_k^c, \gamma_{lk}^c, \eta_{lk}^c, \alpha_k, \Xi_{lk}$ dependent on the characteristics of the received waveforms are provided in (33.24)–(33.28). Thus, the FIM can be computed through

$$\mathbf{J}_{\text{coh}}(\boldsymbol{\Theta}_P) = \begin{bmatrix} \mathbf{A}\mathbf{X}\mathbf{A}^T + \mathbf{B}\mathbf{F}^H\mathbf{A}^T + \mathbf{A}\mathbf{F}\mathbf{B}^T + \mathbf{B}\mathbf{Y}\mathbf{B}^T & \mathbf{A}\mathbf{F}\mathbf{D}^T + \mathbf{B}\mathbf{Y}\mathbf{D}^T & \mathbf{A}\mathbf{P} + \mathbf{B}\mathbf{K} \\ \mathbf{D}\mathbf{F}^H\mathbf{A}^T + \mathbf{D}\mathbf{Y}\mathbf{B}^T & \mathbf{D}\mathbf{Y}\mathbf{D}^T & \mathbf{D}\mathbf{K} \\ \mathbf{P}^H\mathbf{A}^T + \mathbf{K}^H\mathbf{B}^T & \mathbf{K}^H\mathbf{D}^T & \mathbf{Z} \end{bmatrix}. \quad (\text{A.16})$$

After lengthy algebraic manipulations, the expression of the FIM can be obtained as shown in (13.23).

A.4 Calculating FIM for localization with phase errors

First, let us compute the i, j th ($1 \leq i, j \leq NM$) element of the matrix. Referring to the vector in (13.41), let i refer to the index of the element of $\boldsymbol{\vartheta}_P$ corresponding to τ_{lk} , then let j refer to the index of the element of $\boldsymbol{\vartheta}_P$ corresponding to $\tau_{l'k'}$

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{i,j} = -\mathbb{E}_{\mathbf{r}(t); \boldsymbol{\vartheta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \tau_{l'k'}} \right], \quad (\text{A.17})$$

where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \tau_{l'k'}} = \frac{-1}{\Lambda^2(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk}} \frac{\partial \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{l'k'}} + \frac{1}{\Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \tau_{l'k'}}. \quad (\text{A.18})$$

The term $\Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))$ in (A.18) is given in (13.39). Further

$$\frac{\partial \Lambda(\boldsymbol{\vartheta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk}} = C_2 \mathbb{E}_{\boldsymbol{\Omega}} \left\{ \exp \{g\} \frac{\partial g}{\partial \tau_{lk}} \right\}, \quad (\text{A.19})$$

where g is given in (13.38). Assuming an interchange of derivative and integral is valid,¹⁴ then

$$\frac{\partial g}{\partial \tau_{lk}} = \frac{2}{\sigma_w^2} \int_{\mathcal{T}} \Re \left\{ h_l \zeta^* \sqrt{\frac{E}{M}} e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} \left[j2\pi f_c s_k^*(t - \tau_{lk}) + \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} \right] \right\} dt, \quad (\text{A.20})$$

¹⁴The $\Re\{\cdot\}$ means taking the real part of the quantity in the brackets.

where

$$h_l \equiv \tilde{r}_l(t) - \zeta \sqrt{\frac{E}{M}} \sum_{k=1}^M e^{j\tilde{\Omega}_{lk}} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}). \quad (\text{A.21})$$

Finally, the remaining term in (A.18) is

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \tau_{l'k'}} = C_2 \mathbb{E}_{\boldsymbol{\Omega}} \left\{ \exp\{g\} \left[\frac{\partial g}{\partial \tau_{l'k'}} \frac{\partial g}{\partial \tau_{lk}} + \frac{\partial^2 g}{\partial \tau_{lk} \partial \tau_{l'k'}} \right] \right\},$$

where for $l \neq l'$, $\partial^2 g / (\partial \tau_{lk} \partial \tau_{l'k'}) = 0$, and for $l = l'$

$$\begin{aligned} \frac{\partial^2 g}{\partial \tau_{lk} \partial \tau_{lk'}} &= -\frac{2|\zeta|^2 E}{\sigma_w^2 M} \int_{\mathcal{T}} \Re \left\{ e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} \left\{ e^{j\tilde{\Omega}_{lk'}} e^{-j2\pi f_c \tau_{lk'}} \right. \right. \\ &\quad \times \left[4\pi^2 f_c^2 s_{k'}(t - \tau_{lk'}) s_k^*(t - \tau_{lk}) - j2\pi f_c s_{k'}(t - \tau_{lk'}) \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} \right. \\ &\quad \left. \left. + j2\pi f_c \frac{\partial s_{k'}(t - \tau_{lk'})}{\partial \tau_{lk'}} s_k^*(t - \tau_{lk}) + \frac{\partial s_{k'}(t - \tau_{lk'})}{\partial \tau_{lk}} \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} \right] + \rho_{kk'} \right\} dt, \end{aligned}$$

where for $k \neq k'$, $\rho_{kk'} = 0$, and for $k = k'$

$$\rho_{kk} = \frac{1}{|\zeta|^2} \sqrt{\frac{M}{E}} h_l \zeta^* \left[4\pi^2 f_c^2 s_k^*(t - \tau_{lk}) - j4\pi f_c \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} - \frac{\partial^2 s_k^*(t - \tau_{lk})}{\partial \tau_{lk}^2} \right].$$

Now all terms in (A.18) have been obtained. Further taking the expectation $\mathbb{E}_{\mathbf{r}(t); \boldsymbol{\theta}_P}\{\cdot\}$ with respect to $p(\tilde{\mathbf{r}}(t); \boldsymbol{\theta}_P)$ gives (A.17). Next, we compute the elements related to the target complex reflectivity

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+1,i} = [\mathbf{J}(\boldsymbol{\vartheta}_P)]_{i,NM+1} = -\mathbb{E}_{\mathbf{r}(t); \boldsymbol{\theta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_R} \right] \quad (\text{A.22})$$

for $i = 1, \dots, NM$, where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_R} = \frac{-1}{\Lambda^2(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk}} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R} + \frac{1}{\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_R}. \quad (\text{A.23})$$

The new terms introduced by (A.23) are

$$\frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R} = C_2 \mathbb{E}_{\boldsymbol{\Omega}} \left\{ \exp\{g\} \frac{\partial g}{\partial \zeta_R} \right\}, \quad (\text{A.24})$$

where g is given in (13.38) and

$$\frac{\partial g}{\partial \zeta_R} = \frac{2}{\sigma_w^2} \sum_{l=1}^N \int_{\mathcal{T}} \Re \left\{ h_l \sqrt{\frac{E}{M}} \sum_{k=1}^M e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} s_k^*(t - \tau_{lk}) \right\} dt, \quad (\text{A.25})$$

and

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_R} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \left[\frac{\partial g}{\partial \tau_{lk}} \frac{\partial g}{\partial \zeta_R} + \frac{\partial^2 g}{\partial \tau_{lk} \partial \zeta_R} \right] \right\}, \quad (\text{A.26})$$

where g is given in (13.38), $\partial g / \partial \tau_{lk}$ in (A.20), $\partial g / \partial \zeta_R$ in (A.25), and

$$\begin{aligned} \frac{\partial^2 g}{\partial \tau_{lk} \partial \zeta_R} &= \frac{2E}{\sigma_w^2 M} \int_T \Re \left\{ e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} \left[j2\pi f_c s_k^*(t - \tau_{lk}) + \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} \right] \right. \\ &\quad \times \left. \left[\sqrt{\frac{M}{E}} h_l - \zeta^* \sum_{k=1}^M e^{j\tilde{\Omega}_{lk}} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) \right] \right\} dt. \end{aligned}$$

Plugging (13.39), (A.19), (A.24), (A.26) in (A.23) and taking expectation gives (A.22). Then, we compute

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+2,i} = [\mathbf{J}(\boldsymbol{\vartheta}_P)]_{i,NM+2} = -\mathbb{E}_{\mathbf{r}(t); \boldsymbol{\theta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_I} \right] \quad (\text{A.27})$$

for $i = 1, \dots, NM$, where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_I} = \frac{-1}{\Lambda^2(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk}} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I} + \frac{1}{\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_I}. \quad (\text{A.28})$$

The new terms introduced by (A.28) are

$$\frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \frac{\partial g}{\partial \zeta_I} \right\}, \quad (\text{A.29})$$

where g is given in (13.38) and¹⁵

$$\frac{\partial g}{\partial \zeta_I} = \frac{2}{\sigma_w^2} \sum_{l=1}^N \int_T \Im \left\{ j h_l \sqrt{\frac{E}{M}} \sum_{k=1}^M e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} s_k^*(t - \tau_{lk}) \right\} dt, \quad (\text{A.30})$$

and

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \tau_{lk} \partial \zeta_I} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \left[\frac{\partial g}{\partial \tau_{lk}} \frac{\partial g}{\partial \zeta_I} + \frac{\partial^2 g}{\partial \tau_{lk} \partial \zeta_I} \right] \right\}, \quad (\text{A.31})$$

where g is given in (13.38), $\partial g / \partial \tau_{lk}$ in (A.20), $\partial g / \partial \zeta_I$ in (A.30), and

$$\begin{aligned} \frac{\partial^2 g}{\partial \tau_{lk} \partial \zeta_I} &= \frac{2E}{\sigma_w^2 M} \int_T \Im \left\{ e^{-j\tilde{\Omega}_{lk}} e^{j2\pi f_c \tau_{lk}} \left[j2\pi f_c s_k^*(t - \tau_{lk}) + \frac{\partial s_k^*(t - \tau_{lk})}{\partial \tau_{lk}} \right] \right. \\ &\quad \times \left. \left[\sqrt{\frac{M}{E}} h_l + \zeta^* \sum_{k=1}^M e^{j\tilde{\Omega}_{lk}} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) \right] \right\} dt. \end{aligned}$$

¹⁵The $\Im\{\cdot\}$ means taking the imaginary part of the quantity in the brackets.

Plugging (13.39), (A.19), (A.29), (A.31) in (A.28) and taking expectation gives (A.27). Next, we compute

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+1,NM+1} = -\mathbb{E}_{\mathbf{r}(t);\boldsymbol{\theta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R^2} \right], \quad (\text{A.32})$$

where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R^2} = \frac{-1}{\Lambda^2(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \left[\frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R} \right]^2 + \frac{1}{\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R^2}. \quad (\text{A.33})$$

The new term introduced by (A.33) is

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R^2} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \left[\left(\frac{\partial g}{\partial \zeta_R} \right)^2 + \frac{\partial^2 g}{\partial \zeta_R^2} \right] \right\}, \quad (\text{A.34})$$

where g is given in (13.38), $\partial g / \partial \zeta_R$ in (A.25), and

$$\frac{\partial^2 g}{\partial \zeta_R^2} = -\frac{2E}{\sigma_w^2 M} \sum_{l=1}^N \int_{\mathcal{T}} \left| \sum_{k=1}^M e^{j\tilde{\Omega}_{lk}} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) \right|^2 dt.$$

Plugging (13.39), (A.24), (A.34) in (A.33) and taking expectation, we obtain (A.32). Similarly, we compute

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+2,NM+2} = -\mathbb{E}_{\mathbf{r}(t);\boldsymbol{\theta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I^2} \right], \quad (\text{A.35})$$

where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I^2} = \frac{-1}{\Lambda^2(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \left[\frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I} \right]^2 + \frac{1}{\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I^2}. \quad (\text{A.36})$$

The new term introduced by (A.36) is

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I^2} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \left[\left(\frac{\partial g}{\partial \zeta_I} \right)^2 + \frac{\partial^2 g}{\partial \zeta_I^2} \right] \right\}, \quad (\text{A.37})$$

where g is given in (13.38), $\partial g / \partial \zeta_I$ in (A.30) and

$$\frac{\partial^2 g}{\partial \zeta_I^2} = -\frac{2E}{\sigma_w^2 M} \sum_{l=1}^N \int_{\mathcal{T}} \left| \sum_{k=1}^M e^{j\tilde{\Omega}_{lk}} e^{-j2\pi f_c \tau_{lk}} s_k(t - \tau_{lk}) \right|^2 dt.$$

Plugging (13.39), (A.29), (A.37) into (A.36) and taking expectation gives (A.35). Finally, we compute

$$[\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+1,NM+2} = [\mathbf{J}(\boldsymbol{\vartheta}_P)]_{NM+2,NM+1} = -\mathbb{E}_{\mathbf{r}(t);\boldsymbol{\theta}_P} \left[\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R \partial \zeta_I} \right], \quad (\text{A.38})$$

where

$$\frac{\partial^2 \ln \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R \partial \zeta_I} = \frac{-1}{\Lambda^2(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R} \frac{\partial \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_I} + \frac{1}{\Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))} \frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R \partial \zeta_I}. \quad (\text{A.39})$$

The new term introduced by (A.39) is

$$\frac{\partial^2 \Lambda(\boldsymbol{\theta}_P; \tilde{\mathbf{r}}(t))}{\partial \zeta_R \partial \zeta_I} = C_2 \mathbb{E}_{\Omega} \left\{ \exp\{g\} \frac{\partial g}{\partial \zeta_R} \frac{\partial g}{\partial \zeta_I} \right\}, \quad (\text{A.40})$$

where g is given in (13.38), $\partial g / \partial \zeta_R$ in (A.25), and $\partial g / \partial \zeta_I$ in (A.30). Plugging (13.39), (A.24), (A.29), (A.40) into (A.39) and taking expectation we obtain (A.38).

Now we compute the quantity

$$\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T = \left[(\nabla_x \boldsymbol{\vartheta}_P^T)^T, (\nabla_y \boldsymbol{\vartheta}_P^T)^T, (\nabla_{\zeta_R} \boldsymbol{\vartheta}_P^T)^T, (\nabla_{\zeta_I} \boldsymbol{\vartheta}_P^T)^T \right]^T,$$

using τ_{lk} and $\boldsymbol{\vartheta}_P$ respectively given in (13.2) and (13.41)

$$\nabla_{\boldsymbol{\theta}_P} \boldsymbol{\vartheta}_P^T = \begin{bmatrix} \frac{\partial \tau_{11}}{\partial x} & \frac{\partial \tau_{12}}{\partial x} & \cdots & \frac{\partial \tau_{NM}}{\partial x} & 0 & 0 \\ \frac{\partial \tau_{11}}{\partial y} & \frac{\partial \tau_{12}}{\partial y} & \cdots & \frac{\partial \tau_{NM}}{\partial y} & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}, \quad (\text{A.41})$$

where $\partial \tau_{lk} / \partial x = [(x - x_k^t)/d_k^t + (x - x_l^r)/d_l^r]/c$ and $\partial \tau_{lk} / \partial y = [(y - y_k^t)/d_k^t + (y - y_l^r)/d_l^r]/c$.

Relevant Theory: Signal Processing Theory Statistical Signal Processing, and Array Signal Processing

See Vol. 1, Chapter 4 Random Signals and Stochastic Processes

See Vol. 1, Chapter 11 Parametric Estimation

See Vol. 3, Chapter 7 Geolocation—Maps, Measurements, Models, and Methods

See Vol. 3, Chapter 8 Performance Analysis and Bounds

See Vol. 3, Chapter 19 Array Processing in the Face of Nonidealities

References

- [1] H. Bölcseki, A.J. Paulraj, Multiple-input multiple-output (MIMO) wireless systems, in: The Communications Handbook, second ed., CRC Press, 2002, pp. 90.1–90.14.
- [2] E. Fishler, A.M. Haimovich, R.S. Blum, L. Cimini, D. Chizhik, R. Valenzuela, Spatial diversity in radars—models and detection performance, IEEE Trans. Signal Process. 54 (3) (2006) 823–838.
- [3] J. Li, P. Stoica, MIMO radar with colocated antennas, IEEE Signal Process. Mag. 24 (5) (2007) 106–114.
- [4] A.M. Haimovich, R.S. Blum, L. Cimini, MIMO radar with widely separated antennas, IEEE Signal Process. Mag. (1) (2008) 116–129.
- [5] R.S. Blum, Limiting case of a lack of rich scattering environment for MIMO radar diversity, IEEE Signal Process. Lett. 16 (10) (2009) 901–904.
- [6] T. Aittomaki, V. Koivunen, Performance of MIMO radar with angle diversity under Swerling scattering model, IEEE J. Sel. Top. Signal Process. 4 (1) (2010) 101–114 (special issue on MIMO radar).

- [7] E. Grossi, M. Lops, L. Venturino, Robust waveform design for MIMO radars, *IEEE Trans. Signal Process.* 59 (7) (2011) 3262–3271.
- [8] M. Akcakaya, A. Nehorai, MIMO radar sensitivity analysis for target detection, *IEEE Trans. Signal Process.* 59 (7) (2011) 3241–3250.
- [9] Y. Yang, R.S. Blum, MIMO radar waveform design based on mutual information and minimum mean-square error estimation, *IEEE Trans. Aerosp. Electron. Syst.* 43 (2007) 330–343.
- [10] Q. He, N.H. Lehmann, R.S. Blum, A.M. Haimovich, MIMO radar moving target detection in homogeneous clutter, *IEEE Trans. Aerosp. Electron. Syst.* 46 (3) (2010) 1290–1301.
- [11] H. Godrich, A.M. Haimovich, R.S. Blum, Target localization accuracy gain in MIMO radar based systems, *IEEE Trans. Inform. Theory* 56 (6) (2010) 2783–2803.
- [12] S. Sen, N. Nehorai, OFDM MIMO radar with mutual-information waveform design for low-grazing angle tracking, *IEEE Trans. Signal Process.* 58 (6) (2010) 3152–3162.
- [13] R. Niu, R.S. Blum, P.K. Varshney, A.L. Drozd, Target tracking in widely separated non-coherent multiple-input multiple-output radar systems, in: Proceeding of the 43rd Asilomar Conference on Signals, Systems and Computers, November 2009, pp. 1181–1185.
- [14] M.L. Melvin, A STAP overview, *IEEE Aerosp. Electron. Syst. Mag.* 19 (1) (2004) 19–35.
- [15] H.L. Van Trees, *Detection, Estimation, and Modulation Theory III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, John Wiley & Sons Inc., 2001.
- [16] Q. He, R.S. Blum, Noncoherent versus coherent MIMO radar: performance and simplicity analysis, *Signal Process.* 92 (2012) 2454–2463.
- [17] H.V. Poor, *An Introduction to Signal Detection and Estimation*, second ed., Springer-Verlag, New York, 1994.
- [18] Q. He, R.S. Blum, H. Godrich, A.M. Haimovich, Target velocity estimation and antenna placement for MIMO radar with widely separated antennas, *IEEE J. Sel. Top. Signal Process.* 4 (1) (2010) 79–100 (special issue on MIMO radar).
- [19] P.M. Woodward, *Probability and Information Theory, with Applications to Radar*, McGraw-Hill, New York, 1953.
- [20] N. Lehmann, A.M. Haimovich, R.S. Blum, L. Cimini, High resolution capabilities of MIMO radar, in: Proceeding of the 40th Asilomar Conference on Signal, System and Computers, November 2006, pp. 25–30.
- [21] G. San Antonio, D.R. Fuhrmann, F.C. Robey, MIMO radar ambiguity functions, *IEEE J. Sel. Top. Signal Process.* 1 (1) (2007) 167–177.
- [22] C. Chen, P.P. Vaidyanathan, MIMO radar ambiguity properties and optimization using frequency-hopping waveforms, *IEEE Trans. Signal Process.* 56 (12) (2008) 5926–5936.
- [23] H.L. Van Trees, *Detection, Estimation, and Modulation Theory I*, John Wiley & Sons Inc., 2001.
- [24] Q. He, R.S. Blum, A.M. Haimovich, Noncoherent MIMO radar for location and velocity estimation: more antennas means better performance, *IEEE Trans. Signal Process.* 58 (7) (2010) 3661–3680.
- [25] Q. He, R.S. Blum, Diversity gain for MIMO Neyman-Pearson signal detection, *IEEE Trans. Signal Process.* 59 (3) (2011) 869–881.
- [26] Q. He, R.S. Blum, Cramer-Rao bound for MIMO radar target localization with phase errors, *IEEE Signal Process. Lett.* 17 (1) (2010) 83–86.
- [27] H. Godrich, A.M. Haimovich, H.V. Poor, An analysis of phase synchronization mismatch sensitivity for coherent MIMO radar systems, in: Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, Aruba, Dutch Antilles, December 2009, pp. 153–156.
- [28] H. Godrich, A.M. Haimovich, Localization performance of coherent MIMO radar systems subject to phase synchronization errors, in: Proceedings of the 4th International Symposium on Communications, Control and Signal Processing, Limassol, Cyprus, 3–5 March 2010.

- [29] M. Akçakaya, A. Nehorai, MIMO Radar detection and adaptive design under a phase synchronization mismatch, *IEEE Trans. Signal Process.* 58 (10) (2010) 4994–5005.
- [30] V.V. Shakhangulyan, S.S. Sviridenko, Phase synchronization systems studies, *IEEE Trans. Commun.* COM-30 (10) (1982) 2260–2263.
- [31] B.D. Steinberg, Phase synchronizing a nonrigid, distributed, transmit-receive radar antenna array, *IEEE Trans. Aerosp. Electron. Syst. AES-18* (5) (1982) 609–620.
- [32] E.H. Attia, Phase synchronizing large antenna arrays using the spatial correlation properties of radar clutter, Ph.D. Thesis, University of Pennsylvania, 1985.
- [33] E.-A. Lee, C.N. Dorn, A broadcast reference technique for self-calibrating of large antenna phased arrays, *IEEE Trans. Antennas Propag.* 37 (8) (1989) 1003–1010.
- [34] N.J. Willis, *Bistatic Radar*, second ed., SciTech Publishing, 2005.
- [35] P. Lopez-Dekker, J.J. Mallorqui, P. Serra-Morales, J. Sanz-Marcos, Phase synchronization and Doppler centroid estimation in fixed receiver bistatic SAR systems, *IEEE Trans. Geosci. Remote Sens.* 46 (11) (2008) 3459–3471.
- [36] B.P. Ng, J.P. Lie, M.H. Er, A. Feng, A practical simple geometry and gain/phase calibration technique for antenna array processing, *IEEE Trans. Antennas Propag.* 57 (7) (2009) 1963–1972.
- [37] W.C. Lindsey, F. Ghazvinian, W.C. Hagmann, K. Dessouky, Network synchronization, *Proc. IEEE* 73 (10) (1985) 1445–1467.
- [38] B. Sundararaman, U. Buy, A.D. Kshemkalyani, Clock synchronization for wireless sensor networks: a survey, *Ad Hoc Networks* 3 (3) (2005) 281–323.
- [39] Y. Tu, G.J. Pottie, Coherent cooperative transmission from multiple adjacent antennas to a distant stationary antenna through AWGN channels, in: *Proceedings of the IEEE Vehicular Technology Conference*, AL, Spring 2002, vol. 1, pp. 130–134.
- [40] R. Mudumbai, G. Barriac, U. Madhow, On the feasibility of distributed beamforming in wireless networks, *IEEE Trans. Wireless Commun.* 6 (4) (2007) 1–10.
- [41] Q. Li, D. Rus, Global clock synchronization in sensor networks, in: *Proceedings of the IEEE Conference on Computer Communications*, March 2004, pp. 564–574.
- [42] D.R. Brown III, G. Prince, J.A. McNeill, A method for carrier frequency and phase synchronization of two autonomous cooperative transmitters, in: *Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communications*, New York, June 2005, pp. 260–264.
- [43] D.R. Brown III, H.V. Poor, Time-slotted round-trip carrier synchronization for distributed beamforming, *IEEE Trans. Signal Process.* 56 (11) (2008) 5630–5643.
- [44] Q. Wang, K. Ren, Time-slotted round-trip carrier synchronization in large-scale wireless networks, in: *Proceedings of the IEEE International Conference on Communications*, May 2008, pp. 5087–5091.
- [45] Y. Yang, R.S. Blum, Phase synchronization for coherent MIMO radar: algorithms and their analysis, *IEEE Trans. Signal Process.* 59 (11) (2011) 5538–5557.
- [46] R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems, *Proc. IEEE* 95 (1) (2007) 215–233.
- [47] L. Xiao, S. Boyd, Fast linear iterations for distributed averaging, *System Control Lett.* 53 (2004) 65–78.
- [48] S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, Randomized gossip algorithms, *IEEE Trans. Inform. Theory* 52 (2006) 2508–2530.
- [49] L. Xiao, S. Boyd, S.-J. Kim, Distributed average consensus with least-mean-square deviation, *J. Parallel Distr. Comput.* 67 (1) (2007) 33–46.
- [50] F. Fagnani, S. Zampieri, Randomized consensus algorithms over large scale networks, *IEEE J. Sel. Area Commun.* 26 (4) (2008) 634–649.

- [51] T.C. Aysal, M.E. Yildiz, A.D. Sarwate, A. Scaglione, Broadcast gossip algorithms for consensus, *IEEE Trans. Signal Process.* 57 (2009) 2748–2761.
- [52] Y. Yang, R.S. Blum, Broadcast-based consensus with non-zero-mean stochastic perturbations, *IEEE Trans. Inform. Theory* (2013), <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6459027>>.
- [53] L. Schenato, G. Gamba, A distributed consensus protocol for clock synchronization in wireless sensor network, in: Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, LA, December 2007, pp. 2289–2294.
- [54] G. Xiong, S. Kishore, Analysis of distributed consensus time synchronization with Gaussian delay over wireless sensor networks, *EURASIP J. Wireless Commun. Network*. 2009 (2009) 9 (Article ID 528161).
- [55] J. Rutman, Characterization of phase and frequency instabilities in precision frequency sources: fifteen years of progress, *Proc. IEEE* 66 (9) (1978) 1048–1075.
- [56] A. Demir, A. Mehrotra, J. Roychowdhury, Phase noise in oscillators: a unifying theory and numerical methods, *IEEE Trans. Circ. Syst. Fund. Theory Appl.* 47 (5) (2000) 655–674.
- [57] M.I. Skolnik (Ed.), *Radar Handbook*, second ed., McGraw-Hill, New York, 1990.
- [58] N. Levanon, E. Mozeson, *Radar Signals*, John Wiley & Sons, New York, 2004.
- [59] Y. Yang, R.S. Blum, Minimax robust MIMO radar waveform design, *IEEE J. Sel. Top. Signal Process.* 1 (2007) 147–155.
- [60] Y. Yang, R.S. Blum, Z.S. He, D.R. Fuhrmann, MIMO radar waveform design via alternating projection, *IEEE Trans. Signal Process.* 58 (3) (2010) 1140–1445.
- [61] T. Naghibi, F. Behnia, MIMO radar waveform design in the presence of clutter, *IEEE Trans. Aerosp. Electron. Syst.* 47 (2) (2011) 770–781.
- [62] C-Y. Chen, P.P. Vaidyanathan, MIMO radar waveform optimization with prior information of the extended target and clutter, *IEEE Trans. Signal Process.* 57 (9) (2009) 3533–3544.
- [63] B. Tang, J. Tang, Y. Peng, MIMO radar waveform design in colored noise based on information theory, *IEEE Trans. Signal Process.* 58 (9) (2010) 4684–4697.
- [64] J. Zhang, H. Wang, X. Zhu, Adaptive waveform design for separated transmit/receive ULA-MIMO radar, *IEEE Trans. Signal Process.* 58 (9) (2010) 4936–4942.
- [65] A. De Maio, M. Lops, Design principles of MIMO radar detectors, *IEEE Trans. Aerosp. Electron. Syst.* 43 (3) (2007) 886–898.
- [66] A. De Maio, M. Lops, L. Venturino, Diversity-integration tradeoffs in MIMO detection, *IEEE Trans. Signal Process.* 56 (10) (2008) 5051–5061.
- [67] D.R. Fuhrmann, G. San Antonio, Transmit beamforming for MIMO radar systems using signal cross-correlation, *IEEE Trans. Aerosp. Electron. Syst.* 44 (1) (2008) 171–186.
- [68] B. Friedlander, Waveform design for MIMO radars, *IEEE Trans. Aerosp. Electron. Syst.* 43 (3) (2007) 1227–1238.
- [69] J. Li, P. Stoica, Y. Xie, On probing signal design for MIMO radar, *IEEE Trans. Signal Process.* 55 (8) (2007) 4151–4161.
- [70] T. Naghibi, M. Namvara, F. Behniaa, Optimal and robust waveform design for MIMO radars in the presence of clutter, *IEEE Trans. Signal Process.* 90 (4) (2010) 1103–1117.
- [71] G.H. Jajamovich, M. Lops, X. Wang, Space-time coding for MIMO radar detection and ranging, *IEEE Trans. Signal Process.* 58 (12) (2010) 6195–6206.
- [72] A. Aubry, M. Lops, A.M. Tulino, L. Venturino, On MIMO detection under non-Gaussian target scattering, *IEEE Trans. Inform. Theory* 56 (11) (2010) 5822–5838.
- [73] X. Song, S. Zhou, P. Willett, Reducing the waveform cross correlation of MIMO radar with space-time coding, *IEEE Trans. Signal Process.* 58 (8) (2010) 4213–4224.
- [74] H. Wang, G. Liao, H. Liu, J. Li, H. Lv, Joint optimization of MIMO radar waveform and biased estimator with prior information in the presence of clutter, *EURASIP J. Adv. Signal Process.* 2011 (2011) 15.

- [75] M.R. Bell, Information theory and radar waveform design, *IEEE Trans. Inform. Theory* 39 (1993) 1578–1597.
- [76] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
- [77] U. Grenander, G. Szegö, *Toeplitz Forms and Their Applications*, second ed., Chelsea Publishing Company, New York, 1984.
- [78] R.M. Gray, Toeplitz and Circulant Matrices: A Review, 2002. <http://www.ee.stanford.edu/gray/toeplitz.pdf>.
- [79] S.A. Kassam, H.V. Poor, Robust techniques for signal processing: a survey, *Proc. IEEE* 73 (3) (1985) 433–481.

Optimal Radar Waveform Design

14

Joseph R. Guerci

Guerci Consulting LLC, Vienna, VA, USA

Nomenclature

\mathbb{C}^N	N -dimensional complex vector space
$\mathbf{s} \in \mathbb{C}^N$	multidimensional transmit signal
$H_T \in \mathbb{C}^{M \times N}$	multidimensional target transfer function (generally stochastic)
$\mathbf{y}_s \in \mathbb{C}^N$	multidimensional received target “Echo”
$\mathbf{n} \in \mathbb{C}^M$	multidimensional complex additive receiver noise (zero mean)
$R \in \mathbb{C}^{M \times M}$	additive receiver noise covariance matrix
$H_w = R^{-1/2}$	whitening filter
$\mathbf{z}_s \in \mathbb{C}^N$	whitened target echo
$E\{\cdot\}$	expectation operator
$H_T \in \mathbb{C}^{N \times N}$	clutter channel transfer function (generally stochastic)
$p(X H)$	conditional PDF
$S(\omega) \leftrightarrow s(t)$	fourier transform pair
λ	radar wavelength

2.14.1 Introduction

Adaptive processing has long been implemented in the receive chain of radar, beginning with automatic gain control (AGC), cell-averaging constant false alarm rate (CA-CFAR) [1] and culminating in today’s space-time adaptive processing (STAP) [2]. However, adaptivity in the transmit chain is virtually nonexistent, save for mode adaptivity (such as switching in different nonadaptive waveforms such as pulse repetition frequency (PRF) and bandwidth) and adaptive spectral notching to address narrowband interference mitigation. Historically, the reasons for the lack of transmit adaptivity included:

- Inability of radar hardware to transmit arbitrary waveforms (or more generally arbitrary space-time waveforms).
- Lack of sufficient embedded computing to allow for real-time waveform adaptation.
- As a result of the above, there is a lack of basic theory for optimal waveform design as a function of the target-interference channel (except for very specialized cases).

However, as discussed in the abstract, the first two obstacles are no longer the case. All that remains is to revisit the basic radar theoretic problem and derive a new set of design equations allowing for transmit adaptivity as a function of the target-interference channel.

This chapter develops the basic theory of optimal transmit/receive design using a multi-input, multi-output (MIMO) formulation that can account for all potential degrees of freedom (DOFs) such as waveform (fast-time), angle, and polarization. Various applications and examples are provided to further illustrate the potential impact of joint transmit/receive adaptivity. The only mathematical prerequisites for this material is a basic understanding of matrix algebra, basic optimization theory, and stochastic processes.

2.14.2 Optimum transmit-receiver design for the additive colored noise case: detection

Due to the finite bandwidth constraint for all real-world radars, there is no loss in generality in invoking Shannon's sampling theory [3]. This, in turn, allows for a matrix algebra formulation of the basic radar interaction equations which both significantly eases nomenclature and exposition (see below).

Consider the basic radar block diagram in Figure 14.1. A generally complex-valued and multidimensional transmit signal, $\mathbf{s} \in \mathbb{C}^N$, (i.e., an N -dimensional multi-input (MI) signal), interacts with a target denoted by the target transfer matrix $H_T \in \mathbb{C}^{M \times N}$. The resulting M -dimensional multi-output (MO) signal (echo), $\mathbf{y} \in \mathbb{C}^M$, is then received along with ACN $\mathbf{n} \in \mathbb{C}^M$, assumed to be zero mean, wide sense stationary, with corresponding complex valued covariance $R \in \mathbb{C}^{M \times M}$. The vector-matrix formulation is completely general inasmuch as any combination of spatial and temporal dimensions can be represented.

For example, the N -dimensional input vector \mathbf{s} could represent the N complex (i.e., in-phase/quadrature, or "I and Q" [4]) samples of a single-channel transmit waveform $s(t)$, that is,

$$\mathbf{s} = \begin{bmatrix} s(\tau_1) \\ s(\tau_2) \\ \vdots \\ s(\tau_N) \end{bmatrix}. \quad (14.1)$$

The corresponding target transfer matrix, H_T , would thus contain the corresponding samples of the complex target impulse response, $h_T(t)$, which for the causal linear time-invariant (LTI) case would have the form [3]

$$H_T = \begin{bmatrix} h[0] & 0 & 0 & \cdots & 0 \\ h[1] & h[0] & 0 & \cdots & 0 \\ h[2] & h[1] & h[0] & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ h[N-1] & & & h[1] & h[0] \end{bmatrix}. \quad (14.2)$$

Without loss of generality we have assumed uniform time sampling, that is, $\tau_k = (k - 1)T$, where T is a suitably chosen sampling interval [5]. Note also that for convenience and a significant reduction in mathematical nomenclature overhead, $N = M$ is used, (i.e., the same number of transmit and receive

DOFs). The reader is encouraged to, where desired, reinstate the inequality and confirm that the underlying equations derived throughout this chapter have the same basic form except for differing vector and matrix dimensionalities. Also it should be noted that in general H_T is stochastic.

The formalism is readily extensible to the multiple-transmitter, multiple-receiver case. For example, if there are three independent transmit/receive channels (e.g., an AESA), then the input vector \mathbf{s} of Figure 14.1 would have the form

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{bmatrix} \in \mathbb{C}^{3N}, \quad (14.3)$$

where $\mathbf{s}_i \in \mathbb{C}^N$ denotes the samples (as in (14.1)) of the transmitted waveform out of the i th transmit channel. The corresponding target transfer matrix would in general have the form

$$H_T = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \in \mathbb{C}^{3N \times 3N}, \quad (14.4)$$

where the submatrix $H_{i,j} \in \mathbb{C}^{N \times N}$ is the transfer matrix between the i th receive and j th transmit channels for all time samples of the waveform. These examples make clear that the matrix-vector, input-output formalism is completely universal and can accommodate whatever transmit/receive DOF configuration desired.

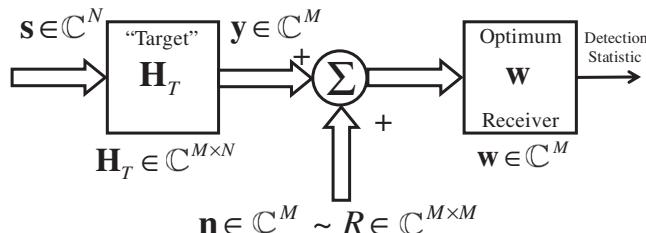
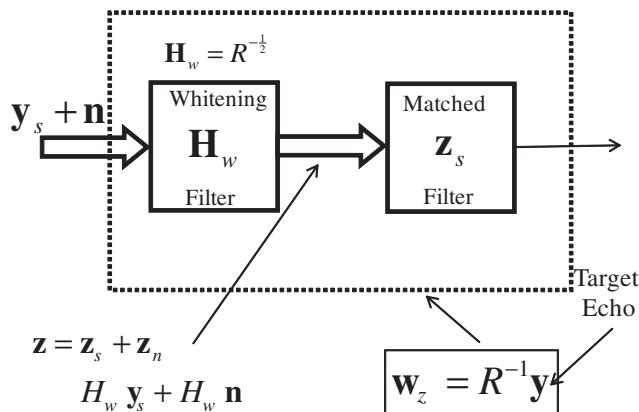


FIGURE 14.1

Fundamental multichannel radar block diagram for the ACN case. The objective is to design both the transmit and receive functions to maximize the output SINR given the channel characteristics.

**FIGURE 14.2**

The optimum (max SINR) receiver for the ACN case consists of a whitening filter followed by a white noise matched filter.

Returning to Figure 14.1, we now wish to jointly optimize the transmit/receive functions to maximize output SINR, which under the Gaussian assumption, would yield a sufficient statistic that maximizes the probability of detection for a prescribed false alarm rate [6]. We will find it convenient to work backward: to begin by optimizing the receiver as a function of the input and then finally optimizing the input and thus the overall output SINR.

For any finite norm input \mathbf{s} , the receiver that maximizes output SINR for the ACN case is the so-called whitening (or colored noise) matched filter, as shown in Figure 14.2 [6]. Note that for the additive Gaussian colored noise (AGCN) case, this receiver is also statistically optimum [6].

$R \in \mathbb{C}^{N \times N}$, which denotes the total interference covariance matrix associated with \mathbf{n} , is further assumed to be independent of \mathbf{s} and Hermitian positive definite [7] (guaranteed in practice due to ever-present receiver noise [6]), then the corresponding whitening filter shown in Figure 14.2 is given by [6]:

$$H_w = R^{-\frac{1}{2}}, \quad (14.5)$$

where the superscript $-\frac{1}{2}$ denotes the “matrix square root inverse” [8].

The output of the linear whitening filter, $\mathbf{z} \in \mathbb{C}^N$, will consist of signal and noise components, $\mathbf{z}_s, \mathbf{z}_n$, respectively, given by

$$\begin{aligned} \mathbf{z} &= \mathbf{z}_s + \mathbf{z}_n \\ &= H_w \mathbf{y}_s + H_w \mathbf{n} \\ &= H_w H_T \mathbf{s} + H_w \mathbf{n}, \end{aligned} \quad (14.6)$$

where $\mathbf{y}_s \in \mathbb{C}^N$ denotes the target echo as shown in Figure 14.2 (i.e., the output of H_T).

Since the noise has been whitened via a linear—in this case full-rank—transformation [6]), the final receiver stage consists of a white noise matched filter of the form (to within a multiplicative scalar)

$$\mathbf{w}_z = \mathbf{z}_s \in \mathbb{C}^N. \quad (14.7)$$

The corresponding output SNR is thus given by

$$\begin{aligned}
 \text{SNR}_o &= \frac{|\mathbf{w}'_z \mathbf{z}_s|^2}{\text{var}(\mathbf{w}'_z \mathbf{z}_n)} \\
 &= \frac{|\mathbf{z}'_s \mathbf{z}_s|^2}{\text{var}(\mathbf{z}'_s \mathbf{z}_n)} \\
 &= \frac{|\mathbf{z}'_s \mathbf{z}_s|^2}{E\{\mathbf{z}'_s \mathbf{z}_n \mathbf{z}'_n \mathbf{z}_s\}} \\
 &= \frac{|\mathbf{z}'_s \mathbf{z}_s|^2}{\mathbf{z}'_s E\{\mathbf{z}_n \mathbf{z}'_n\} \mathbf{z}_s} \\
 &= \frac{|\mathbf{z}'_s \mathbf{z}_s|^2}{\mathbf{z}'_s \mathbf{z}_s} \\
 &= |\mathbf{z}'_s \mathbf{z}_s|,
 \end{aligned} \tag{14.8}$$

where superscript “ $'$ ” denotes complex conjugate transpose, and $\text{var}(\cdot)$ denotes the variance operator. Note that due to the whitening operation $E\{\mathbf{z}_n \mathbf{z}'_n\} = I$.

In words, the output SNR is proportional to the energy in the whitened target echo. This fact is key to optimizing the input function: Choose \mathbf{s} (the input) to maximize the energy in the whitened target echo:

$$\max_{\mathbf{s}} |\mathbf{z}'_s \mathbf{z}_s|. \tag{14.9}$$

Substituting $\mathbf{z}_s = H_w H_T \mathbf{s}$ into (14.9) yields the objective function that explicitly depends on the input

$$\max_{\mathbf{s}} |\mathbf{s}'(H'H)\mathbf{s}|, \tag{14.10}$$

where

$$H \triangleq H_w H_T. \tag{14.11}$$

Recognizing that (14.10) involves the magnitude of the inner product of two vectors \mathbf{s} and $(H'H)\mathbf{s}$, we readily have from the Cauchy-Schwarz theorem [9] the condition that \mathbf{s} must satisfy to yield a maximum, namely, \mathbf{s} must be collinear with $(H'H)\mathbf{s}$:

$$(H'H)\mathbf{s}_{\text{opt}} = \lambda_{\text{max}} \mathbf{s}_{\text{opt}}. \tag{14.12}$$

In other words, the optimum input \mathbf{s}_{opt} must be an eigenvector of $(H'H)$ with associated maximum eigenvalue.

The previous set of input-output design equations represents the absolute optimum that any combination of transmit/receive operations can achieve and thus are fundamentally important in ascertaining the value of advanced adaptive methods (e.g., adaptive waveforms, transmit/receive beamforming). Note

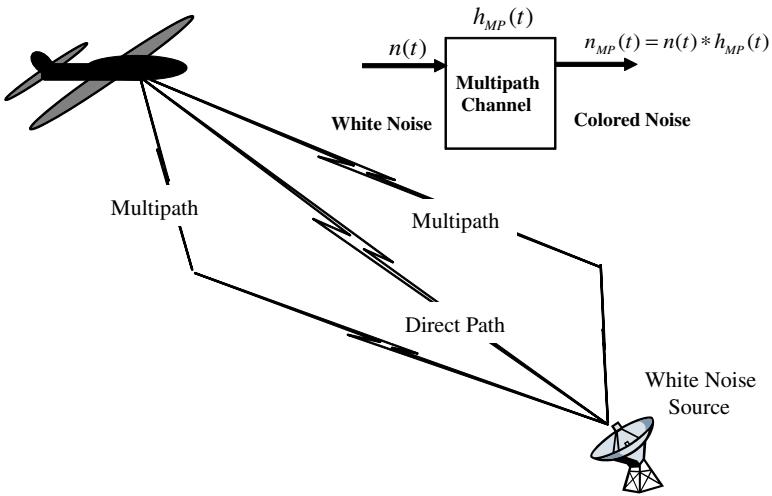
**FIGURE 14.3**

Illustration of colored noise interference resulting from a broadband (i.e., white noise) source undergoing strong discrete multipath reflections. The resulting signal can be approximated as white noise driving a finite impulse response (FIR) filter.

also that (14.12) can be generalized to the case where the target response is random:

$$E\{H'H\}\mathbf{s}_{\text{opt}} = \lambda_{\max}\mathbf{s}_{\text{opt}}. \quad (14.13)$$

In this case, \mathbf{s}_{opt} maximizes the expected value of the output SINR.

Next we illustrate the application of the previously given optimum design equations to the additive colored noise problem arising from a broadband multipath interference source.

2.14.2.1 Application to broadband multipath interference

This application illustrates the optimum transmit/receive configuration for maximizing output SINR in the presence of colored noise interference arising from a multipath broadband noise source. More specifically, for the single transmit/receive channel case, it derives the optimum transmit pulse modulation (i.e., pulse shape/spectrum).

Figure 14.3 illustrates a nominally broadband white noise source undergoing a series of multipath scatterings that in turn colors the noise spectrum [10]. Assuming (for simplicity) that the multipath reflections are dominated by several discrete specular reflections, the resultant signal can be viewed as the output of a causal tapped delay line filter (i.e., an FIR filter [3]) of the form

$$h_{mp}[k] = \alpha_0\delta[k] + \alpha_1\delta[k - 1] + \cdots + \alpha_{q-1}\delta[k - q - 1] \quad (14.14)$$

that is driven by white noise. The corresponding input-output transfer $H_{mp} \in \mathbb{C}^{N \times N}$ is thus given by

$$H_{mp} = \begin{bmatrix} h_{mp}[0] & 0 & \cdots & 0 \\ h_{mp}[1] & h_{mp}[0] & & \vdots \\ \vdots & & \ddots & 0 \\ h_{mp}[N-1] & \cdots & h_{mp}[1] & h_{mp}[0] \end{bmatrix}. \quad (14.15)$$

In terms of the multipath transfer matrix, H_{mp} , the colored noise interference covariance matrix is given by

$$\begin{aligned} E\{\mathbf{n}\mathbf{n}'\} &= E\left\{H_{mp}\mathbf{v}\mathbf{v}'H'_{mp}\right\} \\ &= H_{mp}E\left\{\mathbf{v}\mathbf{v}'\right\}H'_{mp} \\ &= H_{mp}H'_{mp} \\ &= R, \end{aligned} \quad (14.16)$$

where the driving white noise source $\mathbf{v} \in \mathbb{C}^N$ is a zero mean complex vector random variable with an identity covariance matrix:

$$E\{\mathbf{v}\mathbf{v}'\} = I. \quad (14.17)$$

Assuming a unity gain point target at the origin, that is, $h_T[k] = \delta[k]$, yields a target transfer matrix $H_T \in \mathbb{C}^{N \times N}$ given by

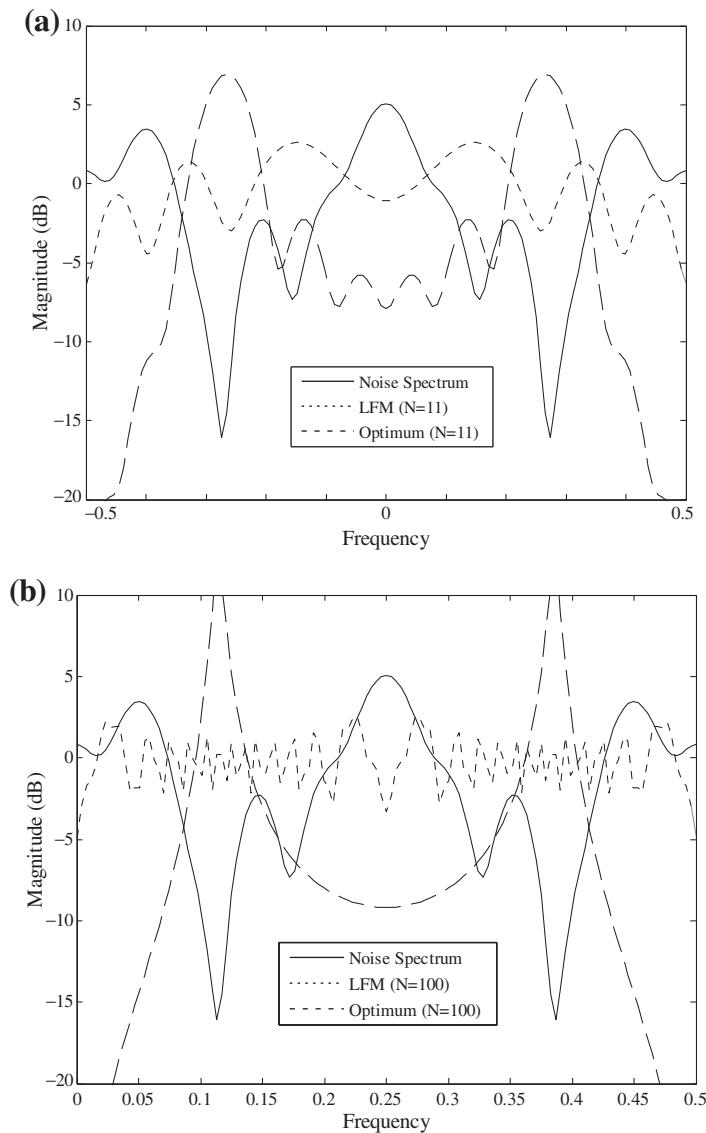
$$\begin{aligned} H_T &= \begin{bmatrix} h_T[0] & 0 & \cdots & 0 \\ h_T[1] & h_T[0] & & \vdots \\ \vdots & & \ddots & 0 \\ h_T[N-1] & \cdots & h_T[1] & h_T[0] \end{bmatrix} \\ &= I. \end{aligned} \quad (14.18)$$

While certainly a more complex (and thus realistic) target model could be assumed, we wish to focus on the impact the colored noise has on shaping the optimum transmit pulse. We will introduce more complex target response models in the target ID section.

Figure 14.4 shows the in-band interference spectrum for the case when $\alpha_0 = 1, \alpha_2 = 0.9, \alpha_5 = 0.5, \alpha_{10} = 0.2$, all other coefficients are set to zero. The total number of fast-time (range bin) samples was set to both a short-pulse case of $N = 11$ (Figure 14.4a) and a long-pulse case of $N = 100$ (Figure 14.4b). Note that the multipath colors the otherwise flat noise spectrum. Also displayed is the spectrum of a conventional (and thus non-optimized) LFM pulse with a time-bandwidth product, $\beta\tau$, of 5 (Figure 14.4a) and 50 (Figure 14.4b), respectively [11, 12].

Given R from (14.16), the corresponding whitening filter H_w is given by

$$H_w = R^{-\frac{1}{2}}. \quad (14.19)$$

**FIGURE 14.4**

Spectra of the colored noise interference along with conventional and optimal pulse modulations. (a) Short-pulse case where total duration for the LFM and optimum pulse are set to 11 range bins (fast-time taps). (b) Long-pulse case where total duration for the LFM and optimum pulse are set to 100 range bins. Note that in both cases the optimum pulse attempts to anti-match to the colored noise spectrum under the frequency resolution constraint set by the total pulse width.

Combining (14.19) with (14.18), the total composite channel transfer matrix H is thus given by

$$H = H_w H_T = H_w = R^{-\frac{1}{2}}. \quad (14.20)$$

Substituting (14.20) into (14.12) yields

$$R^{-1} \mathbf{s}_{\text{opt}} = \lambda \mathbf{s}_{\text{opt}}. \quad (14.21)$$

That is, the optimum transmit waveform is the maximum eigenvector associated with the inverse of the interference covariance matrix.

Displayed in Figures 14.4a and 14.4b are the spectra of the optimum transmit pulses obtained by solving (14.21) for the maximum eigenfunction-eigenvalue pair for the aforementioned short- and long-pulse cases, respectively. Note how the optimum transmit spectrum naturally emphasizes portions of the spectrum where the interference is weak—which is an intuitively satisfying result.

The SINR gain of the optimum short pulse, SINR_{opt} , relative to that of a nonoptimized chirp pulse, SINR_{LFM} , is

$$\text{SINR}_{\text{gain}} \triangleq \frac{\text{SINR}_{\text{opt}}}{\text{SINR}_{\text{LFM}}} = 7.0 \text{ dB}, \quad (14.22)$$

while for the long-pulse case

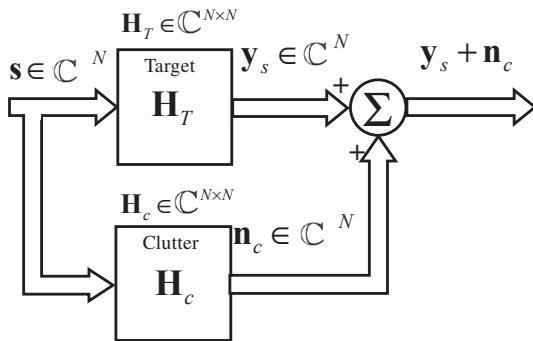
$$\text{SINR}_{\text{gain}} \triangleq \frac{\text{SINR}_{\text{opt}}}{\text{SINR}_{\text{LFM}}} = 24.1 \text{ dB}. \quad (14.23)$$

The increase in SINR for the long-pulse case is to be expected since it has finer spectral resolution and can therefore more precisely shape the transmit modulation to “anti-match” the interference. Of course, the unconstrained optimum pulse has certain practical deficiencies (e.g., poorer resolution, compression sidelobes) compared with a conventional pulse. We will revisit these issues when constrained optimization is introduced.

This application is similar in spirit to the spectral notching waveform design problem that arises when strong co-channel narrowband interferers are present [13]. In this case it is desirable not only to filter out the interference on receive but also to choose a transmit waveform that minimizes energy in the co-channel bands. The reader is encouraged to experiment with different notched spectra and pulse length assumptions and to apply (14.12) as indicated. Non-impulsive target models can also be readily incorporated.

2.14.3 Optimum transmit-receiver design for the clutter case: detection

Radar clutter refers to all unwanted reflections emanating from anything other than the desired target of interest [12]. It is thus a form of transmit-signal dependent interference, in contrast to the previously considered ACN case. Unfortunately, the joint optimization of the transmit and receive functions for the general additive colored noise plus clutter (signal-dependent noise) has been shown to result in a highly nonlinear problem [14] (though efficient iterative methods have been developed to solve these equations [14]). In practice, however, there is often a natural “separation principle” between additive colored noise (signal independent) and clutter (signal dependent). For example, narrowband electromagnetic

**FIGURE 14.5**

Radar signal block diagram for the clutter dominant case illustrating the direct dependency of the clutter signal on the transmitted signal.

interference (EMI) resulting from co-channel interference might require fast-time receiver and transmit spectral notching [13], leaving the slow-time or spatial DOF available for clutter suppression. Similarly, adaptive beamforming for broadband jammer nulling can be separated from the clutter suppression problem in a two-stage approach (see, e.g., [15]). We will thus concentrate in this section on the clutter dominant case and focus solely on maximizing the output signal-to-clutter ratio (SCR).

Unlike the previous additive colored noise case, clutter (i.e., channel reverberations) is a form of signal-dependent noise [16, 17] since the clutter returns depend on the transmit signal characteristics (e.g., transmit antenna pattern and strength, operating frequencies, bandwidths, polarization). Referring to Figure 14.5, the corresponding SCR at the input to the receiver is given by

$$\begin{aligned} \text{SCR} &= \frac{E\{\mathbf{y}_T' \mathbf{y}_T\}}{E\{\mathbf{y}_c' \mathbf{y}_c\}} \\ &= \frac{\mathbf{s}' E\{H_T' H_T\} \mathbf{s}}{\mathbf{s}' E\{H_c' H_c\} \mathbf{s}}. \end{aligned} \quad (14.24)$$

where $H_c \in \mathbb{C}^{N \times N}$ denotes the clutter transfer matrix, which is generally taken to be stochastic. Equation (14.24) is a generalized Rayleigh quotient [7] that is maximized when \mathbf{s} is a solution to the generalized eigenvalue problem

$$E\{H_T' H_T\} \mathbf{s} = \lambda E\{H_c' H_c\} \mathbf{s} \quad (14.25)$$

with corresponding maximum eigenvalue. When $E\{H_c' H_c\}$ is positive definite, (14.25) can be converted to an ordinary eigenvalue problem of the form we have already encountered, specifically,

$$E\{H_c' H_c\}^{-1} E\{H_T' H_T\} \mathbf{s} = \lambda \mathbf{s}. \quad (14.26)$$

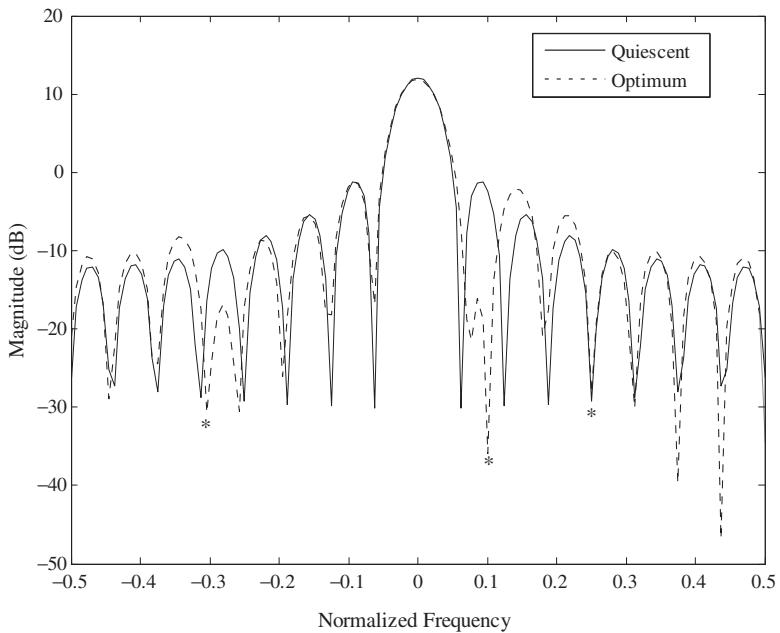
**FIGURE 14.6**

Illustration of proactive sidelobe target blanking on transmit achieved by maximizing the SCR. Note the presence of nulls in the directions of competing targets while preserving the desired mainbeam response.

2.14.3.1 Application: sidelobe clutter discrete suppression

Consider a narrowband $N=16$ element uniform linear array (ULA) with half-wavelength interelement spacing with a boresight pointing quiescent pattern (see Figure 14.6) [18]. In addition to the desired target at a normalized angle of $\bar{\theta} = 0$ (array boresight), there are strong sidelobe targets at $\bar{\theta}_1 = -0.3$, $\bar{\theta}_2 = +0.1$, $\bar{\theta}_3 = +0.25$, where a normalized angle is defined as

$$\bar{\theta} \triangleq \frac{d}{\lambda} \sin \theta. \quad (14.27)$$

In (14.27) d is the interelement spacing of the ULA, and λ is the operating wavelength (consistent units and narrowband operation assumed).

The presence of these targets (possibly large clutter discrete) could have been previously detected, thus making their AOAs known. Also, their strong sidelobes could potentially mask weaker mainlobe targets. With this knowledge, it is desired to minimize any energy from these targets leaking into the mainbeam detection of the target of interest by nulling on transmit, or placing transmit antenna pattern nulls in the directions of the unwanted targets.

For the case at hand, the (m, n) th elements of the target and interferer transfer matrices are given, respectively, by

$$[H_T]_{m,n} = e^{j\phi}(\text{const.}), \quad (14.28)$$

$$[H_c]_{m,n} = \alpha_1 e^{j2\pi(m-n)\bar{\theta}_1} + \alpha_2 e^{j2\pi(m-n)\bar{\theta}_2} + \alpha_3 e^{j2\pi(m-n)\bar{\theta}_3}, \quad (14.29)$$

where φ is an overall bulk delay (two way propagation) that does not affect the solution to (14.25) and will thus be subsequently ignored, and $[H_c]_{m,n}$ is the (m, n) th element of the clutter transfer matrix and consists of the linear superposition of the three target returns resulting from transmitting a narrow-band signal from the n th transmit element and receiving it on the m th receive element of a ULA that uses the same array for transmit and receive [2, 12]. Note that in practice there would be a random relative phase between the signals in (14.29), which for convenience we have ignored but which can easily be accommodated by taking the expected value of the kernel $H'_c H_c$.

Solving (14.25) for the optimum eigenvector yields the transmit pattern that maximizes the SCR, which is the pattern also displayed in Figure 14.6. The competing target amplitudes were set to 40 dB relative to the desired target and 0 dB of diagonal loading was added to $H'_c H_c$ to improve numerical conditioning and allow for its inversion. Although this is somewhat arbitrary, it does provide a mechanism for controlling null depth, that in practice is limited by the amount of transmit channel mismatch [19]. Note the presence of transmit antenna pattern nulls in the directions of the competing targets as desired.

An important caveat to the above was the assumption of perfect knowledge of the transmit array manifold—that is the assumption of a perfectly transmit calibrated AESA. This can only approximately be the case in practice. Indeed knowledge of the receive array manifold, which is generally easier to estimate on-the-fly, is not equivalent to that of the transmit manifold (since slightly different electrical pathways are taken). A MIMO method has been developed for transmit manifold calibration under the DARPA ISAT program and can be found in [20].

2.14.3.2 Application: optimum pulse shaping for maximizing SCR

In this simple example, we rigorously verify an intuitively obvious result regarding pulse shape and detecting a point target in uniform clutter: the best waveform for detecting a point target in distributed i.i.d clutter is itself an impulse (i.e., a waveform with maximal resolution), a well-known result rigorously proven by Manasse [21] using a different method.

Consider a unity point target, arbitrarily chosen to be at the temporal origin. Its corresponding impulse response and transfer matrix are respectively given by

$$h_T[n] = \delta[n] \quad (14.30)$$

and

$$H_T = I_{N \times N}, \quad (14.31)$$

where $I_{N \times N}$ denotes the $N \times N$ identity matrix. For uniformly distributed clutter, the corresponding impulse response is of the form

$$h_c[n] = \sum_{k=0}^{N-1} \tilde{\gamma}_k \delta[n - k], \quad (14.32)$$

where $\tilde{\gamma}_i$ denotes the complex reflectivity random variable of the clutter contained in the i th range cell (i.e., fast-time tap). The corresponding transfer matrix is given by

$$\tilde{H}_c = \begin{bmatrix} \tilde{\gamma}_0 & 0 & 0 & \cdots & 0 \\ \tilde{\gamma}_1 & \tilde{\gamma}_0 & & & \\ \tilde{\gamma}_2 & \tilde{\gamma}_1 & \tilde{\gamma}_0 & & \\ \vdots & & & \ddots & \\ \tilde{\gamma}_{N-1} & \tilde{\gamma}_{N-2} & \tilde{\gamma}_{N-3} & \cdots & \tilde{\gamma}_0 \end{bmatrix}. \quad (14.33)$$

Assuming that the $\tilde{\gamma}_i$ values are i.i.d., we have

$$E \{ \tilde{\gamma}_i^* \tilde{\gamma}_j \} = P_c \delta[i - j] \quad (14.34)$$

and thus

$$E \left\{ [\tilde{H}'_c \tilde{H}_c]_{i,j} \right\} = \begin{cases} 0, & i \neq j, \\ (N+1-i)P_c, & i = j, \end{cases} \quad (14.35)$$

where $[\cdot]_{i,j}$ denotes the (i,j) th element of the transfer matrix. Note that (14.35) is also diagonal (and thus invertible), but with non-equal diagonal elements.

Finally, substituting (14.31) and (14.35) into (14.26) yields

$$E \left\{ \tilde{H}'_c \tilde{H}_c \right\}^{-1} \mathbf{s} = \lambda \mathbf{s}, \quad (14.36)$$

where

$$E \left\{ \tilde{H}'_c \tilde{H}_c \right\}^{-1} = \frac{1}{P_c} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & & \\ & & \ddots & \\ 0 & \cdots & & d_N \end{bmatrix} \quad (14.37)$$

and

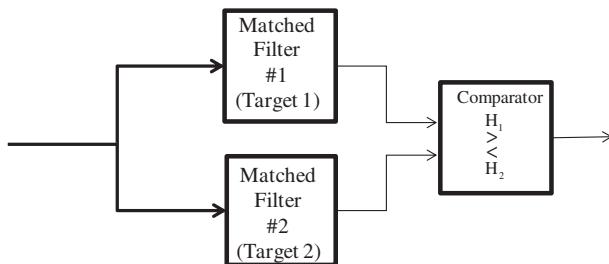
$$d_i \triangleq (N+i-1)^{-1}. \quad (14.38)$$

It is readily verified that the solution to (14.36) yielding the maximum eigenvalue is given by

$$\mathbf{s} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (14.39)$$

Thus the optimum pulse shape for detecting a point target is itself an impulse. This should be immediately obvious since it is the shape that excites the range bin only with the target and zeros out all other range bin returns that contain competing clutter.

Of course, transmitting a short pulse (much less an impulse) is problematic in the real world (e.g., creating extremely high peak power pulses) thus an approximation to a short pulse in the form of a spread spectrum waveform (e.g., LFM) is often employed [11]. This example also makes clear that in

**FIGURE 14.7**

Optimal receiver structure for the binary (two-target) hypothesis testing AGN problem.

uniform random clutter nothing is gained by sophisticated pulse shaping for a point target other than to maximize bandwidth (i.e., range resolution) [21]. The interested reader is referred to [22] for further examples of optimizing other DOF (e.g., angle-Doppler) for the clutter mitigation problem.

Up to this point we have been focused on judiciously choosing the transmit/receive DOF to maximize SINR or SCR (and thus ultimately detection). In the next section we will extend this framework to the target identification problem.

2.14.4 Optimizing the transmit-receive functions for target identification

Consider the problem of determining target type when two possibilities exist (the multitarget case is addressed later in this section). This can be cast as a classical binary hypothesis testing problem [6]:

$$\begin{aligned} \text{(Target 1) } H_1 : \mathbf{y}_1 + \mathbf{n} &= H_{T_1} \mathbf{s} + \mathbf{n}, \\ \text{(Target 2) } H_2 : \mathbf{y}_2 + \mathbf{n} &= H_{T_2} \mathbf{s} + \mathbf{n}, \end{aligned} \quad (14.40)$$

where H_{T_1} , H_{T_2} denote the target transfer matrices for targets 1 and 2, respectively. For the AGCN case, the well-known optimum receiver decision structure consists of a bank of matched filters (generally whitening matched filters), each tuned to a different target assumption, followed by comparator as shown in Figure 14.7 [6]. Note that (14.40) presupposes that either Target 1 or 2 is present, but not both. Also, it has been tacitly assumed that a binary detection test has been conducted to ensure that a target is indeed present [6]. Alternatively, the null hypothesis (no target present) can be included in the test as a separate hypothesis (see below).

Figure 14.8 illustrates the situation at hand. If Target-1 is present, the observed signal $\mathbf{y}_1 + \mathbf{n}$ will tend to cluster about the #1 point in observation space—which could include any number of dimensions relevant to the target ID problem (e.g., fast-time, angle, Doppler, polarization). The uncertainty sphere (generally ellipsoid for ACN case) surrounding #1 in Figure 14.7 represents the 1-sigma probability for the additive noise \mathbf{n} —and similarly for #2. Clearly, if \mathbf{y}_1 and \mathbf{y}_2 are relatively well separated, the probability of correct classification is commensurately high.

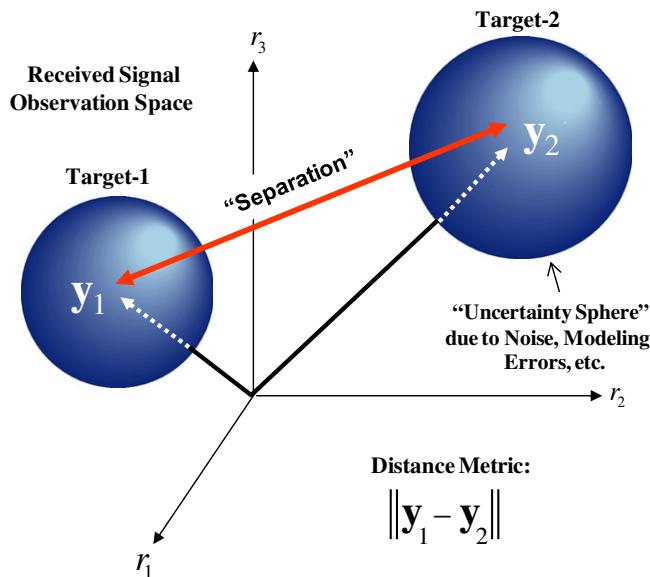
**FIGURE 14.8**

Illustration of the two-target ID problem. The goal of the joint transmitter/receiver design is to maximally separate the received signals in observation space, which in turn maximizes the probability of correct classification for the additive unimodal monotonic distributed noise case under fairly general conditions.

Significantly, \mathbf{y}_1 and \mathbf{y}_2 depend on the transmit signal \mathbf{s} , as shown in (14.40). Consequently, it should be possible to select an \mathbf{s} that maximizes the separation between \mathbf{y}_1 and \mathbf{y}_2 , thereby maximizing the probability of correct classification under modest assumptions regarding the conditional probability density functions (PDFs) (e.g., unimodality), that is,

$$\max_{\mathbf{s}} |\mathbf{d}' \mathbf{d}|, \quad (14.41)$$

where

$$\begin{aligned} \mathbf{d} &\triangleq \mathbf{y}_1 - \mathbf{y}_2 \\ &= H_{T_1}\mathbf{s} - H_{T_2}\mathbf{s} \\ &= (H_{T_1} - H_{T_2})\mathbf{s} \\ &\triangleq H\mathbf{s}. \end{aligned} \quad (14.42)$$

and where

$$H \triangleq H_{T_1} - H_{T_2}. \quad (14.43)$$

Substituting (14.42) into (14.41) yields

$$\max_{\mathbf{s}} |\mathbf{s}' H' H \mathbf{s}|. \quad (14.44)$$

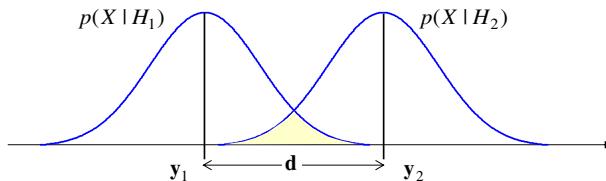
**FIGURE 14.9**

Illustration of the impact of the separation metric \mathbf{d} on the overlap of the conditional PDFs.

This is precisely of the form (14.10) and thus has a solution yielding maximum separation given by

$$(H'H)\mathbf{s}_{\text{opt}} = \lambda_{\max}\mathbf{s}_{\text{opt}}. \quad (14.45)$$

It is noted that (14.45) has an interesting interpretation: \mathbf{s}_{opt} is that transmit input that maximally separates the target responses and is thus the maximum eigenfunction of the transfer kernel $H'H$ formed by the difference between the target transfer matrices (i.e., (14.43)). Again if the composite target transfer matrix is stochastic, $H'H$ is replaced with its expected value $E\{H'H\}$ in (14.45).

Conditions for strict statistical optimality of the above optimization can be established by examining the properties of the conditional PDFs:

$$\begin{aligned} &(\text{Target 1}) H_1 : p(X|H_1, \mathbf{d}), \\ &(\text{Target 2}) H_2 : p(X|H_2, \mathbf{d}). \end{aligned} \quad (14.46)$$

For the case when the two conditional PDFs differ due to a shift in their means (generally multidimensional), an eminently reasonable assumption given the additive noise model of (14.40), the situation at hand can be illustrated as in Figure 14.9. We see that maximizing $\|\mathbf{d}\|$ minimizes the overlap of the PDFs provided:

$$F(X, H_1, H_2, \mathbf{d}_2) \geq F(X, H_1, H_2, \mathbf{d}_1), \quad (14.47)$$

where

$$F(X, H_1, H_2, \mathbf{d}) \triangleq \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |p(X|H_1, \mathbf{d}) - p(X|H_2, \mathbf{d})| dx_1 \dots dx_N \quad (14.48)$$

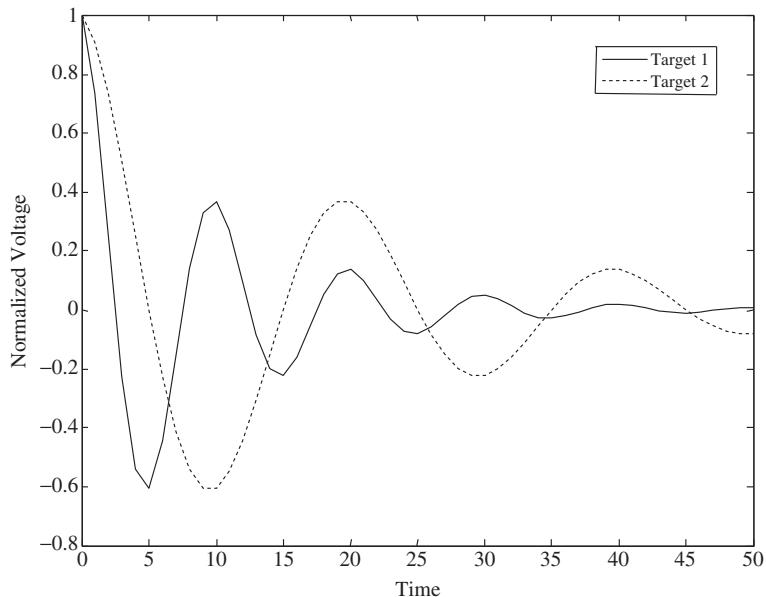
with

$$\|\mathbf{d}_2\| \geq \|\mathbf{d}_1\|. \quad (14.49)$$

Note that $F(X, H_1, H_2, \mathbf{d}) = 0$ when $\mathbf{d} = 0$. Clearly the class of multivariate Gaussian PDFs satisfies the above. Interestingly, even non-Gaussian PDFs can satisfy the above such as the class of unimodal distributions [23], and a broad class of elliptically contoured distributions [24]. This is in stark contrast to receiver optimization where generally the Gaussian assumption is required to ensure strict optimality [6].

2.14.4.1 Application: two-target identification example

Let $h_1[n]$ and $h_2[n]$ denote the impulse responses of targets #1 and #2, respectively (Figure 14.10). Figure 14.11 shows two different (normalized) transmit waveforms—LFM and optimum (per (14.45))—along with their corresponding normalized separation norms of 0.45 and 1, respectively, which

**FIGURE 14.10**

Baseband target impulse responses utilized for the two-target identification problem.

corresponds to 6.9 dB improvement in separation. To determine the relative probabilities of correct classification for the different transmit waveforms, one would first need to set the SNR level, which fixes the conditional PDF herein assumed to be circular Gaussian, and then to measure the amount of overlap to calculate the probability [6].

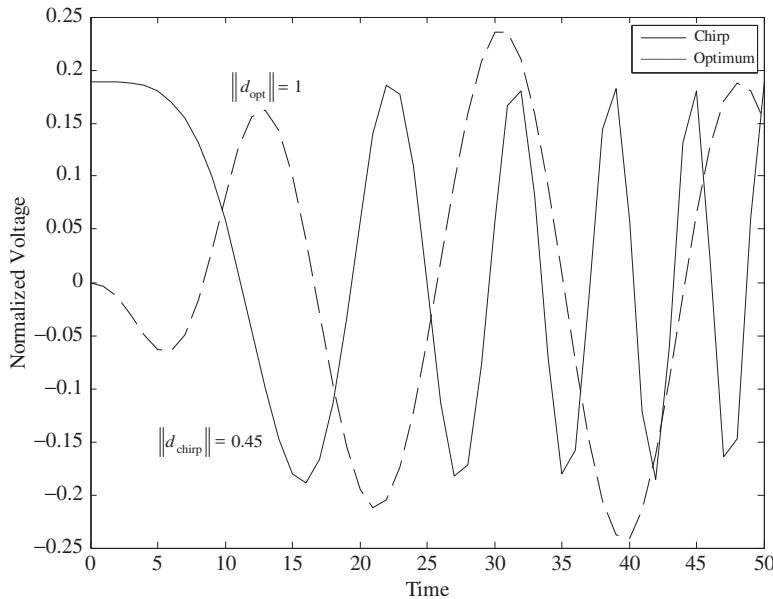
An examination of Figure 14.12 reveals the mechanism by which enhanced separation is achieved. It shows the Fourier spectrum of $H(\omega) = H_{T_1}(\omega) - H_{T_2}(\omega)$, along with that of $S_{\text{opt}}(\omega)$. Note that $S_{\text{opt}}(\omega)$ places more energy in spectral regions where $H(\omega)$ is large (i.e., spectral regions where the difference between targets is large, which is again an intuitively appealing result).

While pulse modulation was used to illustrate the optimum transmit design equations, we could theoretically have used any transmit DOF (e.g., polarization). The choice clearly depends on the application at hand.

2.14.4.2 Multi-target identification case

Given L targets in general, we wish to ensure that the L -target response spheres are maximally separated (an inverse sphere packing problem [25]). To accomplish this, we would like to jointly maximize the norms of the set of separations $\{\|\mathbf{d}_{ij}\| \mid i = 1 : L; j = i + 1 : L\}$:

$$\max_{\mathbf{s}} \quad \sum_{i=1}^L \sum_{j=i+1}^L |\mathbf{d}'_{ij} \mathbf{d}_{ij}|. \quad (14.50)$$

**FIGURE 14.11**

Transmit waveforms employed in the two-target identification example.

Since, by definition, \mathbf{d}_{ij} is given by

$$\mathbf{d}_{ij} \triangleq (H_{T_i} - H_{T_j}) \mathbf{s} \triangleq H_{ij} \mathbf{s}. \quad (14.51)$$

Equation (14.50) can be rewritten as

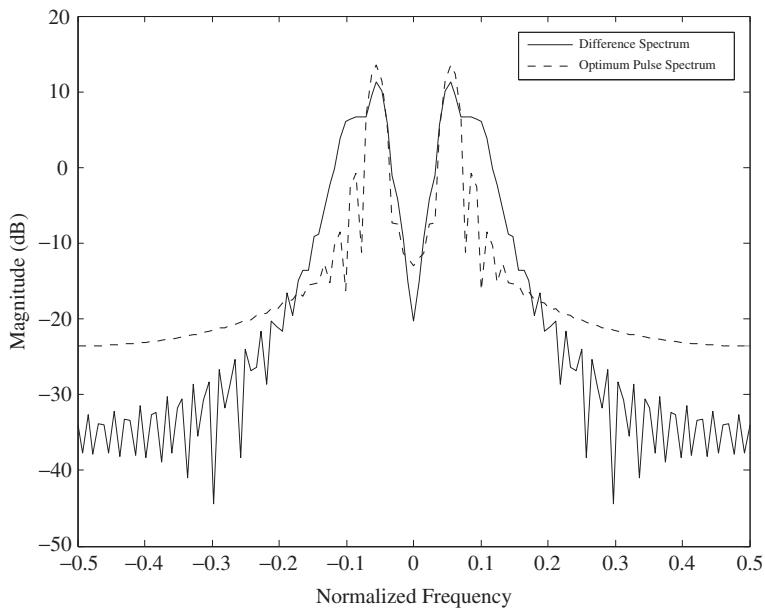
$$\max_{\mathbf{s}} \quad \mathbf{s}' \left(\sum_{i=1}^L \sum_{j=i+1}^L H_{ij}' H_{ij} \right) \mathbf{s} \triangleq \mathbf{s}' K \mathbf{s}. \quad (14.52)$$

Since $K \in C^{N \times N}$ is the sum of positive semidefinite matrices, it shares this same property, and thus the optimum transmit input satisfies

$$K \mathbf{s}_{\text{opt}} = \lambda_{\max} \mathbf{s}_{\text{opt}}. \quad (14.53)$$

2.14.4.3 Application: multitarget ID

Figure 14.13 depicts the impulse responses of three different targets, two of which are the same as in Section 2.14.4.1. Solving (14.52) and (14.53) yields an optimally separating waveform whose average separation defined by (14.50) is 1.0. This is compared with 0.47 for the LFM of Example 3.4, an improvement of 6.5 dB, which is slightly less than the previous two-target example. As expected, the optimum waveform significantly outperforms the unoptimized pulse waveform such as the LFM.

**FIGURE 14.12**

Comparison of the two-target difference spectrum and the optimum pulse spectrum. Note that the optimum pulse emphasizes parts of the spectrum where the two targets differ the most.

2.14.5 Constrained optimum transmit-receiver radar

Often there are a number of additional practical considerations (beyond finite norm) that may preclude transmitting the unconstrained optimum solutions developed so far. We will thus consider two cases of constrained optimization: linear and nonlinear—the latter is generally far more challenging and an area of active research.

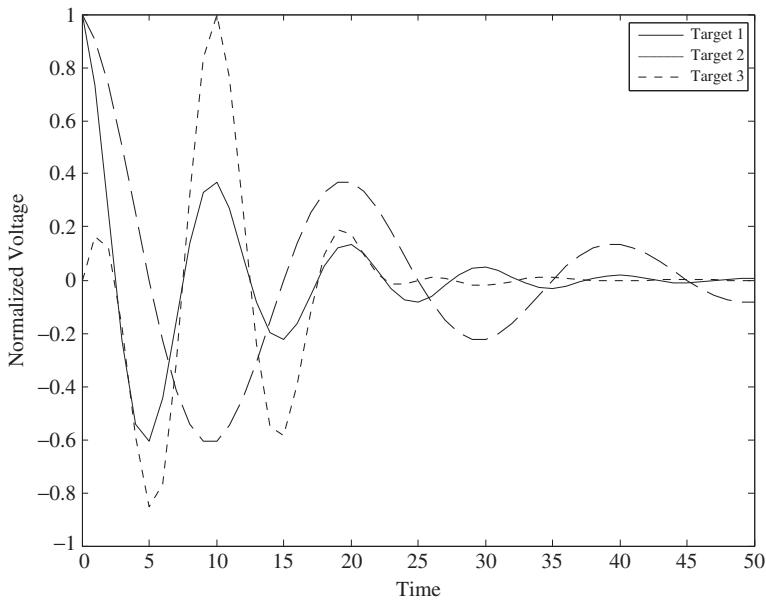
2.14.5.1 Case 1: Linear constraints

Consider the linearly constrained version of the input (transmitter) optimization problem:

$$\max_{\{s\}} |s' H' H s| \quad (14.54)$$

$$\text{subject to: } Gs = \mathbf{0}, \quad (14.55)$$

where $G \in \mathbb{C}^{Q \times N}$. To avoid the overly constrained case, it is assumed that $Q < N$. For example, the rows of G could represent steering vectors associated with known interferers such as unwanted targets or clutter discrete to which we wish to apply transmit nulls.

**FIGURE 14.13**

Target impulse responses used for the three-target identification problem.

Equation (14.55) defines the feasible solution subspace for the constrained optimization problem. It is straightforward to verify that the projection operator

$$P = I - G'(GG')^{-1}G \quad (14.56)$$

projects any $\mathbf{x} \in \mathbb{C}^N$ into the feasible subspace [26]. Thus, we can first apply the projection operator then perform an unconstrained subspace optimization to obtain the solution to (14.54) and (14.55), that is,

$$\max_{\mathbf{s}} |\mathbf{s}' P' H' H \mathbf{p}_s|. \quad (14.57)$$

From (14.53) it is readily apparent that the constrained optimum transmit input satisfies

$$P' H' H \mathbf{p}_{\text{opt}} = \lambda_{\max} \mathbf{s}_{\text{opt}}. \quad (14.58)$$

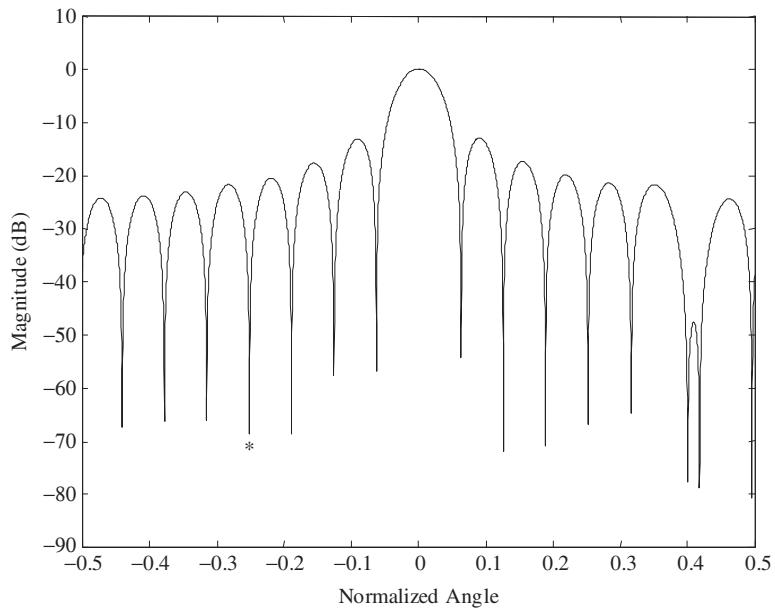
2.14.5.1.1 Application: pre-nulling on transmit

If there are known AOAs for which it is desired not to transmit (e.g., unwanted targets, clutter discrete, keep-out zones), it is possible to formulate a linearly constrained optimization accordingly.

Assume that there is a desired target at $\bar{\theta}_T$ as well as two keep-out angles (normalized) $\bar{\theta}_{I_1}$ and $\bar{\theta}_{I_2}$. The corresponding elements of the target transfer matrix $H_T \in \mathbb{C}^{N \times N}$, assuming an N -element ULA, are thus given by

$$[H_T]_{m,n} = e^{j2\pi(m-n)\bar{\theta}_T}, \quad (14.59)$$

where $[H_T]_{m,n}$ denotes the (m,n) th element of the target transfer matrix.

**FIGURE 14.14**

Example of a linearly constrained optimization in which two interferers are removed via the projection optimization approach.

The keep-out constraints have the form

$$\begin{aligned} \mathbf{0} &= G\mathbf{s} \\ &= \begin{bmatrix} \mathbf{s}'_{I_1} \\ \mathbf{s}'_{I_2} \end{bmatrix} \mathbf{s}, \end{aligned} \quad (14.60)$$

where

$$\mathbf{s}_{I_k} = \begin{bmatrix} 1 \\ e^{j2\pi\bar{\theta}_{I_k}} \\ \vdots \\ e^{j2\pi(N-1)\bar{\theta}_{I_k}} \end{bmatrix}. \quad (14.61)$$

Figure 14.14 shows the resulting constrained optimum transmit pattern for the case where $\bar{\theta}_T = 0$, $\bar{\theta}_{I_1} = -0.25$, $\bar{\theta}_{I_2} = 0.4$. As expected a peak is placed in the desired target direction with nulls simultaneously placed in the keep-out directions.

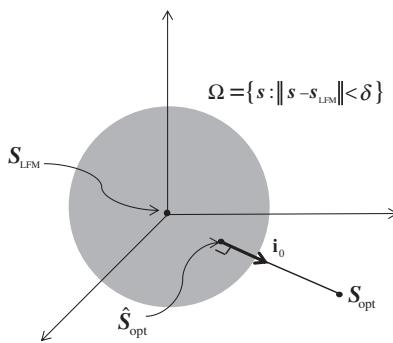
**FIGURE 14.15**

Illustration of a constrained optimization in which the signal should lie within a subspace (in this case convex) defined to be close to a prescribed transmit input (in this case an LFM waveform). The optimum relaxed projection is the point closest to the unconstrained optimum but still residing in the subspace.

2.14.5.2 Case 2: Nonlinear constraints

In practice other generally nonlinear constraints may arise. One family of such constraints relates to the admissibility of transmit waveforms, such as the class of constant modulus and stepped frequency waveforms [11], to name but a few.

For example, if it is desired to transmit a waveform that is nominally of the LFM type (or any other prescribed type) but that is allowed to modestly deviate to better match the channel characteristics, then the nonlinear constrained optimization has the form

$$\max_{\mathbf{s}} \quad |\mathbf{s}' H' H \mathbf{s}| \quad (14.62)$$

$$\text{subject to: } \|\mathbf{s} - \mathbf{s}_{\text{LFM}}\| \leq \delta. \quad (14.63)$$

The previous and similar problems cannot generally be solved in closed form. However, approximate methods can yield satisfactory results, and we will consider two that are based on very different approaches. These simpler methods could form the basis of more complex methods, such as seeding nonlinear search methods.

2.14.5.2.1 Relaxed projection approach

Figure 14.15 depicts the constrained optimization problem in (14.62) and (14.63). It shows the general situation in which the unconstrained optimum solution does not reside within the constrained (i.e., admissible) subspace Ω . In this particular case, the admissible subspace is a convex set [27], defined as

$$\Omega = \{\mathbf{s} : \|\mathbf{s} - \mathbf{s}_{\text{LFM}}\| \leq \delta\}. \quad (14.64)$$

From Figure 14.15 it is also immediately evident that the admissible waveform closest (in a normed sense) to the unconstrained optimum \mathbf{s}_{opt} lies on the surface of Ω along the direction \mathbf{i}_0 , which is the

unit norm vector that points from \mathbf{s}_{LFM} to \mathbf{s}_{opt} , i.e.,

$$\mathbf{i}_o \triangleq \frac{\mathbf{s}_{\text{opt}} - \mathbf{s}_{\text{LFM}}}{\|\mathbf{s}_{\text{opt}} - \mathbf{s}_{\text{LFM}}\|}. \quad (14.65)$$

Thus, the constrained waveform that is closest in norm to \mathbf{s}_{opt} is given by

$$\hat{\mathbf{s}}_{\text{opt}} = \mathbf{s}_{\text{LFM}} + \delta \mathbf{i}_o. \quad (14.66)$$

Note that if δ is allowed to relax to the point where $\delta = \|\mathbf{s}_{\text{opt}} - \mathbf{s}_{\text{LFM}}\|$, then $\hat{\mathbf{s}}_{\text{opt}} = \mathbf{s}_{\text{opt}}$.

2.14.5.2.2 Application: relaxed projection example

Here an LFM similarity constraint is imposed on the multipath interference problem considered previously. Specifically, in Figure 14.16, we plot the loss in SINR relative to the unconstrained long-pulse optimum solution previously obtained as a function of δ , which is varied between $0 \leq \delta \leq \|\mathbf{s}_{\text{opt}} - \mathbf{s}_{\text{LFM}}\|$. Note that for this example improvement generally monotonically increases with increasing δ (except for a very small region near the origin) and that sizeable SINR improvements can be achieved for relatively modest values of the relaxation parameter. In other words, a waveform with LFM-like properties can be constructed that still achieves significant SINR performance gains relative to an unoptimized LFM.

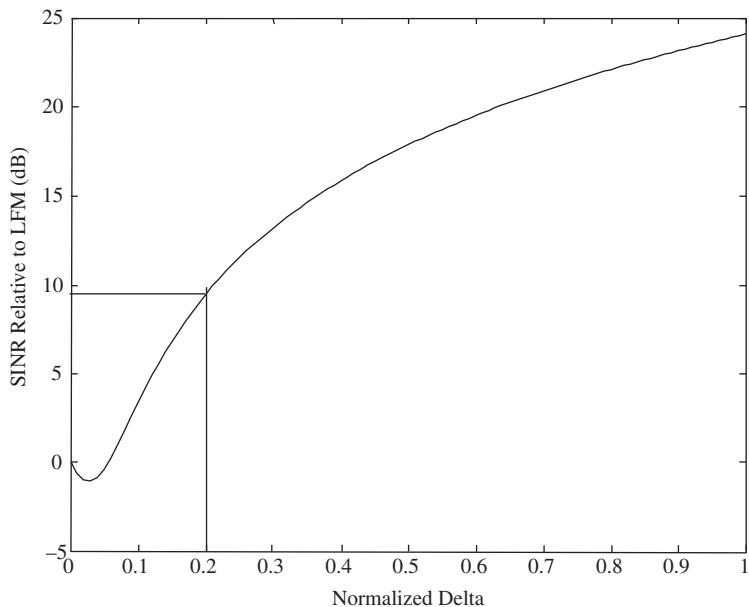
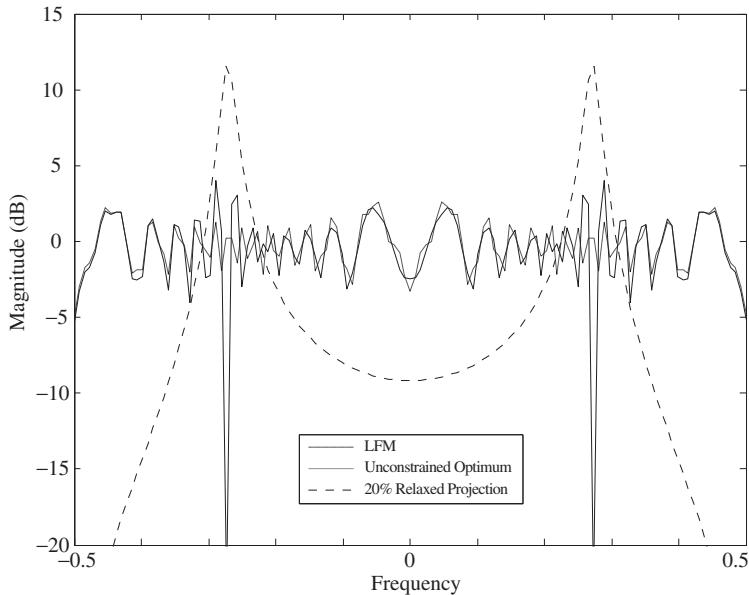


FIGURE 14.16

Illustration of the relaxed projection method for constrained optimization. The plot shows the SINR improvement relative to the unoptimized LFM waveform versus the normalized relaxation parameter δ . Note that for even a modest relaxation of 20% a nearly 10 dB gain in performance is achieved.

**FIGURE 14.17**

Comparison of the pulse spectra for the original LFM, unconstrained optimum, and 20% relaxed projection. Note how the relaxed pulse retains LFM-like spectral characteristics (and thus enhanced resolution for example) yet still achieves a 10 dB SINR improvement.

Figure 14.17 shows the spectra of the unoptimized LFM along with the unconstrained optimum and the relaxed projection pulse with a 20% relaxation parameter. Note how the relaxed pulse is significantly closer to the original LFM spectrum yet still achieves nearly a 10 dB improvement in SINR relative to the LFM waveform.

2.14.5.2.3 Constant modulus and the method of stationary phase

As has become apparent from the previous examples, spectral shaping plays a key role in achieving matching gains. The stationary phase method has been applied to the problem of creating a nonlinear frequency modulated (NLFM) pulse (and thus constant modulus in the sense that the modulus of the baseband complex envelope is constant, i.e., $|s(t)| = \text{constant}$) with a prescribed magnitude spectrum [3, 11].

Specifically, under fairly general conditions [3, 11] it is possible to relate instantaneous frequency ω of a NLFM waveform to time t [3, 11]:

$$\frac{1}{2\pi} \int_{-\infty}^{\omega} |S(\omega)|^2 d\omega = k \int_0^t dt = kt, \quad (14.67)$$

$$t \in [0, T],$$

where $|S(\omega)|$ is the magnitude spectrum of the optimum pulse. Here we have assumed a constant modulus for the NLFM waveform resulting in a integral that is simply proportional to time (see [3,11] for the more general nonconstant modulus case) as well as a finite and causal pulse.

Solving for ω as a function of t in (14.67) yields the frequency modulation that will result in a transmit pulse with a magnitude spectrum equal to $|S(\omega)|$, to within numerical and other theoretical limitations [3,11].

2.14.5.2.4 Application: NLFM to achieve constant modulus

Here we use the method of stationary phase to design a constant modulus NLFM pulse that matches the magnitude spectrum of the optimum pulse derived for the multipath interference problem previously considered.

Figure 14.18 shows the numerically obtained solution to (14.67) (i.e., ω versus t) along with the optimum pulse spectrum (long-pulse case). Note that as one would intuit, the frequency modulation dwells at frequencies where peaks in the optimum pulse spectrum occur and conversely note the regions in which the modulation speeds up to avoid frequencies where the optimum pulse spectrum has nulls or lower energy content.

The constant modulus NLFM waveform so constructed was able to achieve an output SINR that was within 6.0 dB of optimum compared with a 24 dB loss using an LFM waveform of same energy and duration.

It is natural to ask if a NLFM waveform with the same spectral magnitude as the optimum pulse (but not necessarily the same phase) will enjoy some (if not all) of the matching gains. For the steady-state case (infinite time duration) this is indeed true, since from Parseval's [3] theorem the output energy is related to only the spectral magnitudes (i.e., without their phases) of the input pulse and channel transfer function, that is,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |Y(\omega)|^2 d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 |S(\omega)|^2 d\omega, \quad (14.68)$$

where $Y(\omega)$, $H(\omega)$, and $S(\omega)$ denote the Fourier transforms of the channel output, channel impulse response, and input pulse, respectively. Note that the output energy in (14.68) depends on the spectral magnitude of the input pulse (steady-state)—not the phase. Thus, in theory an NLFM waveform that exactly matches the optimum pulse magnitude spectrum will achieve the same matching gains in the steady-state limit (infinite pulse duration) for all square integrable (finite norm) functions.

2.14.6 Open issues and problems

Perhaps amongst the foremost challenges is the development of robust and effective adaptive methods that can accurately estimate the full dimensional transmit-receive channel in real-time and solve for the optimum transmit-receive configuration—no mean feat. A basic outline of methods that include traditional sample statistics methods, orthogonal MIMO waveform concepts for channel probing, and knowledge-aided (KA) methods for leveraging any and all prior and/or externally available channel information can be found in [22] (and references cited therein).

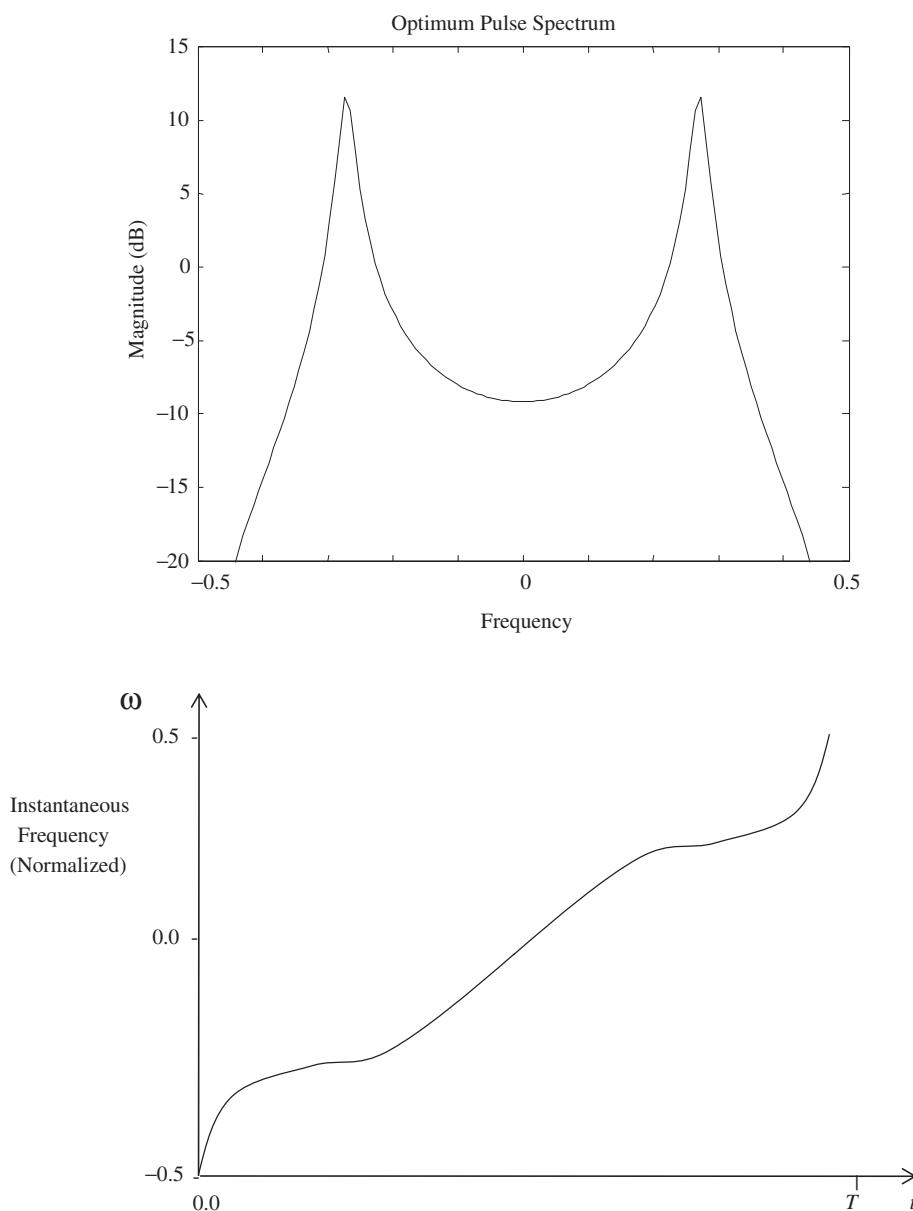
**FIGURE 14.18**

Illustration of the use of the method of stationary phase to create a constant modulus NLFM pulse whose spectral magnitude matches that of the optimum pulse. The NLFM pulse was able to achieve an output SINR that was within 6.0 dB of the optimum compared with a 24 dB loss using an LFM waveform of same energy and duration.

Even if real-time channel characterization methods can be developed, the issue of solving a generally constrained multidimensional transmit-receive waveform optimization problem remains. Here again it is desired to develop a set of efficient and reliable algorithms amenable to real-time implementation. Fortunately, steady advances in HPEC ensure that ever more increasing real-time compute power should be available.

2.14.7 Conclusions and future trends

In this chapter, the fundamental theory for joint optimization of the transmit and receive functions was developed from first principles and applied to the maximization of SINR, SCR, and correct classification for the target ID problem. Constrained optimization was introduced to address additional requirements that often arise in practice, such as the use of constant modulus waveforms to maximize transmitter efficiency. However, many theoretical and practical issues remain before the new field of adaptive transmit technology can be fully exploited.

As mentioned above, many challenges remain in realizing the full potential of joint transmit-receive optimization. Consequently, much of the current research trends in this area are focused on:

- Real-time methods for accurate multidimensional channel characterization.
- Robust and efficient algorithms for solving the generally constrained transmit-receive design equations.
- Leveraging of knowledge-aided (KA) methods such as those developed under the DARPA/AFRL KASSPER project [28,29].
- Combining other waveform diversity concepts with adaptive transmit-receive optimization to allow for enhanced multi-functionality.

Glossary

MIMO	A system with multiple inputs and outputs. In radar it refers to multiple transmit and receive DoFs (usually multiple antennas for example)
SINR	Signal-to-interference-plus-noise ratio. A measure of signal strength relative to all sources of interference (clutter, jamming, receiver noise, etc.)
Colored Noise	A random (noise) process that has a non-identity matrix covariance matrix (usually involving non-zero off diagonal terms)
Whitening Filter	The first stage in an optimum matched filter receiver when the additive noise has some correlation. The filter “whitens” the noise spectrum, which is then followed by a standard matched filter
Cauchy-Schwarz theorem	In its simplest form, a theorem that states that the dot product of two vectors is maximum when they are co-linear
Multipath	Electromagnetic propagation through a non-line-of-sight (NLOS) channel
FIR Filter	Finite impulse response filter. A discrete time LTI tapped delay line filter
LFM	Linear frequency modulation. A type of radar waveform whose instantaneous frequency varies linearly with time over a prescribed bandwidth

Radar Clutter	Radar echo returns from objects other than the desired target(s). Ground reflections is the most common form of clutter
SCR	Signal-to-clutter ratio. A relative measure of signal strength to clutter power
Signal Dependent Noise	Interference whose statistics depend on the radar's transmissions
Constant Modulus	Refers to a class of waveforms whose complex envelope (modulation) has a constant modulus, i.e., $ s(t) = 1$
NLFM	Nonlinear FM. A class of constant modulus waveform whose instantaneous frequency changes in a non-linear fashion

Acronyms

Acronyms that are commonly used in this chapter include the following:

ACN	additive colored noise
AESA	active electronically scanned array
AOA	angle-of-arrival
CA-CFAR	cell-averaging constant false alarm ratio
CNR	clutter-to-noise ratio
DOFs	degrees-of-freedom
FIR	finite impulse response
GMTI	ground moving target indication
KA	knowledge-aided
KASSPER	knowledge-aided sensor signal processing and expert reasoning
LFM	linear frequency modulation
PDF	probability density function
PRF	pulse repetition frequency
Rx	receiver
SAR	synthetic aperture radar
SCR	signal-to-clutter ratio
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
STAP	space-time adaptive processing
Tx	transmitter
ULA	uniform linear array

Relevant Theory: Signal Processing Theory

See Vol. 1, Chapter 2 Continuous-Time Signals and Systems

See Vol. 1, Chapter 3 Discrete-Time Signals and Systems

See Vol. 1, Chapter 4 Random Signals and Stochastic Processes

See Vol. 1, Chapter 12 Adaptive Filters

References

- [1] J.L. Lawson, G.E. Uhlenbeck, Threshold signals, in: MIT Radiation Laboratory Series, vol. 24, McGraw-Hill, New York, 1950.
- [2] J.R. Guerci, Space-Time Adaptive Processing for Radar, Artech House, Norwood, MA, 2003.
- [3] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1984.
- [4] D.K. Barton, Modern Radar System Analysis, Artech House, Norwood, MA, 1988.
- [5] A. Papoulis, Circuits and Systems: A Modern Approach, Holt, Rinehart and Winston, New York, 1980.
- [6] H.L.V. Trees, Detection, Estimation and Modulation Theory, Part I, Wiley, New York, 1968.
- [7] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge (England), New York, 1990.
- [8] G. Strang, Introduction to Linear Algebra, Wellesley Cambridge Press, 2003.
- [9] D.A. Pierre, Optimization Theory With Applications, Courier Dover Publications, 1986.
- [10] J.R. Guerci, S.U. Pillai, Theory and application of optimum transmit-receive radar, in: The Record of the IEEE 2000 International Radar Conference, 2000, pp. 705–710.
- [11] C.E. Cook, M. Bernfeld, Radar Signals, Academic Press New York, 1967.
- [12] M.A. Richards, Fundamentals of Radar Signal Processing, McGraw-Hill, 2005.
- [13] M.J. Lindenfeld, Sparse frequency transmit-and-receive waveform design, IEEE Trans. Aerosp. Electron. Syst. 40 (2004) 851–861.
- [14] S.U. Pillai et al., Optimal transmit-receiver design in the presence of signal-dependent interference and channel noise, IEEE Trans. Inform. Theory 46 (2000) 577–584.
- [15] D.J. Rabideau, Closed loop multistage adaptive beamforming, in: Conference Record of the 33rd Asilomar Conference on Signals, Systems and Computers, 1999, pp. 98–102.
- [16] H.L. Van Trees, E. Detection, Modulation Theory, Part II, John Wiley and Sons, New York, 1971.
- [17] H.L.V. Trees, Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise, Krieger Publishing Co., Inc., 1992.
- [18] S. Pillai, C. Burns, Array Signal Processing, Springer-Verlag, New York, 1989.
- [19] R.A. Monzingo, T.W. Miller, Introduction to Adaptive Arrays, SciTech Publishing, 2003.
- [20] J. Guerci, E. Jaska, ISAT—innovative space-based-radar antenna technology, in: IEEE International Symposium on Phased Array Systems and Technology, 2003, pp. 45–51.
- [21] R. Manasse, The Use of Pulse Coding to Discriminate Against Clutter, vol. AD0260230, Defense Technical Information Center (DTIC), June 7, 1961.
- [22] J.R. Guerci, Cognitive Radar: The Knowledge-Aided Fully Adaptive Approach, Artech House, Norwood, MA, 2010.
- [23] I.A. Ibragimov, On the composition of unimodal distributions, Theory Probab. Appl. 1 (1956) 255.
- [24] K.T. Fang, Y.T. Zhang, Generalized Multivariate Analysis, Science Press, Beijing, 1990.
- [25] W.Y. Hsiang, On the sphere packing problem and the proof of Kepler's conjecture, Int. J. Math. 4 (1993) 739–831.
- [26] W. Gander et al., A constrained eigenvalue problem, Linear Algebra Appl. 114 (1989) 815–839.
- [27] D.C. Youla, H. Webb, Image restoration by the method of convex projections: Part 1. Theory, IEEE Trans. Med. Imag. 1 (1982) 81–94.
- [28] F. Gini, M. Rangaswamy (Eds.), Knowledge Based Radar Detection, Tracking and Classification Adaptive and Learning Systems for Signal Processing, Communications and Control Series, Wiley-IEEE Press, New York, 2008.
- [29] J.R. Guerci, E.J. Baranowski, Knowledge-aided adaptive radar at DARPA: an overview, IEEE Signal Process. Mag. 23 (2006) 41–50.

Multitarget Multisensor Tracking

15

X. Chen, R. Tharmarasa, and T. Kirubarajan

ECE Department, McMaster University, Hamilton, Ontario, Canada

3.15.1 Introduction

Multisensor-multitarget tracking is an emerging technology in which measurements from several sensors are combined such that resulting tracks are significantly better than that obtained when these devices operate individually. Recent advances in sensor technologies, signal processing techniques and improved processor capabilities make it possible for large amounts of data to be fused in real-time. These technical advancements allow the use of many sophisticated algorithms and robust mathematical techniques in multisensor-multitarget tracking. Furthermore, multisensor-multitarget tracking has received significant attention for military applications. Such applications involve a wide range of expertise including filtering, tracking initialization and maintenance, data association, and performance evaluation.

Three major types of architecture, namely, centralized, distributed and decentralized, are commonly used in multisensor-multitarget tracking applications [9, 62, 92, 126]. In the centralized architecture, there are several sensors monitoring the region of interest with only one fusion center. All sensors report their measurements to the fusion center. It is the fusion center's responsibility to process all acquired measurements and update the tracks. The single sensor-multitarget tracking problem can be considered as a special case of the centralized architecture, where only one sensor is deployed to observe the region of interest. In the distributed multisensor-multitarget tracking architecture, there are several fusion centers. One of them is the Central Fusion Center (CFC) and the remaining ones are Local Fusion Centers (LFCs). Measurements generated by the sensors are first processed by the LFCs and local tracks are updated inside each LFC. Then, local tracks from each LFC are reported to the CFC and the track-to-track fusion is accomplished by the CFC to form the global track set. In decentralized tracking architecture, each fusion center (FC) can be considered as a combination of LFC and CFC. Each FC is connected with several sensors and measurements reported by those sensors are used to update the track state inside the FC. Furthermore, each FC will also do track-to-track fusion whenever it receives additional information from its neighboring FCs. Usually, without a major modification, the algorithms developed for the distributed tracking architecture can be used to handle the decentralized tracking architecture. Regardless whether the sensor measurements are processed in the centralized or the distributed architecture, the data of each sensor has to be converted to a common coordinate system before multisensor-multitarget tracking, i.e., sensor registration and data alignment [39, 54, 74].

Filtering plays a vital role in multitarget tracking by obtaining the state estimate from the measurements received from one or more sensors. Tracking filters [10,31] can be broadly categorized as either linear or nonlinear. The Kalman filter [10,53] is a widely known recursive filter that is most suited for linear Gaussian systems. However, most systems are inherently nonlinear. The extensions of Kalman filter, such as extended Kalman filter (EKF) [10] and unscented Kalman filter (UKF) [52,105,128] are applicable to nonlinear systems. Both EKF and UKF are restricted in that the resulting probability densities are approximated as Gaussian. When the system is nonlinear and non-Gaussian, particle filters (or sequential Monte Carlo methods) [3,33,38,100] provide better estimates than many other filtering algorithms. In general, a tracking filter requires a model for target dynamics, and a model mismatch would diminish the performance of the filter. Thus, different models may be required to describe the target dynamics accurately at different times, especially in the case of maneuvering targets, whose kinematic models may evolve in a time-varying manner. Multiple model tracking algorithms such as the Interacting Multiple Model (IMM) [1,10,14–16,18] estimator, which contains a bank of models matched to different modes of possible target dynamics, would perform better in such situations.

Data association [9,13,113] is an essential component in multisensor-multitarget tracking due to the uncertainty in the data origin. Data association refers to the methodology of correctly associating the measurements to tracks, measurements to measurements [58,106] or tracks to tracks [2,6,8,9,19,22,127], depending on the fusion architecture. To address data association, a number of techniques have been developed and two widely used are single-frame assignment algorithm [11,95] and the multi-frame assignment algorithm [11,17,27,32,59,63,95].

Many algorithms have been proposed for the single sensor-multitarget tracking problem, such as the Probability Data Association (PDA) algorithm and the Joint Probability Data Association (JPDA) algorithm [9], the Multiple Hypotheses Tracking (MHT) algorithm [12], and the Probability Hypothesis Density (PHD) filter [78]. In PDA and JPDA algorithms, for each scan the track-to-measurement association events are enumerated and combined probabilistically, while in the MHT algorithm the track-to-measurement association history over several scans are enumerated and updated. In PHD filter, the track-to-measurement association events are not explicitly constructed. Although these algorithms are originally proposed to handle the single sensor-multitarget tracking problem, they are also widely used as the backbone for the multisensor-multitarget tracking.

In many scenarios, after the signal detection process, clutter points provided by the sensor (e.g., sonar, infrared sensor, radar) are not distributed uniformly in the surveillance region as assumed by most tracking algorithms. On the other hand, in order to obtain accurate results, the target tracking filter requires information about clutter's spatial intensity. Thus, nonhomogeneous clutter spatial intensity has to be estimated from the measurement set and the tracking filter's output. Also, in order to take advantage of existing tracking algorithms, it is desirable for the clutter estimation method to be integrated into the tracker itself.

Performance evaluation is very important for multisensor-multitarget tracking problem, especially when the performance of different tracking algorithms needs to be compared. Many measures of performance have been proposed in multisensor-multitarget tracking literatures. These measures can be divided into two different classes: sensor-related measures and tracker-related measures [40]. Sensor-related measures are independent of the tracking algorithm, therefore, most of them are not useful for the performance evaluation of multiple trackers. One exception is the Posterior Cramér-Rao Lower Bound (PCRLB) of tracking [10,44,46], which provides a minimum bound of any tracking estimation.

On the other hand, tracker-related measures have been widely used in the tracker performance evaluation. For general multitarget tracking problem, tracker-related measures have been defined in terms of cardinality, time and accuracy measures [66, 103]. To evaluate a general multitarget tracking problem, a combination of tracker-related measures should be used, because as shown in [29], it has been observed that the tracker may have provided inaccurate results while some measures show correct and satisfactory performances.

In this chapter, various multisensor-multitarget tracking architectures, estimators for spatial clutter intensity, filters for linear and nonlinear systems, algorithms for data associations and multitarget tracking, techniques used in centralized and distributed track-to-track fusion are discussed in detail. In addition, their quantitative and qualitative merits are discussed. Various combinations of these algorithms will provide a complete tracking framework for multisensor networks with application to civilian as well as military problems. For example, the tracking and fusion techniques discussed here are applicable to fields like air traffic control, air/ground/maritime surveillance, mobile communication, transportation, video monitoring and biomedical imaging/signal processing. The tracker performance evaluation, including its guiding principle and several measures of performance, is also discussed in this chapter. A challenging scenario with many closely-spaced targets is used to compare several multitarget tracking algorithms.

3.15.2 Formulation of multisensor-multitarget tracking problems

In a multisensor surveillance systems, several sensors such as radar, infrared (IR), and sonar, report their measurements to the tracker at regular intervals of time (scans or data frames). However, not all measurements are originated from the targets of interest; some measurements may come from physical background objects such as clutter, and others may be generated by thermal noise. In other words, there is a measurement origin ambiguity. Furthermore, in some scans, the target of interest does not produce any measurements at all (i.e., the probability of detection is less than unity). The objective of a typical multisensor-multitarget tracking system is to first partition all sensors' measurements into sets, such that all observations in one set are produced by the same object (i.e., data-association process); then the measurements corresponding to the same object are processed in order to estimate the state of the object (i.e., filtering process) [9, 12].

To handle the multisensor-multitarget tracking problem, usually the Bayesian approach is applied. In the Bayesian approach, the final goal is to construct the posterior probability density function (pdf) of the multitarget state given all the received measurements so far. Since this pdf contains all available statistical information, it is the complete solution to the multisensor-multitarget tracking problem. In principle, given a cost function, it is always possible to obtain the optimal estimate under the given cost function from the posterior probability. Note that, the state of the multitarget system should be a combination of the number of the targets and the state of each target, because in a real scenario both of them are random and unknown [78, 120].

To distinguish the target-originated measurements from the clutter and estimate the state of the multitarget system, the following three models, namely, the target dynamic model, the sensor model, and the clutter model are crucial for all multisensor-multitarget tracking systems.

3.15.2.1 Target dynamic models

Target dynamic model, which is also known as system model, describes the evolution of the state with time, and is given by

$$x(k+1) = f(k, x(k)) + G(k)v(k), \quad (15.1)$$

where f is, in general, a nonlinear function, $x(k)$ is the state of the target, and $v(k)$ is the process noise, which is usually assumed to be Gaussian. The covariance of the process noise multiplied by the gain $G(k)v(k)$ is

$$Q(k) = G(k)q(k)G(k)'. \quad (15.2)$$

The following models are widely used by the multitarget tracker as the target dynamic model: [10]

- *Constant velocity:*

The state vector in one generic coordinate is

$$x = [\xi \ \dot{\xi}]'. \quad (15.3)$$

The F and G in one generic coordinate are

$$F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad (15.4)$$

$$G = \begin{bmatrix} T^2/2 \\ T \end{bmatrix}. \quad (15.5)$$

The process noise standard deviation σ_v is

$$\sigma_v = w_v A_{\max}, \quad (15.6)$$

where w_v is the scaling factor and A_{\max} is the maximum acceleration.

- *Constant acceleration:*

The state vector in one generic coordinate is

$$x = [\xi \ \dot{\xi} \ \ddot{\xi}]'. \quad (15.7)$$

The F and G in one generic coordinate are

$$F = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}, \quad (15.8)$$

$$G = \begin{bmatrix} T^2/2 \\ T \\ 1 \end{bmatrix}. \quad (15.9)$$

The process noise standard deviation σ_v is

$$\sigma_v = w_v J_{\max} T, \quad (15.10)$$

where J_{\max} is the maximum jerk.

- *Coordinated turn:*

The state vector is

$$x = [\xi \ \dot{\xi} \ \eta \ \dot{\eta} \ \omega]'. \quad (15.11)$$

The F and G are

$$F = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & \frac{-(1-\cos(\omega T))}{\omega} & 0 \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) & 0 \\ 0 & \frac{(1-\cos(\omega T))}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} & 0 \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (15.12)$$

$$G = \begin{bmatrix} \frac{T^2}{2} & 0 & 0 \\ T & 0 & 0 \\ 0 & \frac{T^2}{2} & 0 \\ 0 & T & 0 \\ 0 & 0 & T \end{bmatrix}. \quad (15.13)$$

The process noise standard deviations σ_{vv} and σ_{vw} are

$$\sigma_{vv} = w_{vv} A_{\max}, \quad (15.14)$$

$$\sigma_{vw} = w_{vw} T_{\max}, \quad (15.15)$$

where w_{vv} and w_{vw} are the scaling factors for velocity and turn rate, respectively, T_{\max} is the maximum turn rate change in unit time.

3.15.2.2 Sensor models

Sensor model, which is also known as measurement model, is a model relating the noisy measurements to the state and given by

$$z(k) = h(k, x(k)) + w(k), \quad (15.16)$$

where h is, in general, nonlinear functions, $x(k)$ is the state of the target, $z(k)$ is the measurement vector, and $w(k)$ is the measurement noise at measurement time k , which is usually assumed to be Gaussian. The following measurement models are typical for the multisensor-multitarget tracking system:

- x -position

$$h = x. \quad (15.17)$$

- y -position

$$h = y. \quad (15.18)$$

- z -position

$$h = z. \quad (15.19)$$

- Range, r

$$h = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2}. \quad (15.20)$$

For 2D tracking, the terms related to z must be deleted.

- Range rate, \dot{r}

$$h = ((x - x_s)\dot{x} + (y - y_s)\dot{y} + (z - z_s)\dot{z})/r. \quad (15.21)$$

For 2D tracking, the terms related to z must be deleted.

- Azimuth, θ

$$h = \text{atan} \left(\frac{y - y_s}{x - x_s} \right). \quad (15.22)$$

- Azimuth-north θ_n

$$h = \text{atan} \left(\frac{x - x_s}{y - y_s} \right). \quad (15.23)$$

- Elevation γ

$$h = \text{atan} \left(\frac{z - z_s}{\sqrt{(x - x_s)^2 + (y - y_s)^2}} \right). \quad (15.24)$$

- Elevation-vertical, γ_v

$$h = \text{atan} \left(\frac{\sqrt{(x - x_s)^2 + (y - y_s)^2}}{z - z_s} \right). \quad (15.25)$$

- Bi-static range, r_b

$$h = \sqrt{(x - x_r)^2 + (y - y_r)^2} + \sqrt{(x - x_t)^2 + (y - y_t)^2}. \quad (15.26)$$

- Bi-static range rate, \dot{r}_b

$$h = \frac{(x - x_r)}{R_r}\dot{x} + \frac{(y - y_r)}{R_r}\dot{y} + \frac{(x - x_t)}{R_t}\dot{x} + \frac{(y - y_t)}{R_t}\dot{y}. \quad (15.27)$$

In (15.27),

$$R_r = \sqrt{(x - x_r)^2 + (y - y_r)^2}, \quad R_t = \sqrt{(x - x_t)^2 + (y - y_t)^2}$$

3.15.2.3 A clutter model

For a sensor with N resolution cells, sometimes detections will be declared in those cells that are pointed to a region without any targets of interest. The detection, which is not produced by any targets of interest, is known as a false alarm (i.e., clutter). Assume:

- The events of detection in each cell are independent of each other.
- The probability of false alarm is equal to p_{FA} in each cell and $p_{FA} \ll 1$.
- The number of the resolution cells is large, such that $p_{FA} \cdot N > 1$.

Then the probability mass function (pmf) of the number of false alarm in these N resolution cells, $\mu_{\text{FA}}(m)$, is approximately following the Poisson distribution [9]

$$\mu_{\text{FA}}(m) = \exp \{-N \cdot p_{\text{FA}}\} \frac{(N \cdot p_{\text{FA}})^m}{m!}. \quad (15.28)$$

Furthermore, the spatial distribution of the false alarm is uniform based on the above three assumptions. Thus, if the granularity due to the size of the resolution cells can be neglected, the pdf of a false measurement, i.e., the clutter spatial intensity normalized by the expected number of clutter in the measurement space, is [9]

$$\bar{\lambda}_{\text{FA}} = p(z|z \text{ is a false alarm}) = \frac{1}{V}, \quad (15.29)$$

where V represents the volume of the sensor's measurement space. The un-normalized clutter spatial intensity is

$$\lambda_{\text{FA}} = \frac{N \cdot p_{\text{FA}}}{V}. \quad (15.30)$$

3.15.2.4 Spatial clutter intensity estimation

Many target tracking algorithms assume that the clutter background is known or at least homogeneous. However, in real tracking problems, the distribution of clutter is often unknown and spatially non-homogeneous. Thus, there is usually a mismatch between the true spatial distribution of clutter points and the spatial distribution model used in the tracking filter. This mismatch may result in a high false track acceptance rate or a long delay of track initialization. Therefore, it is desirable for the tracking filter to estimate the spatial intensity of clutter from the measurement set. Also, due to the fact that target-originated measurement points and clutter points are indistinguishable before data association in the tracker, the output of tracking filter should also be used in order to get an unbiased estimate of clutter spatial intensity. Furthermore, estimation methods for clutter spatial distribution should be compatible with the existing target tracking algorithms, otherwise their application range would be limited.

One way to estimate clutter spatial intensity is to assume that clutter points are uniformly distributed in the validation gate and then use the sample spatial intensity as the estimate of clutter's spatial intensity [9]. However, this method is based on the current measurement set alone and its performance relies on the volume of validation gate. For example, if the gate is so small that there are only a few measurements falling in it, the estimate of clutter's spatial intensity may suffer from a large variance; on the other hand, if the gate is too large, then the uniform distribution assumption of clutter points may no longer hold. Also, this estimation method is biased, since it does not take into account target-originated measurements in the gate. In [67], in order to obtain an unbiased estimator of clutter's spatial intensity, "track perceivability," the probability that the target exists at the current time given all previous measurement [68], was used to handle target-originated measurements in the current measurement set.

In [43, 87], the surveillance region was divided into sectors and clutter points in each sector were assumed to follow Poisson point processes. Based on the Poisson point processes assumption, three clutter spatial intensity estimators were discussed: the first one was based on the number of measurements falling in each sector, the second was based on each sector's nearest neighbor measurement distance, which is equal to the distance from the center of the sector to its nearest measurement point, while

the third was based on the inter-arrival time between two consecutive measurements falling in the same sector. In all three estimators, after obtaining the clutter intensity estimate based on the current measurement set, a time-averaging filter was used to smooth the clutter intensity estimate over time.

In [81, 82], it was assumed that there are several unknown targets, called clutter generators, in a space which is disjoint from both the state space and the measurement space. All clutter points are generated by the clutter generator and an approximated Bayesian estimation method for the density of the clutter generator and the clutter is proposed. However, the proposed method is intractable and no practical implementation method was given in [81, 82].

In [23], based on Poisson point processes, two methods for joint non-homogeneous clutter background estimation and multitarget tracking were presented. In that paper, non-homogeneous Poisson point processes, whose intensity function are assumed to be a mixture of Gaussian functions, were used to model clutter points. Based on this model, a recursive maximum likelihood method and an approximated Bayesian method using Normal-Wishart conjugate prior-posterior pair were proposed to estimate the non-homogeneous clutter spatial intensity. Both clutter estimation methods were integrated into the Probability Hypothesis Density (PHD) filter, which itself also uses the Poisson point process assumption. The mean and the covariance of each Gaussian function were estimated and used to calculate the clutter density in the update equation of the PHD filter. Simulation results showed that both methods were able to improve the performance of the PHD filter in the presence of slowly time varying non-homogeneous clutter background.

3.15.3 Filters

Filtering is the estimation of the state of a dynamic system from noisy data, based on the predefined target dynamic model and the sensor model. In recursive filtering, the received measurements are processed sequentially rather than as a batch so that it is neither necessary to store the complete measurement set nor to reprocess existing measurement if a new measurement becomes available. The Bayesian recursive filter is widely used in the multisensor-multitarget tracking area and such a filter consists of two stages: prediction and update.

The prediction stage uses the system model to predict the state pdf forward from one measurement time to the next. Suppose that the required pdf $p(x(k)|Z^k)$ at measurement time k is available, where $Z^k = [z(1), z(2), \dots, z(k)]$. The prediction stage involves using the system model (15.1) to obtain the prior pdf of the state at measurement time $k + 1$

$$p(x(k+1)|Z^k) = \int p(x(k+1)|x(k))p(x(k)|Z^k)dx(k). \quad (15.31)$$

The update stage uses the latest measurement $z(k+1)$ to update the prior via Bayes' formula

$$p(x(k+1)|Z^{k+1}) = \frac{p(z(k+1)|x(k+1))p(x(k+1)|Z^k)}{p(z(k+1)|Z^k)}. \quad (15.32)$$

The above recursive propagation of the posterior density is only a conceptual solution. Analytical formulas of the posterior density exist only in a restrictive set of cases.

3.15.3.1 Kalman filter

The Kalman filter assumes that the state and measurement models are linear, i.e., $f(k, x(k)) = F(k)x(k)$; $h(k, x(k)) = H(k)x(k)$. Also, in the Kalman filter, the initial state error and all the noises entering into the system are assumed to be Gaussian, i.e., $v(k)$ is white and Gaussian with zero mean and covariance $Q(k)$, and $\omega(k)$ is white and Gaussian with zero mean and covariance $R(k)$. Under the above assumptions, if $p(x(k)|Z^k)$ is Gaussian, it can be proved that $p(x(k+1)|Z^{k+1})$ is also Gaussian, which can be parameterized by a mean and covariance [10].

The Kalman filter algorithm consists of the following recursive relationship [10]:

$$\hat{x}(k+1|k) = F(k)\hat{x}(k|k), \quad (15.33)$$

$$P(k+1|k) = F(k)P(k|k)F(k)' + Q(k), \quad (15.34)$$

$$\hat{z}(k+1|k) = H(k+1)\hat{x}(k+1|k), \quad (15.35)$$

$$S(k+1) = H(k+1)P(k+1|k)H(k+1)' + R(k+1), \quad (15.36)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + W(k+1)(z(k+1) - \hat{z}(k+1|k)), \quad (15.37)$$

$$P(k+1|k+1) = P(k+1|k) - W(k+1)S(k+1)W(k+1)', \quad (15.38)$$

where

$$W(k+1) = P(k+1|k)H(k+1)'S(k+1)^{-1}. \quad (15.39)$$

The Kalman filter is the optimal solution to the tracking problem if the above assumptions hold because it provides the posterior probability density of targets state.

3.15.3.2 Extended Kalman filter (EKF)

While the Kalman filter assumes linearity, most of the real world problems are nonlinear. The extended Kalman filter is a suboptimal state estimation algorithm for nonlinear systems. In EKF, local linearizations of the equations are used to describe the nonlinearity,

$$\widehat{F}(k) = \left. \frac{df(k)}{dx} \right|_{x=\hat{x}(k|k)}, \quad (15.40)$$

$$\widehat{H}(k) = \left. \frac{dh(k+1)}{dx} \right|_{x=\hat{x}(k+1|k)}. \quad (15.41)$$

The EKF assumes that $p(x(k)|Z^k)$ can be approximated by a Gaussian. Then the equations of the Kalman filter can be used with this approximation and the linearized functions, except the state and measurement prediction are performed using the original nonlinear functions

$$\hat{x}(k+1|k) = f(k, \hat{x}(k|k)), \quad (15.42)$$

$$\hat{z}(k+1|k) = h(k+1, \hat{x}(k+1|k)). \quad (15.43)$$

The above is a first-order EKF based on the first-order series expansion of the nonlinearities. There are several error reduction methods to improve the performance of the EKF [10]. One of them is using the second-order series expansion of the nonlinearities, i.e., higher-order EKFs, but the additional

complexity and little or no benefit has prohibited its widespread use. For continuous-time nonlinear systems, numerical integration on the continuous-time stochastic differential equation of the state from k to $k + 1$ can be used to obtain a better predicted state. The third approach to improve the EKF is using an iteration to compute the updated state as a maximum a posterior (MAP) estimate, rather than an approximate conditional mean. This type of EKF is called the Iterated extended Kalman filter (IEKF). The iteration used by the IEKF amounts to relinearize the measurement equation around the updated state rather than relying only on the predicted state. If the measurement model fully observes the state, then the IEKF is able to handle the non-linear measurement model better than the EKF [61].

3.15.3.3 Unscented Kalman filter (UKF)

When the state transition and observation models are highly nonlinear, the EKF may perform poorly. The unscented Kalman filter does not approximate the nonlinear functions of state and measurement models as required by the EKF. Instead, the UKF uses a deterministic sampling technique known as the unscented transform to pick a minimal set of sample points called sigma points around the mean. Here, the propagated mean and covariance are calculated from the transformed samples [52]. In some UKF implementations, the state random variable is augmented as the concatenation of the original state and noise variables [123]. The steps of UKF are described below.

3.15.3.3.1 Sigma point generation

The state vector $\hat{x}(k)$ with mean $\hat{x}(k|k)$ and covariance $P(k|k)$ is approximated by $2n + 1$ weighted sigma points, where n is the dimension of the state vector, as

$$\chi^0(k|k) = \hat{x}(k|k), \quad w_0 = \frac{\kappa}{(n + \kappa)}, \quad (15.44)$$

$$\chi^i(k|k) = \hat{x}(k|k) + \left(\sqrt{(n + \kappa)P(k|k)} \right)_i, \quad w_i = \frac{1}{2(n + \kappa)}, \quad (15.45)$$

$$\chi^{i+n}(k|k) = \hat{x}(k|k) - \left(\sqrt{(n + \kappa)P(k|k)} \right)_i, \quad w_{i+n} = \frac{1}{2(n + \kappa)}, \quad (15.46)$$

where w_i is the weight associated with the i th point, κ is a scaling parameter, $i = 1, 2, \dots, n$, and $(\sqrt{(n + \kappa)P(k|k)})_i$ is the i th row or column of the matrix square root of $(n + \kappa)P(k|k)$.

3.15.3.3.2 Recursion

1. Find the predicted target state $\hat{x}(k + 1|k)$ and corresponding covariance $P(k + 1|k)$:

- a. Transform the Sigma points using the system model

$$\chi^i(k + 1|k) = f(k, \chi^i(k|k)). \quad (15.47)$$

- b. Find the predicted mean

$$\hat{x}(k + 1|k) = \sum_{i=0}^{2n} w_i \chi^i(k + 1|k). \quad (15.48)$$

- c. Find the predicted covariance

$$P(k+1|k) = Q(k) + \sum_{i=0}^{2n} w_i [\chi_i(k+1|k) - \hat{x}(k+1|k)][\chi_i(k+1|k) - \hat{x}(k+1|k)]'. \quad (15.49)$$

2. Find the predicted measurement $\hat{z}(k+1|k)$ and the corresponding covariance $S(k+1)$:

- a. Regenerate the Sigma points $\chi_i(k+1|k)$ using the mean $\hat{x}(k+1|k)$ and covariance $P(k+1|k)$ in order to incorporate the effect of $Q(k)$. If $Q(k)$ is zero, the resulting $\chi_i(k+1|k)$ will be the same as in (15.47). If the process noise is correlated with the state, then the noise vector must be stacked with the state vector $\hat{x}(k|k)$ before generating the sigma points [52].
- b. Find the predicted measurement mean $\hat{z}(k+1|k)$

$$\hat{z}(k+1|k) = \sum_{i=0}^{2n} w_i \varphi^i(k+1|k), \quad (15.50)$$

where

$$\varphi^i(k+1|k) = h(k, \chi^i(k+1|k)). \quad (15.51)$$

- c. Find the innovation covariance $S(k+1)$ and gain $W(k+1)$

$$\begin{aligned} S(k+1) &= R(k+1) + \sum_{i=0}^{2n} w_i [\varphi^i(k+1|k) - \hat{z}(k+1|k)][\varphi^i(k+1|k) - \hat{z}(k+1|k)]', \\ W(k+1) &= \sum_{i=0}^{2n} w_i [\chi^i(k+1|k) - \hat{x}(k+1|k)][\varphi^i(k+1|k) - \hat{z}(k+1|k)]' S(k+1)^{-1}. \end{aligned} \quad (15.52)$$

3. Update the state $\hat{x}(k+1|k+1)$ and corresponding covariance $P(k+1|k+1)$ using (15.37) and (15.38), respectively.

3.15.3.4 Particle filter

If the true density is substantially non-Gaussian, then a Gaussian model as in the case of the Kalman filter will not yield accurate estimates. In such cases, particle filters will yield an improvement in performance in comparison to the EKF or UKF. The particle filter provides a mechanism for representing the density, $p(x(k)|Z^k)$ of the state vector $x(k)$ at time epoch k as a set of random samples $\{x^p(k) : p = 1, \dots, m\}$, with associated weights $\{w^p(k) : p = 1, \dots, m\}$. That is, the particle filter attempts to represent an arbitrary density function using a finite number of points, instead of a pair of mean vector and covariance matrix that is sufficient for Gaussian distributions. Several variations of particle filters are available and the reader is referred to [3] for detailed description. The Sampling Importance Resampling (SIR) type particle filter, which is arguably the most common technique to implement particle filters, is discussed below. In general, the particles are sampled either from the prior density or likelihood function. Taking

the prior as the importance density, the method of SIR is used to produce a set of equally weighted particles that approximates $p(x(k)|Z^k)$, i.e.,

$$p(x(k)|Z^k) \approx \frac{1}{m} \sum_{p=1}^m \delta(x(k) - x^p(k)), \quad (15.53)$$

where $\delta(\cdot)$ is the Dirac Delta function. The prediction and update steps of the particle filter recursion are given below.

Prediction: Take each existing sample, $x^p(k)$ and generate a sample $x^{*p}(k+1) \sim p(x(k+1)|x^p(k))$, using the system model. The set $\{x^{*p}(k+1) : p = 1, \dots, m\}$ provides an approximation of the prior, $p(x(k+1)|Z^k)$, at time $k+1$.

Update: At each measurement epoch, to account for the fact that the samples, $x^{*p}(k+1)$ are not drawn from $p(x(k+1)|Z^{k+1})$, the weights are modified using the principle of importance sampling. When using the prior as the importance density, it can be shown that the weights are given by

$$w^p(k+1) \propto p(z(k+1)|x(k+1) = x^{*p}(k+1), Z^k). \quad (15.54)$$

A common problem with the above recursion is the degeneracy phenomenon, whereby the particle set quickly collapses to just a single particle. To overcome this problem a regularization can be imposed via reselection as follows.

Reselection: Resample (with replacement) from $\{x^{*p}(k+1) : p = 1, \dots, m\}$, using the weights, $\{w^{*p}(k+1) : p = 1, \dots, m\}$, to generate a new sample, $\{x^p(k+1) : p = 1, \dots, m\}$, then set $w^p(k+1) = 1/m$ for $p = 1, \dots, m$.

The mean of the posterior distribution is used to estimate, $x(k+1|k+1)$ of the target state, $x(k+1)$, i.e.,

$$\hat{x}_{k+1} \approx \frac{1}{m} \sum_{p=1}^m x^p(k+1). \quad (15.55)$$

The accuracy of the particle filter based estimate (15.53) depends on the number of particles employed. A more accurate state estimates can be obtained at the expense of extra computation. The extension of particle filters allows them to be applicable to multitarget tracking problems [60].

3.15.3.5 Interacting multiple-model estimator

Tracking maneuvering targets is a very important task for almost all practical systems. Several schemes and methods have been proposed to track maneuvering targets [69–73]. One of widely used method is the Multiple Mode (MM) approach [10]. In MM approach, the target system is assumed to follow one of predetermined models (or mods) and a number of filters operate in parallel. Thus, the system has both continuous uncertainties brought by the noise, and discrete uncertainties brought by the model uncertainty. Starting from the prior distribution of the target state and the prior probability that the system is in a particular mode, the goal of the MM approach is to obtain the posterior distribution of the target state and posterior mode probability.

Depending on whether mode jumping is allowed, there are static MM estimator and dynamic MM estimator. In static MM algorithms, it is assumed that there is no model switching from one mode to another during the whole estimation process and this assumption is not realistic for many real scenarios.

On the other hand, in dynamic MM algorithms, the target is allowed to switch from one mode to another according to a Markov chain. Optimal dynamic MM estimator is infeasible since it needs to carry all the mode sequence hypotheses and the total number of the mode sequence hypotheses increases exponentially with time. To get a feasible dynamic MM suboptimal estimator, the mode sequences that only differ in “older” modes are combined and the mode sequences are kept to depth n in the Generalized Pseudo-Bayesian (GPB) approaches [9,10]. A GPB algorithm of depth n (GPB_n) requires N_r^n filters in its bank, where N_r is the number of models.

The Interacting Multiple Model (IMM) estimator [9,10,15,16], which mixes the mode sequence hypotheses with depth 1 at the *beginning* of each filtering cycle, requires only N_r number of filters to operate in parallel, (i.e., same as GBP_1 estimator) but is able to perform nearly as well as GPB_2 . Thus, the IMM estimator is very cost-efficient. The main contributor for the cost-efficiency of the IMM estimator is the “mixing/interaction” between its “mode-matched” base state filtering modules at the beginning of each cycle. It has been shown in [7], the same feature is exactly what the IMM has in common with the *optimal* estimator for dynamic MM systems. Besides its cost-efficiency, another advantage of the IMM estimator is that it does not require maneuver detection decision as in the case of Variable State Dimension (VSD) filter [10] algorithms and undergoes a soft switching between models based on the updated mode probabilities.

3.15.3.5.1 Modeling assumptions in dynamic multiple model approach

Base state model:

$$x(k) = f(k, M(k), x(k-1)) + G(k, M(k))v(k), \quad (15.56)$$

$$z(k) = h(k, M(k), x(k)) + w(k, M(k)), \quad (15.57)$$

where $M(k)$ denotes the mode in effect during the sampling period ending at k .

Mode (“model state”)—among the possible r modes:

$$M(k) \in \{M_j\}_{j=1}^r. \quad (15.58)$$

The structure of the system and/or the statistics of the noises can differ from mode to mode:

$$f(M_j) = f_j, \quad (15.59)$$

$$G(k, M(k))v(k) \sim \mathcal{N}(u_j, Q_j). \quad (15.60)$$

Mode jump process: Markov chain with known transition probabilities

$$P\{M(k) = M_j | M(k-1) = M_i\} = p_{ij}. \quad (15.61)$$

3.15.3.5.2 The IMM estimation algorithm

- *Interaction*: Mixing of the previous cycle mode-conditioned state estimates and covariance, using the mixing probabilities $\mu_{i|j}$, to initialize the current cycle of each mode-conditioned filter. For filter j , there is:

$$\hat{x}^{0j}(k-1|k-1) = \sum_{i=1}^r \hat{x}^i(k-1|k-1) \mu_{i|j}(k-1|k-1), \quad (15.62)$$

$$\begin{aligned} P^{0j}(k-1|k-1) &= \sum_{i=1}^r \mu_{i|j}(k-1|k-1) \{P^i(k-1|k-1) \\ &\quad + [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)] \\ &\quad \cdot [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)]'\}, \end{aligned} \quad (15.63)$$

where

$$\mu_{i|j} = \frac{1}{\bar{c}_j} p_{ij} \mu_i(k-1), \quad (15.64)$$

$$\bar{c}_j = \sum_{i=1}^r p_{ij} \mu_i(k-1), \quad (15.65)$$

- *Mode-conditioned filtering:* Calculation of the state estimates and covariances conditioned on a mode being in effect. The KF matched to $M_j(k)$ (filter j) uses $z(k)$ to yield $\hat{x}^j(k|k)$ and $P^j(k|k)$. Likelihood function corresponding to filter j is $\Lambda_j(k) = \mathcal{N}[z(k); \hat{z}^j(k|k-1), S_j(k)]$.
- *Probability evaluation:* Computation of the mixing and the updated mode probabilities. For the mode probabilities of the j th mode ($j = 1, \dots, r$)

$$\mu_j(k) = \frac{1}{c} \Lambda_j(k) \bar{c}_j, \quad (15.66)$$

$$c \triangleq \sum_{j=1}^r \Lambda_j(k) \bar{c}_j. \quad (15.67)$$

- *Overall state estimate and covariance* (for output only): Combination of the latest mode-conditioned state estimates and covariances

$$\hat{x}(k|k) = \sum_{j=1}^r \hat{x}^j(k|k) \mu_j(k), \quad (15.68)$$

$$P(k|k) = \sum_{j=1}^r \mu_j(k) \{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]'\}. \quad (15.69)$$

The IMM estimation algorithm has a *modular structure*.

For non-maneuvering targets, unnecessarily using multiple-model algorithms is not optimal because it might diminish the performance level of the tracker and increase the computational load. A multiple-model approach is only necessary for targets with high maneuverability. Typically, the decision to use a multiple-model estimator should be based on the maneuvering index [10], which quantifies the maneuverability of the target in terms of the process noise, sensor measurement noise and sensor revisit interval. In [56], a study was presented to compare the IMM estimator with the Kalman filter based on the maneuvering index.

3.15.4 Filter initialization

Track initiation is an essential component of all tracking algorithms. Two major types of initialization techniques, namely, the Single-Point (SP) method and the Two-Point-Difference (TPD) method, are commonly used in multisensor-multitarget tracking applications [9, 10, 84].

3.15.4.1 Single-point track initialization

In SP track initialization, every detection (measurement), which is unassociated to any track, is an “initiator.” The unassociated measurement is first converted into Cartesian space through the unbiased conversion method [10] and then a tentative track is declared. In SP algorithm, the position component of the tentative track is initialized using the position component of the converted measurement and the velocity components is set to zero. To compensate the zero velocity assumption made for the tentative track, when initializing the covariance matrix of the tentative track, the maximum possible speed of the target is often used as the standard deviation of the velocity component of the tentative track. For the position component, the variance of the converted measurement is used. Furthermore, for the newly declared tentative track, its velocity and position along different coordinates are usually assumed to be independent. A Kalman filter is then used for subsequent processing of the tentative track.

3.15.4.2 Two-point difference track initialization

In TPD track initialization, every unassociated detection (measurement) is also an “initiator,” but does not immediately yield a tentative track. At the sampling time (scan or frame) following the detection of an initiator, a gate is set up around the initiator based on the assumed maximum and minimum target speed as well as the measurement noise intensities. Thus, it is reasonable to assume that if the initiator is from a target, then the measurement from that target in the second scan (if detected) will fall inside the gate with nearly unity probability. A tentative track will be declared only if there is at least one detection falling in the gate. Since now each tentative track has two measurements, a straight line extrapolation is used to obtain its speed and the covariance matrix. A Kalman filter is then used for the subsequent processing.

3.15.4.3 Issues related to track initialization

It has been demonstrated numerically that the SP method has a smaller mean square error matrix than the TPD method for a 3D radar target tracking problem. Also, it has been analytically shown that, if the process noise approaches zero and the maximum speed of a target used to initialize the velocity variance approaches infinity, then the SP algorithm reduces to the TPD algorithm [84].

In many multi-sensor multitarget tracking algorithms, the track initialization and the track maintenance phases are treated as two independent (and consecutive) stages. However, in a real tracking problems, targets can enter and leave the surveillance region at any time. As a result, track initiation has to be considered at every sampling time. That is, track initialization occurs even after the first few scan. Similarly, the fact that track maintenance stage has been activated does not obviate the need for further track initiations. Both have to be carried out simultaneously throughout the entire tracking interval. Because of this, the track initialization module needs to take into account the

number, states and qualities of the established tracks being retained by the track maintenance module. Otherwise, spurious tracks and track seduction will ensue, damaging the overall quality of the tracker [42].

3.15.5 Data association

In Section 3.15.3, it has been implicitly assumed that there is no measurement origin ambiguity. However, the crux of the multitarget problem is to carry out the association process for measurements whose origins are uncertain due to

- random false alarms in the detection process,
- clutter due to spurious reflectors or radiators near the target of interest,
- interfering targets,
- decoys and countermeasures.

Furthermore, the probability of obtaining a measurement from a target—the target detection probability—is usually less than unity.

Data association problems may be categorized according to the pairs of information that are associated together:

- measurement-to-track association—track maintenance or updating,
- measurement-to-measurement association—parallel updating for centralized tracking,
- track-to-track association—track fusion (for distributed or decentralized tracking).

In this subsection, measurement-to-track association and measurement-to-measurement association techniques are discussed. The track-to-track association is discussed in Section 3.15.9.

3.15.5.0.1 Measurement-to-track association

A multidimensional gate is set up in the measurement space around the predicted measurement in order to avoid searching for the measurement from the target of interest in the entire measurement space. A measurement in the gate, while not guaranteed to have originated from the target the gate pertains to, is a valid association candidate. Thus the name validation region or association region. If there is more than one measurement in the gate, this leads to an association uncertainty.

Figure 15.1a and b illustrates the gating for two well-separated and closely-spaced targets respectively. In the figures, “●” indicates the expected measurement and the “*” indicates the received measurement.

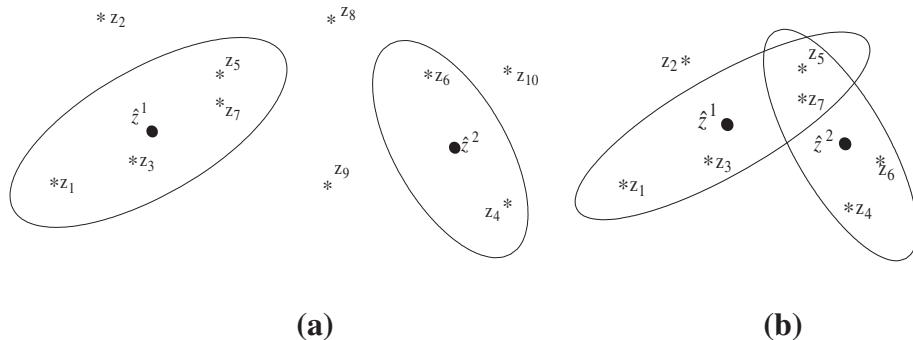
If the true measurement conditioned on the past is normally (Gaussian) distributed with its probability density function given by

$$p(z(k+1)|Z^k) = \mathcal{N}[z(k+1); \hat{z}(k+1|k), S(k+1)], \quad (15.70)$$

then the true measurement will be in the following region:

$$\mathcal{V}(k+1, \gamma) = \{z : [z - \hat{z}(k+1|k)]' S(k+1)^{-1} [z - \hat{z}(k+1|k)] < \gamma\} \quad (15.71)$$

with probability determined by the gate threshold γ . The region defined by (15.71) is called gate or validation region.

**FIGURE 15.1**

Validation regions. (a) Well-separated targets. (b) Closely-spaced targets.

Some well-known approaches for data association in the presence of well-separated targets, where no measurement origin uncertainties exist are discussed below.

Nearest Neighbor (NN): This is the simplest possible approach and uses the measurement nearest to the predicted measurement, assuming that the nearest one is the correct one. The nearest measurement to the predicted measurement is determined according to the distance measure (norm of the innovation squared),

$$D(z) = [z - \hat{z}(k+1|k)]' S(k+1)^{-1} [z - \hat{z}(k+1|k)]. \quad (15.72)$$

Strongest Neighbor (SN): Select the strongest measurement (in terms of signal intensity) among the validated ones—this assumes that signal intensity information is available.

2-D Assignment: This technique is also known as the Global Nearest Neighbor (GNN) method. The fundamental idea behind 2-D assignment is that the measurements from the scan list $\mathcal{M}(k)$ at time k are deemed to have come from the tracks in list $\mathcal{T}(k-1)$ at time $k-1$. To find the best match between $\mathcal{M}(k)$ and $\mathcal{T}(k-1)$, a constrained global optimization problem has to be solved. The optimization is carried out to minimize the “cost” of associating the measurements in $\mathcal{M}(k)$ to tracks predicted from $\mathcal{T}(k-1)$.

To present the 2-D assignment, define a binary assignment variable $a(k, m, n)$ such that

$$a(k, m, n) = \begin{cases} 1 & \text{measurement } \mathbf{z}_m(t_{m_k}) \text{ is assigned to track } \mathcal{T}^n(k-1), \\ 0 & \text{otherwise,} \end{cases} \quad (15.73)$$

where t_{m_k} is the time stamp of the m th measurement from scan or frame k .

A set of complete assignments, which consists of the associations of all the measurements in $\mathcal{M}(k)$ and the tracks in $\mathcal{T}(k-1)$, is denoted by $\mathbf{a}(k)$, i.e.,

$$\mathbf{a}(k) = \{a(k, m, n); m = 0, 1, \dots, M(k); n = 0, 1, \dots, N(k-1)\}, \quad (15.74)$$

where $M(k)$ and $N(k-1)$ are the cardinalities of the measurement and track sets, respectively. The indices $m = 0$ and $n = 0$ correspond to the non-existent (or “dummy”) measurement and track.

The “dummy” notation is used to formulate the assignment problem in a uniform manner, where the non-association possibilities are also considered, making it computer-solvable.

The objective of the assignment is to find the optimal assignment $\mathbf{a}^*(k)$, which minimizes the global cost of association

$$C(k|\mathbf{a}(k)) = \sum_{m=0}^{M(k)} \sum_{n=0}^{N(k-1)} a(k, m, n) c(k, m, n), \quad (15.75)$$

where $c(k, m, n)$ is the cost of the assignment $a(k, m, n)$. That is,

$$\mathbf{a}^*(k) = \arg \min_{\mathbf{a}(k)} C(k|\mathbf{a}(k)). \quad (15.76)$$

The costs $c(k, m, n)$ are the negative of the logarithm of the dimensionless likelihood ratio of the measurement-to-track associations, namely,

$$c(k, m, n) = -\ln \Lambda(k, m, n), \quad (15.77)$$

where

$$\Lambda(k, m, n) = \begin{cases} P_D p[v_m^n(t_{mk})] / \lambda_e & m > 0, n > 0, \\ 1 & m > 0, n = 0, \\ (1 - P_D) & m = 0, n > 0 \end{cases} \quad (15.78)$$

are the following likelihood ratios:

1. that measurement $m > 0$ came from track n for $n > 0$ with the association likelihood function being the probability density function of the corresponding innovation, $p[v_m^n(k)]$ versus from an extraneous source whose spatial density is λ_e ,
2. that measurement $m > 0$ came from none of the tracks (i.e., from the dummy track $n = 0$) versus from an extraneous source. Note that, if measurement $m > 0$ came from none of the tracks, then that measurement must be generated by an extraneous source. Thus, the likelihood ratio must be unity in this case,
3. that the measurement from track n is not in $\mathcal{M}(k)$, i.e., track n is associated with the dummy measurement—the cost of not associating any measurement to a track amounts to the miss probability $1 - P_D$, where the nominal target detection probability is denoted by P_D .

The 2-D assignment is subject to the following constraints.

Validation: A measurement is assigned only to one of the tracks that validated it.

One-to-one constraint: Each track is assigned at most one measurement. The only exception is the dummy track ($n = 0$), which can be associated with any number of measurements. Similarly, a measurement is assigned to at most one track. The dummy measurement ($m = 0$) can be assigned to multiple tracks.

Non-empty association: The association cannot be empty, i.e., the dummy measurement cannot be assigned to the dummy track.

The modified auction algorithm can solve the above constrained optimization problem and that algorithm runs in quasi-polynomial time [94, 95].

Multidimensional (S-D) Assignments: In 2-D assignment only the latest scan is used and information about target evolution through multiple scans is lost. Also it is not possible to change an association later

in light of subsequent measurements. A data-association algorithm may perform better when a number of past scans are utilized. This corresponds to multidimensional assignment for data association. In S -D assignment the latest $S - 1$ scans of measurements are associated with the established track list (from time $k - S + 1$, where k is the current time, i.e., with a sliding window of time depth $S - 1$) in order to update the tracks.

Similarly to the 2-D assignment, define a binary assignment variable $a(k, \{m_s\}_{s=k-S+2}^k, n)$ such that

$$a(k, \{m_s\}_{s=k-S+2}^k, n) = \begin{cases} 1 & \text{measurements } \mathbf{z}_{m_{k-S+2}}(t_{m_{k-S+2}}), \dots, \mathbf{z}_{m_k}(t_{m_k}) \text{ are assigned} \\ & \text{to track } \mathcal{T}^n(k - S + 1), \\ 0 & \text{otherwise,} \end{cases} \quad (15.79)$$

which is the general version of (15.73). The cost associated with (15.79) is denoted as

$$c(k, \{m_s\}_{s=k-S+2}^k, n) = -\ln \Lambda(k, \{m_s\}_{s=k-S+2}^k, n) \quad (15.80)$$

and $\Lambda(k, \{m_s\}_{s=k-S+2}^k, n)$ is the likelihood ratio that the S -tuple of measurements, which is given by $\mathbf{z}_{m_{k-S+2}}(t_{m_{k-S+2}}), \dots, \mathbf{z}_{m_k}(t_{m_k})$, originated from the target represented by track $\mathcal{T}^n(k - S + 1)$ versus some measurements $\mathbf{z}_{m_{k-S+2}}(t_{m_{k-S+2}}), \dots, \mathbf{z}_{m_k}(t_{m_k})$ being extraneous.

The objective of the S -D assignment is to find the S -tuples of measurement-to-track associations $a(k, \{m_s\}_{s=1}^{S-1}, n)$, which minimize the global cost of association given by

$$C(k|\mathbf{a}) = \sum_{n=0}^{N(k-S+1)} \sum_{m_{k-S+2}=0}^{M(k-S+2)} \sum_{m_{k-S+3}=0}^{M(k-S+3)} \cdots \sum_{m_k=0}^{M(k)} a\left(k, \{m_s\}_{s=k-S+2}^k, n\right) c(k, \{m_s\}_{s=k-S+2}^k, n), \quad (15.81)$$

where $M(k)$ is the number of measurements in scan k and \mathbf{a} is the complete set of associations analogous to that defined in (15.74) for the 2-D assignment. The association likelihoods are given by

$$\Lambda\left(k, \{m_s\}_{s=k-S+2}^k, n\right) = \begin{cases} \prod_{s=k-S+2}^k [1 - P_D]^{1-u(m_s)} [P_D p[v_{m_s}^n(s)] \lambda_e]^{u(m_s)} & n > 0, \\ 1 & n = 0, \end{cases} \quad (15.82)$$

where $u(m)$ is a binary function such that

$$u(m) = \begin{cases} 1 & m > 0, \\ 0 & m = 0, \end{cases} \quad (15.83)$$

and $p[v_{m_s}^n(s)]$ is the filter-calculated innovation pdf if the (kinematic) measurement $\mathbf{z}_{m_s}(t_{m_s})$ is associated with track $\mathcal{T}^n(k - S + 1)$ continued with the (kinematic) measurements $\mathbf{z}_{m_{k-S+2}}(t_{m_{k-S+2}}), \dots, \mathbf{z}_{m_k}(t_{m_k})$.

The association costs are given to the generalized S -D assignment algorithm, which uses Lagrangian relaxation, as described in [32,95] to solve the assignment problem in quasi-polynomial time. The feasibility constraints are similar to those from the 2-D assignment.

3.15.5.0.2 Measurement-to-measurement association

Measurement-to-measurement association is the most important step in the parallel updating scheme for the centralized tracking. Without any major modification, multidimensional (S-D) Assignments and 2-D Assignment technique are commonly used for measurement-to-measurement association. A good example of using S-D assignment technique to solve the measurement-to-measurement association problem in a multisensor-multitarget tracking problem is given in [9].

3.15.6 Multitarget tracking algorithms

Three widely used multitarget tracking algorithms, namely, the Probabilistic Data Association (PDA) and Joint Probabilistic Data Association (JPDA) algorithm, the Multiple Hypothesis Tracker (MHT) algorithm, and the Probability Hypothesis Density (PHD) algorithm, are discussed in this section.

3.15.6.1 Probabilistic data association (PDA) and joint probabilistic data association (JPDA)

PDA algorithm is a Bayesian approach that probabilistically associates all the validated measurements to the target of interest [9]. The state update equation of the PDA filter is

$$\hat{x}(k|k) = \hat{x}(k|k-1) + W(k)v(k), \quad (15.84)$$

where

$$v(k) = \sum_{i=1}^{m(k)} \beta_i(k) v_i(k), \quad (15.85)$$

$$v_i(k) = (z_i(k) - \hat{z}(k|k-1)), \quad (15.86)$$

$m(k)$ is the number of validated measurements and

$$\beta_i(k) \triangleq \{\theta_i(k)|Z^k\} \quad (15.87)$$

is the conditional probability of the event that the i th validated measurement is correct.

The covariance associated with the updated state is

$$P(k|k) = \beta_0(k)P(k|k-1) + [1 - \beta_0(k)]P^c(k|k) + \tilde{P}(k), \quad (15.88)$$

where $\beta_0(k)$ is the conditional probability of the event that none of the measurements is correct and the covariance of the state updated with the correct measurement is

$$P^c(k|k) = P(k|k-1) - W(k)S(k)W(k)' \quad (15.89)$$

In equation (15.89), the gain matrix $W(k)$ and the innovation matrix $S(k)$ are calculated using the standard Kalman filter equation (15.36), (15.39). The spread of the innovations term in (15.88) is

$$\tilde{P}(k) \triangleq W(k) \left[\sum_{i=1}^{m(k)} \beta_i(k) v_i(k) v_i(k)' - v(k) v(k)' \right] W(k)'. \quad (15.90)$$

The association of measurements in a multitarget environment with closely-spaced targets must be done while simultaneously considering all the targets. Thus, the Joint Probabilistic Data Association (JPDA) is proposed as an extension of the PDA method to handle the scenario with closely-spaced targets [9]. For a known number of targets, JPDA evaluates the measurement-to-target association events probabilities (for the latest set of measurements) and combines them into the corresponding state estimates.

The JPDA algorithm includes the following steps:

- A validation matrix that indicates all the possible sources of each measurement is set up.
- From the validation matrix, all the feasible joint association events are constructed by the JPDA tracker according to the following two rules:
 - One measurement must be originated from one target or a false alarm.
 - One target can only generate one measurement at most.
- The probabilities of these joint events are evaluated according to the following assumptions:
 - Target-originated measurements are Gaussian distributed around the predicted location of the corresponding target's measurement.
 - False alarms are distributed in the surveillance region according to a Poisson point process model.
- Marginal (individual measurement-to-target) association probabilities are obtained from the joint association probabilities.
- The target states are estimated by separate (uncoupled) PDA filters using these marginal probabilities.

An extension of the PDA algorithm, the Integrated Probabilistic Data Association (IPDA) algorithm is proposed in [85]. The main idea behind IPDA tracker is to introduce the track existence probability, a measure of the quality of the track, to the PDA algorithm. Like PDA tracker, IPDA track can only be used to track well-separated targets in clutter. To handle the multitarget environment with closely-spaced targets, the Joint Integrated Probabilistic Data Association (JIPDA) tracker is proposed in [86], which is actually a JPDA with the track existence probability.

3.15.6.2 Multiple Hypothesis Tracker (MHT)

Multiple Hypothesis Tracking (MHT) is an efficient algorithm for tracking multitargets in a cluttered environment. The algorithm is capable of initiating tracks, accounting false or missing reports, and processing sets of dependent reports. As each measurement is received, probabilities are calculated for the hypotheses that the measurement came from previously existing target, or from a new target, or that the measurement is a false alarm. Target states are estimated from each such data-association hypothesis with a certain filter (e.g., Kalman Filter). As more measurements are received, the probabilities of joint hypotheses are calculated recursively using all available information such as the density of unknown target λ_N , the density of false targets λ_{FA} and the probability of detection P_D . This branching technique allows the correlation of a measurement with its source based on subsequent, as well as previous, data. To keep the number of hypotheses reasonable, unlikely hypotheses are eliminated and hypotheses with similar target estimates are combined. To minimize computational requirements, clustering and pruning techniques are embedded in the MHT tracker. Note that MHT strictly follows the one-to-one assumption

between measurements and targets. There are two types of implementation for MHT, the hypothesis-oriented MHT (HOMHT), which is also known as the measurement-oriented MHT (MOMHT), and the track-oriented MHT (TOMHT).

3.15.6.2.1 Hypothesis-oriented MHT (HOMHT)

In HOMHT, hypotheses are composed of sets of compatible tracks. Multiple tracks are compatible if they have no measurement in common. At every scan, each hypothesis carried over from the previous scan (i.e., parent hypothesis) is expanded into a set of new hypotheses (i.e., offspring hypothesis) by considering all possible track-to-measurement associations for the tracks within the parent hypothesis [9]. The HOMHT includes the following steps:

- *Initialization:* The measurements received at the first scan $\{z_1(1), z_2(1), \dots, z_{m(1)}(1)\}$ have two possible origination: (1) new target with probability $\frac{\lambda_N}{\lambda_N + \lambda_{FA}}$; (2) false alarm with probability $\frac{\lambda_{FA}}{\lambda_N + \lambda_{FA}}$. Therefore, the total number of possible hypotheses is $2^{m(1)}$ where $m(k)$ is the number of measurements received at scan k . The probability of each hypothesis is proportional to

$$P(\mathcal{H}_j(1)|Z(1)) = \lambda_N^n \cdot \lambda_{FA}^f, \quad (15.91)$$

where n and f denote the number of measurements that are assigned as new targets and the number of measurements that are assigned as false alarms give the j th valid hypothesis $\mathcal{H}_j(1)$. Then, select the top K hypotheses based on their probabilities $P(\mathcal{H}_j(1)|Z(1))$. These K hypotheses $\{\mathcal{H}_j(1)\}_{j=1}^K$ are parent hypotheses for the next scan, and their probabilities are normalized at the end of this scan.

- *Update of hypotheses:* In this step, every parent hypothesis will be used to generate a set of offspring hypotheses. The collection of all the offspring hypotheses forms the current set of hypotheses. Then, top K hypotheses will be selected from the current set of hypotheses, and their probabilities will be normalized.
- *Prune hypotheses:* It is infeasible to keep all the hypotheses because the number of hypotheses grows exponentially. Several pruning techniques are embedded to the HOMHT tracker to maintain the number of hypotheses in a suitable level. The first pruning method removes hypothesis whose probability is less than a predefined threshold β . The second pruning method keeps only the top K hypotheses with the greatest probabilities. The third method discards the set of hypotheses with the smallest probabilities such that the total probability of discarded hypotheses does not exceed a user defined threshold γ . For example, the probabilities of all the current hypotheses are $[p_1, p_2, \dots, p_n]$. Assume that $p_1 \geq p_2 \geq \dots \geq p_n$ and $\sum_{i=1}^n p_i = 1$. Then, the third method removes the $(j+1)$ th to n th hypotheses where $\sum_{i=1}^j p_i \geq (1 - \gamma)$ and $\sum_{i=1}^{j-1} p_i < (1 - \gamma)$.
- *Track management:* In HOMHT tracker, two kinds of track management systems are available. The first is based on m/n logic and the second relies on track qualities. Note that, the HOMHT tracker does not probabilistically combine the update states, which are calculated conditioning on the measurement-to-track association events. Thus the equations that used in JIPDA tracker can not be applied here. One method to compute the track quality in the MHT framework is proposed in [88].
- *Find the best hypothesis and output:* The last step of a HOMHT loop for a scan is to find the best current hypothesis and then output it to the user.

3.15.6.2.2 Track oriented MHT (TOMHT)

The track-oriented MHT constructs a target tree for each potential (or postulated) target according to the measurements, and the branches represent the different measurements with which the target may be associated. A trace of successive branches from the root to a leaf of the tree represents a potential measurement history generated by the target. Conventionally, the target trees are referred to as track hypotheses, and a collection of compatible tracks is referred to as global hypothesis. Unlike the HOMHT in which three possible originalities of measurement are considered, usually the TOMHT treats a measurement as either originated from an existing target or a new one [13]. The TOMHT includes the following steps:

- *Initialize tree:* The target trees are initialized on the receipt of the first set of measurements. The root of each tree is a measurement.
- *Build tree:* In this step, the set of trees of previous scan are updated. More specifically, (1) the depth of each previous tree is increased by one and each branch grows to several new branches to account for all possible target-to-measurement associations; (2) each measurement also is used to initiate a new target tree where the measurement is the root of the tree. In addition, the compatibility relation has to be updated in order to find the global hypothesis.
- *Track management:* The track management techniques used for TOMHT are the same as that used for HOMHT.
- *Track level pruning:* The purpose of this step is to remove the branches of low probabilities so that the computation load for finding the global hypothesis is remained in a reasonable level. Two pruning methods can be embedded into the HOMHT: (1) limit the number of branches per tree below B_{\max} , which is defined by user as the maximum number of branches per tree; (2) discard the branches whose probability is lower than a predefined threshold p_{\min} .
- *Find clusters:* All the branches, i.e., the node or potential track, belong to the same target tree are within the same cluster because they share the same root. Therefore, the clustering procedure for TOMHT is done in the tree-to-measurement level. The algorithm and code for HOMHT clustering can be reused for TOMHT by just replacing the track by tree and measurement associated with track by measurements associated with target tree.
- *Find global hypothesis:* This step is used to find the best global hypothesis, which is a collection of compatible trees from all the target trees. Enumeration is the basic method for finding the best global hypothesis. In enumeration method, all valid global hypotheses are constructed and then the best one is chosen according to the costs of hypotheses. However, the total number of valid global hypotheses grows exponentially with respect to the number of trees. For example, assume that 10 trees exist and there are 20 branches for each tree. Then, the total number of global hypotheses is 20^{10} . Although some of the invalid hypotheses will be removed from this 20^{10} hypotheses, the remaining part is still a large amount. Therefore, the pruning method must be used to limit the number of valid hypotheses and then an approximate best global hypothesis is found.

Due to the requirement of node compatibility (or track compatibility), one and only one branch must be selected per tree to form a possible valid global hypothesis. Note that the selection of dummy node means no track is selected from this tree in the global hypothesis.

3.15.6.3 Probability hypothesis density (PHD) method

In tracking multiple targets, if the number of targets is unknown and varying with time, it is not possible to compare states with different dimensions using ordinary Bayesian statistics. However, the problem can be addressed using Finite Set Statistics (FISST) [75] to incorporate comparisons of states of different dimensions. FISST facilitates the construction of “multitarget densities” from multiple target transition functions into computing set derivatives of belief-mass functions [75], which makes it possible to combine states of different dimensions. The main practical difficulty with this approach is that the dimension of state space becomes large when many targets are present, which increases the computational load exponentially with the number of targets. Since the PHD is defined over the state space of one target in contrast to the full posterior distribution, which is defined over the state space of all the targets, the computation cost of propagating the PHD over time is much lower than propagating the full posterior density. A comparison in terms of computation and estimation accuracy of multitarget filtering using FISST particle filter and PHD particle filter is given in [112].

Assume a sensor has monitored an area since time 1 and at time k , the measurement set $Z_k = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m\}$ is provided by that sensor. From time 1 to time k , the union of all available measurement sets from that sensor are $Z_{1:k} = [Z_1, Z_2, \dots, Z_k]$. By definition, the PHD $D_{k|k}(\mathbf{x}_k | Z_{1:k})$, with argument single-target state vector \mathbf{x}_k and all available measurement sets up to time step k , is the density whose integral on any region S of the state space is the expected number of targets $N_{k|k}$ contained in S . That is,

$$N_{k|k} = \int_S D_{k|k}(\mathbf{x}_k | Z_{1:k}) d\mathbf{x}_k. \quad (15.92)$$

Since this property uniquely characterizes the PHD and since the first order statistical moment of the full target posterior distribution possesses this property, the first order statistical moment of the full target posterior is indeed the PHD. The first moment of the full target posterior or the PHD, given all the measurement sets $Z_{1:k}$ up to time step k , is given by [76]

$$D_{k|k}(\mathbf{x}_k | Z_{1:k}) = \int_{X_k \ni \mathbf{x}_k} f_{k|k}(X_k | Z_{1:k}) \delta X_k, \quad (15.93)$$

where X_k is the multitarget state. The approximate expected target states are given by the local maxima of the PHD. The following section demonstrates the prediction and update steps of one cycle of the PHD filter.

Prediction: In a general scenario of interest, there are target disappearances, target spawning and entry of new targets. We denote the probability that a target with state \mathbf{x}_{k-1} at time step $(k-1)$ will survive at time step k by $e_{k|k-1}(\mathbf{x}_{k-1})$, the PHD of spawned targets at time step k from a target with state \mathbf{x}_{k-1} by $b_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1})$ and the PHD of newborn spontaneous targets at time step k by $\gamma_k(\mathbf{x}_k)$. Then, the predicted PHD is given by

$$D_{k|k-1}(\mathbf{x}_k | Z_{1:k-1}) = \gamma_k(\mathbf{x}_k) + \int [e_{k|k-1}(\mathbf{x}_{k-1}) f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}) + b_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1})] D_{k-1|k-1}(\mathbf{x}_{k-1} | Z_{1:k-1}) d\mathbf{x}_{k-1}, \quad (15.94)$$

where $f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1})$ denotes the single-target Markov transition density. The prediction Eq. (15.94) is lossless since there are no approximations.

Update: The predicted PHD can be updated with measurement set Z_k at time step k to get the posterior PHD. Assume that the number of false alarms is Poisson-distributed with the average number λ_{FA} and that the probability density of the spatial distribution of false alarms is $c_{\text{FA}}(\mathbf{z}_k)$. Let the detection probability of a target with state \mathbf{x}_k at time step k be $p_D(\mathbf{x}_k)$. Then, the updated PHD at time step k is given by

$$D_{k|k}(\mathbf{x}_k|Z_{1:k}) \cong \left[\sum_{\mathbf{z}_k \in Z_k} \frac{p_D(\mathbf{x}_k) f_{k|k}(\mathbf{z}_k|\mathbf{x}_k)}{\lambda_{\text{FA}} c_{\text{FA}}(\mathbf{z}_k) + \psi_k(\mathbf{z}_k|Z_{1:k-1})} + (1 - p_D(\mathbf{x}_k)) \right] D_{k|k-1}(\mathbf{x}_k|Z_{1:k-1}), \quad (15.95)$$

where the likelihood function $\psi(\cdot)$ is given by

$$\psi_k(\mathbf{z}_k|Z_{1:k-1}) = \int p_D(\mathbf{x}_k) f_{k|k}(\mathbf{z}_k|\mathbf{x}_k) D_{k|k-1}(\mathbf{x}_k|Z_{1:k-1}) d\mathbf{x}_k \quad (15.96)$$

and $f_{k|k}(\mathbf{z}_k|\mathbf{x}_k)$ denotes the single-sensor/single-target likelihood. The update Eq. (15.95) is not lossless since to derive a closed-formula for the update step, it is necessary to assume that the predicted multitarget target state $f(X_k|Z_{1:k-1})$ is approximately a Poisson point process, where the physical distribution of targets is independent and identically distributed (I.I.D.) with a single probability density $g(\mathbf{x})$ and the target number follows a Poisson distribution. The PHD filter can be implemented thorough the sequential Monte Carlo method [120] or the Gaussian mixture method [121].

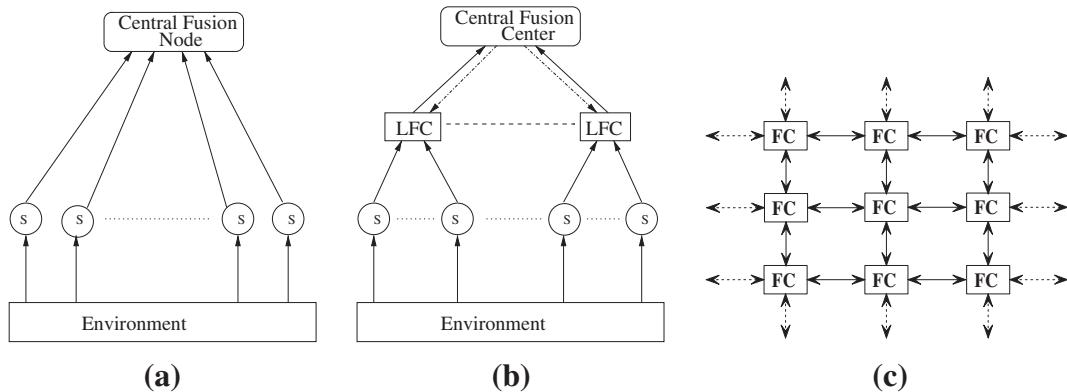
A generalization of PHD filter, so called Cardinalized PHD (CPHD) filter is proposed in [79]. Besides the target PHD function, the CPHD filter also propagates the entirely probability distribution on target number. The core difference between the CPHD filter and the PHD filter is that in the CPHD filter, it is assumed that the predicted multitarget target state $f(X_k|Z_{1:k-1})$ is approximately an I.I.D. cluster point process, where the physical distribution of targets is still I.I.D. with a single probability density $g(\mathbf{x})$ but the target number now can follow any arbitrary distribution $p(n)$. A polynomial running time implementation of the CPHD filer using the Gaussian mixture technique has been proposed in [122].

Both the PHD filter in [78] and the CPHD filter in [79] are unable to maintain the track identity. They only provide identity-free estimates of target states and hence no temporal association of estimates over time. Thus, to use the PHD filter or the CPHD filter as a multitarget tracker, a separate module is required to handle the temporal association for state estimates of individual targets. If the PHD filter or the CPHD filter is implemented thorough the Gaussian mixture technique, then by adding unique tag to each Gaussian component, an association of state estimates to targets over time can be achieved [91].

The multisensor versions of the PHD and CPHD filters are possible-but also computationally intractable [78]. For example, it has been verified in [80] that the measurement-update formula for a two-sensor PHD filter requires a summation over all binary partitions of the current two-sensor measurement set. A computationally tractable approximated multisensor PHD and CPHD filter has been proposed in [83].

3.15.7 Architectures of multisensor-multitarget tracking

Three major types of architecture, namely, centralized, distributed, and decentralized, are commonly used in multisensor-multitarget tracking applications [9, 62, 126]. Here, it is assumed that the data of

**FIGURE 15.2**

Common fusion architectures. (a) Centralized. (b) Distributed. (c) Decentralized.

each sensor has already been converted to a common coordinate system, regardless whether the sensor measurements are processed in the centralized or the distributed architecture.

Centralized Tracking: In the centralized architecture, Figure 15.2a, several sensors are monitoring the region of interest to detect and track the targets therein. All sensors generate measurements at each revisit time and report those measurements to the Central Fusion Center (CFC). It in turn fuses all the acquired measurements and updates the tracks. This is the optimal architecture in terms of tracking performance. However, in a large surveillance region with many sensors, this architecture may not be feasible because of limited resources, e.g., communication bandwidth and computation power. To further improve the reliability and the performance of the centralized architecture, one can use the replicated centralized architecture [24]. In the replicated centralized architecture, there are multiple CFCs operating independently and each sensor reports its measurements to all CFCs. In other words, each CFC processes data from all sensors and there is no communication among CFCs. This architecture has high performance and reliability because the multiple CFCs process the same data. However, it also has higher communication and processing costs.

Two well known schemes used by the CFC to incorporate measurements from multiple sensors are sequential updating and parallel updating:

- In the sequential updating scheme, the updating is carried out with the measurement of one sensor at a time. The posterior state estimate obtained from the measurements reported by the k th sensor is used as the prior information when the measurements from the $k + 1$ th sensor is processed. For one specific sensor, those techniques commonly used in the single sensor-multitarget tracking problem, like one/two point initialization, multi-frame data association, and Kalman filtering, are applied.
- In the parallel updating scheme, the measurements from the various sensors are first associated to yield “supermeasurements,” then these supermeasurements are processed by the CFC as if they were reported by a single sensor. In other words, after the measurement association procedure, the multisensor-multitarget tracking problem has been converted to a single sensor-multitarget tracking problem in the parallel updating scheme.

In the two schemes mentioned above, usually it has been assumed that the sensors are synchronized. In real scenarios, the sensors are seldom perfectly synchronized and too many asynchronous sensors will reduce the optimality of the centralized tracking architecture.

Distributed Tracking: In order to avoid the heavy communication and computational requirement of centralized fusion, distributed or hierarchical architecture, shown in Figure 15.2b, is used alternatively [126]. In this architecture, sensors are connected to Local Fusion Centers (LFCs) and LFCs are in turn connected to a CFC. Each LFC updates its local tracks based on the measurements obtained from the local sensors and sends its tracks state information to CFC. Then, the CFC performs the track-to-track fusion and may send back the updated tracks to the LFCs, if the feedback path is available.

There are two crucial questions in this type of architecture. The CFC needs to decide whether local tracks from different LFCs represent the same target, i.e., the problem of track-to-track association. Furthermore, because of the common process noise or the common prior, the local state estimates for the same target from different LFCs are not independent anymore. Thus, the dependency between the local tracks for the same target from different LFCs has to be handled optimally in the track-track fusion. First, the cross-covariance between the local tracks has to be estimated and then the covariance of the fused track has to be increased correspondingly. Note that, even the above two questions are solved optimally, the distributed tracking architecture is still suboptimal compared to the centralized architecture, because the optimal fusion of the state estimates of local tracks can not provide the same performance as the optimal fusion of the entire data set [9].

To optimally handle the dependency between the local tracks for the same target, the computational cost may be large and the required information may take up too much communication bandwidth, especially when there are many LFCs. Thus, several suboptimal methods are proposed to handle the dependency between the local tracks for the same target. One suboptimal way is using the corresponding tracklets of the local tracks in the CFC for the track-track fusion. A tracklet is a track specially calculated from the local track such that its errors are not cross-correlated with the errors of any other data in the system for the same target [34, 35]. Another suboptimal methods is the covariance intersection technique, which takes a convex combination of the estimate of the mean and the information matrix (i.e., the inverse of the covariance matrix) of the local tracks. The covariance intersection method actually provides an estimate of the upper bound of the covariance matrix of the optimally fused track [25, 51].

Decentralized tracking: When there is no CFC that can communicate with all LFCs in a large surveillance region, neither centralized nor distributed tracking is possible. In such cases, an alternative called decentralized architecture, shown in Figure 15.2c, is used. Decentralized architecture is composed of multiple Fusion Centers (FCs) and no CFC [126]. Here, each FC gets the measurements from one or more sensors that are connected to it, and uses those measurements to update its tracks. In addition, tracks are also updated whenever an FC gets additional information from its neighbors. Note that even though many FCs are available, each FC can communicate only with its neighbors; the FCs within the communication distance every few measurement time steps. There is no common communication facility in decentralized tracking network, i.e., FC cannot broadcast results and communication must be kept on a strictly neighbor-to-neighbor basis [36]. In decentralized tracking architecture, each FC can be considered as a combination of LFC and CFC, because it updates the track state using measurements from connected sensors and it also does track-track fusion whenever neighboring FCs report their track states. Usually, the algorithms developed for the distributed tracking architecture can be used here without any major modification.

Sensor Registration: The benefits afforded by the integration of data from multiple sensors are greatly diminished if sensor biases are present. In practical systems, sensor states are not precisely specified with respect to some common coordinate system and measurements may be subjected to pointing errors, calibration errors or computational errors. As a result, knowledge about sensor location or attitude may not be accurate. Without this knowledge, there may be severe degradation in the performance of data association, filtering and fusion, leading to eventual loss of track quality. Thus, sensor registration and data alignment is the first step in the fusion process. In sensor registration [5], the bias is estimated and the resulting values are used to debias measurements prior to fusion. Assume that local tracks for the same target from different sensors have already been associated, then to estimate the bias vector, the classical approach is to augment the system state to include the bias vector as part of the state and implement an augmented state Kalman filter (ASKF) [39, 54, 74]. However, under some circumstances, the prior track-to-track association may be unnecessary for the sensor registration process [64].

3.15.8 Centralized tracking

It is assumed that there are N_S sensors. The measurement from sensor j at time k is

$$z(k, j) = H(k, j)x(k) + w(k, j), \quad j = 1, \dots, N_S. \quad (15.97)$$

The measurement noise sequences are zero mean, white, independent of the process noise and independent from sensor to sensor with covariances $R(k, j)$. If $N_S = 1$, the above problem becomes a single sensor-multitarget tracking problem. Two widely used techniques to incorporate multiple sensors are sequential updating technique and parallel updating technique.

3.15.8.1 Sequential updating

In sequential updating scheme, the state update is carried out with the measurement of one sensor at a time.

Start the recursion from the predicted state and covariance denoted by

$$\hat{x}(k|k, 0) = \hat{x}(k|k - 1), \quad (15.98)$$

$$P(k|k, 0) = P(k|k - 1). \quad (15.99)$$

The updates with the measurements at time k are

$$\begin{aligned} \hat{x}(k|k, j) &= \hat{x}(k|k, j - 1) + W(k, j)(z(k, j) - H(k, j)\hat{x}(k|k, j - 1)), \quad j = 1, \dots, N_S, \\ P(k|k, j) &= P(k|k, j - 1) - W(k, j)S(k, j)W(k, j)', \quad j = 1, \dots, N_S, \end{aligned} \quad (15.100)$$

where

$$S(k, j) = H(k, j)P(k|k, j - 1)H(k, j)' + R(k, j), \quad (15.101)$$

$$W(k, j) = P(k|k, j - 1)H(k, j)'S(k, j)^{-1}. \quad (15.102)$$

The above update scheme shows that, the update step for the j th sensor is almost same as that for the single sensor-multitarget tracking problem. The only difference is that the posterior track state estimates, $\hat{x}(k|k, j - 1)$, $P(k|k, j - 1)$, which are obtained after the processing of the $j - 1$ th sensor's measurements, is used as the predicted track state $\hat{x}(k|k - 1, j)$, $P(k|k - 1, j)$ for the j th sensor.

For linear measurements, the order of updating in the sequential procedure is immaterial. For non-linear measurements however, measurement from the most accurate sensor should be updated first so as to reduce subsequent linearization errors.

3.15.8.2 Parallel updating

In parallel updating scheme, measurements generated by the same target from each sensor are associated, and simultaneously stacked and updated. After the measurement-to-measurement association process, the multisensor-multitarget tracking problems has been converted into a single sensor-multitarget tracking problem with the following measurement model:

$$z(k) = \begin{bmatrix} z(k, 1) \\ \vdots \\ z(k, N_S) \end{bmatrix} = H(k)x(k) + w(k), \quad (15.103)$$

where

$$H(k) = \begin{bmatrix} H(k, 1) \\ \vdots \\ H(k, N_S) \end{bmatrix},$$

$$w(k) = \begin{bmatrix} w(k, 1) \\ \vdots \\ w(k, N_S) \end{bmatrix},$$

and

$$E[w(k)w(k)^T] = R(k) = \text{diag}[R(k, j)]. \quad (15.104)$$

The most important step of the parallel updating is the measurement-to-measurement association. If S lists of measurements are obtained from S synchronous sensors, then the goal is to group the measurements that could have originated from the same (unknown) target. In one commonly used approach, each feasible S -tuple of measurement $Z_{i_1 i_2 \dots i_S}$, consisting one measurement from each sensor, is assigned a cost (typically, a likelihood ratio similar to (15.80)) and then the set of S -tuples that minimizes the global cost is found. This optimization can be formulated as a multidimensional (S -D) assignment as described in Section 3.15.5.1.

The unknown target state, which is necessary to find the assignment cost, is replaced by its Maximum Likelihood (ML) estimate:

$$X_u = \arg \max_X p(Z_{i_1 i_2 \dots i_S} | X). \quad (15.105)$$

Note that, if the type of measurements does not have a full observability, then the S -tuple in the association needs to have a certain minimum number of measurements from a target in order for the state of the target to be observable.

3.15.9 Distributed tracking

In a distributed or decentralized configuration, each fusion center has a number of tracks. One crucial question is how to handle the dependency between different local tracks generated by the same target.

There are three sources for the cross correlation between any track pairs for the same target [35]:

1. Common prior information. When the same meaningful prior information is used to initialize new tracks for the same target among different Local Fusion Center (LFC) or the Central Fusion Center (CFC), those tracks will be cross-correlated.
2. Updated tracks including duplicate measurement sequence. Considering a multisensor-multitarget problem with two LFCs and one CFC. For target \mathcal{T} , LFC i , $i = \{1, 2\}$ has formed its local track $x_{\mathcal{T}}(k)^{(i)}$ only based on its own measurement set $\{z^{(i)}(1), \dots, z^{(i)}(k)\}$, respectively. At the end of time k , a global track $x_{\mathcal{T}}(k)^G$ for \mathcal{T} has been built inside the CFC by fusing $x_{\mathcal{T}}(k)^{(1)}$, $x_{\mathcal{T}}(k)^{(2)}$. At the end of time $k + m$, the CFC fuses local track $x_{\mathcal{T}}(k+m)^{(i)}$ and the predicted global track $x_{\mathcal{T}}(k+m|k)^G$. However, $x_{\mathcal{T}}(k+m|k)^G$, which is predicted from $x_{\mathcal{T}}(k)^G$, and $x_{\mathcal{T}}(k+m)^{(i)}$, which is estimated from measurement sequence $\{z^{(i)}(1), \dots, z^{(i)}(k+m)\}$, both contain the information in measurement sequence $\{z^{(i)}(1), \dots, z^{(i)}(k)\}$, so $x_{\mathcal{T}}(k+m|k)^G$ and $x_{\mathcal{T}}(k+m)^{(i)}$ are correlated.
3. Common process noise. If target is maneuvering or the state model used by trackers includes the process noise, then those measurements from the same target will be correlated because they contain the same process noise component.

Thus, if the target is strictly moving with a constant velocity, no process noise is used in LFCs and CFC, each LFC form its tracks strictly based on its own measurement sets (i.e., no feedback from the CFC to the LFC), and the CFC is restarted whenever a new or updated local track was available (i.e., memoryless fusion center), then there will not be any cross-correlation between any track pairs for the same target.

One useful tool to explain and analyze the information dependence due to communication is information graph [26,62]. Information events, such as observation by a sensor at a given time or fusion by a FC at a specific time, are represented by the nodes of the graph. The flow of information is represented by a directed link between the nodes. Thus, a node that is a common predecessor to two nodes contains the common information of those two nodes and the common information of any two or more nodes can be found out by identifying their common predecessors. Information graph is especially useful when the communication structure among FCs is complicated since in this case the identification of common information may not be easy.

3.15.9.1 Cross-covariance of the estimation errors

The dependency between the local state estimation errors for the same target is characterized by the cross-covariances of the local estimation errors, so it is important to estimate the cross-covariance between local track pair [9].

Assume sensors are synchronized, then the recursion for the cross-covariance between local track i and j is given by

$$\begin{aligned} P^{ij}(k|k) = & [I - W^i(k)H^i(k)] \left[F(k-1)P^{ij}(k-1|k-1)F(k-1)' \right. \\ & \left. + Q(k-1) \right] [I - W^j(k)H^j(k)]. \end{aligned} \quad (15.106)$$

This is a linear recursion and its initial condition is assuming the initial errors to be uncorrelated, i.e., $P^{ij}(0|0) = 0$ (no meaningful prior information). This is a reasonable assumption in view of the fact that the initial estimates are usually based on the initial measurements, which are assumed to have independent errors.

Assume the difference of the estimates of local track i and j is

$$\hat{\Delta}^{ij}(k) = \hat{x}^i(k|k) - \hat{x}^j(k|k) \quad (15.107)$$

then its corresponding covariance is

$$E \left\{ \hat{\Delta}^{ij}(k)\hat{\Delta}^{ij}(k)^T \right\} = P^i(k|k) + P^j(k|k) - P^{ij}(k|k) - P^{ji}(k|k). \quad (15.108)$$

The above equations indicate that the effect of the dependency between the estimation error is to reduce the covariance of the difference of the estimates of local track i and j , because the common process noise and the common prior reduces the positive correlation between the estimation errors. In addition, the dependency between the local tracks from the same target leads to a larger covariance of the fused estimate state than in the case of independent errors. The above calculation of the cross-correlation is optimal only when it has been done synchronously.

3.15.9.2 Association for tracks with dependent errors

In this algorithm, the problem of associating tracks represented by their local estimates and covariances from S fusion centers is considered [8].

Consider the assignment formulation for track-to-track association from S fusion centers. Assume fusion center S_i has a list of N_i tracks. Define a binary assignment variable $\chi_{i_1 i_2 \dots i_S}$ as

$$\chi_{i_1 i_2 \dots i_S} = \begin{cases} 1 & \text{tracks } i_1, i_2, \dots, i_S \text{ are from the same target,} \\ 0 & \text{otherwise.} \end{cases} \quad (15.109)$$

A subset of indices $\{i_1, i_2, \dots, i_S\}$ could be zero in the assignment variable meaning that no track will be from the target in the corresponding list of the fusion centers.

The S-D assignment formulation finds the most likely hypothesis by solving the following constrained optimization.

$$\min_{\chi_{i_1 i_2 \dots i_S}} \sum_{i_1=0}^{N_1} \sum_{i_2=0}^{N_2} \dots \sum_{i_S=0}^{N_S} c_{i_1 i_2 \dots i_S} \chi_{i_1 i_2 \dots i_S} \quad (15.110)$$

subject to the constraints

$$\sum_{i_2=0}^{N_2} \dots \sum_{i_S=0}^{N_S} \chi_{j i_1 i_2 \dots i_S} = 1, \quad j = 1, 2, \dots, N_1 \quad (15.111)$$

$$\sum_{i_1=0}^{N_1} \sum_{i_3=0}^{N_3} \dots \sum_{i_S=0}^{N_S} \chi_{i_1 j i_3 \dots i_S} = 1, \quad j = 1, 2, \dots, N_2 \quad (15.112)$$

$$\vdots \quad (15.113)$$

$$\sum_{i_1=0}^{N_1} \cdots \sum_{i_{S-1}=0}^{N_{S-1}} \chi_{i_1 i_2 \cdots i_{S-1} j} = 1, \quad j = 1, 2, \dots, N_S \quad (15.114)$$

and

$$\chi_{i_1 i_2 \cdots i_S} \in \{0, 1\}, \quad i_1 = 1, \dots, N_1, \quad i_2 = 1, \dots, N_2, \quad i_S = 1, \dots, N_S. \quad (15.115)$$

In (15.110) the assignment cost is

$$c_{i_1 i_2 \cdots i_S} = -\log \lambda_{i_1 i_2 \cdots i_S} \quad (15.116)$$

where $\lambda_{i_1 i_2 \cdots i_S}$ is the likelihood ratio of the track association hypothesis vs. all tracks from extraneous targets. The following equation can be used to calculate $\lambda_{i_1 i_2 \cdots i_S}$

$$\lambda_{i_1, i_2 \cdots i_S} = V^{M-1} \mathcal{N}[\hat{x}_{\mathcal{S}_i}; 0, P_{\mathcal{S}_i}] \left[\prod_{s \in \mathcal{S}_i} P_{D_s} \right] \left[\prod_{s \in \{1, 2, \dots, S\} \setminus \mathcal{S}_i} 1 - P_{D_s} \right] \quad (15.117)$$

where $1/V$ is the diffuse pdf of track density, $\mathcal{S}_i = \{j | i_j > 0, j = 1 : \dots, S\} = \{s_1, \dots, s_M\}$, M is the number of elements in \mathcal{S}_i and

$$\hat{x}_{\mathcal{S}_i} = \left[\hat{x}_{s_2}^{i_{s_2}} - \hat{x}_{s_1}^{i_{s_1}}, \dots, \hat{x}_{s_M}^{i_{s_M}} - \hat{x}_{s_1}^{i_{s_1}} \right]^T \quad (15.118)$$

and $P_{\mathcal{S}_i}$ is its covariance matrix. From (15.108), the diagonal blocks in $P_{\mathcal{S}_i}$ are

$$(P_{\mathcal{S}_i})_{j-1, j-1} = P_{s_1}^{i_{s_1}} + P_{s_j}^{i_{s_j}} - P_{s_1, s_j}^{i_{s_1} i_{s_j}} - \left(P_{s_1, s_j}^{i_{s_1} i_{s_j}} \right)' \quad j = 2, \dots, M \quad (15.119)$$

and off-diagonal blocks are

$$(P_{\mathcal{S}_i})_{j-1, g-1} = P_{s_1}^{i_{s_1}} - P_{s_1, s_j}^{i_{s_1} i_{s_j}} - \left(P_{s_1, s_j}^{i_{s_1} i_{s_j}} \right)' + P_{s_j, s_g}^{i_{s_j} i_{s_g}} \quad j, g = 2, \dots, M. \quad (15.120)$$

The Maximum Likelihood (ML) estimate of the track states obtained by fusing the set of tracks $\{i_{s_1}, \dots, i_{s_M}\}$ is given by

$$x_{\mathcal{S}_i}^{\text{fused}} = \left(E' \bar{P}_{\mathcal{S}_i}^{-1} E \right)^{-1} E' \bar{P}_{\mathcal{S}_i} \bar{x}_{\mathcal{S}_i}, \quad (15.121)$$

where $E = [I_{n_x} \ I_{n_x} \ \cdots \ I_{n_x}]$ is $(M \times n_x) \times n_x$ matrix and n_x is the dimension of the state vector. Also,

$$\bar{x}_{\mathcal{S}_i} = \left[\hat{x}_{s_1}^{i_{s_1}}, \dots, \hat{x}_{s_M}^{i_{s_M}} \right], \quad (15.122)$$

$$\bar{P}_{S_i} = \begin{bmatrix} P_{s_1}^{i_{s_1}} & P_{s_1, s_2}^{i_{s_1} i_{s_2}} & \dots & P_{s_1, s_M}^{i_{s_1} i_{s_M}} \\ P_{s_2, s_1}^{i_{s_2} i_{s_1}} & P_{s_2}^{i_{s_2}} & \dots & P_{s_2, s_M}^{i_{s_2} i_{s_M}} \\ \vdots & \vdots & \ddots & \vdots \\ P_{s_M, s_1}^{i_{s_2} i_{s_1}} & P_{s_M, s_2}^{i_{s_M} i_{s_2}} & \dots & P_{s_M}^{i_{s_M}} \end{bmatrix}. \quad (15.123)$$

The covariance matrix of the fused track is given by

$$P_{S_i}^{\text{fused}} = \left(E' \bar{P}_{S_i}^{-1} E \right)^{-1}. \quad (15.124)$$

3.15.9.3 Tracklet fusion

A tracklet is a track computed so that its state estimation error is not cross-correlated with the state estimation errors of any other data in the system for the same target. There are two types of methods to build tracklet [35]:

1. Form a tracklet by decorrelating a local track.
2. Form a tracklet directly from a sequence of measurements.

Both of them ignore the cross correlation caused by the common process noise. Furthermore, as being pointed out in [35], only the tracklet-from-track type of methods can be used to decorrelate the cross correlation from the common prior information. Thus, in this chapter, only the tracklet-from-track type of methods are discussed. In the following, $\hat{x}_{n|m}^j$ and $P_{n|m}^j$ represent the predicted track state and covariance matrix in the j th LFC. \hat{x}_n^j and P_n^j represent the posterior track state and covariance matrix in the j th LFC. $F_{n|m}$ is the state transfer matrix from the m th scan to the n th scan.

There are three methods falling in tracklet-from-track category [35]:

1. Inverse Kalman filter (Frenkel's Method No. 1). It outputs u_n^j for a local track \hat{x}_n^j from LFC j at time n , given that the local track was previously reported by the same LFC to the CFC at frame m . Tracklet u_n^j and its covariance are calculated by:

$$P_{n|m}^j = F_{n|m}^T P_m^j F_{n|m}, \quad (15.125)$$

$$A_n^j = P_{n|m}^j \left[P_{n|m}^j - P_n^j \right]^{-1}, \quad (15.126)$$

$$\hat{x}_{n|m}^j = F_{n|m}^T \hat{x}_m, \quad (15.127)$$

$$u_n^j = \hat{x}_{n|m}^j + A_n^j \left[\hat{x}_n^j - \hat{x}_{n|m}^j \right], \quad (15.128)$$

$$U_n^j = A_n^j P_n^j. \quad (15.129)$$

The input to the CFC is u_n^j , the equivalent measurement, and U_n^j , its corresponding measurement error covariance matrix. The equivalent measurement equation matrix H for u_n^j in the CFC is an identity matrix. The global track estimate \hat{x}_m^G has to be predicted to frame n . This method cannot be used if there is only one measurement in the tracklet interval $n - m$.

2. Tracklet with decorrelated state estimate (Frenkel's Method No. 2). It outputs y_n^j for a local track \hat{x}_n^j from LFC j at time n , given that the track data was previously reported by the same LFC to the CFC at frame m . Tracklet y_n^j and its covariance are calculated by:

$$P_{n|m}^j = F_{n|m}^T P_m^j F_{n|m}, \quad (15.130)$$

$$\hat{x}_{n|m}^j = F_{n|m}^T \hat{x}_m, \quad (15.131)$$

$$C_n^j = P_n^j \left[P_{n|m}^j \right]^{-1}, \quad (15.132)$$

$$B_n^j = I - C_n^j, \quad (15.133)$$

$$y_n^j = \hat{x}_n^j - C_n^j \hat{x}_{n|m}^j, \quad (15.134)$$

$$Y_n^j = B_n^j P_n^j. \quad (15.135)$$

The input to the CFC is y_n^j and its corresponding covariance matrix Y_n^j . The equivalent measurement equation matrix H for y_n^j in the CFC is B_n^j . Still, the estimate of global track \hat{x}_m^G has to be predicted to n . This method can be used even when there is only one measurement in the tracklet interval $n - m$.

3. Inverse information filter [9, 36, 104]. It outputs \underline{y}_n^j for a local track \hat{x}_n^j from local tracker j at time n , given that the track data was previously sent by the same LFC to the CFC at frame m . Tracklet \underline{y}_n^j and its covariance are calculated by:

$$P_{n|m}^j = F_{n|m}^T P_m^j F_{n|m}, \quad (15.136)$$

$$\hat{x}_{n|m}^j = F_{n|m}^T \hat{x}_m, \quad (15.137)$$

$$\underline{y}_n^j = \left(P_n^j \right)^{-1} \hat{x}_n - \left(P_{n|m}^j \right)^{-1} \hat{x}_{n|m}, \quad (15.138)$$

$$\underline{Y}_n^j = \left(P_n^j \right)^{-1} - \left(P_{n|m}^j \right)^{-1}. \quad (15.139)$$

The input to the CFC is \underline{y}_n^j and its corresponding covariance matrix \underline{Y}_n^j . Note, to use \underline{y}_n^j and \underline{Y}_n^j directly, the CFC should run the information filter (i.e., inverse Kalman filter).

Since the received tracklets are assumed to be independent of each other and independent of the local tracks, tracklets association problem can be solved by using an *S-D* association technique, where the set of tracklets are treated as the set of measurements with independent measurement noise.

3.15.9.4 Covariance intersection

When cross-correlation between two estimations is unknown and hard to calculate, a sub-optimal information fusion algorithm, called Covariance Intersection (CI) algorithm, is proposed in [51]. It takes a convex combination of the estimation of the mean and the information matrix (the inverse of the covariance matrix) and can provide a consistent fused estimate. Here the consistency for the estimate \hat{x} , \hat{P}_x of unknown state x is defined as:

$$\hat{P}_x - E[(\hat{x} - x)(\hat{x} - x)^T] \geq 0 \quad (15.140)$$

Assume there are two independent sensors, a and b , to observe one target. Each sensor is equipped with a Kalman filter and based on its measurement sequence, at time k , two consistent estimate of target state $\hat{x}(k)_a$, $\hat{P}(k)_a$, $\hat{x}(k)_b$, $\hat{P}(k)_b$ are obtained from sensor a and sensor b , respectively. Now the information from $\hat{x}(k)_a$, $\hat{P}(k)_a$, $\hat{x}(k)_b$, $\hat{P}(k)_b$, has to be fused to get $\hat{x}_f(k)$, $\hat{P}_f(k)$, a better estimation of target state. Furthermore, $\hat{x}_f(k)$, $\hat{P}_f(k)$ has to be consistent.

In the CI algorithm, a linear convex combination formula is used to fuse $\hat{x}(k)_a$, $\hat{P}(k)_a$ and $\hat{x}(k)_b$, $\hat{P}(k)_b$:

$$\hat{P}_f(k)^{-1} = w \hat{P}_a(k)^{-1} + (1 - w) \hat{P}_b(k)^{-1} \quad (15.141)$$

$$\hat{P}_f(k)^{-1} \hat{x}_f(k) = w \hat{P}_a(k)^{-1} \hat{x}_a(k) + (1 - w) \hat{P}_b(k)^{-1} \hat{x}_b(k) \quad (15.142)$$

where $w \in [0, 1]$ is determined by minimizing the trace or the determinant of $\hat{P}_f(k)$. In both cases, the cost function is convex respect to w and the semidefinite convex programming technique can be used to solve this minimization problem. To avoid the time consuming optimization procedure in the CI algorithm, two fast algorithms to calculate the weight w have been proposed in [37, 89].

The intuition behind the CI algorithm comes from the geometric interpretation of covariance matrixes $\hat{P}(k)_a$, $\hat{P}(k)_b$ and $\hat{P}(k)_f$. Just as shown in Fig. 1 of [51], no matter what the cross-correlation between $\hat{x}(k)_a$ and $\hat{x}(k)_b$ is, the theoretical optimal $\hat{P}_f(k)$ always lies within the intersection of two ellipsoids that represent $\hat{P}(k)_a$, $\hat{P}(k)_b$ respectively. Thus, if the fused covariance matrix, $\hat{P}(k)_f$, encloses the intersection region, it must be consistent even if there is no knowledge about the cross-correlation $\hat{P}(k)_{ab}$. Actually, the fused covariance from (15.142) can be thought as a Gaussian approximation of the actual covariance intersection. Obviously, the tighter the updated covariance encloses the intersection region, the greater the amount of information which can be used. The consistent of the fused covariance has guaranteed that the estimation is non-divergent.

From the intuition behind the CI algorithm, it is clear that a significant drawback of the CI algorithm in [51] is that it will inevitably over estimate the fused covariance matrix and will lead to unnecessary loss in its calculated fusion accuracy. As pointed out in [25], if all local trackers give the same covariance matrix estimations, i.e., $\hat{P}_1 = \dots = \hat{P}_n = \hat{P}$, then there is $P_{CI} = \hat{P}$ no matter how many estimates have been fused. In other words, under this circumstance, the error bound provided by P_{CI} does not represent any possible reduction in the uncertainty of the fused estimate. This phenomenon has been explained

in [25] using the set estimation theory. Based on the unknown but bounded set estimation theory, a tighter bound for the fused covariance matrix is given in [25]. There, given two local tracker's estimation $\hat{x}_1, \hat{P}_1, \hat{x}_2, \hat{P}_2$, it assumes that the unknown target state x lies inside two ellipsoids $\Omega(\hat{x}_1, \hat{P}_1), \Omega(\hat{x}_2, \hat{P}_2)$, respectively, and the ellipsoid is defined as

$$\Omega(\hat{x}_i, \hat{P}_i) = \{x : (x - \hat{x}_i)\hat{P}_i^{-1}(x - \hat{x}_i) \leq 1\} \quad (15.143)$$

Given $\Omega(\hat{x}_1, \hat{P}_1), \Omega(\hat{x}_2, \hat{P}_2)$, x can only be in their intersection area, which may not be ellipsoid. However, a bounding ellipsoid for this intersection, which satisfy $\Omega(w) \supset \{\Omega(\hat{x}_1, \hat{P}_1) \cap \Omega(\hat{x}_2, \hat{P}_2)\}$, can be defined as

$$\Omega(w) = \{x : w(x - \hat{x}_1)\hat{P}_1^{-1}(x - \hat{x}_1) + (1 - w)(x - \hat{x}_2)\hat{P}_2^{-1}(x - \hat{x}_2) \leq 1\} \quad (15.144)$$

where $w \in [0, 1]$. With some algebra simplification, the above formula comes into:

$$(x - x_{CI})'[(1 - \alpha^2)P_{CI}]^{-1}(x - x_{CI}) \leq 1 \quad (15.145)$$

in which x_{CI}, P_{CI} follows (15.141), (15.142), respectively and α is given by:

$$\alpha^2 = (\hat{x}_1 - \hat{x}_2)'(w^{-1}\hat{P}_1 + (1 - w)^{-1}\hat{P}_2)^{-1}(\hat{x}_1 - \hat{x}_2) \quad (15.146)$$

Thus, the fused covariance matrix has been shrink from P_{CI} to $(1 - \alpha^2)P_{CI}$.

3.15.10 Performance evaluation

In this section, the Posterior Cramér-Rao Lower Bound (PCRLB) of tracking [10, 44, 46, 47, 115], which provides a minimum bound of any tracking estimation, and several tracker-related metrics, which measure the performance of multitarget trackers in cardinality, time and accuracy [40, 66, 103], are discussed.

3.15.10.1 Posterior Cramér-Rao Lower Bound (PCRLB)

3.15.10.1.1 Background

Let X_k be an unknown and random state vector, and let $\hat{X}_k(Z_k)$ be an unbiased estimate of X_k based on the measurement data, Z_k . The PCRLB, which is defined to be the inverse of the Fisher Information Matrix (FIM), $J(k)$ [117], then gives a lower bound of the error covariance matrix, i.e.,

$$C(k) \triangleq E \{[\hat{X}_k(Z_k) - X_k][\hat{X}_k(Z_k) - X_k]'\} \geq J(k)^{-1}, \quad (15.147)$$

where E denotes expectation over (X_k, Z_k) and $[\cdot]'$ denotes the transpose. The inequality in (15.147) means that $C(k) - J(k)^{-1}$ is a positive semi-definite matrix.

A recursive formula for the evaluation of the posterior FIM, $J(k)$, is given by [116]:

$$J(k+1) = \underbrace{D_k^{22} - D_k^{21}(J(k) + D_k^{11})^{-1}D_k^{12}}_{J_X(k+1)} + J_Z(k+1), \quad (15.148)$$

where

$$D_k^{11} = E \left\{ -\Delta_{X_k}^{X_k} \ln p(X_{k+1}|X_k) \right\}, \quad (15.149)$$

$$D_k^{12} = E \left\{ -\Delta_{X_k}^{X_{k+1}} \ln p(X_{k+1}|X_k) \right\}, \quad (15.150)$$

$$D_k^{21} = (D_k^{12})', \quad (15.151)$$

$$D_k^{22} = E \left\{ -\Delta_{X_{k+1}}^{X_{k+1}} \ln p(X_{k+1}|X_k) \right\}, \quad (15.152)$$

$$J_Z(k+1) = E \left\{ -\Delta_{X_{k+1}}^{X_{k+1}} \ln p(Z_{k+1}|X_{k+1}) \right\}, \quad (15.153)$$

and Δ_α^β is a second-order partial derivative operator whose (i, j) th term is given by

$$\Delta_\alpha^\beta(i, j) = \frac{\partial^2}{\partial \alpha(i) \partial \beta(j)}, \quad (15.154)$$

$\alpha(i)$ and $\beta(i)$ are the i th components of vectors α and β , respectively. In the above, $Z_{k+1} = [Z_{k+1}(1), Z_{k+1}(2), \dots, Z_{k+1}(n)]$, where $Z_{k+1}(i)$ is the measurement vector at sensor i at sampling time $k+1$ and n is the number of sensors utilized at sampling time $k+1$.

3.15.10.1.2 Dynamic system contribution to the PCRLB

Let the state vector at time k , obtained by stacking the state vectors of all targets, be denoted by $X_k = [X_k^1, X_k^2, \dots, X_k^T]'$, where X_k^t is the state vector of target t and T is the total number of targets in the surveillance region. If we assume that targets are moving independently and the state equation of each target is linear, then the overall state equation is given by

$$X_{k+1} = F_k X_k + v_k, \quad (15.155)$$

where

$$F_k = \text{diag}(F_k^1, F_k^2, \dots, F_k^T), \quad (15.156)$$

$$v_k = [v_k^1, v_k^2, \dots, v_k^T]'. \quad (15.157)$$

In the above, F_k^t is the state transition matrix and v_k^t is the process noise of target t . If v_k^t is Gaussian with zero mean and covariance Γ_k^t , then the covariance matrix of v_k , Γ_k , is given by

$$\Gamma_k = \text{diag}(\Gamma_k^1, \Gamma_k^2, \dots, \Gamma_k^T). \quad (15.158)$$

It can be shown that in the case of linear, Gaussian dynamics (e.g., [101]) we have

$$D_k^{11} = F_k' \Gamma_k^{-1} F_k, \quad (15.159)$$

$$D_k^{12} = -F_k' \Gamma_k^{-1}, \quad (15.160)$$

$$D_k^{22} = \Gamma_k^{-1}. \quad (15.161)$$

Using the Matrix Inversion Lemma and (15.159)–(15.161), there is

$$J_X(k) = \left[\Gamma_{k-1} + F_{k-1} J(k-1)^{-1} F'_{k-1} \right]^{-1}. \quad (15.162)$$

The matrix $J_X(k)$ gives the prior information regarding the target states at time k .

Besides process noise, sometimes the dynamic model uncertainty also contributes to the PCRLB, especially in the case of maneuvering targets, whose kinematics model may evolve in a time-varying manner. Thus, it is necessary to consider the PCRLB for a filtering problem with multiple switching dynamic models and additive Gaussian noise. Some related works can be found in [48, 109].

3.15.10.1.3 Measurement contribution to the PCRLB

The measurement contribution to the PCRLB is given by $J_Z(k)$. Consider the general case in which there is measurement origin uncertainty, with measurements originating from one of the targets or from clutter. The j th measurement at the i th sensor is given by

$$Z_k(j, i) = \begin{cases} h_k^i(X_k^t) + \omega_k^i(j) & \text{if originated from target } t, \\ v_k^i(j) & \text{if false alarm,} \end{cases} \quad (15.163)$$

where h_k^i is (in general) a nonlinear function, $\omega_k^i(j)$ is a zero mean Gaussian random variable with covariance Σ_k and $v_k^i(j)$ is uniformly distributed across the surveillance region A (with hyper-volume, V). The probability mass function of the number of false alarms, $\mu_{\text{FA}}(m)$, which is Poisson-distributed with mean λV , is given by

$$\mu_{\text{FA}}(m) = \frac{e^{-\lambda V} (\lambda V)^m}{m!}, \quad (15.164)$$

where m is the number of false alarms and λ is the spatial density of the false alarms.

When multiple targets are present, the association between the measurements and the targets is not known and must be considered in the PCRLB calculation. The following assumptions are made regarding the measurements:

- Each measurement can be generated by one of the targets or the clutter.
- Each target can produce zero or one measurement at any one time.

If sensors have independent measurement processes, J_Z can be written as [114]

$$J_Z(k) = \sum_{i=1}^n J_{Z_i}(k), \quad (15.165)$$

where

$$J_{Z_i}(k) = \sum_{m_k(i)=0}^{\infty} p(m_k(i)) J_{Z_i}(m_k(i), k), \quad (15.166)$$

$$J_{Z_i}(m_k(i), k) = E \left\{ -\Delta_{X_k}^{X_k} \ln p(Z_k(i)|X_k, m_k(i)) | m_k(i) \right\}. \quad (15.167)$$

In the above equations, n is the number of sensors used at time k and $m_k(i)$ is the number of measurements at sensor i at time k .

The probability of receiving $m_k(i)$ measurements, $p(m_k(i))$, from sensor i is given by

$$p(m_k(i)) = \sum_{d=0}^{\min(T, m_k(i))} \left(\mu_{\text{FA}}(m_k(i) - d) \sum_{D_k^i} \left(\prod_{t=1}^T (P_D^t(i))^{D_k^i(t)} (1 - P_D^t(i))^{(1-D_k^i(t))} \right) \right), \quad (15.168)$$

where D_k^i is the detection vector that indicates which targets are detected at sensor i (at time k).¹ The total number of targets that are detected is d , i.e., $\sum_{t=1}^T D_k^i(t) = d$. $P_D^t(i)$ is the probability of detection of target t by sensor i .

The probability density function of the measurement $Z_k(i)$ conditioned on X_k and $m_k(i)$ is given by [49]

$$p(Z_k(i)|X_k, m_k(i)) = \sum_{a_k(i)} p(Z_k(i)|X_k, m_k(i), a_k(i)) p(a_k(i)|m_k(i)), \quad (15.169)$$

where $a_k(i)$ is the association vector that indicates which measurement originated from which target. Each element $a_k(j, i)$ of $a_k(i)$ is a random variable that takes a value in $[0, 1, \dots, T]$, with 0 indicating a false alarm. $a_k(j, i) = t$ indicates that the measurement j originates from target t . If the targets are well separated in the measurement space, there is no measurement origin uncertainty in terms of targets and any one measurement can be originated from a known target or clutter [114]. However, if the targets are closely-spaced or cross one another, it is hard to find the association vector and all possible associations must be considered in the calculation of measurement information $J_Z(k)$.

It has been shown in [115] that (t_1, t_2) th block of $J_{Z_i}(k)$ is given as follows:

$$[J_{Z_i}(k)]_{t_1 t_2} = E \left\{ [H_k^i]'_{t_1} [Q_k^i]_{t_1 t_2} \Sigma_k^{-1} [H_k^i]_{t_2} | X_k \right\}, \quad (15.170)$$

where

$$[H_k^i]_t(\alpha, \beta) = \frac{\partial h_k^i(\alpha, X_k^t)}{\partial X_k^t(\beta)}, \quad (15.171)$$

$$[Q_k^i]_{t_1 t_2} = \sum_{m_k(i)=0}^{\infty} p(m_k(i)) [Q_k^i(m_k(i))]_{t_1 t_2}, \quad (15.172)$$

and $[H_k^i]_t(\alpha, \beta)$ denotes the (α, β) th element of matrix $[H_k^i]_t$. Q_k^i is the Information Reduction Matrix (IRM) for sensor i and $[Q_k^i]_{t_1 t_2}$ is the (t_1, t_2) th block of the IRM. $[Q_k^i(m_k(i))]_{t_1 t_2}$ is also given in [115]. No closed form analytical solution exists for the IRM Q_k^i , which must then be calculated using a numerical integration technique.

¹ $D_k^i(t)$ takes the value 1 if target t is detected and 0 otherwise.

3.15.10.2 Tracker-related measures for performance evaluation

Generally, a tracker-related measure is defined as a function getting \mathcal{T} , \mathcal{L} , and \mathcal{M} as inputs and providing some measures that evaluate the quality of the tracking algorithm. Here, \mathcal{T} is defined as the collection of all tracks, \mathcal{L} is the total information of truth, and \mathcal{M} is the set of all measurements. The following categories are considered for the tracker-related measures [40]:

- *Tracker-dependent:*

These measures are developed for individual types of trackers. For example, performance measures have been specially defined for IMM filters [65, 124], for assignment based tracking algorithms [46], for dynamic programming [50], MHT trackers [20] and IPDA algorithm [125]. This group of metrics is also called *algorithm based* metrics because they are defined for a specific type of tracking and filtering method, or a specific application, and may not be applicable to other tracking methods.

- *Tracker-independent:*

These types of measures can be applied to every tracker. Various categories can be again defined based on the availability of truths and tracks. When truths and tracks are both available, a large class of performance metrics can be defined after finding an association between the estimated tracks and available truths. In real scenarios, sometimes truths are not available. In this case, statistical tests are done on the estimated tracking results to check the consistency of the estimates. The type of metrics is also called *algorithm free* metrics.

In this subsection, only the tracker-independent metrics are reviewed, because only these type of the tracker-related measures can be applied to every tracking algorithms. Different categories of algorithm free metrics may be defined based on the availability of truths and tracking results as

- *Available truths and tracks:*

This is the most popular and applicable case in performance evaluation when the goal is to evaluate tracking results with the known truths. There are three classes of metrics summarized as follows:

- *Track Cardinality Measures:*

This metric measures numerical characteristics of obtained results. For example, the number of confirmed tracks associated with truths, and number of missed and false tracks can be considered as cardinality measures. The major limitation of these measures is that they do not provide any information about the performance of individual tracks, such as the consistency of track and the accuracy of estimation. Also, no information is available about the time characteristics of the estimated tracks.

- *Time (Durational) Measures:*

Time performance of estimated tracks is evaluated by this class of metrics. These performance measures provide information about the persistency of a track. For example, track probability of detection is represented as a metric evaluating the detection ability of a tracker in estimating every truth. Unlike cardinality measures, time metrics can provide more useful information about the duration or persistency of estimated tracks.

- *Accuracy Measures:*

This is the most common measure evaluating the closeness of estimated values to the truths. Several measures can be defined based on the type of distance between the set of truths and tracks. For example, the root mean squared error (RMSE) of target estimation is the most common

criterion used in the literature in which a traditional Mahanabolis distance is derived to compute the error. Some other measures might be also defined based on other types of distances.

To evaluate the quality of the tracking algorithm in track cardinality and state accuracy jointly, the Optimal Subpattern Assignment (OSPA) metric has been proposed in [107]. The OSPA metric in [107] is the summation of two terms, one term measures the cardinality error when the number of tracks is not equal to the number of truth while the other term gives the localization error. However, the track's label is not considered in [107], so the OSPA metric in [107] is unable to measure the track labeling error or the track identity swap. To take into consideration that each track normally has a label and identity, an extension of the OSPA metric has been proposed in [102]. Furthermore, by replacing the criterion from minimizing the RMSE to minimizing the mean OSPA, a new type of multitarget filters, the set-JPDA filter and the set-MHT algorithm, have been proposed in [30,110].

- *Available tracks and unknown truths:*

This case is very common in real scenarios when there is no information about truths. In this situation, the consistency of tracking results may be checked. The innovation of tracking is used as the main source of information and, then, common statistical tests may be made on the received information. There are also other scoring metrics defined in [21].

Above performance evaluation is mainly discussed for a single sensor case. For a multiple sensor problem, performance measures may be found separately for each sensor and, afterwards, a final metric is extracted by fusing individual measures. This method is used for the distributed tracking methods when sensors estimate the state of targets independently. For the centralized case [4], the multiple sensor can be treated as a single sensor case with a group of unique estimates for targets of interest. The measure of performance is then found for the estimated states [4].

3.15.11 Simulations—a multiple closely-spaced target scenario

In this subsection, several tracker-related measures of performance are applied to a simulated scenario in order to demonstrate the application of several multitarget tracking algorithms [40]. The following tracking algorithms are applied in this subsection:

- Interactive Multiple Model-Hypothesis-Oriented Multiple Hypothesis Tracker (IMM-HOMHT) [12].
- Interactive Multiple Model-Track Oriented Multiple Hypothesis Tracker (IMM-TOMHT) [12].
- Multiple Model-Gaussian Mixture Probability Hypothesis Density Filter (IMM-GM-PHD) [93].
- Multiple Model-Cardinalized Gaussian Mixture Probability Hypothesis Density Filter (IMM-GM-CPHD).
- Interactive Multiple Model-SD Assignment (IMM-SD) [97].
- Interactive Multiple Model-Joint Integrated Probability Data Association (IMM-JIPDA) [119].
- Interactive Multiple Model-2D Assignment (IMM-SD-2D) [97].

Parameters of every tracker are adjusted according to the scenario used for the simulation. An EKF algorithm is used in the filtering stage of all algorithms. Three different types of motion models are

used for the IMM filter: the Constant Velocity (CV) model, the Constant Acceleration (CA) model and the Constant Turn (CT) [10] model. Although it is possible that every tracker uses its own selection of motion models, in order to provide a good comparison here the same models are used as IMM modes in every trackers. In order to deal with time varying number of targets, m/n logic and track quality methods are used to initialize new-born targets and delete dead ones [4]. The values of m and n depend on the scenario and tracking algorithm. Except JIPDA algorithm and PHD filters that utilize quality-based method for track management, other methods use m/n logic. Once the estimated tracks are obtained, the tracking results as well as the truths are both used to calculate performance metrics. For the PHD and the CPHD filter, unique tag is attached to each Gaussian component and a JIPDA type algorithm is used for the track management. For the TOMHT, an enumeration based algorithm is used to approximately search the best global hypothesis. Table 15.1 summarizes the general parameters that are common among all scenarios.

A challenging multiple target scenario is used in this section in order to evaluate the performance of trackers in dealing with closely-spaced targets [55].

Parameters of the simulated scenario are described as follows:

- *Sensors and scenario parameters:*

A single sensor is used to generate the measurements of targets. Measurements are gathered in terms of range and bearing [10]. The variance of the range and bearing measurements are chosen to be $\sigma_r = 1.96 \times 10^{-4}$ (nmi) and $\sigma_\theta = 1 \times 10^{-3}$ (rad), respectively. Parameters of the underlying scenario are presented in Table 15.2.

Table 15.1 Common Parameters of Scenarios

Parameter	Description
N_{MCMC}	Number of Monte Carlo runs
N_K	Number of scans
N_T	Total number of truths
$X_s = [x_s^{\max} \ y_s^{\max} \ x_s^{\min} \ y_s^{\min}]'$	Coverage area of surveillance region
R_{\max}	Maximum accessible range
P_D	Sensor probability of detection
P_{FA}	Probability of false alarm
$[m/n]_{\text{init}}$	$[m/n]$ logic for confirm or delete
$[m/n]_{\text{maint}}$	$[m/n]$ logic for keep or delete
P_{dc}^q	Deleting confirmed tracks threshold (quality-based)
P_{dt}^q	Deleting tentative tracks threshold (quality-based)
P_c^q	Track confirmation threshold (quality-based)
P_t^q	Tentative track threshold (quality-based)

Table 15.2 Parameters of the Scenario with Closely-Spaced Targets

Parameter	Value
N_{MCMC}	50
N_K	121
N_T	10
X_s	[−400 200–400 400] (nmi)
P_D	.8
P_{FA}	10^{-4}

- *Target parameters:*

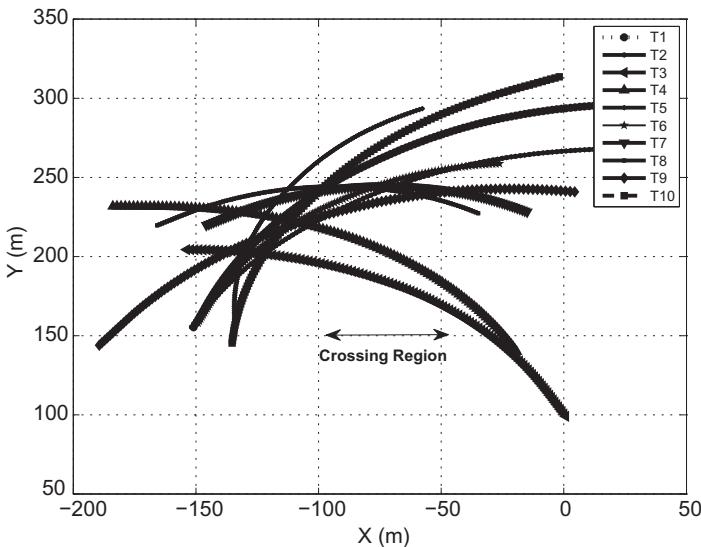
Figure 15.3 shows the generated trajectories of targets in the surveillance region. It can be observed that although all targets are well-separated initially, most of them approach each other in the subsequent scans. In other words, there are several crossing targets in this scenario that makes the tracking problem challenging. Due to the non-maneuvering movement of targets, a CV model is used to characterize the motion of every target.

- *Trackers:*

Table 15.3 presents common parameters for all trackers. Note that PHD algorithm uses quality-based method for track management where $\{P_{dc}^q, P_{dt}^q, P_c^q, P_t^q\}$ are chosen to be $\{.005, .01, .8, .25\}$. Also, IMM-JIPDA tracker utilizes a quality-based approach for track management with $P_t^q = .1$, $P_c^q = .9$, $P_{dc}^q = .0001$, and $p_{dt}^q = .001$ as the parameters.

The tracking results as well as available truths are used to calculate performance metrics. Figure 15.4 shows the performance evaluator, which is designed by the ETF Lab, ECE Department, McMaster University, and this performance evaluator is used in the chapter to calculate and present metrics. Metrics have been classified into time, cardinality and accuracy measures.

For graphical interpretations, the results of three metrics are shown in Figures 15.5–15.7. For the precise definition of each metric, please refer to [40]. Also, the average results are presented in Tables 15.4–15.6 for the analysis over every individual tracker. In all tables, T1 to T6 stand for IMM-SD, IMM-JIPDA, IMM-HOMHT, IMM-TOMHT, IMM-GM-PHD, IMM-GM-CPHD, respectively. It can be seen that all trackers achieve, relatively the same detection performance in terms of track probability of detection. From Figure 15.6, it can be concluded that all targets have been well detected by trackers because measure of completeness shows a value very close to one for all trackers. Note that, there is no dense clutter area in this scenario. Therefore, the average number of false tracks shows a small value for all trackers. Nevertheless, the main difficulty of this scenario is the presence of crossing and closely-spaced targets that may affect the continuity measure and number of breaks and swaps. As the Table 15.4 shows, MHT trackers provide the lowest measure of continuity. It can be seen that IMM-HOMHT generates the most number of breaks in tracks leading to the lowest continuity measure. IMM-JIPDA tracker provides results with acceptable number of breaks even though there are many swaps in the estimated tracks that results in representing IMM-JIPDA as the third-worst tracker in terms of the continuity measure. These results show that, when using the parameters listed in Table 15.3, MHT and JIPDA trackers cannot provide satisfactory results compared to SD and PHD trackers, because the

**FIGURE 15.3**

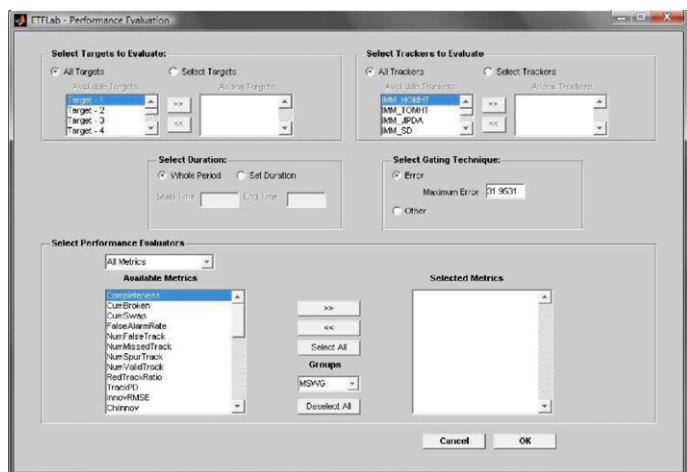
Simulated trajectories for closely-spaced target scenario.

Table 15.3 Parameters of Trackers

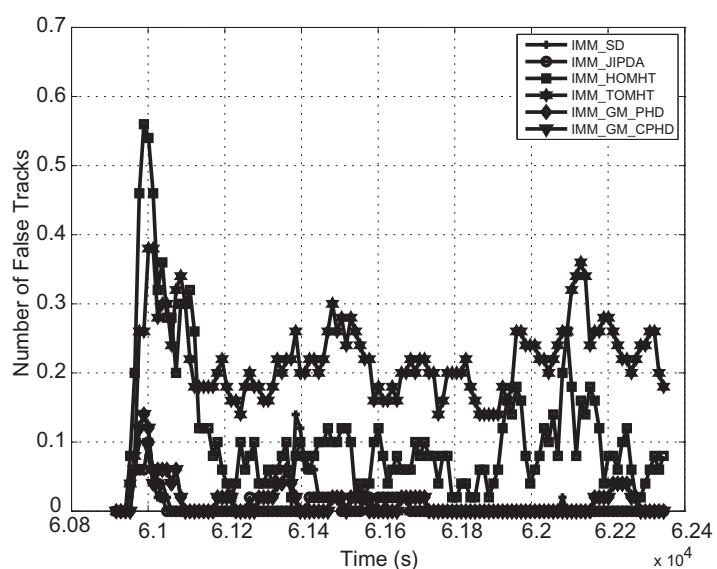
Parameter	IMM-SD	IMM-HOMHT	IMM-TOMHT	IMM-GM-PHD	IMM-GM-CPHD	IMM-JIPDA
$[m/n]_{\text{init}}$	[3, 4]	[3, 4]	[3, 4]	NAN	NAN	NAN
$[m/n]_{\text{maint}}$	[1, 5]	[1, 5]	[1, 5]	NAN	NAN	NAN
v_{\max}	.6 (m/s)	.6 (m/s)	.6 (m/s)	20 (m/s)	20 (m/s)	20 (m/s)
a_{\max}	.02 (m/s ²)	.5 (m/s ²)	.5 (m/s ²)	1 (m/s ²)	1 (m/s ²)	.5 (m/s ²)
v_{\max}^{θ}	$\frac{\pi}{200}$ (rad/s)	$\frac{\pi}{20}$ (rad/s)	$\frac{\pi}{20}$ (rad/s)	$\frac{\pi}{5}$ (rad/s)	$\frac{\pi}{5}$ (rad/s)	$\frac{\pi}{5}$ (rad/s)
P_G	.99	.75	.75	NAN	NAN	.99

probability of breaks and swaps gets higher due to the presence of closely-spaced targets. Also, the tables show the superiority of IMM-SD tracker in terms of the majority of metrics over other trackers.

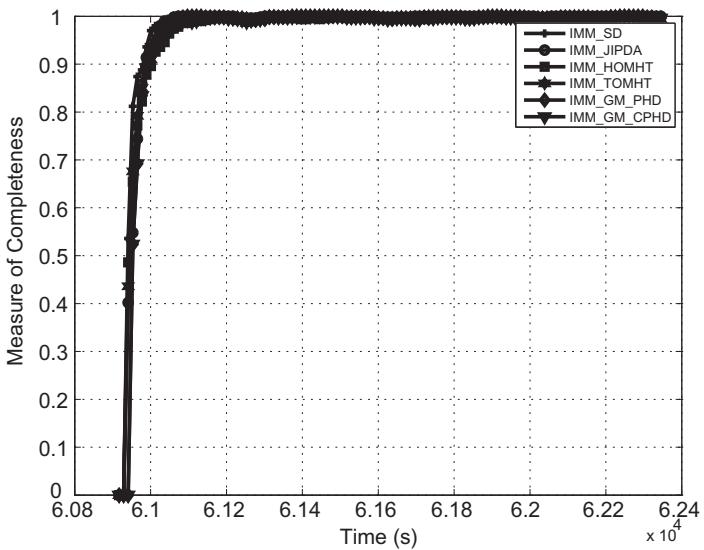
As a conclusion, it can be observed that, with parameters listed in Table 15.3, IMM-SD tracker has been able to provide the most accurate results with the least number of breaks and swaps in the tested tracks. MHT and JIPDA cannot track available targets continually although the detection capability of trackers are still comparable with other trackers. PHD filters stand between the best tracker and MHT and JIPDA trackers with satisfactory detection and continuity measures. It should be noted that

**FIGURE 15.4**

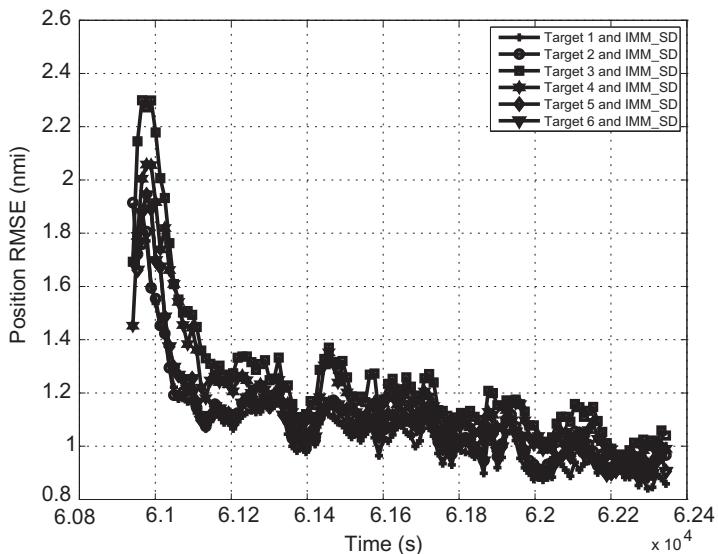
Performance evaluator.

**FIGURE 15.5**

Average number of false tracks for the closely-spaced targets scenario.

**FIGURE 15.6**

Measure of completeness for the closely-spaced targets scenario.

**FIGURE 15.7**

Position RMSE for the closely-spaced targets scenario.

Table 15.4 Cardinality Performance Metrics for the Closely-Spaced Targets Scenario

Metric	T1	T2	T3	T4	T5	T6
Measure of completeness	0.973	0.969	0.966	0.967	0.962	0.962
Average number of breaks in tracks	0.5	0.22	6.06	3.86	0.88	0.88
Average number of swaps in tracks	2.02	4.76	3.14	2.62	3.16	3.16
Rate of false alarm	0.00077	0.00054	0.009	0.018	0.0012	0.0012
Average number of false tracks	0.0093	0.0065	0.111	0.214	0.014	0.014
Average number of missed tracks	0.264	0.31	0.344	0.333	0.379	0.379
Average number of spurious tracks	0.003	0.0028	0.086	0.197	0.01	0.01
Average number of valid tracks	9.73	9.69	9.66	9.667	9.62	9.62
Track continuity (%)	88.98	79.44	67.05	74.49	82.13	82.13

Table 15.5 Time Performance Metrics for the Closely-Spaced Targets Scenario

Metric	Unit	T1	T2	T3	T4	T5	T6
Rate of false alarm	1/s	0.00078	0.00054	0.0092	0.018	0.0012	0.0012
Redundant track ratio	NU	0.983	0.983	0.975	0.965	0.974	0.974
Track probability of detection	NU	0.973	0.969	0.966	0.967	0.962	0.962
Track fragmentation	1/s	0.772	1.92	4.042	1.934	1.344	1.344
Confirmed tracks latency (MAX)	s	72	72	72	72	84	84
Confirmed tracks latency (Average)	s	36	44.4	42	42	55.2	55.2
Tracks deletion latency (MAX)	s	0	0	0	0	0	0
Tracks deletion latency (Average)	s	0	0	0	0	0	0
Tentative tracks latency (MAX)	s	48	60	48	48	72	72
Tentative tracks latency (Average)	s	6	27.6	15.6	15.6	39.6	39.6
Total execution time	s	53.621	51.967	47.459	128.555	32.541	32.725

theoretically the IMM-SD tracker can be considered as an implementation of IMM-TOMHT and both two trackers should provide the same performance. However, the simulation results show that the IMM-SD tracker works better than the IMM-TOMHT. One explanation is that the global optimal hypothesis is obtained thorough the enumeration in the current IMM-TOMHT tracker, so there is no guarantee that the hypothesis found by the current IMM-TOMHT is the real optimal one. On the other hand, in IMM-SD tracker, Lagrange Relaxation method is used to get an approximated optimal global hypothesis.

Table 15.6 Accuracy Performance Metrics for the Closely-Spaced Targets Scenario

Metric	Unit	T1	T2	T3	T4	T5	T6
Root mean squared error (position)	m	1.133	1.172	1.36	1.285	1.178	1.178
Root mean squared error (velocity)	m/s	0.0906	0.0891	0.0904	0.0905	0.0907	0.0907
Average euclidean error (Position)	m	0.838	0.871	0.973	0.927	0.865	0.865
Average euclidean error (velocity)	m/s	0.0895	0.088	0.0892	0.0893	0.0898	0.0897
Average harmonic error (position)	m	0.239	0.257	0.2706	0.259	0.246	0.246
Average harmonic error (velocity)	m/s	0.0867	0.0852	0.086	0.0864	0.0874	0.0874
Average geometric error (position)	m	0.533	0.556	0.602	0.578	0.547	0.547
Average geometric error (velocity)	m/s	0.0883	0.0868	0.0878	0.088	0.0887	0.0887
Bayesian estimation error quotient	NU	0.432	0.436	0.482	0.464	0.443	0.443
Estimate measurement error ratio	NU	0.00131		0.00131	0.0013	NAN	NAN
Normalized estimation error squared	NU	114.43	107.78	109.55	109.15	112.96	112.96

3.15.12 Summary

In this chapter, various filters, data-association techniques, multitarget tracking algorithms, multisensor-multitarget architectures and measures of performance were discussed in detail for the multisensor-multitarget tracking problem. Various combinations of these algorithms provide a complete tracking and fusion framework for multisensor networks with application to civilian as well as military problems. For example, the tracking and fusion techniques discussed here are applicable to fields like air traffic control, air/ground/maritime surveillance, mobile communication, transportation, video monitoring and biomedical imaging/signal processing. Using a scenario with many closely-spaced targets, it was also shown that the algorithms discussed here are all capable of handling the challenging multitarget tracking problem.

Relevant Theory: Signal Processing Theory and Machine Learning

See [Vol. 1, Chapter 4](#) Random Signals and Stochastic Processes

See [Vol. 1, Chapter 11](#) Parametric Estimation

See [Vol. 1, Chapter 12](#) Adaptive Filters

See [Vol. 1, Chapter 19](#) A Tutorial Introduction to Monte Carlo Methods, Markov Chain Monte Carlo and Particle Filtering

References

- [1] G.A. Ackerson, K.S. Fu, On state estimation in switching environments, *IEEE Trans. Automat. Control* 15 (1) (1970) 10–17.
- [2] A.T. Alouani, J.E. Gray, D.H. McCabe, Theory of distributed estimation using multiple asynchronous sensors, *IEEE Trans. Aerosp. Electron. Syst.* 41 (2) (2005) 717–722.

- [3] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Process.* 50 (2) (2002) 174–188.
- [4] Y. Bar-Shalom, W.D. Blair, *Multitarget/Multisensor Tracking: Applications and Advances*, Artech House, October 2000.
- [5] Y. Bar-Shalom, Airborne GMTI radar position bias estimation using static-rotator targets of opportunity, *IEEE Trans. Aerosp. Electron. Syst.* 37 (2) (2001) 695–699.
- [6] Y. Bar-Shalom, Dimensionless score function for multiple hypothesis tracking, *IEEE Trans. Aerosp. Electron. Syst.* 43 (1) (2007) 392–400.
- [7] Y. Bar-Shalom, S. Challa, H.A.P. Blom, IMM estimator versus optimal estimator for hybrid systems, *IEEE Trans. Aerosp. Electron. Syst.* 41 (3) (2005) 986–991.
- [8] Y. Bar-Shalom, H. Chen, Multisensor track-to-track association for tracks with dependent errors, *J. Adv. Inform. Fusion* 1 (1) (2006) 3–14.
- [9] Y. Bar-Shalom, P. Willett, X. Tian, *Tracking and Data Fusion: A Handbook of Algorithm*, YBS Publishing, Storrs, CT, 2011.
- [10] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, 2001.
- [11] Y. Bar-Shalom, T. Kirubarajan, C. Gokberk, Tracking with classification-aided multiframe data association, *IEEE Trans. Aerosp. Electron. Syst.* 41 (3) (2005) 868–878.
- [12] S.S. Blackman, Multiple hypothesis tracking for multiple target tracking, *IEEE Aerosp. Electron. Syst. Mag.* 19 (1, Part 2) (2004) 5–18.
- [13] S. S. Blackman, R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.
- [14] W.D. Blair, G.A. Watson, T. Kirubarajan, Y. Bar-Shalom, Benchmark for radar resource allocation and tracking in the presence of ECM, *IEEE Trans. Aerosp. Electron. Syst.* 34 (4) (1998) 1097–1114.
- [15] H.A.P. Blom, A sophisticated tracking algorithm for ATC surveillance data, in: *Proceedings of the International Radar Conference*, Paris, May 1984, pp. 393–398.
- [16] H.A.P. Blom, Y. Bar-Shalom, The interacting multiple model algorithm for systems with Markovian switching coefficients, *IEEE Trans. Automat. Control* 33 (8) (1988) 780–783.
- [17] A. Capponi, H.W. De Waard, A mean track approach applied to the multidimensional assignment problem, *IEEE Trans. Aerosp. Electron. Syst.* 43 (2) (2007) 450–471.
- [18] C.B. Chang, M. Athans, State estimation for discrete system with switching parameters, *IEEE Trans. Aerosp. Electron. Syst.* 14 (3) (1978) 418–425.
- [19] H. Chen, T. Kirubarajan, Y. Bar-Shalom, Performance limits of track-to-track fusion versus centralized estimation: theory and application, *IEEE Trans. Aerosp. Electron. Syst.* 39 (2) (2003) 386–400.
- [20] K.C. Chang, S. Mori, C.Y. Chong, Evaluating a multiple-hypothesis multitarget tracking algorithm, *IEEE Trans. Aerosp. Electron. Syst.* 30 (2) (1994) 578–590.
- [21] K.C. Chang, X. Zhao, A greedy assignment algorithm and its performance evaluation, in: *Proceedings of American Control Conference*, Seattle, USA, June 1995.
- [22] K.C. Chang, R.K. Saha, Y. Bar-Shalom, On optimal track-to-track fusion, *IEEE Trans. Aerosp. Electron. Syst.* 33 (4) (1997) 1271–1276.
- [23] X. Chen, R. Tharmarasa, M. Pelletier, T. Kirubarajan, Integrated clutter estimation and target tracking using Poisson point processes, *IEEE Trans. Aerosp. Electron. Syst.* 48 (2) (2012) 1210–1235.
- [24] C.Y. Chong, Distributed architectures for data fusion, in: *Proceedings of the First International Conference on Information Fusion*, Las Vegas, NV, August 1998.
- [25] C.Y. Chong, S. Mori, Convex combination and covariance intersection algorithms in distributed fusion, in: *Proceedings of the Fourth International Conference on Information Fusion*, Montreal, QC, Canada, August 2001.

- [26] C.Y. Chong, S. Mori, Distributed fusion and communication management for target identification, in: Proceedings of the Eighth International Conference on Information Fusion, Philadelphia, PA, July 2005.
- [27] M.R. Chummun, T. Kirubarajan, K.R. Pattipati, Y. Bar-Shalom, Fast data association using multidimensional assignment with clustering, *IEEE Trans. Aerosp. Electron. Syst.* 37 (3) (2001) 898–913.
- [28] C.I. Coman, T. Kreitmair, Evaluation of the tracking process in ground surveillance applications, in: Proceedings of the Sixth European Radar Conference, Rome, Italy, September 2009.
- [29] S. Coraluppi, D. Grimmett, P. de Theije, Benchmark evaluation of multistatic trackers, in: Proceedings of the Sixth International Conference on Information Fusion, Florence, Italy, July 2006.
- [30] D.F. Crouse, P. Willett, L. Svensson, D. Svensson, M. Guerriero, The set MHT, in: Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, July 2011.
- [31] F. Daum, Nonlinear filters: beyond the Kalman filter, *IEEE Aerosp. Electron. Syst. Mag.* 20 (8, Part 2) (2005) 57–69.
- [32] S. Deb, M. Yeddanapudi, K.R. Pattipati, Y. Bar-Shalom, A generalized S-dimensional assignment for multisensor-multitarget state estimation, *IEEE Trans. Aerosp. Electron. Syst.* 33 (2) (1997) 523–538.
- [33] A. Doucet, N. de Freitas, N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, New York, 2001.
- [34] O.E. Drummond, A hybrid sensor fusion algorithm architecture and tracklets, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, vol. 3163, July 1997, pp. 485–502.
- [35] O.E. Drummond, Track and tracklet fusion filtering using data from distributed sensors, in: Proceedings of Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, May 2001, pp. 167–186.
- [36] H. Durrant-Whyte, M. Stevens, Data fusion in decentralised sensing networks, in: Proceedings of the Fourth International Conference on Information Fusion, Montreal, QC, Canada, August 2001.
- [37] D. Fränken, A. Hüpper, Improved fast covariance intersection for distributed data fusion, in: Proceedings of the Seventh International Conference on Information Fusion, Stockholm, Sweden, July 2005.
- [38] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc. Radar Signal Process.* 140 (2) (1993) 107–113.
- [39] N.J. Gordon, B. Ristic, M. Robinson, Performance bounds for recursive sensor registration, in: Proceedings of the Sixth International Conference on Information Fusion, Cairns, Australia, July 2003.
- [40] A.A. Gorji, R. Tharmarasa, T. Kirubarajan, Performance measures for multiple target tracking problems, in: Proceedings of 14th International Conference on Information Fusion, Chicago, US, July 2011.
- [41] D. Grimmett, S. Coraluppi, B.R. La Cour, C.G. Hempel, T. Lang, P.A.M. de Theije, P. Willett, MSTWG multistatic tracker evaluation using simulated scenario data sets, in: Proceedings of the 11th International Conference on Information Fusion, Cologne, Germany, September 2008.
- [42] K. Harishan, R. Tharmarasa, T. Kirubarajan, T. Thayaparan, Automatic track initialization and maintenance in heavy clutter using integrated JPDA and ML-PDA algorithms, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, vol. 8137, September 2011.
- [43] T. Hanselmann, D. Mušicki, M. Palaniswami, Adaptive target tracking in slowly changing clutter, in: Proceedings of the Ninth International Conference on Information Fusion, Florence, Italy, July 2006.
- [44] M.L. Hernandez, A.D. Marrs, N.J. Gordon, S.R. Maskell, C.M. Reed, Cramér-Rao bounds for non-linear filtering with measurement origin uncertainty, in: Proceedings of the Fifth International Conference on Information Fusion, Annapolis, USA, 2002.
- [45] M.L. Hernandez, T. Kirubarajan, Y. Bar-Shalom, Multisensor resource deployment using posterior Cramér-Rao bounds, *IEEE Trans. Aerosp. Electron. Syst.* 40 (2) (2004) 399–416.
- [46] M.L. Hernandez, B. Ristic, A. Farina, L. Timmoneri, A comparison of two Cramér-Rao lower bounds for nonlinear filtering with $P_d < 1$, *IEEE Trans. Signal Process.* 52 (9) (2004) 2361–2370.

- [47] M.L. Hernandez, A. Farina, B. Ristic, A PCRLB for tracking in cluttered environments: measurement sequence conditioning approach, *IEEE Trans. Aerosp. Electron. Syst.* 42 (2) (2006) 680–704.
- [48] M.L. Hernandez, B. Ristic, A. Farina, T. Sathyam, T. Kirubarajan, Performance measure for Markovian switching systems using best-fitting Gaussian distributions, *IEEE Trans. Aerosp. Electron. Syst.* 44 (2) (2008) 724–747.
- [49] C. Hue, J.P. Le Cadre, P. Pérez, Performance analysis of two sequential Monte Carlo methods and posterior Cramér-Rao bounds for multitarget tracking, in: Proceedings of the Fifth International Conference on Information Fusion, Annapolis, MD, vol. 1, July 2002, pp. 464–473.
- [50] L.A. Johnston, V. Krishnamurthy, Performance evaluation of a dynamic programming track before detect algorithm, *IEEE Trans. Aerosp. Electron. Syst.* 38 (1) (2002) 228–242.
- [51] S.S. Julier, J.K. Uhlmann, A non-divergent estimation algorithm in the presence of unknown correlations, in: Proceedings of American Control Conference, Albuquerque, New Mexico, June 1997.
- [52] S.J. Julier, J.K. Uhlmann, A new extension of the Kalman filter to nonlinear systems, in: Proceedings of SPIE Conference on Signal Processing, Sensor Fusion and Target Recognition VI, Orlando, FL, vol. 3068, April 1997, pp. 182–193.
- [53] R.E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME — J. Basic Eng.* 82 (1960) 34–45.
- [54] K. Kastella, B. Yeary, T. Zadra, R. Brouillard, E. Frangione, Bias modeling and estimation for GMTI applications, in: Proceedings of the Third International Conference on Information Fusion, Paris, France, July 2000.
- [55] T. Kirubarajan, Y. Bar-Shalom, R. McAllister, R. Schutz, B. Engelberg, Multitarget tracking using an IMM estimator with debiased E-2C measurements for AEW systems, in: Proceedings of the Second International Conference on Information Fusion, Sunnyvale, CA, July 1999.
- [56] T. Kirubarajan, Y. Bar-Shalom, Kalman filter versus IMM estimator: when do we need the latter? *IEEE Trans. Aerosp. Electron. Syst.* 39 (4) (2003) 1452–1457.
- [57] T. Kirubarajan, Y. Bar-Shalom, Probabilistic data association techniques for target tracking in clutter, *Proc. IEEE* 92 (3) (2004) 536–557.
- [58] T. Kirubarajan, Y. Bar-Shalom, D. Lerro, Bearings-only tracking of maneuvering targets using a batch-recursive estimator, *IEEE Trans. Aerosp. Electron. Syst.* 37 (3) (2001) 770–780.
- [59] T. Kirubarajan, H. Wang, Y. Bar-Shalom, K.R. Pattipati, Efficient multisensor fusion using multidimensional data association, *IEEE Trans. Aerosp. Electron. Syst.* 37 (2) (2001) 386–400.
- [60] C. Kreucher, K. Kastella, A.O. Hero III, Multitarget tracking using the joint multitarget probability density, *IEEE Trans. Aerosp. Electron. Syst.* 41 (4) (2005) 1396–1414.
- [61] T. Lefebvre, H. Bruyninckx, J. De Schutter, Kalman filters for non-linear systems: a comparison of performance, *Int. J. Control.* 77 (2004) 639–653.
- [62] M.E. Liggins, C.Y. Chong, I. Kadar, M.G. Alford, V. Vannicola, S. Thomopoulos, Distributed fusion architectures and algorithms for target tracking, *Proc. IEEE* 85 (1) (1997) 95–107.
- [63] L. Lin, Y. Bar-Shalom, T. Kirubarajan, New assignment-based data association for tracking move-stop-move targets, *IEEE Trans. Aerosp. Electron. Syst.* 40 (2) (2004) 714–725.
- [64] X. Lin, T. Kirubarajan, Y. Bar-Shalom, Multisensor bias estimation with local tracks without a priori association, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, vol. 5204, August 2003.
- [65] X.R. Li, Y. Bar-shalom, Performance prediction of interacting multiple model algorithm, *IEEE Trans. Aerosp. Electron. Syst.* 29 (3) (1993).
- [66] X.R. Li, Z. Zhao, Evaluation of estimation algorithms. Part I: Incomprehensive measures of performance, *IEEE Trans. Aerosp. Electron. Syst.* 42 (5) (2006) 1340–1358.

- [67] X.R. Li, N. Li, Integrated real-time estimation of clutter density for tracking, *IEEE Trans. Signal Process.* 48 (10) (2000) 2797–2805.
- [68] N. Li, X.R. Li, Target perceivability and its applications, *IEEE Trans. Signal Process.* 49 (11) (2001) 2588–2604.
- [69] X.R. Li, V. Jilkov, Survey of maneuvering target tracking. Part III. Measurement models, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, vol. 4473, July 2001.
- [70] X.R. Li, V. Jilkov, Survey of maneuvering target tracking. Part IV. Decision-based methods, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, Orlando, FL, vol. 4728, April 2002.
- [71] X.R. Li, V. Jilkov, Survey of maneuvering target tracking. Part I. Dynamic models, *IEEE Trans. Aerosp. Electron. Syst.* 39 (4) (2003) 1333–1364.
- [72] X.R. Li, V. Jilkov, Survey of maneuvering target tracking. Part V. Multiple-model methods, *IEEE Trans. Aerosp. Electron. Syst.* 41 (4) (2005) 1255–1321.
- [73] X.R. Li, V. Jilkov, Survey of maneuvering target tracking. Part II. Motion models of ballistic and space targets, *IEEE Trans. Aerosp. Electron. Syst.* 46 (1) (2010) 96–119.
- [74] X. Lin, Y. Bar-Shalom, T. Kirubarajan, Multisensor multitarget bias estimation for general asynchronous sensors, *IEEE Trans. Aerosp. Electron. Syst.* 41 (3) (2005) 899–921.
- [75] R. Mahler, An Introduction to Multisensor-Multitarget Statistics and its Application, Lockheed Martin Technical Monograph, 2000.
- [76] R. Mahler, Multi-target moments and their application to multi-target tracking, in: Proceedings of the Workshop on Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, 2001, pp. 134–166.
- [77] R. Mahler, Random set theory for target tracking and identification, in: D.L. Hall, J. Lindas (Eds.), *Handbook of Multisensor Data Fusion*, CRC Press, Boca Raton, FL, 2002 (Chapter 14).
- [78] R. Mahler, Multitarget bayes filtering via first-order multitarget moments, *IEEE Trans. Aerosp. Electron. Syst.* 39 (4) (2003) 1152–1178.
- [79] R. Mahler, PHD filter of higher order in target number, *IEEE Trans. Aerosp. Electron. Syst.* 43 (4) (2007) 1523–1541.
- [80] R. Mahler, The multisensor PHD filter: I. General solution via multitarget calculus, in: Proceedings of SPIE Conference on Signal Processing, Sensor Fusion and Target Recognition XVIII, Orlando, FL, vol. 7336, April 2009.
- [81] R. Mahler, CPHD and PHD filters for unknown backgrounds. I: Dynamic data clustering, in: Proceedings of SPIE Conference on Sensors and Systems for Space Applications III, Orlando, FL, USA, vol. 7330, 2009.
- [82] R. Mahler, CPHD and PHD filters for unknown backgrounds. II: Multitarget filtering in dynamic clutter, in: Proceedings of SPIE Conference on Sensors and Systems for Space Applications III, Orlando, FL, USA, vol. 7330, 2009.
- [83] R. Mahler, Approximate multisensor CPHD and PHD filters, in: Proceedings of 10th International Conference on Information Fusion, Edinburgh, UK, July 2010.
- [84] M. Mallick, B.L. Scala, Comparison of single-point and two-point difference track initiation algorithms using position measurements, *Acta Automat. Sinica* 34 (3) (2008) 258–265.
- [85] D. Mušicki, R. Evans, S. Stankovic, Integrated probabilistic data association, *IEEE Trans. Automat. Control* 39 (6) (1994) 1237–1241.
- [86] D. Mušicki, R. Evans, Joint integrated probabilistic data association: JIPDA, *IEEE Trans. Aerosp. Electron. Syst.* 40 (3) (2004) 1093–1099.
- [87] D. Mušicki, S. Suvorova, M. Morelande, B. Mora, Clutter map and target tracking, in: Proceedings of Eighth International Conference on Information Fusion, Wyndham Philadelphia, PA, USA, July 2005, pp. 69–76.
- [88] D. Mušicki, R. Evans, Multiscan multitarget tracking in clutter with integrated track splitting filter, *IEEE Trans. Aerosp. Electron. Syst.* 45 (4) (2009) 1432–1447.

- [89] W. Niehsen, Information fusion based on fast covariance intersection filtering, in: Proceedings of the Fifth International Conference on Information Fusion, Annapolis, Washington DC Area, July 2002.
- [90] N. Okello, B. Ristic, Maximum likelihood registration for multiple dissimilar sensors, IEEE Trans. Aerosp. Electron. Syst. 39 (3) (2003) 1074–1083.
- [91] K. Panta, D.E. Clarak, B.N. Vo, Data association and track management for the Gaussian mixture probability hypothesis density filter, IEEE Trans. Aerosp. Electron. Syst. 45 (3) (2009) 1003–1016.
- [92] G.W. Pulford, Taxonomy of multiple target tracking methods, Proc. IEE Radar Sonar Navig. 152 (5) (2005) 291–304.
- [93] S.A. Pasha, B.N. Vo, W.K. Ma, A Gaussian mixture PHD filter for jump Markov system models, IEEE Trans. Aerosp. Electron. Syst. 45 (3) (2009) 919–936.
- [94] K.R. Pattipati, S. Deb, Y. Bar-Shalom, R.B. Washburn, A new relaxation algorithm and passive sensor data association, IEEE Trans. Automat. Control 37 (2) (1992) 198–213.
- [95] K.R. Pattipati, T. Kirubarajan, R.L. Popp, Survey of assignment techniques for multitarget tracking, in: Proceedings of the Workshop on Estimation, Tracking, and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, May 2001.
- [96] R.L. Popp, K.R. Pattipati, Y. Bar-Shalom, Dynamically adaptable m -best 2D assignment and multi-level parallelization, IEEE Trans. Aerosp. Electron. Syst. 35 (4) (1999) 1145–1160.
- [97] R.L. Popp, T. Kirubarajan, K.R. Pattipati, Survey of assignment techniques for multitarget tracking, in: Y. Bar-Shalom, W.D. Blair (Eds.), Multitarget/Multisensor Tracking: Applications and Advances III, Artech House, 2000 (Chapter 2).
- [98] K. Punithakumar, T. Kirubarajan, A. Sinha, Multiple-model probability hypothesis density filter for tracking maneuvering targets, IEEE Trans. Aerosp. Electron. Syst. 44 (1) (2008) 87–98.
- [99] K. Panta, V. Ba-Ngu, S. Singh, Novel data association schemes for the probability hypothesis density filter, IEEE Trans. Aerosp. Electron. Syst. 43 (2) (2007) 556–570.
- [100] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman Filter: Particle Filters for Tracking Applications, Artech House Publishers, 2004.
- [101] B. Ristic, S. Zollo, S. Arulampalam, Performance bounds for maneuvering target tracking using asynchronous multi-platform angle-only measurements, in: Proceedings of the Fourth International Conference on Information Fusion, Montreal, Quebec, August 2001.
- [102] B. Ristic, B.-N. Vo, D. Clark, B.-T. Vo, A metric for performance evaluation of multi-target tracking algorithms, IEEE Trans. Signal Process. 59 (7) (2011) 3452–3457.
- [103] R. Rothrock, O.E. Drummond, Performance metrics for multiple-sensor, multiple-target tracking, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, Orlando, USA, July 2000.
- [104] M. Ridley, E. Nettleton, S. Sukkarieh, H. Durrant-Whyte, Tracking in decentralised air-ground sensing, in: Proceedings of the Fifth International Conference on Information Fusion, Annapolis, MD, August 2002.
- [105] S. Sarkka, On unscented Kalman filtering for state estimation of continuous-time nonlinear systems, IEEE Trans. Automat. Control 52 (9) (2007) 1631–1641.
- [106] T. Sathyam, A. Sinha, T. Kirubarajan, Computationally efficient assignment-based algorithms for data association for tracking with angle-only sensors, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, vol. 6699, August 2007.
- [107] D. Schuhmacher, B.-T. Vo, B.-N. Vo, A consistent metric for performance evaluation of multi-object filters, IEEE Trans. Signal Process. 56 (8) (2008) 3447–3457.
- [108] D. Schuhmacher, B.-T. Vo, B.-N. Vo, On performance evaluation of multi-object filters, in: Proceedings of the 13th International Conference on Information Fusion, Edinburgh, Scotland, July 2010.
- [109] L. Svensson, On Bayesian Cramér-Rao bound for Markovian switching systems, IEEE Trans. Signal Process. 58 (9) (2010) 4507–4516.
- [110] L. Svensson, D. Svensson, M. Guerriero, P. Willett, Set JPDA filter for multitarget tracking, IEEE Trans. Signal Process. 59 (10) (2011) 4677–4691.

- [111] P.J. Shea, T. Zadra, D. Klamer, E. Frangione, R. Brouillard, K. Kastella, Precision tracking of ground targets, in: Proceedings of IEEE Aerospace Conference, Big Sky, MT, March 2000.
- [112] H. Sidenbladh, Multi-target particle filtering for the probability hypothesis density, in: Proceedings of the Sixth International Conference on Information Fusion, vol. 2, July 2003, pp. 800–806.
- [113] D. Smith, S. Singh, Approaches to multisensor data fusion in target tracking: a survey, *IEEE Trans. Knowledge Data Eng.* 18 (12) (2006) 1696–1710.
- [114] R. Tharmarasa, T. Kirubarajan, M.L. Hernandez, Large-scale optimal sensor array management for multi-target tracking, *IEEE Trans. Syst. Man Cybernet.* 37 (5) (2007) 803–814.
- [115] R. Tharmarasa, T. Kirubarajan, M.L. Hernandez, A. Sinha, PCRLB-based multisensor array management for multitarget tracking, *IEEE Trans. Aerosp. Electron. Syst.* 43 (2) (2007) 539–555.
- [116] P. Tichavsky, C.H. Muravchik, A. Nehorai, Posterior Cramér-Rao bounds for discrete-time nonlinear filtering, *IEEE Trans. Signal Process.* 46 (5) (1998) 1386–1396.
- [117] H. Van Trees, *Detection, Estimation and Modulation Theory*, vol. I, Wiley, New York, 1968.
- [118] B.A. van Doorn, H.A.P. Blom, Systematic error estimation, in multisensor fusion systems, in: Proceedings of SPIE Conference on Signal and Data Processing of Small Targets, Orlando, FL, April 1993.
- [119] J. Vermaak, S.J. Godsill, P. Perez, Monte Carlo filtering for multi-target tracking and data association, *IEEE Trans. Aerosp. Electron. Syst.* 41 (1) (2005) 309–322.
- [120] B.-N. Vo, S. Singh, A. Doucet, Sequential Monte Carlo implementation of the PHD filter for multi-target tracking, in: Proceedings of the Sixth International Conference on Information Fusion, vol. 2, July 2003, pp. 792–799.
- [121] B.-N. Vo, W.K. Ma, The Gaussian mixture probability hypothesis density filter, *IEEE Trans. Aerosp. Electron. Syst.* 54 (11) (2006) 4091–4104.
- [122] B.-T. Vo, B.-N. Vo, A. Cantoni, Analytic implementations of the cardinalized probability hypothesis density filter, *IEEE Trans. Aerosp. Electron. Syst.* 55 (7) (2007) 3553–3567.
- [123] E.A. Wan, R. van der Merwe, The unscented Kalman filter, for nonlinear estimation, in: Proceedings of IEEE, Adaptive Systems for Signal Processing, Communications, and Control Symposium, Alberta, Canada, October 2000.
- [124] H. Wang, T. Kirubarajan, Y. Bar-Shalom, Precision large scale air traffic surveillance using an IMM estimator with assignment, *IEEE Trans. Aerosp. Electron. Syst.* 35 (1) (1999) 255–266.
- [125] X. Wang, D. Musicki, Evaluation of IPDA type filters with a low elevation sea-surface target tracking, in: Proceedings of the Sixth International Conference on Information Fusion, Queensland, Australia, May 2003.
- [126] N. Xiong, P. Svensson, Multi-sensor management for information fusion: issues and approaches, *Inform. Fusion* 3 (1) (2002) 163–186.
- [127] H. You, Z. Jingwei, New track correlation algorithms in a multisensor data fusion system, *IEEE Trans. Aerosp. Electron. Syst.* 42 (4) (2006) 1359–1371.
- [128] R. Zhan, J. Wan, Iterated unscented Kalman filter for passive target tracking, *IEEE Trans. Aerosp. Electron. Syst.* 43 (3) (2007) 1155–1163.

Passive Bistatic Radar

16

Hugh Griffiths

University College London, London, UK

Nomenclature

Principal symbols

A	target silhouette area (m^2)
A_R	radial component of target acceleration (m s^{-2})
a	target dimension (m)
B	receiver effective bandwidth (Hz)
b	target dimension (m)
c	scale parameter of compound K -distribution clutter model
F	receiver effective noise figure (dB)
F_t	pattern propagation factor for transmitter-to-target path
f_D	Doppler shift of received echo (Hz)
G_p	processing gain (dB)
G_r	receive antenna gain (dBi)
G_t	transmit antenna gain (dBi)
k	Boltzmann's constant ($1.38 \times 10^{-23} \text{ J K}^{-1}$)
L	bistatic baseline (m)
N	number of expected targets
P_d	probability of detection
P_{direct}	direct signal power (W)
P_n	receiver noise power (W)
P_r	received target echo power (W)
P_t	transmit power (W)
R_T	transmitter-to-target range (m)
R_R	target-to-receiver range (m)
$s(t)$	transmitted signal
T	update interval (s)
T_0	noise reference temperature, 290 K

T_R	range delay (s)
v	target velocity (m s^{-1})
$w_{i,k}$	particle weight
$(x_i \ y_i)$	particle location
$(\dot{x}_i \ \dot{y}_i)$	particle velocity
β	bistatic angle
δ	angle of target velocity with respect to bistatic bisector
θ_1	elevation angle of transmitter
θ_2	elevation angle of receiver
θ_B	angular width of forward scatter
θ_R	angle of arrival of received echo
λ	wavelength (m)
v	shape parameter of compound K -distribution clutter model
σ_B	target bistatic radar cross section (m^2)
σ_B°	bistatic scattering coefficient (m^2/m^2)
σ_{FS}	forward scatter radar cross section (m^2)
ξ	particle vector
Φ	power density at target (W m^{-2})
Ψ	output of matched filter

2.16.1 Introduction

Passive bistatic radar (PBR) is the name given to a bistatic radar system that makes use of emissions from broadcast, communications or radionavigation transmitters rather than a dedicated, co-operative radar transmitter. The history of the subject goes back a long way; some of the very first radar experiments (such as the famous Daventry Experiment in 1935 [2]) were of this kind, and in fact the “ghosting” effect on analog television receivers due to aircraft targets, which is a manifestation of PBR, was noticed as early as 1938 [3].

The term passive bistatic radar deserves some explanation. Other terms that have been used include *Passive Coherent Location* (PCL), *Parasitic Radar*, *Passive Radar*, *Passive Covert Radar*, and *Piggyback Radar*. For different reasons none of these seems completely satisfactory, so we have used passive bistatic radar throughout.

To complete the definitions, a *hitchhiker* is a bistatic radar which uses an existing monostatic radar as a transmitter of opportunity. There are some significant differences from passive bistatic radars, since the signals are usually pulsed and the transmit radiation pattern is usually directional, and scanning in azimuth. Nevertheless, some of the systems described in this chapter will be examples of hitchhikers rather than strictly passive bistatic radars.

Modern interest in the subject dates back to the 1980s, when a number of systems based on UHF television or VHF FM radio transmissions were built and evaluated [4,5]. Since then interest has grown steadily, as evidenced by the increasing number of publications in research journals and at conferences.

The subject has been especially suitable for research by university groups, since the hardware required for an experimental system is simple and low cost, and there are no licensing issues because the transmitter sources already exist. As well as this, such sources tend to be high power and sited to give wide coverage. PBR may also allow VHF and UHF frequencies to be used which are not normally available for radar, and where in a defense context there may be an advantage against stealthy targets compared to conventional microwave radar frequencies. Finally, provided the receiver antennas are inconspicuous the system is completely covert, and so it is difficult to deploy countermeasures against PBR receivers, since if the location of the receiver is not known any jamming has to be spread over a range of angles, diluting its effectiveness.

PBR can be thought of as fitting into the overall subject of *Waveform Diversity*, since the waveforms of broadcast, communications and radionavigation sources are not explicitly designed for radar use and so may be far from optimum for radar purposes. It is therefore necessary to understand the effect of the waveform on the performance of the passive bistatic radar, so as to be able to choose the most appropriate illuminator, and to process the waveform in the optimal way [6,7].

As well as waveform selection and processing, current areas of research include the development of processing techniques to exploit digital modulation formats, which are replacing analog modulation in many applications and in many countries. Also, it is found that the signal and interference environment, especially in cities, where the density of transmitters is high, is severe in terms of the number and level of signals. This means that appropriate processing techniques are required to suppress this background to a level sufficient to allow target detection. Finally, advances are being made in tracking algorithms, and in radar imaging. These themes are investigated further in the sections that follow.

Despite the attractions noted above, there have been rather few applications where PBR has been able to demonstrate sufficient advantage to justify practical deployment. Most applications, though not all, have been concerned with air surveillance.

Some applications where PBR may be able to demonstrate an advantage include low-cost scientific measurements of the ionosphere [8,9], planets [10], wind [11,12], or thunderstorms [13]. PBR has also been proposed as a gap filler where coverage of conventional air surveillance radars is compromised, for example by wind turbines [14]. Two commercial PBR systems that have been built are Lockheed Martin's *Silent Sentry*TM [15] and Thales's HA-100 *Homeland Alerter* [82]. Nevertheless, the challenge to PBR systems remains to identify and exploit applications where there is a clear advantage in terms of performance and/or cost.

Furthermore, the electromagnetic spectrum represents a finite resource with many different applications, of which radar is only one, competing for frequency allocation. The only thing that can be said with certainty is that this problem of spectrum congestion will only get worse. So there is great motivation and scope for techniques which use the spectrum in a more intelligent and adaptive manner [16], and PBR represents one such technique. In that sense PBR has been described as "green radar," since it does not generate any electromagnetic pollution. It may ultimately be that broadcast and communications signals may be devised also to optimize their performance as PBR illuminators, as well as for their primary purpose.

This chapter assumes familiarity with basic signal waveform and processing theory. Suitable background may be found in Refs. [17–19]. More detailed information on bistatic radar in general may be found in Refs. [20–24].

2.16.2 Bistatic radar

Bistatic radar is defined as a radar in which the transmitter and receiver are at separate locations, sufficiently separated that the properties are significantly different to those of a monostatic radar. This section provides a review of these properties. Many are a consequence of the bistatic geometry [25], which is shown in plane form in Figure 16.1. The transmitter and receiver are separated by a distance L , known as the bistatic baseline. The angle at the target subtended by the transmitter and receiver is the bistatic angle, β . The range from the transmitter to the target is R_T and the range from the target to the receiver is R_R . The target direction at the transmitter and receiver, measured with respect to North, are given by θ_T and θ_R respectively. The target velocity is v , in a direction which makes an angle δ with respect to the bisector of the bistatic angle.

Usually the locations of the transmitter and receiver, and hence the baseline L , are known and the quantities that can be measured for a given target at the receiver are (i) the difference in delay between receipt of the direct signal from the transmitter and the target echo, which provides the bistatic range sum ($R_T + R_R$), (ii) the direction of arrival of the target echo θ_R , and (iii) the Doppler shift f_D of the target echo. The range of the target from the receiver can be deduced from

$$R_R = \frac{(R_T + R_R)^2 - L^2}{2(R_T + R_R + L \sin \theta_R)} \quad (16.1)$$

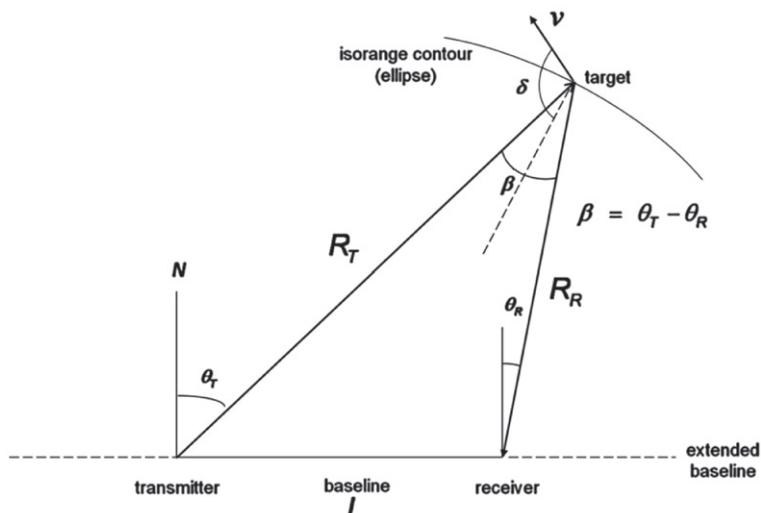


FIGURE 16.1

Bistatic radar geometry.

The Doppler shift f_D is given by the rate of change of the bistatic range sum

$$f_D = \frac{1}{\lambda} \frac{d}{dt} (R_T + R_R) \quad (16.2)$$

In the simplest case where the transmitter and receiver are both stationary and only the target is moving

$$\frac{dR_T}{dt} = v \cos \left(\delta - \frac{\beta}{2} \right) \quad \text{and} \quad \frac{dR_R}{dt} = v \cos \left(\delta + \frac{\beta}{2} \right) \quad (16.3)$$

so that

$$f_D = \frac{2v}{\lambda} \left(\cos \delta \cos \frac{\beta}{2} \right) \quad (16.4)$$

It can be seen that at the point at which the target is crossing the baseline, $\beta = 180^\circ$ and so $f_D = 0$. This can be appreciated since for such a target the transmitter-to-target range is changing in an equal and opposite way to the target-to-receiver range. Furthermore, the echo arrives at the receiver at the same instant as the direct signal from the transmitter, irrespective of the location of the target on the baseline. This means that for a target on the baseline the bistatic radar provides neither range nor Doppler information. Nevertheless, this configuration, which is known as forward scatter, has some attractive properties in respect of target RCS enhancement. These are explained in Section 2.16.4.

2.16.3 Bistatic radar equation

In its simplest form the radar equation for a bistatic radar is derived in the same way as that for a monostatic radar:

$$\frac{P_r}{P_n} = \frac{P_t G_t}{4\pi R_T^2} \cdot \sigma_B \cdot \frac{1}{4\pi R_R^2} \cdot \frac{G_r \lambda^2}{4\pi} \cdot \frac{1}{k T_0 B F} \quad (16.5)$$

in which P_r is the received target echo power; P_n is the receiver noise power; P_t is the transmit power; G_t is the transmit antenna gain; R_T is the transmitter-to-target range; σ_B is the target bistatic radar cross section; R_R is the target-to-receiver range; G_r is the receive antenna gain; λ is the signal wavelength; k is Boltzmann's constant $= 1.38 \times 10^{-23} \text{ J K}^{-1}$; T_0 is the noise reference temperature, 290 K; B is the receiver effective bandwidth; and F is the receiver effective noise figure.

This basic equation can be modified to include the effect of losses, the pattern propagation factors associated with the transmitter-to-target and target-to-receiver paths, and of integration gain.

We can note that the numerator includes the transmit and receive antenna gains individually, and that the denominator includes the factor $1/R_T^2 R_R^2$. This latter factor means that, for omnidirectional transmit and receive antenna patterns, contours of constant signal-to-noise ratio are defined by $R_T R_R = \text{const}$, which are *Ovals of Cassini*, but when directional antenna patterns are used the contours are weighted accordingly and no longer have the same shape. The $1/R_T^2 R_R^2$ factor also means that the signal-to-noise ratio has a minimum value when $R_T = R_R$ (i.e., when the target is equidistant from the transmitter and receiver) and is maximum when the target is either close to the transmitter or close to the receiver.

2.16.4 Target signatures

One of the advantages sometimes ascribed to bistatic radar is that it may offer a counter-stealth capability, since targets that are shaped and/or treated to minimize their monostatic radar cross section (RCS) may nevertheless have higher bistatic RCS. This is not easy to verify, both because bistatic measurements of targets are difficult to make [26], and because the values—certainly those of military targets—are likely to be classified.

One of the few sets of published measurements [27] shows the RCS at X-band of three small ship targets as a function of bistatic angle. Here, the ratio of bistatic to monostatic RCS falls off with increasing bistatic angle, so that as β approaches 45° the RCS is of order 10 dB below the monostatic value. This may be explained by the fact that such targets are likely to include pronounced right-angled features which will form dihedral or trihedral corner reflectors. These will give strong monostatic scattering, but somewhat weaker bistatic scattering.

Targets without such dihedral features will show different behavior. Early theoretical work on bistatic electromagnetic scattering from targets led to the *bistatic equivalence theorem* [28], which states that the bistatic RCS of a given target at a bistatic angle β will be the same as the monostatic RCS measured at the bisector of the bistatic angle, and scaled in frequency by the factor $\cos \beta/2$. This depends on a number of assumptions: (i) the target is sufficiently smooth, (ii) there is no shadowing of one part of the target by another, and (iii) retroreflectors persist as a function of angle. In practice these conditions are not always met, so the theorem should be used with some care, particularly for complex targets and at large values of β .

Enhancement of bistatic target RCS at frequencies where the dimensions of target features are comparable to the radar wavelength (typically at VHF or HF frequencies for aircraft targets) will occur, just as with monostatic signatures. This resonance effect occurs when contributions from different scatterers comprising the target (such as the nose, cockpit, tailplane, engine intakes, etc. of an aircraft) add in phase at a particular radar frequency and geometry.

2.16.4.1 Forward scatter

There is, however, one mechanism by which the bistatic RCS of a target may be enhanced substantially, which is the forward scatter geometry introduced at the end of Section 2.16.2. The enhancement in RCS can be understood by making reference to Babinet's Principle from physical optics [29]. Suppose that an infinite screen is placed between the transmitter and receiver, so that the signal received is zero. Now suppose that a target-shaped hole is cut in the screen, between the transmitter and receiver. Babinet's Principle states that the signal that would be diffracted through the target-shaped hole must be equal and opposite to the signal diffracted around the target, since the two contributions must add to zero (Figure 16.2).

Determination of the signal diffracted through an aperture of a given size and shape is a standard electromagnetics problem, and for simple shapes the results are well known. For example, for a rectangular aperture of sides a and b , the pattern in each plane has a sinc shape with main lobes whose angular widths are

$$\frac{\lambda}{a} \quad \text{and} \quad \frac{\lambda}{b} \quad (\text{radians}) \quad (16.6)$$

and the peak scattered signal corresponds to a forward scatter RCS of

$$\sigma_{FS} = \frac{4\pi a^2 b^2}{\lambda^2} \quad (16.7)$$

Figure 16.3 plots these quantities as a function of frequency for an idealized small aircraft target for which $a = 10$ m and $b = 1$ m and hence $A = 10$ m 2 .

In comparing this with the monostatic RCS (which for a small aircraft might be of the order of 0 dBm 2), it is seen that the forward scatter RCS can be several tens of dB higher, particularly at

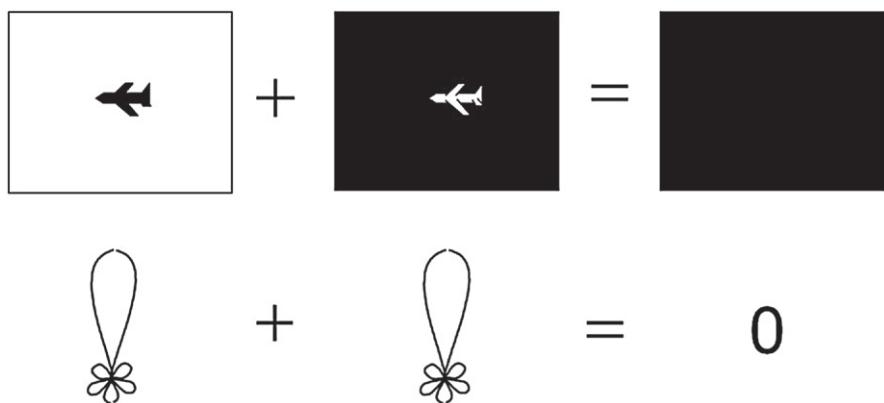


FIGURE 16.2

Babinet's Principle applied to determine the forward scatter radar signature of a target.

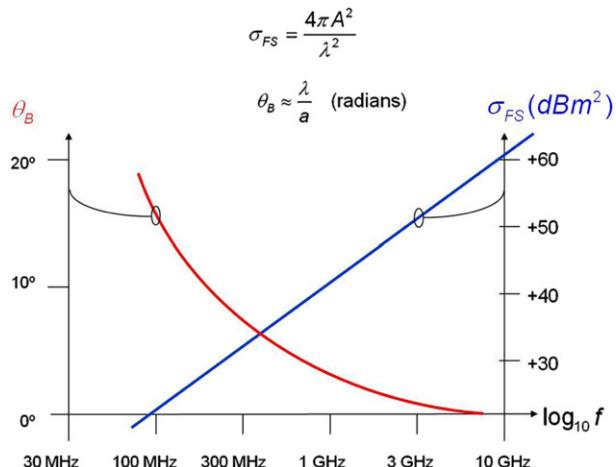
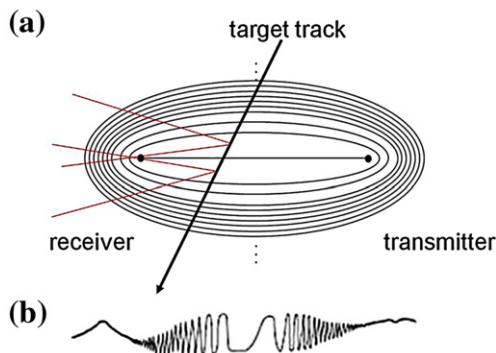


FIGURE 16.3

Forward scatter RCS σ_{FS} and angular width of scatter θ_B for an idealized small aircraft target with $A = 10$ m 2 and $d = 10$ m.

**FIGURE 16.4**

(a) Iso-phase contours. Each contour represents an additional wavelength in the transmitter-target-receiver path; (b) Doppler history of forward scatter echoes from a target crossing the baseline between transmitter and receiver.

microwave frequencies, although the scatter is concentrated in a fairly narrow beam. At lower frequencies the forward scatter RCS is lower, but the scattered energy is spread over a wider range of angles, which might be preferable for the purpose of detection. This suggests that the optimum frequency range to exploit this phenomenon will be at VHF or UHF, which are precisely the bands in which many high-power PBR Illumination sources are to be found. However, while forward scatter may provide very good target detection performance, target localization will be poor. This point is developed in Section 2.16.6 in the context of the ambiguity function for bistatic radar. Finally, we can note that the clutter power will in general be large in forward scatter, both because of the large clutter cell area and because the backscatter coefficient will be large.

The variation of Doppler frequency as a target crosses the baseline can be visualized by imagining a set of elliptical iso-range (or equivalently, iso-phase) contours, with transmitter and receiver as the two focal points, and with each contour corresponding to an extra wavelength in the transmitter-target-receiver path (Figure 16.4). Each time the target crosses one of these iso-phase contours the bistatic echo will undergo one cycle of Doppler shift. The red¹ lines indicate the angular width λ/a of the scattering, so the Doppler variation of the forward-scattered target echo will be detectable at the receiver while the position of the target lies between the two limits shown.

As an example of this, the embedded audio clip is of the signal from a palm-sized VHF FM radio receiver located at the author's home in Camden Town, just to the north of the center of London, listening to the 93.5 MHz BBC Radio 4 transmission from Wrotham, to the south-east of London at a range of approximately 50 km (Figure 16.5).

As the flight paths of aircraft on their way to land at Heathrow airport to the west of London pass over the bistatic baseline formed by the VHF transmitter and the receiver, the Doppler shift of the target echo will follow this form, and the net received signal is the sum of the direct signal and the Doppler-shifted forward scatter echo. This gives a characteristic beat frequency signature, lasting for several seconds and passing through zero as the target crosses the bistatic baseline (Figure 16.3b). This phenomenon is

¹For interpretation of color in Figure 16.4, the reader is referred to the web version of this book.

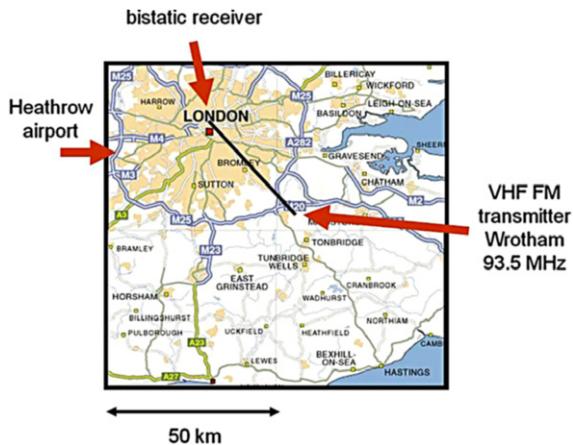


FIGURE 16.5

Forward scatter geometry used for the experiment of the audio recording. Click the [audio clip 1](#) to hear this transmission.

readily observed with simple receiving equipment; indeed the “ghosting” effect that may be observed on analog television receivers due to strong multipath or aircraft targets, referred to in the opening paragraph of this chapter, is another manifestation. It is usually regarded as a nuisance, but of course to us it forms the basis of a radar.

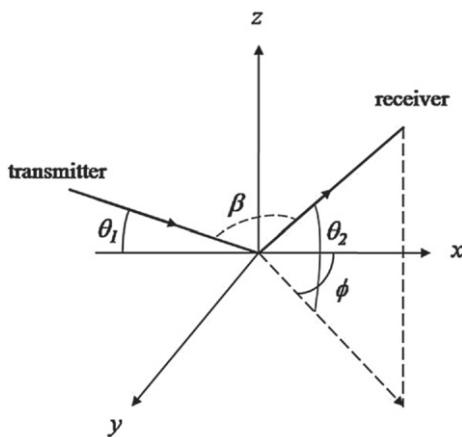
2.16.5 Bistatic radar clutter

A knowledge of the properties of radar clutter, in terms of the mean reflectivity (σ^0), the higher-order (statistical) properties, and the Doppler spectrum, is necessary for performance prediction, comparative performance assessment, design of detection processing, and measurement of performance for acceptance. Bistatic radar clutter depends on all of the same parameters as monostatic clutter, both of the radar (frequency, incidence angle, polarization, resolution cell size, etc.) and of the surface, but also all of those associated with the bistatic geometry [30]. In PBR, just as with conventional radar, clutter will provide a limit to the detection sensitivity, and this point is developed further in the discussion of the signal and interference environment in PBR.

Figure 16.6 shows the bistatic geometry and defines the associated angles. The incidence and scattering angles are θ_1 and θ_2 respectively. In some treatments θ_1 is denoted as θ_i and θ_2 as θ_s ; however, the meaning of the two angles can be interchanged without changing the result because of the reciprocity theorem of electromagnetics. β is the bistatic angle, related to θ_1 and θ_2 by

$$\beta = \cos^{-1} (\sin \theta_1 \sin \theta_2 - \cos \theta_1 \cos \theta_2 \cos \phi) \quad (16.8)$$

and which is zero for the monostatic case and 180° for forward scattering. $\phi = 0^\circ$ or 180° corresponds to the special case referred to as in-plane scattering; the general case of $\phi \neq 0^\circ$ or 180° is referred to as out-of-plane scattering.

**FIGURE 16.6**

Bistatic clutter geometry.

Probably the best summary of present knowledge is the chapter by Weiner [31], which lists nine bistatic clutter measurement campaigns from the past 40 years, of which five include sea clutter, and only one of which used out-of-plane geometries. Domville [32] reported the results of an extensive set of trials in the UK in the mid-1960s of X-band radar scattering from a range of different surface types (sea, urban land, rural land, forest, semi-desert). These used an aircraft-borne CW transmitter and a receiver in a second aircraft. An example of these results is shown in Figure 16.7, corresponding to horizontally polarized in-plane scattering from the sea with a 20 kt wind, plotting the mean bistatic scattering coefficient σ_B^o in terms of the incidence and reflection angles. Thus $\theta_2 = \theta_1$ corresponds to monostatic scattering; $\theta_2 = (180^\circ - \theta_1)$ corresponds to specular scattering (the “specular ridge”) with a maximum at grazing incidence (forward scatter).

Yates investigated the difference in statistical distributions of simultaneously-acquired monostatic and bistatic X-band SAR imagery of land scenes [33]. She used the compound K -distribution model [34], which models the clutter as Rayleigh-distributed speckle modulated by a gamma-distributed texture. This has been found to give a good representation of the properties of high-resolution low grazing angle sea clutter, as well as high-resolution land clutter, and even high-resolution sonar scattering from the seabed. The usual form of the distribution is:

$$f(x) = \frac{4c}{\Gamma(\nu)} (cx)^\nu K_{\nu-1}(2cx) \quad (16.9)$$

in which c is a scale parameter, ν is a shape parameter and $K_\nu(\cdot)$ is the modified Bessel function of the third kind of order ν . The value of the shape parameter is a measure of the length of the tail of the distribution or degree of “spikiness” of the clutter, such that a low value of ν (<1) indicates spiky clutter. For $\nu = \infty$ the distribution reduces to Rayleigh.

Her results showed that the K -distribution gave a good fit to the measured data, but more importantly, that the shape parameter ν was consistently higher for the bistatic imagery than for the monostatic imagery—in other words, that the distribution corresponding to the bistatic clutter was shorter-tailed.

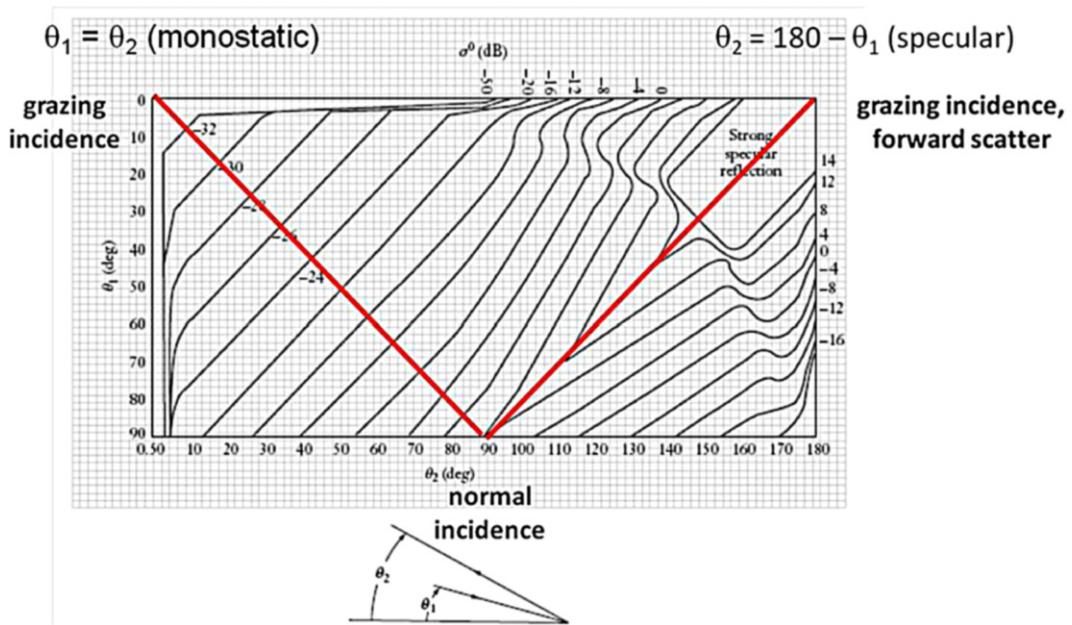


FIGURE 16.7

Measured dependence of in-plane bistatic scattering coefficient of X-band sea clutter at HH polarization on angles θ_1 and θ_2 (adapted from [32]).

A physical explanation for this is that the target scene contains dihedral or trihedral features, such as those formed by horizontal ground and vertical building walls or tree trunks or stalks of vegetation, which give strong monostatic scatter but weaker bistatic scatter.

Recent measurements of the statistical properties of simultaneously-acquired monostatic and bistatic S-band sea clutter have shown essentially similar behavior [35, 36]. This suggests that there should be an advantage in using the bistatic geometry for the detection of small targets against a clutter background. In the examples measured in this work the improvement in detection sensitivity is a function of the bistatic geometry, but in some configurations can be more than 6 dB. Thus in any practical application it will be important to choose the appropriate bistatic geometry in order to best exploit this effect, which has been termed *clutter diversity*.

Clearly there is a great deal more work to be done to explore and understand these effects as a function of all of the parameters: frequency, bistatic geometry, polarization, resolution and surface properties, and hence to develop reliable models.

2.16.6 PBR waveforms

A key feature of PBR is that the waveforms of the signals are not explicitly designed for radar purposes, and therefore that their performance when used for PBR may be far from optimal. It is important,

therefore, to understand the nature of the wide range of waveforms that may be used for PBR, so as to be able to choose the most suitable ones and to process them in the most appropriate way. The classical tool for analyzing and displaying waveform properties is the *ambiguity function*, introduced by Woodward [18] in the early 1950s:

$$|\psi(T_R, f_D)|^2 = \left| \int_{-\infty}^{\infty} s_t(t) s_t^*(t + T_R) \exp[j2\pi f_D t] dt \right|^2 \quad (16.10)$$

The ambiguity function is the square magnitude of the output from a filter matched to the signal $s_t(t)$, and represents the point target response of the radar as a function of delay T_R and Doppler shift f_D . A plot of this (the *ambiguity diagram*) shows the resolution, sidelobe pattern, and presence of ambiguities in range and in Doppler in an elegant and easily-visualized manner.

In the case of bistatic radar it has already been noted that the range and Doppler resolution for a target on or close to the bistatic baseline are degraded, no matter what the waveform. Thus it can be seen that the ambiguity function for a bistatic radar depends not only on the waveform properties, but also on the location of the target with respect to the transmitter and receiver. This was analyzed by Tsao et al. [37], who also explained the effect by pointing out that the simple linear relationships between target range and echo delay and between target velocity and Doppler shift that exist for a monostatic radar no longer hold for a bistatic radar, and hence that the ambiguity function has to be expressed as a function of more variables than just T_R and f_D . They express the ambiguity function for a bistatic radar as:

$$\begin{aligned} & |\psi(R_{RH}, R_{Ra}, V_H, V_a, \theta_R, L)|^2 \\ &= \left| \int_{-\infty}^{\infty} s_t(t - \tau_a(R_{Ra} \cdot \theta_R, L)) s_t^*(t + \tau_R(R_{RH}, \theta_R, L)) \right. \\ &\quad \times \left. \exp[j2\pi f_{DH}(R_{RH}, V_H, \theta_R, L) - 2\pi f_{Da}(R_{Ra}, V_a, \theta_R, L)t] dt \right|^2 \end{aligned} \quad (16.11)$$

where R_{RH} and R_{Ra} are the hypothesized and actual ranges (delays) from the receiver to the target, V_H and V_a are the hypothesized and actual target radial velocities with respect to the receiver, f_{DH} and f_{Da} are the hypothesized and actual Doppler frequencies, and θ_R and L are as defined in Figure 16.1.

To illustrate this effect, Figure 16.8 shows the ambiguity function for four different target locations and directions of motion. The waveform in each case consists of a short sequence of three rectangular pulses. In Figure 16.8a the target is on the baseline approaching the receiver and the ambiguity function is essentially the same as for the monostatic configuration. In Figure 16.8b the target approaches the baseline obliquely, but the ambiguity function is little altered. In Figure 16.8c the target approaches the baseline from a perpendicular direction, which broadens the main peak and alters the position of the sidelobes in the ambiguity function. Finally in Figure 16.8d the target crosses the baseline, at which point there is effectively no resolution in either range or Doppler.

As well as the distortion of the ambiguity function due to the bistatic geometry, we also need to consider the form and variation of the ambiguity function due to the waveform itself. Several workers have investigated the performance of the waveforms of different PBR sources by digitizing off-air

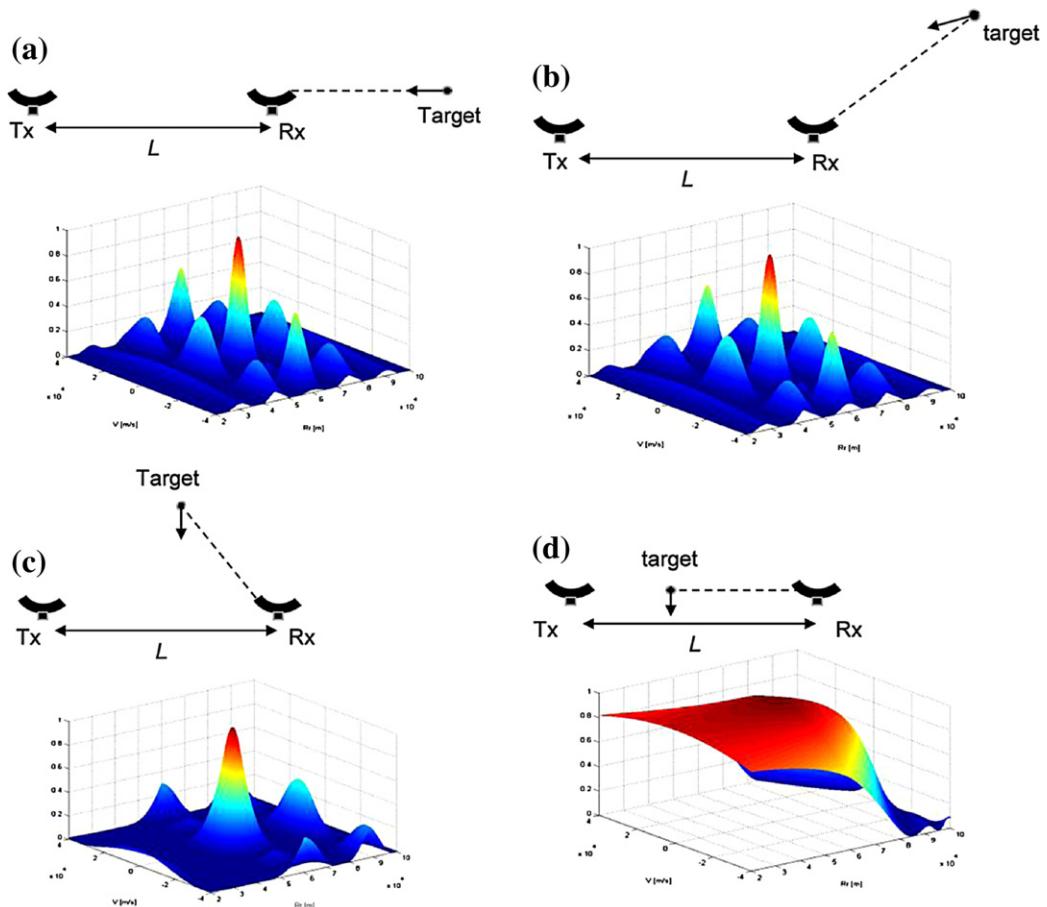


FIGURE 16.8

Bistatic ambiguity functions for a sequence of three rectangular pulses and for four different target locations and directions of motion.

samples of the signals and computing and plotting their (monostatic) ambiguity functions [38–40]. Of course, these results represent a best case, since the range and Doppler resolution may be degraded if the bistatic geometry is less favorable. Some interesting and important points emerge from this.

Firstly, if there are periodic features associated with the modulation, there are corresponding ambiguities. As an example of this, in the PAL analog TV format used in the UK the picture information is encoded as 625 successive lines, each of $64 \mu\text{s}$ duration in two interlaced scans at a 25 Hz rate, with a $5.5 \mu\text{s}$ sync pulse at the beginning of each line. Because of the sync pulses and because in general there will be strong correlation between adjacent lines of a TV picture, there will be pronounced bistatic range ambiguities at multiples of 9.6 km, corresponding to the $64 \mu\text{s}$ line repetition period. Other analog

modulation formats that have been used in other countries, such as NTSC (in the USA) and SECAM (in France) are broadly similar.

Secondly, it is found that for analog modulation formats the ambiguity performance depends significantly on the instantaneous program content. For example, with FM radio the ambiguity function depends strongly on whether the modulation is speech or music, and even on the type of music. With speech, during pauses between words the bandwidth of the modulated signal is narrow and hence the range resolution is poor. With music that has a high spectral content the bandwidth of the modulated signal is broad and hence the range resolution rather better, so (for example) a cacophony of rock music is likely to be better for PBR purposes than speech or music from a solo instrument. Thus in general PBR signals with analog modulation have an ambiguity performance that is time-varying.

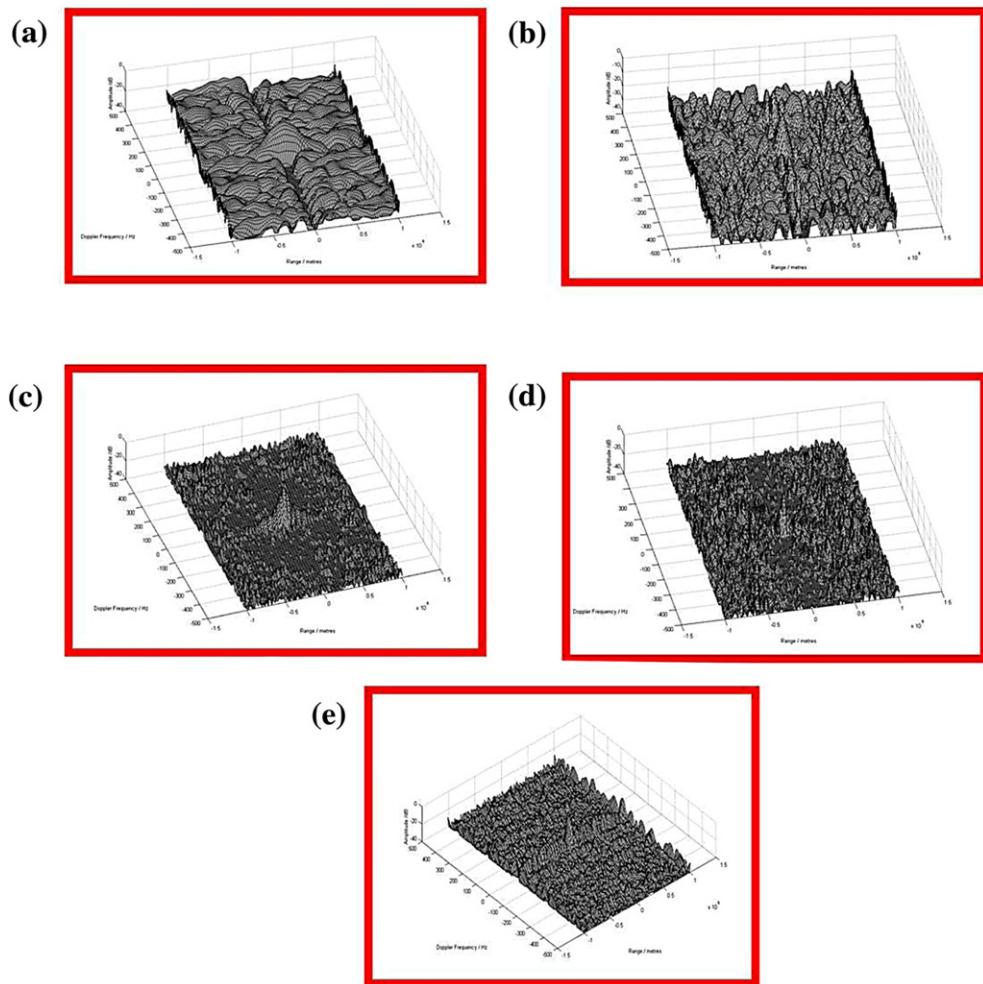
In contrast, digital modulation formats such as DAB and DVB give a signal spectrum that is flatter and more noise-like, and which does not depend on instantaneous program content, so the ambiguity performance does not vary with time and does not depend on the program content, and better approximates to the thumbtack ideal. In many countries analog television and radio are being phased out and replaced by their digital counterparts. In the USA, for example, analog television was discontinued in 2009. France switched off all analog services on 29 November, 2011, to be followed by Japan on 31 March, 2012.

Some examples of the measured ambiguity functions are shown in Figure 16.9. Figure 16.9a shows the ambiguity function of a VHF FM radio station with speech modulation (BBC Radio 4). The peak and the sidelobe structure are well defined, although the peak is relatively broad, as a consequence of the low spectral content of the modulation. Figure 16.9b shows the equivalent result for a station with fast-tempo jazz music modulation (Jazz FM). The peak and the sidelobe structure are correspondingly sharper because of the higher spectral content of the modulation. In both cases, the floor of the ambiguity function is down by a factor of $(B\tau)^{1/2}$, rather than by $(B\tau)$, which would be expected for coherent waveforms. Figure 16.9c–e shows typical ambiguity functions for digital transmissions (DAB, DBV-TV, and GSM, respectively). These functions are more favorable for PBR purposes than signals with analog modulation, such as Figure 16.9a and b since the peak of the ambiguity function is narrower and the sidelobes are lower. Also, they are time-invariant and do not depend on the program content.

A subtlety associated with some digital transmissions is the use of single frequency networks. Here, all transmitters in a given network use the same frequency, with Coded Orthogonal Frequency Division Multiplexing (COFDM) modulation, and each information block is preceded by a guard interval of duration greater than the maximum expected multipath delay [42].

Radio and television broadcast transmitters are in general high-powered and favorably-sited to give wide coverage. In practice the transmit power used by broadcast transmitters has to take into account the fact that the antennas of domestic receivers may be inefficient and/or poorly-sited, so the transmitter power will be several tens of dB greater than would be needed for a dedicated, optimized link. As an example, the highest-power analog television transmitters in the UK have an EIRP of 1 MW, with the antenna on a tower as much as 330 m high, which is itself generally located on top of a high hill.

There is a huge variety of sources that may be used as PBR illuminators. In addition to the analog and digital radio and television transmitters mentioned above, we can add cellphone basestations, HF short-wave broadcast transmitters, WiFi and WiMAX transmitters and various satellite-borne sources in a variety of different orbit patterns. Table 16.1 gives a list (which is by no means exhaustive), showing for each the form of modulation and evaluating the power density $\Phi = P_t G_t / (4\pi R_T^2)$ at representative values of transmitter-to-target range.

**FIGURE 16.9**

Measured ambiguity functions from (a) VHF FM BBC Radio 4 (speech), (b) VHF FM Jazz FM (fast-tempo jazz music), (c) Digital Audio Broadcast (DAB) at 222.4 MHz, (d) Digital Video Broadcasting (DVB-T) at 505 MHz, and GSM900 at 944.6 MHz.

2.16.7 The signal and interference environment [43]

It has already been remarked in Section 2.16.5 that the direct signal, multipath (i.e., clutter) and interference may limit the detection performance of a passive bistatic radar, especially in an urban environment.

Table 16.1 Summary of Typical Parameters of PBR Illuminators of Opportunity (from [41])

Transmission	Frequency	Modulation, Bandwidth	$P_t G_t$	Power Density ^a $\Phi = \frac{P_t G_t}{4\pi R_T^2}$
HF broadcast	10–30 MHz	DSB AM, 9 kHz	50 MW	–67 to –53 dBW/m ² at $R_T = 1000$ km
VHF FM	88–108 MHz	FM, 50 kHz	250 kW	–57 dBW/m ² at $R_T = 100$ km
Analog TV	~550 MHz	PAL, SECAM, NTSC 5.5 MHz	1 MW	–51 dBW/m ² at $R_T = 100$ km
DAB	~220 MHz	digital, OFDM 220 kHz	10 kW	–71 dBW/m ² at $R_T = 100$ km
Digital TV	~750 MHz	digital, 6 MHz	8 kW	–71 dBW/m ² at $R_T = 100$ km
Cellphone base station (GSM)	900 MHz, 1.8 GHz	GMSK, FDMA/TDMA/FDD 200 kHz	10 W	–81 dBW/m ² at $R_T = 10$ km
Cellphone base station (3G)	2 GHz	CDMA, 5 MHz	10 W	–81 dBW/m ² at $R_T = 10$ km
WiFi 802.11	2.4 GHz	DSSS/OFDM, 5 MHz	100 mW	–41 dBW/m ² at $R_T = 10$ m
WiMAX 802.16	2.4 GHz	QAM, 20 MHz	20 W	–88 dBW/m ² at $R_T = 10$ km
GNSS	L-band	CDMA, FDMA 1–10 MHz	200 W	–134 dBW/m ² at Earth's surface
DBS TV	Ku-band 11–12 GHz	analog and digital	55 dBW	–107 dBW/m ² at Earth's surface
Satellite SAR ^c	5.3 GHz	chirp pulse, 15 MHz	68 MW	–55 dBW/m ² at Earth's surface

^a Assuming free-space line-of-sight propagation.
^b Would be subject to additional attenuation due to propagation through walls.
^c Parameters from ASAR instrument carried by ESA's ENVISAT satellite.

In a similar way to Eq. (16.5), the ratio of the direct signal (DSI) from the transmitter at the PBR receiver to the thermal noise level is given by

$$\frac{P_{\text{direct}}}{P_n} = \frac{P_t G_t}{4\pi L^2} \cdot \frac{G_r \lambda^2}{4\pi} \cdot \frac{1}{k T_0 BF} \quad (16.12)$$

under the assumption of free-space line-of-sight propagation.

Inserting typical values shows that this ratio may easily be as much as 100 dB, or even more, and this direct signal will be compounded by multipath versions of the same signal, other co-channel transmissions, broadband noise from sources such as computers or imperfectly-suppressed vehicle ignition, and out-of-band emissions from other transmitters (the spectral mask specification for such transmitters extends down typically to -60 dB; at this level the out-of-band emissions have no effect on reception by conventional FM radio receivers). The signal and interference environment is shown schematically in Figure 16.10, plotted as a function of frequency and of direction ($-1 \leq \sin \theta \leq +1$). The environment may also be non-stationary, since some of the reflections may come from moving vehicles or other clutter sources. The problem of detecting target echoes against such a background has been likened to “listening for a whisper in a cocktail party” [44].

Figure 16.11 provides another illustration of the same phenomenon. This shows noise levels in a 50 kHz bandwidth in the FM radio band from 95 to 100 MHz, measured from the 11th floor of a building at UCL in central London. Figure 16.11a shows thermal noise level (approximately -115 dBm) corresponding to a 5 dB receiver noise figure, obtained with the antenna input terminated in a 50Ω load. Figure 16.11b shows the equivalent spectrum when a vertically-polarized dipole is connected. It is offset vertically by 10 dB with respect to Figure 16.11a to allow a proper side-by-side comparison between the signal levels. In this case the strongest direct signals are some 85 dB above thermal noise.

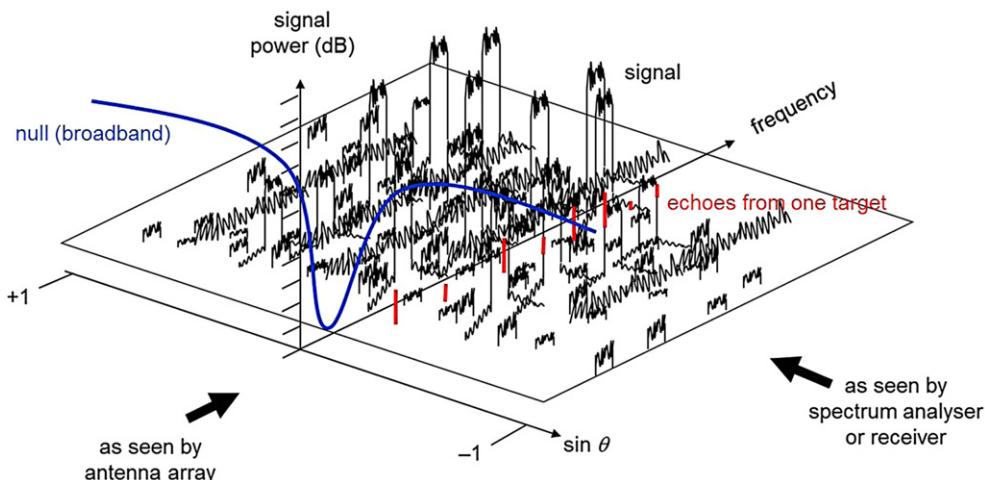


FIGURE 16.10

The signal and interference environment at the PBR receiver.

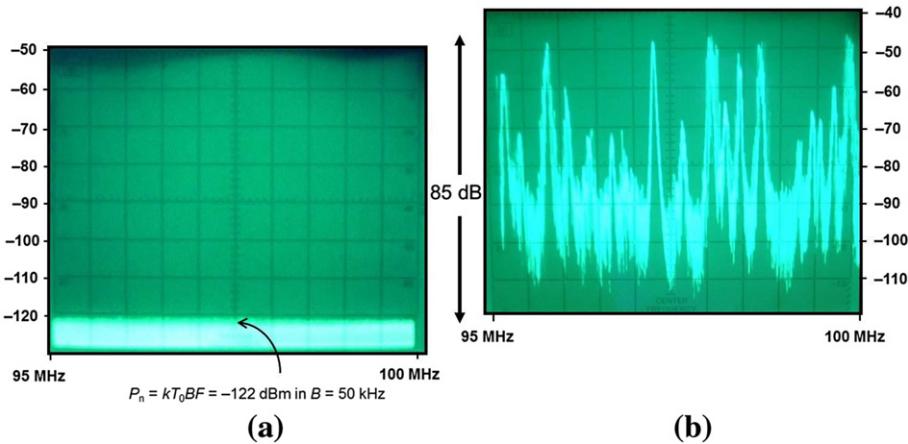


FIGURE 16.11

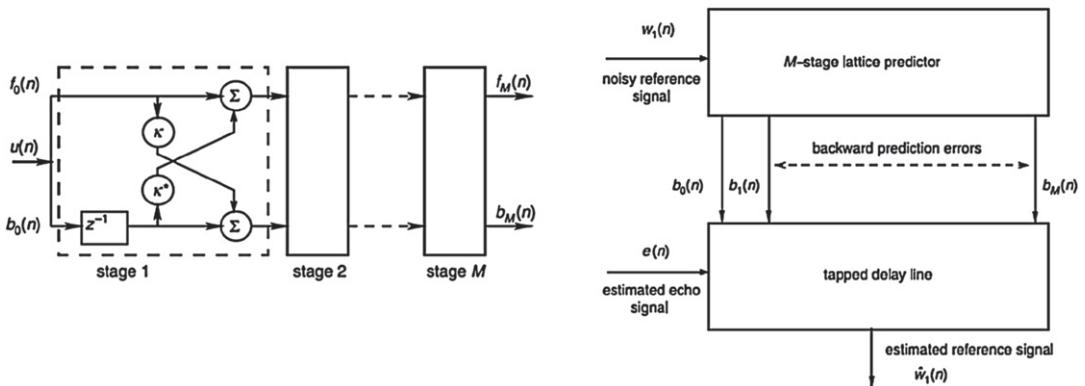
Noise levels in a passive radar receiver in FM radio band: (a) thermal noise in 50 kHz bandwidth corresponding to 5 dB receiver noise figure; top of screen = -50 dBm (b) signal and noise levels measured from 10th floor of UCL in Central London on vertically-polarized dipole antenna; top of screen = -40 dBm .

2.16.7.1 Suppression techniques

A range of techniques may be used to attempt to suppress these signals and interference. First, and most attractive because of effectiveness, simplicity, and cost, is to site the receive antenna so that it is physically shielded from the direct path signal, using topography, buildings or shrouds. This technique alone can often provide adequate suppression, but in turn limits the receive antenna's field of view (FOV). In some cases, this trade-off is entirely acceptable. For air surveillance, terrain blockage of the receiver's FOV is nearly always unacceptable, especially for low-altitude surveillance. However, judicious siting of the receive antenna in some geometries can still provide acceptable physical shielding. The classic example is to site the receiver (a) between the transmitter and the surveillance region with the transmitter looking *over the receiver's shoulder* and (b) with a large building or other structure located directly behind the receiver, i.e., along the baseline, to shield the direct path. As an extreme example, in the case of the Manastash Ridge Radar [8,9] the receiver was located remotely, behind a mountain range and at a distance of 150 km from the transmitter, which provided more than adequate suppression of the direct transmitter signal.

A second technique is spatial cancellation of the direct path signal. An array antenna at the receiver can be configured to steer a null at the direct path signal, and null depths of several tens of dB are achievable. Inspection of Figure 16.10 shows that radiation pattern nulls should ideally be broadband (in direction and depth) so that multiple signals from a given transmitter may be suppressed by a single null. Of course, an adaptive antenna and its associated signal processing introduces significant complexity, and, to some extent, negates one of the claimed advantages of PBR—that the receiver is simple and cheap.

A third approach is to use a transversal filter equaliser method to suppress multipath in a similar way to that used in mobile communications. Two architectures that have been used with success for this purpose have been described by Howland et al. [46] and by Colone et al. [47]. The former uses

**FIGURE 16.12**

Two-stage adaptive noise canceller used by Howland [46]. The first stage (a) is an adaptive M -stage lattice predictor with prediction order $M = 50$, equivalent to the Gram-Schmidt algorithm, and the second (b) is equivalent to a multiple regression filter.

a two-stage adaptive noise canceller, of which the first stage is an adaptive M -stage lattice predictor, with prediction order $M = 50$, and the second an adaptive tapped delay line (Figure 16.12). The first is equivalent to the Gram-Schmidt algorithm and the second to a multiple regression filter. This approach was able to suppress the direct path interference by approximately 75 dB.

The approach of Colone et al. [47] may be understood by visualizing the signal and multipath environment in the range-Doppler plane (as in Figure 16.14). It is assumed that a clean reference version of the direct signal is obtained via a separate antenna and receiver which is free of multipath. The algorithm successively nulls the multipath signals via appropriately-weighted and delayed versions of the direct signal. In the example given the algorithm was halted when 60 dB of cancellation had been achieved.

Filtering techniques based on canceling the direct signal and its multipath using a reference version of the direct signal will, of course, not be able to suppress other co-channel transmissions or noise sources. In practice a combination of techniques will be necessary, and even then it is unlikely to be possible to suppress the multipath and interference right down to the level of receiver thermal noise, so this will need to be taken into account in performance prediction calculations. Furthermore, to reduce the dynamic range of the signal to be digitized, it may be helpful to employ a stage of analog adaptive null steering, of the type described in [48], prior to digitization.

2.16.7.2 Performance prediction

As an example of such a performance prediction, the bistatic radar equation (16.5) can be written in the form

$$(R_R)_{\max} = \left(\frac{\Phi \sigma_B G_r \lambda^2 L G_p}{(4\pi)^2 (S/N)_{\min} k T_0 B F} \right)^{1/2} \quad (16.13)$$

where the appropriate value of Φ is used according to the illuminator source and the transmitter-to-target range, $(S/N)_{\min}$ is the minimum signal-to-noise ratio for adequate probability of detection, and

G_p is the processing gain. The maximum value of this is set by the maximum time T_{\max} for which the target echo remains coherent, which in turn depends on the extent to which the target is maneuvering or accelerating. A rule of thumb value for this time, obtained by setting the change in one-way pathlength due to the radial component of target acceleration A_R to be equal to half a wavelength, is

$$\frac{A_R T_{\max}^2}{2} = \frac{\lambda}{2} \quad (16.14)$$

hence

$$T_{\max} = \left(\frac{\lambda}{A_R} \right)^{1/2} \quad (16.15)$$

For a VHF FM radio waveform of bandwidth 50 kHz and an integration time of 1 s, this gives $G_p = 47$ dB. Adopting a suitable value of F (perhaps 25 dB) to take account of the finite suppression of the direct signal, multipath and other signals, and inserting appropriate values for Φ , σ_B and for G_r , allows contours of detection sensitivity to be plotted, and an example of this is shown in Figure 16.13. These have the “Ovals of Cassini” form discussed in Section 2.16.3.

It can be seen that aircraft targets of this kind should be detectable at ranges well in excess of 100 km, and results have been reported in the literature from several practical systems [45, 49, 50] which confirm this. An example showing aircraft detections of civilian aircraft targets at slightly closer ranges, presented in terms of bistatic range—Doppler plots, is shown in Figure 16.14.

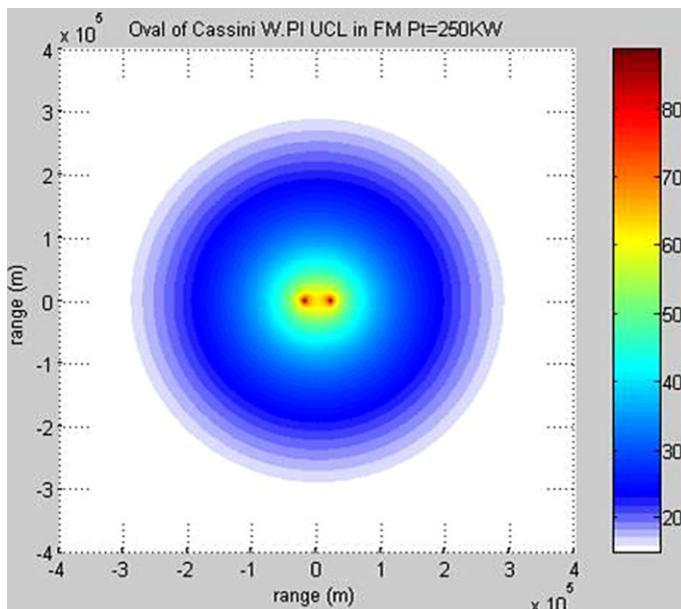
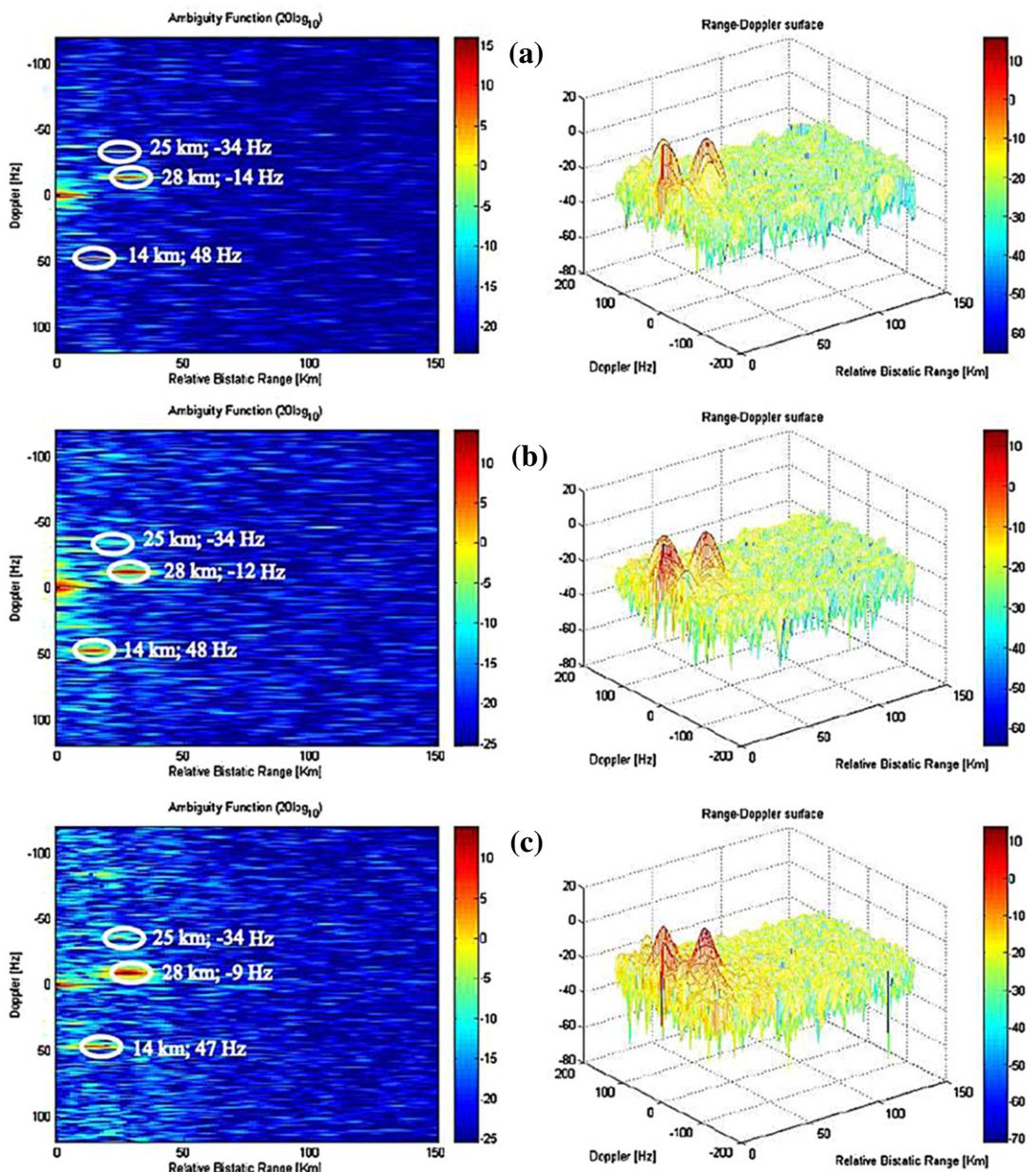
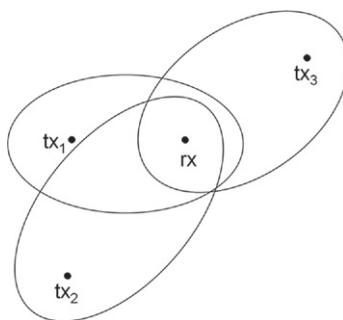


FIGURE 16.13

Predicted detection range performance of a VHF FM-based PBR receiver system, with $P_t = 250$ kW, $\sigma_B = 100$ m 2 , $F = 25$ dB [49].

**FIGURE 16.14**

2-D bistatic range-Doppler display plots for the BBC 91.3 MHz Wrotham FM transmitter. (a)–(c) show three successive 1 s blocks of data [52].

**FIGURE 16.15**

Triangulation using three transmit-receive pairs. The differential range at each transmitter-receiver pair defines an ellipse, and the target is located at the point where the three ellipses intersect.

2.16.7.3 Multilateration and tracking

In practice the processing required to form reliable tracks of multiple targets is more difficult, at least partly due to the dependence of the ambiguity function on instantaneous modulation (for analog transmissions) and on bistatic geometry. Use of multiple transmitters and/or receivers can be useful, since for a given target the dependence of ambiguity functions on instantaneous modulation and on geometry is deterministic, so the transmissions that give inadequate ambiguity performance can be ignored.

2.16.7.3.1 Multilateration

Figure 16.15 shows how multilateration can be used to determine the location of a target. In this example there is a single receiver and three transmitters, all with omnidirectional antennas. Each transmitter-receiver pair provides a measurement of bistatic range from the time difference of arrival (TDOA) of the echo signal with respect to the direct signal. Each of these defines an ellipse, and the three ellipses intersect at various points, of which only one is the true target location. The false locations (which have been termed “Casper’s ghosts” [53]) need to be identified and excised (a process known, inevitably, as “ghostbusting”).

2.16.7.3.2 Extended Kalman filter (EKF) tracking

Differential range is not the only information available at the bistatic receiver. In principle, Doppler shift and direction of arrival (DOA) of each target echo may also be measured, so these pieces of information can be exploited as well. One of the early PBR experiments by Howland [5] making use of the vision carrier of analog television transmissions, used a single receiver to measure Doppler and DOA (but not differential range). This approach has been termed narrowband PBR, since the detection bandwidth is narrow (and hence the sensitivity is improved) and only small a portion of the total signal spectrum is used. The receiver was located at Pershore, UK, and the transmitter was at Crystal Palace in the south of London, at a range of some 150 km from Pershore. The Doppler and DOA measurements were input to an extended Kalman filter (EKF) algorithm to form tracks of the detected targets.

The conversion of Doppler and DOA measurements into target state estimates was improved by initializing the process with a genetic algorithm. Howland found that while nearly all high-altitude targets were detected, only about one third could be tracked, with the remaining two thirds either lost in the CFAR and initial Kalman filtering, or having ambiguous or adequately-accurate bearing estimates, which underlines the difficulty of target tracking with PBR, especially if only a single transmit-receive pair is used.

2.16.7.3.3 Probability hypothesis density (PHD) filtering and tracking

Tobias and Lanterman [54] tackled the problem of ghost excision and target state estimation using the probability hypothesis density (PHD) approach, which was originally developed by Mahler [55]. The PHD is defined as being any function that, when integrated over any given area, specifies the expected number of targets present in the area.

They use the particle filter implementation of the update equations in which the PHD is represented by a collection of particles and their corresponding weights. Using the same notation as [54], at time-step k each particle in the filter is a vector of the form

$$\xi_i = [x_i \ y_i \ \dot{x}_i \ \dot{y}_i]^T \quad (16.16)$$

and has a weight $w_{i,k}$, where $(x_i \ y_i)$ specify the location of the particle and $(\dot{x}_i \ \dot{y}_i)$ specify its velocity components. As per the defining property of the PHD,

$$\tilde{N} = E[\text{no. of targets}] = [N_{k|k}]_{\text{nearest integer}}, \quad (16.17)$$

where

$$N_{k|k} = \sum_i w_{i,k} \quad (16.18)$$

Specifically, the PHD is expected to (a) automatically estimate the number of targets, (b) resolve ghost targets, and (c) fuse sensor (i.e., bistatic transmit-receive pair) data without the need for any explicit report-to-track association [56].

Results were presented from a simulation using three bistatic transmit-receive pairs measuring first range and then range/Doppler on two aircraft targets flying in the Washington DC area. The transmitters were three local VHF FM stations and the receiver was based on that used by Lockheed Martin's Silent Sentry™, located 30–50 km from the transmitters. The simulations assumed adequate target visibility, overlapping coverage, and no multipath, and calculated SNRs ranged from 12.2 dB to 32.5 dB.

In its simplest form, the simulation began by independently and randomly assigning the particles' two-dimensional position and velocity components to fall within the FOV of each transmit-receive pair. Particle weights were initially set to zero. These particles were then propagated forward in one-second steps. Birth particles with random positions and velocities were added at each time step to model new targets. One new target and hence one birth particle was assumed to appear at each step. The PHD then assigned (and updated) particle weights $w_{i,k+1}$ at each time step by incorporating range/Doppler observations, calculated probability of detection, Poisson-distributed false alarms, and a single-target likelihood function. Finally, the expected number of targets in the FOV was calculated by means of Eq. (16.18). The locations of the N expected targets were found by extracting the N highest peaks from the PHD represented by these weights.

The results of these preliminary simulations were described as “encouraging.” It was observed that in areas of low SNR the number of targets was overestimated. Subsequently an improved method was developed which removed the need to restrict particles to areas of high SNR, though at the expense of greater computational load [56]. This also had the effect of reducing the number of particles needed from a few thousand to a few hundred.

2.16.8 Applications

In this section, a presentation of practical applications of PBR is provided. We start by acknowledging that although PBR is a vibrant and lively subject with a steady stream of new publications in research journals and at conferences, there have been rather few applications in which PBR systems have been operationally deployed, as distinct from experimental demonstration.

2.16.8.1 Low-cost scientific measurements

The principal applications in which PBR has been deployed thus far have been for low-cost scientific remote sensing measurements. As a general comment, radar remote sensing relies on a *model* to relate the desired remotely-sensed quantity to some parameter of the radar echo. The model may be the result of rigorous electromagnetic scattering calculations, or empirical measurements, or a mixture of the two. An example of this is the relationship in a satellite-borne radar altimeter between the duration of the leading edge of the echo from the ocean surface and the wave height. This will work best when the relationship between the remotely-sensed quantity and the parameter of the radar echo is maintained over a wide dynamic range, and certainly the relationship must be monotonic. The extra degrees of freedom associated with the bistatic geometry may offer a better chance of finding a regime where such a relationship exists.

2.16.8.1.1 The Manastash Ridge Radar

As has already been mentioned, examples of low-cost remote sensing applications have included measurements of the ionosphere, planets, wind or thunderstorms. The first of these was developed and demonstrated by Sahr and his co-workers at the University of Washington, Seattle, in a system known as the Manastash Ridge Radar (MRR), as a means of studying turbulence in the ionosphere. Although dedicated monostatic radars have been successfully used for this application, PBR has the advantages (especially for a university group) of simplicity and low cost, as well as avoiding the complication and safety issues associated with a dedicated transmitter. Furthermore, frequencies around 100 MHz (using VHF FM broadcast transmissions) are very suitable for this purpose. Two noteworthy features were the separation of the receiver behind a range of mountains to suppress the direct signal in the receiver (already identified as an issue in Section 2.16.7), and the use of GPS and the Internet for timing and data transfer. These resulted in a low-cost but high-performance system.

The resulting data products approximate those from the current monostatic radars, in terms of fine range, time, and Doppler resolution. Ground clutter and aircraft (at shorter range), and meteor trails are also routinely detected. The data was made available in the form of range-Doppler plots (out to ranges of 1200 km) every half hour via the Internet. A photograph of the receiver, and an example of the data product, are shown in Figure 16.16a and b respectively.

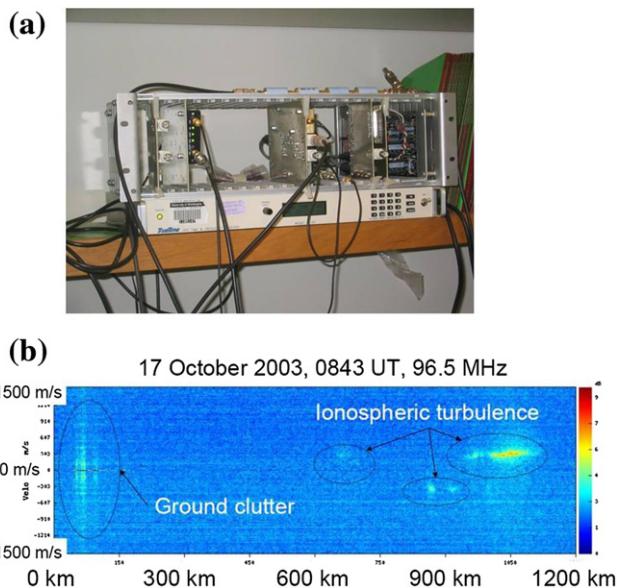


FIGURE 16.16

(a) PBR receiver for the Manastash Ridge Radar; (b) example of the data product, made available on the Internet every 30 min, and consisting of 10 s of data sampled every 4 min.

2.16.8.1.2 Bistatic adjunct to weather radar

A second such application is a bistatic adjunct to the NEXRAD meteorological radars in the USA to measure the vector wind field [12,57]. This approach has the following advantages and disadvantages compared to the conventional dual- or multiple-Doppler method using two or more monostatic radars:

Advantages:

- (1) Simultaneous measurement of all Doppler velocities in individual weather volumes, thus minimizing storm evolution-induced errors, and (2) less expensive and simpler installation by using only one transmitter and scanning antenna. The network can also be used for (3) hail detection by comparing the intensity of obliquely (bistatically) scattered energy to Rayleigh scattering predictions and (4) rapid scanning of localized weather, which typically requires multiple monostatic weather radars.

Disadvantages:

- (1) Less sensitivity to weak weather returns due to the low-gain receiving antenna; (2) less sensitivity to weather returns near ground level when using horizontal transmit polarization, which are used by many Doppler weather radars; (3) more sensitivity to contamination from transmitter sidelobes (i.e., direct path breakthrough) and to secondary scattering from weather echoes via transmit antenna sidelobe illumination, both due to the broad-beam receiving antenna; and (4)

increased measurement errors by about a factor of 2 when operating at large oblique scattering angles. Contamination is acceptable in most situations and can be reduced by using higher-gain antennas when necessary. The lower sensitivity and increased errors can be offset with more BNR sites, which represents a practical solution due to their low cost.

Between 1993 (when the first BNR prototype was introduced) and 2002, nine sets of BNRs have been sold and deployed at monostatic Doppler weather radar sites for research, testing, and operation: one in the United Kingdom, one in Canada, one in Germany, two in Japan, and four in the United States.

2.16.8.1.3 Ocean remote sensing using GPS signals

A third application is the use of GPS satellite signals scattered from the ocean surface to measure surface winds [58]. The power density at the target (Table 16.1) from such illumination sources is low, but since the phenomena being sensed vary on relatively long timescales, long integration times can be used.

This idea was first put forward, along with some preliminary experimental measurements with an aircraft-borne receiver, by Garrison et al. [59], and came from a realization that the scattered multipath GPS signal from the sea surface that for conventional GPS purposes is a nuisance actually contains useful information about the sea surface roughness, and hence the wind speed (yet another example of the principle that “one person’s interference is another person’s signal”). Specifically, the width of the cross-correlation function between the scattered signal and the locally generated PRN code provides a measure of the surface wind speed. Subsequent work [60,61] has carried out more detailed experiments, comparing the results with other satellite remote sensing data and with surface truth measurements from buoys, and confirmed the viability of the technique as a simple, low-cost approach to ocean remote sensing. Of course, signals from the Russian GLONASS and the European GALILEO systems are equally usable in this application.

2.16.8.2 Border or perimeter surveillance

Forward scatter fences based on PBR illuminators may be used for border or perimeter surveillance, to detect intruders or smugglers, or for maritime surveillance, for example for harbor protection, fisheries protection or counter-piracy. Success here will depend on the availability of illuminating sources with appropriate coverage, and suitable locations for receivers (whether on land or at sea), as well as understanding the forward scatter signature of maritime targets and of sea clutter. It will exploit the significant enhancement of target RCS that occurs in forward scatter, but will also need to take account of the high clutter level in forward scatter, so the performance is likely to be clutter-limited rather than noise-limited.

It may also be desirable to attempt to exploit the micro-Doppler signature of targets, so as to be able to distinguish between different types of target (for example, between a human and an animal, between a human that is walking and one that is running, or between a human carrying a weapon and one without) on the basis of the Doppler modulation due to the swinging motion of arms and legs [62]. In the same way, a tracked vehicle may be distinguished from a wheeled one. The poor Doppler resolution associated with the forward scatter geometry means that this technique is not compatible with forward scatter, but for other geometries it may well be usable. Greneker [63] has demonstrated a PBR system based on HDTV signals around 548 MHz, with a bandwidth of 6 MHz, and detected and characterized the micro-Doppler signatures of a walking human target and of an accelerating vehicle.

2.16.8.3 Air surveillance

Air surveillance has already been mentioned as a potential application of PBR, and examples of experimental results from a number of systems have been shown above. These generally exploit high-power analog or digital broadcast transmissions, which have appropriate power and coverage [46, 50–52].

Some other results of this kind were presented as early as 1995 by Koch and Westphal, who used transmissions from GPS satellites to detect a variety of air targets including military and civil aircraft, an anti-tank missile, and the MIR Space Station [64, 65]. As with remote sensing applications using GPS signals, the power density at the target is comparatively low. The range from the GPS satellite to the target is of order 20,000 km, and the signal level at the input to a GPS receiver is of order –160 dBW. However, the experiments exploited long integration times (several seconds, giving 60–70 dB of processing gain) and the substantial enhancement of target RCS in forward scatter (Figure 16.3) to compensate.

2.16.8.4 Airborne PBR

Another potentially valuable set of applications comes if the PBR receiver is located in an aircraft. This might allow (for example) airborne early warning (AEW) or detection of ground vehicle targets. Some first results from a system of this kind, developed by Woodbridge and co-workers at University College London (UCL) have been reported in [66, 67].

For these experiments a two-channel VHF receiver was mounted in a Piper PA 28-181 light aircraft with simple dipole antennas for the reference channel and signal channels taped to the inside of the window (Figure 16.17), flying from Shoreham airport on the south coast of England. An issue with such a system is that the receivers will certainly be line-of-sight to the transmitters, so the degree of suppression of the direct signal in the radar receiver may need to be even higher than for a terrestrial system.

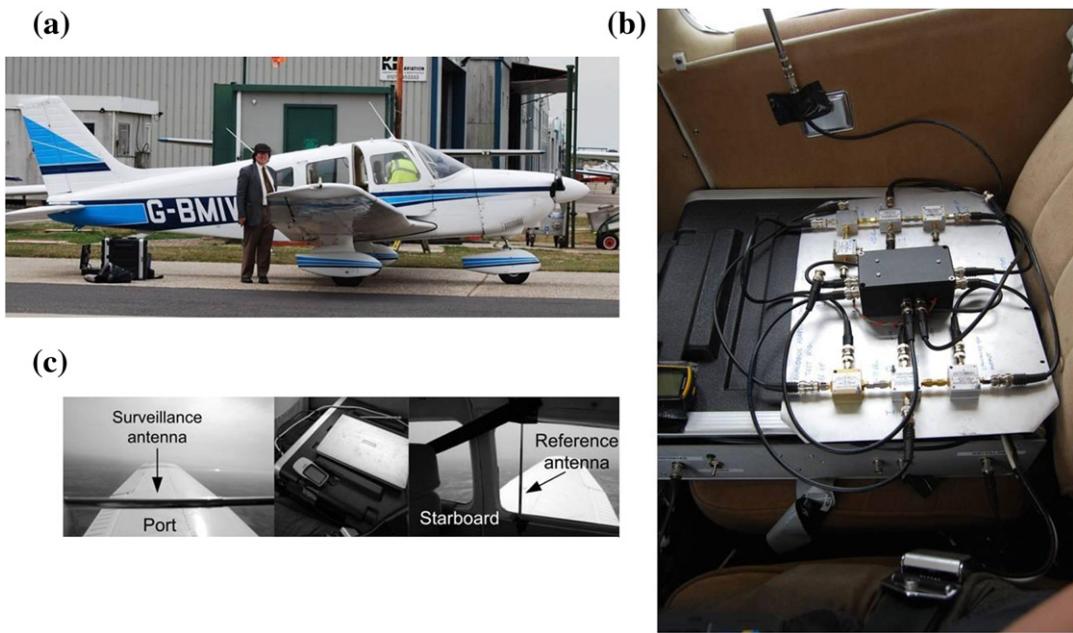
Figure 16.18 shows the results of detection of civil aircraft targets using VHF FM transmitters at Wrotham, Crystal Palace, Guildford and Oxford, giving in each case (from the differential TDOA) an ellipse on which the target must lie. The target velocity vector is also computed from the measured Doppler shift and the known velocity of the aircraft carrying the receiver, and these are shown around the locus of the ellipses. This information can be used to resolve the correct target location—in other words to excise the ghosts—since it can be seen that the velocity vectors agree at one intersection point but not the other.

Although these results must be regarded as preliminary, they indicate the potential of the idea and point the way for further work.

2.16.8.5 PBR imaging

PBR may also be used as the basis for bistatic radar imaging. The synthetic aperture may be formed by a moving transmitter or a moving receiver, or if both transmitter and receiver are fixed, by the motion of the target itself (bistatic ISAR). Bistatic images and High Resolution Range Profiles (HRRPs) may be used in target classification and Non-Cooperative Target Recognition (NCTR) processing, just as with monostatic imaging.

There have been a number of experiments using satellite-borne illuminators of opportunity as the basis for bistatic SAR. The azimuth resolution of such a system may be derived in the same way as for a monostatic SAR, recognizing that the phase history of the sequence of echoes is determined by

**FIGURE 16.17**

Airborne PBR experiments: (a) Piper PA 28-181 aircraft, (b) receiver system, (c) antennas (photos: James Brown).

the change in transmitter-to-target path (if the receiver is stationary) or target-to-receiver path (if the transmitter is stationary) rather than the two-way path in a monostatic SAR. Thus in general the azimuth resolution is a factor 2 coarser than the equivalent monostatic SAR.

An example of an image obtained in a system of this kind is shown in Figure 16.19. The illuminator in this case is the European Space Agency's ASAR synthetic aperture radar carried by the ENVISAT satellite, whose principal parameters are given in Table 16.2.

The receiver system consisted of two channels: one pointed upwards for the direct signal from the satellite and one for the echoes from the target scene [68,69]. The antenna used for the target scene channel was a 1.2 m diameter parabolic dish, which gave a gain of approximately 35 dBi and a beamwidth of approximately 3.5° (which limits the angular field of view in the image). The image shows that short-range imaging of this kind is certainly feasible, but the practical applicability is limited by the satellite orbit, such that the target scene is illuminated for approximately 1 s with a repeat interval of the order of days.

2.16.8.6 Short-range surveillance using WiFi and WiMAX

Another class of signal that has received significant attention for short-range surveillance using PBR are the wireless transmissions for WiFi Local Area Networks (LANs)—IEEE Std 802.11 [70–72] and

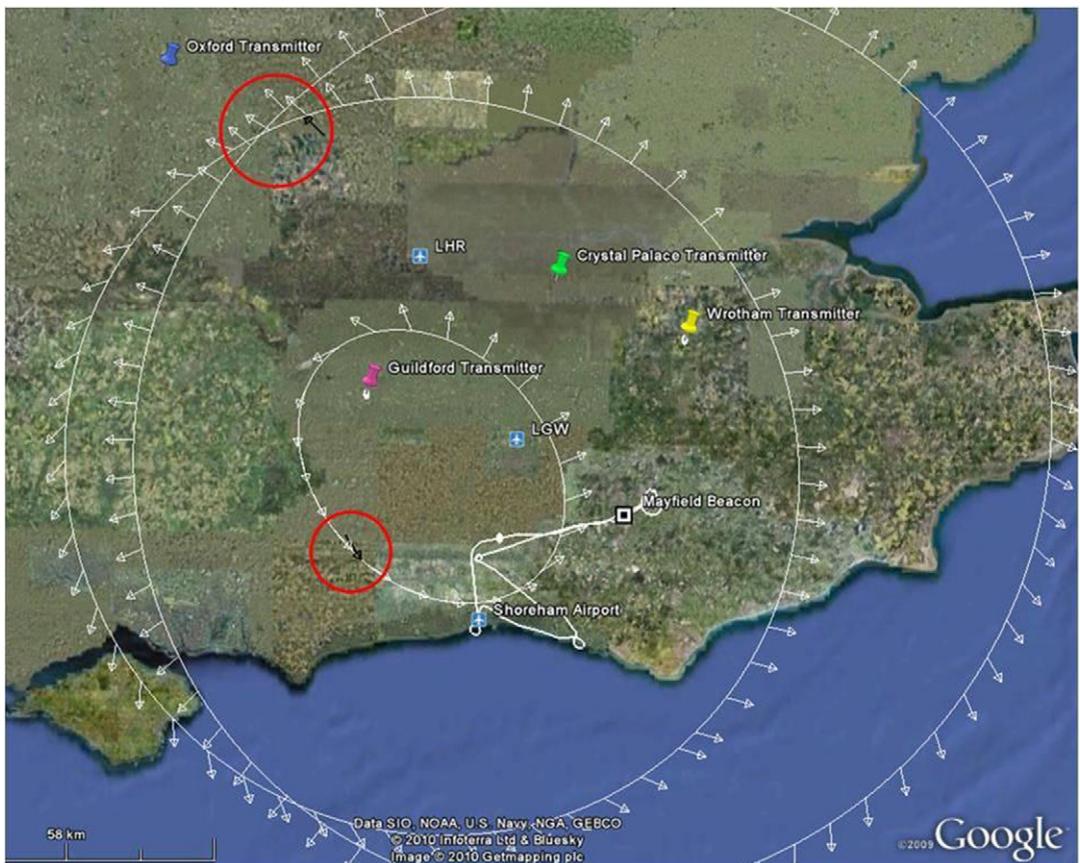


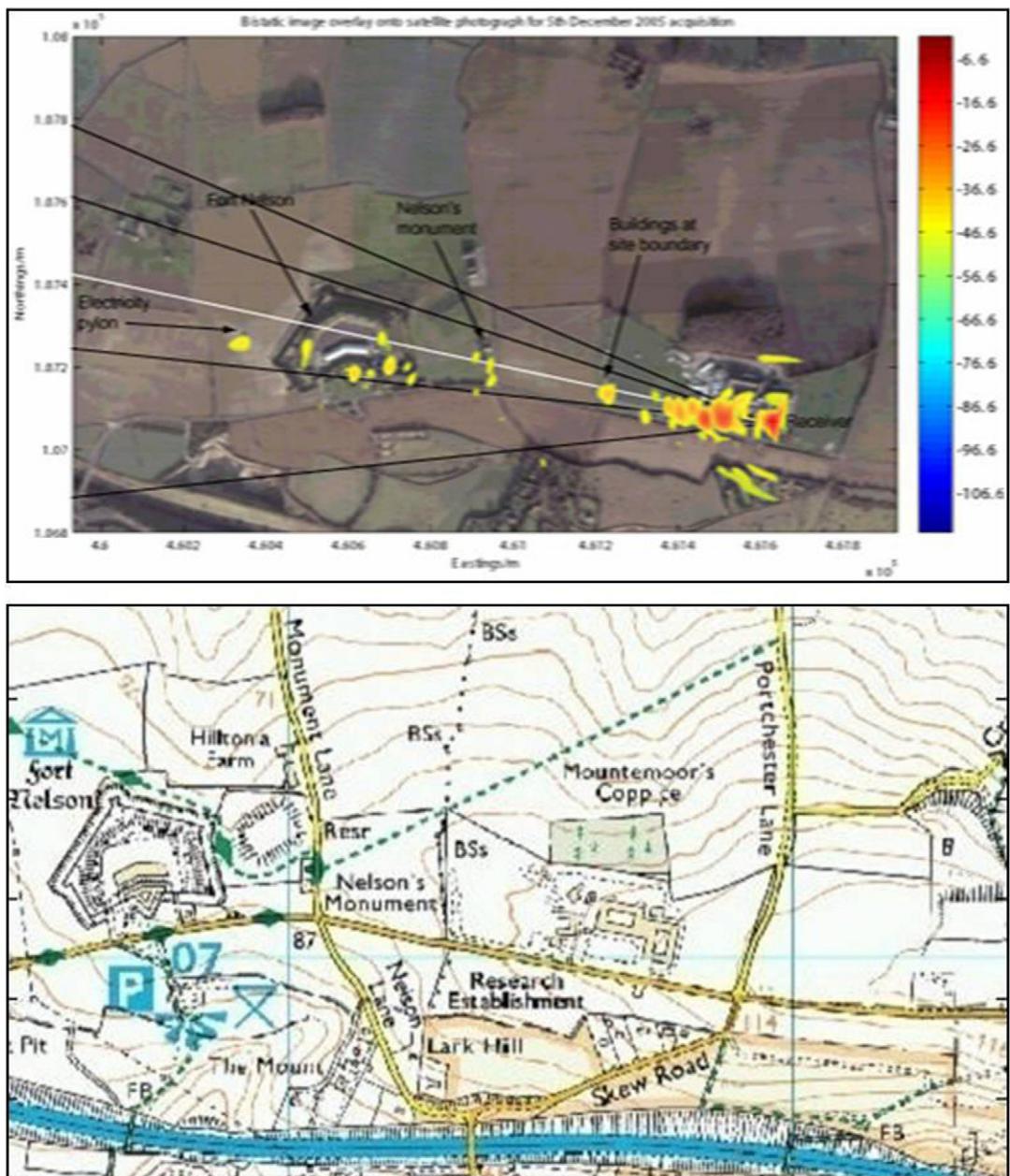
FIGURE 16.18

Results from trials with UCL aircraft-borne PBR receiver. The ambiguities associated with the intersection of the differential range ellipses may be resolved using the target velocity vectors.

WiMAX Metropolitan Area Networks (MANs)—IEEE Std 802.16 [73–75]. The WiFi standard is low-power and short-range, intended primarily for indoor use, and therefore potentially usable for surveillance within buildings or for short-range outdoor applications; the WiMAX standard provides broader coverage (up to several tens of km) so may be useful for applications such as port or harbor surveillance.

The overall IEEE 802.11 WiFi standard uses either Direct Sequence Spread Spectrum (DSSS) or Orthogonal Frequency Division Multiplexing (OFDM). The 802.11b and 802.11g standards operate in the 2.4 GHz band, while 802.11a uses the 5 GHz band. The transmitters tend to use dynamic power management according to the number of users, but a maximum value for the transmit power might be 100 mW.

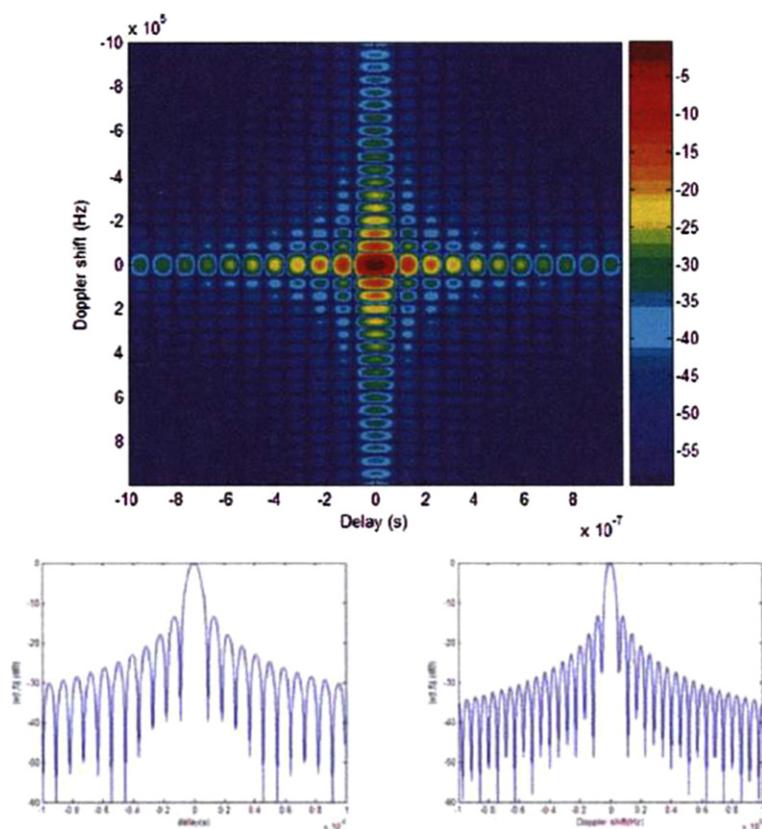
A typical ambiguity function, showing cuts at zero range and zero Doppler, is given in Figure 16.20. It can be seen that the ambiguity function is well behaved. The range resolution is of the order of 30 m

**FIGURE 16.19**

Experimental results from satellite bistatic experiments: imaged targets overlaid on aerial photograph (upper); corresponding map data (lower).

Table 16.2 Principal Parameters of ASAR Radar and ENVISAT Satellite

Frequency	5.331 GHz
Antenna	320-element electronically-scanned array, 10 m × 1.3 m
Transmit power	1365 W
PRF	1650–2100 Hz
Pulse bandwidth	Up to 15 MHz
Pulse length	20 µs
Polarization	VV or HH
Swath width	Up to 100 km
Orbit	Low Earth Orbit (LEO) Sun-synchronous, 800 km mean altitude, 1, 3, or 35-day repeat cycle

**FIGURE 16.20**

Ambiguity function for WiFi 802.11b preamble/header signal, and cuts at zero range and zero Doppler [76].

Table 16.3 Experimental Parameters

Transmit power	+6 dBm
transmit antenna gain	0 dBi
Receive antenna gain	18 dBi
Wavelength	0.123 m
Target RCS	Variable
Transmitter-to-target range	2.2 m
Target-to-receiver range	1–4 m
Total losses	11 dB
Effective bandwidth	11 MHz
Receiver noise Figure	3 dB

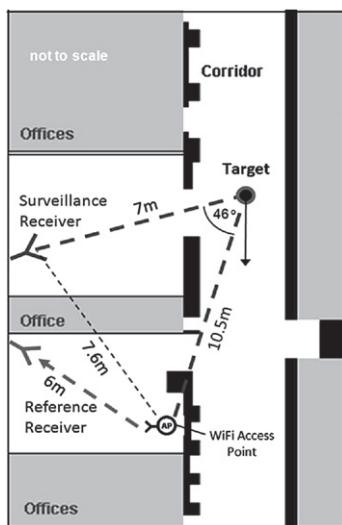
which is rather coarse for indoor applications, but the Doppler resolution can be quite good, particularly if long integration times are used, suggesting that micro-Doppler information on echoes from moving targets, such as walking humans, may be extracted and exploited.

The first reported example of such a system is in [70], using an 802.11 wireless network transmitter as the illumination source in a simple short-range configuration in an indoor environment. Table 16.3 shows some of the operating parameters of the system.

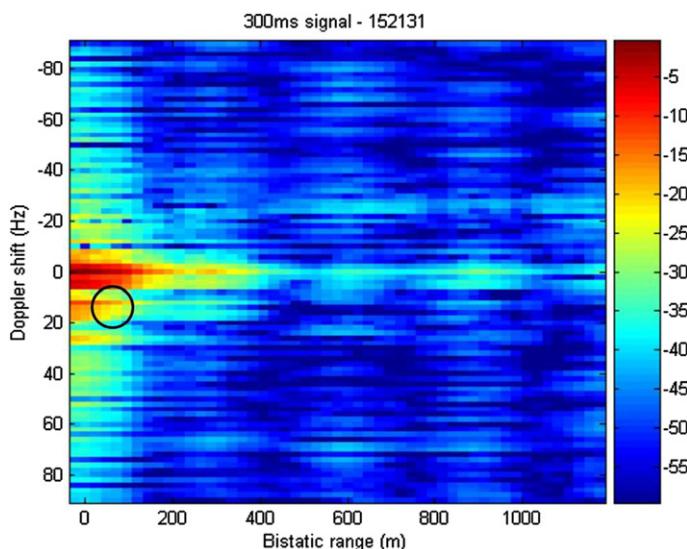
The results show power levels consistent with a simple application of the bistatic radar equation. This form of system, and in particular the type of waveform transmissions and their radar properties, have been analyzed in more detail in [76]. Here the WiFi transmission sequence has been found to be quite sophisticated and dependent on the user environment, but is dominated by the Direct Sequence Spread Spectrum (DSSS) and Orthogonal Frequency Division Multiplexing (OFDM) signals. Ambiguity function analysis of the DSSS based WiFi signal shows that range and Doppler performance is comparable with that expected from theoretical predictions. Detection of moving human targets was demonstrated for the first time. This work shows that this technique has considerable promise for a low cost and widely deployable short-range detection and tracking system.

In [72] the passive bistatic WiFi radar approach is explored further to determine performance limits, by examining the detection of targets in a dense clutter background. Through-wall detections of personnel targets moving at differing velocities within an indoor environment were presented for the first time in this publication. Figure 16.21 shows the geometry employed. As an example, Figure 16.22 shows a “raw” range Doppler map integrated over a time interval of 300 ms in which a target is circled against quite a dense clutter background. The target is corroborated with ground truth and provides a sound basis for subsequent detection and tracking. The complexity of the clutter environment with multiple reflections occurring in both the direct and indirect channels can be appreciated. Overall these papers demonstrate that it is feasible to use WiFi transmissions to make passive bistatic radar systems that can be used over short ranges in both indoor and outdoor scenarios.

As indicated in Section 2.16.6, the “big brother” to WiFi, WiMAX has considerably higher transmission power levels (perhaps 20 W compared to 100 mW) and represents a potentially attractive alternative to DAB and DVB. In the UK there are currently two bands licensed at 3.5 and 3.6 GHz. The bandwidth is variable from 1.25 MHz to 20 MHz and hence higher range resolutions than WiFi are possible.

**FIGURE 16.21**

Schematic of the indoor dense clutter WiFi radar experimental set up.

**FIGURE 16.22**

"Raw" range Doppler map integrated over a time span of 300 ms in which a target is circled against quite a dense clutter background.

This means that imaging techniques start to be more readily addressable, although, to date, application has concentrated on detection.

In [75] Chetty et al. have considered the use of WiMAX transmissions in the problem of providing maritime radar surveillance. Their study shows that a 4000 m^2 target should be detected out to a range of 45 km with a transmitter-to-receiver baseline of between 10 and 32 km. This level of coverage is suitable for busy commercial maritime areas such as that covered by the Port of London. Range and Doppler resolution are calculated to be 5 m and 3.5 Hz respectively which is consistent with parameters appropriate for the detection of marine vessels. They conclude by recommending field trials of an experimental system.

2.16.8.7 PBR using HF skywave signals

A final application is long-range surveillance using over-the-horizon (OTH) HF signals. Skywaves can propagate over very long distances around the curvature of the Earth by means of a series of reflections and refractions occurring successively between the Earth's surface and reflecting layers in the ionosphere (the D-layer at a height of about 60 km, the E-layer at about 100 km and the F-layer from 250 to 400 km). The range provided by one propagation hop is typically from about 1000 km to 3000 km, although it can be as great as 4000 km. Surface wave propagation occurs near the ground for short distances, up to 100 km over land and 300 km over the sea. Attenuation of the wave is a function of the antenna height, polarization, frequency, and surface type. Also, this surface wave decays with distance more rapidly than the normal inverse square law of line-of-sight propagation [77].

Such frequencies are used for long-range communication (though less so now than in their heyday in the twentieth century), for HF broadcast, and for HF OTH radar. The ionosphere is a region of ionized gas (plasma) that surrounds the Earth at a height of 50–2000 km. The ionization is a consequence of solar UV and X-ray radiation, and hence is a function of the time of day (since the ions recombine during darkness), season of the year (since there are more hours of daylight during summer) and of solar activity, which varies on a period of approximately 11 years, known as the sunspot cycle. The distribution of ionization is modified by the Earth's magnetic field, generally enhancing and raising the ion content near the equator. The coverage therefore depends on all of these factors, and (most importantly) on frequency, so it is necessary to choose the correct frequency to provide the desired coverage at any given date and time of day.

In general, HF broadcast signals will be of low bandwidth (for example, a DSB AM broadcast signal will typically have a bandwidth of about 12 kHz, corresponding to a radar range resolution of 12.5 km). Also, although HF transmitters may use physically large antennas, their beamwidths at such low frequencies are relatively broad, and their footprints at such long ranges are correspondingly broad. The resolution both in range and in azimuth, when such signals are used for radar purposes, is therefore coarse. Another consequence of the low radar frequency is that the Doppler shift for a given target velocity is lower than would be the case at higher radar frequencies.

As with other forms of broadcast signals used for PBR, digital modulation formats, such as the Digital Radio Mondiale (DRM) standard, gives better performance than analog signals, since their ambiguity performance is constant with time and does not depend on the program material [78].

Bistatic radar experiments have been conducted using HF OTH radars (such as the Australian JORN network [83] or the US ROTHR (AN/TPS-71) systems [84] in Virginia, Puerto Rico, and Texas) as

illumination sources, for applications such as detection of air and maritime targets. The geometry employed here is usually with the receiver close to the target scene, and receiving both the direct (sky-wave) signal from the transmitter and the target echoes.

Another variation is to use HF radio broadcast signals (such as the BBC's World Service or the US Voice of America) as the illumination source. A website [85] shows results obtained from some amateur radio experiments in which a communications receiver listens to a CW carrier from a transmitter at a frequency of about 26 MHz and at a range from the receiver of approximately 100 km. The audio output from the receiver is digitized and displayed in the form of a spectrogram (frequency vs time). Typical results show the carrier over a period of several minutes, as well as Doppler-shifted echoes from aircraft undergoing particular maneuvers. The author of the website states that this processing is very straightforward and urges: "do try this at home"; the cost of the hardware (essentially just a radio receiver and a PC) is low and the software to perform the processing can be downloaded from the website.

Lesturgie and Poullin [79] have published results of experiments using an HF broadcast signal in Kiev, Ukraine and a receiver located either on a ship or on the shore, detecting aircraft targets at ranges typically from 50 to 200 km.

All of these results indicate that PBR using HF skywave illuminators has considerable potential for target detection and tracking, at ranges much greater than achieved by conventional radars.

2.16.9 Open issues and problems

We provide here a list of open issues and problems, in addition to those already alluded to above.

2.16.9.1 Identifying applications

As noted in the introduction to this chapter, one of the key open issues with PBR is to identify appropriate applications in which there is a demonstrable advantage over conventional approaches in terms of performance and/or cost. To those listed in the introduction and in the previous section we might add use as an adjunct to ATC (Air Traffic Control) radar in regions where coverage is poor. This is under active investigation, but it has yet to be shown to be 100% reliable.

2.16.9.2 Detection of and countermeasures against PBR

One of the claimed advantages of PBR in a defense/security context is that it is potentially completely covert. As long as the receive antennas are inconspicuous, an adversary will have no indication that a PBR system has been deployed. Turning that on its head, there will be no indication that a PBR system may have been deployed by an adversary against our own assets. A British WW2 intelligence document [80] describing the German *Klein Heidelberg* bistatic radar [81], which used the British Chain Home radar transmissions as its illumination source, considers countermeasures against a bistatic radar of this kind, and lists four possibilities:

- i. Physical attack of the receiver (if its location is known);
- ii. Turn off, or otherwise interrupt the transmitter;
- iii. Alter the form of the transmitted signal to make synchronization more difficult;

- iv. Use a repeater jammer to generate multiple false targets which will overload and confuse the detection and tracking processing.

Nearly 70 years later, these approaches seem just as relevant.

2.16.9.3 Characterization of bistatic radar clutter

It has already been remarked that bistatic radar clutter is poorly understood, both in terms of its mean reflectivity and of its statistical properties. There is therefore great scope for conducting trials to make measurements of bistatic radar clutter, as a function of surface type (sea, land, etc.) and conditions (wave height, wind speed, etc.), bistatic geometry, frequency and polarization, resolution, and hence for developing models both for the mean clutter RCS and for the amplitude statistics, as a function of these parameters. The experimental data should also be useful in developing a better understanding of the physical scattering mechanisms involved.

In the case of PBR, and particularly for urban environments, it should be possible to use the wealth of data and analysis that has been gathered of mobile radio multipath propagation to model the properties of PBR clutter. This should in turn allow improved direct signal and multipath suppression algorithms to be designed and evaluated.

2.16.10 Conclusions and future trends

This chapter has attempted to provide an account of the fundamentals of passive bistatic radar. We can conclude that the subject has come a long way since the first modern experiments in the early 1980s—and certainly since the first radar experiments using broadcast signals more than 50 years prior to that. We have seen that PBR offers the potential of covert operation with simple and low-cost equipment and with no need for a transmitting licence, and the ability to use parts of the EM spectrum that are not normally available for radar use. The wide variety of broadcast, communications and radionavigation sources, and their excellent spatial coverage and (in many cases) high power gives great scope for PBR. In common with all bistatic radars it may allow advantage to be taken of mechanisms such as forward scatter which significantly enhance the radar signature of targets. The fact that PBR systems can be simple and low cost has meant that they have been very suitable for research by university groups, and there have been numerous publications on the subject.

Against this must be balanced the fact that the waveforms that are used for PBR are not designed explicitly for radar, so their performance in radar applications is not optimal. Care must therefore be taken to understand how best to choose which transmissions to use and how best to process them. It is found that analog modulation formats tend to give time-varying ambiguity performance, but more modern digital modulation formats are much better in this respect. In addition, as with all bistatic radars, the ambiguity performance is a function of the bistatic geometry, so for targets on or close to the bistatic baseline the range and Doppler resolution are both poor, no matter what the waveform.

Since most PBR modulation sources are continuous and high-powered, and operate in bands that are already congested, the level of direct signal and other noise sources is inevitably high, and considerable signal processing effort has to be expended to suppress these signals to allow the target echoes to be reliably detected.

The applications to which PBR systems may be put therefore need careful thought. It is no good claiming to do “almost as well” as conventional radar approaches. It is important to understand very thoroughly the relationship between the desired application (surveillance, remote sensing, etc.) and the performance requirements that follow, and the properties of the illuminator source that might be used—for example, in terms of coverage (spatial and temporal), power level, bandwidth (resolution), integration time (scene stationarity) and so on.

Amongst the applications identified, ones which show particular promise are:

- Scientific measurements (remote sensing) where long integration times can be employed.
- Border or perimeter surveillance (perhaps exploiting forward scatter) and/or protection of critical assets.
- Gap filling in areas where operation of conventional sensors is incomplete.

We may also comment that because the ambiguity functions of PBR illuminators depend fundamentally on bistatic geometry, there will always be regions where the performance of a bistatic sensor is compromised. It makes sense, therefore, to think in multistatic terms rather than purely bistatic.

Finally, given the ever-greater degree of spectral congestion, it may make sense in future deliberately to design the waveforms used for broadcast, communications and radionavigation purposes so that their performance in PBR applications is in some sense optimized, as well as for their primary purpose. The ability of modern digital signal processing to generate precise, wideband waveforms and to vary them adaptively is key to this. Professor Mike Inggs has coined the term “commensal” (literally, “sharing the same table”) for this approach.

Acronyms

Commonly used acronyms:

AM	Amplitude Modulation
ASAR	Advanced Synthetic Aperture Radar
ATC	Air Traffic Control
BBC	British Broadcasting Corporation
BNR	Bistatic Network Radar
CDMA	Code Division Multiple Access
CFAR	Constant False Alarm Rate
DAB	Digital Audio Broadcasting
DBS	Direct Broadcast Satellite
DOA	Direction of Arrival
DRM	Digital Radio Mondiale
DSB	Double Sideband
DSI	Direct Signal Interference
DSSS	Direct Sequence Spread Spectrum
DVB	Digital Video Broadcasting
EIRP	Effective Isotropic Radiated Power

EKF	Extended Kalman Filter
ESA	European Space Agency
FDMA	Frequency Division Multiple Access
FM	Frequency Modulation
FOV	Field of View
GLONASS	<i>GLObal'naya NAVigatsionnaya Sputnikovaya Sistema</i>
GMSK	Gaussian Minimum-Shift Keying
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GSM	<i>Groupe Spéciale Mobile</i> ; Global System for Mobile Communications
HDTV	High Definition Television
HF	High Frequency
HRRP	High Resolution Range Profile
ISAR	Inverse Synthetic Aperture Radar
JORN	Jindalee Operational Radar Network
LAN	Local Area Network
LEO	Low Earth Orbit
MAN	Metropolitan Area Network
MRR	Manastash Ridge Radar
NCTR	Non-Cooperative Target Recognition
NEXRAD	Next Generation Radar
NTSC	National Television System Committee
OFDM	Orthogonal Frequency Division Multiplexing
OTH	Over-The-Horizon
PAL	Phase Alternating Line
PBR	Passive Bistatic Radar
PCL	Passive Coherent Location
PHD	Probability Hypothesis Density
PRN	Pseudo Random Noise
QAM	Quadrature Amplitude Modulation
RCS	Radar Cross Section
ROTHR	Relocatable Over-The-Horizon Radar
SAR	Synthetic Aperture Radar
SECAM	<i>Séquentiel Couleur à Mémoire</i>
SNR	Signal-to-Noise Ratio
TDMA	Time Division Multiple Access
TDOA	Time Difference Of Arrival
UCL	University College London
UHF	Ultra High Frequency
UV	Ultra Violet
VHF	Very High Frequency
WiMAX	Worldwide Interoperability for Microwave Access

Glossary

Bistatic radar	a radar using antennas for transmission and reception at sufficiently different locations that the angles or ranges to the target are significantly different [1]
Clutter diversity	exploitation of the variation of clutter properties as a function of bistatic geometry, so as to be able to choose the optimum bistatic radar geometry to best exploit the variation of target and clutter behavior so as to obtain optimal detection performance
Forward scatter	a bistatic radar configuration in which the target lies on or close to the baseline between the transmitter and receiver
Illuminator of opportunity	a transmitter source whose primary purpose is not part of a bistatic radar, but which is used, opportunistically, by a bistatic radar as its transmitter
Passive bistatic radar (PBR)	a set of techniques using broadcast, communications or radionavigation signals as illumination sources rather than using a dedicated radar transmitter. Other terms that have been used include <i>passive coherent location</i> (PCL), <i>passive and covert radar</i> (PCR), <i>covert radar</i> , <i>non-cooperative radar</i> , <i>broadcast radar</i> , <i>parasitic radar</i> , and <i>opportunistic radar</i> [1]
Waveform diversity	adaptivity of the radar waveform to dynamically optimize the radar performance for the particular scenario and tasks. May also exploit adaptivity in other domains, including the antenna radiation pattern (both on transmit and receive), time domain, frequency domain, coding domain and polarization domain [1]

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/B978-0-12-411597-2.00016-3>.

Acknowledgments

I express my thanks to all of the people from whom I have learned about bistatic radar, and who have helped by providing results and by their discussion and constructive criticism. In particular I would like to mention (in alphabetical order) Chris Baker, James Brown, Mike Cherniakov, Mike Inggs, Krzysztof Kulpa, Aaron Lanterman, Daniel O'Hagan, John Sahr, Simon Watts, Keith Ward, Aric Whitewood, Nick Willis and Karl Woodbridge. Their inspiration and help has been enormously appreciated.

References

- [1] IEEE Standard Radar Definitions, IEEE Std-2008, 26 March 2008.
- [2] R.A. Watson-Watt, Three Steps to Victory, Odhams Press, London, 1957, pp. 107–117 (Chapter 20).

- [3] Science News Letter, April 23, 1938.
- [4] H.D. Griffiths, N.R.W. Long, Television-based bistatic radar, IEE Proc. Pt.-F, 133 (7) (1986) 649–657.
- [5] P.E. Howland, Target tracking using television-based bistatic radar, IEE Proc. Radar Sonar Navig. 146 (3) (1999) 166–174.
- [6] C.J. Baker, H.D. Griffiths, I. Papoutsis, Passive coherent radar systems – Part II: Waveform properties, IEE Proc. Radar Sonar Navig. 152 (3) (2005) 160–168.
- [7] M.C. Wicks, E.L. Mokole, S.D. Blunt, R.S. Schneible, V.J. Amuso (Eds.), Principles of Waveform Diversity and Design, SciTech Publishing Inc., Raleigh, NC, 2010.
- [8] J.D. Sahr, F.D. Lind, The Manastash Ridge radar: a passive bistatic radar for upper atmospheric radio science, Radio Sci. 32 (6) (1997) 2345–2358.
- [9] J.D. Sahr, Passive radar observation of ionospheric turbulence, in: N.J. Willis, H.D. Griffiths (Eds.), Advances in Bistatic Radar, SciTech, Raleigh, NC, 2007 (Chapter 10).
- [10] R.A. Simpson, Spacecraft studies of planetary surfaces using bistatic radar, IEEE Trans. Geosci. Remote Sens. 31 (2) (1993) 465–482.
- [11] J. Wurman, M. Randall, C.L. Frush, E. Loew, C.L. Holloway, Design of a bistatic dual-Doppler radar for retrieving vector winds using one transmitter and a remote low-gain passive receiver, Proc. IEEE 82 (12) (1994) 1861–1871.
- [12] S. Satoh, J. Wurman, Accuracy of wind fields observed by a bistatic Doppler radar network, J. Ocean. Atmos. Technol. 20 (2003) 1077–1091.
- [13] E.F. Greneker, J.L. Geisheimer, The use of passive radar for mapping lightning channels in a thunderstorm, in: IEEE Radar Conference, Huntsville, AL, USA, 5–8 May 2003, pp. 28–33.
- [14] D.J. Bannister, Radar in-fill for Greater Wash area: feasibility study final report, Department for Business, Enterprise and Regulatory Reform, UK, 31 August 2007.
- [15] J. Baniak, G. Baker, A.M. Cunningham, L. Martin, Silent Sentry passive surveillance, Aviation Week and Space Technology, 7 June 1999.
- [16] H.D. Griffiths, L. Cohen, S. Watts, E. Mokole, C.J. Baker, M.C. Wicks, S.D. Blunt, Radar spectrum management and engineering issues: regulatory and technical approaches, IEEE Proc.
- [17] H.L. Van Trees, Detection, Modulation and Estimation Theory, Parts I, II, III, and IV, Wiley, 2001.
- [18] P.M. Woodward, Probability and Information Theory, with Applications to Radar, Pergamon Press, 1953; Reprinted by Artech House, 1980.
- [19] F. Le Chevalier, Principles of Radar and Sonar Signal Processing, Artech House, 2002.
- [20] V.S. Chernyak, Fundamentals of Multisite Radar Systems: Multistatic Radars and Multiradar Systems, Gordon & Breach, 1998.
- [21] N.J. Bistatic Willis, Radar, second ed., Technology Service Corp., Silver Spring, MD, 1995, corrected and republished by SciTech Publishing, Inc, Raleigh, NC, 2005.
- [22] N.J. Willis, H.D. Griffiths, Advances in Bistatic Radar, SciTech Publishing Inc., Raleigh, NC, 2007.
- [23] N.J. Willis, Bistatic Radar, in: M.I. Skolnik (Ed.), Radar Handbook, third ed., McGraw-Hill, 2008 (Chapter 23).
- [24] M. Cherniakov (Ed.), Bistatic Radar: Principles and Practice, (2007) and Bistatic Radar: Emerging Technology, Wiley, 2008.
- [25] M.C. Jackson, The geometry of bistatic radar systems, IEE Proc. Pt. F. 133 (7) (1986) 604–612.
- [26] G.W. Ewell, Bistatic radar cross section measurements, in: N.C. Currie (Ed.), Techniques of Radar Reflectivity Measurement, Artech House, 1989 (Chapter 7).
- [27] G.W. Ewell, S.P. Zehner, Bistatic radar cross section of ship targets, IEEE J. Ocean. Eng. OE-5 (4) (1980) 211–215.
- [28] R.E. Kell, On the derivation of bistatic RCS from monostatic measurements, Proc. IEEE 53 (1965) 983–988.
- [29] M. Born, E. Wolf, Principles of Optics, sixth ed., Pergamon Press, 1980, p. 559.

- [30] H.D. Griffiths, W.A. Al-Ashwal, K.D. Ward, R.J.A. Tough, C.J. Baker, K. Woodbridge, Measurement and modelling of bistatic radar sea clutter, *IET Radar Sonar Navig.* 4 (2) (2010) 280–292.
- [31] M. Weiner, Clutter, in: N.J. Willis, H.D. Griffiths (Eds.), *Advances in Bistatic Radar* Scitech Inc., Raleigh, NC, 2008 (Chapter 9).
- [32] A.R. Domville, The bistatic reflection from land and sea of X-band radio waves, GEC (Electronics) Ltd., Memo SLM 1802 , Stanmore, England, July 1967.
- [33] G.A. Yates, *Bistatic Synthetic Aperture Radar*, PhD Thesis, University College London, January 2005.
- [34] K.D. Ward, R.J.A. Tough, S. Watts, *Sea Clutter: Scattering, the K-Distribution and Radar Performance*, IET, London, 2006.
- [35] W.A. Al-Ashwal, C.J. Baker, A. Balleri, H.D. Griffiths, R. Harmann, M. Inggs, W.J. Miceli, M. Ritchie, J.S. Sandenbergh, A. Stove, R.J.A. Tough, K.D. Ward, S. Watts, K. Woodbridge, Statistical analysis of simultaneous monostatic and bistatic sea clutter at low grazing angles, *Electron. Lett.* 47 (10) (2011) 621–622.
- [36] W.A. Al-Ashwal, C.J. Baker, A. Balleri, H.D. Griffiths, R. Harmann, M. Inggs, W.J. Miceli, M. Ritchie, J.S. Sandenbergh, A. Stove, R.J.A. Tough, K.D. Ward, S. Watts, K. Woodbridge, Measurements of bistatic radar sea clutter, in: *IEEE Radar Conference 2011*, Kansas City, MO, 23–27 May 2011, pp. 217–221.
- [37] T. Tsao, M. Slamani, P. Varshney, D. Weiner, H. Schwarzlander, S. Borek, Ambiguity function for a bistatic radar, *IEEE Trans. Aerosp. Electron. Syst.* 33 (3) (1997) 1041–1051.
- [38] M.A. Ringer, G.J. Frazer, Waveform analysis of transmissions of opportunity for passive radar, *Proc. ISSPA'99*, Brisbane, Australia, 22–25 August 1999, pp. 511–514.
- [39] H.D. Griffiths, C.J. Baker, H. Ghaleb, R. Ramakrishnan, E. Willman, Measurement and analysis of ambiguity functions of off-air signals for passive coherent location, *Electron. Lett.* 39 (13) (2003) 1005–1007.
- [40] C.J. Baker, H.D. Griffiths, I. Papoutsis, Passive coherent radar systems – Part II: Waveform properties, *IEE Proc. Radar Sonar Navig.* 152 (3) (2005) 160–168.
- [41] H.D. Griffiths, C.J. Baker, Passive Bistatic Radar, in: W. Melvin (Ed.), *Principles of Modern Radar SciTech Publishing Inc.*, vol. 3 Raleigh, NC, (Chapter 11) (in press).
- [42] D. Poullin, Passive detection using digital broadcasters (DAB, DVB) with COFDM modulation, *IEE Proc. Radar Sonar Navig.* 152 (3) (2005) 143–152.
- [43] H.D. Griffiths, C.J. Baker, The signal and interference environment in passive bistatic radar information, in: *Decision and Control Symposium*, Adelaide, February 2007, pp. 12–14.
- [44] A. Farina, A. De Maio, T. Volpi, *Waveform diversity: past, present and future*, NATO Lecture Series SET-119 *Waveform Diversity for Advanced Radar Systems*, 2006.
- [45] S. Haykin, *Communication Systems*, fourth ed., John Wiley, New York, 2001.
- [46] P.E. Howland, D. Maksymiuk, G. Reitsma, FM radio based bistatic radar, *IEE Proc. Radar Sonar Navig.* 152 (3) (2005) 107–115.
- [47] F. Colone, R. Cardinali, P. Lombardo, Cancellation of clutter and multipath in passive radar using a sequential approach, in: *IEEE Radar Conference*, Verona, NY, USA, 24–27 April 2006, pp. 393–399.
- [48] H.D. Griffiths, A four-element VHF adaptive array processor, in: *2nd IEE International Conference on Antennas and Propagation*, York, IEE Conference Publication No. 195, Part I, 13–16 April 1981, pp. 185–189.
- [49] H.D. Griffiths, C.J. Baker, Passive Coherent Radar systems – Part I: Performance prediction, *IEE Proc. Radar Sonar Navig.* 152 (3) (2005) 153–159.
- [50] A.P. Andrews, HDTV-based passive radar, in: *AOC 4th Multinational PCR Conference*, Syracuse, NY, 6 October, 2005.
- [51] D. O'Hagan, Passive bistatic radar performance using FM radio illuminators of opportunity, PhD Thesis, University College London, March 2009.
- [52] C.J. Coleman, R.A. Watson, H. Yardley, A practical bistatic passive radar system for use with DAB and DRM illuminators, in: *Proceedings International Conference RADAR 2008*, Adelaide, Australia, 2–5 September 2008, pp. 1–7.

- [53] J.M. Caspers, Bistatic and Multistatic Radar, in: M.I. Skolnik (Ed.), *Radar Handbook*, first ed., McGraw-Hill, New York, 1970 (Chapter 36).
- [54] M. Tobias, A.D. Lanterman, Probability hypothesis density-based multitarget tracking with bistatic range and Doppler observations, *IEE Proc. Radar Sonar Navig.* 152 (3) (2005) 195–205.
- [55] R.P.S. Mahler, Multitarget Bayes filtering via first-order multitarget moments, *IEEE Trans. Aerosp. Electron. Syst.* 39 (4) (2003) 1152–1178.
- [56] M. Tobias, Probability hypothesis densities for multitarget, multisensor tracking with application to passive radar, PhD thesis, Georgia Institute of Technology, 2006.
- [57] J. Wurman, Wind Measurements, in: N.J. Willis, H.D. Griffiths, (Eds.), *Advances in Bistatic Radar*, SciTech Publishing Inc., Raleigh, NC, 2007 (Chapter 8).
- [58] J.L. Garrison, A. Komjathy, V.U. Zavorotny, S.J. Katzberg, Wind speed measurement using forward scattered GPS signals, *IEEE Trans. Geosci. Remote Sens.* 40 (1) (2002) 50–65.
- [59] J.L. Garrison, S.J. Katzberg, M.I. Hill, Effect of sea roughness on bistatically scattered range coded signals from the global positioning system, *Geophys. Res. Lett.*, 25 (13) (1998) 2257–2260.
- [60] Y. Huiazu, J.L. Garrison, G. Heckler, V.U. Zavorotny, Stochastic voltage model and experimental measurement of ocean-scattered GPS signal statistics, *IEEE Trans. Geosci. Remote Sens.* 42 (10) (2004) 2160–2169.
- [61] J.L. Garrison, J.K. Voo, S.H. Yueh, M.S. Grant, A.G. Fore, J.S. Haase, Estimation of sea surface roughness effects in microwave radiometric measurements of salinity using reflected Global Navigation Satellite System signals, *IEEE Geosci. Remote Sens. Lett.* 8 (6) (2011) 1170–1174.
- [62] V.C. Chen, *The Micro-Doppler Effect in Radar*, Artech House, 2011.
- [63] E.F. Greneker, Detection and tracking of humans and vehicle targets using high definition television signals in urban areas, *Proc. SPIE* (2007).
- [64] V. Koch, R. Westphal, A new approach to a multistatic passive radar sensor for air defense, in: IEEE International Radar Conference RADAR 1995, Washington, DC, IEEE Conf. Publ. No. 95CH3571-0, 8–11 May 1995, pp. 22–28.
- [65] V. Koch, R. Westphal, New approach to a multistatic passive radar sensor for air/space defense, *IEEE AES Mag.* 10 (11) (1995) 24–32.
- [66] J. Brown, K. Woodbridge, A.G. Stove, S. Watts, Air target detection using airborne passive bistatic radar, *Electron. Lett.* 46 (20) (2010) 1396–1397.
- [67] J. Brown, H.D. Griffiths, A.G. Stove, S. Watts, K. Woodbridge, Passive bistatic radar location experiments from an airborne platform, *IEEE AES Mag.* (in press).
- [68] A.P. Whitewood, Bistatic radar using a spaceborne illuminator, PhD thesis, University College London, June 2006.
- [69] A. Whitewood, C.J. Baker, H.D. Griffiths, Bistatic radar using a spaceborne illuminator, in: Proceedings IET International Conference RADAR 2007, Edinburgh, 15–18 October 2007, pp. 1–5.
- [70] H. Guo, S. Coetzee, D. Mason, K. Woodbridge, C.J. Baker, Passive radar detection using wireless networks, in: IET Radar Conference RADAR 2007, Edinburgh, 15–18 September 2007, pp. 1–4.
- [71] F. Colone, K. Woodbridge, H. Guo, D. Mason, C.J. Baker, Ambiguity function analysis of wireless LAN transmissions for passive radar, *IEEE Trans. Aerosp. Electron. Syst.* 47 (1) (2011) 240–264.
- [72] K. Chetty, G. Smith, H. Guo, K. Woodbridge, Target detection in high clutter using passive bistatic WiFi radar, in: IEEE Radar Conference, Pasadena, CA, USA, 4–8 May 2009, pp. 1–5.
- [73] Q. Wang, Y. Lu, C. Hou, Evaluation of WiMAX transmission for passive radar applications, *Microw. Opt. Technol. Lett.* 52 (7) (2010) 1507–1509.
- [74] F. Colone, P. Falcone, P. Lombardo, Ambiguity function analysis of WiMAX transmissions for passive radar, in: IEEE International Radar Conference RADAR 2010, Washington, DC, 10–14 May 2010, pp. 689–694.
- [75] K. Chetty, K. Woodbridge, H. Guo, G.E. Smith, Passive bistatic WiMAX radar for marine surveillance, in: IEEE International Radar Conference RADAR 2010, Washington, DC, 10–14 May 2010.

- [76] H. Guo, K. Woodbridge, C.J. Baker, Evaluation of WiFi beacon transmissions for wireless based passive radar, in: IEEE Radar Conference, Rome, Italy, 1–6 May 2008, pp. 26–30.
- [77] J.M. Headrick, J.F. Thompson, Applications of high-frequency radar, *Radio Sci.* 33 (4) (1998) 1045–1054.
- [78] J.M. Thomas, H.D. Griffiths, C.J. Baker, Ambiguity function analysis of digital radio mondiale signals for HF passive bistatic radar, *Electron. Lett.* 42 (25) (2006) 1482–1483.
- [79] M. Lesturgie, D. Poullin, Frequency allocation in radar: solutions and compromise for low frequency band, in: SEE International Radar Conference RADAR 99, Paris, France, 18–20 May 1999.
- [80] Air Scientific Intelligence Interim Report, Heidelberg, A.D.I. (Science), IIE/79/22, Public Records Office, Kew, London, 24 November 1944.
- [81] H.D. Griffiths, N.J. Willis, Klein Heidelberg – the first modern bistatic radar system, *IEEE Trans. Aerosp. Electron. Syst.* 46 (4) (2010) 1571–1588.

Websites

- [82] <<http://www.air-defense.net/forum/index.php?topic=11376.0>>.
- [83] <<http://www.defence.gov.au/dmo/esd/jp2025/jp2025.cfm>>.
- [84] <http://www.raytheon.com/capabilities/products/stellent/groups/public/documents/legacy_site/cms01_049201.pdf>.
- [85] <<http://www.qsl.net/g3cwi/doppler.htm>>.

Through-the-Wall Radar Imaging: Theory and Applications

17

Moeness G. Amin* and Fauzia Ahmad†

*Center for Advanced Communications, Villanova University, Villanova, PA, USA

†Radar Imaging Lab, Center for Advanced Communications, Villanova University, Villanova, PA, USA

2.17.1 Introduction

The field of remote sensing has developed a range of interesting imaging approaches for a variety of applications. Through-the-wall sensing is a relatively new area that addresses the desire to see inside buildings for various purposes, including determining the room layouts, discerning the building intent and nature of activities, locating the occupants, and even identifying and classifying objects within the building. Through-the-wall sensing is highly desired by police, fire and rescue, emergency relief workers, and military operations. Accurate sensing and imaging can allow a police force to obtain an accurate description of a building in a hostage crisis, or allow firefighters to locate people trapped inside a burning structure. In essence, the goals of through-the-wall sensing technology are to provide vision into otherwise obscured areas [1].

Each remote sensing application area has driven different sensing modalities and imaging algorithm development based upon propagation characteristics, sensor positioning, and safety issues. Traditional optical, radar, and sonar image processing all begin with basic wave physics equations to provide focusing to individual points. In many radar applications, for example, data sampled from many sensors are mathematically integrated to provide equivalent focusing using free-space propagation assumptions. Free-space imaging is commonly seen in synthetic aperture radar (SAR) techniques since atmospheric distortions are often negligible and can be safely ignored in first-order calculations [2].

Conventional imaging approaches exploit the wave equation to compute the expected phase at each point in space and time over which the data are collected. The complex phase front is similar to the spatial representation of the wavefront captured by a hologram [3]. The complex returns can then be compared against the predicted returns from points in the imaging target space to focus on each point in that space. The focusing is analogous to image reconstruction in holography, where a spatial pattern is projected back into the originating target image space. In true free-space conditions, this focusing approach represents a mathematically accurate way to perform imaging. More sophisticated approaches extend beyond free-space assumptions to allow for more complicated propagation effects, such as adaptive optics [4], atmospheric correction for radar [2], and matched field processing for sonar [5]. Correction approaches range from simple wavefront calibration to more sophisticated volumetric propagation corrections.

Free-space propagation does not apply to several applications where transmission through scattering media is encountered, including many modern imaging approaches such as geophysical sensing, medical imaging, and more recently through-the-wall imaging and sensing. In these applications, propagating signals diffract through a volume. As examples, geophysical imaging techniques generally measure seismic propagation through the earth to look for discontinuities that are often indicators of oil, gas, water, or mineral deposits [6]. In medical imaging, ultrasound tomographic approaches account for propagation through different tissue classes [7].

Non-freespace scattering applications are more representative of the through-the-wall sensing problem, albeit each has its own distinct challenges and approaches [1]. In geophysical and medical applications, the propagation medium is discontinuous but still respectively fills the sensing volume of the earth and tissue. To better propagate into the volume, sensors are placed in direct contact with the medium (e.g., seismometers for geophysical sensing, ultrasound transducers for medical imaging). In through-the-wall sensing, there are many air-material interfaces that dramatically change the wavefront. Through-the-wall sensors may be placed against the front walls or located some distance away from the structure. In either case, attenuation is largely seen only in the building materials and contents rather than in the large volumes of air that occupy most of the space outside or within a building. The rich through-the-wall scattering environment makes volumetric tomographic imaging approaches most relevant for through-the-wall sensing. Rather than using free-space focusing assumptions, correcting for propagation effects may greatly improve the imaging solution [8–10].

Through-the-wall sensing is best motivated by examining the applications primarily driving its development. Through-the-wall sensing grew from the application of ground-penetrating radar systems to walls, with specific applications documented in the literature since the late 1990s showing abilities to sense beyond a single wall from near range [11–25]. Applications can be divided based upon whether information is sought on motions within a structure or on imaging the structure and its stationary contents.

The need to detect motion is highly desired in situations like firefighters finding a child in a burning building or law enforcement officers locating hostages and their captors. Such applications can resort to Doppler discrimination of movement from background clutter [26–31]. Motion detection and localization can be decomposed into zero (0-D), one (1-D), two (2-D), or three-dimensional (3-D) systems. A 0-D system is simply a motion detector and will report any motion in the scene. Such systems may be useful to detect whether a room is occupied or not, often useful in monitoring applications such as security systems or intruder detection. Interior motion may be ample information for a firefighter to decide whether to enter a room or building. Since range or angle is not required, 0-D systems may use continuous wave (CW) tones as waveforms. While such systems have exquisite sensitivity, it is difficult to confine the sensitivity to desired areas, so care must be taken to prevent undesired detections beyond the intended range or even from the system operator himself due to multipath reflections. 0-D systems are not particularly useful in cases where other moving individuals may be within the sensor field of view.

One-dimensional systems provide a range to a target but not an angle. The extra dimension provides the ability to separate and possibly discriminate multiple targets. The range information can help bound whether a detection is in the adjacent room, or perhaps deeper within the building. The systems can obviously gate out detections from ranges beyond the ranges of interest, and may more easily discriminate

desired target motion from the motion of the operator. The systems often consist of a single receiving antenna in addition to a transmitting antenna (although sometimes a single antenna can be used for both functions). Motion of the operator can be gated out in range, or a reference antenna pointed towards the radar operator can be used for cancellation. Ranging is historically one of the principle reasons for the development of radar technology, but the short stand-off ranges of most through-the-wall systems provide some particular challenges. Three methods are often employed to obtain target range, namely, ultrawideband pulses [22,32–36], dual frequencies [28,37–39], or stepped CW [11,20].

Two-dimensional systems provide a slice through the scene in the range and angle dimensions [11,22,23]. The 2-D representation provides better localization of the mover, at the expense of a larger antenna array whose elements are collinear. 2-D systems will be subject to layover of objects, meaning that objects out of plane will appear to be rotated to the imaging plane, which can lead to distortion of the field of view. 3-D systems attempt to represent a volumetric representation of the field of view in the range, azimuth, and elevation dimensions [40]. The third dimension avoids the layover issue of targets being projected into a 2-D plane; at the expense of a 2-D array and higher processing requirements. Potentially, the third dimension can provide additional localization for identification of targets. Height information may allow discrimination between people and animals such as household pets, since radar cross-section alone can be unreliable in the presence of through-the-wall multipath and other issues. The ultimate goal is higher resolution for even better moving target classification. Both 2-D and 3-D processing techniques have generally used either multilateration [8,17] or SAR techniques [41,42] using either ultrawideband pulses or stepped CW radars to localize features in the scene.

Imaging of structural features and contents of buildings requires at least 2-D (and preferably 3-D) systems. It cannot rely on Doppler processing for separation of desired features, so multilateration or SAR approaches have been the most commonly used approaches. The general idea behind multilateration is to correlate range measurements from multiple sensors to specific points in the image. With sufficient spatial diversity from a large set of transmit/receive combinations, specific reflection points will start to integrate above the background interference. However, ambiguities will arise as the number of reflection points increases which can provide an over-determined system relative to the transmit/receive signal pairs which can detract from image quality. SAR-based systems can be thought of as a coherent extension of the multilateration concept. Instead of incoherent combinations of range returns from multiple transmit/receive pairs, coherent algorithms are used to provide a complex matched filter to specific points in the target space. SAR algorithms are well established for stand-off imaging applications [2]. Stand-off applications generally assume free-space propagation to each point in the target scene, although platform motion compensation and atmospheric effects are removed with autofocusing algorithms.

Most of the SAR and multilateration techniques usually neglect propagation distortions such as those encountered by signals passing through walls and objects. Distortions degrade the performance and can lead to ambiguities in target and wall localizations. Free-space assumptions no longer apply after the electromagnetic waves propagate through the first wall. As a result, free-space approximations may carry imaging systems through to the first wall, but propagation effects will then affect further imaging results. Shadowing, attenuation, multipath, reflection, refraction, diffraction, and dispersion, all play a role in how the signals will propagate after the first interface. Without factoring in propagation effects, imaging of contents within buildings will be severely impacted. As such, image formation methods,

array processing techniques, target detection, image sharpening, and clutter and multipath identification and rejection paradigms must work in concert and be reexamined in view of the nature and specificities of the underlying sensing problem.

A common occurrence of incorrect localization is objects outside the building being illuminated by the reflection off the first wall and subsequently creating an ambiguous image visible inside the building. Moreover, the strong front wall reflection causes nearby indoor weak targets to go undetected. Multipath propagation introduces ghosts or false targets in the image. Uncompensated refraction through walls can lead to localization or focusing errors, leading to offsets and blurring of imaged targets [41,43–45]. Bragg scattering off repeating structural elements, such as rebar in concrete walls or repetitive voids in concrete block walls, can cause image ambiguities and modulation of subsequent wavefronts. Some of the wall propagation effects can be compensated for using image-focusing techniques incorporating proper wavefront corrections [41,44]. SAR techniques and tomographic algorithms, specifically tailored for through-the-wall imaging, are capable of making some of the adjustments for wave propagation through solid materials. While such approaches are well suited for shadowing, attenuation, and refraction effects, they do not account for multipath, Bragg scattering, as well as strong reflections from the front wall.

In this chapter, we consider the recent algorithmic advances that address some of the aforementioned unique challenges in through-the-wall sensing operations. More specifically, Section 2.17.2 deals with techniques for the mitigation of the wall clutter. Front wall reflections are often stronger than target reflections, and they tend to persist over a long duration of time. Therefore, weak and close by targets behind walls become obscured and invisible in the image. Approaches based on both electromagnetic modeling and signal processing are advocated to significantly mitigate the front wall clutter. Section 2.17.3 presents an approach to exploit the rich indoor multipath environment for improved target detection. It uses a ray tracing model to implement a multipath exploitation algorithm, which maps each multipath ghost to its corresponding true target location. In doing so, the algorithm improves the radar system performance by aiding in ameliorating the false positives in the original SAR image as well as increasing the SNR at the target locations, culminating in enhanced behind the wall target detection and localization. Section 2.17.4 discusses a change detection approach to moving target indication for through-the-wall applications. Change detection is used to mitigate the heavy clutter that is caused by strong reflections from exterior and interior walls. Both coherent and noncoherent change detection techniques are examined and their performance is compared using real data collected in a semi-controlled laboratory environment. Section 2.17.5 deals with fast data acquisition schemes to provide timely actionable intelligence in through-the-wall operations. The demand for high degree of situational awareness to be provided by through-the-wall radar systems requires the use of wideband signals and large antenna arrays. As a result, large amounts of data are generated, which presents challenges in both data acquisition and processing. The emerging field of compressive sensing is employed to provide the means to circumvent possible logistic difficulties in collecting measurements in time and space and provide fast data acquisition and scene reconstruction for moving target indication.

It is noted that the chapter provides a signal processing perspective to through-the-wall radar imaging. Outside the scope of this chapter, there has been significant work performed for TWRI applications in the areas of antenna and system design [36,46–48], wall attenuation and dispersion [49–51], electromagnetic modeling [52–55], inverse scattering approaches [24,56,57], and polarization exploitation [58–60].

The progress reported in this chapter is substantial and noteworthy. However, many challenging scenarios and situations remain unresolved using the current techniques and, as such, further research and development is required. However, with the advent of technology that brings about better hardware and improved system architectures, opportunities for handling more complex building scenarios will definitely increase.

2.17.2 Wall clutter mitigation

Scattering from the exterior front wall is typically stronger than that from targets of interest, such as a human or a rifle, which have relatively small radar cross sections (RCS). This makes imaging of stationary targets behind walls difficult and challenging. The problem is further compounded when targets are in close vicinity of walls of either a high dielectric constant or with layered structures. In particular, hollow cinder block walls contain an air-gap void within the cinder block with disparate dielectric constants. This establishes a periodic structure resonance cavity that traps electromagnetic (EM) modes. The consequence of this layered composite structure on radar target imaging is to introduce long time constant relaxations on target detections in radar range profiles. Therefore, weak and close by targets behind walls become obscured and sometimes totally invisible in the image. Thus, wall reflections should be suppressed, or significantly mitigated, prior to applying image formation methods. One of the simple but effective methods is based on background subtraction. If the received signals can be approximated as the superposition of the wall and the target reflections, then subtracting the raw complex data without target (empty scene) from that with the target would remove the wall contributions and eliminate its potential overwhelming signature in the image. Availability of the empty scene, however, is not possible in many applications, and one must resort to other means to deal with strong and persistent wall reflections.

In the past few years, a number of approaches have been proposed to mitigate the front wall returns and effectively increase the signal-to-clutter ratio. For example, the wall reflections can be gated out, the corresponding parameters can be estimated, and then used to model and subtract the wall contributions from the received data [44]. This wall-dependent technique is effective, but its performance is subject to wall estimation and modeling errors. Another method is proposed in [42], which employs three parallel antenna arrays at different heights. The difference between the received signals at two different arrays is used for imaging. Due to the receiver symmetry with respect to the transmitter, a simple subtraction of the radar returns leads to wall reflection attenuation and image improvement. However, this technique requires two additional arrays and the effect of the subtraction operation on the target reflection is unknown and cannot be controlled. In [54], the walls reflections are eliminated by operating the imaging radar in cross-polarization as the cross-polarization returns from a planar interface such as the wall surface are theoretically zero, unlike the returns from humans. However, the cross-polarization returns from the targets behind walls are often very weak compared to their co-polarization counterparts, and as such, the radar performance may be limited by noise.

An eigen-structure technique is developed in [61] to decompose the received radar signal into three subspaces: clutter (including the front wall), target, and noise. Singular value decomposition (SVD) is applied on the data matrix to extract the target signatures. SVD has also been used as a wall clutter reduction method for TWRI in [62,63]. Another approach for wall clutter mitigation is based on spatial filtering. It utilizes the strong similarity between wall EM responses, as viewed by different

antennas along the entire physical or synthesized array aperture, is proposed in [64]. The idea is that under monostatic radar operation, the wall returns approximately assume equal values and identical signal characteristics across the array elements, provided that the extent of the wall is much greater than the antenna beamwidth. On the other hand, returns from targets with limited spatial extent, such as humans, vary from sensor to sensor. The dc component corresponding to the constant-type return, which is typical of walls, can be significantly reduced while preserving the target returns by applying an appropriate spatial filter along the array axis. However, care must be exercised in the choice of the spatial filter so that its characteristics cause minimum distortion to the target returns. The spatial-domain preprocessing scheme is analogous, in its objective, to the moving target indication (MTI) clutter filter operation in the time domain.

In this section, we present in details two of the aforementioned methods, namely, the wall parameter estimation method and the spatial filtering technique. For the spatial filtering technique, we consider two different types of filters, namely, moving average subtraction filter and the infinite impulse response notch filter, and compare their effects on the target image.

2.17.2.1 Spatial filtering

2.17.2.1.1 Characteristic of the wall reflection

Assume a synthetic aperture radar (SAR) is used in which a single antenna at one location transmits and receives the radar signal, then moves to the next location, and repeats the same operation along the axis parallel to the wall. Assume N antenna locations. With the wall reflection, the received signal at the n th antenna due to a single point target of complex reflectivity σ is given by

$$z(n, t) = s_w(t - \tau_w) + \sigma s(t - \tau_n), \quad (17.1)$$

where $s_w(t)$ is the signal reflected from the wall, τ_w is the two-way traveling time of the signal from the n th sensor to the wall, $s(t)$ is the transmitted signal convolved with the two-way transmission transfer function of the wall, and τ_n is the two-way traveling time between the n th antenna and the target. Note that σ is assumed to be independent of frequency and τ_w does not vary with the sensor location since the wall is parallel to the array. Furthermore, if the wall is homogeneous and much larger than the beamwidth of the antenna, the wall reflection $s_w(t)$ will be the same for all antenna locations. This implies that the first term in (17.1) assumes the same value across the array aperture. When the wall is not parallel to the array axis, τ_w should be calculated for each sensor location. Compensations for signal attenuations due to differences in the two-way traveling distances for different antenna positions should be performed prior to filtering.

Unlike τ_w , the time delay τ_n in (17.1) changes with each antenna location, since the signal path from the antenna to the target is different. For $t = t_0$ and $\sigma = 1$, the received signal is

$$z(n, t_0) = s_w(t_0 - \tau_w) + s(t_0 - \tau_n) \quad (17.2)$$

for $n = 0, 1, \dots, N - 1$. Since the time t is fixed, the received signal is now a function of n via the variable τ_n . We can rewrite (17.2) as a discrete function of n such that

$$z_{t_0}(n) = v_{t_0} + u_{t_0}[n], \quad (17.3)$$

where $v_{t_0} = s_w(t_0 - \tau_w)$ and $u_{t_0}[n] = s(t_0 - \tau_n)$. Since the delay τ_n is not linear in n , $u_{t_0}[n]$ for $n = 0, 1, \dots, N - 1$ is a nonuniformly sampled version of $s(t)$. Figure 17.1a shows the signals received

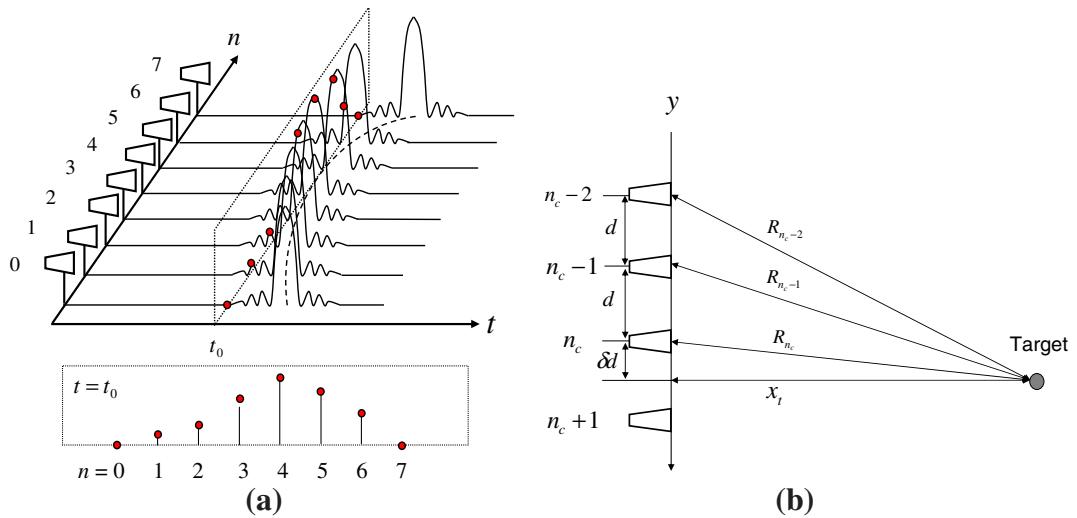


FIGURE 17.1

Received signals using SAR data collection scheme. (a) Array signal at $t = t_0$. (b) Range between antenna locations and the target.

by the N antennas at a given time instant. For example, suppose that the n_c th sensor is the closest to the target (see Figure 17.1b). If the wall diffraction is negligible, then

$$\begin{aligned} R_{n_c-m} &= \sqrt{x_t^2 + (\delta d + md)^2} = \sqrt{(x_t^2 + \delta d^2) + 2md \delta d + (md)^2} \\ &= \sqrt{R_{n_c}^2 + 2md \delta d + (md)^2}, \end{aligned} \quad (17.4)$$

where $m = |n - n_c|$. In most cases, the target's range R_{n_c} is much larger than the antenna spacing d . Then, $R_{n_c}^2 \gg 2md \delta d + (md)^2$. Using Taylor series expansion, we can approximate (17.4) as,

$$R_{n_c-m} \approx R_{n_c} \left(1 + \frac{1}{2} \frac{2md \delta d + (md)^2}{R_{n_c}^2} \right). \quad (17.5)$$

Therefore,

$$R_{n_c-m} - R_{n_c} \approx \frac{1}{2} \frac{2md \delta d + (md)^2}{R_{n_c}}. \quad (17.6)$$

The received signal at $t = \tau_{n_c}$ is given by

$$z(n, \tau_{n_c}) = s_w(\tau_{n_c} - \tau_w) + s(\tau_{n_c} - \tau_n) \approx v_{\tau_{n_c}} + s \left(-\frac{2md \delta d + (md)^2}{R_{n_c} c} \right). \quad (17.7)$$

The spatial frequency transform of (17.7) is given by

$$Z(\kappa, \tau_{n_c}) = \sum_{n=0}^{N-1} z(n, \tau_{n_c}) e^{-j\kappa n} \approx v_{\tau_{n_c}} \cdot \delta(\kappa) + S_{\tau_{n_c}}(\kappa), \quad (17.8)$$

where κ denotes the spatial frequency, and $S_{\tau_{n_c}}(\kappa)$ is the target spatial signature at $t = \tau_{n_c}$ such as

$$S_{\tau_{n_c}}(\kappa) = \sum_{n=0}^{N-1} s \left(-\frac{2md\delta d + (md)^2}{R_{n_c}c} \right) e^{-j\kappa n}. \quad (17.9)$$

The above analysis shows that separating wall reflections from target reflections amounts to basically separating constant (zero-frequency signal) from nonconstant valued signals across antennas, which can be performed by applying a proper spatial filter.

2.17.2.1.2 Target spatial signature

In the previous section, it is shown that the target signal across the antennas at a fixed time (fixed down-range) is a nonuniformly sampled version of the target reflection. In order to minimize the effect of the wall removal, which is done by spatial filtering, on the target signal, it is necessary to find the target spatial frequency bandwidth as viewed by the array aperture. This is a critical factor in designing the proposed spatial filters for the purpose of wall reflection removal. The target spatial frequency bandwidth depends on the target downrange, as shown in (17.9), and is also inversely proportional to the extent of the target signal return along the array. At any given time, the latter depends on the array length, target down range, and the target signal temporal duration.

Considering a point target, which reflects the radar signal without modification, the target spatial signature can be analytically computed. The minimum bandwidth occurs when the extent of the target return is maximum. Two different cases can be identified depending on the target location in cross-range relative to the edge of the array. Figure 17.2a shows the target signal return in the time-space domain when the target is located within the array aperture, whereas Figure 17.2b shows the case in which the target is outside the array aperture. In these figures, L represents the array length, τ is the signal temporal duration, and R is the target range. Consider the case of Figure 17.2a, where l_0 is the cross-range of the target location from the edge of the array. Without loss of generality, assume that l_0 is larger than or equal to $L/2$. The maximum width of the target return across the array for the case of Figure 17.2a is [64]

$$\mu \approx \sqrt{2R\delta_r} + \min\{\sqrt{2R\delta_r}, L - l_0\} \quad (17.10)$$

and for the case of Figure 17.2b,

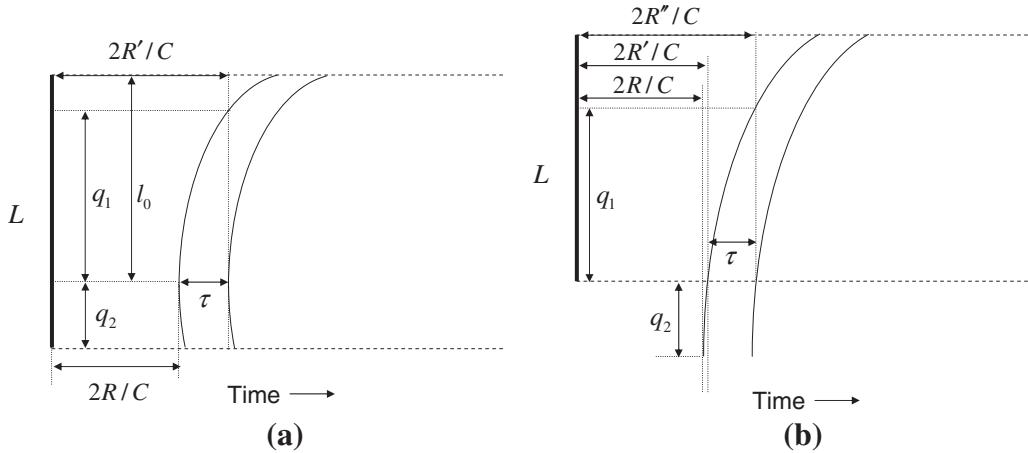
$$\mu = \sqrt{q_2^2 + 2W\delta_r} - q_2. \quad (17.11)$$

In the above equations, δ_r is the radar down-range resolution and

$$W = \sqrt{R^2 + q_2^2}. \quad (17.12)$$

The maximum width is achieved when the target is located along the centerline of the array

$$\mu_{\max} = \min\{2\sqrt{2R\delta_r}, L\}. \quad (17.13)$$

**FIGURE 17.2**

Target return signal in the time-space domain. (a) Target is within the array aperture. (b) Target is outside the array aperture.

Define $\mu/\sqrt{2}$ as the effective width. Then, the maximum effective width is

$$\mu_{\max,\text{eff}} = \min\{2\sqrt{R\delta_r}, L\} \quad (17.14)$$

and the corresponding minimum frequency bandwidth is given by

$$\kappa_{B,\min} = \frac{1}{\mu_{\max,\text{eff}}}. \quad (17.15)$$

Note that $\kappa_{B,\min}$ decreases as R increases. This implies that the spatial filter should have a narrower stopband to process the more distant target.

2.17.2.1.3 Moving average wall removal

Moving Average (MA) filter background subtraction is a noncausal Finite impulse response (FIR) filter, which notches out the zero spatial frequency component. The MA subtraction method has been effectively used for removal of the ground reflections in ground penetrating radar [65]. The effect of the MA on the target image can be viewed in the context of the corresponding change in the point spread function (PSF). With the transmitted signal chosen as a stepped-frequency waveform consisting of Q narrowband signals, let $z(n, f_q)$ be the signal received at the n th antenna at frequency f_q . Let $\bar{z}(n, f_q)$ be the signal after MA subtraction. That is

$$\tilde{z}(n, f_q) = z(n, f_q) - \bar{z}(f_q), \quad (17.16)$$

where $\bar{z}(f_q)$ is the MA of $z(n, f_q)$ such as

$$\bar{z}(f_q) = \frac{1}{2D+1} \sum_{m=n-D}^{m=n+D} z(m, f_q), \quad (17.17)$$

where $2D+1$ is the filter length. When the sum is taken over the entire length of the array, the moving average becomes just an average over the entire aperture and $2D+1=N$. The k th pixel value of the image obtained by applying delay-and-sum (DS) beamforming to the received data $z(n, f_q)$ is given by

$$b(k, l) = \sum_{n=0}^{N-1} \sum_{q=0}^{Q-1} z(n, f_q) \exp\{j2\pi f_q \tau_{n,(k,l)}\}, \quad (17.18)$$

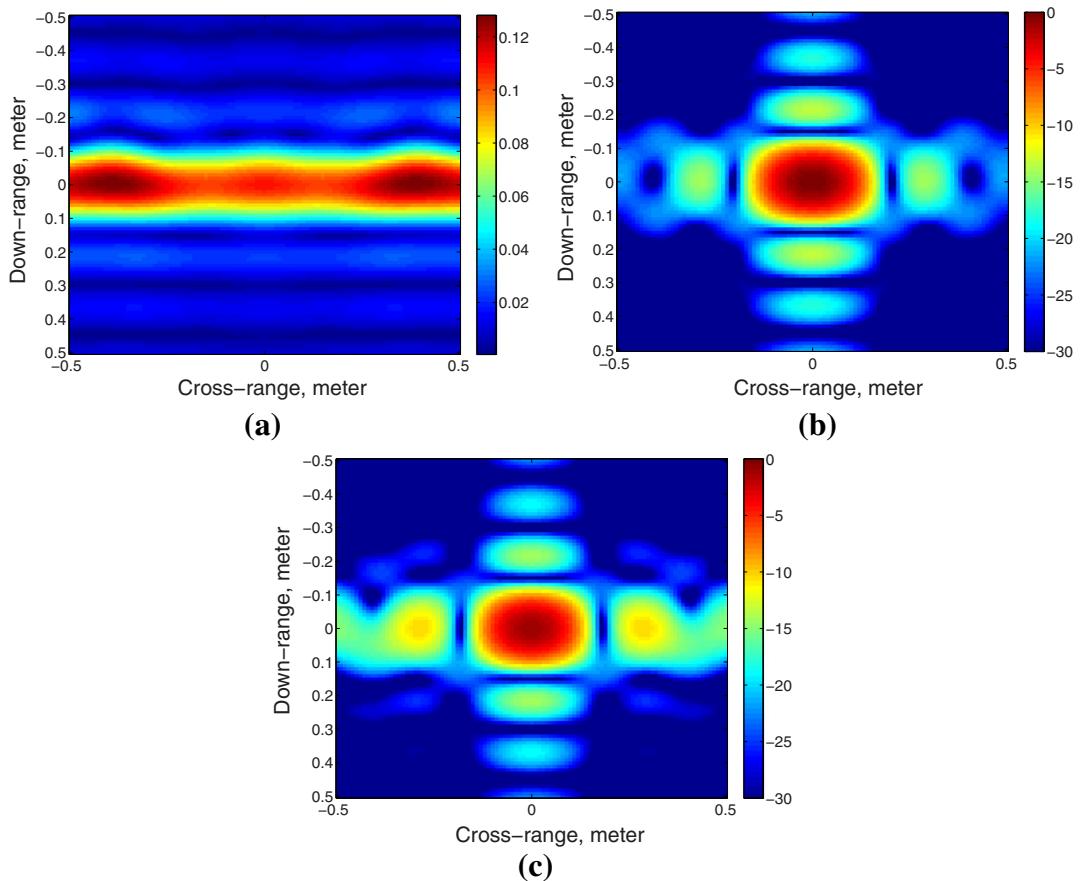
where $\tau_{n,(k,l)}$ is the two-way traveling time, through the air and the wall, between the n th antenna and the k th pixel location. The new DS beamforming image after filtering is given by

$$\begin{aligned} \tilde{b}(k, l) &= \sum_{n=0}^{N-1} \sum_{q=0}^{Q-1} \tilde{z}(n, f_q) \exp\{j2\pi f_q \tau_{n,(k,l)}\} \\ &= b(k, l) - \sum_{n=0}^{N-1} \sum_{q=0}^{Q-1} \bar{z}(f_q) \exp\{j2\pi f_q \tau_{n,(k,l)}\}. \end{aligned} \quad (17.19)$$

The second term in the above equation is due to the average subtraction. This additional term depends on the signal waveform, target location, and the antenna positions, implying that the effect of average subtraction varies according to these parameters. The PSF is the DS image of a point source as a function of the down-range and the cross-range. Figure 17.3a shows the image of the second term in (17.19) when a point target is located 6 m away from the center of a 3 m long antenna array. In this case, the maximum value is about 0.13 (The maximum value in PSF is one). The signal is a step-frequency waveform covering the 2–3 GHz band with 5 MHz steps. It is noted that the magnitude of this image assumes a high value over a large cross-range area, whereas the down-range spread of the function is small. Figure 17.3b is the PSF without spatial filtering. The modified PSF of the spatial filtered signal, which is the difference between Figure 17.3a and b, is shown in Figure 17.3c. We observe that the size of the mainlobe of the modified PSF remains unchanged with about 0.8 dB loss in the maximum value. However, the sidelobes are higher compared to the original PSF. Although the average subtraction seems to work well in this simulation, the limits of using this subtraction method when handling the distant targets can be shown as follows.

The discrete Fourier transform of $\tilde{z}(f_q)$ is

$$\begin{aligned} \tilde{Z}(\kappa, f_q) &= \sum_{n=0}^{N-1} \tilde{z}(n, f_q) \exp\{-j2\pi \kappa n/N\} = \sum_{n=0}^{N-1} (z(n, f_q) - \bar{z}(f_q)) \exp\{-j2\pi \kappa n/N\} \\ &= Z(\kappa, f_q) - \bar{z}(f_q) \sum_{n=0}^{N-1} \exp\{-j2\pi \kappa n/N\} = Z(\kappa, f_q) - N\bar{z}(f_q)\delta[\kappa]. \end{aligned} \quad (17.20)$$

**FIGURE 17.3**

Modification in the PSF due to MA subtraction. (a) Second term in (17.19). (b) PSF without filtering. (c) PSF after MA subtraction.

Note that

$$Z(\kappa, f_q)|_{\kappa=0} = \sum_{n=0}^{N-1} z(n, f_q) = N \bar{z}(f_q). \quad (17.21)$$

Therefore,

$$\tilde{Z}(\kappa, f_q) = \begin{cases} Z(\kappa, f_q) & \kappa \neq 0, \\ 0 & \kappa = 0. \end{cases} \quad (17.22)$$

As shown above, the subtraction removes the single spatial frequency component ($\kappa = 0$) without changing other components. The actual spatial frequency band that is filtered out is always $1/dN = 1/L$,

where L is the array length. When $\mu_{\max, \text{eff}}$ is equal to L , the target bandwidth is $1/L$. In this case, most of the target signal return will be removed by the process of the average subtraction. The range at which $\mu_{\max, \text{eff}} = L$ is

$$R_L = \frac{L^2}{4\delta_r}. \quad (17.23)$$

For example, for $L = 2$ m and $\delta_r = 0.075$ m (2 GHz bandwidth), R_L is equal to 13.33 m, implying that when the target's down-range is larger than 13.33 m, the MA subtraction method will eliminate most of the target signature, which is undesirable. Thus, the MA subtraction is simple but produces spatial filter characteristics, which may not be applicable to all wall types and target locations.

2.17.2.1.4 Notch filtering

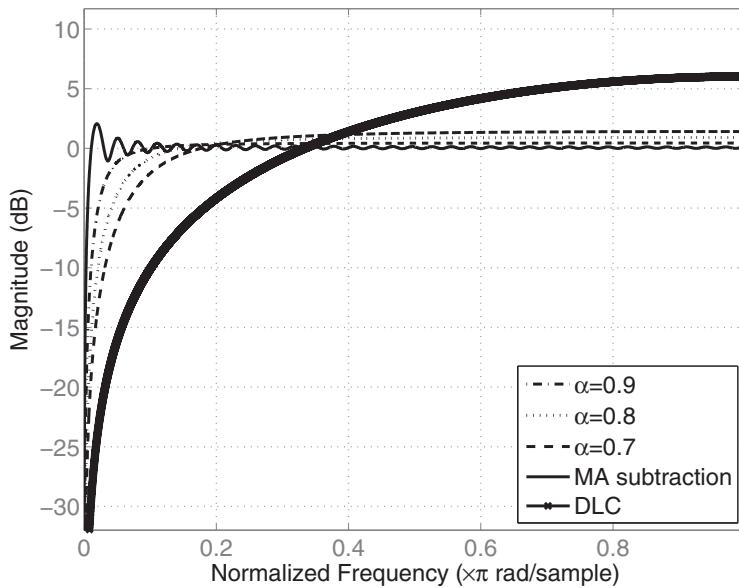
Zero-phase spatial IIR filters, such as notch filters, are good candidates for wall clutter mitigation. The IIR filter provides a flexible design. It overcomes the problem of the fixed MA (average) filter characteristics that could be unsuitable for various wall and target types and locations. Although the IIR filter spatial extent is truncated because of the finite number of antennas, the filter is still capable of delivering good wall suppression performance without a significant change in the target signal when the number of antennas is moderate. A simple IIR notch filter that rejects zero frequency component is given by

$$H_{\text{NF}}(\omega) = \frac{1 - e^{-j\omega}}{1 - \alpha e^{-j\omega}}. \quad (17.24)$$

The positive parameter $\alpha (< 1)$ determines the width of the filter notch. Figure 17.4 shows the frequency response of the filter for various values of α and that of MA subtraction with 67 antenna locations. Note the difference between the notch filters and MA. The notch filter does not have ripples and can change its stopband depending on the parameter α . In the underlying problem, ω is the spatial frequency and the parameter α provides a compromise between wall suppression and target signal conservation. This flexibility is an advantage of the notch filtering over MA subtraction method. As shown in the figure, as α moves closer to one, the stopband becomes narrower such that most of the signal except dc will pass through the filtering.

The spatial spectrum of wall reflections may have a nonzero width due to unstable antenna path and local inhomogeneities. In this case, a wider filter stopband should be applied to remove most of the wall reflections around dc, but it would also reduce the target reflection. Therefore, the spatial filter should be adjusted to the environment. It is noted that the frequency spectrum of the target reflection depends on the range to the target, the transmitted waveform, and the distance between antenna locations. These parameters determine the sampling point of the received signal and, subsequently, the spatial frequency characteristics of the target. If the sampled target signal does not have a dc component, it will not be severely affected by the spatial filter and, as such, the target image after notch filtering will approximately remain unaltered. It is noted that the desire to have the target spatial spectrum least attenuated by the spatial filtering could play an important role in designing signal waveforms for TWRI.

The filter zero-phase property, which is required for a focused image formation, can be achieved by applying the notch filter twice in the forward and backward directions [66]. By two-way filtering, group delay can be removed. One way to avoid the loss of the target signal returns by filtering, particularly if these returns vary slowly across the array aperture antennas, is to increase the delay between antenna

**FIGURE 17.4**

Frequency response of the notch filter for various α , DLC with two pulses, and MA subtraction with 67 antennas.

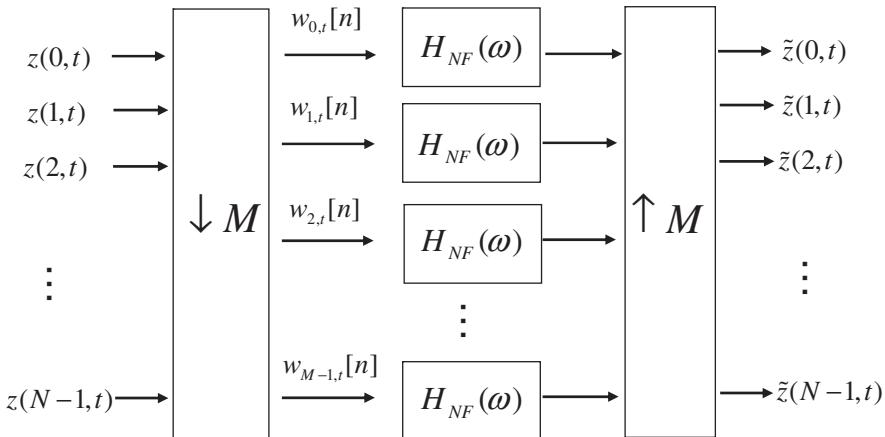
outputs. This is true whether the MA, the notch filter, or any other filter is employed. The increased delay can be achieved by down sampling of sensors without making the sensor spacing too large to cause aliasing. When the antenna positions are down sampled by M , there will be M different sets with N/M antennas. These M sets are filtered independently and the results are used by the delay and sum beamforming. Figure 17.5 shows the diagram of notch filter processing. The downsampled signal $w_{m,t}$ for $n = 0, 1, \dots, (N/M) - 1$ and $m = 0, 1, \dots, M - 1$

$$w_{m,t}[n] = z(nM + m - 1, t). \quad (17.25)$$

The effective filter length of the IIR filter should be considered when downsampling since the length of the downsampled signal is now N/M . A highly truncated IIR will lead to undesired filter characteristics.

2.17.2.1.5 Imaging results

A through-the-wall SAR system was set up in the Radar Imaging Lab at Villanova University. A line array of length 1.2446 m with 0.0187 m inter-element spacing was synthesized, parallel to a 0.14 m thick solid concrete wall, at a standoff distance of 1.01 m. The back and the side walls of the room were covered with RF absorbing material to reduce clutter. A stepped-frequency signal covering 1–3 GHz band with 2.75 MHz frequency step was employed. A vertical dihedral was used as the target and was placed 3.04 m away on other side of the front wall. The size of each face of the dihedral is 0.39 m by 0.28 m. The empty scene without the dihedral target present was also measured for comparison.

**FIGURE 17.5**

Block diagram of notch filtering.

First, we examine the validity of the wall EM return assumption that the spatial filtering approach is based on. Figure 17.6a shows the signal reflected from the concrete wall. It clearly shows the first few reflections from the wall and as expected, the wall return assumes almost constant values across the antenna positions. In order to quantify the performance of the images, we apply the target-to-clutter ratio (TCR) which is commonly adopted in SAR image evaluations [67, 68]. TCR is similar to the MTI improvement factor, except that the latter is for the time-domain, whereas TCR is for the image domain. TCR is calculated as

$$\text{TCR} = 20 \log_{10} \left(\frac{\max_{(k,l) \in A_t} |b(k,l)|}{\frac{1}{N_c} \sum_{(k,l) \in A_c} |b(k,l)|} \right), \quad (17.26)$$

where A_t is the target area, A_c is the clutter area, and N_c is the number of pixels in the clutter area. TCR, in essence, is the ratio between the maximum pixel value of the target to the average pixel value in the clutter region. The clutter region is the area where no target is present and the wall reflection is severe. The clutter area is manually selected in close vicinity to the wall where wall reflections are most pronounced. The rectangles depicted in the figures represent the clutter area and the target area in our example. Figure 17.6b shows the image without any preprocessing and the target is masked by the wall response. Figure 17.6c shows the result of applying background subtraction, and the target is clearly visible. Figure 17.7a and b demonstrates that the spatial filtering approach is effective in reducing the wall reflections without significantly compromising the target image. Figure 17.7a shows the DS image after MA subtraction, and Figure 17.7b shows the DS image after the notch filtering. TCRs in these figures are 2.2 dB (with wall), 26.2 dB (background subtraction), 13.4 dB (notch filtering), and 15.3 dB (MA subtraction). The background subtracted image provides the highest TCR, while MA subtraction and notch filtering provide comparable performance.

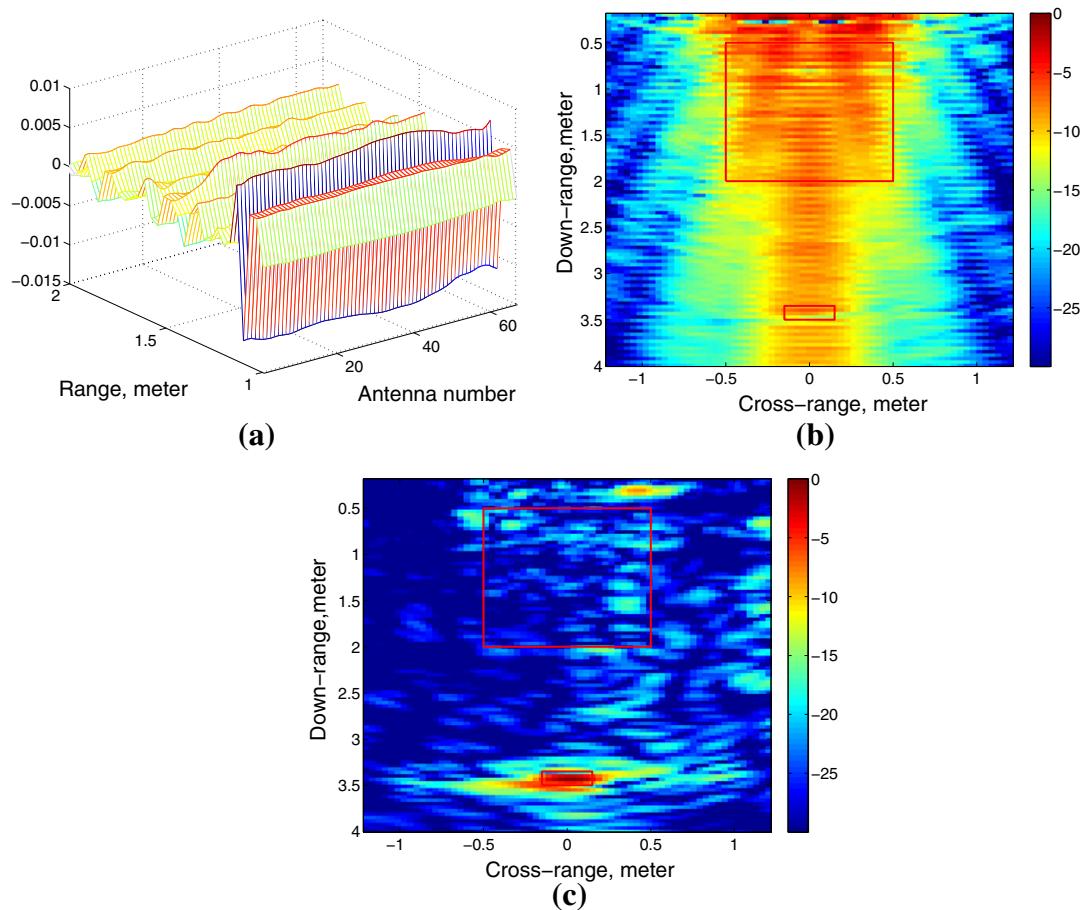


FIGURE 17.6

(a) Wall reflections of the solid concrete wall in time domain across the antenna positions. (b) DS image with wall reflections. (c) DS image with background subtraction.

2.17.2.2 Wall parameter estimation, modeling, and subtraction

2.17.2.2.1 Approach

In this approach, the effect of the wall on EM wave propagation is achieved through three steps. The wall parameters are accurately estimated. Then, the reflected signal from the wall is properly modeled. Finally, the modeled signal is subtracted from the measured signal.

Building walls, such as brick, adobe, and poured concrete walls, can be modeled by homogeneous dielectric slabs. Using the geometric optics approach, the reflection coefficient γ for vertically and

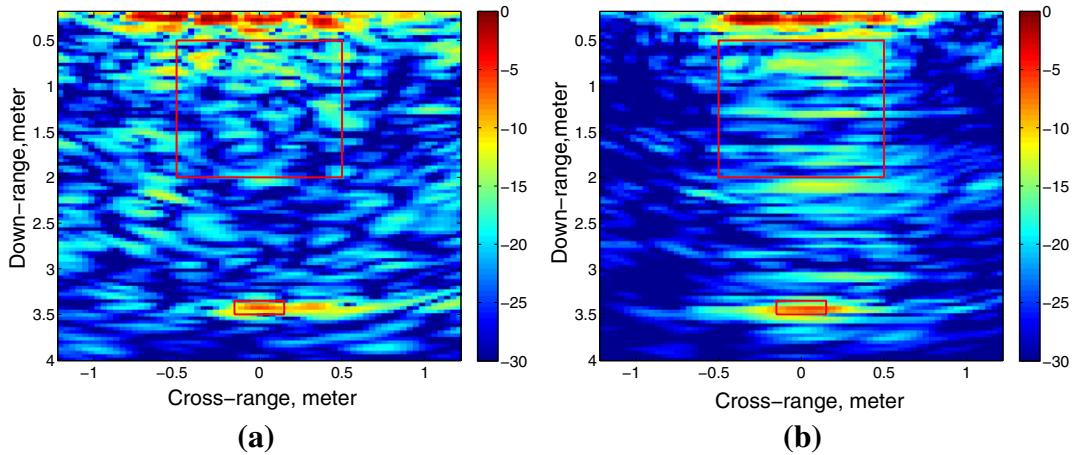


FIGURE 17.7

DS images for solid concrete wall. (a) MA subtraction. (b) Spatial notch filtering.

horizontally polarized incident waves through the wall is given by [44,69]

$$\gamma = \frac{R_{01} + R_{10}e^{j2k_{1x}d}}{1 + R_{01}R_{10}e^{j2k_{1x}d}}, \quad R_{10} = -R_{01} = \frac{1 - p_{10}}{1 + p_{10}}, \quad (17.27)$$

where $p_{10} = k_{0x}/k_{1x}$ for horizontal polarization and $p_{10} = \varepsilon_r k_{0x}/k_{1x}$ for vertical polarization of the incident fields. In (17.27), ε_r is the relative dielectric constant, d is the thickness of the wall, and k_{0x} and k_{1x} are the normal components of the propagation constants in air and in the dielectric, respectively. The expressions for the vertical/transverse magnetic and horizontal/transverse electric unit vectors are given in [69].

The wall parameters can be estimated from the time-domain backscatter at a given location [44,70]. Basically, the response of the wall is gated from the total backscatter signal and transferred to the frequency domain. The mean-squared error between the calculated reflection coefficient γ of (17.27) and the measured reflection coefficient γ_m given by

$$\delta = \frac{1}{N_f} \sum_{n=1}^{N_f} |\gamma_m(f_n) - \gamma(f_n)|^2 \quad (17.28)$$

is computed, and its minimum is searched for the wall thickness and permittivity [44,71]. In (17.28), N_f is the number of frequency points and f_n is the n th frequency.

The estimated wall parameters are then used to compute the wall EM reflection. This could be accomplished either numerically using EM modeling software, or in case of far-field conditions, analytically using the following expression [44]

$$E_r^w(k_n) = \frac{G\lambda}{8\pi} \frac{e^{j2k_n x_0}}{x_0} \Gamma(\theta_i = 0, k_n), \quad (17.29)$$

where x_0 is the standoff distance of the radar from the wall, λ is the wavelength, G is the antenna gain, $k_n = 2\pi f_n/c$ is the n th wave constant, c is the speed of light in free space, and $\Gamma(\theta_i = 0, k_n)$ is the reflectivity matrix from the wall at normal incidence angle and is given by,

$$\Gamma = \begin{bmatrix} \gamma_v & 0 \\ 0 & \gamma_h \end{bmatrix}. \quad (17.30)$$

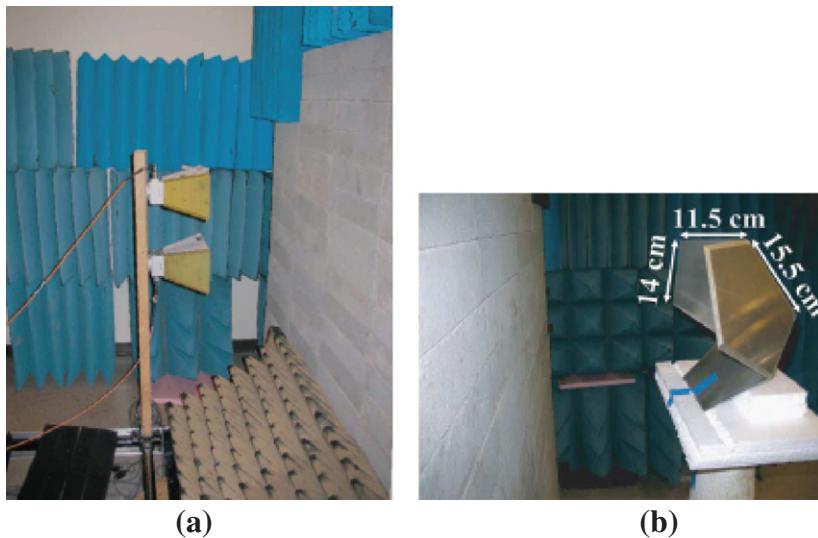
In (17.30), the subscripts “v” and “h,” respectively, denote the vertical and horizontal polarization and $\gamma_{v/h}$ is given by (17.27). This estimated wall reflection should be coherently subtracted from the total received signal in order to isolate the signature of the target on the other side of the wall.

The isolated target return can then be processed to generate an image of the scene. The effects of the wall transmissivity on the target image are the following: (1) dislocation of the target from its actual position in range and (2) significant degradation in cross-range resolution [41,44]. To compensate for the effects of transmission through the wall, a compensation factor proportional to the inverse of the squared of the wall transmittivity can be used in the conventional free-space image formation methods. In cases where transmissivity is low and the signal is noisy, division by a small noise-affected number may cause significant distortion. In such cases, one may use only the phase of transmission coefficients as a compensation factor [44].

2.17.2.2 Imaging results

An experimental study was conducted at the Radiation Laboratory in University of Michigan. The measurement setup consists of an HP8753D vector network analyzer, two double ridge horn antennas with operational frequency range of 1–18 GHz, an XY table along with a control unit, and a personal computer. The vertically aligned antennas, one for transmission and the other for reception, are mounted on a vertical wooden rod (along the z -axis), which is attached to the carriage of the XY table. A $2.37\text{ m} \times 1.77\text{ m}$ wall composed of poured concrete blocks is made on top of a layer of cinder blocks inside the laboratory. The underlying cinder block layer is used to line up the antennas in height at approximately the middle of the concrete block wall. Figure 17.8a shows the side view of the measurement setup. A small trihedral corner reflector with pentagonal panel geometry is used as a point target behind the wall, as shown in Figure 17.8b. The back corner of the trihedral (scattering phase center) is at $x = -0.71\text{ m}$ and at the same height as that of the receiver antenna, i.e., 1.28 m above the ground plane. The transmitting antenna is attached about 0.25 m below the receiving antenna on the wooden rod. The antennas are moved along a scan line of length 95.88 cm , parallel to the wall, with a spacing of 2.04 cm . The apertures of both antennas are about 0.45 m away from the wall. The frequency of operation is from 1 to 2.5 GHz and the frequency step is 12.5 MHz . It is noted that the antennas are in the far-field region of the target.

An image is first constructed under assumption of free-space propagation by using the total received signal, and is shown in Figure 17.9a. Here, in addition to the target, the wall is also imaged, primarily as two parallel lines showing the front and back boundaries. Since backscatter from the wall is very strong, the sidelobes generated by image formation spill over into desired image domain, which is manifested as multiple lines parallel to the wall surfaces. Also, the target appears blurred and its location is biased. Figure 17.9b shows the image after the estimate, model, and subtract approach was applied. The wall thickness and the permittivity were estimated using (17.28) as 20 cm and $5.7 + j0.6$, respectively. The wall reflection was computed using (17.29) and subtracted from the total received signal,

**FIGURE 17.8**

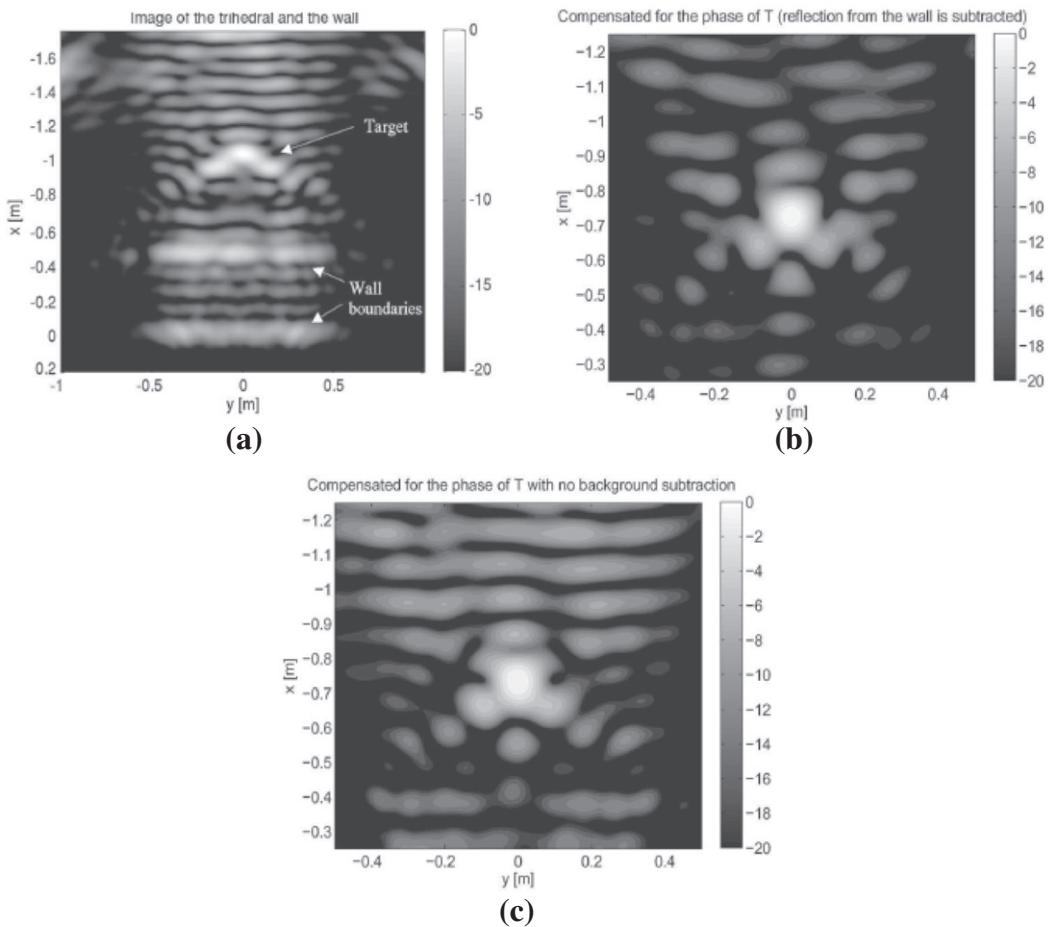
(a) Side view of the measurement setup. (b) Trihedral target.

and the resulting signal was used for image formation. For comparison, the signal prior to wall return subtraction was used for imaging wherein compensation for the effect of transmission through the wall was applied and is shown in Figure 17.9c. Although the imaged target is at the correct location and is refocused in Figure 17.9c, existence of the wall sidelobes in the image is still evident. By comparing Figure 17.9b and c, a substantial improvement in clutter rejection is observed after subtraction of the modeled wall return.

2.17.3 Multipath exploitation

The existence of targets inside a room or in an enclosed structure introduces multipath in the radar return, which results in false targets or ghosts in the radar images. These ghosts lie on or near the vicinity of the back and side walls, leading to a cluttered image with several false positives. Without a reasonable through-the-wall multipath model, it becomes difficult to associate a ghost to a particular target. Identifying the nature of the targets in the image and tagging the ghosts with their respective target according to a developed multipath model, although important to reduce false alarms, is not the final goal of a high-performance through-the-wall imaging system. Since each multipath provides some information about the target, it becomes prudent to utilize the multipath rather than ignoring it, once identified. The utilization or exploitation of the multipath to one's advantage represents a paradigm shift when compared to the classical approach of either ignoring or mitigating it.

For through-the-wall radar imaging applications, the existence of multipath has been recently demonstrated for stationary targets in [52, 72, 73]. In [72], the authors use distributed fusion to remove the false

**FIGURE 17.9**

(a) Image of the trihedral behind wall using full received data and free-space image formation. (b) Image of trihedral after estimated wall return subtraction and wall compensation. (c) Image of trihedral after wall compensation.

targets caused either from multipath or target interactions for stationary scenes after suitable image registration. In [52], time reversal techniques are used to alleviate ghosts and clutter from the target scene and in [73], a synthetic aperture radar (SAR) based image of a human inside a room is shown along with possible multipath ghosts. However, no more rigorous multipath modeling and analysis are presented in the aforementioned references [52, 72, 73] and the references therein.

In this section, we derive a model for the multipath in an enclosed room of four walls. The model considers propagation through a front wall and specular reflections at interior walls. A SAR system is considered and stationary or slow moving targets are assumed. Although the multipath model

presented deals with walls, reflections from the ceiling and floor can be readily handled. We demonstrate analytically that the multipath as seen by each sensor is displaced, and, therefore, we derive the actual focusing positions of the multipath ghosts in downrange and crossrange. The multipath model permits an implementation of a multipath exploitation algorithm, which associates as well as maps each target ghost back to its corresponding true target location. In so doing, the exploitation algorithm improves the radar system performance by ameliorating the false positives in the original SAR image as well as increasing the signal-to-clutter ratio (SCR) at the target locations, culminating in enhanced behind the wall target detection and localization. It is noted that the exploitation algorithm only maps back target-wall interactions; target-target interactions are left untreated. The multipath exploitation algorithm is inspired by the work in [74], wherein the shadowed regions of a target are revealed via its multipath returns. The difference, however, is that we are not striving to reveal the shadowed regions of the target. In our case, we deal with targets with arbitrary dielectric constants, and wish to associate and map each multipath to its true target location.

2.17.3.1 Image formation algorithm

In this section, we describe the through-the-wall delay-and-sum beamforming approach [40, 41, 75, 76]. We consider a SAR system in which a single antenna at one location transmits a wideband signal and receives the radar return, and then moves to the next location and repeats the same operation along the axis parallel to the front wall. Assume N monostatic antenna locations. The setup is as depicted in Figure 17.10. Consider a single point target, located at $\mathbf{x}_t = [x_t, y_t]^T$. At each antenna location, the radar transmits a pulsed waveform $s(t)$, where “ t ” indexes the time within the pulse, and measures the reflected signal. The target return at the n th antenna location is given by,

$$r_n(t) = T_n^2(\mathbf{x}_t)a(\mathbf{x}_t)s(t - 2\tau_n(\mathbf{x}_t)), \quad (17.31)$$

where $a(\mathbf{x}_t)$ represents the target reflectivity, $T_n(\mathbf{x}_t)$ is the complex amplitude associated with the one-way propagation through a dielectric wall for the n th antenna location [77], and $\tau_n(\mathbf{x}_t)$ represents the one-way through-the-wall propagation delay from the n th antenna location to the target.

The scene of interest comprises several pixels indexed by the downrange and the crossrange. The complex composite return from the k th pixel location \mathbf{x}_k is obtained by applying time delays to the N

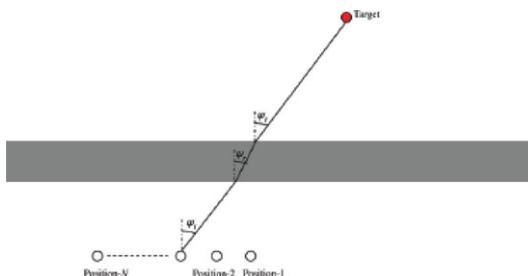


FIGURE 17.10

Through-the-wall synthetic aperture radar model.

received signals, followed by weighting and summing the results. That is,

$$r(t, \mathbf{x}_k) = \sum_{n=0}^{N-1} w_n r_n(t + 2\tau_n(\mathbf{x}_k)). \quad (17.32)$$

The signal $r(t, \mathbf{x}_k)$ is passed through a matched filter, with impulse response $h(t) = s^*(-t)$. The complex amplitude assumed by the pixel \mathbf{x}_k in the image $I(\cdot)$ is obtained by sampling the matched filter output at time $t = 0$,

$$I(\mathbf{x}_k) = \{r(t, \mathbf{x}_k) * h(t)\}|_{t=0}, \quad (17.33)$$

where “*” denotes the convolution operation. Equations (17.32) and (17.33) describe the standard beamforming approach in through-the-wall radar imaging. It is noted that if the imaged pixel is in the vicinity of or at the true target location, then the complex amplitude in (17.33) assumes a high value as given by the system’s point spread function. The process described in (17.32) and (17.33) is carried out for all pixels in the scene of interest to generate the image of the scene.

The beamforming approach as described above is pertinent to a point target in a two-dimensional (2D) scene. Extensions to three-dimensional (3D) scenes and spatially extended targets are discussed in [40, 78]. It is noted that, in the above beamforming description, we have not considered the multipath returns; neither have we addressed the specificities of calculating the delays, $\tau_n(\cdot)$, which will be treated in detail in the following sections. Without loss of generality, we now assume that the weights $w_n = 1, \forall n$ and the target has unit reflectivity.

2.17.3.2 Multipath model

Consider a room under surveillance using a SAR system. A priori knowledge of the room layout, i.e., wall locations and material properties, is assumed. The scene being observed relative to the n th sensor location is as shown in Figure 17.11. The origin is marked as point “O” in the figure, and the standard convention

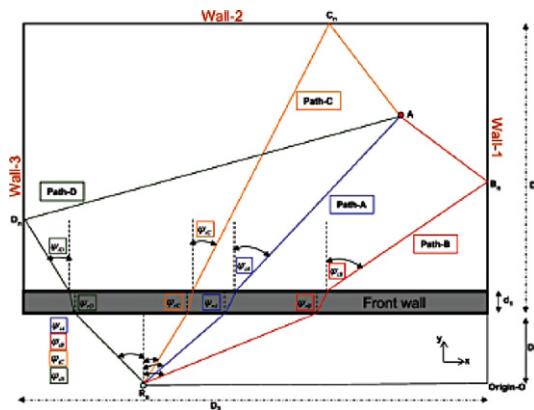


FIGURE 17.11

Multipath model.

for the positive x - and y -axes is assumed. The n th sensor location is given by $\mathbf{R}_n = [-D_{xn}, 0]^T$. The front wall has a thickness d_1 and dielectric constant ε_1 . For notational simplicity, the back and the side walls are also assumed to have ε_1 as the dielectric constant. The side walls, labeled as wall-1 and wall-3, have a length D_1 , whereas the front wall and the back wall (wall-2) have a length D_2 . The target is stationary and at location $\mathbf{A} = [-x_t, y_t]^T$. The standoff distance from the front wall is constant for each sensor location constituting the synthetic aperture, and is denoted as D_y . In the figure, we consider the direct path, referred to as path-A, and three additional paths, namely, paths-B, C, and D, which correspond to single-bounce multipath. In general, there exist other paths which can contribute to the multipath returns; these include multiple bounces from the back and side walls as well as paths that include multiple reflections within each of the four walls themselves. Examples of the former and the latter are provided in Figure 17.12. Such paths are defined as higher-order multipath. Hence, in Figure 17.11, we have considered only first-order multipath. It is noted that higher-order multipath are in general weaker compared to the first-order multipath due to the secondary reflections and refractions at the various air-wall interfaces and high attenuation in the wall material. Therefore, we choose to exclude these paths from the model.

The walls are assumed smooth with specular reflections. The smoothness assumption is valid, since for through-the-wall radars, the wavelength of operation is much larger than the roughness of the walls. Specular reflections are a direct consequence of the wall smoothness, and necessitate that the angle of incidence be equal to the angle of reflection. Note that, in general, the back and side walls may each be of a different material (interior or exterior grade), and thus may assume different values for the wall thickness and dielectric constant. If interior reflections inside these walls are considered, then the thickness of these walls is required [52, 73, 74]. However, since we have ignored such higher-order paths, we only require the dielectric constant of the back and side walls to be known. As the EM wave propagates through the front wall, it bends at the medium discontinuity as dictated by Snell's law, i.e., each of the paths, as seen in Figure 17.11, has an associated angle of incidence and an angle of refraction. For example, the angles of incidence and refraction for path-B are denoted as $\psi_{iB}^{(n)}$ and $\psi_{rB}^{(n)}$, respectively. Similar nomenclature follows for the remaining paths.

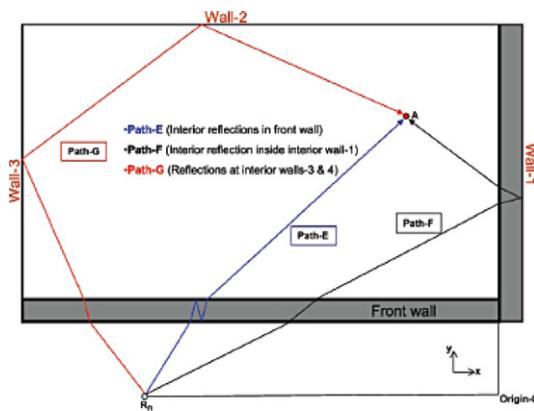


FIGURE 17.12

Other higher order paths.

Let the reflection points on wall-1, 2, and 3 be denoted by B_n , C_n , and D_n , with respective position vectors $\mathbf{B}_n = [0, y_{Bn}]^T$, $\mathbf{C}_n = [-x_{Cn}, D_1 + D_y]^T$, and $\mathbf{D}_n = [-D_2, y_{Dn}]^T$. It is clear that these position vectors are dependent on the sensor location. The one-way path delays for the four paths, with the antenna at the n th location, are denoted by $\tau_p^{(n)}$, $p \in \{A, B, C, D\}$, and can be derived from the geometry as

$$\begin{aligned}\tau_A^{(n)} &= (d_1 \sqrt{\varepsilon_1} \sec(\psi_{rA}^{(n)}) + (y_t - d_1) \sec(\psi_{iA}^{(n)})) / c, \\ \tau_B^{(n)} &= (d_1 \sqrt{\varepsilon_1} \sec(\psi_{rB}^{(n)}) + (y_{Bn} - d_1) \sec(\psi_{iB}^{(n)}) + x_t \operatorname{cosec}(\psi_{iB}^{(n)})) / c, \\ \tau_C^{(n)} &= (d_1 (\sqrt{\varepsilon_1} \sec(\psi_{rC}^{(n)}) - \sec(\psi_{iC}^{(n)})) + (2D_1 + 2D_y - y_t) \sec(\psi_{iC}^{(n)})) / c, \\ \tau_D^{(n)} &= (d_1 \sqrt{\varepsilon_1} \sec(\psi_{rD}^{(n)}) + (y_{Dn} - d_1) \sec(\psi_{iD}^{(n)}) + (D_2 - x_t) \operatorname{cosec}(\psi_{iD}^{(n)})) / c,\end{aligned}\quad (17.34)$$

with c being the speed of light and the coordinates of B_n , C_n , and D_n given by

$$\begin{aligned}y_{Bn} &= y_t - x_t \cot(\psi_{iB}^{(n)}), \\ x_{Cn} &= x_t + (D_1 + D_y - y_t) \tan(\psi_{iC}^{(n)}), \\ y_{Dn} &= y_t - (D_2 - x_t) \cot(\psi_{iB}^{(n)}).\end{aligned}\quad (17.35)$$

Equations (17.34) and (17.35) depend on the angles of incidence and refraction. Since the standoff distance is a known constant, by projecting the various paths to the x -axis, we obtain the following equations, which are useful in calculating the various angles.

$$\begin{aligned}d_1 \tan(\psi_{rA}^{(n)}) + (y_t - d_1) \tan(\psi_{iA}^{(n)}) - D_{xn} + x_t &= 0, \\ d_1 \tan(\psi_{rB}^{(n)}) + (y_{Bn} - d_1) \tan(\psi_{iB}^{(n)}) - D_{xn} &= 0, \\ (2D_1 + 2D_y - y_t) \tan(\psi_{iC}^{(n)}) + d_1 (\tan(\psi_{rC}^{(n)}) - \tan(\psi_{iC}^{(n)})) - D_{xn} + x_t &= 0, \\ d_1 \tan(\psi_{rD}^{(n)}) + (y_{Dn} - d_1) \tan(\psi_{iD}^{(n)}) - D_2 + D_{xn} &= 0.\end{aligned}\quad (17.36)$$

The angles of refraction can be determined from Snell's law as,

$$\psi_{rp}^{(n)} = \sin^{-1} \left(\frac{\sin(\psi_{ip}^{(n)})}{\sqrt{\varepsilon_1}} \right), \quad p \in \{A, B, C, D\}. \quad (17.37)$$

Substituting the angles of refraction from (17.37) in (17.36), and using (17.35), we obtain a set of equations that can be solved numerically for the angles of incidence by using the Newton method. The angles of refraction can then be obtained using (17.37).

We are now in a position to write the radar signal return from the single target scene as a superposition of the direct path and the multipath returns. For the n th sensor location, we obtain

$$r_n(t) = \sum_{p \in \{A, B, C, D\}} T_{pn}^2 s(t - 2\tau_p^{(n)}) + \sum_{\substack{(p, q) \in \{A, B, C, D\} \\ p \neq q}} T_{pn} T_{qn} s(t - \tau_p^{(n)} - \tau_q^{(n)}), \quad (17.38)$$

where T_{pn} is the complex amplitude associated with reflection and transmission coefficients for the one-way propagation along path- p , and depends on the angles of incidence, the angles of refraction, and

dielectric properties of the walls [77]. In (17.38), the first summation captures the two-way returns along the direct path and each multipath. More specifically, the signal propagates along a particular path- p , $p \in \{A, B, C, D\}$, reaches the target and retraces the same path back to the radar, i.e., path- p . The multipath returns due to the combination paths are captured by the second summation, i.e., the wave initially travels to the target via path- p and returns to the radar through a different path- q , $q \in \{A, B, C, D\}$.

From Figure 17.11 and (17.35), we observe that the coordinates of the points of reflection at the back and side walls are sensor dependent, implying that the multipath presents itself at different locations to the different sensors. Location of the multipath as seen by each sensor and the actual focusing position of the resulting ghost in the image are discussed next.

2.17.3.2.1 Multipath locations

For simplicity of analysis, consider the scenario in Figure 17.11, but without the front wall. We first focus on the multipath originating from wall-1, i.e., the signal traveling to the target via path-A, and following path-B back to the radar or vice versa. The goal is to find the location of the multipath corresponding to the above combination path as seen by each of the N sensors. Let the point $P_n^{w_1}$, represented by the vector $[-x_{tn}^{w_1}, y_{tn}^{w_1}]^T$, denote the multipath location as seen by the n th sensor. The superscript “ w_1 ” stresses that the multipath is associated with the reflection at wall-1 and the subscript “ t ” indicates that the multipath is due to the target located at $[-x_t, y_t]^T$. Reflecting path-B about wall-1 yields an alternate radar-target geometry, as shown in Figure 17.13. We obtain the virtual target denoted by point A^1 with coordinates $[x_t, y_t]^T$ such that the distance $AB_n = A^1B_n$, which implies that

$$AR_n + AB_n + B_nR_n = AR_n + A^1R_n. \quad (17.39)$$

In other words, the combination path comprising of path-A and path-B has the same length as path-A and the path connecting the n th sensor to the virtual target at A^1 . Due to the monostatic nature of the radar system, this combination path appears to the n th sensor as the traditional two way path of length $2R_n P_n^{w_1}$. Therefore, we have

$$R_n P_n^{w_1} = \frac{AR_n + A^1R_n}{2}. \quad (17.40)$$

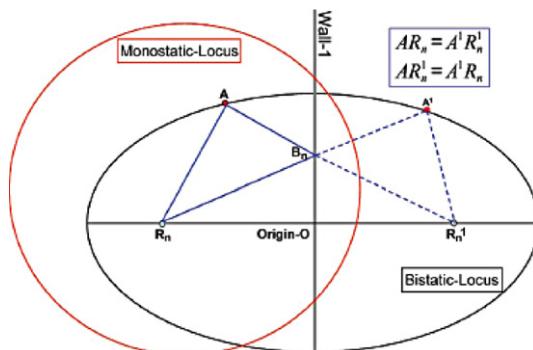


FIGURE 17.13

The equivalence between the bistatic and monostatic geometries.

Additionally, we obtain the virtual radar at $\mathbf{R}_n^1 = [D_{xn}, 0]^T$ on the other side of the wall, as shown in Figure 17.13. Considering Figure 17.13, it is readily observed that the first order multipath has the same time delay as a bistatic configuration, comprising the radar and the virtual radar. That is, in terms of range, we have

$$R_n A + R_n A^1 = R_n A^1 + R_n^1 A^1, \quad (17.41)$$

where $R_n A^1 + R_n^1 A^1$ represents the bistatic configuration. In the bistatic case, the constant range contour corresponding to the multipath is an ellipse, which has foci at the radar and virtual radar and passes through the target and the virtual target locations. On the other hand, the monostatic measurement scenario described by (17.40) has a circular constant range contour, centered at the radar. Since the two measurements are equivalent, the location of the multipath corresponds to the point of intersection of the bistatic elliptical and the monostatic circular constant range contours. The equations for the ellipse and circle are given by

$$\begin{aligned} \frac{(x_{tn}^{w_1})^2}{(R_n P_n^{w_1})^2} + \frac{(y_{tn}^{w_1})^2}{(R_n P_n^{w_1})^2 - D_{xn}^2} &= 1, \\ (x_{tn}^{w_1} + D_{xn})^2 + (y_{tn}^{w_1})^2 &= (R_n P_n^{w_1})^2. \end{aligned} \quad (17.42)$$

Solving (17.42) for an intersection point, we can readily see that $x_{tn}^{w_1} = 0$ is the only solution. In other words, regardless of the location of the target, its first-order multipath always falls on the wall. The y-coordinate of $P_n^{w_1}$ can then be derived as

$$y_{tn}^{w_1} = \pm \sqrt{\frac{(AR_n + A^1 R_n)^2}{4} - D_{xn}^2}. \quad (17.43)$$

In (17.43), the positive y-coordinate is the desired value as the other solution lies behind the radar. It is clear from (17.43) that the multipath location is sensor dependent. Therefore, the locations of the target multipath corresponding to the various sensor positions will be displaced from one another. In the presence of the front wall, the multipath remains on the wall, but the equations for the ellipse and circle are different than that presented in (17.42). Hence, for the scenario comprising of the front-wall and wall-1, the multipath appears on wall-1, with its y-coordinate given by

$$y_{tn}^{w_1} = \left(\frac{c\tau_A^{(n)} + c\tau_B^{(n)}}{2} + d_1(\sec(\psi_{it}^{(n)w_1}) - \sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_1})) \right) \cos(\psi_{rt}^{(n)w_1}), \quad (17.44)$$

where $\psi_{it}^{(n)w_1}$ and $\psi_{rt}^{(n)w_1}$ are, respectively, the angles of incidence and refraction for the ghost. The solution in (17.44) depends on the angle $\psi_{it}^{(n)w_1}$, which can be obtained by solving

$$\begin{aligned} d_1 \tan(\psi_{rt}^{(n)w_1}) + \left(\frac{c\tau_A^{(n)} + c\tau_B^{(n)}}{2} + d_1 \sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_1}) \right) \sin(\psi_{it}^{(n)w_1}) - D_x &= 0, \\ \psi_{rt}^{(n)w_1} &= \sin^{-1} \left(\frac{\sin(\psi_{it}^{(n)w_1})}{\sqrt{\varepsilon_1}} \right). \end{aligned} \quad (17.45)$$

So far, we have considered the multipath corresponding to wall-1; there exist two other multipath returns tagged to the remaining walls, namely, wall-2, and wall-3. We can readily show, using similar analysis, that the multipath associated with wall-2 appears on wall-2, i.e., $y_{tn}^{w_2} = D_y + D_1$, at an x -coordinate given by

$$x_{tn}^{w_2} = -\sin(\psi_{it}^{(n)w_2}) \times \left(\begin{array}{l} d_1(\sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_2}) - \tan(\psi_{rt}^{(n)w_2}) \operatorname{cosec}(\psi_{it}^{(n)w_2})) \\ + D_x \operatorname{cosec}(\psi_{it}^{(n)w_2}) - \frac{c\tau_A^{(n)} + c\tau_C^{(n)}}{2} \end{array} \right). \quad (17.46)$$

The angles in the above equation are determined by solving the equations,

$$\begin{aligned} d_1 \sec(\psi_{it}^{(n)w_2}) + (D_1 - d_1) \sec(\psi_{it}^{(n)w_2}) + d_1 \sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_2}) - \frac{c\tau_A^{(n)} + c\tau_C^{(n)}}{2} &= 0, \\ \psi_{rt}^{(n)w_2} &= \sin^{-1} \left(\frac{\sin(\psi_{it}^{(n)w_2})}{\sqrt{\varepsilon_1}} \right). \end{aligned} \quad (17.47)$$

Likewise, the multipath associated with wall-3 appears at wall-3, i.e., $x_{tn}^{w_3} = -D_2$, and at a y -coordinate given by,

$$y_{tn}^{w_3} = \left(\frac{c\tau_A^{(n)} + c\tau_D^{(n)}}{2} + d_1(\sec(\psi_{it}^{(n)w_3}) - \sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_3})) \right) \cos(\psi_{rt}^{(n)w_3}). \quad (17.48)$$

The respective angles can be obtained by solving,

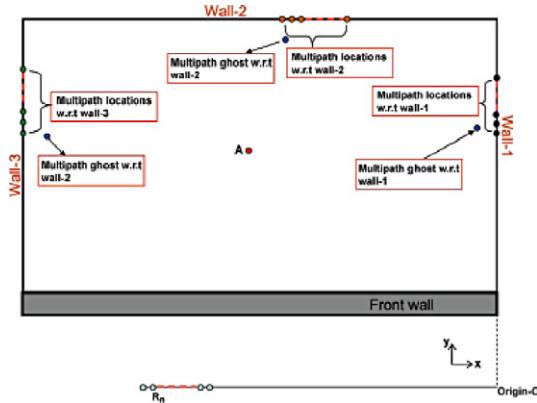
$$\begin{aligned} d_1 \tan(\psi_{rt}^{(n)w_3}) + \left(\frac{c\tau_A^{(n)} + c\tau_D^{(n)}}{2} + d_1 \sqrt{\varepsilon_1} \sec(\psi_{rt}^{(n)w_3}) \right) \sin(\psi_{it}^{(n)w_3}) - D_2 + D_x &= 0, \\ \psi_{rt}^{(n)w_3} &= \sin^{-1} \left(\frac{\sin(\psi_{it}^{(n)w_3})}{\sqrt{\varepsilon_1}} \right). \end{aligned} \quad (17.49)$$

From the above equations, it is again observed that the N sensors view the multipath, resulting from the combination paths associated with a particular wall, at different locations on that wall. That is, the multipath maybe regarded as a moving target. As a result, when applying beamforming, the multipath ghost appears at a different pixel in the vicinity of the true multipath locations. We note that the multipath ghost will lie inside the room except when the target is near the corners of the room. In this case, as the multipath corresponding to the N sensors may appear along an extrapolation of the wall, the multipath ghost may appear outside the room. We further note that the virtual target corresponds to two-way propagation along the single-bounce multipath, and is readily seen to lie outside the room perimeter. The multipath focusing pixel analysis for the combination paths is discussed next.

2.17.3.2.2 Multipath focusing analysis

Consider Figure 17.14, which shows the multipath locations w.r.t to walls-1–3, and the focusing pixels w.r.t to these walls. Consider the multipath locations associated with wall-1; further assume that the focused multipath ghost appears at a pixel location given by

$$\mathbf{x}_t^{w_1} = [-\Delta x_1 + x_{t1}^{w_1}, \Delta y_1 + y_{t1}^{w_1}]^T, \quad (17.50)$$

**FIGURE 17.14**

Multipath locations and imaged location of multipath ghosts.

where $\mathbf{x}_{t1}^{w_1} = [-x_{t1}^{w_1}, y_{t1}^{w_1}]^T$ is the true multipath location corresponding to the first sensor position. Using a first order Taylor series expansion, which is valid under conditions of a small aperture [79] and when the ghost is in the vicinity of the true multipath locations, we obtain the difference in propagation path length between the multipath ghost location and the true multipath location w.r.t the first sensor position as

$$\begin{aligned} \Delta\zeta_1^{w_1} &= \Delta x_1 \frac{D_{x1} - D_y \tan(\psi_{it}^{(1)w_1}) - d_1 \tan(\psi_{rt}^{(1)w_1}) - x_{t1}^{w_1}}{\sqrt{(D_{x1} - D_y \tan(\psi_{it}^{(1)w_1}) - d_1 \tan(\psi_{rt}^{(1)w_1}) - x_{t1}^{w_1})^2 + (y_{t1}^{w_1} - D_y - d_1)^2}} \\ &\quad + \Delta y_1 \frac{y_{t1}^{w_1} - D_y - d_1}{\sqrt{(D_{x1} - D_y \tan(\psi_{it}^{(1)w_1}) - d_1 \tan(\psi_{rt}^{(1)w_1}) - x_{t1}^{w_1})^2 + (y_{t1}^{w_1} - D_y - d_1)^2}}. \end{aligned} \quad (17.51)$$

Following the analysis in [79], (17.51) can be expressed as

$$\Delta\zeta_1^{w_1} = \Delta x_1 \sin(\psi_{it}^{(1)w_1}) + \Delta y_1 \cos(\psi_{it}^{(1)w_1}). \quad (17.52)$$

In general, for the n th sensor, we have

$$\Delta\zeta_n^{w_1} = \Delta x_1 \sin(\psi_{it}^{(n)w_1}) + (\Delta y_1 - (y_{tn}^{w_1} - y_{t1}^{w_1})) \cos(\psi_{it}^{(n)w_1}), \quad n = 1, \dots, N. \quad (17.53)$$

For the multipath to focus at the location $\mathbf{x}_t^{w_1}$, we must have,

$$\Delta\zeta_n^{w_1} = 0, \quad \text{for } n = 1, \dots, N. \quad (17.54)$$

This yields a least squares (LS) formulation, given by

$$\begin{aligned}\mathbf{A}_1 \mathbf{e}_1 &= \mathbf{b}_1, \mathbf{e}_1 = [\Delta x_1, \Delta y_1]^T, \mathbf{A}_1 := [\mathbf{a}_1^{w_1} \quad \mathbf{a}_2^{w_1}]^T, \\ \mathbf{a}_1^{w_1} &= [\sin(\psi_{it}^{(1)w_1}), \dots, \sin(\psi_{it}^{(N)w_1})]^T, \mathbf{a}_2^{w_1} = [\cos(\psi_{it}^{(1)w_1}), \dots, \cos(\psi_{it}^{(N)w_1})]^T, \\ \mathbf{b}_1 &:= [0, y_{t2}^{w_1} - y_{t1}^{w_1}, \dots, y_{tN}^{w_1} - y_{t1}^{w_1}]^T \circ \mathbf{a}_2^{w_1},\end{aligned}\quad (17.55)$$

where “ \circ ” denotes the Hadamard or element-wise product. The solution of (17.56) is readily obtained by $(\mathbf{A}_1^T \mathbf{A}_1)^{-1} \mathbf{b}_1$.

Now, considering the multipath w.r.t wall-2, we have the following LS formulation for the focused ghost pixel.

$$\begin{aligned}\mathbf{A}_2 \mathbf{e}_2 &= \mathbf{b}_2, \mathbf{e}_2 = [\Delta x_2, \Delta y_2]^T, \mathbf{A}_2 := [\mathbf{a}_1^{w_2} \quad \mathbf{a}_2^{w_2}]^T, \\ \mathbf{a}_1^{w_2} &= [\sin(\psi_{it}^{(1)w_2}), \dots, \sin(\psi_{it}^{(N)w_2})]^T, \mathbf{a}_2^{w_2} = [\cos(\psi_{it}^{(1)w_2}), \dots, \cos(\psi_{it}^{(N)w_2})]^T, \\ \mathbf{b}_2 &:= [0, y_{t2}^{w_2} - y_{t1}^{w_2}, \dots, y_{tN}^{w_2} - y_{t1}^{w_2}]^T \circ \mathbf{a}_2^{w_2}.\end{aligned}\quad (17.56)$$

Similarly, for the multipath from wall-3, we have

$$\begin{aligned}\mathbf{A}_3 \mathbf{e}_3 &= \mathbf{b}_3, \mathbf{e}_3 = [\Delta x_3, \Delta y_3]^T, \mathbf{A}_3 := [\mathbf{a}_1^{w_3} \quad \mathbf{a}_2^{w_3}]^T, \\ \mathbf{a}_1^{w_3} &= [\sin(\psi_{it}^{(1)w_3}), \dots, \sin(\psi_{it}^{(N)w_3})]^T, \mathbf{a}_2^{w_3} = [\cos(\psi_{it}^{(1)w_3}), \dots, \cos(\psi_{it}^{(N)w_3})]^T, \\ \mathbf{b}_3 &:= [0, y_{t2}^{w_3} - y_{t1}^{w_3}, \dots, y_{tN}^{w_3} - y_{t1}^{w_3}]^T \circ \mathbf{a}_2^{w_3}.\end{aligned}\quad (17.57)$$

The formulations in (17.56)–(17.58) assume that the sensor position increases from left to right. In other words, sensor-1 is at the far left of the array whereas sensor- N is at the far right.

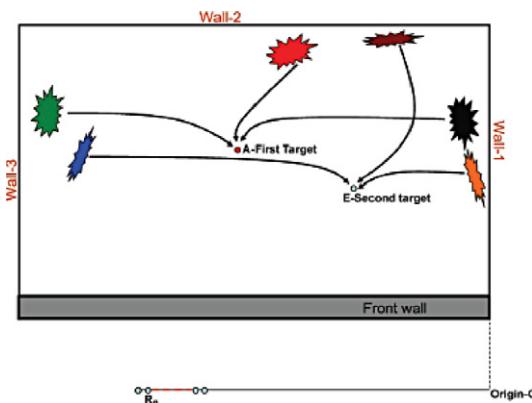


FIGURE 17.15

Multipath exploitation principle.

2.17.3.3 Multipath exploitation algorithm

Noting that the multipath ghosts exist due to the presence of the target, we state our objective as follows. Given the beamformed image $\{I(\mathbf{x}_k)\}_{k=1}^K$, our aim is to associate each multipath ghost with the respective target via the model developed in Section 2.17.3.2. The principle is captured in Figure 17.15, which consists of two targets and six false positives or multipath ghosts. We desire to associate and map these ghosts to their respective true target locations. The main advantages of such an association or mapping are reduction in false positives in the original beamformed image, and an increase in the SCR at the true target coordinates. Note that the first advantage is directly implied in Figure 17.15, whereas the second advantage is explained as follows.

For ease of exposition, the technique is explained considering the focused multipath ghost from wall-1 only. The technique for exploiting all the ghosts w.r.t to all three walls is enumerated subsequently. Consider an arbitrary pixel location $\mathbf{x} = [x, y]^T$, whose focused multipath ghost w.r.t wall-1 presents itself at location $\mathbf{x}^{w_1} = [x^{w_1}, y^{w_1}]^T$. Consider an intermediate image, $I_1(\mathbf{x})$, wherein the association and mapping of the focused multipath ghosts is performed using simple 2D weighting functions. That is,

$$I_1(\mathbf{x}) \in \mathbb{C}^K = \sum_{k=1}^K I(\mathbf{x}_k) \Phi_1(\mathbf{x}^{w_1}, \mathbf{x}_k, \sigma^2) \Phi_2(\mathbf{x}, \mathbf{x}_k, \sigma^2), \quad (17.58)$$

$$\Phi_1(\mathbf{x}^{w_1}, \mathbf{x}_k, \sigma^2) := \exp\left(-|\mathbf{x}^{w_1} - \mathbf{x}_k|^2/\sigma^2\right), \quad \Phi_2(\mathbf{x}, \mathbf{x}_k, \sigma^2) := 1 - \exp\left(-|\mathbf{x} - \mathbf{x}_k|^2/\sigma^2\right),$$

where $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ are the weighting functions and σ^2 is an arbitrary variance. The weighting functions are related to the two dimensional real Gaussian distributions. A natural improvement, albeit not pursued here, is to let the variance in the y -direction to be equal to the system's range resolution, the variance in the x -direction to be equal to the system's crossrange resolution, and introducing a correlation which is a function of the actual ghost location. This would ensure that the ensuing weighting functions are rotated and sheared to match the system's point spread function, nevertheless at a higher computational cost. The role of $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ is explained as follows. Consider the weighting function $\Phi_1(\cdot)$; if the pixel \mathbf{x}_k is in the vicinity of the focused multipath ghost pixel \mathbf{x}^{w_1} associated with the target location \mathbf{x} , then $\Phi_1(\cdot)$ assumes a large value. On the other hand, if \mathbf{x}_k is not in the vicinity of \mathbf{x}^{w_1} , then $\Phi_1(\cdot)$ assumes a low value. Since an exponential function is used, the weights are always between 0 and 1, the lower bound being theoretically unachievable. The vicinity is obviously controlled by the variance parameter, σ^2 . From our extensive simulations, we found that a lower variance must be chosen for smaller dimension rooms. For small rooms, it is logical to assume that the ghosts and higher order multipath are located quite close to one another. Therefore, it becomes prudent to be conservative in selecting the variances. Large variances for smaller rooms lead to false mappings. Thus, variance selection for the problem at hand must be guided by the actual dimensions of the room being imaged. The weighting function $\Phi_2(\cdot)$ rejects pixel locations at and in the vicinity of the imaged pixel, i.e., the location corresponding to the assumed true target location $\mathbf{x}_k = \mathbf{x}$. This is required because, although the focus point is singular, but due to the system's point spread function, some of the energy is spread to the neighboring pixels.

From our extensive simulations, we found that as the original image $I(\cdot)$ is complex-valued, summing or integrating across the multipath pixels may result in destructive interference yielding a poor SCR. This loss of SCR is against our principle of multipath exploitation. To alleviate this problem, we use

the intensity of the original beamformed image in (17.58) as

$$I_1(\mathbf{x}) \in \mathbb{R}^K = \sum_{k=1}^K |I(\mathbf{x}_k)|^2 \Phi_1(\mathbf{x}^{w_1}, \mathbf{x}_k, \sigma^2) \Phi_2(\mathbf{x}, \mathbf{x}_k, \sigma^2), \quad (17.59)$$

which implies an incoherent summation. Now considering the ghosts w.r.t the three walls, we can readily see that (17.59) is modified as

$$I_1(\mathbf{x}) = \sum_{w_i=w_1}^{w_3} \sum_{k=1}^K |I(\mathbf{x}_k)|^2 \Phi_1(\mathbf{x}^{w_i}, \mathbf{x}_k, \sigma^2) \Phi_2(\mathbf{x}, \mathbf{x}_k, \sigma^2). \quad (17.60)$$

Image $I_1(\cdot)$ will have no intensity at and near the vicinity of the focused multipath pixels. Hence, consider the following composite image obtained by simple pixel-wise multiplication.

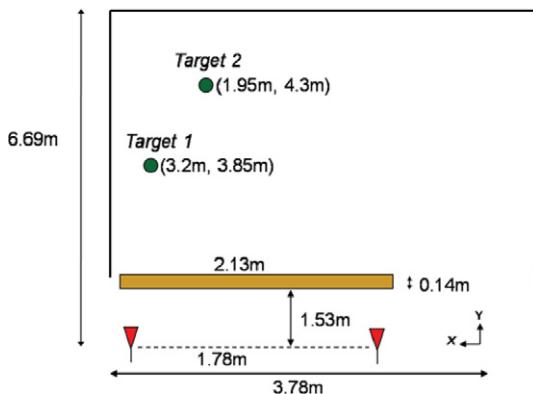
$$I_f(\mathbf{x}) = I_1(\mathbf{x}) \times |I(\mathbf{x})|^2. \quad (17.61)$$

The effect of pixel-wise multiplication in (17.61) is as follows. The image $I_1(\cdot)$ will have deep nulls at the hypothesized ghosts locations, whereas a large peak at the true target location. Thus, (17.61) is essentially a masking operation, nulling the ghosts in the original image $I(\cdot)$, and simultaneously enhancing the energy at the true target location. Hence, the composite image $I_f(\cdot)$ will alleviate the false positives, i.e., the multipath ghosts and increase the intensity of the true target pixel. In the approach described, the multipath ghost locations are readily obtained using the multipath focusing analysis in Section 2.17.3.2.2.

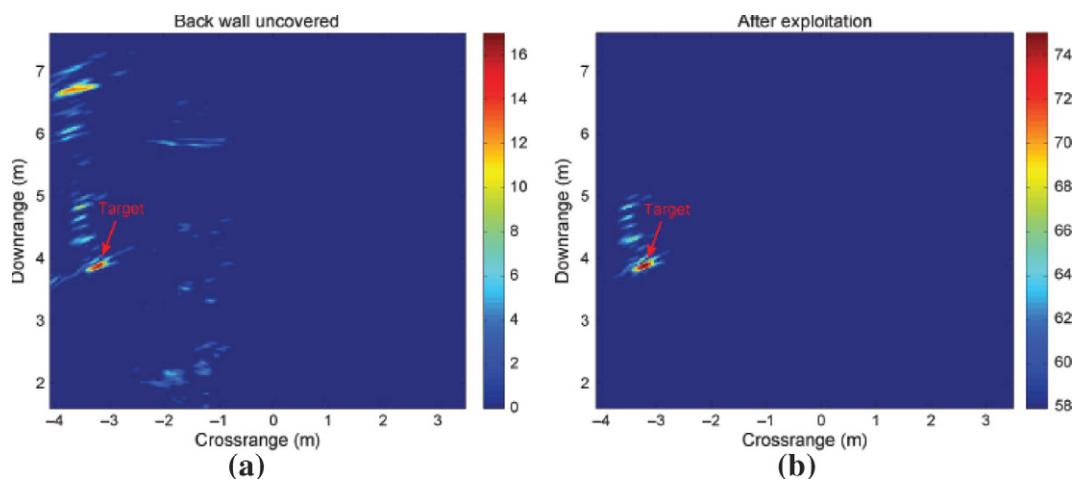
It is noted that, in practice, different targets may exhibit significantly weaker bistatic radar cross sections (RCS) relative to their monostatic RCS. Accordingly, some or maybe all of the target ghosts may go unnoticed. In this case, the proposed multipath exploitation approach will not prove beneficial as there may not be any increase in the target SCR. Also, consider a scenario where one target presents itself at one of the other target's focused multipath ghost. Since multipath is unexploited in the original beamformed image, we would declare the presence of the target at the pixel in question. On the other hand, the proposed approach will map this target and of course the collocated multipath ghost back to the other target's location causing the second target to go undetected. The question then arises whether one should ignore the multipath ghosts and consider them as true targets or instead exploit the multipath ghosts. The former has an unreasonably high false alarm rate, whereas the latter yields poor target detection for the considered scenario, which, although rare, can occur in practice. However, noting that indoor targets of interest are typically in motion, it is reasonable to assume that a genuine target would not persist on a multipath ghost throughout the surveillance period. Hence, multipath exploitation should be preferred. Another scenario of interest is when the direct path to the target is blocked. In that case, although the ghosts will be mapped to the true target coordinates in the image $I_1(\cdot)$, the masking operation in (17.61) will cause the target to still go undetected as the original image $I(\cdot)$ will have a null at the target location.

2.17.3.4 Experimental results

A through-the-wall SAR system was set up in the Radar Imaging Lab at Villanova University. A stepped-frequency signal with 696 steps covering the 1–3 GHz frequency band was used. This provides a range resolution of 0.15 m and a maximum unambiguous range of 40 m. The room dimensions

**FIGURE 17.16**

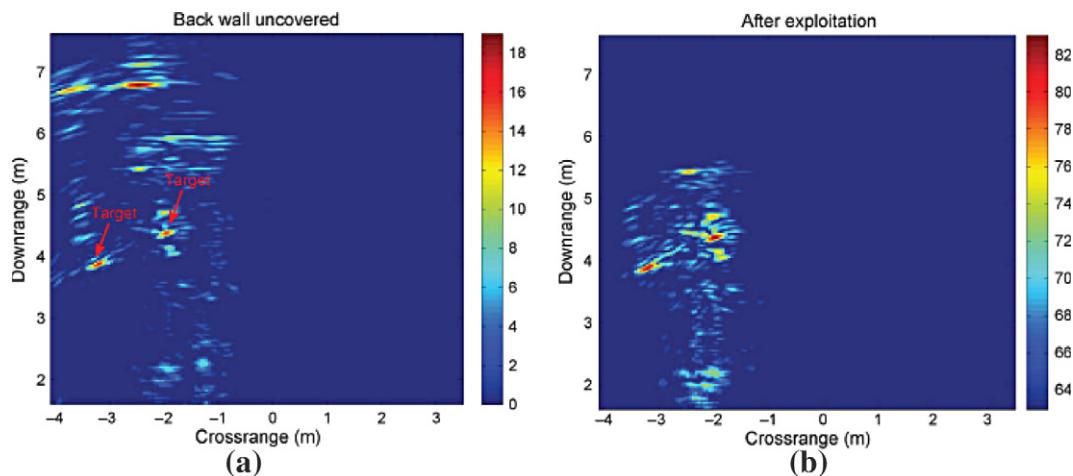
Scene Layout.

**FIGURE 17.17**

Single target experiment: (a) original, (b) after multipath exploitation.

are $D_1 = 5.09\text{m}$ and $D_2 = 3.78\text{m}$. The side walls and part of the floor were covered with RF absorbers and only multipath due to the back wall was considered. The front wall is made of 0.15m thick solid concrete blocks with a dielectric constant of 7.66. The synthetic aperture consisted of an 81-element monostatic linear array with an inter-element spacing of 2.22cm . The standoff distance of the array from the front wall is 1.53m . The experimental set up is depicted in Figure 17.16.

Two different experiments were conducted. In the first experiment, a 0.35m diameter metal sphere, labeled as "Target 1" in Figure 17.16, was used as a target. The second experiment consisted of

**FIGURE 17.18**

Two target experiment: (a) original, (b) after multipath exploitation.

two targets, the first being the sphere used in the single target case, while the second, labeled as “Target 2” in Figure 17.16, was a smaller metallic sphere of 0.2032 m diameter. An empty scene measurement without the targets present in the scene was also made and was coherently subtracted from the target scene data to reduce the clutter. The beamformed image corresponding to the single target scene is depicted in Figure 17.17a; both the target and the ghost are clearly visible in the image. In addition, other target-room interactions, such as those resulting from multiple bounces within the front wall, can be observed. The result after multipath exploitation is provided in Figure 17.17b. The variance of the weighting functions in the exploitation algorithm was chosen to be equal to the system’s range resolution. We observe that as expected, the ghost has been correctly mapped back to the target, while the other target-room interactions are left as is in the exploited image. Figure 17.18a shows the beamformed image corresponding to the two target scene. Both targets as well as the two ghosts with respect to the back wall are clearly seen. The multipath exploited image is provided in Figure 17.18b, in which one can see that the ghosts have been mapped back to the target locations. We note that, similar to the single target case, the other target-room and target-target interactions are retained in Figure 17.18b.

2.17.4 Change detection based MTI approach

Detection of humans is one of the most important objectives in urban sensing and through-the-wall radar technology [13, 14, 27, 73, 80–82]. Humans belong to the class of animate objects which are characterized by motion of the limbs, breathing, and heartbeat. These features separate animate and inanimate objects and allow the detection of targets of interest to proceed based on changes in the phase of the scattered radar signals over successive probing and data observations.

Change detection techniques have been recently used to detect moving targets in the presence of heavy clutter that is caused by strong reflections from exterior and interior walls. In the case of moving targets, the subtraction of two consecutive or non-consecutive data frames or images enjoys the same benefits as the background subtraction process for the case of stationary targets. That is, it results in effective removal of all wall returns and non-target clutter. The subtraction operation performed for either case of stationary or moving targets is referred to as Change Detection (CD). The length of the time period elapsing between the two datasets to be subtracted may differ for the two cases. While the stationary scene permits long time periods, the moving target case necessitates short periods. When operating on data frames, both operations can be described by what is known as delay line cancelers (DLC) [83].

It is known that moving target indication (MTI) processing applies clutter filters to remove radar returns scattered from stationary objects. Delay line cancelers can be designed such that their frequency responses place a notch at DC and concurrently meet other desirable passband and stopband filter characteristics [83]. Doppler filter banks typically follow the delay line canceler, and provide benefits of signal separations, radial velocity measurements, and noise reduction [83]. It is important to note that for urban sensing environments, changes in the backscattered signal phase due to motion do not necessarily lend themselves to Doppler frequency shifts. The human motion can be abrupt and highly nonstationary, producing a time-dependent phase whose rate of change may fail to translate into a single shift or multi-component sinusoids that can be captured by different Doppler filters. Instead, the corresponding wide spectrum of human motions becomes non-localizable and can span the entire radar frequency band. In lieu of Doppler filters, time-frequency processing can be applied to reveal the instantaneous frequency signatures, as it is the case with Human gait. Human gait classification and biometric radar have been discussed in [30,84–87]. However, apart from regularized motions, such as walking and running, time-frequency Doppler signal representations are very complex and difficult to interpret, especially when dealing with non-homogeneous walls. Therefore, the application of Doppler and Microdoppler filters for indoor target surveillance may not significantly add to target motion detection that is already achieved by delay line cancelers.

Change detection in through-the-wall radar imaging was first discussed concurrently in [88] and [89]. In [88], the Synchronous Impulse Reconstruction (SIRE) radar system, operating in the frequency band of 300 MHz–3 GHz and employing a multi transmit/receive design, was used for MTI. The change detection approach discussed in [88] subtracts the down range profiles over consecutive data frames, emulating a derivative operator. It was shown that the SIRE successfully detects a human target moving within an enclosed structure. On the other hand, change detection was applied in [89] in the context of background subtraction to detect stationary targets using data acquired during interrogations of a scene at two different time instants. Image intensity subtraction was employed and examples based on ray tracing data covering 0.7–3.1 GHz frequency range were provided. More recently [90], applied change detection for slow-moving target detection in through-the-wall radar imaging (TWRI) by subtracting the image intensity corresponding to different data frames, followed by a zero threshold operation to suppress the reference image. Examples of target detection in [90] included both EM modeling data and real experiments. Change detection using first- and second-order motion detection images for TWRI application were discussed in [91,92]. The former has better sensitivity to motion, whereas the latter provides better signal-to-noise ratio (SNR). Experimental data, reported in [91,92] using four antennas and a signal bandwidth of 500 MHz–1 GHz, demonstrated that change detection techniques

are capable of detecting walking human and simulated human breathing using machine motion. Moving target indicator filtering was also used in [93] for detection of humans in motion inside buildings. An ultrawideband frequency-modulated continuous-wave radar with an extended frequency sweep from 0.5 to 8 GHz was used in the differential mode to track human motion behind a brick wall.

In this section, we examine change detection formulation and performance when clutter removal is performed either coherently or noncoherently. Rather than operating on successive pulses, delay line canceler equivalences are applied to different data frames for each range bin (or equivalently different complex amplitude images for each pixel) (coherent CD) or different intensity images for each pixel (noncoherent CD). The frames can be consecutive, dealing with targets exhibiting sudden short motions, or nonconsecutive with relatively long time difference, for the case in which the target changes its range gate position.

2.17.4.1 Signal model

We develop the signal model for wideband operation with multiple transmitters and receivers under coherent integration. Sequential use of the transmitters with simultaneous reception at all receivers is assumed. It is noted that time-multiplexing of the transmitters is a viable option for coherent radar operation in through-the-wall radar sensing applications. This is due to the following two main reasons: (a) Ground-based urban imaging radar systems are typically vehicle portable, which restricts the available real estate for deploying multiple antennas on the same platform; this, combined with the important constraint of low cost, results in deployment of far more receivers than wideband transmitters, and (b) Indoor targets of interest, such as humans, move at low velocities. As a result, the time required to collect one data frame with sequential use of the transmitters is small enough so that the target remains stationary during the measurement. The sequential transmit operation is a salient feature of two known through-the-wall radar imaging systems; one is built by the Army Research Lab [82,88], and the other by the Defense Research and Development Canada [48].

Consider an imaging radar with an M -element linear transmit array and an N -element linear receive array. Let $s(t)$ be the wideband signal used for interrogating the scene. The signal duration T_s is assumed to be less than the pulse repetition interval T_r . Assuming coherent integration of K pulses, the coherent integration interval becomes $K T_r$. For the case of a single point target of reflectivity σ_p , located at $\mathbf{x}_p = (x_p, y_p)$, the k th emitted pulse with the m th transmitter active is received at the n th receiver in the form

$$z_{mn}(t, k) = \sigma_p s(t - m K T_r - k T_r - \tau_{p,mn}), \quad (17.62)$$

where $\tau_{p,mn}$ is the propagation delay for the signal to travel between the m th transmitter, the target at \mathbf{x}_p , and the n th receiver. For through-the-wall propagation, $\tau_{p,mn}$ will comprise the components corresponding to traveling distances before, through, and after the wall [40]. The term $m K T_r$ in Eq. (17.62) indicates the additional time required for data collection when the m th transmitter becomes active in a sequential order. We assume that the transmitters transmit their pulses in contiguous time intervals, each of length $K T_r$. It is further assumed that we deal with targets of constant radar cross section (RCS) over the range of the used frequencies and target aspect angles. The received signal is passed through a filter matched to the transmit waveform and the output can be expressed as

$$z_{mn}^{MF}(t, k) = z_{mn}(t, k) * h(t) = \int_0^{T_r} h(u) z_{mn}(t - u, k) du, \quad (17.63)$$

where $h(t) = s^*(-t)$. With K pulses coherently integrated, we obtain

$$z_{mn}^{CI}(t) = \sum_{k=0}^{K-1} z_{mn}^{MF}(t, k). \quad (17.64)$$

In order to generate an image of the scene being interrogated, the MN outputs corresponding to the operation of M transmitters and N receivers are processed as follows. The region of interest is divided into a finite number of pixels in x and y , where x and y represent the crossrange and downrange, respectively. The composite signal corresponding to the pixel, located at $\mathbf{x}_q = (x_q, y_q)$, is obtained by summing time delayed versions of the MN filtered signals,

$$z_q(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} z_{mn}^{CI}(t + \tau_{q,mn}), \quad (17.65)$$

where $\tau_{q,mn}$ is the focusing delay applied to the matched filtered output of the n th receiver with the m th transmitter active. It is noted that additional weighting can be applied during the summation operations of (17.65) to control the sidelobe level of the transmit-receive array pattern of the imaging radar system. Substituting (17.62)–(17.64) in (17.65) yields

$$z_q(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} \int_0^{T_r} h(u) z_{mn}(t + \tau_{q,mn} - u, k) du. \quad (17.66)$$

The complex amplitude image value $I(\mathbf{x}_q)$ for the pixel at \mathbf{x}_q is obtained by sampling the signal $z_q(t)$ at time $t = 0$. That is,

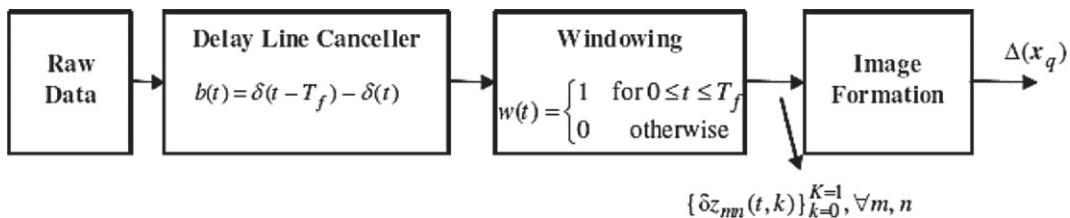
$$I(\mathbf{x}_q) = z_q(t)|_{t=0}. \quad (17.67)$$

The process described by (17.65)–(17.67) is repeated for all pixels in the image to generate the composite image of the scene. The general case of multiple targets can be obtained by superposition of target reflections [40, 41].

2.17.4.2 Change detection

In its simplest form, the MTI processing is achieved by using the radar images corresponding to two data frames, which may be consecutive or separated by one or more data frames. Due to time multiplexing of transmitters, a data frame length T_f is equal to M coherent integration intervals, i.e., $T_f = MKT_r$. Let $\{I^{(t)}(\mathbf{x}_q)\}_{q=0}^{Q-1}$ and $\{I^{(t+LT_f)}(\mathbf{x}_q)\}_{q=0}^{Q-1}$ be the radar images corresponding to the data collected over the two frames, where Q is the number of pixels in each image and L denotes the number of frames between the two acquisitions. It is noted that $L = 1$ represents the case when the two acquisitions are performed over consecutive frames. The timing interval for each data frame is assumed to be a fraction of a second so that the moving target appears stationary during data collection.

Change detection can be performed in two ways, namely, coherent CD and noncoherent CD [88–92]. In the former approach, complex amplitude image values corresponding to the two data frames are subtracted on a pixel by pixel basis. The latter approach involves subtraction of image pixel intensity

**FIGURE 17.19**

Block diagram of the data domain CD approach, equivalent to complex amplitude image subtraction (coherent CD).

values corresponding to different data frames, followed by a zero threshold operation to suppress the reference image. For the coherent approach, the motion detection image can be expressed as follows,

$$\Delta(\mathbf{x}_q) = I^{(t+LT_f)}(\mathbf{x}_q) - I^{(t)}(\mathbf{x}_q), \quad q = 0, 1, \dots, Q-1. \quad (17.68)$$

Using (17.66) and (17.67), we can rewrite (17.68) as

$$\Delta(\mathbf{x}_q) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} \int_0^{T_r} h(u) \delta z_{mn}(t + \tau_{q,mn} - u, k) du \Big|_{t=0}, \quad (17.69)$$

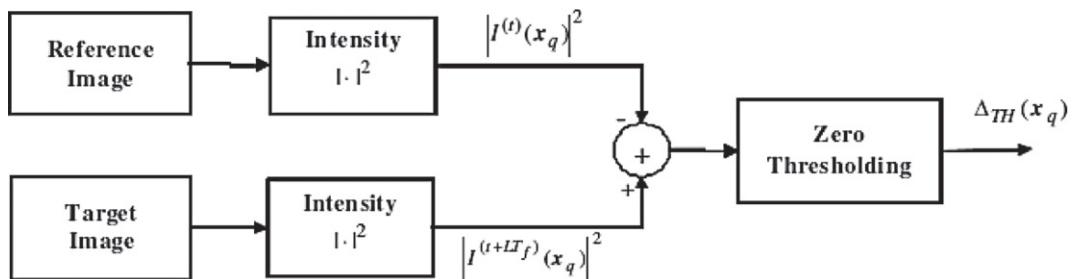
where

$$\delta z_{mn}(t, k) = z_{mn}(t - LMKT_r, k) - z_{mn}(t, k) = z_{mn}(t, k) * [\delta(t - LMKT_r - kT_r) - \delta(t - kT_r)] \quad (17.70)$$

and $z_{mn}(t, k)$ is defined in (17.62). From (17.68)–(17.70), we observe that the subtraction of complex amplitude image values for each pixel amounts to subtraction of the received data corresponding to the two frames for each range bin, followed by image formation. Equation (17.70) provides the equivalent delay line canceler representation for change detection directly in the data domain. A block diagram of the data domain CD algorithm implementation is provided Figure 17.19. Note that the windowing operation in the figure ensures application of the algorithm over data frames of length T_f . The block diagram of Figure 17.19 is a two-frame delay line canceler, reminiscent of the commonly used two-pulse DLC for clutter filter in range-Doppler radars [83]. More frames can be used in CD leading to multiple frames delay line canceler, which is similar to three- or higher-pulse DLC. However, unlike range-Doppler radar systems, in which higher order clutter filters provide flexibility in trading off clutter suppression with Doppler frequency magnitude, there are no clear advantages of using multiple frames in CD for the underlying problem. Hunt suggested averaging multiple frames, only to emphasize the stationary targets [92].

The noncoherent CD approach involving intensity images is depicted in Figure 17.20. In this case, the intensity images corresponding to the two data frames are first computed. Then, a difference image is generated by a pixel by pixel subtraction of the intensity images as

$$\Delta(\mathbf{x}_q) = |I^{(t+LT_f)}(\mathbf{x}_q)|^2 - |I^{(t)}(\mathbf{x}_q)|^2, \quad q = 0, 1, \dots, Q-1. \quad (17.71)$$

**FIGURE 17.20**

Block diagram of the noncoherent change detection approach.

The final motion detection image is obtained by a zero threshold operation

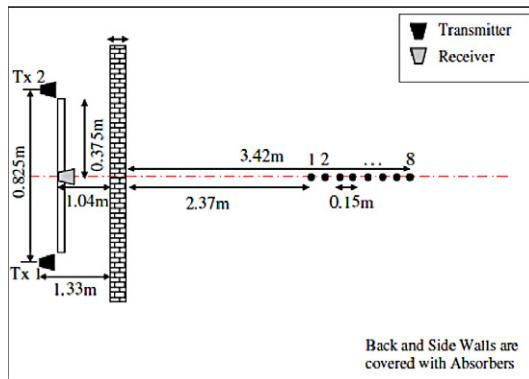
$$\Delta_{TH}(\mathbf{x}_q) = \begin{cases} \Delta(\mathbf{x}_q), & \text{if } \Delta(\mathbf{x}_q) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (17.72)$$

This threshold results in the target appearing in the final motion detection image only at the position it was occupying during acquisition of the second (test) data frame, since the target image at position corresponding to the first (reference) data frame appears as a negative value after intensity subtraction and is thus removed. It is noted that, unlike the complex amplitude image subtraction, there is no equivalent data domain representation for the intensity image subtraction or noncoherent CD approach due to the nonlinearity of the modulus function in (17.71). However, the delay line canceler framework can still be applied. The image intensity at each pixel can be viewed as a nonlinear function of the data in the respective frame, and the subtraction, implied by the DLC, is performed between two corresponding pixels separated by one or multiple frames.

The aforementioned change detection schemes, though presented for motion detection, can also be applied for detection of stationary targets in the presence of clutter and noise [89]. If a target-free background image of the scene being interrogated is available (feasible when the radar system is being used for long term surveillance), it can be subtracted (complex amplitude or intensity pixel subtraction) from an image of the same scene acquired at a later time that may contain a target of interest, thus rendering background subtraction as a special case of change detection.

2.17.4.3 Experimental results

A wideband through-the-wall imaging radar system was set up in the Radar Imaging Lab at Villanova University. A stepped-frequency CW signal, consisting of 101 frequency steps of size 10 MHz, covering 1.5–2.5 GHz band was chosen for imaging. The bandwidth of 1 GHz provides a range resolution of 15cm and the 10 MHz frequency step allows 15m unambiguous range, which is roughly three times the length of the room being imaged. An Agilent network analyzer, model ENA 5071B, was used for signal synthesis and data collection. A horn antenna, model ETS-Lindgren 3164-04, with an operational bandwidth from 0.7 to 6 GHz, was mounted on a Field Probe Scanner to synthesize an 11-element uniform receive line array with an inter-element spacing of 7.5cm. Two horn antennas (model ETS-Lindgren 3164-04), mounted on tripods and placed slightly above and on either side of the receive array,

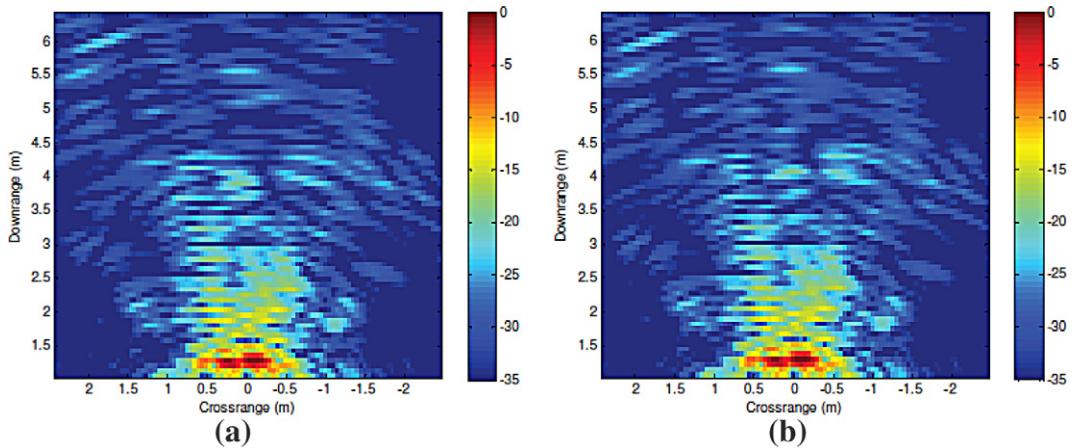
**FIGURE 17.21**

Layout of the scene and target trajectory.

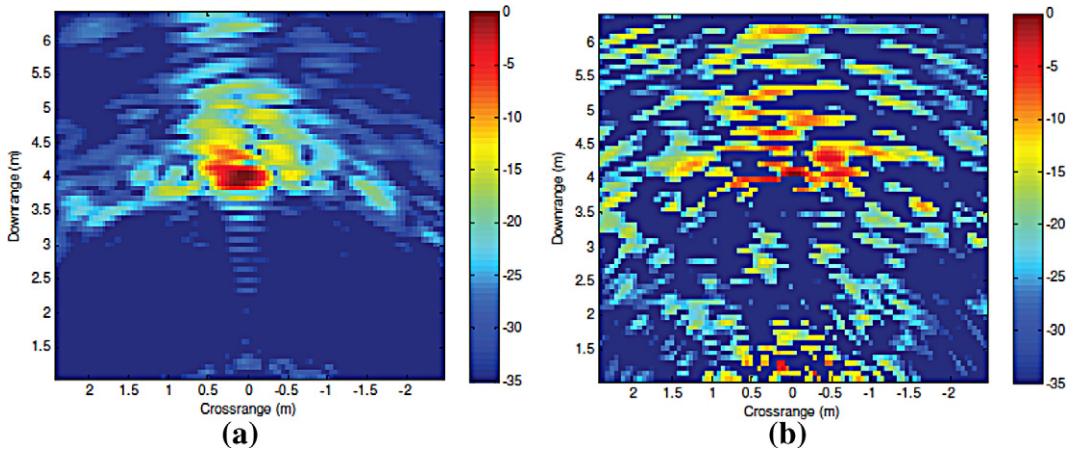
were used as transmitters. Each transmitter was spaced 3.75 cm away from the nearest receiver. Thus, for the system used, $M = 2$ and $N = 11$. A 10 ft \times 8 ft wall segment was constructed utilizing 0.14 m thick solid concrete blocks with a dielectric constant of 7.66. The receive array was at a standoff distance of 1.05m from the wall, whereas the transmitters were positioned 1.33m in downrange from the wall.

In order to illustrate the performance of change detection under complex and intensity image delay line cancelers, we consider the following experiment. A human walks away from the front wall in an empty room, following a straight line path. Since simultaneous data cannot be collected at all receive locations due to the synthetic nature of the receive aperture, eight different target positions were considered, separated by 15 cm, along the trajectory, as illustrated in Figure 17.21. Data was collected for all eight positions, with the target stationary at each position, using time multiplexing of the transmitters. At each position, coherent integration was performed over 16 measurements. That is, $K = 16$ for the experiment.

We first consider the case where the person moves from position 3 to position 4. That is, the two acquisitions are performed over consecutive frames and $L = 1$. Figure 17.22a and b show the beamformed images using the data corresponding to positions 3 and 4, respectively. Both images show strong reflections from the front wall, rendering target returns hard to detect in either image. Figure 17.23a shows the result of coherent change detection wherein the complex radar image of Figure 17.22a is subtracted from the complex image of Figure 17.22b. We observe that all stationary background including the front wall has been eliminated from the image, and the target undergoing motion has been detected. Since the target appears in two consecutive range cells, no clear target displacement is observed and the target image appears blurred. The corresponding noncoherent CD results under intensity image DLC are presented in Figure 17.23b. We observe that the intensity image DLC removes most of the imaged target in the case of consecutive range cell acquisitions. Moreover, due to the nonlinearity of the modulus function, the stationary background in the image is in general different for each data acquisition, thereby causing the background to be only partially canceled out under intensity image DLC [90]. As a result, stronger artifacts than those of the complex image DLC can be clearly observed. The ability

**FIGURE 17.22**

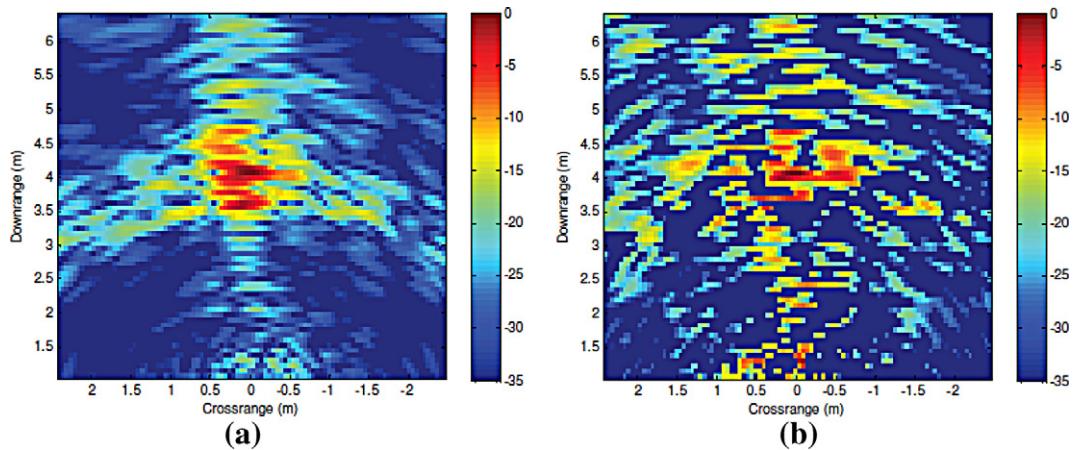
Beamformed images for the target at position (a) 3 and (b) 4 (consecutive range cells).

**FIGURE 17.23**

Images after change detection under (a) coherent CD and (b) noncoherent CD for the human walking from position 3 to position 4.

to detect the target in the change detected image is, therefore, reduced under intensity image DLC as compared to complex image DLC.

Next, we investigate the performance of the various schemes under large target displacement between consecutive acquisitions. Consider the case where the person moves from position 1 to position 4. In this case, $L = 3$. Figure 17.24a shows the result of coherent change detection. Since the target

**FIGURE 17.24**

Images after change detection under (a) coherent CD and (b) noncoherent CD for the human walking from position 1 to position 4.

undergoes large range migration, the moving target presents itself as two targets in the image. The noncoherent change detection results for the large target displacement case are presented in Figure 17.24b. We observe that the intensity image DLC retains the target at position 4 and removes the target at position 1. However, similar to the small target displacement case, the artifacts resulting from the background not being totally canceled out render target detection difficult.

2.17.5 Compressive sensing approach to moving target indication for urban sensing

The most desirable goal of a through-the-wall radar system is to provide situational awareness in a fast and reliable manner. This goal is primarily challenged due to increasing demands on the radar system to deliver high resolution images in both range and cross-range, which requires use of wideband signals and large array apertures, respectively. In addition, the presence of multipath and clutter can significantly contaminate the radar data and compromise the main intent of providing enhanced system capabilities for imaging of building interiors and tracking of targets behind walls.

Most radar data acquisition systems acquire samples in frequency (or time) and space, and then apply compression to reduce the amount of stored information. This approach has three inherent inefficiencies. First, as the demands for high resolution and more accurate information increase, so does the number of data samples to be recorded, stored, and subsequently processed. Second, there are significant data redundancies not exploited by the traditional sampling process. Third, it is wasteful to acquire and process data samples that will be discarded later. Further, in radar imaging systems, whether using time or stepped-frequency based pulsing, along data collection leads to image degradation. This is because,

unless all objects in the scene are stationary during the entire data collection operations, image smearing of targets would occur.

Compressive sensing (CS) is a very effective technique for scene reconstruction from a relatively small number of data samples without compromising the imaging quality [94–98]. In general, the minimum number of data samples or sampling rate that is required for scene image formation is governed by the Nyquist theorem. However, when the scene is sparse, compressed sensing provides very efficient sampling, thereby significantly decreasing the required volume of data collected. Towards the objective of providing persistent surveillance in urban environments, such techniques will yield reduced cost, simplified hardware, and efficient sensing operations that allow super-resolution imaging of sparse behind-the-wall scenes. CS has been successfully applied to TWRI of moving and stationary targets [97, 99].

In this section, we exploit full benefits of the compressive sensing to aid in fast data acquisition in wideband through-the-wall radar imaging systems for moving target detection and localization. Change detection is first used for removal of stationary background (clutter and stationary targets). Since the removal of stationary background via change detection converts a populated scene into a sparse scene of moving targets, reduction in data volume is then pursued under the framework of compressive sensing.

2.17.5.1 Signal model

Assume that there are N antenna locations constituting the synthetic aperture and a stepped-frequency signal of M narrowband signals. The signal received at the n th antenna location with the m th frequency is given by,

$$y(n, m) = \sum_{p=0}^{P-1} \sigma_p \exp(-j2\pi f_m \tau_{n,p}) + B(n, m), \quad (17.73)$$

where σ_p is the reflection coefficient of the p th moving target, $\tau_{n,p}$ is the two-way propagation delay for the signal to travel from the n th antenna to the p th moving target, and $B(n, m)$ represents the contribution of the stationary background. For free-space propagation, $\tau_{n,p}$ is proportional to the Cartesian distance between the n th antenna and the p th target, whereas in the presence of the wall, it contains contributions from paths in air and through the wall [100].

The MTI processing is achieved by collecting multiple data intervals over a period of time and buffering the corresponding stepped-frequency data measured at each antenna location [101]. Let I be the number of collected data intervals and $y^{(i)}(n, m)$, $n = 1, \dots, N$, $m = 1, \dots, M$, define the measured stepped-frequency dataset for the i th interval. It is noted that the timing interval for each dataset is a fraction of a second so that the moving target appears stationary during data collection. Change detection is applied to the measured datasets as follows,

$$\Delta^{(i)}(n, m) = y^{(i+1)}(n, m) - y^{(i)}(n, m), \quad (17.74)$$

where $i = 1, \dots, I - 1$, forming the set of difference signals

$$\{\Delta^{(1)}(n, m), \Delta^{(2)}(n, m), \dots, \Delta^{(I-1)}(n, m)\}. \quad (17.75)$$

Using (17.73) and (17.74), the i th difference signal can be expressed as

$$\Delta^{(i)}(n, m) = \sum_{p=0}^{P-1} \left\{ \sigma_p \exp \left(-j 2\pi f_m \tau_{n,p}^{(i+1)} \right) - \exp \left(-j 2\pi f_m \tau_{n,p}^{(i)} \right) \right\} \quad (17.76)$$

The component of the radar return from the stationary background is the same over the i th and $(i + 1)$ th time intervals, and is thus removed from the difference signal. For moving targets, two situations may arise. In the first case, the target undergoes large range migration, whereas in the second case, the target undergoes small range migration, i.e., moves into the next range cell. In the first case, as seen from (17.76), the moving target will present itself as two targets, whereas in the second case, the target will appear as spread. We will focus on the former situation in this chapter.

2.17.5.2 Compressive sensing for MTI

Consider the i th difference dataset. For notational simplicity, we drop the superscript “ (i) .” Let Δ_n be the $M \times 1$ difference signal vector corresponding to the n th antenna location. For large range migration, an equivalent signal representation with $Q = 2P$ targets can be used. Thus, with the region of interest divided into L pixels, the relationship between the difference signal Δ_n and the pixel locations can be expressed as [96, 97]

$$\Delta_n = \mathbf{A}_n s, \quad (17.77)$$

where

$$\begin{aligned} \Delta_n &= [\Delta(n, 0) \Delta(n, 1) \cdots \Delta(n, M-1)]^T, \quad [\mathbf{A}_n]_{ml} = \exp(-j 2\pi f_m \tau_{n,l}), \\ [\mathbf{s}]_l &= \begin{cases} \tilde{\sigma}_q, & \tau_{n,l} = \tau_{n,q}, \\ 0, & \tau_{n,l} \neq \tau_{n,q}. \end{cases} \end{aligned} \quad (17.78)$$

In (17.78), $\tilde{\sigma}_q$ is the reflectivity of the q th target and $\tau_{n,l}(\tau_{n,q})$ is the two-way signal traveling time between the n th antenna location and l th pixel’s (q th target’s) location. The \mathbf{A}_n matrix is similar to the Fourier matrix which relates the frequency and time representations [97].

Given $\bar{\Delta}_n$ for $n = 1, \dots, N$, which is a $K (< M)$ dimensional vector consisting of elements randomly chosen from Δ_n , we can recover s by solving the following equation:

$$\hat{s} = \arg \min_s \|s\|_{l_1} \text{ subject to } \|\bar{\mathbf{A}}s - \bar{\Delta}\|_{l_2} < \delta, \quad (17.79)$$

where δ represents the tolerance error, and

$$\begin{aligned} \bar{\mathbf{A}} &= [\bar{\mathbf{A}}_0 \bar{\mathbf{A}}_1 \dots \bar{\mathbf{A}}_{N-1}]^T, \quad \bar{\Delta} = [\bar{\Delta}_0 \bar{\Delta}_1 \dots \bar{\Delta}_{N-1}]^T, \\ \bar{\Delta}_n &= [\Delta(n, i_0) \Delta(n, i_1) \cdots \Delta(n, i_{K-1})]^T, \quad [\bar{\mathbf{A}}_n]_{ml} = [\mathbf{A}_n]_{i_k l}. \end{aligned} \quad (17.80)$$

In the above equation, $\bar{\Delta}$ and $\bar{\mathbf{A}}$ represent respective concatenations of $\bar{\Delta}_n$ and $\bar{\mathbf{A}}_n$, and $i_k \in [0, 1, \dots, M-1]$ for $k = 0, \dots, K-1$, are the indices of the randomly measured frequencies.

A stable solution of the sparse signal recovery problem in (17.79) is guaranteed provided that $\bar{\mathbf{A}}$ satisfies the Restricted Isometry Property (RIP), which states that all subsets of r columns taken from

$\bar{\mathbf{A}}$ are, in fact, nearly orthogonal, r being the sparsity of the signal $\bar{\mathbf{A}}$ [102]. In practice, it is difficult to check this property. Other related measures on the matrix $\bar{\mathbf{A}}$, such as mutual coherence, are often used to guarantee stable recovery through l_1 -minimization [102]. Mutual coherence of the columns of $\bar{\mathbf{A}}$ can be viewed as the largest off-diagonal entry of the Gram matrix $\bar{\mathbf{A}}^H \bar{\mathbf{A}}$, where the columns of $\bar{\mathbf{A}}$ have been normalized. The matrix $\bar{\mathbf{A}}$ is considered to be incoherent if the value of the mutual coherence is small. For stepped-frequency systems, the coherence between the columns of $\bar{\mathbf{A}}$ is reduced by randomly distributing the selected frequency bins over the entire frequency band [96, 97, 103].

Selecting a proper tolerance error δ is also very important for stable recovery of the scene being imaged. The tolerance error can be chosen using the cross-validation strategy [96, 104], which does not

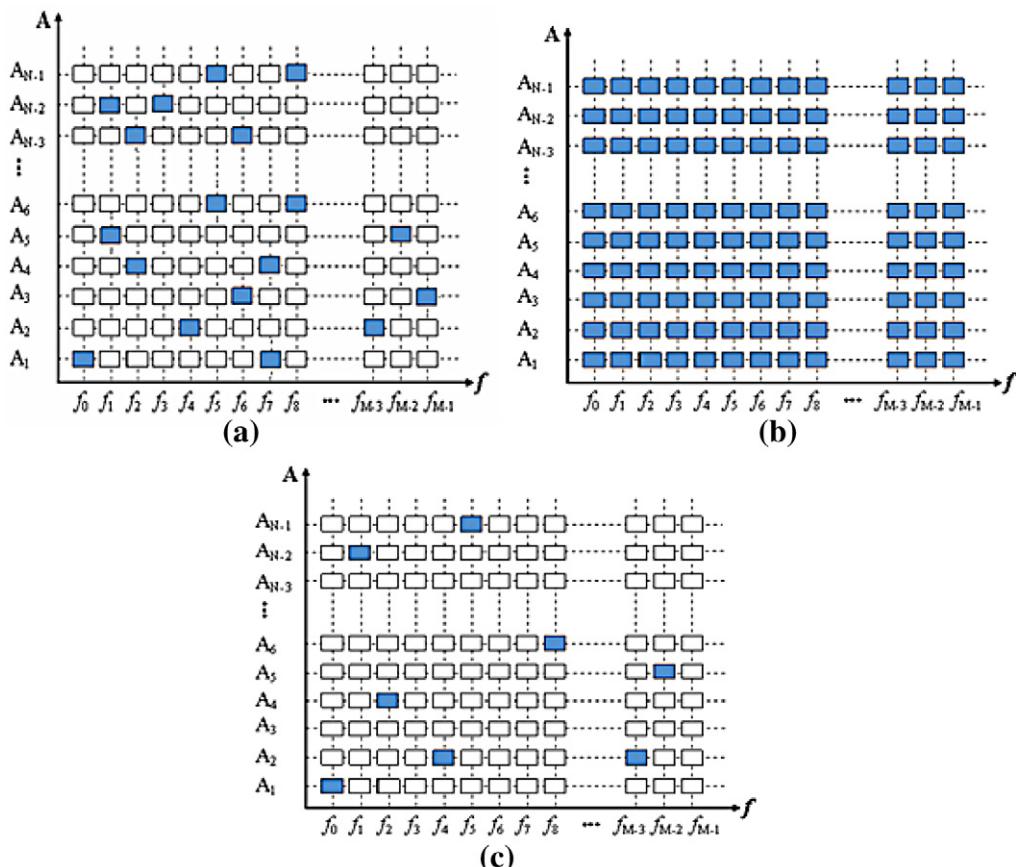


FIGURE 17.25

Data collection for (a) compressed sensing along frequency only, (b) conventional high-resolution radar, and (c) compressive sensing along both antennas and frequency. The filled rectangles indicate the measured data samples.

require any knowledge or estimates of the noise statistics. This method depends on separating the data measurements into estimation and cross-validation sets. The CS method is applied to the estimation dataset with an initial selection of δ and the method's result is tested on the cross-validation dataset. As the algorithm iterates, the prediction performance in the cross-validation set increases. When the method starts to overfit the estimation dataset, which means estimating part of the noise, performance in the cross-validation set decreases. Further decrease in δ is not beneficial and the algorithm should be terminated. The cross-validation based algorithm consists of the following steps, where the subscript "E" denotes the estimation set, and "CV" the cross-validation set:

- i. *Initialize:* Set $\delta = \|\bar{\Delta}_{CV}\|_{l_2}$ and $i = 1$.
- ii. *Estimate:* Solve (17.79) to estimate the target locations $\hat{s}(i)$ with the estimation data set $\bar{\Delta}_E$, $\bar{\Delta}_E$.
- iii. *Cross-validate:* If $\|\bar{\Delta}_{CV} - \bar{\Delta}_{CV}\hat{s}(i)\|_{l_2} < \delta$, then set $\delta = \|\bar{\Delta}_{CV} - \bar{\Delta}_{CV}\hat{s}(i)\|_{l_2}$, increment i and iterate from Step ii. Otherwise, terminate the iteration.

Equations (17.79) and (17.80) represent one strategy that can be adopted for compressive sensing based MTI approach, wherein frequencies are chosen randomly for all the antenna locations constituting the array aperture. Figure 17.25a depicts the data collection scheme for the employed CS imaging strategy, wherein the horizontal axis is the frequency, the vertical axis represents the antenna position, and the dark boxes represent the sampled data. For comparison, the data collection scheme for conventional high-resolution radar is also provided in Figure 17.25b, which reiterates that all the frequency bins are transmitted and received at all antenna locations. Equations (17.79) and (17.80) can also be extended so that random measurements include both antennas and frequencies [96,97]. That is, a few frequency bins are collected at few antenna locations, as shown in Figure 17.25c.

2.17.5.3 Experimental results

A wideband SAR system was set up in the Radar Imaging Lab at Villanova University. A stepped-frequency CW signal, consisting of 201 frequency steps of size 10 MHz, covering 1–3 GHz band was chosen for imaging. An Agilent network analyzer, model ENA 5071B, was used for signal synthesis and data collection. A horn antenna, model ETS-Lindgren 3164-04, with an operational bandwidth from 0.7 to 6 GHz, was mounted on a Field Probe Scanner to synthesize a 41-element uniform linear array with an inter-element spacing of 3.745 cm, as shown in Figure 17.26a. A 3.05 m × 2.44 m wall segment was constructed utilizing 0.14 m thick solid concrete blocks with a dielectric constant of 7.66. The array was at a standoff distance of 1.06 m from the wall. A 5 Gallon jug filled with salt water (emulating a human) was used as the target, as shown in Figure 17.26b. In order to emulate a moving target, two datasets were collected with the target at two different positions. Since simultaneous data cannot be collected at all antenna locations due to the synthetic nature of the array aperture, data were collected for the two positions, with the target stationary at each position. The target positions were selected as (0.8, 1.83) m and (0.25, 2.2) m, so as to emulate large range migration of the target between the two measurements.

The region to be imaged is chosen to be 4.8 m × 5.3 m centered in the middle of the scene beyond the wall and divided into 33 × 57 pixels in crossrange and downrange. Figure 17.27a shows the result of change detection wherein the full dataset (201 × 41 samples) corresponding to target position 1 is subtracted from that for position 2, followed by conventional image formation. We observe that all background returns have been eliminated from the image and the target undergoing motion has been

detected. Moreover, as expected, the target has presented itself at the two locations corresponding to positions 1 and 2.

For the CS approach, instead of measuring all 201 frequencies, we use a random subset of 7 frequencies only at each of the 41 antenna locations. According to CS theory, an r -sparse L -dimensional signal can be recovered from $O(r \log(L))$ measurements [95]. Considering the system resolution and for the salt water jug extent of 0.27 m and 0.27 m in crossrange and downrange, respectively, each target will approximately occupy only 6 out of 1881 image pixels. For the large range migration case under consideration, the 41×7 measured data points well exceed this requirement ($(2 \times 6) \log(1881) \approx 90$). The CS reconstructed image is shown in Figure 17.27b. For comparison, the beamformed image with the same reduced random dataset as CS is provided in Figure 17.27c. From Figure 17.27b, we observe that the CS approach localizes the target at positions 1 and 2 accurately. On the other hand, as seen in Figure 17.27c, beamforming with the reduced dataset has degraded image quality with significantly more false targets. A similar experiment without the absorbers on the side and back walls was conducted and shown to reduce the sparsity of the scene due to the presence of “ghosts.” The latter stems from the target-wall reflections and multipath returns. Compressive sensing can still be effective in localizing the target and its ghosts, but requires an increase in the number of data samples compared to the case of “no ghosts.”

2.17.6 Applications

Through-the-wall radar imaging has a variety of applications in both civil and military paradigms. It has been successfully sought out for surveillance and reconnaissance in urban environments, requiring not only the layout of the building, including types and locations of walls, but also detection and localization of both moving and stationary targets within enclosed structures [1, 80, 105]. This technology can also be used in rescue missions, searching for fire, earthquakes, and avalanche victims and survivors, behind-

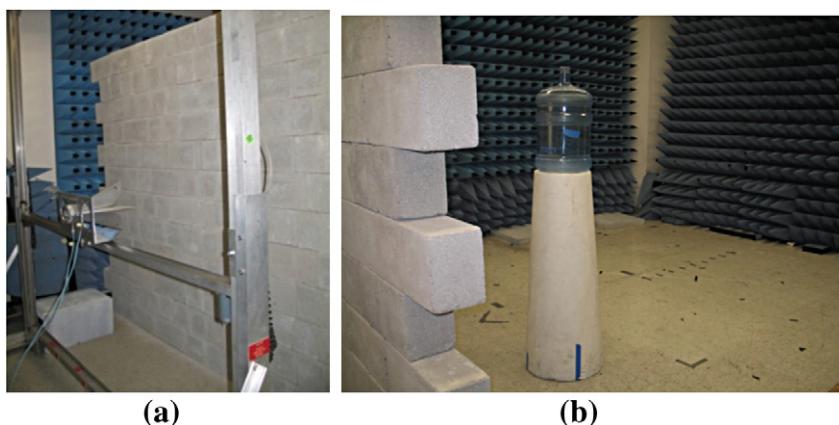
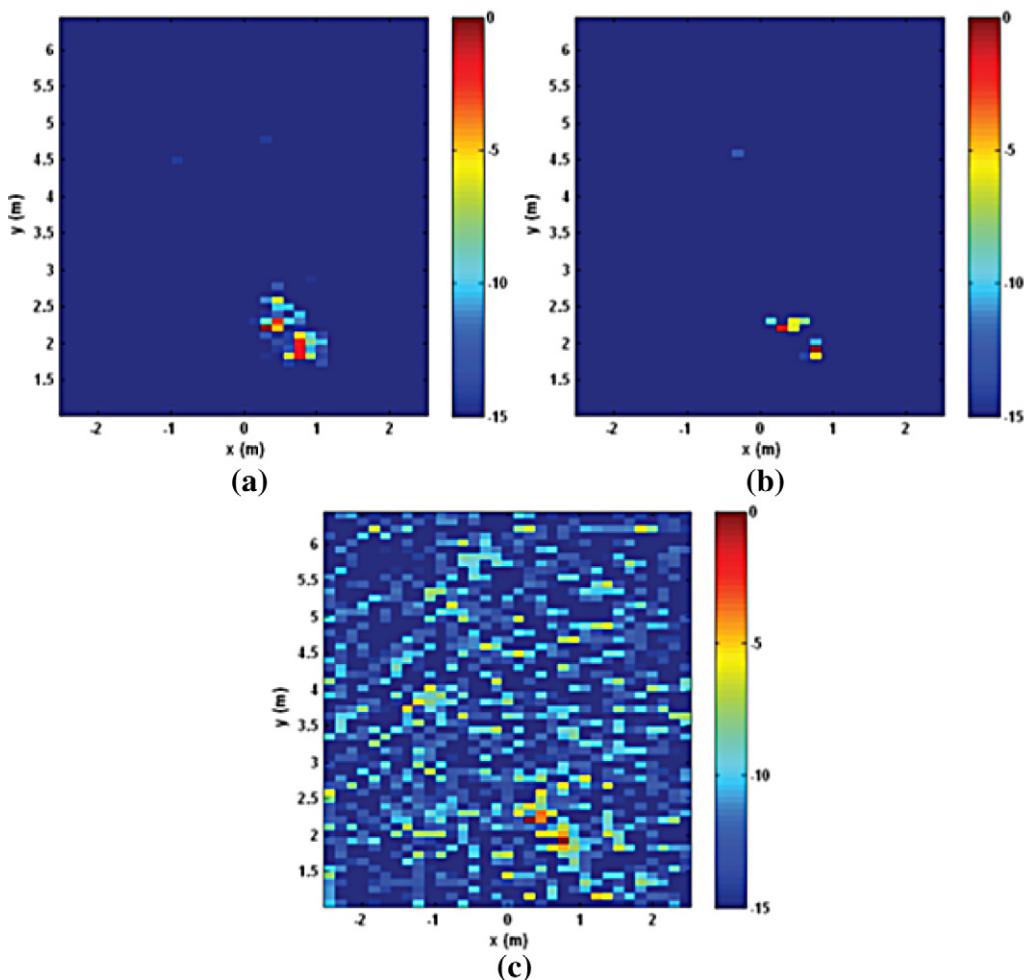


FIGURE 17.26

(a) Wideband SAR system, (b) scene layout.

**FIGURE 17.27**

Change detection results using data from position 1 and position 2. (a) Beamforming with full difference dataset; (b) CS result with limited difference dataset; (c) beamforming with limited difference dataset.

the-wall detection and surveillance of suspected criminals and outlaws, and by law enforcement officers for locating hostages and their captors [20, 106].

2.17.7 Open issues and problems

There are many challenges still facing the TWRI technology. Imaging of stationary scenes using SAR is considered more difficult than MTI for moving targets. The difficulty arises from false positives and

ghosts, both are products of multipath. Further, when targets are close to walls, imaging algorithms may not be able to separate targets from wall responses, as they both tend to merge. Another important TWRI goal is target classification. The relatively limited bandwidth does not lead to sufficient resolution of key scatterers, rendering the target image as a cluster of contagious high value pixels, a situation that does not lend itself to effective feature extraction. Classifications of Human gait take advantage of the motion of the limbs and their respective microDoppler signatures. When performed behind walls, these signatures blur and it can be difficult to achieve low classification errors. Complex buildings, made of multiple floors, each with populated scenes present the ultimate challenge of TWRI technology, specifically with long standoff distances whether it is associated with ground-based or airborne systems. The latter, due to covertness and operation logistics, is becoming the cornerstone of recent research and development in this area, especially for defense applications [107].

2.17.8 Data sets

The Center for Advanced Communications at Villanova University has conducted several through-the-wall imaging experiments and collected a variety of datasets in a laboratory environment. The datasets include full-polarization free-space and through-the-wall collections under semi-controlled conditions with a stepped-frequency radar system. Targets consist of both calibrated reflectors (dihedrals, trihedrals, and spheres) as well as a number of common indoor objects (phone, computer, tables, chair, filing cabinet). In addition, a jug of saline solution is also used as a target to crudely approximate a human. These datasets are available for download at the following website: <http://www.villanova.edu/engineering/centers/cac/twri.htm>.

2.17.9 Conclusions

In this chapter, we presented recent algorithmic advances in through-the-wall radar imaging. First, considering ground-based EM sensing, we discussed two methods, namely, spatial filtering and EM modeling based solution, for mitigating the front wall return prior to application of the image formation methods. The spatial filtering technique builds on the strong correlation of wall EM responses across antenna array elements to reduce constant-type return that is typical of walls in monostatic illuminations. The spatial filter, thus, allows the follow-on beamforming to unmask and image behind-the-wall targets that have limited spatial extent, such as humans, compared to walls without the need for *a priori* knowledge of the wall characteristics. A simple IIR notch filter with flexible design was compared with the fixed design MA subtraction filtering, commonly used in GPR. On the other hand, the EM modeling approach is wall-dependent and relies on accurate estimation of the wall parameters for suppression of the wall return. Extraction of the wall parameters and coherent subtraction of the modeled wall reflection was shown to significantly improve the signal-to- wall-clutter ratio.

Second, an approach to exploit the rich indoor multipath environment for improved target detection was described. A ray tracing approach was used to derive a multipath model, considering reflections inside an enclosed room comprising four homogeneous walls. Using the model, it was demonstrated analytically that the multipath corresponding to each sensor appeared on the wall but changes position from one sensor to another. Hence, a least squares technique was used to estimate its actual focusing

location in both downrange and crossrange. The model was utilized to develop a multipath exploitation technique which associates multipath ghosts with their respective targets and maps them to their true target locations. This technique reduced the false positives in the original beamformed image as well as increased the signal-to-noise ratio at the true target locations.

Third, we discussed a change detection approach to moving target indication for through-the-wall applications. Change detection was used to mitigate the heavy clutter that is caused by strong reflections from exterior and interior walls. Both coherent and noncoherent change detection techniques were examined and their performance was compared under both consecutive and non-consecutive acquisitions. For non-consecutive acquisitions, the coherent CD scheme showed two sets of imaged targets corresponding to the positions of a single target at the two data acquisitions, whereas the noncoherent CD retained only one set of imaged targets, though with significantly more artifacts. For consecutive acquisitions, the coherent change detection provided better performance than the noncoherent CD, which removed most of the imaged target during the zero-thresholding step.

Finally, we identified and localized moving targets behind walls and inside enclosed structures using an approach that combines sparsity-driven radar imaging and moving target indication. The removal of stationary background via change detection resulted in a sparse scene of moving targets, thereby inviting application of compressed sensing techniques for fast data acquisition. Using stepped-frequency based radar imaging, it was demonstrated that a sizable reduction in the number of frequency samples was provided by compressive sensing without degradation in system performance.

Relevant Theory: Radar Signal Processing and Statistical Signal Processing

See this Volume, [Chapter 10](#) Introduction: Radar Signal Processing

See Vol. 3, [Chapter 6](#) Quickest Change Detection

References

- [1] E. Baranowski, Through-wall imaging: historical perspective and future directions, *J. Franklin Inst.* 345 (6) (2008) 556–569.
- [2] J.C. Curlander, R.N. McDonough, *Synthetic Aperture Radar-Systems and Signal Processing*, Wiley, New York, 1991.
- [3] J.W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1996.
- [4] R.K. Tyson, *Principles of Adaptive Optics*, second ed., Academic Press, Boston, MA, 1998, p. 1998.
- [5] R.J. Vaccaro, The past, present, and the future of underwater acoustic signal processing, *IEEE Signal Proc. Mag.* 15 (4) (1998) 21–51.
- [6] E.A. Robinson, S. Treitel, *Geophysical Signal Analysis*, Englewood Cliffs, NJ, Prentice-Hall, 1980.
- [7] A.C. Kak, M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, 1988.
- [8] F. Ahmad, M.G. Amin, Noncoherent approach to through-the-wall radar localization, *IEEE Trans. Aerosp. Electron. Syst.* 42 (4) (2006) 1405–1419.
- [9] F. Ahmad, M.G. Amin, A noncoherent radar system approach for through-the-wall imaging, in: Proceedings of the SPIE—Sensors, and Command, Control, Communications, and Intelligence Technologies IV Conference, Orlando, FL, vol. 5778, 2005, pp. 196–207.
- [10] F. Ahmad, M.G. Amin, Performance of autofocusing schemes for single target and populated scenes behind unknown walls, in: Proceedings of the SPIE—Radar Sensor Technology XI, Orlando, FL, April 2007, vol. 6547, 2007.

- [11] F. Ahmad, G. Frazer, S.A. Kassam, M.G. Amin, Design and implementation of near-field, wideband synthetic aperture beamformers, *IEEE Trans. Aerosp. Electron. Syst.* 40 (1) (2004) 206–220.
- [12] F. Ahmad, M.G. Amin, P.D. Zemany, Performance analysis of dual-frequency CW radars for motion detection and ranging in urban sensing applications, in: Proceedings of SPIE—Radar Sensor Technology XI, Orlando, FL, April 2007, vol. 6547, 2007.
- [13] M.G. Amin (Ed.), Special Issue on Advances in Indoor Radar Imaging, *J. Franklin Inst.* 345 (6) (2008) 556–722.
- [14] M. Amin, K. Sarabandi (Eds.), Special Issue on Remote Sensing of Building Interior, *IEEE Trans. Geosci. Remote Sens.* 47(5) (2009) 1270–1420.
- [15] S.E. Borek, An overview of through the wall surveillance for homeland security, in: Proceedings of the 34th, Applied Imagery and Pattern Recognition, Workshop 6, October 2005, pp. 19–21.
- [16] D.G. Falconer, R.W. Ficklin, K.G. Konolige, Robot-mounted through-wall radar for detecting, locating, and identifying building occupants, in: Proceedings of the IEEE International Conference on Robotics and Automation, 2000, pp. 1868–1875.
- [17] D.D. Ferris Jr., N.C. Currie, A survey of current technologies for through-the-wall surveillance (TWS), in: Proceedings of the SPIE—Sensors, C31 Information, and Training Technologies for Law Enforcement, Boston, MA, November 1998, vol. 3577, pp. 62–72.
- [18] R.J. Fontana, Recent system applications of short-pulse ultra-wideband (UWB) technology, in: *IEEE Trans. Microwave Theory Tech.* 52 (9) (2004).
- [19] L.M. Frazier, Surveillance through walls and other opaque materials, in: Proceedings of the IEEE National Radar Conference 1996, 1996, pp. 27–31.
- [20] E.F. Greneker, RADAR flashlight for through-the-wall detection of humans, in: Proceedings of the SPIE—Targets Backgrounds: Characterization and Representation IV, vol. 3375, 1998, 280–285.
- [21] A.R. Hunt, Image formation through walls using a distributed radar sensor network, in: Proceedings of the SPIE—Laser Physics, Photonics, Spectroscopy and Molecular Modeling V, vol. 5778, 2005, pp. 169–174.
- [22] S. Nag, M.A. Barnes, T. Payment, G. Holladay, An ultra-wideband through-the-wall radar for detecting the motion of people in real time, in: Proceedings of the SPIE—Radar Sensor Technology and Data Visualization, Orlando, FL, April 2002, vol. 4744, 2002, pp. 48–57.
- [23] L.P. Song, C. Yu, Q.H. Liu, Through-wall imaging (TWI) by radar: 2-D tomographic results and analyses, *IEEE Trans. Geosci. Remote Sens.* 43 (12) (2005) 2793–2798.
- [24] F. Soldovieri, R. Solimene, Through-wall imaging via a linear inverse scattering algorithm, *IEEE Geosci. Remote Sens. Lett.* 4 (4) (2007) 513–517.
- [25] F. Soldovieri, F. Ahmad, R. Solimene, Validation of microwave tomographic inverse scattering approach via through-the-wall experiments in semi-controlled conditions, *IEEE Geosci. Remote Sens. Lett.* 8 (1) (2011) 123–127.
- [26] L.M. Frazier, Motion detection radar (MDR) for law enforcement, *IEEE Potentials* 16 (5) (2004) 3079–3082.
- [27] S.S. Ram, H. Ling, Through-wall tracking of human movers using joint Doppler and array processing, *IEEE Geosci. Remote Sens. Lett.* 5 (3) (2008) 537–541.
- [28] P. Setlur, F. Ahmad, M.G. Amin, P.D. Zemany, Experiments on through-the-wall motion detection and ranging, in: Proceedings of the SPIE—Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VI, Orlando, FL, April 2007, vol. 6538, 2007.
- [29] P. Setlur, M.G. Amin, T. Thayaparan, Micro-doppler signal estimation for vibrating and rotating targets, in: Proceedings of the Eighth International Symposium on Signal Processing and its Applications, Sydney, Australia, August 2005, 2005.
- [30] T. Thayaparan, L. Stankovic, I. Djurovic, Micro-doppler human signature detection and its application to gait recognition and indoor imaging, *J. Franklin Inst.* 345 (6) (2008) 700–722.

- [31] P. Setlur, F. Ahmad, M.G. Amin, Analysis of Micro-Doppler signals using linear FM basis decomposition, in: Proceedings of the SPIE Symposium on Defense and Security, Radar Sensor Technology X Conference, Vol. 6210, Orlando, FL, April 2006.
- [32] C.P. Lai, R.M. Narayanan, Ultrawideband random noise radar design for through-wall surveillance, *IEEE Trans. Aerosp. Electron. Syst.* 46 (4) (2010) 1716–1730.
- [33] K. Lukin, V. Konovalov, Through wall detection and recognition of human beings using noise radar sensors, in: Proceedings of the NATO RTO SET Symposium on Target Identification and Recognition Using RF Systems, Oslo, Norway, October 2004, P15.1–P15.12.
- [34] T.E. McEwan, Ultra-Wideband Radar Motion Sensor, US Patent 5,361,070, 1994.
- [35] V. Venkatasubramanian, H. Leung, Chaos UWB radar for through-the-wall imaging, *IEEE Trans. Image Process.* 18 (6) (2009) 1255–1265.
- [36] Y. Yang, A.E. Fathy, Development and implementation of a real-time see-through-wall radar system based on FPGA, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1270–1280.
- [37] F. Ahmad, M.G. Amin, P.D. Zemany, Dual-frequency radars for target localization in urban sensing, *IEEE Trans. Aerosp. Electron. Syst.* 45 (4) (2009).
- [38] F. Ahmad, M.G. Amin, P. Setlur, Through-the-wall target localization using dual-frequency CW radars, in: Proceedings of the SPIE—Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies Conference, Orlando, FL, April 2006, vol. 6201, 2006.
- [39] Y. Zhang, M.G. Amin, F. Ahmad, Time-frequency analysis for the localization of multiple moving targets using dual-frequency radars, *IEEE Signal Process. Lett.* 15 (2008) 777–780.
- [40] F. Ahmad, Y. Zhang, M. Amin, Three-dimensional wideband beamforming for imaging through a single wall, *IEEE Geosci. Remote Sens. Lett.* 5 (2) (2008) 176–179.
- [41] F. Ahmad, M.G. Amin, S.A. Kassam, Synthetic aperture beamformer for imaging through a dielectric wall, *IEEE Trans. Aerosp. Electron. Syst.* 41 (1) (2005) 271–283.
- [42] M. Dehmollaian, M. Thiel, K. Sarabandi, Through-the-wall imaging using differential SAR, *IEEE Trans. Geosci. Remote Sens.* 47(5) (2009) 1289–1296.
- [43] F. Ahmad, M.G. Amin, G. Mandapati, Autofocusing of through-the-wall radar imagery under unknown wall characteristics, *IEEE Trans. Image Process.* 16 (7) (2007) 1785–1795.
- [44] M. Dehmollaian, K. Sarabandi, Refocusing through building walls using synthetic aperture radar, *IEEE Trans. Geosci. Remote Sens.* 46 (6) (2008) 1589–1599.
- [45] Y. Yoon, M.G. Amin, F. Ahmad, MVDR beamforming for through-the-wall radar imaging, *IEEE Trans. Aerosp. Electron. Syst.* 47 (1) (2011) 347–366.
- [46] G. Barrie, J. Tunaley, An Analysis of Through- and in-the-Wall UWB Impulse Radar: System Design Considerations, Technical Memorandum TM 2003-134, Defence Research and Development Canada (DRDC), 2003.
- [47] V.R. Komanduri, A. Hoofar, N. Engheta, Low-profile array design considerations for through-the-wall microwave imaging applications, in: Proceedings of the IEEE International AP-S Symposium, Washington DC, July 2005.
- [48] P. Sévigny et al., Concept of operation and preliminary experimental results of the DRDC through-wall SAR system, in: Proceedings of the SPIE—Radar Sensor Technology XIV, Orlando, FL, April 2010, vol. 7669, 2010, pp. 766907-1–766907-11.
- [49] E.F. Greneker, E.O. Rausch, Wall characterization for through-the-wall radar applications, in: Proceedings of the SPIE—Radar Sensor Technology XII Conference, Orlando, FL, April 2008, vol. 6947.
- [50] A. Muqaibel, Ahmad Safaai-Jazi, Characterization of wall dispersive and attenuative effects on UWB radar signals, *J. Franklin Inst.* 345 (2008) 640–658.

- [51] C. Thajudeen, A. Hoofar, F. Ahmad, T. Dogaru, Measured complex permittivity of walls with different hydration levels and the effect on power estimation of TWRI target returns, *Prog. Electromagn. Res. B* 30 (2011) 177–199.
- [52] R.J. Burkholder, Electromagnetic models for exploiting multi-path propagation in through-wall radar imaging, in: Proceedings of the International Conference on Electromagnetics in Advanced Applications, IEEE, Torino, Italy, 2008, pp. 572–575.
- [53] P.C. Chang, R.J. Burkholder, J.L. Volakis, R.J. Marhefka, Y. Bayram, High-frequency EM characterization of through-wall building imaging, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1375–1387.
- [54] T. Dogaru, C. Le, SAR images of rooms and buildings based on FDTD computer models, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1388–1401.
- [55] M. Thiel, K. Sarabandi, Ultrawideband multi-static scattering analysis of human movement within buildings for the purpose of stand-off detection and localization, *IEEE Trans. Antenn. Propag.* 59 (4) (2011) 1261–1268.
- [56] M. Dehmollaian, K. Sarabandi, An approximate solution of scattering from reinforced concrete walls using an analytic 1-D periodic Green's function, *IEEE Trans. Antenn. Prop.* 56 (8) (2008) 2681–2690.
- [57] R. Solimene, F. Ahmad, F. Soldovieri, A novel CS-SVD strategy to perform data reduction in linear inverse scattering problems, *IEEE Geosci. Remote Sens. Lett.* 9 (5) (2012) 881–885.
- [58] T. Dogaru, C. Le, Through-the-Wall Small Weapon Detection Based on Polarimetric Radar Techniques, US ARL Technical Report ARL-TR-5041, Adelphi, MD, 2009.
- [59] K.M. Yemelyanov, N. Engheta, A. Hoofar, J.A. McVay, Adaptive polarization contrast techniques for through-wall microwave imaging applications, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1362–1374.
- [60] W. Zhang, A. Hoofar, C. Thajudeen, F. Ahmad, Full polarimetric beamforming algorithm for through-the-wall radar imaging, *Radio Sci.* 46 (RS0E16) (2010), <http://dx.doi.org/10.1029/2010RS004631>.
- [61] F.H.C. Tivive, A. Bouzerdoum, M.G. Amin, An SVD-based approach for mitigating wall reflections in through-the-wall radar imaging, in: Proceedings of the IEEE Radar Conference, Kansas City, MO, May 23–27, 2011, 2011, pp. 519–524.
- [62] R. Chandra, A.N. Gaikwad, D. Singh, M.J. Nigam, An approach to remove the clutter and detect the target for ultra-wideband through wall imaging, *J. Geophys. Eng.* 5 (4) (2008) 412–419.
- [63] P.K. Verma, A.N. Gaikwad, D. Singh, M.J. Nigam, Analysis of clutter reduction techniques for through wall imaging in UWB range, *Prog. Electromagn. Res. B* 17 (2009) 29–48.
- [64] Y.-S. Yoon, M.G. Amin, Spatial filtering for wall-clutter mitigation in through-the-wall radar imaging, *IEEE Trans. Geosci. Remote Sens.* 47 (9) (2009) 3192–3208.
- [65] H. Brunzell, Detection of shallowly buried objects using impulse radar, *IEEE Trans. Geosci. Remote Sens.* 37 (2) (1999) 875–886.
- [66] F. Gustafsson, Determining the initial states in forward-backward filtering, *IEEE Trans. Signal Process.* 44 (4) (1996) 988–992.
- [67] M. Çetin, W. Karl, D. Castañon, Evaluation of a regularized SAR imaging technique based on recognition-oriented features, in: Proceedings of the SPIE—Algorithms for SAR Imagery VII, Orlando, FL, April 2000, vol. 4053, pp. 40–51.
- [68] M. Dehmollaian, K. Sarabandi, Analytical, numerical, and experimental methods for through-the-wall radar imaging, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, NV, April 2008, 2008, pp. 5181–5184.
- [69] J.A. Kong, *Electromagnetic Wave Theory*, EMW Publishing, Cambridge, MA, 2000.
- [70] C. Thajudeen, A. Hoofar, A Comparative study of wall parameter estimation and mitigation techniques, in: Proceedings of the USNC/URSI Meeting, Toronto, Canada, July 2010, 2010.

- [71] J. Zhang, Y. Huang, Extraction of dielectric properties of building materials from free-space time-domain measurement, in: Proceedings of High Frequency Postgraduate Student Colloquium, September 1999, 1999, pp. 127–132.
- [72] F. Ahmad, M.G. Amin, Multi-location wideband synthetic aperture imaging for urban sensing applications, *J. Franklin Inst.* 345 (6) (2008) 618–639.
- [73] T. Dogaru, C. Le, Validation of Xpatch Computer Models for Human Body Radar Signature, US ARL Technical Report ARL-TR-4403, Adelphi, MD, 2008.
- [74] S. Kidera, T. Sakamoto, T. Sato, Experimental study of shadow region imaging algorithm with multiple scattered waves for UWB radars, *Prog. Electromagn. Res.* 5 (4) (2009) 393–296.
- [75] G. Alli, D. DiFilippo, Beamforming for through-the-wall radar imaging, in: M.G. Amin (Ed.), *Through-the-Wall Radar Imaging*, CRC Press, Boca Raton, FL, 2010, pp. 81–120.
- [76] M.G. Amin, F. Ahmad, Wideband synthetic aperture beamforming for through-the-wall imaging, *IEEE Signal Process. Mag.* 25 (4) (2008) 110–113.
- [77] P. Setlur, M.G. Amin, F. Ahmad, Multipath model and exploitation in through-the-wall and urban radar sensing, *IEEE Trans. Geosci. Remote Sens.* 49 (10) (2011) 4021–4034.
- [78] Y.-S. Yoon, M.G. Amin, High-resolution through-the-wall radar imaging using beamspace MUSIC, *IEEE Trans. Antenn. Propag.* 56 (6) (2008) 1763–1774.
- [79] G. Wang, M.G. Amin, Y. Zhang, A new approach for target locations in the presence of wall ambiguity, *IEEE Trans. Aerosp. and Electron. Syst.* 42 (1) (2006) 301–315.
- [80] M.G. Amin (Ed.), *Through-the-Wall Radar Imaging*, CRC Press, Boca Raton, FL, 2010.
- [81] C.P. Lai, R.M. Narayanan, Through-wall imaging and characterization of human activity using ultrawideband (UWB) random noise radar, in: Proceedings of the SPIE Sensors and C3I Technologies for Homeland Security and Homeland Defense, Orlando, FL, May 2005, vol. 5778, 2005, pp. 186–195.
- [82] A. Martone, K. Ranney, R. Innocenti, Through-the-wall detection of slow-moving personnel, in: Proceedings of the SPIERadar Sensor Technology XIII, Orlando, FL, April 13–17, 2009, vol. 7308, 2009, pp. 73080Q1–73080Q12.
- [83] M. Skolnik, *Introduction to Radar Systems*, third ed., McGraw Hill, New York, 2001, p. 2001.
- [84] R.M. Narayanan, M.C. Shastry, P.H. Chen, M. Levi, Through-the-wall detection of stationary human targets using Doppler radar, *Prog. Electromagn. Res. B* 20 (2010) 147–166.
- [85] I. Orovic, S. Stankovic, M. Amin, A new approach for classification of human gait based on time-frequency feature representations, *Signal Process.* 91 (6) (2011) 1448–1456.
- [86] S.S. Ram, Y. Li, A. Lin, H. Ling, Doppler-based detection and tracking of humans in indoor environments, *J. Franklin Inst.* 345 (6) (2008) 679–699.
- [87] D. Tahmoush, J. Silvious, Stride rate in radar micro-doppler images, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, October 11–14, 2009, 2009, pp. 4218–4223.
- [88] K. Ranney et al. Recent MTI experiments using ARL's synchronous impulse reconstruction (SIRE) radar, in: Proceedings of the SPIE—Radar Sensor Technology XII, Orlando, FL, April 2008, vol. 6947, 2008, pp. 694708-1–694708-9.
- [89] J. Moulton, S.A. Kassam, F. Ahmad, M.G. Amin, K. Yemelyanov, Target and change detection in synthetic aperture radar sensing of urban structures, in: Proceedings of the IEEE Radar Conference, Rome, Italy, May 26–30, 2008.
- [90] F. Soldovieri, R. Solimene, R. Pierri, A simple strategy to detect changes in through the wall imaging, *Prog. Electromagn. Res. M7* (2009) 1–13.
- [91] A.R. Hunt, Use of a frequency-hopping radar for imaging and motion detection through walls, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1402–1408.

- [92] Y.D. Zhang, A. Hunt, Image and localization of behind-the-wall targets using collocated and distributed apertures, in: M.G. Amin (Ed.), *Through-the-Wall Radar Imaging*, CRC Press, Boca Raton, FL, 2010, pp. 121–156 (Chapter 4).
- [93] N. Maaref, P. Millot, C. Pichot, O. Picon, A study of UWB FM-CW Radar for the detection of human beings in motion inside a building, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1297–1300.
- [94] F. Ahmad (Ed.), *Compressive Sensing*, in: Proceedings of SPIE, vol. 8365, SPIE, Bellingham, WA, 2012.
- [95] E. Candes, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.* 59 (2006) 1207–1223.
- [96] A. Gurbuz, J. McClellan, W. Scott, A compressive sensing data acquisition and imaging method for stepped frequency GPR, *IEEE Trans. Signal Process.* 57 (7) (2009) 2640–2650.
- [97] Y.-S. Yoon, M.G. Amin, Compressed sensing technique for high-resolution radar imaging, in: Proceedings of the SPIE, Orlando, FL, April 2008, vol. 6968, 2008, pp. 69681A-1–69681A-10.
- [98] D. Donoho, M. Elad, V. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inform. Theory* 52 (1) (2006) 6–18.
- [99] F. Ahmad, M. Amin, Compressive urban sensing, SPIE Newsroom, Published Online: 27 September 2012, <http://dx.doi.org/10.1117/2.1201209.004413>, <<http://spie.org/x90448.xml>>, 2012.
- [100] G. Wang, M.G. Amin, Imaging through unknown walls using different standoff distances, *IEEE Trans. Signal Process.* 54 (10) (2006) 4015–4025.
- [101] A. Martone, K. Ranney, R. Innocenti, Automatic through the wall detection of moving targets using low-frequency ultra-wideband radar, in: Proceedings of the IEEE International Radar Conference, Washington D.C., May 2010, 2010, pp. 39–43.
- [102] L.C. Potter, E. Ertin, J.T. Parker, M. Cetin, Sparsity and compressed sensing in radar imaging, *Proc. IEEE* 98 (6) (2010) 1006–1020.
- [103] Q. Huang, L. Qu, B. Wu, G. Fang, UWB through-wall imaging based on compressive sensing, *IEEE Trans. Geosci. Remote Sens.* 48 (3) (2010) 1408–1415.
- [104] P. Boufounos, M.F. Duarte, R.G. Baraniuk, Sparse signal reconstruction from noisy compressive measurements using cross-validation, in: Proceedings of the 2007 IEEE 14th Workshop on Statistical Signal Processing, 2007.
- [105] M. Farwell et al., Sense through the wall system development and design considerations, *J. Franklin Inst.* 345 (6) (2008) 570–591.
- [106] M. Pieraccini, G. Luzi, D. Dei, L. Pieri, C. Atzeni, Detection of breathing and heartbeat through snow using a microwave transceiver, *IEEE Geosci. Remote Sens. Lett.* 5 (1) (2008) 57–59.
- [107] E. Ertin, R.L. Moses, Through-the-wall SAR attributed scattering center feature estimation, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1338–1348.

18

Multi-Channel SAR for Ground Moving Target Indication

Stefan V. Baumgartner and Gerhard Krieger

German Aerospace Center (DLR), Microwaves and Radar Institute, Oberpfaffenhofen, Germany

Nomenclature

*	complex conjugated
<i>a</i>	acceleration magnitude of the target (m/s^2)
A_i	complex coefficient containing free-space attenuation, backscattering coefficient, and two-way antenna pattern weighting
a_x	along-track acceleration (m/s^2)
a_y	across-track acceleration (m/s^2)
B_C	bandwidth of the clutter Doppler spectrum (Hz)
B_r	bandwidth of the transmitted pulse (Hz)
c	clutter vector in time domain
C	clutter in Doppler domain
c_0	speed of light in vacuum (m/s)
d	multi-channel steering vector
d_a	physical along-track antenna separation (m)
d_{pc}	effective along-track phase center separation (m)
f_a	Doppler or azimuth frequency (Hz)
$f_C(x)$	probability density function of cluster-plus-noise
f_{DC}	Doppler shift of the moving target signal (Hz)
$f_{DC,st}$	Doppler shift of the stationary target signal (Hz)
$f_T(x)$	probability density function of target-plus-noise
H	reference function or matched filter in frequency domain
H	vector complex conjugate transposition; Hermitian transpose
h	reference function or matched filter in time domain
h_a	azimuth reference function of the matched filter in time domain
H_a	azimuth reference function of the matched filter in Doppler domain
$h_{a,st}$	azimuth reference function for a stationary signal in time domain
h_r	range reference function in time domain
I	compressed signal or impulse response function (IRF)
j	imaginary unit

K	total number of training cells
k	training cell number/index
k_a	Doppler slope of the moving target signal (Hz/s)
$k_{a,st}$	Doppler slope of the stationary world matched filter or stationary signal (Hz/s)
k_r	slope of the range chirp in range-frequency domain (Hz/s)
k_ϕ	slope of the unregistered ATI phase in time domain (rad/s)
L	antenna length or height (m)
L_a	antenna length in azimuth direction (m)
L_{SA}	length of the synthetic aperture (m)
m	receiving channel number/ index
M	total number of receiving channels
m_c	constant for scaling purposes
\mathbf{n}	noise vector
n	temporal sample number/index
N	integer number or total number of azimuth samples
N	Noise in Doppler domain
P_d	probability of detection
P_{fa}	probability of false alarm; false alarm rate
PRF	pulse repetition frequency (Hz)
$PRF_{min,3\text{ dB}}$	minimum required pulse repetition frequency (Hz)
q	quadratic Doppler coefficient (Hz/s^2)
\mathbf{Q}	multi-channel signal interference (= clutter + noise) in Doppler domain
r	range or slant range (m)
r_0	minimum range (m)
r_{10}	range at time $t = 0$ (m)
$\text{rect}[\cdot]$	rectangular function
\mathbf{r}_p	position vector of the antenna phase center in Cartesian $\{x, y, z\}$ coordinate system (m)
\mathbf{R}_W	clutter covariance matrix
r_{st}	range to a stationary target (m)
\mathbf{r}_t	position vector of the target in Cartesian $\{x,y,z\}$ coordinate system (m)
s	azimuth signal in baseband
\mathbf{s}	multi-channel signal vector of dimension $MN \times 1$ (M = number of antennas, N = number of temporal samples)
\mathbf{S}	multi-channel signal matrix of dimension $M \times N$ (M = number of antennas, N = number of temporal samples)
s_{ATI}	ATI signal
s_{DPCA}	DPCA signal
s_i	azimuth signal in baseband received by channel i
$s_{i,\text{reg}}$	co-registered or aligned baseband signal received by channel i
$S_{k\varphi}$	phase ramp in Doppler; needed for co-registration
s_{RX}	received signal
$s_{RX,b}$	received signal in baseband
s_{TX}	transmitted pulse or waveform

$s_{\text{TX,b}}$	transmitted pulse or waveform in baseband
t	azimuth or slow time (s)
T	sample interval, $T = 1/\text{PRF}$ (s)
T	vector or matrix transposition
t_{Fr}	fractional time of fractional Fourier transform
T_{SA}	synthetic aperture time (s)
U	input signal in frequency domain
u	input signal
u_t	directional cosine; measured from the azimuth-axis (x -axis)
v_0	velocity magnitude of the target at $t = 0$ (m/s)
v_{x0}	along-track velocity of the target at $t = 0$ (m/s)
v_{y0}	across-track velocity of the target at $t = 0$ (m/s)
v_p	platform velocity (m/s)
v_{r0}	line-of-sight velocity (m/s)
$v_{r0,\text{blind}}$	line-of-sight blind velocity (m/s)
$v_{r0,\text{max}}$	maximum line-of-sight velocity (m/s)
w	weight vector
w_r	envelope of the pulse
x	x -axis, along-track or azimuth direction (m)
x_0	along-track position of the target at time $t = 0$ (m)
x_t	azimuth or along-track position of the target (m)
y	y -axis or across-track direction (m)
y_{out}	signal after filtering
y_0	across-track position of the target at time $t = 0$ (m)
y_p	across-track position of the platform (m)
y_t	across-track position of the target (m)
z	z -axis or altitude (m)
z	space-time snapshot of the noise and clutter contaminated multi-channel signal
Z	measured multi-channel signal in Doppler domain
z_0	altitude of the target at time $t = 0$ (m)
z_p	altitude of the platform (m)
z_t	altitude of the target (m)
α	moving angle of the target or road angle (rad)
α_{Fr}	rotation angle of the fractional Fourier transform
$\alpha_{\text{Fr,opt}}$	optimum rotation angle of the fractional Fourier transform giving the highest SCNR
Δf_a	Doppler bandwidth of the moving target signal (Hz)
Δr	residual range cell migration (m)
Δr_{blur}	spread of the blur in range direction (m)
Δt	time difference corresponding to along-track baseline (s)
$\Delta t_{\text{img},0}$	azimuth time corresponding to azimuth imaging position of signal (s)
$\Delta t_{\text{img},1}$	azimuth time corresponding to azimuth imaging position of first ambiguity (s)
Δt_{shift}	time difference relevant for co-registration (s)
Δx	along-track difference (m)

Δx_{blur}	spread of the blur in azimuth direction (m)
$\Delta x_{\text{img},0}$	along-track or azimuth displacement (m)
$\Delta\phi$	ATI phase (rad)
$\delta_{3\text{dB}}$	spatial resolution given by the one-way 3 dB antenna beamwidth (m)
δ_a	azimuth resolution of the SAR image (m)
δ_r	range resolution of the SAR image (m)
$\delta_r(t)$	range difference (m)
δ_φ	phase error of aliased clutter signals (rad)
γ_{SNR}	complex correlation coefficient
$\theta_{3 \text{ dB}}$	one-way 3 dB antenna beamwidth (rad)
$\theta_{3 \text{ dB,a}}$	one-way 3 dB antenna beamwidth in azimuth (rad)
θ_D	depression angle of the antenna (rad)
θ_i	incidence angle of the radar pulse (rad)
φ_a	phase of azimuth signal (rad)
φ_i	phase of azimuth signal of channel i (rad)
$\varphi_{i,\text{reg}}$	phase of co-registered azimuth signal of channel i (rad)
λ	radar wavelength (m)
$\sigma_{\Delta r_0}$	standard deviation of the line-of-sight velocity (m/s)
$\sigma_{\Delta x_{\text{img}}}$	standard deviation of the azimuth position (m)
$\sigma_{\Delta\phi}$	standard deviation of the ATI phase (rad)
τ	“range time” or “fast time” (s)
τ_p	pulse duration (s)
ψ	squint angle (rad)
ψ_{DOA}	direction-of-arrival angle (rad)

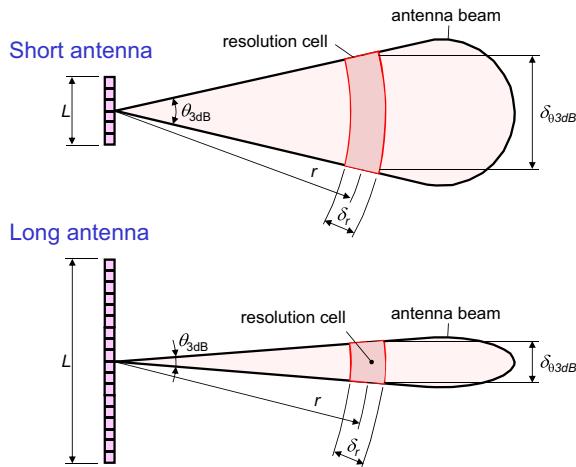
2.18.1 Introduction

Moving target indication (MTI) originated in the military field with the aim to detect approaching sea and air targets. Originally, stationary radar stations with a rotating antenna installed on the earth's surface were used for this task. The signal processing was, at least from today's point of view, quite simple: the range measured by the traveling time of the transmitted and backscattered pulse, and the information about the angular position of the radar antenna were used for determining the position of the detected target. The achievable angular resolution was limited by the antenna beamwidth. For plane antennas the one-way 3 dB beamwidth is given as [1]

$$\theta_{3\text{dB}} \cong 0.886 \cdot \frac{\lambda}{L}, \quad (18.1)$$

where λ is the radar wavelength and L the antenna length or height, depending if the beamwidth in azimuth or elevation shall be computed. It is obvious that a longer antenna has a smaller beamwidth which results in an improved spatial resolution

$$\delta_{\theta_{3\text{dB}}} \cong \theta_{3\text{dB}} \cdot r = 0.886 \cdot \frac{\lambda}{L} \cdot r, \quad (18.2)$$

**FIGURE 18.1**

Influence of the antenna length L on the beamwidth $\theta_{3\text{dB}}$ and on the achievable spatial resolution $\delta_{\theta3\text{dB}}$.

where r is the distance or range between the antenna phase center and the target. The relationships given in (18.1) and (18.2) are visualized in Figure 18.1. From the spatial resolution point of view large antennas are preferred for classical MTI. The achievable range resolution δ_r is determined by the transmitted pulse waveform. It is independent of the beamwidth and discussed later in Section 2.18.2.3.

With more modern pulse Doppler MTI radars the Doppler shifts of the backscattered target signals were exploited for deciding if a target was moving or not (the signal backscattered from a moving target in contrast to a stationary target is shifted in Doppler due to its range change during the illumination time). For suppressing unwanted radar echos (= clutter) backscattered from stationary or slowly moving unwanted targets (buildings, hills, trees, sea, rain, etc.) a notch around zero Doppler frequency or more sophisticated Doppler filter banks were used [2].

In the context with flying radar platforms additionally to MTI the notation GMTI can be found in the literature. GMTI stands for “ground moving target indication” and is strictly speaking a special case of MTI. GMTI focuses on targets moving on the earth’s surface (land and ocean).

The implementation of GMTI capabilities to radars flying at high altitude is more sophisticated since the platform carrying the radar system moves by itself. This motion causes a spread of the clutter Doppler spectrum so that especially signals backscattered from slowly moving targets are masked and, hence, cannot be detected. For side-looking radars where the antenna beam points perpendicular to the flight direction the bandwidth of the unwanted clutter Doppler spectrum (in the following denoted as clutter bandwidth B_C) is proportional to the platform velocity as well as to the azimuth antenna beamwidth $\theta_{3\text{dB},a}$ [1]:

$$B_C \cong \theta_{3\text{dB},a} \cdot \frac{2v_p}{\lambda} = 0.886 \cdot \frac{2v_p}{L_a}, \quad (18.3)$$

where L_a is the length of a flat antenna in azimuth or flight direction and v_p is the velocity of the radar platform. Since the velocity v_p of a given platform is more or less fixed, at the beginning for GMTI

large stabilised antennas (i.e., large antenna lengths L_a) with narrow beams and low sidelobe levels were used to narrow down the clutter bandwidth. Thus, in the classical single-channel GMTI case where only one antenna is available, the GMTI detection performance is mainly limited by the antenna length. Single-channel GMTI is restricted either to fast moving targets whose Doppler shifted signals lie outside the clutter bandwidth, or to targets with high reflectivity or radar cross section (RCS), resulting in high signal-to-clutter-plus-noise ratios (SCNRs) so that even a velocity independent detection is possible [4].

In contrast to pure GMTI systems air- and spaceborne synthetic aperture radar (SAR) systems were primarily developed for imaging the stationary world but not for detecting moving targets [1]. To achieve high resolution with a SAR system a long illumination time and, hence, a small antenna with a wide azimuth beam is required. The best achievable azimuth resolution of a SAR system operated in the so called stripmap mode is given as

$$\delta_a \cong \frac{L_a}{2}. \quad (18.4)$$

The resolution is independent of the range (that's one of the reasons why with spaceborne SAR systems high resolution can be achieved). The smaller the azimuth antenna length L_a , the better is the resolution. However, just this requirement is in contradiction with the need for large antennas and narrow beams for classical single-channel GMTI (remember that the shorter the antenna, the larger the clutter bandwidth given in (18.3) and the worse the detection capability of slowly moving targets embedded in the clutter). The desired signal for SAR imaging, i.e., the radar echos from the stationary non-moving scene which shall be imaged, can be considered as unwanted clutter for GMTI.

Owing to the nature of conventional SAR processing moving targets in general are depicted blurred and displaced from their actual positions [5]. The reason is the additional Doppler shift of moving target signals. Two examples are shown in Figure 18.2 where a slowly moving ship and a moving train are imaged. The so called “ship-of-the-wake” or “train-of-the-track” effects (i.e., the azimuth displacements of the targets) can clearly be recognized. The magnitude of the displacement depends on the target velocity. If for instance a typical imaging geometry of the German TerraSAR-X satellite [6] flying at an altitude of 514 km is considered, a comparatively slowly moving ship with a velocity of 30 km/h is displaced up to 600 m. The displacement of fast road vehicles traveling with 130 km/h

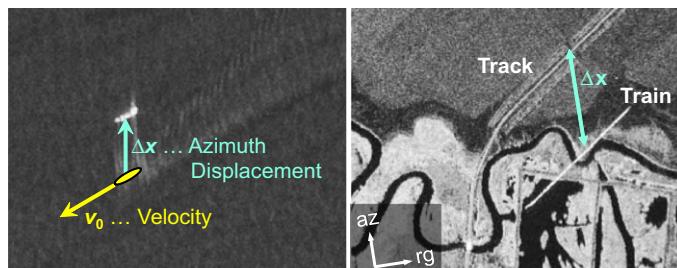


FIGURE 18.2

TerraSAR-X images of a moving ship in the Strait of Gibraltar (left) and a moving train near Wolgograd, Russia (right). The azimuth displacements can clearly be recognized.

may be already in the order of 2500 m. A suitable GMTI algorithm should not only be able to detect the “displaced” moving targets, but also to estimate their true (non-displaced) geographical positions, their velocities and moving directions.

For adding GMTI capabilities to SAR systems without preventing high resolution imaging appropriate techniques for suppressing the clutter are necessary. This can be achieved by implementing more than one receiving antenna or receiving channel. The signals received by multiple antennas, which are arranged in flight direction and which are separated by a certain baseline, can be combined in different ways: once for suppressing the clutter and so enabling the detection of even slowly moving targets, and once for estimating the motion and position parameters of the targets.

The question how many receiving antennas are needed can briefly be answered: for suppressing the clutter at least two receiving antennas separated by a certain baseline in azimuth or flight direction are necessary. A third antenna allows for a more robust estimation of the moving target’s true position and motion parameters. Additional antennas incorporate further degrees of freedom which for instance can be used for suppressing jammers [7]. This may be of importance for military applications. However, more than three antennas not necessarily improve the detection and parameter estimation performance significantly [8].

Today, GMTI is no longer limited to military applications. A SAR-GMTI system flying at high altitude can also be used for civilian wide area traffic monitoring, which has evolved into an important research topic during the last years [9–12]. Real-time traffic monitoring data for instance are used by traffic monitoring centers for ensuring the mobility and safety of the road users. Nowadays these data are mainly collected operationally from stationary sensors mounted along the major roads. Outside of these roads still a severe data lack exists, which even in short-term cannot fully be stuffed by the additional use of floating car data [13] and signaling information generated by the phone network. However, SAR-GMTI systems might be used in near future to fill these information gaps, especially if the information is required on a non-regular basis as in the case of major events and catastrophes.

Modern SAR-GMTI systems are expected to have at least the following capabilities:

- Detection of even slowly moving targets with low reflectivity (low RCS) against a strong clutter interference.
- Estimation of the moving targets’ true geographical positions.
- Estimation of the moving targets’ velocities and driving directions.

Furthermore, especially for military applications, it may also be required to

- Have real-time capability.
- Track the moving targets during the (increased) observation time.
- Refocus the “blurred” images of extended moving targets (an extended target is a target occupying more than one SAR resolution cell) like ships and larger land vehicles to high resolution for recognition purposes.

It has to be pointed out here that the latter three points will not be treated in this tutorial. Information about target tracking can e.g., be found in [14]. Refocusing of extended moving targets can be performed with inverse SAR (ISAR) imaging techniques. Adequate information and references can be found in [15–17] and also in the ISAR chapter of the ELSEVIER e-reference.

Principally two groups of multi-channel GMTI algorithms can be discriminated. The first group is based on the classical dual-channel techniques along-track interferometry (ATI) and displaced phase center antenna (DPCA). State-of-the-art spaceborne SAR systems limited to two physical receiving (RX) channels such as the German TerraSAR-X [6] and the Canadian RADARSAT-2 apply these GMTI techniques successfully. Thus, a special focus on these classical techniques is given in this tutorial. The second group is based on space-time adaptive processing (STAP) techniques for which a separate tutorial/chapter can be found in the ELSEVIER e-reference. For that reason only a short introduction to STAP is given in Sections 2.18.6.2.2 and 2.18.8.

For understanding the following sections of the tutorial the reader shall be familiar with the basic principles of SAR imaging. Good extended tutorials on SAR can for instance be found in [1] and in the SAR chapter of the ELSEVIER e-reference. As mathematical background mainly linear algebra (vectors and matrices), the understanding of the convolution and the Fourier transform and its inverse are required.

The remainder of this tutorial is organized as follows: in Section 2.18.2 the SAR principle is explained before in Section 2.18.3 the moving target single- and multi-channel signal model is derived. The effects caused by moving target signals are discussed in Section 2.18.4. They are fundamental for understanding the parameter estimation principles discussed afterwards. The classical dual-channel techniques are presented in detail in Section 2.18.5. In Section 2.18.6 the general GMTI processing chain is discussed and in Section 2.18.7 the basic Doppler parameter estimation methods are introduced. A short introduction to STAP is given in Section 2.18.8 before the tutorial is concluded with Section 2.18.9.

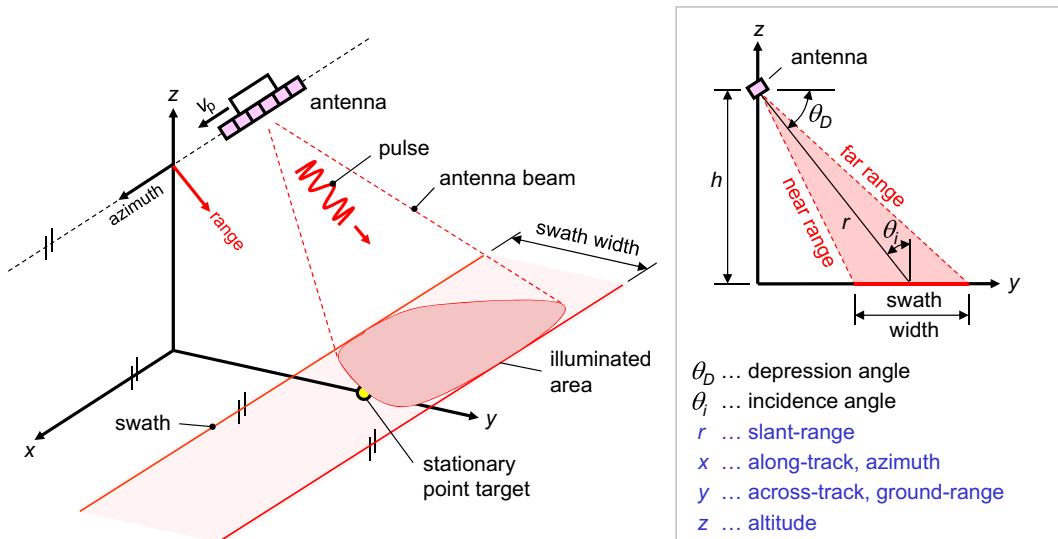
2.18.2 Synthetic aperture radar principle

For the following investigations a flat earth surface and a straight flight path of the SAR platform parallel to the earth surface are assumed. Although SAR and GMTI are not restricted to these assumptions (e.g., for spaceborne SAR-GMTI curved orbits have to be considered [18]), the explanations and equations given in the tutorial can be simplified to a certain degree and presented in a way better understandable by the interested reader who is not a SAR expert.

2.18.2.1 SAR acquisition geometry and operation

A SAR instrument consists of a pulsed transmitter, at least of one antenna which is used both for transmitting and receiving, and of a phase coherent receiver [1]. The typical side-looking imaging geometry of a SAR system is shown in Figure 18.3.

The platform carrying the radar instrument moves at constant altitude with constant velocity v_p parallel to the x -axis. The moving direction of the radar is also denoted as along-track or azimuth direction. The antenna is mounted in a way so that the antenna beam with a certain depression angle θ_D points perpendicular to the azimuth direction towards the ground (the system in Figure 18.3 is left-looking with respect to the flight path). An area on ground with a certain swath width is illuminated by the beam. The radar is periodically emitting radar pulses of duration τ_p with the so called Pulse Repetition Frequency (PRF). The PRF typically is in the order of several 100 Hz (airborne systems) to several 1000 Hz (spaceborne systems). The pulses are backscattered from the illuminated area on

**FIGURE 18.3**

SAR acquisition geometry.

ground, coherently received, down converted, digitized and stored in the mass memory of the SAR instrument. SAR processing is carried out afterwards, either onboard the platform or on ground after downloading the data.

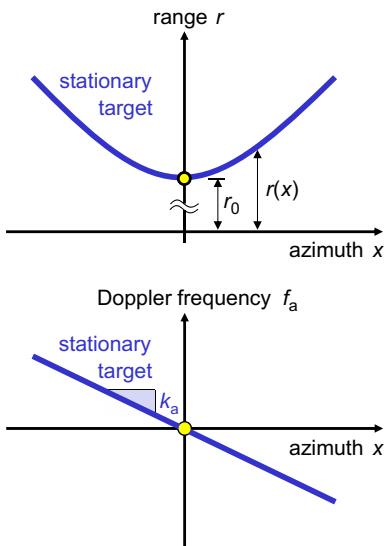
What the SAR measures are the backscattered signal energy and the time interval between the emitted and received pulses. The pulse travel time is proportional to the two-way range, i.e., the range from the antenna phase center to the target and back. A side-looking geometry is necessary so that for each measured slant range r the corresponding ground range or across-track position y can be computed unambiguously (cf. Figure 18.3).

The SAR principle is based on a movement of the sensor with respect to the illuminated targets on ground. Due to the motion of the platform the range r between the platform and a specific stationary point target on ground changes as shown in Figure 18.4. This range change causes a Doppler frequency shift of the received signal which during SAR processing is exploited for synthesizing a large antenna along azimuth direction, resulting in a narrow synthetic azimuth beam width and, hence, in high azimuth resolution.

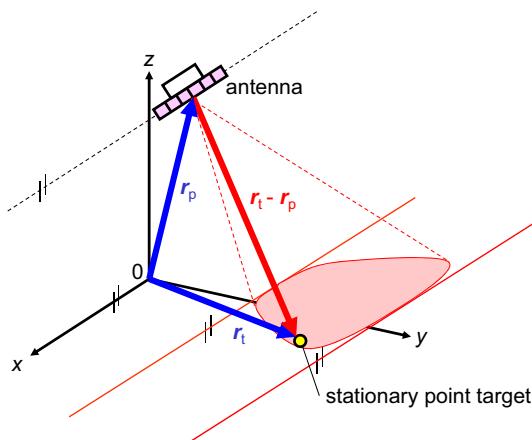
Due to the importance for GMTI processing discussed later, in the following the range and Doppler histories of the signal backscattered from a particular non-moving point target are derived.

The position of the antenna phase center located at the moving SAR platform with respect to the origin of the Cartesian $\{x, y, z\}$ coordinate system can be written as (see also sketch in Figure 18.5)

$$\mathbf{r}_p(t) = \begin{bmatrix} x_p(t) \\ y_p \\ z_p \end{bmatrix} = \begin{bmatrix} v_p t \\ 0 \\ z_p \end{bmatrix}, \quad (18.5)$$

**FIGURE 18.4**

Range (top) and Doppler frequency history (bottom) of a stationary point target.

**FIGURE 18.5**

SAR acquisition geometry with point target and platform position vectors r_t and r_p .

where v_p is the constant platform velocity and t the time. At time $t = 0$ the platform is at altitude z_p above the origin of the Cartesian coordinate system. Let now the position of a certain non-moving "stationary" target be

$$\mathbf{r}_t(t) = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}. \quad (18.6)$$

The indices “0” indicate time independent parameters. The distance between the antenna and the target can simply be computed as

$$r(t) = \|\mathbf{r}_t - \mathbf{r}_p\| = \sqrt{(x_t - x_p)^2 + (y_t - y_p)^2 + (z_t - z_p)^2}, \quad (18.7)$$

where $\|\cdot\|$ denotes the L_2 norm. For describing the SAR principle in common literature it is often assumed, without restriction of generality, that $x_t = x_0 = 0$ so that at $t = 0$ the point target is broadside the SAR platform. The minimum range r_0 at $t = 0$ in this case can be written as

$$r_0 = \sqrt{y_0^2 + (z_t - z_p)^2}, \quad (18.8)$$

so that (18.7) simplifies to

$$r(t) = \sqrt{r_0^2 + x_p^2} = \sqrt{r_0^2 + (v_p t)^2} \cong r_0 + \frac{v_p^2}{2r_0} t^2. \quad (18.9)$$

The quadratic approximation given after the “ \cong ” sign is obtained by a second-order Taylor expansion about $t = 0$. The time t is proportional to the azimuth position of the platform x_p :

$$t = \frac{x_p}{v_p}. \quad (18.10)$$

2.18.2.2 Stationary point target signal model

One single pulse transmitted by the SAR system can be expressed as

$$s_{TX}(\tau) = s_{TX,b}(\tau) \cdot \exp\left[j \frac{2\pi}{\lambda} c_0 \tau\right], \quad (18.11)$$

where $s_{TX,b}(\tau)$ represents the pulse waveform in baseband, τ is the so called “fast time,” j is the imaginary unit, λ is the radar wavelength given by the carrier frequency and c_0 denotes the speed of light in vacuum. Conventionally in SAR a linear frequency modulated (LFM) waveform, a so called “range chirp,” with a certain bandwidth B_r and a certain duration τ_p (in the order of microseconds) is transmitted (although SAR is not limited to such waveforms). The range chirp in baseband is given as

$$s_{TX,b}(\tau) = w_r(\tau) \cdot \exp\left[j\pi \frac{B_r}{\tau_p} \tau^2\right] = w_r(\tau) \cdot \exp[j\pi k_r \tau^2], \quad (18.12)$$

where k_r denotes the chirp slope and w_r the envelope. The signal received from a point-like target is then a delayed and attenuated copy of the transmitted signal which can be written as

$$s_{RX}(\tau) = A \cdot s_{TX}\left(\tau - \frac{2r(\tau)}{c_0}\right). \quad (18.13)$$

The free-space attenuation, the backscattering coefficient, the elevation and the azimuth angles to the target as well as the weighting of the two-way antenna pattern are covered by the coefficient A . After coherent down-conversion to baseband using, e.g., a phase preserving quadrature demodulator, the received signal is given as

$$s_{\text{RX,b}}(\tau) = A \cdot s_{\text{TX}} \left(\tau - \frac{2r(\tau)}{c_0} \right) \cdot \exp \left[-j \underbrace{\frac{2\pi}{\lambda} 2r(\tau)}_{=-\varphi_a(\tau)} \right]. \quad (18.14)$$

This signal can be separated into two parts:

1. The transmitted and delayed waveform s_{TX} whose delay is given by the two-way path $2r(\tau)$ between the antenna and the target.
2. An exponential term with phase φ_a representing the azimuth modulation of the signal, which is independent of the transmitted waveform.

The raw data signal given in (18.14) in its one-dimensional representation is in fact stored in a two-dimensional arrangement in the mass memory of the radar system according to the range and azimuth dimension. To get a better insight in this storage procedure one can make use of the start-stop-approximation. It assumes that the antenna and, hence, the SAR platform is motionless when a pulse is emitted and the scattered signal received. Afterwards the antenna moves to its next sending/receiving position along the flight track. This approximation can be made since the pulse travel time is much smaller than the time needed for the antenna to move to the next position. In range dimension the signal is sampled when the antenna is “motionless.” The range sampling frequency is determined by the analog-digital converter. For a complex signal this sampling frequency has to be at least as large as the chirp bandwidth B_r so that the Nyquist criterion is not violated. In azimuth dimension the sampling frequency and, hence, the imaginary antenna “stops” are determined by the PRF.

The signal can therefore be written in a two-dimensional form as

$$s_{\text{RX,b}}(\tau, t) = A \cdot s_{\text{TX}} \left(\tau - \frac{2r(\tau)}{c_0} \right) \cdot \exp \left[j \frac{4\pi}{\lambda} r(t) \right], \quad (18.15)$$

where τ is the “fast time” representing the range direction and t is the “slow time” representing the azimuth direction.

Due to the importance for GMTI we will put the main focus on the azimuth signal

$$s(t) = A \cdot \exp \left[-j \frac{4\pi}{\lambda} r(t) \right] \cdot \text{rect} \left[\frac{t}{T_{\text{SA}}} \right] = A \cdot \exp[j\varphi_a(t)] \cdot \text{rect} \left[\frac{t}{T_{\text{SA}}} \right], \quad (18.16)$$

where the rectangular function $\text{rect}[\cdot]$, defined e.g., in [3], is introduced for pointing out that the signal duration is limited by the illumination time T_{SA} given by the azimuth beamwidth of the antenna pattern. A small azimuth antenna length results in a wide beam (see also (18.1) and long illumination or synthetic aperture time, respectively. For typical airborne systems this time is in the order of several seconds, for

state-of-the-art spaceborne systems around one second or smaller. The longer this time, the better is the achievable azimuth resolution after SAR processing.

The phase φ_a within the exponential term can according to (18.9) also be approximated using a second-order Taylor expansion, so that

$$\varphi_a(t) = -\frac{4\pi}{\lambda} r(t) \cong -\frac{4\pi}{\lambda} \left(r_0 + \frac{v_p^2}{2r_0} t^2 \right). \quad (18.17)$$

The azimuth phase modulation furthermore can be interpreted as azimuth frequency or Doppler frequency variation f_a if the time derivative is taken in the following way:

$$f_a(t) = \frac{1}{2\pi} \cdot \frac{\partial \varphi_a(t)}{\partial t}. \quad (18.18)$$

If for the phase φ_a the quadratic approximation given in the second part of (18.17) is inserted, the linear approximation of the Doppler frequency history is obtained:

$$f_a(t) \cong -\frac{2v_p^2}{\lambda r_0} t = k_{a,st} t. \quad (18.19)$$

It is obvious that the azimuth signal in the first approximation has the shape of a LFM signal with $k_{a,st}$ denoting the signal's azimuth chirp slope or Doppler slope (see also Figure 18.4 bottom).

2.18.2.3 Pulse compression and image formation

As mentioned in the previous section, the transmitted pulse typically has a time duration in the order of a few microseconds, whereas the illumination time of a particular point target is in the order of seconds. Thus, for achieving a high range and azimuth resolution pulse compression has to be employed. Pulse compression generally can be performed by convolving an uncompressed input signal $u(t)$ with a proper reference function $h(t)$. The pulse compressed signal $I(t)$ in its general form is then given as

$$I(t) = u(t) * h(t) = \int_{-\infty}^{+\infty} u(\tau)h(t - \tau)d\tau, \quad (18.20)$$

where $*$ denotes convolution.

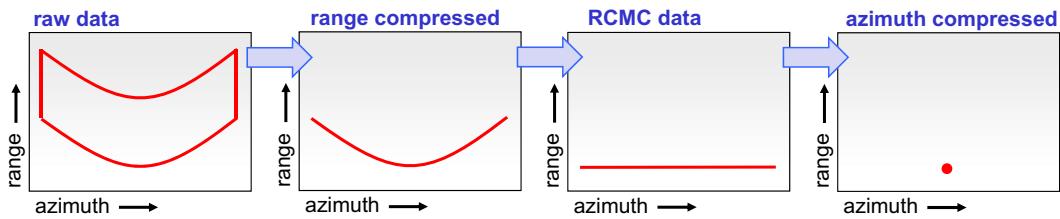
The optimal filter theory says that for signals embedded in white Gaussian noise the best signal-to-noise ratio (SNR) after convolution is achieved if the reference function $h(t)$ is the complex conjugated and time reverted version of the “expected” input signal $u(t)$:

$$h(t) = m_c \cdot u^*(-t), \quad (18.21)$$

where m_c is a constant which may be used for scaling purposes and $*$ denotes the complex conjugation. If the reference function is constructed in this way, it is denoted as “Matched Filter” [19]. The resolution improvement by applying a matched filter is sketched in Figure 18.6. The comparatively long input signal $u(t)$ after matched filtering is compressed to a pulse $I(t)$ of short duration. The time resolution and, hence, the spatial resolution is significantly improved compared to the input signal.

**FIGURE 18.6**

Pulse compression with a matched filter.

**FIGURE 18.7**

Simplified SAR processing steps for a single non-moving point target.

It is known that a cyclic convolution in time domain corresponds to a simple multiplication in frequency domain. For that reason the convolution in (18.20) can equivalently be written as

$$I(t) = u(t) * h(t) = F^{-1}\{F\{u(t)\} \cdot F\{h(t)\}\} = F^{-1}\{U(f) \cdot H(f)\}, \quad (18.22)$$

where F and F^{-1} denote the Fourier and inverse Fourier transforms, respectively, and $U(f)$ and $H(f)$ are the frequency domain representations of $u(t)$ and $h(t)$, respectively.

SAR processing or SAR image formation, within the GMTI community often denoted as “Stationary World Matched Filtering,” can be described briefly by the following three steps:

- 1. Range compression:* A pulse compression along range dimension is performed. For the range chirp in (18.12) the reference function is

$$h_r(t) = \exp[-j\pi k_r t^2]. \quad (18.23)$$

- 2. Range cell migration correction (RCMC):* The curvature of the range history is eliminated (see third image from left in Figure 18.7).
- 3. Azimuth compression:* A pulse compression along azimuth is performed. For the azimuth signal in (18.16) the reference function is

$$h_{a,st}(t) = \exp\left[j\frac{4\pi}{\lambda}r(-t)\right] \cong \exp[-j\pi k_{a,st} t^2], \quad (18.24)$$

where the approximation after the “ \cong ” sign is obtained by inserting the Taylor expansion from (18.17) into (18.16), substituting $k_{a,st} = (-2v_p^2)/(\lambda r_0)$ and dropping the constant phase term $(-4\pi/\lambda)r_0$ which is unimportant for pulse compression.

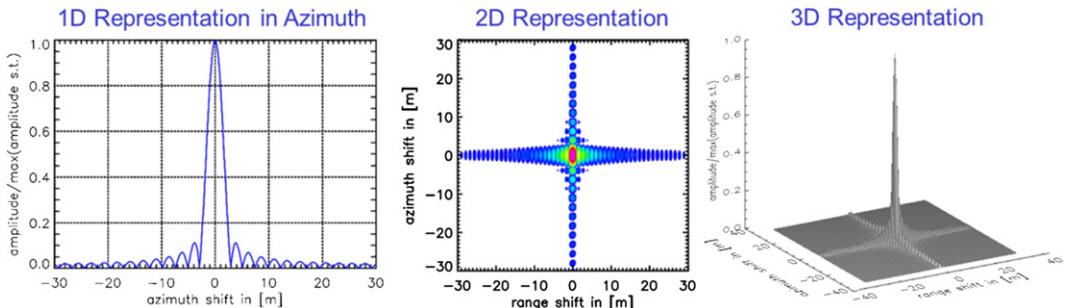


FIGURE 18.8

Impulse response function of a stationary point target (left: cut along azimuth direction; middle: two-dimensional representation; right: three-dimensional representation; system parameters: $v_p = 7300$ m/s, $y_0 = 514$ km, $r_{10} = 726.9$ km, $B_r = 150$ MHz, PRF = 3000 Hz, $\lambda = 0.0311$ m, $T_{SA} = 0.64$ s).

After performing these three steps a focused SAR image is obtained. These steps are visualized for a single point target in Figure 18.7. Details on state-of-the-art SAR processing algorithms and on the RCMC can e.g., be found in [3, 20].

The focused image of a single point target is also denoted as impulse response function (IRF). The simulated IRF of a perfectly focused stationary point target is shown in Figure 18.8.

The IRF has the shape of a two-dimensional sinc function. The geometric resolution is determined by the 3 dB width of the IRF. The best achievable azimuth resolution is given in (18.4). If as transmitted waveform the LFM pulse given in (18.12) with a rectangular envelope w_r is used, the best achievable range resolution is [3]

$$\delta_r \cong \frac{c_0}{2B_r}. \quad (18.25)$$

The larger the chirp bandwidth B_r , the better is the range resolution.

Due to the nature of SAR processing, the IRF of any target (independent whether it is moving or not) always is imaged at the position where the Doppler frequency of its uncompressed azimuth signal is zero. For a stationary target this position corresponds to the minimum range r_0 and to the actual azimuth position x_t . In contrast a moving point target, which is discussed in the next section, is displaced in azimuth and to a little extend in range, depending on the motion parameters. For investigating the displacements and additional effects the azimuth and range axis of the IRF plots shown in Figure 18.8, and in some of the Figures provided later in Section 2.18.4, are labeled as azimuth shift and range shift, respectively. The origins of the axes are centered around the positions x_t and r_0 . This has the advantage that the displacement quantities easily can be read off.

2.18.3 Moving point target signal model

The obvious difference between a stationary and a moving point target is the position on ground, which varies over time depending on the target's motion parameters. This time varying position difference

results in a change of the range and Doppler histories and furthermore in some peculiar effects observable in the SAR images. For developing and understanding the fundamental moving target detection and parameter estimation principles and algorithms discussed later, it is first necessary to discuss suitable moving point target signal models. In the next two sections a single-channel as well as a multi-channel signal model are derived.

2.18.3.1 Single-channel signal model

In many of the publications related to GMTI a target in linear motion with constant acceleration during the observation time is assumed. This is especially for short observation times a valid assumption. Furthermore, it is assumed that the target does not change its altitude, i.e., that it does not move in z -direction. This is reasonable since the slopes of common roads only may cause a z -velocity component negligibly small compared to the large x - and y -components. The general acquisition geometry to consider is similar to that sketched in Figure 18.5, apart from the “stationary point target” which has to be replaced by a “moving point target.”

Under the afore mentioned assumption of linear motion the position of the moving point target (i.e., the motion equation) can in contrast to the stationary point target in (18.6) be written as

$$\mathbf{r}_t(t) = \begin{bmatrix} x_t(t) \\ y_t(t) \\ z_t \end{bmatrix} = \begin{bmatrix} x_0 + v_{x0}t + \frac{1}{2}a_x t^2 \\ y_0 + v_{y0}t + \frac{1}{2}a_y t^2 \\ z_0 \end{bmatrix} = \begin{bmatrix} x_0 + v_0 \cos \alpha \cdot t + \frac{1}{2}a \cos \alpha \cdot t^2 \\ y_0 + v_0 \sin \alpha \cdot t + \frac{1}{2}a \sin \alpha \cdot t^2 \\ z_0 \end{bmatrix}, \quad (18.26)$$

where x_0 , y_0 , and z_0 are the positions at $t = 0$, v_{x0} and v_{y0} are the along-track and across-track velocity components at $t = 0$, and a_x and a_y are the constant along-track and across-track acceleration components. If the target moves during the illumination time along a straight line its position can also be expressed using the moving direction or road angle α and the velocity and acceleration magnitudes v_0 and a (see right part of (18.26)). In this case the velocity and acceleration magnitudes are given as $v_0 = \sqrt{v_{x0}^2 + v_{y0}^2}$ and $a = \sqrt{a_x^2 + a_y^2}$. The moving direction α is measured counter-clockwise from the x -axis towards the y -axis as depicted in Figure 18.9.

Even an acceleration change \dot{a} might be considered in the motion equations [21–23]. However, changing accelerations are neglected in this tutorial. They only play a significant role at long illumination times in the order of several seconds and for ISAR imaging purposes.

The position vector of the platform is the same as in (18.5). The range history $r(t) = \|\mathbf{r}_t - \mathbf{r}_p\|$ of the moving target can then be written as

$$r(t) = \sqrt{\left(x_0 + v_{x0}t + \frac{1}{2}a_x t^2 - v_p t\right)^2 + \left(y_0 + v_{y0}t + \frac{1}{2}a_y t^2\right)^2 + (z_0 + z_p)^2}. \quad (18.27)$$

An analytical treatment of this expression is difficult because of the square root. For investigating effects on SAR imagery caused by moving targets it is appropriate to use the third-order Taylor expansion of the range history about $t = 0$ which is given as

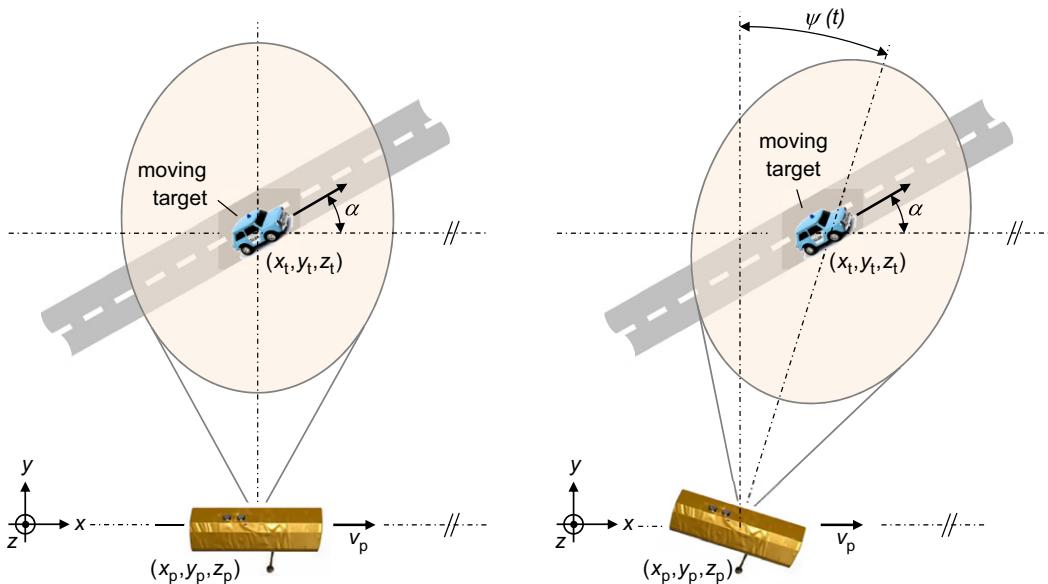


FIGURE 18.9

Target moving along a straight road section (left: non-squinted acquisition geometry; right: squinted acquisition geometry; the squint angle $\psi(t)$ is measured from broadside direction in the slant range plane and its positive counting direction is clockwise).

$$\begin{aligned}
 r(t) \cong & r_{10} + \frac{1}{r_{10}} \{ y_0 v_{y0} + x_0 (v_{x0} - v_p) \} t \\
 & + \frac{1}{2r_{10}} \left\{ x_0 a_x + (v_{x0} - v_p)^2 + v_{y0}^2 + y_0 a_y - \frac{1}{r_{10}^2} [x_0 (v_{x0} - v_p) + y_0 v_{y0}]^2 \right\} t^2 \\
 & + \frac{1}{2r_{10}} \left\{ a_x (v_{x0} - v_p) + v_{y0} a_y \left(1 - \frac{y_0^2}{r_{10}^2} \right) - \frac{x_0}{r_{10}^2} (v_{x0} - v_p)^3 \left[1 + \frac{x_0 a_x + y_0 a_y + v_{y0}^2}{(v_{x0} - v_p)^2} \right] \right. \\
 & \quad \left. - \frac{y_0}{r_{10}^2} [x_0 a_x + (v_{x0} - v_p)^2 + v_{y0}^2] \right\} t^3,
 \end{aligned} \tag{18.28}$$

where the terms in the order of $1/r_{10}^4$ have been dropped. The range between the antenna and the target at $t = 0$ is represented by r_{10} . In contrast to r_0 in (18.8) the range r_{10} corresponds not to the range of closest approach since now the target is in motion and in the most general case not located at broadside position at $t = 0$ (i.e., $x_0 \neq 0$):

$$r_{10} = \sqrt{x_0^2 + y_0^2 + (z_t - z_p)^2}. \tag{18.29}$$

With $x_0 \neq 0$ either a target track not centered in the azimuth beam or a squinted geometry as depicted in Figure 18.9 on the right can be considered. The target position x_0 at time $t = 0$ in this case can be

expressed in terms of the squint angle ψ of the antenna beam: $x_0 = r_{10} \sin \psi$. If the squint angle ψ is zero the antenna points perpendicular to the flight path so that $x_0 = 0$. A squint angle is either caused by a platform yaw due to crosswind or due to antenna beam steering. Typical squint angles caused by platform yaw are in the order of a few degrees. Thus, the x_0/r_{10}^2 terms in (18.28) as well as the $x_0 a_x$ term have no significant contribution and therefore can be neglected. Also the y_0/r_{10}^2 term can be dropped so that the range equation for the moving target simplifies to

$$\begin{aligned} r(t) \cong & r_{10} + \frac{1}{r_{10}} \{y_0 v_{y0} + x_0(v_{x0} - v_p)\}t \\ & + \frac{1}{2r_{10}} \left\{ (v_{x0} - v_p)^2 + v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2}\right) + y_0 a_y \right\} t^2 \\ & + \frac{1}{2r_{10}} \left\{ a_x(v_{x0} - v_p) + v_{y0} a_y \left(1 - \frac{y_0^2}{r_{10}^2}\right) \right\} t^3. \end{aligned} \quad (18.30)$$

The azimuth phase of the moving target signal in the monostatic case (one common transmit (TX) and receiving (RX) antenna) is given by $\varphi_a = -\frac{4\pi}{\lambda} r(t)$. The third-order Taylor expansion of the moving target's Doppler frequency computed with (18.18) is then

$$\begin{aligned} f_a \cong & \underbrace{-\frac{2}{\lambda r_{10}} \{y_0 v_{y0} + x_0(v_{x0} - v_p)\}}_{f_{DC}} \\ & - \underbrace{\frac{2}{\lambda r_{10}} \left\{ (v_{x0} - v_p)^2 + v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2}\right) + y_0 a_y \right\} t}_{k_a} \\ & - \underbrace{\frac{3}{\lambda r_{10}} \left\{ a_x(v_{x0} - v_p) + v_{y0} a_y \left(1 - \frac{y_0^2}{r_{10}^2}\right) \right\} t^2}_{q} \\ = & f_{DC} + k_a t + q t^2, \end{aligned} \quad (18.31)$$

where f_{DC} denotes the Doppler shift, k_a the Doppler slope and q the quadratic Doppler coefficient.

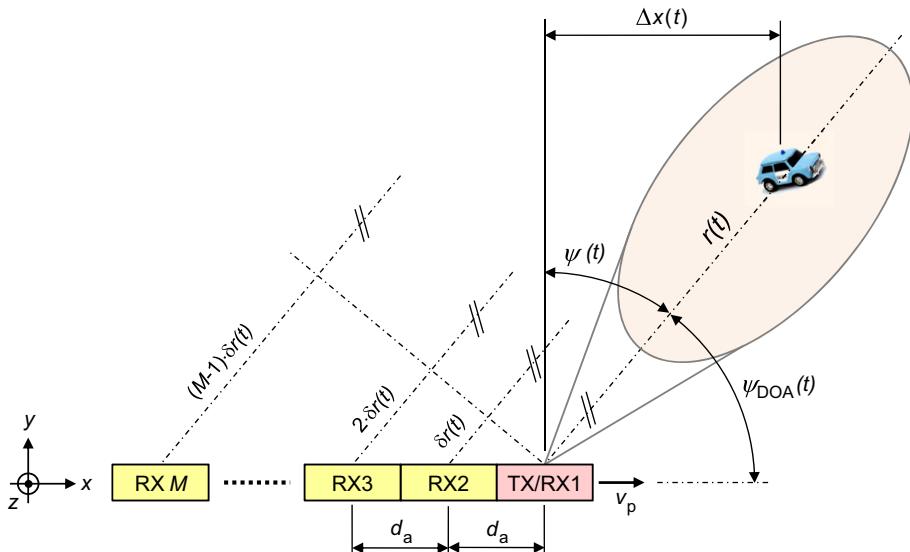
By comparing (18.31) with (18.30) the approximated range history also can be expressed in terms of Doppler parameters:

$$r(t) \cong r_{10} - \frac{\lambda}{2} f_{DC} t - \frac{\lambda}{4} k_a t^2 - \frac{\lambda}{6} q t^3. \quad (18.32)$$

Using this range approximation the single-channel moving target azimuth signal can be written as

$$\begin{aligned} s(t) &= A(t) \cdot \exp \left[-j \frac{4\pi}{\lambda} r(t) \right] \cdot \text{rect} \left[\frac{t}{T_{SA}} \right] \\ &\cong A(t) \cdot \exp \left[j \left(2\pi f_{DC} t + \pi k_a t^2 + \frac{2}{3} \pi q t^3 \right) \right] \cdot \text{rect} \left[\frac{t}{T_{SA}} \right], \end{aligned} \quad (18.33)$$

where the constant phase determined by r_{10} due to its unimportance has been dropped.

**FIGURE 18.10**

Multi-channel SAR geometry with uniformly distributed antennas.

2.18.3.2 Multi-channel signal model

So far the range and Doppler histories of a moving target signal have been derived for the single-channel (monostatic) case where a common antenna is used for both TX and RX of the radar pulses. Now a system with M antennas is considered, where each antenna is separated from its neighbor in the along-track direction by a certain along-track baseline d_a as depicted in Figure 18.10.

The first antenna is used for TX and RX, antenna 2 and all others for receive only. Following the derivation in [23] the ranges $r_m(t)$, $m \in \{1, 2, \dots, M\}$, for $d_a \ll r$ can be written as

$$\begin{aligned} r_1(t) &= r(t), \\ r_2(t) &\cong r_1(t) + \delta r(t), \\ r_m(t) &\cong r_1(t) + (m-1) \cdot \delta r(t). \end{aligned} \quad (18.34)$$

The range difference $\delta r(t)$ is given by [24]

$$\begin{aligned} \delta r(t) &= d_a \cos \psi_{\text{DOA}}(t) = d_a \frac{\Delta x(t)}{r_1(t)} \\ &\cong d_a \left[\frac{x_0}{r_{10}} + \frac{v_{x0} - v_p}{r_{10}} t \right], \end{aligned} \quad (18.35)$$

where $\cos \psi_{\text{DOA}} = \cos(\pi/2 - \psi)$ is the directional cosine measured from the x -axis. The expression after the “ \cong ” sign is the result of a first-order Taylor expansion. The multi-channel azimuth signals

corresponding to the ranges in (18.34) are then

$$\begin{aligned}s_1(t) &= A_1(t) \cdot \exp\left[-j\frac{2\pi}{\lambda}2r_1(t)\right] \cdot \text{rect}\left[\frac{t}{T_{SA}}\right], \\ s_2(t) &= A_2(t) \cdot \exp\left[-j\frac{2\pi}{\lambda}(2r_1(t) + \delta r(t))\right] \cdot \text{rect}\left[\frac{t}{T_{SA}}\right], \\ s_m(t) &= A_m(t) \cdot \exp\left[-j\frac{2\pi}{\lambda}(2r_1(t) + (m-1) \cdot \delta r(t))\right] \cdot \text{rect}\left[\frac{t}{T_{SA}}\right].\end{aligned}\quad (18.36)$$

For the classical GMTI techniques ATI and DPCA treated in Section 2.18.5, the multi-channel azimuth signals in (18.36) need to be aligned or co-registered with the fore channel (=RX1) so that the antenna phase centers are at the same spatial location at different times. Thus, a shift of the signals is necessary. For co-registering with the signal $s_1(t)$ the signal $s_2(t)$ needs to be shifted by Δt , and $s_m(t)$ by $(m-1) \cdot \Delta t$.

For bistatic operation (only the fore antenna transmits, all others receive) the effective phase center separation is $d_{pc} = d_a/2$ so that the time difference can be written as

$$\Delta t = \frac{d_{pc}}{v_p} = \frac{d_a}{2v_p}. \quad (18.37)$$

The co-registered signals are then (to keep the equations shorter the $\text{rect}[\cdot]$ functions have been omitted in the following)

$$\begin{aligned}s_{2,\text{reg}}(t) &= A_2(t + \Delta t) \cdot \exp\left[-j\frac{2\pi}{\lambda}(2r_1(t + \Delta t) + \delta r(t + \Delta t))\right], \\ s_{m,\text{reg}}(t) &= A_m(t + (m-1) \cdot \Delta t) \\ &\quad \cdot \exp\left[-j\frac{2\pi}{\lambda}(2r_1(t + (m-1) \cdot \Delta t) + (m-1) \cdot \delta r(t + (m-1) \cdot \Delta t))\right].\end{aligned}\quad (18.38)$$

It can be shown that the range difference $\delta r(t + (m-1) \cdot \Delta t)$ can be approximated as

$$\delta r(t + (m-1) \cdot \Delta t) \cong \delta r(t) + \frac{v_{x0} - v_p}{r_{10}} d_a \cdot (m-1) \cdot \Delta t \quad (18.39)$$

and the range $r_1(t + (m-1) \cdot \Delta t)$ also can be written as [24]

$$r_1(t + (m-1) \cdot \Delta t) \cong r_1(t) + h(t, \Delta t), \quad (18.40)$$

with

$$h(t, \Delta t) = -\frac{\lambda}{2} f_{DC} \cdot (m-1) \cdot \Delta t - \frac{\lambda}{2} k_a \cdot (m-1) \cdot \Delta t \cdot \left[t + \frac{1}{2} \cdot (m-1) \cdot \Delta t\right]. \quad (18.41)$$

In practice all azimuth signals are sampled with a frequency given by the PRF. The sampling interval corresponds to $T = 1/\text{PRF}$ so that e.g., the time discrete representation of the signals in (18.36) can be written as

$$s_m[n] = s_m(nT), \quad -\frac{N}{2} \leq n < \frac{N}{2} - 1, \quad n \in \mathbb{Z}, \quad (18.42)$$

where N is the total number of azimuth samples. The received discrete azimuth signals (either unregistered or co-registered) can also be collected in a data matrix given as

$$\mathbf{S} = \begin{pmatrix} s_1[n] \\ s_2[n] \\ \vdots \\ s_M[n] \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1N} \\ s_{21} & s_{22} & \cdots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{M1} & s_{M2} & \cdots & s_{MN} \end{pmatrix}, \quad \in \mathbb{C}^{M \times N}, \quad (18.43)$$

where M is the total number of receiving channels. Vectorizing (18.42) by stacking each succeeding column beneath the other yields

$$\mathbf{s} = [s_{11}, s_{21}, \dots, s_{M1}, \dots, s_{1N}, \dots, s_{MN}]^T, \quad \in \mathbb{C}^{MN \times 1}, \quad (18.44)$$

where T means vector transposition.

2.18.4 Effects on SAR imagery

The first who has investigated the GMTI capabilities of SAR was Raney. In his fundamental paper from 1971 he already investigated the basic effects on SAR imagery caused by moving targets [5]. He found that a target motion parallel to the flight path of the radar results in a defocusing of the IRF and, hence, in a decreased peak amplitude and in a decreased signal-to-clutter-plus-noise ratio (SCNR). A motion perpendicular to the flight path causes an azimuth displacement of the target image proportional to the target's across-track velocity. The understanding of these effects is fundamental for deriving appropriate moving target motion and position parameter estimation methods.

In the following sections the range cell migration, the residual range cell migration and the major effects caused by moving targets are discussed. As a starting point it is assumed that the signals are already range compressed in a perfect manner.

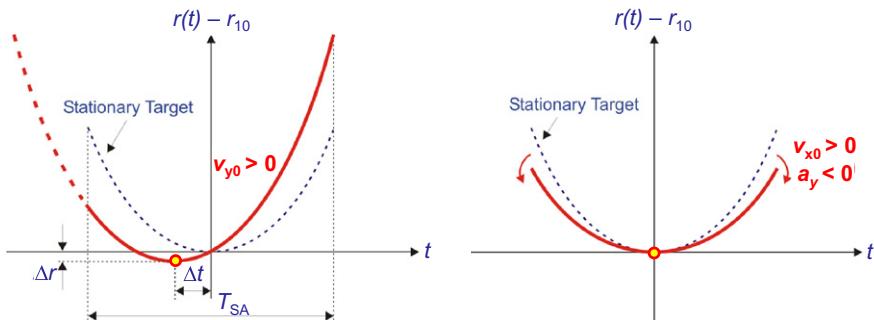
2.18.4.1 Residual range cell migration

Depending on the motion parameters the range histories of a moving target (18.30) and a stationary target (18.7) located at the same position at $t = 0$ are quite different. Examples are shown in Figure 18.11, where the range history of a moving target signal¹ (red) is compared with that of a stationary target (blue).

If the target travels in across-track direction with a certain across-track velocity v_{y0} the range history is shifted in azimuth by Δt and in range by Δr (Figure 18.11, left). The curvature itself is not changed significantly.

When the target travels in along-track direction or accelerates in across-track direction (Figure 18.11, right) the curvature is changed but the range history is not shifted. The range curvature change is equivalent to a quadratic phase error which after conventional SAR processing results in a blurred IRF.

¹For interpretation of color in Figure 18.11, the reader is referred to the web version of this book.

**FIGURE 18.11**

Influence of some motion parameters on the range history (left: influence of across-track velocity; right: influence of along-track velocity and across-track acceleration; the range history of the moving target signal is depicted in red color and the circle marks the moving target position after SAR focusing). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

A SAR processor performs a RCMC adapted only for stationary targets as depicted in the third image from the left in Figure 18.7. This RCMC is conventionally performed in the frequency domain [20]. For computing the residual range cell migration of the moving target signals it is necessary to express the range history as a function of Doppler frequency. The relation between time and Doppler frequency may be easily obtained from (18.31) if the quadratic Doppler coefficient q is neglected (for small illumination times the introduced error is negligibly small):

$$t = \frac{f_a - f_{DC}}{k_a}. \quad (18.45)$$

By inserting this relationship in (18.32) the quadratic approximation of the moving target range history as a function of Doppler frequency is obtained:

$$r(f_a) \cong r_{10} + \frac{\lambda f_{DC}^2}{4k_a} - \frac{\lambda}{4k_a} f_a^2. \quad (18.46)$$

For a stationary target located at the same position as the moving target at time $t = 0$ the quadratic approximation is

$$r_{st}(f_a) \cong r_{10} + \frac{\lambda f_{DC,st}^2}{4k_{a,st}} - \frac{\lambda}{4k_{a,st}} f_a^2 \cong r_0 - \frac{\lambda}{4k_{a,st}} f_a^2, \quad (18.47)$$

where $f_{DC,st}$ and $k_{a,st}$ are the Doppler parameters of the stationary target. These are obtained from (18.31) by setting the motion parameters v_{x0} , v_{y0} , a_x , and a_y to zero so that

$$f_{DC,st} = \frac{2x_0}{\lambda r_{10}} v_p \quad (18.48)$$

and

$$k_{a,st} = -\frac{2v_p^2}{\lambda r_{10}}. \quad (18.49)$$

For a non-squinted acquisition geometry $f_{DC,st}$ and, hence, x_0 are zero.

Since the SAR processor only performs the RCMC correctly for stationary targets a residual range cell migration $\Delta r(f_a)$ remains for moving target signals. Its quadratic approximation is given as

$$\Delta r(f_a) = r(f_a) - r_{st}(f_a) \cong \frac{\lambda}{4} \left(\frac{f_{DC}^2}{k_a} - \frac{f_{DC,st}^2}{k_{a,st}} \right) + \frac{\lambda}{4} \left(\frac{1}{k_{a,st}} - \frac{1}{k_a} \right) f_a^2 \quad \text{for } |f_a| \leq \frac{\text{PRF}}{2}. \quad (18.50)$$

This expression is only valid for signals which are not aliased in Doppler. The residual range cell migration is the reason why the IRF of the moving target also may be blurred in range direction after azimuth compression. A detailed explanation on the range blur is given in Section 2.18.4.2.

An example for the residual range migration for an airborne system is shown in Figure 18.12.

The results show that for targets accelerating in along-track direction (Figure 18.12, third image from the left) and for targets moving with constant velocity in across-track direction (second image from the right) almost no residual range cell migration exists. Thus, the major part of the signal energy is distributed along a single azimuth line. Such signals can easily be extracted from the range-compressed and RCMC data array for parameter estimation purposes discussed later.

A couple of years ago the so called “Keystone Transform” has been introduced with the aim to remove the linear range cell migration of the moving target signals, independent of their motion parameters [25]. However, the final result is the same as obtained by a conventional SAR processor based on chirp scaling [20] with omitted azimuth compression: the linear range cell migration of moving target signals is removed. Thus, if anyhow SAR processing is carried out the application of the Keystone transform is not necessary.

If the Doppler parameters f_{DC} and k_a of a particular moving target signal are known (e.g., after estimation using proper techniques) a RCMC adapted to this target can be performed. An example is shown in Figure 18.13, where the same signals as for Figure 18.12 are used. In this case almost no

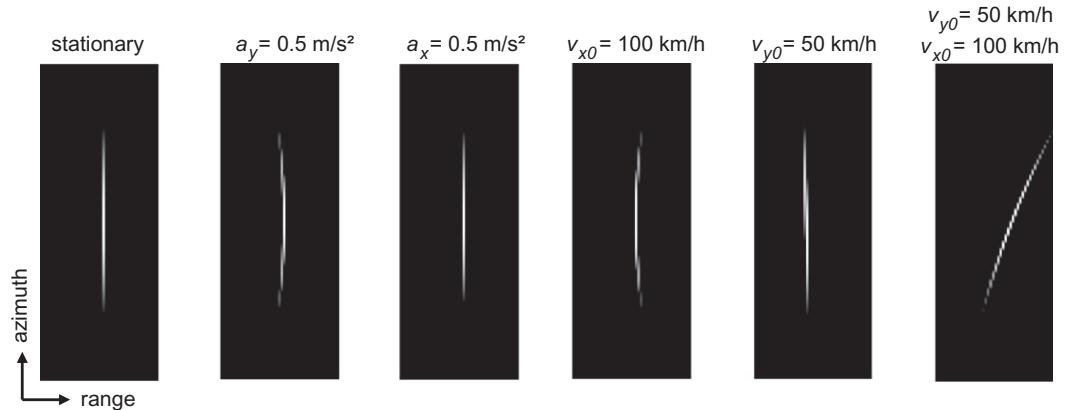
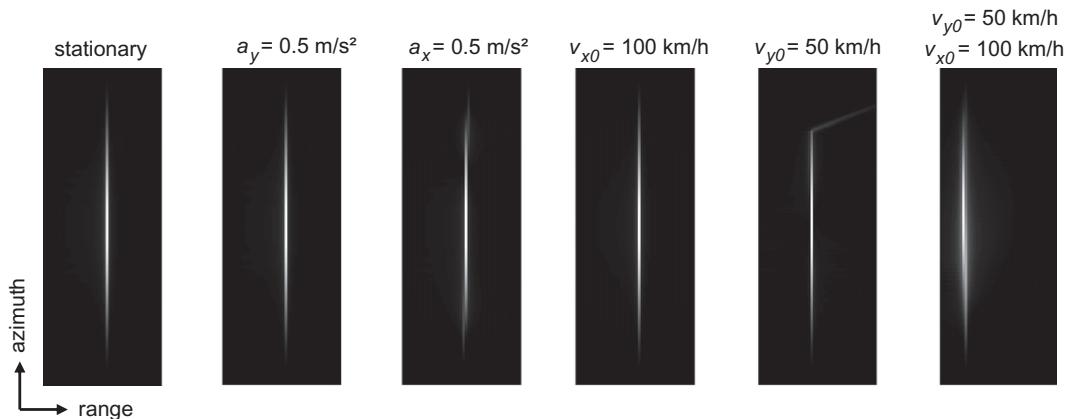


FIGURE 18.12

Simulated range compressed data after conventional RCMC adapted for stationary targets (simulation parameters: $v_p = 90 \text{ m/s}$, $h = |z_0 - z_p| = 2200 \text{ m}$, $r_{10} = 3810 \text{ m}$, $\lambda = 0.0312 \text{ m}$, PRF = 2500 Hz, $B_r = 100 \text{ MHz}$).

**FIGURE 18.13**

Simulated range compressed data after RCMC adapted for each target's Doppler parameters (simulation parameters: $v_p = 90 \text{ m/s}$, $h = |z_0 - z_p| = 2200 \text{ m}$, $r_{10} = 3810 \text{ m}$, $\lambda = 0.0312 \text{ m}$, PRF = 2500 Hz, $B_r = 100 \text{ MHz}$).

residual range cell migration for the moving target signal remains. However, each moving target signal has different Doppler parameters and therefore requires the application of a different adapted RCMC which requires high computational power.

With the proper moving target parameters additionally an adapted azimuth matched filter can be constructed so that a perfectly focused moving target image is obtained [26].

2.18.4.2 Along-track velocity

Equation (18.31) can be used for investigating the effects on the Doppler history. The major effect caused by the along-track velocity v_{x0} is a change of the Doppler slope k_a with respect to the stationary target as sketched in Figure 18.14. Here B_C is the clutter bandwidth and T_{SA} is the illumination time.

The Doppler slope change is equivalent to a change of the quadratic part of the range history. After azimuth compression using the SWMF with Doppler slope $k_{a,st}$ the mismatch $k_{a,st} - k_a$ corresponds to a quadratic phase error in time domain. The IRF of the moving target is therefore defocused in the azimuth direction as shown in Figure 18.15. Unfortunately no analytical description of the defocused IRF exists.

The spread of the blur in azimuth direction can be approximated as

$$\Delta x_{\text{blur}} \cong \left| 1 - \frac{k_a}{k_{a,st}} \right| \cdot T_{SA} v_p = \left| 1 - \frac{(v_{x0} - v_p)^2 + v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2} \right) + y_0 a_y}{v_p^2} \right| \cdot T_{SA} v_p. \quad (18.51)$$

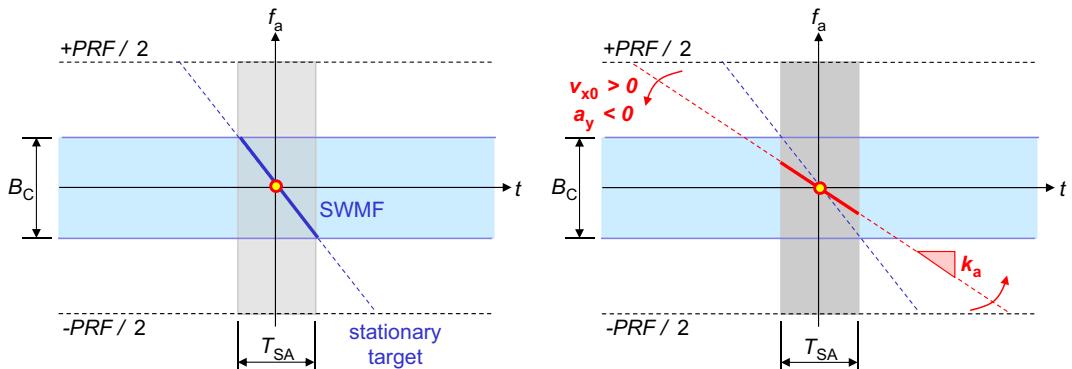


FIGURE 18.14

Doppler history of a stationary target and a SWMF (left), and of a target moving either in along-track direction or accelerating in across-track direction (right). The Doppler history of the stationary target is shown in blue and of the moving target in red color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

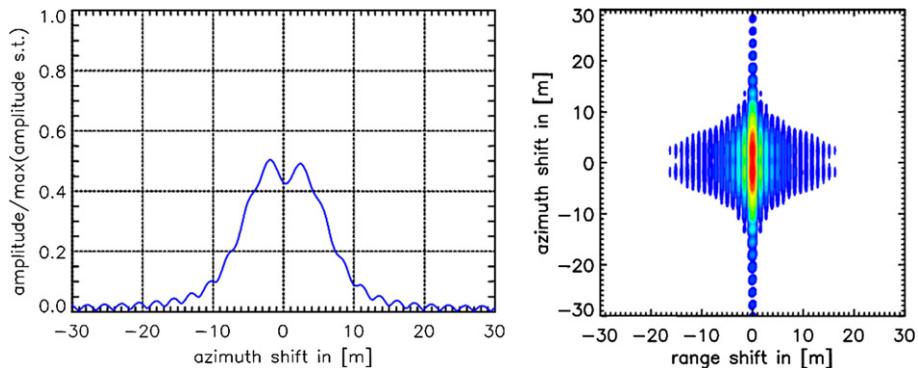


FIGURE 18.15

Impulse response function of a simulated point target moving with constant velocity of $v_{x0} = 50$ km/h in along-track direction, focused with SWMF (left: cut along azimuth; right: 2D representation; system parameters: $v_p = 7300$ m/s, $y_0 = 514$ km, $r_{10} = 726.9$ km, $B_r = 150$ MHz, PRF = 3000 Hz, $\lambda = 0.0311$ m).

If the target only has an along-track velocity component (all other motion parameters are zero) for a system with $v_p \gg v_{x0}$ (i.e., a spaceborne system) the simpler equation

$$\Delta x_{\text{blur}} \cong 2T_{\text{SA}}v_{x0} \quad (18.52)$$

can be used [27]. This equation shows that the moving target is smeared by twice the distance it has moved in the along-track direction during the illumination time T_{SA} . The backscattered signal energy

is distributed over a larger area. With increasing v_{x0} the signal amplitude decreases as shown in Figure 18.15 left. A decreased signal amplitude leads to a decreased $SCNR$ and a lower probability of detection. The target's IRF additionally may be blurred in range due to the residual range cell migration.

For targets with $f_{DC} = 0$ (i.e., zero across-track velocity component) the spread of the blur in range direction can be computed using (18.50) and the relation $B_C \cong k_{a,st} \cdot T_{SA}$ (non-squinted acquisition geometry assumed so that $x_0 = 0$ and $f_{DC,st} = 0$):

$$\Delta r_{blur} \cong 2 \left| \Delta r \left(f_a = \frac{B_C}{2} \right) \right| = \frac{\lambda}{8} \cdot \left| \frac{1}{k_{a,st}} - \frac{1}{k_a} \right| \cdot B_C^2 = \frac{\lambda}{8} \cdot \left| k_{a,st} - \frac{k_{a,st}^2}{k_a} \right| \cdot T_{SA}^2. \quad (18.53)$$

After inserting the Doppler parameters given in (18.31) and (18.49) as result

$$\Delta r_{blur} \cong \frac{v_p^2}{4r_{10}} \cdot \left| 1 - \frac{v_p^2}{(v_{x0} - v_p)^2 + y_0 a_y} \right| \cdot T_{SA}^2 \quad (18.54)$$

is obtained. Especially for airborne systems the range blur may become significant. Let us assume for instance a system with $v_p = 90$ m/s, $y_0 = 2200$ m, $r_{10} = 3810$ m and an extremely long illumination time $T_{SA} = 6$ s. In this case the range blur of the IRF of a target moving in along-track direction with 100 km/h (all other motion parameters are assumed to be zero) is 21 m. However, for spaceborne systems with $v_p \ll v_{x0}$ and $v_p^2 \ll y_0 a_y$ the range blur given in (18.53) and (18.54) can be neglected, especially under the aspect that the typical illumination time is in the order of one second.

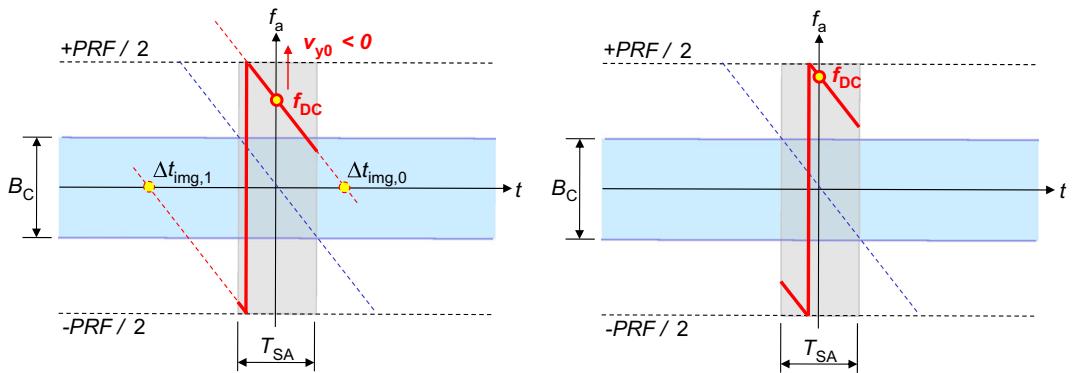
The relation between the range blur, the residual range cell migration and the motion parameters clearly can be recognized by looking again at Figure 18.12. Especially the across-track acceleration and the along-track velocity are the dominant motion parameters responsible for the quadratic phase errors (i.e., the mismatch with the Doppler slope of the SWMF) and, hence, for the residual range cell migration and the azimuth and range blur.

2.18.4.3 Across-track velocity

The major effect caused by the across-track velocity v_{y0} is a change of the Doppler shift f_{DC} given in (18.31). A secondary effect is a slight change of the Doppler slope k_a . In Figure 18.16 the Doppler history of a target moving in across-track direction is sketched. After SAR processing using the SWMF all targets, independent if stationary or moving, are imaged at the positions corresponding to their zero Doppler frequencies. For the moving target signal in Figure 18.16 this may either be the position marked with $\Delta t_{img,0}$ or $\Delta t_{img,1}$ ($\Delta t_{img,1}$ corresponds to the position of an ambiguity that is caused by an aliasing of the signal sampled by the PRF).

The azimuth time $\Delta t_{img,0}$ where the target is imaged can be computed by setting (18.31) to zero and substituting k_a with $k_{a,st}$ for taking into account the Doppler slope of the SWMF. If only the quadratic approximation is used (the cubic coefficient q in (18.31) has been dropped since its contribution compared to the Doppler shift and Doppler slope is negligibly small) the following imaging time is obtained:

$$\Delta t_{img,0} = -\frac{f_{DC}}{k_{a,st}} \quad \text{for } |f_{DC}| \leq \frac{\text{PRF}}{2}. \quad (18.55)$$

**FIGURE 18.16**

Doppler histories of a target moving in across-track direction (in red) and of the SWMF (in blue) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

The imaging time corresponds to an along-track or azimuth displacement of

$$\Delta x_{\text{img},0} = -\frac{f_{\text{DC}}}{k_{\text{a,st}}} v_p. \quad (18.56)$$

For a non-squinted acquisition geometry (i.e., $x_0 = 0$) the equation simplifies to

$$\Delta x_{\text{img},0} \cong -y_0 \frac{v_{\text{r}0}}{v_p} = -r_{10} \frac{v_{\text{r}0}}{v_p}, \quad (18.57)$$

where $v_{\text{r}0}$, denoted as line-of-sight velocity, is the projection of the across-track velocity to the slant range direction. The relationship between the line-of-sight and the across-track velocity for a non-squinted geometry is given as

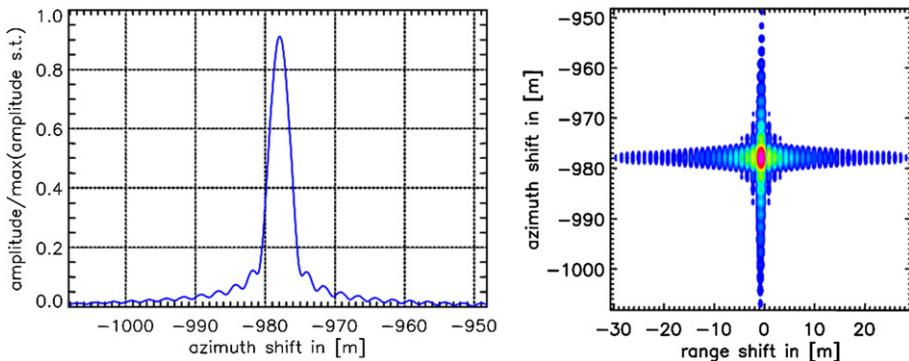
$$v_{\text{r}0} = \frac{y_0}{r_{10}} v_{y0} = v_{y0} \sin \theta_i, \quad (18.58)$$

with θ_i being the incidence angle.

The target is displaced in the flying direction (i.e., $\Delta x_{\text{img},0} > 0$) if it moves towards the radar (i.e., $v_{y0} < 0$). It is displaced in opposite direction (i.e., $\Delta x_{\text{img},0} < 0$) if it moves away from the radar (i.e., $v_{y0} > 0$). In Figure 18.17 an IRF of a simulated point target moving with a constant velocity of 50 km/h in across-track direction is shown. It has a large azimuth displacement of about -978 m. Additionally the IRF is slightly shifted by -0.66 m in range direction (right image).

The reason for the range displacement is the residual range cell migration. It can be computed with (18.50) by taking into account the fact that the major part of moving target's signal energy is located around f_{DC} :

$$\Delta r_{\text{img},0} \cong \Delta r(f_a = f_{\text{DC}}) = \frac{\lambda}{4} \cdot \frac{f_{\text{DC}}^2 - f_{\text{DC,st}}^2}{k_{\text{a,st}}} \quad \text{for } |f_{\text{DC}}| \leq \frac{\text{PRF}}{2}. \quad (18.59)$$

**FIGURE 18.17**

Impulse response function of a simulated point target moving with constant velocity $v_{y0} = 50 \text{ km/h}$ in across-track direction, focused with SWMF (left: cut along azimuth; right: 2D representation; system parameters: $v_p = 7300 \text{ m/s}$, $y_0 = 514 \text{ km}$, $r_{10} = 726.9 \text{ km}$, $B_r = 150 \text{ MHz}$, PRF = 3000 Hz, $\lambda = 0.0311 \text{ m}$).

For a non-squinted acquisition geometry (i.e., $x_0 = 0$ and $f_{DC,st} = 0$) the simplified expression

$$\Delta r_{\text{img},0} \cong -r_{10} \frac{v_{r0}^2}{2v_p^2} = -\frac{y_0^2}{r_{10}} \frac{v_{y0}^2}{2v_p^2} \quad (18.60)$$

can be used. The range displacement is always negative. Thus, a target with a non-zero across-track velocity component is always displaced towards the radar. The range displacements are small if compared to the azimuth displacements, especially for spaceborne systems. For instance, a fast target moving with a line-of-sight velocity of 100 km/h is displaced by only -5 m (typical low-earth orbit platform parameters $r_{10} = 700 \text{ km}$ and $v_p = 7300 \text{ m/s}$ assumed). For airborne systems the displacement is larger. Here the same target is displaced by -143 m ($r_{10} = 3000 \text{ m}$ and $v_p = 90 \text{ m/s}$ assumed).

As shown in Figure 18.17 the target's IRF is severely displaced in azimuth but it is well focused. The slightly decreased peak amplitude is caused by the reduced spectral overlap of the SWMF with the moving target signal. The SWMF acts as a bandpass filter. Its Doppler bandwidth is conventionally limited to the clutter bandwidth B_C given in (18.3). The reason for that is that the clutter for SAR imaging is the wanted signal. Everything outside the clutter bandwidth is uninteresting and therefore is filtered out. The extreme case is shown on the right side of Figure 18.16. Here no spectral overlap between the moving target signal shown in red and the bandwidth B_C of the SWMF exists. As a consequence, fast moving targets are often not visible in conventional processed SAR images. To enable imaging and detection of fast moving targets the processing bandwidth of the SWMF has to be increased to the maximum possible bandwidth determined by the PRF. This is an important point to remember.

Due to the Doppler shift a part of the signal energy may be backfolded (i.e., aliased) so that ambiguities of the IRF appear at certain positions in the SAR image. On the left side of Figure 18.16, the primary IRF containing most of the signal energy is imaged at position $\Delta t_{\text{img},0}$ whereas the ambiguity is imaged

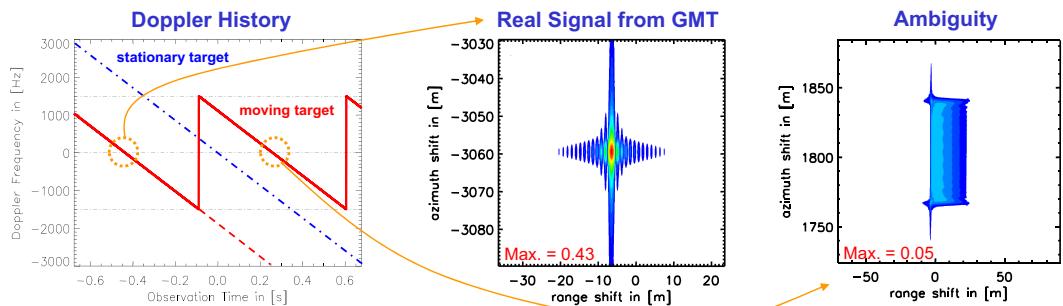


FIGURE 18.18

Impulse response of a simulated point target moving with high velocity in across-track direction with $v_{y0} = 150 \text{ km/h}$ (left: Doppler history; middle: primary impulse response containing most of the energy; right: first ambiguity (note the different scale); system parameters: $v_p = 7300 \text{ m/s}$, $y_0 = 514 \text{ km}$, $r_{10} = 726.9 \text{ km}$, $B_r = 150 \text{ MHz}$, PRF = 3000 Hz, $\lambda = 0.0311 \text{ m}$).

at $\Delta t_{\text{img},1}$. To compute also the ambiguous target positions (18.56) can be extended to

$$\Delta x_{\text{img},N} \cong \frac{N \cdot \text{PRF} - f_{\text{DC}}}{k_{\text{a,st}}} v_p, \quad \text{for } N \in \mathbb{Z}. \quad (18.61)$$

An example of an ambiguous signal is shown in Figure 18.18.

To keep the ambiguities at a low level commonly a high PRF should be chosen. To ensure that at least half of the signal energy of a point target lies within the PRF band its Doppler shift has to be smaller than $|f_{\text{DC}}| \leq \text{PRF}/2$. Using this requirement with (18.31) and again assuming for simplicity a non-squinted geometry (i.e., $x_0 = 0$) the upper bound for the across-track velocity is

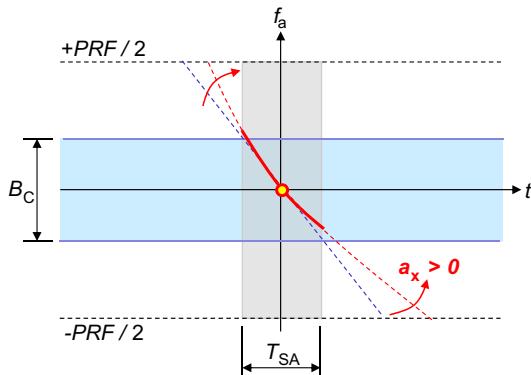
$$|v_{y0}| \leq \lambda \frac{r_{10}}{y_0} \frac{\text{PRF}}{4}. \quad (18.62)$$

This bound ensures that at least half of the signal energy is unambiguously available. The equation can be considered as a criterion for selecting the minimum required PRF of a SAR-GMTI system depending on the highest “expected” target across-track velocity:

$$\text{PRF}_{\min,3 \text{ dB}} = \frac{4y_0}{\lambda r_{10}} |v_{y0,\max}|. \quad (18.63)$$

For targets with $f_{\text{DC}} \gg 0$ (i.e., for targets moving with a certain across-track velocity) the spread of the range blur can be approximated as ($x_0 = 0$ and $f_{\text{DC,st}} = 0$ assumed)

$$\Delta r_{\text{blur}} \cong \left| \Delta r \left(f_a = f_{\text{DC}} - \frac{B_C}{2} \right) - \Delta r \left(f_a = f_{\text{DC}} + \frac{B_C}{2} \right) \right| = \frac{\lambda}{2} \cdot \left| 1 - \frac{k_{\text{a,st}}}{k_a} \right| \cdot f_{\text{DC}} \cdot T_{\text{SA}}. \quad (18.64)$$

**FIGURE 18.19**

Doppler history of a target accelerating in along-track direction.

With the inserted Doppler parameters given in (18.31) this results in

$$\Delta r_{\text{blur}} \approx \frac{y_0}{r_{10}} \cdot \left| 1 - \frac{v_p^2}{(v_{x0} - v_p)^2 + v_{y0} \left(1 - \frac{y_0^2}{r_{10}^2} \right) + y_0 a_y} \right| \cdot |v_{y0}| \cdot T_{\text{SA}}. \quad (18.65)$$

If for instance an airborne system with $v_p = 90 \text{ m/s}$, $y_0 = 2200 \text{ m}$, $r_{10} = 3111 \text{ m}$, and $T_{\text{SA}} = 6 \text{ s}$ would be used the IRF of a target moving in across-track direction with 50 km/h (all other motion parameters are assumed to be zero) has a range blur of 0.05 m . For state-of-the-art airborne SAR systems this is below the achievable range resolution and therefore can be neglected. However, if aside of the across-track velocity of 50 km/h also an along-track velocity component of 100 km/h is considered, the range blur increases to 52 m .

2.18.4.4 Along-track acceleration

The major effect of the along-track acceleration a_x is a change of the quadratic coefficient q in (18.31) which causes a deflection of the Doppler history as sketched in Figure 18.19.

After azimuth compression the IRF of the moving target has a decreased peak amplitude and shows non-symmetric (unbalanced) sidelobes. The strength of this effect depends mainly on the synthetic aperture time T_{SA} . The longer this time the larger are the third-order phase errors in time domain and the more severe the effect. An example for an airborne system is shown in Figure 18.20.

For spaceborne SAR systems with typical illumination times below one second the effects caused by the along-track acceleration are negligibly small.

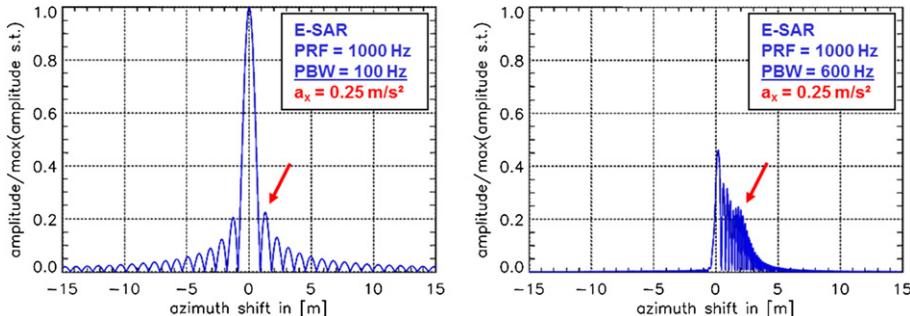


FIGURE 18.20

Impulse response functions of a simulated point target accelerating in along-track direction, focused with SWMF (left: the bandwidth of the SWMF was 100 Hz corresponding to an integration time of about 1 s; right: bandwidth of 600 Hz corresponding to about 6 s integration time; system parameters: $v_p = 90 \text{ m/s}$, $y_0 = 3500 \text{ m}$, $r_{10} = 4950 \text{ m}$, $B_r = 100 \text{ MHz}$, PRF = 1000 Hz, $\lambda = 0.0311 \text{ m}$).

2.18.4.5 Across-track acceleration

The across-track acceleration a_y causes a similar major effect as the along-track velocity v_{y0} : a change of the Doppler slope k_a given in (18.31) (see also Figure 18.14). This results again in a blurred IRF with decreased peak amplitude. The IRF looks similar as the response shown in Figure 18.15 and is therefore not separately depicted. The azimuth blur caused by the across-track acceleration can be derived with (18.51). If all other target motion parameters are zero it is given as

$$\Delta x_{\text{blur}} \cong \frac{y_0}{v_p} a_y T_{\text{SA}} \quad (18.66)$$

Even an allegedly small acceleration may cause an azimuth blur in the order of several meters. For an airborne system with $y_0 = 2200 \text{ m}$, $r_{10} = 3111 \text{ m}$, $v_p = 90 \text{ m/s}$, and $T_{\text{SA}} = 6 \text{ s}$ an acceleration of 0.1 m/s^2 results in an azimuth blur of approximately 15 m. For a spaceborne system with $y_0 = 514 \text{ km}$, $r_{10} = 729 \text{ km}$, $v_p = 90 \text{ m/s}$, and $T_{\text{SA}} = 0.64 \text{ s}$ the azimuth blur is 4.5 m. Neither for airborne nor for high resolution spaceborne SAR systems this effect should be neglected. By measuring the spread of the azimuth blur of the moving target's IRF the effects caused by across-track accelerations and along-track velocities cannot be separated.

Additionally to the azimuth blur the IRF is also blurred in range. The range blur caused by the across-track acceleration can be computed with (18.54). Using the same parameters for an airborne and spaceborne system as before, the range blurs caused by an acceleration of 0.1 m/s^2 are 0.6 m and 7 mm. Thus, at least for the spaceborne system, the range blur can be neglected.

2.18.4.6 Summary of effects

The effects treated so far caused by moving targets on SAR imagery are summarized in Table 18.1. The knowledge of these effects and their origins are fundamental for developing suitable GMTI systems and algorithms [9].

Table 18.1 Major Effects on SAR Imagery Caused by Moving Targets [9]

Motion parameter	Major effects on SAR imagery/IRF	System/GMTI algorithm aspects
Across-track velocity	<ul style="list-style-type: none"> Shift of Doppler frequency leads to azimuth displacement of moving vehicles, since in common SAR processing all vehicles are imaged at their corresponding zero Doppler positions Small shift of impulse response in range Decrease of peak amplitude since the spectral overlap between the moving vehicle's Doppler history and the Doppler history of the SWMF decreases with higher across-track velocities 	<ul style="list-style-type: none"> To avoid ambiguities due to Doppler aliasing (= backfolding), the PRF has to be increased For a non-squinted SAR acquisition geometry the Doppler shift is only determined by the across-track velocity \Rightarrow if the Doppler shift can be estimated also the across-track velocity can be computed During SAR processing the full Doppler bandwidth given by the PRF has to be considered, otherwise fast moving targets will be filtered and will not appear in the SAR image
Along-track velocity	<ul style="list-style-type: none"> Change of Doppler slope and hence azimuth and range defocusing of impulse response due to mismatch with Doppler slope of SWMF and residual range cell migration 	<ul style="list-style-type: none"> The Doppler slope can be estimated with proper methods but along-track velocity and across-track acceleration cannot be separated using only the Doppler slope information
Across-track acceleration	<ul style="list-style-type: none"> Causes same effect as along-track velocity 	
Along-track acceleration	<ul style="list-style-type: none"> Deflection of Doppler history; this leads to unbalanced sidelobes of the impulse response function and to a decrease of peak amplitude. 	<ul style="list-style-type: none"> At systems with high platform velocities and short observation times (typical spaceborne systems) this effect can be ignored

2.18.5 Classical dual-channel techniques

In the following two sections the classical dual-channel techniques are discussed. They principally can be applied on raw data, range compressed data and fully focused data.

2.18.5.1 Along-track interferometry

For along-track interferometry (ATI) two receiving antennas displaced in azimuth direction by a certain baseline d_a are necessary. The receiving antennas can either be mounted on the same platform or on separate platforms flying in formation along the same track (separate platforms allow for larger baselines). Anyhow, each antenna observes the scene from the same point in space at slightly different times t_1 and t_2 . During the time lag $\Delta t = t_2 - t_1$ the signals from stationary targets remain the same whereas the signals from targets moving in range direction experience a phase shift $\Delta\phi$. This phase shift is called ATI phase. It is proportional to the line-of-sight velocity of the moving target. The time lag Δt has to be sufficiently short to avoid the effect of temporal decorrelation caused by slight changes in the scene (e.g., due to wind) between the acquisitions of an interferometric image pair. The first application of ATI was oceanography where it was used to measure tidal currents with a velocity estimation accuracy in the order of several cm/s [28].

The ATI principle is sketched in Figure 18.21. At time t_1 the target is observed by the fore antenna at range r_1 . At time t_2 it is observed by the aft antenna at a different range r_2 . The range difference $\Delta r(t) = r_2(t + \Delta t) - r_1(t)$ is proportional to the ATI phase. The ATI signal is computed by multiplying the signal of the fore antenna with the complex conjugate and co-registered signal of the aft antenna (remember the co-registered multi-channel signals given in (18.38)):

$$s_{\text{ATI}}(t) = s_1(t) \cdot s_{2,\text{reg}}^*(t) = A_1(t) \cdot A_{2,\text{reg}}^*(t) \cdot \exp \left\{ j \underbrace{\left(\varphi_1(t) - \varphi_{2,\text{reg}}(t) \right)}_{= \Delta\phi(t)} \right\}. \quad (18.67)$$

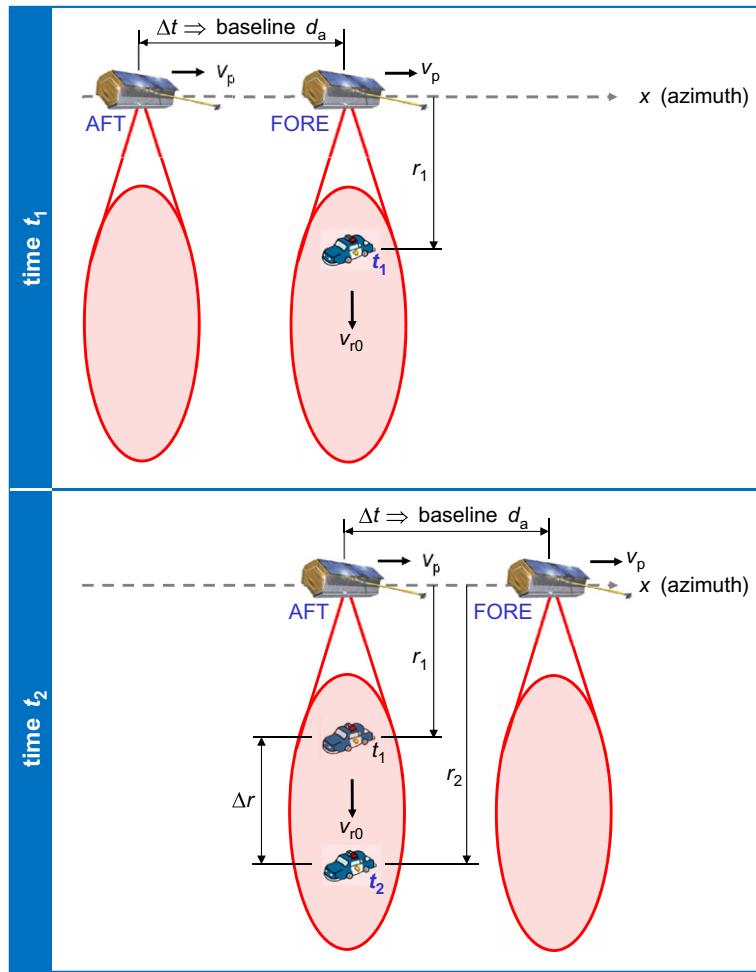
If both channels are well calibrated and the RCS of the target does not change between both observations the complex coefficients A_1 and $A_{2,\text{reg}}$ are identical (i.e., same phases and amplitudes), so that the ATI phase may be computed as:

$$\Delta\phi(t) = \arg\{s_{\text{ATI}}(t)\} = \varphi_1(t) - \varphi_{2,\text{reg}}(t). \quad (18.68)$$

The phases of both signals are given as

$$\begin{aligned} \varphi_1(t) &= -\frac{4\pi}{\lambda} r_1(t), \\ \varphi_{2,\text{reg}}(t) &= -\frac{4\pi}{\lambda} r_2(t + \Delta t) = -\frac{4\pi}{\lambda} r_{2,\text{reg}}(t), \end{aligned} \quad (18.69)$$

with $\Delta t = d_{\text{pc}}/v_p$ where d_{pc} is the antenna phase center separation in along-track direction (= effective along-track baseline). If both antennas receive and transmit independently (= monostatic operation) the phase center separation corresponds to the physical antenna separation, i.e., $d_{\text{pc}} = d_a$. In case of bistatic

**FIGURE 18.21**

Along-track interferometry principle.

operation where only the fore antenna is used for transmission but both antennas for signal reception the effective along-track baseline reduces to half the physical distance, i.e., $d_{pc} = d_a/2$.

By inserting (18.69) in (18.68) the ATI phase may be written as

$$\Delta\phi(t) = \frac{4\pi}{\lambda} [r_{2,\text{reg}}(t) - r_1(t)] = \frac{4\pi}{\lambda} \Delta r(t). \quad (18.70)$$

If the target moves with constant line-of-sight velocity v_{r0} the ATI phase of the target at broadside position for monostatic operation in the simplest case can be approximated as (the squint angle and all

other motion parameters of the target are assumed to be zero) [28]

$$\Delta\phi \cong \frac{4\pi}{\lambda} v_{r0} \Delta t. \quad (18.71)$$

If the line-of-sight velocity is computed from the measured ATI phase blind velocities and ambiguities may occur. The reason is that the ATI phase can only be measured in fractions of 2π . A blind velocity occurs if the measured ATI phase becomes zero. This is the case for

$$\Delta\phi = N \cdot 2\pi, \quad \text{for } N \in \mathbb{Z}. \quad (18.72)$$

The blind velocities can be computed by inserting (18.71) in the previous equation:

$$v_{r0,\text{blind}} = \frac{\lambda}{2\Delta t} N. \quad (18.73)$$

If the target moves with a blind velocity its ATI phase cannot be discriminated from the ATI phase of a stationary target and, hence, it cannot be detected in the interferogram.

An ambiguity occurs if the ATI phase is equal or larger than π . In this case neither the motion direction (away or towards the radar) nor the velocity can be determined unambiguously. The maximum line-of-sight velocity causing no ambiguities is given as (again derived with (18.71))

$$|v_{r0,\text{max}}| < \frac{\lambda}{4\Delta t}. \quad (18.74)$$

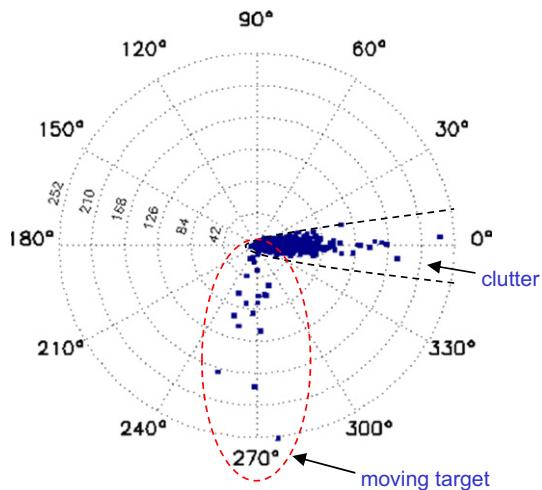
The maximum line-of-sight velocity can be increased by using larger radar wavelengths λ or smaller time lags Δt (i.e., smaller baselines).

In Figure 18.22 a polar plot of an azimuth line of interest containing a moving target signal is shown. The azimuth line was focused with the SWMF before the polar plot was generated. Several dots in the polar plot belong to the target since it spans multiple resolution cells. It also can be seen that the target dots have slightly different phases. The reason is that the target is not perfectly focused (remember the potential mismatch of the Doppler slope of the target and the SWMF). Furthermore, since ATI performs no clutter suppression also the clutter contribution disturbing the ATI phase clearly can be recognized. As a consequence the ATI phases of moving target signals are biased towards zero so that velocities are generally underestimated. For low SCNR values this bias is more significant.

Even if there is no temporal decorrelation and if the clutter would be suppressed somehow prior to ATI phase computation, the minimum detectable velocity (MDV) and the accuracy of the line-of-sight velocity estimate is limited by the thermal receiver noise. The receiver noise in both RX channels causes a phase decorrelation which results in a noisy estimate of the ATI phase. The amplitude of the complex correlation coefficient due to thermal receiver noise between both channels is given as [29]

$$|\gamma_{\text{SNR}}| = \frac{1}{1 + \text{SNR}^{-1}}, \quad (18.75)$$

where SNR is the signal-to-noise ratio after SAR focusing (i.e., after pulse compression). The standard deviation of the ATI phase for point-like scatterers for values of γ_{SNR} close to one can be approximated as [30,31]

**FIGURE 18.22**

Polar plot of ATI phases of an extracted azimuth line containing a moving target signal. The data was acquired with DLR's airborne sensor E-SAR.

$$\sigma_{\Delta\phi} \cong \sqrt{\frac{1 - \gamma_{\text{SNR}}^2}{2\gamma_{\text{SNR}}^2}}. \quad (18.76)$$

This equation can be considered as a lower bound of the ATI phase standard deviation. For extended objects (i.e., targets larger than a single SAR resolution cell) the standard deviation is larger but it can be reduced to a certain degree by ATI phase averaging (=multi-looking).

The lower bound of the standard deviation of the line-of-sight velocity computed with (18.71) and (18.75) can be expressed as

$$\sigma_{v_{r0}} = \left| \frac{\lambda}{4\pi \Delta t} \right| \sigma_{\Delta\phi}. \quad (18.77)$$

The standard deviation of the azimuth position computed using (18.56) is then

$$\sigma_{\Delta x_{\text{img}}} = \left| -\frac{r_{10}}{v_p} \right| \sigma_{v_{r0}} = \left| -\frac{\lambda r_{10}}{4\pi v_p \Delta t} \right| \sigma_{\Delta\phi}. \quad (18.78)$$

By having a closer look at (18.77) and (18.78) it is obvious that the standard deviation can be decreased by increasing the time lag Δt and, hence, the antenna phase center separation d_{pc} . However, it has to be kept in mind that by increasing Δt the maximum unambiguously detectable line-of-sight velocity in (18.74) and the blind velocities in (18.73) decrease. Techniques for phase unwrapping are necessary if the motion and position parameters of faster moving targets shall be estimated with high accuracy. Promising candidates are for instance multi-baseline ATI techniques [32,33]. Ambiguities might also be resolved if the range cell migration of the target is exploited for a rough line-of-sight

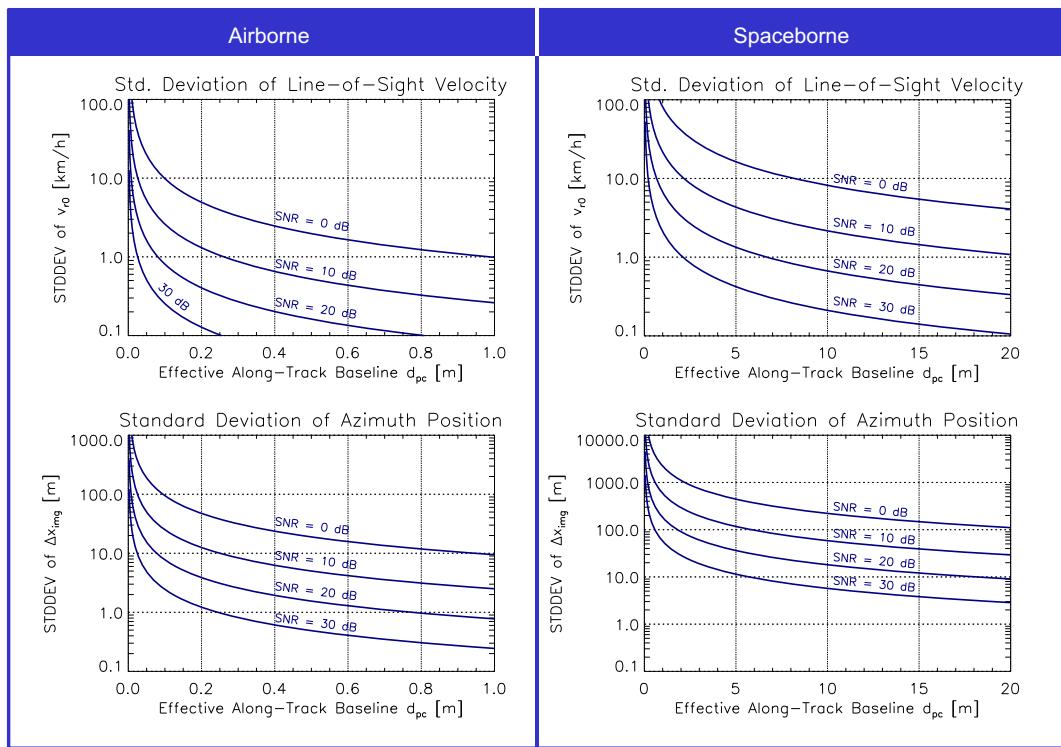


FIGURE 18.23

Standard deviations of the estimates of line-of-sight velocity and azimuth position for an exemplary airborne (left) and spaceborne (right) SAR acquisition geometry (airborne parameters: $\lambda = 3.12 \text{ cm}$, $r_{10} = 3111 \text{ m}$, $v_p = 90 \text{ m/s}$; spaceborne parameters: $\lambda = 3.12 \text{ cm}$, $r_{10} = 726.9 \text{ km}$, $v_p = 7450 \text{ m/s}$).

velocity estimation which afterwards is refined with the information obtained from the ATI phase [21]. For large time lags, temporal decorrelation may also have a negative influence, but this aspect is not treated here.

Figure 18.23 shows the standard deviations of the line-of-sight velocity and the azimuth position for exemplary airborne and spaceborne acquisition geometries. It is obvious that with airborne systems even with small baselines in the fraction of a meter good estimation results can be achieved. For spaceborne systems comparatively large baselines are required to push down the standard deviations and, hence, to improve the parameter estimation accuracy. However, the size and weight of SAR satellites are limited by launcher restrictions. For instance, the German TerraSAR-X satellite has an antenna length of 4.8 m which can be split into two RX sub-apertures of 2.4 m each. In bistatic operation (i.e., TX with full antenna and RX with both halves simultaneously) the phase center separation or the effective along-track baseline, respectively, is only 1.2 m. For high SNR values of 20 dB the azimuth position estimation accuracy is in the order of 150 m. The accuracy decreases to 490 m if the SNR decreases to 10 dB. Compared to airborne systems, state-of-the-art spaceborne systems generally achieve worse

position and motion parameter estimation accuracies. The two main reasons are the too short baselines and the worse SNR values.

For an in-depth ATI performance analysis also the clutter contribution has to be considered [34, 35]. However, for understanding the basic ATI principles and performance limitations the previous explanations are sufficient.

In reality a moving target seldom has only a line-of-sight velocity component with all other motion components being zero as assumed in the derivation of (18.71). Although ATI is primarily sensitive to the line-of-sight velocity v_{r0} also the other motion components contribute to the ATI phase. Unfortunately there exists no analytical description for the ATI signal focused with a SWMF. However, if the Doppler slope of the matched filter used for azimuth compression is adapted to the Doppler slope k_a of the target an analytical solution exists (matched filters adapted to target motion's parameters are discussed in Section 2.18.7.1). In case of a non-squinted geometry (i.e., $x_0 = 0$ and $f_{DC,st} = 0$) the ATI phase of the peak of the refocused target signal is given as [23]

$$\Delta\phi(t = \Delta t_{img,0}) = \frac{\pi d_a y_0}{\lambda r_{10}} \cdot \left[\frac{1}{v_p} + \frac{2(v_{x0} - v_p)}{\lambda r_{10} k_a} \right] \cdot v_{y0}. \quad (18.79)$$

The ATI phase depends not only on the across-track velocity v_{y0} but also on the along-track velocity v_{x0} . This has to be taken into account for accurate parameter estimation.

2.18.5.2 Displaced phase center antenna technique

The displaced phase center antenna (DPCA) technique is one of the simplest clutter suppression techniques in radar [36]. The antenna configuration is exactly the same as in ATI: at least two receiving antennas are needed (see also Figure 18.21). The only difference is the signal processing. Instead of a complex conjugate multiplication the co-registered signal received by the aft antenna is subtracted from the signal received by the fore antenna for obtaining the DPCA signal

$$s_{DPCA}(t) = s_1(t) - s_{2,reg}(t). \quad (18.80)$$

Signals from stationary targets are canceled since they are identical in both successive observations. Targets moving with sufficient line-of-sight velocity cause a certain phase shift and, thus, will not be canceled. As a consequence even slowly moving targets otherwise masked by the clutter can be detected in the “clutter-suppressed” DPCA image. If the clutter is homogeneous DPCA can be considered as the optimal linear filter for target detection [37]. Although DPCA originally was developed for dual-channel systems, it also can be used with more than two channels [38–40].

In older literature it is often stated that the so called “DPCA Condition” has to be fulfilled. That means that the PRF has to be selected in a way that the spatial location of the fore antenna phase center when an echo is received is the same as the location of the aft antenna phase center when the next echo is received:

$$PRF = \frac{1}{\Delta t}. \quad (18.81)$$

Nowadays this restriction is relaxed. By modern signal processing techniques co-registration can be performed with sub-sample accuracy using interpolation or resampling, respectively. Often the co-registration is done by applying a phase ramp on the signal after transforming it via an FFT to

the Doppler domain (remember the displacement law of the Fourier transform). The phase ramp to be applied in Doppler domain for co-registration along azimuth generally is given as

$$S_{k_\varphi}(f_a) = \exp(-j2\pi f_a \Delta t). \quad (18.82)$$

This leads to accurate co-registration results as long as the Nyquist sample theorem is fulfilled for both clutter and moving target signals. In other words, a sufficiently high PRF is required for avoiding aliasing of signals in Doppler. However, if the PRF is too low the aliased clutter signals after co-registration have a constant phase error [18]. A coarse co-registration without interpolation can be done by shifting the data by an integer amount of samples. The time in (18.82) for fine co-registration with sub-sample accuracy can then be decreased to $\Delta t_{\text{shift}} (\Delta t_{\text{shift}} < \Delta t)$. The constant phase error derived in [18] in this case can be written as

$$\delta_\varphi = \pm 2\pi \cdot \Delta t_{\text{shift}} \cdot \text{PRF}. \quad (18.83)$$

Due to this constant phase error aliased clutter might be mistaken as false moving targets. Especially for spaceborne systems the Doppler bandwidth is large compared to the PRF so that Doppler aliasing cannot be excluded. It generally is more severe than in airborne systems. Therefore, for spaceborne systems it is advisable to meet the DPCA condition in (18.81) as accurate as possible so that no phase ramp in Doppler for fine-co-registration needs to be applied. In this case Δt_{shift} is zero and the phase error in (18.83) vanishes.

For achieving a good detection performance it is essential that the receiving channels are well calibrated. For calibration the “Adaptive 2D Channel Balancing” method originally proposed by Ender [7] has established itself in the GMTI community. The method operates in the 2D frequency domain. The channel transfer functions are adapted to a reference channel, conventionally the fore channel. The method is not limited to two channels. Also co-registration is performed between the channels since any phase ramps in frequency domain are removed [41]. In Figure 18.24 an example with three moving targets is shown. It can be seen that after channel balancing (right column) the SCNR as well as the correlation coefficient γ_{SNR} is increased. Even targets with lower RCS are clearly visible in the balanced DPCA image. Theoretically the clutter can be suppressed down to noise level. The application of a channel balancing method is absolutely crucial, especially for airborne GMTI algorithms based on classical ATI and DPCA.

If the channels are well calibrated so that $A_1(t) \cong A_{2,\text{reg.}}(t)$, the DPCA signal can also be expressed as

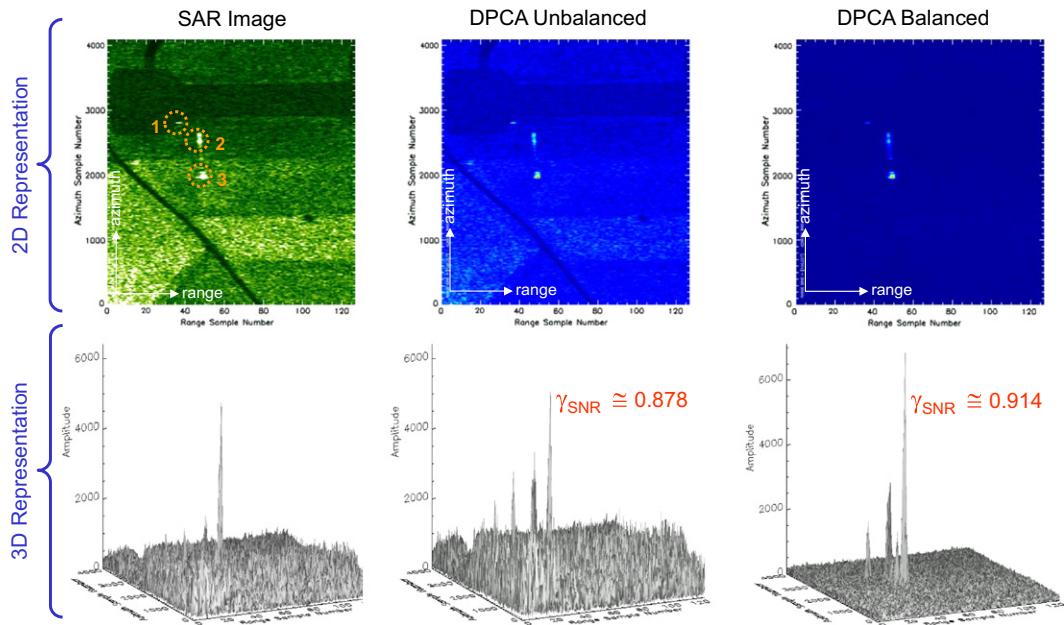
$$s_{\text{DPCA}}(t) = 2|A_1(t)| \cdot \left| \sin \frac{\varphi_1(t) - \varphi_{2,\text{reg.}}(t)}{2} \right| \cdot \exp \left\{ j \left(\frac{\varphi_1(t) + \varphi_{2,\text{reg.}}(t)}{2} + \varphi_{A1}(t) + \frac{\pi}{2} \right) \right\}, \quad (18.84)$$

where $\varphi_{A1}(t)$ is the phase of the complex coefficient $A_1(t)$. For detecting a moving target the magnitude of the DPCA signal is of interest:

$$|s_{\text{DPCA}}(t)| = 2|A_1(t)| \cdot \left| \sin \frac{\varphi_1(t) - \varphi_{2,\text{reg.}}(t)}{2} \right|. \quad (18.85)$$

If the target moves with constant line-of-sight velocity v_{r0} (all other motion parameters are zero) the DPCA magnitude can be approximated as

$$|s_{\text{DPCA}}(t)| \cong 2|A_1(t)| \cdot \left| \sin \left(\frac{2\pi}{\lambda} v_{r0} \Delta t \right) \right|. \quad (18.86)$$

**FIGURE 18.24**

Clutter suppression capabilities of unbalanced DPCA and balanced DPCA (left column: single-channel airborne SAR image containing three targets moving in range direction; middle: DPCA image generated from unbalanced two-channel data; right: DPCA image generated from balanced data).

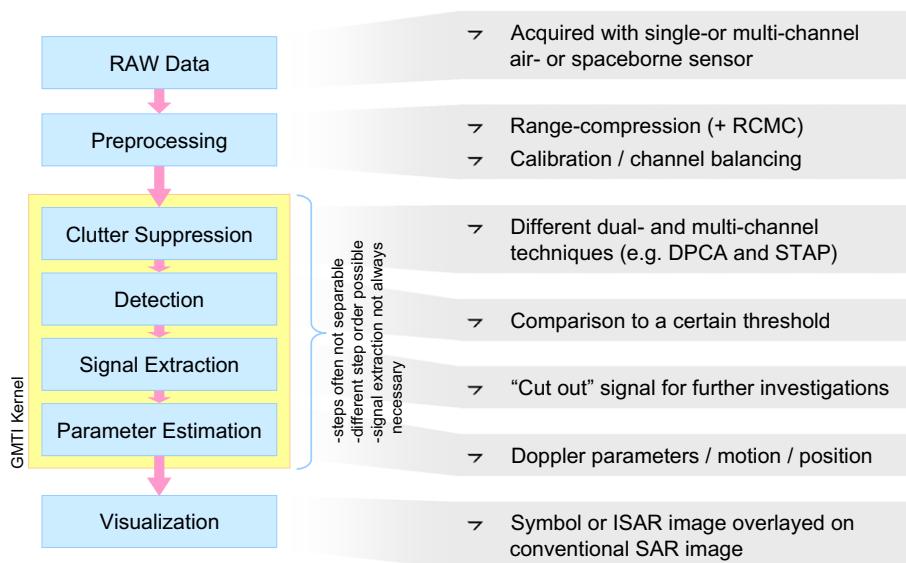
It is obvious that the DPCA magnitude drops to zero if the sine becomes zero. This is the case if the target moves with a blind velocity given in (18.73). Thus, neither with ATI nor with DPCA targets moving with blind velocities can be detected.

2.18.6 General GMTI processing chain

In this section the general GMTI processing chain is presented and discussed. Figure 18.25 shows a simplified flow chart containing the typical GMTI processing steps. At the very first beginning the acquired radar data in their raw form are necessary as input. These data somehow need to be preprocessed before the “GMTI Kernel” can be applied.

The “GMTI Kernel” represents a general GMTI algorithm with the processing steps: clutter suppression, detection, signal extraction, and parameter estimation. Principally every state-of-the-art GMTI algorithm consists of these steps. However, the steps are not always clearly separable and the step order may be different. Especially “Signal Extraction” sometimes is not necessary.

In the following sections the different steps are explained.

**FIGURE 18.25**

Simplified flow chart of general GMTI processing steps.

2.18.6.1 Preprocessing

As input for a SAR-GMTI processor single- or multi-channel raw data acquired with air- or spaceborne sensors are used. These data need to be preprocessed (step 2 in Figure 18.25) before the “GMTI Kernel” can be applied. Depending on the particular GMTI algorithm to be used in the “GMTI Kernel” several preprocessing options are conceivable (no claim to completeness):

- Range compression.
- Range compression with RCMC.
- Conventional SAR processing.
- Motion adapted SAR processing.

2.18.6.1.1 Range compression

The acquired data are only compressed in range, no RCMC and no azimuth compression is carried out. This kind of preprocessing is used especially for airborne GMTI algorithms based on STAP [4,37]. Most of these algorithms use only short integration times so that range cell migration can be neglected. Large SNR values are required since due to the low sample support the compression gain obtained by coherent integration is limited. However, with state-of-the-art airborne systems this is not a problem in contrast to spaceborne systems which have to cope with comparatively low SNR values.

2.18.6.1.2 Range compression with RCMC

The data are range compressed and a RCMC is performed. The RCMC may either be adapted to stationary targets, as done during conventional SAR processing, or in an iterative way to some of the moving target's motion parameters. A detailed discussion on RCMC and residual range cell migration was already given in Section 2.18.4.1 and shall not be repeated here.

2.18.6.1.3 Conventional SAR processing

The data are processed with a conventional SAR processor, e.g., a Range/Doppler [3] or a Chirp Scaling Processor [20]. The full Doppler bandwidth given by the PRF has to be considered to avoid filtering of signal energy of fast moving targets (cf. Section 2.18.4.3).

2.18.6.1.4 Motion adapted SAR processing

The data are processed iteratively using different assumptions of target's motion parameters. This kind of preprocessing requires high computational power. However, the advantage is that high SCNR values may be achieved. This especially is crucial for spaceborne systems which generally suffer from low SNR and SCNR values (cf. Section 2.18.5.1).

2.18.6.2 Clutter suppression

Most of the multi-channel GMTI algorithms found in the literature use either DPCA or STAP techniques for clutter suppression.

2.18.6.2.1 DPCA

Dual-channel algorithms mainly use the DPCA technique discussed in Section 2.18.5.2 for clutter suppression. This technique can also be used if more than two RX channels are available [38, 39]. For instance several clutter suppressed image pairs can be generated using DPCA. The ATI phase between these image pairs is then less influenced by clutter [42]. This improves the moving target parameter estimation accuracy significantly. In the ideal case the clutter is suppressed down to noise level so that the curves shown in Figure 18.23 apply for the velocity and position estimation accuracy.

In the general case the clutter filtered signal y_{out} is obtained by a linear combination of the received multi-channel signal or space-time snapshot z with the clutter filter or weight vector w

$$y_{\text{out}} = w^H z, \quad (18.87)$$

where H denotes complex conjugate transposition (i.e., Hermitian transpose) [43]. The clutter and noise contaminated space-time snapshot z has the same array structure as the multi-channel signal vector s given in (18.44). The dimension is $MN \times 1$, where M is the number of receiving antennas and N the number of considered temporal samples.

If multi-channel DPCA using M antennas shall be applied for clutter suppression, the weight vector is given as [43]

$$w = \begin{bmatrix} d_1 \\ -d_2 \end{bmatrix}, \quad (18.88)$$

where \mathbf{d}_1 and \mathbf{d}_2 are the M -dimensional spatial-only steering vectors for the first and second pulses (i.e., $N = 2$), respectively. The steering vectors are then

$$\mathbf{d}_1 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (\text{1st element off}), \quad \mathbf{d}_2 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} \quad (\text{Mth element off}). \quad (18.89)$$

It has to be noted that in this case the DPCA condition has to be fulfilled. The data must not be co-registered. The output of the clutter filter is a scalar if only two temporal samples are used. To filter an entire azimuth line the filter has to be applied successively, each time using two temporal samples.

2.18.6.2.2 Adaptive clutter suppression

In the multi-channel case adaptive clutter filtering can be performed by multiplying the acquired and preprocessed multi-channel signal with the inverse of the clutter covariance matrix and a steering vector [4,37]. The adaptive weight vector is formed in the following way

$$\mathbf{w} = \mathbf{R}_W^{-1} \mathbf{d}, \quad (18.90)$$

where \mathbf{R}_W is the clutter covariance matrix of dimension $MN \times MN$ and \mathbf{d} is the steering vector.

The clutter covariance matrix itself can either be computed analytically using the known system, antenna and geometry parameters or, more powerfully, be estimated from the real data (it is then called “Empirical Clutter Covariance Matrix” or “Sample Clutter Covariance Matrix”). In the latter case the clutter covariance matrix is “adapted” to the real “clutter” data. The processing is then called “Space-Time Adaptive Processing (STAP)” (see also Section 2.18.9). The empirical clutter covariance matrix can be computed by averaging training data which shall not contain moving target signals [4]:

$$\widehat{\mathbf{R}}_W = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H, \quad (18.91)$$

where \mathbf{z}_k is the $MN \times 1$ vector of training data from training cell k , and K is the total number of used training cells. The averaging for instance could be performed over range as exemplarily sketched in Figure 18.26. To exclude potential moving target signals a guard zone should be used [43,44].

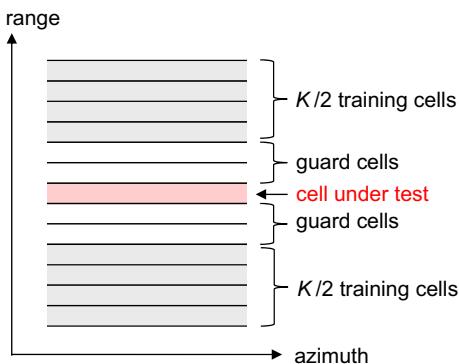
If again two successive temporal samples ($N = 2$) and as steering vector

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} \quad (18.92)$$

are used, then the clutter suppression method is called “Adaptive DPCA” (ADPCA) [44].

2.18.6.3 Detection

Conventionally moving targets are detected on a pixel by pixel basis. A threshold is set to discriminate between one of the two hypotheses H_0 (moving target signal not present) and H_1 (moving target signal present):

**FIGURE 18.26**

Use of training data for computing the empirical clutter covariance matrix.

$$\begin{aligned} H_0 : z &= c + n, \\ H_1 : z &= s + c + n, \end{aligned} \quad (18.93)$$

where z denotes the vector of the received data, c is the clutter vector and n the noise vector.

Generally a so called ‘‘Constant False Alarm Rate (CFAR)’’ detector is envisaged. The threshold is computed in a way that the percentage of the image pixels which lie above the threshold is constant. For computation the clutter statistics have to be known precisely.

If the probability density function (PDF) of the clutter is known, the false alarm rate or the probability of a false alarm is obtained by integrating the PDF:

$$P_{\text{fa}} = 1 - \int_{-\infty}^{s_{\text{th}}} f_C(x) dx = \int_{s_{\text{th}}}^{\infty} f_C(x) dx \quad (18.94)$$

where s_{th} is the threshold and $f_C(x)$ the PDF of the clutter metric to be tested (e.g., the clutter DPCA amplitude, the clutter ATI phase, the clutter ATI phase combined with the clutter ATI amplitude, etc.). If a certain false alarm rate is desired this equation needs to be solved for the threshold s_{th} . Depending on $f_C(x)$ an analytical solution is not always possible so that numerical methods may be necessary.

The probability of detection can be expressed as

$$P_d = \int_{s_{\text{th}}}^{\infty} f_T(x) dx, \quad (18.95)$$

where $f_T(x)$ is the target plus interference PDF.

Analytical descriptions of the PDFs of the clutter multi-look ATI phase and ATI amplitude can be found in [34]. In that paper different clutter types are modeled and verified with real data. A discrimination between homogeneous (Gaussian), heterogeneous (non-Gaussian) and extremely heterogeneous clutter is made. A novel polynomial PDF called p-distribution is introduced. This PDF matches the real data much more accurately, particularly for heterogeneous composite terrain. All clutter parameters

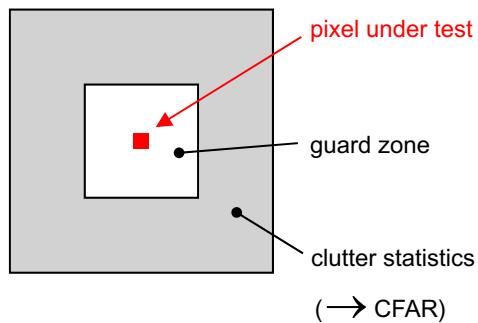


FIGURE 18.27

Principle for estimating the clutter statistics for the pixel under test.

for determining the detection thresholds are estimated from the real data. Further detection metrics are discussed in [45]. Here also a so called “Hyperbolic Detector” well suited for heterogeneous terrain such as urban areas is introduced.

For estimating the clutter statistics from the real data the moving target itself has to be excluded. This can be achieved by introducing a guard zone around the pixel under test as depicted in Figure 18.27. The clutter statistics is then estimated from the data surrounding the guard zone. The purpose of the guard zone is to exclude that moving target signal components disturb the clutter and, hence, lead to a wrong clutter PDF estimate. The size of the guard zone shall be chosen in accordance with the expected sizes of the moving vehicles to be detected. If the guard zone is too small or if a lot of targets move close together, the clutter PDF estimate is biased and the false alarm rate will not remain constant anymore.

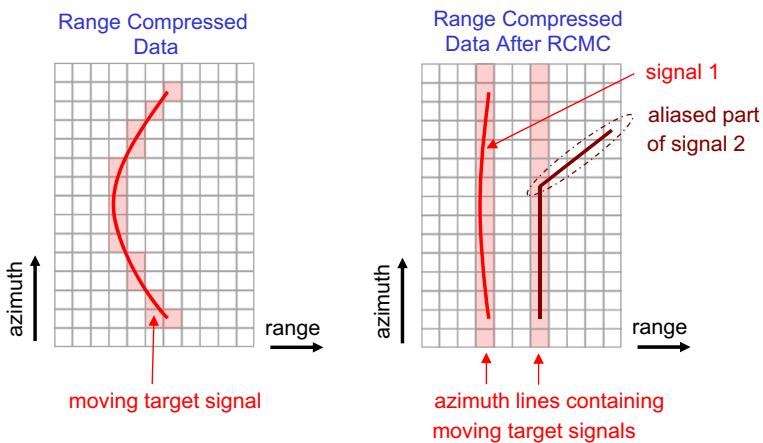
For studying and comparing the performance of different detector types as quality measures the probability of detection P_d versus varying SCR and P_d versus P_{fa} can be used [46]. The latter quality measure is known as receiver operating characteristics (ROC) of a detector. Since the exact PDF of the target signal is generally unknown, it is conventionally assumed to be either deterministic or Gaussian distributed.

2.18.6.4 Signal extraction

Once a moving target has been detected it is of interest to estimate its motion and position parameters. For that purpose the moving target signal may be extracted from the data. An extraction also may be required if for instance the target image shall be refocused to high resolution using ISAR imaging techniques [15–17]. In the following two practicable extraction methods are discussed.

2.18.6.4.1 Range history tracking

The principle is depicted in Figure 18.28 on the left. The method operates on range compressed data where the clutter already has been canceled. Each pixel corresponding to the clutter suppressed moving target signal has to be detected (the detected pixels are represented by the pink squares in the image). The aligned pixels represent the moving target azimuth signal $s(t)$ which afterwards can be passed to

**FIGURE 18.28**

Moving target signals in the range compressed SAR data array (left: range compressed data without RCMC; right: range compressed data after conventional RCMC for stationary targets).

the parameter estimation stage. The range history tracking method was introduced in [47]. A detailed description of the “tracking” algorithm is given in [21].

The method has the advantage that even signals backfolded in Doppler can be tracked and that information about the shape of the range history is obtained (remember in particular the influence of the across-track velocity on the range history shape discussed in Section 2.18.4.1). The disadvantage is that the clutter suppressed signal must have high SCNR values to be detectable on a pixel-by-pixel basis. This restricts the application to airborne GMTI. Tracking might also become problematic if several targets with overlapping range histories are in the data array.

2.18.6.4.2 Extraction of azimuth lines

The principle of the azimuth line extraction method is sketched in Figure 18.28 on the right. The data have to be range compressed and also the range cell migration has to be corrected, at least with respect to stationary targets. Depending on the motion parameters and on the range pixel spacing (which is determined by the range sampling rate) a major part of the signal energy may be distributed along a single azimuth line (see also discussion on residual range cell migration in Section 2.18.4.1). Thus, by extracting the proper azimuth line the signal $s(t)$ is obtained. The method can also be applied prior to the detection step. Then all azimuth lines of the data are successively extracted. Afterwards detection and parameter estimation are carried out using different methods discussed later.

From signals backfolded in Doppler only a part of the signal energy can be recovered. The “aliased part” is lost if the RCMC was adapted for stationary targets (cf. Figure 18.28, right). Successive pre-processing with different motion adapted RCMC may help to catch more signal energy and avoid aliasing [39,48]. Anyhow, the simple extraction of azimuth lines is a very fast method which also works for signals with low SCNR as in the case of spaceborne GMTI. Furthermore, even multi-target scenarios

(i.e., if more than one moving target signal is contained in the extracted azimuth line) can be resolved by using e.g., a matched filter bank or the fractional Fourier transform discussed later in Sections 2.18.7.1 and 2.18.7.2, respectively.

Also from fully focused SAR images the azimuth lines can be extracted. However, before applying a matched filter bank the azimuth compression has to be removed (this can simply be done by adapting the reference function of the matched filter).

2.18.6.5 Parameter estimation

In (18.31) it is shown that the target's motion parameters are related to the Doppler parameters shift f_{DC} , slope k_a , and quadratic coefficient q . Thus, motion parameter estimation can be reduced to Doppler parameter estimation. From the known Doppler parameters the position and motion parameters of the moving target can be computed. The intention of this section is to show how this computation is principally performed by different state-of-the-art GMTI algorithms. In the following it is supposed that the Doppler slope k_a and shift f_{DC} already have been estimated. Basic Doppler parameter estimation methods are discussed afterwards in Section 2.18.7.

It is also supposed that the radar parameters and the acquisition geometry are known accurately so that the parameters y_0 , r_{10} , and λ need not to be estimated. This is a valid assumption, particularly if the range is large compared to a potential range displacement $\Delta r_{img,0}$ (i.e., $r_{10} \gg \Delta r_{img,0}$). Additionally a non-squinted acquisition geometry is assumed (i.e., $x_0 = 0$ and $f_{DC,st} = 0$).

2.18.6.5.1 Across-track velocity

The relation between the Doppler shift and the across-track velocity is given in (18.31). In case of a non-squinted acquisition geometry the simple relationship

$$f_{DC} = -\frac{2y_0}{\lambda r_{10}} v_{y0} \quad (18.96)$$

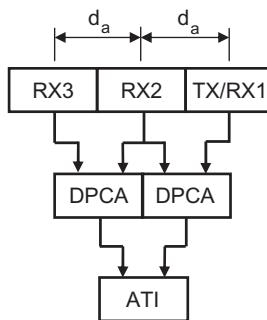
applies. However, in practice the across-velocity is computed from the estimated ATI phase $\Delta\phi$. In the simplest case the relationship given in (18.71) is used so that the across-track velocity computes to

$$v_{y0} \cong \frac{\lambda r_{10}}{4\pi y_0 \Delta t} \Delta\phi. \quad (18.97)$$

To mitigate the influence of clutter and to keep the estimation bias at a low level the moving target's IRF should be refocused before estimating the ATI phase (refocusing increases the SCNR). More robust algorithms exploit three RX channels for ATI phase estimation as sketched in Figure 18.29. They generate two DPCA images: the first by subtracting channel 2 from 1, and the second by subtracting channel 3 from 2. The corresponding clutter suppressed ATI signal is then obtained by

$$s_{ATI}(t) = s_{DPCA,12}(t) \cdot s_{DPCA,23,reg}^*(t). \quad (18.98)$$

This has the advantage that the clutter is suppressed theoretically down to noise level so that in the best case the performance limits depicted in Figure 18.23 can be reached.

**FIGURE 18.29**

Principle of the DPCA-ATI technique.

One promising state-of-the-art GMTI algorithm which additionally uses the fractional Fourier transform (explained in detail in Section 2.18.7.2) for target refocusing and additional parameter estimation is presented in [35]. An extension to spaceborne geometries can be found in [18]. Here also different transmit and receiver aperture switching strategies for generating additional receiving channels with larger baselines are discussed. Especially for spaceborne systems larger baselines are of importance for improving the parameter estimation accuracy.

2.18.6.5.2 Along-track velocity

Remember (18.31) where it is shown that the Doppler slope k_a depends on the along-track velocity v_{x0} , the across-track acceleration a_y and on the across-track velocity v_{y0} in the following way:

$$k_a = -\frac{2}{\lambda r_{10}} \left\{ (v_{x0} - v_p)^2 + v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2} \right) + y_0 a_y \right\}. \quad (18.99)$$

Under the assumption that the across-track velocity v_{y0} can be estimated with ATI still two unknowns remain: v_{x0} and a_y . Without further information the two unknowns cannot be computed. This is just the crux of many GMTI algorithms which exploit only the Doppler slope. Since the required additional information cannot be gained accelerations are assumed to be negligibly small. They are simply set to zero. The along-track velocity is derived by rearranging (18.99)

$$v_{x0} = v_p - \sqrt{-\frac{\lambda r_{10}}{2} k_a - v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2} \right) - y_0 a_y}, \quad (18.100)$$

where incorrectly the term $y_0 a_y$ is ignored by many GMTI algorithms. However, this assumption is in reality only valid if an across-track motion can be excluded a priori. Thus, the target must move only in along-track direction for avoiding estimation errors.

In [49] it is shown that the typical standard deviation of the acceleration of a common passenger car is in the order of 0.5 m/s^2 . For TerraSAR-X a comparatively small across-track acceleration of 0.25 m/s^2

would lead, if neglected in (18.100), to an along-track velocity estimation error of -32 km/h . This demonstrates clearly that accelerations should not be neglected if highly accurate velocity parameter estimates are desired.

In the literature so far only two algorithms can be found which take into account the accelerations for parameter estimation. Both of them don't use the estimated Doppler slope for along-track velocity estimation. They use a different measure not influenced by across-track accelerations.

The algorithm in [50] was developed for a single-channel SAR but can principally also be used in the multi-channel case. It exploits the fact that the effective synthetic aperture time and, hence, the Doppler bandwidth varies with the moving target's along-track velocity. Thus, by estimating the Doppler bandwidth of the extracted moving target signal the along-track velocity can be computed without the negative influence of accelerations in the following way:

$$v_{x0} = v_p - \frac{L_{\text{SA}}|k_a|}{\Delta f_a}, \quad (18.101)$$

where L_{SA} is the length of the synthetic aperture for a stationary target located at the same range, and Δf_a is the estimated Doppler bandwidth of the moving target signal. The drawback of the algorithm is that it only works properly for strong point-like targets with a RCS that is independent of the aspect angle. Furthermore, the whole target signal energy has to be collected since otherwise the Doppler bandwidth cannot be estimated accurately.

In [24] the ATI phase history of range compressed moving target signals between two or more receiving channels which are not co-registered is exploited. The slope of the ATI phase history is given as

$$k_\phi = \frac{2\pi(v_{x0} - v_p)}{\lambda r_{10}} d_a. \quad (18.102)$$

It is not influenced by accelerations. Thus, by estimating the slope the along-track velocity v_{x0} can be computed as

$$v_{x0} = v_p + \frac{\lambda r_{10}}{2\pi d_a} k_\phi. \quad (18.103)$$

The method is suitable for airborne systems where the SNR is generally much higher than in satellite systems. A good performance can be achieved for signals with SCNR values larger than 10 dB.

Both methods discussed before require large SCNR values to achieve a good along-track velocity estimation accuracy. Therefore they are not well suited for spaceborne systems. They should always be used supplementary to methods exploiting the Doppler slope.

2.18.6.5.3 Accelerations

The estimation of target accelerations is generally of minor interest. The reason why accelerations are still of importance is that they may improve the velocity estimation accuracy.

The across-track acceleration can be computed by rearranging (18.99):

$$a_y = -\frac{1}{y_0} \left[\frac{\lambda r_{10}}{2} k_a + (v_{x0} - v_p)^2 + v_{y0}^2 \left(1 - \frac{y_0^2}{r_{10}^2} \right) \right]. \quad (18.104)$$

The along-track acceleration principally can be computed by rearranging the equation of the quadratic Doppler coefficient which is given as:

$$q = -\frac{3}{\lambda r_{10}} \left\{ a_x(v_{x0} - v_p) + v_{y0} a_y \left(1 - \frac{y_0^2}{r_{10}^2} \right) \right\}. \quad (18.105)$$

However, for estimating q long observation times are necessary. At short observation times q is negligibly small and cannot be estimated precisely (cf. also Section 2.18.4.4 where the effect caused by the along-track acceleration is discussed).

2.18.6.5.4 True target position

As discussed in Section 2.18.4.3 a target moving in across-track direction appears displaced from its actual position in the SAR image. It is displaced in azimuth as well as in range direction.

It is shown in (18.57) that the along-track displacement is directly related to the across-track velocity v_{y0} in the following way:

$$\Delta x_{\text{img},0} \cong -y_0 \frac{v_{y0}}{v_p}. \quad (18.106)$$

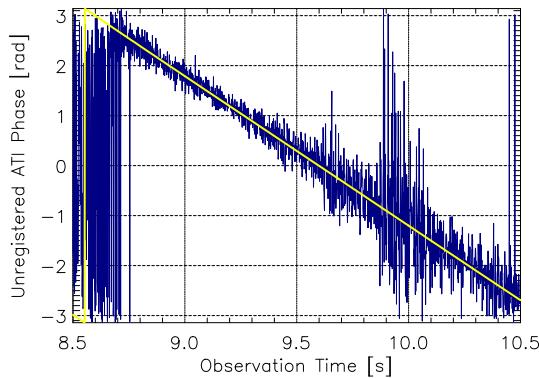
Thus, once the across-track velocity is estimated (e.g., with ATI) the azimuth displacement can be computed. The same is valid for the range displacement given in (18.60).

Another interesting dual-channel method for estimating the broadside position, which corresponds to the true target position in the non-squinted case, is presented in [51]. Here again the ATI phase history of range compressed signals which are not co-registered is exploited. The distances between the target and the fore and the aft antenna phase centers are exactly equal if the target is located broadside. Thus, by estimating the zero crossing of the ATI phase history the broadside position and, hence, the true target position can be computed. Since the ATI phase history is strongly influence by clutter bandpass filtering in the fractional Fourier domain is performed. The effectiveness of this filtering is illustrated in Figure 18.30 (the zero crossing is at $t = 9.6$ s). The method can also be used with more than two channels by additionally using DPCA for clutter suppression [24]. Also the ATI phase slope for computing the along-track velocity can be estimated as discussed in Section 2.18.6.5.2.

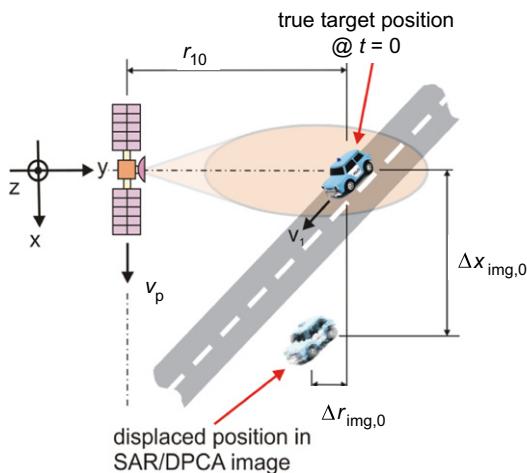
A third method for estimating the true target positions are displacement based algorithms. These algorithms require the incorporation of a priori knowledge about the road network. The along-track displacement $\Delta x_{\text{img},0}$ can then simply be measured: it is the distance between the position where the target is imaged and the road axis (cf. Figure 18.31). For a spaceborne system the range displacement $\Delta r_{\text{img},0}$ is only in the order of a few meters and therefore can be neglected (cf. Section 2.18.4.3). The along-track displacement estimation accuracy is mainly related to the position accuracy of the road axis and the road width. From the estimated displacement easily the across-track velocity can be derived:

$$v_{y0} \cong -\frac{v_p}{y_0} \Delta x_{\text{img},0}. \quad (18.107)$$

Since also the road angle α is available from the road database the along-track velocity can be computed as $v_{x0} = v_{y0} / \tan \alpha$. The across-track acceleration is obtained with (18.104) after the Doppler slope k_a has been estimated. For computing the along-track acceleration again the road angle can be used so that $a_x = a_y / \tan \alpha$. Although all important motion parameters can be derived the displacement based

**FIGURE 18.30**

Unregistered ATI phase history (blue: without clutter suppression; yellow: clutter suppressed using bandpass filtering in fractional Fourier domain) [24] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

**FIGURE 18.31**

Azimuth and range displacement of a moving target.

algorithms have one major disadvantage: problems may arise if a lot of adjacent roads are located in the scene. Then the detected targets under circumstances cannot be assigned to the correct roads.

An operational traffic monitoring algorithm for TerraSAR-X based on the displacement technique is presented in [11, 52].

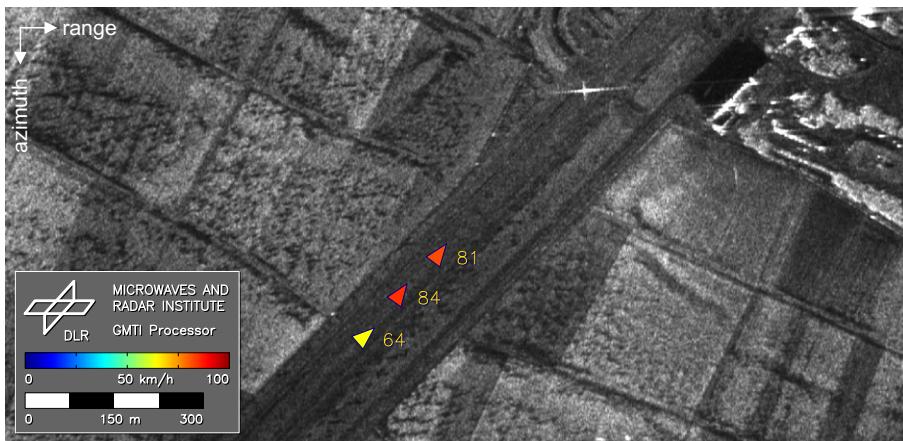


FIGURE 18.32

Single-channel SAR image with three detected moving targets (location: airfield in Memmingen, Germany; data acquired with DLR's E-SAR sensor) [53].

2.18.6.6 Visualization

Finally, the detected target can be visualized. Often a symbol representing the moving target is overlaid on the corresponding SAR image on its estimated position. For instance, the symbol may be an arrow pointing into the estimated moving direction and the color of this symbol may correspond to the velocity. An example for such a visualization is shown in Figure 18.32.

Another elegant possibility for visualization is Google Earth. The parameters of the detected moving targets are written into a Keyhole Markup Language (KML) file. Also links to target symbols and SAR images can be incorporated easily. The KML file can then be visualized with Google Earth as shown in Figure 18.33.

Furthermore, additional information about the detected target can be provided in an interactive way. The user or operator can click on the moving target symbols to open a window. This window may contain more information about the detected target. Also data from additional data sources can be incorporated and fused with the SAR data. For instance, the estimated parameters of detected ships could be fused with “ground truth” data obtained from the automatic identification system (AIS) [54]. An example is depicted in Figure 18.34. Such ground truth data can also be used for evaluating the parameter estimation accuracy of GMTI algorithms.

2.18.7 Doppler parameter estimation: basic methods

The intention of this section is to introduce the fundamental and most promising Doppler parameter estimation methods which are used in state-of-the-art GMTI algorithms. Most of these methods can even be used for both target detection and parameter estimation in one step.



FIGURE 18.33

SAR image and detected moving targets as Google Earth overlay (location: near Chiemsee, Germany; scene size $6.0 \times 1.9 \text{ km}^2$; data acquired with DLR's F-SAR sensor). Some of the slower moving targets are false detections.

For estimating the Doppler shift f_{DC} which is mainly related to the across-track velocity the ATI technique discussed in Sections 2.18.5.1 and 2.18.6.5.1 can be used. The Doppler slope k_a for instance can be estimated by using a matched filter bank, the fractional Fourier transform or the Wigner-Ville distribution in combination with the Radon transform. Due to their importance they are discussed in detail.

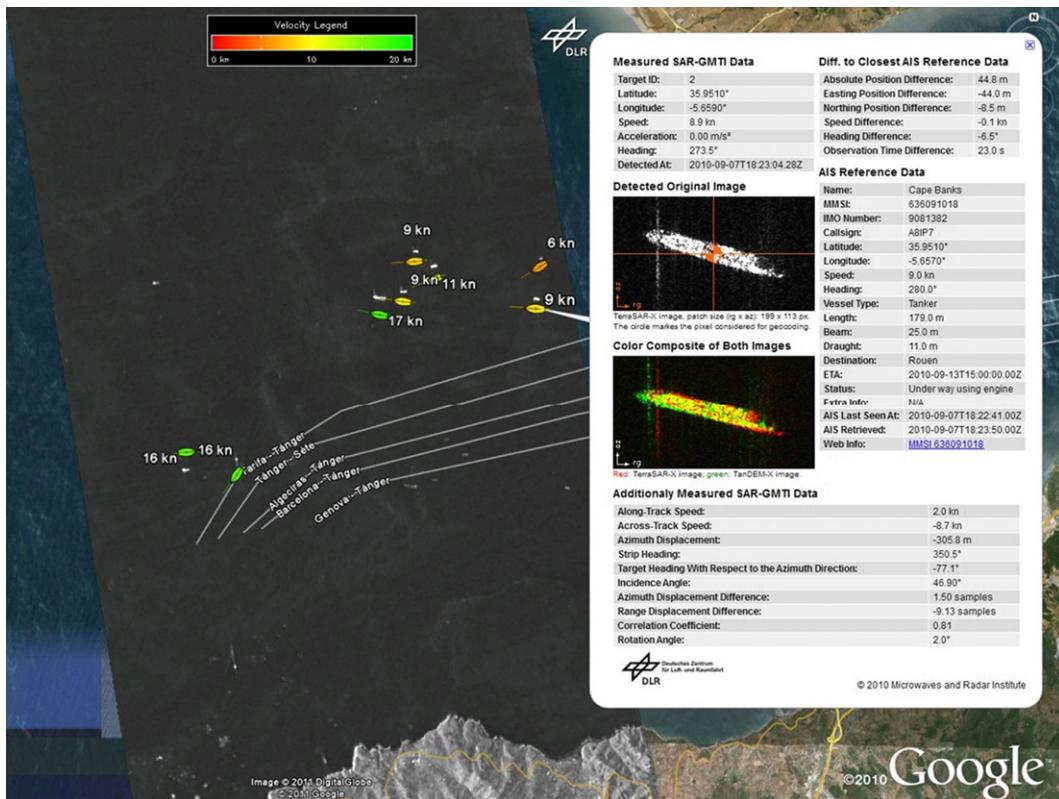
In the following explanations for simplicity the single-channel moving point target signal model is used. However, the application of the presented detection and parameter estimation methods is not limited to the single-channel case. The methods also can be applied on clutter suppressed multi-channel signals almost without restrictions.

2.18.7.1 Matched filter bank

The basic principle of a matched filter was discussed in Section 2.18.2.3. It was shown that a signal can be focused by performing a convolution with a proper reference function. Also a moving target signal can perfectly be focused if the proper reference function is applied [21]. The proper azimuth reference function for focusing the moving target signal is given as

$$h_a(t) \cong \exp \left[-j \left(-2\pi f_{DC} t + \pi k_a t^2 - \frac{2}{3}\pi q t^3 \right) \right]. \quad (18.108)$$

However, the Doppler parameters f_{DC} , k_a , and q are unknown in advance. They need to be estimated. This can be done by convolving the range compressed moving target signal successively with different azimuth reference functions. For each of these functions different Doppler parameter are assumed. The estimated Doppler parameters \hat{f}_{DC} , \hat{k}_a , and \hat{q} are obtained from the reference function that after convolution results in the highest peak value of the IRF. The successive convolution with different

**FIGURE 18.34**

Google Earth image of the Strait of Gibraltar overlaid with the KML file obtained from the TerraSAR-X/TanDEM-X large along-track baseline GMTI processor [55]. The color coded symbols (color is velocity dependent) mark the estimated “true” geographical positions of the automatically detected ships, also the displaced ship images in white color are visible.

reference functions is called “Matched Filter Bank” and can be considered as a maximum likelihood estimator for the Doppler or motion parameters [56].

Principally the moving target signal has to be extracted from the range compressed data array before the matched filter bank can be applied [47]. For this task the range history tracking method presented in Section 2.18.6.4.1 can be used for example. However, as a brute force method also two-dimensional matched filters can be constructed taking additionally into account the range history of the target signal. Each (pre-detected) pixel of the range compressed data array has then to be convolved with a set of two-dimensional matched filters. One can conceive that this method has a high computational load and is rather time consuming. However, the method may be of importance for spaceborne systems which suffer from low SNR values. The advantages are the higher compression gain and the possibility to estimate

the parameters from fast moving targets aliased in Doppler [57]. The minimum PRF requirement given in (18.62) is not applicable in this case. A lower PRF resulting in less range ambiguities can be used. Additionally, the swath width (cf. Figure 18.3) of the SAR system can be increased [1] by lowering the PRF.

The moving target signal can also be extracted “by chance” using the method discussed in Section 2.18.6.4.2. With this method successive azimuth lines are extracted. The disadvantage is that no direct information about the range history is preserved. In this case the matched filter bank is not very sensitive to the Doppler shift f_{DC} . However, the Doppler shift or equivalently the across-track velocity can be estimated using ATI. If additionally the quadratic coefficient q is neglected (it only has a significant contribution at large observation times) a much simpler azimuth reference function can be used:

$$h_a(t) = \exp[-j\pi k_a t^2]. \quad (18.109)$$

The equivalent in Doppler domain is given as

$$H_a(f_a) \cong \exp\left[j\pi \frac{1}{k_a} f_a^2\right]. \quad (18.110)$$

It has to be kept in mind that with this reference function only the Doppler slope \hat{k}_a of the moving target signal can be estimated. After focusing with this reference function the target still appears displaced from its actual position.

An example of the application of a matched filter bank on real dual-channel airborne data is shown in Figure 18.35. On the left side a SAR image containing four targets moving on a runway with different along-track velocities is shown. Target 2 moved with 10 km/h. The azimuth line containing “Target 2” was extracted from the clutter suppressed DPCA image and the azimuth compression was removed. Afterwards the matched filter bank was applied. The matched filter bank output, denoted as “Matched Filter Map,” is shown on the right, once in its two-dimensional representation and once in its 3D representation. The target is represented by the focused peak. The peak position corresponds to a certain estimated Doppler slope and to a certain azimuth position. In the example the target moved only in along-track direction ($v_{y0} = 0$ and $a_y = 0$) so that with (18.100) the along-track velocity v_{x0} can be computed easily (the negative sign in the image is due to the antiparallel motion with respect to the flight path).

As already explained, the application of the matched filter bank is not limited to azimuth lines containing only one moving target signal. In fact a multi-target scenario can be resolved as shown in the example depicted in Figure 18.36. Here an azimuth line of interest containing a part of the Autobahn A8 near Chiemsee, Germany, where several targets have moved was extracted. In the SAR image on the left the targets are severely blurred in azimuth and cannot be recognized and separated. However, in the matched filter map of the DPCA signal each target appears as a focused peak at a certain azimuth and Doppler slope position. Thus, a matched filter bank not only can be used for parameter estimation, but also for target detection and target separation. It can deal with multi-component LFM signals. Target detection in the matched filer map either can be performed by applying a certain amplitude threshold or by comparing the sharpness of the peaks with analytical sharpness functions [27].

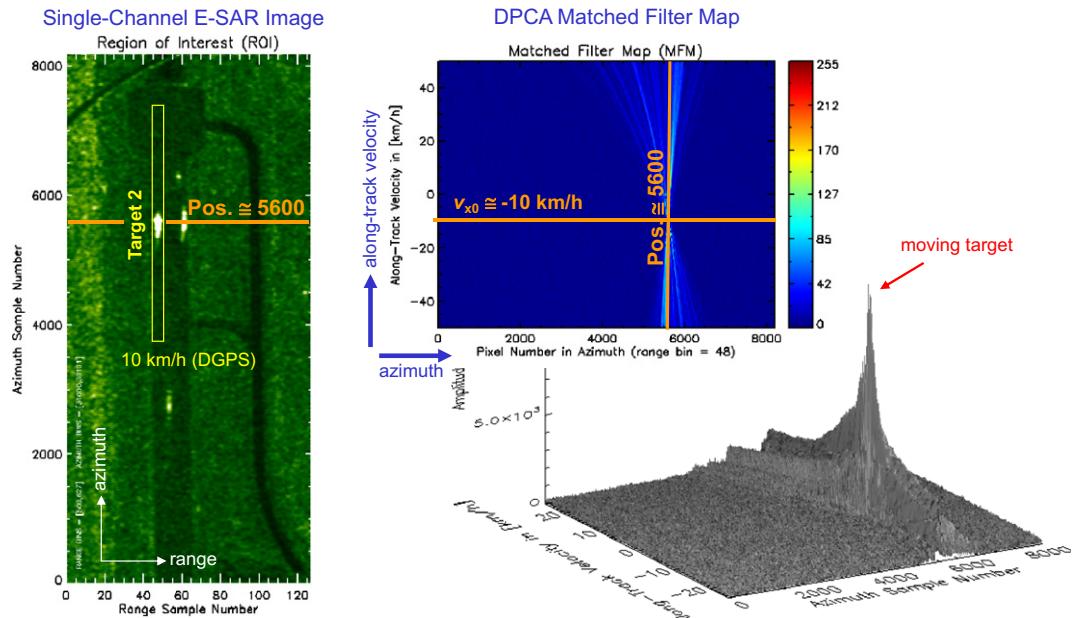


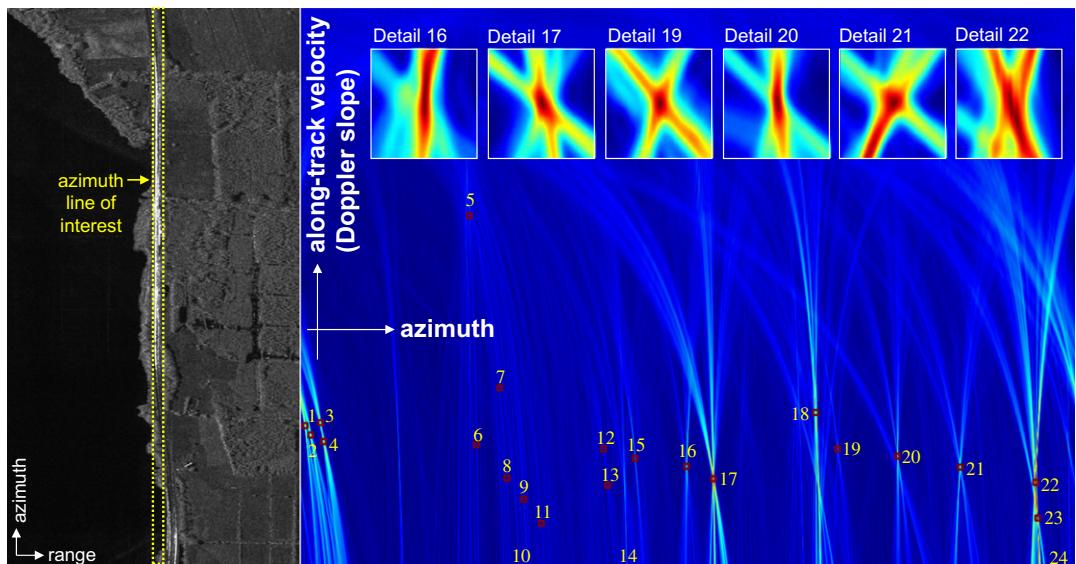
FIGURE 18.35

Application of a matched filter bank on an extracted azimuth line (left: single-channel SAR image acquired with DLR's E-SAR sensor [58]; top right: 2D representation of the matched filter bank output; bottom: 3D representation of the output).

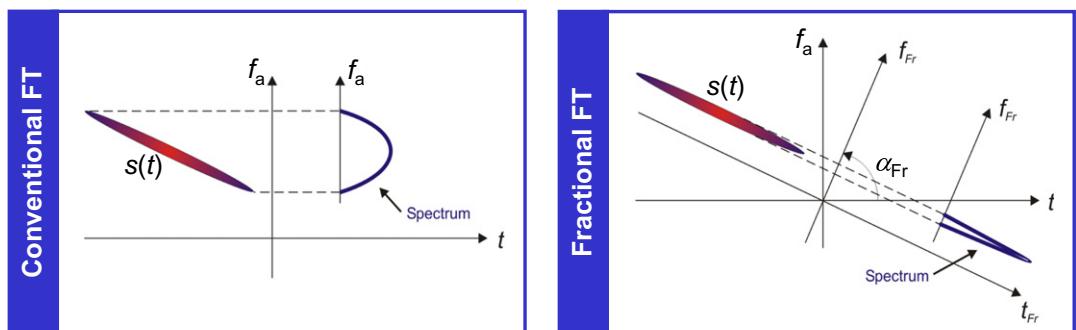
2.18.7.2 Fractional Fourier transform

The fractional Fourier transform (FrFT) is a linear operator (no cross-terms in case of multi-component LFM signals) and can be considered as a generalization (or rotation with rotation angle α_{Fr}) of the conventional Fourier transform [59]. The FrFT has found many applications like swept-frequency filters, time-variant filtering and multiplexing, pattern recognition and the study of time-frequency distributions. It is also known, that the Radon transform [60] of the Wigner spectrum is equal to the magnitude square of the fractional Fourier transform. Filtering in the fractional Fourier domain, rather than in the ordinary Fourier domain, allows one to decrease the mean square error in the estimate of a distorted and noisy signal [61].

The principle of the FrFT is sketched in Figure 18.37. After application of the FrFT the signal energy in the fractional Fourier domain is integrated along the fractional time axis t_{Fr} . For LFM signals the application of the FrFT with the optimum rotation angle $\alpha_{\text{Fr, opt}}$ (in this case the fractional time axis t_{Fr} is parallel to time-frequency history of the LFM signal) results in a sharp spectrum with maximum peak amplitude. Due to this behavior, which for LFM signals is equivalent to pulse compression and matched filtering, the FrFT is also attractive for SAR-GMTI [62].

**FIGURE 18.36**

Single-channel SAR image acquired with E-SAR with marked azimuth line of interest (left) and corresponding DPCA matched filter bank output (right).

**FIGURE 18.37**

Comparison between conventional Fourier transform (left) and Fractional Fourier transform (right) for a LFM signal $s(t)$.

The application of the FrFT for estimating the Doppler slope k_a is quite similar to the matched filter bank approach. Instead of applying different matched filters for focusing now the FrFT is applied successively on the extracted range compressed data whereby each time a different rotation angle α_{Fr} is used. The rotation angle corresponds to the Doppler slope in the following way [35]:

$$\alpha_{Fr} = \text{acot} \left(k_a \frac{N}{\text{PRF}^2} \right) = \tan^{-1} \left(-\frac{\text{PRF}^2}{k_a N} \right), \quad (18.111)$$

where the ratio between the number of samples N and the PRF is used for normalization purposes. The estimate of the Doppler slope \hat{k}_a is then computed with this equation using the optimum rotation angle $\alpha_{Fr,\text{opt}}$ which maximizes the peak amplitude of the spectrum.

As with the matched filter bank also the FrFT can cope with multi-component LFM signals as shown in Figure 18.38. On the top the “Fractional Spectra Map” is depicted. This map is obtained by successive

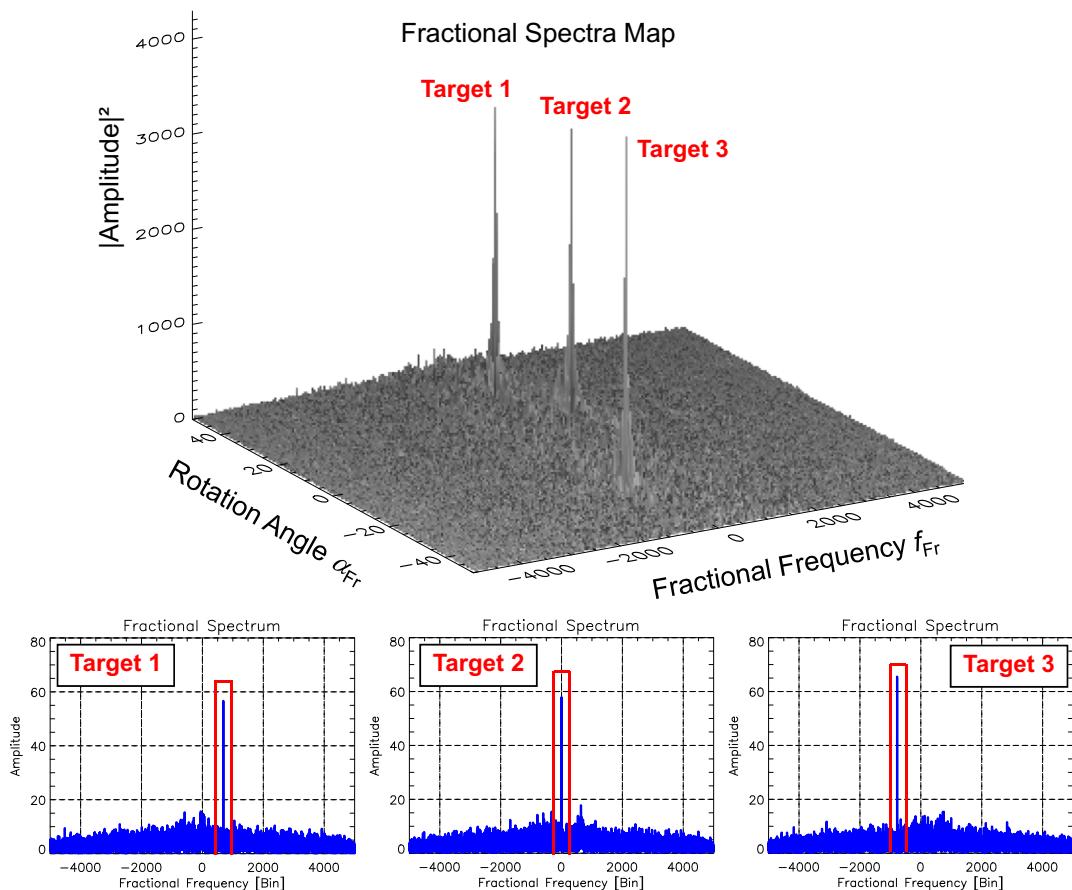


FIGURE 18.38

Fractional spectra map (top) containing three simulated moving target signals with different Doppler slopes embedded in white Gaussian noise. At the bottom the cuts along the optimum rotation angle $\alpha_{Fr,\text{opt}}$ for each of the signals are shown.

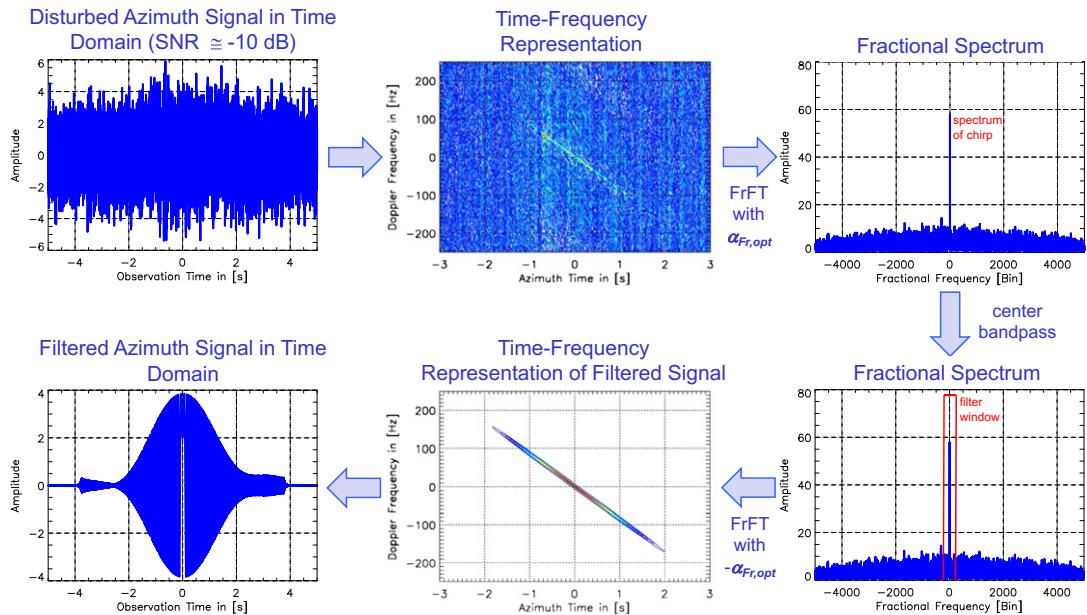


FIGURE 18.39

Bandpass filtering in the fractional Fourier domain.

application of the fractional Fourier transform with different rotation angles. At the bottom for each of the three moving target signals the spectrum obtained by using the optimum rotation angle is shown.

In contrast to the matched filter bank the application of the FrFT has higher computational cost but also some advantages. For instance, bandpass filtering can be applied directly in the fractional Fourier domain as depicted in Figure 18.39. By this orthogonal interferences can be filtered out and multi-component LFM signals can be separated from each other and extracted from the data. After applying the inverse FrFT on the bandpass filtered signal using $-\alpha_{Fr,opt}$ as rotation angle, the clutter and noise suppressed moving target signal is obtained in the time domain.

2.18.7.3 Time-frequency analysis

One application of time-frequency (TF) analysis in the field of GMTI is the estimation of the instantaneous Doppler frequency f_a of the moving target signal. From the estimated instantaneous Doppler frequency (cf. (18.18)) the phase history φ_a of the range compressed moving target signal can be computed by integration [63]:

$$\varphi_a(t) = 2\pi \int_{t'=t_0}^t f_a(t') dt' + \varphi(t_0), \quad (18.112)$$

where $\varphi(t_0)$ is an unimportant constant phase term. With the estimated phase a proper reference function $h_a(t)$ can be constructed. With this reference function even targets in non-linear motion can be properly

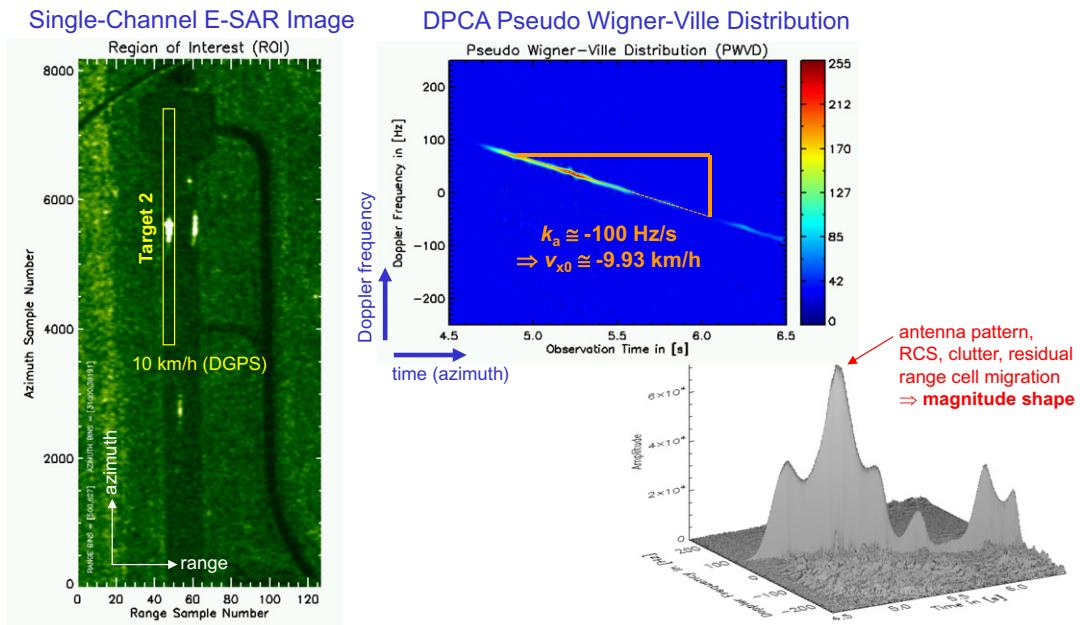


FIGURE 18.40

Application of the Pseudo Wigner-Ville Distribution (PWVD) on an extracted azimuth line (left: single-channel SAR image; top right: time–frequency map as output from the PWVD; bottom right: 3D representation of the output).

focused. Thus, the application of the reference function can already be considered as a simplified ISAR imaging method.

Before the phase history can be computed with (18.112) it is necessary to estimate the Doppler frequency f_a . For this purpose a suitable TF transform has to be applied on the range compressed moving target signal. Probably the best known linear TF transform is the short-time Fourier transform. As disadvantage a trade-off between time and frequency resolution has to be made: either a small time or a small frequency resolution can be obtained. It is impossible to get both simultaneously. In contrast the Wigner-Ville distribution (WVD) is the TF distribution having the best time-frequency resolution [64, 65]. However, since it is a quadratic TF transform cross-terms and interferences may occur if the WVD is applied on multi-component LFM signals. To avoid these issues or to keep at least the negative influences at a low level, clutter suppression and signal separation has to be performed prior to its use, e.g., by applying DPCA and range history tracking. In [17] it is shown how the Pseudo Wigner-Ville Distribution (PWVD) can be used for Doppler parameter estimation and ISAR imaging.

In Figure 18.40 an example is shown where the PWVD is used for estimating the Doppler slope k_a of the moving target signal. The PWVD was applied on the same data set used in Section 2.18.7.1 for the explanation of the matched filter bank principle.

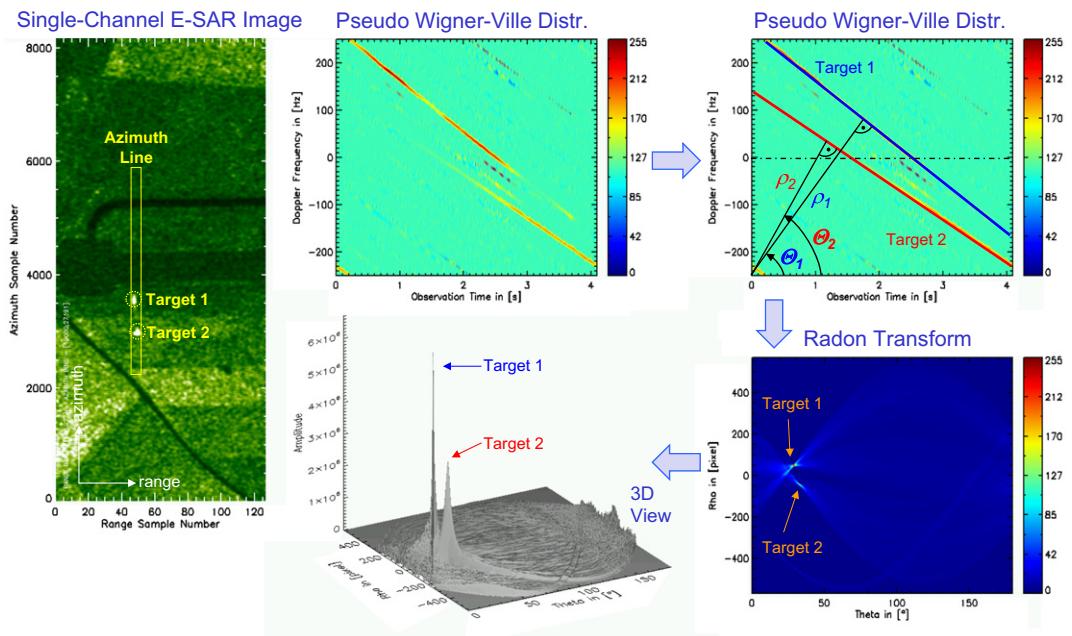


FIGURE 18.41

Application of the Radon transform on the Pseudo Wigner-Ville Distribution of moving target signals (left: single-channel SAR image with azimuth line containing two moving targets; top middle and right: PWVD of the azimuth line; bottom right and middle: result of the Radon transform).

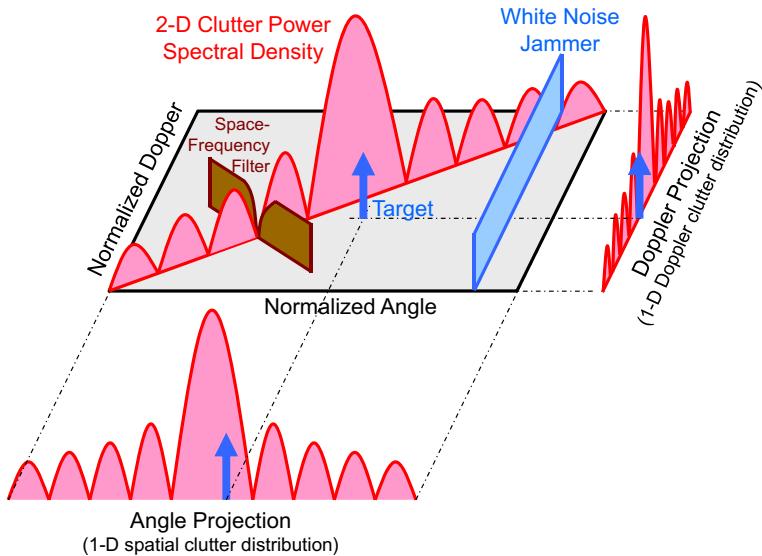
2.18.7.4 Radon transform

Also the Radon transform [60] applied on the TF representation (i.e., the Wigner-Ville spectrum) of the moving target signal can be used for target detection and Doppler slope estimation [66]. An example is shown in Figure 18.41. However, nowadays algorithms relying on the Radon transform of the Wigner-Ville spectrum are replaced increasingly by algorithms using the fractional Fourier transform.

2.18.8 Space-time adaptive processing

The intention of this section is to provide a short introduction to space-time adaptive processing (STAP). A comprehensive tutorial is given in the STAP chapter of the ELSEVIER e-reference. Further reading can be found in [4,37,43,67].

STAP is a two-dimensional filtering technique. It operates in the angle-Doppler plane sketched exemplarily in Figure 18.42. The clutter suppression is performed by beamformers that take into account the space-time nature of the received signals. With the beamforms nulls in the angle-Doppler beam-pattern can be placed. Doppler histories and angular directions of moving target signals are no longer

**FIGURE 18.42**

Angle-Doppler plane.

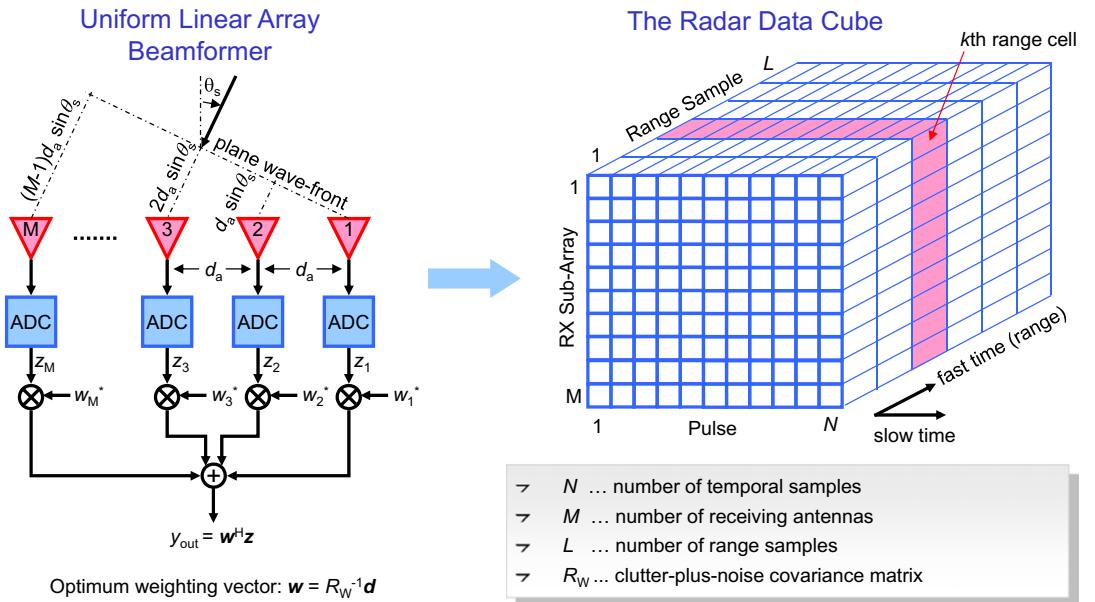
rigidly coupled. The optimum beamformer or weight vector, respectively, is given in (18.90) in Section 2.18.6.2.2. The difference to adaptive clutter suppression treated in Section 2.18.6.2.2 is that the steering vector \mathbf{d} can be adapted to certain signal properties for parameter estimation purposes. However, most of the state-of-the-art STAP algorithms only estimate the line-of-sight velocity v_{r0} (related to the Doppler shift f_{DC}) and the direction-of-arrival angle ψ_{DOA} , from which the true target position can be computed. The target's along-track velocity v_{x0} is not estimated.

2.18.8.1 Joint domain STAP

Originally STAP was designed to operate in the time domain (note that the data must not be co-registered). STAP processing operates on the radar data cube depicted on the right in Figure 18.43. The data cube is convenient for visualizing subsequent space-time processing although the radar processor does not store the data in the format shown in the figure [68]. Each data cube corresponds to a single coherent processing interval (CPI).

The page of the data cube corresponding to the k th range cell is

$$\mathbf{Z} = \begin{pmatrix} [\mathbf{Z}]_{11} & [\mathbf{Z}]_{12} & \dots & [\mathbf{Z}]_{1N} \\ [\mathbf{Z}]_{21} & [\mathbf{Z}]_{22} & \dots & [\mathbf{Z}]_{2N} \\ \vdots & \vdots & \dots & \vdots \\ [\mathbf{Z}]_{M1} & [\mathbf{Z}]_{M2} & \dots & [\mathbf{Z}]_{MN} \end{pmatrix}, \quad \mathbf{Z} \in \mathbb{C}^{M \times N}, \quad (18.113)$$

**FIGURE 18.43**

Uniform linear array beamformer (left) and radar data cube (right).

where \$M\$ is the number of RX channels and \$N\$ is the number of temporal samples per RX channel. The matrix in previous equation can be vectorized by stacking each succeeding column one beneath the other. This yields the space-time snapshot for the \$k\$th range cell, i.e.,

$$\mathbf{z} = [z_{11}, \dots, z_{M1}, \dots, z_{1N}, \dots, z_{MN}]^T = [z_1^T, \dots, z_N^T]^T, \quad \mathbf{z} \in \mathbb{C}^{MN \times 1}, \quad (18.114)$$

where the elements \$[\mathbf{Z}]_{mn}\$ are abbreviated as \$z_{mn}\$. The space-time snapshot in general is composed of

$$\mathbf{z} = \mathbf{s} + \mathbf{c} + \mathbf{n}, \quad (18.115)$$

where \$\mathbf{s}\$ denotes the moving target signal, \$\mathbf{c}\$ the clutter and \$\mathbf{n}\$ the uncorrelated component due to thermal receiver noise or sky noise. The multi-channel signal for STAP often is modeled as [8]

$$s(u(t)) = ae^{-j\frac{4\pi}{\lambda}r(t)} D_t(u(t)) \begin{bmatrix} D_{r,1}(u(t))e^{j\frac{2\pi}{\lambda}u(t)x_1} \\ D_{r,2}(u(t))e^{j\frac{2\pi}{\lambda}u(t)x_2} \\ \vdots \\ D_{r,M}(u(t))e^{j\frac{2\pi}{\lambda}u(t)x_M} \end{bmatrix} = ae^{-j\frac{4\pi}{\lambda}r(t)} D_t(u) \mathbf{d}(u), \quad (18.116)$$

where \$a\$ is a complex amplitude describing the reflectivity of the scatterer, \$r(t)\$ denotes the range to the antenna array center, \$D_t(u)\$ and \$D_{r,m}(u)\$ are the transmit and receive antenna characteristics of the

m th channel, x_m denotes the antenna phase center positions in azimuth direction with respect to the array origin, and $u(t)$ is the directional cosine. It is assumed that the antenna array center origin is freely chosen at the center of the array, so that $\sum_m x_m = 0$. The common phase multiplier $e^{-j\frac{4\pi}{\lambda}r(t)}$ represents the conventional azimuth chirp used for classical SAR imaging via azimuth compression or matched filtering.

The space-time processor linearly combines the elements of the space-time snapshot by applying a weight vector. As result at the output of the space-time processor a scalar is obtained

$$y_{\text{out}} = \mathbf{w}^H \mathbf{z}, \quad (18.117)$$

where H denotes conjugate transposition and \mathbf{w} is the complex weight vector of dimension $MN \times 1$. The optimal weight vector maximizes the output SCNR and takes, under the assumption of homogeneous Gaussian clutter, the form

$$\mathbf{w} = \mathbf{R}_W^{-1} \mathbf{d}, \quad (18.118)$$

where \mathbf{R}_W is the interference covariance matrix of dimension $MN \times MN$ and \mathbf{d} is the steering vector of dimension $MN \times 1$.

In practice both \mathbf{R}_W and \mathbf{d} are unknown and need to be estimated. Thus, instead of (18.118) an estimate of the weight vector in the form of

$$\hat{\mathbf{w}} = \hat{\mathbf{R}}_W^{-1} \mathbf{v}, \quad \hat{\mathbf{w}} \in \mathbb{C}^{MN \times 1}, \quad (18.119)$$

is applied, where $\hat{\mathbf{R}}_W$ is an estimate of \mathbf{R}_W and \mathbf{v} is a surrogate for \mathbf{d} . This approach is known as sample matrix inversion (SMI). It is common to compute the covariance matrix estimate as

$$\hat{\mathbf{R}}_W = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H, \quad \hat{\mathbf{R}}_W \in \mathbb{C}^{MN \times MN}, \quad (18.120)$$

where \mathbf{z}_k of dimension $MN \times 1$ are known as secondary training data. Conventionally K data ranges are used for training (i.e., the averaging is performed along range). To avoid target self-whitening the cell under test as well as cells where already a target has been detected should be excluded from the data.

2.18.8.2 Post-doppler STAP

Since the classical STAP is computationally inefficient and additionally requires a large number of training cells, STAP algorithms needing lower computational power and less training cells have been developed.

With a linear transformation the space-time snapshot \mathbf{z} can be projected into a lower dimensional subspace (= Reduced-Dimension STAP) [68]:

$$\tilde{\mathbf{z}} = \mathbf{T}^H \mathbf{z}, \quad \mathbf{T} \in \mathbb{C}^{MN \times J}, \quad (18.121)$$

where \mathbf{T} is the transformation matrix, $\tilde{\mathbf{z}}$ has dimension $J \times 1$ with $J \ll MN$. The transformation matrix \mathbf{T} is independent of the data. In contrast to the joint-domain STAP the computational burden for matrix

inversion drops from $O(N^3M^3)$ to $O(J^3)$. The transformation of the steering vector and the optimal weight vector are given as $\tilde{\mathbf{d}} = \mathbf{T}^H \mathbf{d}$ and $\tilde{\mathbf{w}} = \tilde{\mathbf{R}}_W^{-1} \tilde{\mathbf{d}}$, respectively. The estimate of the covariance matrix again can be computed in the same manner as before, but now using the transformed data $\tilde{\mathbf{z}}$ instead of \mathbf{z} . The adaptive weight vector using the estimated covariance matrix and a surrogate of the steering vector is given as $\widehat{\mathbf{w}} = \widehat{\mathbf{R}}_W^{-1} \tilde{\mathbf{v}}$.

A practically implementable and efficient Reduced-Dimension STAP algorithm is Post-Doppler STAP [4]. The space-time snapshot given in (18.114) is transformed to Doppler domain before STAP processing takes place. For the period where target and clutter remain in the same range-Doppler resolution cell, the measured space-time snapshot can be expressed by the random vector [4,8]

$$\mathbf{Z}(u_t, f_a) = a\mathbf{s}(u_t, f_a) + \mathbf{C}(f_a) + \mathbf{N}(f_a), \quad \mathbf{Z} \in \mathbb{C}^{M \times 1}, \quad (18.122)$$

where a is a complex constant comprising the target's RCS among others and $\mathbf{C} + \mathbf{N}$ is the clutter-plus-noise interference.

The optimum detection for one range-Doppler resolution cell, under the assumption of homogeneous Gaussian clutter, is achieved by comparing [4]

$$T_{\text{opt}}(\mathbf{z}) = \left| \mathbf{d}^H(u_t, f_a) \mathbf{R}_W^{-1}(f_a) \mathbf{z}(f_a) \right|^2 \quad (18.123)$$

to a threshold. Clutter suppression is performed by multiplying each Doppler frequency bin f_a of the signal vector \mathbf{z} (i.e., the space-time snapshot transformed to Doppler domain) with the inverse of the clutter-plus-noise covariance matrix $\mathbf{R}_W^{-1}(f_a)$. A target match (i.e., matched filtering) is performed afterwards by multiplying the intermediate result with the complex conjugated and transposed steering vector \mathbf{d}^H (i.e., the expected moving target signal). With this matched filtering operation the Doppler shift f_{DC} (which is proportional to the line-of-sight velocity v_{r0} , cf. (18.31)) and the direction-of-arrival angle ψ_{DOA} of the moving target signal (from which the true azimuth position can be computed).

The clutter covariance matrix for each Doppler frequency bin can be estimated from the data by performing averaging in range:

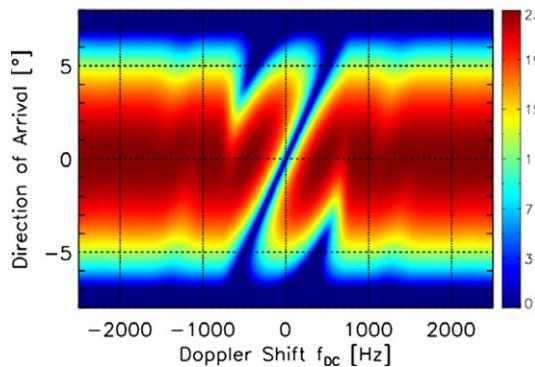
$$\widehat{\mathbf{R}}_W(f_a) = \frac{1}{k} \sum_{k=1}^k \mathbf{z}(r_k, f_a) \mathbf{z}^H(r_k, f_a), \quad \widehat{\mathbf{R}}_W(f_a) \in \mathbb{C}^{M \times M}. \quad (18.124)$$

The resulting detection performance in the optimum case is directly proportional to the remaining output SCNR of the optimum filter given as

$$\text{SCNR}(u_t, v_{r0}) = |a|^2 \mathbf{d}^H(u_t, v_{r0}) \mathbf{R}_W^{-1}(u_t, v_{r0}) \mathbf{d}(u_t, v_{r0}). \quad (18.125)$$

It can be used for analyzing the detection performance of a given system design [4,8]. Maximizing the SCNR results in a maximized probability of detection P_d for a fixed false alarm rate P_{fa} .

A visualization example of the SCNR is given in Figure 18.44. Here u_t and v_{r0} are substituted by the direction-of-arrival angle ψ_{DOA} and the Doppler shift f_{DC} .

**FIGURE 18.44**

SCNR (in dB) as a function of Doppler frequency and direction-of-arrival angle.

Under assumption of homogeneous Gaussian clutter and non-fluctuating target RCS (Swerling-0 case) the probability of detection can be computed analytically as [68]

$$\begin{aligned} P_d &= \int_{\beta_T}^{\infty} x \exp \left[-\frac{(x^2 + a^2)}{2} \right] I_0(\alpha x) dx \\ &= 1 - \int_0^{\beta_T} x \exp \left[-\frac{(x^2 + a^2)}{2} \right] I_0(\alpha x) dx, \end{aligned} \quad (18.126)$$

where $I_0(\cdot)$ is the modified zero-order Bessel function of the first kind and α^2 is given as

$$\alpha^2 = 2 \cdot \text{SCNR}. \quad (18.127)$$

The normalized detection threshold β_T calculates to

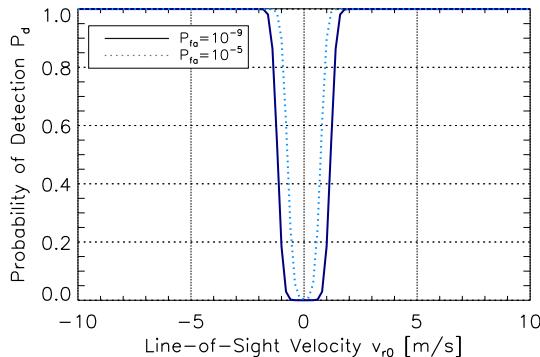
$$\beta_T = \sqrt{-2 \ln P_{fa}}. \quad (18.128)$$

where P_{fa} is the desired false alarm rate.

The probability of detection for a given false alarm rate can either be plotted as a function of SCNR, or, more relevant in the field of GMTI, as a function of line-of-sight velocity. An example is shown in Figure 18.45. The minimum detectable velocity (MDV) can directly be read off. For achieving a probability of detection of 0.9 for given SNR and CNR values of 20 dB, the MDVs are approx. 1 m/s ($=3.6$ km/h) and 1.5 m/s ($=5.4$ km/h) for false alarm rates of 10^{-5} and 10^{-9} , respectively.

Most STAP techniques were developed to be used with multi-channel airborne systems. Airborne STAP techniques have in contrast to spaceborne techniques the following advantages:

- High SNR values so that only a small amount of temporal samples has to be considered (at state-of-the-art airborne systems 64–256 azimuth samples of each RX channel are used).

**FIGURE 18.45**

Probability of detection for an exemplary system with assumed values of $\text{SNR} = \text{CNR} = 20 \text{ dB}$ as a function of line-of-sight velocity for two different false alarm rates.

- Due to the small amount of samples quadratic and higher order phase terms of the moving target signals can be neglected (but that also implies that the motion parameters normally causing these quadratic and higher order phase terms cannot be estimated with the methods presented in Sections 2.18.6.5 and 2.18.7) so that in most cases a simple FFT is sufficient for focusing without applying a matched filter bank.
- Due to the small amount of samples the range cell migration can be neglected.

The successful application of STAP techniques on multi-channel spaceborne systems is more challenging. Since spaceborne systems suffer from low SNR significantly more temporal samples of the moving target signal need to be integrated coherently for ending up with a sufficient target detection performance. Thus, especially for systems with high range resolution, the range cell migration of the target signal has to be taken into account (cf. Section 2.18.4.1) and a matched filter bank has to be applied (cf. Section 2.18.7.1). Promising novel techniques are the ISTAP (imaging STAP) and EDPCA (extended DPCA) techniques introduced recently in [39,40].

2.18.8.3 EDPCA

The Extended Displaced Phase Center Antenna Technique (EDPCA) is an extension of the DPCA and ATI methods to three or more channels [39,40]. The flow chart of the algorithm is shown in Figure 18.46. The SAR compression filter is matched to the moving target parameters for maximizing target's SNR. For this task, the application of a bank of SAR processing filters is necessary. That means that for each of the M complex SAR images several times an adapted range cell migration correction and azimuth compression has to be performed. The number of necessary iterations depends on the range resolution and the accepted loss in SCNR compared to the optimum case.

The clutter cancellation filter is either derived from the estimated clutter-plus-noise covariance matrix or pre-computed using the known system, instrument and geometry parameters. EDPCA is partially adaptive and can be used with an arbitrary number of RX channels.

The empirical clutter-plus-noise covariance matrix is estimated by averaging the measured data vector over N_r range cells and N_a azimuth cells (note that EDPCA operates in contrast to Post-Doppler STAP on fully compressed data) [69]

$$\widehat{\mathbf{R}}_E(\boldsymbol{\vartheta}) = \frac{1}{N_r N_a} \sum_{r=R_{\min}}^{R_{\max}} \sum_{T=T_{\min}}^{T_{\max}} \mathbf{Z}_E(r, T, \boldsymbol{\vartheta}_s) \mathbf{Z}_E^H(r, T, \boldsymbol{\vartheta}_s) \in \mathbb{C}^{M \times M}, \quad (18.129)$$

where in the parameter vector $\boldsymbol{\vartheta}_s$ all moving target parameters (i.e., along-track, across-track or line-of-sight velocity) used for adapted SAR are condensed. The clutter-plus-noise covariance matrix is only of dimension $M \times M$ (M = number of RX channels) and is thus invertible with low computational power ($\cong O(M^3)$).

For achieving a good performance the clutter-plus-noise covariance matrix should be estimated for each image pixel under test. A large number of range and azimuth cells should be used for averaging. The pixel under test and a guard zone should be excluded from training. Already detected targets should also be excluded.

The normalized test statistics is given as [69]

$$T_{EDPCA}(r, T, \boldsymbol{\vartheta}_s) = \frac{\left| \mathbf{d}_E^H(\boldsymbol{\vartheta}_s) \mathbf{R}_E^{-1}(\boldsymbol{\vartheta}_s) \mathbf{Z}_E(\boldsymbol{\vartheta}_s) \right|^2}{\mathbf{d}_E^H(\boldsymbol{\vartheta}_s) \mathbf{R}_E^{-1}(\boldsymbol{\vartheta}_s) \mathbf{d}_E(\boldsymbol{\vartheta}_s)} \geq \eta_E, \quad (18.130)$$

where \mathbf{d}_E is the steering vector and η_E is a CFAR threshold.

What is not shown in the flow chart in Figure 18.46 is a clustering stage, where several pixel-based detections of the same target are clustered to only one physical target.

2.18.8.4 ISTAP

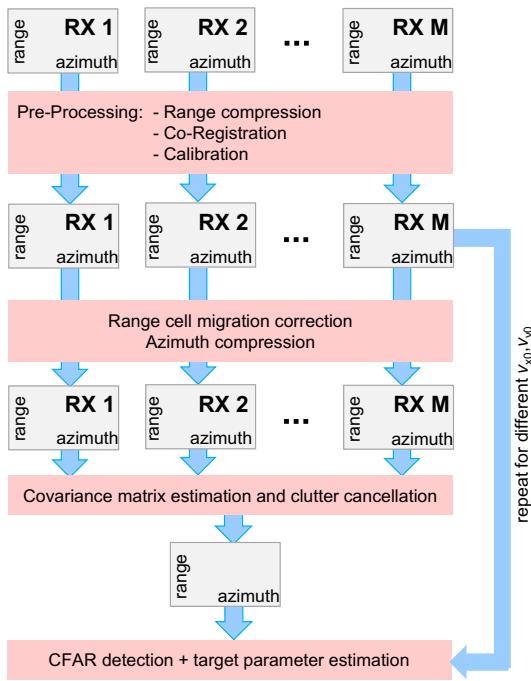
Imaging space-time adaptive processing (ISTAP) is a combination of Post-Doppler STAP and SAR [39, 40]. As with Post-Doppler STAP clutter cancellation is performed in the Doppler domain. However, there is no segmentation into short coherent processing intervals so that all the data are coherently processed. Thus, high SCNR values can be achieved so that the ISTAP technique is well suited for spaceborne systems.

The flow chart of ISTAP is shown in Figure 18.47. As with EDPCA the SAR compression filter is matched to the moving target parameters for maximizing the target SNR. Again a bank of SAR processing filters is necessary. However, ISTAP requires less computational power since the clutter-plus-noise covariance needs only to be estimated once and not for every iteration. Furthermore, also clutter cancellation needs to be performed only once.

The empirical Doppler-dependent clutter-plus-noise covariance matrix is estimated by averaging the measured data vector over N_r range cells, identical to conventional Post-Doppler STAP [70]:

$$\widehat{\mathbf{R}}_I(f_a) = \frac{1}{N_r} \sum_{r=R_{\min}}^{R_{\max}} \mathbf{Z}_I(r, f_a) \mathbf{Z}_I^H(r, f_a) \in \mathbb{C}^{M \times M}. \quad (18.131)$$

Again, the clutter-plus-noise covariance matrix is only of dimension $M \times M$ and is thus invertible with low computational power ($\cong O(M^3)$).

**FIGURE 18.46**

Processing chain for EDPCA.

The clutter-plus-noise covariance matrix needs to be estimated for each Doppler bin.

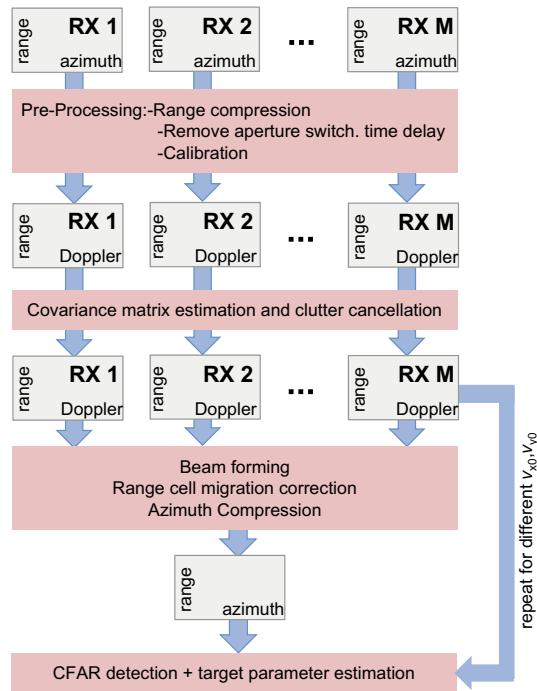
The normalized test statistics for ISTAP is given as [70]

$$T_{\text{ISTAP}}(r, T, \vartheta_s) = \frac{\left| \int H_a(f_a, \vartheta_s) d_I^H(f_a, \vartheta_s) R_I^{-1}(f_a) Z_I(f_a, \vartheta_s) d f_a \right|^2}{\int d_I^H(f_a, \vartheta_s) R_I^{-1}(f_a) d_I(f_a, \vartheta_s) d f_a} \geq \eta_I, \quad (18.132)$$

where H_a is the SAR transfer function (only the Doppler slope is considered but not the shift) which maximizes the SNR of the target with parameter ϑ_s , d_I is the steering vector and η_I is the CFAR threshold.

2.18.9 Conclusion and future trends

A tutorial for GMTI with multi-channel SAR systems was provided. The SAR principle was explained and a single- and multi-channel moving point target signal model was derived. A special focus was laid on the effects caused by moving target signals in the SAR images. The knowledge about these effects is essential for understanding the different parameter estimation methods. The classical dual-channel techniques ATI and DPCA and their limitations were discussed in detail, since they are currently of

**FIGURE 18.47**

Processing chain for ISTAP.

special importance for state-of-the-art spaceborne SAR-GMTI systems. A general GMTI processing chain was presented and basic Doppler parameter estimation methods were introduced. Also a short introduction to STAP, which is covered in detail by a separate chapter in the ELSEVIER e-reference, was given.

The ongoing research in GMTI can be summarized as:

- Improvement of the target detection performance and the reduction of false alarms.
- Accurate and robust SAR-GMTI algorithms for spaceborne SAR-MTI systems.
- Improvement of the position and motion parameter estimation performance. This is of special importance for spaceborne SAR-GMTI where due to low SNR and SCNR values large estimation errors may arise. First promising techniques are ISTAP and EDPCA introduced in [39,40]. A different method requiring a large along-track baseline was presented in [71].
- Reduction of the system complexity and processing time. This is of special importance for affordable real-time systems to be used for civilian traffic monitoring. A suitable technique which also can be combined with classical STAP was introduced in [72].
- Trend to shorter wavelengths (e.g., Ka-band with wavelengths in the order of 8 mm): this allows for smaller instruments enabling both high resolution and good GMTI performance.

Table 18.2 Summary of GMTI Algorithms

Algorithm	Clutter Cancell.	Adaptive	# RX Ch.	Position	Estimation accuracy		Prob. of detect.	Comput. load
					Along-track velocity	Across-track velocity		
Single-channel (general)	No	No	1	Very low	Good	Very low	Very low	Low/ Medium
Matched filter bank	No	No	≥ 1	No	Good	No	Low	Medium
Fractional fourier transform	Yes	No	≥ 1	No	Good	No	Low	Medium
DPCA	Good	No	2	No	No	No	Good	Low
ATI	No	No	2	Low	No	Low	Good	Low
DPCA-ATI	Good	No	≥ 3	Good	Good	Good	Good	Medium
Post-doppler STAP	Very good	Yes	≥ 2	Good	No	Good	Good	Medium
EDPCA	Very good	Partially	≥ 2	Good	Good	Good	Very good	Highest
ISTAP	Very good	Yes	≥ 2	Very good	Good	Very good	Very good	High

- Low PRF GMTI algorithms which can be used together with high-resolution wide swath (HRWS) SAR imaging.

2.18.9.1 Summary of algorithms

In Table 18.2 a summary of the algorithms treated in this tutorial is given. Also the parameter estimation accuracy and the required computational load are assessed in a qualitative way.

Applications:

- Detection and parameter estimation of moving targets in the military field.
- Civilian traffic monitoring of road vehicles and ships.

Open Issues and Problems:

- Accurate estimation of moving target's along-track velocity, especially without the negative influence of accelerations.
- Improvement of the moving target parameter estimation accuracy of spaceborne GMTI algorithms.
- Reduction of false alarms.

Glossary

ATI	along-track interferometry; the phase differences of the signals received by two antennas separated in along-track or flight direction are measured. The phase difference is related to moving target parameters.
Clutter	unwanted radar echos; for GMTI echos from the stationary background scene are considered as clutter
DPCA	displaced phase center antenna; similar antenna arrangement as with ATI but instead of the phases the amplitudes are evaluated. Commonly DPCA is used for suppressing the clutter, i.e., the signals of stationary targets.
IRF	impulse response function; focused signal of a point target
(G)MTI	(ground) moving target indication; the detection of targets moving on ground and the estimation of their geographical positions, their velocities and moving directions.
SAR	synthetic aperture radar; a side-looking imaging radar system exploiting the Doppler effect due to platform motion for imaging
SCR	signal-to-clutter ratio
SCNR	signal-to-clutter-plus-noise ratio
SNR	signal-to-noise ratio
STAP	space-time adaptive processing
SWMF	stationary window matched filter; a two-dimensional filter used for SAR processing; every conventional SAR processor is considered as a SWMF in this tutorial

Relevant Theory: Signal Processing Theory and Array Signal Processing

See [Vol. 1, Chapter 3](#) Discrete-Time Signals and Systems

See [Vol. 1, Chapter 12](#) Adaptive Filters

See [Vol. 3, Chapter 19](#) Array Processing in the Face of Nonidealities

References

- [1] K. Tomiyasu, Tutorial review of synthetic-aperture radar SAR with applications to imaging of the ocean surface, Proc. IEEE 66 (1978) 563–583.
- [2] M.I. Skolnik (Ed.), Radar Handbook, second ed., McGraw-Hill, 1990.
- [3] I.G. Cumming, F.H. Wong, Digital Signal Processing of Synthetic Aperture Radar: Algorithms and Implementation, Artech House, 2005.
- [4] J.H.G. Ender, Space-time processing for multichannel synthetic aperture radar, Electron. Commun. Eng. J. 11 (1999) 29–38.
- [5] R.K. Raney, Synthetic aperture imaging radar and moving targets, IEEE Trans. Aerosp. Electro. Syst. AES-7 (3) (1971) 499–505.
- [6] R. Werninghaus, S. Buckreuss, The terraSAR-X mission and system design, IEEE Trans. Geosci. Remote Sens. 48 (2) (2010) 606–614.
- [7] J.H.G. Ender, The airborne experimental multi-channel SAR system AER-II, in: 1st European Conference on Synthetic Aperture Radar (EUSAR), Königswinter, Germany, March 1996, pp. 49–52.
- [8] J.H.G. Ender, C.H. Gierull, D. Cerutti-Maori, Improved space-based moving target indication via alternate transmission and receiver switching, IEEE Trans. Geosci. Remote Sens. 46 (12) (2008) 3960–3974.
- [9] K. Bethke, S. Baumgartner, M. Gabele, D. Hounam, E. Kemptner, D. Klement, G. Krieger, R. Erxleben, Air-and spaceborne monitoring of road traffic using SAR moving target indication—Project TRAMRAD, ISPRS J. Photogramm. Remote Sens. 61 (3–4) (2006) 243–259.
- [10] D. Cerutti-Maori, J. Klare, A.R. Brenner, J.G.H. Ender, Wide area traffic monitoring with the SAR/GMTI system PAMIR, Trans. Geosci. Remote Sens. 46 (10) (2008) 3019–3030.
- [11] S. Suchandt, H. Runge, H. Breit, U. Steinbrecher, A. Kotenkov, U. Balss, Automatic extraction of traffic flows using terraSAR-X along-track interferometry, IEEE Trans. Geosci. Remote Sens. 48 (2) (2010) 807–819.
- [12] S.V. Baumgartner, G. Krieger, Real-time road traffic monitoring using a fast a priori knowledge based SAR-GMTI Algorithm, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, Hawaii, USA, July 2010.
- [13] (2012, March) Floating car data. <http://en.wikipedia.org/wiki/Floating_car_data>.
- [14] W. Koch, R. Klemm, Ground target tracking with stap radar, IEE Proc. -Radar Sonar Navig. 148 (3) (2001) 173–185.
- [15] V. Chen, H. Ling, Time-Frequency Transforms for Radar Imaging and Signal Analysis, Artech House Publishers, 2002.
- [16] P. Berens, U. Gebhardt, J. Holzner, ISAR imaging of ground moving vehicles using PAMIR data, in: International Radar Conference—Surveillance for a Safer World (RADAR), Bordeaux, France, October 2009, pp. 1–5.
- [17] W. Rieck, Zeit-Frequenz-Signal-Analyse für Radaranwendungen mit synthetischer Apertur (SAR), Ph.D. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 1997.
- [18] S. Chiou, M.V. Dragosevic, Moving target indication via RADARSAT-2 multichannel synthetic aperture radar, EURASIP J. Adv. Signal Process. 2010 (2010).

- [19] J.C. Curlander, R.N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*, Wiley-Interscience, New York, 1991.
- [20] A. Moreira, J. Mittermayer, R. Scheiber, Extended chirp scaling algorithm for air- and spaceborne SAR data processing in stripmap and scanSAR imaging modes, *IEEE Trans. Geosci. Remote Sens.* 34 (5) (1996) 1123–1136.
- [21] J.J. Sharma, The Influence of Target Acceleration on Dual-Channel SAR-GMTI (Synthetic Aperture Radar Ground Moving Target Indication) Data, University of Calgary, Calgary, Alberta, Canada, Master Thesis, October 2004.
- [22] J.J. Sharma, C.H. Gierull, M.J. Collins, Compensating the effects of target acceleration in dual-channel SAR-GMTI, *IEE Proc. Radar Sonar Navig.* 153 (1) (2006) 53–62.
- [23] J.J. Sharma, The influence of target acceleration on velocity estimation in dual-channel SAR-GMTI, *IEEE Trans. Geosci. Remote Sens.* 44 (1) (2006) 134–147.
- [24] S. Baumgartner, G. Krieger, Acceleration-independent along-track velocity estimation of moving targets, *IET Radar Sonar Navig.* 4 (3) (2010) 474–487.
- [25] R.P. Perry, R.C. DiPietro, R.L. Fante, SAR imaging of moving targets, *IEEE Aerosp. Electron. Syst.* 35 (1) (1999) 188–200.
- [26] J. Jao, Theory of synthetic aperture radar imaging of a moving target, *IEEE Trans. Geosci. Remote Sens.* 39 (9) (2001) 1984–1992.
- [27] F. Meyer, S. Hinz, A. Laika, R. Bamler, A-priori information driven detection of moving objects for traffic monitoring by sar, in: *Proc. IEEE Int. Geosci. Remote Sens. Symp. IGARSS '05*, vol. 4, 2005, pp. 2932–2936.
- [28] R.M. Goldstein, H.A. Zebker, Interferometric radar measurements of ocean surface currents, *Nature* 328 (1987) 707–709.
- [29] H.A. Zebker, J. Villasenor, Decorrelation in interferometric radar echoes, *IEEE Trans. Geosci. Remote Sens.* 30 (5) (1992) 950–959.
- [30] E. Rodriguez, J.M. Martin, Theory and design of interferometric synthetic aperture radars, *IEE Proc. F Radar Signal Process.* 139 (2) (1992) 147–159.
- [31] D. Just, R. Bamler, Phase statistics of interferograms with applications to synthetic aperture radar, *Appl. Optics* 33 (20) (1994) 4361–4368.
- [32] A. Budillon, G. Ferraiuolo, V. Pascazio, G. Schirinzi, Multichannel SAR interferometry via classical and bayesian estimation techniques, *EURASIP J. Appl. Signal Process.* 2005 (20) (2005) 3180–3193.
- [33] A. Budillon, V. Pascazio, G. Schirinzi, Estimation of radial velocity of moving targets by along-track interferometric sar systems, *IEEE Geosci. Remote Sens. Lett.* 5 (3) (2008) 349–353.
- [34] C.H. Gierull, Statistical analysis of multilook SAR interferograms for CFAR detection of ground moving targets, *IEEE Trans. Geosci. Remote Sens.* 42 (4) (2004) 691–701.
- [35] S. Chiu, Application of fractional fourier transform to moving target indication via along-track interferometry, *EURASIP J. Appl. Signal Process.* 20 (2005) 3293–3303.
- [36] F.R. Dickey, M.M. Sante, Final report on anti-clutter techniques, General Electric Co., Heavy Military Electron. Dept., Syracuse, NY, Rep. No. R65EMH37, 1953.
- [37] R. Klemm, *Space-Time Adaptive Processing: Principles and Applications*, IEE, UK, 1998.
- [38] P. Lombardo, F. Colone, D. Pastina, Monitoring and surveillance potentialities obtained by splitting the antenna of the COSMO-SkyMed SAR into multiple sub-apertures, *IEEE Proc. Radar Sonar Navig.* 153 (2) (2006) 104–116.
- [39] D. Cerutti-Maori, I. Sikaneta, Optimum GMTI processing for space-based SAR/GMTI systems—theoretical deviation, in: European Conference on Synthetic Aperture Radar (EUSAR), Aachen, Germany, June 2010, pp. 1–4.

- [40] D. Cerutti-Maori, I. Sikaneta, Optimum GMTI processing for space-based SAR/GMTI systems—simulation results, in: European Conference on Synthetic Aperture Radar (EUSAR), Aachen, Germany, June 2010, pp. 1–4.
- [41] C.H. Gierull, Digital Channel Balancing of Along-Track Interferometric SAR Data, Defence R&D Canada, Ottawa, Canada, Technical Memorandum DRDC Ottawa TM 2003–024, March 2003.
- [42] S. Chiu, C. Gierull, A. Durak, Clutter effects on ground moving targets' interferometric phase, in: IEEE Geoscience and Remote Sensing Symposium (IGARSS), vol. 4, Seoul, Korea, July 2005, pp. 2928–2931.
- [43] J.R. Guerci, Space-Time Adaptive Processing for Radar, Artech House, 2003.
- [44] C.P. Banahan, D. Perks, C. Baker, K. Woodbridge, GMTI clutter cancellation using real non-ideal data, *IET Radar Sonar Navig.* 4 (2) (2010) 302–314.
- [45] I.C. Sikaneta, J.-Y. Chouinard, Eigendecomposition of the multi-channel covariance matrix with applications to SAR-GMTI, *Signal Process.* 84 (9) (2004) 1501–1535.
- [46] C.H. Gierull, Moving Target Detection with Along-Track SAR Interferometry, Defence RD Canada–Ottawa, Tech. Rep. DRDC Ottawa TR 2002–084, August 2002.
- [47] C.H. Gierull, I.C. Sikaneta, Raw data based two-aperture SAR ground moving target indication, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Toulouse, France, vol. 2, July 2003, pp. 1032–1034.
- [48] S.V. Baumgartner, G. Krieger, SAR traffic monitoring using time-frequency analysis for detection and parameter estimation, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Boston, USA, vol. 2, July 2008, pp. II-25–II-28.
- [49] S.V. Baumgartner, M. Gabele, G. Krieger, K.-H. Bethke, S. Zuev, Traffic monitoring with SAR: implications of target acceleration, in: 6th European Conference on Synthetic Aperture Radar (EUSAR), Dresden, Germany, May 2006.
- [50] F. Zhou, R. Wu, M. Xing, Z. Bao, Approach for single channel SAR ground moving target imaging and motion parameter estimation, *IET Radar Sonar Navig.* 1 (1) (2007) 59–66.
- [51] C.H. Gierull, Azimuth positioning of moving targets in two-channel SAR by direction-of-arrival estimation, *Electron. Lett.* 40 (21) (2004) 1380–1381.
- [52] F. Meyer, S. Hinz, A. Laika, D. Weihsing, R. Bamler, Performance analysis of the TerraSAR-X traffic monitoring concept, *ISPRS J. Photogramm. Remote Sens.* 61 (3–4) (2006) 225–242.
- [53] S. Baumgartner, G. Krieger, Acceleration independent along-track velocity estimation of moving targets, in: 7th European Conference on Synthetic Aperture Radar (EUSAR), Friedrichshafen, Germany, vol. 4, June, 2008, pp. 207–210.
- [54] Automatic Identification System, June 2011. <http://en.wikipedia.org/wiki/Automatic_Identification_System>.
- [55] S.V. Baumgartner, G. Krieger, Large along-track baseline SAR-GMTI: first results with the TerraSAR-X/TanDEM-X satellite constellation, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, Canada, July 2011, pp. 1319–1322.
- [56] S. Barbarossa, Detection and imaging of moving objects with synthetic aperture radar -Part 1: Optimal detection and parameter estimation theory, *IEE Proc. F* 139 (1) (1992) 79–88.
- [57] M. Gabele, I. Sikaneta, Motion parameter estimation of Doppler-Ambiguous moving targets in SAR-GMTI, in: International Radar Symposium (IRS), Cologne, Germany, September 2007.
- [58] R. Scheiber, P. Prats, M. Nannini, K. Macedo, C. Andres, J. Fischer, R. Horn, Advances in airborne SAR interferometry using the experimental SAR system of DLR, in: European Microwave Week (EuMW), Horizon House, Munich, Germany, October 2007, pp. 91–94.
- [59] L.B. Almeida, The fractional Fourier transform and time-frequency representations, *IEEE Trans. Signal Process.* 42 (11) (1994) 3084–3091.

- [60] J. Radon, Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, Berichte Sächsische Akademie der Wissenschaften, Leipzig, Mathematisch-Physikalische Klasse 69 (1917) 262–277.
- [61] A. Kutay, H. Ozaktas, O. Ankan, L. Onural, Optimal filtering in fractional Fourier domains, IEEE Trans. Signal Process. 45 (5) (1997) 1129–1143.
- [62] H.-B. Sun, G.S. Liu, H. Gu, W.-M. Su, Application of the fractional Fourier Transform to moving target detection in airborne SAR, IEEE Trans. Aerosp. Electron. Syst. 38 (4) (2002) 1416–1424.
- [63] S. Barbarossa, A. Farina, Detection and imaging of moving objects with synthetic aperture radar. 2. Joint time-frequency analysis by Wigner-Ville distribution, IEE Proc. F. 139 (1) (1992) 89–97.
- [64] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, The Wigner distribution tool for time-frequency signal analysis, Part I: Continuous-time signals, Philips J. Res. 35 (3) (1980) 217–250.
- [65] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, The Wigner distribution tool for time-frequency signal analysis, Part II: Discrete time signals, Philips J. Res. 35 (4–5) (1980) 276–300.
- [66] J.C. Wood, D.T. Barry, Radon transformation of time-frequency distributions for analysis of multicomponent signals, IEEE Trans. Signal Process. 42 (11) (1994) 3166–3177.
- [67] S. Barbarossa, A. Farina, Space-time-frequency processing of synthetic aperture radar signals, IEEE Trans. Aerosp. Electron. Syst. 30 (2) (1994) 341–358.
- [68] W.L. Melvin, A STAP overview, IEEE A&E Syst. Mag. 19 (1) (2004) 19–35.
- [69] D. Cerutti-Maori, I. Sikaneta, A generalization of DPCA processing for multi-channel SAR/GMTI radars, IEEE Trans. Geosci. Remote Sens. 51 (1) (2013) 560–572.
- [70] D. Cerutti-Maori, I. Sikaneta, C.H. Gierull, Optimum SAR/GMTI processing and its application to the radar satellite RADARSAT-2 for traffic monitoring, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 3868–3881.
- [71] S.V. Baumgartner, G. Krieger, K.-H. Bethke, A Large along-track baseline approach for ground moving target indication using TanDEM-X, in: International Radar Symposium (IRS), Cologne, Germany, September 2007.
- [72] S.V. Baumgartner, G. Krieger, Fast GMTI Algorithm for traffic monitoring based on a priori knowledge, IEEE Trans. Geosci. Remote Sens. 50 (11) (2012) 4626–4641.

Introduction to Inverse Synthetic Aperture Radar

19

Marco Martorella

*Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy
CNIT—Radar and Surveillance Systems (RaSS) National Laboratory, Pisa, Italy*

2.19.1 Introduction

Radar Imaging refers to the ability to form images of natural or man-made objects using Electro-Magnetic echo location. As will become clearer later, coherent radars may have suitable specifications that allow implementation of special features using specific signal processing. We may argue that, given a suitable coherent radar, radar imaging can be provided by adding some “special” signal processing to the received signal. Conventional radar images are typically represented as two-dimensional (2D) images where a mapping function transforms the three-dimensional (3D) world into a 2D image. An obvious comparison could be formulated with photographic images, as these are also the result of some mapping from the 3D world into a 2D photographic image. However, there are several differences that may be pointed out in regards of the type of mapping and image features. Radar images, as well as other type of images (e.g., photographic, infra-red, X-ray images) are usually characterized by means of some quality measurements, such as geometrical and radiometric resolution and signal-to-noise (SNR) ratio.

Pushed by the need to form high quality radar images that can be used in applications such as automatic target recognition and classification, researchers have designed a variety of radar imaging processors. In this tutorial we will introduce the fundamental concepts at the base of radar imaging and we will provide an overview of the most commonly used radar imaging techniques. Examples will be also used throughout this chapter to clarify concepts and to show some radar imaging results.

2.19.2 Historical overview

We have to distinguish two starting points when considering the origins of radar imaging: one for Synthetic Aperture Radar (SAR) and one for Inverse Synthetic Aperture Radar (ISAR). Although the two approaches to radar imaging have quite a lot in common, there are some significant differences that mark a line between them. As mentioned in [1], the SAR concept was conceived in 1951 by Carl Wiley, although the first operational system (classified) was built in 1957 by the Willow Run Laboratories of the University of Michigan for the US Department of Defense (DoD). Unclassified SAR systems were successfully built by NASA in the 1960s. The first spaceborne SAR system, SEASAT-A, was launched in 1978. Although this spaceborne system was specifically designed for oceanographic purposes, it also produced important results in other fields, such as in ice and land studies. The results obtained with

SEASAT-A demonstrated the importance of radar imaging for the observation of the earth. Since then, several spaceborne SAR systems have been launched that provide improved resolution, wider coverage and faster revisit times. Several airborne SAR system have also been developed to overcome limitations of spaceborne SAR systems, such as cost, revisiting time and resolution. After the first experiments on the 1960s operated by the NASA, other important missions have been accomplished, such as the SIR-A, SIR-B and SIR-C missions, which flew in 1981, 1984, and 1994, respectively. The history of ISAR began later, when Walker and Aushermann, with their pioneering work [2,3] introduced the concept of radar imaging of rotating objects with fixed antennas. The main insight in their work was to exploit Doppler information generated by the rotation of an object to separate echoes returning from different parts of the object along a cross-range axis. Such Doppler separation, together with the time-delay separation (along the radar range), produces a two-dimensional (2D) image, which can be mapped onto an image plane.

2.19.3 High resolution radar and radar imaging

The definition of an image as given by the IEEE is “a spatial distribution of a physical property such as radiation, electric charge, conductivity, or reflectivity, mapped from another distribution of either the same or another physical property.” Narrowing this definition to a radar image, we may define a radar image as “a spatial distribution of the electromagnetic (EM) reflectivity of an object mapped from a distribution of currents on the object’s surface.” The latter concept of mapping an object’s current distribution onto an image of the object’s reflectivity function derives from the fact that a radar is able to capture echoes of e.m. energy irradiated by the radar itself and backscattered from the object. The backscattering effect is produced by the object because the incident e.m. field induces a set of currents on the object’s surface, which in turns produces a scattered e.m. field that (in part) propagates back to the radar. If a radar image has fine resolution, the object’s reflectivity may be observed in fine detail with the result that it would be possible to spatially separate reflectivity contributions from different parts of the object. It is quite intuitive that the finer the resolution the more the detail that may be visible in the image. A desirable radar imaging system would provide finer resolution to allow characterizing smaller and smaller scale objects.

Although 3D radar imaging is nowadays possible, we will consider the more usual concept of 2D radar images. Therefore, we will be conscious that the mapping function that links the object of interest with its image is mathematically representable with a function $f : \mathbb{C}^3 \rightarrow \mathbb{C}^2$, where the symbol \mathbb{C} represents the set of complex numbers, as radar images are represented by complex numbers (I & Q or magnitude and phase).

Radar images, as will be clearer later, are typically represented in cartesian coordinates, where one axis is aligned with the radar range direction and the other with the cross-range direction (otherwise indicated as azimuth direction). It is worth pointing out that the range direction is uniquely identified by the orientation of the radar antenna (typically coincident with the antenna maximum gain direction). In a 3D world, there are an infinite number of directions that are orthogonal to the range direction, therefore, the concept of cross-range direction becomes highly ambiguous if some constraints are not applied that uniquely define such a direction. Nevertheless, a radar image will be identified as a mapping of some physical quantity that is defined in a 3D coordinate system onto a plane identified by the range direction and a cross-range direction, which, as it will be clearer later, will depend on the radar-target geometry and dynamics.

Nevertheless, in order to enable radar imaging capabilities in a radar, sufficiently fine resolution must be achievable in both the range and cross-range directions. A desirable radar image will show fine and possibly similar resolution in both axes.

2.19.3.1 Resolution

The resolution is the minimum distance between two quantities at which a measurement system is said to be able to distinguish two separate contributions. This general definition may be applied to all sorts of measurements and the resolution may be expressed in terms of a specific measurement unit, e.g., meters for spatial resolution, Hertz for frequency resolution, and so on.

By applying the general definition of resolution to radar, this can be defined a little more specifically as the minimum distance along a given direction between two point-like scatterers of equal magnitude such that the two scatterers can be distinguished by the radar. This concept is further explained in Figure 19.1. The two scatterers may be distinguished if the two received echoes are sufficiently separated such that a signal processing is able to detect both contributions separately. According to Rayleigh criterion, the minimum distance that allows separating two echoes is equal to the echo half-power width.

Traditionally, in radar, we talk about range resolution to indicate the ability of the radar to distinguish scatterers along the range direction. The same concept applies to azimuth and elevation resolution in the azimuth and elevation planes.

We will now pay special attention to the Doppler resolution. The Doppler resolution is the ability of the radar to distinguish two point-like scatterers where each scatterer produces a Doppler component due to their radial motion with respect to the radar. The principle behind the Doppler frequency generation can be explained in simple terms in the radar case. With reference to Figure 19.2, let a point like target move with a given velocity v with respect to the radar. Therefore, the radar-target distance can be approximated with the linear function $R(t) = R_0 + v_R t$. Assuming that a radar transmits a pure tone $s_T(t) = A \cos(2\pi f_0 t)$, the received signal can be written by taking into account the round-trip delay and an amplitude attenuation:

$$s_R(t) = B \cos[2\pi f_0(t - \tau(t))], \quad (19.1)$$

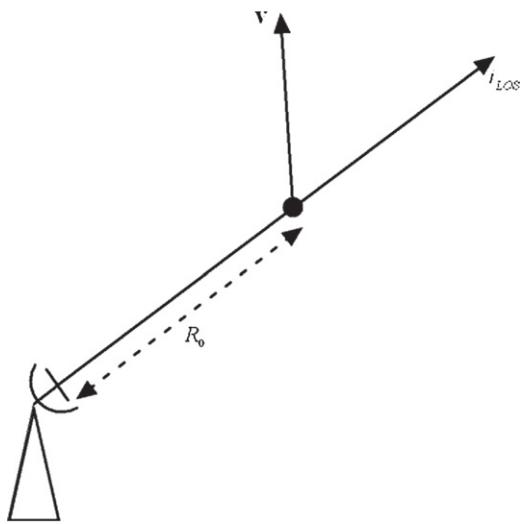
where

$$\tau(t) = \frac{2}{c} R(t) = \frac{2}{c}(R_0 + v_R t) \quad (19.2)$$



FIGURE 19.1

Radar resolution.

**FIGURE 19.2**

Doppler frequency.

and $B < A$. By substituting (19.2) in (19.1) we obtain

$$s_R(t) = B \cos [2\pi(f_0 + f_D)t + \varphi_0],$$

where $\varphi_0 = -\frac{4\pi R_0}{\lambda}$ with $\lambda = \frac{c}{f_0}$ and the Doppler frequency can be defined as $f_D = -\frac{2v_R}{\lambda}$, where v_R is the scatterer's radial velocity and λ is the radar wavelength. The radar Doppler resolution, as will be clear later, becomes important it is directly related to the cross-range resolution in Inverse Synthetic Aperture Radar (ISAR) systems.

In radar imaging, the concept of resolution applies to range and cross-range, as they represent the two coordinates in a radar image. Therefore, we can define the concept of radar image resolution as a pair of values that indicate the range and cross-range resolution in the image.

In the following subsections, the concepts of range and cross-range resolution will be detailed and methods for obtaining high range and cross-range resolution will be addressed.

2.19.3.2 High range resolution

In pulsed radar where the phase is not measured (or used), the range resolution is typically associated with the transmitted pulse duration. In fact, as a first approximation, a signal echo due to an ideal scatterer, persists for a time interval equal to the length of the transmitted pulse. Therefore, we should expect to be able to detect a second ideal scatterer when the echo of the first scatterer vanishes, leading to an expression for the range resolution equal to $\Delta r = \frac{c\Delta T}{2}$, where ΔT is the pulse duration [4]. Such a concept is roughly represented in Figure 19.1, where the echoes relative to two ideal scatterers are partly overlapped. It becomes evident that a minimum distance between two ideal scatterers exists, below

which, the two echoes are largely overlapped and therefore the two contributions are not distinguishable any longer. As finite bandwidth signals have infinite duration pulses, we will use the definition of duration at the point where peak power has reduced by 3 dB, which is consistent with the Rayleigh criterion.

In practice, all radar systems employ a matched filter, which ensures that a maximum SNR is obtained at the filter output. It must be recalled that a matched filter produces the transmitted signal autocorrelation function at its output when an echo is present at its input. The autocorrelation function has the following properties:

1. It is the Fourier Transform of the Energy Spectral Density (ESD),
2. $B\Delta\tau = 1$,

where B is the transmitted signal bandwidth and $\Delta\tau$ is the relation signal duration at the output of the matched filter.

The second property, which represents the uncertainty relationship of a Fourier Transform pair (in this case, it is given by the autocorrelation and the ESD), indicates that to obtain short duration pulses at the output of the matched filter wide bandwidth signals should be transmitted.

Although, short duration pulses provide a straightforward way to increase the bandwidth, this should be discouraged as it has the drawback of reducing the transmitted energy (and hence the system sensitivity), unless higher peak power pulses are used in transmission to compensate. Typically, phase modulations are used to increase the signal bandwidth rather than amplitude modulation since they are able to effectively increase the signal bandwidth without having to decrease the pulse duration, e.g., linear frequency modulated signals (or chirp signals) are typical phase modulated signals.

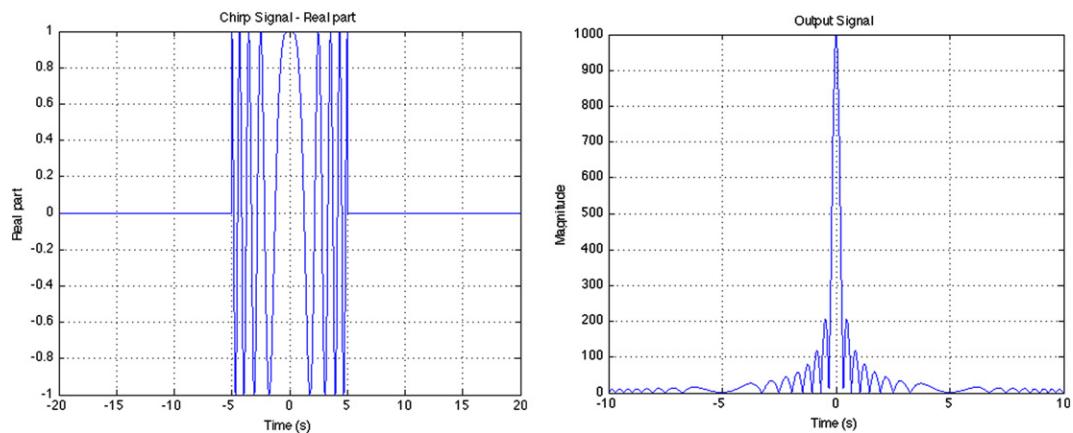
If we agree that two echoes relative to two separate scatterers can be distinguished only if the echoes are separated by a delay equal to the pulse duration at the output of the matched filter ($\Delta\tau$), then we can say that the range resolution is related to the transmitted signal bandwidth by means of the following expression:

$$\Delta R = \frac{c}{2} \Delta\tau = \frac{c}{2B}. \quad (19.3)$$

The effect of using wide bandwidth transmitted signals in conjunction with a matched filter is called *pulse compression* and it is a standard way to proceed to enable higher range resolution. The term pulse compression comes from the simple fact that the duration of the pulse at the matched filter output is shortened by a factor $\rho = \frac{\Delta T}{\Delta\tau}$, namely compression factor, with respect the duration of the transmitted pulse (ΔT). As an example, if we want to achieve a range resolution of 1 m, we have to generate a signal with a bandwidth of 150 MHz, independently of its duration, which corresponds to a non-modulated signal of 1.5 ns of duration.

A side effect of the pulse compression is the presence of sidelobes in the compressed signal. Sidelobes are unwanted signal peaks that may mask other echoes and they must be suppressed or attenuated as much as possible. A performance indicator in terms of sidelobes is the Side Lobe Level (SLL), which is defined as ratio between the pulse peak and the highest sidelobe peak. For an unweighted signal, the SLL is typically around -13 dB.

Both analog and digital modulations are typically used to obtain wide bandwidth waveforms. Chirp signals are the most commonly used among analog modulations as they are easy to generate and they show desired characteristics in terms of SLL and robustness to noise and Doppler effect. Alternatively,

**FIGURE 19.3**

Pulse compression—Chirp signal—Uncompressed (transmitted) pulse on the left and Compressed pulse (Matched filter output) on the right.

Barker and pseudo-random digital codes can be used to generate digital phase modulated signals, which have similar characteristics to those of chirp signals. The big advantage of pulse compression is that high resolution can be achieved by transmitting long pulses, therefore maintaining high sensitivity at long ranges.

An example of pulse compression is provided in Figure 19.3, where a chirp signal is considered as transmitted signal.

2.19.3.3 High cross-range resolution

The ability to resolve scatterers in a cross-range direction is related to the angular separation in the same direction. Traditionally, azimuth and elevation resolution were achieved by building wide antennas since, approximately, the angular resolution is inversely proportional to the antenna size. For example, the angular resolution of a rectangular antenna can approximately be calculated as follows:

$$\alpha_x \simeq \frac{\lambda}{L_x}, \quad (19.4)$$

where λ is the radar wavelength, L_x is the size of the antenna along a given cross-range direction (usually named as azimuth and elevation) and α_x is the angular resolution along the same cross-range direction (expressed in radians). Although it is directly related to system parameters, specifically the antenna size and the radar wavelength, the angular resolution is not sufficient to provide high resolution needed for radar imaging. The main reasons are that:

- it refers to an angular domain and not a spatial domain (images should be scaled in spatial coordinates and not angular coordinates),

- the cross-range resolution, which can be calculated from the angular resolution as shown in (19.5), becomes range (R) dependent

$$\delta_x = R\alpha_x = \frac{R\lambda}{L_x}. \quad (19.5)$$

At long ranges, the cross-range resolution may become coarse, even in the case of fine angular resolution. To give an example, if we consider a distance $R = 10^4$ m, a wavelength $\lambda = 3$ cm and an antenna size $L_x = 3$ m, the cross-range resolution would be equal to $\delta_x = 100$ m. Conversely, if we wanted to obtain a cross-range resolution equal to 1 m at a distance of 10 km, we would have to build a 300 m wide antenna.

It is evident that this problem cannot be solved by building very wide antennas as there are practical limits to that. Another solution could be to build an antenna array, which may relax the previous problem. Nevertheless, at very long ranges the problem would still be too hard to solve. Inverse Synthetic Aperture Radar provides a solution to this problem enabling high cross-range resolution as if a wide aperture had been used. This concept will be detailed in the next section.

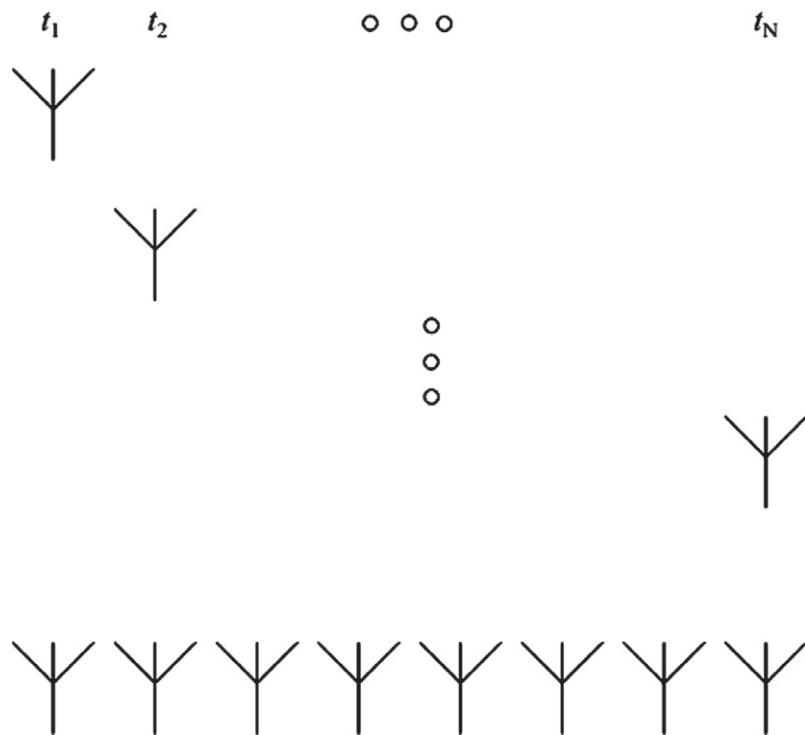
2.19.4 Inverse synthetic aperture radar

The concept of Inverse Synthetic Aperture Radar will be introduced in a modern way. Rather than traditionally considering an ISAR system as a configuration where the radar is static and an object moves with respect to it, we will migrate from the SAR concept and geometry.

2.19.4.1 From SAR to ISAR

As pointed out in Section 2.19.3, real aperture antennas or antenna arrays do not provide a viable solution for radar imaging systems. Nevertheless, high cross-range resolution can only be enabled if an antenna aperture can be formed. The first idea of SAR was conceived by thinking of a single element that moves along a given trajectory, therefore providing the means for forming a virtual array in a given time interval. Such concept is depicted in Figure 19.4, where a synthetic array formation is compared with a real array. As the formation of the synthetic aperture is not instantaneous, any equivalence between the synthetic aperture and a real array can be stated only if the illuminated scene is static during the synthetic aperture formation (from t_1 to t_N). If such an assumption can be made, there would be no physical difference between the signal acquired by a synthetic aperture radar and a real aperture radar (which makes use of a real array). The condition under which the effect of the element motion can be neglected is known as the *stop & go* assumption, which implies that the transmission of the signal and the reception of its echo occur instantaneously at a particular position. Obviously, this assumption cannot be perfectly matched unless the platform that carries the radar stops every time a pulse is transmitted and received before moving to the next position. Nevertheless, in practical scenarios, such an assumption can be considered satisfied since the round trip delay (the time for the e.m. wave to propagate from the radar to the illuminated scene and back) is short enough to neglect the element offset created by the platform motion during such a time interval.

Attention should now be paid to the “relative motion” that there is between the platform and the target, as such a motion is not necessarily produced by a moving platform that carries the radar. Relatively speaking, if the sensor is stationary and the target moves with respect to it inducing a relative motion,

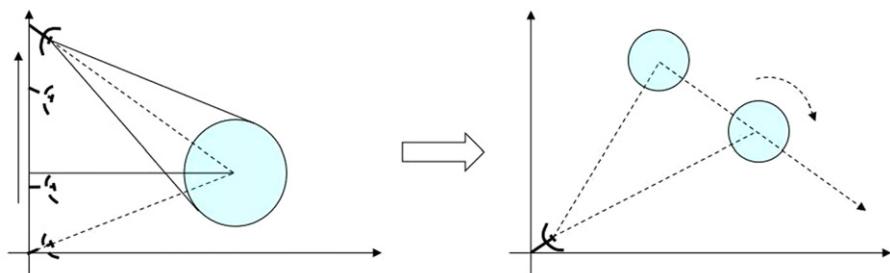
**FIGURE 19.4**

Synthetic aperture—virtual array.

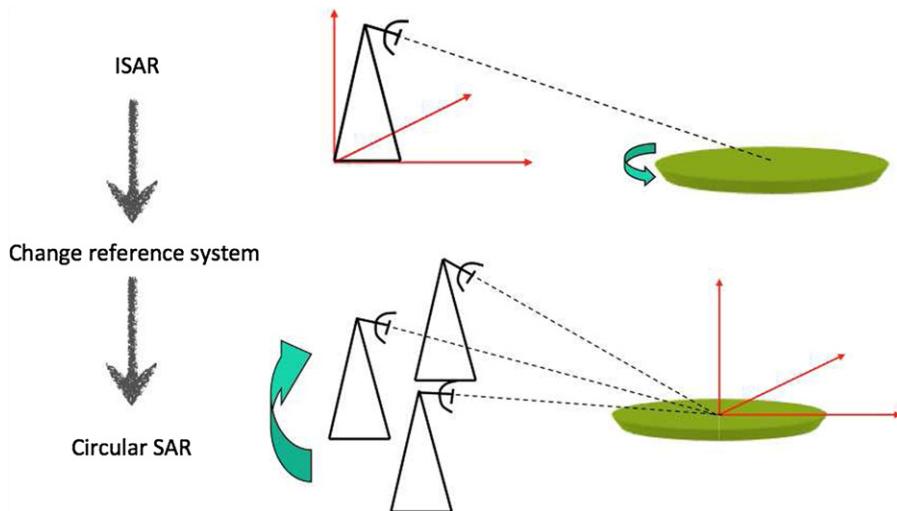
a synthetic aperture would be created at the same way. To strengthen this concept, one could argue that the cases of stationary target and moving platform and the case of stationary platform and moving target can only be stated once the reference system has been chosen. In fact, by placing the reference system on the target, the first case is enabled whereas, by placing the reference system on the radar, the latter is obtained. According with this last view, the differences between synthetic aperture and inverse synthetic aperture would only depend on where the reference system is positioned. Such a concept is depicted in Figure 19.5 where a Spot-light SAR configuration is transformed into an ISAR configuration by moving the reference system from the target to the radar.

Conversely, the same concept may be argued by starting with a controlled ISAR configuration such as that of a turntable experiment. In a turntable configuration, the antenna is fixed on the ground (typically mounted on a turret) and the target is positioned on a turntable, which rotates as the radar takes measurements of the target. By moving the reference system from the radar to the target, a circular SAR geometry can be enabled, as depicted in Figure 19.6.

In truth, a subtle but significant detail exists that substantially defines the difference between SAR and ISAR. Such a detail is the cooperation of the illuminated target. To better explain this concept, one may place the reference system on the target. If such a target moves (with unknown motions) with

**FIGURE 19.5**

From Spot-light SAR to ISAR.

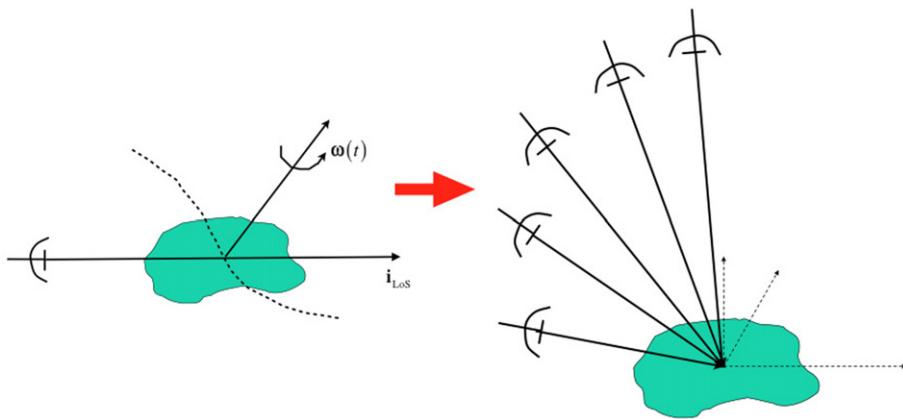
**FIGURE 19.6**

From ISAR to circular SAR.

respect to the radar, the synthetic aperture formed during the CPI differs from the expected one (which is formed by controlled platform motion). Any SAR image formation that follows would be then based on the erroneously predicted synthetic aperture and therefore would lead to defocused images. A pictorial example is shown in Figure 19.7.

2.19.4.2 Geometry

Figure 19.8 shows the ISAR Geometry. The reference system T_ξ is embedded in the radar with the axis ξ_2 oriented along the line of sight (LOS). Without losing generality, it is assumed that the target moves along a trajectory that intersects the axis ξ_2 at the central instant $t = 0$. The target rotation due to the translation

**FIGURE 19.7**

Synthetic array formed by a non-cooperative target.

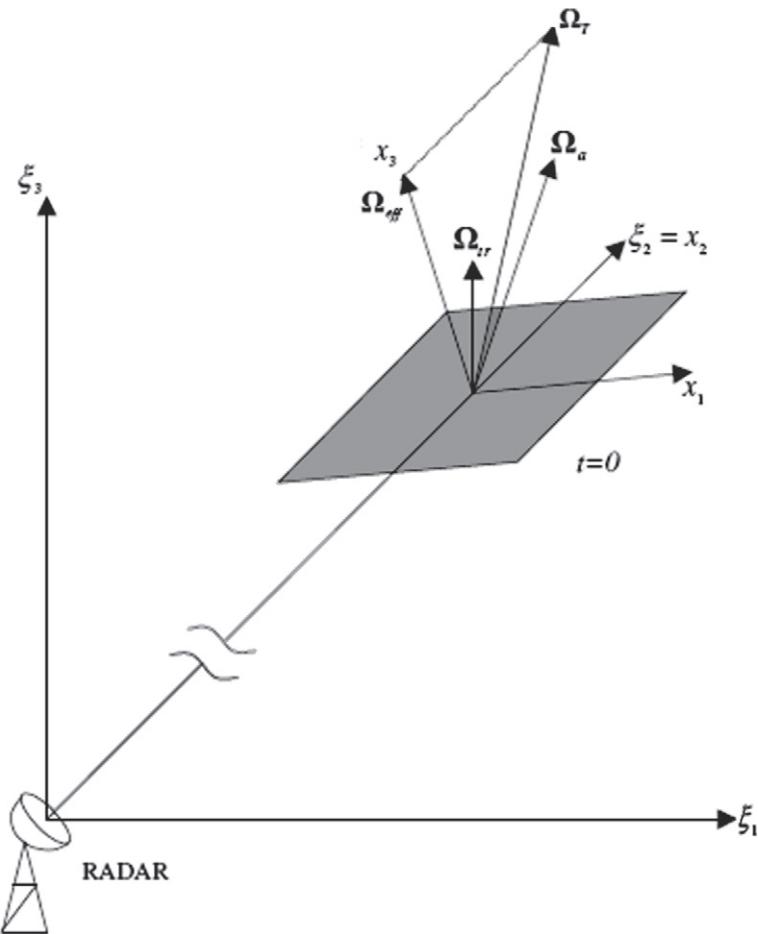
motion is denoted as the translational rotation vector $\Omega_{\text{tr}}(t)$. In practical conditions, external forces produce angular motions that are represented by the angular rotation vector $\Omega_a(t)$, which is applied to the center O of the target (from now on “point O”). The sum of these two rotation vectors yields the total angular rotation vector $\Omega_T(t)$. The projection of $\Omega_T(t)$ on the plane orthogonal to the LOS is called effective rotation vector $\Omega_{\text{eff}}(t)$, which is the rotation vector component that contributes to the target aspect angle variation. The imaging plane (x_1, x_2) is orthogonal to the effective rotation vector and is represented in Figure 19.8 (this will be demonstrated in Section 2.19.5). The time varying coordinate system T_x is chosen so to have the x_2 axis oriented along the LOS, the x_3 axis along the effective rotation vector and the origin in the point “O” at time $t = 0$. With this choice, the x_1 and the x_2 axes become the cross-range and range coordinates of the imaging plane, respectively. It is worth noting that, in general, the imaging plane (x_1, x_2) is time-varying because the effective rotation vector varies with respect to the time.

2.19.4.3 Signal modeling

A convenient way to represent the received signal when dealing with ISAR processing is by using a time-frequency format. In this representation, the frequency coordinate is represented by the variable f , whereas the time coordinate is represented by the variable t . When written in these terms, the received signal can be seen as a time varying signal spectrum, where the time variance is typically introduced by the target relative motions and not by the transmitter (unless time-varying modulations are used, e.g., in adaptive systems).

Therefore, the complex base-band received signal, in free space conditions, can be written in a time-frequency format as follows:

$$S_R(f, t) = W(f, t) \int \xi(\mathbf{x}) \exp \left(-j \frac{4\pi f}{c} R(t) \right) d\mathbf{x}, \quad (19.6)$$

**FIGURE 19.8**

ISAR geometry.

where

$$W(f, t) = \text{rect}\left(\frac{t}{\Delta t}\right) \text{rect}\left(\frac{f - f_0}{B}\right), \quad (19.7)$$

and where f_0 represents the carrier frequency, B is the transmitted signal bandwidth, Δt is the observation time, c is the speed of light in vacuum, $R(t)$ is the distance between the radar and the generic point \mathbf{x} on the target and $\xi(\mathbf{x})$ is the target reflectivity function. Function $\text{rect}(\cdot)$ yields 1 when $|\cdot| < 0.5$ and 0 otherwise.

2.19.4.4 Image formation

ISAR image formation was initially introduced in two pioneering works [2,3], where a Range-Doppler approach was presented to describe how an e.m. image could be formed by exploiting a target's rotation with respect to a radar. Since then, several ISAR image reconstruction techniques have been presented to form well focussed and high resolution images. Most of the techniques presented in the literature are based on a two-step approach. A first step is taken to eliminate the target's radial motion (often addressed with the term autofocus) and a second step is applied to the motion compensated data which aims at forming the image. In the following subsections, the problem of radial motion compensation is firstly addressed and a few techniques are presented that implement this step and then image formation techniques are presented that make use of the classical Fourier approach and Time-Frequency Transforms (TFTs).

2.19.4.4.1 Radial motion compensation

The *straight-iso-range* approximation can be applied when the target is much smaller than the radar-target distance. In practice this is equivalent to effectively approximating the radar-target distance as follows:

$$R(t) \simeq R_0(t) + \mathbf{x}^T \cdot \mathbf{i}_{\text{LoS}}(t), \quad (19.8)$$

where $R_0(t)$ is the residual from point “O” to the radar at the time t , \mathbf{x} is the column vector that identifies a scatterer on the target and \mathbf{i}_{LoS} is the column unit vector that identifies the radar Line of Sight. To further clarify this concept, the reader may refer to Figure 19.9.

By substituting (19.8) into (19.6) we obtain the expression in (19.9)

$$S_R(f, t) = W(f, t) \exp \left(-j \frac{4\pi f}{c} R_0(t) \right) \int \xi(\mathbf{x}) \exp \left(-j \frac{4\pi f}{c} \mathbf{x}^T \cdot \mathbf{i}_{\text{LoS}}(t) \right) d\mathbf{x}. \quad (19.9)$$

The radial motion compensation aims at eliminating the phase modulation produced by the phase term outside the integral, namely

$$\phi_0(t) = \exp \left(-j \frac{4\pi f}{c} R_0(t) \right). \quad (19.10)$$

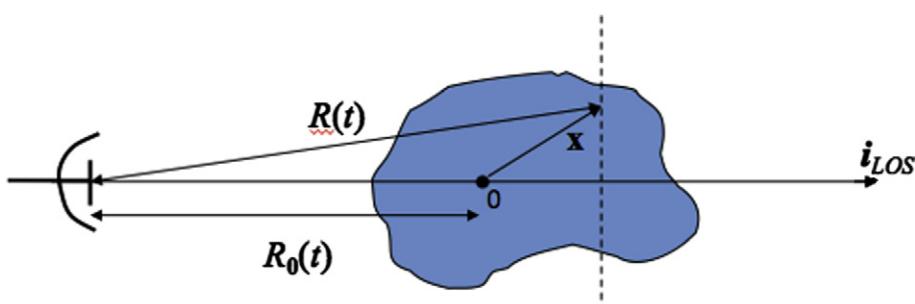


FIGURE 19.9

Straight iso-range approximation.

In ISAR applications the target is typically non-cooperative, therefore, the term in (19.10) is not known a priori. The direct consequence is that such a motion compensation must be performed blindly. Techniques used in ISAR imaging to perform this task are usually referred to as *autofocusing techniques*, as they aim at focusing ISAR images by trying to compensate for the phase modulation introduced by $\phi_0(t)$, which typically provoke an image defocusing effect.

After perfect radial motion compensation, the compensated signal can be written as follows:

$$S_C(f, t) = W(f, t) \int \xi(\mathbf{x}) \exp\left(-j \frac{4\pi f}{c} \mathbf{x}^T \cdot \mathbf{i}_{\text{LoS}}\right) d\mathbf{x}. \quad (19.11)$$

2.19.4.4.2 Range-doppler image formation

We will now concentrate on the phase term inside the integral in (19.11). It is worth noting that the scalar product in the phase term is actually the radial coordinate $x_2(t)$. After radial motion compensation, any scatterer's radial coordinate variation ($x_2(t)$) would be generated by the target's rotation with respect to point "O." It is worth pointing out that only a component of the target's rotation vector produces an effective target's aspect angle variation. To explain this concept a bit more clearly let the target's rotation vector be represented by the sum of two components, one aligned with the radar LoS and the other orthogonal to it. This can be expressed mathematically as follows:

$$\Omega_T = \Omega_{\text{LoS}} + \Omega_{\text{eff}}. \quad (19.12)$$

We will now demonstrate that the component aligned with the radar LoS (Ω_{LoS}) does not produce any range variation for any of the target's scatterers.

To demonstrate this concept, we will refer to Figure 19.10, where a radar is depicted together with a target with an arbitrary scatterer that rotates with respect to a reference system defined by three cartesian

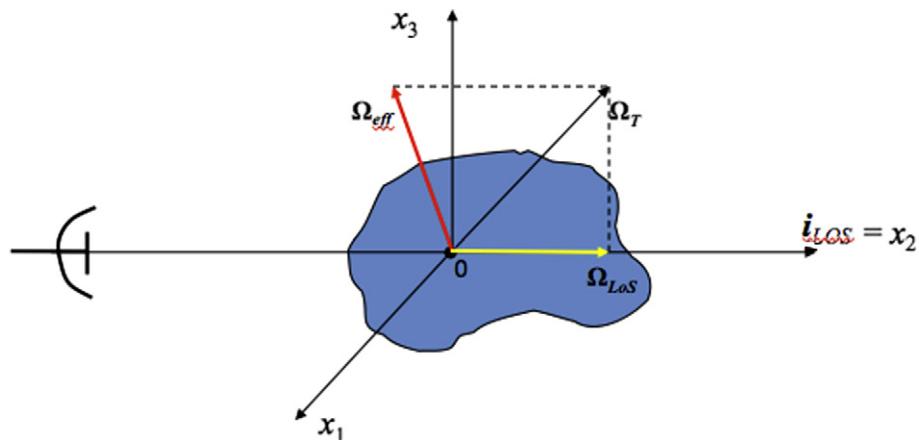


FIGURE 19.10

Effective rotation vector.

coordinates (x_1, x_2, x_3) . By assuming that the radar-target distance is much greater than the size of the target (same assumption made for the straight iso-range approximation), and that the radar is aligned with the x_2 axis, we can write the differential equation systems that rules the rotational motion of a rigid body given a rotation vector $\omega(t)$

$$\dot{\mathbf{x}}(t) = \boldsymbol{\Omega}_T \times \mathbf{x}(t)$$

with the initial condition $\mathbf{x}(0) = \mathbf{x}_0$. Therefore, radar-target distance changes can be measured as changes along the LoS direction (x_2 axis), which can be calculated as follows:

$$\dot{x}_2(t) = \Omega_{T_3}x_1(t) - \Omega_{T_1}x_3(t),$$

which indicates that the rotation vector component along the LoS (x_2 axis) does not produce any target-radar distance changes.

This means that the radar is actually blind to any motion induced by rotation vector component aligned with the LoS. The effective target rotation vector is the orthogonal component that produces scatterer's range variations. Therefore, variations will be observed in the image plane, which means that only the coordinates (x_1, x_2) will play a role. More specifically, the effective rotation vector causes radial variations that are the result of a rotation around the effective rotation axis. The effect on the compensated signal phase can mathematically be expressed as follows:

$$S_C(f, t) = W(f, t) \int \xi(\mathbf{x}) \exp\left(-j \frac{4\pi f}{c} (x_1 \sin(\Omega_{\text{eff}} t) + x_2 \cos(\Omega_{\text{eff}} t))\right) d\mathbf{x}. \quad (19.13)$$

We can now demonstrate that when the effective rotation vector can be considered constant within an observation time $\Delta\tau$, the ISAR image can be interpreted as the image of the target's projection onto a plane, namely the *Image Projection Plane* (IPP). This can be done by simply manipulating the Eq. (19.13), as follows:

$$S_C(f, t) = W(f, t) \int \xi'(x_1, x_2) \exp\left(-j \frac{4\pi f}{c} (x_1 \sin(\Omega_{\text{eff}} t) + x_2 \cos(\Omega_{\text{eff}} t))\right) dx_1 dx_2, \quad (19.14)$$

where

$$\xi'(x_1, x_2) = \int \xi(\mathbf{x}) dx_3 \quad (19.15)$$

is the target's reflectivity function projected onto the image plane.

When small aspect angles are spanned, the signal in (19.13) can be approximated with the following:

$$S_C(f, t) \simeq W(f, t) \int \xi'(x_1, x_2) \exp\left(-j \frac{4\pi f}{c} (x_1 \Omega_{\text{eff}} t + x_2)\right) dx_1 dx_2. \quad (19.16)$$

After making the following substitutions:

$$\begin{aligned} v &= \frac{2f}{c} \Omega_{\text{eff}} x_1, \\ \eta &= \frac{2}{c} x_2. \end{aligned} \quad (19.17)$$

Equation (19.16) can be rewritten as follows:

$$S_C(f, t) \simeq K W(f, t) \int \xi'(\eta, v) \exp(-j2\pi(vt + f\eta)) d\eta d\tau. \quad (19.18)$$

It should be pointed out that the compensated signal in (19.18) can be read as a windowed Fourier Transform (FT) of the target's projected reflectivity function.

Therefore, an image of the projected reflectivity function $\xi'(\eta, v)$ can be obtained by simply applying an Inverse Fourier Transform (IFT) to the compensated signal. The result can be written as follows:

$$I_C(\eta, v) = \text{2D-IFT}[S_C(f, t)] = K w(\eta, v) \otimes \otimes \xi'(\eta, v). \quad (19.19)$$

A direct interpretation of the result in (19.19) suggests that the target's ISAR image obtained by applying a 2D-IFT to the radial motion compensated signal is a filtered version of the target's reflectivity function projected onto the ISAR image plane.

2.19.5 ISAR image evaluation

The interpretation of an ISAR image is not as straightforward as that of a SAR image. This is the direct consequence of the target of interest being non-cooperative. Specifically, the following issues must be considered to give a correct interpretation of an ISAR image:

- the cross-range coordinate is represented in the Doppler frequency domain and not in spatial coordinates,
- the IPP, and therefore the target projection shown in the image, is not known a priori,
- the imaging system response, namely the *Point Spread Function* (PSF), is not known a priori and cannot be entirely controlled,
- the cross-range resolution is not known a priori and it is not a controllable parameter.

Each of these aspects is analyzed in the following subsections.

2.19.5.1 ISAR image coordinates

The coordinates η and v , which appear in (19.19) can be interpreted as the delay-time and Doppler coordinates. This is a direct consequence of the substitutions that were made in (19.17). In fact the coordinate η is the exact calculation of the round-trip delay relative to a scatterer located in the range coordinate x_2 , whereas the coordinate v can be interpreted as the Doppler frequency generated by a scatterer with a radial velocity equal to $v_r = \Omega_{\text{eff}}x_1$.

As ISAR images are typically used for target's classification and recognition, a desired output would be an ISAR image represented in terms of spatial coordinates (x_1, x_2) . An operation of coordinate scaling would involve an inversion of the eqs. in (19.17). Such an operation will be detailed in Section 2.19.9.

Theoretically, we may rewrite the result in (19.19) in terms of spatial coordinates, as follows:

$$I_C(x_1, x_2) = K w(x_1, x_2) \otimes \otimes \xi'(x_1, x_2). \quad (19.20)$$

2.19.5.2 Additional considerations on the image projection plane and limitations of the use of the RD technique

As demonstrated in Eqs. (19.14) and (19.15), the function that maps the target's 3D complex reflectivity function onto the 2D image domain is a projection along the target's effective rotation vector. Nevertheless, it should be pointed out that such a result is valid under the following assumptions:

- far field (validity of the straight iso-range approximation),
- constant target's rotation vector.

With reference to the geometry described in Section 2.19.4.2, it should be noted that the range position in the ISAR image only depends on the x_2 coordinate whereas the Doppler position only depends on the x_1 coordinate. The dependence of the range position in the ISAR image on the x_2 coordinate is the direct consequence of the choice of the reference system T_x , which has the x_2 axis aligned with the radar range coordinate ξ_2 . The dependence of the Doppler coordinate on the x_1 axis is demonstrated in (19.17), where a linear relationship holds between the two quantities. It is then evident that the x_3 coordinate does not affect either the range or the Doppler position of a scatterer in the ISAR image. This can be interpreted in the following terms: two scatterers that are located in two positions such that they only differ by the third coordinate, namely x_3 , are mapped onto the same range and Doppler. Therefore, their contributions are coherently summed. By extending this concept, we can say that a range-Doppler bin in an ISAR image will show a complex intensity that is the result of a sum of all contributions from those scatterers that have the same coordinates in terms of (x_1, x_2) and differ for their x_3 coordinate. This is equivalent to saying that the target is projected onto the (x_1, x_2) -plane, which can be interpreted as the IPP.

It is worth pointing out that such a result is critical when interpreting ISAR images, in facts:

- ISAR images are images of the target's projection onto a plane (IPP),
- the IPP is not known a priori and therefore, the interpretation of the ISAR image is consequentially difficult.

2.19.5.3 Point spread function

ISAR image quality can be assessed by defining the PSF. The PSF is the image of an ideal point-like scatterer located in a generic position \mathbf{x} . Therefore, we can easily calculate the ISAR image PSF when the RD is used to form the image. This can be done mathematically as follows:

$$\text{PSF}(x_1, x_2) = K w(x_1, x_2) \otimes A \delta(x_1 - x_{10}, x_2 - x_{20}) = K' w(x_1 - x_{10}, x_2 - x_{20}). \quad (19.21)$$

It should be noted that

- the PSF is space-invariant: an ideal scatterer is imaged in the same way independently of its position,
- the image geometrical resolution only depends on the characteristics of the window $W(f, t)$.

2.19.5.4 ISAR image resolution

The ISAR image resolution can be obtained by analyzing the PSF. In order to analyze the PSF in terms of spatial coordinates we will rewrite (19.14) by introducing the concept of spatial frequency coordinates.

Mathematically, this can be done as follows:

$$S_C(f, t) \simeq W(f, t) \int \xi'(x_1, x_2) \exp(-j2\pi(x_1 X_1 + x_2 X_2)) dx_1 dx_2,$$

where

$$\begin{aligned} X_1 &= \frac{2f}{c} \sin(\Omega_{\text{eff}} t), \\ X_2 &= \frac{2f}{c} \cos(\Omega_{\text{eff}} t). \end{aligned} \quad (19.22)$$

In particular, in the case of small aspect angle variations, i.e., when $\Omega_{\text{eff}} t \ll 1$, $|t| < \frac{\Delta t}{2}$, the polar domain described by the parametric functions in (19.22) can be approximated by a rectangular domain, as shown in Figure 19.11. Under the same assumption, the spatial frequency coordinates defined in (19.22) can be approximated with the following:

$$\begin{aligned} X_1 &= \frac{2f_0}{c} \Omega_{\text{eff}} t, \\ X_2 &= \frac{2f}{c}. \end{aligned} \quad (19.23)$$

It is worth noting that for X_1 the frequency f has been substituted by the central frequency f_0 , as a result of the approximation of the polar domain by a rectangular window, which intersects the angular sector at the coordinate $\bar{X}_2 = \frac{2f_0}{c}$. It should also be noted that this approximation is the one that leads to the minimum error, as it can be inferred by examining Figure 19.11 where the two polar and rectangular windows are superposed.

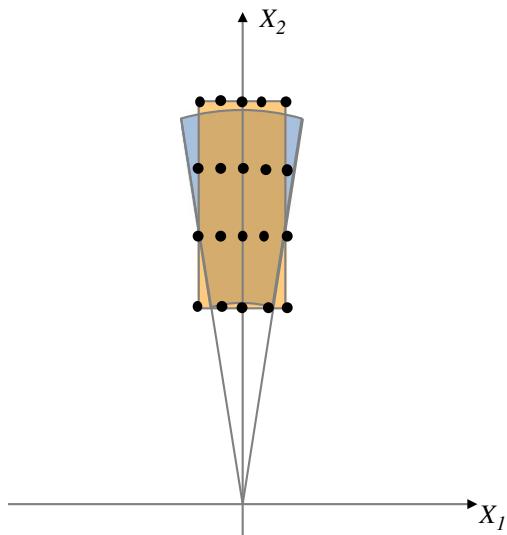


FIGURE 19.11

Fourier domain—rectangular approximation.

By considering the rectangular approximation, we can rewrite the acquisition window $W(f, t)$ in terms of spatial frequencies. This can be done mathematically by substituting (19.23) into (19.7) as follows:

$$W(X_1, X_2) = \text{rect}\left(\frac{X_1}{\frac{2f_0\Omega_{\text{eff}}\Delta t}{c}}\right) \text{rect}\left(\frac{X_2 - X_{20}}{\frac{2B}{c}}\right), \quad (19.24)$$

where $X_{20} = \frac{2f_0}{c}$.

The PSF can be obtained by calculating the IFT of (19.24), as follows:

$$w(x_1, x_2) = K \text{sinc}\left(\frac{x_1}{\delta x_1}\right) \text{sinc}\left(\frac{x_2}{\delta x_2}\right),$$

where

$$\begin{aligned}\delta x_1 &= \frac{c}{2f_0\Omega_{\text{eff}}\Delta t}, \\ \delta x_2 &= \frac{c}{2B}.\end{aligned}$$

are the cross-range and range resolution respectively. The following remarks can be made.

- The range resolution depends on the transmitted signal bandwidth. The direct consequence of this is that the range resolution can be directly controlled (by setting the desired bandwidth B) or, at least, it can be calculated. Therefore, it can be considered as a known parameter.
- The cross-range resolution depends on the target's motions. This results in a parameter that is not controllable which means that it cannot be defined by setting one or more radar parameters. For the same reason, the cross-range resolution is an unknown parameter.
- An ISAR imaging system performance in terms of resolution cannot be predicted unless the target's motions are known. This uncertainty may reflect onto uncertainties in the performance of classifiers based on ISAR images.

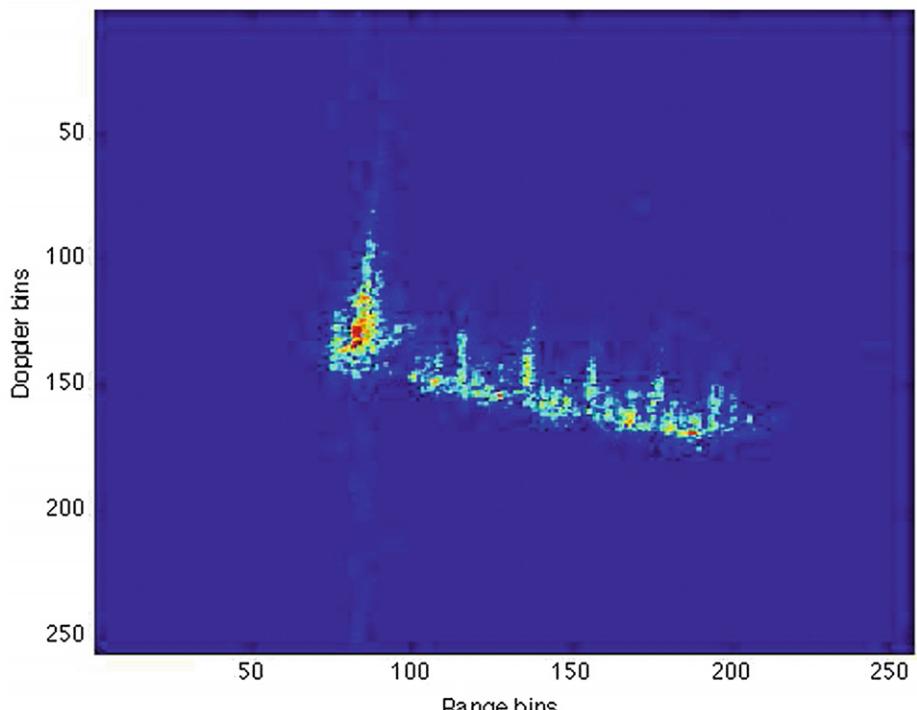
It must be pointed out that the uncertainty about the cross-range resolution is exactly the same problem of cross-range scaling. A technique for estimating the effective rotation vector magnitude will be presented in Section 2.19.9, where a technique will be presented that aims to solve this problem.

2.19.6 Examples of ISAR images

Depending on the type of targets to be imaged, ISAR systems and platform may change to allow covering areas of interest and improving performance. For instance, ISAR imaging of sea vessels is often carried out by means of coastal radars, airborne radars and also by spaceborne radar, whereas, images of aircraft are more likely to be obtained by means of ground-based radar. Some specific ISAR imaging systems are also employed to image spaceborne objects, such as the Fraunhofer TIRA system. An example of a sea vessel ISAR image obtained by processing data collected by an airborne radar is shown in Figure 19.12, together with an aerial picture of the same target. It is interesting to note some important features, such as the ship's cranes, which may be used by a classifier to recognize the type of target. Another example of an aircraft ISAR image is shown in Figure 19.13. This ISAR image was obtained by processing data collected by a ground based radar located near an airport. The target, a Boeing 737



Aerial photo of the target



ISAR image

FIGURE 19.12

Airborne radar imaging of a moving sea vessel (bulk loader).



Aerial photo of the target

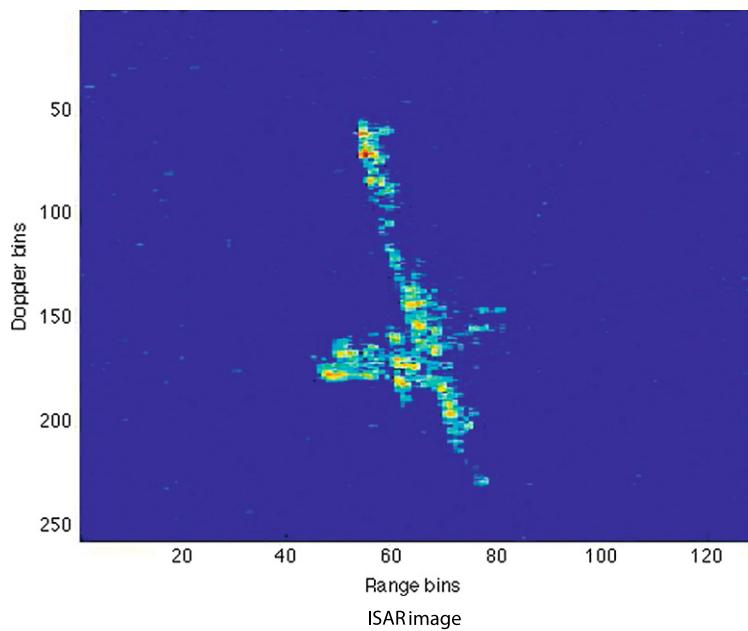


FIGURE 19.13

Ground-based radar imaging of a flying aircraft (Boeing 737).

was taking off during the acquisition. Also in this case features such as length and wing-span may be used to classify this target. The last example is about a spaceborne SAR image of a non-cooperative sea vessel together with the same target's image obtained by using an ISAR processor. It is worth noting that the SAR processor is unable to focus the target as this is moving with respect to the rest of the illuminated scene (see Figure 19.14).

2.19.7 Image autofocus

Radial motion compensation is accomplished by means of a step in the ISAR processing chain that is typically termed *image autofocus*. This operation is accomplished by removing the phase term $\phi_0(t)$ (see Eq. (19.10)). When no external data are available, the motion compensation must be performed only by using the radar received signal. For this specific reason, the image focusing process is called ISAR image autofocus. In many years of research in the field of ISAR autofocus, several techniques have been developed, each of them showing pros and cons. Autofocus algorithms can be classified as parametric and non-parametric [5]. Parametric methods need a parametric model of the radar received signal, whereas non-parametric techniques do not make use of any model. Two parametric and two of the non-parametric techniques are detailed in the following subsections.

2.19.7.1 ICBA

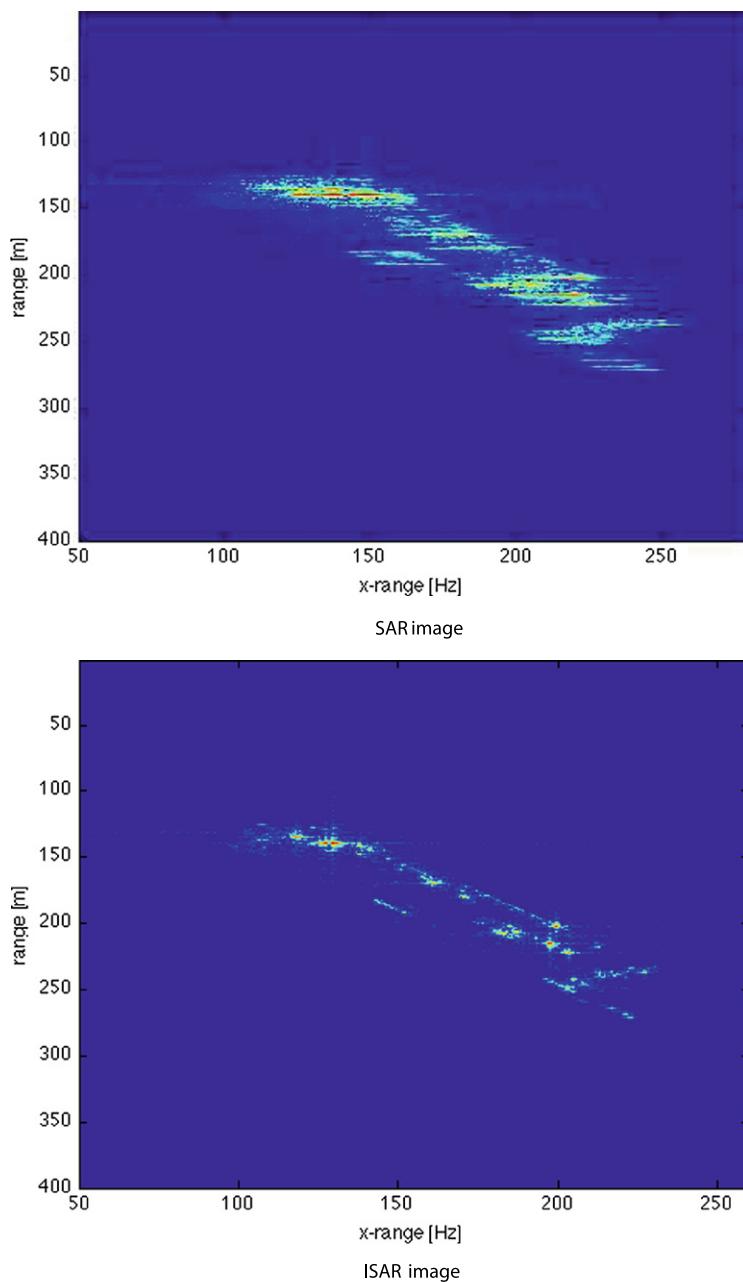
The autofocus technique adopted here is the Image Contrast Based Autofocus (ICBA) algorithm [6]. The ICBA is a parametric technique and it is based on Image Contrast (IC) maximization. The idea behind this approach is based on the simple concept that an image will more focussed when the value of the IC is larger.

The autofocus technique aims at removing the term $R_0(t)$ due to the target's residual translational motion. For a relatively short observation time interval Δt and relatively smooth target motions, the radar-target residual distance can be expressed by means of an N th order polynomial, as defined in (19.25). It is worth pointing out that a smooth target motion does not imply that the target undergoes slow motions but that the function that represents the target's motion is a continuous and differentiable function

$$R_0(t) = \sum_{n=0}^N \frac{1}{n!} \alpha_n t^n. \quad (19.25)$$

In Eq. (19.25), α_n are the *focusing parameters*, which can be denoted by means of a vector format, namely $\boldsymbol{\alpha}$. The estimation of $R_0(t)$ resorts to the estimation of the target radial motion parameters. By naming the radial motion model with $R_0(\boldsymbol{\alpha}, t)$, the radial motion compensation problem can be recast as an optimization problem where the Image Contrast (IC) is maximized with respect to the unknown vector $\boldsymbol{\alpha}$, as defined in (19.26)

$$\hat{\boldsymbol{\alpha}} = \arg \max_{\boldsymbol{\alpha}} \{ \text{IC}(\boldsymbol{\alpha}) \}, \quad (19.26)$$

**FIGURE 19.14**

Spaceborne SAR and ISAR images of a moving sea vessel—ISAR imaging is required to focus non-cooperative moving targets.

where

$$\text{IC}(\boldsymbol{\alpha}) = \frac{\sqrt{A \left\{ [I(\eta, \nu, \boldsymbol{\alpha}) - A[I(\eta, \nu, \boldsymbol{\alpha})]]^2 \right\}}}{A[I(\eta, \nu, \boldsymbol{\alpha})]}, \quad (19.27)$$

and where $A[\cdot]$ indicates an average operation over the variables η and ν , $I(\eta, \nu, \boldsymbol{\alpha})$ can be defined as the ISAR image magnitude or intensity (power) after compensating the target's translational motion by using $\boldsymbol{\alpha}$ as focusing parameters. This can be expressed mathematically as follows:

$$I(\eta, \nu, \boldsymbol{\alpha}) = \left| 2\text{D-IFT} \left[S_R(f, t) \exp \left(j \frac{4\pi f}{c} R_0(\boldsymbol{\alpha}, t) \right) \right] \right|^p, \quad (19.28)$$

where $p = 1$ in the case of image magnitude and $p = 2$ in the case of image intensity.

The optimization problem can be solved numerically by using classical methods, such as the Nelder-Mead [7] or more recent Genetic Algorithms (GA) [8].

2.19.7.2 IEBA

An approach similar to the one proposed in Section 2.19.7.1 can be devised by substituting the IC with the Image Entropy (IE). As for the IC, the IE is a good indicator of the image focus. Unlike the IC, the IE has small values when the image is well focussed whereas it reaches large values when an image is not well focussed. The mathematical expression that may be used to calculate the IE follows in (19.29):

$$IE(\boldsymbol{\alpha}) = - \iint \ln(\bar{I}(\eta, \nu, \boldsymbol{\alpha})) \bar{I}(\eta, \nu, \boldsymbol{\alpha}) d\eta d\nu, \quad (19.29)$$

where $\bar{I}(\eta, \nu, \boldsymbol{\alpha}) = \frac{I(\eta, \nu, \boldsymbol{\alpha})}{A[I(\eta, \nu, \boldsymbol{\alpha})]}$.

2.19.7.3 DSA

The Dominant Scatterer Autofocus (DSA), also known as Hot Spot (HS), is a two-stage technique: the first is to set up a rough alignment of all the profiles, before applying a form of phase compensation in the second step. The principles for this technique were obtained by delving into two other areas of research, namely time delay estimation [9] and adaptive beamforming [10]. A brief overview of the algorithm follows, whilst a more complete description may be found in [11].

After measuring and storing the complex envelopes of the echo samples, high resolution range profiles can be generated. Let $s_R(\eta, t)$ be a range profile acquired at time t . A cross-correlation and shift operation can be performed between successive range profiles in order to obtain a rough range profile alignment. Let us refer to the roughly aligned profiles as:

$$s_1(\eta, t) = A(\eta, t) \exp[j\phi_i(\eta, t)], \quad (19.30)$$

where τ indicates a range cell and t is the slow time.

A search along the range coordinate is then performed in order to find a dominant and stable scatterer. The range cell where such a scatterer is found is called the reference range η_0

$$s_0(t) = A(\eta_0, t) \exp[j\phi_0(t)], \quad A(\eta_0, t) \simeq A. \quad (19.31)$$

The value of η_0 is found by measuring the normalized echo variance in each range cell, and is determined to be the range for which the variance value is a minimum. This approach relies on the assumption that a dominant scatterer with large radar cross-section exists and therefore, the measured phase can be attributed to the phase generated by one point scatterer.

The next step is to perform a phase conjugation using the phase history of $s_0(t)$. In particular, by applying it to the range cell data corresponding to η_0 , the following result is obtained:

$$s_2(t) = A_0(\eta_0, t) \simeq A. \quad (19.32)$$

By applying the same operation to all other range cells motion compensation is achieved:

$$s_C(\eta, t) = A(\eta, t) \exp\{j[\phi(\eta, t) - \phi_0(t)]\}. \quad (19.33)$$

This algorithm is known as the Minimum Variance Algorithm (MVA) or the dominant scatterer algorithm (DSA) due to the criterion used to choose the reference range cell. A more robust version of this algorithm, which combines the echoes from several reference range cells, will be now considered. Called the multiple scatterer algorithm (MSA) [11], this modified algorithm essentially averages the phase differences of M reference scatterers (after unwrapping) to provide the phase correction. Typically three reference range bins ($M = 3$) are sufficient to produce focussed images.

This concept is translated into mathematical details as follows. Let the m th reference cell be represented as:

$$s_m(\eta_m, t) = A(\eta_m, t) \exp[j\phi_m(t)], \quad A_m(t) \simeq A \quad (19.34)$$

and let M be the number of selected range cells. Therefore, the estimation of the phase history $\phi_0(t)$ can be obtained by averaging the phase histories of the selected range cells, as shown in Eq. (19.35)

$$\phi_0(t) = \frac{1}{M} \sum_{m=1}^M \phi_m(t). \quad (19.35)$$

Phase conjugation is then carried out as before (c.f. Eqs. (19.32) and (19.33)). The DSA algorithm is summarized in Figure 19.15.

2.19.7.4 PGA

The extension of the original DSA proposed in [11] leads to a general question about how much information remains in those range cells that are discarded after selecting the MV range cell that could

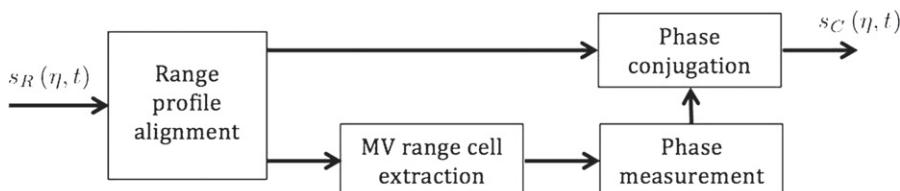
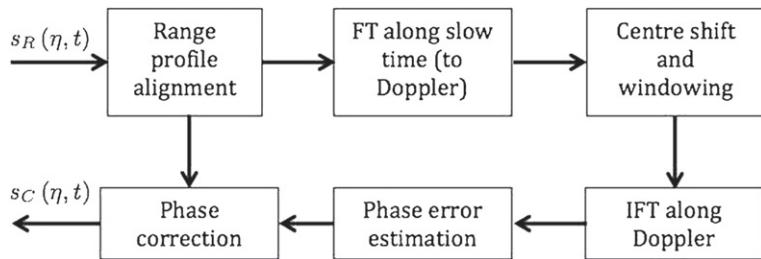


FIGURE 19.15

DSA algorithm flow chart.

**FIGURE 19.16**

PGA flow chart.

be still used to improve the phase error estimation. An answer to this question may be found in the solution proposed by Jakowatz et al. [12], namely the Phase Gradient Algorithm (PGA).

The PGA algorithm substitutes the phase conjugation approach used in the DSA with a solution based on a Maximum Likelihood estimator. The ML approach theoretically uses the information contained in all range cells, although in practice, the reduction in those range cells where the SNR is high enough improves the PGA performance.

As shown in Figure 19.16, a Range-Doppler ISAR image is formed with the roughly aligned range profiles. The peak value in each range cell, which is supposed to be the return from a dominant scatterer, is firstly found along the Doppler coordinate and then center-shifted and windowed in the Doppler domain (Low-pass filter). Each range cell is then transformed back via an Inverse Fourier Transform (IFT) to obtain phase shifted and filtered time histories. This operation corresponds to isolating a scatterer contribution and zero-Doppler forcing, which can be interpreted as a way to remove scatterer's radial motion.

Let an arbitrary k th range cell time history be represented as in (19.36), where two consecutive samples are considered.

$$\begin{aligned} g(k, m-1) &= a(k) \exp[j\phi(m-1)] + n(k, m-1), \\ g(k, m) &= a(k) \exp[j\phi(m)] + n(k, m), \end{aligned} \quad (19.36)$$

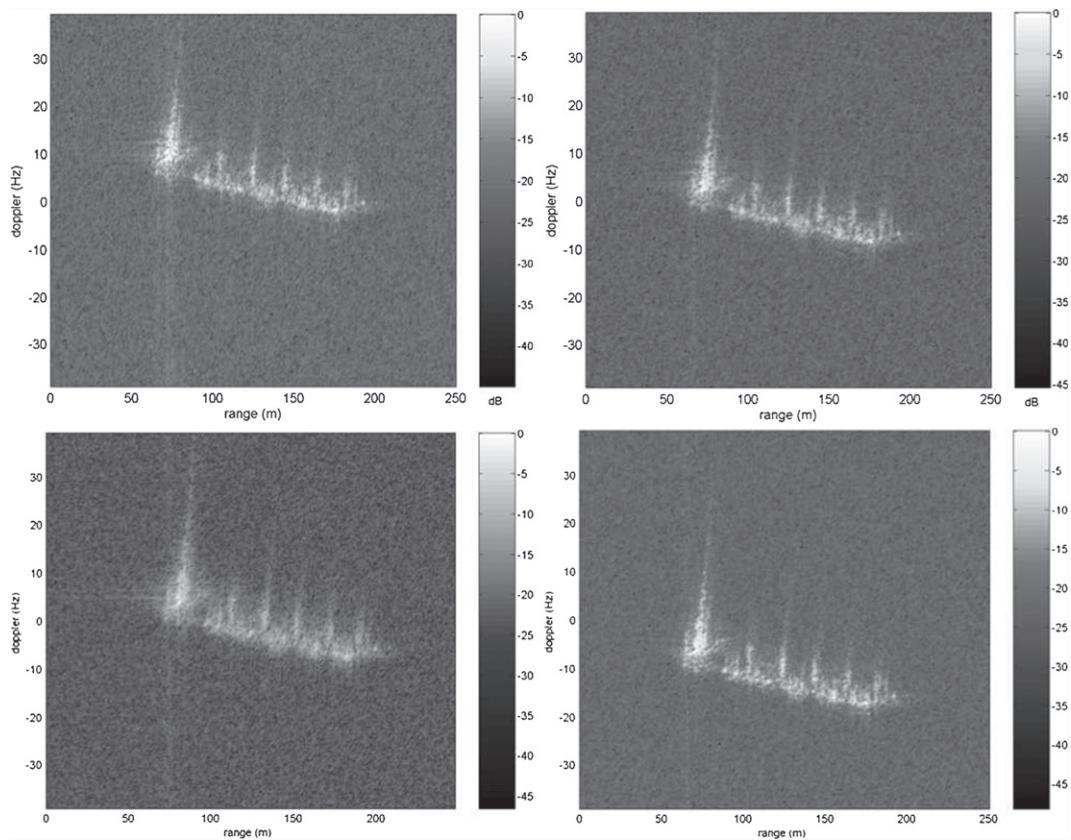
where k indicates the range cell number, m indicates the m th time sample, $a(k)$ is the amplitude, which is assumed to be constant for two consecutive samples, $\phi(m)$ is the phase at time m and $n(k, m)$ is an additive white Gaussian noise sample.

The phase gradient between two consecutive samples can be defined as

$$\Delta\phi = \phi(m) - \phi(m-1). \quad (19.37)$$

The PGA aims to estimate the phase gradient expressed in (19.37) by using the ML approach. The theoretical derivation of the ML estimator and its performance are detailed in [12]. The solution is shown in (19.38)

$$\hat{\Delta\phi}(m) = \angle \left(\sum_{k=1}^N g(k, m) g^*(k, m-1) \right), \quad (19.38)$$

**FIGURE 19.17**

ISAR images formed by applying autofocusing algorithms: (upper-left) DSA, (upper-right) PGA, (bottom-left) IEBA, (bottom-right) ICBA.

where N is the number of range cells used to estimate the phase error and where the symbol $\angle(\cdot)$ indicates the phase of a complex number. The phase correction term can be then calculated by integrating the estimated phase gradient, as follows:

$$\hat{\phi}(m) = \sum_{n=2}^m \Delta\hat{\phi}(n), \quad \hat{\phi}(0) = 0. \quad (19.39)$$

Examples of the application of the DSA, PGA, IEBA and ICBA to a ship target are shown in Figure 19.17. Results show that all techniques are able to correctly focus the ISAR image although small differences can be noticed among the images.

2.19.8 Time-windowing

As already stated, the RD technique can be successfully applied when the effective rotation vector does not change significantly during the CPI. However, the target's own motion may induce a non-uniform target's rotation vector. In order to minimize target's rotation variations, the CPI can be controlled via a time-windowing approach. Typically, an operator would define a fixed time window length (CPI) and would process the entire dataset by sliding the window and forming ISAR images. The same operator would then select the ISAR images that would be suitable for target classification or recognition. One of the most important requirements for such images is a good level of focus, as the target's details would then be sharper than in defocused images. Short CPIs are more likely to provide well focussed images although the resolution may be poor due to a small total aspect angle variation. On the other hand, long CPIs are more likely to provide wider aspect angle variations, although they would increase the chance that the target's effective rotation vector would be time-varying and therefore produce defocused images. It is quite evident that a trade-off must be identified to obtain an optimal result that would be based on obtaining both a high resolution and a well focussed image (See [Video Files 1, 2, 3](#)).

The technique described here is an automatic time-windowing technique, originally proposed in [13]. Specifically, the time window position across the data and its length are automatically chosen in order to obtain one or more images with the highest focus. To better explain this concept, we will refer to Figure 19.18 where the data is referred to as distributed along the time axis and a temporal window is defined by two parameters, namely τ and Δt .

The criterion used to define the highest focused image is based on the IC. Basically, the IC is used as an indicator of the image focus, which is assumed to be a maximum when the position (τ) and length (Δt) identify a time window that selects a data subset where the condition of constant rotation vector and resolution are optimal. It is worth highlighting that the criterion of optimality here adopted is based on best image focus.

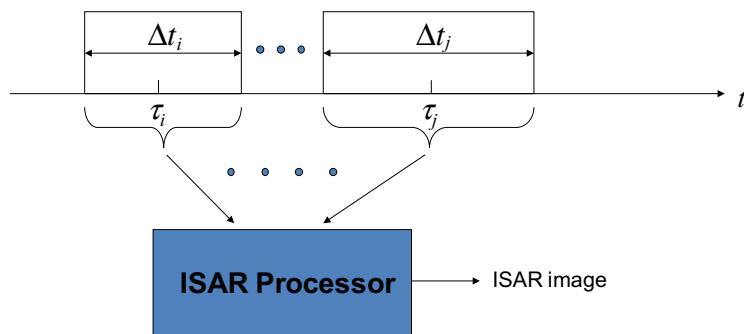


FIGURE 19.18

Time-windowing concept.

Therefore, the optimal time window position and length are obtained by maximizing the IC with respect to the couple $(\tau, \Delta t)$. Therefore, the following optimization problem can be formulated

$$(\tau_{\text{opt}}, \Delta t_{\text{opt}}) = \arg \max_{(\tau, \Delta t)} [IC(\tau, \Delta t)], \quad (19.40)$$

where the $IC(\tau, \Delta t)$ is obtained by redefining (19.27) with respect the two new time variables.

It must be noticed that the variables $(\tau, \Delta t)$ are discrete variables as the data in input to an ISAR processor is digitised. Therefore, the problem in (19.40) is a discrete optimization problem. Specifically, such a problem can be classified as a non-linear Knapsack Problem [14]. The solution proposed in [13] is based on a double linear search, which can be briefly defined as follows:

1. Maximization of the contrast with respect to τ for a given guess $\Delta t^{(\text{in})}$. Let τ_{opt} be the solution of such maximization.
2. Optimization with respect to Δt with $\tau = \tau_{\text{opt}}$.

This procedure is depicted in Figure 19.19 for the sake of clarity.

The justification of this procedure is heuristic. It can be observed both in simulated and real ISAR data of several types of targets that the position of the optimal time-window is quite independent of the window length. This means that the IC peak position along the τ axis does not change by changing the window length. Physically, this can be explained by the fact that the target's own motion will be characterized by regular motions at given times and less regular motions at other times. The reader may think of a ship that undergoes pitch and roll due to the sea surface waves. A regular motion may be disturbed by an incoming wave and generate complex motions which cause the rotation vector to rapidly change.

An example is provided in Figure 19.20 where the IC is calculated by moving four fixed length windows along the time axis τ . It can be noticed that the position of the peaks is practically the same for all the four windows.

As an example, we apply the technique of optimal time-windowing to an ISAR dataset of an airplane, namely a boeing 737. The results are shown in Figure 19.21, where three images are displayed. The first image (Figure 19.21a) is obtained by processing a short CPI dataset ($\Delta t = 0.41$ s). It is quite evident that, although the image is well focussed, the resolution is poor. The second image (Figure 19.21b) shows the case of long CPI ($\Delta t = 3.27$ s). The result shows potential finer resolution but at the same time a strong defocusing effect produced by the target's rotation vector time-variance. The third image (Figure 19.21c) shows the result obtained by automatically selecting the optimal time-window, which resulted in a value of the CPI equal to $\Delta t = 0.87$ s. The conclusion is that the image produced by means of the automatic optimal time window selection shows fine details due to a fine resolution whilst retaining a good level of image focus.

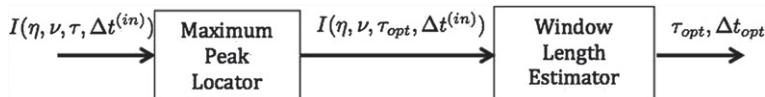
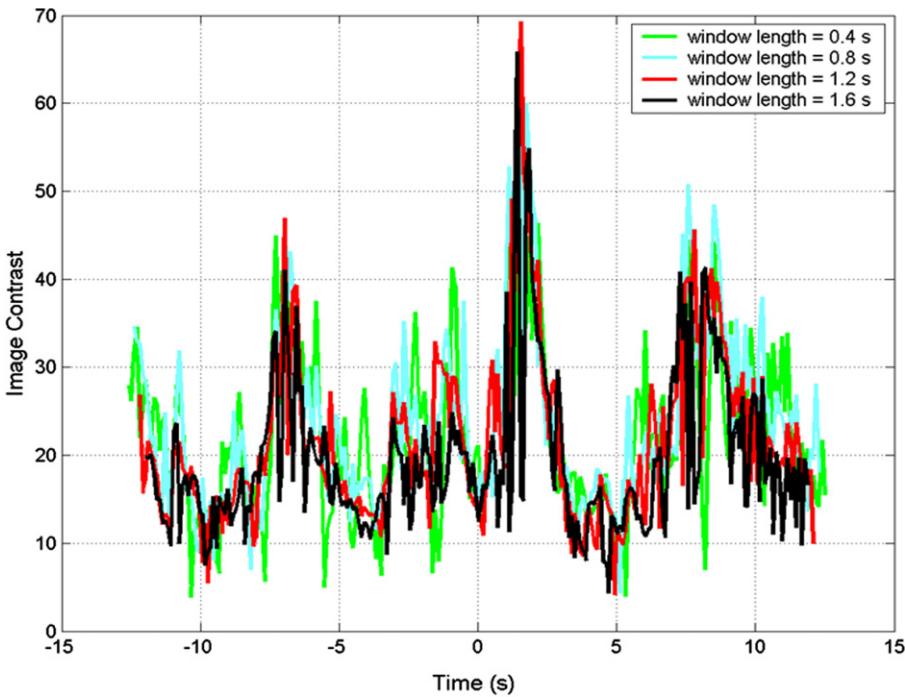


FIGURE 19.19

Optimal time window estimator.

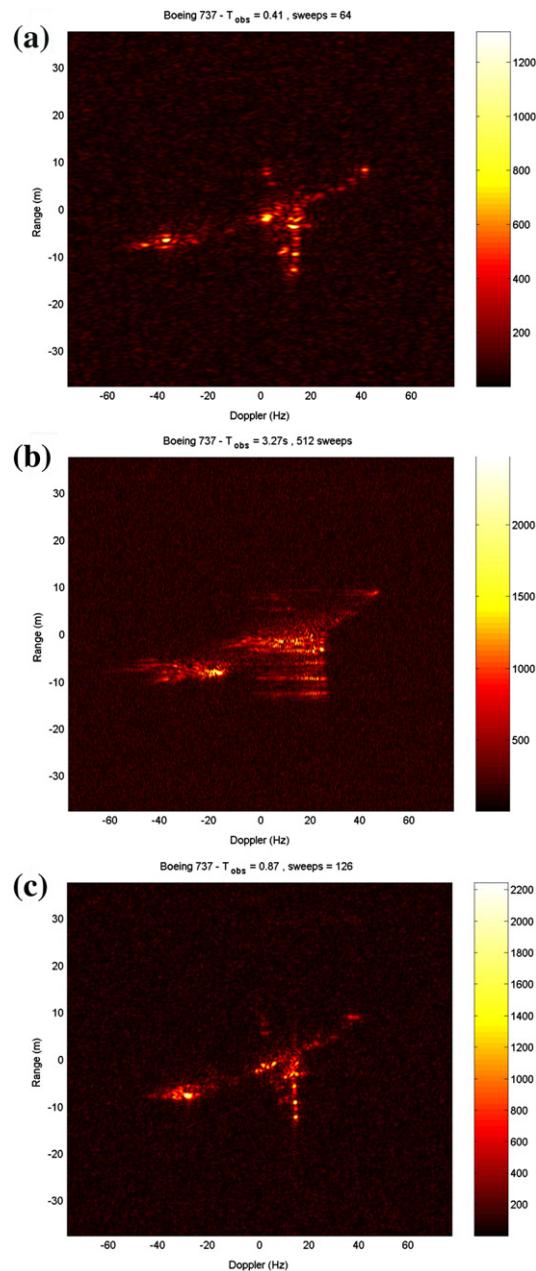
**FIGURE 19.20**

IC as a function of time by using four different time window lengths.

Often, a sequence of ISAR images is used for classifying or recognizing a target, as the information about the target increases when the number of images used increases. A direct consequence of the use of multiple images, especially in the case of highly maneuvering or oscillating targets, is the change in the IPP, which leads to different target projections. This typically leads to a better indication about the geometrical features of the target, which can then be used for classification purposes. To this purpose, it would be desirable to obtain a number of well focussed ISAR images out of a long target observation time. An extension of the time-windowing approach that provides multiple well-focussed images can be obtained by selecting multiple peaks from the maximum peak locator and iteratively find optimal time window length for each of the selected peaks. The implementation of this extension is straightforward and is not worth describing in details.

2.19.9 Image scaling

Inverse Synthetic Aperture Radar generates two dimensional, high resolution images of targets in the time delay—Doppler domain. In order to determine the size of the target, it is required to have fully scaled image. The range scaling can be performed by using the well known relationship $r = c\eta/2$, where r is the slant range coordinate and η is the time delay. On the other hand, cross range scaling requires the estimation of the modulus of the target effective rotation vector.

**FIGURE 19.21**

ISAR image obtained by processing (a) a short CPI dataset (0.41 s), (b) a long CPI dataset (3.27 s), (c) the optimal CPI (0.87 s).

In this subsection, we will illustrate an algorithm that is based on the assumption of quasi-constant target rotation [15]. When the target's rotation vector can be assumed constant within the CPI, the chirp rate produced by the scattering centers can be related to the modulus of the target's effective rotation vector by means of an analytical expression. Therefore, each scattering center carries information about the modulus of the target's rotation vector through its chirp rate.

To demonstrate this, we will consider the radial motion compensated signal in (19.14). The phase term inside the integral function can be approximated with a second order Taylor polynomial, as follows:

$$\varphi(f, t; x_1, x_2) = -\frac{4\pi f}{c} \left(x_2 + x_1 \Omega_{\text{eff}} t - \frac{1}{2} x_2 \Omega_{\text{eff}}^2 t^2 \right). \quad (19.41)$$

It should be noted that the phase in (19.41), represents the phase of a chirp signal, where the second order term coefficient is usually referred to as *chirp rate*. The echo of an ideal scatterer located in (\bar{x}_1, \bar{x}_2) , with finite reflectivity function $\tilde{\xi}(\bar{x}_1, \bar{x}_2)\delta(x_1 - \bar{x}_1, x_2 - \bar{x}_2)$ can be written by approximating its phase with the expression in (19.41), as follows:

$$S_C(f, t; \bar{x}_1, \bar{x}_2) = W(f, t) \tilde{\xi}(\bar{x}_1, \bar{x}_2) \exp \left(-j \frac{4\pi f}{c} \left(\bar{x}_2 + \bar{x}_1 \Omega_{\text{eff}} t - \frac{1}{2} \bar{x}_2 \Omega_{\text{eff}}^2 t^2 \right) \right). \quad (19.42)$$

The range compressed profile can be obtained by applying a FT to the signal in (19.42) along the coordinate f . This can be mathematically expressed as follows:

$$R_p(\eta, t; \bar{x}_1, \bar{x}_2) = B \tilde{\xi}(\bar{x}_1, \bar{x}_2) \text{sinc} \left(B \left(\eta - \frac{2}{c} \bar{x}_2 \right) \right) \text{rect} \left(\frac{t}{\Delta t} \right) \exp \left(-j \frac{4\pi f_0}{c} \left(\bar{x}_2 + \bar{x}_1 \Omega_{\text{eff}} t - \frac{1}{2} \bar{x}_2 \Omega_{\text{eff}}^2 t^2 \right) \right). \quad (19.43)$$

If a method for perfectly estimating the chirp rate of a given scatterer was directly available, the following equation could be written:

$$m = \frac{2f_0}{c} x_2 \Omega_{\text{eff}}^2. \quad (19.44)$$

Therefore, as the scatterer's range coordinate x_2 can be readily obtained by measuring the delay-time η , the effective rotation vector can be obtained by inverting Eq. (19.44), as follows:

$$\Omega_{\text{eff}} = \sqrt{\frac{c}{2f_0 x_2}} m. \quad (19.45)$$

In practice, a scatterer's chirp rate, as well as its range, must be estimated from the received data. Therefore, the estimation of the effective rotation vector magnitude would in general be affected by an error.

Techniques for estimating target's scattering center chirp rates have been proposed that make use of atomic decomposition [16], CLEAN technique [17, 18] and based on the IC method proposed in [15].

To make the estimation more accurate and robust the chirp rates of a number of target's scatterers can be measured together with their ranges. The problem of estimating the effective rotation vector

magnitude is then transformed into a problem of estimating the slope of a straight line that fits the scatterplot generated by the set of range and chirp rate estimates. One way of solving this problem is to apply a LSE approach [19]. The mathematical problem can be set as follows:

$$m_k = ax_{2k} + \epsilon_k, \quad (19.46)$$

where $a = \frac{2f_0}{c}\Omega_{\text{eff}}^2$ and m_k , x_{2k} , and ϵ_k are the chirp rate estimate, the range estimate and the estimation error for the k th scatterer, respectively. The LSE problem and its solution for the estimation of a is given in (19.47)

$$\hat{a} = \arg \min_a \sum_{k=1}^N \epsilon_k^2 = \frac{N \sum_{k=1}^N \hat{m}_k x_{2k} - \sum_{k=1}^N \hat{m}_k \sum_{k=1}^N x_{2k}}{N \sum_{k=1}^N x_{2k}^2 - \left(\sum_{k=1}^N x_{2k} \right)^2}. \quad (19.47)$$

An example is provided in Figure 19.22, where the range vs chirp rate scatterplot is shown for a large ship. The linear relationship between range and chirp rate, theoretically predicted by Eq. (19.46), is quite evident.

The fully scaled ISAR image of the ship is shown in Figure 19.23. The full representation of the ISAR image in spatial coordinates allows measuring important features, e.g., ship's length, directly from the ISAR picture.

To conclude this section, we provide a few remarks regarding effective applicability of this cross-range scaling technique. First of all, as any other cross-range scaling technique, its application is effective only if a constant effective rotation vector applies during the CPI. In fact, such a condition is necessary to establish a linear relationship between Doppler and cross-range coordinates. Moreover, in the specific case of the technique discussed in this section, a well focussed image must be produced in order to make sure that quadratic phases are associated with single scatterer's motions. The number of scattering centers also play an important role as the accuracy of the effective rotation vector estimation generally improves when the number of independent scattering centers increases.

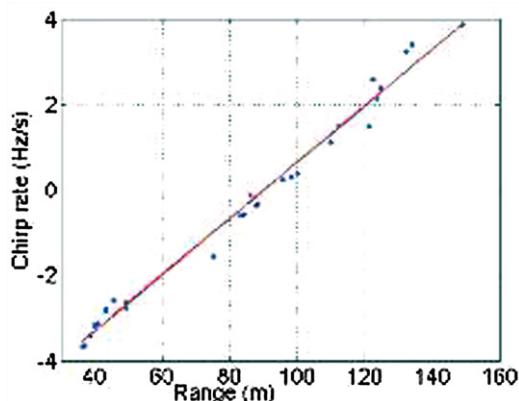
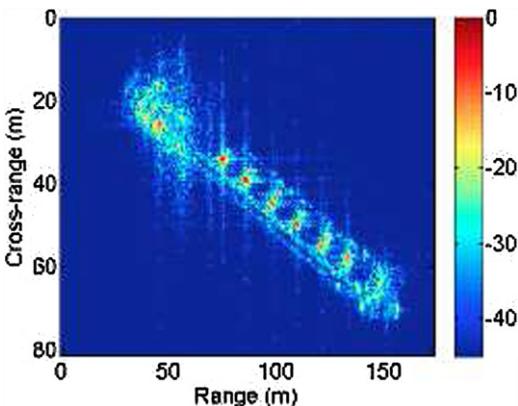


FIGURE 19.22

Range vs. chirp rate scatterplot.

**FIGURE 19.23**

Fully scaled ISAR image.

2.19.10 Time-frequency image formation

The RD technique is based on the assumption that the Doppler frequency of each scatterer, relative to point O, is constant during the observation time. This hypothesis is usually valid for low spatial resolution (of the order of one meter) and when the target does not undergo fast maneuvers and/or is affected by significant oscillating motions, such as pitch, roll and yaw (typical of sea vessels). When very high spatial resolutions are required (of the order of ten centimeters), typically, a longer integration time is needed and the Doppler frequency associated with each scatterer becomes time-varying. The situation is also aggravated when the target maneuvers or when it undergoes angular motions, as in the case of ships. In these cases, the target's rotation vector is not constant and the RD technique fails because of the spreading effects due to the time-varying frequency of each scatterer's contribution. To solve this problem, the Fourier approach employed by the RD technique is replaced by Time Frequency Transforms (TFTs), which are suitable for the analysis of non-stationary signals.

Specifically, bilinear TFTs, such as those described by the Cohen's class [20], prove effective when dealing with ISAR signals, which, up to some extent, can be approximated with chirp-like signals (second order phase terms).

Pioneering work in this sense was delivered by Victor Chen, which is mainly collected in [21], although more work in this area followed [22,23]. In this subsection, we will recall the work done in [23], as it aims to analytically derive the ISAR image PSF in when using bilinear TFTs. The PSF is derived in the case of Wigner-Ville transform, which is the basic Cohen's class TFT. The derivation in the case of all the other TFTs can be obtained by simply filtering the data in the adjunct Fourier domain, which is the Ambiguity Function domain, as demonstrated in [20].

In order to derive the PSF, we will consider the signal model in (19.42), where the motion compensated received signal relative to a single ideal scatterer located in the coordinates (\bar{x}_1, \bar{x}_2) is considered. It is

worth noting that a simple RD approach would lead to a smeared image reconstruction along the cross-range (Doppler) coordinate, due to the presence of a quadratic phase (chirp term).

The Cohen's class generic TFT can be analytically expressed as follows:

$$\text{CTFT}(t, v) = \iiint K(\theta, \tau) s^* \left(u - \frac{\tau}{2} \right) s \left(u + \frac{\tau}{2} \right) \cdot \exp(-j\theta t) \exp(-j2\pi v \tau) \exp(-j\theta u) du d\tau d\theta, \quad (19.48)$$

where, $s(t)$ is the signal to be transformed and $K(\theta, \tau)$ is the transform kernel, which defines any specific bilinear TFT belonging to the Cohen's class.

The Wigner-Ville (WV) is a particular TFT that can be obtained from the Cohen's class by posing $K(\theta, \tau) = 1$. The WV is analytically expressed in (19.49)

$$\text{WV}(t, v) = \iiint s^* \left(u - \frac{\tau}{2} \right) s \left(u + \frac{\tau}{2} \right) \exp(-j2\pi v \tau) d\tau. \quad (19.49)$$

After applying (19.49) to the range compressed signal in (19.43) we obtain a data cube where for each range cell a time-frequency representation of the data is obtained. With respect to the Fourier Transform, this approach has the advantage of capturing the time-varying signal spectrum. From an ISAR imaging perspective, it is important to note that for each time instant t , an ISAR image can be obtained in the time-delay (η) and Doppler (v) domain. The ISAR image obtained for the time instant $t = 0$ is shown in (19.50)

$$\begin{aligned} I_{\text{TFT}}(\eta, v) &= I_{\text{FFT}}(\eta, t, v)|_{t=0} \\ &= \left| B \bar{\xi}(\bar{x}_1, \bar{x}_2) \text{sinc} \left(B \left(\eta - \frac{2}{c} \bar{x}_2 \right) \right) \right|^2 \exp(-j2\pi f_0 \eta) \text{sinc} \left(2\Delta t \left(v - \frac{2f_0}{c} \Omega_{\text{eff}} \bar{x}_1 \right) \right). \end{aligned} \quad (19.50)$$

By looking at the analytical expression in (19.50), it is worth pointing out that

- the effect of chirp-like terms is completely canceled out as an effect of the application of a bilinear transform, which has the ability of canceling all even terms, including the quadratic term,
- the bilinear transform produces a square value in the sinc function along the delay-time coordinate,
- the Doppler resolution is halved with respect to the Fourier approach.

As a side effect, cross-terms are typically introduced by bilinear transforms, therefore resulting in fake target's scatterer images. This effect could be very detrimental in target recognition by means of ISAR images as false scatterers may appear in critical positions. Specific transform kernels $K(\theta, \tau)$ have the property of removing or attenuating the cross-terms. Nevertheless, a trade-off between cross-terms removal and Doppler resolution loss must be found in the selection of the kernel function. A suitable kernel function can be designed as a product of two single variable functions, as follows:

$$K(\theta, \tau) = F(\theta)G(\tau), \quad (19.51)$$

where $F(\theta)$ and $G(\tau)$ may be two weighting windows, such as Hamming, Kaiser, and so on. The shape parameters in such weighting functions can be determined based on the required level of cross-term suppression and Doppler resolution loss [23]. The choice of the kernel function such as expressed in (19.51) leads to the definition of the Smoothed Pseudo Wigner Ville (SPWV), which is typically

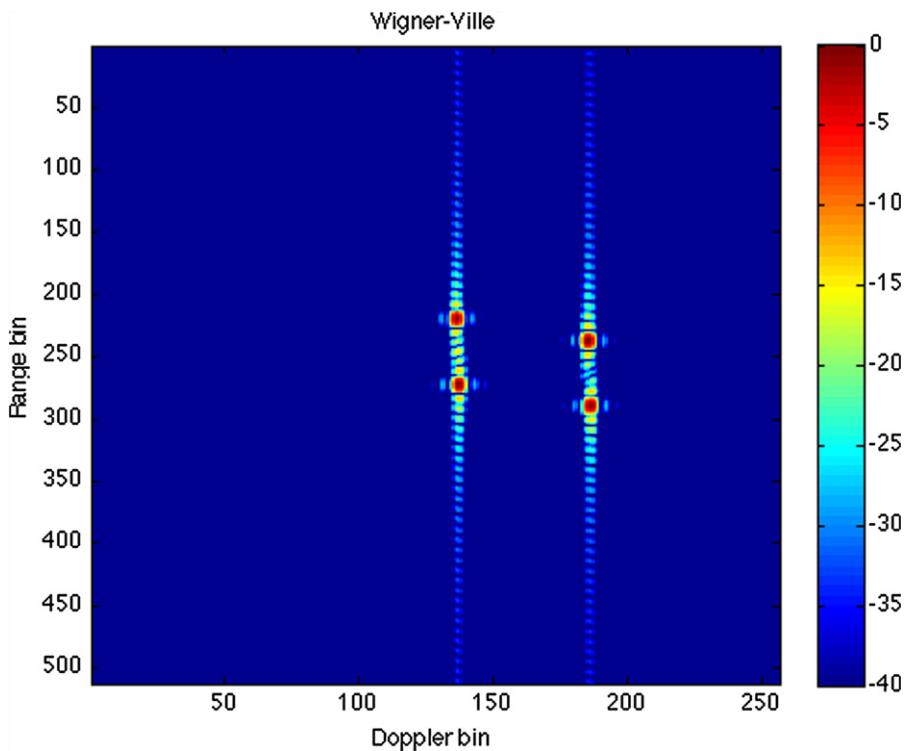


FIGURE 19.24

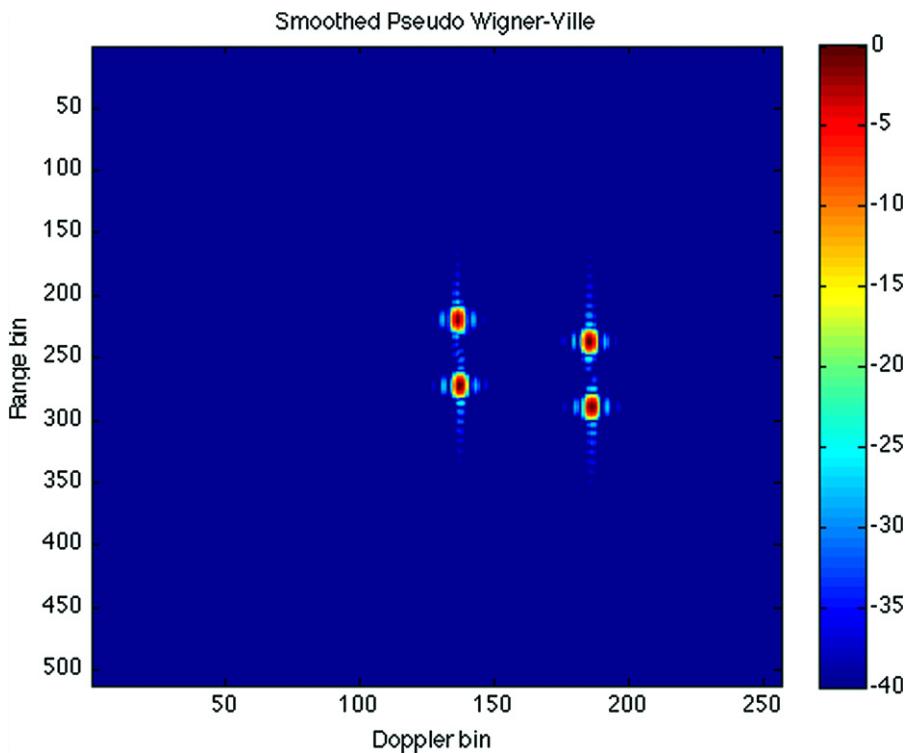
ISAR image obtained by applying the Wigner-Ville TFT.

implemented in ISAR applications as it provides a flexible solution to the problem of cross-terms and resolution.

An example of the application of weighting windows is shown in Figures 19.24 and 19.25, where the results obtained by applying the WV and the SPWV are shown. In particular, the SPWV is applied by selecting a Kaiser window with shape parameter equal to $K = 3$. The reader should easily note that cross-terms present in Figure 19.24 tend to disappear in Figure 19.25, although the resolution tends to get worse. A trade-off between cross-terms reduction and resolution loss must always be taken into account when using TFT applied to ISAR imaging.

2.19.11 Polarimetric ISAR (Pol-ISAR)

Firstly a brief introduction on polarimetric SAR systems will be given to provide some motivations for using polarimetric radar imaging. As polarimetric radar imaging systems are necessarily more expensive to build than traditional single polarization radar imaging system, their reason to exist must be supported

**FIGURE 19.25**

ISAR image obtained by applying the SPWV.

by an increase in performance or in their ability to overcome physical limitations of single polarization sensors.

2.19.11.1 Polarimetric SAR imaging

Polarimetric Synthetic Aperture Radar (Pol-SAR) has been widely used for classifying imaged areas. More specifically, Polarimetric Synthetic Aperture Radar (Pol-SAR) systems are used in land, ice and ocean remote sensing to obtain extra information about scattering mechanisms that can be exploited for extracting physical parameters of interest [24–26] as well as in target classification applications [27–32]. Pol-SAR systems can be realized by exploiting fully polarimetric radars. A fully polarimetric radar is able to measure a scattering matrix rather than a reflectivity function, as in the case of single polarization radar. The canonical representation of the scattering matrix is as follows:

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix}, \quad (19.52)$$

where each term S_{ij} is the scattering coefficient obtained by transmitting polarization i and by receiving polarization j , where i and j may be horizontal (H) or vertical (V) polarization. Since a generic target

responds differently depending on which polarization is used in transmission and reception, the information contained in a scattering matrix is more complete than that contained in a single polarization scattering coefficient. The fully polarimetric information contained in the scattering matrix may be used to estimate physical parameters from PolSAR images, which cannot be estimated with a single polarization radar.

A PolSAR image can be interpreted as the result of a multichannel SAR system where the number of available channels is equal to four, one for each element of the scattering matrix. The result, is that for each image pixel, a complex four-element vector is available that can be seen as a polarimetric signature of the imaged area represented by the same pixel.

PolSAR image formation does not differ from single polarization SAR image formation as all four images must be formed by using the same processing. In this way, the geometrical co-registration among all channels is automatically sorted out. Nevertheless, some improvement in the image auto-focus processing can be attained by optimally processing all polarimetric data, as it will be shown in Section 2.19.11.2.

As the image formation does not require extra attention with respect to single polarization SAR, most of the effort has been spent in finding an optimal use of the polarimetric information contained in a PolSAR image. For this reason, several polarimetric decomposition have been introduced that aim at helping to extract useful information that may be used to estimate physical parameters of interest. A direct information that can be inferred from the fully polarimetric signature relates to the target's shape. Therefore, having a polarimetric signature for each image scatterer, allows identifying shapes with a resolution that can be equal to the SAR image resolution. This allows separating signal components that are the result of different types of scattering that originate from different scatterer's shapes.

Under the validity of the reciprocity theorem, which corresponds to the physical condition of reciprocal media, it can be demonstrated that a fully polarimetric signal can be represented by means of three orthogonal complex vectors (three dimensional complex base). Generally, PolSAR images are represented by coding each decomposed polarimetric channel with a color. A typical color base is the Red-Green-Blue (RGB).

The polarimetric decompositions that have been introduced can be categorized into two groups: Coherent Decomposition (CD) and Incoherent Decomposition (ID). CDs are employed in those cases where the imaged target is characterized by coherent scattering. This typically happens when single dominant scatterers are present in a resolution cell. In this case, the scattering matrix \mathbf{S} in (19.52) is able to completely characterize the scattering mechanism. On the other hand, IDs are employed to characterize incoherent scattering, which is typical of distributed scatterers. This is the case of the presence of a number of non-dominant scatterers in a resolution cell. The most used CDs are the Pauli's decomposition [33], the Krogager's decomposition [34] and the Cameron's decomposition [27], whereas the most commonly used IDs are the Huynen's Phenomenological decomposition [35], the Cloude-Pottier decomposition [36] and the Freeman's decomposition [37]. A comprehensive tutorial on PolSAR imaging is available at [<http://earth.eo.esa.int/polsarpro/tutorial.html>].

2.19.11.2 Polarimetric ISAR imaging

As in the case of SAR, fully polarimetric radar can be exploited to obtain polarimetric ISAR imagery, with the aim of improving target classification and recognition performance. The first idea to exploit

fully polarimetric information in ISAR imaging was developed in [38], where the ISAR image autofocus is improved by introducing full polarization in the derivation of the autofocus technique. Two algorithms from the class of parametric autofocusing techniques, namely the Image Contrast Based Autofocusing (ICBA) and Image Entropy Based Autofocusing (IEBA) [6,39–44] have been extended by introducing the full polarimetric information contained in the received data. The Image Contrast (IC) and the Image Entropy (IE) represent two ways of measuring the focus of an ISAR image. In the case of a single polarization ISAR system, the success of maximizing the IC or minimizing the IE, which is also used to achieve motion compensation, depends on the polarization used by the system. Evidently, specific scatterers may produce a more stable signal return in a given polarization. Often, in SAR applications, the HH and VV co-polarization channels offer a higher SNR with respect to the cross-polarizations (cross-pol) HV and VH. Nevertheless, in ISAR applications, such a priori knowledge cannot be taken for granted. In any case, the availability of all polarizations provides the basis for optimizing the autofocusing algorithm with respect to the polarimetric space.

Polarimetric radars may also maximize the Signal to Noise Ratio (SNR) with respect to the polarization space in order to improve detection performance [45]. The concept of increasing the performance by finding an optimal polarization can be extended to the ISAR image autofocus problem. Similarly to the maximization of the SNR with respect to the radar polarization, one may argue that a polarization exists that maximizes the image focus. This insight can be justified by considering that the image focus strongly depends on the time invariance of the scatterer contributions. Moreover, the Doppler components for each scatterer are generally modulated due to several causes. Among such causes, are target-radar dynamics, modulation induced by scatterer scintillation and the effect of noise. These causes can be reduced by exploiting full polarization. In fact, both the SNR and the modulation effects induced by scatterers when illuminated from different aspect angles can be jointly reduced by finding the optimal polarization. The definition of such an optimality criterion will be defined in Section 2.19.11.2.3. To set the scene, we firstly introduce the signal model and calculate the polarimetric ISAR image PSF in the next subsection.

2.19.11.2.1 Signal model

The polarimetric matrix of the received signal, in free space conditions, can be written in the time-frequency domain by extending the signal model defined in [6]:

$$\begin{aligned} \mathbf{S}_R(f, t) = & W(f, t) \exp \left[-j \frac{4\pi f}{c} R_0(t) \right] \\ & \cdot \int_{\text{Target}} \boldsymbol{\xi}(\mathbf{x}) \exp \left\{ -j \frac{4\pi f}{c} [\mathbf{x}^T \cdot \mathbf{i}_{LoS}(t)] \right\} d\mathbf{x}. \end{aligned} \quad (19.53)$$

The polarimetric matrix of the received signal can be expressed as

$$\mathbf{S}_R(f, t) = \begin{bmatrix} S_R^{HH}(f, t) & S_R^{HV}(f, t) \\ S_R^{VH}(f, t) & S_R^{VV}(f, t) \end{bmatrix}$$

and the scattering matrix

$$\boldsymbol{\xi}(\mathbf{x}) = \begin{bmatrix} \xi^{HH}(\mathbf{x}) & \xi^{HV}(\mathbf{x}) \\ \xi^{VH}(\mathbf{x}) & \xi^{VV}(\mathbf{x}) \end{bmatrix}.$$

Before proceeding, it is convenient to use a different notation, as detailed in [33], and exploit the characteristics of isotropic media that are encountered in ISAR applications. Therefore, the polarimetric data that represents the received signal can be written according to Pauli's decomposition:

$$\mathbf{S}_R = \frac{1}{\sqrt{2}} \begin{bmatrix} S_R^{VV} + S_R^{HH}, & S_R^{VV} - S_R^{HH}, & 2S_R^{HV} \end{bmatrix}^T, \quad (19.54)$$

where the dependence on (f,t) is omitted for notation simplicity.

The same decomposition applies for the target scattering matrix. Therefore, the scattering vector obtained from the scattering matrix is

$$\boldsymbol{\xi}(\mathbf{x}) = [\xi^{VV}(\mathbf{x}) + \xi^{HH}(\mathbf{x}), \xi^{VV}(\mathbf{x}) - \xi^{HH}(\mathbf{x}), 2\xi^{HV}(\mathbf{x})]. \quad (19.55)$$

Thus, the received signal can be seen as a vector in a complex three-dimensional polarimetric space. All possible projections can be obtained by means of an internal product between the received signal vector and a generic polarization vector \mathbf{p} :

$$S_R^{(\mathbf{p})} = \mathbf{S}_R \cdot \mathbf{p}, \quad (19.56)$$

where vector \mathbf{p} can be expressed according to the decomposition introduced by Cloude and Papathanassiou in [46]:

$$\begin{aligned} \mathbf{p} &= \frac{1}{\sqrt{2}} \begin{bmatrix} p^{VV} + p^{HH}, & p^{VV} - p^{HH}, & 2p^{HV} \end{bmatrix} \\ &= \begin{bmatrix} \cos \alpha \exp(j\varphi) \\ \sin \alpha \cos \beta \exp(j\delta) \\ \sin \alpha \sin \beta \exp(j\gamma) \end{bmatrix}, \end{aligned} \quad (19.57)$$

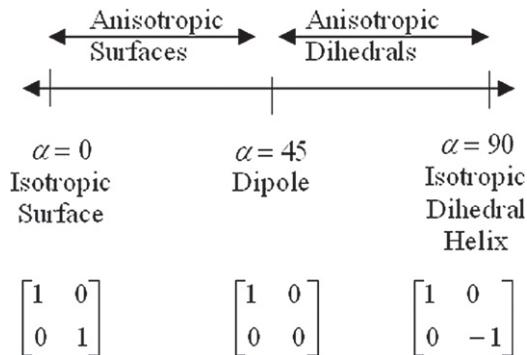
where:

- α is the scatterer's internal degree of freedom, which ranges in the interval $[0, 90]^\circ$. The meaning of such an angle is related to the scattering properties of the target, e.g., for an ideal dipole the value of α is equal to 45° (see Figure 19.26).
- β represents a physical rotation of the scatterer on the plane perpendicular to the e.m. wave propagation direction.
- φ, δ, γ are the scatterer phases of the three polarimetric components.

It is worth noting that such a representation is meant to highlight the physical properties of the scattering mechanism induced by a given scatterer. Therefore, by defining the unit vector \mathbf{p} , it is possible to define a specific polarization that resonates with a scatterer with given physical properties. Moreover, the decomposition proposed in [46] provides, among other polarimetric decompositions, a suitable domain for IC max and/or IE min search. Nevertheless, other types of decomposition could also be used for the same purpose.

2.19.11.2.2 Image formation

The definition of the steps that lead to the image formation will be carried out without including the noise contribution, which will be added subsequently for performance analysis. It is worth pointing out that this is as a standard procedure in SAR/ISAR processing.

**FIGURE 19.26**

Interpretation of the internal degree of freedom α .

Motion compensation consists of removing the phase term $\exp\left[-j\frac{4\pi f}{c}R_0(t)\right]$ due to the radial movement of the focusing point O. The noiseless received signal after perfect motion compensation can be written as follows:

$$\mathbf{S}_C^{(\mathbf{p})}(f, t) = W(f, t) \int_V \xi^{(\mathbf{p})}(\mathbf{x}) \exp\left\{-j\frac{4\pi f}{c}\left[\mathbf{x}^T \cdot \mathbf{i}_{LoS}(t)\right]\right\} d\mathbf{x}, \quad (19.58)$$

where $\zeta^{(\mathbf{p})}(\mathbf{z}) = \zeta(\mathbf{z}) \cdot \mathbf{p}$.

It is worth pointing out that the signal in (19.58) is a scalar signal and therefore, it has the same characteristics of the single polarization signal expressed in (19.11). Therefore, ISAR image formation can be carried out exactly at the same way as is done for single polarization ISAR. The Pol-ISAR image can be written as:

$$I_C^{(\mathbf{p})}(\eta, v) = 2D - IFT\left[S_C^{(\mathbf{p})}(f, t)\right] = K w(\eta, v) \otimes \otimes \xi'^{(\mathbf{p})}(\eta, v).$$

2.19.11.2.3 Polarimetric autofocus

The idea of jointly processing all polarimetric channels for obtaining highly focussed ISAR images was firstly introduced in [47]. Such an insight relied on the concept of enhancing the image focus by maximizing the IC over the joint space of the focusing parameters α and of the polarization \mathbf{p} . In formula:

$$(\hat{\alpha}_{IC}, \hat{\mathbf{p}}_{IC}) = \arg \max_{\alpha, \mathbf{p}} \{IC(\alpha, \mathbf{p})\}, \quad (19.59)$$

where

$$IC(\alpha, \mathbf{p}) = \frac{\sqrt{A \left\{ [I^{(\mathbf{p})}(\eta, v, \alpha) - A \{I^{(\mathbf{p})}(\eta, v, \alpha)\}]^2 \right\}}}{A \{I^{(\mathbf{p})}(\eta, v, \alpha)\}} \quad (19.60)$$

and where $\xi = [\alpha_1, \dots, \alpha_N]$, with α_i the model polynomial coefficients.

Equation (19.60) represents the new image contrast function defined in the joint domain and where $A(\cdot)$ is the mean operator. It is worth noting that the IC can be interpreted as a normalized standard deviation. Therefore, higher values of the IC mean sharper images.

In the same way, image focus can be enhanced by minimizing the Image Entropy (IE), as follows:

$$(\hat{\alpha}_{\text{IE}}, \hat{\mathbf{p}}_{\text{IE}}) = \arg \min_{\alpha, \mathbf{p}} \{\text{IE}(\alpha, \mathbf{p})\}, \quad (19.61)$$

where

$$\text{IE}(\alpha, \mathbf{p}) = \iint \ln[\bar{I}^{(\mathbf{p})}(\eta, v, \alpha)] \bar{I}^{(\mathbf{p})}(\eta, v, \alpha) d\nu d\eta, \quad (19.62)$$

$$\bar{I}^{(\mathbf{p})}(\eta, v, \alpha) = \frac{|I^{(\mathbf{p})}(\eta, v, \alpha)|^2}{A[|I^{(\mathbf{p})}(\eta, v, \alpha)|^2]}. \quad (19.63)$$

2.19.11.2.4 Initialisation

In order to proceed with the application of the Pol-ICBA and Pol-IEBA to the fully polarimetric ISAR data, a solution for the initial polarization vector must also be provided. The problem can be solved by means of the following algorithm:

1. The polarization vector that provides the maximum SNR is obtained by solving the optimization problem stated by Eq. (19.64).

$$\hat{\mathbf{p}}_M = \arg \max_{\mathbf{p}} \left(\frac{\iint |S_R^{(\mathbf{p})}(f, t)|^2 df dt}{\iint |N_R^{(\mathbf{p})}(f, t)|^2 df dt} \right). \quad (19.64)$$

The SNR can be assumed maximum when the signal energy reaches its maximum, provided that the noise level is the same in all the polarization channels (basically when the noise level in the H and V receiving channels are the same). Therefore, Eq. (19.64) can be simplified as follows:

$$\hat{\mathbf{p}}_M = \arg \max_{\mathbf{p}} \left(\iint |S_R^{(\mathbf{p})}(f, t)|^2 df dt \right) \quad (19.65)$$

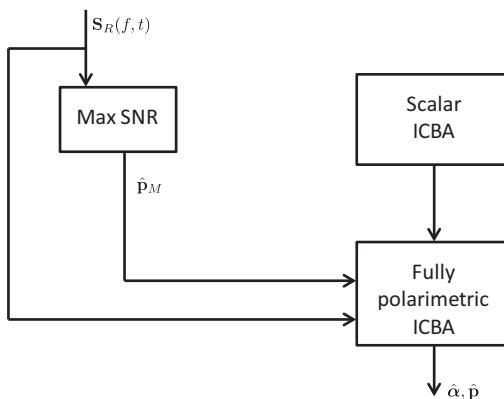
An initial guess for the focusing parameter vector ξ can be obtained by applying the Radon Transform and by running a one-dimensional optimization problem (see [6,8]). Specifically, the scalar ICBA and IEBA must be applied to the received signal with polarization $\hat{\mathbf{p}}_M$ as found at step 1:

$$\left(\hat{\alpha}_{\text{IC}}^{(\hat{\mathbf{p}}_M)} \right) = \arg \max_{\alpha} \{ \text{IC}(\alpha, \hat{\mathbf{p}}_M) \}, \quad (19.66)$$

$$\left(\hat{\alpha}_{\text{IE}}^{(\hat{\mathbf{p}}_M)} \right) = \arg \min_{\alpha} \{ \text{IE}(\alpha, \hat{\mathbf{p}}_M) \}, \quad (19.67)$$

where $\text{IC}(\alpha, \hat{\mathbf{p}}_M)$ and $\text{IE}(\alpha, \hat{\mathbf{p}}_M)$ can be obtained from (19.60) and (19.62)

Therefore, the initial guess can be formed by adjoining the polarization vector $\hat{\mathbf{p}}_M$ to the focusing parameter vector $\hat{\alpha}_{\text{IE}}^{(\hat{\mathbf{p}}_M)}$, which can be expressed as $(\hat{\alpha}_{\text{IE}}^{(\hat{\mathbf{p}}_M)}, \hat{\mathbf{p}}_M)$.

**FIGURE 19.27**

Polarimetric ICBA flow chart.

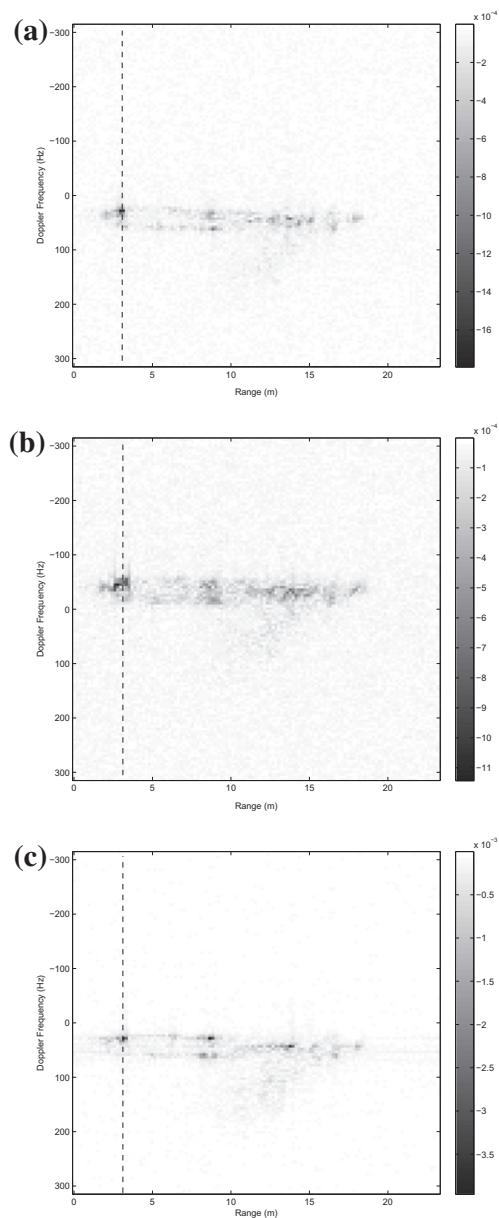
2.19.11.2.5 Optimization

Once the initial guess is estimated, the optimization problems as stated in (19.59) and (19.61) can be solved iteratively by using numerical methods for maximum (or minimum) search. Several methods for solving optimization problems have been proposed in the literature. Such methods can be grouped into deterministic and statistical methods. The first type, which will be adopted in this paper, make use of the characteristics of the cost function, such as gradient, hessian, etc. to determine the next step in order to converge to the local maximum or minimum. In this case, convergence to the global maximum (minimum) must be ensured by a suitable choice of the initial guess. The method used here is the simplex method, proposed by Nelder and Mead in [7]. Nevertheless, other solutions may be obtained by using statistical methods, such as Genetic Algorithms (GA) [8]. For the sake of clarity the algorithm flow chart is shown in Figure 19.27.

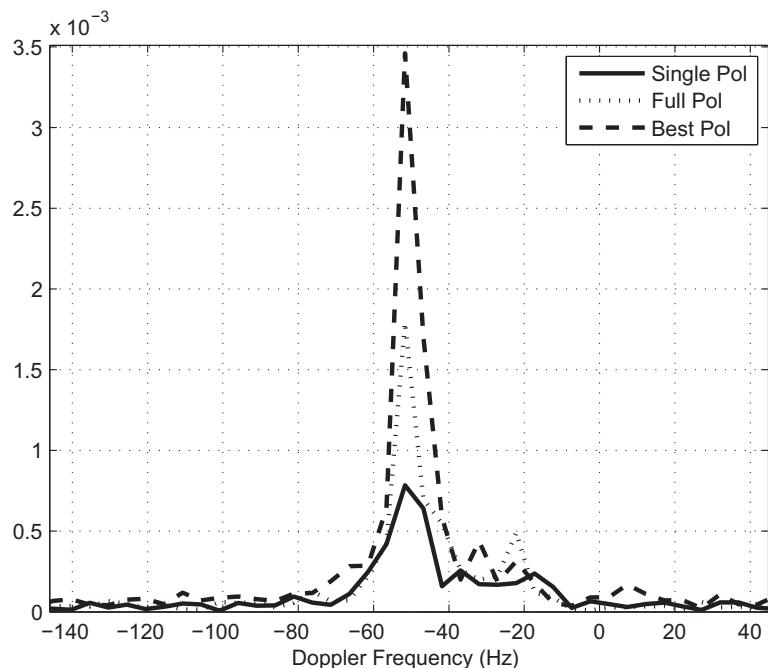
2.19.11.2.6 Most focussed ISAR image

The proposed algorithm also provides the most focused ISAR image. The polarization \hat{p}_{IC} , which maximizes (19.59) or equivalently, the polarization \hat{p}_{IE} , which minimizes (19.61), are obtained as part of the solution of the optimization problems. Therefore, ISAR images obtained by processing the received data in polarizations \hat{p}_{IC} and \hat{p}_{IE} represent the best focused images according to the IC and IE focus indicators, respectively.

Results relative to the use of polarimetric ISAR image autofocus are provided in Figure 19.28. Specifically, the VV polarization ISAR image obtained by means of Pol-ICBA (in Figure 19.28a) is compared with the ISAR image obtained by applying a single polarization ICBA algorithm (in Figure 19.28b). It is quite evident that the image obtained by using the Pol-ICBA is sharper than that obtained with the single polarization ICBA. For the sake of completeness, also the ISAR image projected onto the polarization that maximizes the IC is shown in Figure 19.28c. A section cut along the cross-range in

**FIGURE 19.28**

Pol-ISAR images obtained by applying (a) Pol-ICBA (VV-Channel), (b) single polarization ICBA (VV-Channel), (c) Pol-ICBA (polarimetric channel that maximizes the IC).

**FIGURE 19.29**

Cross-range sections obtained from Figure 19.28.

correspondence of a scatterer's peak further demonstrate that there is a SNR ratio increase due to a better image focus, as shown in Figure 19.29.

2.19.12 Bistatic radar imaging

A renewed interest in bistatic radar, with specific attention to bistatic radar imaging has led the research effort to the development and implementation of Bistatic Synthetic Aperture Radar (BiSAR) and Bistatic Inverse Synthetic Aperture Radar (B-ISAR) systems.

Bistatic SAR algorithms have been proposed in the literature to solve the problem of the bistatic radar geometry. The main idea behind those algorithms has been that of extending monostatic SAR processing techniques to the bistatic case. Such extensions are practically obtained by rewriting the received signal phase model in order to account for the bistatic geometry. Some BiSAR algorithms can be found in [48–52].

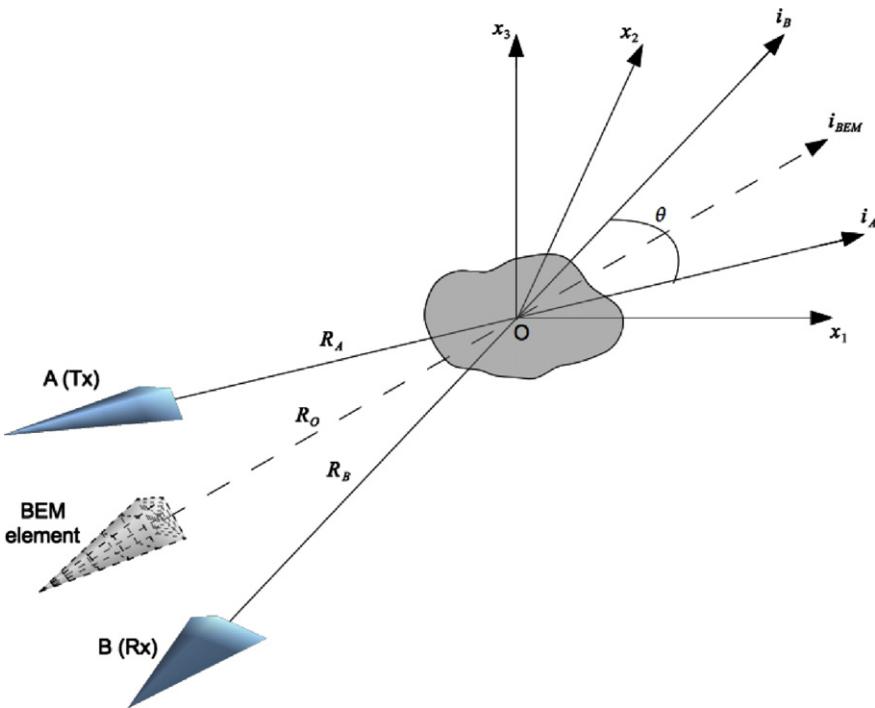
Some more details will be provided in Section 2.19.12.1 concerning B-ISAR, in order to demonstrate the usability of monostatic ISAR processors in bistatic geometries.

2.19.12.1 Bistatic ISAR

There are a number of reasons why bistatic radar imaging may be of interest in non-cooperative target imaging applications. The main ones are summarized below:

- *Geometrical limitations of monostatic ISAR:* In order to obtain ISAR images with a significant Doppler spread, it is necessary that the target changes its aspect angle with respect to the radar during the Coherent Processing Interval (CPI). This produces a set of geometrical cases where even if the target is moving with respect to the radar, an ISAR image cannot be obtained. A simple case is given by a target moving along the radar's Line of Sight (LOS). In this case the target aspect angle does not change in time and hence an ISAR image cannot be produced. This case is important because a target moving directly towards a radar may be hostile.
- *ISAR imaging of stealthy targets:* Military targets may be constructed to minimize the energy backscattered towards the radar. This makes them almost invisible to a radar. One approach to achieve this is by reflecting the electromagnetic energy towards directions other than that of the radar. Therefore, stealthiness usually refers only to monostatic radars. The use of a bistatic radar may enable the detection and therefore the imaging of stealthy targets.
- *Exploitation of bistatic SAR systems:* A number of bistatic SAR experiments have been conducted in the recent years to prove the effectiveness of bistatic radar imaging. The data collected by bistatic SAR systems could be processed as bistatic ISAR data and therefore, non-cooperative targets could be imaged by using a bistatic ISAR processor.
- *Multistatic ISAR imaging:* Multistatic ISAR imaging may be achieved by using one or more transmitters and a number of receivers. To maximize the gain out of such configurations, each receiver would benefit from acquiring the signal transmitted by other transmitters. This enables several bistatic configurations where the transmitters and the receivers are not co-located. In order to fully understand multistatic configurations, the bistatic configuration must be studied first.
- *Passive ISAR imaging:* There is an increasing interest in the passive radar field, as existing illuminators of opportunity can be used to detect and track targets. Although with limited bistatic range resolution, it has been recently demonstrated that Passive ISAR (P-ISAR) imaging can be enabled by exploiting modern digital broadcast communications [53].

Although several techniques for image reconstruction have been proposed for BiSAR that provide effective tools for radar imaging of static scenes, they do not apply to the ISAR case targets of interest are non-cooperative moving objects. Nevertheless, ISAR imaging usually aims to provide images of relatively small targets when compared to SAR images, where the imaged area can reach the size of hundreds of square kilometers. This advantage of ISAR with respect to SAR allows consideration of the use of monostatic ISAR processors even when the geometry is bistatic. In this subsection, the limits of applicability of a monostatic ISAR processor to a bistatic ISAR configuration are analyzed. The analysis here is limited to the case of no synchronization errors. The effect of synchronization errors on a bistatic ISAR (B-ISAR) have been analyzed in [54]. A well established method for analyzing radar imaging systems is the calculation of the Point Spread Function (PSF) which in our case will be addressed as bistatic ISAR image PSF. It will be shown that the bistatic ISAR image PSF depends on the bistatic angle, which introduces distortions in the bistatic ISAR image. The PSF of an imaging system also depends on the image formation processing adopted. Radial motion compensation followed by a

**FIGURE 19.30**

Bistatic geometry.

Range-Doppler technique will be considered as the ISAR image formation method, as it represents the standard procedure for obtaining ISAR images.

In this section the terminology and theory for Bistatic ISAR imaging is introduced. The geometry of the problem is illustrated in Figure 19.30.

The bistatic configuration produces an almost geometrically equivalent monostatic configuration which can be seen as a virtual transmitting/receiving element that lies at the bisector between the transmitter and the receiver. Such an equivalence and its limitation will be the subject of the next subsections.

2.19.12.1.1 Signal modeling

After motion compensation, as for the monostatic case (19.6), the received signal, can be written in a time-frequency format as follows:

$$S_R(f, t) = W(f, t) \int_V \xi(\mathbf{x}) \exp(-j\varphi(\mathbf{x}, f, t)) d\mathbf{x}, \quad (19.68)$$

where $\xi(\mathbf{x})$ represents, in this case, a bistatic reflectivity function and where the phase term ($\varphi(\mathbf{x}, f, t)$), which takes into account the bistatic configuration, can be written as follows:

$$\begin{aligned}\varphi(\mathbf{x}, f, t) &= \frac{4\pi f}{c} [R_A(t) + R_B(t) + \mathbf{x} \cdot \mathbf{i}_A(t) + \mathbf{x} \cdot \mathbf{i}_B(t)] \\ &= \frac{4\pi f}{c} [R_0(t) + K(t)\mathbf{x} \cdot \mathbf{i}_{\text{BEM}}(t)],\end{aligned}\quad (19.69)$$

where

$$R_0(t) = \frac{R_A(t) + R_B(t)}{2}, \quad (19.70)$$

$$\mathbf{i}_{\text{BEM}}(t) \frac{\mathbf{i}_A(t) + \mathbf{i}_B(t)}{|\mathbf{i}_A(t) + \mathbf{i}_B(t)|}, \quad (19.71)$$

$$K(t) = \left| \frac{\mathbf{i}_A(t) + \mathbf{i}_B(t)}{2} \right|, \quad (19.72)$$

and where $R_A(t)$ and $R_B(t)$ are the distances between point O on the target and the transmitter and the receiver, respectively, $\mathbf{i}_A(t)$ and $\mathbf{i}_B(t)$ are the unit vectors that indicate the LOS for the transmitter and the receiver, and \mathbf{x} is the vector that locates a generic point on the target.

An analysis of the effects of the bistatic geometry on the ISAR image Point Spread Function (PSF) follows.

2.19.12.1.2 PSF of the bistatic ISAR image

The term $K(t)$ carries information about the change in time of the bistatic geometry. However, what significantly affects the ISAR image PSF is the change of the bistatic angle during the coherent integration time. In this section the ISAR image PSF will be derived for the bistatic case, and the distortion introduced by the bistatic geometry will be related to the bistatic angle variation.

In deriving the PSF, two assumptions are made that will allow the application of the Range Doppler technique when reconstructing the ISAR image (following motion compensation). These two assumptions are: (1) the far field condition and (2) a short integration time. These assumptions avoid the need for consideration of non-constant target rotation vectors and the use of polar reformatting, and are generally satisfied in typical ISAR scenarios where the resolutions required are not exceptionally high.

When the target rotation vector is constant, the received signal backscattered by a single ideal scatterer located at a generic point may be rewritten (after motion compensation) in the following way:

$$S_R(f, t) = W(f, t) \iint \xi'(x_1, x_2) \exp(-j\varphi(x_{10}, x_{20}, f, t)) dx_1 dx_2, \quad (19.73)$$

where the phase $\varphi(x_{10}, x_{20}, f, t)$ may be written as:

$$\varphi(x_{10}, x_{20}, f, t) = \frac{4\pi f}{c} [K(t)(x_{10} \sin \Omega_{\text{eff}} t + x_{20} \cos \Omega_{\text{eff}} t)]. \quad (19.74)$$

In Eqs. (19.73) and (19.74), (x_{10}, x_{20}) are the coordinates of a generic scatterer on the target with respect to a reference system centered on the target itself (see Figure 19.30). Note: the third coordinate (x_{30}) does not appear in Eqs. (19.73) and (19.74), as in the monostatic case.

Under Assumptions (1) and (2) the bistatic angle changes are relatively small, even when a target covers relatively large distances within the integration time. As a result of this, the bistatic angle can be approximated by a first order Taylor (Maclaurin) polynomial:

$$\theta(t) \cong \theta(0) + \dot{\theta}(0)t, \quad (19.75)$$

where $-T_{\text{obs}}/2 \leq t \leq T_{\text{obs}}/2$ and $\dot{\theta} = \frac{d\theta}{dt}$.

Therefore, the term $K(t)$ can also be approximated with a first order Taylor (Maclaurin) polynomial, and by using Eq. (19.72) the following equation may be obtained:

$$K(t) \cong K(0) + \dot{K}(0)t = \cos\left[\frac{\theta(0)}{2}\right] - \frac{\dot{\theta}(0)}{2} \sin\left[\frac{\theta(0)}{2}\right]t = K_0 + K_1 t. \quad (19.76)$$

Therefore, Eq. (19.74) becomes:

$$\varphi(x_{10}, x_{20}, f, t) = \frac{4\pi f}{c} [(K_0 + K_1 t) \cdot (x_{10} \sin \Omega_{\text{eff}} t + x_{20} \cos \Omega_{\text{eff}} t)]. \quad (19.77)$$

For small integration angles (short integration time hypothesis) the sinusoids can be approximated by means of linear terms, as follows:

$$\varphi(x_{10}, x_{20}, f, t) \approx \frac{4\pi f}{c} [(K_0 + K_1 t)\Omega_{\text{eff}} t x_{10}] + \frac{4\pi f}{c} [(K_0 + K_1 t)x_{20}]. \quad (19.78)$$

An ISAR image reconstruction consists of:

1. Radial motion compensation.
2. Image formation.

When performing the radial motion compensation, any of the available techniques may be used. In fact, the phase term $\frac{4\pi f}{c} \left[\frac{R_A(t) + R_B(t)}{2} \right]$ may be removed as in monostatic configurations. This is because the term $K(t)$ does not affect the target radial motion compensation.

After motion compensation, the image formation (by means of the RD technique) makes use of two Fourier Transforms (FT): one along the frequency coordinate f (range compression) and one along the time coordinate t (cross-range image formation). In order to obtain the PSF of the bistatic ISAR system, we calculate the two FTs analytically.

2.19.12.1.3 Range compression

The range compression is obtained by Fourier transforming equation (19.73) along the variable f .

$$\begin{aligned} S'_R(\eta, t) &= \int W(f, t) \delta(x_1 - x_{10}, x_2 - x_{20}) \exp(-j\varphi(x_{10}, x_{20}, f, t)) \exp(-j2\pi f \eta) dx_1 dx_2 df \\ &= \exp\left(-j\frac{4\pi f_0}{c}(K_0 + K_1 t)\Omega_{\text{eff}} t x_{10}\right) \delta\left[\eta - \frac{2}{c}(K_0 + K_1 t)x_{20}\right] \otimes_\eta W'(\eta, t), \end{aligned} \quad (19.79)$$

where

$$W'(\eta, t) = \text{FT} [W(f, t)] = B \exp(-j2\pi f_0 \eta) \text{rect}\left[\frac{t}{\Delta t}\right] \text{sinc}[B\eta]$$

and \otimes_η is the convolution operator over the variable η .

Two effects are induced by the bistatic geometry:

1. the range position x_2 is scaled by a factor K_0 ;
2. a range migration is induced by the bistatic angle variation within the integration time.

Whilst the first effect can be corrected a posteriori by rescaling the range coordinate, the second effect can be significantly detrimental. If range migration occurs, the position of one scatterer can be moved from one range cell to another during the integration time, thereby resulting in a distortion of the PSF.

In order to avoid range migration the following equation must be satisfied:

$$|K_1| \Delta t x_{20}^M < \Delta r. \quad (19.80)$$

In Eq. (19.80), Δr is the range resolution of the radar and x_{20}^M is the scatterer with maximum distance from the target's zero range (focusing center). By substituting the expression of K_1 in (19.80) and by expressing the limitation with respect to the bistatic angle variation, the following relationship is obtained:

$$|\dot{\theta}(0)| < \frac{2\Delta r}{\Delta t x_{20}^M \left| \sin\left[\frac{\theta(0)}{2}\right] \right|}. \quad (19.81)$$

When the constraint (19.81) is satisfied, Eq. (19.79) can be rewritten as follows:

$$\begin{aligned} S'_R(\tau, t) &= \exp\left(-j\frac{4\pi f_0}{c}(K_0 + K_1 t)\Omega_{\text{eff}} t x_{10}\right) \\ &\times \delta\left[\eta - \frac{2}{c}K_0 x_{20}\right] \otimes_\tau W'(\eta, t). \end{aligned} \quad (19.82)$$

2.19.12.1.4 Cross-range image formation

Cross-range image formation is achieved by FT (19.82) along the coordinate t . The result is a complex image in the time-delay (range) and Doppler domains.

$$\begin{aligned} \text{PSF}(\eta, v) &= \int \exp\left(-j\frac{4\pi f_0}{c}(K_0 + K_1 t)\Omega_{\text{eff}} t x_{10}\right) \delta\left[\eta - \frac{2}{c}K_0 x_{20}\right] \otimes_\tau \\ &\times W'(\eta, t) \exp(-j2\pi t v) dt \\ &= \text{CH}[\nu, \alpha_0, \alpha_1] \delta\left[\eta - \frac{2}{c}K_0 x_{20}\right] \otimes_\eta \otimes_v w(\eta, v), \end{aligned} \quad (19.83)$$

where

$$w(\eta, v) = B \Delta t \exp(-j2\pi f_0 \eta) \text{sinc}[\Delta t v] \text{sinc}[B\eta], \quad (19.84)$$

$$\text{CH}[\nu, \alpha_0, \alpha_1] = \text{FT}\{\text{ch}[t, \alpha_0, \alpha_1]\}, \quad (19.85)$$

$$\text{ch}[t, \alpha_0, \alpha_1] = \exp\left(-j2\pi(\alpha_0 t + \alpha_1 t^2)\right), \quad (19.86)$$

$$\alpha_0 = \frac{2f_0\Omega_{\text{eff}}x_{10}}{c} \cos\left[\frac{\theta(0)}{2}\right], \quad (19.87)$$

$$\alpha_1 = -\frac{2f_0\Omega_{\text{eff}}x_{10}}{c} \sin\left[\frac{\theta(0)}{2}\right] \frac{\dot{\theta}(0)}{2}, \quad (19.88)$$

and \otimes_v is the convolution operator over the variable v .

Therefore the PSF of Eq. 19.83 can be rewritten as:

$$\text{PSF}(\eta, v) = \text{CH}[v, \alpha_0, \alpha_1] \otimes_v w\left(\eta - \frac{2}{c}K_0x_{20}, v\right). \quad (19.89)$$

It is worth recalling that a convolution between an infinite duration chirp and a sinc function is equivalent to a FT of a finite duration chirp, where the parameter of the sinc function is equivalent to the duration of the chirp.

As can be seen from Eqs. (19.86–19.88), the chirp rate depends on the position of the scatterer along the cross-range direction. By following [55] as a rule of thumb for negating the chirp effect (when a RD technique is used), the chirp rate must satisfy the following relationship:

$$|\alpha_1| < \frac{1}{\Delta t^2}. \quad (19.90)$$

By substituting Eq. (19.90) into (19.88), and by expressing it with respect to $\dot{\theta}(0)$, a rule for determining the maximum bistatic angle variation is obtained

$$|\dot{\theta}(0)| < \left| \frac{c}{f_0\Delta t^2\Omega_{\text{eff}}x_{10}^M \sin\left[\frac{\theta(0)}{2}\right]} \right|. \quad (19.91)$$

Bistatic ISAR scenarios which do not satisfy the constraint in (19.91) are likely to provoke image distortion. Nevertheless, such scenarios are particular ones where the ISAR system is pushed to the limit.

An example of strong bistatic angle variations is shown to highlight the distortion effects that may be encountered when the constraint in (19.91) is not satisfied. The geometry considered in this simulation is depicted in Figure 19.31, where the target moves along a rectilinear trajectory that is almost aligned with the BEM LoS. In this case the bistatic angle changes are the most severe. The image displayed in Figure 19.32 is obtained by applying the monostatic ISAR processor to the data generated with the bistatic geometry depicted in Figure 19.31. A non-distorted image is also produced by considering a monostatic radar located in the position of the BEM element and displayed in Figure 19.33. From a comparison between the two images it is possible to appreciate the distortion effects.

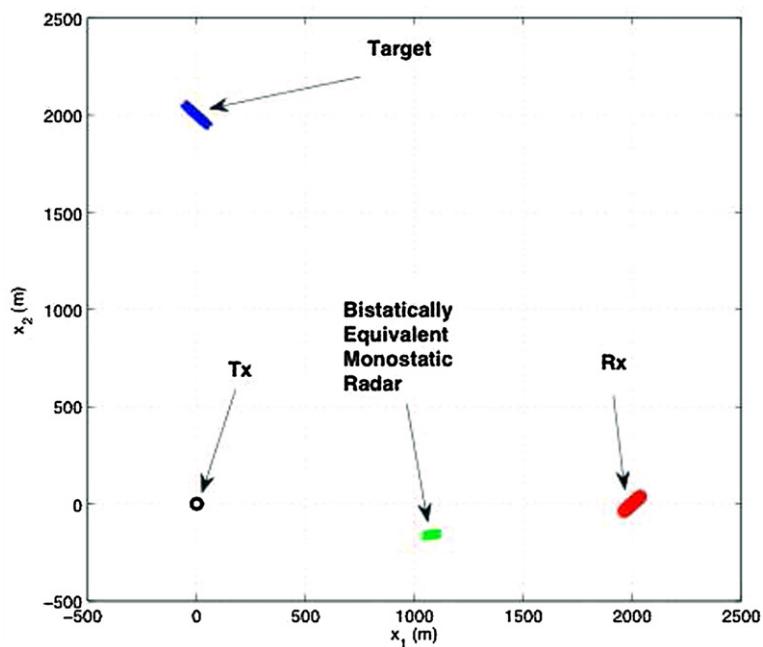


FIGURE 19.31

Simulated bistatic geometry.

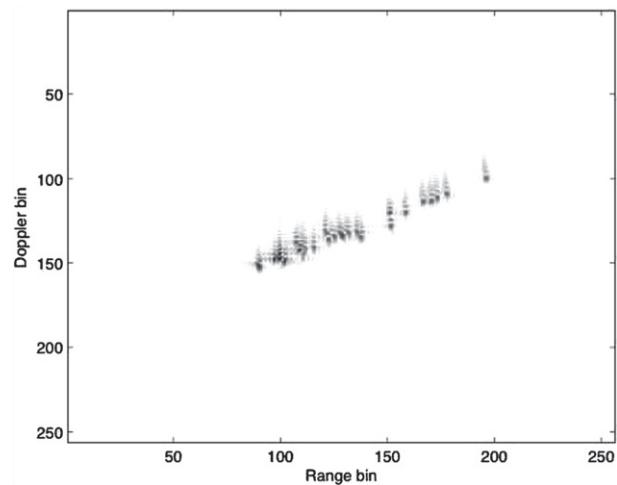
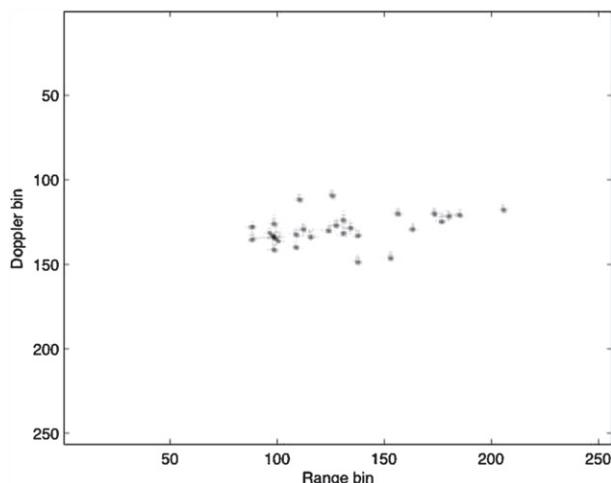


FIGURE 19.32

Distorted B-ISAR image.

**FIGURE 19.33**

Non-distorted BEM ISAR image.

2.19.13 Conclusion

The main concepts and algorithms relative to ISAR imaging have been treated in this tutorial. Specifically, the concept of high resolution applied to radar has been used to introduce ISAR imaging. A model based approach has been proposed as a method to derive the ISAR processor. Specifically, ISAR geometry and received signal modeling have been defined. Some physical and mathematical details have been included in this tutorial with the purpose of helping the reader understanding concepts on one side and providing the basis for implementing basic ISAR imaging algorithms. Polarimetric and bistatic ISAR imaging have also been discussed as they represent more recent advances in ISAR imaging that are opening the doors to the use of ISAR when polarimetric radars are employed or when the radar configuration is not monostatic (bistatic and multistatic). Examples have also been shown in a variety of scenarios.

Acronyms

2D	two-dimensional
3D	three-dimensional
BiSAR	bistatic SAR
B-ISAR	bistatic ISAR
CD	coherent decomposition
CPI	coherent processing interval

CTFT	Cohen's class time frequency transforms
DSA	sar = synthetic aperture radar dominant scatterer autofocus
EM	electro-magnetic
ESD	energy spectral density
FT	Fourier transform
GA	genetic algorithm
HS	hot spot
IC	image contrast
ICBA	image contrast based autofocus
ID	incoherent decomposition
IE	image entropy
IEBA	image entropy based autofocus
IFT	inverse Fourier transform
IPP	image projection plane
ISAR	inverse synthetic aperture radar
LoS	radar line of sight
ML	maximum likelihood
MVA	minimum variance algorithm
PGA	phase gradient algorithm
P-ISAR	passive ISAR
Pol-ISAR	polarimetric ISAR
Pol-SAR	polarimetric SAR
PSF	point spread function
RD	range-doppler
SLL	side lobe level
SNR	signal-to-noise ratio
TFT	time-frequency-transform
WV	Wigner-Ville (time frequency transform)

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/B978-0-12-411597-2.00019-9>.

Acknowledgment

Special thanks go to DSTO for releasing data that has been processed to form the ISAR images displayed in this chapter.

Relevant Theory: Signal Processing Theory, Statistical Signal Processing, and Array Signal Processing

See [Vol.1, Chapter 2](#) Continuous-Time Signals and Systems

See [Vol. 1, Chapter 3](#) Discrete-Time Signals and Systems

See [Vol. 1, Chapter 9](#) Discrete Multi-Scale Transforms in Signal Processing

See Vol. 3, Chapter 3 Non-Stationary Signal Analysis Time-Frequency Approach

See Vol. 3, Chapter 19 Array Processing in the Face of Nonidealities

References

- [1] G. Franceschetti, R. Lanari, *Synthetic Aperture Radar Processing*, CRC Press, 1999.
- [2] J.L. Walker, Range-doppler imaging of rotating objects, *IEEE Trans. Aerosp. Electron. Syst.* 16 (1980) 23–52.
- [3] D.A. Ausherman, A. Kozma, J.L. Walker, H.M. Jones, E.C. Poggio, Developments in radar imaging, *IEEE Trans. Aerosp. Electron. Syst.* 20 (1984) 363–400.
- [4] M.A. Richards, J.A. Scheer, W.A. Holm, *Principles of Modern Radar*, Scitech Publishing, 2010.
- [5] F. Berizzi, M. Martorella, B. Haywood, E.D. Mese, S. Bruscoli, A survey on ISAR autofocusing techniques, in: *Proceedings of the IEEE ICIP 2004*, Singapore, 2004.
- [6] M. Martorella, F. Berizzi, B. Haywood, A contrast maximization based technique for 2D ISAR autofocusing, *IEE Proc. Radar Sonar Navig.* 152 (4) (2005) 253–262.
- [7] J. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, Convergence properties of the nelder-mead simplex method in low dimensions, *SIAM J. Optim.* 9 (1998) 112–147.
- [8] M. Martorella, F. Berizzi, S. Bruscoli, Use of genetic algorithms for contrast maximization and entropy minimization in ISAR autofocusing, *J. Appl. Signal Process.* 2006 (2006) 1–11 (special issue on Inverse Synthetic Aperture Radar).
- [9] G.C. Carter, Time delay estimation for passive sonar signal processing, *IEEE Trans. Acoust. Speech Signal Process.* 29 (1981) 463–470.
- [10] B.D. Steinberg, Radar imaging from a distorted array: the radio camera algorithm and experiments, *IEEE Trans. Antennas Propag.* 29 (1981) 740–748.
- [11] B. Haywood, R.J. Evans, Motion compensation for ISAR imaging, in: *Proceedings of ASSPA 89*, Adelaide, Australia, 1989, pp. 113–117.
- [12] C.V. Jakowatz, D.E. Wahl, P.H. Eichel, D.C. Ghiglia, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*, Springer, 1996.
- [13] M. Martorella, F. Berizzi, Time windowing for highly focused ISAR image reconstruction, *IEEE Trans. Aerosp. Electron. Syst.* 41 (2005) 992–1007.
- [14] R.G. Parker, *Discrete Optimisation*, Academic Press, 1988.
- [15] M. Martorella, A novel approach for ISAR image cross-range scaling, *IEEE Trans. Aerosp. Electron. Syst.* 44 (1) (2008) 281–294.
- [16] O. Yeste-Ojeda, J. Grajal, G. Lopez-Risueno, Atomic decomposition for radar applications, *IEEE Trans. Aerosp. Electron. Syst.* 44 (2008) 187–200.
- [17] J. Tsao, B. Steinberg, Reduction of sidelobe and speckle artifacts in microwave imaging: the CLEAN technique, *IEEE Trans. Antennas Propag.* 36 (1988) 543–556.
- [18] M. Martorella, N. Acito, F. Berizzi, Statistical clean technique for ISAR imaging, *IEEE Trans. Geosci. Remote Sens.* 45 (11) (2007) 3552–3560.
- [19] S.M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Signal Processing, Prentice Hall, 1993.
- [20] L. Cohen, Time-frequency distributions—a review, *IEEE Proc.* 77 (7) (1989) 941–980.
- [21] V. Chen, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*, Artech House, 2002.
- [22] T. Thayaparan, L. Stankovic, C. Wernik, M. Dakovic, Real-time motion compensation, image formation and image enhancement of moving targets in ISAR and SAR using S-method based approach, *IET Signal Process.* 2 (2008) 247–264.

- [23] F. Berizzi, E. Mese, M. Diani, M. Martorella, High-resolution ISAR imaging of maneuvering targets by means of the range instantaneous doppler technique: modeling and performance analysis, *IEEE Trans. Image Process.* 10 (2001) 1880–1890.
- [24] S. Cloude, E. Pottier, An entropy based classification scheme for land applications of polarimetric SAR, *IEEE Trans. Geosci. Remote Sens.* 35 (1997) 68–78.
- [25] S. Jiancheng, J. Dozier, An entropy based classification scheme for land applications of polarimetric SAR, *IEEE Trans. Geosci. Remote Sens.* 33 (1995) 905–914.
- [26] M. Migliaccio, A. Gambardella, M. Tranfaglia, SAR polarimetry to observe oil spills, *IEEE Trans. Geosci. Remote Sens.* 45 (2007) 506–511.
- [27] W.L. Cameron, N.N. Youssef, L.K. Leung, Simulated polarimetric signatures of primitive geometrical shapes, *IEEE Trans. Geosci. Remote Sens.* 34 (3) (1996) 793–803.
- [28] R. Touzi, F. Charbonneau, Characterization of target symmetric scattering using polarimetric SARs, *IEEE Trans. Geosci. Remote Sens.* 40 (11) (2002) 2507–2516.
- [29] R. Touzi, F. Charbonneau, Characterization of target symmetric scattering using polarimetric SARs, *IEEE Trans. Geosci. Remote Sens.* 42 (10) (2004) 2039–2045.
- [30] W.L. Cameron, H. Rais, Conservative polarimetric scatterers and their role in incorrect extensions of the Cameron decomposition, *IEEE Trans. Geosci. Remote Sens.* 44 (12) (2006) 3506–3516.
- [31] M. Martorella, E. Giusti, A. Capria, F. Berizzi, B. Bates, Automatic target recognition by means of polarimetric ISAR images and neural networks, *IEEE Trans. Geosci. Remote Sens.* 47 (2009) 3786–3794.
- [32] M. Martorella, E. Giusti, L. Demi, Z. Zhou, A. Cacciama, F. Berizzi, B. Bates, Target recognition by means of polarimetric ISAR images, *IEEE Trans. Aerosp. Electron. Syst.* 47 (2011) 225–239.
- [33] S.R. Cloude, E. Pottier, A review of target decomposition theorems in radar polarimetry, *IEEE Trans. Geosci. Remote Sens.* 34 (2) (1996) 498–518.
- [34] E. Krogager, New decomposition of the radar target scattering matrix, *Electron. Lett.* 26 (18) (1990) 1525–1526.
- [35] J.R. Huynen, Measurement of the target scattering matrix, *Proc. IEEE* 53 (2) (1965) 936–946.
- [36] S.R. Cloude, The characterization of polarization effect in EM scattering, PhD dissertation, Univ. Birmingham, Fac. Eng., Birmingham, UK, 1986.
- [37] A. Freeman, S.T. Durden, A three component scattering model for polarimetric SAR data, *IEEE Trans. Geosci. Remote Sens.* 36 (3) (1998) 963–973.
- [38] M. Martorella, J. Palmer, F. Berizzi, B. Haywood, B. Bates, Polarimetric ISAR autofocusing, *IET Signal Process.* 2 (3) (2008) 312–324.
- [39] L. Xi, L. Gousui, J. Ni, Autofocusing of ISAR images based on entropy minimization, *IEEE Trans. Aerosp. Electron. Syst.* 35 (1999) 1240–1252.
- [40] F. Berizzi, G. Corsini, Autofocusing of inverse synthetic aperture radar images using contrast optimization, *IEEE Trans. Aerosp. Electron. Syst.* 32 (1996) 1185–1191.
- [41] J.R. Fienup, J.J. Miller, Aberration correction by maximising generalised sharpness metrics, *J. Opt. Soc. Am. A* 20 (2003) 609–620.
- [42] M. Iwamoto, T. Fujisaka, M. Kondoh, Autofocusing algorithm of inverse synthetic aperture radar using entropy, *Electron. Commun. Jpn.* 83 (1999) 97–106.
- [43] R. Morrison, M.N. Do, D.C. Munson Jr, SAR image autofocus by sharpness optimisation: a theoretical study, *IEEE Trans. Image Process.* 16 (2007) 2309–2321.
- [44] R. Morrison, D.C. Munson Jr, An experiemntal study of a new entropy-based SAR autofocusing technique, in: Proceedings of 2002 International Conference on Image Processing, vol. 2, 2002, pp. 441–444.
- [45] L.M. Novak, Studies of target detection algorithms that use polarimetric radar data, *IEEE Trans. Aerosp. Electron. Syst.* 25 (1989) 150–165.

- [46] S.R. Cloude, K.P. Papathanassiou, Polarimetric SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 36 (5) (1998) 1551–1565.
- [47] M. Martorella, L. Cantini, F. Berizzi, B. Haywood, E. Dalle Mese, Optimised image autofocusing for polarimetric ISAR, in: *Proceedings of the Eusipco Conference 2006*, 22–25 April 2002, Florence, Italy, Hindawi, 2006.
- [48] H. Nies, O. Loffeld, K. Natroshvili, Analysis and focusing of bistatic airborne SAR data, *IEEE Trans. Geosci. Remote Sens.* 45 (2007) 3342–3349.
- [49] M. Soumekh, Bistatic synthetic aperture radar inversion with application in dynamic object imaging, *IEEE Trans. Signal Process.* 39 (9) (1991) 2044–2055.
- [50] M. Soumekh, *Synthetic Aperture Radar Signal Processing with MATLAB Algorithms*, Wiley, 1999.
- [51] J.H.G. Ender, I. Walterscheid, A.R. Brenner, New aspects of bistatic SAR: processing and experiments, in: *IEEE IGARSS Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS)*, Anchorage, AK, United States, vol. 3, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ, United States, 2004, pp. 1758–1762.
- [52] D. D'Aria, A.M. Guarneri, F. Rocca, Focusing bistatic synthetic aperture radar using dip move out, *IEEE Trans. Geosci. Remote Sens.* 42 (7) (2004) 1362–1376.
- [53] D. Olivadese, M. Martorella, E. Giusti, D. Petri, F. Berizzi, Passive ISAR with dvb-t signal, in: *Proceedings of the EUSAR 2012*, 2012.
- [54] M. Martorella, Bistatic ISAR image formation in presence of bistatic angle changes and phase synchronization errors, in: *Proceedings of the EUSAR 2008 Conference*, Friedrichshafen, Germany, 2008.
- [55] F. Berizzi, E. Dalle Mese, Sea-wave fractal spectrum for SAR remote sensing, *IEE Proc. Radar Sonar Navig.* 148 (2) (2001) 56–66.

20

SAR Interferometry and Tomography: Theory and Applications

Gianfranco Fornaro^{*} and Vito Pascazio^{†,‡}

^{*}Istituto per il Rilevamento Elettromagnetico Ambientale (IREA), Consiglio Nazionale delle Ricerche (CNR), Napoli, Italy

[†]Dipartimento di Ingegneria, Università degli Studi di Napoli Parthenope, Napoli, Italy

[‡]Laboratorio Nazionale di Comunicazioni Multimediali, CNIT Complesso Universitario di Monte S. Angelo, Edificio Centri Comuni, Napoli, Italy

2.20.1 Introduction

During the 19th century, the theory of electromagnetic fields became a firmly established science: Maxwell's equations accurately described the propagation of the fields, Marconi's wireless experiments demonstrated the possibility of wireless communication at large distances. Nevertheless, even since that early time it was evident that electromagnetic waves could be not only used for communication but also to obtain information, or better to "sense" the environment and the objects without being in contact with them.

Remote sensing is today well established and intensively used for acquiring information about the Earth's surface [1]: among the most used remote sensing systems, active microwave sensors and particularly Synthetic Aperture Radar (SAR) have gained an increasing interest from both a scientific and industrial viewpoint. This success is a consequence of the capability of the sensor to operate independently of an external illumination source (day and night) and practically in almost any meteorological condition.

2.20.1.1 Microwave high resolution imaging

Active sensors make use of radars, typically installed on spacecraft, or on aircraft, or even on ground. They transmit a coherent (i.e., well controlled at the level of a single oscillation) signal and record the echoes scattered back to the sensor from the observed area. Accordingly, they are independent from any external illumination sources: this peculiarity, together with the fact that they work at wavelength that are, differently from optical and infrared sensors, almost immune to the presence of clouds and fog, provide the system with the possibility to operate day and night and also in adverse weather conditions.

Modern SAR sensors transmit signals whose bandwidth is of the order of tens/hundreds MHz, thus leading to spatial resolution along the range (typically called across track direction) of the order of meters or fraction of meter.

For comparable optical aperture and antenna size, the spatial resolution along the flight track of images acquired by microwave sensors should be several order of magnitude worse than the optical images.

This drawback is however overcome by the possibility to synthesize a very large antenna (of the order of few kilometers) by moving a much smaller real antenna along a straight trajectory corresponding to the platform flight track. This possibility, first postulated by Wiley [2] with the Doppler beam sharpening concept, is a direct consequence of the coherent nature of the system.

The operation of synthesizing a large antenna is today carried out typically off-line, after the downlink to a ground station, via a digital processing operation, usually referred to as **SAR focusing** operation, that, coherently combine on a 2D domain the echoes received from the radar at different positions. The obtained images are characterized by a resolution along the direction of array synthesis (typically referred to as along track direction) of the order of the length of the physical antenna dimension, independently from the wavelength and the height of the platform. Depending on the operative mode (scan, strip, and spot-mode) and on the fact that the platform can be disturbed during its motion by turbulences as in the airborne case, this operation can be in some cases more problematic.

SAR images are complex entities where the intensity basically measures the energy backscattered by the ground targets toward the sensor, which depends on the geometric (shape, roughness, and slope) and on physical (conductivity and permittivity) properties of the observed scene.

SAR sensors provide information about the observed scene complementary to that provided by optical systems. SAR images are nowadays used in many areas of interest: In glaciology they are used for glaciers monitoring and snow mapping, in agriculture for crop classification and soil moisture monitoring, in forestry for biomass estimation, etc. They are also used in environment monitoring for the detection of oil spills, flooding, as well as to monitor the urban growth or moving targets.

The technique that has opened probably the widest range of application is **SAR Interferometry** [3,4].

As in any electromagnetic coherent system the phase information is related to the travelled path, that is the distance target-sensor (**range**). Radar measurements therefore embed the information of distances with extreme high accuracy, on the order of wavelength fraction: due to the randomness of the scattering mechanism this information can be however extracted only as a relative measurement between different images. SAR Interferometry (InSAR) is a technique that, by exploiting at least two SAR images acquired from slightly different angles, allows retrieving the topography of the observed scene. A single SAR image provides measurement of measure the backscattering scene property only on a 2D domain, i.e., by performing a projection onto the plane containing the flight direction and the radar line of sight. Similarly to the human eyes system, height sensitivity can be achieved by combining two images of the same area acquired from two slightly different positions. The key principle of SAR interferometry is the use of the phase difference between SAR images for the accurate measurement of the distances of a target from two sensors displaced in location in order to create a parallax.

The two images can be acquired simultaneously if two antenna are present at the same time on the platform (*single-pass* interferometry) or through different passes of the same antenna (*repeat-pass* interferometry). In the latter case changes of the scene backscattering properties and variations of the phase delay contribution from the atmosphere may strongly impair the accuracy of the results. The accuracy of the topography estimation depends on the component orthogonal to the line of sight of the antenna vector separation commonly referred to as (spatial) baseline: for this reason this technique is also referred to as **across-track interferometry**.

As an alternative to topographic mapping, when the two antennas are present on the same platform and are separated along the flight direction, they acquire images repeated with a revisit of a few milliseconds.

This is the case of the **along-track interferometry** which allows monitoring fast movements of targets on the ground. Applications concern for example the estimations of ocean currents or moving detection and velocity estimation [5].

An interesting extension of across and along-track InSAR is the **Differential SAR Interferometry** (DInSAR): by exploiting phase difference of images acquired at times (epochs), separated typically by some days, it allows accurately monitoring slow displacements over the epoch sequence. Differential interferometric data can be acquired by radar observations separated in time either from a single radar on one platform (e.g., ERS-1, JERS-1, ENVISAT, TerraSAR-X) or from multiple radars on different platforms provided the radar have similar radar operating a viewing parameters (e.g., Cosmo Skymed constellation). Since the precision of radar in estimating distance is in the order of fraction of wavelength, DInSAR can estimate movements with sub-centimetric accuracy using L-, C-, or X-Band radars.

Major applications of this technique regard the natural hazard and security area. Besides, by exploiting archives of past images multipass techniques are also extremely useful to provide a past monitoring.

Interferometry applications have dramatically increased the use of microwave remote sensing for the environment monitoring: This is also testified by the growing interest of the major international space agencies in the development and launch of spaceborne SAR sensors satellites. The twin satellites ERS-1 and ERS-2 [6] of the European Space Agency (ESA), operative since 1992 and 1996, respectively, each one with a revisiting time of 35 days each, were characterized by the possibility to acquire a pairs of *tandem* images, i.e., interferometric images separated by only one day. ERS sensors from nineties to the first decade of 2000 were the very first systems used for the operative demonstration and routinely application of interferometry. Their acquisitions have been deeply exploited for years to develop most of the interferometric processing algorithm currently used and to demonstrate the potentials of the application of SAR Interferometry in several natural risk areas. Recently, the Italian Cosmo Skymed [7] and German TerraSAR-X [8] missions improved the quality of SAR product by providing images with spatial resolution up to one meter. Together with its twin satellite TanDEM-X, TerraSAR-X is going to provide the most accurate Digital Elevation Model (DEM), i.e., the topography, of the Earth on a global scale with a relative accuracy of 2 m for slope lower than 20° and 4 m for higher slopes with a spatial grid of 12 m. From the other side, the Italian COSMO-Skymed mission [7] is, worldwide, the unique constellation of more than two SAR sensors exploited also for civilian applications. It is composed of four medium-size satellites, each one equipped with an X-band high-resolution SAR system, allowing acquiring images on the same area every 4 days on average, thus both reducing the effects of decorrelation and allowing a more frequent imaging which is useful for interferometric application to cases of emergency.

Polarimetry [9–16] and **Polarimetric SAR interferometry** [17,18] and are techniques that use multi-polarization channels to extract further information on the scattering mechanism. Polarimetric information allows separating different scattering mechanisms. Whereas SAR polarimetry is a technique that use single antenna data, Polarimetric SAR interferometry use data acquired by two antennas. The former has a wide use in the field of classification, the latter allows generating interferograms corresponding to different scattering mechanisms and has an area of application in the field of forest height retrieval and biomass estimation.

The advances in the SAR hardware has allowed to reach very high imaging resolutions at microwaves on the order of a meter and has in parallel stimulated the development of advanced processing techniques able to extract from the data the highest possible information content.

One of the most important and recent innovations in SAR processing is associated with the extension of the imaging process from a 2D domain to a multi-dimensional domain. The so-called **SAR Tomography** has been among the first examples giving to SAR the ability of reconstructing images of the backscattering property of the scene also along the direction (elevation) orthogonal to the two classical dimension (azimuth and range). The key aspect of this technique is the possibility to synthesize, similarly to the flight direction (azimuth) an array also along the “height” direction to sharpen and steer the beam in such a way to measure the backscattering characteristics of the scene along the elevation direction and hence to generate full 3D images.

SAR Tomography allows vertically profiling (3D imaging) the backscattering to detect targets which interfere in the same pixel of a single SAR image and even monitoring, with the extension of the imaging properties to the time direction (4D Imaging), their individual deformation. Beside the application to a distributed scenario such as forest where the scattering is distributed along the height, the tomographic technique also provides significant advances in the application to the imaging and monitoring of areas characterized by a high density of scatterers, such as urban areas, opening the possibility to achieve dense imaging and monitoring of single buildings and individual structures from the space, for the first time comparable to what obtainable with in situ systems like laser scanner [19,20].

Polarimetric SAR tomography [21–24] takes benefits of both polarimetry and tomography: by accessing the multibaseline information on different polarization channels, it allows retrieving scattering profiles along the elevation direction associated with different scattering mechanisms such as single bounce, double bounce and volume scattering. This work concentrates on the development of SAR interferometry (including multipass Differential SAR Interferometry) and Tomography for 3D reconstruction and target deformation monitoring.

2.20.2 Basics concepts in SAR imaging and SAR interferometry

2.20.2.1 High resolution image formation

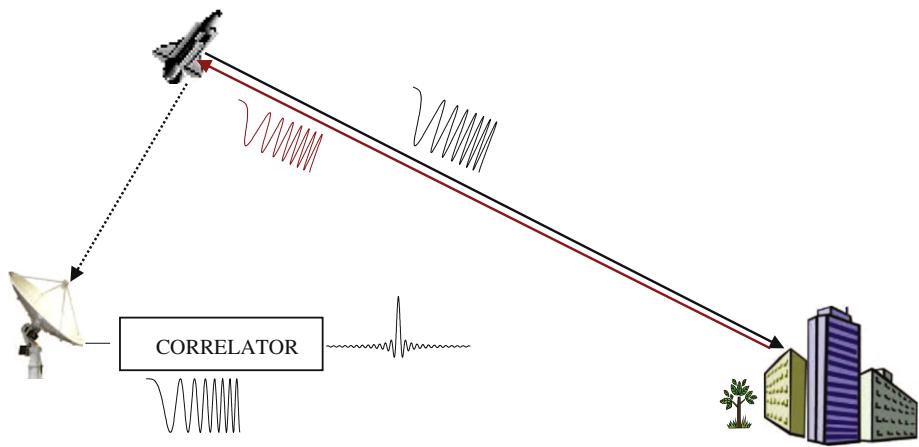
Among the several parameters characterizing an image, resolution plays certainly a major role. In the radar case the resolution along the range coordinate depends on the system bandwidth [25,26]. Large bandwidths are obtained, with simplified (i.e., low peak power) hardware, by transmitting long duration linear frequency modulated (chirp) pulses which are, after echo reception, compressed (typically on the ground) via correlation techniques: this operation is commonly referred to as range pulse-compression or range focusing (see Figure 20.1).

The transmitted chirp pulse has the following expression:

$$s(t) = \text{rect}\left[\frac{t}{\tau}\right] e^{j\pi\alpha_t t^2}, \quad (20.1)$$

wherein $\text{rect}[\cdot]$ is the window function, τ is the pulse duration, and α_t is the Chirp rate (Hz^2). The correlation of the response of a target at range r with the transmitted pulse replica provides the expression of the range impulse response function (IRF), also known as range Point Spread Function (PRF):

$$\hat{y}(t) = s\left(t - \frac{2r}{c}\right) * s^*(-t) \cong B \text{sinc}(Bt), \quad (20.2)$$

**FIGURE 20.1**

System geometry in the range direction.

with c being the light-speed and $B = \alpha_t \tau$ the bandwidth of the transmitted pulse. The (3 dB) range resolution is numerically given by [27]

$$\delta_r \cong 0.88 \frac{c}{2B}. \quad (20.3)$$

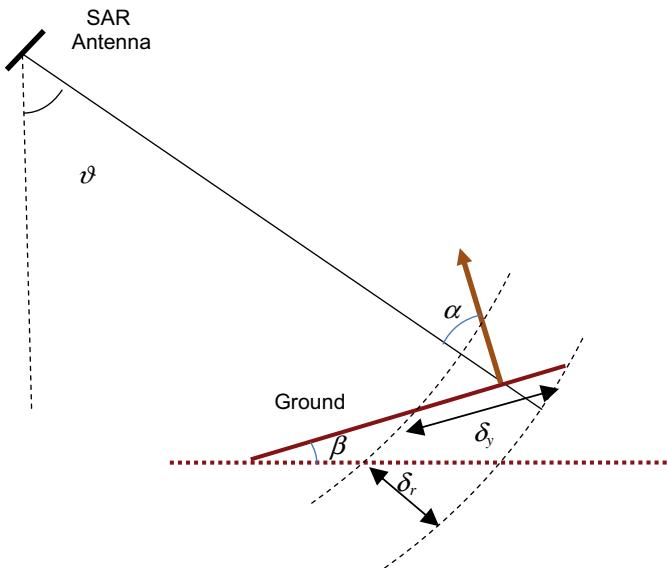
Such a resolution value is also referred to as *slant range resolution* to highlight that it refers to the line-of-sight (LOS) direction. Scaling factors derived by standard trigonometry should be applied to achieve the resolution along the main scene direction, for instance along the direction corresponding to the projection of range onto the local cartographic reference system usually referred to as *ground range resolution*:

$$\delta_y \cong \frac{\delta_r}{\sin(\alpha - \beta)}, \quad (20.4)$$

where α is the so-called incident angle, defined as the angle between the radar LOS and the local normal to the surface at the point of the reflection on the ground (see Figure 20.2), and β is the terrain slope. The ground range resolution is, of course, coarser than the slant range resolution.

Note that in the absence of terrain slope, the incidence angle is equal to the angle θ (known as look angle and defined with respect to the nadir direction) only when the Earth curvature can be neglected, i.e., as in the case of airborne sensors operating at low altitude.

In case of flat ground and if the SAR antenna beamwidth in the range direction is not too big, as usually occur for instance for new generation X-band SAR sensors, the ground resolution is almost constant along the footprint. Differently, in case of non-flat ground, as shown in Figure 20.3, the ground resolution can change significantly, as effect of topography, giving rise to the well known effects of *foreshortening*, *layover*, and *shadowing*. Under conditions of foreshortening different resolution cells can contain contribution of very different, in terms of dimension, ground areas (see Figure 20.3b). Layover is beyond the limit case just described: in this case, points more distant in the ground can appear

**FIGURE 20.2**

SAR range resolution. δ_r is the slant range resolution, δ_y is the ground range resolution, ϑ is the look angle, α is the incidence angle, and β is the local terrain slope. In absence of topography ($\beta = 0$), it results that $\alpha + \vartheta = \pi/2$.

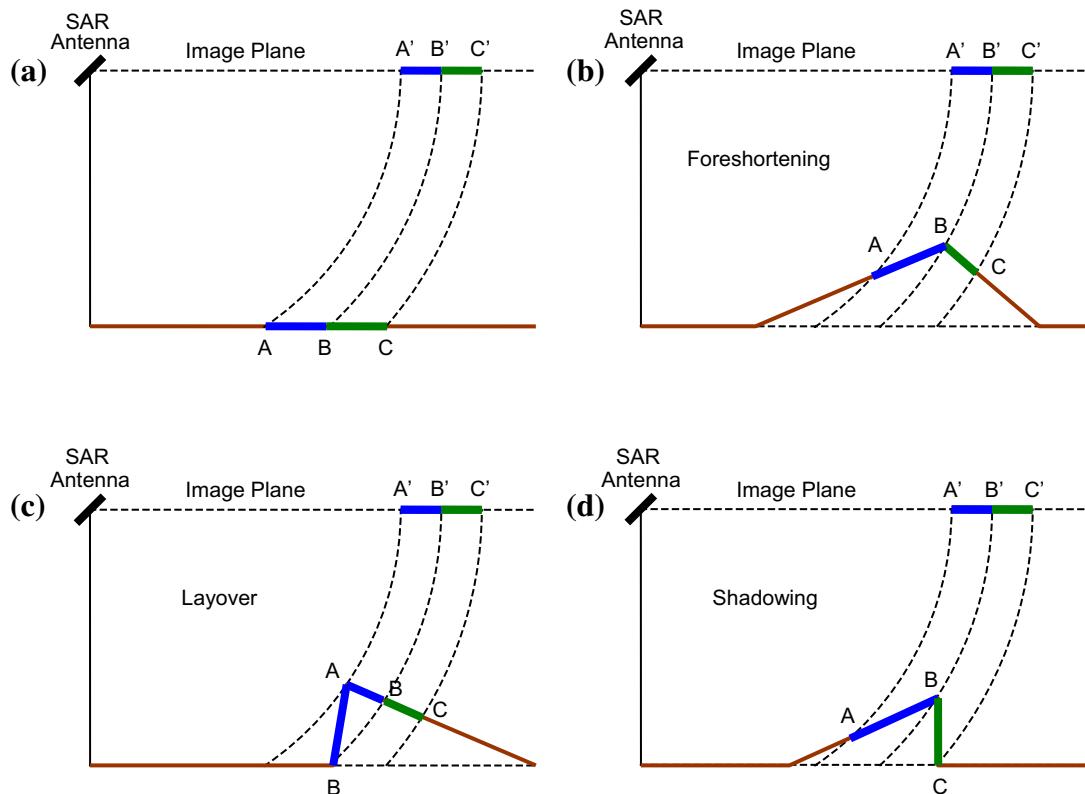
closer to the SAR radar sensor, and are mapped erroneously in the SAR image (see Figure 20.3c). Such effect is very common in mountainous areas with steep slopes, and in urban areas [4]. The shadowing effect occurs when ground area are masked by reliefs, as in Figure 20.3d. In this last case, a slant range resolution cell does not map any ground area (the ground area BC in Figure 20.3d is not seen from the SAR antenna). The above presented distortion effects makes that SAR images of mountain and urban areas can look very different from optical images, as an effect of geometrical distortions.

In the azimuth direction the focusing operation is necessary to synthesize a long antenna with higher resolution capabilities, that is for achieving the beam sharpening.

With reference to Figure 20.4 where the system imaging geometry represented in the flight direction, the system “senses” the scene by transmitting pulses at regular time instants, regulated by the pulse repetition frequency. The echoes collected in each position may be coherently processed in such a way to synthesize (digitally) an antenna whose dimension is equal to the footprint (X) of the real antenna [25,28]:

$$X = \frac{\lambda}{L} r, \quad (20.5)$$

where λ is the wavelength, L is the azimuth length of the real antenna, and r is the range of the target (range). Note that λ/L is the angular aperture of the real SAR antenna in the azimuth direction. The

**FIGURE 20.3**

SAR distortion effects: (a) normal conditions, (b) foreshortening, (c) layover, and (d) shadowing.

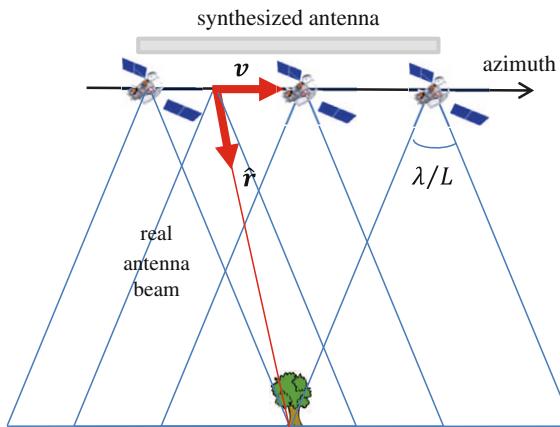
final resolution of the image, provided by the synthetic antenna is [25]:

$$\delta_{az} \cong 0.88 \frac{\lambda}{2X} r \cong 0.88 \frac{L}{2}, \quad (20.6)$$

where the resolution gain factor 2 in the first equality is associated with the capability of the array to transmit and receive the radiation from each position of the real antenna during the synthetic antenna formation.

The time interval in which the scatterer is illuminated is referred to as *integration time*: for a standard operating mode, as that illustrated in Figure 20.4 (referred to as *stripmap mode*), it is trivially given by the ratio between the real antenna footprint and the platform velocity (v).

A dual approach used for the computation of the azimuth resolution of the focused image is provided by the so called Doppler analysis which states that when either transmitters or receiver are subject to a uniform motion, the received radiation is subject to a frequency shift (called Doppler shift)

**FIGURE 20.4**

System geometry in the azimuth direction. The movement of the platform allows to synthesize a larger antenna thus achieving a sharpening of the beam of the real antenna.

equal $2v \cdot \hat{r}/\lambda$, with v being the velocity vector (in our case of the platform) and \hat{r} being the versor of the receiving radiation (in our case the direction locating the scatterer from the platform). It is therefore clear that, during the integration interval, the echo from the target sweeps an interval from $v \cdot \hat{r}|_{\min} = -v \sin[\lambda/(2L)] \cong -v\lambda/(2L)$ to $v \cdot \hat{r}|_{\max} = v \sin[\lambda/(2L)] \cong v\lambda/(2L)$, which corresponds to a Doppler frequency interval from $2v \cdot \hat{r}/\lambda|_{\min} \cong -v/L$ to $2v \cdot \hat{r}/\lambda|_{\max} \cong v/L$. Accordingly, the Doppler bandwidth amounts to:

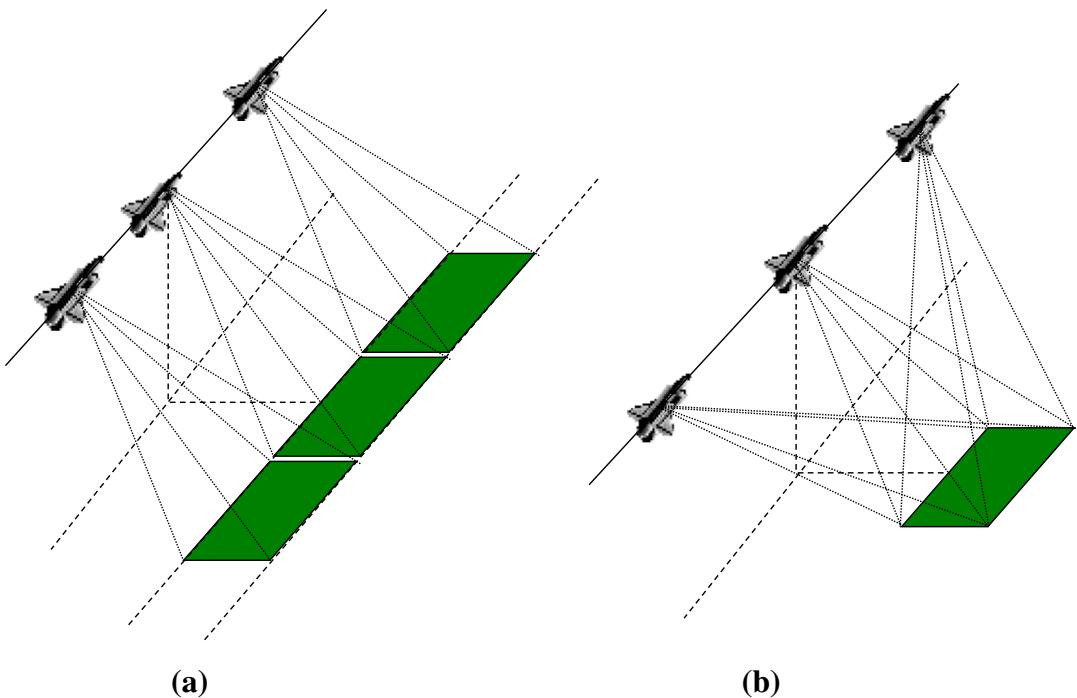
$$B_d = 2 \frac{v}{L}. \quad (20.7)$$

Such a bandwidth is able to provide pulses with time duration equals $L/(2v)$, which corresponds a spatial extent of about $L/2$. Unfortunately, the relative motion between the sensor and the target introduces also linear (phase) distortions as well as motion through resolution cell. Therefore, to provide short duration pulses the phase distortions affecting the available bandwidth must be compensated at the azimuth focusing level with the use of filters which are intrinsically 2D and also space variant only (for rectilinear tracks) with the range. The SAR focusing topic is out of the scope of this work, readers can refer to [25,28].

2.20.2.2 Operational modes

The classical operational mode of a SAR system considers the antenna pointing with a fixed offset from the flight direction, non-necessary orthogonal (i.e., broadside) pointing: this is referred to as **Stripmap mode** to highlight the fact that the scene is illuminated along a strip. The stripmap imaging mode geometry is depicted in Figures 20.4 and 20.5a.

In this way the integration time for forming the image of a target is, as discussed in the previous section, limited by the ratio between the real antenna azimuth footprint dimension and the platform velocity. This poses a limitation on the maximum achievable resolution. Another limitation of this imaging mode is associated with the coverage of the imaged strip in the slant range direction, which is

**FIGURE 20.5**

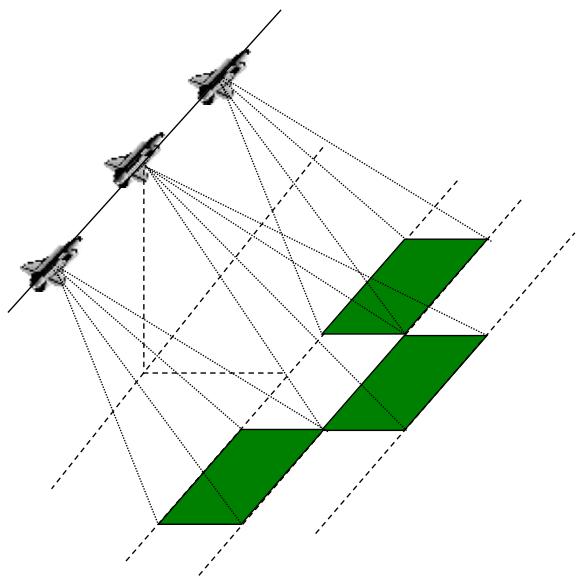
Stripmap (a) and Spotlight (b) operational modes.

in this case provided by the ground range extent of the real antenna footprint [25]. Slant range coverage and imaging resolution can be traded-off by operating a beam steering during the acquisition. Beside the classical stripmap mode, the two most known operational mode are spotlight and scan mode.

In the **Spotlight mode** the antenna beam is steered backward with respect to the antenna flight direction in such a way to collect data from a fixed area on ground on a longer (compared with the classical stripmap mode) flight segment [26,29,30]. The spotlight imaging mode geometry is depicted in Figure 20.5b.

In particular with respect to the stripmap mode (Figure 20.5) where the beam orientation is fixed, in the spotlight mode what is fixed is the illuminated area: this spotlight configuration is usually referred to as staring spotlight mode. A configuration that allow obtaining illumination interval and hence resolutions between the stripmap and *staring* spotlight mode is the so called *sliding* spotlight in which the angular beam steering rate is reduced in such a way to allow the footprint to slide on the ground [31,32]. With respect to the staring spotlight, the resolution loss is compensated by an increase of the azimuth coverage.

A mode complementary to the spotlight is the **ScanSAR mode** [33] whose geometry is shown in Figure 20.6. In this case the antenna is steered in the range direction to increase the range coverage.

**FIGURE 20.6**

ScanSAR operational mode.

During the aperture synthesis in azimuth, the beam is regularly steered in range to sweep among a fixed number (typically 2–4) of adjacent range subswaths. The sweep mechanism is carried out in such a way to avoid gaps along the azimuth direction over the subswaths, i.e., to avoid the presence of areas which are not illuminated in the azimuth direction. The data collected in an illumination sub-interval for a generic subswath is called burst. In the stripmap case each target is imaged from the whole antenna beam and therefore the radiometric accuracy is preserved, that is homogenous area are imaged in a (average) constant backscattering level. In the ScanSAR case, as the target is seen only from a, or from a few small portions of the azimuth beam during the burst acquisition, not only we have a reduction of the azimuth resolution (this is the price for the increase in the range cover), but also different areas can be seen by different portions of the azimuth antenna beam. The latter effect produces radiometric losses seen as stripes along the azimuth (scalloping): homogenous area are imaged in a variable backscattering level. A mitigation of the scalloping problem is achieved by the adopting a TOPS (literally reverse of SPOT) acquisition mode [34]. In this case, in addition to the range steering, a steering in azimuth, with a forward rotation (i.e., opposite to that of the SPOT mode), is carried out to allow the azimuth beam to run forward, faster than the platform, in such a way that almost all scatterers in azimuth are imaged by the highest possible beam portion.

The azimuth resolution for the different modes can be evaluated by referring to the Doppler bandwidth, evaluated in (20.7), which can be written as a product of the Doppler rate f_M and the integration time T_{int} :

$$B_d = f_M T_{\text{int}}. \quad (20.8)$$

Equation (20.8) follows directly from the fact that the signal collected along the azimuth direction (slow time) is with a good approximation a linear frequency modulated pulse; the associated Doppler rate f_M equals:

$$f_M = \frac{2v^2}{\lambda r}. \quad (20.9)$$

In the Stripmap case the integration time is fixed by the real antenna beamwidth:

$$T_{\text{int}} = \frac{\lambda}{L} r \frac{1}{v}. \quad (20.10)$$

By substituting Eqs. (20.10) and (20.9) in Eq. (20.8), Eq. (20.7) is obtained. In the Scansar and Spotlight cases, due to the range or azimuth antenna sweep, the integration time is fixed to a value which is lower and higher than the limit in (20.10), respectively, in order to select the wanted azimuth resolution.

It is important to point out also that, while in the Stripmap case the spectral properties of the received signal are time invariant along the azimuth, in the case of Scansar and Spotlight acquisition, due to the antenna steering, the spectral properties show an azimuth space variance. For instance, in the spotlight mode, and particularly in the sliding spotlight configuration, due to the difference between the platform and footprint velocity, the angular view of the system to the scene is azimuth dependent and accordingly, the received Doppler bandwidth is progressively translated from positive to negative frequencies. All these aspects must be accounted during post processing, such as for instance image resampling and/or image filtering as in the case of SAR interferometry [35].

2.20.3 SAR interferometry

SAR image is a 2D complex signal, resulting from the coherent processing of raw data acquired from the synthetic antenna [26]. The amplitude of the SAR images represents the reflectivity of the ground area under view while the phase of the SAR images is randomly distributed [26,36]. In addition to electromagnetic scattering properties of the targets, the latter embeds also very important geometrical measurements.

Such information can be extracted exploiting two [37–39] (or more than two [40–42] SAR complex images in the framework of SAR interferometry. In particular, the term SAR Interferometry (InSAR) is referred to all methods that employ at least two complex SAR images, exploiting mainly their phases, in order to derive more information about a ground scene respect to the information provided by a single SAR image. The additional information is provided when at least one among the key acquisition parameters of the SAR system is different from acquisition to acquisition.

There exist two possible main configurations of SAR Interferometry: across track interferometry [38] and along track interferometry [43]. In the across track configuration, two (or more) SAR sensors fly on two parallel flight lines and look at the ground from slightly different look angles. In the along-track configuration, two (or more) sensors fly on the same path, looking the scene from the same position but with a very small temporal gap. The across track InSAR configuration allows recovering the height profile of the ground area under observation, while the InSAR along track configuration is mainly used for measurement of fast displacements such as ocean currents [44] and for moving target detection and velocity estimation [45,46].

In all interferometric processing the starting point are the SAR complex images, that can be obtained by means of a two dimensional (2D) processing of the raw data acquired from the SAR sensors [47]. The SAR complex images $z(x,r)$ are representative of the reflectivity of the ground scene, in the sense that they are 2D discrete complex signals of the azimuth (x) and range (r) coordinates, where each sample (an image pixel) embeds the mean reflectivity characteristics of a sampling cell of the ground scene.

2.20.3.1 Across-track SAR interferometry for measuring the surface topography

The regular and controlled oscillation of coherent radiation used in SAR systems allows determining with high accuracy the variations of the propagation distance: such a basic property is the key principle of interferometric techniques.

A single SAR image provides measurement about the backscattering scene property along two directions: The target range (i.e., the distance of the target from the illumination track) and the position of the target along the track (the azimuth direction). Hence, no information is provided on the angle under which the target is imaged (look angle). Knowledge of the latter information completes the set of coordinates in a cylindrical reference system with the axis coincident with the track, thus allowing a full localization of the scatterers in 3D and therefore an estimation of topography.

Similar to the mechanism used in human visual system for the determination of the depth, SAR interferometry is a technique that exploits the parallax in the view of the scene to allows extending the capability of a single SAR system to the reconstruction of the scene elevation profile: as SAR is sensitive to distance whereas the visual system is sensitive to angles, the mechanism is indeed slightly different.

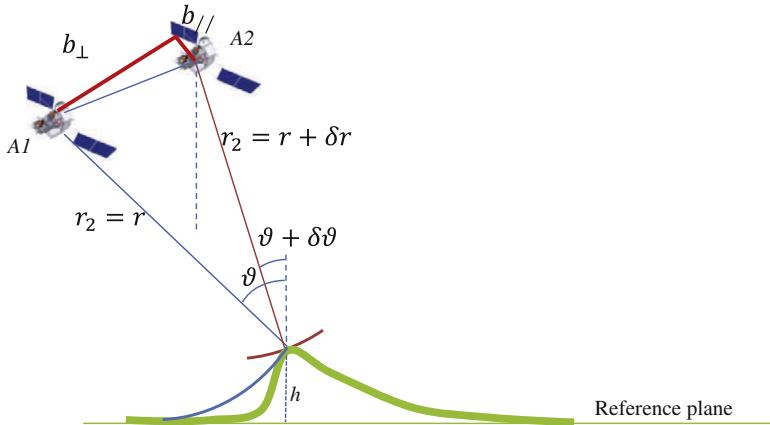
Figure 20.7 shows the geometry of a basic (two-antenna) interferometric acquisition in the plane orthogonal to the flight track: it is clear that by measuring the range of the target with a single (master) SAR system (say the blue line) it is not possible to uniquely localize the position of the scatterer because at the same range would be located all of the points distributed on a equi-range curve (the blue¹ one) in the elevation beam (dash line).

By using a second (slave) antenna that images the scene from a different look angle the system is able to measure also the range from a second location [48]: there is then only one point (the intersection of the two equi-range, i.e., blue and red lines in Figure 20.7) that obey to the distance measurements pair. The larger the separation between the two antennas, the sharper the crossing and hence the higher the height accuracy. As in the visual system the 3D sensitivity is given by the difference in the location of the object in the two images at the different eyes [37], the accuracy of the stereometric system in Figure 20.7 is related to the variation of the distance of the target from one to the other antenna (range difference). To provide sufficient accuracy in such a range variation, SAR interferometry uses the phase difference between the two SAR images: the path difference is hence measured to an accuracy which is a fraction of the wavelength (centimeters at microwaves).

Specifically, the terms that plays the key role in the determination of the height is the path difference δr . In particular, at large distances it can be shown that the (variation of) the path difference (with respect to a reference point for instance located on a plane) is [37,39]:

$$\delta r \cong \delta\vartheta \frac{h}{\sin\vartheta} = \frac{b_{\perp}}{r \sin\vartheta} h, \quad (20.11)$$

¹For interpretation of color in Figures 20.7 and 20.12 the reader is referred to the web version of this book.

**FIGURE 20.7**

Interferometric geometry.

where (see Figure 20.7) $\delta\vartheta$ is the variation of the look angle between the master and slave antenna, b_{\perp} is the orthogonal baseline component, that is the component orthogonal to the (master) line of sight of the vector (baseline) connecting the two satellites, and ϑ is the incidence angle, that is the angle between the vertical direction related to the target and the direction of the incoming radiation (line of sight). It is however important to note that Eq. (20.11) represent a simplification of the true scenario which is useful to understand the key principles of across-track interferometry. In the reality, the height is referred to a geographic or cartographic reference system and determined by measuring δr and by knowing the orbital state vectors [49].

In SAR interferometry the path difference is measured to accuracy of the order of the wavelength by using the phase difference signal:

$$\varphi_{12} = \frac{4\pi}{\lambda} (r_2 - r_1) = \frac{4\pi}{\lambda} \delta r \cong \frac{4\pi}{\lambda} \frac{b_{\perp}}{r \sin \vartheta} h. \quad (20.12)$$

For application to topographic mapping the two interferometric images can, or better should be acquired at the same time by two antennas on the same platform (bistatic system). This is because changes in the scattering properties, as well as differences in the propagation phase delay through the atmosphere strongly impact the quality of the retrieved DEM. In the case of the Shuttle Radar Topography Mission (SRTM) in 2001 an extensible boom 60 m long was mounted on-board the Shuttle to separate the slave from the master antenna available in the fuselage [50]. The German TanDEM-X mission of 2010 is instead the first example of a bistatic system composed by two twin satellites (TerrSAR-X and Tandem-X) orbiting in a close (500 m one behind the other) formation [51].

2.20.3.2 Statistical characterization of across-track SAR interferometric signals

As mentioned before, an Across Track SAR interferometric (InSAR) system is used to reconstruct earth topography, providing high precision Digital Elevation Model (DEM) of Earth surface. The geometry

of an InSAR system has already been shown in Figure 20.7, where two SAR systems look at the scene from two slightly different tracks. As already introduced, the distance b between the two SAR tracks is called baseline, its component orthogonal to the look direction b_{\perp} is the orthogonal baseline, while its component parallel to the look direction (to the slant range) b_{\parallel} is the parallel baseline.

In order to understand how an InSAR system works, consider the distance r_1 between the first SAR antenna A_1 and a point target T on the ground, and the distance r_2 between the second SAR antenna A_2 and the same point target T , as shown in Figure 20.7, while ϑ_1 and ϑ_2 denote the angles at which the two antennas look at the point target on the ground (slightly different from each other).

Consider now the two complex (envelope of the) images $Z_1(n, m)$ and $Z_2(n, m)$ obtained processing raw data collected by the two SAR sensors, where (n, m) are the discrete coordinates corresponding to the continuous azimuth and range coordinates (x, r) . Such images can be considered as random processes whose expression is [52]:

$$Z_k(n, m) = \rho(n, m) e^{j\varphi_\rho(n, m)} e^{j\frac{4\pi}{\lambda} r_k(n, m)} X_k(n, m), \quad k = 1, 2, \quad (20.13)$$

where $\rho(n, m) e^{j\varphi_\rho(n, m)}$ is the complex envelope of the (deterministic) ground reflectivity function (which, in first approximation, can be assumed to be constant with the antenna position $k = 1, 2$, since, as commented above, the view angles change only slightly from one position to another), $[4\pi r_k(n, m)/\lambda]$ are phase factors related to the different propagation paths between the two antenna positions and the point target, and

$$X_k(n, m) = A_k e^{j\Omega_k(n, m)}, \quad k = 1, 2 \quad (20.14)$$

is the random process representing the multiplicative *speckle* noise at the k th antenna, typical of any coherent system, which is commonly assumed to be a complex Gaussian correlated process with zero mean and unit variance [53]. Of course, $Z_k(n, m)$ are also random processes.

As result of the SAR signal model given by (20.13) and (20.14), the phase of a SAR image pixel (n, m) is given by three main contributions:

- a first term φ_ρ , representing the phase shift induced by the scattering mechanism; it is deterministic, and it is the same for the two antennas;
- a second term $[4\pi r_k/\lambda]$, representing the phase shift due to the different propagation; it is deterministic, and it depends on the antenna;
- a third term Ω_k , induced by the coherent nature of the SAR processing; it is **random**, and it depends on the antenna.

Other phase terms related to geometrical uncertainties, to random propagation effects, or to the changes of the scattering mechanism between the two SAR image acquisitions (for instance, due to the time delay between the two acquisitions) can be also present in Eq. (20.13).

After a processing called image registration which aims to locate the response of a target in the azimuth range pixel in the two images same pixel [37, 54], the two SAR images Z_1 and Z_2 are used to build the so-called multi-look SAR interferogram:

$$\Psi_{12} = \arg \left(\sum_{n=1}^{N_L} Z_1^{(n)} Z_2^{*(n)} \right), \quad (20.15)$$

where $\arg(\cdot)$ denotes the principal value of the phase, N_L is the number of looks [26], and the explicit dependence on (n,m) has been understood (the same will be made in the following). Equation (20.15) represents, for homogeneous targets, the Maximum Likelihood Estimator (MLE) of the interferometric phase [39]. In the following, we will consider the single look case, with $N_L = 1$.

From (20.12) and (20.13), it is easy to show that the interferometric phases are related to the observed scene height profile through the well known mapping [3,53]:

$$\Psi_{12} = \left\langle \frac{4\pi}{\lambda} \delta r + \Omega_{12} \right\rangle_{2\pi} = \langle \varphi_{12} + \Omega_{12} \rangle_{2\pi}, \quad (20.16)$$

where $\langle \cdot \rangle_{2\pi}$ denotes the modulo 2π operation and

$$\Omega_{12} = \arg(X_1 X_2^*) = \langle \Omega_1 - \Omega_2 \rangle_{2\pi} \quad (20.17)$$

is the decorrelation phase noise related to the phase differences of the speckle. In Eq. (20.16) it has been assumed, as commented before, that the scattering phase term φ_ρ , is constant in the two SAR images, so their differences vanishes.

The problem to be solved in across track InSAR consists of estimating the height values $h(n, m)$, starting from the measured (then, noisy) wrapped phases $\Psi_{12}(n, m)$. Such problem is worldwide known as phase-unwrapping problem, as it amounts to find the unwrapped phase $\hat{\Psi}_{12}$ (not constrained to belong to the interval $[0, 2\pi)$) corresponding to the measured wrapped phase Ψ_{12} (constrained to belong to the interval $[0, 2\pi)$) [55]:

$$\hat{\Psi}_{12} \cong \Psi_{12} + 2i\pi, \quad i = \dots, -2, -1, 0, 1, 2, \dots \quad (20.18)$$

The unwrapped phase will be proportional to an estimate of the height h , according to the model given by Eq. (20.12):

$$\hat{\Psi}_{12} = \frac{4\pi b_\perp}{\lambda r \sin \vartheta} \hat{h} = \alpha_{12} \hat{h}. \quad (20.19)$$

Once that the phase unwrapping problem (20.18) has been solved, an estimate of the height profile can be obtained from Eq. (20.19).

Equation (20.19) shows how much sensitive is the unwrapped phase with the height. Considering a fixed geometry for the satellite (or airplane) carrying the SAR antennas (R_0 and ϑ are fixed), it is easy to understand that the larger the orthogonal baseline and the larger the frequency, the more sensitive is the SAR interferometer. In other words, if we want to measure the height profile h , it could seem more convenient to use a larger baseline and a higher frequency, because for a given variance of the phase noise, the corresponding height variance (inaccuracy) decreases (see Eq. (20.19)). However, an increase of the baseline value may contributes to decorrelate the two speckle phase contributions (Ω_1 and Ω_2), thus increasing the interferometric phase noise (geometrical or spatial decorrelation) [56,57]. For a distributed scattering the correlation between the two speckle contributions decreases linearly with the increase of the baseline [37,39], for a point scatterer the decorrelation disappear because such targets are not affected by speckle. The difference between such two scattering mechanisms influences also the multipass interferometric processing chains (see Section 2.20.4). In any case an increase in the baseline impact also the degree of complexity of the phase unwrapping step described in the following sections.

Before describing in the following sub-section part the several methods that have been proposed in the scientific literature to solve the phase unwrapping problem [42, 58–63], it is important to describe the random nature of the SAR complex signal and of the SAR phase terms.

Consider the random terms X_k , $k = 1, 2$ given by (20.14), representing the multiplicative speckle noise present in the SAR complex signals Z_k , $k = 1, 2$ given by (20.13). Such random terms can be modeled as zero mean, mutually correlated Gaussian complex variables with unit variance, which assume uncorrelated values in adjacent pixels, and have real and imaginary parts mutually uncorrelated [52]. By understanding the dependence on range and azimuth discrete co-ordinates n and m , for the sake of simplicity of notation, we can consider the vector:

$$\mathbf{X} = [X_{c1} \ X_{c2} \ X_{s1} \ X_{s2}]^T, \quad (20.20)$$

whose (real valued) elements X_{ck} and X_{sk} , with $k = 1, 2$, denoting the cosine and sine components of the speckle signal X_k , are zero mean Gaussian random variables. The assumed statistical model implies that [52]:

- a. X_{ci} and X_{sk} are independent $\forall i, k$: $E[X_{ci}X_{sk}] = E[X_{ci}]E[X_{sk}] = 0$;
- b. the cross-correlation between X_{ci} and X_{ck} is equal to the cross-correlation between X_{si} and X_{sk} , $\forall i, k$: $E[X_{ci}X_{ck}] = E[X_{si}X_{sk}]$.

Note that the X_{ci}, X_{sk} independence is true if the speckle band-pass spectrum is Hermitian respect to the central frequency, assumption which can be considered always satisfied since the speckle vector \mathbf{X} is related to the complex envelope of a modulated real signal [64]. Moreover, from the assumption (b), it stems that $E[X_{ck}^2] = E[X_{sk}^2]$, $k = 1, 2$.

In these assumptions, the probability density function of the vector \mathbf{X} is given by [65]:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^3 |\mathbf{C}|^{1/2}} \exp \left\{ -\frac{1}{2} \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} \right\}, \quad (20.21)$$

where $x = [x_{c1} \ x_{c2} \ x_{s1} \ x_{s2}]^T$, and \mathbf{C} is the covariance matrix given by:

$$\mathbf{C} = \begin{bmatrix} 1 & \gamma_{12} & 0 & 0 \\ \gamma_{12} & 1 & 0 & 0 \\ 0 & 0 & 1 & \gamma_{12} \\ 0 & 0 & \gamma_{12} & 1 \end{bmatrix}, \quad (20.22)$$

where γ_{12} is the *correlation coefficient* of X_{c1} and X_{c2} , which by virtue of assumption (b) assume the same value of the correlation coefficient of X_{s1} and X_{s2} , given by:

$$\gamma_{12} = \frac{E[X_{c1}X_{c2}]}{(E[X_{c1}^2]E[X_{c2}^2])^{1/2}} = \frac{E[X_{s1}X_{s2}]}{(E[X_{s1}^2]E[X_{s2}^2])^{1/2}}, \quad (20.23)$$

where $E[\cdot]$ denotes expectation. Note that in Eq. (20.23) the terms present at the denominator are equal to one, as they represent the unit variance of the considered processes, so that their explicit presence should

not be necessary. Nonetheless, we use such definition, as it is valid also in the case of not-normalized processes.

We note that, according to the assumptions (a) and (b), γ_{12} given by (20.23) is equal to the interferometric *coherence* usually employed in InSAR systems [57], defined starting from complex signals [3]:

$$\begin{aligned}\eta_{12} &= \frac{E[X_1 X_2^*]}{(E[|X_1|^2] E[|X_2|^2])^{1/2}} = \frac{E[X_{c1} X_{c2}]}{(E[X_{c1}^2] E[X_{c2}^2])^{1/2}} = \gamma_{12}, \\ \eta_{C,12} &= \frac{E[Z_1 Z_2^*]}{(E[|Z_1|^2] E[|Z_2|^2])^{1/2}} = \frac{E[X_1 X_2^*]}{(E[|X_1|^2] E[|X_2|^2])^{1/2}} e^{j\frac{4\pi}{\lambda}(r_1 - r_2)} \\ &= \eta_{12} e^{j\frac{4\pi}{\lambda}(r_1 - r_2)} = \gamma_{12} e^{j\frac{4\pi}{\lambda}(r_1 - r_2)}.\end{aligned}\quad (20.24)$$

Note that first result of Eq. (20.24) ($\eta_{12} = \gamma_{21}$) implies that the coherence η_{12} of the (complex) speckle noise is *real valued*, due to the above assumptions (a) and (b). Note also that the module of the coherence $\eta_{C,12}$ of the complex received signals Z_1 and Z_2 is, in module, equal to the coherence η_{12} of the (complex) speckle noise.

Starting from Eq. (20.21), it is possible to derive, by a change of variable from Cartesian to polar (from X to A_1 , A_2 , Ω_1 , Ω_2), and following integration with respect to A_1 and A_2 , the pdf of the single-look speckle phase difference $\Omega_{12} = (\Omega_1 - \Omega_2)$ [53]:

$$\begin{aligned}f_{\Omega_{12}}(\omega_{12}) &= \frac{1}{2\pi} \frac{1 - |\gamma_{12}|^2}{1 - |\gamma_{12}|^2 \cos^2(\omega_{12})} \\ &\times \left\{ 1 + \frac{|\gamma_{12}| \cos(\omega_{12}) \cos^{-1}[-|\gamma_{12}| \cos(\omega_{12})]}{[1 - |\gamma_{12}|^2 \cos^2(\omega_{12})]^{1/2}} \right\}, \quad \omega_{12} \in (-\pi, \pi],\end{aligned}\quad (20.25)$$

where the dependence of phase difference ω_{12} and γ_{12} on (n,m) has been, as before, understood.

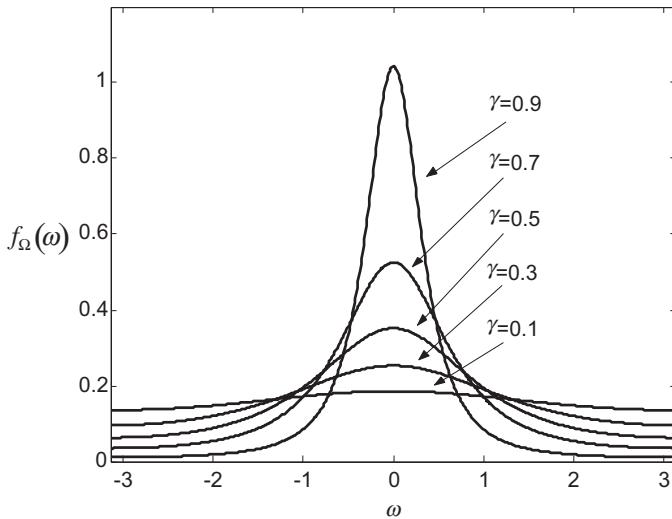
The coherence coefficient γ_{12} is influenced by all factors that cause differences between the two complex speckle images X_1 and X_2 . The larger these differences, the smaller the coherence coefficient's value. A coherence reduction can be induced by actual physical changes occurring between the acquisition times of the two data sets (temporal decorrelation) and/or by changes of the ground reflectivity when it is seen from different angles (spatial decorrelation) [57]. Note that also the coherence is a function of the ground coordinate pair (n,m) , so that it may change across the image.

The plot of the speckle phase difference pdf (20.25) for different values of the coherence coefficient is given in Figure 20.8. It has to be noted that the pdf become less peaked by reducing the coherence value. The smaller the coherence value, the larger the variance of the speckle phase noise.

The joint pdf $f_{\Psi_{12}}(\psi_{12}; h)$ of the interferometric phases can be obtained from the joint pdf of the speckle noise phase $f_{\Omega_{12}}(\omega_{12})$ given by Eq. (20.25) and plotted in Figure 20.8, exploiting the random variable transformation given by Eq. (20.16), which leads to:

$$f_{\Psi_{12}}(\psi_{12}; h) = f_{\Omega_{12}}(\psi_{12} - \alpha_{12}h), \quad (20.26)$$

where the dependence of phase difference ψ_{12} , h , and γ_{12} on (n,m) has been again understood. The pdf's (20.26) have the same shape of pdf of Eq. (20.25), but they are centered on the value $\alpha_{12}h$.

**FIGURE 20.8**

Interferometric phase pdf plotted for different coherence values (0.01, 0.1, 0.25, 0.5, 0.7, and 0.9).

Such pdf family, that as it can be noted is parametrized by the height h , will be used as starting point of some of the phase unwrapping methods described in the following section.

The pdf (20.26) of the interferometric phase is strongly influenced by the *coherence*. Such parameter can be written as the product of four main contributions [3,39,57]:

$$\gamma = \gamma_{\text{SNR}} \gamma_G \gamma_Z \gamma_T, \quad (20.27)$$

where γ_{SNR} represents the influence of thermal noise in the receiver, γ_G represents the decorrelation effects due to the different SAR view acquisition angles, depending upon the spatial baseline, γ_Z represents the decorrelation effects due to volume scattering mechanisms, and γ_T represents the so called temporal decorrelation effects [3,4].

The first factor in Eq. (20.27) can be computed starting from the circular Gaussian and independent nature of thermal noise and is given by:

$$\gamma_{\text{SNR}} = \frac{1}{\sqrt{1 + 1/\text{SNR}_1} \sqrt{1 + 1/\text{SNR}_2}}, \quad (20.28)$$

where SNR_1 and SNR_2 are the signal to noise ratio on the two receiving SAR interferometric antennas [4,53,57].

The second factor in Eq. (20.27) is the so-called geometric coherence, also referred to as angular or baseline coherence; it is present for all scattering situations, it depends on the system parameters and on the overall observation geometry, including the different SAR view acquisition angles, and depending upon the spatial baseline.

Geometric coherence values can be easily computed, in case of flat terrain, and for a white scattering process, leading to [3]:

$$\gamma_G = \begin{cases} \left(1 - \frac{b_{\perp}}{b_c}\right) & \text{for } |b_{\perp}| \leq b_c, \\ 0 & \text{elsewhere,} \end{cases} \quad (20.29)$$

where b_{\perp} is the orthogonal baseline and b_c is the orthogonal critical baseline given by:

$$b_c = \frac{\lambda r}{2\delta_r \cot(\vartheta - \beta)}, \quad (20.30)$$

where all symbols in Eq. (20.30) have been previously introduced.

Geometric decorrelation effects, for flat terrain geometry, can be explained also from the so-called spectral shift effect [55]. This interpretation allows also to derive a filtering strategy of the interferometric channels aimed at mitigating such decorrelation. Such proper processing is also called “common band filtering” [55, 66], because requires to process, to maximize the geometric coherence value, the common (overlapped) part of the spectrum of the two interferometric signals. The larger the baseline, the larger the terrain slope, the less the common part of the two spectra, and the larger the decorrelation effects. For a non-flat topography the approach in [67] can be adopted. For the ERS and ASAR-ENVISAT sensors, the critical baseline is of about 1100 m for $\beta = 0$, while for the last generations high resolution systems such as COSMO-Skymed and TerraSAR-X, this value is significantly enlarged. Therefore, for these new generation sensors, thanks to fact that the distribution of the baseline values is bounded by an “orbital tubes” significantly smaller than the critical baseline, such common band filtering is typically not required.

Similar effects can be induced also in the azimuth direction by the presence of variations of the azimuth antenna pointing [67]. This effect can be critical in some acquisition modes where antenna steering is present, such as in ScanSAR and Spotlight modes. Also in this case, the larger the acquisition geometric diversity, the less the spectra overlapping, and the larger the decorrelation effects.

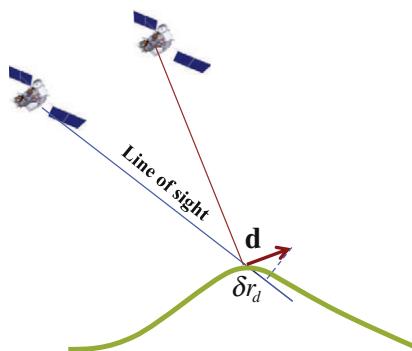
The third factor in Eq. (20.27) γ_c , is the volume coherence, it is due to volume scattering, and it is the effect induced by the scattering layer to increase the size projected range cell, and consequently, to decrease the correlation distance. As in the case of γ_G , it is dependent on the spatial baseline.

The last factor in Eq. (20.27) γ_T represents the so called temporal decorrelation [57] due to the instability of scattering mechanisms in the two different acquisition times, as the structure of the scatterer can change in the meantime between the two acquisitions. Such effect can be very important when the two SAR images are acquired at distance of several days or months in the case of vegetated or agricultural areas, or in presence of different climatic conditions.

2.20.3.3 The differential SAR interferometry technique for measuring displacements

Differential Interferometry (DInSAR) is a particular configuration of SAR interferometry. The reference geometry is the same of the classical InSAR case, but the target on the ground is allowed to move, displacing of say d , between the two successive passes (see Figure 20.9).

In the following, for sake of simplicity, we indicate deterministic and stochastic terms all with non-capital symbols: the nature is specified whenever ambiguous. In this case, the interferometric phase is

**FIGURE 20.9**

Differential interferometric geometry.

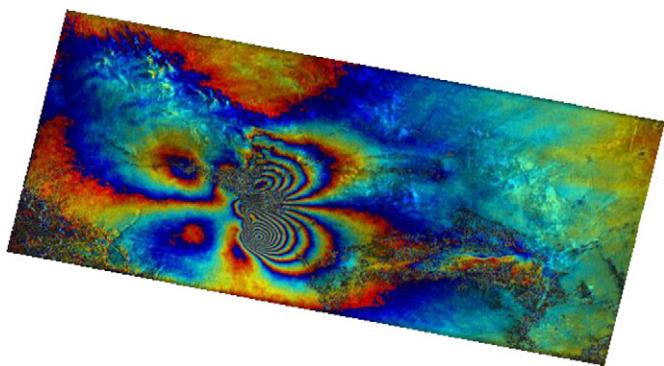
composed by three main factors:

$$\varphi = \frac{4\pi}{\lambda} \delta r = \frac{4\pi}{\lambda} \delta r_d + \varphi_z + \varphi_a + \varphi_n, \quad (20.31)$$

where δr_d is the measured range variation that, in the far field observation approximation, is equal to the component of the displacement along the line of sight, φ_z is the phase contribution corresponding to the target height as in Eq. (20.12), φ_a is a stochastic term associated with the variation, between the two passes, of the wave propagation delay through the atmosphere, φ_n is the phase noise which in this case includes also the temporal decorrelation effects in addition to the decorrelation noise due to the speckle (ω in Section 2.20.3.2). In the cases in which the topographic contribution is limited, that is if the baseline is negligible and/or an external DEM is available to compute and cancel out φ_z from Eq. (20.31), and in the case of a predominance of the deformation component and/or limited effects of atmosphere, displacements can be measured to accuracy which are on the order of the wavelength. By using this classical two passes DInSAR configuration in the past Scientists have been able to capture the surface deformation field generated by major earthquakes, or highlight deformation associated with volcanic activities.

The idea of mapping ground deformation via the interference of signals acquired by SAR systems was demonstrated for airborne systems in [43] and for the very first time using real data from the European Remote Sensing Satellite (ERS) in keystone experiments by [68], for ice-stream velocity measures in Antarctica, and by [69] for the co-seismic deformation field generated by the Landers earthquake (CA-USA). The Landers result received cover of Nature (vol. 364, 8 July 1993, Issue No. 6433) with a title “The image of an Earthquake” that translates the importance of the achievements and of the DInSAR technology with reference to application to seismic and to geo-hazards in general.

Today, with the availability of many SAR sensors with interferometric capabilities orbiting around the Earth, co-seismic, i.e., before and after main seismic events (see Figure 20.10), DInSAR data are almost analyzed routinely by scientists to study the displacements induced by known and unknown geological faults that are the causes of catastrophic events all over the world.

**FIGURE 20.10**

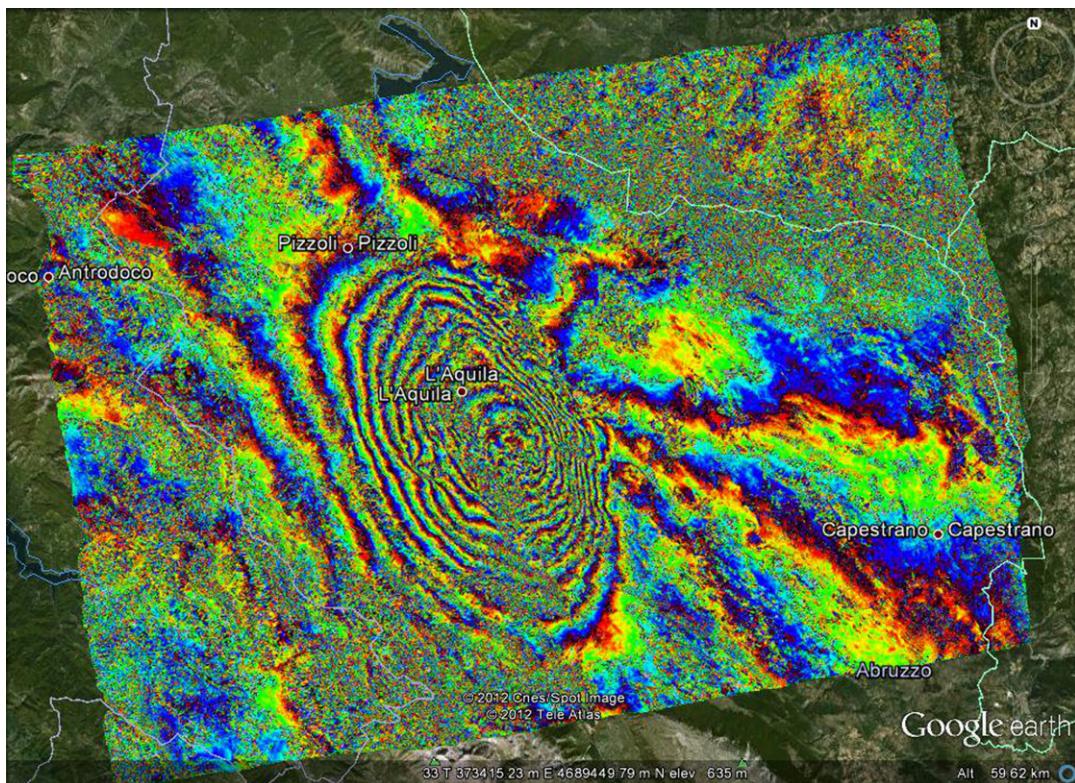
Co-seismic interferogram of the Bam earthquake obtained by a combination of Envisat Advanced Synthetic Aperture Radar (ASAR) Wide Swath Mode (WSM) image with an Image Mode (IM) image.

Polimi/Poliba.

An example of measurement of DInSAR co-seismic displacement is reported to provide an idea of the powerfulness of this technology with reference to the 6.6 Mw Iran earthquake in 2003 that stroke the city of Bam. The image Figure 20.10, shows the co-seismic interferogram obtained by interferometric combination of the SCANSAR (Wide Swath Mode) acquisition of September 24th, 2003 and the STRIPMAP (Image Mode) acquisition of December 3rd, 2003: each color-cycle correspond to 1.55 cm in the line of sight. The coherence of the data and therefore the quality of the interferogram is very high due to the arid nature of the region: each color-cycle correspond to 2.8 cm displacement in the line of sight.

A key factor in such application is associated with the revisiting time. The retired satellites ERS-1, ERS-2, and ENVISAT of the European Space Agency have been the satellites on which repeat pass interferometry has been experimented for the very first time. These satellites were characterized in the normal operational situation by a strip width of approximately 100 Km EW and a revisiting time (i.e., the time necessary to repeat approximately the same orbit) of 35 days. The revisiting time poses a limitation to the minimum number of days necessary to generate an interferogram. The new generation of sensors operating at slightly lower orbits with respect to ERS and ENVISAT, such as TerraSAR-X and Tandem-X allows reducing the revisiting time to 11 days. The Italian COSMO-SkyMed (CSK) constellation is a constellation formed by four SAR satellites that acquires data for interferometric use, regardless of the specific satellite. This peculiarity allows CSK to provide the highest maximum revisit rate of an area of interest, that is one acquisition every 4 days (on average) for the whole constellation, instead of one acquisition every 16 days for a single satellite. CSK and TerraSAR-X provide spatial resolution (between 1×1 and 3×3 m 2) one order of magnitude better than the previous available C-Band satellite SAR data.

These systems operates in X-band and are characterized by higher spatial resolution with respect to the past ESA C-Band satellites; the counterpart of these advantages are however the reduced swath coverage which in the classical stripmap imaging mode reduces to 40 Km in the EW direction. An example of co-seismic interferogram obtained with a 8 days temporal baseline form COSMO-Skymed data is discussed in the following. On 6 April, 2009 the MW 6.3 L'Aquila earthquake occurred in the Central Apennines

**FIGURE 20.11**

Co-seismic interferogram of 6 April 2009 6.0 Mw L'Aquila earthquake in Italy. COSMO-Skymed acquisitions of 4 April 2009 and 12 April 2009.

(Italy) causing extensive damage to the town of L'Aquila and killing 300 inhabitants. The event epicenter was located few kilometers southwest of the town of L'Aquila, the main shock nucleated at a depth of ≈ 9 km, was preceded by a preseismic sequence with the largest shock having a ML 4 magnitude, and was followed by a vigorous aftershock sequence. In Figure 20.11 it is shown the interferogram evaluated by the CSK pair of April 4th and April 12th: the activated fault is located NW-SE emerging to the right of the dense fringes area: each color-cycle correspond to 1.55 cm in the line of sight.

Since then, many experiments showed the potentiality of the technique in detecting deformation phenomena not only associated with earthquakes [70] but also in volcanic areas [71, 72] and of glaciers [73].

However, in order to fully exploit the potentiality of the SAR technology in order to measure deformation with a centimetric/millimetric accuracy two or few images are typically not sufficient. At such accuracy level the presence of the atmospheric component and additional disturbing contributions such as orbital inaccuracies cannot be in fact neglected. For sake of simplicity we indicate still with φ the

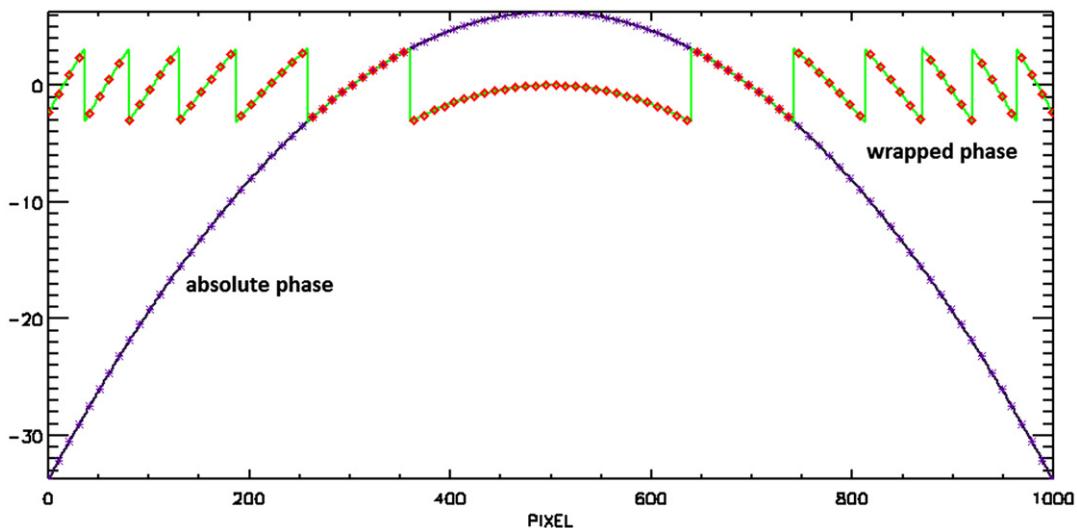
differential interferometric phase obtained after the subtraction of the contribution associated with the DEM (differential interferogram). We have therefore:

$$\varphi = \frac{4\pi}{\lambda} \delta r = \frac{4\pi}{\lambda} \delta r_d + \varphi_z + \varphi_a + \varphi_o + \varphi_n, \quad (20.32)$$

where φ_o is associated with orbital inaccuracies which affect the computation of the phase contribution associated with the topography (baseline error): as for Eq. (20.31), the noise term is associated with the noise contributions such as decorrelation of the response not only due to variation of the speckle contribution due to the angular imaging diversity (spatial baseline decorrelation), but also to changes of the backscattering response over the time (temporal decorrelation). Availability of on board GPS systems allows significantly mitigating the effects of orbital errors; for airborne systems which are subject to trajectory deviations due to turbulence the GPS must be integrated with accurate inertial navigation systems. Due to the DEM subtraction φ_z is now the residual target height, i.e., the height of the target with respect to the reference DEM. Accordingly, to have the possibility to achieve measurement of small deformation components, use of accurate external DEM as well as small baseline separations, is mandatory. In any case the atmospheric component plays a major role because it causes the presence of errors which are spatially correlated and therefore may be mixed with possible deformations.

The atmospheric contribution is typically separated in two components: a turbulent component which is associated with the air inhomogeneity and that causes a spatial variation of the Atmospheric Phase Delay (APD) and a stratified component which is associated with the vertical stratification of the atmosphere. Both these terms occur in the lower part of the atmosphere, the troposphere, whereas contribution in the upper part of the atmosphere mainly show contributions which are very low spatial variable and can be misinterpreted as orbital inaccuracies. The former is commonly referred to as wet component and is dependent on the relative humidity, the latter, also referred to as hydrostatic or improperly as dry component, is responsible of a contribution which is highly correlated with the topography and therefore almost negligible on quasi flat areas. A model that describes the statistical behavior of the turbulent component is due to Kolmogorov. In this case the turbulence is assumed spatially stationary and isotropic. The refractivity index, i.e., the excess in parts per million of the reflectivity index with respect to the vacuum, that provided the increase in the path difference due to the crossing of the atmosphere can be modeled in terms of the variogram, i.e., the variance of the difference of the refractivity contributions between two points. For separations below the order of a kilometer, the variance of the difference of the refractivity between two points small and decreasing with a power of 2/3 of the distance. A more thoughtful characterization and analysis of the tropospheric contribution can be found in [74].

The temporal correlation of the atmosphere is however typically low: this means that APD contributions over different epochs can be averaged together to diminish its contribution on the path difference. Therefore to measure small (up to millimeter per year) displacements and to handle the problem of monitoring at high resolution ground scatterers, techniques based on the use of several images acquired over the same scene have been developed. In fact, by exploiting a higher dimensional acquisition space, the “phase firms” of the different components such as DEM error, displacement, and APD can be deterministically or stochastically characterized and estimated directly from the received data. This topic is treated in more details in Section 2.20.5.

**FIGURE 20.12**

Phase unwrapping problem in 1D: example of the effects of wrapping.

2.20.3.4 Phase unwrapping

Unwrapping aims at reconstructing an unrestricted (absolute²) signal starting from a measured wrapped signal version restricted a reference interval. In the case of interferometry, the phase is intrinsically wrapped in the $(-\pi, \pi)$ interval being extracted from complex values. Accordingly, we have:

$$\psi = \langle \varphi \rangle_{2\pi} = [(\varphi + \pi) \bmod 2\pi] - \pi, \quad (20.33)$$

where ψ and φ are the restricted (wrapped) and unrestricted (absolute) phase. Phase unwrapping is a step necessary to reconstruct a phase signal $\hat{\varphi}$ which is an estimate of φ .

Figure 20.12 reports an example in a 1D case: phase unwrapping aims to estimate the absolute phase (shown in blue) starting from the measured restricted phase (shown in red).

It should also be noted that the term “absolute” phase in the interferometric context is typically used to refer to the phase corresponding to δr which has been corrected by an offset to account for the correct number of global cycles which are lost due to the wrapping operator and for the timing errors. In SAR interferometry such an offset is commonly evaluated after phase unwrapping by using one or more reference points with a known topography.

The problem of unwrapping is ambiguous as it admits in principle infinitely many solutions (the wrapped phase can be itself a possible absolute phase), and a reasonable solution can be obtained by imposing a certain degree of continuity: the absolute phase in Figure 20.12 is, among all possible absolute phase functions corresponding to the measured wrapped phase, the continuous one. Unfortunately,

²The term “absolute” is in the phase unwrapping context related to the restriction operator and does not concern the absolute value operator.

in real cases, in addition to the noise, the problem is further complicated by the presence of a finite sampling rate, see the measured red diamond samples in Figure 20.12 and the absolute black star samples corresponding to the absolute phase. Moreover, the actual solution could be locally not continuous.

In the following, some popular approaches to solve the phase unwrapping problem and to recover the height profile of the ground scene will be described.

2.20.3.4.1 Residue cut algorithms for PhU

Residue cut [75], also known as branch-cut algorithms have been a workhorse for the PhU for many years prior to the advent of more effective solutions based on Least Square and optimizations based on Linear Programming algorithms. The starting point of this approach is the estimation of the absolute phase variations along arcs $\Delta\varphi$ wrapped phase variation $\Delta\psi$. In particular it results that

$$u = \langle \Delta\psi \rangle_{2\pi} = \langle \Delta\varphi \rangle_{2\pi} \underset{\substack{|\Delta\varphi| < \pi \\ \equiv}}{=} \Delta\varphi, \quad (20.34)$$

where $\Delta f = f(n-1) - f(n)$ and where the first equality is taken as the definition of u that is of wrapped variation of the wrapped phase. Equation (20.34) states that, by wrapping the variation from sample to sample of the measured wrapped phase an estimate of the absolute phase variations can be retrieved provided that the latter is limited to the restriction interval, i.e., within the $(-\pi, \pi)$ interval.

Following the stage of estimation of the absolute phase differences, an integration step of the estimated variations must be implemented to pass from $\Delta\varphi$ to φ . Possible errors on a phase variation estimate, either due to an intrinsic variation of the absolute phase or to a missing wrapping jump due to the noise, propagates in all the subsequent samples during the integration process.

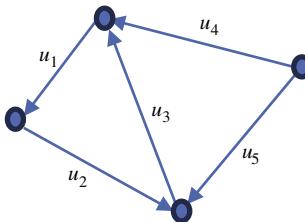
A way to control such errors is by measuring redundant variations: in the 1D case this is only possible if we measure the variations over non adjacent samples: this operation is however critical because of the higher probability to variations of the absolute phase larger than π , thus impacting the validity of the approximation in Eq. (20.34). Fortunately, the 2D case allows having higher redundancy degree of even limiting the measurement of spatial variations over adjacent samples [76].

The basic idea of the residue cut algorithm is to follow elementary closed circuits defined on the set of image pixel and check for the presence of inconsistencies in the estimate of the absolute phase variation.

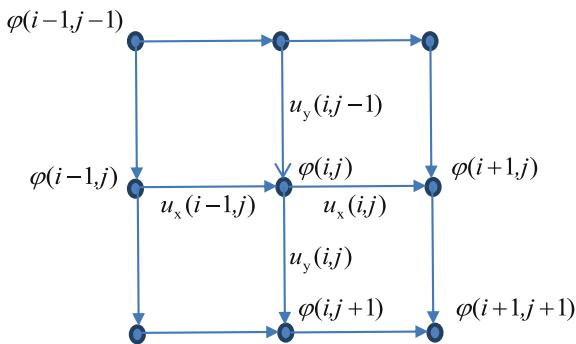
Following the stage of estimation of the absolute phase differences, the algorithm proceeds with the integration of the estimated variations but carries out a consistency check on the loops. By referring to Figure 20.13, we refer for simplicity to two elementary loops: the triangular shape of the circuits is a choice that is typically carried out when the phase unwrapping has to be carried out on a sparse grid. Assuming for instance an anticlockwise direction of the loops, it is clear that if the estimates of the true phase variation are correct, i.e., corresponding to the absolute phase we have:

$$\begin{aligned} u_1 + u_2 + u_3 &= \Delta\varphi_1 + \Delta\varphi_2 + \Delta\varphi_3 = 0, \\ u_5 + u_3 - u_4 &= \Delta\varphi_5 + \Delta\varphi_3 - \Delta\varphi_4 = 0. \end{aligned} \quad (20.35)$$

If, however, due to the wrapping operator in (20.34), at least one of the estimate is affected by errors, say u_3 for which we have that $u_3 = \Delta\varphi_3 - 2\pi$ then the closed loops on the left in Eq. (20.35) do not

**FIGURE 20.13**

The integration of the estimated absolute phase differences over closed paths.

**FIGURE 20.14**

Representation of the phase and phase variations on elementary closed loops for the phase unwrapping on a regular 2D grid.

sum up to zero:

$$\begin{aligned} u_1 + u_2 + u_3 &= \Delta\varphi_1 + \Delta\varphi_2 + \Delta\varphi_3 = 2\pi, \\ u_5 + u_3 - u_4 &= \Delta\varphi_5 + \Delta\varphi_3 - \Delta\varphi_4 = -2\pi. \end{aligned} \quad (20.36)$$

In this situation, it is the loops are said to be affected by residues. It is clear that due to a choice of the loop orientation, residues appear always in pair. The line connecting the two residues, which are supposed to be located at the center of the loops, is referred to as cut and intersects the wrong absolute phase estimate: this phase variation is eliminated during the final integration patch covering all pixels.

2.20.3.4.2 Least squares and Green's solution to PhU problem

We refer, for sake of simplicity, to a regular 2D grid (see Figure 20.14): it is possible to measure the horizontal $u_x(i,j)$ and vertical $u_y(i,j)$ variations between adjacent samples, leading to the measurement from the wrapped phase values of a 2D field $\mathbf{u} = u_x \hat{\mathbf{x}} + u_y \hat{\mathbf{y}}$, with $u_x(i,j) = \langle \psi(i+1,j) - \psi(i,j) \rangle_{2\pi}$ and $u_y(i,j) = \langle \psi(i,j+1) - \psi(i,j) \rangle_{2\pi}$, that represents an estimate of the gradient of the absolute phase.

As already explained in the previous section, the variation field should be zero curl, $\nabla \times \mathbf{u} = 0$. Obviously, due to the 2π (multiple) errors that affect the variation field \mathbf{u} , this condition is not always

satisfied and therefore the result of the process of integration of the spatial variations to retrieve the unwrapped phase is dependent on the specific path. Instead of using circulation over elementary square paths to locate the presence of vortex (residues) and to trace cuts between opposite residues to avoid the integration path crossing areas over which summation over closed path is not zero, Least Squares and Green's algorithm try to find a solution that carries out a global integration in such a way to mitigate the error propagation.

In particular the Least Squares approach looks for φ such that $\nabla\varphi$ is the “closest” (in some norm) to \mathbf{u} ; more specifically:

$$\varphi = \arg \min \| \nabla\varphi - \mathbf{u} \|_{L^k} \quad (20.37)$$

with typically $k = 0, 1, 2$. The least square solution ($k = 2$) can be easily achieved by solving the Poisson equation [62, 77]:

$$\nabla^2 \varphi(\mathbf{r}) = \nabla \cdot \mathbf{u}(\mathbf{r}), \quad (20.38)$$

for any \mathbf{r} internal to the domain of interest (say S), with the following boundary condition:

$$\nabla\varphi(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}) = \mathbf{0}, \quad (20.39)$$

for any \mathbf{r} on the boundary (say C), $\hat{\mathbf{n}}(\mathbf{r})$ being the normal unit vector on the boundary curve.

An effective iterative implementation on a regular 2D grid can be achieved by iterating on each (internal) pixel the following equation:

$$\begin{aligned} 4\varphi(i, j) &= \varphi(i-1, j) + u_x(i-1, j) + \varphi(i, j-1) + u_y(i, j-1) \\ &\quad + \varphi(i+1, j) - u_x(i, j) + \varphi(i, j+1) - u_y(i, j). \end{aligned} \quad (20.40)$$

assuming left and right values of the unknown to be corresponding to the current and previous iteration, respectively. The iterative solution of (20.40) is appealing because it can be extended also to a sparse 2D grid via the use of triangulations.

Iterative algorithms might be however time demanding depending on the size of the processed image: a more time-effective solution for a regular 2D grid is provided by frequency domain analysis. In particular, by using the Green's function formulation it can be shown that the solution to (20.38) and (20.39) on a surface S with boundary C can be explicitly written as:

$$\varphi(\mathbf{r}') = - \int \int_S dS \nabla g(\mathbf{r}' - \mathbf{r}) \cdot \nabla\varphi(\mathbf{r}) + \oint_C dc\varphi(\mathbf{r}) \frac{\partial g(\mathbf{r}' - \mathbf{r})}{\partial \mathbf{n}}, \quad (20.41)$$

wherein $\nabla\varphi$ is approximated (in the discrete space case) with the variations measured from the wrapped data \mathbf{u} and where $g(\mathbf{r})$ is the free-space Green solution of the Laplace equation, i.e., :

$$\nabla^2 g(\mathbf{r}) = \delta(\mathbf{r}), \quad (20.42)$$

that is

$$g(\mathbf{r}) = \frac{1}{2\pi} \ln |\mathbf{r}|. \quad (20.43)$$

The solution in (20.41) is attractive because it allows to investigate analytically the spreading of errors related to the errors related to the substitution of the true gradient with the measured gradient, i.e., $\nabla\varphi \cong \mathbf{u}$ [61].

The Green solution requires, on the other hand, the knowledge of the true phase on the border, this problem can be solved by adopting an iterative scheme in which starting from $\varphi = 0$, $\varphi(\mathbf{r})$ is evaluated in

all the points, including the boundary C . In the case of sparse grid of measurements where the boundary may be characterized by multiple curves, it is preferable to use the FD approximation to the differential equation.

2.20.3.4.3 Minimum cost flow solution to PhU problem

In both methods, FD approximation and Green, the solution do not honour the data, in the sense that φ is a multiple of 2π because the Least Square solution intrinsically diffuse possible error in \mathbf{u} over different pixels to achieve a global square minimum. Methods like the Minimum Discontinuity (MD) [78] and Minimum Cost Flow (MCF) allow “honouring” the data, by directly search in a discrete space the 2π multiple correction. The MD approach seeks for the $2\pi\mathbf{k}$ field that provides $\varphi = \psi + 2\pi\mathbf{k}$, with ψ collecting all the measured wrapped phase on the selected (full or sparse) grid, showing the minimum discontinuity between adjacent points. The MCF algorithm [60, 79], operates in a similar way by seeking for the field of integers \mathbf{k} of correction to \mathbf{u} (with minimum norm) such that the corrected variation field on a network defined over the set of available pixel is characterized by a null rotational component, i.e., it solves the problem:

$$\begin{cases} \hat{\mathbf{k}} = \arg \min_{\mathbf{k}} J(\mathbf{k}), \\ J(\mathbf{k}) = \|\mathbf{k}\|_{L^p}, \\ \text{subject to } \nabla \times (\mathbf{u} + 2\pi\mathbf{k}) = \mathbf{0} \text{ on all closed loop,} \end{cases} \quad (20.44)$$

where L^p is the (typically weighted) LP norm.

The case $p = 1$ is chosen for the capability to limit the number of points in which the correction is carried out, and at the same time allows using very efficient Linear Programming solvers [60].

More specifically typically a triangulation is carried out to define a network over the sparse grid of pixels selected according to a sufficient level of coherence. Letting N_{aS} be number of arcs of the network, the objective function to minimize is written as:

$$J(\mathbf{k}) = \sum_{j=1}^{N_{aS}} w(j) |k_j|, \quad (20.45)$$

where $w(j)$ is the weight associated with the j th arc, the PhU amounts to solve the following problem

$$\begin{cases} \hat{k}_j = \arg \min_{k_j} \sum_{j=1}^{N_{aS}} w(j) |k_j|, \\ \text{subject to } 2\pi \sum_{j \in C_l} k_j = - \sum_{j \in C_l} u(j) \text{ for all out of } L \text{ loops} \end{cases} \quad (20.46)$$

with k_j integer, and being C_l the generic elementary closed loop. Such a problem can be recast in a linear form as described in Ref. [60]. In particular the following change of variables is implemented:

$$\begin{cases} k_j^+ = \max \{0, k_j\} \\ k_j^- = -\min \{0, k_j\} \end{cases} \quad j = 1, \dots, N_{aS}, \quad (20.47)$$

thus leading to the following expression for the unknown vector and the objective function, respectively:

$$k_j = k_j^+ - k_j^-. \quad (20.48)$$

Therefore the problem (20.46) can be recast in the more feasible form:

$$\begin{cases} \left\{ \hat{k}_j^+, \hat{k}_j^-, j = 1, \dots, N_{aS} \right\} = \arg \min \left(\sum_{j=1}^{N_{aS}} w(j) k_j^+ - w(j) k_j^- \right), \\ \text{subject to } 2\pi \sum_{j \in C_l} k^+(j) - 2\pi \sum_{j \in C_l} k^-(j) = - \sum_{j \in C_l} u(j) \text{ for all } l \text{ out of } L \text{ loops,} \end{cases} \quad (20.49)$$

which is now linear with respect to the all the (positive integer) unknown k_j^+ and k_j^- . It is a typical problem of Integer Linear Programming (ILP) solvable with computationally efficient techniques [60, 79]. Thanks to the particular structure of the network, which is based on a triangulation, the ILP problem in (20.49) can be cast as a problem of flow optimization problem for which very efficient MCF solvers can be used. For all j , after that the integers k_j^+ and k_j^- are estimated, k_j is evaluated via (20.48) and used to correct the estimate of the variation over the arc of the network, u_j , spatial integration (which is no more dependent on the integration path) is applied to retrieve the unwrapped interferogram.

Reference [76] provides a complete comparison between the above PhU approaches.

2.20.4 Multibaseline SAR interferometry

The statistical approaches developed to solve the problem of Phase Unwrapping and SAR interferometry are based on the exploitation of the statistical model of the SAR interferometric signal [42]. The statistical techniques are in general based on the use of more than two SAR complex images (at least two interferograms) [40–42, 80–82], and are often referred as multi-channel SAR Interferometry.

The basis geometry of a multi-channel SAR system is reported in Figure 20.15.

First consider the dual channel case (classical SAR Interferometry). As it has been shown in the previous sections, the actual measured values of the interferometric phase Ψ_{12} differ from the nominal ones by virtue of phase noise effects, as modeled by pdf described by Eq. (20.26). Once the data (Ψ_{12}) has been observed, Eq. (20.26) can be seen as a function of the unknown parameter h , providing the single-interferogram likelihood function.

The plot of the likelihood function (Eq. (20.26) as function of h) for a given measured value of the interferometric phase Ψ_{12} , and consequently of h , is shown in Figure 20.16, for two different coherence values ($\gamma = 0.8$ for the solid line, $\gamma = 0.4$ for the dotted line). It shows very clearly that the likelihood function, due to its periodic nature (the period is $\alpha_{12}/2\pi$), exhibits an infinite number of global maxima. Note also that the effect of different coherence values (the smaller the coherence value, the larger the variance of the pdf of the data) is to change the amplitude and the curvature of the likelihood function.

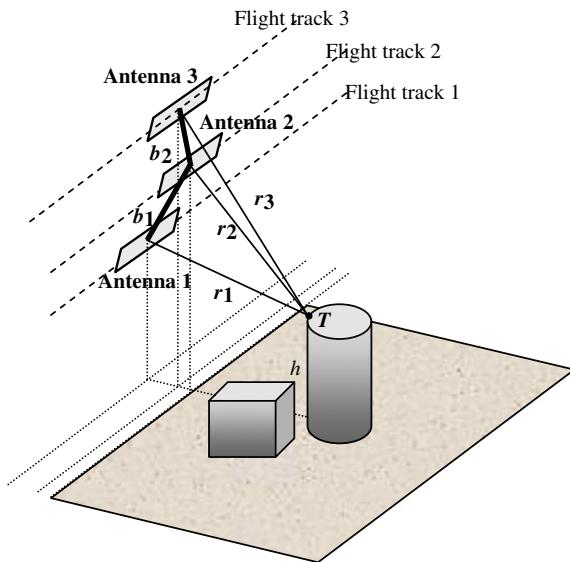
The Maximum Likelihood solution of the Single Interferogram InSAR problem is given by:

$$\hat{h}_{ML}(n, m) = \arg \max_h f_{\Psi_{12}}(\psi_{12}(n, m); h(n, m)). \quad (20.50)$$

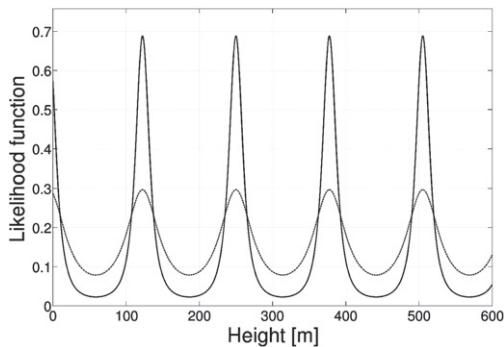
Problem (20.50) admits the following infinite solutions [40, 42, 83]:

$$\hat{h}_{ML} = \frac{\psi_{12}}{\alpha_{12}} \pm k \frac{2\pi}{\alpha_{12}}, \quad k = 0, 1, 2, \dots, \quad (20.51)$$

as a direct consequence of the periodic nature of the likelihood function.

**FIGURE 20.15**

Interferometric SAR *multi-channel geometry*.

**FIGURE 20.16**

Single-interferogram likelihood function for different coherence values ($\gamma = 0.8$ for the solid line, $\gamma = 0.4$ for the dotted line).

In order to resolve this ambiguity, we can introduce additional independent phase measures. Suppose we have multiple measures of wrapped phase Ψ_{1i} , $i = 2, \dots, N$, obtained in N different acquisition conditions. Different wrapped phase values can be obtained, for instance, from SAR raw data acquisition from N_b different flight track (baseline-diversity) as it is shown in Figure 20.15, or in N_b sensor

bands (frequency-diversity). For example, in the SIR-C/X-SAR mission X-, C-, and L-band data were acquired ($N_b = 3$), while in the SRTM mission X- and C-band data were acquired ($N_b = 2$) [50]. Each band could also be divided into sub-bands, so that $N_r \geq N_b$ sets of data at different frequencies could be available [42,80]. The band partition can be operated both along the range frequency band and the azimuth frequency band [42].

Hence, the different wrapped phases are given by:

$$\begin{aligned}\Psi_{1i}(n, m) &= \left\langle \frac{4\pi}{\lambda} (r_1 - r_i) + \Omega_{1i}(n, m) \right\rangle_{2\pi}, \\ \Omega_{1i}(n, m) &= \arg(X_1(n, m)X_i^*(n, m)) = \langle \Omega_1(n, m) - \Omega_i(n, m) \rangle_{2\pi},\end{aligned}\quad (20.52)$$

where Ω_i is the random process representing the phase of *speckle* noise at the i th antenna or in the i th band.

The choice of the number and kind of wrapped phase data sets, is a crucial points. A proper choice of the different flight tracks (hence, of the baselines), and/or of the different bands and sub-bands could allow to obtaining statistically independent wrapped phase data sets, as explained in Ref. [56,83]. In such case, the likelihood function relative to the single data set $\Psi_{1i}(n,m)$ is given by Eq. (20.26), evaluated at $\gamma = \gamma_{1i}$ and $\alpha = \alpha_{1i}$:

$$\begin{aligned}f_{\Psi_{1i}}(\psi_{1i}; h) &= \frac{1}{2\pi} \frac{1 - |\gamma_{1i}|^2}{1 - |\gamma_{1i}|^2 \cos^2(\psi_{1i} - \alpha_{1i}h)} \left\{ 1 + \frac{|\gamma_{1i}| \cos(\psi_{1i} - \alpha_{1i}h) \cos^{-1}[-|\gamma_{1i}| \cos(\psi_{1i} - \alpha_{1i}h)]}{[1 - |\gamma_{1i}|^2 \cos^2(\psi_{1i} - \alpha_{1i}h)]^{1/2}} \right\},\end{aligned}\quad (20.53)$$

where the dependence on (n,m) has been understood, and, in the case of N statistically independent wrapped phase data sets, the overall multi-interferogram likelihood function will be given by:

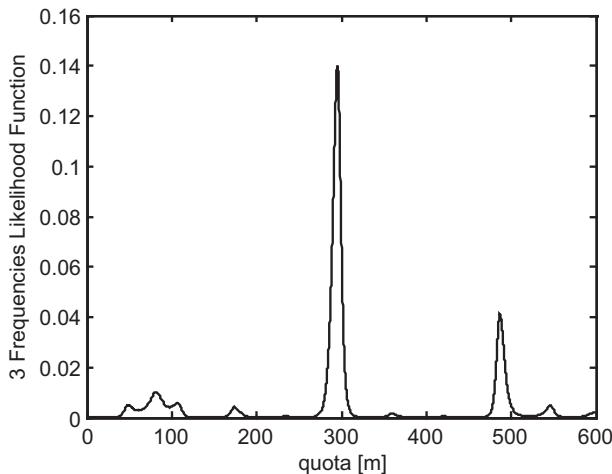
$$F_m(\Psi; h) = \prod_{i=2}^N f_{\Psi_{1i}}(\psi_{1i}; h),\quad (20.54)$$

where $\Psi = [\psi_{11}(n, m) \psi_{12}(n, m) \cdots \psi_{1N}(n, m)]^T$ is the measured wrapped phase data vector. Note that we have a wrapped phase data vector for each position (n,m) on the ground. The plot of Eq. (20.54) for a given height value ($h = 300$ m) for three values of the measured interferometric phases (suppose to consider four SAR image channels and then three interferograms), and for $\gamma_{1i} = 0.8, i = 1,2,3$, is shown in Figure 20.17. Note that the multiplication of the three single-interferogram likelihood functions, allows to avoid multiple global maxima present in each single-interferogram likelihood function, at least in the range of interest for h [83].

The ML estimate can be obtained by finding the unique value of h that maximize, for each position (n,m) , the multi-interferogram likelihood function:

$$\hat{h}_{\text{ML}}^{(N)} = \arg \max_h F_m(\Psi; h).\quad (20.55)$$

The uniqueness results shown above can, in principle, also be obtained using only two interferograms: it suffices that the ratio between the period of the two single-interferogram likelihood functions is

**FIGURE 20.17**

Multiple-interferogram likelihood function.

not rational, so that the overall double-frequency likelihood function is not periodic [84]. Hence, in the multi-interferogram case, the likelihood function will exhibit a single global maximum (the ML estimate).

In order to correctly implement (20.53) and (20.54), accurate knowledge of the single-interferogram likelihood functions is very important. While it can be easily obtained in the frequency-diversity case (frequencies are known with a very high precision), the same is not true for the baseline-diversity case, where baselines are known only to the precision of the inertial navigation systems, which are usually not able to achieve the required accuracy of fraction of wavelengths. A second difference is related to the fact that the different image pairs used to obtain the different wrapped phase values to be used in (20.53) are acquired in such a way that the independence of the wrapped phase values is seldom satisfied, so that model (20.54) is not valid. In this case, the determination of the multi-interferogram likelihood function would require joint statistical characterization of the different interferograms.

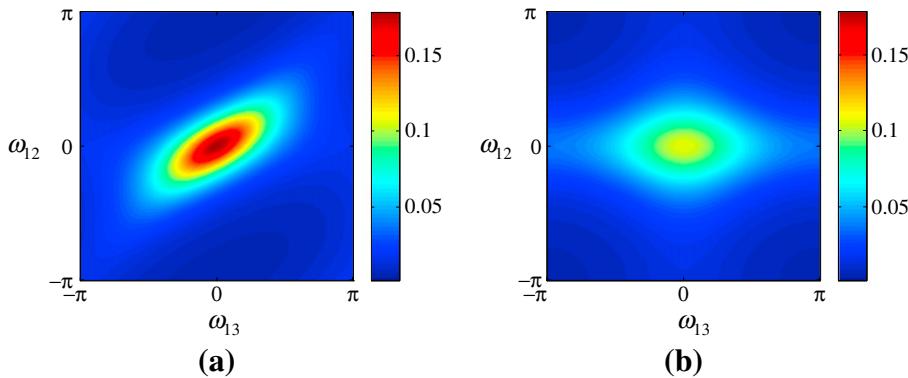
In Ref. [52], the joint pdf for the general case of correlated interferograms has been derived. The model given by Eqs. (20.13) and (20.14) can be generalized to the case of $k = 1, 2, \dots, K$, $K \geq 2$ complex images $Z_k(n, m)$.

With reference to the case of $K = 3$, whose geometry is shown in Figure 20.15, the interferometric phases, obtained by beating two of the three images given by Eq. (20.13) for $k = 1, 2, 3$, are:

$$\psi_{12} = \arg(Z_1 Z_2^*), \quad \psi_{13} = \arg(Z_1 Z_3^*), \quad \psi_{23} = \arg(Z_2 Z_3^*). \quad (20.56)$$

Generalizing results of Eqs. (20.16) and (20.17), the interferometric phases can be related to the observed scene height profile through:

$$\psi_{ik} = \left\langle \frac{4\pi}{\lambda} (r_j - r_k) + \Omega_{jk} \right\rangle_{2\pi}, \quad i, k = 1, 2, 3, \quad (20.57)$$

**FIGURE 20.18**

Second order pdf of dual baseline phase interferograms with $B_{12} = 307$ m, $B_{13} = 637$ m, and $\gamma_0 = 0.8$: (a) statistically dependent interferograms; (b) statistically independent interferograms.

where, also in this case,

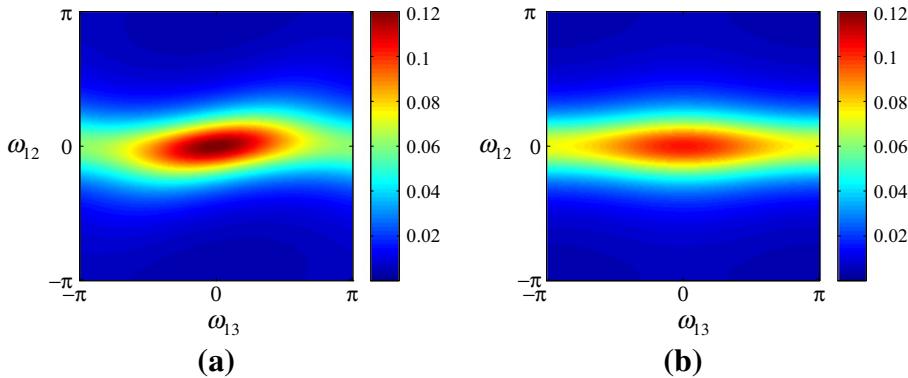
$$\begin{aligned} X_k &= A_k e^{j\Phi_k}, \quad k = 1, 2, 3, \\ \Omega_{ik} &= \arg(X_i X_k^*) = \langle \Omega_i - \Omega_k \rangle_{2\pi}. \end{aligned} \quad (20.58)$$

Note that the three interferometric phases in Eqs. 20.56 and 20.57 are not independent each other, as each of them is univocally determined when the other two are known (f.i. $\Psi_{23} = \langle \Psi_{13} - \Psi_{12} \rangle_{2\pi}$). Hence, it suffices to consider only two of them. It is usually convenient considering the two interferograms obtained by referring the phase to the same master antenna, f.i. the antenna 1, then getting Ψ_{12} and Ψ_{13} .

Also in this more general multi-channel case, generalizing results relevant to two channels [52], it is possible to derive a closed form for the pdf $f_{\Omega_{12}, \Omega_{13}}(\omega_{12}, \omega_{13})$ of the interferometric speckle phase differences $\Omega_{12} = (\Omega_1 - \Omega_2)$ and $\Omega_{13} = (\Omega_1 - \Omega_3)$ [52]. Such pdf is completely symmetric respect to the interferometric phase values $\omega_{12} = \omega_1 - \omega_2$, $\omega_{13} = \omega_1 - \omega_3$, and $\omega_{23} = \omega_{13} - \omega_{12} = \omega_2 - \omega_3$, as shown in Ref. [52].

The 2D representation of pdf $f_{\Omega_{12}, \Omega_{13}}$ is shown in Figure 20.18a, obtained with the values $B_{12} = 307$ m and $B_{13} = 637$ m ($B_{23} = B_{13} - B_{12} = 330$ m), where the images coherence values have been computed according to the first order spatial decorrelation model introduced in Eq. (20.29), assuming $\gamma_0 = 0.8$, and consequently $\gamma_{12} = 0.57$, $\gamma_{13} = 0.33$, and $\gamma_{23} = 0.55$. We can notice the typical behavior met for a couple of correlated random variables. The joint pdf that would be obtained with independent interferograms, for the same parameters values used for the pdf of Figure 20.18a, is shown for comparison in Figure 20.18b.

The 2D representation of the joint pdf $f_{\Omega_{12}, \Omega_{13}}$ for different baseline values ($B_{12} = 111$ m, $B_{13} = 937$ m, and $B_{23} = B_{13} - B_{12} = 826$ m), for $\gamma_0 = 0.8$, and for the corresponding coherence values $\gamma_{12} = 0.72$, $\gamma_{13} = 0.11$, and $\gamma_{23} = 0.19$, is reported in Figure 20.19a. The joint pdf that would be obtained with independent interferograms, for the same parameters values used for the pdf of Figure 20.19a, is shown for comparison in Figure 20.19b. Note that the difference between pdfs of Figure 20.19a and b is less pronounced than the one between pdfs of Figure 20.18a and b, since in this case the

**FIGURE 20.19**

Second order pdf of dual baseline phase interferograms with $B_{12} = 111$ m, $B_{13} = 937$ m, and $\gamma_0 = 0.8$: (a) statistically dependent interferograms; (b) statistically independent interferograms.

statistical independence assumption is approximately met, because two of the three coherence values are very low.

Also in this case the interferometric phases pdf can be obtained through a change of variables, generalizing what described for one variable in Eq. (20.26):

$$f_{\Psi_{12}, \Psi_{13}}(\psi_{12}, \psi_{13}; h) = f_{\Omega_{12}, \Omega_{13}}(\psi_{12} - \alpha_{12}h, \psi_{13} - \alpha_{13}h). \quad (20.59)$$

The differences evidenced in Figures 20.18 and 20.19 for the speckle phases, with and without correlation, result also in the case of the interferometric phases Ψ_{12} and Ψ_{13} . In some conditions, the differences between the two cases pdfs can be more pronounced, in other less. Of course, in case of significantly correlated interferograms, the height estimation according (20.55) by using the joint pdf of independent interferograms (Eq. (20.54) in place of (20.59)) would lead to larger estimation errors both in terms of quadratic errors and bias, especially when the two joint pdfs assumes very different shapes [85]. Equation (20.59), which can be easily generalized to the case of more than two interferograms [52], is very general, and reduce to (20.54) in the case of independent (and hence, uncorrelated) interferograms.

Use of pdf (20.59) in place of (20.54) allows to obtain better results in terms of accuracy (Bias and Cramer-Rao Lower Bounds) of the reconstructed ground height profiles [52, 85].

2.20.4.1 Bayesian statistical solution to PhU problem

To solve the problem of the estimation of the ground elevation profile also Bayesian statistical techniques have been proposed [86–88]. In the framework of such techniques, in particular if a Maximum a Posteriori (MAP) estimation scheme is adopted, an *a-priori* joint pdf of the unknown height profile has to be introduced. It is based on the use of Markov Random Fields (MRF) as *a-priori* statistical term modeling pixels contextual statistical information in the 2D unknown height profile to be reconstructed. The MRF allows to describe the *local* spatial interaction between couples of pixels, through a set of model parameters (hyperparameters) which can be tuned following unsupervised procedures.

Consider a discrete (lexicographically ordered) two-dimensional (2D) points lattice $\mathcal{L} = \{k, k = 1, \dots, N_p\}$, where N_p is the number of pixels of the SAR image, and let $\mathbf{h} = [h_1 \ h_2 \ \dots \ h_{N_p}]^T$ the corresponding ground elevation values. Consider now an InSAR system, and let ψ_{kn} the wrapped phase values (single sample of a discrete interferogram) measured at the lattice point k and at n th interferogram. The wrapped phase values ψ_{kn} relative to the position k can be structured and ordered in the following way: let $\boldsymbol{\psi}_k = [\psi_{k1} \ \psi_{k2} \ \dots \ \psi_{kN}]^T$ be the vector of the wrapped phases measured in k position at the N different interferograms, and $\boldsymbol{\psi} = [\boldsymbol{\psi}_1^T \ \boldsymbol{\psi}_2^T \ \dots \ \boldsymbol{\psi}_{N_p}^T]^T$ be the vector collecting all available wrapped phase values. Then, \mathbf{h} is the vector of the unknown height values, and $\boldsymbol{\psi}$ the vector of the all available data (multi-interferogram).

The MAP estimation can be formulated as:

$$\hat{\mathbf{h}}_{\text{MAP}} = \arg \max_{\mathbf{h}} f_{\text{Post}}(\mathbf{h} | \boldsymbol{\psi}) = \arg \max_{\mathbf{h}} \ln f_{\boldsymbol{\psi}}(\boldsymbol{\psi} | \mathbf{h}) g_{\mathbf{H}}(\mathbf{h}; \sigma), \quad (20.60)$$

where $f_{\text{Post}}(\mathbf{h} | \boldsymbol{\psi})$ is the *a posteriori* joint pdf of the unknown image, $f_{\boldsymbol{\psi}}(\boldsymbol{\psi} | \mathbf{h})$ is the Bayesian likelihood function [89], and $g_{\mathbf{H}}(\mathbf{h}, \boldsymbol{\psi})$ is the *a-priori* pdf of the unknown image.

The Bayesian likelihood function can be easily obtained by Eq. (20.59):

$$f_{\boldsymbol{\psi}}(\boldsymbol{\psi} | \mathbf{h}) = \prod_{k=1}^{N_p} f_{\boldsymbol{\psi}_k}(\psi_k | \mathbf{h}), \quad (20.61)$$

where the statistical independence of the interferograms in the different ground positions $k = 1, \dots, N_p$ has been exploited. The Bayesian likelihood functions are formally equal to the likelihood function that can be obtained in the classical statistical case; in the Bayesian framework, differently from the classical one, the unknown image is seen as a sample of a random vector \mathbf{H} .

The *a-priori* pdf $g_{\mathbf{H}}(\cdot)$ is usually defined in such a way to express *a-priori* information about the unknown image, assigning high probability to particular pixels configurations. In the SAR interferometry case, being the unknown image representative of the ground elevation map of a geographic area, a strong contextual pixel information is very likely to be. In particular, it can be assumed that the unknown image can be modeled as a *Markov Random Field* (MRF) [90], a general image model able to represent contextual pixel information extending the 1-D Markov property to the 2-D case, whose corresponding joint pdf is given by a Gibbs distribution:

$$g_{\mathbf{H}}(\mathbf{h}; \sigma) = \frac{1}{Z(\sigma)} \exp[-\mathcal{E}(\mathbf{h}, \sigma)] = \frac{1}{Z(\sigma)} \exp \left[- \sum_{k=1}^{N_p} \sum_{j \in \mathcal{N}_k} U(h_k, h_j, \sigma_{kj}) \right], \quad (20.62)$$

where $Z(\sigma) = \int \exp(-\mathcal{E}(\mathbf{h}; \sigma)) d\mathbf{h}$ is the so-called partition function (a constant factor needed to normalize the integral of the pdf to one), $\mathcal{E}(\cdot)$ is the energy function, $U(\cdot)$ is the potential function between h_k and h_j , \mathcal{N}_k is the *neighbourhood system* of k th pixel [91] (usually, the eight pixels around the k th one), σ is the *hyperparameter* vector, and σ_{kj} are the *hyperparameters*. With such definition, the Gibbs-MRF model in Eq. (20.62), by means of the *hyperparameters*, well adapts to describe the image *local* nature, leading to a powerful and general model, well-suited to represent a very wide class of height profiles. The hyperparameter values are not known, of course, and they have to be estimated from the available interferometric data $\boldsymbol{\psi}$ data.

The solution procedure essentially consists of two steps: ML estimation (σ) of the hyperparameter vector σ and MAP estimation ($\hat{\mathbf{h}}_{\text{MAP}}$) of the actual realization of height profile process \mathbf{H} .

This approach can work very well [87], even if the method is always limited by the coherence value and by the corresponding number of available independent acquisitions of the same scene. The lower are the coherence values over the entire image, the larger is the total number N of needed interferograms to obtain good quality reconstructions.

Performance of the Bayesian method can significantly outperform the one relevant to classic approaches [85, 92] depending on the capability to estimate the *a-priori* model of the unknowns.

2.20.4.2 Graph cuts solution to PhU problem

Statistical approaches, especially the Bayesian ones, have proved to be effective dealing with noisy data and big discontinuities. However these algorithms can be time consuming and computationally heavy due to the *a-priori* model (estimation of hyperparameters) and to the optimization step. It is possible to overcome these limits by introducing a fast and efficient (in term of global optimization) algorithm to unwrap the interferometric phase in the multichannel configuration [85].

To reduce the computational time needed to unwrap the multichannel interferometric phase, two aspects can be taken in consideration: first, a non-local *a-priori* energy function, the Total Variation (TV) model [93], and secondly an optimization algorithm based on graph cuts, Ishikawa algorithm [94], have been exploited.

The *a-priori* energy corresponding to the TV model can be written as follows:

$$\mathcal{E}(\mathbf{h}, \sigma) = \sigma \sum_{K=1}^{N_p} \sum_{j \in \mathcal{N}_k} |h_k - h_j|. \quad (20.63)$$

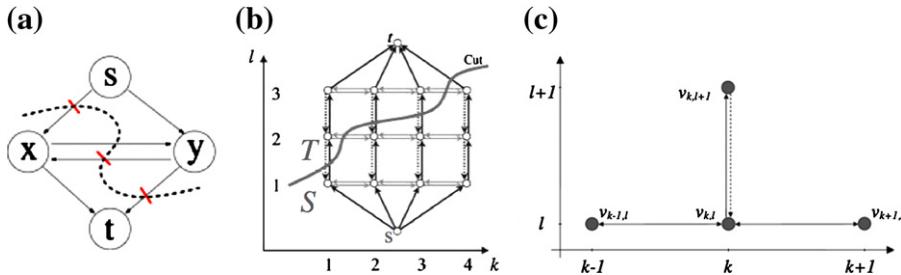
Note that in this expression, σ is a scalar, making model in Eq. (20.63) a non local one, differently from the Bayesian model (20.62) presented in the previous section. This choice is done in order to make the algorithm faster. A non local *a-priori* model avoids the estimation of local hyperparameters σ_{kj} , as only one parameter σ for the whole image has to be estimated.

Between the existing non local *a-priori* energy models, TV has been chosen, due to its main advantage: TV does not penalize discontinuities in the image while simultaneously not penalizing smooth functions either [93].

Given the TV *a-priori* energy model (20.63), the MAP estimation can be obtained from the minimization of the following function:

$$\hat{\mathbf{h}}_{\text{MAP}} = \arg \max_{\mathbf{h}} f_{\text{Post}}(\mathbf{h} | \boldsymbol{\psi}) = \arg \min_{\mathbf{h}} \left(-\ln f_{\boldsymbol{\psi}}(\boldsymbol{\psi} | \mathbf{h}) + \sigma \sum_{K=1}^{N_p} \sum_{j \in \mathcal{N}_k} |h_k - h_j| \right). \quad (20.64)$$

In order to minimize this energy function, graph-cut based optimization algorithms are used. Graph-cut optimization [95, 96] is successful because the exact minimum or an approximate minimum with certain guarantees of quality can be found. Compared to the classical optimization algorithm, it provides comparable results with much less computational time and compared to the deterministic algorithm ICM [90], it avoids the risk of being trapped in local minima solution which can be far from the global one.

**FIGURE 20.20**

(a) Example of a graph construction and of a cut on it (dashed line); (b) Ishikawa Graph Construction—On the axes there are the pixels and the labels; data edges are depicted as black arrows; constraint edges are represented by horizontal arrows and penalty edges are depicted as dotted arrows; (c) Ishikawa edges weights—Representation of the edges for a vertex.

Let us, first, define a graph and a graph-cut problem. Suppose $G = (V, E)$ is a directed graph with non negative edge weights, where V is the set of vertexes and E the set of edges. This graph has two special vertexes (terminals) called the *source* s and the *sink* t . A s - t -cut $C = \{S, T\}$ is defined as a partition of the vertexes into two disjoint sets S and T such that $s \in S$ and $t \in T$. A cost of this cut is defined as the sum of weights of all edges that go from S to T . Figure 20.20a shows a simple graph made of four vertexes, two vertexes (x,y) plus the sink and the source (t,s). The edges are the links between the vertexes. An example of a cut is represented with the dashed line.

The minimum s - t -cut problem is to find a cut C with the smallest cost. This problem is exactly equivalent to its dual problem, which consists in computing the maximum flow from the source to the sink. Between the several algorithms proposed to solve the maximum flow problem, the one proposed in [97] turns to be the most adapted to computer vision problems.

To solve the MAP unwrapping problem, which means find the value of \mathbf{h} that minimizes (20.64) the optimization procedure proposed by Ishikawa is implemented [94]. The interesting aspect of Ishikawa algorithm is that under some hypothesis on the energy to be minimized and if the graph is correctly constructed, the algorithm provides the global optimum of the considered energy.

Two hypothesis are at the base of Ishikawa algorithm: convexity of the *a-priori* energy and a linear order on the label set. Using the TV model the first hypothesis is satisfied. For the second hypothesis, in the case of phase unwrapping, the labels are the height of the image pixels. The heights are supposed to be represented as integers in the range $\{0,1,2,\dots,L-1\}$, where L is the size of the label sets. This condition satisfies the second hypothesis necessary for the Ishikawa graph construction.

Ishikawa method is based on computing a minimum cut in a particular graph. Ishikawa graph $G = (V, E)$ contains $N_p \times L$ nodes (N_p is the size of the image and L is the size of the label set) denoted by $\{v_{k,l}\}$, plus two special nodes s and t . For each pixel k , we associate L nodes, that represent all the possible heights that the pixel k can take. The construction of Ishikawa graph, in case of a 1D image for legibility, is shown in Figure 20.20b.

Ishikawa graph contains three families of edges $E = E_D \cup E_C \cup E_I$. E_D is a set of directed edges called data edges (black arrows of Figure 20.20b). It represents the data energy term. E_C is a set of

directed edges called penalty edges (dotted arrows of Figure 20.20b). It ensures that only one data edge is in the minimum cut for each pixel k . Finally, E_I is a set of constraint edges between all neighbor pixels (horizontal arrows in Figure 20.20b). It represents the *a-priori* energy term.

To better understand how to set the edges weight for the multichannel phase unwrapping problem let us consider Figure 20.20c. The vertex $v_{k,l}$ is the vertex identified by the pixel k and the label l . The cost of the edges in Ishikawa graph are reported in Eq. (20.66). For more details on the graph construction see [94].

$$\begin{aligned} c(v_{k,l}, v_{k,l+1}) &= -\ln f_\psi(\psi_k | h_k = l), \\ c(s, v_{k,0}) &= \infty, \\ c(v_{k,L}, t) &= -\ln f_\psi(\psi_k | h_k = L - 1), \\ c(v_{k,l+1}, v_{k,l}) &= \infty, \\ c(v_{k-1,l}, v_{k,l}) &= 1. \end{aligned} \quad (20.65)$$

Constructing the Ishikawa graph as shown and finding the cut with a minimum cost on this particular graph allows to find the exact (global) optimum solution for the multichannel phase unwrapping problem given by Eq. (20.64). The main disadvantage of Ishikawa method is related to the memory load. In fact, the algorithm stores the whole graph and then performs the cut. This can be a problem when the size of the images largely increase.

Before using Ishikawa algorithm, the hyperparameter σ has to be estimated. The hyperparameter σ depends both on data energy and on *a-priori* energy term. A method to perform the hyperparameter parameter is the analysis of the so called *L-curve*. To find automatically the corner of the *L-curve*, the triangular method described in [98] can be used, providing good results in limited time.

2.20.5 Multipass interferometry

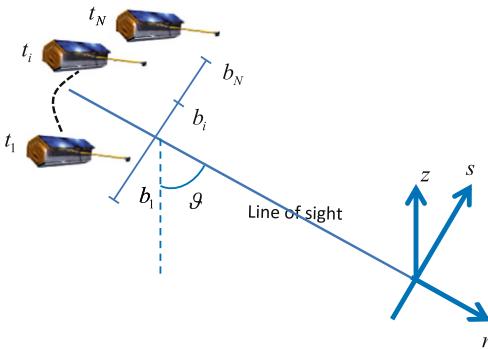
Advanced differential SAR interferometric (A-DInSAR) techniques process multitemporal data, acquired over repeat passes, to generate very accurate deformation time series and therefore to achieve a regular monitoring of the deformation of the observed scene. A-DInSAR techniques also mitigate most of the limitations of the standard single-interferogram approaches, such as temporal and geometric decorrelation and the atmospheric phase delay and, unlike conventional Interferometry, they allow increasing the measurement accuracy from centimeter up to millimeter. These techniques also improve standard approaches both in terms of deformation modeling capabilities and quality measurements.

We refer to the geometry depicted in Figure 20.21, where we assume that the satellite is collecting the data at time instants t_1, t_2, \dots, t_N , with baselines b_1, b_2, \dots, b_N , with respect to a reference image.

Let φ be the vector that collects the phase measurements, $\varphi_1, \varphi_2, \dots, \varphi_N$. The following data model is assumed:

$$\varphi = \frac{4\pi}{\lambda} (\xi s + d) + \varphi_a + \varphi_o + \varphi_n, \quad (20.66)$$

where ξ is a vector of a-dimensional parameters that collect the angular variations corresponding to the baseline distribution $\xi_i = b_i/r$, d is the vector collecting the displacements at the different times,

**FIGURE 20.21**

Multipass interferometric geometry.

φ_a , φ_o , and φ_n are respectively the unwanted vectors collecting the APD, also known as Atmospheric Phase Screen (APS) [74], the orbit errors and the noise, which, in this case, includes also the temporal decorrelation effects in addition to the spatial decorrelation contribution due to the speckle. Note that the elevations can be related to the target height z as $s = h/\sin(\vartheta)$.

The final aim of multipass DinSAR techniques is to reconstruct d , and especially for some specific applications to estimate s to correctly localize the target to which the displacements refer. In doing this the cancellation of the unwanted disturbance terms is carried out by using a proper modeling, of deterministic nature (such as for instance the baseline dependent term related to the target elevation and the orbit error which show a predictable spatial variability) and stochastic models, such as for the characterization and filtering of the atmospheric component.

Some remarks are now in order. First of all the phase of each single SAR image depends on the phase of the backscattering coefficient that, for distributed scatterers, can be strongly variable from pixel to pixel. Therefore interferograms must be formed in such a way to mitigate, as much as possible the effects of the backscattering coefficient. Secondly, the interferograms are extracted by complex numbers and therefore the phase values are wrapped, i.e., they are determined only on a basic interval of 2π size.

With regards to the process of determining the set of interferogram to be used for the displacement estimation process, there are essentially two alternative options: the (temporal) Persistent Scatterers Interferometry (PSI) and the (spatial) Coherent Scatterers Interferometry (CSI) approaches. Almost all implementations of these class of approaches makes use of the model in Eq. (20.66).

The main difference between the two approaches is the basic assumption underlying the typology of the scattering which impacts the strategy followed in the interferogram generation and in the scale of analysis.

In the PSI case the interferometric analysis is carried out at the highest possible spatial resolution, the set of interferograms is generated with respect to the master image without taking any care to avoid large spatial and temporal baselines. This strategy is chosen in PSI because the technique is aimed at monitoring the deformation of dominant scatterers, i.e., of scatterers showing a persistency of the scattering mechanism over the time. Many natural reflectors, typically present over man-made

structures, satisfy this requirement. In the PSI approach the use of large spatial baselines are functional to the achievement of a high accuracy in the estimation of the target elevation which provides a high accuracy in the localization of ground scatterers and therefore in the identification of the scatterers subject to possible deformations.

Most of PSI algorithms rely on the use of simplified version of the model in (20.66) in which the deformation is assumed to be linear i.e., $\mathbf{d} = \mathbf{d}_L = vt$, where v is the deformation mean velocity and t is the vector collecting the time instants.

In this case the phase vector $\boldsymbol{\varphi}$ is expressed as:

$$\boldsymbol{\varphi} = \frac{4\pi}{\lambda} (\xi s + vt) + \boldsymbol{\varphi}_n. \quad (20.67)$$

From Eqs. (20.66) and (20.67) it is evident that, besides the linear deformation approximation, other two assumptions are fundamental: that is the possibility to neglect the atmospheric contribution and the orbital inaccuracy. While the latter assumption does not play a critical role because the orbital errors are rather accurately modeled as spatial (azimuth-range) planar phase contributions that can be accurately estimated and compensated directly on the measured complex data in wide areas, neglecting the atmospheric phase pattern is a main issue. One possibility to handle the atmospheric contribution is by exploiting its the spatial correlation properties, specifically the fact that contributions on a scale of a few hundreds of meters are strongly correlated as explained in Section 2.20.3.3. Accordingly, this contribution can be estimated on a coarse grid and subtracted from the data. In the case of the original Persistent Scatterers algorithm of [99, 100], the grid of pixels used for the estimation of the atmospheric phase pattern is selected by measuring the amplitude dispersion index, which is a proxy of the phase stability.

Following this compensation each pixel is investigated for the presence of scatterers with “persistent” properties, i.e., a parameter that describes the fit between the model and the real phase is considered:

$$C(s, v) = \frac{1}{N} \left| \sum_{i=1, \dots, N} e^{j\varphi_i - j \frac{4\pi}{\lambda} (\xi_i s + t_i v)} \right|_n. \quad (20.68)$$

Equation (20.68) provides a normalized (in the [0,1] interval) measure, of the temporal (i.e., persistent) coherence property of scatterers: s (scatterer elevation) and v (velocity) are unknowns that are determined via a maximization

$$(\hat{s}, \hat{v}) = \arg \max C(s, v). \quad (20.69)$$

It is interesting to note that by letting:

$$\mathbf{y}_{\boldsymbol{\varphi}} = [e^{j\varphi_1} \ e^{j\varphi_2} \ \dots \ e^{j\varphi_N}], \quad (20.70)$$

where φ_j is the j th component of $\boldsymbol{\varphi}$.

Now, introducing:

$$\mathbf{a}(s, v) = \left[e^{j \frac{4\pi}{\lambda} (\xi_1 s + t_1 v)} \ e^{j \frac{4\pi}{\lambda} (\xi_2 s + t_2 v)} \ \dots \ e^{j \frac{4\pi}{\lambda} (\xi_N s + t_N v)} \right], \quad (20.71)$$

where ξ_j is the j th component of $\boldsymbol{\xi}$, and t_j is the j th component of \mathbf{t} .

Then Eq. (20.68) can be rewritten as:

$$C = \frac{1}{N} \left| \mathbf{y}_\varphi^H \mathbf{a}(s, v) \right| = \frac{1}{N} \left(\left| \mathbf{a}^H(s, v) \mathbf{y}_\varphi \mathbf{y}_\varphi^H \mathbf{a}(s, v) \right| \right)^{-1/2}. \quad (20.72)$$

This equation highlights the fact that the PS approach uses indeed with unitary weighting all the interferograms $\mathbf{y}_\varphi \mathbf{y}_\varphi^H$, and therefore the whole information available in the data is used in the determination of the target height and velocity.

Since the proposal of the PS technology, many other PSI techniques have been proposed: SPN [101], PSI-GENESIS [102], IPTA [103], SPINUA [104]. In most of the cases they differ essentially in the selection of the sparse grid over which the interferometric analysis is carried out. A solution that extensively uses the model in (20.67) is the Persistent Scatterers Pairs (PSP) algorithm [105]. In this case the mitigation of the atmospheric contribution is carried out by considering the phase variation over spatial arcs: the atmospheric contribution as well as non-linear effects spatially slowly variable are automatically canceled provided that the length of the arc is sufficiently small. In the PSP case, the selection of the sparse grid of analysis is carried out by moving from a reference starting point to the adjacent points using the spatial arcs and the model in (20.67).

The capability of SAR to monitor deformation phenomenon up to centimeters by compensating topography and atmospheric disturbances via processing of stacks of data was for the first time provided by the group of Politecnico of Milan with the persistent scatterers technique. The examples included in the seminal papers [99, 100] that demonstrated for the very first time the possibility to monitor demonstration with millimeter per year accuracy of the Pomona (CA-USA). Among the first examples of application of the Persistent Scatterers is the monitoring of the Etna volcano (see Figure 20.22).

Currently Multipass DInSAR techniques, such as small baseline and persistent scatterers are applied to monitor deformation phenomena associated in seismogenetic areas [106], landslides [107], areas affected by underground excavation or water level changes [108] gas pumping and storage, etc. over wide areas.

An example of application of PSI on a national scale is provided in Figure 20.23.

Multipass DinSAR techniques allows also image and monitor ground scatterers with high spatial resolution. Persistent Scatterers Interferometry (PSI) techniques exploits all the available acquisitions in such a way to improve the performances with respect to this specific case.

PSI approach are designed to monitor deformation of dominant scatterers. PSI may however loose the capability of monitoring scatterers which shows non negligible temporal or spatial decorrelation: such scatterers are corresponding to areas in which the scattering is not concentrated in a dominant point but rather is distributed over the resolution cell.

Techniques complementary to PSI that uses only interferograms showing a sufficient high degree of spatial coherence [109–111], in this work referred to as coherent stacking interferometric (CSI) techniques, aims to overcome the problems of PSI related to distributed scattering.

In fact, following the lines of the classical (single-pair) DInSAR analysis, CSI techniques carry out an analysis of multilook interferograms: the multilook operation, i.e., a spatial averaging which is carried out by exploiting the hypothesis that the scattering is distributed, allows on one side accessing a measure of quality of the signal in each interferogram, i.e., the spatial coherence, and on the other side it allows increasing the signal phase quality through the averaging operation.

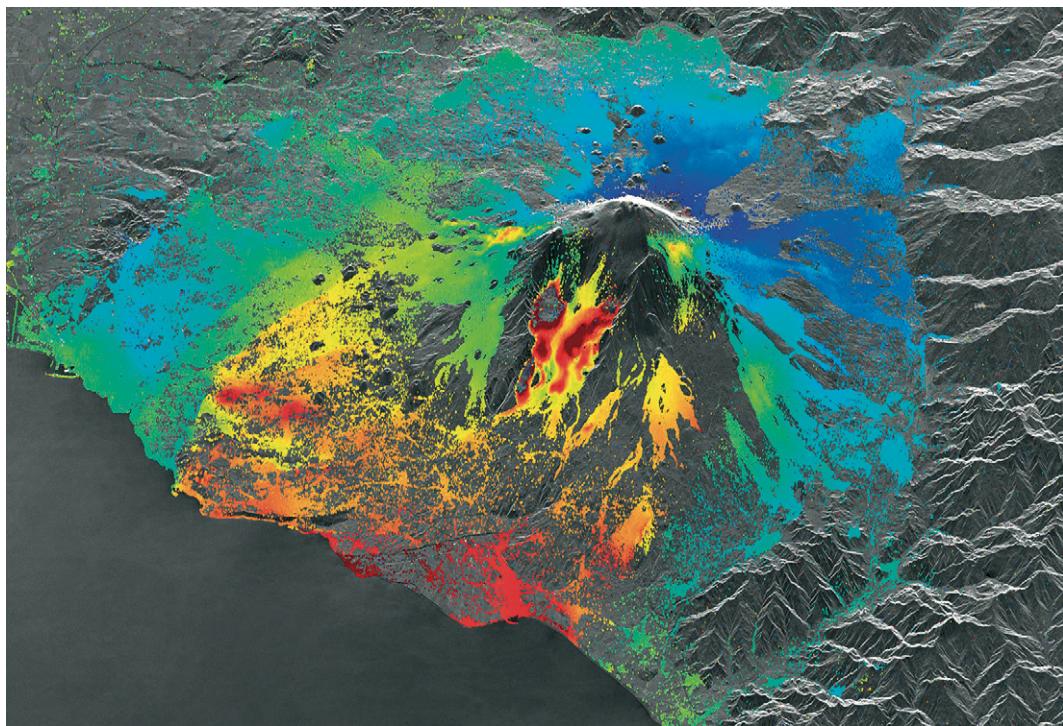


FIGURE 20.22

The persistent scatterer technique applied to the monitoring of the Etna volcano in the period 1995–1999, colorbar is saturated in the range (−30,30) mm/yr.

Tele-Rilevamento Europa (T.R.E. s.r.l.) and Politecnico of Milano.

The first interferometric stacking technique proposed in the literature has been the Small Baseline Subset (SBAS) technique [109]. This technique makes use of a limitation on the spatial and temporal baselines to control the temporal and geometric (also known as angular) decorrelation phenomena which is more critical in the presence of distributed scattering.

As described in Section 2.20.3.2 geometrical decorrelation is a phenomenon which is associated with the fact that even small variations of the radar line of sight direction in the elevation direction induced by the orbital separation (spatial baseline decorrelation) and/or aspect angle (induced by doppler centroid variation); temporal decorrelation is due to changes over the time of the scattering. The SBAS technique uses only interferograms generated by choosing thresholds on the spatial and temporal baselines, that is the spatial orbital and temporal separations, respectively, and on the Doppler centroid difference (for systems suffering of large variation of the Doppler centroid) thus limiting the effects of angular and temporal decorrelation. Mathematically, letting $\Delta\bar{\varphi} = [\Delta\bar{\varphi}_1, \Delta\bar{\varphi}_2, \dots, \Delta\bar{\varphi}_M]$ the vector that collects the M (multilook) interferograms values in a generic pixel, we have:

$$\Delta\bar{\varphi} = \mathbf{A}\bar{\varphi}, \quad (20.73)$$

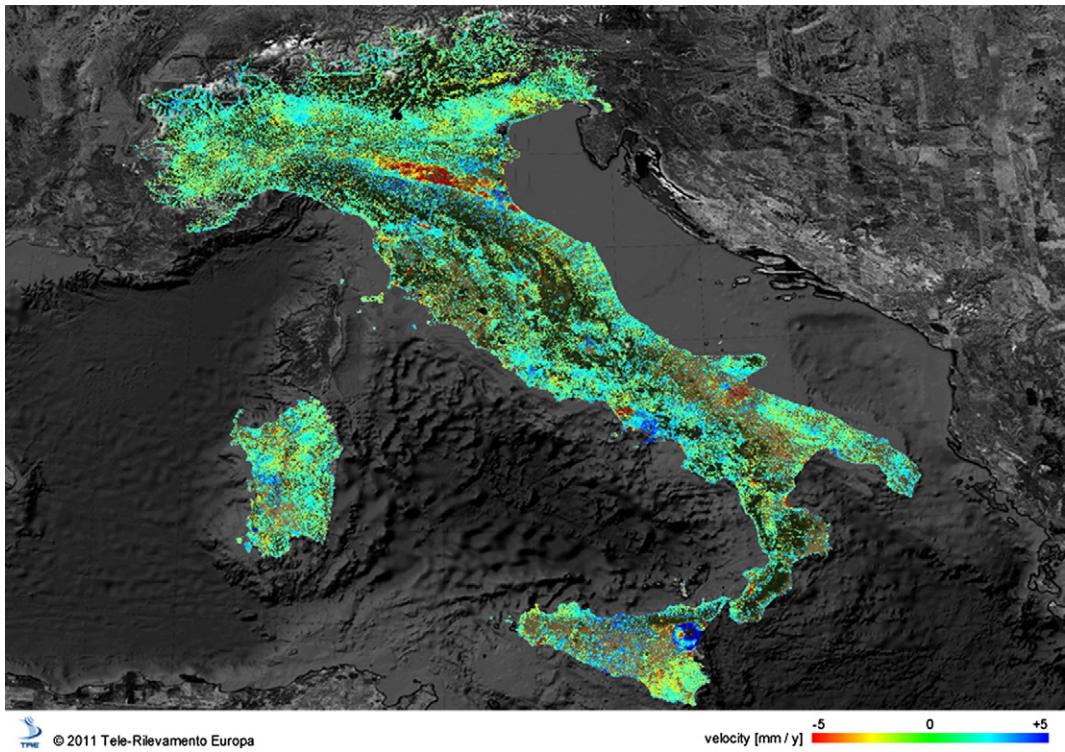


FIGURE 20.23

Result of the PST project funded by the Italian Ministry of the Environment for monitoring the displacements in Italy via the Persistent Scatterers Interferometry technique. Data of the ESA ERS-1 and ERS-2 satellites.

TREuropa.

where $\bar{\phi}$ is the multilook version of ϕ and \mathbf{A} is a $M \times N$ matrix ($M > N$) incidence matrix that describes the structure of the interferogram stack used in the processing: it is dependent on the pairing of the acquisitions in the interferogram generation. Equation (20.73) refers to the absolute phase values: hence the set of interferogram is unwrapped, (commonly via the Minimum Cost Flow algorithm [58–60, 79]) and then (20.73) is inverted pixel by pixel to retrieve the phase signal over the stack of acquisitions.

The SBAS approach, by using the Singular Value Decomposition [112] technique allows also handling the case in which, due to the limitations on the baselines, the acquisitions are grouped in different independent subsets leading to the case of ill-conditioning of the matrix \mathbf{A} in (20.73). This latter feature allows also performing a “semi-coherent” combination of data of different sensors (f.i. ERS and ENVISAT), i.e., the combination of sets of coherent data for which are incoherent each other.

It is worth pointing out that the system in (20.73) is intrinsically invertible for each pixel up to a constant: typically this constant is set in such a way to have a null deformation at time t_1 . Moreover, as each interferogram is known up to a constant phase value, the solution of (20.73) for all pixels is known up

to the deformation of a known point which must be set, as in classical leveling technology, as a reference point. After the inversion the deformation time series are separated from atmospheric contribution by using (20.66): in particular the deformation mean velocity and the residual topography are estimated by assuming for $\bar{\varphi}$ the model in (20.67). The linear motion and topography errors are subtracted from the data and the residuals are then filtered to separate the non-linear deformation component from the APD contribution: this separation relies on the assumption that the deformations are typically correlated spatially and temporally (slow deformations) whereas the atmosphere is spatially correlated but almost independent from epoch to epoch. The latter assumption is certainly valid for the turbulent component of the APD whereas it may be critical in case of dense temporal sampling (i.e., low revisiting time) for the stratified contribution (depending on the topography) which exhibit seasonal dependencies.

Differently from the SBAS approach where typically a degree of redundancy is imposed at the interferogram generation level to control the presence of possible unwrapping errors, the CPT technique uses a Minimum Spanning Three [113] strategy to connect the acquisition stack, and therefore to generate the interferograms. Another special feature of the CPT approach is related to the implementation of a step which allows, via modeling, to retrieve the linear deformation component and DEM error contribution prior to extraction of the non-linear component. In this sense the technique is very similar to the ESD approach proposed in [114].

The A-DinSAR techniques working at lower resolution, such as SBAS and CPT, allows to easily implement a two-stage, low-resolution and high-resolution, interferometric processing strategy for the evaluation of the deformation over at small and a high scale, respectively.

In other words, once the atmospheric phase component and the non-linear deformation are extracted, they are subtracted from the original data corresponding to each acquisition at full resolution. Such data can be used in conjunction with the model in (20.67) to perform, as in standard PSI, the estimation of the deformation at full resolution. It is worth noting that, due to the fact that the increase of resolution is aimed at analyzing dominant scatterers, it is not convenient at this stage the introductions of limitation on the baselines. Moreover, whereas common band filtering may be beneficial to counteract the geometrical decorrelation effect in case of distributed scattering, for the full resolution analysis tuned to dominant scattering such filtering can be deleterious because in this case it cannot be assumed that spectral components of the scattering mechanism are still independent.

A possibility that provides improved performances in the full resolution analysis is given by the use of imaging techniques. Based on a tomographic processing, multidimensional SAR imaging use both the amplitude and phase values of the received signal to perform a full resolution analysis that allows increasing the detection performances of persistent scatterers, as well as the estimation of their associated parameters, i.e., the velocity and the elevation. Furthermore, by analyzing the scattering distribution in the elevation/velocity plane, they allows identifying and separating the contributions of scatterers which are interfering in the same pixel. The interference of scatterers is particularly critical in the case of complex scenarios such as urban areas, in which the phenomenon of layover affect all the vertical structures. These topics are the subject of the SAR Tomography and Differential SAR Tomography techniques which are discussed in details later on.

Figure 20.24 provides an example of application of the SBAS based approach to the monitoring of wide areas [115]. Many deformation phenomena corresponding to natural hazard are evident (the Campi Flegrei caldera subsidence in Naples, the Colli Albani uplift in the south of Rome, etc.). Besides providing the measurement of deformation associated with progressive and regular hazard sources, such as subsidence due to water pumping, regular volcanic activities, etc., the multipass SAR technique is

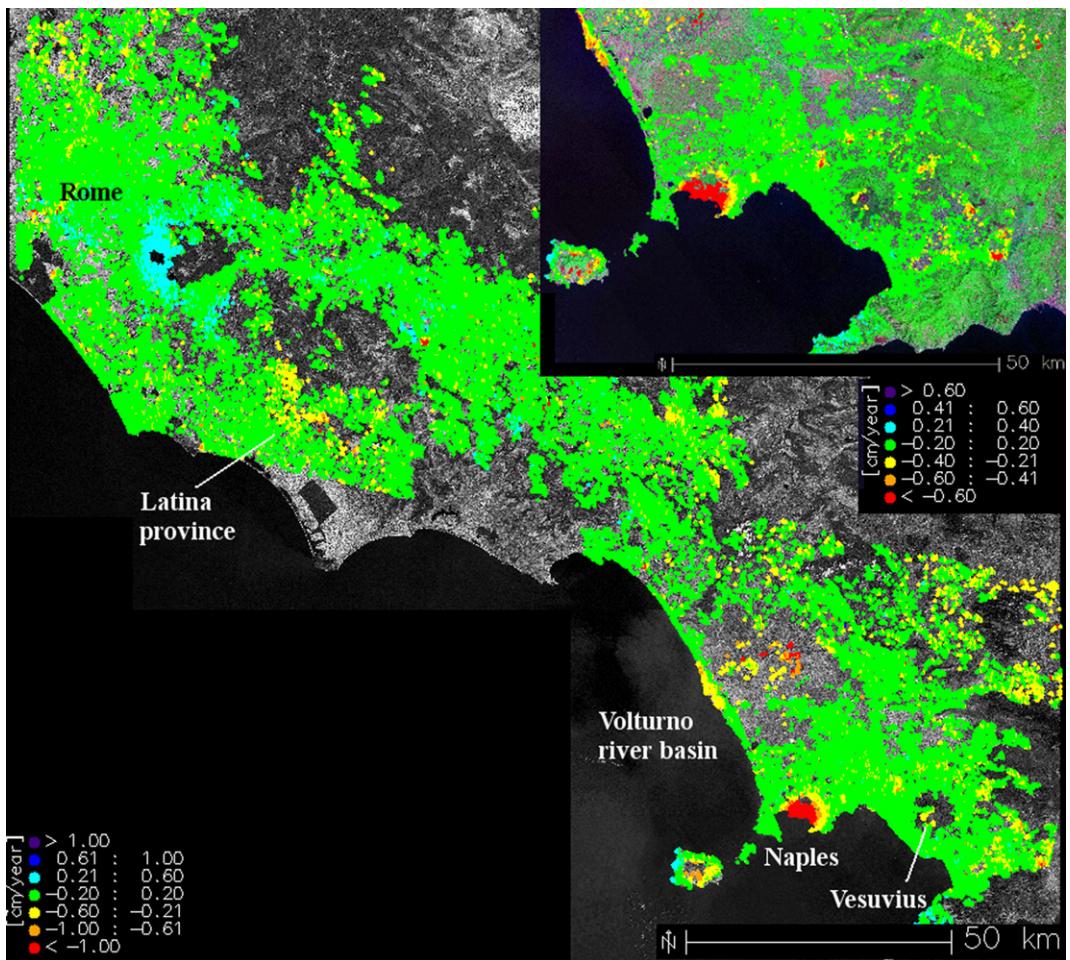


FIGURE 20.24

Example of a wide area monitoring using by combining the results of 3 ERS-Envisat frames from Naples to Rome.

also particularly important in the post crisis phase. The recent literature includes many examples of application of the multipass DInSAR technique to the monitoring of post-seismic deformation. One of the most effective example of monitoring of post-seismic deformation is provided by the CSK constellation with reference to the l'Aquila earthquake. Starting from April 6th, 2009 date the CSK constellation (at that time operative with three satellites) was capable in six months to acquire different dataset with a rate that, for the best dataset, was the highest possible, i.e., about 1 acquisition each 5 days on average.

Figure 20.25 shows the image corresponding to the mean post-seismic velocity: the red spot shows the deformation in correspondence of the Paganica San Demetrio fault. The time series also shows very clearly the typical exponentially decreasing post-seismic deformations.

Particularly relevant are the effects of uplift which are observed on the mountainous areas in the North and South area. Such effects are an evidence of the effects of the non-turbulent (stratified) component of the troposphere corresponding to the hydrostatic path delay, and show a seasonally variability. It is worth to note that the reduction of the wavelength with respect to C-Band data X-Band systems such as COSMO/SKYMED and TerraSAR-X, on one side provides a higher sensitivity to deformation, on the other side it introduces an amplification of the range variation component associated with ADP. The tropospheric component under investigation cannot be filtered via a simple approach based on the temporal independence of APD, as done in the classical multitemporal interferometric processing scheme described above. Such effects, whose compensation is of main importance for the application in emergencies situations (in which typically a few acquisitions are available), can be handled either by using external tropospheric delay measurements, provided for instance by GPS networks, or by subtracting from the interferograms the components highly correlated to the topography: the latter solution may be critical in cases in which the deformation signal is itself correlated with the topography (for instance, volcanoes inflation).

A final example of application of multitemporal DInSAR analysis is provided by the use of PSI on data acquired by the last generation, high resolution satellites. In Figure 20.26 it is shown the result of PSI applied to four ascending and descending tracks acquired by the TerraSAR-X system operating in the spotlight mode over the Berlin station. The figure provides the measurement of deformation caused by thermal dilations. At certain positions along the rail track, there is a 180° phase change. At these locations rail expansion joints are installed and the rail tracks dilate horizontally and in opposite directions.

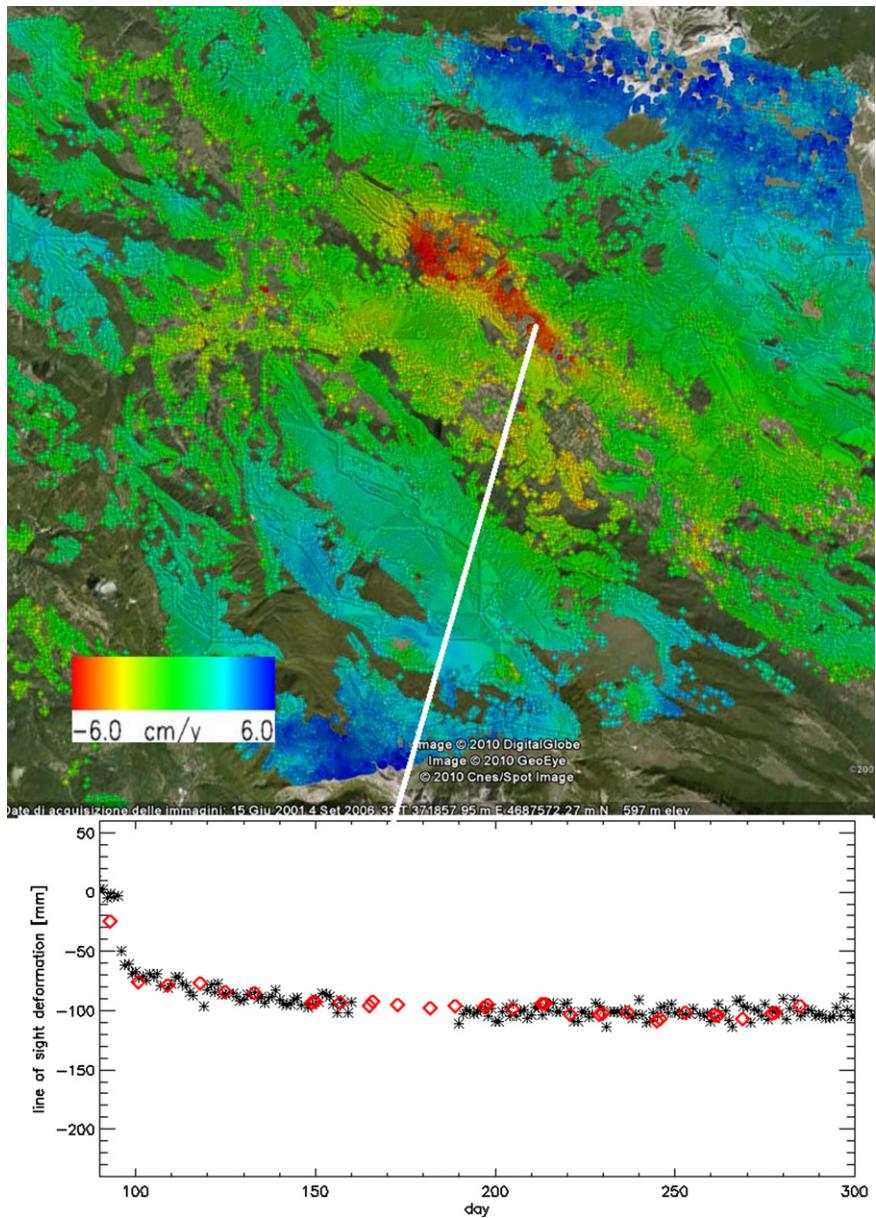
2.20.5.1 Multipass phase unwrapping

As already stated in Section 2.20.4, Phase Unwrapping is necessary to extract the APD component and to estimate the non-linear deformation signal at either small or large scale: PhU is by far the most critical step of any A-DInSAR technique.

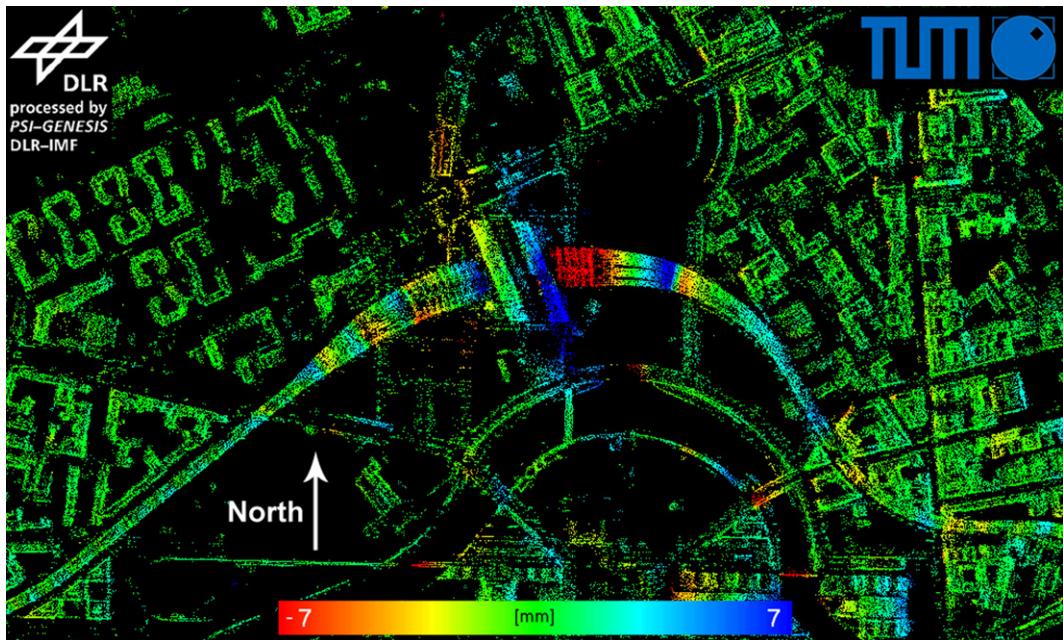
Most of PhU algorithms have been developed in the context of SAR Interferometry, and make use of the pixel-to-pixels spatial variations. One of the most used algorithm is the Minimum Cost Flow (MCF) PhU [60, 79]: it is based on the use of triangulations defined on the sparse grid of useful pixels and the PhU is cast as the problem to minimize the L1 norm of a vector that corrects the measured phase differences subject to the zero-curl (i.e., non rotational) constraint, that is the constraint that forces to zero the circulation of the unwrapped phase differences on all the triangles. Very efficient MCF solvers are available in the framework of network-flow algorithms [60].

The MCF technique, typically exploited for the unwrapping of single interferograms, has been also extensively used in the multitemporal DInSAR context where commonly several interferograms must be unwrapped. More specifically, phase unwrapping of DInSAR data stacks is carried out by independently unwrapping each interferogram with the MCF optimization approach (single-step PhU).

The single-step PhU discards any relationship between the different interferograms, which are indeed all related to the same wanted signals; particularly it neglects the typical inherent redundant nature of the interferogram stack.

**FIGURE 20.25**

Post-seismic deformation monitoring over the L'Aquila area. Upper image: post-seismic mean deformation velocity. Lower image: comparison between GPS (black stars) and SAR measurements (red diamonds) [116]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

**FIGURE 20.26**

Geocoded persistent scatterers obtained from a fusion of 4 different stacks (ascending & descending) of TerraSAR-X high resolution spotlight data [117]. Colors represent estimated amplitudes of seasonal deformation in west-eastern direction in the interval $[+7, -7]$ mm.

DLR and Technical University of Munich (TUM), Germany.

To overcome this limitation, the single-step PhU has been recently integrated with Model Based (MB) phase unwrapping algorithm that exploits the model in (20.67) also at the phase unwrapping stage. More specifically, starting from the measured interferograms, the MB PhU estimates the variations of the topography ($\Delta\hat{s}$) and deformation mean velocity ($\Delta\hat{v}$) over the set of spatial arcs defined on a grid of reliable pixels, i.e., with a coherence degree [118] which is above a threshold for a fixed percentage of the total number of interferograms:

$$(\Delta\hat{s}_j, \Delta\hat{v}_j) = \arg \max_{\Delta s_j, \Delta v_j} \frac{1}{N} \left| \sum_{i=1, \dots, N_{as}} e^{j\Delta\psi_i - j\frac{4\pi}{\lambda}(\xi_i \Delta s_j + t_i \Delta v_j)} \right|, \quad (20.74)$$

where $\Delta\psi_i$ is the variation of the wrapped phase measured on the j th (out of N_{as}) spatial arc.

The set of all measured topography and deformation mean velocity deformation mean velocity, $\Delta\hat{s}_j$ and $\Delta\hat{v}_j$, are then spatially integrated (generally in a Least-Square sense) over the network associated with the spatial arcs to retrieve the topography and deformation mean velocity at the pixel level. The phase signal corresponding to the signal component model is subtracted from each interferogram to ease the subsequent un-modeled unwrapping step, generally carried out with the application of the MCF algorithm applied independently to the each of the available interferograms [115, 119].

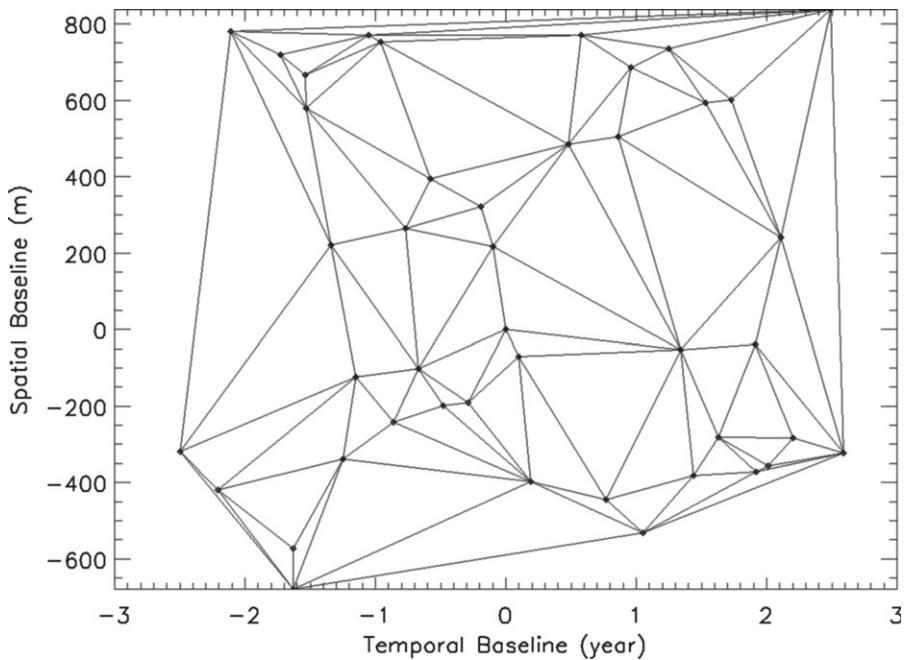


FIGURE 20.27

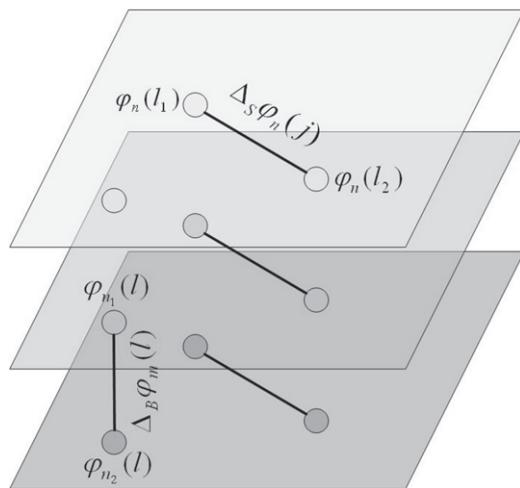
Example of triangulation in the acquisition (spatial and temporal baseline) domain.

Recently new algorithms have been proposed to improve the un-modeled PhU stage by exploiting the redundant nature of the interferograms [120–122].

A first contribution along this line is the two step PhU algorithm [122], where the use of a triangulation is in the spatial temporal baseline domain, see Figure 20.27, is used: each point represents an acquisition and arcs represent the interferograms. For each spatial arc, the interferometric measurements available in all the interferograms provide an estimate of the phase variations over the acquisition arcs in Figure 20.27. It is clear therefore that such phase variations must sum up to zero over closed circuits and therefore an MCF step can be therefore carried out in the acquisition domain to provide corrections of the phase variation over the selected spatial arc in each interferograms which are used for the subsequent MCF PhU implemented, as usual, in the spatial domain.

This two-step approach has the advantage to use the redundancy of the available interferograms, *phase summation of interferograms over close loop should be zero*, but imply a rigid scheme in the interferogram generation. The use of a triangulation does not allow fixing a limit to the baseline separation, as usually done in the SBAS approach, because the maximum baseline is defined by the triangulation scheme.

A solution to tackle problem is obtained by address the PhU problem in a generalized framework which makes use of the over-determined nature of the operator that relates the phase differences to the absolute phase values, [119]. More specifically, by referring to Figure 20.28 where a 3D representation of the stack of acquisition is shown, we let l and n indicating the pixel and the acquisition, respectively, and j

**FIGURE 20.28**

3D structure of the data corresponding to a multi-temporal multi-baseline acquisition data set: any horizontal layer corresponds to the spatial grid of a single acquisition. Vertical arcs are interferometric phase values in a pixel whereas horizontal arcs are spatial variations of interferometric phase.

the spatial arc. It is possible to define two operators implementing a differentiation along the acquisition (i.e., the interferogram generation) and along the space (spatial variation). A generic estimation of the absolute phase variation over the m th interferogram obtained by wrapping the difference of the interferogram phase values at the end point of a spatial arc can be seen as a result of a double differentiation (along the acquisitions and along the space) of the phase values:

$$u_m(j) \cong \varphi_{n_2}(l_2) - \varphi_{n_1}(l_2) - \varphi_{n_2}(l_1) + \varphi_{n_1}(l_1), \quad (20.75)$$

where n_1 and n_2 are the indexes of the acquisitions of the m th interferogram on the j th spatial arc, l_1 and l_2 are the indexes of the acquisitions of the j th arc. Collecting all the measurements u_m and unknowns φ_n for any spatial arc and interferogram in the vectors \mathbf{u} and $\boldsymbol{\varphi}$. Accordingly the following linear system can be written:

$$\mathbf{u} + 2\pi\mathbf{k} = \mathbf{A}_{SB}\boldsymbol{\varphi}, \quad (20.76)$$

where \mathbf{A}_{SB} is the matrix (typically with a number of rows larger than the number of columns due to the redundant selection of interferograms) that computes the differences at the right hand side of (20.75) along the acquisitions and space for all the interferograms (acquisition arcs) and spatial arcs; \mathbf{k} is the vector of unknown 2π multiples. From (20.76) it is evident that errors in the measurement of \mathbf{u} move the vector out of the range of the operator \mathbf{A}_{SB} , $\mathcal{R}\{\mathbf{A}_{SB}\}$, and therefore the vector \mathbf{k} must be selected as the vector that brings \mathbf{u} in $\mathcal{R}\{\mathbf{A}_{SB}\}$. It is natural, as in the minimum cost flow approach to look for a correction vector \mathbf{k} to be with integer values and with a convenient weighted minimum norm. An effective, and less computational demanding approach, is by exploiting the null-space matrix associated

with \mathbf{A}_{SB} , i.e., the matrix \mathbf{Z} whose rows are the vectors of the null space of \mathbf{A}_{SB} , $\mathcal{N}\{\mathbf{A}_{SB}\}$, i.e., such that $\mathbf{Z}\mathbf{A}_{SB} = \mathbf{0}$. By applying \mathbf{Z} to the (20.76) we determine a set of equations involving only the measurement \mathbf{u} and the unknown vector \mathbf{k} of 2π multiples. It is worth noting that in case triangulations are carried out in both the space and baseline domain \mathbf{Z} is the matrix that evaluates the circulations over the elementary triangles. In conclusion, the vector \mathbf{k} is determined by solving the following optimization problem:

$$\begin{cases} \hat{\mathbf{k}} = \arg \min_{\mathbf{k}} J(\mathbf{k}), \\ J(\mathbf{k}) = \|\mathbf{k}\|_{L^p}, \\ \text{subject to } \mathbf{Z}(\mathbf{u} + 2\pi\mathbf{k}) = \mathbf{0} \end{cases} \quad (20.77)$$

where \mathbf{Z} is the Left Null Space matrix associated with \mathbf{A}_{SB} .

The null-space approach is characterized by the desirable feature of being an approach in which the degree and the typology of redundancy of the measured interferometric phase variations (double differences) shall not follow any specific constraint, such as the triangulation scheme in the baseline domain used in [122] that imposes a critical constraint on the generation of interferograms. This characteristic allow to freely generate the set of interferograms.

2.20.6 SAR tomography

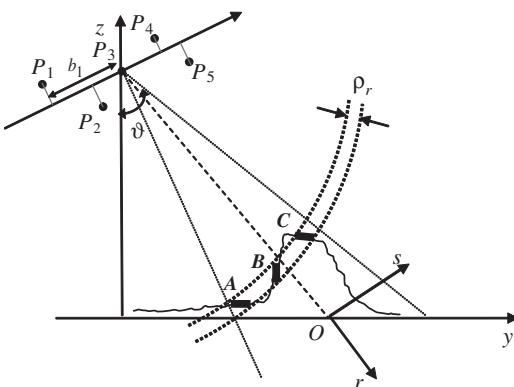
The classical SAR imaging allows high resolution capability in 2D, i.e., azimuth and range. Nevertheless, the reality is 3D and, either due to the coupling of the intrinsic side-looking geometry and the vertical structure of the targets, or to the penetration of the radiation below the imaged surface, the resulting image represents only a “projection” along the elevation direction of the illuminated 3D scene backscattering properties over the azimuth slant-range plane.

SAR tomography aims to achieving a 3D reconstruction by using the imaging diversity along the elevation direction (orthogonal to the azimuth and range plane) of (spatial) multibaseline acquisitions. In particular, as it is difficult and economically disadvantageous using several satellites to provide multibaseline acquisitions from spaceborne, SAR Tomography is typically implemented by using data corresponding to multiple passes of a single antenna SAR system over the same area. It is worth pointing out that 3D imaging capability is already available with interferometric SAR systems, relying only on the use of the phase difference between the signal acquired in at least two passes. However, InSAR implicitly assumes the presence of only a single scattering mechanism, i.e., do not comply with a possible integration (overlay) of the scattering along the elevation direction.

SAR Tomography is the extension of SAR interferometry to allow a full 3D imaging: it simply synthesizes, as in the case of the azimuth direction where a array is (digitally) formed, an array also in the elevation direction by exploiting the different baselines available over repeated passes.

Most appreciable differences with respect to the azimuth synthesis is that passes are unevenly spaced in elevation and that the number of baselines is sensibly lower than the number of echoes collected in azimuth direction. The overall effect is a generally poor elevation resolution and the presence of significant distortions associated with possible excessive non-uniformity of passes.

Let us consider a multi-pass configuration exploiting N images acquired along N different orbits, not necessarily co-planar and not uniformly-spaced (see Figure 20.29). We denote with P_n denote the n th orbit locations, b_n , $n = 1, \dots, N$ the orthogonal baselines of the n th orbit measured with respect a

**FIGURE 20.29**

Multi-Pass SAR geometry in the range-elevation (r, s) plane (case $N = 5$).

reference “master” orbit (P_3 in Figure 20.29), and ϑ is the look angle. A ground height profile with three point scatterers (A, B, and C) lying in the same range-azimuth resolution cell is depicted.

Letting h_n be the signal in a fixed azimuth and range pixel at the generic n th antenna, we have [123–125]:

$$h_n = \int \gamma(s) e^{-j \frac{4\pi}{\lambda} r_n(s)} ds, \quad (20.78)$$

where $r_n(s)$ is the distance of the scatterer at elevation s from the n th antenna. The phase term is supposed to be not affected by APD component because the acquisitions are supposed to be simultaneous. In case of repeat pass acquisitions, the APD must be compensated. Moreover, in the latter case it is assumed that the target does not exhibit any deformation: this assumption will be relaxed for the description of the differential tomography.

The data are processed by via deramping, i.e., the distance $r_n(0)$ of a reference elevation point (0 in Figure 20.29) is subtracted from the data at each antenna: expanding $r_n(s) - r_n(0)$ and absorbing s -dependent terms in $\gamma(s)$, we have:

$$g_n = h_n e^{j \frac{4\pi}{\lambda} r_n(s)} = \int \gamma(s) e^{j 2\pi \xi_n s} ds = \text{FT} [\gamma(s)]|_{\xi_n}, \quad (20.79)$$

where the integral is limited to an interval of typically from few meters to hundreds of meters, and where:

$$\xi_n = \frac{2b_n}{\lambda r}. \quad (20.80)$$

Equation (20.79) shows that the received data at the different antennas, in any fixed azimuth and range position, are samples of the Fourier Transform (FT) of the reflectivity function along the elevation direction at the frequencies described by Eq. (20.80), and follows the (LOS orthogonal component of) baseline distribution. SAR tomography processing aims at processing the data vector $\mathbf{g}^T = [g_1, g_2, \dots, g_N]$ in such a way to achieve a reconstruction of the backscattering distribution $\gamma(s)$. This is carried out by

inverting a discretized version of (20.79) at M_s elevation samples (bins) s_1, s_2, \dots, s_{M_s} :

$$\mathbf{g} = [\mathbf{a}(s_1) \ \mathbf{a}(s_2) \ \cdots \ \mathbf{a}(s_{M_s})] \boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\gamma}, \quad (20.81)$$

where vector $\boldsymbol{\gamma}^T = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_{M_s}]$. Equation (20.81) provides the discrete model to be inverted in the framework of SAR tomography. Many solutions are available [124, 126, 127]

2.20.6.1 Linear non-adaptive inversion for SAR tomography

A simple way to invert Eq. (20.81) and to recover the backscattering distribution is by applying the beamforming, i.e.,:

$$\hat{\boldsymbol{\gamma}} = \mathbf{A}^H \mathbf{g}, \quad (20.82)$$

that is:

$$\hat{\gamma}_n = \mathbf{a}^H(s_n) \mathbf{g}. \quad (20.83)$$

Note that the reconstruction of the backscattering sample $\hat{\gamma}_n$ is achieved via a filter $\mathbf{a}(s_n)$ which does not depend on the data (non-adaptive inversion).

Another possibility is offered by the Singular Value Decomposition (SVD) analysis of the operator in Eq. (20.79), i.e., by the SVD decomposition of the matrix \mathbf{A} in the discrete case. In this latter case the assumption on the support can lead to a slight degree of super-resolution with respect to the Rayleigh resolution limit for elevation, given by:

$$\rho_s = \frac{\lambda r}{2(b_{\max} - b_{\min})}, \quad (20.84)$$

where b_{\max} and b_{\min} are the maximum and minimum of $\{b_n, n = 1, \dots, N\}$. The corresponding height resolution is given by:

$$\rho_z = \rho_s \sin \vartheta. \quad (20.85)$$

SVD decomposition provides the following fundamental equations pair:

$$\mathbf{g} = \sum_{k=0}^N \sigma_k \mathbf{u}_k \mathbf{v}_k^H \boldsymbol{\gamma}, \quad (20.86)$$

$$\boldsymbol{\gamma} = \sum_{k=0}^N \sigma_k^{-1} \mathbf{v}_k \mathbf{u}_k^H \mathbf{g}, \quad (20.87)$$

where \mathbf{u}_k ($N \times 1$ vector) and \mathbf{v}_k ($M_s \times 1$ vector) are the left (data) and right (unknown) singular vectors of \mathbf{A} and σ_k are the singular values. Note that we have assumed $M_s \geq N$, i.e., a sampling of the elevation interval in a number of points which is larger than the number of acquisitions thus translating in a underdetermined characteristic of the matrix \mathbf{A} .

Equations (20.86) and (20.87) represent the fundamental result of the SVD analysis: in particulars Eq. (20.86) states that in principle all the different vectors \mathbf{u}_k concur in the composition of the observed vector \mathbf{g} ; their contributions are however weighted by the singular values. Accordingly, low singular

values σ_k indicates that the component along the corresponding data singular vector is attenuated by the imaging operator in the data formation and should be carefully treated when reconstructing the unknown via Eq. (20.87) because this data component may be overwhelmed by the unavoidable noise. As a result, typically the reconstruction in Eq. (20.87) is limited to the data and unknown singular vector pairs corresponding to significant singular values; the associated inversion scheme is referred to as Truncated SVD (TSVD). In the case in which the acquisitions are uniformly spaced with a separation of b_\perp if the output sampling is chosen verifying the Nyquist conditions, i.e., $\Delta s = \lambda r / (2Nb_\perp) = \rho_s$, i.e., N samples in the Nyquist interval $(-a, a)$ with $a = \lambda r / (4b_\perp)$ it can be shown that the operator \mathbf{A} in (20.81) is a Discrete Fourier Transform (DFT) matrix that is characterized by constant singular values. In such a case $\sigma_k = \sigma$ and therefore the direct and inverse operator in (20.86) and (20.87) becomes Hermitian conjugate pairs. In cases in which $\Delta s < \lambda r / (2Nb_\perp)$ the singular values shows a decay which translate the presence of a redundancy in the acquired spectral samples that can be exploited to increase the resolution of the reconstruction below the Rayleigh limit in (20.84).

In other words, by progressively restricting the elevation ROI with respect to $N\rho_s$, a smooth decay in the singular values is observed. For a reduction factor $F < 1$, if the noise is small enough, the truncation can be arrested to a number of singular values which is larger than FN : in this condition a degree of super resolution is achieved [124].

In addition to the super-resolution degree, the SVD decomposition allows also to reduce the level of sidelobes with respect to a reconstruction carried out by simple Beamforming.

2.20.6.2 Linear adaptive inversion for SAR tomography

A more general expression for the evaluation of $\hat{\gamma}_n$ is given by:

$$\hat{\gamma}_n = \mathbf{f}_n^H \mathbf{g}, \quad (20.88)$$

where \mathbf{f}_n^H is the filter for the estimation of $\gamma_n = \gamma(s_n)$. Letting $\mathbf{C}_g = E[\mathbf{g}\mathbf{g}^H]$, that is the data covariance matrix, a solution obtained *mutatis mutandis* from the spectral estimation theory such that:

$$\begin{aligned} \hat{\mathbf{f}}_n &= \arg \min_{\mathbf{f}_n} E[|\mathbf{f}_n^H \mathbf{g}|^2] = \arg \min_{\mathbf{f}_n} \mathbf{f}_n^H \mathbf{C}_g \mathbf{f}_n, \\ &\text{subject to } \mathbf{f}_n^H \mathbf{a}(s_n) = 1, \end{aligned} \quad (20.89)$$

i.e., a solution that achieves the minimum output power, subject to unitary gain at the ‘frequency’ (s_n) of interest (Capon filter), is provided by [128]:

$$\mathbf{f}_n = \frac{\mathbf{C}_g^{-1} \mathbf{a}(s_n)}{\mathbf{a}^H(s_n) \mathbf{C}_g^{-1} \mathbf{a}(s_n)}. \quad (20.90)$$

Substituting (20.90) in (20.88) provides:

$$E |\hat{\gamma}(s_i)|^2 = \frac{1}{\mathbf{a}^H(s_i) \mathbf{C}_g^{-1} \mathbf{a}(s_i)}. \quad (20.91)$$

It is interesting to note that $\mathbf{C}_g = \mathbf{I}$, white data spectrum, leads to $\mathbf{f}_n = \mathbf{a}(s_n)/N$. The advantage of the Capon filer is the achievement of higher super-resolution for line spectra (i.e., concentrated scatterers

along s) compared to the SVD. However a disadvantage of the Capon filter is the need to estimate the data Covariance matrix. This estimation is carried out via spatial averaging, and operation known in the SAR processing as multi-look:

$$\mathbf{C}_g = \frac{1}{L} \sum_{l=1}^L \mathbf{g}(l) \mathbf{g}^H(l), \quad (20.92)$$

where $\mathbf{g}(l)$ is the data vector in the l th pixel located close to the pixels in which the tomographic processing aims to provide an estimate of the backscattering distribution.

Note that this operation inevitably leads to a loss of azimuth-range resolution. Note also that the reconstruction of the backscattering sample $\hat{\gamma}_n$ is in this case data dependent (adaptive inversion).

An example of reconstruction of a 3D SAR image on real data is provided in Figure 20.30: in this case the results are the first one obtained by using spaceborne SAR data, specifically data acquired

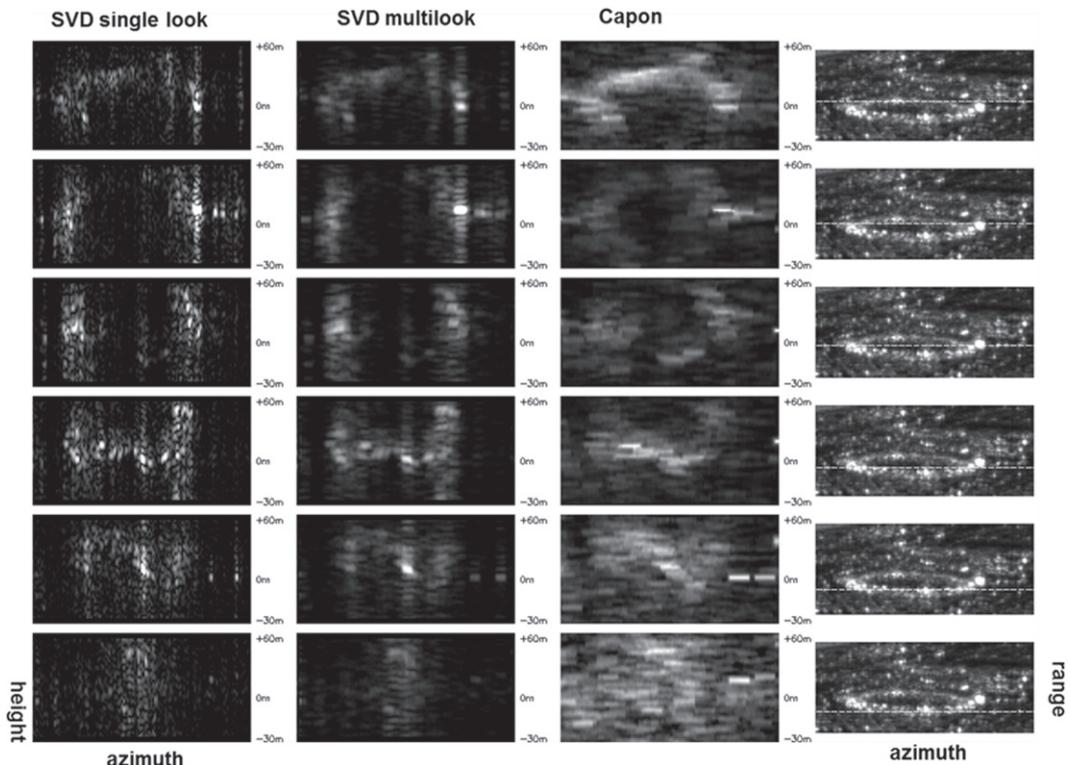


FIGURE 20.30

3D reconstruction of the San Paolo stadium in Naples via SVD. Left most column: elevation sections obtained by applying the SVD to the single look data. Second column: SVD with three looks. Third column Capon with five looks. In all cases the height scale is reported. Right column: azimuth range (averaged) image: the white horizontal lines indicate the position of the sections with the range.

by the ERS sensor [124]. In the first three columns the reconstructions obtained by SVD and Capon are reported. In particular six sections in the azimuth-elevation plane (i.e., at constant slant range) of the San Paolo stadium in Naples are shown: the locations of the constant range are indicated by the horizontal white line in the last column showing the classical azimuth-range representation. It is evident the capability of the tomographic technique to capture the 3D shape of the structure. It is also evident for the Capon approach that the increase in the height resolution is paid in terms of spatial (azimuth and/or range) resolution losses.

2.20.6.3 Compressive sensing inversion for SAR tomography

A recent approach to solve the SAR tomography problem is based on Compressive Sensing [123, 126, 129]. Compressive Sensing CS is a model-based framework for data acquisition and signal recovery based on the premise that a signal having a sparse representation in one basis, can be reconstructed from a small number of measurements collected in a second basis, that is incoherent with the first [130]. In our case sparseness requires a small number of stable targets in the same range-azimuth resolution cell [123].

In SAR Tomography, this approaches is particularly effective [123, 131] because on data acquired high frequency systems (C- or X-band sensors), the scattering in the same range-azimuth cell along the elevation occurs typically on a few concentrated scattering points, as shown in Figure 20.29.

Considering the tomographic model (20.81), the matrix \mathbf{A} in the context of CS is called the *measurement matrix* [130]. As already commented before, the inversion of Eq. (20.81) is equivalent to an inverse FFT operation, and would provide an estimate of the reflectivity function with a nominal 3-dB elevation resolution given by Eq. (20.84).

In practical cases, the orbits are usually not uniformly spaced and the number N of the available acquisitions is generally much lower than the number of unknown samples M_s . As mentioned before in Section 7.1, truncated SVD (TSVD) can be used [125] for the inversion of (20.52). An alternative technique is based on Compressive Sampling [126]. It exploits the sparsity property of the unknown vector $\boldsymbol{\gamma}$ and allows obtaining very satisfactory reconstructions, even when a reduced number of acquisitions, almost randomly spaced within the overall elevation aperture, is available. Moreover, increased resolutions can be obtained adopting CS.

In the considered sparsity hypothesis, the sampled reflectivity function can be written as:

$$\boldsymbol{\gamma} = \boldsymbol{\Psi}\boldsymbol{\alpha}, \quad (20.93)$$

where $\boldsymbol{\Psi}$ is the *sparsity matrix* and $\boldsymbol{\alpha}$ is a sparse M_s -dimensional vector. It has to be noted that in this case, by choosing sampling points to represent the unknown scattering distribution, and considering the circumstance that such function is sparse just in the domain of spatial samples (see also Figure 20.29, where only $K=3$ scattering contributions are present in the whole scattering distribution along the elevation), the matrix $\boldsymbol{\Psi}$ becomes the identity matrix [123].

An estimate of the vector $\boldsymbol{\alpha}$ can be found solving the L_1 -norm minimization problem [130]:

$$\hat{\boldsymbol{\alpha}} = \arg \min \|\boldsymbol{\alpha}\|_1 \text{ subject to } \mathbf{g} = \mathbf{A}\boldsymbol{\Psi}\boldsymbol{\alpha}. \quad (20.94)$$

The optimization problem in (20.94) is valid for the noiseless case. In the more realistic case, some noise is present in addition on the measurements:

$$\mathbf{g} = \mathbf{A}\boldsymbol{\Psi}\boldsymbol{\alpha} + \mathbf{w} \quad (20.95)$$

with w a complex Gaussian vector with zero mean and uncorrelated elements. In this case, the solution can be found by solving the linear programming problem [132]:

$$\hat{\alpha} = \arg \min \|\alpha\|_1 \text{ subject to } \|\mathbf{A}\Psi\alpha - g\|_2 \leq \varepsilon, \quad (20.96)$$

where ε is a small positive number.

For discussing the super resolution issue, we can refer to the case of high signal to noise ratio, so that the maximum resolution that can be achieved depends only on the acquisition configuration and not on the noise level. The presence of an higher noise level can only degrade the evaluated performance.

In practical cases $N < M_s$, so that Eq. (20.81) expresses an underdetermined system, which will admit a not-unique solution. Anyway, CS theory [130, 132] ensures that if it is satisfied an incoherency property between \mathbf{A} and Ψ , it is indeed possible to recover the K largest components α_i of α from the N measurements provided that the following inequality is satisfied:

$$N \geq C\mu^2(\mathbf{A}, \Psi) K M_s (\log M_s)^4, \quad (20.97)$$

where C is a small constant depending on the measurement and sparsity matrixes \mathbf{A} and Ψ , which can be empirically estimated by numerical simulations, K denotes the number of non-zero coefficients of α , and $\mu(\mathbf{A}, \Psi)$ is the mutual coherence between the measurement basis \mathbf{A} and the representation (sparsity) basis Ψ , defined in Ref. [132].

As far as the improvement or resolution, in Ref. [133] it has been shown that the elevation resolution (null-to-null) which can be obtained adopting a CS approach is given by:

$$\rho_{sCS} = \rho_s / \eta_{\text{sup}}, \quad (20.98)$$

where ρ_s is the standard Fourier techniques elevation resolution and η_{sup} is the super-resolution factor (greater than one).

For a fixed number of acquisitions N , extension of illuminated scene in the elevation direction S , a number of scatterers K , and for a given standard resolution ρ_s , there is an upper limit for the super-resolution factor given by [133]:

$$\eta_{\text{sup}} = \exp \left[\left(\frac{N}{CK} \right)^{1/4} \right] \frac{\rho_s}{2S}. \quad (20.99)$$

Combining Eqs. (20.98) and (20.99), a limit to the maximum resolution can be expressed as:

$$\rho_{sCS} = 2S \exp \left[- \left(\frac{N}{CK} \right)^{1/4} \right], \quad N \leq M_s. \quad (20.100)$$

Equations (20.99) and (20.100) provide the super-resolution limits of the CS approach applied to SAR tomography.

Some experiments about the performance of the Compressive Sensing approach to solve the tomographic problem in terms of resolution capabilities, are presented in Ref. [123], with reference to simulated data using COSMO-Skymed system parameters (see Table 20.1), and to the real data acquired by the sensors ERS1–2, over the city of Naples, between 1998 and 2001.

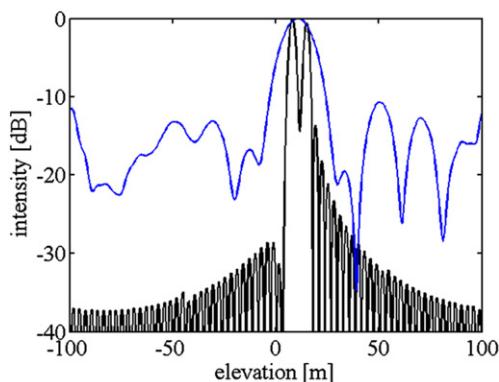
Table 20.1 COSMO-Skymed System Parameters

COSMO-Skymed Parameters	
Wavelength	0.031 m
View angle	23°
Range distance	670 km
Ground range resolution	3 m
Azimuth resolution	3 m

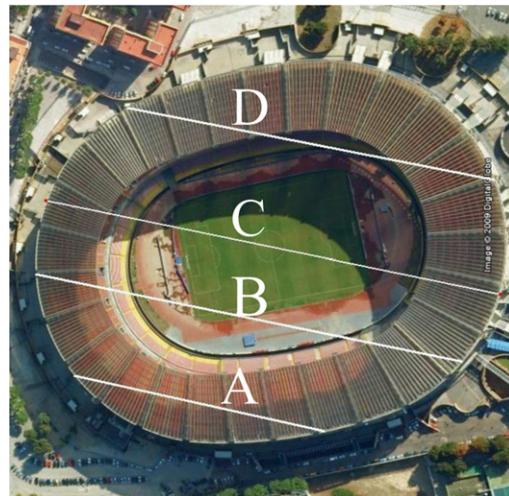
The simulated observed scene is composed of two stable and coherent scatterers located in the same range-azimuth resolution cell at different elevations and responding with the same radar cross section.

In the simulation presented in Figure 20.31, it has been assumed that the SAR signals are acquired along multiple passes with a total orthogonal baseline span ($b_{\max} - b_{\min}$) = 567 m. The theoretical resolution in the elevation direction is given, according to Eq. (20.84), by $\rho_s = 36.6$ m. Assuming that the elevation extension of the ground scene is $S = 200$ m, the orbit spacing in the elevation direction corresponding to the Nyquist sampling rate is equal to 52 m [133].

Two scatterers at the elevation values $s_1 = 7$ m and $s_2 = 14$ m are considered, so that they cannot be resolved using standard FFT based techniques, since their distance is below the system configuration resolution. Moreover, $N = 9$ orbits with orthogonal baselines -271 m, -238 m, -154 m, -68 m, 0 m, 64 m, 118 m, 209 m, 295 m, are considered, and the acquired signals have been corrupted with additive Gaussian noise, whose average power is 30 dB below the signal power.

**FIGURE 20.31**

TCS interpolated and normalized reflectivity profile (black) of two scatterers at $s_1 = 7$ m and $s_2 = 14$ m obtained from nine orbits with $\eta_{\text{sup}} = 6$, compared with the TSVD reconstructed profile (blue) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

**FIGURE 20.32**

Stadium San Paolo in Naples (ITALY).

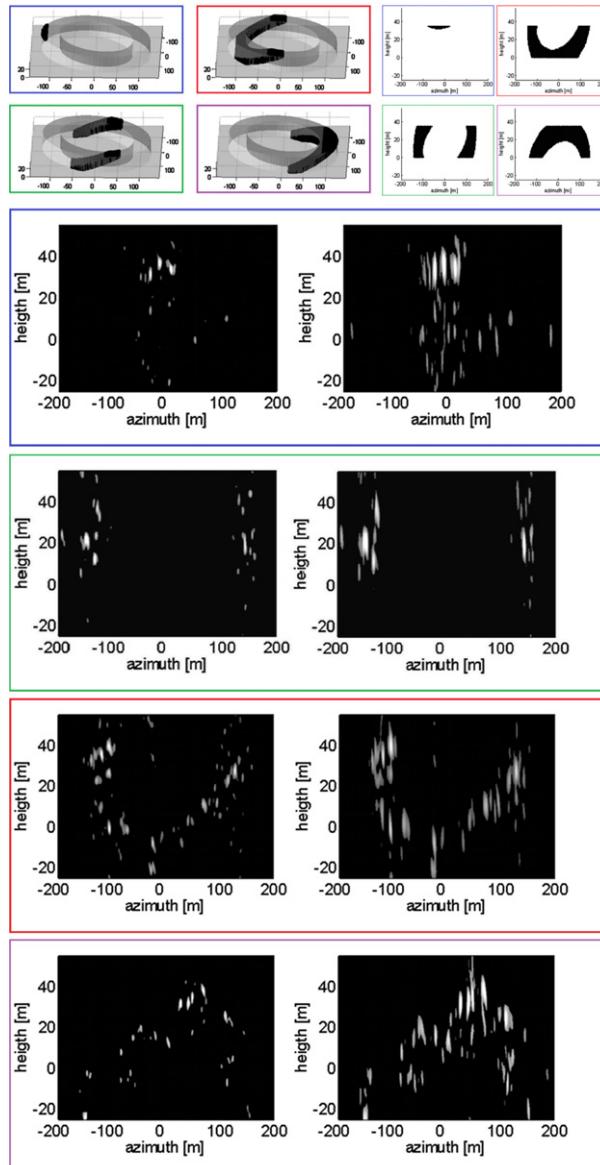
The elevation profile obtained by the CS approach starting from the nine available acquisitions and with $\eta_{\text{sup}} = 6$, corresponding to the upper limit for the resolution $\rho_{s\text{CS}} = 5 \text{ m}$ is shown in Figure 20.3 (black line) where it is compared with the TSVD reconstruction (blue line), this latter able to achieve conventional “Fourier” resolutions.

In the real data experiment, 15 passes, whose orthogonal baselines span a total baseline $S_T = 893 \text{ m}$, are used. The resulting theoretical 3 dB elevation resolution is $\rho_s = 26.6 \text{ m}$, which corresponds to a height resolution $\rho_z = 10.39 \text{ m}$. The tomographic processing of the four azimuth-height constant range sections, indicated with four segments in the ortophoto image shown in Figure 20.32, are focused and shown in Figure 20.33 using CS (left) and TSVD (right). For the CS reconstruction, a super-resolution factor $\eta = 3$, corresponding to elevation and height resolutions of 8.87 m and 3.49 m , has been assumed. Also in this case, the sections of the San Paolo stadium in Naples in elevation (i.e., at constant slant range) shows, at different slant ranges, the capability of the technique to capture the 3D shape of the structure. A resolution improvement of the CS method with respect to the TSVD one is evident.

2.20.6.4 Performance comparison of SAR tomography methods

In the previous sections four SAR tomography approaches were described: Beamforming and TSVD, which attain an elevation resolution related and limited by the overall baseline span (see Eq. (20.84)), Capon, which allows super-resolution along the elevation direction, but requires the use of looks and therefore a loss of resolution, and Compressive Sensing based techniques, which exploits the sparsity of the scattering profile in elevation, and allows a strong improvement in elevation resolution.

The four method are compared with a data set simulating an observed scene composed by three stable and coherent scatterers, located in the same range-azimuth resolution cell at different elevations,

**FIGURE 20.33**

Azimuth height sections of tomographic reconstruction over the San Paolo stadium in Naples (ITALY), obtained with TCS (left) and TSVD (right). (Blue) section A, (Red) section B, (Green) section C, (Purple) section D of Figure 20.32. Credit: Univ. Napoli Parthenope; © IEEE. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

and responding with the same radar cross section. More precisely the three scatterers are located at the height values $z_1 = -6.83$ m, $z_2 = 0$ m, and $z_3 = 6.83$ m.

In the simulation presented in Figure 20.34, it has been assumed that the SAR signals are acquired along 30 multiple passes with a total orthogonal baseline span ($b_{\max} - b_{\min}$) = 1066 m and therefore a mean baseline separation of 36.7 m: such a baseline distribution has been selected according to a real dataset of the Envisat satellite. According to Eq. (20.84), the theoretical resolution in the elevation direction is given by $\rho_s = 22.3$ m, corresponding to a height resolution $\rho_z = 8.7$ m. The three targets have been therefore located in two resolution cells (in elevation). A total of 15 independent looks have been generated by superimposing to the complex scatterer reflectivity a uniform phase independent from look to look. Moreover, the acquired signals have been corrupted with additive Gaussian noise, whose average power is 10 dB below the signal power.

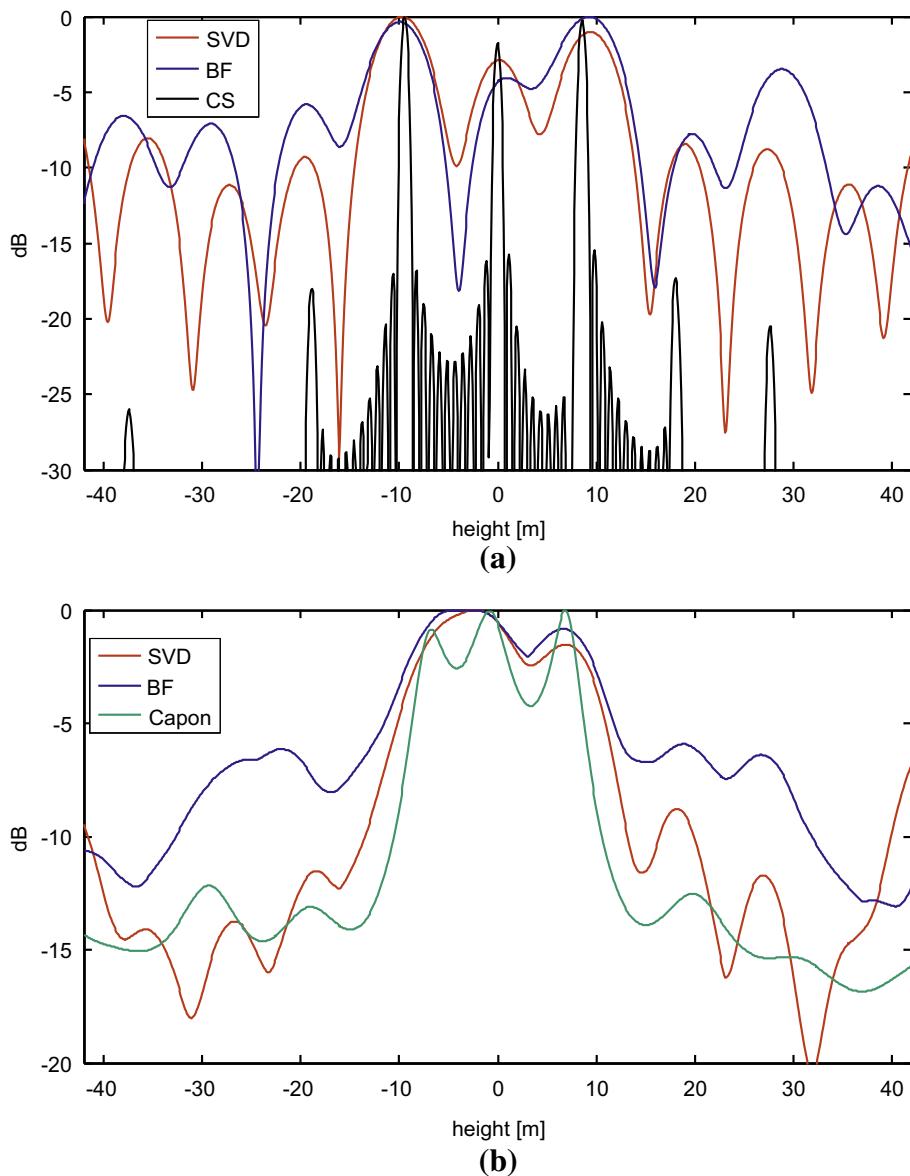
The elevation interval $2a$ corresponding to the mean baseline is 645 m: this interval has been restricted with a reduction factor given by $F = 1/3$, thus leading to an investigated interval of 215 m.

The reconstruction of the height profile obtained by the CS, SVD and Beamforming tomographic approaches, starting from the 30 available acquisitions have been generated separately for all the available looks. The result for a particular look are shown in Figure 20.34a; all approaches are able to recover correctly the presence of three scatterer present in the profile and to resolve them from each other although the super-resolution capability of the CS based approach appear to be very evident. Due to the particular realization of the noise the scatterers separation is slightly overestimated and therefore the three targets are easily visible also from the Beamforming and SVD algorithms: the latter has slightly better performances on the resolution and on the sidelobe ratio.

Beamforming, SVD, and Capon were applied to the multilook data: the result are shown in Figure 20.34b. On one hand it is evident that on average the Beamforming and SVD are not able to resolve the scatterers, although the SVD provide generally better performances with respect to the Beamforming, particularly on the sidelobe ratio. On the other hand, it is evident that in this case the Capon provide better resolution capability with respect to the other two processing approaches, by resolving the presence of the three distinct scatterers which are also located at the three correct heights. In any case differently from the CS the super-resolution degree is paid by the necessity to have different looks. A more detailed analysis of the performances of the CS tomography in terms of resolution degree can be found in [129].

2.20.7 4D Imaging

The 3D SAR focusing technique, also known as SAR Tomography allows to profile the scattering distribution along the elevation direction. Differential SAR Tomography, also referred to as 4D (3D space + velocity) SAR imaging (focusing) is a natural extension of SAR Tomography to targets that exhibit displacements. It allows measuring the scattering distribution in an elevation–velocity (EV) plane, also known as tomo-doppler plane: locations of peaks in the EV plane allows identifying and measuring elevation and velocity of scatterers, even interfering in the same resolution cell. The original idea in [134] is framed in the statistical context and makes use of multilook data to estimate the data covariance matrix and then to apply adaptive (non-linear) estimation, namely Capon filtering to achieve superresolution and sidelobe reduction.

**FIGURE 20.34**

Results of SAR tomography on simulated data. (a) Reconstructions obtained by Beamforming, SVD and Compressive Sensing on single look data. (b) Results of Beamforming, SVD and Capon for multilook data.

Again referring to 0 (the deramping reference point) we let the scatterer located at elevation s to have a deformation in the line of sight equal to $d(s, t_n)$, much smaller than the resolution cell; accordingly, in place of Eq. (20.79) we have the following direct model for the signal collected at the generic antenna following the deramping step:

$$g_n = \int \gamma(s) e^{j2\pi\xi_n s + j\frac{4\pi}{\lambda}d(s, t_n)} ds. \quad (20.101)$$

Moreover, the Fourier expansion of the second exponential term with respect to t_n provides [114]:

$$e^{j\frac{4\pi}{\lambda}d(s, t_n)} = \int a(s, v) e^{j2\pi\eta_n v} dv, \quad (20.102)$$

with:

$$\eta_n = \frac{2t_n}{\lambda}. \quad (20.103)$$

Accordingly, Eq. (20.101) becomes:

$$g_n = \iint \gamma(s, v) e^{j2\pi\xi_n s + j2\pi\eta_n v} dv ds, \quad (20.104)$$

where $\gamma(s, v) = \gamma(s)a(s, v)$.

Equation (20.104) shows that the received data are samples of the 2D Fourier Transform of the scattering distribution in the EV plane: as in the 3D case it can be inverted to estimate the backscattering distribution in the EV plane.

It is worth to note that, according to Eq. (20.102), v assumes the meaning of a spectral velocity. In analogy to the concept of instantaneous and spectral frequency in frequency modulation, there is strictly speaking no relation between v and the instantaneous temporal velocity. An exception is represented by the case in which the targets exhibit a linear temporal deformation; in this case instantaneous and spectral velocities are equal. However, under quite general conditions [135] the spectral support and the instantaneous frequency variation interval turn out to be equal. Accordingly, the appearance of focused energy in v in the elevation-velocity plane indicates the presence of scatterers with regular phase variation that can be exploited to monitor the scatterer.

MultiD (3D and 4D) focusing technique relies on the assumption that the data stack is accurately calibrated in amplitude and phase. While amplitude calibration can be easily carried out either via the use of auxiliary information provided with the data, or via simple equalization of the power on selected uniform areas or stable strong scatterers; phase calibration may be a much more complex issue due to the presence of decorrelation phenomena, APD variation and non linear deformation: the latter allows higher focusing of peaks corresponding to persistent scatterers in the EV plane.

The phase calibration step, designed to remove the atmospheric phase and low resolution (background) deformation from the signal data stack, a easy solution is by using a low resolution (small scale) processing with SBAS based interferometry approaches, to estimate and compensate for APD variation and possible non linear deformation.

Finally, a key problem associated with Multi-D imaging is the detection of scatterers that maintain persistent scattering characteristics over the time. Following the key idea of the Persistent Scatterers

(PS) technique, in classical PSI techniques this step is carried out in a sequence of two stages. In the first the so called Persistent Scatterers Candidates [99, 100] are selected by analyzing the stability of the amplitude response, i.e., by up-threshold, in each pixel, of the ratio between the standard deviation and the mean of the amplitude response, named amplitude dispersion index:

$$D_A = \sqrt{N} \frac{\left\{ \sum_{n=1}^N [|g_n| - E(|g_n|)]^2 \right\}^{1/2}}{\sum_{n=1}^N |g_n|}. \quad (20.105)$$

The amplitude dispersion index is close to zero over targets showing a high stable backscattering and is proxy of the phase stability. PSC are used in PSI techniques to carry the phase calibration of the data, i.e., the estimation and removal of the Atmospheric Phase Delay (APD) contribution. In the second step, the detection of Persistent Scatterers is refined by down-threshold of the temporal coherence index in Eq. (20.68): all pixels that show a temporal coherence above the threshold are selected as PS and topography and motion parameters are provided.

In Multi-D SAR imaging the amplitude information is directly involved in the formation of the elevation profile (3D imaging) or elevation-velocity image. As far as dominant scatterers are concerned, a detector (the Generalized Maximum Likelihood Ratio Test, briefly GLRT) [136] which is derived from an optimal detection scheme (the Neuman-Pearson criterion) which maximizes the detection probability for a fixed false alarm probability is provided by [137]:

$$\max \left[\frac{|\mathbf{a}^H \mathbf{g}|}{\|\mathbf{a}\| \|\mathbf{g}\|} \right] < T, \quad (20.106)$$

where the maximum has to be considered over s for the 3D case, or (s, v) for the 4D case, and T is a threshold belonging to the interval $(0, 1)$, set accordingly to the desired false alarm probability.

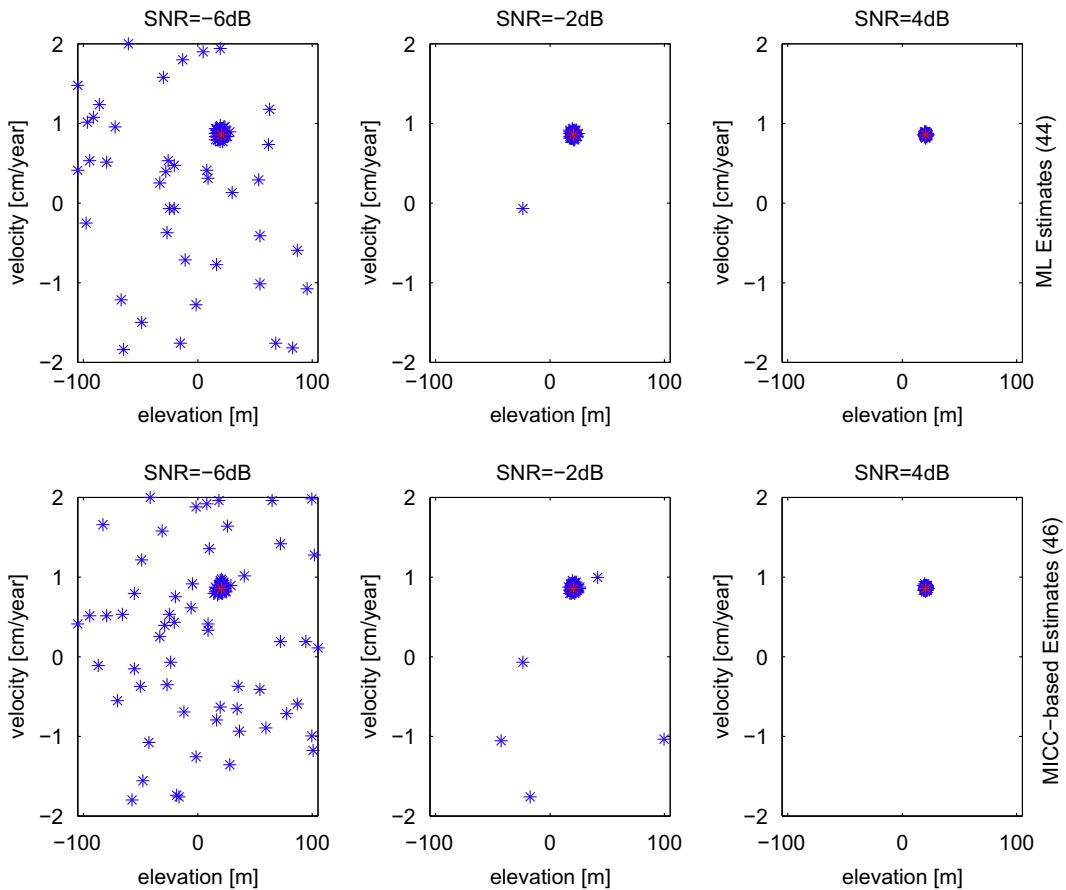
It is instructive to note that in the PS case, the test is obtained by letting $\mathbf{g} = \mathbf{y}_\varphi$, i.e., by eliminating from the data the amplitude information. It has been shown that, for a fixed false alarm probability, the tomographic based test allows achieving typically a gain of 1 dB in the detection of dominant scatterers [137]. Note that Eq. (20.105) is based on the selection of the peaks of the (normalized) tomographic response ($\mathbf{a} = \mathbf{a}(s)$ for 3D, $\mathbf{a} = \mathbf{a}(s, v)$ for 4D) achieved by simply applying a beamforming inversion scheme, i.e., by multiplying by the data by the transpose conjugate the direct matrix.

Also in terms of accuracy of the estimated mean deformation velocity and residual topography contributions the adoption of the detection scheme based on tomographic processing provides improvements. In Figure 20.35 it is shown a comparison, in terms of scatter plot of the estimated elevation for a tomographic processing (top) and for a PSI estimation.

Further studies have been carried out for the development of detections schemes able to perform a higher order analysis to detect and separate the contribution of scatterers interfering in the same resolution cell [138–141].

An example of the results obtained via 4D imaging are shown in Figure 20.36: in this case the dataset was acquired by the ERS sensor the city of over Rome. All detected single and double scatterers where geolocalized: due to the geolocalization process, scatterers located at different elevations and interfering in the same azimuth-range pixels are projected at two different ground positions, see Figure 20.37.

The detected single scatters (shown in the uppermost image) show the presence of deformations affecting the downtown of Rome [142]. In the two images on the left, a zoom on the results of single

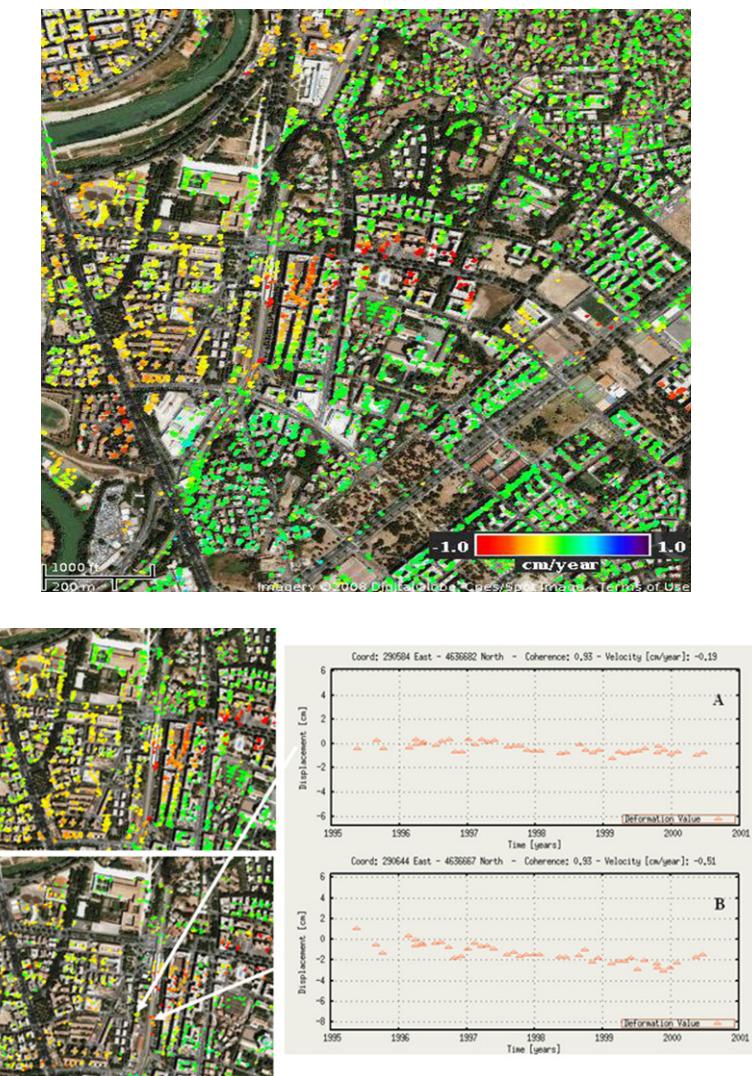
**FIGURE 20.35**

Scatter plots of the Tomographic (top) and PSI-based (bottom) estimators for mean deformation velocity and topography for different values of the SNR.

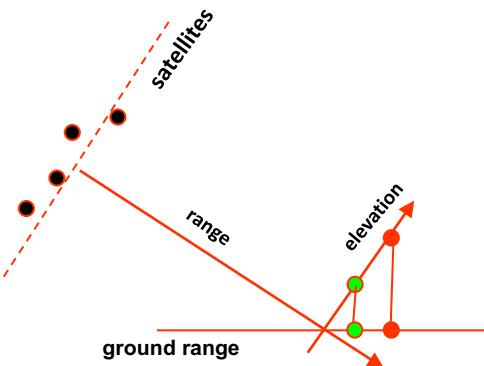
and double scatterers are shown to demonstrate the capability of the detected double scatterers in 4D imaging to provide important information in an area where layover impacts the monitoring capabilities, see the railway tracks. The right plots show the capability of 4D imaging to separate time series from interfering scatterers.

Even when only a dominant scattering mechanism, it has been shown that the use of both amplitude and phase information in MDI allows to improve the performance in terms of dominant persistent scatterers detection with respect to classical PSI that uses only phase information [137].

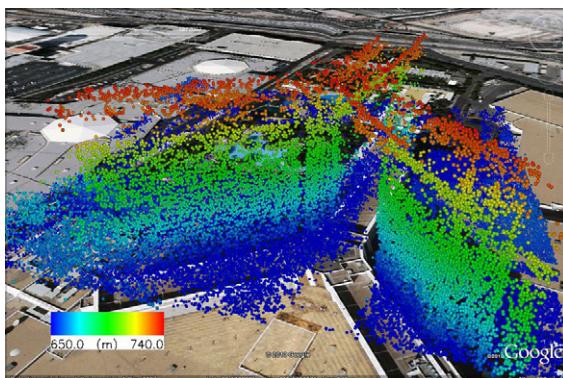
In the following the result of a MDI processing of a set of 25 TerraSAR-X (TSX) spotlight data acquired over ascending orbits over the city of Las Vegas, Nevada, USA is shown [19]: in this case the

**FIGURE 20.36**

Example of separation of scatterers interfering in the same azimuth range pixel via multi-dimensional imaging and determination of velocity and time series. Top image: estimated velocity of dominant scatterers in the area of Rome. Bottom left: zoom on in an area exhibiting deformation due to alluvial deposits of the Tevere river: representation of dominant (upper image) and double (down image) scatterers. Right images: time series extracted by the double scatterers analysis.

**FIGURE 20.37**

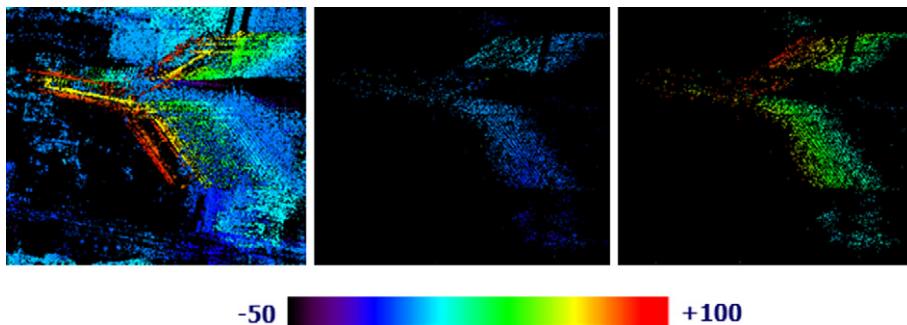
Graphical explanation of the mechanism of projection onto the ground range showing that two different scatterers located at the same range and resolved by the tomographic technique are located at different positions.

**FIGURE 20.38**

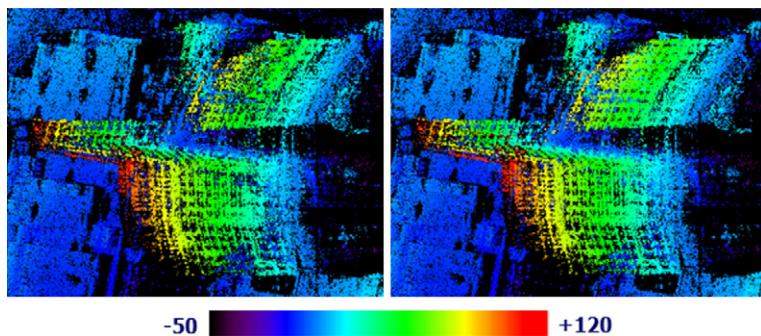
Example of reconstruction at full resolution of the Mirage Hotel in Las Vegas, USA obtained by processing 1 m resolution spotlight data with a 4D imaging approach. Credit IREA-CNR and DLR, © IEEE.

improvement of resolution from $5 \text{ m} \times 20 \text{ m}$ of ERS to $1 \text{ m} \times 1 \text{ m}$ of TerraSAR-X makes the layover of buildings much more evident.

In particular, in Figure 20.38 it is shown the reconstruction of the Mirage Hotel obtained by detecting persistent scatterers on a 1st order (single scatterers) and 2nd order (double scatterers interfering in the same pixel) analysis. The possibility to synthesize a fine beam also in the elevation direction provide the TerraSAR-X sensor with the capability of reconstructing the building in 3D. In Figure 20.39 it is shown the result of the multidimensional analysis, specifically the topography for the single and double scatterers: the mid and right images shows the capability of the tomographic based processing of solving the layover and hence separating the contribution of the two layers of the ground and façade interfering due to the mechanism of folding of the vertical building toward the left (near range). The measured

**FIGURE 20.39**

Detected scatterers for the formation of image in Figure 20.38. From left to right, single scatterers, lower layer of double scatterers and top layer of double scatterers. Credit IREA-CNR and DLR, © IEEE.

**FIGURE 20.40**

Results on the Bellagio hotel for the Las Vegas dataset: single scatterers (left), single and double scatterers (right). Colors are associated with the estimated topography.

deformations show essentially only the presence of thermal dilations for the roof area. Another example of the effectiveness of multi-dimensional imaging is provided in Figure 20.40 for the Bellagio hotel. The image on the right shows the final set of detected single and double scatterers and the results, especially on the upper façade of the building where interference with the ground is much critical, shows a significant improvement in the reconstruction of the building structure.

2.20.8 Conclusion

This review work has addressed the topics related to SAR interferometry and SAR Tomography. For the first topic we have described in details both from a deterministic and stochastic viewpoint the single

and multibaseline interferometric technique for the estimation of the topographic height. Furthermore, the differential interferometry technique has been addressed for the description of the possibility to accurately monitor small deformation of the Earth surface, especially with reference to the application to volcanic and seismic risk. Persistent scatterers interferometry, the differential interferometry technique that has for the first time shown the possibility to accurately monitor deformation, finds it “killer application” to monitoring of man-made structures and has been also addressed. Furthermore whole sections have been devoted to SAR tomography, also known as multi-dimensional imaging, which is a technique based on an imaging approach and extends the concepts of interferometry. By using the amplitude and phase information collected over multi-baseline multipass data it provides the best technology currently used for the imaging and monitoring of urban areas and infrastructures. SAR tomography is also applied for volume scattering profiling with application to forest mapping [125], particularly in conjunction with polarimetry with a technique known as Polarimetric SAR Tomography. In this review work we have limited our analysis to surface scattering single polarization: Further description of SAR Tomography and Polarimetric SAR tomography can be found in [21, 22, 143]. Finally application examples including results obtained with the recent very high resolution SAR sensors have been provided to give the reader the possibility to have information about the current applicability, potential and limits of the coherent SAR processing.

Acknowledgments

The Authors wish to thank the anonymous Reviewers for their comments which contributed to improve the quality of the paper, and Prof. Gilda Schirinzi, Prof. Alessandra Budillon, Prof. Giampaolo Ferraioli, and Dr. Fabio Baselice, of the Università di Napoli Parthenope, Italy, and Dr. Diego Reale of IREA-CNR, Napoli, Italy, for the valuable discussions about the main topics of the paper. Moreover, the Authors wish to thank Prof. Richard Bamler of DLR and Technical University of Munich, Germany, and Dr. Michael Eineder of DLR, Germany, and Dr. Alessandro Ferretti, of Telerilevamento Europa (TRE), Italy, for providing some of the images included in this paper. The Authors wish to thank also Prof. Fabrizio Lombardini of University of Pisa for providing the Capon results relevant to the San Paolo Stadium data set, and Dr. Nicola D’Agostino of the Istituto Nazionale di Geofisica e Vulcanologia for providing the GPS data used in Figure 20.25.

References

- [1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, K.P. Papathanassiou, A tutorial on synthetic aperture radar, *IEEE Geosci. Rem. Sens. Mag.* 1 (1) (2013) 6–43.
- [2] C.A. Wiley, Synthetic aperture radars: a paradigm for technology evolution, *IEEE Trans. Aerosp. Electron. Syst.* 21 (3) (1985) 440–443.
- [3] R. Bamler, P. Hartl, Synthetic aperture radar interferometry, *Inverse Probl.* 14 (1998) R1–R54.
- [4] P.A. Rosen, S. Hensley, I.R. Joughin, F.K. Li, S.N. Madser, E. Rodriguez, R.M. Goldstein, Synthetic aperture radar interferometry, *Proc. IEEE* 88 (3) (2000) 333–382.
- [5] R. Romeiser, H. Breit, M. Eineder, H. Runge, P. Flament, K. de Jong, J. Vogelzang, Current measurements by SAR along-track interferometry from a Space Shuttle, *IEEE Trans. Geosci. Remote Sens.* 43 (10) (1995) 2315–2324.
- [6] ERS. <<https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/ers>>.

- [7] CSK. <http://www.asi.it/en/activity/earth_observation/cosmoskymed>.
- [8] TerraSAR. Available from: <http://www.dlr.de/dlr/en/desktopdefault.aspx/tabcid-10377/565_read-436/>.
- [9] J.S. Lee, E. Pottier, Polarimetric Radar Imaging: From Basics to Applications, CRC Press, 2009.
- [10] S.R. Cloude, E. Pottier, A review of target decomposition theorems in radar polarimetry, *IEEE Trans. Geosci. Remote Sens.* 34 (2) (1996) 498–518.
- [11] S.R. Cloude, E. Pottier, An entropy-based classification scheme for land applications of polarimetric SAR, *IEEE Trans. Geosci. Remote Sens.* 35 (1) (1997) 68–78.
- [12] A. Freeman, S.T. Durden., A three-component scattering model for polarimetric SAR data, *IEEE Trans. Geosci. Remote Sens.* 36 (3) (1998) 963–973.
- [13] R. Touzi, W.M Boerner, J.S Lee, E. Lueneburg, A review of polarimetry in the context of synthetic aperture radar: concepts and information extraction, *Can. J. Remote Sens.* 30 (3) (2004) 380–407.
- [14] T. Le Toan, A. Beaudoin, D. Guyon., Relating forest biomass to SAR data. *IEEE Trans. Geosci. Remote Sens.* 30 (2) (1992) 403–411.
- [15] I. Hajnsek, E. Pottier, and S. Claude., Inversion of surface parameters from polarimetric SAR, *IEEE Trans. Geosci. Remote Sens.* 41 (4) (2003) 727–744.
- [16] I. Hajnsek, T. Jagdhuber, H. Shon, and K. P. Papathanassiou, Potential of estimating soil moisture under vegetation cover by means of PolSAR, *IEEE Trans. Geosci. Remote Sens.* 42 (2) (2009) 442–454.
- [17] S.R. Cloude, K. Papathanassiou, Polarimetric SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 36 (4) (1998) 1551–1565.
- [18] K.P. Papathanassiou, S.R. Cloude, Single-baseline polarimetric SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 39 (6) (2001) 2352–2363.
- [19] D. Reale, G. Fornaro, A. Pauciullo, X. Zhu, R. Bamler, Tomographic imaging and monitoring of buildings with very high resolution SAR data, *IEEE Geosci. Remote Sens. Lett.* 8 (4) (2011) 661–665.
- [20] G. Fornaro, A. Pauciullo, D. Reale, X. Zhu, R. Bamler, SAR tomography: an advanced tool for 4D spaceborne radar scanning with application to imaging and monitoring of cities and single Buildings, *IEEE Geosci. Remote Sens. Newslett.* (2012) 10–18.
- [21] S. Tebaldini, Single and multipolarimetric SAR tomography of forested areas: a parametric approach, *IEEE Trans. Geosci. Remote Sens.* 48 (5) (2010) 2375–2387.
- [22] S. Tebaldini, F. Rocca, Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-Bands, *IEEE Trans. Geosci. Remote Sens.* 50 (1) (2012) 232–246.
- [23] S.R. Cloude, Polarization coherence tomography, *Radio Sci.* 41 (4) (2006) <<http://dx.doi.org/10.1029/2005RSE003436>>.
- [24] S.R. Cloude, Dual-baseline coherence tomography, *IEEE Geosci. Remote Sens. Lett.* 4 (1) (2007) 127–131.
- [25] J.C. Curlander, R.N. McDonough, Synthetic Aperture Radar: Systems and Signal Processing, Wiley-Interscience, 1991.
- [26] C. Elachi, Spaceborne Radar Remote Sensing: Applications and Techniques, IEEE, 1988.
- [27] M.R. Ducoff, B.W. Titjen, Pulse Compression Radar, in: M. Skolnik (Ed.), *Radar Handbook*, third ed., McGraw-Hill Professional, 2008, ISBN 978-0071485470 (Chapter 8).
- [28] I.G. Cumming, F.H. Wong, Digital Processing of Synthetic Aperture Radar Data: Algorithms And Implementation, Artech House Remote Sensing Library, 2005.
- [29] W.G. Carrara, R.M. Majewski, R.S. Goodman, Spotlight Synthetic Aperture Radar: Signal Processing Algorithms, Artech House Remote Sensing Library, 1995.
- [30] L. Mittermayer, A. Moreira, O. Loffeld, Spotlight SAR data processing using the frequency scaling algorithm, *IEEE Trans. Geosci. Remote Sens.* 37 (5) (1999) 2198–2214.
- [31] P. Prats, R. Scheiber, J. Mittermayer, A. Meta, A. Moreira, Processing of sliding spotlight and TOPS SAR data using baseband azimuth scaling, *IEEE Trans. Geosci. Remote Sens.* 48 (2) (2010) 770–780.

- [32] R. Lanari, S. Zoffoli, E. Sansosti, G. Fornaro, F. Serafino, A new approach for hybrid stripmap/spotlight SAR data focusing, *IEE Proc. Radar Sonar Navig.* 148 (2001) 363–372.
- [33] R. Bamler, M. Eineder, ScanSAR processing using standard high precision SAR algorithms, *IEEE Trans. Geosci. Remote Sens.* 34 (1) (1996) 212–218.
- [34] F. De Zan, A. Monti Guarneri, TOPSAR: Terrain observation by progressive scans, *IEEE Trans. Geosci. Remote Sens.* 44 (9) (2006) 2352–2360.
- [35] M. Eineder, N. Adam, R. Bamler, N. Yague-Martinez, H. Breit, Spaceborne spotlight SAR interferometry with terraSAR-X, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1524–1535.
- [36] F.T. Ulaby, R.K. Moore, A. Fung, *Microwave Remote Sensing: Active and Passive, Radar Remote Sensing and Surface Scattering and Emission Theory*, vol. II, Artech House Publishers, 1986.
- [37] G. Franceschetti, G. Fornaro, Synthetic aperture radar interferometry, in: G. Franceschetti, R. Lanari (Eds.), *Synthetic Aperture Radar Processing*, CRC Press, 1999, 167–223.
- [38] A.K. Gabriel, R.M. Goldstein, Crossed orbit interferometry: theory and experimental results from SIR-B, *Int. J. Remote Sens.* 9 (1988) 857–872.
- [39] E. Rodriguez, J.M. Martin, Theory and design of interferometric synthetic aperture radars, *Proc. Inst. Elect. Eng. Part. F* 139 (1992) 147–159.
- [40] A. Ferretti, C. Prati, F. Rocca, Multibaseline DEM reconstruction: the wavelet approach, *IEEE Trans. Geosci. Remote Sens.* 37 (1999) 705–715.
- [41] F. Lombardini, Optimum absolute phase retrieval in a three-element SAR interferometer, *Electron. Lett.* 34 (1998) 1522–1524.
- [42] V. Pascazio, G. Schirinzi, Estimation of terrain elevation by multi-frequency interferometric wide band SAR Data, *IEEE Signal Process. Lett.* 8 (2001) 7–9.
- [43] R.M. Goldstein, H.A. Zebker, Interferometric radar measurements of ocean surface currents, *Nature* (1987) 707–709.
- [44] R. Romeiser, H. Breit, M. Eineder, H. Runge, P. Flament, K. de Jong, J. Vogelzang, Current measurements by SAR along-track interferometry from a space shuttle, *IEEE Trans. Geosci. Remote Sens.* 43 (10) (2005) 2315–2324.
- [45] A. Budillon, V. Pascazio, G. Schirinzi, Estimation of radial velocity of moving targets by along-track interferometric SAR systems, *IEEE Geosci. Rem. Sens. Lett.* 5 (3) (2008) 349–353.
- [46] A. Budillon, V. Pascazio, G. Schirinzi, Multi-channel along-track interferometric SAR systems: moving targets detection and velocity estimation, *Int. J. Navig. Obs.* Q3 (2008) 1–16.
- [47] G. Franceschetti, R. Lanari, V. Pascazio, G. Schirinzi, WASAR: a Wide angle SAR processor, *IEE Proc. F* 139 (2) (1992) 107–114.
- [48] L.C. Graham, Synthetic interferometer for topographic mapping, *Proc. IEEE* 62 (1974) 763–768.
- [49] O. Hellwijk, H. Ebner, Geocoding SAR interferograms by least squares adjustment, *J. Photogramm. Remote Sens.* 55 (4) (2000) 277–288.
- [50] J.J. Van Zyl, The shuttle radar topography mission (SRTM): a breakthrough in remote sensing of topography, *Acta Astronaut.* 48 (5–12) (2001) 559–565.
- [51] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, M. Zink, TanDEM-X: a satellite formation for high-resolution SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 45 (11) (2007) 3317–3341.
- [52] M. Lucido, F. Meglio, V. Pascazio, G. Schirinzi, Closed form evaluation of the second order statistical distribution of the interferometric phases in dual-baseline SAR systems, *IEEE Trans. Signal Proc.* 58 (3) (2010) 1698–1707.
- [53] D. Just, R. Bamler, phase statistics of interferograms with applications to synthetic aperture radar, *App. Opt.* 33 (20) (1994) 4361–4368.

- [54] G. Fornaro, G. Franceschetti, Image registration in interferometric SAR processing, IEE Proc. Radar. Sonar Navig. 142 (1995) 313–320.
- [55] F. Gatelli, Guarneri A. Monti, F. Parizzi, P. Pasquali, C. Prati, F. Rocca, The wave-number shift in SAR interferometry, IEEE Trans. Geosci. Remote Sens. 32 (1994) 855–865.
- [56] G. Ferraiuolo, F. Meglio, V. Pascazio, G. Schirinzi, DEM reconstruction accuracy in multi-channel SAR interferometry, IEEE Trans. Geosci. Rem. Sens. 47 (1) (2009) 191–201.
- [57] S.H.A. Zebker, J. Villasenor, Decorrelation in interferometric radar echoes, IEEE Trans. Geosci. Remote Sens. 30 (1992) 950–959.
- [58] C.W. Chen, H.A. Zebker, Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization, J. Opt. Soc. Am. A 18 (2001) 338–351.
- [59] C.W. Chen, H.A. Zebker, Phase unwrapping for large SAR interferograms: statistical segmentation and generalized network models, IEEE Trans. Geosci. Remote Sens. 40 (8) (2002) 1709–1719.
- [60] M. Costantini, A novel phase unwrapping method based on network programming, IEEE Trans. Geosci. Remote Sens. 36 (3) (1998) 813–821.
- [61] G. Fornaro, G. Franceschetti, R. Lanari, Interferometric SAR phase unwrapping using Green's formulation, IEEE Trans. Geosci. Remote Sens. 34 (3) (1996) 720–727.
- [62] D.C. Ghiglia, L.A. Romero, Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods, J. Opt. Soc. Amer. A 11 (1) (1994) 107–117.
- [63] R.M. Goldstein, H.A. Zebker, C.L. Werner, Satellite radar interferometry: two-dimensional phase unwrapping, Radio Sci. 23 (4) (1998) 713–720.
- [64] J.G. Proakis, M. Salehi, Communication Systems Engineering, second ed., Prentice Hall, 2001.
- [65] W.B. Davenport, W.L. Root, An Introduction to the Theory of Random Signal and Noise, IEEE Communications Society Press, 1987.
- [66] C. Prati, F. Rocca, Improving Slant-Range Resolution with Multiple SAR Surveys, IEEE Trans. Aerosp. Electron. Syst. 29 (1993) 135–44.
- [67] G. Fornaro, Monti A. Guarneri, Minimum mean square error space-varying filtering of interferometric SAR data, IEEE Trans. Geosci. Remote Sens. 40 (2002) 11–21.
- [68] R.M. Goldstein, H. Engelhardt, B. Kamb, R.M. Frolich, Satellite RADAR interferometry for monitoring ice-sheet motion—application to an Antarctic ice stream, Science 262 (5139) (1993) 1525–1530.
- [69] D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, T. Rabaute, The displacement field of the Landers earthquake mapped by radar interferometry, Nature 364 (6433) (1993) 138–142.
- [70] G. Peltzer, P.A. Rosen, Surface displacement of the 17 May 1993 Eureka valley, California, Earthquake observed by SAR interferometry, Science 268 (1995) 1333–1336.
- [71] R. Lanari, P. Lundgren, E. Sansosti, Dynamic deformation of etna volcano observed by satellite radar interferometry, Geophy. Res. Lett. 25 (10) (1998) 1541–1544.
- [72] D. Massonnet, P. Briole, A. Arnaud, Deflation of Mount Etna monitored by spaceborne radar interferometry, Nature 375 (1995) 567–570.
- [73] E. Rignot, Fast recession of a west antarctic glacier, Science 281 (1998) 549–551.
- [74] R.F. Hanssen, Radar Interferometry: Data Interpretation and Error Analysis (Remote Sensing and Digital Image Processing), Springer, 2001.
- [75] R.M. Goldstein, H.A. Zebker, C.L. Werner, Satellite radar interferometry: two dimensional phase unwrapping, Radio Sci. 23 (4) (1998) 713–720.
- [76] D.C. Ghiglia, M.D. Pitt, Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software, Wiley-Interscience, 1998.
- [77] D.C. Ghiglia, L.A. Romero, Minimum L^p -norm two-dimensional phase unwrapping, J. Opt. Soc. Amer. A 13 (1996) 1999–2007.

- [78] T.J. Flynn, Two-dimensional phase unwrapping with minimum weighted discontinuity, *J. Opt. Soc. Am. A* 14 (1997) 2692–2701.
- [79] M. Costantini, P.A. Rosen, A Generalized Phase Unwrapping Approach for Sparse Data, in: Proc. Int. Geosci. Remote Sens. Symp. (IGARSS 1999) Hamburg, Germany, 1999, pp. 267–269.
- [80] R. Bamler, M. Eineder, Accuracy of differential shift estimation by correlation and split-bandwidth interferometry for wideband and delta-k SAR systems, *IEEE Geosci. Remote Sens. Lett.* 2 (2) (2005) 151–155.
- [81] G. Fornaro, A. Pauciullo, E. Sansosti, Phase difference-based multichannel phase unwrapping, *IEEE Trans. Image Proc.* 14 (2005) 960–972.
- [82] G. Fornaro, A. Monti Guarnieri, A. Pauciullo, F. De-Zan, Maximum likelihood multi-baseline SAR interferometry, *IEE Proc. Radar Sonar Navigat.* 153 (2006) 279–288.
- [83] V. Pascazio, G. Schirinzi, Multi-frequency InSAR height reconstruction through maximum likelihood estimation of local planes parameters, *IEEE Trans. Image Process.* 11 (2002) 1478–1489.
- [84] W. Xu, E. Chang, L. Kwoh, H. Lim, W. Cheng, Phase-unwrapping of SAR interferogram with multi-frequency or multi-baseline, in: Proc. Int. Geosci. Remote Sens. Symp. (IGARSS 1994), Pasadena, USA, 1994, pp. 730–732.
- [85] G. Ferraioli, A. Shabou, F. Tupin, V. Pascazio, Multichannel phase unwrapping with graph-cuts, *IEEE Geosci. Remote Sens. Lett.* 6 (3) (2009) 562–566.
- [86] J. Dias, J. Leitao, The $z\pi m$ algorithm: a method for interferometric image reconstruction in SAR/SAS, *IEEE Trans. Image Process.* 11 (2002) 408–422.
- [87] G. Ferraiuolo, V. Pascazio, G. Schirinzi, Maximum a posteriori estimation of Height profiles in InSAR imaging, *IEEE Geosci. Remote Sens. Lett.* 1 (2004) 66–70.
- [88] G. Nico, G. Palubinskas, M. Datcu, Bayesian approach to phase unwrapping: theoretical study, *IEEE Trans. Signal Process.* 48 (2000) 2545–2556.
- [89] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1993.
- [90] S.Z. Li, Markov Random Field Modelling in Computer Vision, Springer, Computer Science Workbench, 2001.
- [91] S. Geman, D. Geman, Stochastic relaxation, Gibbs' distributions, and Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (1984) 721–741.
- [92] A. Shabou, F. Baselice, G. Ferraioli, Urban digital elevation model reconstruction using very high resolution multi-channel InSAR data, *IEEE Trans. Geosci. Remote Sens.* 50 (2012) 4748–4758.
- [93] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1–4) (1992) 259–268.
- [94] H. Ishikawa, Exact optimization for markov random fields with convex priors, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (2003) 1333–1336.
- [95] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (11) (2001) 1222–1239.
- [96] V. Kolmogorov, R. Zabih, Multi-camera scene reconstruction via graph cuts, in: Proceedings of the 7th European Conference on Computer Vision, 3, 2002, 82–96.
- [97] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *Energy Minimization Methods in Computer Vision and Pattern Recognition* (2001) 359–374.
- [98] J. Castellanos, S. Gomez, V. Guerra, The triangle method for finding the corner of the L-curve, *Appl. Numer. Math.* 43 (4) (2002) 359–373.
- [99] A. Ferretti, C. Prati, F. Rocca, Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 38 (2000) 2202–2212.
- [100] A. Ferretti, C. Prati, F. Rocca, Permanent scatterers in SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 39 (1) (2001) 8–20.

- [101] M. Crosetto, E. Biescas, J. Duro, J. Closa, A. Arnaud, Generation of advanced ERS and envisat interferometric SAR products using the stable point network technique, *Photogramm. Eng.* 74 (4) (2008) 443–451.
- [102] N. Adam, B. Kampes, M. Eineder, J. Worawattanamateekul, M. Kircher, in: *The Development of a Scientific Permanent Scatterer System*, ISPRS Hannover Workshop, Hannover, Germany, 2003.
- [103] C. Werner, U. Wegmüller, T. Strozzi, A. Wiesmann, Interferometric point target analysis for deformation mapping, in: *Proc. Int. Geosci. Remote Sens. Symp.*, (IGARSS 2003), Toulouse, France, 2003.
- [104] F. Bovenga, R. Nutricato, A. Refice, J. Wasowski, Application of multi-temporal differential interferometry to slope instability detection in Urban/Peri-Urban areas, *Eng. Geol.* 88 (2006) 219–240.
- [105] M. Costantini, S. Falco, F. Malvarosa, F. Minati, F. Trillo (2009), Method of persistent scatterers pairs (PSP) and high resolution SAR interferometry, in: *Proc. Int. Geosci. Remote Sensing Symp.* (IGARSS 2009) Cape Town, South Africa, 2009.
- [106] A. Hooper, P. Segall, H. Zebker, Persistent scatterer InSAR for crustal deformation analysis with application to volcán Alcedo, Galápagos, *J. Geophys. Res.* 112 (2007) B07407.
- [107] G.E. Hilley, R. Burgmann, A. Ferretti, F. Novali, F. Rocca, Dynamics of Slow-Moving Landslides from Permanent Scatterer Analysis, *Science* 304 (5679) (2004) 1952–1955.
- [108] G. Herrera, J.C. Davalillo, J. Mulas, G. Cooksley, O. Monserrat, V. Pancioli, Mapping and monitoring geomorphological processes in mountainous areas using PSI data: central pyrenees case study, *Nat. Hazard Earth Syst. Sci.* 9 (2009) 1587–1598.
- [109] P. Berardino, G. Fornaro, R. Lanari, E. Sansosti, A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms, *IEEE Trans. Geosci. Remote Sens.* 40 (11) (2002) 2375–2383.
- [110] P. Blanco-Sánchez, J. Mallorqui, S. Duque, D. Monnells, The Coherent pixels technique (CPT): an advanced DInSAR technique for nonlinear deformation monitoring, *Pure Appl. Geophys.* 165 (6) (2008) 1167–1193.
- [111] O. Mora, J. Mallorqui, A. Broquetas, Linear and non-linear terrain deformation maps from a reduced set of interferometric SAR images, *IEEE Trans. Geosci. Remote Sens.* 41 (10) (2003) 2243–2253.
- [112] G.H. Golub, C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.
- [113] A. Refice, F. Bovenga, R. Nutricato, MST-based stepwise connection strategies for multipass Radar data, with application to coregistration and equalization, *IEEE Trans. Geosci. Remote Sens.* 44 (8) (2006) 2029–2040.
- [114] G. Fornaro, D. Reale, F. Serafino, Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers, *IEEE Trans. Geosci. Remote Sens.* 47 (1) (2009) 224–237.
- [115] G. Fornaro, A. Pauciullo, F. Serafino, Deformation monitoring over large areas with multipass differential sar interferometry: a new approach based on the use of spatial differences, *Int. J. Remote Sens.* 30 (6) (2009) 1455–1478.
- [116] N. D'Agostino, D. Cheloni, G. Fornaro, R. Giuliani, D. Reale, Space-time distribution of afterslip following the 2009 L'Aquila earthquake, *J. Geophys. Res.* 117 (2012) B02402, doi: <<http://dx.doi.org/10.1029/2011JB008523>>.
- [117] S. Gernhardt, R. Bamler, Deformation monitoring of single buildings using meter-resolution SAR data in PSI, *ISPRS J. Photogramm. Remote Sens.* (2012) (in print).
- [118] R. Touzi, A. Lopes, J. Bruniquel, P.W. Vachon, Coherence estimation for SAR imagery, *IEEE Trans. Geosci. Remote Sens.* 37 (1999) 135–149.
- [119] G. Fornaro, A. Pauciullo, D. Reale, A null-space method for the phase unwrapping of multi-temporal sar interferometric stacks, *IEEE Trans. Geosci. Remote Sens.* 49 (6) (2011) 2323–2334.
- [120] P. Agram, H. Zebker, Edgelist phase unwrapping algorithm for time-series InSAR analysis, *J. Opt. Soc. Am. A* 27 (3) (2010) 605–612.
- [121] M. Costantini, F. Malvarosa, F. Minati, A general formulation for robust and efficient integration of finite differences and phase unwrapping on sparse multidimensional domains, in: *ESA Fringe 2009 Workshop*, Frascati, Italy, 2009.

- [122] A. Pepe, L. Lanari, On the extension of the minimum cost flow algorithm for phase unwrapping of multi-temporal differential SAR interferograms, *IEEE Trans. Geosci. Remote Sens.* 44 (9) (2006) 2374–2383.
- [123] A. Budillon, A. Evangelista, G. Schirinzi, Three-dimensional SAR focusing from multipass signals using compressive sampling, *IEEE Trans. Geosci. Remote Sens.* 49(1) (2011) 488–499.
- [124] G. Fornaro, F. Lombardini, F. Serafino, Three-dimensional multipass SAR focusing: experiments with long-term spaceborne data, *IEEE Trans. Geosci. Remote Sens.* 43 (4) (2005) 702–714.
- [125] A. Reigber, A. Moreira, First demonstration of airborne SAR tomography using multibaseline L-band data, *IEEE Trans. Geosci. Remote Sens.* 38 (5) (2000) 2142–2152.
- [126] A. Budillon, A. Evangelista, G. Schirinzi, SAR Tomography from Sparse Samples, in: Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS 2009), 2009, pp. IV-865–IV-868.
- [127] F. Gini, F. Lombardini, Multibaseline cross-track SAR interferometry: a signal processing perspective, *IEEE Aerosp. Electron. Sys. Mag.* 20 (8) (2005) 71–93.
- [128] F. Lombardini, M. Montanari, F. Gini, Reflectivity estimation for multibaseline interferometric radar imaging of layover extended sources, *IEEE Trans. Signal Process.* 51 (2003) 1508–1519.
- [129] X.X. Zhu, R. Bamler, Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR, *IEEE Trans. Geosci. Remote Sens.* 50 (1) (2012) 247–258.
- [130] E.J. Candes, M.B. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.* 25 (2) (2008) 21–30.
- [131] X.X. Zhu, R. Bamler, Demonstration of super-resolution for tomographic SAR imaging in urban environment, *IEEE Trans. Geosci. Remote Sens.* (2012), <<http://dx.doi.org/10.1109/TGRS.2011.2177843>>.
- [132] E.J. Candes, T. Tao, The Dantzig selector: statistical estimation when p is much larger than n , *Ann. Statist.* 35 (6) (2007) 2313–2351.
- [133] A. Budillon, G. Schirinzi, Artifact Reduction in SAR Compressive Sampling Tomography, in: Proc. Int. Geosci. Remote Sens. Symp. 2011 (IGARSS 2011) (2011) 2700–2703.
- [134] F. Lombardini, Differential tomography: a new framework for SAR interferometry, *IEEE Trans. Geosci. Remote Sens.* 43 (1) (2005) 37–44.
- [135] A.B. Carlson, P.B. Crilly, *Communication Systems*, fifth ed., Mc Graw Hill Higher Education, 2009.
- [136] S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice Hall, 1998.
- [137] A. De Maio, G. Fornaro, A. Pauciullo, Detection of single scatterers in multidimensional SAR imaging, *IEEE Trans. Geosci. Remote Sens.* 47 (7) (2009) 2284–2297.
- [138] A. Pauciullo, D. Reale, A. De Maio, G. Fornaro, Detection of double scatterers in SAR tomography, *IEEE Trans. Geosci. Remote Sens.* 50 (9) (2012) 3567–3586.
- [139] X. Zhu, R. Bamler, Very high resolution spaceborne SAR tomography in urban environment, *IEEE Trans. Geosci. Remote Sens.* 48 (12) (2010) 4296–4308.
- [140] F. Lombardini, M. Pardini, Multiple scatterers identification in complex scenarios with adaptive differential tomography, in: Proceedings of the International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, July 2009.
- [141] F. Lombardini, F. Gini, Model order selection in multibaseline interferometric radar systems, *EURASIP J. Adv. Signal Proc.* 2005 (20) (2005), 3206–3219.
- [142] G. Fornaro, F. Serafino, D. Reale, 4-D SAR imaging: the case study of Rome, *IEEE Geosci. Remote Sens. Lett.* 7 (2010) 236–240.
- [143] S.R. Cloude, Dual-baseline Coherence Tomography, *IEEE Geosci. Remote Sens. Lett.* 4 (1) (2007) 127–131.

Radar Polarimetry Basics and Selected Earth Remote Sensing Applications

21

Laurent Ferro-Famil and Eric Pottier

IETR, University of Rennes 1, Rennes, France

2.21.1 Introduction

Polarimetry (Polar: polarization, Metry: measure) is the science of acquiring, processing, and analyzing the polarization state of an electromagnetic field [1]. SAR polarimetry is concerned with the utilization of polarimetry in radar remote sensing applications, whose objectives are to remotely detect and characterize objects and environments. Due to its wide coverage, high resolution, and very low sensitivity to weather conditions, Synthetic Aperture Radar (SAR) is by itself a unique tool for studying large areas in a very short amount of time. The combination of SAR with polarimetric diversity has shown to greatly improve the characterization capabilities with respect to scalar systems. Polarization permits to increase the observation space but most of all, to relate the amplitude and phase of reflected signal to the electromagnetic properties, shape, orientation, structure ... of observed objects, i.e., polarimetry offers a high potential for estimating physical characteristics of objects and environments.

2.21.1.1 Historical overview of radar polarimetry

Studies on wave polarization were first led in optics and adapted to radar during the 1940s. During three decades, scientists like Sinclair [2,3], Kennaugh [4] Deschamps [5] ... developed theoretical tools needed to handle coherent vector signals and techniques that used polarization as a way to discriminate radar returns or to identify objects according to specific polarimetric behaviors. During the 1970s Huynen [6] deeply modified the way to process polarimetric quantities, by combining rigorous techniques, based on mathematical and signal processing concepts, with phenomenology, which considers polarimetric information as a whole, rather than a sum of contributions. He identified structural aspects of polarimetric returns, that could be related to basic properties of observed objects, like their shapes, orientations ... Some parts of this work were used and extended during the 1980s by Boerner [1] in the frame of electromagnetic inverse problems, and by Cloude [7] who was preparing a new formalism that would, once again, strongly increase the possibilities of analysis of polarimetric measurements. The first airborne AIRSAR polarimetric SAR campaign, led in 1988 by the JPL highly promoted SAR polarimetry.

The 1990s were an auspicious decade for SAR polarimetry, and modern pioneers, like Freeman [8,9], Van Zyl [10] and their colleagues from JPL, Lee [11] and his research team, from the ONR/NRL, Cloude and Pottier [12,13], the DLR HR department in Germany, Yamaguchi [14] in Japan and many others produced and stimulated a large amount of high quality scientific work. During this period,

clear and rigorous polarimetric formalism and operators were defined under the supervision of Boerner, original polarimetric decomposition schemes and interpretations were introduced, specific statistical techniques were adapted from the signal processing community to develop specific detection, estimation and classification techniques. Several polarimetric airborne sensors were developed and occasional space shuttle missions were flown, that provided data sets over which scientists could compare their approaches.

At the end of the 1990s, Cloude and Papathassiou investigated a new diversity mode combining polarimetry with interferometry [15] that provide information about the elevation of a scatterer. They found out that polarimetric SAR interferometry (PolInSAR) was particularly well adapted to the characterization of volumetric environments, like forest or other kinds of vegetated environments. This discovery opened a new era for polarimetric remote sensing, since when they are used separately polarimetry and interferometry might not be well adapted for such an application, depending on density properties of the observed medium. Most of the developments in SAR polarimetry performed during the 1990s were oriented towards a qualitative characterization of polarimetric scattering, i.e., the main objectives concerned the understanding and accurate description of scattering rather than its utilization for estimating physical quantities. Due to a much easier access to polarimetric SAR data, current challenges involve more systematic approaches whose objectives concern the quantitative estimation of physical parameters.

2.21.1.2 Organization of the work

The main objective of this chapter is to provide a brief and, unfortunately, restrictive panorama of some characteristic techniques and applications of modern SAR polarimetry. This chapter is organized into three main parts introducing first basic definitions, representations and processing concepts of radar polarimetry. The second section is dedicated to more advanced processing techniques oriented towards specific applications like volumetric media characterization, soil moisture estimation ... The last part of the chapter shows how polarimetry can be combined with other kinds of signal diversity, like Time-Frequency, interferometry, in order to significantly increase the amount of information as well as the possibilities of characterization and estimation of complex scattering phenomena.

The first part presents some basics of SAR polarimetry, i.e., a summary of the work done by our colleagues mentioned above. It is intended for beginners in polarimetry or scientists who wish to clarify some specific principles, notations or basic concepts in radar polarimetry. The polarization of a wave is first described in terms of vector representations. From the general expression of a plane wave electric field is presented the concept of polarization ellipse which perfectly describes the polarimetric properties of a wave. It is then shown that a two-element Jones vector can be defined to parametrize this ellipse and that simple Special Unitary operators can be used to represent a wave in any arbitrary polarization basis. Once a wave is acquired in a polarization basis, i.e., with a given set of receiving antennas, mathematical tools may be used to compute the signal that would be delivered by any arbitrary set of receiving antennas, i.e., without actually measuring the signal for each arbitrary set. It is then shown that for given acquisition conditions, the polarimetric properties of a scatterer can be represented under the form of a (2×2) coherent scattering matrix. The concept of change of polarization basis is extended to this kind of representation and the structure of this matrix is given for different kinds of canonical scatterers, i.e., helix, plane, dipole ... Over distributed environments, Synthetic Aperture Radar (SAR) images are

known to be affected by the speckle effect. Similarly to the scalar case, speckled polarimetric responses cannot be reliably characterized using single-look representations, i.e., from the coherent scattering matrix. Incoherent second order representations, i.e., specific polarimetric covariance matrices, are then introduced and characterized. It is then shown that polarimetric decomposition techniques are needed to provide an interpretation of the polarimetric scattering response of a medium from second order quantities. Such decomposition approaches may be based on scattering models or use specific mathematical properties, in order to extract and characterize canonical scattering behaviors from a distributed averaged representation. The necessity and usefulness, as well as the domain of application, of several decomposition techniques are discussed. Finally multivariate statistics, adapted to polarimetric quantities are introduced and applied to the classification and segmentation of polarimetric SAR images.

The second section presents advanced decomposition techniques and their application to the estimation of physical quantities. The proposed approaches are all directly based on the basics introduced in the first part that are improved or adapted to specific situations using physical interpretation and simple mathematics. Different model based techniques are introduced that aim to characterize a volumetric medium lying over a natural soil. Several hypotheses or levels of adaptivity lead to different interpretations and solutions. It is also shown that polarimetry may be efficiently used to roughly estimate the azimuthal topography of a scene or to introduce physical considerations into a statistical segmentation process. Finally, early studies on soil moisture and roughness estimation from specific polarimetric indicators are presented. This part illustrates the fact that SAR polarimetry is an interactive multi-disciplinary subject open to a wide range of scientists and engineers wishing to investigate physical phenomena using interesting mathematical tools.

The last part presents selected topics in multidimensional polarimetric SAR signal processing. The dimensionality of polarimetric signal being rather low, the use of other kinds of diversity is often required to solve an estimation problem. This part shows how three different applications, Time-Frequency characterization of complex environments, Pol-inSAR characterization of forested areas and polarimetric SAR tomography, i.e., polarimetric 3-D imaging, rely on the same physical considerations. All these techniques are based on the modeling of the correlation properties of polarimetric signals acquired from different positions, which may vary with the position at which is performed the acquisition, the considered polarization channel and the nature of the observed object or medium.

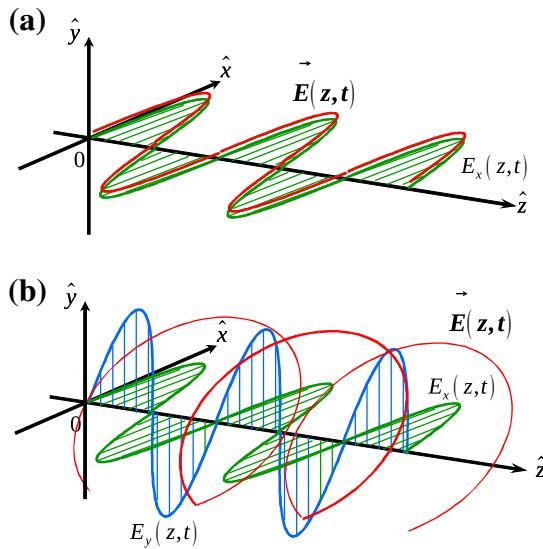
2.21.2 SAR polarimetry basics

2.21.2.1 Polarimetric scattering: formalism, descriptors, and operators

2.21.2.1.1 Polarization of a monochromatic plane wave

2.21.2.1.1.1 Polarization ellipse

The polarization of a wave is studied here in the particular case of constant amplitude monochromatic plane waves [16] propagating in a homogeneous, loss less medium, free of charges, with a celerity c . This wave is perfectly defined by its electric field, $\mathbf{E}(\mathbf{r}, t)$, where \mathbf{r} and t represent spatial and time coordinates, respectively. Without any loss of generality, the electric field may be represented in an orthogonal basis $(\hat{x}, \hat{y}, \hat{z})$ defined so that the direction of propagation is aligned with one of the basis

**FIGURE 21.1**

Spatial trajectories of (a) horizontally and (b) circularly polarized planes waves.

axes, $\hat{k} = \hat{z}$, with $\hat{k} = \mathbf{k}/\|\mathbf{k}\|$ and \mathbf{k} the wave vector. The expression of the electric is given by

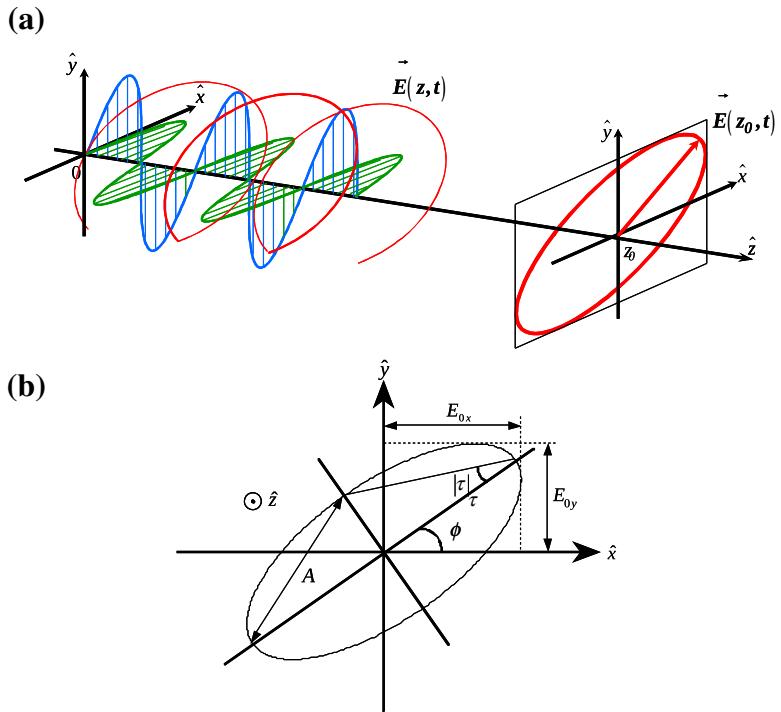
$$\mathbf{E}(z, t) = \begin{bmatrix} E_x(z, t) \\ E_y(z, t) \\ 0 \end{bmatrix} = \begin{bmatrix} E_{0x} \cos(\omega t - kz + \delta_x) \\ E_{0y} \cos(\omega t - kz + \delta_y) \\ 0 \end{bmatrix}. \quad (21.1)$$

Due to the planar shape of the wave, the electric field is composed of two orthogonal sinusoidal signals, oriented along directions \hat{x} and \hat{y} , that are orthogonal to the direction of propagation \hat{z} . The term $\omega t - kz$, related to the space-time duality of propagating waves being common to both components, the polarimetric properties of $\mathbf{E}(z, t)$ are defined by the amplitude and phase of each component, E_{0x} , E_{0y} , and δ_x , δ_y [17]. As shown in Figure 21.1, for a fixed temporal coordinate, t_0 , the spatial evolution of a plane monochromatic wave follows a helical trajectory along the z axis, whose characteristics depend on the amplitude and phase terms of (21.1).

Three-dimensional helical curves being difficult to represent and to analyze, one generally makes use of the space-time duality in (21.1), i.e., shifts in time and space are equivalent, to characterize the polarization of the wave in the time domain, at a fixed position, $z = z_0$ [17]. As time evolves, the wave propagates through the equi-phase plane, orthogonal to the direction of propagation and located at coordinate z_0 and describes a characteristic elliptical locus as shown in Figure 21.2.

The nature of the wave temporal trajectory may be determined from the following parametric relation between the components of $\mathbf{E}(z_0, t)$:

$$\frac{E_x^2(z_0, t)}{E_{0x}^2} - 2 \frac{E_x(z_0, t)E_y(z_0, t)}{E_{0x}E_{0y}} \cos(\delta_y - \delta_x) + \frac{E_y^2(z_0, t)}{E_{0y}^2} = \sin(\delta_y - \delta_x). \quad (21.2)$$

**FIGURE 21.2**

(a) propagation of the electric field through an orthogonal plane at $z = z_0$ (b) parametrized polarization ellipse.

The expression in (21.2) is the equation of an ellipse, called the polarization ellipse that describes the wave polarization, and may be characterized using three parameters as shown in Figure 21.2.

- A is called the ellipse amplitude and is determined from the ellipse axis as:

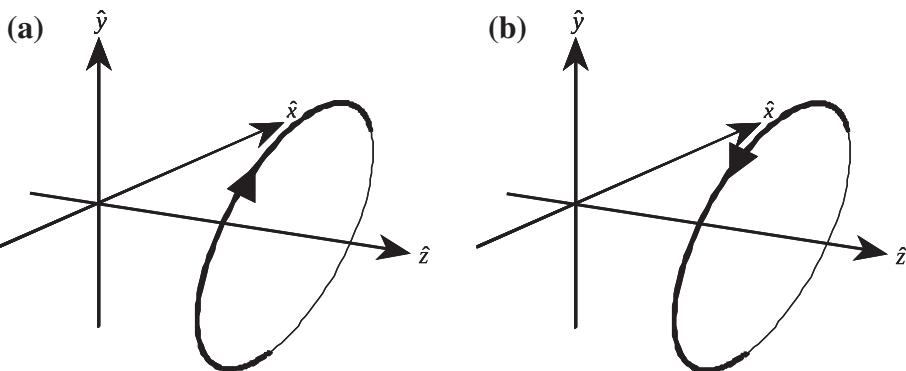
$$A = \sqrt{E_{0x}^2 + E_{0y}^2}. \quad (21.3)$$

- ϕ is the ellipse orientation and is defined as the angle between the ellipse major axis and \hat{x} :

$$\tan(2\phi) = 2 \frac{E_{0x} E_{0y}}{E_{0x}^2 + E_{0y}^2} |\cos(\delta_y - \delta_x)|, \quad \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \quad (21.4)$$

- τ the ellipse aperture, also called ellipticity, defined as:

$$|\sin 2\tau| = 2 \frac{E_{0x} E_{0y}}{E_{0x}^2 + E_{0y}^2} |\sin(\delta_y - \delta_x)|, \quad |\tau| \in \left[0, \frac{\pi}{4}\right]. \quad (21.5)$$

**FIGURE 21.3**

Definition of the sense of rotation along the polarization ellipse (a) left handed rotation (b) right handed rotation.

As time goes, the electric field vector $\mathbf{E}(z_0, t)$ rotates in the (x, y) plane to describe the polarization ellipse. By convention, the sense of rotation is determined while looking in the direction of propagation is associated to the sign of τ . A right hand rotation corresponds to $\frac{\partial \xi(t)}{\partial t} > 0 \Rightarrow \tau < 0$ whereas a left hand rotation is characterized by $\frac{\partial \xi(t)}{\partial t} < 0, \Rightarrow \tau < 0$. Figure 21.3 provides a graphical description of the rotation sense convention [16, 18].

Three types of polarizations can be specified:

- *Linear polarization:* $\delta = \delta_y - \delta_x = 0$.

The electric field is a sine wave lying on a plane oriented with an angle ϕ with respect to the \hat{x} axis, with:

$$\mathbf{E}(z_0, t) = \sqrt{E_{0x}^2 + E_{0y}^2} \begin{bmatrix} \cos \phi \\ \sin \phi \\ 0 \end{bmatrix} \cos(\omega t_0 - kz + \delta_x)$$

and the polarization ellipse reduces to a line segment.

- *Circular polarization:* $\delta = \delta_y - \delta_x = \frac{\pi}{2} + k\pi$ and $E_{0x} = E_{0y}$.

In this case, the wave rotates circularly around the \hat{z} axis as shown in Figure 21.1, and has a constant modulus $|\mathbf{E}(z, t)|^2 = E_{0x}^2 + E_{0y}^2$. The polarization ellipse is then a circle.

- *Elliptic polarization:* In this case, wave describes a helical trajectory around the \hat{z} axis.

The polarimetric quantities introduced in this section have been defined under a specific and arbitrary convention, that may be associated to the way a fixed antenna measures a field: the wave is observed at a fixed position, z_0 , as time goes. One could have chosen another approach: observing the wave at a fixed time t_0 , for a varying position along the \hat{z} axis. Such an approach would have led to different expressions for the polarimetric parameters, since derivatives of the phase term in (21.1) with respect to z and t have opposed signs. This fact is taken into account when defining the change of polarimetric

Table 21.1 Unitary Jones Vectors and Associated Polarization Ellipse Parameters for Some Canonical Polarization States

Polarization State	$\mathbf{u}(x,y)$	ϕ	τ
Horizontal (H)	$\mathbf{u}_H = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$	0	0
Vertical (V)	$\mathbf{u}_V = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\pi/2$	0
Linear +45	$\mathbf{u}_{+45} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\pi/4$	0
Linear -45	$\mathbf{u}_{-45} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$-\pi/4$	0
Left circular	$\mathbf{u}_L = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ j \end{bmatrix}$	$[-\pi/2, \dots, \pi/2]$	$\pi/4$
Right circular	$\mathbf{u}_R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -j \end{bmatrix}$	$[-\pi/2, \dots, \pi/2]$	$-\pi/4$

basis for a scattering matrix (21.32). Readers interested by this aspect are invited to look at previous reference works [1, 16, 17, 19].

2.21.2.1.1.2 Jones vector

In the case of a constant amplitude monochromatic plane wave, the space-time expression of the electric field given in (21.1) is redundant and may be simplified using complex notations. The Jones vector, \mathbf{E} , which completely describes the wave polarization is obtained from the complex envelope of the field as:

$$\mathbf{E} = \mathbf{E}(z)|_{z=0} = \begin{bmatrix} E_{0_x} e^{j\delta_x} \\ E_{0_y} e^{j\delta_y} \end{bmatrix}, \quad (21.6)$$

where $\mathbf{E}(z)$ is the time-independent complex envelope of the electric field, and is obtained from its real expression as $\mathbf{E}(z, t) = \Re(\mathbf{E}(z)e^{j\omega t})$. The definitions of a polarization state from the polarization ellipse descriptors or from a Jones vector are equivalent. A Jones vector can be formulated as a two-dimensional complex vector function of the polarization ellipse characteristics as follows [16]:

$$\mathbf{E} = A e^{j\alpha} \begin{bmatrix} \cos \phi \cos \tau - j \sin \phi \sin \tau \\ \sin \phi \cos \tau + j \cos \phi \sin \tau \end{bmatrix} = A e^{j\alpha} \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \tau \\ j \sin \tau \end{bmatrix}, \quad (21.7)$$

where α is an absolute phase term.

Some examples of Jones vectors and polarization ellipse parameters are given in Table 21.1.

Jones vectors may also be written using particular matrix operators belonging to the group of special unitary (2×2) matrices, $SU(2)$. Such a group is called a Lie group and is generated through the matrix exponentiation of a, generally non unique, set of matrix generators. Readers interested by this topic are referred to [20]. The definition of the generators using the Pauli matrix leads to three $SU(2)$ operators, that may be viewed as complex rotation matrices, given by [7]:

$$\mathbf{U}_2(\phi) = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}, \quad \mathbf{U}_2(\tau) = \begin{bmatrix} \cos \tau & j \sin \tau \\ j \sin \tau & \cos \tau \end{bmatrix}, \quad \mathbf{U}_2(\alpha) = \begin{bmatrix} e^{+j\alpha} & 0 \\ 0 & e^{-j\alpha} \end{bmatrix}. \quad (21.8)$$

As special unitary matrices, these operators verify $\mathbf{U}_2^{-1} = \mathbf{U}_2^\dagger$ and $|\mathbf{U}_2| = 1$ where \dagger stands for the transpose conjugate operator and $|\mathbf{U}|$ is the determinant of \mathbf{U} . From the definition given in (21.7), a Jones vector $\mathbf{E}_{(x,y)}$, describing a general elliptical polarization state in the Cartesian basis, may be expressed using $SU(2)$ operators as [1]:

$$\begin{aligned} \mathbf{E}_{(x,y)} &= A \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \tau & j \sin \tau \\ j \sin \tau & \cos \tau \end{bmatrix} \begin{bmatrix} e^{+j\alpha} & 0 \\ 0 & e^{-j\alpha} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= A \mathbf{U}_2(\phi) \mathbf{U}_2(\tau) \mathbf{U}_2(\alpha) \mathbf{x} \\ &= A \mathbf{U}_2(\phi, \tau, \alpha) \mathbf{x}, \end{aligned} \quad (21.9)$$

where $\mathbf{x} = \mathbf{u}_H$ corresponds to the unit Jones vector associated to the horizontal polarization state.

2.21.2.1.1.3 Change of polarization basis

One of the main advantages of radar polarimetry, linked to the vector nature of polarization, resides in the fact that a target response acquired in a polarization basis, may be expressed in any basis using a simple mathematical transformation, without any additional measurements. An orthonormal polarization basis is composed of two unitary norm Jones vectors, \mathbf{u} and \mathbf{u}_\perp , whose hermitian scalar product is null:

$$\mathbf{u}^\dagger \mathbf{u}_\perp = 0 \quad \text{with} \quad \mathbf{u}^\dagger \mathbf{u} = \mathbf{u}_\perp^\dagger \mathbf{u}_\perp = 1. \quad (21.10)$$

For a given Jones vector, \mathbf{u} , there exist an infinite number of orthogonal unitary vectors satisfying (21.10). The $SU(2)$ formalism introduced in the former paragraph may then be used to pick a unique solution. Both vectors may be defined from their ellipse parameters in the (\hat{x}, \hat{y}) basis:

$$\mathbf{u} = \mathbf{U}_2(\phi, \tau, \alpha) \mathbf{x} \quad \text{and} \quad \mathbf{u}_\perp = \mathbf{U}_2(\phi_\perp, \tau_\perp, \alpha_\perp) \mathbf{x}. \quad (21.11)$$

The orthogonality is then introduced by rewriting \mathbf{u}_\perp as a function of $\mathbf{y} = [0, 1]^T$, the second element of the Cartesian basis:

$$\mathbf{u}_\perp = \mathbf{U}_2(\phi_\perp, \tau_\perp, \alpha_\perp) \mathbf{x} = \mathbf{U}_2(\phi, \tau, \alpha) \mathbf{y}. \quad (21.12)$$

It is thus straightforward to remark that the orthogonality condition implies that \mathbf{u} and \mathbf{u}_\perp are associated with ellipse parameters satisfying:

$$\phi_\perp = \phi + \frac{\pi}{2}, \quad \tau_\perp = -\tau, \quad \alpha_\perp = -\alpha. \quad (21.13)$$

A Jones vector, $\mathbf{E}_{(\mathbf{x},\mathbf{y})} = E_x \mathbf{x} + E_y \mathbf{y}$ expressed in the Cartesian (\mathbf{x}, \mathbf{y}) basis, transforms to $\mathbf{E}_{(\mathbf{u},\mathbf{u}_\perp)} = E_u \mathbf{u} + E_{u_\perp} \mathbf{u}_\perp$ in the orthonormal $(\mathbf{u}, \mathbf{u}_\perp)$ polarimetric basis, by the way of a $SU(2)$ transformation. The coordinates E_u and E_{u_\perp} can be determined according to the following expression

$$\begin{aligned}\mathbf{E}_{(\mathbf{u},\mathbf{u}_\perp)} &= E_u \mathbf{u} + E_{u_\perp} \mathbf{u}_\perp \\ \Rightarrow \mathbf{E}_{(\mathbf{x},\mathbf{y})} &= E_u \mathbf{U}_2(\phi, \tau, \alpha) \mathbf{x} + E_{u_\perp} \mathbf{U}_2(\phi, \tau, \alpha) \mathbf{y} = E_x \mathbf{x} + E_y \mathbf{y}.\end{aligned}\quad (21.14)$$

It then follows:

$$\begin{bmatrix} E_u \\ E_{u_\perp} \end{bmatrix} = \mathbf{U}_2(\phi, \tau, \alpha)^{-1} \begin{bmatrix} E_x \\ E_y \end{bmatrix} \Rightarrow \mathbf{E}_{(\mathbf{u},\mathbf{u}_\perp)} = \mathbf{U}_{2(\mathbf{x},\mathbf{y}) \rightarrow (\mathbf{u},\mathbf{u}_\perp)} \mathbf{E}_{(\mathbf{x},\mathbf{y})} \quad (21.15)$$

with $\mathbf{U}_{2(\mathbf{x},\mathbf{y}) \rightarrow (\mathbf{u},\mathbf{u}_\perp)} = \mathbf{U}_2(\phi, \tau, \alpha)^{-1} = \mathbf{U}_2(-\alpha) \mathbf{U}_2(-\tau) \mathbf{U}_2(-\phi)$.

To summarize, the special unitary $SU(2)$ matrix, corresponding to any elliptical basis change, is defined by:

$$\mathbf{U}_2(\phi, \tau, \alpha) = \mathbf{U}_2(\alpha) \mathbf{U}_2(\tau) \mathbf{U}_2(\phi) = \frac{1}{\sqrt{1+|\rho|^2}} \begin{bmatrix} 1 & -\rho^* \\ \rho & 1 \end{bmatrix} \begin{bmatrix} e^{+j\xi} & 0 \\ 0 & e^{-j\xi} \end{bmatrix}, \quad (21.16)$$

where α, ϕ, τ correspond to the three geometric parameters of the polarization ellipse described by the first or principal Jones vector of the new basis. This special unitary $SU(2)$ basis change matrix can also be described using the parameters ρ and ξ which correspond to the polarization ratio of the first or principal Jones vector of the new basis and are given by [16]:

$$\rho = \frac{\tan \phi + j \tan \tau}{1 - j \tan \phi \tan \tau} \quad \text{and} \quad \xi = \alpha - a \tan(\tan \phi \tan \tau). \quad (21.17)$$

2.21.2.1.2 Polarimetric scattering descriptors

2.21.2.1.2.1 Radar measurements with full polarimetric diversity

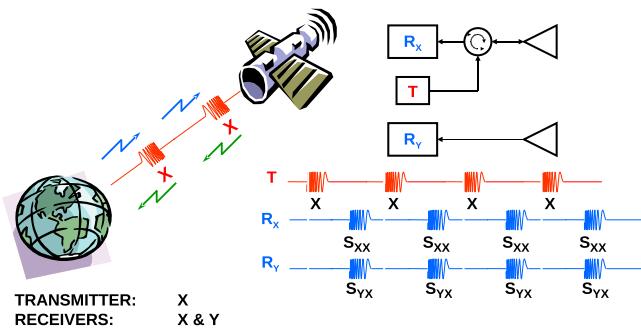
A synopsis of radar acquisitions operated with a polarization diversity on receive only is shown in Figure 21.4. An electric field with a given polarization state, $\mathbf{E}_T \propto \mathbf{u}$ is transmitted and the received signal is measured over both elements of the polarization basis $(\mathbf{u}, \mathbf{u}_\perp)$:

$$\mathbf{E}_T \propto \mathbf{u} \rightarrow \mathbf{E}_R = E_{\mathbf{u}\mathbf{u}} \mathbf{u} + E_{\mathbf{u}_\perp\mathbf{u}_\perp} \mathbf{u}_\perp. \quad (21.18)$$

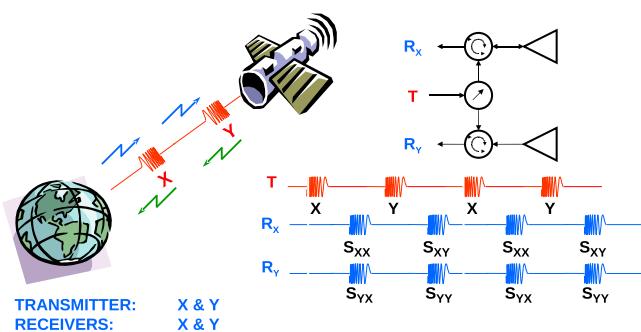
The response of the object may then be represented in any arbitrary polarization basis at the reception using $SU(2)$ transformations introduced earlier. This kind of acquisition requires twice the amount of storage of a single-polarization measurement and changes to particular polarization bases at the reception may be used to detect or characterize scatterers with a particular scattering behavior. Nevertheless, such a diversity is incomplete, since the received information depends on the emitted wave polarization.

As shown in Figure 21.5, a measurement with full polarimetric diversity requires to emit two signals with orthogonal polarization states:

$$\begin{aligned}\mathbf{E}_T \propto \mathbf{u} &\rightarrow \mathbf{E}_{R_u} = E_{\mathbf{u}\mathbf{u}} \mathbf{u} + E_{\mathbf{u}_\perp\mathbf{u}_\perp} \mathbf{u}_\perp, \\ \mathbf{E}_T \propto \mathbf{u}_\perp &\rightarrow \mathbf{E}_{R_{u_\perp}} = E_{\mathbf{u}_\perp\mathbf{u}_\perp} \mathbf{u}_\perp + E_{\mathbf{u}\mathbf{u}} \mathbf{u}.\end{aligned}\quad (21.19)$$

**FIGURE 21.4**

Synopsis of a radar operating with a polarization diversity of receive.

**FIGURE 21.5**

Synopsis of a radar operating with full polarization diversity.

This kind of acquisition requires four times the amount of storage of a single-polarization measurement and twice the amount of emitted pulses, but particular polarization bases may be tuned at both emission and reception to adapt polarization to the features of a measured scatterer in an optimal way. All the measured complex amplitudes E_{pq} , with $p, q = \mathbf{u}$ or \mathbf{u}_\perp , may be gathered into a (2×2) matrix, \mathbf{S} , called scattering matrix.

$$\mathbf{S} \propto [E_{R_u} E_{R_{u_\perp}}] = \begin{bmatrix} E_{\mathbf{u}\mathbf{u}} & E_{\mathbf{u}\mathbf{u}_\perp} \\ E_{\mathbf{u}_\perp\mathbf{u}} & E_{\mathbf{u}_\perp\mathbf{u}_\perp} \end{bmatrix}. \quad (21.20)$$

2.21.2.1.2.2 Scattering matrix and target vector

As a reflection occurs over an object, the incident and scattered Jones vectors of these fields are related by a (2×2) complex scattering matrix, \mathbf{S} , as:

$$\mathbf{E}^s = \frac{e^{-jkr}}{r} \mathbf{S} \mathbf{E}^i = \frac{e^{-jkr}}{r} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \mathbf{E}^i, \quad (21.21)$$

where \mathbf{S} is named scattering matrix and the S_{ij} , are the so-called complex scattering coefficients or complex scattering amplitudes. The diagonal elements of the scattering matrix receive the name of co-pol terms, since they relate the same polarization for the incident and the scattered fields. The off-diagonal elements are known as cross-pol terms as they relate orthogonal polarization states. Finally, the term $\frac{e^{-jkr}}{r}$ accounts for propagation effects both in amplitude and phase. The relation expressed in (21.21) relies on plane wave assumptions for both the incident and scattered fields. The elements of the scattering matrix can be related with the radar cross section of a given target as follows:

$$\sigma_{qp} = 4\pi |S_{qp}|^2. \quad (21.22)$$

The scattering matrix is generally measured in the cartesian HV basis, i.e., $\mathbf{u} = \hat{x}$ and $\mathbf{u}_\perp = \hat{y}$. Moreover, in the monostatic backscattering case, where the transmitting and receiving antennas are located at the same position, placed at the same location, the role of the transmitting and the receiving antennas can be interchanged, in case of a reciprocal propagation medium, and the scattering matrix is symmetric, i.e., $\mathbf{S} = \mathbf{S}^T$. In such a case, the \mathbf{S} may be written in the HV basis as:

$$\mathbf{S}_{xy} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{HV} & S_{VV} \end{bmatrix} = \begin{bmatrix} |S_{HH}|e^{j\phi_{HH}} & |S_{HV}|e^{j\phi_{HV}} \\ |S_{HV}|e^{j\phi_{HV}} & |S_{VV}|e^{j\phi_{VV}} \end{bmatrix}. \quad (21.23)$$

The elements S_{HH} and S_{VV} produce the power return in the co-polarized channels and the elements S_{HV} and S_{VH} produce the power return in the cross-polarized channels. The structure of the scattering matrix may be further simplified by getting rid of an absolute phase term that reveal useful for interferometric applications but does not bring any polarimetric information. This reference phase term is generally taken as the argument of the HH channel scattering term, but any other choice could be used. The resulting relative scattering matrix is the given by:

$$\mathbf{S}_{xy\text{ rel}} = e^{-j\phi_{HH}} \mathbf{S}_{xy} = \begin{bmatrix} |S_{HH}| & |S_{HV}|e^{j(\phi_{HV}-\phi_{HH})} \\ |S_{HV}|e^{j(\phi_{HV}-\phi_{HH})} & |S_{VV}|e^{j(\phi_{VV}-\phi_{HH})} \end{bmatrix}. \quad (21.24)$$

In consequence, in the monostatic case, the fully polarimetric response of a scatterer can be characterized by 5 real parameters (5 degrees of freedom): 3 amplitudes and 2 relative phases. The total scattered power, in the case of a polarimetric radar system is called the Span, and is defined as the L^2 norm of \mathbf{S} :

$$\text{Span} = \text{tr} (\mathbf{S} \mathbf{S}^{*T}) = |S_{11}|^2 + 2|S_{12}|^2 + |S_{22}|^2, \quad (21.25)$$

where $\text{tr}(\mathbf{A})$ is the trace of \mathbf{A} .

In its original form, the classical (2×2) coherent Sinclair matrix \mathbf{S} is not well adapted to manipulations required by modern signal processing techniques. In order to overcome this limitation, target or scattering vectors have been defined from linear combinations of the elements of \mathbf{S} . Among the infinite number of possibilities, two representations are generally preferred, the Pauli target vector, $\mathbf{k} = \mathbf{k}_P$, obtained from a projection onto the Pauli matrix set, and the lexicographic one, \mathbf{k}_L , defined as:

$$\mathbf{k} = \mathbf{k}_P = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} & S_{HH} - S_{VV} & 2S_{HV} \end{bmatrix}^T \quad (21.26)$$

and

$$\mathbf{k}_L = [S_{HH} \ \sqrt{2}S_{HV} \ S_{VV}]^T. \quad (21.27)$$

The norms of these scattering vectors equal the Span of the \mathbf{S} matrix

$$||\mathbf{k}||^2 = ||\mathbf{k}_L||^2 = |S_{HH}|^2 + 2|S_{HV}|^2 + |S_{VV}|^2 = \text{Span}(\mathbf{S}) \quad (21.28)$$

and the Pauli and lexicographic vectors correspond through a (3×3) special unitary matrix, $\mathbf{U}_{3(L \rightarrow P)}$

$$\mathbf{k} = \mathbf{U}_{3(L \rightarrow P)} \mathbf{k}_L \quad \text{with} \quad \mathbf{U}_{3(L \rightarrow P)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{bmatrix}. \quad (21.29)$$

2.21.2.1.2.3 Change of polarization basis

Consider a monostatic back-scattering $\mathbf{S}_{(\hat{x}, \hat{y})}$ matrix defined in the Cartesian (\hat{x}, \hat{y}) basis [17, 18], with:

$$\mathbf{E}_{(\hat{x}, \hat{y})}^s = \mathbf{S}_{(\hat{x}, \hat{y})} \mathbf{E}_{(\hat{x}, \hat{y})}^i. \quad (21.30)$$

As introduced earlier, the incident Jones vector, $\mathbf{E}_{(\hat{x}, \hat{y})}^i$ expressed in the Cartesian (\hat{x}, \hat{y}) basis, transforms to $\mathbf{E}_{(\hat{u}, \hat{u}_\perp)}^i$ in the orthonormal (\hat{u}, \hat{u}_\perp) polarimetric basis, by the way of a $SU(2)$ transformation:

$$\mathbf{E}_{(\hat{u}, \hat{u}_\perp)}^i = \mathbf{U}_{(\hat{x}, \hat{y}) \rightarrow (\hat{u}, \hat{u}_\perp)} \mathbf{E}_{(\hat{x}, \hat{y})}^i \quad (21.31)$$

with $\mathbf{U}_{(\hat{x}, \hat{y}) \rightarrow (\hat{u}, \hat{u}_\perp)} = \mathbf{U}_2(\phi, \tau, \alpha)^{-1}$, whose parameters are obtained from the expression of \hat{u} in the (\hat{x}, \hat{y}) basis. In order to apply the same polarimetric basis change to the scattered Jones vector $\mathbf{E}_{(\hat{x}, \hat{y})}^s$, one has to take into account that, in the monostatic case, the incident and scattered Jones vector represent fields propagating in opposite directions, i.e., $\mathbf{k}^s = -\mathbf{k}^i$, by an additional complex conjugation on the $SU(2)$ operator applied onto the scattered Jones vector [16]. Injecting (21.31) into (21.30) and comparing the result to $\mathbf{E}_{(\hat{u}, \hat{u}_\perp)}^s = \mathbf{S}_{(\hat{u}, \hat{u}_\perp)} \mathbf{E}_{(\hat{u}, \hat{u}_\perp)}^i$, the monostatic back-scattering matrix $\mathbf{S}_{(\hat{u}, \hat{u}_\perp)}$ expressed in the in the orthonormal (\hat{u}, \hat{u}_\perp) polarimetric basis is then given by:

$$\begin{aligned} \mathbf{S}_{(\hat{u}, \hat{u}_\perp)} &= \mathbf{U}_{(\hat{x}, \hat{y}) \rightarrow (\hat{u}, \hat{u}_\perp)}^* \mathbf{S}_{(\hat{x}, \hat{y})} \mathbf{U}_{(\hat{x}, \hat{y}) \rightarrow (\hat{u}, \hat{u}_\perp)}^{-1} \\ &= \mathbf{U}_2(\phi, \tau, \alpha)^T \mathbf{S}_{(\hat{x}, \hat{y})} \mathbf{U}_2(\phi, \tau, \alpha) \end{aligned} \quad (21.32)$$

The transformation expressed in Eq. (21.32) is named a con-similarity transformation and allows synthesizing the monostatic back-scattering matrix \mathbf{S} in any elliptical polarization basis when measured in the Cartesian $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ basis.

The change of basis may of (21.32) may be adapted to target vectors using (3×3) special unitary matrices. Despite the fact that there is no noticeable relation between $SU(2)$ and $SU(3)$ matrix groups, one may derive corresponding operators by identifying (21.32) with the definition of a target vector. In the case of scattering vectors defined in the Pauli basis, the $SU(2) - SU(3)$ correspondence, may be defined as:

$$\begin{aligned} \mathbf{U}_2(\phi) &= \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \Rightarrow \mathbf{U}_{3T}(2\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\phi & \sin 2\phi \\ 0 & -\sin 2\phi & \cos 2\phi \end{bmatrix}, \\ \mathbf{U}_2(\tau) &= \begin{bmatrix} \cos \tau & j \sin \tau \\ j \sin \tau & \cos \tau \end{bmatrix} \Rightarrow \mathbf{U}_{3T}(2\tau) = \begin{bmatrix} \cos 2\tau & 0 & j \sin 2\tau \\ 0 & 1 & 0 \\ j \sin 2\tau & 0 & \cos 2\tau \end{bmatrix}, \\ \mathbf{U}_2(\alpha) &= \begin{bmatrix} e^{+j\alpha} & 0 \\ 0 & e^{-j\alpha} \end{bmatrix} \Rightarrow \mathbf{U}_{3T}(2\alpha) = \begin{bmatrix} \cos 2\alpha & j \sin 2\alpha & 0 \\ j \sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (21.33)$$

The monostatic target vector $\mathbf{k}_{(\hat{\mathbf{x}}, \hat{\mathbf{y}})}$ expressed in the Cartesian $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ basis transforms to $\mathbf{k}_{(\hat{\mathbf{u}}, \hat{\mathbf{u}}_\perp)}$, expressed in the orthonormal $(\hat{\mathbf{u}}, \hat{\mathbf{u}}_\perp)$ polarimetric basis by the way of a Special Unitary transformation given by [20]:

$$\mathbf{k}_{(\hat{\mathbf{u}}, \hat{\mathbf{u}}_\perp)} = \mathbf{U}_{3T}(2\phi, 2\tau, 2\alpha) \mathbf{k}_{(\hat{\mathbf{x}}, \hat{\mathbf{y}})}. \quad (21.34)$$

2.21.2.1.2.4 Canonical scattering mechanisms

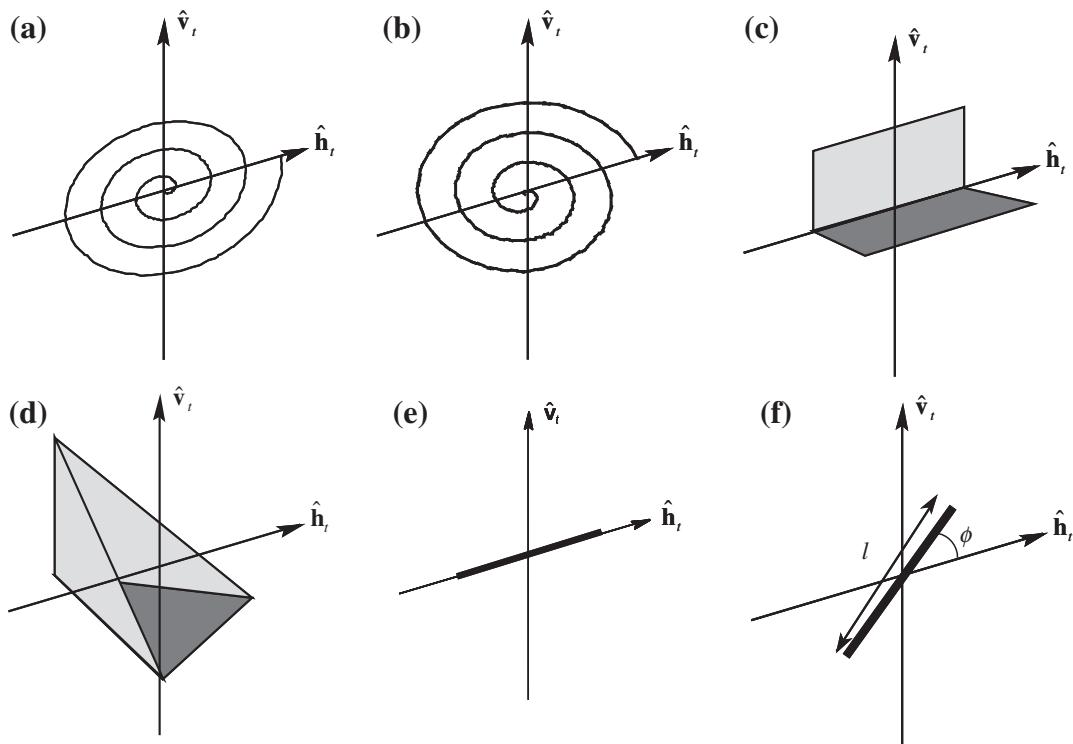
A real target always presents a complex scattering response as a consequence of its complex geometrical structure and its reflectivity properties. Consequently, the interpretation of this response is rather difficult. This section lists some elementary targets presenting canonical scattering mechanisms characterized by \mathbf{S} expressed in three canonical orthogonal polarimetric bases:

- Cartesian polarization basis (\mathbf{h}, \mathbf{v}) where \mathbf{h} stands for the horizontal polarization and \mathbf{v} for the vertical polarization.
- Linear rotated basis $(\mathbf{a}, \mathbf{a}_\perp)$ where \mathbf{a} indicates the linear polarization at 45° and \mathbf{a}_\perp the orthogonal linear polarization at -45° .
- Circular polarization basis $(\mathbf{l}, \mathbf{l}_\perp)$ where \mathbf{l} refers to the left circular polarization and $\mathbf{l}_\perp = \mathbf{r}$ to the orthogonal left circular polarization or equivalent to the right circular polarization.

The shapes of the different canonical scatterers are illustrated in Figure 21.6.

Sphere, flat plate, trihedral. Scattering matrices of a sphere, a plane or a trihedral in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis $(\mathbf{a}, \mathbf{a}_\perp)$	Circular basis $(\mathbf{l}, \mathbf{l}_\perp)$
$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}$

**FIGURE 21.6**

Canonical scatterers (a) left helix, (b) right helix, (c) dihedral, (d) trihedral, (e) horizontal dipole, (f) oriented dipole with angle ϕ .

Horizontal dipole. Scattering matrices of a dipole along the horizontal axis in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\mathbf{S} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$	$\mathbf{S} = \frac{1}{2} \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}$

Oriented dipole. Scattering matrices of a dipole oriented with an angle ϕ in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \begin{bmatrix} \cos^2 \phi & \frac{1}{2} \sin 2\phi \\ \frac{1}{2} \sin 2\phi & \sin^2 \phi \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} \frac{1+\sin 2\phi}{2} & \frac{1}{2} - \cos^2 \phi \\ \frac{1}{2} - \cos^2 \phi & \frac{1-\sin 2\phi}{2} \end{bmatrix}$	$\mathbf{S} = \frac{1}{2} \begin{bmatrix} e^{j2\phi} & -j \\ -j & e^{-j2\phi} \end{bmatrix}$

Dihedral. Scattering matrices of a horizontal dihedral in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Scattering matrices of a dihedral oriented with an angle ϕ in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos^2 \phi \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} \sin 2\phi & -\cos 2\phi \\ -\cos 2\phi & -\sin 2\phi \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} e^{j2\phi} & 0 \\ 0 & e^{-j2\phi} \end{bmatrix}$

Right helix. Scattering matrices of a right helix oriented with an angle ϕ in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \frac{e^{-j2\phi}}{2} \begin{bmatrix} 1 & -j \\ -j & -1 \end{bmatrix}$	$\mathbf{S} = \frac{e^{-j2\phi}}{2} \begin{bmatrix} -j & -1 \\ 1 & j \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} 0 & 0 \\ 0 & -e^{-j2\phi} \end{bmatrix}$

Left helix. Scattering matrices of a left helix oriented with an angle ϕ in the three polarization basis:

Cartesian basis (\mathbf{h}, \mathbf{v})	Linear rotated basis ($\mathbf{a}, \mathbf{a}_\perp$)	Circular basis ($\mathbf{l}, \mathbf{l}_\perp$)
$\mathbf{S} = \frac{e^{-j2\phi}}{2} \begin{bmatrix} 1 & j \\ j & -1 \end{bmatrix}$	$\mathbf{S} = \frac{e^{-j2\phi}}{2} \begin{bmatrix} j & -1 \\ -1 & -j \end{bmatrix}$	$\mathbf{S} = \begin{bmatrix} e^{-j2\phi} & 0 \\ 0 & 0 \end{bmatrix}$

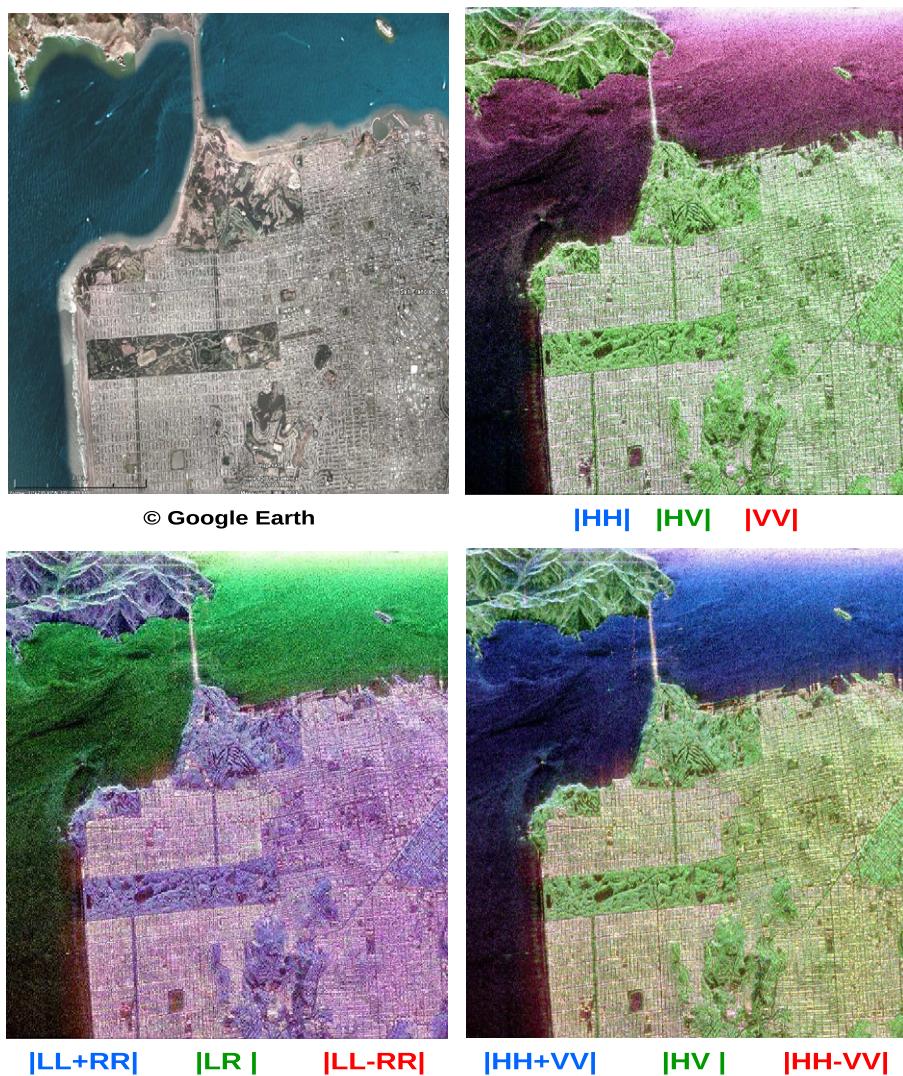
Color-coded polarimetric images are displayed for three different polarimetric bases on Figure 21.7 that can be easily interpreted in terms of canonical scattering mechanisms

2.21.2.1.3 Second order incoherent representations

2.21.2.1.3.1 Single polarization SAR: intensity

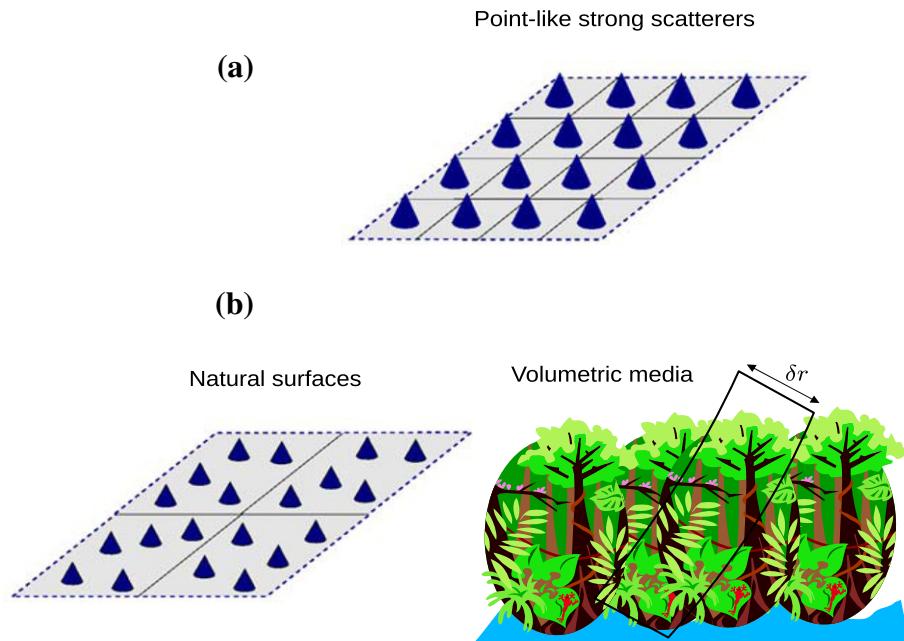
SAR imaging and the speckle effect. The objective of Synthetic Aperture Radar imaging is to provide a high resolution map of the reflectivity of a scene. This technique based on the use of both spatial and frequency diversity. The response of the scene to signals having a large bandwidth is measured from different positions along an acquisition track. Signals are generally focused using a 2-D matched filter that uses both spectral and spatial diversities to create a 2-D azimuth-range image, $s(x, r)$, with improved resolutions in both directions (Figure 21.46). The signal focused at coordinates (x, r) , may be represented as a 2-D convolution [21]:

$$s(x, r) = \int a_c(x', r') e^{-j2kr'} h(x - x', r - r') dv' + n(x, r), \quad (21.35)$$

**FIGURE 21.7**

Optical and color-coded polarimetric SAR images for three different polarimetric bases of the San-Francisco bay. Data acquired by the NASA/JPL AirSAR sensor.

where $a_c(x, r)$ represents the density of coherent reflectivity of the scene, i.e., its normalized coefficient of reflection and $h(x, r)$ is the 2-D impulse response of the SAR acquisition-processing chain, that steers the resolution properties of the image. The term $n(x, r)$ corresponds to additive noise induced by the acquisition system and may be omitted in the following for the sake of simplicity. Two extreme cases, illustrated in Figure 21.8, may be commented:

**FIGURE 21.8**

(a) deterministic (b) random SAR scenes.

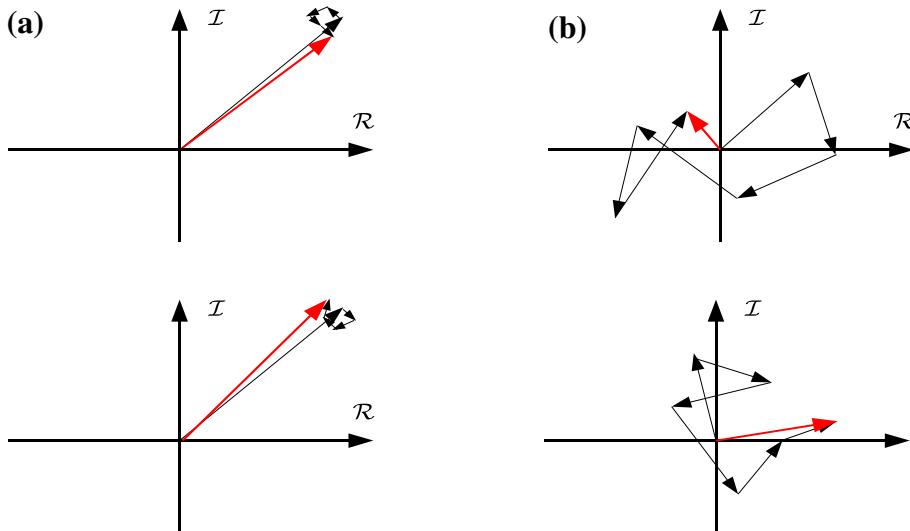
- The resolution cell under observation contains a dominant point-like scatterer with a strong response plus some negligible, clutter contributions. In this case, the coherent reflectivity density and the SAR response are given by

$$\begin{aligned} a_c(x, r) &= A e^{j\phi_0} \delta(x - x_0, r - r_0) + \epsilon \approx A e^{j\phi_0} \delta(x - x_0, r - r_0) \\ \Rightarrow s(x, r) &\approx A e^{j\phi_0} e^{-j2kr_0} h(x - x_0, r - r_0). \end{aligned} \quad (21.36)$$

- The resolution cell under observation contains a large number of independent scatterers, uniformly distributed, with random amplitude and phase and having equivalent reflectivities. In this case, the coherent reflectivity is a continuum of independent, random, contributions.

$$E(a_c(x, r) a_c^*(x', r')) = \sigma \delta(x - x', r'). \quad (21.37)$$

According to (21.35) the SAR response is the sum of all the contributions from scatterers located in the resolution cell. As it shown in Figure 21.9, in the case of a point-like scatterer with a strong response, the SAR response follows a quasi deterministic behavior dominated by a main contribution. In the case of numerous equally uniformly distributed contributions, SAR focusing may be associated to a random walk in the complex plane, and the resulting SAR response is a random variable.

**FIGURE 21.9**

Two realizations of the coherent summation over resolutions containing (a) a dominant deterministic scatterer (b) uniformly distributed random scatterers.

Over homogeneous, untextured, areas where a SAR resolution cell contains a large number of independent contributions, the SAR response has a highly random behavior, i.e., the amplitude and phase resulting from the random walk described in Figure 21.9 varies considerably from one resolution cell to the other, due to slight changes in the geometric or radiometric configurations of the elementary scatterers. This effect, known as speckle, is inherent to the use of a coherent imaging system and should not be misinterpreted as noise. Speckle is a physical phenomenon that confers SAR data a random behavior. Over areas having a homogeneous scattering behavior, the central limit theorem applies, and the SAR response is well modeled by the following product model

$$s = \sqrt{I} \eta \Rightarrow s(l) = \sqrt{I} \eta(l), \quad (21.38)$$

where I is the intensity of the response and may be related to σ in (21.37), the reflectivity of the observed medium. The parameter l represents the index of a realization of the complex random variable $\eta \in \mathbb{C}$ whose first and second moments are [22]

$$\mathbb{E}(\eta(l)) = 0, \text{var}(\eta(l)) = \mathbb{E}(|\eta(l)|^2) = 1 \quad \text{and} \quad \mathbb{E}(\eta(l)\eta(m)^*) = \delta_{m-l}. \quad (21.39)$$

The speckle component is generally assumed to follow a centered complex normal distribution, $\eta \sim \mathcal{N}_{\mathbb{C}}(0, 1)$, with independent real η_r and imaginary η_i parts, i.e., with a uniform phase distribution and identical normal distributions, $\mathcal{N}_{\mathbb{R}}(0, \frac{1}{2})$, for its real and imaginary parts.

From (21.38), the probability density function (pdf) of $s = x + jy$ is given by

$$f_{x,y}(x, y) = f_x(x)f_y(y) = \frac{1}{\pi I} e^{-\frac{x^2+y^2}{I}} \Rightarrow f_s(s) = \frac{1}{\pi I} e^{-\frac{|s|^2}{I}} \quad (21.40)$$

which may be rewritten as $s \sim \mathcal{N}_{\mathbb{C}}(0, I)$. The pdfs of the argument and squared modulus of s are

$$f_{\arg(s)}(\phi) = \frac{1}{2\pi} \quad \text{and} \quad f_{|s|^2}(\rho^2) = \frac{1}{I} e^{-\frac{\rho^2}{I}}. \quad (21.41)$$

It may be observed that, due to its uniform distribution, $\arg(s)$ cannot be used to characterize the observed environment, while $|s|^2$, the Maximum Likelihood (ML) estimate of I , shows an important variability, since $E(|s|^2) = I$ and $\text{var}(|s|^2) = I^2$, i.e., the standard deviation of the positive quantity $|s|^2$ is equal to its mean!

Second order estimates using multilooking. A solution to obtain more stable statistical estimates consists in considering L simultaneously several independent realizations, or looks, of s , $\{s(l)\}_{l=1}^L$. The joint pdf of these L independent looks is [22]

$$f(s(1) \dots s(L)) = \prod_{l=1}^L f_{s(l)}(s(l)) = \frac{1}{\pi^L I^L} e^{-\frac{1}{I} \sum_{l=1}^L |s(l)|^2} = \frac{1}{\pi^L I^L} e^{-L \hat{I}}, \quad (21.42)$$

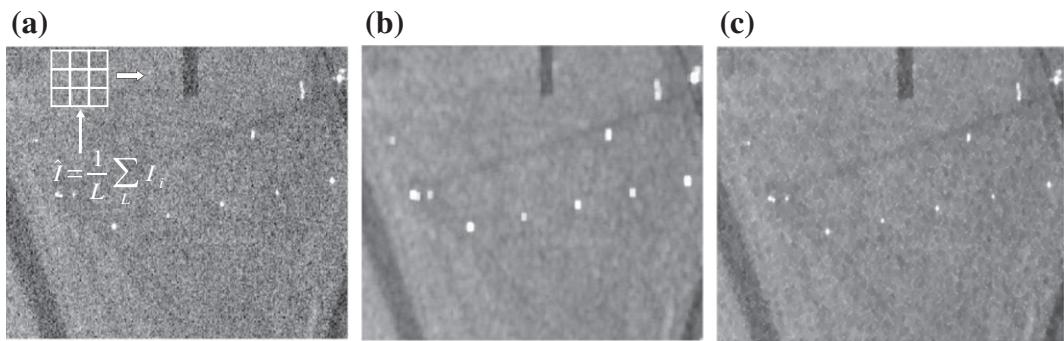
where $\hat{I} = \frac{1}{L} \sum_{l=1}^L |s(l)|^2$, the L-look sample intensity is the ML estimate of I and its statistics are given by

$$f(\hat{I}) = \frac{L^L \hat{I}^{L-1}}{(L-1)! I^L} e^{-L \hat{I}} \quad \text{with} \quad E(\hat{I}) = I, \quad \text{var}(\hat{I}) = \frac{I}{N}. \quad (21.43)$$

As illustrated, the use of a sufficiently high number of looks, L , permits to significantly reduce the standard deviation of the estimate \hat{I} , i.e., \hat{I} may be considered as a reliable estimate of the true intensity I .

Speckle filtering. In summary, a SAR response s may have very different statistical behaviors, depending on the kind of scatterer under observation. Over distributed environments, composed of a continuum of randomly distributed scatterers, the focused SAR signal, consisting of a coherent sum of a large number of independent random contributions, is affected by the speckle phenomenon characterized by a highly random aspect. The resulting SAR signal is modeled as product of an amplitude term, \sqrt{I} , related to the average scattering intensity of the illuminated medium, and of a random component, η following a circular complex normal distribution (central limit theorem) and playing the role of a speckle generator. Due to the high variability of both its phase and amplitude, s cannot be directly used to characterize the scattering properties of the illuminated medium in a reliable way. Instead, several independent realizations of the SAR signal, $s(l)$ with $l = 1, \dots, L$, may be used to build an incoherent representation, i.e., that does not depend on the uniformly distributed phase of s , called the sample intensity and defined as $\hat{I} = \frac{1}{L} \sum_{l=1}^L |s(l)|^2$, which is the ML estimate of the true intensity I over homogeneous environments.

The optimality of the sample intensity requires that the L samples $s(l)$ belong to the same distribution, given in (21.40). Among the many ways to create L samples of a SAR response, the approach generally

**FIGURE 21.10**

Intensities of images acquired by the DLR ESAR sensor at L band over Oberpfaffenhofen (a) single look, (b) (7×7) boxcar filtered, (c) (7×7) Lee filtered.

used in practice consists in selecting responses over resolution cells located in the azimuth-range vicinity of the cell under study. Under the hypothesis of a local stationary scattering behavior, this techniques permits to obtain identically distributed independent looks. Numerous studies have been led on the problem of efficient and accurate speckle filtering [11,23]. An example using a simple $(N \times N)$ sliding box and a more refined Lee filter [18], is given in Figure 21.10.

2.21.2.1.3.2 Dual polarization SAR: wave covariance matrix

Similarly to single-polarization acquisitions, measurements over several polarimetric channels may be affected by speckle. In this case, Jones vectors $\mathbf{E} = [E_x \ E_y]^T$ may be characterized using their first and second order moments

$$\mathbf{J} = \mathbb{E}(\mathbf{EE}^\dagger) = \mathbb{E}\left(\begin{bmatrix} E_x E_x^* & E_x E_y^* \\ E_y E_x^* & E_y E_y^* \end{bmatrix}\right) = \begin{bmatrix} \mathbb{E}(\mathbf{E}) = \mathbf{0}, \\ I_x & \sqrt{I_x I_y} \rho \\ \sqrt{I_x I_y} \rho^* & I_y \end{bmatrix}. \quad (21.44)$$

The (2×2) wave covariance matrix, \mathbf{J} , contains the information about the second orders moments of \mathbf{E} . Its diagonal elements provide the intensity for each polarimetric channel, whereas the normalized cross-correlation factor indicates the stability of the measured ellipse parameters, along the direction of observation (time or space). If $|\rho| = 1$ the components of the measured Jones vectors samples, $\mathbf{E}(l)$, are maximally correlated, i.e., $\mathbf{E}(l) \propto \mathbf{E} \forall l$. Oppositely, if $|\rho| = 0$ the uncorrelated components of the measured Jones vectors are very differently affected by speckle, denoting large variations of the polarimetric properties over the different samples $\mathbf{E}(l)$. A speckled Jones vector follows a complex multivariate normal distribution, $\mathbf{E}(l) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{J})$, whose expression is given by [22]

$$f_{\mathbf{E}}(\mathbf{E}(l)) = \frac{1}{\pi^2 |\mathbf{J}|} e^{-\mathbf{E}^\dagger(l) \mathbf{J}^{-1} \mathbf{E}(l)} = \frac{1}{\pi^2 |\mathbf{J}|} e^{-\text{tr}(\mathbf{J}^{-1} \mathbf{E}(l) \mathbf{E}^\dagger(l))}. \quad (21.45)$$

Similarly to the single-polarization case, the variance of the second order estimates may be reduced using L independent looks.

$$f(\{\mathbf{E}(l)\}_{l=1}^L) = \frac{1}{\pi^{2L} |\mathbf{J}|^L} e^{-\text{tr}(\mathbf{J}^{-1} \sum_{l=1}^L \mathbf{E}(l) \mathbf{E}^\dagger(l))} = \frac{1}{\pi^{2L} |\mathbf{J}|^L} e^{-L \text{tr}(\mathbf{J}^{-1} \hat{\mathbf{J}})}, \quad (21.46)$$

where $\hat{\mathbf{J}} = \frac{1}{L} \sum_{l=1}^L \mathbf{E}(l) \mathbf{E}^\dagger(l)$ is the L-look ML estimate of \mathbf{J} , and follows a Wishart complex pdf [18]

$$f(\hat{\mathbf{J}}) = \frac{L^{2L} |\hat{\mathbf{J}}|^{L-2}}{\tilde{\Gamma}_2(L) |\mathbf{J}|^L} e^{-L \text{tr}(\mathbf{J}^{-1} \hat{\mathbf{J}})}, \quad (21.47)$$

where $\tilde{\Gamma}_2(L)$ is the complex Gamma function.

In order to characterize the variability of the received Jones vector samples, an indicator, called degree of polarization has been defined:

$$\text{DoP} = \sqrt{1 - 4 \frac{|\mathbf{J}|}{\text{tr}(\mathbf{J})^2}}. \quad (21.48)$$

A received wave is said to be fully polarized if $\text{DoP} = 1$, i.e., $|\rho| = 1$. For depolarized waves, $\text{DoP} = 0$ corresponds to completely uncorrelated Jones vectors, with $|\rho| = 0$, and whose parameters vary in such a random way that the backscattered intensity is equally distributed over the polarimetric channels, and \mathbf{J} is proportional to the identity matrix. Between these two extreme cases lies the general case of partial polarization, with $0 < \text{DoP} < 1$. It is however important to note that the elements of the wave covariance matrix \mathbf{J} depend on the choice of the polarization basis. Let $\mathbf{J}_{(x,y)}$ the wave covariance matrix expressed in the Cartesian (\mathbf{x}, \mathbf{y}) basis, transforms to $\mathbf{J}_{(u,u_\perp)}$ in the orthogonal $(\mathbf{u}, \mathbf{u}_\perp)$ polarimetric basis, by the way of a Special Unitary similarity transformation as [16]:

$$\begin{aligned} \mathbf{J}_{(u,u_\perp)} &= E(\mathbf{E}_{(\mathbf{u},\mathbf{u}_\perp)} \mathbf{E}_{(\mathbf{u},\mathbf{u}_\perp)}^\dagger) \\ &= \mathbf{U}_{(x,y) \rightarrow (u,u_\perp)} E(\mathbf{E}_{(x,y)} \mathbf{E}_{(x,y)}^\dagger) \mathbf{U}_{(x,y) \rightarrow (u,u_\perp)}^\dagger \\ &= \mathbf{U}_{(x,y) \rightarrow (u,u_\perp)} \mathbf{J}_{(x,y)} \mathbf{U}_{(x,y) \rightarrow (u,u_\perp)}^\dagger, \end{aligned} \quad (21.49)$$

where $\mathbf{U}_{(x,y) \rightarrow (u,u_\perp)}$ corresponds to the elliptical basis transformation $SU(2)$ matrix. The trace and determinant of a Hermitian matrix being invariant under unitary similarity transformations, the wave degree of polarization (DoP) is independent of the polarization basis chosen to represent the covariance matrix.

2.21.2.1.3.3 Fully polarimetric SAR: coherency or covariance matrices

Data statistics. In order to easily compute and exploit second order statistics, the elements of speckle-affected scattering matrix samples, $\mathbf{S}(l)$ are re-arranged into target vectors, $\mathbf{k}(l)$ or $\mathbf{k}_L(l)$, introduced in (21.26) and (21.27). The first and second order moments of these vectors are given by

$$E(\mathbf{k}) = E(\mathbf{k}_L) = \mathbf{0}, \quad \mathbf{T} = E(\mathbf{k}\mathbf{k}^\dagger), \quad \text{and} \quad \mathbf{C} = E(\mathbf{k}_L\mathbf{k}_L^\dagger), \quad (21.50)$$

where the (3×3) covariance matrices, \mathbf{T} and \mathbf{C} are called coherency and covariance matrices and are related through a $SU(3)$ transformation:

$$\mathbf{k} = \mathbf{U}_{3(L \rightarrow P)} \mathbf{k}_L \Rightarrow \mathbf{T} = \mathbf{U}_{3(L \rightarrow P)} \mathbf{C} \mathbf{U}_{3(L \rightarrow P)}^\dagger. \quad (21.51)$$

For a fully developed speckle, when the distributed scattering mentioned previously are satisfied, a target vector follows a complex normal pdf:

$$f_{\mathbf{k}}(\mathbf{k}(l)) = \frac{1}{\pi^3 |\mathbf{T}|} e^{-\mathbf{k}^\dagger(l) \mathbf{T}^{-1} \mathbf{k}(l)} = \frac{1}{\pi^3 |\mathbf{T}|} e^{-\text{tr}(\mathbf{T}^{-1} \mathbf{k}(l) \mathbf{k}^\dagger(l))}. \quad (21.52)$$

The joint pdf of L independent realizations is given by

$$f(\{\mathbf{k}(l)\}_{l=1}^L) = \frac{1}{\pi^{3L} |\mathbf{T}|^L} e^{-\text{tr}(\mathbf{T}^{-1} \sum_{l=1}^L \mathbf{k}(l) \mathbf{k}^\dagger(l))} = \frac{1}{\pi^{3L} |\mathbf{T}|^L} e^{-L \text{tr}(\mathbf{T}^{-1} \widehat{\mathbf{T}})}, \quad (21.53)$$

where $\widehat{\mathbf{T}} = \frac{1}{L} \sum_{l=1}^L \mathbf{k}(l) \mathbf{k}^\dagger(l)$ is the L-look ML estimate of \mathbf{T} , and follows a Wishart complex pdf

$$f(\widehat{\mathbf{T}}) = \frac{L^{3L} |\widehat{\mathbf{T}}|^{L-2}}{\tilde{\Gamma}_2(L) |\mathbf{T}|^L} e^{-L \text{tr}(\mathbf{T}^{-1} \widehat{\mathbf{T}})}, \quad (21.54)$$

where $\tilde{\Gamma}_3(L)$ is the complex Gamma function. Similar expressions may be obtained for the lexicographic definition, by replacing \mathbf{k} with \mathbf{k}_L and \mathbf{T} with \mathbf{C} .

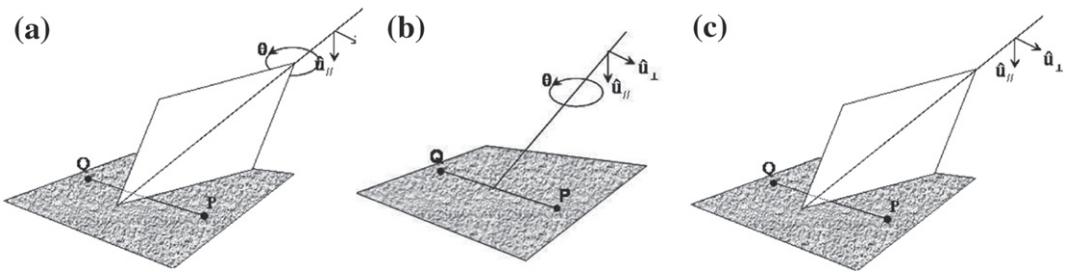
Properties of the coherency and covariance matrices. Developed as well as parametrized expressions of \mathbf{T} and \mathbf{C} are given hereafter.

The coherency matrix may be written as:

$$\begin{aligned} \mathbf{T} &= \frac{1}{2} \mathbf{E} \left(\begin{bmatrix} |S_{xx} + S_{yy}|^2 & (S_{xx} + S_{yy})(S_{xx} - S_{yy})^* & 2(S_{xx} + S_{yy})S_{xy}^* \\ (S_{xx} - S_{yy})(S_{xx} + S_{yy})^* & |S_{xx} - S_{yy}|^2 & 2(S_{xx} - S_{yy})S_{xy}^* \\ 2S_{xy}(S_{xx} + S_{yy})^* & 2S_{xy}(S_{xx} - S_{yy})^* & 4|S_{xy}|^2 \end{bmatrix} \right) \\ &= \begin{bmatrix} 2A_0 & C - jD & H + jG \\ C + jD & B_0 + B & E + jF \\ H - jG & E - jF & B_0 - B \end{bmatrix}, \end{aligned} \quad (21.55)$$

where $A_0, B_0, B, C, D, E, F, G, H$ are 9 real coefficients, called Huynen parameters, used by Huynen [6] to characterize second order polarimetric representations. The covariance matrix \mathbf{C} may be expressed as

$$\begin{aligned} \mathbf{C} &= \mathbf{E} \left(\begin{bmatrix} |S_{XX}|^2 & \sqrt{2}S_{XX}S_{XY}^* & S_{XX}S_{YY}^* \\ \sqrt{2}S_{XY}S_{XX}^* & 2|S_{XY}|^2 & \sqrt{2}S_{XY}S_{YY}^* \\ S_{YY}S_{XX}^* & \sqrt{2}S_{YY}S_{XY}^* & |S_{YY}|^2 \end{bmatrix} \right) \\ &= \sigma \begin{bmatrix} 1 & \beta\sqrt{\delta} & \rho\sqrt{\gamma} \\ \beta^*\sqrt{\delta} & \delta & \varepsilon\sqrt{\gamma\delta} \\ \rho^*\sqrt{\gamma} & \varepsilon^*\sqrt{\gamma\delta} & \gamma \end{bmatrix}, \end{aligned} \quad (21.56)$$

**FIGURE 21.11**

Geometrical configuration of (a) azimuthal symmetry, (b) rotation symmetry, (c) reflection symmetry.

where σ, δ, γ are real parameters and β, ρ, ϵ are complex and form a 9-element set of real coefficients.

Coherency and covariance matrices may be represented in any arbitrary polarimetric basis using a Special Unitary similarity transformation:

$$\begin{aligned}\mathbf{T}_{(u,u_\perp)} &= E(\mathbf{k}_{(u,u_\perp)}\mathbf{k}_{(u,u_\perp)}^\dagger) \\ &= \mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)} E(\mathbf{k}_{(x,y)}\mathbf{k}_{(x,y)}^\dagger) \mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)}^\dagger \\ &= \mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)} \mathbf{T}_{(x,y)} \mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)}^\dagger\end{aligned}\quad (21.57)$$

where $\mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)}$ is an elliptical basis transformation $SU(3)$ matrix adapted to the coherency matrix formalism. For covariances matrices equivalent representations may be obtained by combining $\mathbf{U}_{T(x,y) \rightarrow (u,u_\perp)}$ with $\mathbf{U}_{3(L \rightarrow P)}$.

Coherency and covariance matrices have specific structures [12] over media showing particular symmetry properties shown in Figure 21.11. Some media exhibit reflection symmetry when the distribution of the polarimetric responses is symmetric with respect to the plane of incidence. Such a behavior is mainly observed over horizontal natural environments and is characterized by a lack of correlation between co- and cross-polarization channels, i.e., $E(S_{hh}S_{hv}^*) = E(S_{vv}S_{hv}^*) = 0$. Other media may have covariance matrices that remain invariant by rotation around the radar line of sight, i.e., invariant under transformation through the $U_3(2\phi)$ $SU(3)$ operator. Finally, media exhibiting both reflection and rotation symmetries are said to be azimuth symmetric. Such a response is generally encountered over dense volumetric environments.

The specific shapes of the coherency and covariance matrices corresponding to the three different scattering symmetry configurations are:

- *Reflection symmetry case:*

$$\mathbf{T} = \begin{bmatrix} a & b & 0 \\ b^* & c & 0 \\ 0 & 0 & d \end{bmatrix} \Leftrightarrow \mathbf{C} = \begin{bmatrix} \alpha & 0 & \beta \\ 0 & \delta & 0 \\ \beta^* & 0 & \gamma \end{bmatrix}. \quad (21.58)$$

- *Rotation symmetry case:*

$$\mathbf{T} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & c^* & b \end{bmatrix} \Leftrightarrow \mathbf{C} = \begin{bmatrix} \alpha & \beta & \delta \\ \beta^* & \alpha - \delta & -\beta^* \\ \delta & -\beta & \alpha \end{bmatrix}. \quad (21.59)$$

- *Azimuthal symmetry case:*

$$\mathbf{T} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix} \Leftrightarrow \mathbf{C} = \begin{bmatrix} \alpha & 0 & \beta \\ 0 & \alpha - \beta & 0 \\ \beta & 0 & \alpha \end{bmatrix}. \quad (21.60)$$

2.21.2.2 Polarimetric decomposition techniques

2.21.2.2.1 Necessity of polarimetric decompositions

As it has been shown previously, polarimetric scattering behaviors may be analyzed by comparing relative scattering matrices, \mathbf{S}_{rel} , to canonical scattering mechanisms, i.e., plate, dipole, helix ...

The use of second order incoherent polarimetric representations does not represent a limitation since a relative scattering matrix can be reconstructed from the hermitian external product of a target vector:

$$\mathbf{S}(l) \equiv \mathbf{k}(l) \Rightarrow \mathbf{S}_{\text{rel}}(l) \equiv \mathbf{k}(l)\mathbf{k}^\dagger(l). \quad (21.61)$$

Nevertheless, incoherent averaging involved in the speckle filtering stage may perturb significantly the identification of a polarimetric scattering mechanism from a coherency matrix. The only configuration for which it is possible to establish a unique relation between \mathbf{S}_{rel} and \mathbf{T} may be expressed as:

$$\mathbf{k}(l) = \alpha(l)\mathbf{k}_0 \Rightarrow \mathbf{T} = \mathbf{E}(\mathbf{k}\mathbf{k}^\dagger) = A\mathbf{k}_0\mathbf{k}_0^\dagger \equiv \mathbf{S}_{\text{rel}0} \quad (21.62)$$

In such a case, speckle only affects the Span of the polarimetric responses, relative phases and amplitudes remain unaffected and polarimetric scattering is said to be deterministic.

Other more likely configurations may be considered. that involve several independent, and hence in the Gaussian case uncorrelated, scattering mechanisms occurring within each resolution cell:

$$\left. \begin{array}{l} \mathbf{k}(l) = \sum_{n=1}^{N_s} \mathbf{k}_n(l) \\ \mathbf{E}(\mathbf{k}_n(l)\mathbf{k}_m^\dagger(l)) = \mathbf{T}_n\delta_{n-m} \end{array} \right\} \Rightarrow \mathbf{T} = \mathbf{E}(\mathbf{k}\mathbf{k}^\dagger) = \sum_{n=1}^{N_s} \mathbf{T}_n \neq \mathbf{S}_{\text{rel}}. \quad (21.63)$$

In this very frequently encountered case, a coherency matrix cannot be related to a single relative scattering matrix.

2.21.2.2.1.1 Concept of distributed coherency matrix

A coherency matrix is said to be distributed if it cannot be related to a single relative scattering matrix [12]. This property may be revealed using various approaches. First one may consider the decomposition of a coherency matrix onto its eigenvector basis:

$$\mathbf{T} = \mathbf{V}\Lambda\mathbf{V}^\dagger = \sum_{i=1}^3 \lambda_i \mathbf{v}_i \mathbf{v}_i^\dagger \quad \text{with } \mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3], \ \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3), \quad (21.64)$$

where \mathbf{V} is composed of orthogonal eigenvectors, i.e., $\mathbf{V}^{-1} = \mathbf{V}^\dagger$, and the real eigenvalues may be sorted as $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$. It is obvious in (21.64) that if \mathbf{T} has a unitary rank, i.e., $\lambda_2 = \lambda_3 = 0$, then a unique relative scattering matrix may be associated to $\lambda_1 \mathbf{v}_1 \mathbf{v}_1^\dagger$. For other configurations, a simplification of the eigenstructure in (21.64), i.e., a selection of dominant terms, may be a way to reconstruct a representative equivalent deterministic polarimetric scattering mechanism.

The concept of Degrees of Freedom (DoF) may also be used to illustrate the concept of distributed coherency matrix. A coherency matrix may be represented using normalized correlation coefficients as follows:

$$\mathbf{T} = \begin{bmatrix} \sqrt{T_{11}} & 0 & 0 \\ 0 & \sqrt{T_{22}} & 0 \\ 0 & 0 & \sqrt{T_{33}} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12}^* & 1 & \rho_{23} \\ \rho_{13}^* & \rho_{23}^* & 1 \end{bmatrix} \begin{bmatrix} \sqrt{T_{11}} & 0 & 0 \\ 0 & \sqrt{T_{22}} & 0 \\ 0 & 0 & \sqrt{T_{33}} \end{bmatrix}. \quad (21.65)$$

A “pure” coherency matrix, i.e., composed of a single scattering mechanism, possesses the following properties:

$$\begin{aligned} \mathbf{T} &= \lambda_1 \mathbf{v}_1 \mathbf{v}_1^\dagger \quad \text{with } |\rho_{ij}| = 1, \rho_{23} = \rho_{12}^* \rho_{13} \\ \Rightarrow \mathbf{T} &\equiv \{T_{11}, T_{22}, T_{33}, \arg(\rho_{12}), \arg(\rho_{13})\} \end{aligned} \quad (21.66)$$

Then $\text{DoF}(\mathbf{T}) = 5 = \text{DoF}(\mathbf{S}_{\text{rel}})$.

For a distributed matrix, a similar approach gives

$$\begin{aligned} \mathbf{T} &= \sum_{i=1}^3 \lambda_i \mathbf{v}_i \mathbf{v}_i^\dagger \quad \text{with } |\rho_{ij}| \leq 1, \rho_{23} = \rho_{23} \\ \Rightarrow \mathbf{T} &\equiv \{T_{11}, T_{22}, T_{33}, |\rho_{12}|, |\rho_{13}|, |\rho_{23}|, \arg(\rho_{12}), \arg(\rho_{13}), \arg(\rho_{23})\}. \end{aligned} \quad (21.67)$$

Then $\text{DoF}(\mathbf{T}) = 9 > \text{DoF}(\mathbf{S}_{\text{rel}}) = 5$.

2.21.2.2.1.2 Types of polarimetric decomposition techniques

The main objective of polarimetric decomposition techniques is to characterize a polarimetric response using one or a set of scattering mechanisms. Four main types of decompositions may be considered:

- The coherent decomposition techniques: they operate over scattering matrices \mathbf{S} in order to express to express it as a combination of components corresponding to canonical scattering mechanisms. Such an approach is valid only over deterministic environments whose response is not affected by speckle. For distributed responses, coherent decomposition results may be meaningless.
- The incoherent decompositions based on dichotomy: the objective is to rewrite a coherency matrix under the form of a pure response plus a remaining term, $\mathbf{T} = \mathbf{T}_0 + \mathbf{T}_{\text{rem}}$. This dichotomy is driven by physical considerations.
- The decompositions based on the eigenvector decomposition of \mathbf{T} .
- Those based on a scattering model that aim to identify specific full-rank scattering mechanisms from \mathbf{T} .

2.21.2.2.2 Some eigenvector based decomposition technique

An important class of decomposition techniques is based on the eigenvector decomposition of the coherency matrix recalled hereafter:

$$\mathbf{T} = \mathbf{V} \Lambda \mathbf{V}^\dagger = \sum_{i=1}^3 \lambda_i \mathbf{v}_i \mathbf{v}_i^\dagger \quad \text{with } \mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3], \ \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3), \quad (21.68)$$

where \mathbf{V} is a $SU(3)$ operator. The real eigenvalues may be sorted as $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and are independent of the polarization basis used to represent \mathbf{T} . The decomposition leads to three scattering mechanisms, that are orthogonal and hence uncorrelated. The weight of each contribution of each deterministic mechanism to the global distributed scattering is specified by the eigenvalue λ_i , whereas the type of scattering is related to the unitary eigenvector \mathbf{v}_i . If only one eigenvalue is nonzero then the unitary rank coherency matrix \mathbf{T} corresponds to a pure target and can be related to a single scattering matrix. On the other hand, if all eigenvalues are equal, the coherency matrix is composed of three orthogonal scattering mechanisms with equal amplitudes, the target is said to be random and there is no correlated polarized structure at all. Between these two extremes, there exists the case of partial targets where the coherency matrix has non-null and non-equal eigenvalues. The analysis of its polarimetric properties requires a study of the eigenvalues distribution as well as a characterization of each scattering mechanism of the expansion.

2.21.2.2.2.1 Cloude dominant scattering mechanism decomposition

Cloude was the first to consider eigenvector-based decomposition [24], a well-known technique in signal processing. The dominant dominant scattering mechanism is associated to the eigenvector corresponding to the largest eigenvalue

$$\mathbf{S}_{\text{rel}} \equiv \lambda_1 \mathbf{v}_1 = \mathbf{k}_1. \quad (21.69)$$

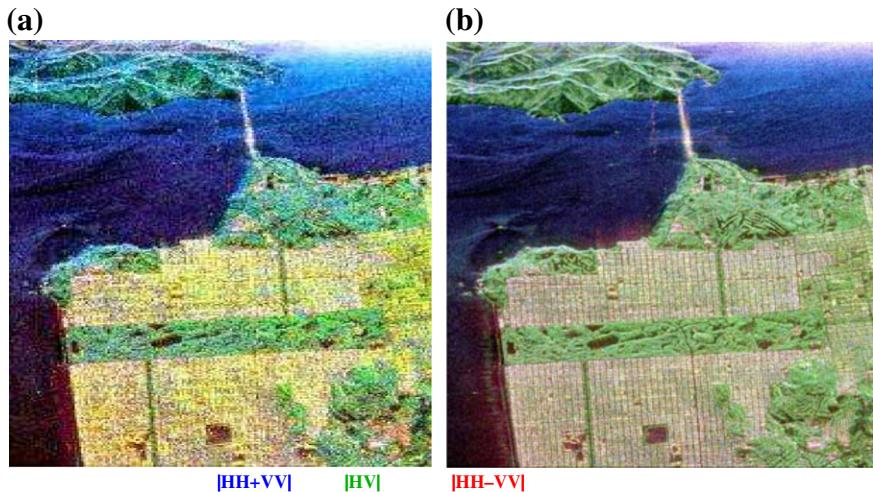
The resulting target vector may be rewritten using Huynen parameters [6] as

$$\mathbf{k}_1 = \frac{1}{\sqrt{2A_0}} \begin{bmatrix} 2A_0 \\ C + jD \\ C - jD \end{bmatrix} = \begin{bmatrix} \sqrt{2A_0} \\ \sqrt{B_0 + B} e^{+j \tan(D/C)} \\ \sqrt{B_0 - B} e^{-j \tan(G/H)} \end{bmatrix}. \quad (21.70)$$

This polarimetric parametrization of the target vector \mathbf{k}_1 involves the fit of a combination of three simple scattering mechanisms: surface scattering, dihedral scattering and volume scattering, which are characterized from the three components of the target vector such as:

- *Surface scattering: $A_0 \gg B_0 + B, B_0 - B$.*
- *Dihedral scattering: $B_0 + B \gg A_0, B_0 - B$.*
- *Volume scattering: $B_0 - B \gg A_0, B_0 + B$.*

The polarimetric color-coded images built using the first eigenvector only is displayed in Figure 21.12 and is compared to the original Pauli basis image.

**FIGURE 21.12**

Polarimetric color-coded images in the Pauli basis (a) after Cloude decomposition (b) original image.

2.21.2.2.2.2 The $H/A/\bar{\alpha}$ decomposition

Mean scattering mechanism. The dominant scattering mechanism approach mentioned above remains valid when $\lambda_1 \gg \lambda_2, \lambda_3$, but may lead to a significant loss of information when the importance secondary mechanisms cannot be neglected. An alternative technique has been proposed for extracting average parameters that is based on a particular parametrization of the eigenvectors and that account for the magnitude (eigenvalue) of each scattering mechanism (eigenvector) in order to define an average pure scattering mechanism [12, 13].

Each unitary eigenvector of \mathbf{T} is parametrized using four angular parameters and an arbitrary absolute phase:

$$\mathbf{v}_i = e^{j\phi_i} [\cos \alpha_i \quad \sin \alpha_i \cos \beta_i e^{j\delta_i} \quad \sin \alpha_i \sin \beta_i e^{j\gamma_i}]^T. \quad (21.71)$$

The eigenvector matrix may then be written as:

$$\mathbf{V} = \begin{bmatrix} \cos \alpha_1 e^{j\phi_1} & \cos \alpha_2 e^{j\phi_2} & \cos \alpha_3 e^{j\phi_3} \\ \sin \alpha_1 \cos \beta_1 e^{j(\delta_1+\phi_1)} & \sin \alpha_2 \cos \beta_2 e^{j(\delta_2+\phi_2)} & \sin \alpha_3 \cos \beta_3 e^{j(\delta_3+\phi_3)} \\ \sin \alpha_1 \sin \beta_1 e^{j(\gamma_1+\phi_1)} & \sin \alpha_2 \sin \beta_2 e^{j(\gamma_2+\phi_2)} & \sin \alpha_3 \sin \beta_3 e^{j(\gamma_3+\phi_3)} \end{bmatrix}. \quad (21.72)$$

The parametrization of a \mathbf{V} matrix is made so as to enable a probabilistic interpretation of the scattering process. Considering the eigenvector decomposition as a statistical model of the polarimetric response under the form of a 3 symbol Bernoulli process i.e., the response is modeled as the sum of three uncorrelated \mathbf{S} matrices, represented by the columns of the \mathbf{V} matrix, occurring with pseudo-probabilities

p_i , given by:

$$p_i = \frac{\lambda_i}{\sum_{i=1}^3 \lambda_i} = \frac{\lambda_i}{\text{Span}} \quad \text{with} \quad \sum_{i=1}^3 p_i = 1, \quad (21.73)$$

i.e., p_i represents the fraction of the Span that weights the i th scattering mechanism. The mean pure scattering mechanism is extracted under the form of unitary target vector \mathbf{v}_0 , made of mean scattering parameters:

$$\mathbf{v}_0 = e^{j\phi} \begin{bmatrix} \cos \bar{\alpha} \\ \sin \bar{\alpha} \cos \bar{\beta} e^{j\bar{\delta}} \\ \sin \bar{\alpha} \sin \bar{\beta} e^{j\bar{\gamma}} \end{bmatrix}, \quad (21.74)$$

where ϕ is physically equivalent to an absolute phase and where mean parameters are obtained as:

$$\bar{\alpha} = \sum_{i=1}^3 P_i \alpha_i, \bar{\beta} = \sum_{i=1}^3 P_i \beta_i, \bar{\delta} = \sum_{i=1}^3 P_i \delta_i, \bar{\gamma} = \sum_{i=1}^3 P_i \gamma_i. \quad (21.75)$$

A mean target vector is the defined as [12, 13]:

$$\mathbf{k}_0 = \sqrt{\bar{\lambda}} \mathbf{v}_0 = \sqrt{\bar{\lambda}} e^{j\phi} \begin{bmatrix} \cos \bar{\alpha} \\ \sin \bar{\alpha} \cos \bar{\beta} e^{j\bar{\delta}} \\ \sin \bar{\alpha} \sin \bar{\beta} e^{j\bar{\gamma}} \end{bmatrix}, \quad (21.76)$$

where the parameter $\bar{\lambda}$ corresponds to the mean target power (Span) and is defined by:

$$\bar{\lambda} = \sum_{i=1}^3 p_i \lambda_i. \quad (21.77)$$

The reconstructed mean scattering mechanism may then be associated to a unique scattering matrix and in order to be characterized. The polarimetric color-coded images built using the mean parameters is displayed in Figure 21.13 and is compared to the original Pauli basis image.

Decomposition parameters. The decomposition parameters are related to underlying physical scattering mechanisms and hence may be used to associate observables with physical properties of the medium. The parameter $\bar{\alpha}$ is roll invariant, i.e., remains constant under any rotation around the radar line of sight, and is an indicator of the nature of the scattering mechanism [12]. The three others parameters ($\bar{\beta}$, $\bar{\gamma}$, and $\bar{\delta}$) can be used to define the orientation of the observed object [18].

The useful range of the parameter $\bar{\alpha}$ corresponds to a continuous change from surface scattering in the geometrical optics limit $\bar{\alpha} = 0^\circ$ through surface scattering under physical optics to the Bragg surface model, encompassing dipole scattering or single scattering by a cloud of anisotropic particles $\bar{\alpha} = 45^\circ$, moving into double bounce scattering mechanisms between two dielectric surfaces and finally reaching dihedral scatter from metallic surfaces $\bar{\alpha} = 90^\circ$.

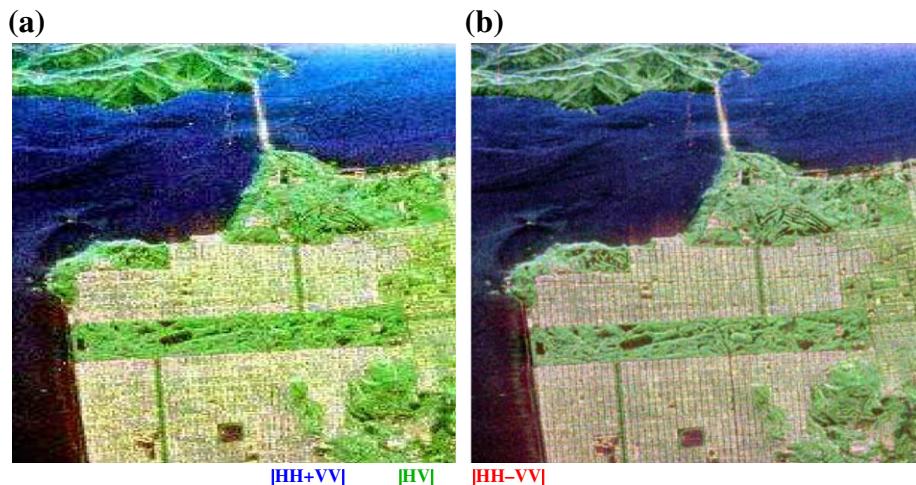


FIGURE 21.13

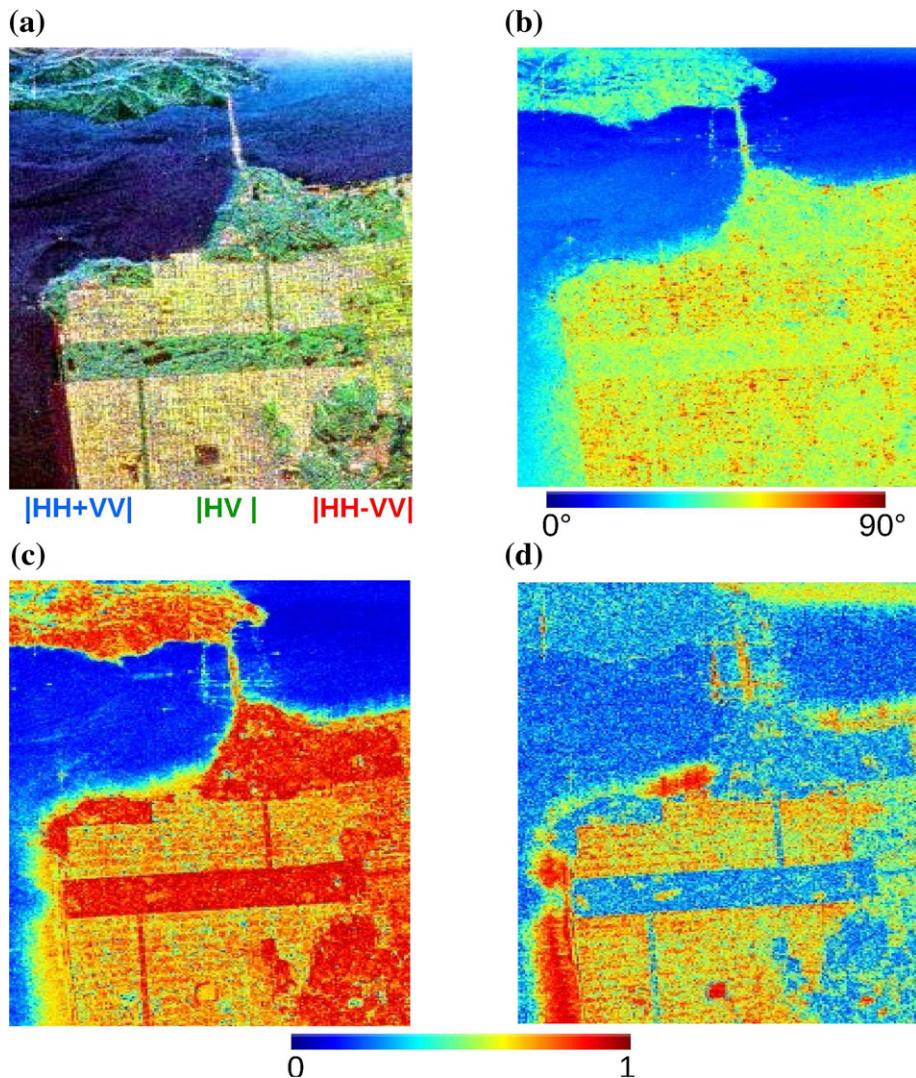
Polarimetric color-coded images in the Pauli basis (a) after reconstruction of a mean scattering mechanism (b) original image.

The image in Figure 21.14 shows that the $\bar{\alpha}$ parameter is related directly to underlying average physical scattering mechanism, and hence may be used to associate observables with physical properties of the medium. Low value occurs over the ocean region, indicative of dominant single scattering $\bar{\alpha} = 0^\circ$. Urban area and parkland areas consist of medium and high parameter values ($45^\circ < \bar{\alpha} < 90^\circ$).

The set of eigenvalues can be completely defined by the Span (sum of the eigenvalues) and two parameters: the polarimetric entropy, which an indicator of the randomness of the polarimetric scattering phenomenon and the anisotropy, that indicates the relative importance of secondary mechanisms. It was shown previously that for a unitary rank coherency matrix, i.e., $\lambda_2 = \lambda_3 = 0$, the scattering is deterministic and \mathbf{T} may be associated to a single scattering matrix, whereas if all eigenvalues are identical ($\lambda_1 = \lambda_2 = \lambda_3$) then the averaged coherency \mathbf{T}_3 matrix represents a completely uncorrelated, non-polarized random scattering structure. In between these two extreme behaviors, the case of distributed or partially polarized scatterers prevails. In order to define the degree of statistical disorder of each distinct scatter type within the ensemble, the polarimetric entropy H , provides an efficient and suitable basis-invariant parameter, and is given by [12, 13]:

$$H = - \sum_{i=1}^N p_i \log_3 (p_i) \quad \text{with } 0 \leq H \leq 1. \quad (21.78)$$

For a low entropy value, $H < 0.3$, then the observed medium may be considered as weakly depolarizing and the dominant scattering mechanism may be recovered, using the eigenvector corresponding to the largest eigenvalue, and the other components may be neglected. However, if the entropy is high, this

**FIGURE 21.14**

$H/A/\bar{\alpha}$ decomposition parameters (a) original Pauli image, (b) $\bar{\alpha}$, (c) H , (d) A .

approximation is not valid anymore and a mean scattering mechanism may be extracted. When $H = 1$, T is proportional to the identity matrix, the polarization information is a white noise process.

The image in Figure 21.14 shows that low entropy scattering occurs over the ocean (scattering by a slightly rough surface). High entropy occurs over the parkland areas. At this resolution, the urban area

consists of a mixture of low and high entropy processes, which are due to the different street/building classes that are aligned along the radar look direction, or aligned somewhat off bore sight, or 45° aligned.

Another eigenvalue parameter, defined as the polarimetric anisotropy A can be introduced to completely describe the set of eigenvalues [18]:

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \quad \text{with } 0 \leq A \leq 1. \quad (21.79)$$

The polarimetric anisotropy A , shown in Figure 21.14, is complementary to the polarimetric entropy H and measures the relative importance of the second and the third eigenvalues of the eigendecomposition. From a practical point of view, the anisotropy A can be employed as a source of discrimination mainly when $H > 0.7$. The reason is that for lower entropy values, the second and third eigenvalues are highly affected by noise. Consequently, the anisotropy A is also very noisy. Inherent of the spatial averaging, however, the entropy H increases, and the number of distinguishable classes identifiable from polarimetric observations reduces. As example, an entropy $H = 0.9$ can correspond to two limit types of scattering process with associated eigenvalues spectra given by $(\lambda_1 = 1, \lambda_2 = 0.4, \lambda_3 = 0.3)$ and $(\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 0.3)$.

2.21.2.2.3 The model based Freeman-Durden decomposition

The Freeman-Durden decomposition [9] is a technique for fitting a physically based three-component scattering mechanism model to polarimetric SAR observations. The involved mechanisms are volumetric scattering from a cloud of randomly oriented dipoles, even- or double-bounce reflection from a pair of orthogonal surfaces with different dielectric constants, and surface scattering from a moderately rough surface. This composite scattering model is used to describe the polarimetric backscattering of natural environments.

The first component of the Freeman-Durden decomposition consists of a slightly rough surface scattering response modeled as:

$$\mathbf{S} = \begin{bmatrix} R_{hh} & 0 \\ 0 & R_{vv} \end{bmatrix} \Rightarrow \mathbf{C}_s = f_s \begin{bmatrix} |\beta|^2 & 0 & \beta \\ 0 & 0 & 0 \\ \beta^* & 0 & 1 \end{bmatrix}$$

$$\text{with } f_s = |R_{vv}|^2, \beta = \frac{R_{hh}}{R_{vv}}, \quad (21.80)$$

where the reflection coefficients depend on the chosen scattering model.

The double-bounce scattering component, representing successive reflections on a ground and on a tree trunk for instance, is modeled as:

$$\mathbf{S} = \begin{bmatrix} e^{j2\gamma_{hh}} R_{T_{hh}} R_{G_{hh}} & 0 \\ 0 & e^{j2\gamma_{vv}} R_{T_{vv}} R_{G_{vv}} \end{bmatrix} \Rightarrow \mathbf{C}_d = f_d \begin{bmatrix} |\alpha|^2 & 0 & \alpha \\ 0 & 0 & 0 \\ \alpha^* & 0 & 1 \end{bmatrix}$$

$$\text{with } f_d = |R_{T_{vv}} R_{G_{vv}}|^2, \alpha = e^{j2(\gamma_{hh}-\gamma_{vv})} \frac{R_{T_{hh}} R_{G_{hh}}}{R_{T_{vv}} R_{G_{vv}}}, \quad (21.81)$$

where R_T and R_G represent forward reflection coefficient on the trunk and ground, respectively, and the exponential terms account for path propagation effects. The volume scattering from a forest canopy is modeled as the contribution from a cloud of randomly oriented cylinder-like scatterers.

$$\begin{aligned} \mathbf{S} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \Rightarrow \mathbf{C}_v &= \lim_{a \text{ or } b \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} \mathbf{U}_C(\phi) \begin{bmatrix} |a|^2 & 0 & ab^* \\ 0 & 0 & 0 \\ a^*b & 0 & |b|^2 \end{bmatrix} \mathbf{U}_C^\dagger(\phi) d\phi \\ &\Rightarrow \mathbf{C}_v = \frac{f_v}{8} \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \end{aligned} \quad (21.82)$$

where f_v corresponds to the contribution of the volume scattering component. Assuming that the volume, double-bounce, and surface scatter components are uncorrelated, the total second-order statistics are the sum of the above statistics for the individual mechanisms. Thus, the model for the total backscattering is:

$$\mathbf{C} = \mathbf{C}_s + \mathbf{C}_d + \mathbf{C}_v = \begin{bmatrix} f_s|\beta|^2 + f_d|\alpha|^2 + \frac{3f_v}{8} & 0 & f_s\beta + f_d\alpha + \frac{f_v}{8} \\ 0 & \frac{f_v}{4} & 0 \\ f_s\beta^* + f_d\alpha^* + \frac{f_v}{8} & 0 & f_s2 + f_d + \frac{3f_v}{8} \end{bmatrix}. \quad (21.83)$$

This model gives four equations in five unknowns. However, after having subtracted the volume contribution, a test on the sign of the real part of $E(S_{hh}S_{vv}^*)$ may be used to determine which of the single and double bounce contributions is the dominant. If this term is positive, surface scattering is considered as dominant and the parameter α is set to -1 , whereas β is set to $+1$ otherwise. The Freeman-Durden decomposition is illustrated in Figure 21.15. The Freeman-Durden model-fitting approach is based on physical considerations and not purely on mathematical concepts. Nevertheless it presents some limitations:

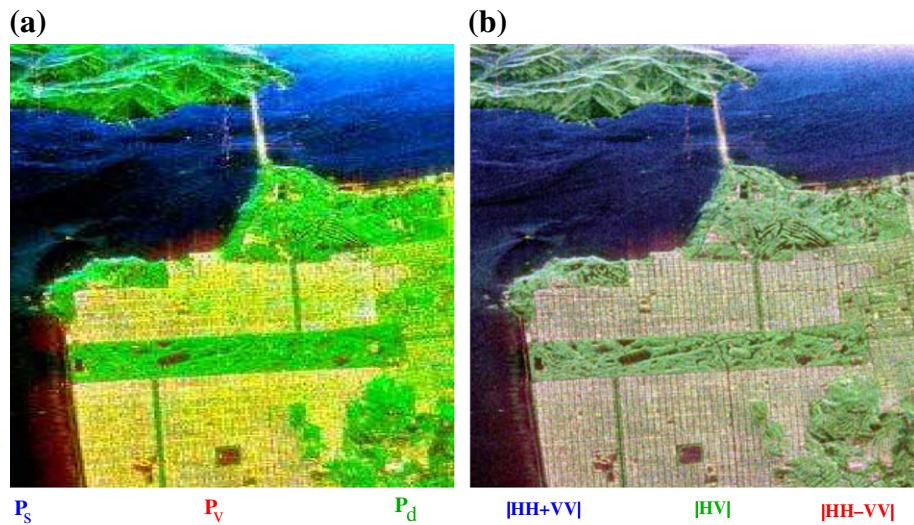
- Reflection symmetry is assumed, i.e., co- and cross-polarization channels are assumed to be uncorrelated.
- Single and double bounce reflections cannot be fully estimated simultaneously, i.e., some of their parameters are fixed

2.21.2.3 A simple polarimetric classification scheme

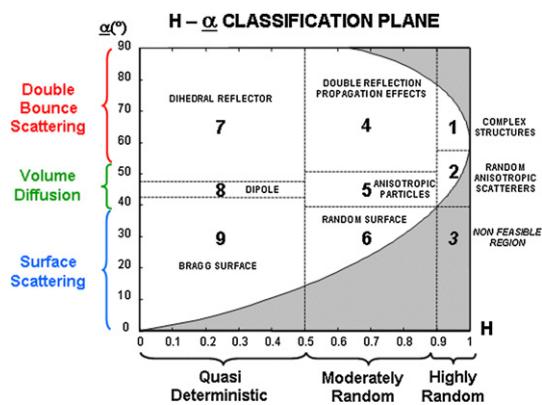
2.21.2.3.1 Unsupervised classification in the 3-D $H/A/\bar{\alpha}$ space

2.21.2.3.1.1 $H/\bar{\alpha}$ classification

In 1997, Cloude and Pottier [13] proposed an unsupervised classification scheme based on the use of the two-dimensional $H/\bar{\alpha}$ plane, where all random scattering mechanisms can be represented. The key idea is that entropy arises as a natural measure of the inherent reversibility of the scattering data and that the $\bar{\alpha}$ angle can be used to identify the underlying average scattering mechanisms. The $H/\bar{\alpha}$ plane is sub-divided into nine basic zones characteristic of classes of different scattering behavior, in order to separate the data into basic scattering mechanisms, as shown in Figure 21.16. The location of the boundaries within the feasible combinations of H and $\bar{\alpha}$ values is set based on the general properties of the scattering mechanisms. There is of course some degree of arbitrariness on the setting of these

**FIGURE 21.15**

Polarimetric color-coded images (a) after the Freeman-Durden decomposition (b) original image in the Pauli basis.

**FIGURE 21.16**

H/α segmentation plane.

boundaries which are not dependent on a particular data set. In Figure 21.16, nine zones are specified, related to specific scattering characteristics that can be measured via the coherency \mathbf{T}_3 matrix:

- Zone 9: Low Entropy Surface Scatter

In this zone occur low entropy scattering processes with $\bar{\alpha}$ values less than 42.5° . These include GO (Geometrical Optics) and PO (Physical Optics) surface scattering—Physical surfaces such as water at L and P-Bands, sea-ice at L-Band, as well as very smooth land surfaces, all fall into this category.

- Zone 8: Low Entropy Dipole Scattering

In this zone occur strongly correlated mechanisms which have a large imbalance between S_{HH} and S_{VV} in amplitude. An isolated dipole scatterer would appear here, as would scattering from vegetation with strongly correlated orientation of anisotropic scattering elements. The width of this zone is determined by the ability of the Radar to measure the S_{HH}/S_{VV} ratio i.e., on the quality of the calibration.

- Zone 7: Low Entropy Multiple Scattering Events

This zone corresponds to low entropy double, or even, bounce scattering events, such as provided by isolated dielectric and metallic dihedral scatterers. These are characterized by $\bar{\alpha}$ values superior to 47.5° .

- Zone 6: Medium Entropy Surface Scatter

This zone reflects the increase in entropy H due to changes in surface roughness and due to canopy propagation effects. In surface scattering theory the entropy H of low frequency theories like Bragg scatter is zero. Likewise, the entropy of high frequency theories like GO is also zero. However, in between these two extremes, there is an increase in entropy H due to the physics of secondary wave propagation and scattering mechanisms. Thus as the roughness/correlation length of a surface changes, its entropy H will increase. Further, a surface cover comprising oblate ellipsoidal scatterers (leaves or disks for example) will generate an entropy $0.6 < H < 0.7$.

- Zone 5: Medium Entropy Vegetation Scattering

Here again we have moderate entropy H but with a dominant dipole type scattering mechanism. The increased entropy H is due to a central statistical distribution of orientation angle. Such a zone would include scattering from vegetated surfaces with anisotropic scatterers and moderate correlation of scatterer orientations.

- Zone 4: Medium Entropy Multiple Scattering

This zone accounts for dihedral scattering with moderate entropy H . This occurs for example in forestry applications, where double bounce mechanisms occur at P and L bands following propagation through a canopy. The effect of the canopy is to increase the entropy H of the scattering process. This class is typical for urban areas, where dense packing of localized scattering centers can generate moderate entropy H with low order multiple scattering dominant. The boundary between zones 4, 5, 6, and 1, 2, 3, is set as $H = 0.9$. This is chosen on the basis of the upper limit for surface, volume, and dihedral scattering before random distributions apply.

- Zone 3: High Entropy Surface Scatter

This class is a non-feasible region in the $H/\bar{\alpha}$ plane.

- Zone 2: High Entropy Vegetation Scattering

High entropy volume scattering arises when $\bar{\alpha} = 45^\circ$ and $H > 0.9$. This can arise for single scattering from a cloud of anisotropic needle like particles or from multiple scattering from a cloud of low

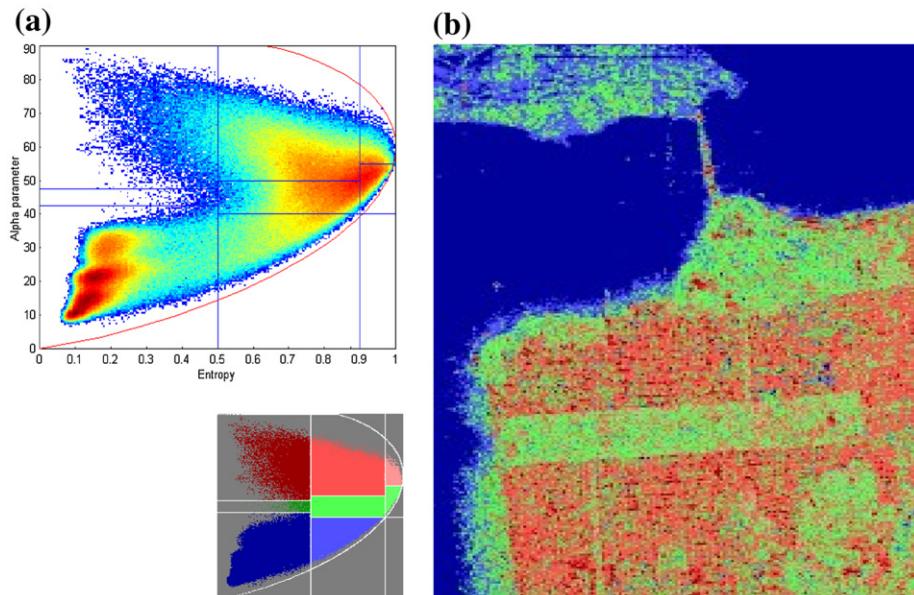


FIGURE 21.17

Classification of the San Francisco bay image in the $H/\bar{\alpha}$ plane (a) pixel density (b) classification results.

loss symmetric particles. In both cases, however, the entropy H lies above 0.9, where the feasible region of $H/\bar{\alpha}$ plane is rapidly shrinking. Scattering from forest canopies lies in this region, as does the scattering from some types of vegetated surfaces with random highly anisotropic scattering elements. The extreme behavior in this class is random noise i.e., no polarization dependence, a point which lies at the extreme right of Zone 2.

- Zone 1: High Entropy Multiple Scattering

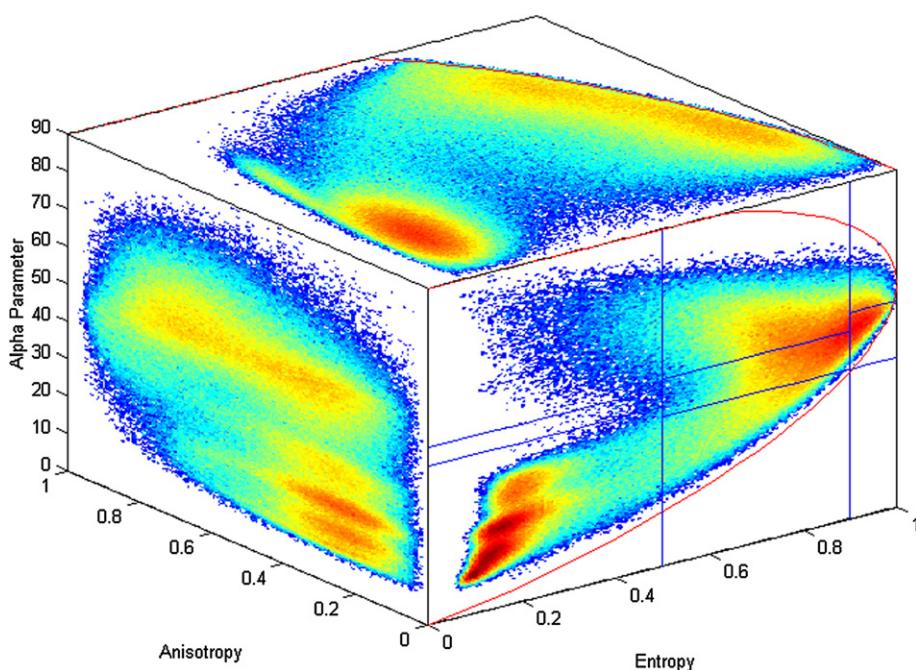
In the $H > 0.9$ region, it is still possible to distinguish double bounce mechanisms in a high entropy environment. Again such mechanisms can be observed in forestry applications or over vegetation having well developed branch and crown structure.

The distribution of the San Francisco Bay PolSAR data on the $H/\bar{\alpha}$ plane is shown in Figure 21.17.

This segmentation of the $H/\bar{\alpha}$ plane is offered merely to illustrate a simple unsupervised classification strategy and to emphasize the geometrical segmentation of physical scattering processes. The corresponding result is shown in Figure 21.17. It is this key feature which makes this an unsupervised, measurement-data-independent approach to the scatter feature classification problem.

2.21.2.3.1.2 Including the polarimetric anisotropy A

Inherent of the spatial averaging, the entropy H may increase, and the number of distinguishable classes identifiable from polarimetric observations is reduced. For example, the feasible region of the $H/\bar{\alpha}$ plane is rapidly shrinking for high values of entropy ($H > 0.7$), where parameter reaches the limited value of 60° .

**FIGURE 21.18**

Pixel density in the $H/A/\bar{\alpha}$ space over the San Francisco bay image.

This remark is confirmed by the analysis of the distribution of the San Francisco Bay PolSAR data in the extended and complemented three-dimensional $H/A/\bar{\alpha}$ space, as shown in Figure 21.18. This representation shows that it is possible to discriminate new classes using the anisotropy value.

For example, it is now possible to notice that there exists in the Low Entropy Surface Scattering area (Z9) a second class associated with a high anisotropy value and which corresponds to the presence of a second physical mechanism which is not negligible.

Identical remarks can be made concerning the Medium Entropy Vegetation Scattering area (Z5) and the Medium Entropy Multiple Scattering area (Z4). Due to the spread of the PolSAR data along the anisotropy axis, it is now possible to improve the capability to distinguish different types of scattering process which have quite the same high entropy value:

- High entropy and low anisotropy correspond to random scattering.
- High entropy and high anisotropy correspond to the presence of two scattering mechanisms with the same probability.

It is thus possible to sub-divide each plane of the $H/A/\bar{\alpha}$ space into basic zones characteristic of classes of different scattering behavior, in order to separate the data into basic scattering mechanisms. There still exists some degree of arbitrariness on the setting of these boundaries which are not dependent on a particular data set.

The corresponding result is shown in Figure 21.18 for each plane of the $H/A/\bar{\alpha}$ space.

In order to extend the classification scheme and to improve the capability to distinguish different types of scattering process, it is proposed to use some combinations between entropy (H) and anisotropy (A) information, as shown in Figure 21.19. The (\cdot) operation represents the element by element multiplication of two matrices.

The examination of the different figures corresponding to the different combinations between entropy (H) and anisotropy (A) images leads to the following interesting remarks:

1. The $(1 - H)(1 - A)$ image corresponds to the presence of a single dominant scattering process (low entropy and low anisotropy with $\lambda_2 \approx \lambda_3 \approx 0$).
2. The $H(1 - A)$ image characterizes a random scattering process (high entropy and low anisotropy with $\lambda_2 \approx \lambda_3 \approx \lambda_1$).
3. The HA image relates to the presence of two scattering mechanisms with the same probability (high entropy and high anisotropy with $\lambda_3 \approx 0$).
4. The $(1 - H)A$ image corresponds to the presence of two scattering mechanisms with a dominant process (low to medium entropy) and a second one with medium probability (high anisotropy with $\lambda_3 \approx 0$).

From the analysis of the different images shown in Figure 21.20 and from the distribution of the San Francisco Bay PolSAR data in the $H/A/\bar{\alpha}$ classification space shown in Figure 21.18, it can be concluded that these three parameters have to be considered now as key parameters in the polarimetric analysis and/or inversion of PolSAR data.

The information contained in these three roll-invariant parameters extracted from the local estimate of the coherency \mathbf{T} matrix, corresponds to the type of scattering process which occurs within the pixel to be classified (combination of entropy H and anisotropy A) and to the corresponding physical scattering mechanism ($\bar{\alpha}$ parameter).

2.21.2.3.2 Unsupervised statistical segmentation using Wishart statistics

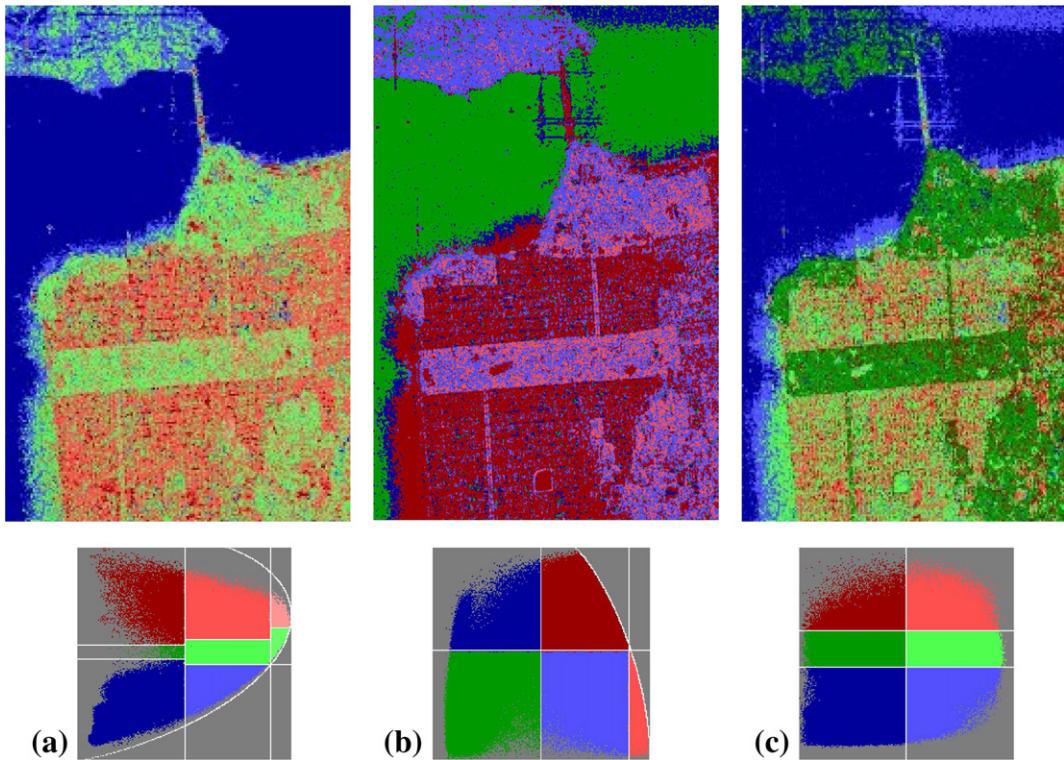
2.21.2.3.2.1 Principles of the k-mean segmentation technique

The segmentation of a SAR image over a fixed number of clusters aims to assign, in an optimal way, a SAR image pixel, p , to one of the M possible clusters, $\{C_1, \dots, C_M\}$ according to a SAR observable, \mathbf{x} . A solution is provided by Bayes optimal decision rule using the a posteriori probability of the different clusters

$$\text{Decide } p \in C_i \text{ if } P(C_i|\mathbf{x}) > P(C_j|\mathbf{x}) \quad \forall j \neq i. \quad (21.84)$$

A pixel is assigned to the most probable cluster conditionally to the observation of \mathbf{x} over the pixel under consideration. The risk, or probability of error, related to this decision is obtained from the a posteriori probabilities of the unselected clusters:

$$P(\text{error}|\mathbf{x}) = \sum_{j \neq i} P(C_j|\mathbf{x}). \quad (21.85)$$

**FIGURE 21.19**

Classification results in the (a) $H/\bar{\alpha}$ (b) H/A (c) $A/\bar{\alpha}$ planes over the San Francisco bay image.

It is obvious, that assigning a pixel to the cluster with the highest a posteriori probability minimizes the conditional error probability. The total error probability is then computed using

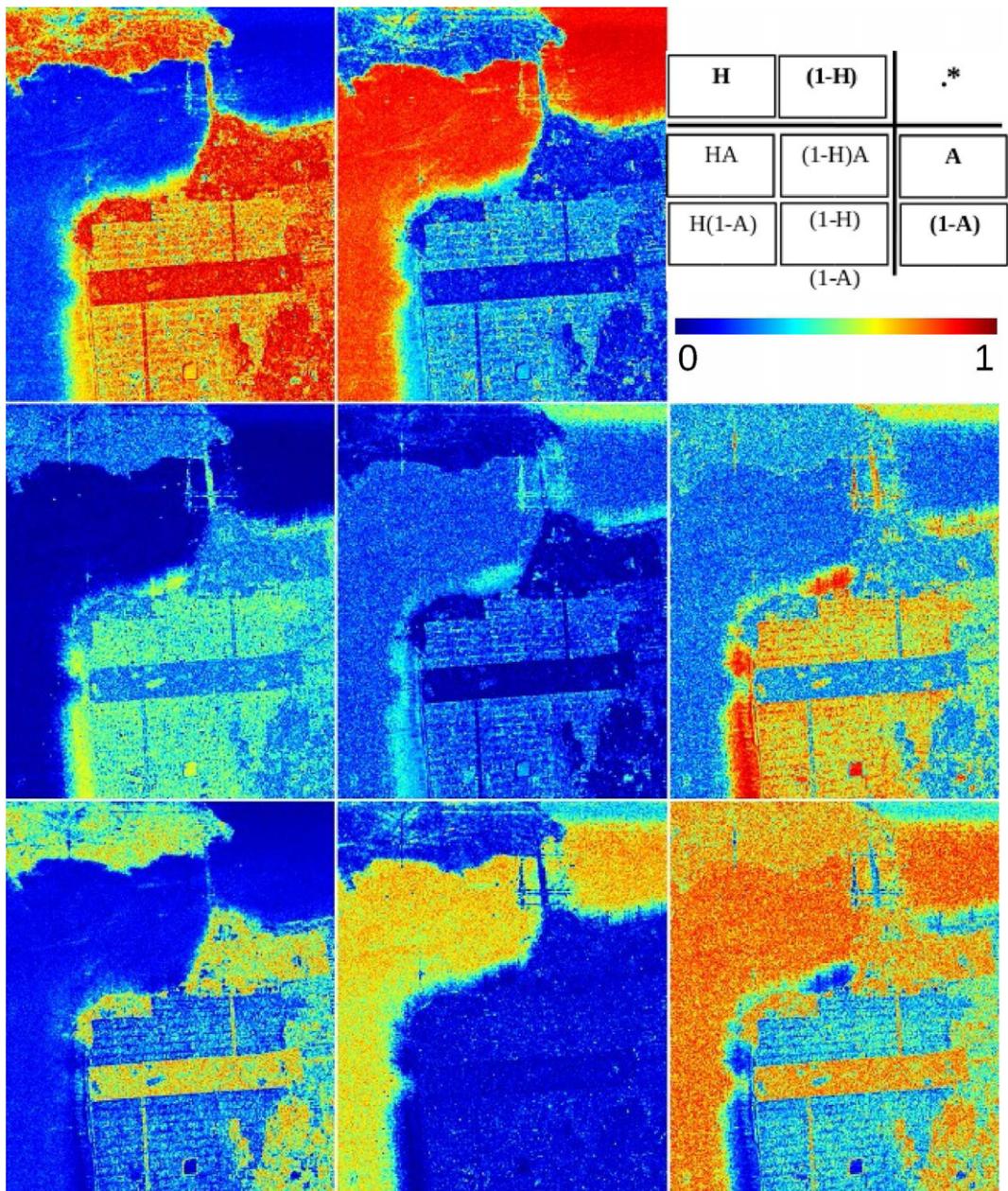
$$P(\text{error}) = \int P(\text{error}|\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \quad (21.86)$$

The set of a posteriori probabilities generally being uneasy to derive, one prefers then to express such quantities using the Bayes rule

$$P(C_i|\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) = f_{\mathbf{x}}(\mathbf{x}|C_i) P(C_i), \quad (21.87)$$

where $f_{\mathbf{x}}(\mathbf{x}|C_i)$ is the likelihood of cluster C_i with respect to \mathbf{x} . Inserting this expression in (21.84), the expression of the optimal decision rule becomes

$$\text{Decide } p \in C_i \text{ if } f_{\mathbf{x}}(\mathbf{x}|C_i) P(C_i) > f_{\mathbf{x}}(\mathbf{x}|C_j) P(C_j) \quad \forall j \neq i. \quad (21.88)$$

**FIGURE 21.20**

H and A combinations.

Additionally, if the prior probabilities are unknown, or supposed to be equal, (21.88) may written as a maximum likelihood decision rule:

$$\text{Decide } p \in C_i \text{ if } f_{\mathbf{x}}(\mathbf{x}|C_i) > f_{\mathbf{x}}(\mathbf{x}|C_j) \quad \forall j \neq i. \quad (21.89)$$

The likelihood of a cluster, $f_{\mathbf{x}}(\mathbf{x}|C_i)$, can be estimated from the statistical parameters of the samples it contains, and an optimal solution to the unsupervised segmentation problem consists in distributing the pixels of an image over the set of M clusters so as to maximize the global likelihood defined as the product of all individual likelihood functions. A rigorous resolution requires to test all the possible combinations and to select the one corresponding to the maximum joint likelihood value. This optimal solution cannot be applied due to the unrealistic computational load it involves. Alternative solutions based on sub-optimal iterative optimization procedures are generally preferred. The k-means procedure is an iterative optimization algorithm described by the synopsis of Figure 21.21. The algorithm begins with the initialization of the image pixel distribution over the N clusters. This distribution may be done in a random way or according to user specifications. Once all pixels are assigned, the different cluster centers are computed according to the processed data type. Each pixel is then assigned to the closest cluster according to a distance measure. The convergence of the algorithm is then tested using stability metrics. If a termination criterion is met, the segmentation stops, otherwise a new iteration starts over from the class center computation step. The k-means algorithm aims to optimize a global function by iteratively optimizing local expressions. It is known that this type of techniques may get stuck into locally stable states and fail to determine the optimal pixel distribution. The initialization of

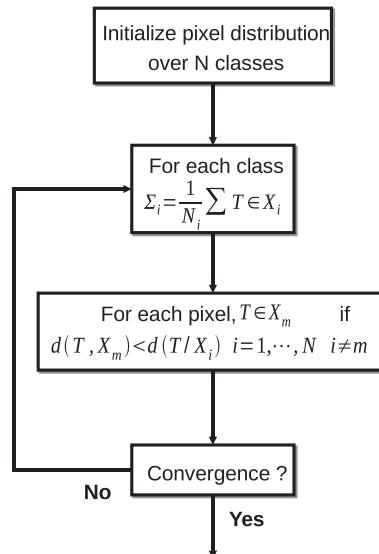


FIGURE 21.21

Synopsis of the k-means algorithm.

the pixel distribution into N classes is then a critical stage of such a segmentation algorithm. An adequate initialization permits a fast convergence and provides correctly segmented clusters. The convergence of the algorithm is evaluated by testing a condition of termination, that may be based on an estimate of the classification quality, or consist in a maximum number of iterations or in a sufficiently low number of pixels that are differently classified from one iteration to the other.

2.21.2.3.2.2 Unsupervised $H/A/\bar{\alpha}$ segmentation

The k-means clustering algorithm may be adapted to the segmentation of \mathbf{T} or \mathbf{C} polarimetric representations using the Wishart pdf of L-look sample matrix estimates expressed in (21.54):

$$f_{\widehat{\mathbf{T}}}(\widehat{\mathbf{T}}|\mathbf{T}_i) = \frac{L^{3L} |\widehat{\mathbf{T}}|^{L-2}}{\tilde{\Gamma}_2(L) |\mathbf{T}_i|^L} e^{-L \text{tr}(\mathbf{T}_i^{-1} \widehat{\mathbf{T}})} \quad (21.90)$$

with

$$\widehat{\mathbf{T}} = \frac{1}{L} \sum_{l=1}^L \mathbf{k}(l) \mathbf{k}(l)^\dagger \quad \text{and} \quad \mathbf{T}_i = \frac{1}{N_i} \sum_{n=1}^{N_i} \widehat{\mathbf{T}}(n) \in C_i. \quad (21.91)$$

The ML decision rule may be simplified as:

$$f_{\widehat{\mathbf{T}}}(\widehat{\mathbf{T}}|\mathbf{T}_i) > f_{\widehat{\mathbf{T}}}(\widehat{\mathbf{T}}|\mathbf{T}_j) \equiv d(\widehat{\mathbf{T}}, \mathbf{T}_i) < d(\widehat{\mathbf{T}}, \mathbf{T}_j) \quad (21.92)$$

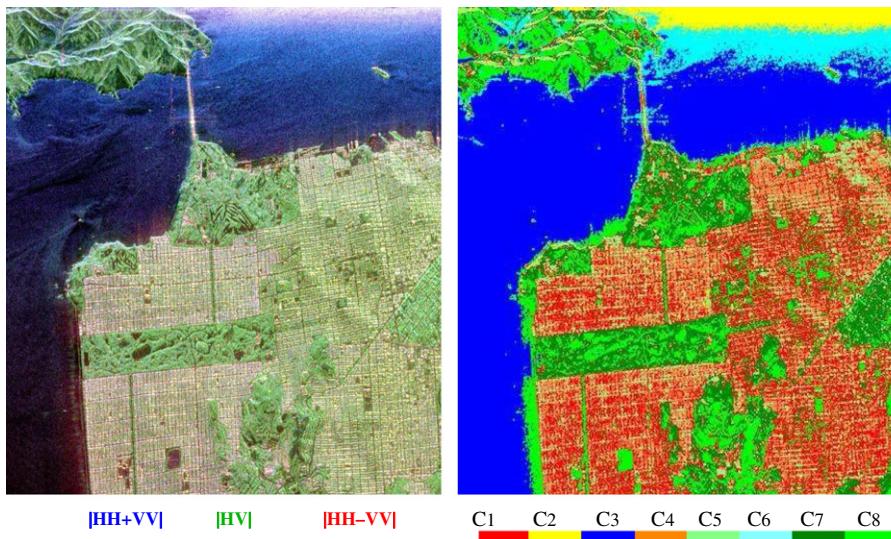
with $d(\widehat{\mathbf{T}}, \mathbf{T}_i)$ a statistical proximity, and not distance, between the sample matrix $\widehat{\mathbf{T}}$ and the estimate of the class center coherency matrix \mathbf{T}_i , defined as

$$d(\widehat{\mathbf{T}}, \mathbf{T}_i) = \log |\mathbf{T}_i| + \text{tr}(\mathbf{T}_i^{-1} \widehat{\mathbf{T}}). \quad (21.93)$$

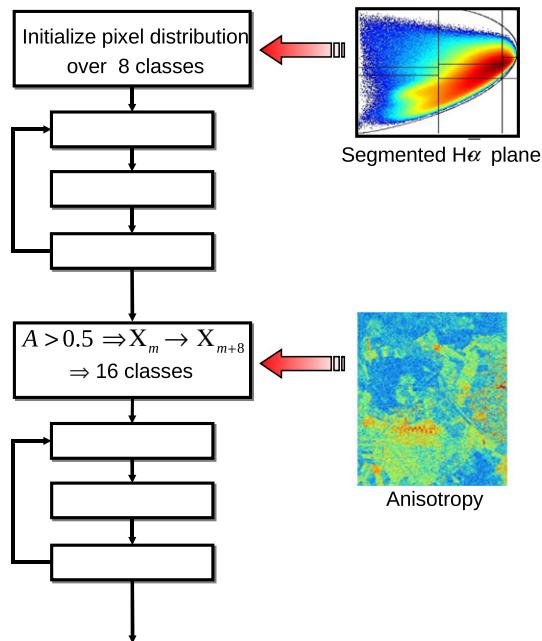
The particularity of the identification procedure based on H and $\bar{\alpha}$, introduced earlier, resides in the estimation of the type of observed media from a physical interpretation of canonical scattering mechanisms using robust indicators. Nevertheless, the analysis of natural scenes using this unsupervised approach may reach some limitations:

- The arbitrarily fixed linear decision boundaries in the $H - \bar{\alpha}$ plane may not fit data distribution. A natural cluster corresponding to similar targets may lie across a frontier in the decision plane. In this case, pixels with very similar characteristics may be assigned, in an almost random way, to different classes due slightly different locations in the $H - \bar{\alpha}$.
- Even if the computation of H and $\bar{\alpha}$ requires fully polarimetric data, these two parameters do not represent the whole polarimetric information. The use of other indicators such as the Span or specific correlations coefficients may improve the classification results in a significant way.

Segmentation procedures based on the whole coherency matrix statistics permit to overcome the limitations mentioned above. Nevertheless, it is shown in the following, that the physical interpretation of the scattering phenomenon permits to enhance in a significant way the performance of statistical segmentation schemes. The unsupervised segmentation scheme mentioned above is initialized in an efficient way with the results of the unsupervised identification of a scattering mechanism, using H and $\bar{\alpha}$. This initialization provides 8 stable initial clusters relating to the underlying physical scattering

**FIGURE 21.22**

Wishart $H - \bar{\alpha}$ segmentation results into eight classes.

**FIGURE 21.23**

Synopsis of the $H - A - \bar{\alpha}$ segmentation.

mechanisms. An important improvement in the segmentation accuracy can be observed in the image presented in Figure 21.22. The main kinds of natural media are clearly discriminated by the Wishart $H/\bar{\alpha}$ segmentation scheme. This unsupervised classification algorithm modifies the decision boundaries in an adaptive way to better fit the natural distribution of the scattering mechanisms and takes into account information related to the back-scattered power. The ML Wishart segmentation may be further improved by explicitly including the anisotropy information during the segmentation procedure. As mentioned previously, the anisotropy indicates the relative importance of secondary mechanisms obtained from the expansion of a coherency matrix. This polarimetric indicator is particularly useful to discriminate scattering mechanisms with different eigenvalue distributions but with similar intermediate entropy values. In such cases, a high anisotropy value indicates two dominant scattering mechanisms with equal probability and a less significant third mechanism, while a low anisotropy value corresponds to a dominant first scattering mechanism and two non-negligible secondary mechanisms with equal importance. Among the different approaches tested, the best way to introduce the anisotropy information in the classification algorithm consists in implementing two successive segmentation procedures as shown in Figure 21.23.

Polarimetric data are first segmented according to the algorithm presented in the former paragraph. Once this procedure has converged, the eight resulting clusters are split into 16 ones by comparing the anisotropy of each pixel to a threshold fixed to 0.5. The 16 segments are then used to initialize a second Wishart ML segmentation procedure. The segmentation results presented in Figure 21.24 show an enhanced description of the San Francisco scene. The introduction of the anisotropy in the clustering

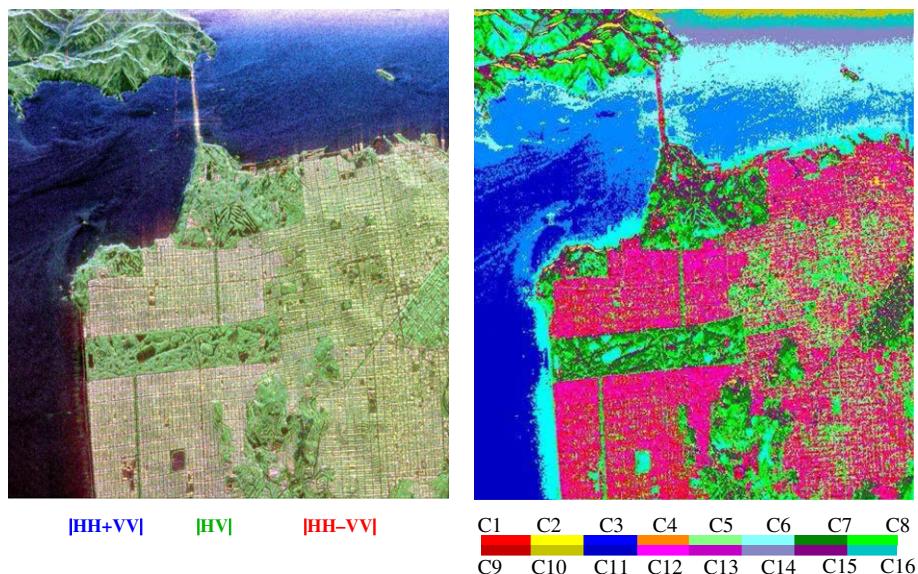


FIGURE 21.24

Wishart $H - A - \bar{\alpha}$ segmentation results into 16 classes.

process permits to split large segments into smaller clusters discriminating small disparities in a refined way. The Wishart $H/A/\bar{\alpha}$ classification scheme gathers into segments pixels with similar statistical properties, but does not provide any information concerning the nature of the scattering mechanism associated to each cluster.

2.21.3 Advanced polarimetric decomposition techniques and some of their applications

2.21.3.1 Advanced decomposition techniques

2.21.3.1.1 The Yamaguchi 4-component decomposition

As seen previously, the 3-component scattering power model proposed by Freeman and Durden [9] can be successfully applied to decompose SAR observations under the reflection symmetry condition. However, it can be possible to find some areas in a SAR image for which the reflection symmetry condition does not hold. Based on the 3-component scattering model approach, Yamaguchi et al. proposed a 4-component scattering model by introducing an additional term corresponding to non-reflection symmetric cases [25]. In order to accommodate the decomposition scheme for the more general scattering case encountered in complicated geometric scattering structures, the fourth component introduced is equivalent to a helix scattering power. This helix scattering power term, that corresponds to responses having correlated co- and cross-polarization channels, and appears over heterogeneous areas (complicated shape targets or man-made structures) whereas it is quasi null over almost all natural environments. The scattering matrices, corresponding to a left-helix target or to a right-helix target, have the form:

$$\mathbf{S}_{LH} = \frac{1}{2} \begin{bmatrix} 1 & j \\ j & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{RH} = \frac{1}{2} \begin{bmatrix} 1 & -j \\ -j & -1 \end{bmatrix} \quad (21.94)$$

and the corresponding covariance matrices are given by:

$$\mathbf{C}_{LH} = \frac{f_c}{4} \begin{bmatrix} 1 & -j\sqrt{2} & -1 \\ j\sqrt{2} & 2 & -j\sqrt{2} \\ -1 & j\sqrt{2} & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{RH} = \frac{f_c}{4} \begin{bmatrix} 1 & j\sqrt{2} & -1 \\ -j\sqrt{2} & 2 & j\sqrt{2} \\ -1 & -j\sqrt{2} & -1 \end{bmatrix}, \quad (21.95)$$

where f_c corresponds to the contribution of the helix scattering component.

The second important contribution proposed by Yamaguchi et al. in the 4-component decomposition model approach, concerns the modification of the volume scattering matrix in the decomposition according to the relative backscattering magnitudes of $E(|S_{hh}|^2)$ versus $E(|S_{vv}|^2)$ [25]. In the theoretical modeling of volume scattering, a cloud of randomly oriented dipoles is implemented with a probability function being uniform for the orientation angles. However, for vegetated areas where vertical structure seems to be rather dominant, the scattering from tree trunks and branches displays a non-uniform angle distribution. The proposed new angle pdf is given by:

$$f_\phi(\phi) = \frac{\cos \phi}{2} \text{ for } |\theta| < \frac{\pi}{2} \quad \text{and null otherwise.} \quad (21.96)$$

Assuming a cloud of randomly oriented, very thin horizontal cylinder-like scatterers, i.e., $b \rightarrow 0$ in (21.82), the volume scattering covariance matrix is given by:

$$\mathbf{C}_{vh} = \frac{f_v}{15} \begin{bmatrix} 8 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 3 \end{bmatrix}. \quad (21.97)$$

Assuming now a cloud of randomly oriented, very thin vertical cylinder-like scatterers, i.e., $a \rightarrow 0$ in (21.82), the volume scattering covariance matrix is given by:

$$\mathbf{C}_{vv} = \frac{f_v}{15} \begin{bmatrix} 3 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 8 \end{bmatrix}. \quad (21.98)$$

The asymmetric form of the two volume scattering covariance matrices is of considerable use since it can be adjusted to the measured data according to the ratio $CR = 10 \log_{10} (\mathbb{E}(|S_{hh}|^2)/\mathbb{E}(|S_{vv}|^2))$. Depending on the scene co-polarization ratio, the appropriate volume scattering covariance matrix is chosen according to the following rules

$$\begin{aligned} CR < -2dB &\Rightarrow \mathbf{C}_{vh}, \\ -2dB \leq CR < +2dB &\Rightarrow \mathbf{C}_{vFreeman}, \\ +2dB \leq CR &\Rightarrow \mathbf{C}_{vv}, \end{aligned} \quad (21.99)$$

Assuming that the volume, double-bounce, surface, and helix scatter components are uncorrelated, the total second-order statistics are the sum of the above statistics for the individual mechanisms. Thus, the model for the total backscatter is:

$$\mathbf{C} = \mathbf{C}_s + \mathbf{C}_d + \mathbf{C}_{lh/rh} + \mathbf{C}_v \quad (21.100)$$

Figure 21.25 shows the algorithm for the four-component scattering power decomposition. The Yamaguchi decomposition is intended to apply to non-reflection symmetry case, the scheme automatically includes the reflection symmetry condition, thus proposing a decomposition scheme for the more general scattering case encountered in complicated geometric scattering structures. The Yamaguchi decomposition is illustrated in Figure 21.26.

2.21.3.1.2 The Freeman 2-component decomposition

In 2007, Freeman has proposed a new and original 2-component scattering model to polarimetric SAR observations of forests [26]. The selected mechanisms are canopy scatter from a reciprocal medium with reflection symmetry and a ground scatter term representing either a double-bounce scatter from a pair of orthogonal surfaces with different dielectric constants (ground-trunk interaction) or a Bragg scatter from a moderately rough surface, which is seen through a layer of vertically oriented scatterers [26].

The volume scattering from a forest canopy is modeled as the contribution from a cloud of randomly oriented cylinder-like scatterers. The second-order statistics covariance matrix \mathbf{C}_v for scatterers from a

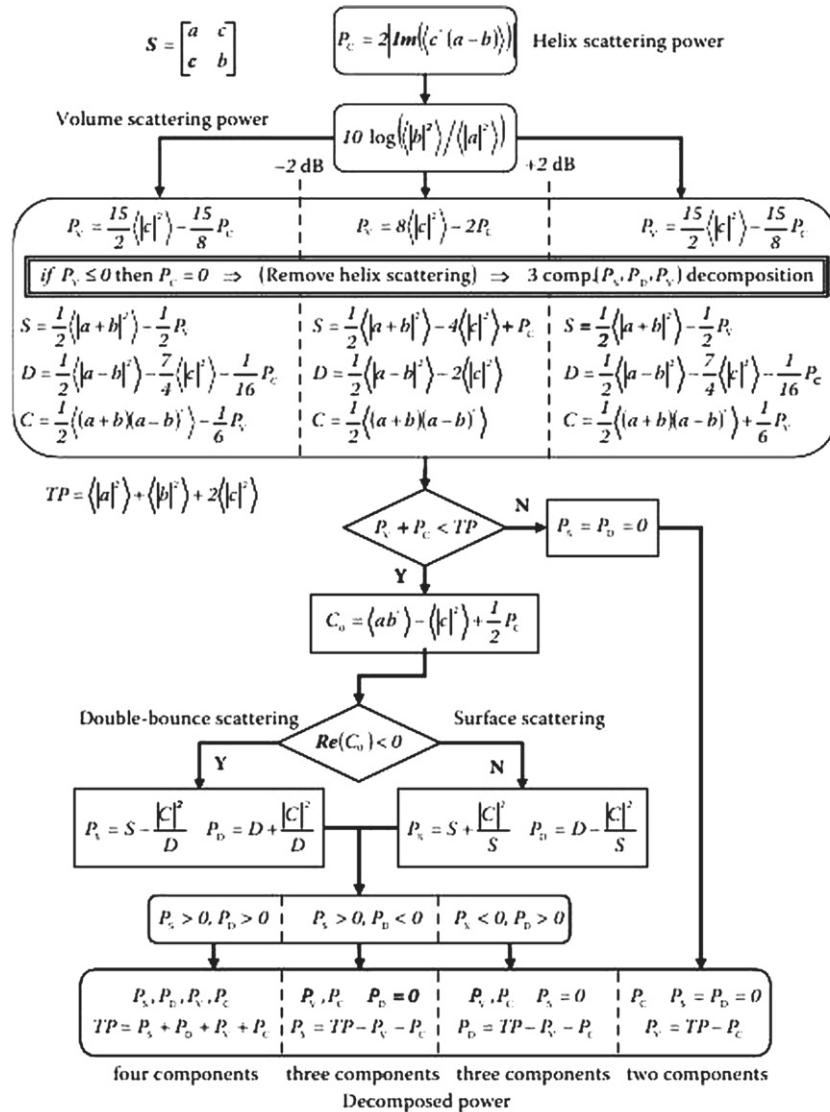
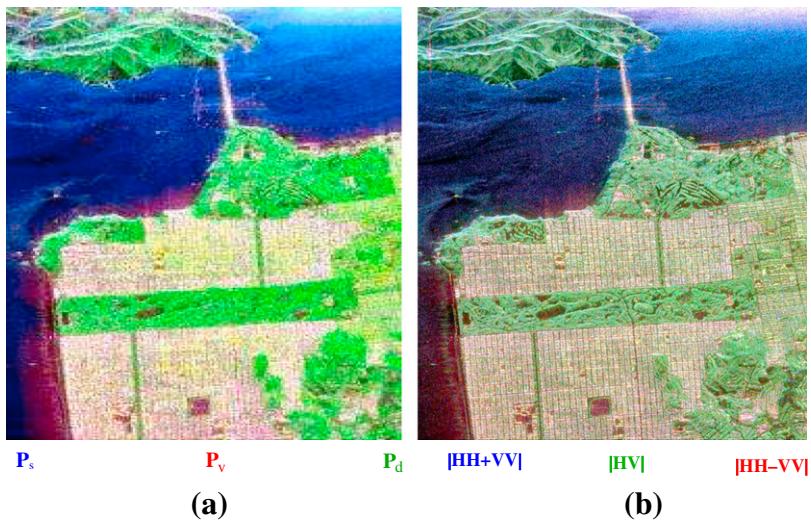


FIGURE 21.25

Algorithm of the four-component decomposition (Courtesy of Prof. Yoshio Yamaguchi).

reciprocal medium with reflection symmetry is given by:

$$\mathbf{C}_v = f_v \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1-\rho & 0 \\ \rho & 0 & 1 \end{bmatrix}, \quad (21.101)$$

**FIGURE 21.26**

Polarimetric color-coded images (a) after the Yamaguchi decomposition and (b) original image in the Pauli basis.

where $\rho \leq 1$. The second scattering mechanism, called ground, corresponds to double-bounce scattering or surface scattering or a combination of both. The resulting covariance matrix is given by:

$$\mathbf{C}_g = f_g \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 0 & 0 \\ \alpha^* & 0 & |\alpha|^2 \end{bmatrix}. \quad (21.102)$$

The model for the total backscattering is:

$$\mathbf{C} = \mathbf{C}_g + \mathbf{C}_v. \quad (21.103)$$

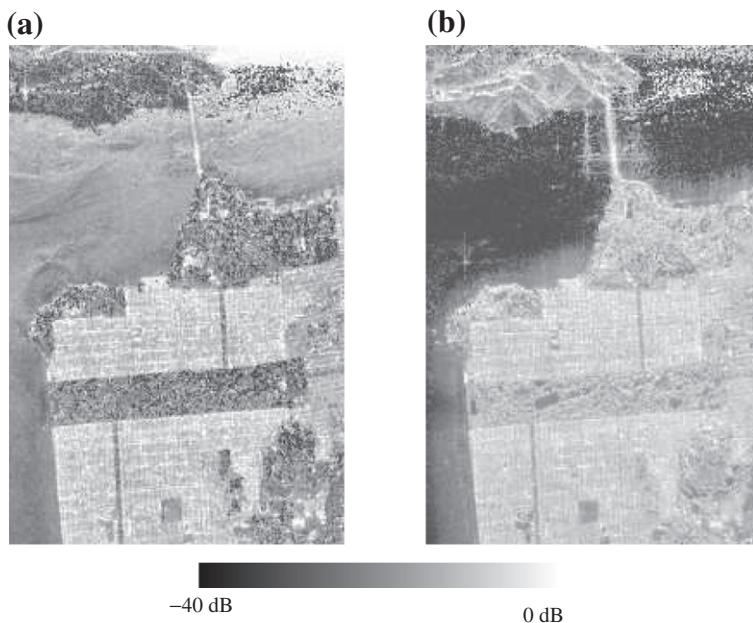
In contrast to the Freeman-Durden 3-component decomposition, the new Freeman 2-component decomposition presents an equal number of input and output parameters and thus can be easily solved without any a priori assumption [26]. The Freeman 2-component decomposition is illustrated in Figure 21.27.

2.21.3.1.3 A 2-Component decomposition with improved volume description

This decomposition, proposed in [27], is based on an improved description of volume scattering and is included in a global Pol-inSAR (SAR polarimetric interferometry) characterization approach.

Similarly to the Freeman and Durden, Freeman II and Yamaguchi decompositions, a volume is described as a cloud of anisotropic particles, whose scattering matrix is given as:

$$\mathbf{S} = \mathbf{R}^T(\psi) \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mathbf{R}(\psi) = A \mathbf{R}^T(\psi) \begin{bmatrix} 1 + \delta^* & 0 \\ 0 & 1 - \delta^* \end{bmatrix} \mathbf{R}(\psi), \quad (21.104)$$

**FIGURE 21.27**

Images of the span of the (a) Ground (b) Double Bounce contributions of the Freeman decomposition.

where $\mathbf{R}(\psi)$ is a $SU(2)$ operator accounting for a rotation around the radar line of sight and δ is the particle scattering anisotropy, defined as:

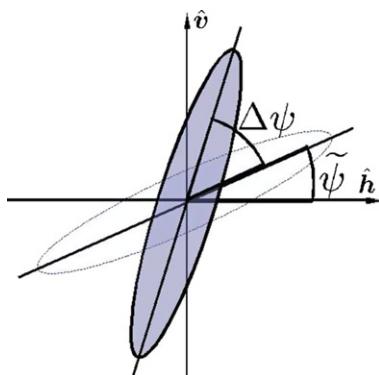
$$\delta = \left(\frac{a - b}{a + b} \right)^*. \quad (21.105)$$

The particle orientation angles ψ in the polarization plane shown in Figure 21.28 are assumed to follow a unimodal circular distribution and to be independent from other vegetation characteristics. Under the central limit theorem condition, given a large number of scatterers, the orientations of these scatterers are normally distributed and follow the circular normal distribution (also known as the von Mises distribution) which is the circular analog of the Gaussian distribution

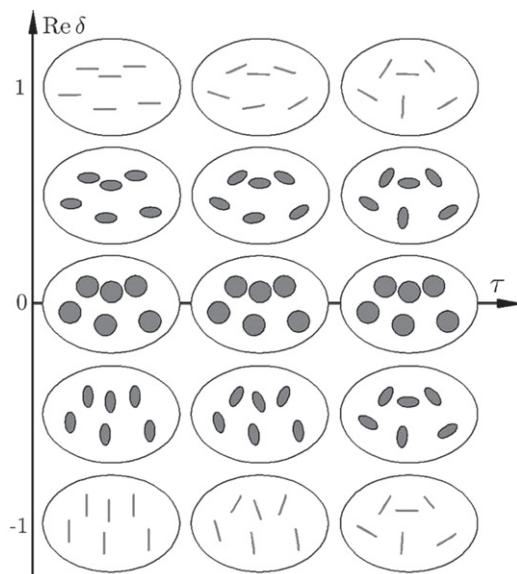
$$f_\psi(\psi) = \frac{e^\kappa \cos(2(\psi - \tilde{\psi}))}{\pi I_0(\kappa)}, \quad (21.106)$$

where κ is the degree of concentration, analogous to the inverse of the standard deviation, $\tilde{\psi} \in [-\pi/2, \pi/2]$, is the mean particle orientation angle, and $I_0()$ is the modified Bessel function of order 0. For the sake of interpretation, the normalized degree of orientation randomness τ is introduced as

$$\tau = I_0(\kappa)e^{-\kappa}. \quad (21.107)$$

**FIGURE 21.28**

Particle orientation: mean orientation $\tilde{\psi}$ and the orientation of an individual particle $\psi = \tilde{\psi} + \Delta\psi$.

**FIGURE 21.29**

Schematic representation of the effective shapes and orientations of particles in the LOS for different degrees of particle scattering anisotropy δ and orientation randomness τ .

The physical meaning of δ and τ is illustrated on Figure 21.29

The particle scattering anisotropy, δ , characterizes the effective shape of the average particle in dependence of the particle and background permittivities and tilt angle distribution [27]. Assuming

simple spheroidal particles, one can relate δ to the effective particle shapes: As $|\delta| \rightarrow 0$, the average effective particle shape approaches an isotropic sphere/disk, whereas, as $|\delta| \rightarrow 1$, the effective shape tends toward a dipole. If the phases of the scattering coefficients a and b are similar, then δ is a function only of their moduli. Then, in the line-of-sight (LOS) direction, the particle axis of symmetry tends to be horizontal if $\Re(\delta) > 0$ and vertical if $\Re(\delta) < 0$, with respect to the polarization basis of the particle scattering amplitude matrix. As $\tau \rightarrow 0$, the volume becomes strongly aligned in the preferred orientation direction, whereas, for $\tau \rightarrow 1$, the particle orientations become completely random. The degree of orientation randomness becomes meaningless for (effectively) isotropic scatterers, i.e., when $\delta = 0$.

The effective volume coherency matrix is the given by

$$\mathbf{T}_v = \mathbf{R}(\tilde{\psi}) \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} f_\psi(\psi) \mathbf{T}_v d\psi \mathbf{R}(\psi)^\dagger \quad (21.108)$$

which may be rewritten as [27]

$$\mathbf{T}_v = \mathbf{R}(\tilde{\psi}) \begin{bmatrix} 1 & g_c \delta & 0 \\ g_c \delta^* & \frac{1+g}{2} |\delta|^2 & 0 \\ 0 & 0 & \frac{1-g}{2} |\delta|^2 \end{bmatrix} \mathbf{R}(\psi)^\dagger, \quad (21.109)$$

where $g = I_2(\kappa)/I_0(\kappa)$ and $g_c = I_1(\kappa)/I_0(\kappa)$. The ground contribution is given by a general reflection symmetric matrix

$$\mathbf{T}_g = \begin{bmatrix} 1 & \beta & 0 \\ \beta^* & \beta_{22} & 0 \\ 0 & 0 & \beta_{33} \end{bmatrix} \quad (21.110)$$

and the global response is modeled as

$$\mathbf{T} = f_v \mathbf{T}_v + f_g \mathbf{T}_g. \quad (21.111)$$

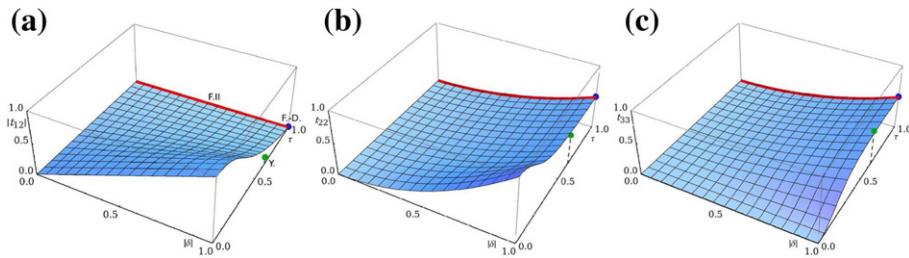
The generality of this approach compared to Freeman and Yamaguchi ones and illustrated in Figure 21.30 is related to the very high number of parameters used to describe the ground and volumes contributions. Such an approach can be unambiguously applied, using the NNED approach [28] described in next section, or when the volume and ground component can be separated using multiple polarimetric and interferometric data sets for instance [27].

2.21.3.1.4 Another adaptive model-based decomposition

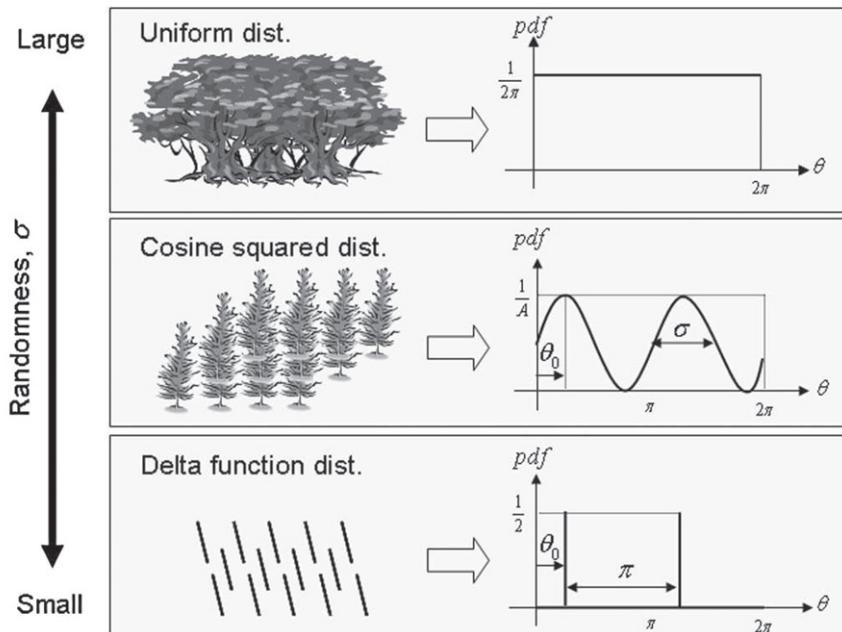
Similarly to the former approach, this decomposition technique, presented in [29], describes a natural volume as a cloud of anisotropic scatterers. The distribution chosen to model the orientation of the particles is a n th order squared cosine distribution:

$$f_\psi(\psi, \tilde{\psi}, n) = \frac{\cos^{2n}(\psi - \tilde{\psi})}{\int_0^{2\pi} \cos^{2n}(\psi - \tilde{\psi})}. \quad (21.112)$$

As illustrated in Figure 21.31, such a function may be used to represent distributions ranging from deterministic, with $n \rightarrow +\infty$, to uniform, $n = 0$ [29, 30]. Similarly to the former decomposition, an

**FIGURE 21.30**

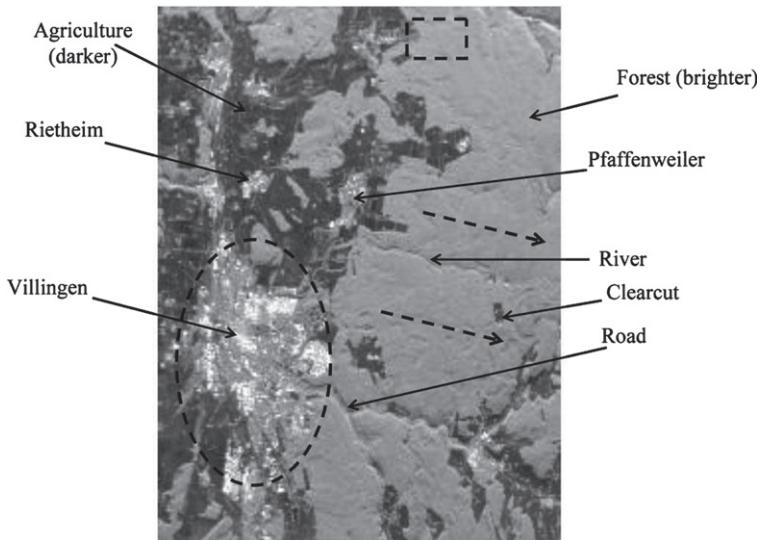
Normalized volume coherency matrix elements (a) $|T(1, 2)|$, (b) $|T(2, 2)|$, (c) $|T(3, 3)|$ as functions of the degree of orientation randomness τ and the particle scattering anisotropy $|\delta|$. The blue point, the red line, and the green point correspond to the parameter ranges of the Freeman Durden, the Freeman II, and the Yamaguchi models, respectively (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

**FIGURE 21.31**

Various pdfs in terms of randomness that can be modeled by a squared cosine distribution.

Indicator of the degree of randomness of the particle orientation can be derived as:

$$\tau(n) = \sqrt{\frac{\pi^2}{12} + \sum_{k=0}^{n-1} \frac{n! n!}{k!(2n-k)!} \frac{(-1)^{n-k}}{n-k}}. \quad (21.113)$$

**FIGURE 21.32**

NASA/JPL AIRSAR L-band image over the Black Forest, dotted lines specify the direction of topographic change. A urban area is outlined by the dotted circle. From [29].

Using $f_\psi(\psi, \tilde{\psi}, n)$, the covariance matrix of an anisotropic dipole with mean orientation $\tilde{\psi}$ is given by

$$\mathbf{C}_v(\tilde{\psi}, n) = \mathbf{C}_\alpha + \frac{2n}{n+1} \mathbf{C}_\beta(2\tilde{\psi}) + \frac{n(n-1)}{(n+1)(n+2)} \mathbf{C}_\gamma(4\tilde{\psi}), \quad (21.114)$$

where \mathbf{C}_α is the volume component of the Freeman approach, and $\mathbf{C}_\beta, \mathbf{C}_\gamma$ are other constant matrices.

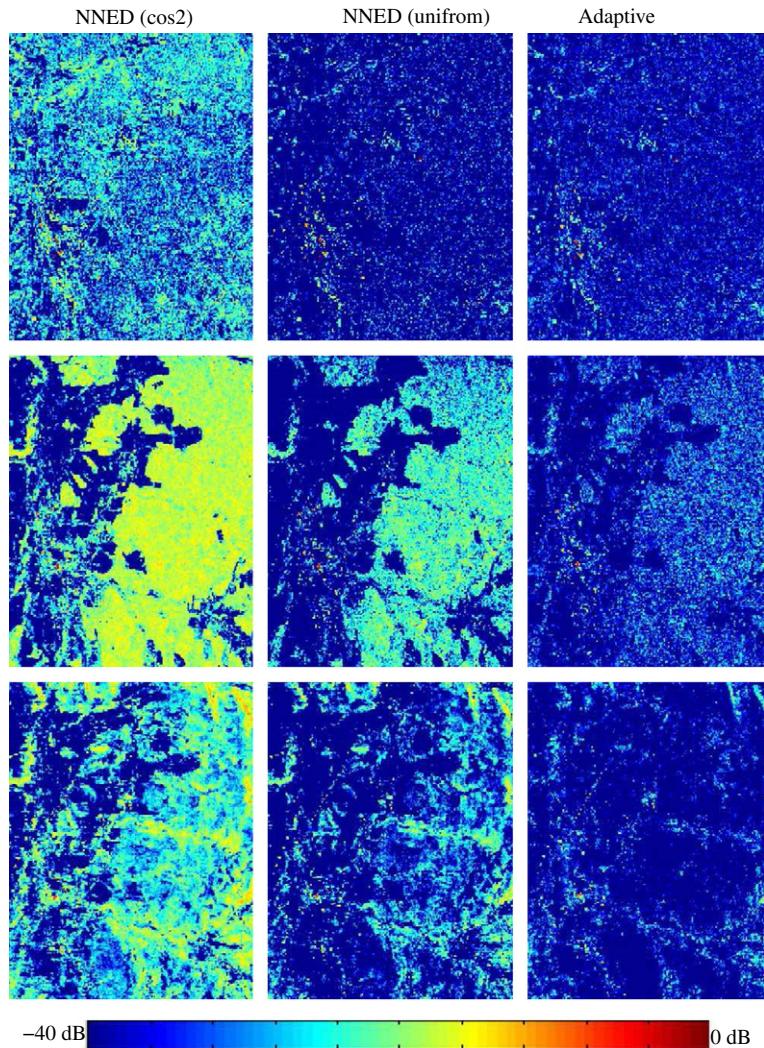
The global covariance matrix is composed of a set of canonical contributions

$$\mathbf{C} = f_v \mathbf{C}_v(\tilde{\psi}, n) + f_d \mathbf{C}_d + f_s \mathbf{C}_s + \mathbf{C}_{\text{rem}}, \quad (21.115)$$

where \mathbf{C}_{rem} is a remaining element that account for terms that do not fit the multi-component model and $\mathbf{C}_d, \mathbf{C}_s$ are identical to the corresponding terms of the Freeman decomposition. The parameters of the volume component are then estimated using a Non Negative Eigenvalue Decomposition approach [28]. For each set of parameters $(\tilde{\psi}, n)$ under test, f_v is estimated as the maximal real value so that

$$\mathbf{C}'_{\text{rem}} = \mathbf{C} - f_v \mathbf{C}_v(\tilde{\psi}, n) \geq 0, \quad (21.116)$$

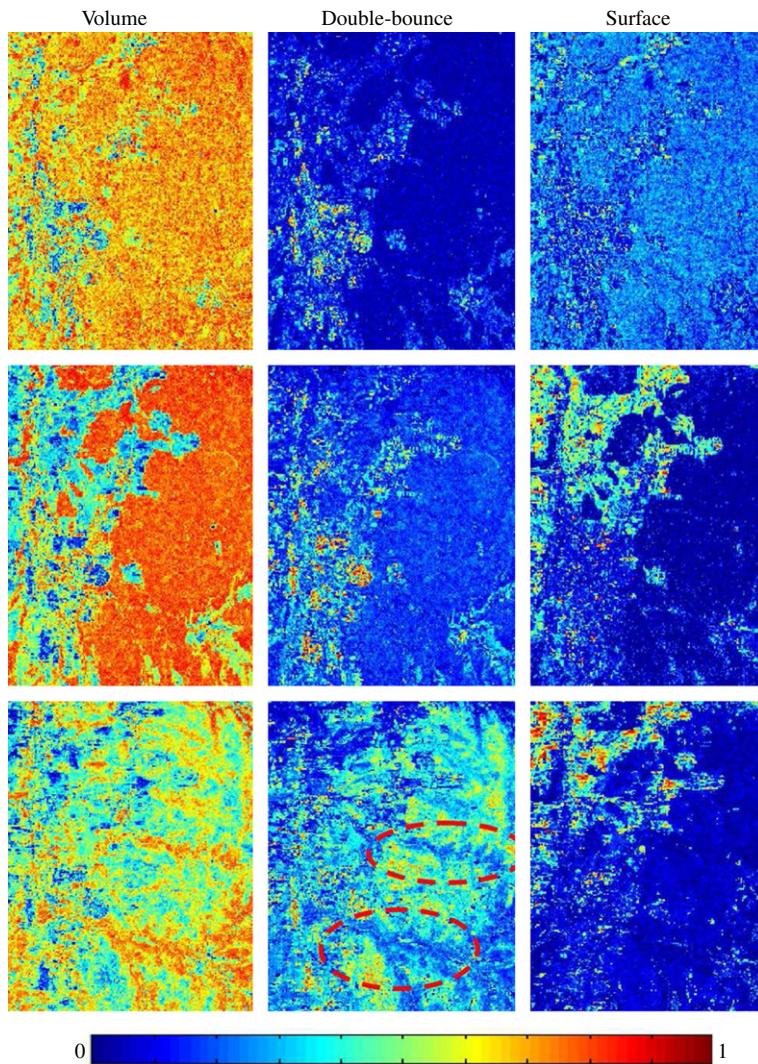
where $\mathbf{C}'_{\text{rem}} \geq 0$ means that \mathbf{C}'_{rem} is positive semi definite, i.e., the eigenvalues of \mathbf{C}'_{rem} are non-negative. The optimal $(f_v, \tilde{\psi}, n)$ is the one for which the remaining polarimetric span $\text{tr}(\mathbf{C}'_{\text{rem}})$ is minimal [29]. Single and double bounce contributions are then estimated for the optimal \mathbf{C}'_{rem} matrix in (21.116). The effectiveness of this approach is demonstrated on Figures 21.32–21.34, where the span of the global remaining term \mathbf{C}_{rem} of (21.115) is used as an indicator of the quality of the decomposition.

**FIGURE 21.33**

$\text{Span}(\mathbf{C}_{\text{rem}})/\text{Span}(\mathbf{C})$ (top) C-band, (middle) L-band, and (bottom) P-band. Results obtained using NNED and two distributions: (Left) Cosine-squared and (center) uniform distributions. From [29].

2.21.3.1.5 An efficient eigenvector based decomposition

The eigenvector based polarimetric decomposition technique proposed by Cloude and Pottier has been widely used for remote sensing applications and tested over various kinds of scenes. The most used parameters of the decomposition are those related to the set of eigenvalues, H and A , as well as $\bar{\alpha}$,

**FIGURE 21.34**

Normalized (Left) Volume, (center) double-bounce, and (right) surface components for the (top) C-band, (middle) L-band, and (bottom) P-band Black Forest images. Dotted lines in red specify (upper) river and (lower) road. From [29] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

the indicator of the nature of polarimetric scattering. Nevertheless this approach reaches some limitations [31]:

- The definition of a mean equivalent scattering mechanism through the averaging of angular parameters may involve important biases, [32].
- Among the four angular parameters used to describe an eigenvector, α is the only roll-invariant. As a consequence all the remain parameters are affected by azimuthal rotations that might be induced by the topography of a scene.

A new approach, presented in [31], proposes an original solutions to both limitations.

2.21.3.1.5.1 Target vector in the circular basis

Target vectors in the Pauli and circular bases are given by

$$\mathbf{k}_P = \frac{1}{2} \begin{bmatrix} S_{hh} + S_{vv} \\ S_{hh} - S_{vv} \\ 2S_{hv} \end{bmatrix} \quad \mathbf{k}_C = \begin{bmatrix} S_{ll} \\ \sqrt{2}S_{l\perp} \\ S_{l\perp l\perp} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} S_{hh} - S_{vv} + j2S_{hv} \\ j\sqrt{2}(S_{hh} + S_{vv}) \\ S_{vv} - S_{hh} + j2S_{hv} \end{bmatrix}. \quad (21.117)$$

In the circular basis, a rotation around the radar line of sight may be represented as

$$\mathbf{k}_C(\psi) = \mathbf{R}(\psi)\mathbf{k}_C \quad \text{with} \quad \mathbf{R}(\psi) = \begin{bmatrix} e^{j2\psi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-j2\psi} \end{bmatrix}, \quad (21.118)$$

where $\mathbf{R}(\psi)$ is a $SU(3)$ operator. This expression shows that, in the circular basis, azimuthal rotation only affects symmetrically the phase of two elements, whereas it mixes up terms in the Pauli basis.

In the circular basis, any target vector may be represented using a set of 4 $SU(3)$ operators [31]:

$$\mathbf{k}_C(\psi) = \mathbf{R}(\psi)\mathbf{R}(\Upsilon)\mathbf{R}(\beta_C)\mathbf{R}(\alpha) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \sin(\alpha)\cos(\beta_C)e^{j(-\frac{4}{3}\Upsilon-2\psi)} \\ \cos(\alpha)e^{j(\frac{8}{3}\Upsilon)} \\ -\sin(\alpha)\sin(\beta_C)e^{j(-\frac{4}{3}\Upsilon+2\psi)} \end{bmatrix}, \quad (21.119)$$

where the parameters may be interpreted as follows [31]

- $\psi = \arg(-S_{l\perp l\perp}^* S_{ll}^*)/4$ measures the rotation around the LOS of the scattering vector in term of the circular co-polarization phase difference. A disoriented scattering vector, i.e., whose parameters do not depend on the azimuthal orientation, may be obtained by compensating ψ according to:

$$\mathbf{k}_{C_{\psi=0}} = \mathbf{R}(-\psi)\mathbf{k}_C(\psi). \quad (21.120)$$

- $\Upsilon = \arg(\mathbf{k}_{C_{\psi=0}}(2)\mathbf{k}_{C_{\psi=0}}^*(2))/4$ is the phase difference between co- and cross-polarization channels of $\mathbf{k}_{C_{\psi=0}}$. This parameters permits to discriminate anisotropic objects like dipoles and quarter waves. It is worth noting that Υ becomes meaningless for $\alpha = 0$ or $\pi/2$
- $\alpha = \alpha_P$ is the original angle defined by Cloude and Pottier [12].

- $\beta_C = \text{acos}(|\mathbf{k}_C(\mathbf{1})|/\sin(\alpha)|\mathbf{k}_C|) \neq \beta_P$, introduced by Corr and Rodrigues in 2002, is the degree of imbalance between the co-polarized coefficients of the circular polarization scattering vector representing the scattering of left-right helices. Unlike β_P , β_C is ψ -invariant.
- $\text{Hel} = \sin^2(\alpha)(\cos^2(\beta_C) - \cos^2(\beta_C)) \in [-1, 1]$, is the helicity parameter and corresponds to the normalized energetic difference between circular co-polarized returns and may be used to estimate the left-right degree of asymmetry of the target vector.

2.21.3.15.2 Parametrization using the exact number of degrees of freedom

The eigendecomposition of a coherency matrix is recalled here:

$$\mathbf{T} = \mathbf{V}\Lambda\mathbf{V}^\dagger = \sum_{i=1}^3 \lambda_i \mathbf{v}_i \mathbf{v}_i^\dagger \quad \text{with } \mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3], \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3). \quad (21.121)$$

The parametrization of an eigenvector proposed by Cloude and Pottier is:

$$\mathbf{v}_i = e^{j\phi_i} [\cos \alpha_i \ \sin \alpha_i \cos \beta_i e^{j\delta_i} \ \sin \alpha_i \sin \beta_i e^{j\gamma_i}]^T. \quad (21.122)$$

Ignoring absolute phase terms, the Number of Parameters (NoP) of such a parametrization may be compared to the number of Degrees of Freedom of a general coherency matrix:

$$\text{NoP}(\mathbf{V}) = 4, \text{NoP}(\Lambda) = 3 \Rightarrow \text{NoP}(\mathbf{V}\Lambda\mathbf{V}^\dagger) = 15 \text{ and } \text{DoP}(\mathbf{T}) = 9. \quad (21.123)$$

In order to overcome this profusion of parameters, Cloude and Pottier proposed to average parameters.

Another approach is proposed in [31] that uses a minimal number of parameters. The polarimetric covariance matrix in the circular basis may be expressed as:

$$\mathbf{C}_C = E(\mathbf{k}_C \mathbf{k}_C^\dagger) = \mathbf{V}_C \Lambda_C \mathbf{V}_C^\dagger. \quad (21.124)$$

Parametrizing the first eigenvector as shown above, one obtains

$$\mathbf{v}_{C_1} = \begin{bmatrix} \sin(\alpha_1) \cos(\beta_{C_1}) e^{j(-\frac{4}{3}\Upsilon_1 - 2\psi_1)} \\ \cos(\alpha_1) e^{j(\frac{8}{3}\Upsilon_1)} \\ -\sin(\alpha_1) \sin(\beta_{C_1}) e^{j(-\frac{4}{3}\Upsilon_1 + 2\psi_1)} \end{bmatrix} \quad (21.125)$$

which can be rewritten as

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{R}_4(-\alpha_1) \mathbf{R}_3(-\beta_{C_1}) \mathbf{R}_2(-\Upsilon_1) \mathbf{R}_1(-\psi_1) \mathbf{v}_{C_1} = \mathbf{R}_{1-4}^{-1}(\psi_1, \Upsilon_1, \beta_{C_1}, \alpha_1) \mathbf{v}_{C_1}. \quad (21.126)$$

Using the orthogonality condition between the first and second eigenvectors, one obtains

$$\mathbf{v}_{C_1}^\dagger \mathbf{v}_{C_2} = 0 \Rightarrow \begin{bmatrix} 0 \\ a \\ b \end{bmatrix} = \mathbf{R}_{1-4}^{-1}(\psi_1, \Upsilon_1, \beta_{C_1}, \alpha_1) \mathbf{v}_{C_2}. \quad (21.127)$$

The second eigenvector may then be described using only two additional parameters, ξ and ϵ

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{R}_{5-6}^{-1}(\xi, \epsilon) \mathbf{R}_{1-4}^{-1}(\psi_1, \Upsilon_1, \beta_{C_1}, \alpha_1) \mathbf{u}_{C_2} \quad (21.128)$$

and then, by definition

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{R}_{1-6}^{-1}(\psi_1, \Upsilon_1, \beta_{C_1}, \alpha_1, \xi, \epsilon) \mathbf{u}_{C_3}, \quad (21.129)$$

where [31]

- $\epsilon = \text{acos}|a| = \text{acos}\left(\frac{\cos\alpha_2}{\sin\alpha_1}\right)$ measures the predictability of α_2 knowing α_1 .
- $\xi = \frac{1}{4} \arg(ab^*)$ describes the direction of the second and third eigenvectors in term of real rotation. This may be more sensitive to the impact of noise and target fluctuations.

The number of parameters needed to describe the eigendecomposition is given by:

$$\text{NoP}(\mathbf{v}_{C_1}) = 4, \text{NoP}(\mathbf{v}_{C_2}) = 2, \text{NoP}(\mathbf{v}_{C_3}) = 0 \Rightarrow \text{NoP}(\mathbf{V}_C \Lambda_C \mathbf{V}_C^\dagger) = 9 = \text{DoP}(\mathbf{T}). \quad (21.130)$$

This decomposition is applied to a C-band data set acquired by the DDRE EMISAR sensor. Images of the main parameters are given on Figure 21.36 and compared to the usual $H/A/\bar{\alpha}$ indicators in Figure 21.35.

2.21.3.2 Advanced characterization of scattering mechanisms

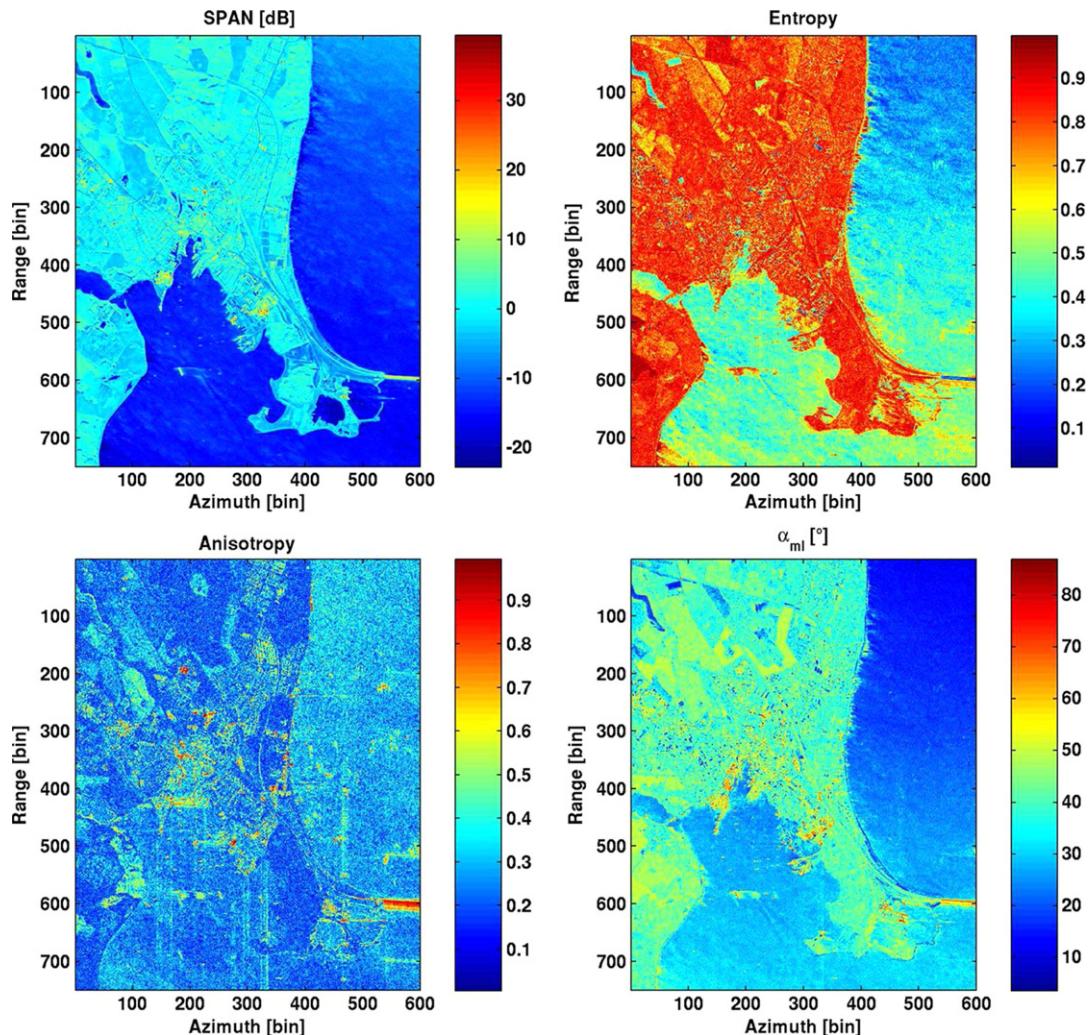
2.21.3.2.1 Estimation of azimuthal orientation

Rotation around the radar line of sight plays an important role in SAR polarimetry. In a series of publications, Lee and his colleagues [33–35] developed an original technique to estimate the azimuthal component of the topography of scene using polarimetric SAR data, and applied it for characterizing different natural environments. As it has been mentioned previously, the circular Right-Left polarimetric basis is particularly adapted to the estimation of azimuthal rotation. The expressions of a target vector in the circular basis, before and after rotation are recalled here:

$$\mathbf{k}_{rl} = \begin{bmatrix} S_{rr} \\ \sqrt{2}S_{rl} \\ S_{ll} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} S_{hh} - S_{vv} + j2S_{hv} \\ j(S_{hh} + S_{vv}) \\ S_{vv} - S_{hh} + j2S_{hv} \end{bmatrix} \mathbf{k}_{rl}(\psi) = \begin{bmatrix} S_{rr}e^{-j2\psi} \\ \sqrt{2}S_{rl} \\ S_{ll}e^{j2\psi} \end{bmatrix}. \quad (21.131)$$

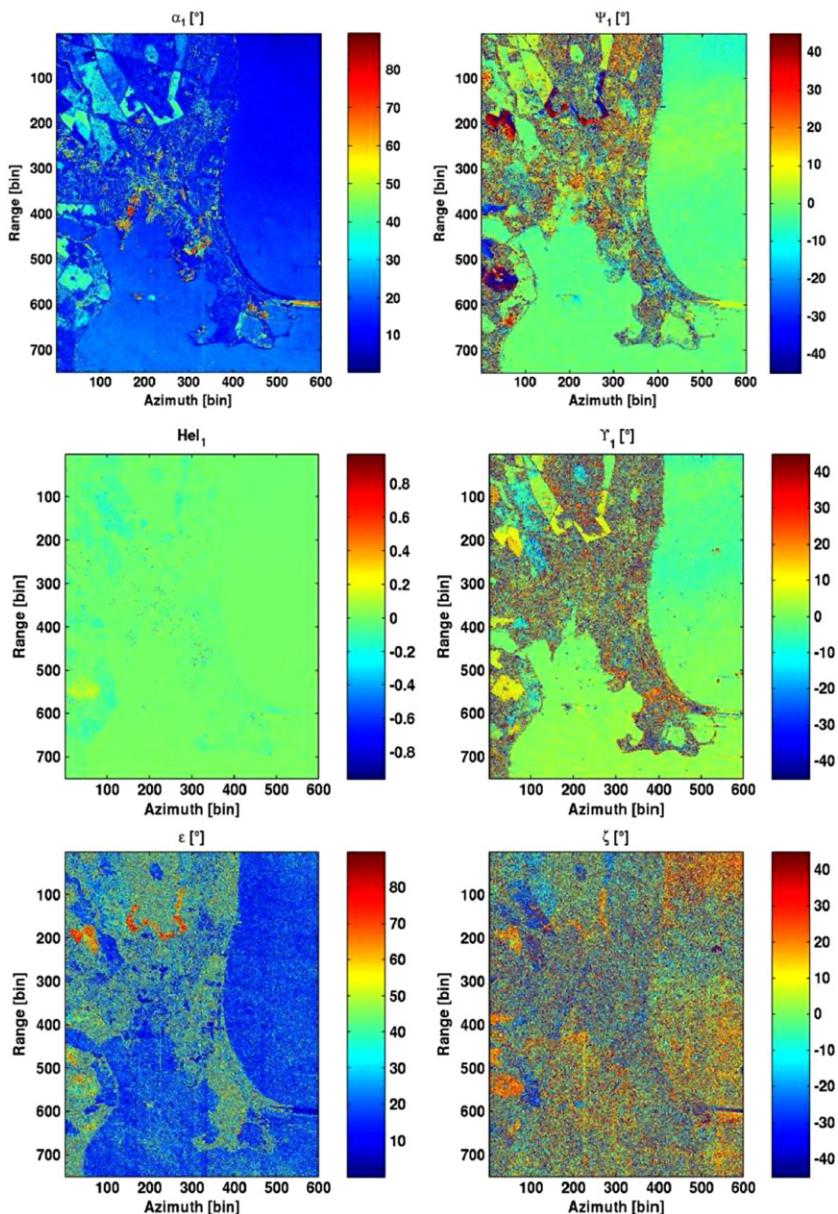
The covariance matrix of the oriented vector, $\mathbf{C}_{rl} = E(\mathbf{k}_{rl}(\psi)\mathbf{k}_{rl}^\dagger(\psi))$, may be expressed as a function of $\mathbf{C}_{rl} = E(\mathbf{k}_{rl}\mathbf{k}_{rl}^\dagger)$ as

$$\mathbf{C}_{rl}(\psi) = \begin{bmatrix} \mathbf{C}_{rl}(1, 1) & \mathbf{C}_{rl}(1, 2)e^{-j2\psi} & \mathbf{C}_{rl}(1, 3)e^{-j4\psi} \\ \mathbf{C}_{rl}(2, 1)e^{+j2\psi} & \mathbf{C}_{rl}(2, 2) & \mathbf{C}_{rl}(2, 3)e^{-j2\psi} \\ \mathbf{C}_{rl}(3, 1)e^{+j4\psi} & \mathbf{C}_{rl}(3, 2)e^{+j2\psi} & \mathbf{C}_{rl}(3, 3) \end{bmatrix}. \quad (21.132)$$

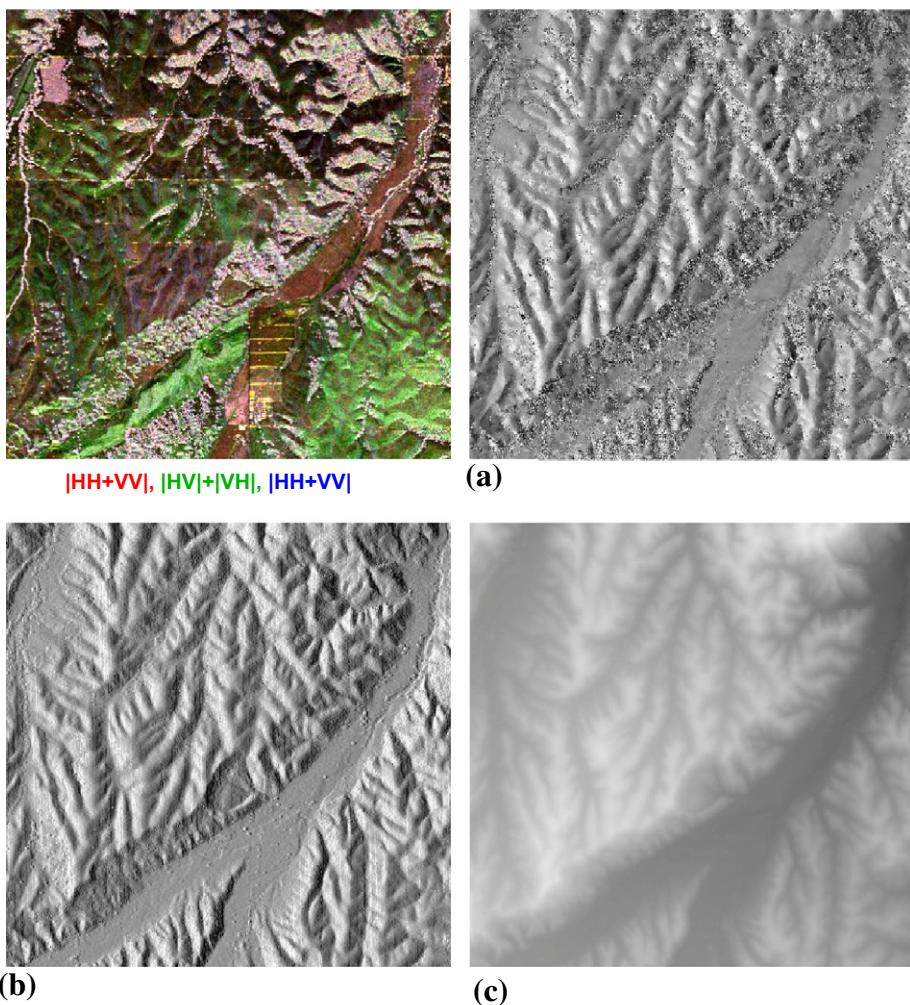
**FIGURE 21.35**

Classic ψ -invariant $H - A - \bar{\alpha}$ parameters characterizing the area around Nyborg port.

This expression shows that the rotation angle may be extracted from three coefficients of $\mathbf{C}_{rl}(\psi)$. Unfortunately, the correlation terms $\mathbf{C}_{rl}(1, 2)$, $\mathbf{C}_{rl}(1, 3)$, $\mathbf{C}_{rl}(2, 3)$ are in general complex, and unrelated phase terms may significantly perturb the estimation of ψ . However, horizontal natural environments generally satisfy the hypothesis of reflection symmetry, i.e., $E(S_{hh}S_{hv}^*) = E(S_{vv}S_{hv}^*) = 0$. In such a case, $\mathbf{C}_{rl}(1, 3) \in \mathbb{R}$, and ψ can be estimated from the argument of $\mathbf{C}_{rl}(\psi)(1, 3)$ [34]. Figure 21.37

**FIGURE 21.36**

Main parameters of the eigendecomposition proposed in [31].

**FIGURE 21.37**

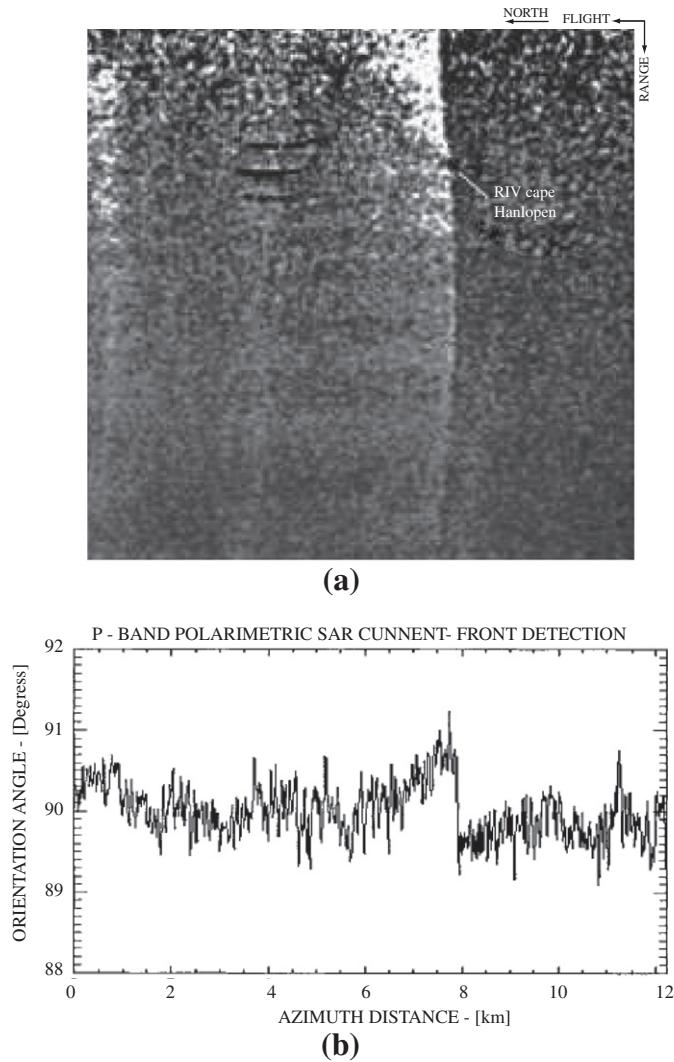
(a) polarization orientation angles over Camp Roberts, CA, USA at Land, (b) angle computed from a C band DEM, (c) C band DEM. From [34].

compares azimuthal slopes estimated using a single polarimetric image at L band, to those derived from a C band DEM.

A similar technique was applied to characterize current-induced sea topography at P band, Figure 21.38.

2.21.3.2.2 A statistical segmentation that preserves scattering mechanisms

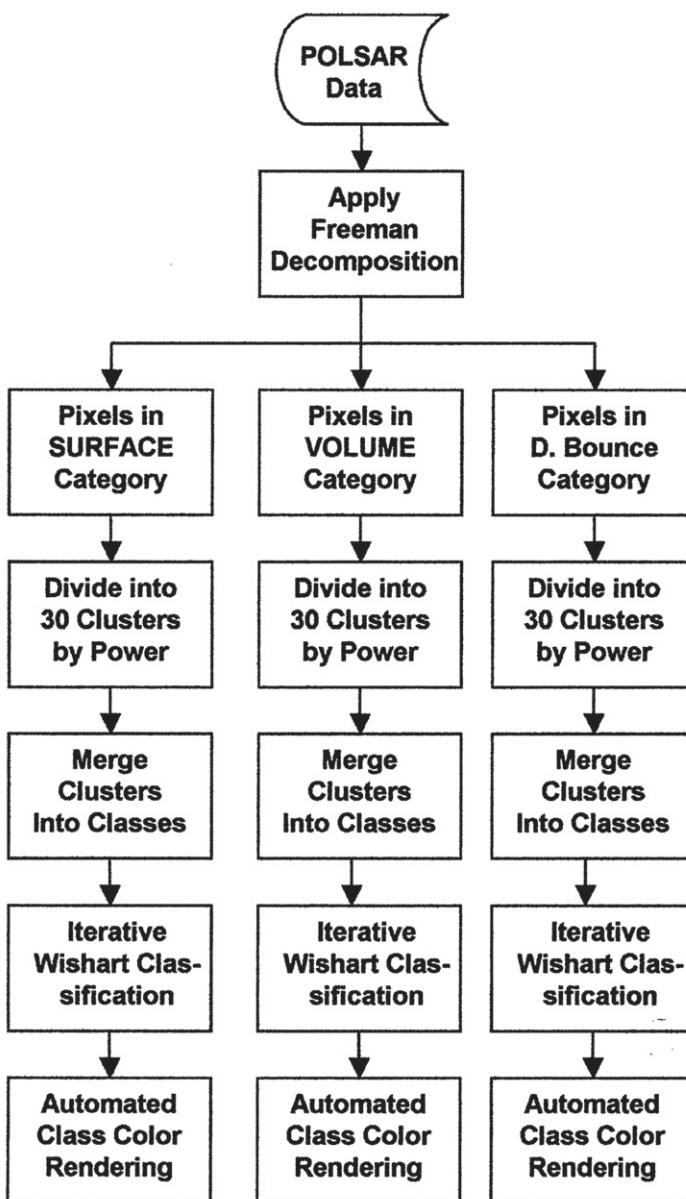
The statistical Wishart segmentation scheme mentioned previously permits to cluster into compact groups pixels having similar statistical polarimetric behavior. In practice, this classifier is highly sensitive

**FIGURE 21.38**

(a) Orientation angle image computed at P band over the Gulf stream (b) a profile. From [33].

to the intensity driven polarimetric span [32,36]. In consequence, small variations of polarimetric parameters between media having similar spans may be ignored pixels having different scattering behaviors may be affected to the same class. In order to overcome this problem, a scattering mechanism preserving segmentation [36] illustrated in Figure 21.39 can be applied.

Pixels are first separated according to their scattering mechanism using the Freeman decomposition. each group is then separately segmented using a Wishart segmentation. Results are shown in Figure 21.40.

**FIGURE 21.39**

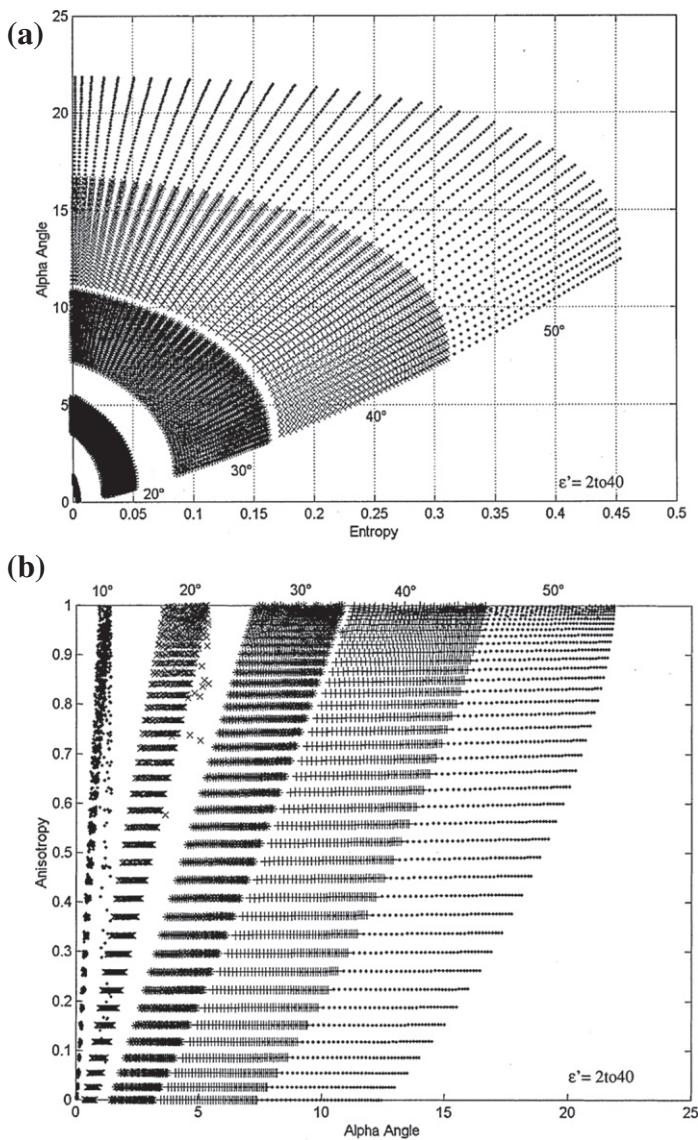
Synopsis of scattering mechanism preserving classifier.

**FIGURE 21.40**

Classification results over San Francisco.

2.21.3.3 Characterization of soils using polarimetric decomposition techniques

Polarimetric scattering being highly sensitive to the geometry and electromagnetic properties of a medium, polarimetric SAR data have been intensively used for estimating the moisture and roughness of soils. Some of the techniques proposed in the literature are directly based on polarimetric decomposition techniques.

**FIGURE 21.41**

(a) $H - \bar{\alpha}$ Entropy/Alpha (b) $A - \bar{\alpha}$ diagrams for different local incidence angles and different values of $\Delta\psi$.

2.21.3.4 A methods using H , A , and $\bar{\alpha}$ and the SPM model

This work, based on the $H/A/\bar{\alpha}$ decomposition has been presented in [37] and is based on the use of the Small Perturbation surface scattering Model (SPM) [38]. At order 1, this model permits to simulate the polarimetric response of a slightly rough surface:

$$\mathbf{T}_{\text{SPM}}(k\sigma, W, \theta, \epsilon) = m_s^2(k\sigma, W)\mathbf{T}_0(\theta, \epsilon), \quad (21.133)$$

where k stands for the carrier wavenumber, σ represents the standard of the rough surface heights, W is the stationary surface spectrum, θ is the local angle of incidence, ϵ is the surface dielectric constant, and \mathbf{T}_0 is a reflection symmetric coherency matrix. Due to the restriction of the model to the first order, roughness and dielectric properties can be easily separated, since roughness only affects the polarimetric span whereas the rest of polarimetric information parameters, \mathbf{T}_0 in (21.133), depends on the incidence angle and dielectric constant only. At order 1, the cross-polarization is null, $E(|S_{hv}|^2)$ and the co-polarization channel are perfectly correlated. As a consequence, the entropy of \mathbf{T}_{SPM} is very close to 1 and does not fit observations at L or C bands. In order to overcome this limitation, it is proposed in [37] to introduce depolarization within the SPM model by artificially rotating \mathbf{T}_{SPM} around the radar line of sight:

$$\mathbf{T}(\Delta\psi) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \mathbf{R}(\psi)\mathbf{T}_{\text{SPM}}(k\sigma, W, \theta, \epsilon)\mathbf{R}^\dagger(\psi)f_\psi(\psi)d\psi, \quad (21.134)$$

where $f_\psi(\psi)$ is a uniform pdf centered around 0 and whose width equals $\Delta\psi$.

The eigendecomposition of Cloude and Pottier is then applied to $\mathbf{T}(\Delta\psi)$ to produce $H(\Delta\psi)$, $\bar{\alpha}(\Delta\psi)$, $A(\Delta\psi)$. Simulated values presented in Figure 21.41 show that $H(\Delta\psi)$, $\bar{\alpha}(\Delta\psi)$ may be used to estimate the dielectric constant, whereas the $A(\Delta\psi)$ is not correlated with α . A look-up table procedure is used in [37] to estimate ϵ as the value that minimizes the distance between measured and simulated H , $\bar{\alpha}$ values, whereas an empirical relation, $A = 1 - k\sigma$, is used to estimate the surface roughness. A polynomial relation $m_v = f_{\text{pol}}(\epsilon)$ is used to convert dielectric permittivity to volumetric moisture content. Results obtained using L band acquisitions over two sites are given in Figure 21.42.

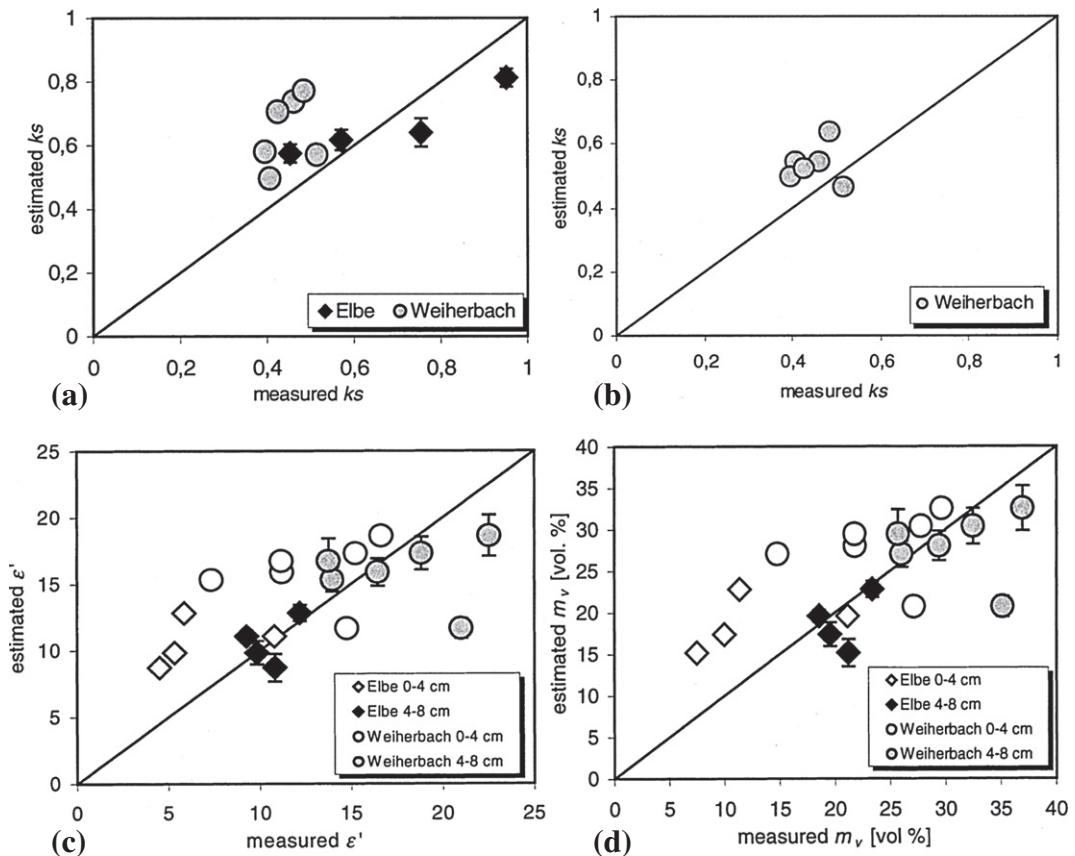
2.21.3.5 A method using reflection symmetry and the IEM model

Another technique, introduced in [39,40], proposes an alternative approach, based on the IEM surface scattering model [41] and on the analytical computation of the eigenvalues on a reflection symmetric coherency matrix:

2.21.3.5.1 The SERD and the DERD parameters

Two eigenvalue-based parameters, the Single bounce Eigenvalue Relative Difference (SERD) and the Double bounce Eigenvalue Relative Difference (SERD) have been introduced in [39,40] to characterize natural media. Both parameters are derived from the eigenvalues of a coherency \mathbf{T} considering the reflection symmetry hypothesis:

$$\mathbf{T} = \frac{1}{2}E \left(\begin{bmatrix} |S_{hh} + S_{vv}|^2 & (S_{hh} + S_{vv})(S_{hh} - S_{vv})^* & 0 \\ (S_{hh} - S_{vv})(S_{HH} + S_{vv})^* & |S_{hh} - S_{vv}|^2 & 0 \\ 0 & 0 & 4|S_{hh}|^2 \end{bmatrix} \right). \quad (21.135)$$

**FIGURE 21.42**

Estimation results at L band (a) roughness without additive noise filtering, (b) roughness with additive noise filtering, (c) dielectric constant (d) volumetric soil moisture content.

In such a case, it is possible to derive the analytical expressions of the corresponding unsorted (NOS) eigenvalues given by [39,40]:

$$\begin{aligned}\lambda_{1\text{NOS}} &= \frac{\mathbb{E}(|S_{hh}|^2 + |S_{vv}|^2) + \sqrt{\mathbb{E}(|S_{hh}|^2 - |S_{vv}|^2)^2 + 4\mathbb{E}(|S_{hh}S_{vv}^*|^2)}}{2}, \\ \lambda_{2\text{NOS}} &= \frac{\mathbb{E}(|S_{hh}|^2 + |S_{vv}|^2) - \sqrt{\mathbb{E}(|S_{hh}|^2 - |S_{vv}|^2)^2 + 4\mathbb{E}(|S_{hh}S_{vv}^*|^2)}}{2}, \\ \lambda_{3\text{NOS}} &= 2\mathbb{E}(|S_{hv}|^2).\end{aligned}\quad (21.136)$$

The first and second eigenvalues depend on the co-polarized backscattering coefficients and on the correlation between the vertical and horizontal channels (ρ_{hhvv}). In this case, the relation $\lambda_{1\text{NOS}} > \lambda_{2\text{NOS}}$

always holds. The third eigenvalue corresponds to cross-polarized channel and is related to multiple scattering for rough surfaces. One may remark that the entropy remains invariant against any permutation within the set of eigenvalues.

In order to determine the scattering mechanisms, an analysis is led on the angles extracted from the first two eigenvectors \mathbf{u}_1 and \mathbf{u}_2 , associated to $\lambda_{1\text{NOS}}$ and $\lambda_{2\text{NOS}}$, respectively. The corresponding α angles verify the following property

$$\alpha_1 + \alpha_2 = \frac{\pi}{2} \Rightarrow \begin{cases} \alpha_1 \leq \frac{\pi}{4} \Leftrightarrow \text{SB}, \alpha_2 > \frac{\pi}{4} \Leftrightarrow \text{DB} \\ \quad \text{or} \\ \alpha_1 > \frac{\pi}{4} \Leftrightarrow \text{DB}, \alpha_2 \leq \frac{\pi}{4} \Leftrightarrow \text{SB}, \end{cases} \quad (21.137)$$

where SB stands for Single Bounce reflection and DB for Double Bounce reflection. The two eigenvalue-based parameters called the Single bounce Eigenvalue Relative Difference (SERD) and the Double bounce Eigenvalue Relative Difference (DERD) are built up to compare the relative importance of the different scattering mechanisms and are defined as:

$$\text{SERD} = \frac{\lambda_S - \lambda_{3\text{NOS}}}{\lambda_S + \lambda_{3\text{NOS}}}, \quad \text{DERD} = \frac{\lambda_D - \lambda_{3\text{NOS}}}{\lambda_D + \lambda_{3\text{NOS}}}, \quad (21.138)$$

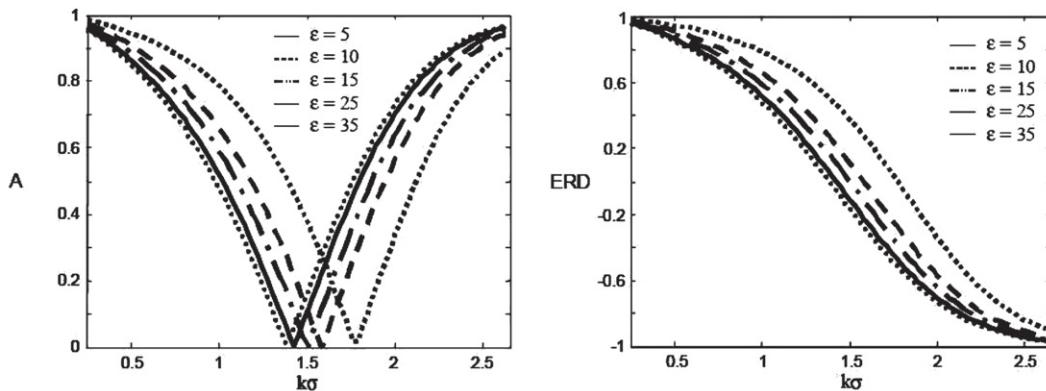
where λ_S and λ_D are the two eigenvalues respectively associated to the single bounce and to the double bounce scattering mechanisms. Due to the lack of sorting, SERD and DERD, cover a wider range than the anisotropy, $\text{SERD}, \text{DERD} \in [-1, 1]$, and are associated to specific scattering mechanisms.

The DERD parameter can be compared with the anisotropy A , whereas the SERD parameter reveals useful with large entropy, in order to determine the nature and the importance of the different scattering mechanisms.

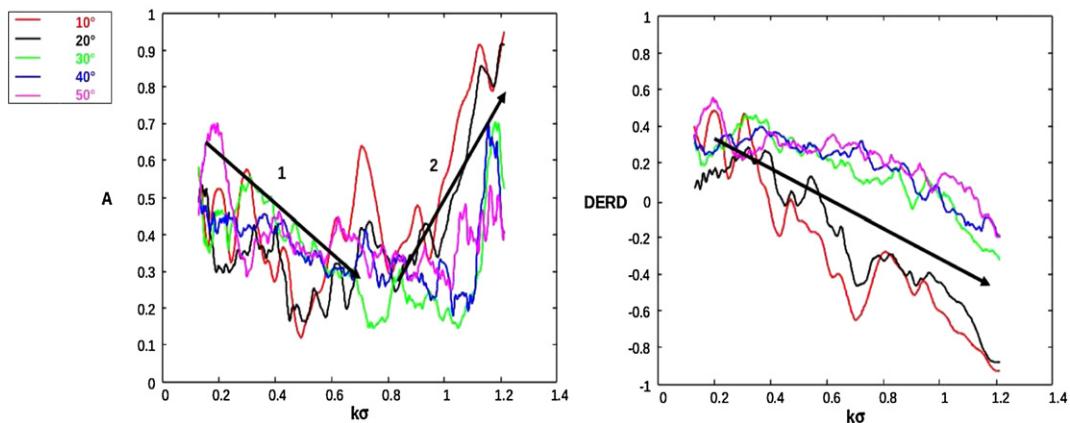
2.21.3.5.2 Roughness and moisture retrieval

In order to characterize natural surfaces, the Integral Equation Model (IEM) is employed to derive a polarimetric coherency matrix [39,40]. This model natural accounts for depolarization and produce general reflection symmetric second order representations. Using this model, the DERD parameter can be compared to the polarimetric anisotropy A that is usually employed as a surface roughness descriptor [37]. Figure 21.43 shows, both parameters variations versus the roughness indicator $k\sigma$ obtained using the IEM model for various dielectric constants, ϵ , where k is radar wave number and σ is the surface root mean square height.

The DERD parameter is similar to the anisotropy A for small roughness values, but presents a different behavior for high frequencies. These parameters are very sensitive to surface roughness relative to frequency, whereas the dependence on the dielectric constant ϵ is less significant. For each dielectric constant ϵ value, one anisotropy A value corresponds to two different values of $k\sigma$, thus introducing an ambiguity for surface roughness extraction, whereas the DERD is strictly monotonic with $k\sigma$. An important difference between these two parameters is that the dynamic range of the DERD parameter is larger $[-1; +1]$ than the anisotropy range $[0; +1]$. It follows that the DERD parameter has now to

**FIGURE 21.43**

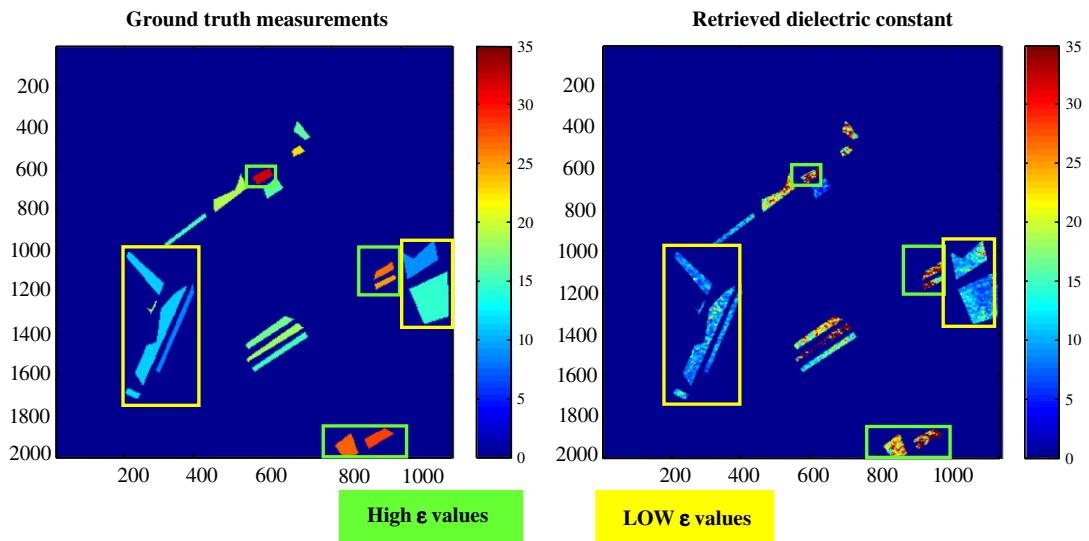
Values of A and DERD simulated using the IEM and various surface conditions.

**FIGURE 21.44**

Values of A and DERD obtained from data sets acquired by the JRC, for different roughness conditions.

be considered as a better surface roughness discriminator. Similar conclusions can be done on indoor measurements realized at the European JRC laboratory [40] and depicted in Figure 21.44.

In order to avoid a bias on the mean $\bar{\alpha}$ [REF] that needs to be compensated using the entropy, only α_{SB} is used to retrieve ϵ by comparison with simulated $\alpha_{SB,IEM}$ values. Results obtained over the test site of Alling are shown in Figure 21.45.

**FIGURE 21.45**

Dielectric constant estimation results obtained over the Aling test site using data acquired at L band the DLR ESAR sensor, Note that α_{IEM} has been corrected by a constant over the whole image.

2.21.4 Selected topics in multidimensional polarimetric SAR signal processing

2.21.4.1 Coherent Time-Frequency characterization of complex polarimetric features

Conventional SAR image analysis and geophysical parameter retrieval techniques from strip-map (SAR) data generally assume that scenes are formed of static scatterers observed in the direction perpendicular to the flight track and at a fixed frequency, equal to the emitted signal carrier's one. When imaging complex objects and media, potential variations of the signal measured during the SAR acquisition may strongly affect feature estimates derived from the resulting SAR data and may lead to erroneous interpretations. It is well known [42] that perturbations induced by the motion of a scatterer may highly modify its SAR response in terms of reflectivity, spatial localization and focusing. Static objects with anisotropic geometrical structures or having a frequency selective response may show a varying electromagnetic behavior as they are illuminated from different positions or at different frequency components during SAR integration. The resulting SAR response being well described by the spatial convolution of an conventional scene SAR image with specific functions accounting for each effect, non ideal features can be easily detected and characterized in the spectral domain.

This section presents different techniques for detecting scatterers having a varying response during the SAR acquisition and characterizing the underlying physical phenomenon that generates these variations.

These approaches are based on specific coherent Time-Frequency (TF) decompositions that analyze the response of scatterers by locally estimating the spectral content of their SAR response. These TF techniques can be applied on already focused SAR images and not only on raw SAR signal, and were designed in order to deal with already focused SAR images and not only raw SAR signals and to perform a two-dimensional (2D) rangeazimuth coherent and revertible analysis, i.e., the original SAR image can be reconstructed from local spectral estimates. A physical interpretation of the spectral decomposition in the azimuth and range directions as well as the use of polarimetric SAR data permit to characterize the backscattering properties complex media in a refined way compared to classical techniques and estimate some of their physical features.

2.21.4.1.1 Principles of coherent Time-Frequency analysis of SAR data

2.21.4.1.1.1 SAR image spectral content

As depicted in Figure 21.46, a SAR measurement consists in repeatedly emitting a signal, $s_e(t)$, in the across track direction and receiving the echo from the observed scene, $s_r(x, t)$, at different locations x along the acquisition track. A scatterer P_0 located at coordinates $(x_0, y_0, 0)$ is observed for different values of the azimuth look angle, ϕ , defined by $\tan(\phi) = (x - x_0)/d_0(x)$, defined $d_0(x) = \sqrt{r_0^2 + (x - x_0)^2}$ being the varying radar-scatterer distance. The range of the azimuth angle is defined by the antenna aperture, whereas the emitted signal is characterized by a bandpass spectrum centered around a carrier frequency f_c , with bandwidth B_f .

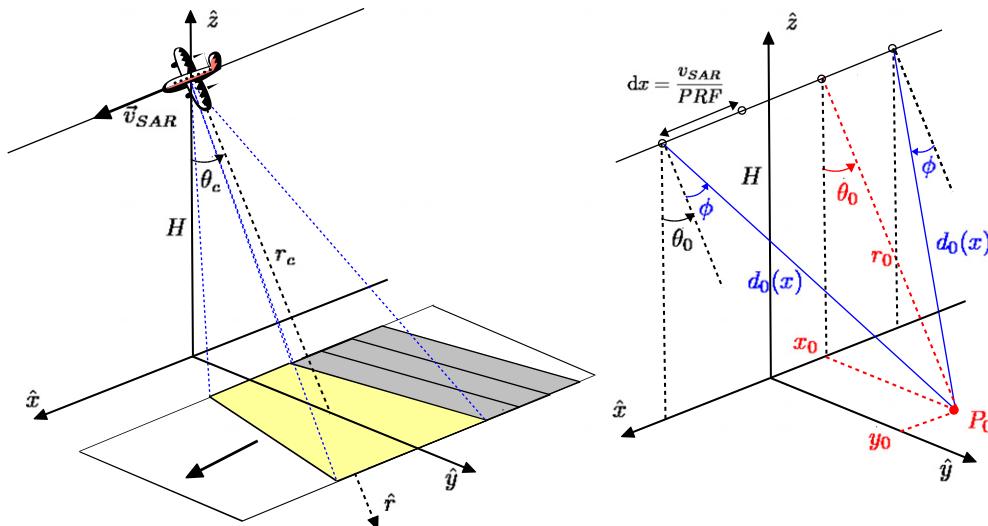
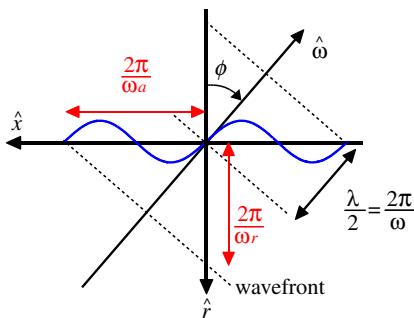


FIGURE 21.46

Geometrical configuration of a SAR acquisition. Repeated emission-reception of signals, with a rectangular antenna aperture (left). Observation of a scatterer P_0 from different positions along the acquisition track, with corresponding azimuth look angles.

**FIGURE 21.47**

Decomposition of a plane wave propagating along $\hat{\omega}$ into azimuth and range components. A spherical wave may be represented as a sum of plane waves with varying azimuth orientation ϕ .

After base-band conversion, the received signal can be focused to produce a 2-D coherent image in the azimuth-range domain, $s(x, r)$, that represents a reconstruction of the coherent reflectivity of the scene. Under simplifying assumptions and considering ideal acquisition conditions [21], a coherent SAR image may be formulated as convolution of the scene coherent reflectivity, $r(x, r)$ and the SAR 2-D impulse response, $h(x, r)$ and may be represented in both spatial and spectral domains as:

$$s(x, r) = r(x, r) \otimes h(x, r) \equiv S(\omega_a, \omega_r) = R(\omega_a, \omega_r)H(\omega_a, \omega_r), \quad (21.139)$$

where $S(\omega_a, \omega_r) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s(x, r) e^{-j\omega_a x} e^{-j\omega_r r} dx dr$ is the 2-D Fourier transform of the SAR image $s(x, r)$ and \otimes represents the convolution operator. The spectral coordinates (ω_a, ω_r) , representing two-way wavenumbers in range and azimuth, are illustrated in Figure 21.47 and can be formulated as [43]:

$$\omega_a = \omega \sin \phi, \quad \omega_r = \omega \cos \phi - \omega_c, \quad (21.140)$$

where ω is the emitted signal wavenumber and can be related to the electrical frequency using, $f \in f_c + \left[-\frac{B_f}{2}, \frac{B_f}{2}\right]$, through the wave propagation velocity, c , as $\omega = \frac{4\pi f}{c}$.

For an ideal scene, whose reflectivity is uniformly distributed over the spectral domain, the resolutions of the SAR image are driven by the impulse response of the SAR system and are given by

$$\delta x = \frac{2\pi}{\Delta \omega_a} = \frac{c}{4f_c \sin(\Delta\phi/2)} \quad \text{and} \quad \delta r = \frac{2\pi}{\Delta \omega_r} = \frac{c}{2B_f}, \quad (21.141)$$

where $\Delta\phi$ stands for the processed azimuthal aperture, whose maximal value is set by the acquisition antenna characteristics.

It is worth noting that the direct relation between a SAR image and the reflectivity of scene in (21.139), as well as the physical meaning of the spectral coordinates in (21.140) are valid when dealing with coherent Single Look Complex (SLC) SAR data sets, i.e., each pixel of the SAR image corresponds to a complex number whose modulus is proportional to the focused reflectivity and whose absolute phase depends on the observed medium as well as on the measurement phase history. Transforming an SLC

image to an incoherent one, like an intensity image $I(x, r) = |s(x, r)|^2$, involves an irremediable loss of information and interpretation.

2.21.4.1.1.2 Time frequency decomposition

The T-F decompositions technique selected here is based on the use of a 2-D windowed Fourier transform, or 2-D Gabor transform. This kind of transformation permits to decompose a two-dimensional signal, $s(\mathbf{l})$, with $\mathbf{l} = [x, y]^T$ a 2-D location, into different spectral components, using a convolution with an analyzing function $g(\mathbf{l})$, as follows [44]:

$$s(\mathbf{l}_0; \omega_0) = \int s(\mathbf{l})g(\mathbf{l} - \mathbf{l}_0) \exp(-j\omega_0^T(\mathbf{l} - \mathbf{l}_0))d\mathbf{l}, \quad (21.142)$$

where $\omega_0 = [\omega_x, \omega_y]^T$ indicates a position in frequency, and $s(\mathbf{l}_0; \omega_0)$ represents the decomposition result around the spatial and frequency locations \mathbf{l}_0 and ω_0 . The application of a Fourier transform to (21.142) shows that the spectrum of $s(\mathbf{l}_0; \omega_0)$ is given by the product of the original signal spectrum and the transform of the analyzing function g shifted around the frequency vector ω_0 :

$$S(\omega; \omega_0) = S(\omega)G(\omega - \omega_0). \quad (21.143)$$

It is clear from Eqs. (21.142) and (21.143) that this time frequency approach may be used to characterize, in the spatial domain, behaviors corresponding to particular spectral components of the signal under analysis, selected by the analyzing function g . Among the wide variety of existing TF analysis methods, the simple atomic decomposition selected in this study presents some interesting properties. It is linear, and hence preserves the coherence and energy of signals, it is not affected by artifacts related to cross-terms and may be inverted, i.e., depending on the analyzing function g , $s(\mathbf{l})$ may be reconstructed from a set of TF samples $s(\mathbf{l}; \omega_0)$, provided that some sampling conditions in spatial and spectral domains are satisfied. The resolutions of the analysis in space and frequency are not independent, and their product is fixed by the Heisenberg-Gabor uncertainty relation, given in 1-D by [44]

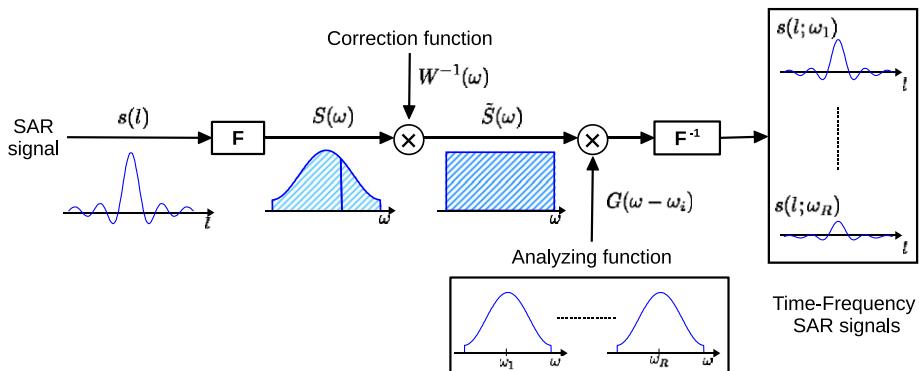
$$\delta l \delta \omega = u_g. \quad (21.144)$$

This relation specifies that the space-frequency resolution product equals a constant u_g , determined by g . An analyzing function with an excessively narrow spectral bandwidth would involve an excellent spectral resolution but might then lead to a meaningless analysis in the space domain owing to a bad localization, i.e., increasing the description ability of the analysis in one domain worsens its accuracy in the dual one. The nature of the analyzing function is generally chosen so as to preserve resolution while maintaining sufficiently low side-lobe amplitudes in the space domain.

2.21.4.1.1.3 Time-Frequency decomposition of SAR images

The timefrequency approach presented here deals with processed SAR images, denoted $s(\mathbf{l})$ with $\mathbf{l} = [x, r]^T$, rather than raw data. This type of data is better accessible to common users and is generally processed through compensation procedures to reduce the effects of acquisition errors. In practice the simple SAR image model given in (21.139) needs to be completed in order to account for additional weighting terms, mainly due to the antenna pattern and side-lobe reduction functions, as [45]

$$S(\omega) = R(\omega)H(\omega)W(\omega) \quad (21.145)$$

**FIGURE 21.48**

Synopsis of the proposed Time-Frequency decomposition of a mono-dimensional.

with $\omega = [\omega_a, \omega_r]^T$. The synopsis of the TF decomposition based on the spectral definition on (21.143) is given in Figure 21.48.

The first step consists of correcting for potential spectral imbalances, represented by $W(\omega)$ in Eq. (21.145), in the original, full-resolution SAR image. This can be achieved by calculating average image spectra in range and azimuth and then multiplying the full-resolution spectrum $S(\omega)$ by the inverse of the estimated 2-D weighting function. The TF decomposition is then conducted by multiplying the corrected spectrum by the Fourier transform of the analyzing function and going back to the spatial domain. The resulting still focused SAR image $s(l; \omega_0)$ has a lower resolution than the original SAR data and depicts the scene behavior over the 2D frequency domain located in the neighborhood of ω_0 . As it might be observed in (21.145), the use of processed data limits the spectral exploration range to the one of the reference function used for raw data processing and focusing. In order to emphasize the physical interpretation of coherent SAR image analysis, one may simplify the wavenumber expressions given in (21.140) using narrow beamwidth and bandwidth approximations

$$\omega_a \approx \omega_c \sin \phi, \quad \omega_r \approx \omega - \omega_c. \quad (21.146)$$

Time-Frequency decomposition in the azimuth direction consists of deriving a set of images containing different parts of the SAR Doppler spectrum with a reduced resolution, but corresponding to different azimuth look angles. This kind of analysis can be applied to detect objects or media with anisotropic behaviors, like scatterers with complex geometrical structures, human-made objects, or natural media having periodic structures in the case of agricultural areas, or linear alignments of strong scatterers [45]. In the range direction, TF analysis permits to compare the response of a scene observed at different frequencies, contained within the emitted signal spectral domain, and can be used to detect and characterize media with frequency-sensitive responses, like resonating spherical or cylindrical objects, periodic structures, or coupled scatterers with interfering characteristics.

Polarimetric SLC SAR images can be easily decomposed by applying the presented approach independently over each polarization channel. Usual polarimetric representations may then be reconstructed

to study the polarimetric behavior of a scene around specific spectral locations.

$$\mathbf{k}(\omega_0) = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh}(\omega_0) + S_{vv}(\omega_0) \\ S_{hh}(\omega_0) - S_{vv}(\omega_0) \\ 2S_{hv}(\omega_0) \end{bmatrix} \quad \text{and} \quad \mathbf{T}(\omega_0) = \mathbb{E}[\mathbf{k}(\omega_0)\mathbf{k}^H(\omega_0)], \quad (21.147)$$

where spatial locations, \mathbf{l} , have been omitted.

2.21.4.1.2 Characterization of natural environments with non-stationary polarimetric SAR responses

2.21.4.1.2.1 Discrimination of non-stationary POLSAR responses

Discrete Time-Frequency decomposition in range and azimuth. As shown in (21.147), the timefrequency decomposition can be applied around any frequency location inscribed within the rangeazimuth frequency range defined by the SAR transfer function $H(\omega)$. Nevertheless, it is often useful to first process the analysis around a limited (discrete) set of frequency locations to appreciate the global behavior of the scene under observation and emphasize changes from one subspectral image to another by minimizing their correlation. The polarimetric SAR data set under study has been acquired by the DLR E-SAR sensor, at L band, over the Alling test site in Germany. The original image resolution is 2 m in range and 1 m in azimuth, corresponding to an azimuthal variation of the look angle of approximately 7.5° and a chirp bandwidth of 75 MHz. Figure 21.49 shows the full- resolution span image corresponding to the total polarimetric backscattered power. The considered scene is mainly composed of agricultural fields and forest. An urban area is located at the bottom left corner of the image.

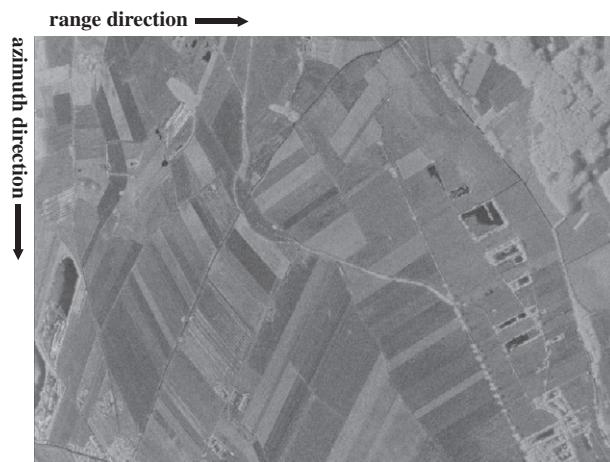
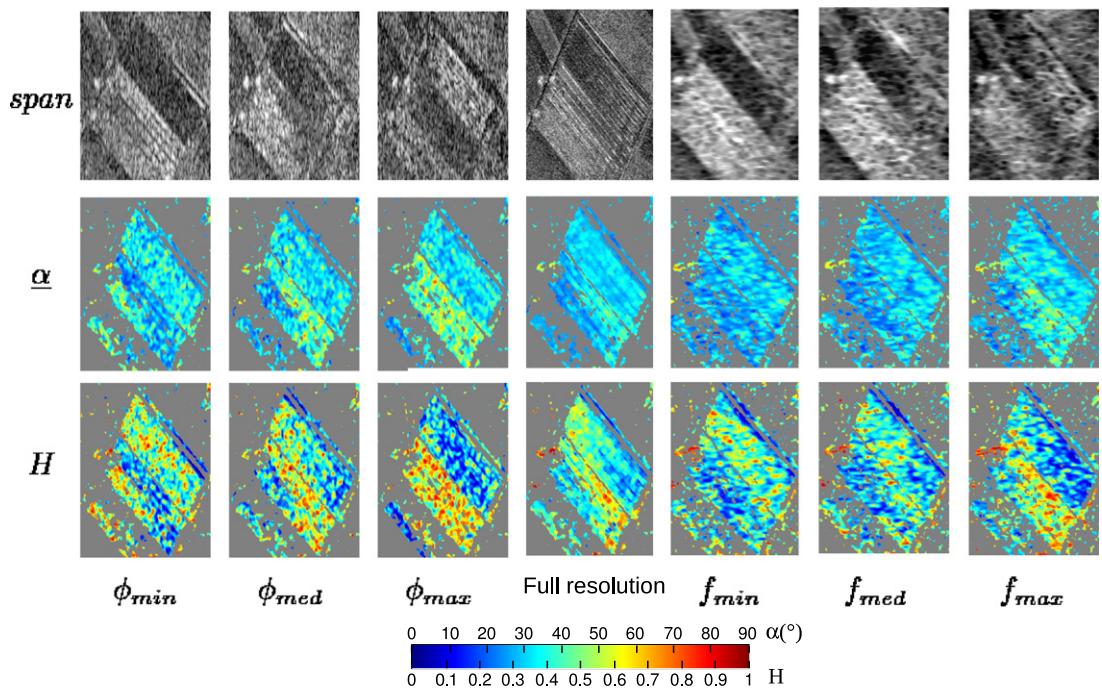


FIGURE 21.49

Full resolution span image of the Alling test site.

**FIGURE 21.50**

Polarimetric parameters at full resolution (center) and after 1-D TF analysis in the azimuth (left) and range (right) directions.

The decomposition in the azimuth direction is performed using independent subspectra and keeping the range resolution to its original value. Figure 21.50 shows results obtained over an area corresponding to plowed fields. Images of the *span*, *H*, and α parameters are represented for different azimuthal look angles and for the full-resolution case. It can be observed in Figure 21.50 that large variations in the scattering mechanism nature, α , and degree of randomness, *H*, occur while the azimuth look angle changes. For particular azimuth look angles, some fields show a sudden change of behavior: the *span* reaches a maximum value, whereas the polarimetric indicators *H* and α are characterized by low values. The stripes in the *span* image, indicating that coherent constructive and destructive interferences occur within the pixels, are characteristic for Bragg resonant scattering over periodic surfaces [45].

Other types of media may also have nonstationary polarimetric features during the azimuthal integration. It was observed that some point targets and linear structures, such as diffracting edges or road berms, have significant backscattering pattern variations as the look angle changes. In particular, metallic link fencing was found to present a scattering mechanism ranging from single-bounce up to double-bounce scattering, depending on the SAR azimuthal look angle. In general, nonstationary targets have strongly anisotropic shapes, or facets acting like directional scatterers, involving changes in the underlying scattering mechanism and in the total backscattered power.

In contrast, forested areas have a stationary behavior during the SAR integration. Backscattering from forested areas at L band is known to be dominated by volume diffusion, which corresponds to the scattering over randomly distributed anisotropic constituents. The coherent integration of the randomly scattered waves leads to a response that is characterized by a high intensity and low degree of polarization, but with isotropic behavior.

Detection of non-stationary polarimetric TF behaviors. Each pixel of the SAR scene is associated with a set of R independent target vectors, $\mathbf{k}(\omega_i)$ with $i = 1, \dots, R$, derived from independent range-azimuth subspectra, i.e., subspectra selected using non-overlapping functions $G(\omega_i)$. Under the classical speckle affected scattering hypothesis, these target vectors follow independent complex Gaussian multivariate distributions, $f(\mathbf{k}(\omega_i)) = \mathcal{N}_{\mathcal{C}}(\mathbf{0}, \Sigma_i)$. The stationary aspect of the scattering behavior of each pixel may then be studied by comparing the second order statistics of $\mathbf{k}(\omega_i)$ for different spectral locations, i.e., by testing the following hypothesis:

$$H : \Sigma_1 = \dots = \Sigma_R = \Sigma. \quad (21.148)$$

As it is shown in [46] this hypothesis can be tested easily using sample coherency matrices obtained from n_i independent realizations or looks of each TF target vector, $\mathbf{k}(\omega_i)$:

$$\mathbf{T}_i = \frac{1}{n_i} \sum_{l=1}^{n_i} \mathbf{k}_l(\omega_i) \mathbf{k}_l^H(\omega_i) \quad f(\mathbf{T}_i | \Sigma_i) = \mathcal{W}_{\mathcal{C}}(n_i, \Sigma_i). \quad (21.149)$$

The TF stationary behavior of $\mathbf{k}(\omega_i)$ is evaluated by means of a Maximum Likelihood (ML) ratio, Λ , built from the R independent sample coherency matrices as follows:

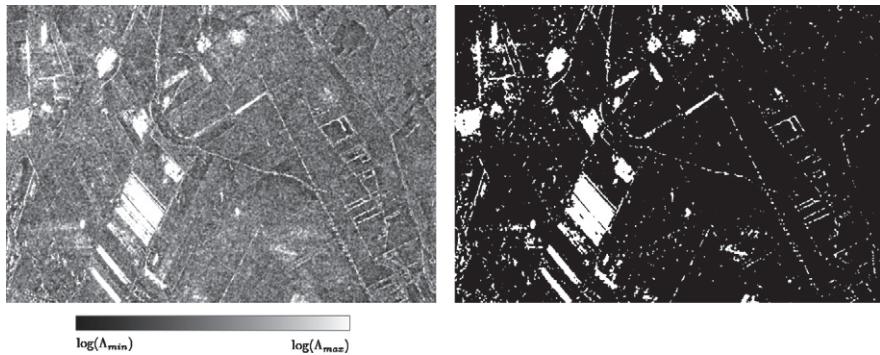
$$\Lambda = \frac{\max p(\mathbf{T}_1, \dots, \mathbf{T}_R | H \text{ true})}{\max p(\mathbf{T}_1, \dots, \mathbf{T}_R | H \text{ false})} = \frac{\max_{\Sigma} p(\mathbf{T}_1 | \Sigma, \dots, \mathbf{T}_R | \Sigma)}{\max_{\Sigma_1, \dots, \Sigma_R} p(\mathbf{T}_1 | \Sigma_1, \dots, \mathbf{T}_R | \Sigma_R)}. \quad (21.150)$$

Replacing the likelihoods in (21.150) by their expression and the expectations by their ML estimates, one gets the following simple expression [45]

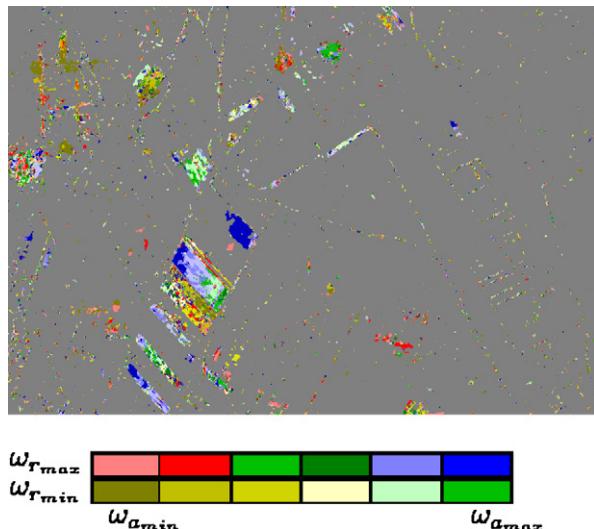
$$\Lambda = \frac{\prod_{i=1}^R \det(\mathbf{T}_i)^{n_i}}{\det(\mathbf{T}_t)^{n_t}} \quad \text{with} \quad \mathbf{T}_t = \frac{1}{n_t} \sum_{i=1}^R n_i \mathbf{T}_i \quad \text{and} \quad n_t = \sum_{i=1}^R n_i. \quad (21.151)$$

The hypothesis is accepted and the pixel under test is considered to have a stationary polarimetric TF behavior, with an arbitrarily chosen probability of false alarm P_{fa} , if $\Lambda > c_{\beta}$, where the relation between the threshold value and the probability of false alarm, $P_{fa}(c_{\beta}) = \beta$, has been derived in [45]. The ML ratio and nonstationary pixel map shown in Figure 21.51 indicate that an important number of pixels have a nonstationary behavior during the duration of the SAR acquisition. Most of the varying scatterers belong to agricultural fields affected by Bragg resonance. Complex targets and diffracting edges, whose scattering characteristics highly depend on the observation position, are discriminated over built-up areas and linear alignment of scatterers.

The ML ratio based detection approach may be further developed to determine nonstationary scattering behavior position in the range Doppler spectrum by comparing the contributions of each subspectrum

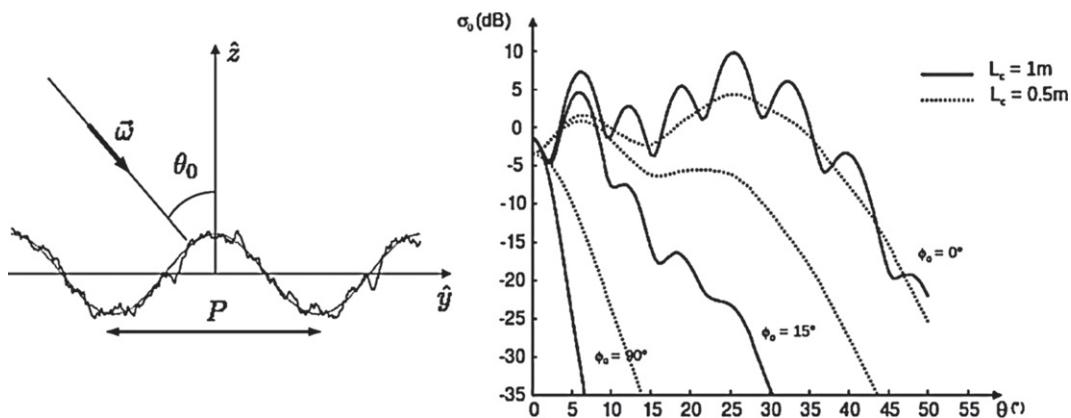
**FIGURE 21.51**

Discrimination of non-stationary scatterers. Image of the ML ratio in log-scale (left), non-stationary pixel map (right). Non stationary pixels are represented in white (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

**FIGURE 21.52**

Spectral location of the less stationary component among 12 range-azimuth subspectra for each nonstationary pixel.

image in the global ML ratio information [45]. It can be observed from the localization results displayed in Figure 21.52 on many fields affected by Bragg resonance that some groups of pixels, belonging to the same field, have a maximum anisotropic behavior in different subapertures. This is a consequence of the sliding effects of Bragg resonance on periodic structures, that is described in the next section.

**FIGURE 21.53**

Example of randomly perturbed periodic surface (left) and its associated backscattering coefficient, at L-band, with $B = 10 \text{ cm}$, $P = 1 \text{ m}$, $\epsilon = 9$, and $\sigma_h = 1 \text{ mm}$.

2.21.4.1.2.2 Analysis of Bragg resonant scattering over natural soils using a polarimetric 2-D TF continuous decomposition

In the scene under examination, Bragg resonance over agricultural fields is an important source of polarimetric variations. Bragg resonance is due to the coherent summation of simultaneously constructive contributions from a set of scatterers and is likely to happen during the observation of periodic surfaces or randomly irregular surfaces with a strong periodic component, as described in 1-D in Figure 21.53 A random surface, $h(x, y)$, with a quasi-periodic component in the y direction, can be described as

$$h(x, y) = B \cos\left(\frac{2\pi}{P}y\right) + \psi(x, y), \quad (21.152)$$

where P and B are the spatial period and amplitude, respectively, of the periodic component of $h(x, y)$, and the random perturbation term, $\psi(x, y)$, corresponds to an isotropic stationary random rough surface. This component is fully described by σ_h , the standard deviation of its zero mean Gaussian height probability density function, and φ_ψ , its correlation function. The Bragg resonance condition can be written as a function of the incident wavelength, λ , as

$$\omega_y = n \frac{2\pi}{P} \quad \text{or} \quad \sin \theta_0 \cos \phi_0 = \frac{n\lambda}{2P}, \quad (21.153)$$

where $\omega_y = \omega \sin \theta_0$ corresponds to the local amplitude of the ground wave vector at the, n is an unknown integer number indicating the mode of the resonance, θ_0 is the local angle of incidence and ϕ_0 the azimuthal angular difference between the observation position and the normal to the rows of the periodic surface. In the case of SAR measurements, ϕ_0 can be decomposed as $\phi_0 = \phi_t + \phi$, where ϕ_t represents the orientation of the surface with respect to the normal to the SAR platform flight track, and ϕ with the angle of observation in the azimuthal spectrum, as defined in (21.140). Yueh et al. [47] developed several approaches to model the scattering of electromagnetic waves from randomly

perturbed periodic surfaces. Their study reports that the influence of the resonating modes on the total backscattering response varies significantly with the surface parameters. As it can be seen in Figure 21.53, for a large surface correlation length, l_c , and for low values of the azimuth orientation angle, ϕ_0 , almost all the intensity peaks corresponding to different resonance modes can be discriminated. As l_c increases, the scattering pattern becomes smoother and only a few dominant resonance peaks can be observed. In the presence of resonance, the co-polarization returns S_{HH} and S_{VV} have almost identical values, characterized by a high intensity. As the resonant effect decreases, i.e., for high values of ϕ_0 , these polarimetric channels have distinct responses, with a significantly reduced amplitude. According to the resonance condition enounced in (21.153) similar anisotropic fields with different locations in range, θ_0 , or differently oriented, i.e., with different ϕ_t values may resonate at different azimuthal frequencies. If the resonance conditions cannot be satisfied for any azimuthal angle within the antenna aperture or if the surface scattering characteristics do not show a resonance peak, they also might not resonate at all [48].

Moreover, some fields may have parts resonating at different positions in the azimuthal frequency domain due to the joint dependence of the resonance condition on the incidence and azimuth angles, as seen in (21.153). This phenomenon is illustrated in Figure 21.54, where the location of a resonance peak is plotted as a function of the range and azimuth frequencies. As the azimuthal look angle, ϕ , varies, the set of incidence angles θ_0 satisfying Eq. (21.153) changes, leading to the apparition of sliding resonance stripes in the (f_{rg}, f_{az}) . The width of the resonance stripes is fixed by the width of the analyzing function in the azimuth and range directions, $(\Delta_g \omega_r, \Delta_g \omega_a)$ or equivalently $(\Delta_g f, \Delta_g \phi)$.

A range-azimuth continuous Time-Frequency analysis is performed over three points (P_1, P_2, P_3), located at different range positions inside a plowed field [48]. As depicted in Figure 21.55, results can

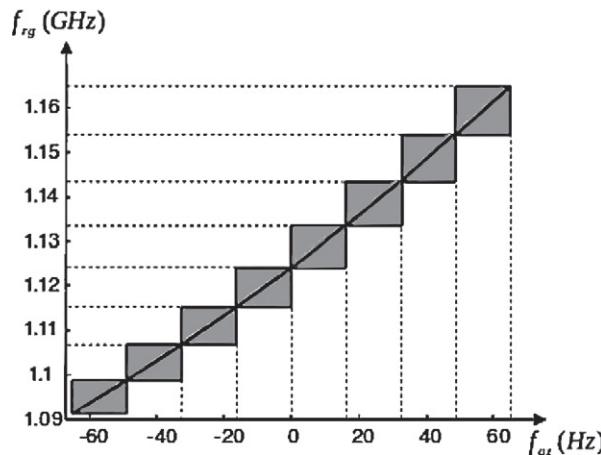
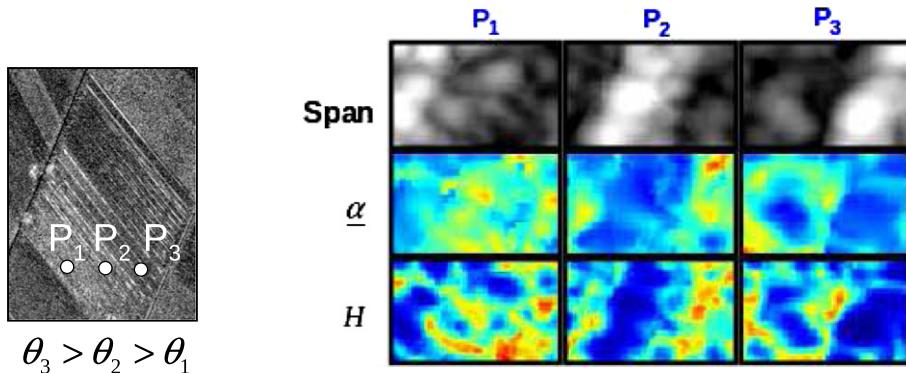


FIGURE 21.54

Location of resonance peaks in the (f_{rg}, f_{az}) plane. The solid line indicates the location of a resonance peak for a periodic surface characterized by $P = 0.6$ m and observed at L band. Gray areas indicate the location of potential resonance areas for each range-azimuth sub-spectrum.

**FIGURE 21.55**

Location of test points P_1 , P_2 , and P_3 and range azimuth frequency representation plane (left) and representation of polarimetric characteristics in this domain.

be represented, for each point, in the rangeazimuth frequency plane. The results of the timefrequency analysis, as shown in Figure 21.55, demonstrate that all three points under investigation do not have a stationary range and azimuth scattering behavior. Some (ω_r, ω_a) couples show high span values corresponding to low H and α . These observations agree with the predictions of the scattering model developed by Yueh et al. [47]. As the surface resonates, the co-polarization signals tend to be similar, involving a low α value, typical for surface reflection. This scattering mechanism is weighted by a strong intensity and dominates secondary intensities, potentially corresponding to multiple scattering terms, and results in a very low entropy value. This nonstationary behavior was found to have a preponderant influence on the polarimetric properties of resonating field at full resolution. Here, α and H values are significantly lower than those for similar fields that remained unaffected by Bragg resonance. The oblique resonating stripes, shown in the different range frequency planes in Figure 21.55, illustrate well the dependence of the resonance condition on both range and azimuth frequencies, as shown in Figure 21.54.

It can also be observed that as the incidence angle increases, from P_1 to P_3 , the oblique resonating stripe slides from low azimuth frequencies to higher ones. This displacement of the resonance locations is due to the dependence of the Bragg condition on the incidence angle and corroborates the analysis of the Bragg resonance as presented in Figure 21.54. Polarimetric indicators of pixels that do not belong to resonating stripes are unaffected by the Bragg resonance and have values similar to those observed over stationary fields.

2.21.4.1.3 TF polarimetric characterization of specific scatterers

2.21.4.1.3.1 Detection of point-like scatterers using the Internal Hermitian Product (IHP)

This approach, developed by Souyris et al. [49], was the first of a series of works on coherent TF analysis that fully exploit the physical nature of SAR signal together with polarimetric diversity. Its objective is to assess the joint use of the magnitude and the phase of a SAR polarimetric image for point target detection

and analysis. The detection principle is based on the fact that scattering by point like scatterers is a coherent process, i.e., during the SAR integration waves follow a specular path. Their T-F response, after an adequate phase compensation should ideally be constant, in practice remain coherent. Over speckle affected environments, scattering is due to wave diffusion by a large number of scatterers in a random and spatially uncorrelated way. In consequence their spectral response should show a low level of correlation.

Single polarization IHP. A natural way to detect the presence of a correlated signal, i.e., a target, embedded in uncorrelated noise, i.e., the clutter, is to compute the normalized correlation between two samples:

$$\rho_{TF} = \frac{E(s(\omega_1)s^*(\omega_2))}{\sqrt{E(|s(\omega_1)|^2)E(|s(\omega_2)|^2)}}. \quad (21.154)$$

In an ideal configuration, $s(\omega) = t + c(\omega)$, where t represents the ideally constant target response and $c(\omega)$ the uncorrelated clutter contribution $E(c(\omega_1)c(\omega_2)^*)$ if the samples originate from independent spectra. In this case the TF normalized correlation writes

$$\rho_{TF} = \frac{|t|^2}{|t|^2 + \sigma_c^2} = \frac{1}{1 + SCR^{-1}} \quad \text{with} \quad SCR = \frac{|t|^2}{\sigma_c^2}. \quad (21.155)$$

where it was assumed that the variance of the clutter response σ_c^2 was constant over the spectral domain. The correlation is than a function of the Signal to Clutter power Ratio (SCR), i.e., when a coherent object with a strong responses is illuminated, it can be detected by thresholding $|\rho_{TF}|$. As reported in [49], such an approach may lead to poor result if the SCR is not high enough and another approach is proposed, based on the IHP, defined as

$$IHP = E(s(\omega_1)s^*(\omega_2)). \quad (21.156)$$

It is simply the un-normalized correlation between two signal samples. As it is shown on Figure 21.56 this approach permits to improve the contrast between a target and its background. One may argue that

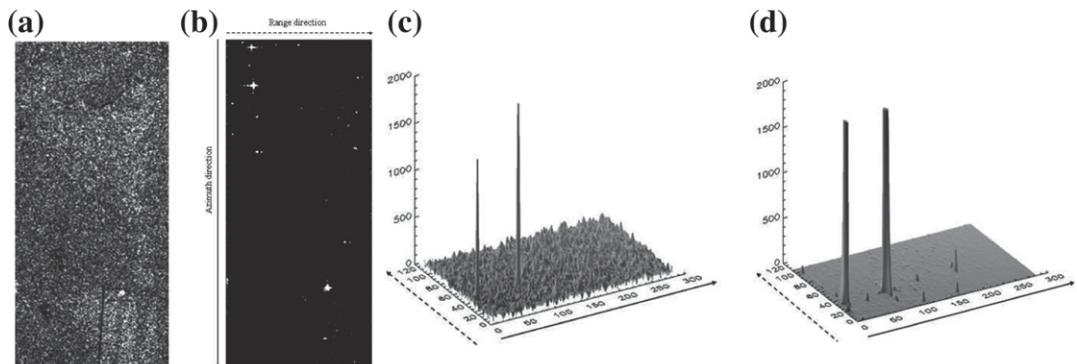


FIGURE 21.56

(a) original L-band SAR image acquired by the ONERA/RAMSES sensor at L band, (b) 2-look IHP image, and (c) 3-D view of the upper part of the image, (d) 3-D view of the upper part of the IHP image. From [49].

such an approach may be more associated to strong scatterer selection, based on their TF coherence properties, rather than detection, since the amplitude IHP is not normalized.

Polarimetric IHP. The IHP concept is exported to the polarimetric case [49] using the definition of the polarimetric correlation used in polarimetric SAR interferometry (POL-inSAR) [15]

$$\text{IHP}(\mathbf{w}) = \mathbb{E}(s_{\mathbf{w}}(\omega_1)s_{\mathbf{w}}^*(\omega_2)) \quad (21.157)$$

where $s_{\mathbf{w}}(\omega) = \mathbf{w}^\dagger \mathbf{k}(\omega)$, $\mathbf{k}(\omega) = [S_{hh}(\omega) \quad \sqrt{2}S_{hv}(\omega) \quad S_{vv}(\omega)]^T$ and \mathbf{w} is a projection unitary vector that permits to select one or a linear combination of polarization channels. This vector is then tuned over the space of unitary 3-element complex vectors to find the optimal value, ω_{opt} , that maximizes $|\text{IHP}(\mathbf{w})|$. The polarimetric properties of the target response are then given by $\mathbf{k}_t = \omega_{\text{opt}}$ that may be processed through classical decomposition techniques to obtain specific polarimetric indicators. An example of application to polarimetric SAR data is given in Figure 21.57.

2.21.4.1.3.2 Detection of coherent scatterers and their polarimetric characteristics

Persistent Scatterers (PS), whose response remain constant along time, are of particular interest for differential interferometry applications to subsidence monitoring [50]. Since it is likely that scatterers having stable TF responses, i.e., when observed from different positions and different frequencies, remain stable in time too, studies were led to investigate which types of constructions or objects actually behave as Coherent Scatterers.

2.21.4.1.3.3 CS detection based on TF entropy

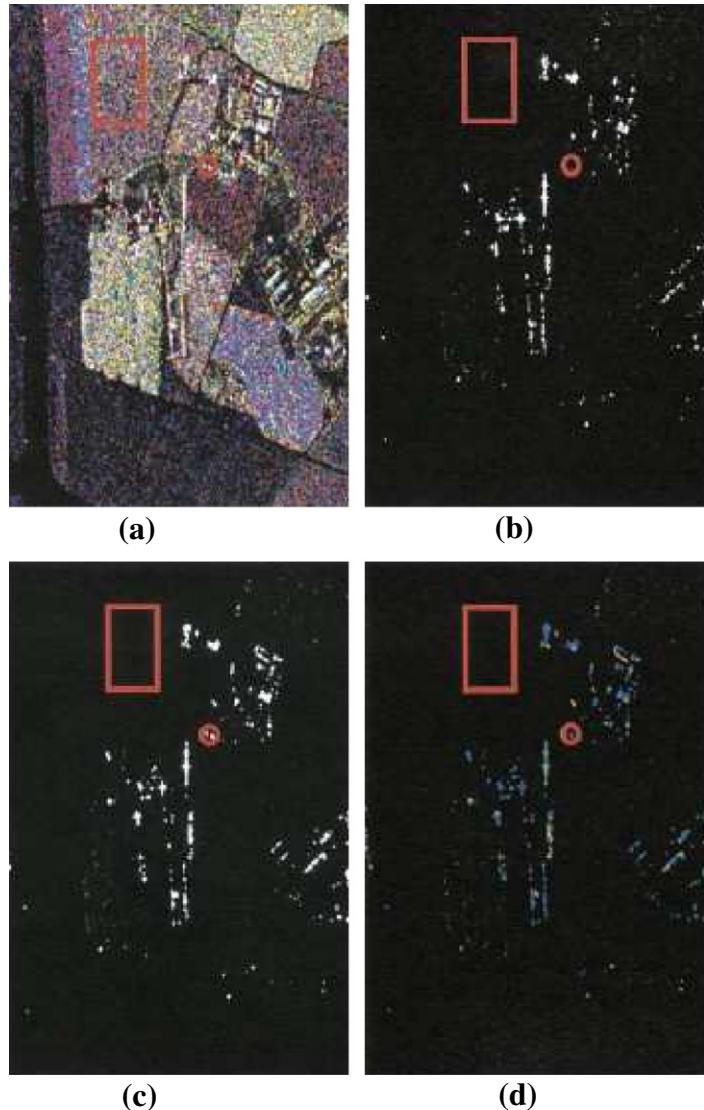
In [51] the normalized correlation approach of (21.154) is compared to a new stability indicator, the TF entropy, H_{TF} . Unlike (21.154), this indicator can handle more than two samples at a time. The signal of a given polarization channel is sampled at R different spectral locations, over independent or slightly overlapping spectral domains to create a TF signal vector, \mathbf{s}_{TF} and estimate its $(R \times R)$ covariance matrix, \mathbf{C}_{TF} :

$$\mathbf{s}_{\text{TF}} = [s(\omega_1), \dots, s(\omega_R)]^T \quad \mathbf{C}_{\text{TF}} = \mathbb{E}(\mathbf{s}_{\text{TF}}\mathbf{s}_{\text{TF}}^\dagger). \quad (21.158)$$

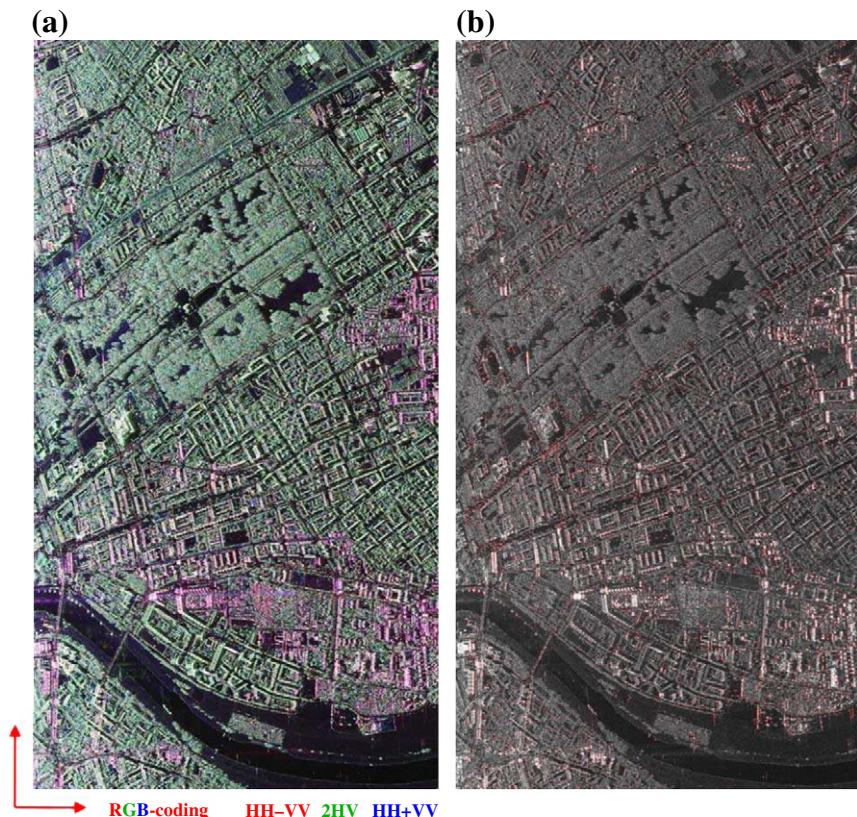
Similarly to the single-image polarimetric case, a TF entropy can be computed as [51]

$$H_{\text{TF}} = - \sum_{i=1}^R p_i \log_R(p_i) \quad \text{with} \quad p_i = \frac{\lambda_{\text{TF}_i}}{\sum_{i=1}^R \lambda_{\text{TF}_i}}, \quad (21.159)$$

where λ_{TF_i} is the i th eigenvalue of \mathbf{C}_{TF} . If different TF samples are maximally correlated, the TF entropy equals 0, whereas for totally uncorrelated signals having the same amplitude, H_{TF} tends toward 1. The analysis conducted over the city of Dresden indicates that many CS can be found over all the polarimetric channels (Figures 21.58 and 21.59). In the central part of the image is located the city center of Dresden, whereas in the central upper part a park with forested and vegetated areas can be recognized. Note that individual buildings and building blocks oriented parallel to the azimuth direction are characterized by a strong dihedral component, while buildings/blocks that have an orientation angle with respect to

**FIGURE 21.57**

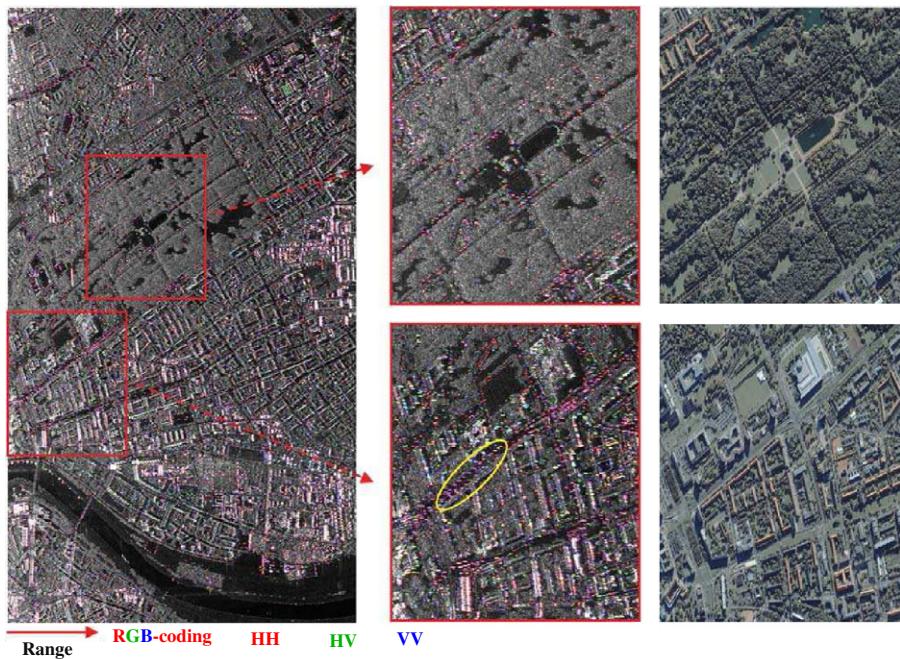
Polarimetric 2L-IHP. (Red circle) Test over point targets. (Red square) Test over speckle. (a) ONERA/RAMSES polarimetric SAR scene (5-m resolution, L-band) Red is HH . Green is HV . Blue is VV . (b) 2L-IHP, polarization HH . (c) POL-2L-IHP of the SAR scene. (d) Detected target analysis. Red indicates a dominant odd number of reflections. Green a dipole-like scattering. Blue a dominant double-bounce effects. From [49] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

**FIGURE 21.58**

(a) Color-coded polarimetric image of Dresden. (b) Detected coherent scatterers (in red) in HH polarization, superimposed to the intensity image. From [51] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

azimuth flight direction are characterized by a strong cross-polarized scattering component. One may note the position of the CSs in the region of the park (in the walking promenades and on the central place within the park) and in the dense urbanized areas (on the corners of buildings, along the streets, and on the two bridges over the Elbe River).

The detected CSs are characterized, in general, by high amplitudes, and low polarimetric entropy values. Their majority has a man-made character, a fact that predicts a relative high temporal stability. Because of their strong polarized behavior polarimetric acquisition diversity increases significantly the performance of CSs detection. Indeed for the Dresden data set the number of detected CSs by using fully polarimetric data, and an optimization similar to (21.157), is by a factor up to ten higher

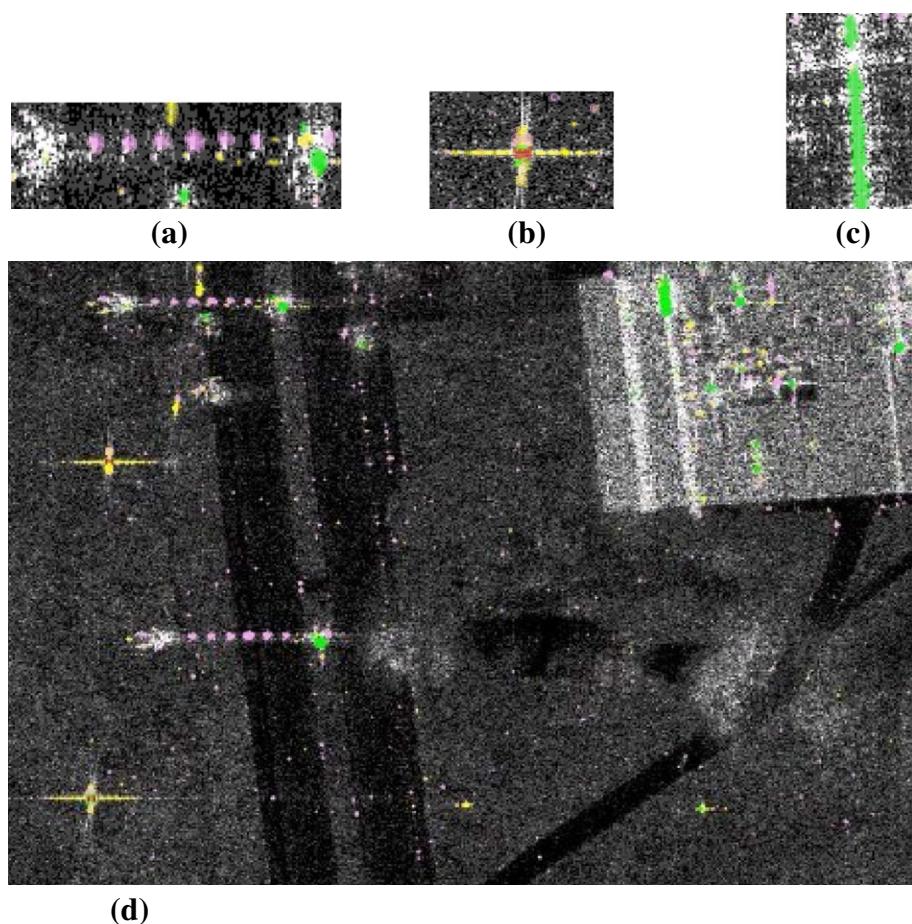
**FIGURE 21.59**

(Left) Detected CSs in lexicographic polarimetric basis. (Central upper) Details of the park region. (Central lower) Details of a dense urban area. (Right upper) Ikonos image of the Groer Garten Park. (Right lower) Ikonos image of the dense urban area. From [51].

compared to the number of CSs detected by using single-polarization data. The fact that the majority of the CSs is not depolarizing, and can be described by its scattering matrix, makes polarimetric information essential for the characterization, interpretation, and information extraction from individual CSs [51].

2.21.4.1.3.4 CS detection based on multiple criteria

In [52], a study of High Resolution X-band images is conducted over urban structures. The stability of the TF responses using a feature vector containing the coefficient of variation of the signal intensity, the TF entropy and gradient-based indicators of the stability of the continuous TF response. A statistical analysis of this feature vector is led to discriminate scatters having a coherent or unstable behavior in frequency, in azimuth or in both directions. A polarimetric analysis is led over specific objects. At such a high resolution, the TF behavior of objects could be used for their recognition using automatic techniques. An illustration is given in Figure 21.60.

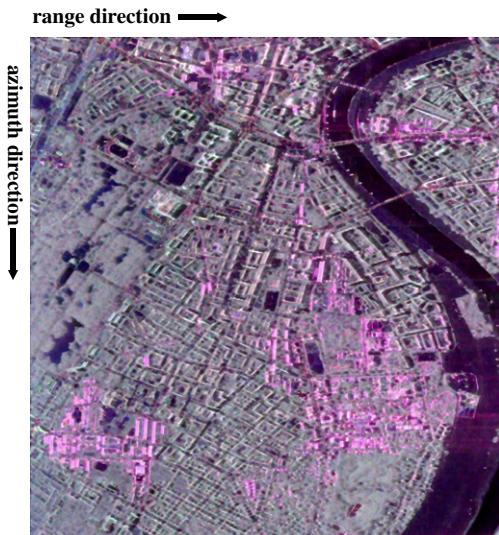
**FIGURE 21.60**

Examples of target point classification. Resolution 10 cm. (Red) Frequency invariant. (Purple) Range variant. (Green) Azimuth variant. (Gray levels) 2-D variant. (a) Resonant in range (ladder). (b) Frequency invariant (triangular). (c) Azimuth variant (building corner). (d) Classification on the entire scene. From [52] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

2.21.4.1.4 *Polarimetric Time-Frequency characterization of a dense urban area*

2.21.4.1.4.1 Polarimetric Time-Frequency features

Figure 21.61 shows a color-coded polarimetric SAR image of the city of Dresden acquired by DLR's E-SAR sensor data at L-band. The scene is mainly composed of built-up areas including vegetation spots. A forest and a park can be seen on the left part of the image and a river with smooth banks is located on the right part.

**FIGURE 21.61**

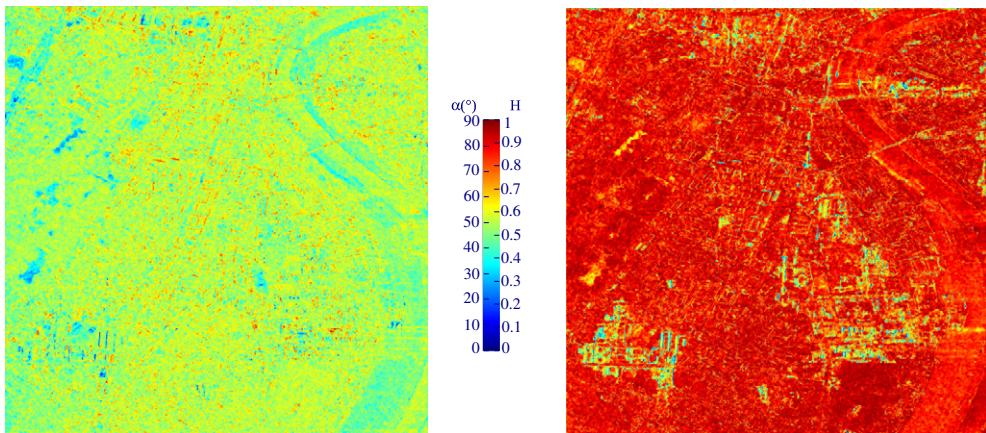
Color coded image of the Dresden test site (Pauli basis).

Polarimetric properties of media are generally investigated through a decomposition of second order multivariate polarimetric representations. The resulting parameters provide information on the media geometrical structure and on the underlying scattering mechanisms. Two parameters, obtained from the well known eigenvector-based decomposition introduced in [13] are displayed in Figure 21.62.

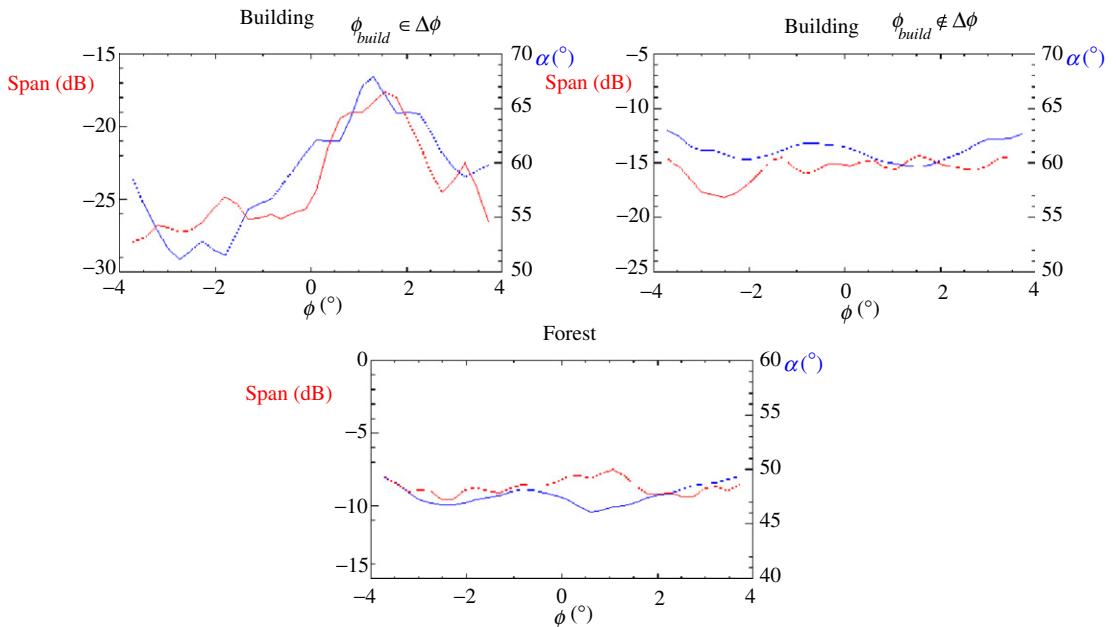
The entropy image shown in Figure 21.62 reveals that the polarimetric behavior of most of the scene is highly random. Over urban areas, the polarimetric response is composed of a large number of different polarimetric contributions originating from complex building structures as wall as from surrounding vegetation. The resulting high entropy involves that an interpretation of polarimetric indicators may not be relevant. Over buildings aligned with the flight track direction, the entropy has intermediate values and the α parameter reveals the presence of dominant single and double bounce reflexions. Buildings which do not face the radar track are characterized by a strong cross-polarization component and high entropy and can hardly be discriminated from vegetated areas.

Figure 21.63 presents a continuous TF analysis in the azimuth direction of three different media: a building facing the radar, an oriented building, and a forested area. The SPAN, corresponding to the sum of the intensities in all polarimetric channels, and the polarimetric α angle are computed for the different media at each frequency location and mean values are then estimated over pixels belonging to the object.

A non stationary behavior is clearly visible in Figure 21.63 with a sudden a large variation of both SPAN and α levels with the observation angle in azimuth ϕ . This anisotropic behavior is due to the highly directional patterns of coherent scattering mechanisms which may occur as the radar faces a large

**FIGURE 21.62**

Polarimetric parameter images, $\underline{\alpha}$ (left) and H (right).

**FIGURE 21.63**

Continuous TF analysis in the azimuth direction (SPAN in red, $\underline{\alpha}$ in blue) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

artificial structure, such as a building [53]. This particular effect can only be observed if the building orientation with respect to the radar flight track falls within the processed antenna azimuth aperture.

On the contrary, oriented buildings and vegetated areas, like forest patches, show stationary behaviors. The identification of buildings from their TF response thus requires an additional criterion to complement the stationarity information. It is known that man-made objects are likely to have a coherent response, whereas natural media may be considered as random. The discrimination of such responses can be achieved by studying the coherence of the backscattered polarimetric signal in the Time-Frequency domain and requires the use of an adequate TF polarimetric SAR (PolSAR) signal model.

2.21.4.1.4.2 PolSAR TF signal modeling and analysis

PolSAR data TF model. The proposed TF signal model [54] is given by the following expression, where the spatial coordinates, \mathbf{I} , have been omitted:

$$\mathbf{s}(\omega) = \mathbf{t}(\omega) + \mathbf{c}(\omega). \quad (21.160)$$

The signal $\mathbf{s}(\omega)$ contains the full coherent polarimetric polarization information and can be associated to a well-known scattering vector [13]:

$$\mathbf{k}(\omega) = \frac{1}{\sqrt{2}} [S_{hh}(\omega) + S_{vv}(\omega), S_{hh}(\omega) - S_{vv}(\omega), 2S_{hv}(\omega)]^T, \quad (21.161)$$

where $S_{pq}(\omega)$ represents an element of the (2×2) scattering matrix \mathbf{S} sampled at the frequency coordinates ω .

The signal described in (21.160) is composed of two contributions:

- The term $\mathbf{t}(\omega)$ is highly coherent and can be associated to a deterministic or almost deterministic target response. Depending on the structure of the observed object, the response can remain constant during the SAR acquisition, or can be non-stationary if the backscattering behavior is sensitive to the azimuth angle of observation or illumination frequency.
- The second term, $\mathbf{c}(\omega)$, represents the response of distributed environments. It is uncorrelated, but may follow a non-stationary behavior in particular cases, e.g., vegetated terrains with a strong topography, very dense environments whose response results from the sum of a large number of uncorrelated contributions.

This composite model may be tested using $\mathbf{s}(\omega)$ second order statistics:

- The coherence of $\mathbf{s}(\omega)$ can be used to determine the dominant component within the pixel under consideration. A high value indicates that $\mathbf{t}(\omega)$ is the most important term in (21.160), and a low one corresponds to scattering from an incoherent, distributed, medium.
- The stability of the dominant component can then be tested by studying the stationarity of the variance of $\mathbf{s}(\omega)$.

Second order statistics. Due to the signal high dimensionality, usual scalar tools are not well adapted to the study of second-order TF polarimetric statistics. A polarimetric TF target vector is built by gathering

the PolSAR information sampled at R spectral coordinates $\omega_i, i = 1, \dots, R$.

$$\mathbf{k}_{\text{TF}} = \left[\mathbf{k}^T(\omega_1), \dots, \mathbf{k}^T(\omega_R) \right]^T. \quad (21.162)$$

The sampling coordinates, ω_i , and the frequency domain resolution of the analyzing function g are chosen so that the R sub-spectra do not overlap and span the whole full resolution spectrum [48]. A polarimetric TF sample covariance matrix, $\mathbf{T}_{\text{TF-Pol}}$, is then computed as follows

$$\mathbf{T}_{\text{TF-Pol}} = \left\langle \mathbf{k}_{\text{TF}} \mathbf{k}_{\text{TF}}^\dagger \right\rangle = \begin{bmatrix} \mathbf{T}_{11} & \cdots & \mathbf{T}_{1R} \\ \vdots & \ddots & \vdots \\ \mathbf{T}_{R1} & \cdots & \mathbf{T}_{RR} \end{bmatrix},$$

where $\mathbf{T}_{ij} = \left\langle \mathbf{k}(\omega_i) \mathbf{k}(\omega_j)^\dagger \right\rangle.$ (21.163)

Non-stationary pixel discrimination. Stationarity is assessed by testing the fluctuations of the variance of the signal at the different spectral locations [46,48]. In the polarimetric case, the signal sample variance is given by a (3×3) polarimetric coherency matrix, i.e., by the diagonal terms of the $\mathbf{T}_{\text{TF-Pol}}$ matrix: $\{\mathbf{T}_{ii}\}_{i=1,\dots,R}$. The polarimetric TF response is considered as stationary if the sample \mathbf{T}_{ii} matrices, assumed to follow independent complex Wishart distributions $\mathbf{T}_{ii} \sim W_C(n_i, \Sigma_{ii})$ with n_i looks, have the same expectation Σ . The corresponding hypothesis is given by:

$$H_0 : \Sigma_{11} = \cdots = \Sigma_{RR} = \Sigma. \quad (21.164)$$

The corresponding Maximum Likelihood (ML) ratio is:

$$\Lambda = \frac{\max_{\Sigma} L(\Sigma, \dots, \Sigma)}{\max_{\Sigma_{ii}} L(\Sigma_{11}, \dots, \Sigma_{RR})}. \quad (21.165)$$

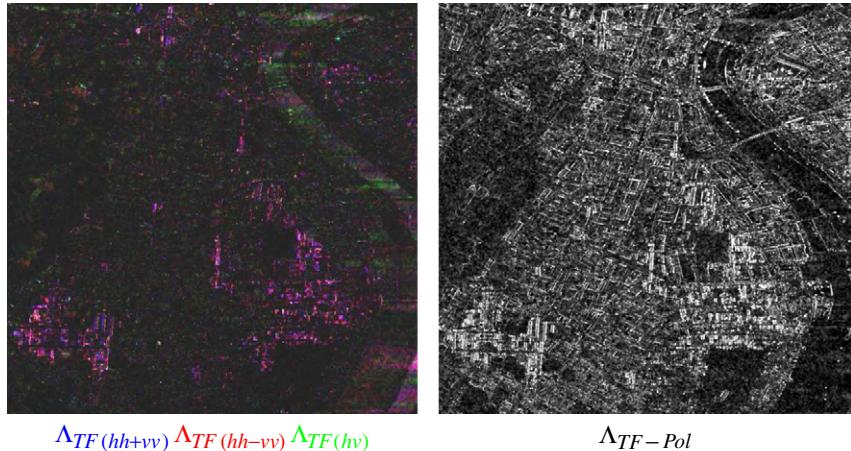
The likelihood terms in (21.165) are maximized by replacing the expectation matrices by the ML estimates and the hypothesis is tested using the resulting ML test [45], which takes the following form:

$$\Lambda = \frac{\prod_{i=1}^R |\mathbf{T}_i|^{n_i}}{|\mathbf{T}_t|^{n_t}}, \quad (21.166)$$

where $\mathbf{T}_t = \frac{\sum_{i=1}^R n_i \mathbf{T}_{ii}}{\sum_{i=1}^R n_i}$ and $n_t = \sum_{i=1}^R n_i.$

Figure 21.64 presents a log-image of the Λ parameter on the Dresden test site, obtained with four spectral coordinates in the azimuth direction over the Dresden test site.

The Λ parameter reach high values over natural areas indicating a stationary spectral behavior. Over buildings, Λ decreases, pointing out the invalidity of the stationary hypothesis over such objects. Highly anisotropic pixels, such as those corresponding to the wall-ground dihedral reflection or specular reflection from oriented roofs are clearly identified in Figure 21.64 due to their very low stationary aspect.

**FIGURE 21.64**

Non stationary TF behavior indicator, Δ_{TF} , computed separately for different polarimetric channels (left) simultaneously using the whole polarimetric information (right).

Coherent pixel discrimination. In [51], the eigenvalues of a single-polarization covariance matrix have been used to derive a coherency indicator. These eigenvalues carry information on the correlation structure, but are also sensitive to potential PolSAR fluctuations due to non-stationarity. A solution has been proposed in order to overcome this limitation and to jointly use all the polarimetric channels [54]. Under the hypothesis of uncorrelated spectral responses, the off-diagonal terms of the TF covariance matrix verify:

$$H_0 : \Sigma_{ij} = \mathbf{0} \quad \forall i \neq j. \quad (21.167)$$

The corresponding ML ratio is given by:

$$\Theta = \frac{\max_{\Sigma_{ii}} L(\Sigma_{11}, \dots, \Sigma_{RR})}{\max_{\Sigma_{TF}} L(\Sigma_{TF})} = \frac{|\mathbf{T}_{TF}|^{n_i}}{\prod_{i=1}^R |\mathbf{T}_{ii}|^{n_i}}. \quad (21.168)$$

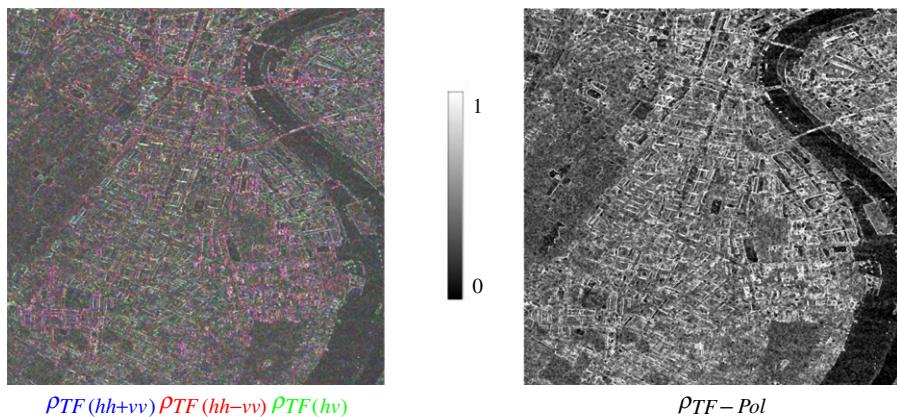
This ML ratio expression can be rewritten as:

$$\Theta = \left| \tilde{\mathbf{T}}_{TF} \right|^{n_i} \quad (21.169)$$

with

$$\tilde{\mathbf{T}}_{TF-Pol} = \begin{bmatrix} \mathbf{I} & \boldsymbol{\Gamma}_{12} & \cdots & \boldsymbol{\Gamma}_{1R} \\ \boldsymbol{\Gamma}_{12}^\dagger & \mathbf{I} & & \vdots \\ \vdots & & \ddots & \vdots \\ \boldsymbol{\Gamma}_{1R}^\dagger & \cdots & \cdots & \mathbf{I} \end{bmatrix}, \quad (21.170)$$

where $\boldsymbol{\Gamma}_{ij} = \mathbf{T}_{ii}^{-1/2} \mathbf{T}_{ij} \mathbf{T}_{jj}^{-1/2}$.

**FIGURE 21.65**

TF coherence indicator, ρ_{TF} , computed separately for different polarimetric channels (left) simultaneously using the whole polarimetric information (right).

The normalized covariance matrix, $\tilde{\mathbf{T}}_{\text{TF-Pol}}$ results from the whitening of the TF polarimetric covariance matrix by the separate polarimetric information at each frequency location. This representation is then insensitive to spectral polarimetric intensity variations and is characterized by its off-diagonal matrices Γ_{ij} which can be viewed as an extension of the scalar normalized correlation coefficient to the polarimetric case. The ML ratio in (21.168) is a function of the eigenvalues of $\tilde{\mathbf{T}}_{\text{TF-Pol}}$, which reflect the correlation structure: flat for decorrelated responses ($\tilde{\mathbf{T}}_{\text{TF-Pol}} \rightarrow \mathbf{I}_d$), heterogeneous for correlated ones. Taking into account $\tilde{\mathbf{T}}_{\text{TF-Pol}}$ peculiar form, a correlation indicator, named TF-Pol coherence, can be defined as [54]:

$$\rho_{\text{TF-Pol}} = 1 - \left| \tilde{\mathbf{T}}_{\text{TF-Pol}} \right|^{\frac{1}{3K}}. \quad (21.171)$$

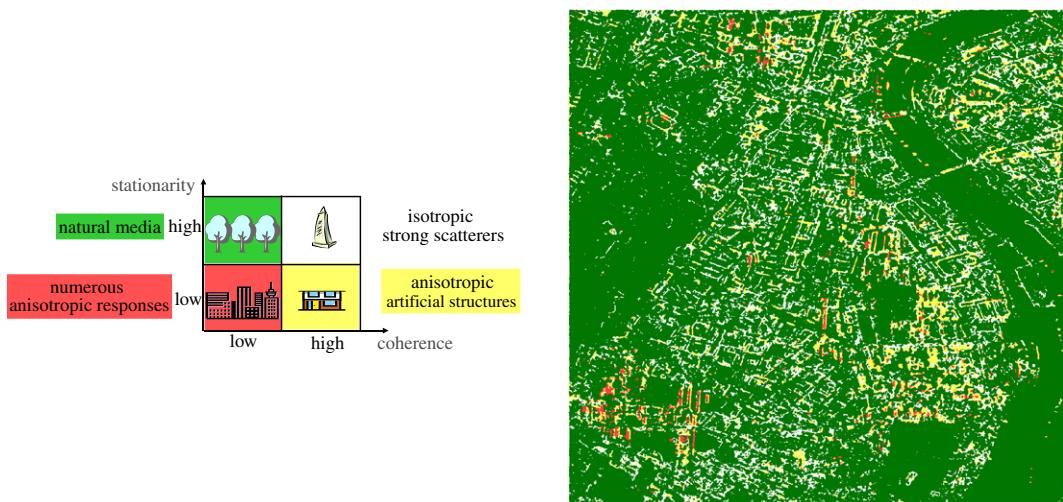
Figure 21.65 presents an image of $\rho_{\text{TF-Pol}}$ over the Dresden test site, computed from four spectral locations in the azimuth direction.

As expected, the TF-Pol coherence is high over buildings due to the presence of strong coherent reflectors. It can be also noticed that buildings are identified independently of their orientation.

2.21.4.1.4.3 PolSAR TF analysis

PolSAR TF classification. The stationarity and coherence indicators derived above can be merged to classify the scene. Both $\rho_{\text{TF-Pol}}$ and Λ parameters are thresholded and combined into four classes. The application of the fusion strategy over the Dresden site is shown in Figure 21.66.

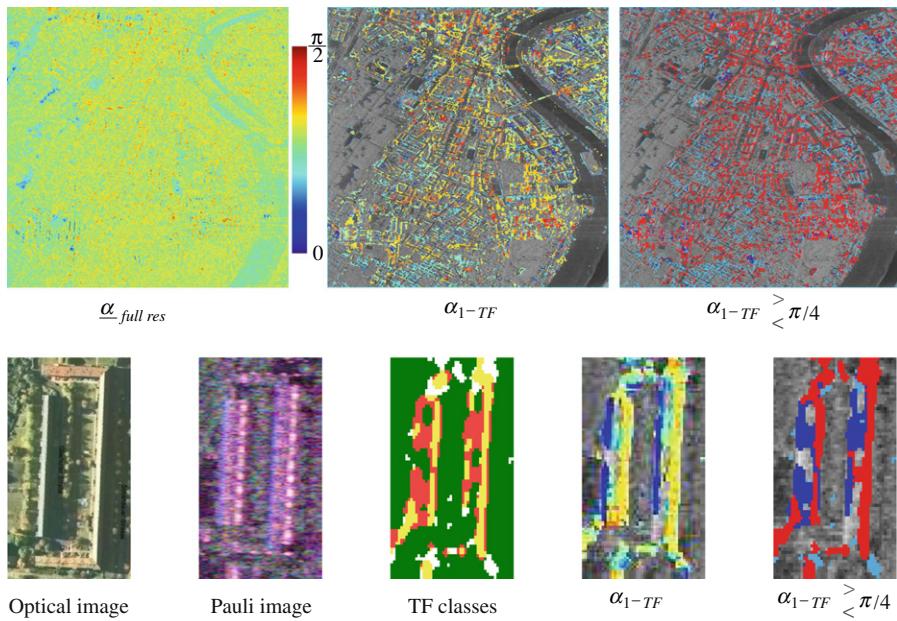
The resulting map permits a good estimation of building locations. A physical interpretation can be given for each of the four classes:

**FIGURE 21.66**

TF polarimetric classification. Classification scheme (left) results obtained over Dresden (right).

- *Coherent and stationary pixels (white class):* The \mathbf{t} term in (21.160) is dominant and constant during the SAR acquisition. This kind of behavior corresponds to strong scatterers with an isotropic response, like oriented buildings, lamp-posts,
- *Coherent and non stationary pixels (yellow class):* The \mathbf{t} contribution dominates but varies during the measure, causing fluctuation of the signal with the azimuth angle of observation. This anisotropic effect is characteristic of buildings facing the radar track, whose response is affected by a strong and highly directive pattern, mainly due double bounce reflections or specular single bounce reflection over roofs tilted toward the radar.
- *Incoherent and stationarity pixels (green class):* The uncorrelated component \mathbf{c} dominates and has stable second order statistics. This class corresponds to natural environments (forests, fields, grass areas, ...) of distributed artificial media such as roads, roof tops, terraces, ...
- *Incoherent and non-stationary pixels (red class):* This class indicates the presence of complex scattering contributions, which change during the SAR integration, and sum-up in an incoherent way, like in layover areas.

TF cleaning of PolSAR data. As it was shown in Figure 21.62, the full resolution PolSAR information can hardly be used to analyze the scene geophysical properties due to a very high entropy inherent to the study of dense environments. The proposed PolSAR TF analysis technique can also be used to improve in a significant way the interpretation of polarimetric indicators. The most coherent TF scattering mechanisms is described by the first eigenvector of $\tilde{\mathbf{T}}_{\text{TF-Pol}}$, which can be transformed back to the H-V polarimetric basis using a matrix \mathbf{P} , satisfying $\tilde{\mathbf{T}}_{\text{TF-Pol}} = \mathbf{P} \mathbf{T}_{\text{TF-Pol}} \mathbf{P}^\dagger$. From this eigenvector, one can extract an α_{TF} parameter which shows a much more contrasted and relevant information than

**FIGURE 21.67**

Application of the classification scheme to building characterization: over the whole image (top), over the an isolated building (bottom).

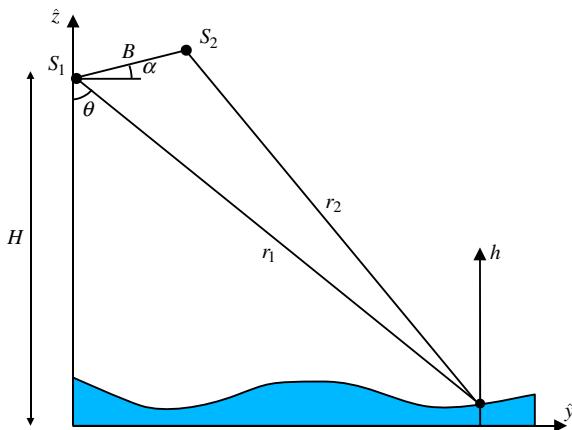
the original full resolution parameter α . Figure 21.67 shows different images of a building of the scene. A comparison between the T-F classification results and the optical image reveals that both the double bounce reflexion is considered as a non-stationary and coherent scattering mechanism, whereas the roof layover is seen as non-stationary and uncorrelated, due to the superposition of the roof and ground contributions. A thresholding of α_{TF} with respect to $\pi/4$, permits to easily separate these two different mechanisms and could be used to get a rough estimate of the building height. Such an information might be useful to interferometric phase unwrapping algorithms which generally face ambiguity issues over urban areas.

2.21.4.2 Analysis of volumetric media using polarimetric SAR interferometry

2.21.4.2.1 A brief introduction to SAR interferometry

2.21.4.2.1.1 Relation between interferometric phase and the height of a scatterer

The geometrical configuration of an interferometric acquisition, depicted in Figure 21.68, shows two sensors, located at slightly different positions, that measure the coherent SAR response of a scatterer. The relative positions of the sensors is measured by the baseline $B \ll H$ and the angle α . After focusing

**FIGURE 21.68**

Basic geometrical configuration of an interferometric SAR acquisition.

of the SAR signals and co-registration of the SAR images, the response of the scatterer is given by

$$s_i = |a_{ci}| \exp\left(-j \frac{4\pi}{\lambda_c} r_i + j\phi_{obji}\right) \quad i = 1, 2, \quad (21.172)$$

where ϕ_{obji} represents the absolute and unknown phase of the response of each object. For a small baseline and if the acquisitions are made within a sufficiently short amount of time, the amplitude and phase of the object responses may be considered identical, $|a_{c1}| = |a_{c2}|$, $\phi_{obj1} = \phi_{obj2}$. In this case, the interferometric phase difference, or interferogram, is given by

$$\Delta\phi_{12} = \arg(s_1 s_2^*) = -\frac{4\pi}{\lambda_c} \Delta r_{12}. \quad (21.173)$$

This interferometric phase being proportional to the path difference, Δr_{12} , it is sensitive to both the elevation of the scatterer with respect to a reference plane and to the range distance r_1 . A local linearization permits to isolate the elevation-dependent part φ_{topo} :

$$\Delta r_{12} \approx \Delta r_{12}(r_1) + \Delta r_{12}(z) \Rightarrow \Delta\phi_{12} \approx \varphi_{fe}(r) + \varphi_{topo}(z), \quad (21.174)$$

where $\varphi_{fe}(r)$ can be computed from the geometrical configuration and hence compensated and $\varphi_{topo}(r_1)$ is given by

$$\varphi_{topo}(z) = k_z z \quad \text{with} \quad k_z = \frac{4\pi B \cos(\theta - \alpha)}{\lambda r_1 \sin \theta}. \quad (21.175)$$

One may note that the estimation of the interferometric phase as the argument of a complex exponential may imply a wrapping due to the limited domain of definition of the \arg operator. As a consequence

$$k_z z \in \mathbb{R} \quad \text{but} \quad \varphi_{topo}(z) \in [-\pi, \pi]. \quad (21.176)$$

This ambiguity can generally be overcome using phase unwrapping techniques.

2.21.4.2.1.2 Interferometric coherence

Over homogeneous distributed environments, the acquired are random variables whose statistics may be studied using the 2-element vector, $\mathbf{s} = [s_1 \ s_2]^T$, following a complex normal distribution, $\mathbf{s} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C})$, where the covariance matrix \mathbf{C} may be written as

$$\mathbf{C} = \mathbb{E}(\mathbf{s}\mathbf{s}^\dagger) = \mathbb{E}\left(\begin{bmatrix} |s_1|^2 & s_1 s_2^* \\ s_1^* s_2 & |s_2|^2 \end{bmatrix}\right) = \begin{bmatrix} I_1 & \sqrt{I_1 I_2} \gamma \\ \sqrt{I_1 I_2} \gamma^* & I_2 \end{bmatrix}, \quad (21.177)$$

where

$$\gamma = \frac{\mathbb{E}(s_1 s_2^*)}{\sqrt{\mathbb{E}(|s_1|^2)\mathbb{E}(|s_2|^2)}} \quad (21.178)$$

is called the interferometric coherence and may be used to estimate the interferometric phase $\Delta\phi_{12} = \varphi = \arg(\gamma)$. Using L independent realizations, or looks, of \mathbf{s} , ML estimates may be computed as [21, 22]

$$\widehat{\mathbf{C}} = \frac{1}{L} \sum_{l=1}^L \mathbf{s}(l)\mathbf{s}^\dagger(l) \quad \widehat{\gamma} = \frac{\widehat{\mathbf{C}}(1, 2)}{\sqrt{\widehat{\mathbf{C}}(1, 1)\widehat{\mathbf{C}}(2, 2)}} \quad \widehat{\varphi} = \arg(\widehat{\gamma}). \quad (21.179)$$

In addition to the need of a large number of looks, L , a good estimate of the interferometric phase requires a large value of the modulus of the coherence, since $\text{var}(\widehat{\varphi}) \rightarrow 0$ when $|\gamma| \rightarrow 1$ and $\text{var}(\widehat{\varphi})$ reaches a maximum as $|\gamma| \rightarrow 0$.

The coherence may be modeled as a product of a large number of contributions, as, for instance:

$$\gamma = \gamma_{\text{SNR}} \gamma_{\text{proc}} \gamma_{\text{temp}} \gamma_{\text{geom}} \gamma_z \quad \text{with} \quad |\gamma_i| \leq 1. \quad (21.180)$$

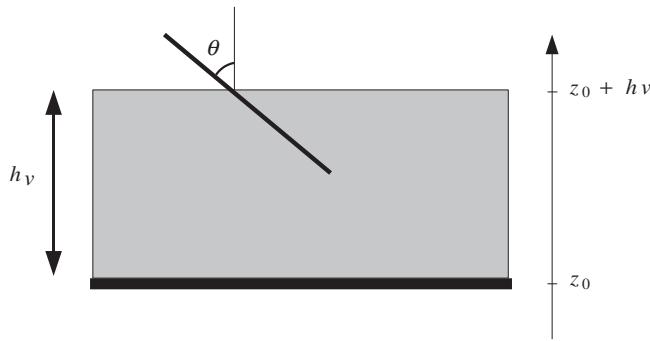
Each element represent a potential source of decorrelation between the acquired signals, due to the presence of additive noise, errors during the acquisition, processing and referencing of the signals, to changes of the scenes between the acquisitions, to spectral shifts induced by acquisition geometry and to vertical decorrelation induced by the presence of a volumic medium. Very few of these can be compensated a posteriori and each of them may influence the phase estimation accuracy since

$$|\gamma| \leq \min_i |\gamma_i|. \quad (21.181)$$

2.21.4.2.1.3 The random volume over ground model for a single polarization channel

The last term of the coherence decomposition in (21.180) is related to the structure of the observed medium through an integral relation [15, 55–58]

$$\gamma_z = \frac{\int_{-\infty}^{+\infty} f(z)e^{ik_z z} dz}{\int_{-\infty}^{+\infty} f(z)dz}, \quad (21.182)$$

**FIGURE 21.69**

Synopsis of a RVOG.

where $f(z)$ represents the equivalent density of coherent reflectivity in the elevation direction. The interferometric coherence may then be used to guess some characteristics of the vertical structure of the observed volume, provided that the other terms in (21.180) are known or compensated. In order to estimate the possibilities of such an approach, one may use a simple model of a homogeneous Random semi-transparent Volume lying over a Ground consisting of a rough surface, the RVOG model, depicted in Figure 21.69 [56,58].

The volume is assumed to be made of particles having the same reflectivity $A_v d_z$ and the same constant extinction properties. i.e.,

$$df_v(z) = -\frac{\kappa_e}{\cos \theta} f_v(z) dz \quad \text{and} \quad f_v(z_0 + h_v) = A \quad (21.183)$$

then

$$\begin{cases} f_v(z) = Ae^{-2\frac{\kappa_e}{\cos \theta}(z_0+h_v-z)}, & \forall z \in [z_0, z_0 + h_v], \\ 0 & \text{otherwise,} \end{cases} \quad (21.184)$$

where the factor 2 account for round-trip attenuation. The ground reflectivity distribution is then given by

$$f_g(z) = f_v(z_0) \sigma_g \delta(z - z_0). \quad (21.185)$$

The intensities backscattered by the ground and the volume are given by

$$I_g = \int_{-\infty}^{+\infty} f_g(z) dz = f_v(z_0) \sigma_g \quad \text{and} \quad I_v = \int_{-\infty}^{+\infty} f_{vol}(z) dz. \quad (21.186)$$

The general unresolved expression of the coherence is then

$$\gamma_z = \frac{\int_{-\infty}^{+\infty} (f_v(z) + f_g) e^{j k_z z} dz}{\int_{-\infty}^{+\infty} f_v(z) + f_g(z) dz} = \frac{\int_{z_0}^{z_0+h_v} f_v(z) e^{j k_z z} dz + I_g e^{j k_z z_0}}{I_v + I_g}. \quad (21.187)$$

In order to provide a physical interpretation for expression (21.187), cases with different levels of complexity can be investigated.

Case 0: Transparent volume. In this extremely simple case, the volume response is null, i.e., $I_v = \kappa_e = 0$. The coherence

$$\gamma_z = e^{j k_z z_0} = e^{j \phi_0} \quad (21.188)$$

has a unitary modulus and its argument provides the elevation of the ground.

Case 1: Lossless volume and no ground. The extinction of waves when they travel through the volume is null, $\kappa_e = 0$ and the ground reflectivity is null, $I_g = 0$. In this case γ_z represents the coherence behavior of the volume only

$$f_v(z) = A \rightarrow \gamma_z = \gamma_{v_0} = e^{j \phi_0} e^{j k_z \frac{h_v}{2}} \operatorname{sinc}\left(\frac{k_z h_v}{2}\right). \quad (21.189)$$

The coherence argument indicates exactly the absolute elevation of the center of the volume, whereas its modulus depends on its width h_v . This case is illustrated in Figure 21.70.

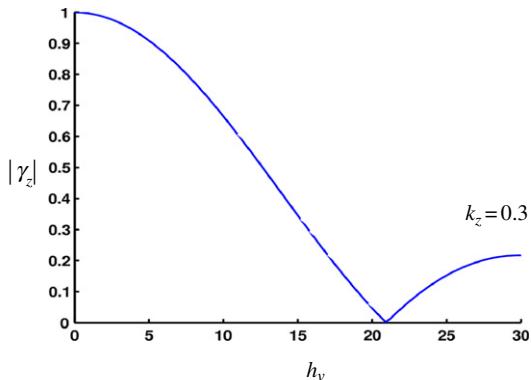


FIGURE 21.70

$|\gamma_z|$ for $\kappa_e = I_g = 0$.

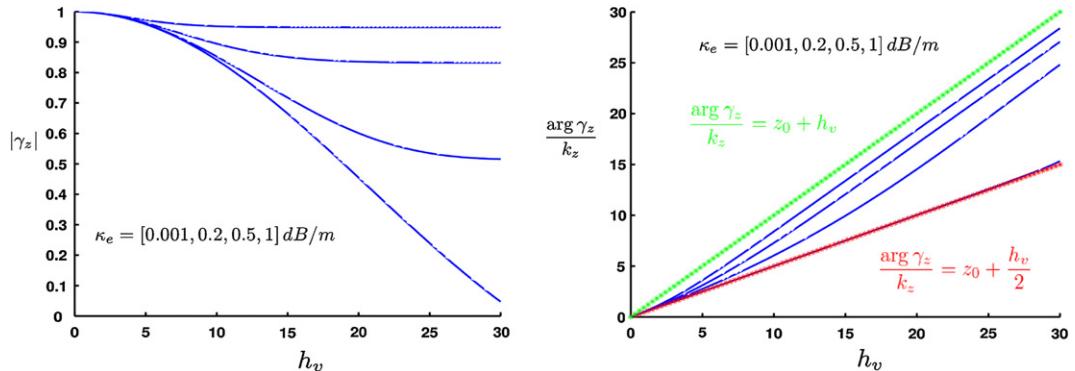


FIGURE 21.71

\$|\gamma_z|\$ for \$\kappa_e \neq 0\$ and \$I_g = 0\$.

Case 2: Volume with constant linear extinction and no ground. The extinction \$\kappa_e\$ is constant over the volume and the ground reflectivity is null, \$I_g = 0\$. The expression of the coherence is

$$\gamma_z = \gamma_v = \frac{\int_{z_0}^{z_0+h_v} e^{p_1 z} dz}{\int_{z_0}^{z_0+h_v} e^{p z} dz} = e^{j\phi_0} \frac{p}{p_1} \frac{e^{p_1 h_v} - 1}{e^{ph_v} - 1} \quad \text{with} \quad p = \frac{2\kappa_e}{\cos \theta}, \quad p_1 = p + jk_z. \quad (21.190)$$

The magnitude and phase behaviors of \$\gamma_v\$ are illustrated in Figure 21.71 for different values of extinction. For a quasi-null extinction, the coherence magnitude has a sinc-like shape and the phase center is located at the center of the volume. As the extinction increases, the magnitude pattern becomes flatter and close to 1 \$\forall h_v\$ when \$\kappa_e\$ is large. The phase center moves then towards the higher interface of the volume. This behavior is due to the fact that as \$\kappa_e\$ increases the wave penetration depth inside the volume decreases and the particles that participate more to the scattering phenomenon are those located near the upper interface. For reduced penetration depths, the decorrelation induced by the numerator of (21.190) becomes negligible.

Case 3: Volume with constant linear extinction and ground. The extinction \$\kappa_e\$ is constant over the volume and the ground reflectivity is non null. The expression of the coherence is [55, 56, 58]

$$\gamma_z = \frac{\int_{z_0}^{z_0+h_v} e^{p_1 z} dz + I_g e^{j\phi_0}}{\int_{z_0}^{z_0+h_v} e^{p z} dz + I_g} = \frac{\gamma_v + \mu e^{j\phi_0}}{1 + \mu} \quad \text{with} \quad \mu = \frac{I_g}{I_v}, \quad (21.191)$$

where \$\mu\$, called the ground to volume intensity ratio, can be considered as a key parameter of the RVOG coherence which steers the coherence \$\gamma_z\$ between extreme behaviors, \$\gamma_z \rightarrow \gamma_v\$ for \$\mu \rightarrow 0\$ and \$\gamma_z \rightarrow e^{j\phi_0}\$ \$\mu \rightarrow +\infty\$. For intermediate values a single coherence term is not sufficient for analyzing the RVOG model.

2.21.4.2.2 Pol-inSAR representations

Polarimetric and interferometric signals consist of two polarimetric target vectors, $\mathbf{k}_1, \mathbf{k}_2$, acquired from slightly different positions. As it has been mentioned for polarimetric TF signal processing, the computation of the correlation between multivariate signals requires to use projectors. The Pol-inSAR coherence has then been defined by Cloude and Papathanassiou as:

$$\gamma(\mathbf{w}_1, \mathbf{w}_2) = \frac{\mathbb{E}(s_1 s_2^*)}{\sqrt{\mathbb{E}(|s_1|^2)\mathbb{E}(|s_2|^2)}} \quad \text{with } s_i = \mathbf{w}_i^\dagger \mathbf{k}_i, \arg(\mathbf{w}_1^\dagger \mathbf{w}_2) = 0 \quad (21.192)$$

and \mathbf{w} is a projection unitary vector that permits to select one or a linear combination of polarization channels. This coherence can be obtained from the (6×6) POL-inSAR coherency matrix, defined as:

$$\mathbf{k}_6 = \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} \Rightarrow \mathbf{T}_6 = \mathbb{E}(\mathbf{k}_6 \mathbf{k}_6^\dagger) = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{12}^\dagger & \mathbf{T}_{22} \end{bmatrix} \quad \text{with } \mathbf{T}_{ij} = \mathbb{E}(\mathbf{k}_i \mathbf{k}_j^\dagger). \quad (21.193)$$

The coherence can then be expressed as

$$\gamma(\mathbf{w}_1, \mathbf{w}_2) = \frac{\mathbf{w}_1^\dagger \mathbf{T}_{12} \mathbf{w}_2}{\sqrt{\mathbf{w}_1^\dagger \mathbf{T}_{11} \mathbf{w}_1 \mathbf{w}_2^\dagger \mathbf{T}_{22} \mathbf{w}_2}}. \quad (21.194)$$

Cloude and Papathanassiou have found the projections defining the optimal scattering mechanisms that maximize $|\gamma(\mathbf{w}_1, \mathbf{w}_2)|$ [15]. Nevertheless using different projection vectors $\mathbf{w}_1 \neq \mathbf{w}_2$ is, in general not recommended. Indeed the phase and magnitude of coherences computed between different polarimetric channels depend on the interferometric configuration, but also on the polarimetric correlation within each separate image. In general, the POL-inSAR coherence is simplified to

$$\gamma(\mathbf{w}, \mathbf{w}) = \gamma(\mathbf{w}) = \frac{\mathbf{w}^\dagger \mathbf{T}_{12} \mathbf{w}}{\sqrt{\mathbf{w}^\dagger \mathbf{T}_{11} \mathbf{w} \mathbf{w}^\dagger \mathbf{T}_{22} \mathbf{w}}}. \quad (21.195)$$

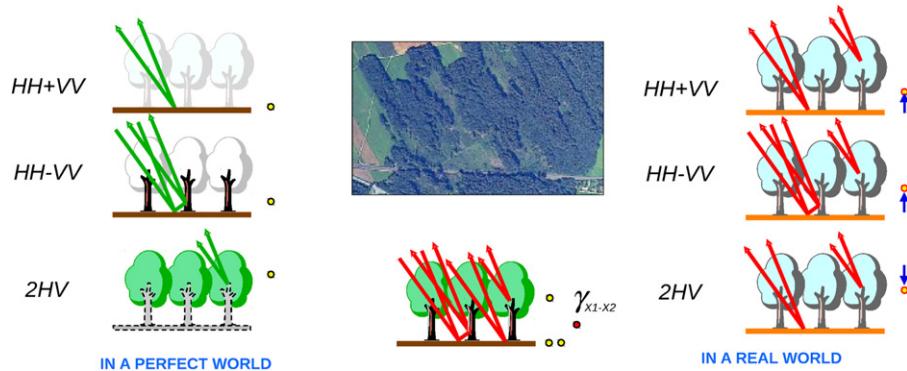
2.21.4.2.3 Polarimetric RVOG model

The principles of a polarimetric solution to the RVOG model are given in Figure 21.72.

The objective is to isolate specific polarimetric scattering mechanisms \mathbf{w}_g and \mathbf{w}_v that null I_v or I_g , respectively. In an ideal world, $\gamma_z(\mathbf{w}_g) = e^{j\phi_0}$ and $\gamma_z(\mathbf{w}_v) = \gamma_v$, and a priori knowledge on the domain of variation of κ_e permits to estimate h_v with a reasonable precision. In general, such RVOG polarimetric eigen-mechanisms do not exist, i.e., $I_v(\mathbf{w}), I_g(\mathbf{w}) > 0, \forall \mathbf{w}$. An original solution, proposed in [58], is summarized in the following. It is based on the hypothesis that the reflectivity and extinction within the volume do not depend on polarization:

$$I_v(\mathbf{w}) = I_v \quad \forall \mathbf{w} \quad \text{and} \quad \kappa_e(\mathbf{w}) = \kappa_e \quad \forall \mathbf{w}. \quad (21.196)$$

This hypothesis correspond to a azimuthal symmetric form for the coherency matrix and is then reasonable for a dense volume of uniformly oriented particles. Under this hypothesis, the only polarization-dependent parameter is the response of the ground $I_g(\mathbf{w})$. The polarimetric version of the coherence

**FIGURE 21.72**

Ideal and realistic scattering configuration over a volumetric natural environment (Courtesy of Dr. Kostas Papathanassiou).

given in (21.191) can be expressed as:

$$\gamma_z(\mathbf{w}) = \frac{\gamma_v + \mu(\mathbf{w})e^{j\phi_0}}{1 + \mu(\mathbf{w})} \quad \text{with} \quad \mu(\mathbf{w}) = \frac{I_g(\mathbf{w})}{I_v}. \quad (21.197)$$

This relation can be written under the following form

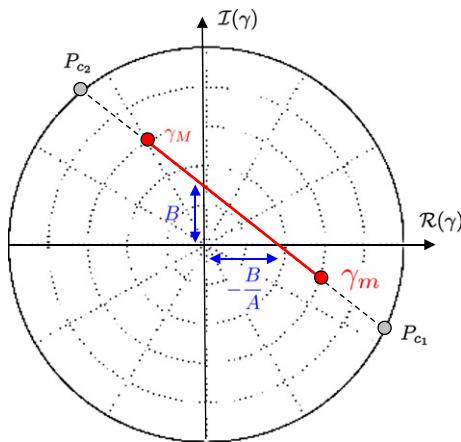
$$\gamma_z(\mathbf{w}) = e^{j\phi_0} + L(\mathbf{w})(\gamma_v - e^{j\phi_0}) \quad \text{with} \quad 0 \leq L(\mathbf{w}) = \frac{1}{1 + \mu(\mathbf{w})} \leq 1 \quad (21.198)$$

which is the parametric equation of a line segment, between $e^{j\phi_0}$ and γ_v in the complex plane. This means that when the RVOG model is valid, complex coherences obtained for arbitrary polarimetric projection vectors are located on a line segment in the complex plane, as illustrated on Figure 21.73.

2.21.4.2.4 Estimation of the height of a RVOG using POL-inSAR data

2.21.4.2.4.1 Vegetation bias removal

In practice the ground to volume intensity ratio $\mu(\mathbf{w})$ never reaches the limits of its domain of definition $0, +\infty$, i.e., $L_{\min} \leq L \leq L_{\max}$. For a ground contribution with a sufficiently low entropy, there exist a scattering mechanism, typically HV , so that $I_g(\mathbf{w}_v) \ll I_g(\mathbf{w}_g)$ and in this case $\gamma_z(\mathbf{w}_v) \approx \gamma_v$. Due to the lack of polarization selectivity of volume scattering, the extremity of the line segment corresponding to L_{\min} cannot be used to estimate the ground elevation. Cloude and Papathanassiou proposed to overcome this problem by extrapolating the line segment until it intercepts the unitary circle, since the RVOG model for a transparent medium, $\gamma_z = \gamma_g = e^{j\phi_0} \rightarrow |\gamma_g| = 1$. This physical interpretation permits to reduce an important elevation bias linked with the completely depolarized volume response.

**FIGURE 21.73**

RVOG line segment within the complex unitary circle.

2.21.4.2.4.2 Practical vegetation height estimation

In [55], Cloude proposes a theoretically sub-optimal but practically robust method to estimate the height of a volume from POL-inSAR data sets.

Line segment. From a L-look sample Pol-inSAR coherency matrix, estimate the line parameters, e.g., $\hat{\gamma}_m$ and $\hat{\gamma}_M$ in Figure 21.73. It is recommended to use the analytical expressions provided in [59] instead of a least-square fit based on sample coherences generated from (21.195) and a set of pre-defined or random projection vectors. Indeed the approach in [59] is immediate and corresponds to the generation of an infinite number of coherences.

Volume and ground coherences. This can be done by selecting one of the segment extremity $\hat{\gamma}_m$ and $\hat{\gamma}_M$ as $\hat{\gamma}_v$, e.g., one may chose the one closer to $\gamma(hv)$. Extrapolate the other extremity of the segment to the unitary circle in order to determine $\hat{\phi}_0$.

Estimate the height assuming an infinite extinction. As mentioned earlier, for an infinite extinction value, the phase center of $\hat{\gamma}_v$ correspond to the upper interface of the volume. Then

$$\hat{h}_\infty = \frac{\arg(\hat{\gamma}_v) - \hat{\phi}_0}{k_z}. \quad (21.199)$$

It is unlikely that $\kappa_e = \infty$, so \hat{h}_∞ will indeed correspond to a phase center located within the volume. This value might underestimate the true volume height.

Estimate the height assuming an null extinction. For a null extinction, the volume coherence magnitude as a sinc-like shape. The corresponding height might be estimated as

$$\hat{h}_0 = 2 \frac{\text{sinc}^{-1}(|\hat{\gamma}_v|)}{k_z}. \quad (21.200)$$

It is also unlikely that $\kappa_e = 0$, i.e., that the phase center is at its lowest possible elevation. This value might underestimate the true volume height.

Empirical linear combination. The final height estimate is

$$\hat{h}_v = \hat{h}_\infty + \epsilon \hat{h}_0. \quad (21.201)$$

Recommended ϵ value: 0.4

This semi-empirical approach reveals more robust than a theoretically more correct one, since it combines uses separate estimates based on the modulus and phase of $\hat{\gamma}_v$, linearly combines estimates with opposed biases and does not use numerical values of the extinction that main cause severe artifacts. Examples of results obtained over forested areas using polarimetric and interferometric SAR processing are given in Figures 21.74 and 21.75.

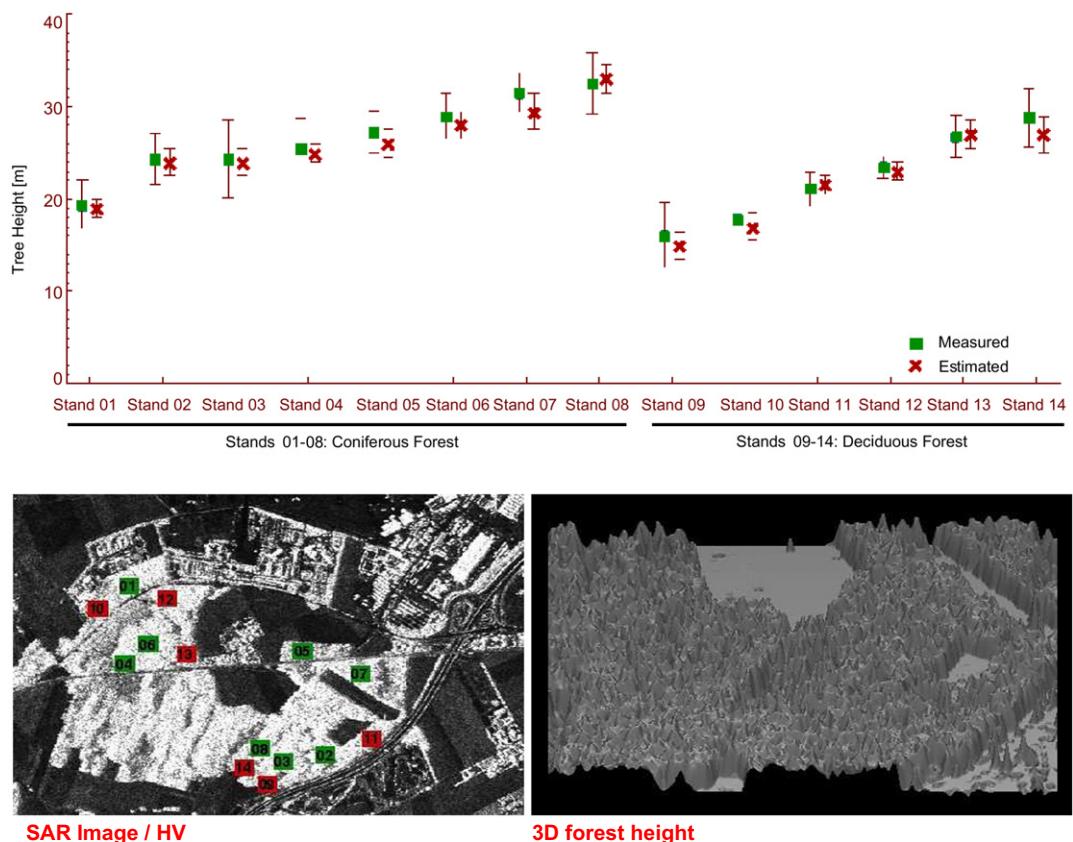
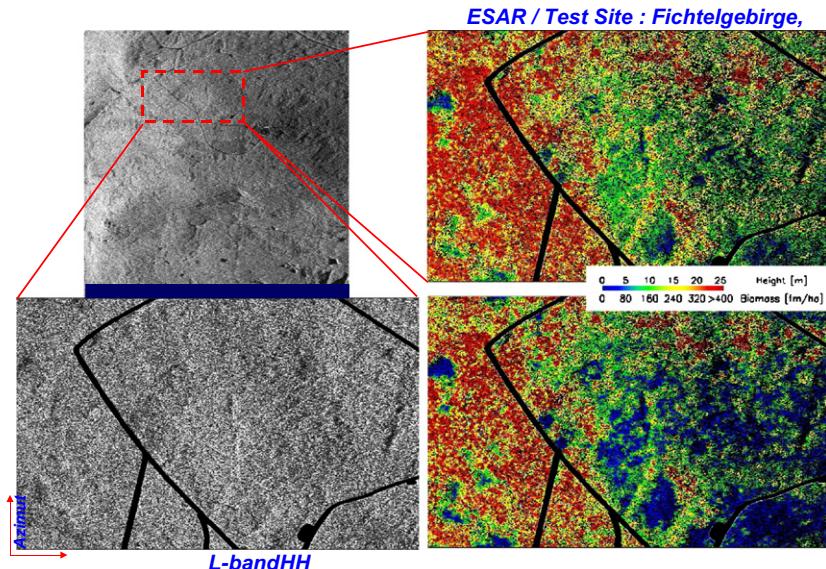


FIGURE 21.74

Forest height estimates over Oberpfaffenhofen, Germany, obtained from POL-inSAR data acquired at L band by the DLR ESAR sensor. Courtesy of Dr. Kostas Papathanassiou.

**FIGURE 21.75**

Forest height and biomass estimates over Fichtelgebirge, Germany obtained at L band. Courtesy of Dr. Kostas Papathanassiou.

2.21.4.3 Analysis of volumetric media using polarimetric SAR tomography

2.21.4.3.1 Basics of SAR Tomographic imaging

SAR tomographic imaging, using Multi-Baseline Interferometry (MB-InSAR), is based on the acquisition of M SAR signals along slightly shifted trajectories t_i , as illustrated on Figure 21.76. Each pair of trajectories is separated by a baseline. After focusing, compensating and georeferencing the acquired signals, M 2-D SAR images are obtained, $s_i(x, r)$, where x represents the azimuth location and $r = \sqrt{x^2 + y^2 + z^2}$, the proximal distance of a given point.

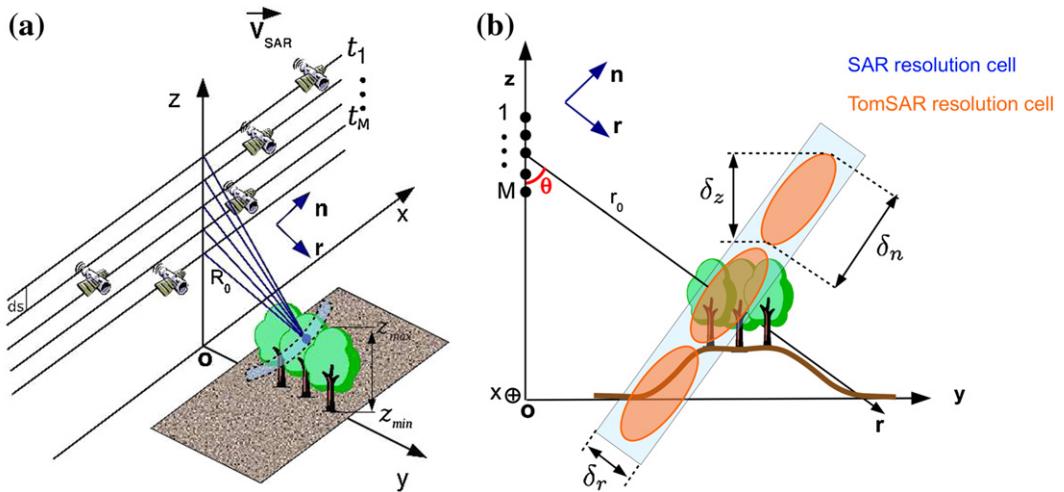
As shown on Figure 21.76, due to its intrinsic cylindrical ambiguity, classical 2-D SAR imaging is not well adapted to the characterization of volumic media. For a given focusing position (x_0, r_0) , the focused SAR results from the coherent integration of the reflectivity density $a(x, r, v)$ of the considered medium.

$$s_i(x_0, r_0) = \int_{\mathcal{C}(x_0, r_0)} a(x, r, z) e^{j k_z z} dz \quad i = 1, \dots, M, \quad (21.202)$$

where $\mathcal{C}(x_0, r_0)$ represents one resolution cell with a cylindrical slice layer shape, whose upper and lower limits are defined by the observed volumic media. The resolution cell limits are given by:

$$|x - x_0| < \frac{\delta x}{2}, \quad |r - r_0| < \frac{\delta r}{2}, \quad z_{v_{\min}} < z < z_{v_{\max}} \quad (21.203)$$

with δx and δr the SAR resolution in azimuth and distance respectively.

**FIGURE 21.76**

Geometrical configuration of an MB-InSAR measurement (a), SAR and TomSAR resolutions cells (b).

The formulation in (21.202) assumes that $a(x, r, z)$ is a function of the elevation position only, i.e., δ_x and δ_r are have sufficiently low values and variations in these directions within a resolution cell may be neglected. The vertical wavenumber, k_{z_i} accounts for variations of the interferometric phase from one image to the other and can be calculated from the geometric configuration of the acquisition. The purpose of Tomographic SAR (TomSAR) imaging, illustrated on Figure 21.76, is to improve the vertical resolution of the SAR measurement in order to estimate the volumic density of the reflectivity $|a(x, r, z)|^2$ in a more accurate way. To do so, the M acquired SAR signals $s_i(x_0, r_0)$, obtained for different numbers of vertical waves k_{z_i} are used in the frame of a classical problem of spectral estimation, for which a series of solutions can be applied [60, 61]. Using Multi-Baseline Polarimetric Interferometry (MB-PolInSAR), Polarimetric Tomographic SAR (PolTomSAR) imaging can provide the observed medium more precise localization and more detailed physical features. The first airborne TOMSAR experiment was conducted by the DLR using their ESAR sensor at L band over the test site of Oberpfaffenhofen [62]. The data set consist of 14 SAR images acquired over quasi-parallel tracks within a short period of time. This study revealed the fantastic power analysis of the combination of 3-D imaging with SAR. Some results of this work are presented in Figure 21.76 where it is shown that tomography can be used to reliably locate scatterers in elevation, estimate building heights and image forest canopies. Tomograms performed over different polarimetric channels can be combined to appreciate the 3-D polarimetric behavior of the media under observation.

Recently, other studies were performed to characterized natural volumetric media, i.e., forested areas, using SAR tomography combined with polarization diversity [63–65]. Similarly to the RVOG model mentioned above, the POLTOMSAR reponse of a forest is assumed to be made of two main components related to the ground and volume contribution. A specific Kronecker structure is then used to extract the vertical and polarimetric features of both components from a set of polarimetric images. Despite its mathematical exactness, the estimated components are affected by an ambiguity and several solutions

are possible. Fortunately, both the spatial and polarimetric domains covered by these possible solutions shrinks considerably as the number of images increases, i.e., as vertical resolution improves. One may note that this ambiguous aspect is common to all POL-inSAR techniques for such an application. Examples of tomograms obtained over volumetric environments are shown in Figures 21.77 and 21.78.

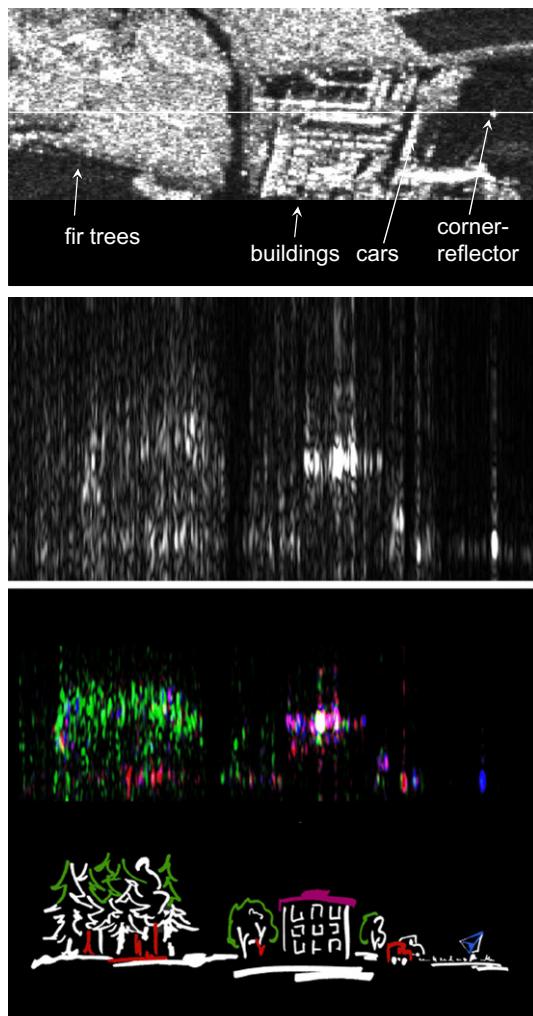
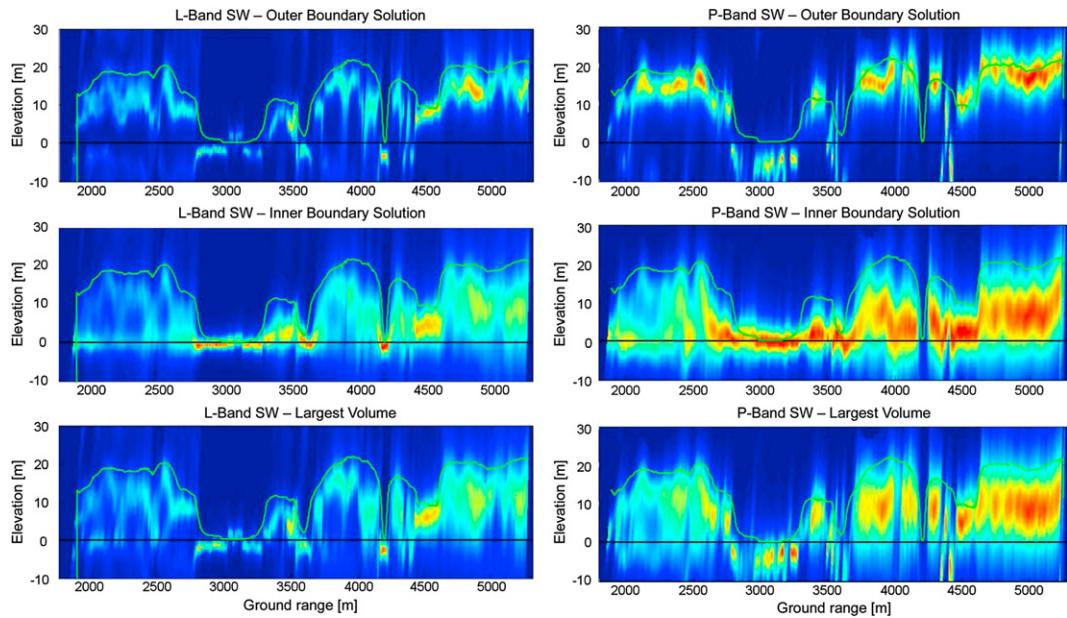


FIGURE 21.77

First airborne tomographic SAR experiment: HH SAR image at L band of the DLR buildings and surroundings (top) HH intensity tomogram estimated over a linear path indicated on the SAR image (middle) Color coded tomogram obtained from estimates performed over the Pauli basis polarimetric channels (bottom). Courtesy of Dr. Andreas Reigber.

**FIGURE 21.78**

Polarimetric SAR tomograms of a boreal forest performed at L (left) and P (right) over data acquired by the DLR within the BioSAR campaign. From top to bottom are represented different solutions that all fit very well with the H100 height obtained through LIDAR measurements and denoted by the green line. Courtesy of Dr. Stefano Tebaldini.

2.21.4.3.2 Tomographic signal models

Glossary of Notation

- \mathbf{A}^T = the transpose of the matrix \mathbf{A} ,
- \mathbf{A}^\dagger = the conjugate transpose of \mathbf{A} ,
- $\text{tr}(\mathbf{A})$ = the trace of $\mathbf{A} \in \mathbb{C}^{m \times m}$,
- $\mathbf{A} \odot \mathbf{B}$ = the Hadamard product of $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, defined by $[\mathbf{A} \odot \mathbf{B}]_{ij} = [\mathbf{A}]_{ij} [\mathbf{B}]_{ij}$,
- $\mathbf{A} \otimes \mathbf{B}$ = the Kronecker product of $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times q}$,
- $\|\mathbf{A}\|_{\mathbf{W}}^2$ = $\text{tr}(\mathbf{A} \mathbf{W} \mathbf{A}^\dagger)$, the weighted Frobenius norm of $\mathbf{A} \in \mathbb{C}^{m \times n}$,
- $\mathbf{1}_{(m \times n)}$ = the $(m \times n)$ matrix of ones,
- $E()$ = the expectation operator,
- $\mathcal{N}(\mathbf{m}, \boldsymbol{\Gamma})$ = complex gaussian distribution with mean \mathbf{m} and covariance matrix $\boldsymbol{\Gamma}$,
- $\hat{\theta} = \arg \max_{\theta} f(\theta)$ is the maximizing argument of $f(\theta)$,
- $\hat{\theta} = \arg \max_{\theta, \text{loc}} f(\theta) \in \mathbb{C}^{n \times 1}$ contains the values of θ corresponding to the n largest local maxima of $f(\theta)$.

2.21.4.3.2.1 TomSAR signal models

Considering an azimuth-range resolution cell that contains n_s backscattering contributions from scatterers located at different heights and assuming no decorrelation between the different acquisitions, the received data vector, $\mathbf{y} \in \mathbb{C}^{M \times 1}$, can be formulated as follows:

$$\mathbf{y}(l) = \sum_{i=1}^{n_s} s_i \mathbf{a}(z_i) + \mathbf{n}(l) = \mathbf{A}(\mathbf{z})\mathbf{s} + \mathbf{n}(l), \quad (21.204)$$

where $l = 1, \dots, L$ indicates one of the L independent realizations of the signal acquisition, also called looks. The source signal vector, $\mathbf{s} = [s_1, \dots, s_{n_s}]^T$, contains the unknown complex reflection coefficient of the n_s scatterers, and $\mathbf{n} \in \mathbb{C}^{M \times 1}$ represents the complex additive noise, assumed to be Gaussianly distributed with zero mean variance σ_n^2 , and to be white in time and space with, i.e., $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{(M \times M)})$ and $E(\mathbf{n}(l)\mathbf{n}^\dagger(k)) = \sigma_n^2 \mathbf{I}_{(M \times M)} \delta_{l,k}$. The steering vector $\mathbf{a}(z)$ contains the interferometric phase information associated to a source located at the elevation position z above the reference focusing plane and is given by:

$$\mathbf{a}(z) = [1, \exp(jk_{z_2}z), \dots, \exp(jk_{z_M}z)]^T, \quad (21.205)$$

where $k_{z_j} = \frac{4\pi}{\lambda} \frac{B_{\perp j}}{r_1 \sin \theta}$ is the two-way vertical wavenumber between the master and the j th acquisition tracks. The corresponding perpendicular baseline $B_{\perp j}$ is aligned with the cross-range direction. The carrier wavelength is represented by λ , whereas θ stands for the incidence angle and r_1 is the slant range distance between the master track and the scatterer. The steering matrix $\mathbf{A}(\mathbf{z})$ consists of n_s steering vectors corresponding each to a backscattering source:

$$\mathbf{A}(\mathbf{z}) = [\mathbf{a}(z_1), \dots, \mathbf{a}(z_{n_s})]^T \quad (21.206)$$

with $\mathbf{z} = [z_1, \dots, z_{n_s}]^T$, the vector of unknown source heights.

Considering now interferometric decorrelation between different acquisitions, the initial model in (21.204) may be reformulated as a sum of contributions from random sources [60]:

$$\mathbf{y}(l) = \sum_{i=1}^{n_s} \mathbf{x}_i(l) \odot \mathbf{a}(z_i) + \mathbf{n}(l), \quad (21.207)$$

where $\mathbf{x}_i \in \mathbb{C}^{M \times 1}$ accounts for both the reflection coefficient of the i th source, s_i , and its potential variations between the interferometric acquisitions or over the L realizations. Depending on the type of scatterer under observation, the source signal \mathbf{x}_i possesses varying statistical properties. The resulting composite signal $\mathbf{y}(l)$ may then follow different scattering behaviors [66].

- *Distributed scatterers*: The media are characterized by a scattering response having a random behavior conferred by the speckle effect. Under some assumptions summarized in [67], the scattered signal may be modeled using a multiplicative random term following a centered complex gaussian distribution with unitary variance

$$\mathbf{x}_i(l) = s_i \mathbf{x}_{u_i}(l) \in \mathbb{C}^{M \times 1} \quad \text{with} \quad \mathbf{x}_{u_i}(l) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_i), \quad (21.208)$$

where $\mathbf{x}_{u_i}(l)$ represents the multiplicative speckle term and s_i is the complex response of the i th scatterer, its reflectivity being given by $\sigma_i = |s_i|^2$. The $(M \times M)$ covariance matrix \mathbf{C}_i describes the interferometric coherence and, for a well calibrated acquisition system, contains unitary diagonal elements, whereas off-diagonal terms depend on the acquisition conditions. This kind of source signal is well adapted to the modeling of scattering over natural environments, like rough surfaces, ground and volumetric environments.

- *Deterministic scatterers:* The backscattered source signal is frozen over all the observations and presents no decorrelation between the different acquisitions. For a well calibrated system, its form is given by

$$\mathbf{x}_i(l) = s_i \mathbf{x}_{c_i}(l) \in \mathbb{C}^{M \times 1} \quad \text{with} \quad \mathbf{x}_{c_i}(l) = \mathbf{1}_{(M \times 1)}. \quad (21.209)$$

This behavior is generally related to specular scattering mechanisms and can be observed over coherent scatterers like calibrators, facets facing the radar, double-bounce reflections over dihedral-like objects having smooth surfaces (like ground-trunk interactions and double bounce reflections between an object and the ground), or may be linked to resonant behaviors over quasi-periodic media [48].

- *Hybrid scatterers:* Consist of a mixture of coherent and distributed scatterers [48]. In such cases, a hybrid MB-InSAR signal model, introduced by Sauer et al. [68], may be used to describe the received signal, as it includes both deterministic and distributed contributions:

$$\begin{aligned} \mathbf{y}(l) &= \mathbf{y}_c(l) + \mathbf{y}_u(l) \\ &= \sum_{i=1}^{n_{s_1}} s_i \mathbf{x}_{c_i}(l) \odot \mathbf{a}(z_i) + \sum_{i=1}^{n_{s_2}} s_i \mathbf{x}_{u_i}(l) \odot \mathbf{a}(z_i) + \mathbf{n}(l). \end{aligned} \quad (21.210)$$

This type of signals can be frequently encountered when dealing with intermediate-resolution SAR images.

2.21.4.3.2.2 PolTomSAR signal model

The polarimetric response of a scatterer is fully described by its (2×2) complex scattering matrix \mathbf{S} . The scattering matrix can be vectorized using, for instance, the Pauli basis matrix set $\{\Psi_p\}_{p=1,\dots,3}$ [12], in order to build a target vector \mathbf{v}

$$\mathbf{v} = \frac{1}{\sqrt{2}} [S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv}] = s \mathbf{k}, \quad (21.211)$$

where $\sigma = |s|^2 = \mathbf{v}^\dagger \mathbf{v}$ represents the polarimetric span [12] of the scatterer response, and $\mathbf{k} = [k_1, k_2, k_3]^T \in \mathbb{C}^{3 \times 1}$ represents a unitary polarimetric target vector, i.e., $\mathbf{k}^\dagger \mathbf{k} = 1$. In an MB-PolInSAR configuration, the array response may be represented by re-arranging the M acquired polarimetric signals, \mathbf{v}_j with $j = 1, \dots, M$ the track index, under the form of a $3M$ element vector, \mathbf{y}_P , composed of 3 MB-inSAR components, each related to a polarization channel:

$$\mathbf{y}_P = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T \in \mathbb{C}^{3M \times 1}, \quad (21.212)$$

where $\mathbf{y}_p \in \mathbb{C}^{3M \times 1}$, with $p = 1, 2$, or, 3 , represents the MB-InSAR response for the p th polarimetric channel, i.e., $[\mathbf{y}_p]_j = [\mathbf{v}_j]_p$. Using this convention of representation, the polarimetric steering vector

and steering matrix are given by:

$$\begin{aligned}\mathbf{a}(z_i, \mathbf{k}_i) &= \mathbf{k}_i \otimes \mathbf{a}(z_i), \\ \mathbf{A}(\mathbf{z}, \mathbf{K}) &= [\mathbf{a}(z_1, \mathbf{k}_1), \dots, \mathbf{a}(z_D, \mathbf{k}_D)],\end{aligned}\quad (21.213)$$

where \mathbf{k}_i is the polarimetric target vector of the i th source and $\mathbf{K} = [\mathbf{k}_1, \dots, \mathbf{k}_{n_s}]$. A polarimetric steering vector can be defined using five real coefficients given by the elevation, z and the real and imaginary parts of two complex numbers defining a unitary 3-element polarimetric complex vector whose absolute phase is arbitrary. Similarly to the Single Polarization (SP) expression given in (21.204), the received MB-PolInSAR signal $\mathbf{y}_P(l)$ may be formulated as

$$\mathbf{y}_P(l) = \mathbf{A}(\mathbf{z}, \mathbf{K})\mathbf{s}(l) + \mathbf{n}(l) \in \mathbb{C}^{3M \times 1}, \quad (21.214)$$

where similarly to the SP case, $\mathbf{s}(l) \in \mathbb{C}^{n_s \times 1}$ represents a realization of the complex amplitude of the i th source. Diverse model assumptions given for SP signals, can be similarly used for PolTomSAR signals.

2.21.4.3.3 Tomographic focusing techniques

The objective of tomographic focusing is to estimate the reflectivity, scattering vector, and height σ , \mathbf{k}_i , z_i of each source using the acquired data sets. The source characteristics can be estimated from the covariance matrix of the received signal $\mathbf{R} = \mathbf{E}(\mathbf{y}\mathbf{y}^\dagger)$. In practice, for a locally Gaussian statistical behavior, a maximum likelihood estimate of \mathbf{R} may be computed from L independent locations surrounding the pixel under analysis, as

$$\widehat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^L \mathbf{y}(l)\mathbf{y}^\dagger(l). \quad (21.215)$$

Under the hypothesis of a uniform true reflectivity and for a sufficient number of looks, L , $\widehat{\mathbf{R}}$ may be used instead of \mathbf{R} to perform tomography. The number of sources, n_s , is in general unknown and needs to be estimated from the measured data. Some commonly used Model Order (MO) selection techniques based on statistical approaches, e.g., ITC, MDL, AIC [69–71], may be used to determine \hat{n}_s . Once n_s is determined, the eigenstructure matrices can be estimated from the sample covariance matrix

$$\widehat{\mathbf{R}} = \widehat{\mathbf{E}}_s \widehat{\Lambda}_s \widehat{\mathbf{E}}_s^\dagger + \widehat{\mathbf{E}}_n \widehat{\Lambda}_n \widehat{\mathbf{E}}_n^\dagger, \quad (21.216)$$

where $\widehat{\mathbf{E}}_s$ and $\widehat{\mathbf{E}}_n$ are respectively estimated signal and noise subspaces.

2.21.4.3.3.1 Single polarization tomography

Using MB-InSAR data sets acquired for a given polarization channel, Single Polarization (SP) tomography can be derived from the data covariance matrix $\widehat{\mathbf{R}}$ using classical mono-dimensional estimators like Beamforming, Capon and MUSIC estimator as well as multi-dimensional methods like maximum likelihood estimators and weighted subspace fitting approaches.

Mono-dimensional estimators. These approaches determine $\hat{\mathbf{z}}$, an estimate of the elevation of the scatterers under observation, as the coordinates of the \hat{n}_s largest local maxima of an continuous objective function $P(z)$:

$$\hat{\mathbf{z}} = \arg \max_{z, \text{loc}} P(z). \quad (21.217)$$

For the classical Beamformer and Capon spectral estimation techniques, the objective function is given by the continuous estimate of the reflectivity, $P_{B,C}(z) = \hat{\sigma}_{B,C}(z)$, defined as

$$\begin{aligned} \hat{\sigma}_B(z) &= \frac{\mathbf{a}^\dagger(z)\hat{\mathbf{R}}\mathbf{a}(z)}{M^2}, \\ \hat{\sigma}_C(z) &= \frac{1}{\mathbf{a}^\dagger(z)\hat{\mathbf{R}}^{-1}\mathbf{a}(z)}. \end{aligned} \quad (21.218)$$

Once the set of scatterer elevation, $\hat{\mathbf{z}}$, is estimated using (21.217), the corresponding reflectivities can be obtained from (21.218). The selection of discrete sources from peaks of the reflectivity spectrum confers to the Beamformer and Capon estimation techniques an important sensitivity to the acquisition configuration, and in particular to the presence of spurious sidelobes related to an irregular baseline sampling. The Beamformer is known to show a low resolution and may then overlook some closely spaced scatterers, whereas Capon's technique possesses an improved resolution but a reduced radiometric accuracy.

MUSIC is a subspace based mono-dimensional technique, whose objective function is a measure of the orthogonality between a steering vector $\mathbf{a}(z)$ and the estimated noise subspace $\hat{\mathbf{E}}_n$ and is given by:

$$P_M(z) = \frac{1}{\mathbf{a}^\dagger(z)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^\dagger\mathbf{a}(z)}. \quad (21.219)$$

Once $\hat{\mathbf{z}}$, is determined by inserting (21.219) in (21.217), a Least-Square (LS) estimate of the complex reflectivity vector $\hat{\mathbf{s}}$ can be obtained [60,61].

Nonparametric approaches like Beamforming and Capon, are generally used to globally appreciate the structure of a volumetric medium and the main trends of the continuous reflectivity distribution in elevation. For the analysis of discrete spectral components, they may fail to discriminate closely spaced scatterers due either to their limited resolution, or to the presence of side lobes that may induce an erroneous estimation of the source location. MUSIC generally presents better performances for the analysis of discrete sources, related to a better resolution. Nevertheless, like all parametric methods, MUSIC is sensitive to data modeling errors, and in particular those related to the estimated number of sources \hat{n}_s . Moreover, MUSIC is known to work well in the case of uncorrelated scatterers, but its performance may degrade significantly in the presence of correlated scatterers since the source signal covariance matrix tends to be singular [72]. One of the main advantages of such techniques resides in the low numerical complexity of the mono-dimensional optimization described in (21.217).

Multi-dimensional estimators. Maximum likelihood (ML) estimators are multi-dimensional techniques, generally given by

$$(\hat{\mathbf{z}}, \hat{\boldsymbol{\sigma}}, \hat{\sigma}_n^2) = \arg \max \mathcal{L}(\mathbf{z}, \boldsymbol{\sigma}, \sigma_n^2) \quad (21.220)$$

with the likelihood function $\mathcal{L}(\mathbf{z}, \sigma, \sigma_n^2)$. In the case of distributed scatterers, the Stochastic ML (SML) can be derived [73]:

$$\begin{aligned}\widehat{\mathbf{z}} &= \arg \min \log |\mathbf{P}_A \widehat{\mathbf{R}} \mathbf{P}_A + \hat{\sigma}_n^2 \mathbf{P}_A^\perp|, \\ \hat{\sigma}_n^2 &= \frac{1}{M - \hat{n}_s} \text{tr}\{\mathbf{P}_A^\perp \widehat{\mathbf{R}}\},\end{aligned}\quad (21.221)$$

where $\mathbf{P}_A^\perp = (\mathbf{I}_{(M \times M)} - \mathbf{A}(\mathbf{z})(\mathbf{A}^\dagger(\mathbf{z})\mathbf{A}(\mathbf{z}))^{-1}\mathbf{A}^\dagger(\mathbf{z}))$ represents the orthogonal projector on the null space of $\mathbf{A}(\mathbf{z})$. Considering deterministic scatterers, the cost function (21.220) results in the Deterministic ML (DML) estimator[74]

$$\widehat{\mathbf{z}} = \arg \min \text{tr}\{\mathbf{P}_A^\perp \widehat{\mathbf{R}}\}. \quad (21.222)$$

Weighted Noise subspace fitting (NSF) estimator is derived by fitting the estimated noise subspace $\widehat{\mathbf{E}}_n$ with $\mathbf{A}(\mathbf{z})$ in the weighted LS sense as follows:

$$Q_{\text{NSF}}(\mathbf{z}) = \|\widehat{\mathbf{E}}_n^\dagger \mathbf{A}(\mathbf{z})\|_{\mathbf{W}}^2 = \text{tr}(\mathbf{A}^\dagger(\mathbf{z}) \widehat{\mathbf{E}}_n \widehat{\mathbf{E}}_n^\dagger \mathbf{A}(\mathbf{z}) \mathbf{W}), \quad (21.223)$$

whereas the weighted Signal Subspace Fitting (SSF) cost function is given by:

$$Q_{\text{SSF}}(\mathbf{z}) = \|\widehat{\mathbf{E}}_s - \mathbf{A}(\mathbf{z}) \mathbf{T}\|_{\mathbf{W}}^2 = \text{tr}(\mathbf{P}_A^\perp(\mathbf{z}) \widehat{\mathbf{E}}_s \mathbf{W} \widehat{\mathbf{E}}_s^\dagger), \quad (21.224)$$

where the fitting matrix \mathbf{T} is replaced by its LS estimate $\widehat{\mathbf{T}} = (\mathbf{A}^\dagger(\mathbf{z})\mathbf{A}(\mathbf{z}))^{-1}\mathbf{A}^\dagger(\mathbf{z})\widehat{\mathbf{E}}_s$. Both cost functions may be used to estimate the elevation of the scatters using a n_s -dimensional minimization:

$$\widehat{\mathbf{z}}_{\text{WSF}} = \arg \max_{\mathbf{z}} Q_{\text{WSF}}(\mathbf{z}), \quad (21.225)$$

where the suffix WSF indicates ones of the methods, NSF or SSF, mentioned above. It has been shown in [75], that any hermitian positive semi definite weighting matrix \mathbf{W} yields consistent parameter estimates. In particular, a consistent estimate of \mathbf{W} permits to obtain minimum variance estimates, which asymptotically reach the Cramr-Rao lower bound [73]. Such a value is given for each case:

$$\begin{aligned}\mathbf{W}_{\text{SSF}_{\text{opt}}} &= (\widehat{\Lambda}_s - \hat{\sigma}_n^2 \mathbf{I}_{(n_s \times n_s)})^2 \widehat{\Lambda}_s^{-1}, \\ \mathbf{W}_{\text{NSF}_{\text{opt}}} &= \mathbf{A}^\dagger(\mathbf{z}) \widehat{\mathbf{E}}_s \mathbf{W}_{\text{SSF}_{\text{opt}}} \widehat{\mathbf{E}}_s^\dagger \mathbf{A}(\mathbf{z}).\end{aligned}\quad (21.226)$$

Compared with the mono-dimensional ones, multi-dimensional methods ML estimators are more robust, leading to global optima, but at an expensive computational cost. Regarding weighted subspace fitting techniques, the appropriate selection of certain weighting matrices can provide these estimators an optimal estimation accuracy as ML techniques at a reduced computational cost.

Numerical examples. In order to demonstrate the performance of the aforementioned tomographic estimators, two scatterers are assumed to locate at $h_1 = 0$ m and $h_2 = 4$ m, acquired respectively by a multibaseline configuration with $M = 5$ acquisitions. The baselines are assumed to be evenly distributed and between each successive acquisitions, the difference of vertical wavenumbers is supposed to be $\Delta k_z = 0.1$. Assuming these two scatterers with equal reflectivity, the covariance matrix of source signals is given by

$$\mathbf{R}_x = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}$$

and $\text{SNR} = 20 \text{ dB}$. Varying the correlation factor ρ , \mathbf{R}_x can respectively represent the source covariance matrix of coherent ($\rho = 1$), distributed ($\rho = 0$), and hybrid scatterers ($0 < \rho < 1$). In order to study the vertical resolution, these estimators are used in the height estimation of uncorrelated signals ($\rho = 0$) with varying the height difference of two scatterers $\Delta h = |h_1 - h_2|$. Figure 21.79a shows when $\Delta h < 2 \text{ m}$, MUSIC estimator degrades significantly due to closely spaced scatterers. SSF and NSF estimators both provide very good resolution with $\Delta h = 0.4 \text{ m}$. Whereas, Capon's method can provide a good height estimate when $\Delta h > 4 \text{ m}$. Now simulating a 256-look sample data covariance matrix with varying ρ from 0 to 0.99, RMSE of the estimated h_2 is plotted with respect to the correlation factor ρ in Figure 21.79b. The NSF estimator provides the most accurate estimate for uncorrelated or partially correlated signals ($\rho < 0.95$), while the SSF estimator copes well with highly correlated signals ($\rho \geq 0.95$). MUSIC cannot deal with highly correlated sources due to the quasi-singularity of \mathbf{R}_x .

2.21.4.3.3.2 Fully polarimetric tomography

Fully Polarimetric (FP) tomography is implemented by jointly using fully polarimetric data sets acquired in a Polarimetric MB-InSAR configuration. It is a useful tool to extract both the scatterer's vertical location and its scattering mechanism. Similar to the SP case, the FP tomography is derived from the data covariance matrix $\widehat{\mathbf{R}}_P = \frac{1}{L} \sum_{l=1}^L (\mathbf{y}_P(l)\mathbf{y}_P^\dagger(l))$ using some mono-dimensional and multi-dimensional FP spectral estimators.

Mono-dimensional estimators. The elevation and scattering mechanisms of the different scatterers are estimated as the coordinates of the \hat{n}_s largest local maxima of a polarimetric objective function $P(z, \mathbf{k})$.

$$\widehat{\mathbf{z}}, \widehat{\mathbf{K}} = \arg \max_{z, \mathbf{k}, \text{loc}} P(z, \mathbf{k}). \quad (21.227)$$

Similarly to the SP case, the Beamformer and Capon objective functions are given by continuous estimates of the reflectivity $P_{B,C}(z, \mathbf{k}) = \hat{\sigma}_{B,C}(z, \mathbf{k})$, with

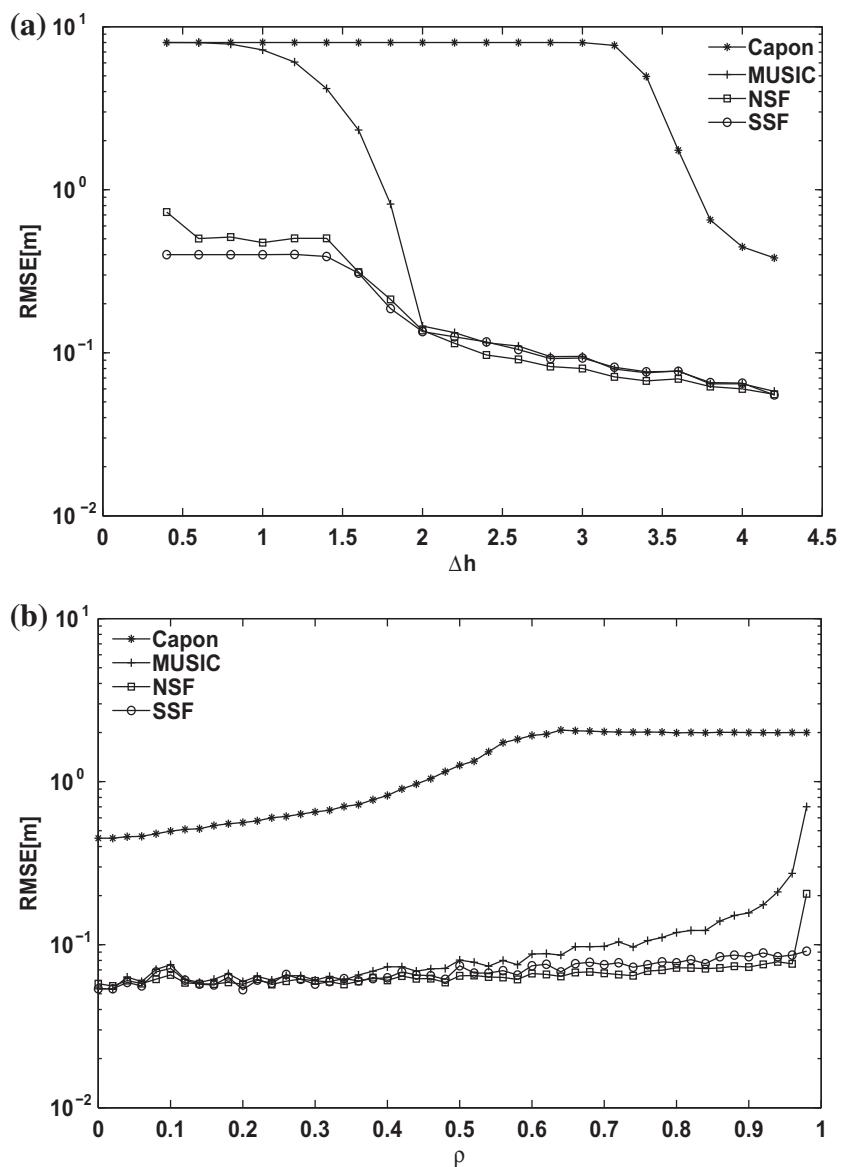
$$\begin{aligned} \hat{\sigma}_B(z, \mathbf{k}) &= \frac{\mathbf{a}^\dagger(z, \mathbf{k})\widehat{\mathbf{R}}_P\mathbf{a}(z, \mathbf{k})}{M^2}, \\ \hat{\sigma}_C(z, \mathbf{k}) &= \frac{1}{\mathbf{a}^\dagger(z, \mathbf{k})\widehat{\mathbf{R}}_P^{-1}\mathbf{a}(z, \mathbf{k})}. \end{aligned} \quad (21.228)$$

The direct local maximization of the FP objective function in (21.227) would require searching solutions in a 5-D argument space. Rewriting the polarimetric steering vector as

$$\mathbf{a}(z_i, \mathbf{k}_i) = (\mathbf{I}_{(3 \times 3)} \otimes \mathbf{a}(z_i))\mathbf{k}_i = \mathbf{B}_a(z_i)\mathbf{k}_i \quad (21.229)$$

permits to formulate reduce the search dimension to one, since

$$\begin{aligned} P_B(z) &= \max_{\mathbf{k}} P_B(z, \mathbf{k}) = \frac{\lambda_{\max}(\mathbf{B}_a^\dagger(z)\widehat{\mathbf{R}}_P\mathbf{B}_a(z))}{M^2}, \\ P_C(z) &= \max_{\mathbf{k}} P_C(z, \mathbf{k}) = \frac{1}{\lambda_{\min}(\mathbf{B}_a^\dagger(z)\widehat{\mathbf{R}}_P^{-1}\mathbf{B}_a(z))}, \end{aligned} \quad (21.230)$$

**FIGURE 21.79**

MB-InSAR: height estimation. RMSE of h estimation VS (a) height difference when $\rho = 0$ (b) source correlation ρ .

where $\lambda_{\min,\max}(\mathbf{M})$ represents the minimum and maximum eigenvalue of the positive semi-definite matrix \mathbf{M} . Once the \hat{n}_s elevations, are determined from the local maxima of (21.230), the corresponding scattering mechanisms can be estimated from:

$$\begin{aligned}\hat{\mathbf{k}}_{B_i} &= \mathbf{e}_{\max}(\mathbf{B}_a^\dagger(\hat{z}_i)\hat{\mathbf{R}}_P\mathbf{B}_a(\hat{z}_i)), \\ \hat{\mathbf{k}}_{C_i} &= \mathbf{e}_{\min}(\mathbf{B}_a^\dagger(\hat{z}_i)\hat{\mathbf{R}}_P^{-1}\mathbf{B}_a(\hat{z}_i)),\end{aligned}\quad (21.231)$$

where $\mathbf{e}_{\min,\max}(\mathbf{M})$ indicates the eigenvector of \mathbf{M} associated with $\lambda_{\min,\max}(\mathbf{M})$, respectively. A similar approach may be applied to concentrate the FP MUSIC objective function:

$$\begin{aligned}P_M(z) &= \max_{\mathbf{k}} P_M(z, \mathbf{k}) = \frac{1}{\lambda_{\min}(\mathbf{B}_a^\dagger(z)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^\dagger\mathbf{B}_a(z))}, \\ \hat{\mathbf{k}}_{M_i} &= \mathbf{e}_{\min}(\mathbf{B}_a^\dagger(\hat{z}_i)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^\dagger\mathbf{B}_a(\hat{z}_i)),\end{aligned}\quad (21.232)$$

where $\hat{\mathbf{E}}_n$ represent here the eigenvectors of $\hat{\mathbf{R}}_P$ spanning its noise space. The reflectivity of the sources can be estimated, using a L-look polarimetric LS approach.

These estimators are considered as mono-dimensional, since the estimation of the elevation and the scattering mechanism is jointly realized by eigendecomposition. Such methods are computationally efficient but may reach some of the limitations mentioned in the SP case.

Multi-dimensional estimators. The ML estimators are formulated in the FP case using the polarimetric steering matrix $\mathbf{A}(\mathbf{z}, \mathbf{K})$, represented by:

- *FP-DML*: $\hat{\mathbf{z}}, \hat{\mathbf{K}} = \arg \max \text{tr}\{\mathbf{P}_A(\mathbf{z}, \mathbf{K})\hat{\mathbf{R}}_P\}$.
- *FP-SML*: $\hat{\mathbf{z}}, \hat{\mathbf{K}} = \arg \min |\mathbf{P}_A(\mathbf{z}, \mathbf{K})\hat{\mathbf{R}}_P\mathbf{P}_A(\mathbf{z}, \mathbf{K}) + \hat{\sigma}_n^2\mathbf{P}_A^\perp(\mathbf{z}, \mathbf{K})|$.

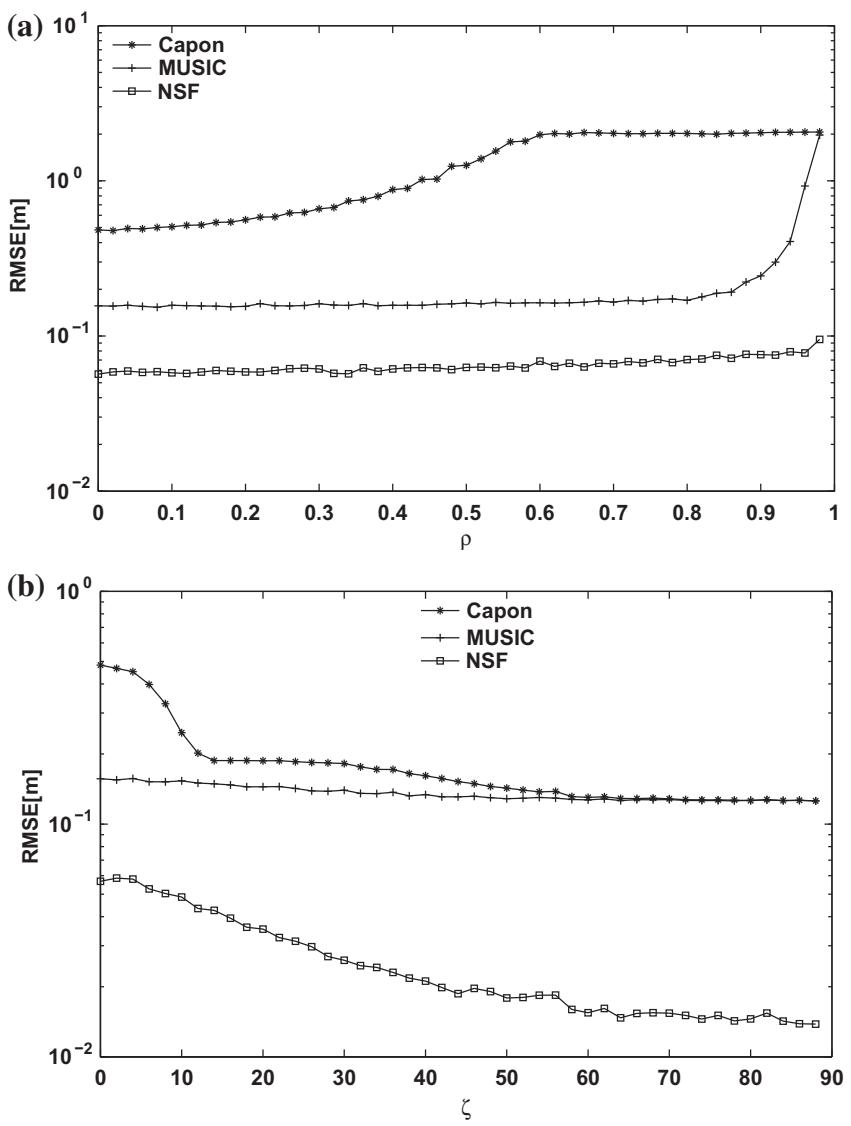
To directly optimize the FP ML estimators, the estimation of \hat{n}_s optimal elevations \hat{z}_i and target vectors \mathbf{k}_i , requires an optimization over a $5\hat{n}_s$ -dimensional space, and implies an excessive computational burden.

Polarimetric WSF estimators are given by:

- *FP-SSF*: $(\hat{\mathbf{z}}, \hat{\mathbf{K}}) = \arg \min \text{tr}\{\mathbf{P}_A^\perp(\mathbf{z}, \mathbf{K})\hat{\mathbf{E}}_s\mathbf{W}_{\text{SSF}_{\text{opt}}}\hat{\mathbf{E}}_s^\dagger\}$.
- *FP-NSF*: $(\hat{\mathbf{z}}, \hat{\mathbf{K}}) = \arg \min \text{tr}\{\mathbf{A}^\dagger(\mathbf{z}, \mathbf{K})\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^\dagger\mathbf{A}(\mathbf{z}, \mathbf{K})\mathbf{W}_{\text{NSF}_{\text{opt}}}\}$.

Directly minimizing the FP-SSF estimator also suffers $5\hat{n}_s$ computational cost, but the FP-NSF estimator can maintain the computational cost as the SP case, using an analytic solution proposed in [66]. Compared with SP tomography, polarimetric tomography can provide more accurate localization of scatterers in the vertical direction due to the polarization diversity. The observed media can be further characterized in terms of both vertical location and scattering mechanism.

Numerical examples. Under the same baseline configuration as the SP case in Section (2.21.4.3.3.1), a 256-look sample covariance matrix is derived from received fully polarimetric data sets in the MB-PolInSAR configuration, which contains both height and polarimetric information. Assuming the above two scatterers with the canonical surface scattering, i.e., $\mathbf{k}_1 = \mathbf{k}_2 = [1, 0, 0]^T$, Figure 21.80a shows RMSE of height estimation w.r.t. ρ : FP-Capon estimator still cannot accurately separate two scatterers due to a limited resolution; FP-NSF estimator provides the most precise height estimate; FP-MUSIC estimator degrades its performance for highly correlated scatterers. To study the role that polarization

**FIGURE 21.80**

MB-PolInSAR height estimation. RMSE of h estimation VS (a) source correlation and (b) polarization difference ζ ($\rho = 0$).

diversity plays in source separation, we denote Polarization Difference Angle ζ with $\cos \zeta = \frac{\mathbf{k}_1^\dagger \mathbf{k}_2}{\sqrt{|\mathbf{k}_1|^2 |\mathbf{k}_2|^2}}$. We keep the same surface scattering for the scatterer at 0 m and vary the scattering mechanism of the scatterer at 4 m to increase the polarization difference angle. As shown in Figure 21.80b, with ζ increasing, polarization diversity between two scatterers improves the height resolution, especially for FP-Capon estimator which reaches the same resolution as MUSIC when $\zeta > 60^\circ$. FP-NSF estimator performs best over all the polarization difference angle.

2.21.4.3.4 Applications

Using MB-PolInSAR data sets, TomSAR and PolTomSAR techniques are applied to retrieve the distribution of the backscattered power in the vertical direction, leading to 3-D images of observed scenes. These techniques permit to analyze volumic structures like forests, sand, snow, ice, etc. Some examples are given to illustrate the potentials of tomographic techniques in the applications of urban, forestry remote sensing and underfoliage imaging.

2.21.4.3.4.1 Urban remote sensing

High-resolution (HR) tomographic techniques are applied to analyze dense urban areas in terms of scatterers' heights and their scattering mechanisms. Diverse scattering patterns can be encountered over dense urban environments, such as double bounce reflection at the wall-ground interaction, surface scattering from the roof and the ground, volumic scattering or a mixture of several scattering patterns. Conventional tomographic estimators may reach some limitations. The proposed model-adaptive FP-NSF estimator in [66,76] illustrates an undeniable performance using L-band intermediate-resolution dualbaseline FP data sets acquired over Dresden, Germany, by the DLR ESAR sensor. Over a test line, the α tomograms obtained by FP-MUSIC and FP-NSF estimators with model order equal to 2 are shown in Figure 21.81. The heights of buildings are validated against LIDAR data as shown in [77]. FP-MUSIC

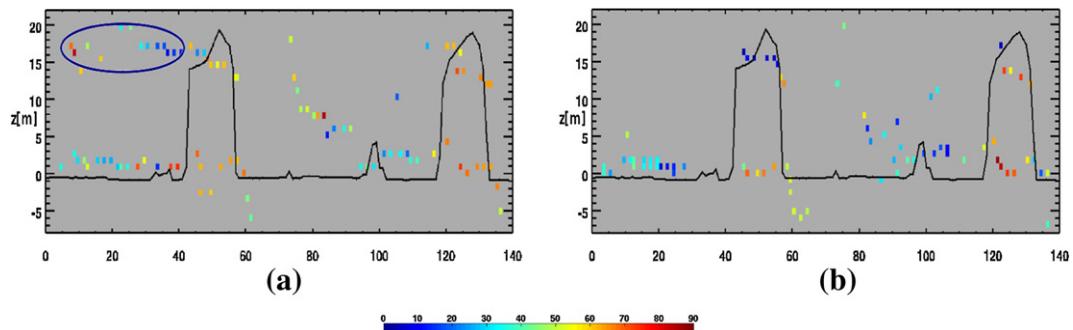
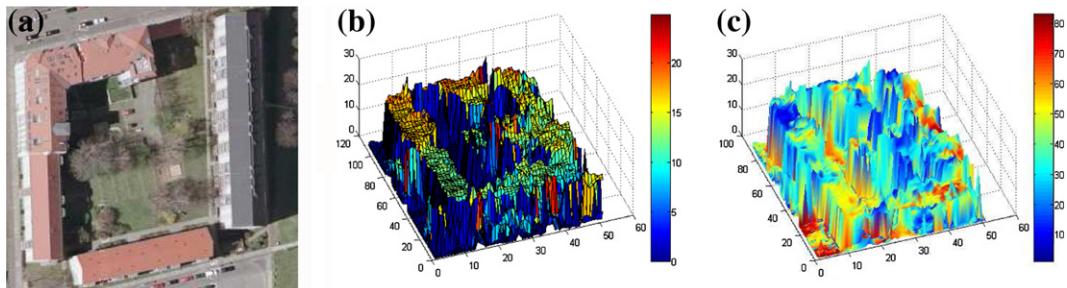
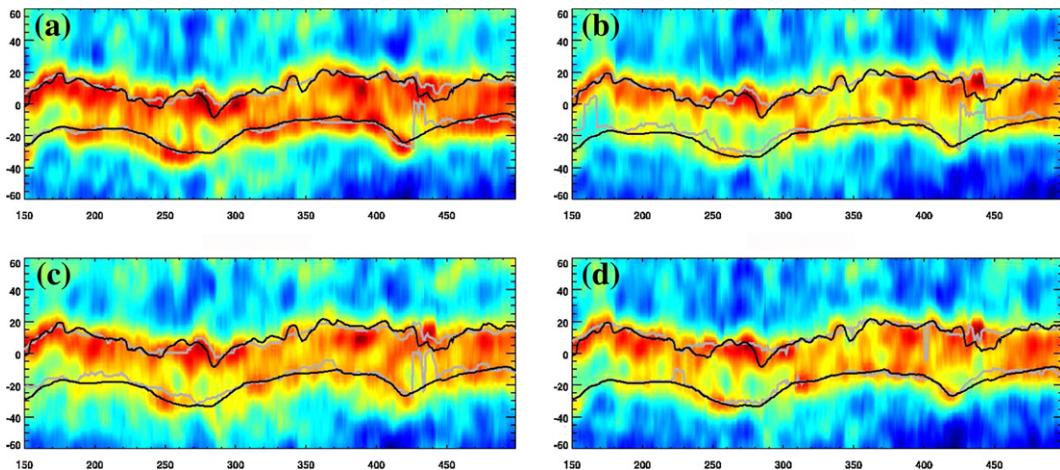


FIGURE 21.81

α tomograms ($n_s = 2$). Black line: LIDAR data. (a) FP-MUSIC and (b) FP-NSF (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

**FIGURE 21.82**

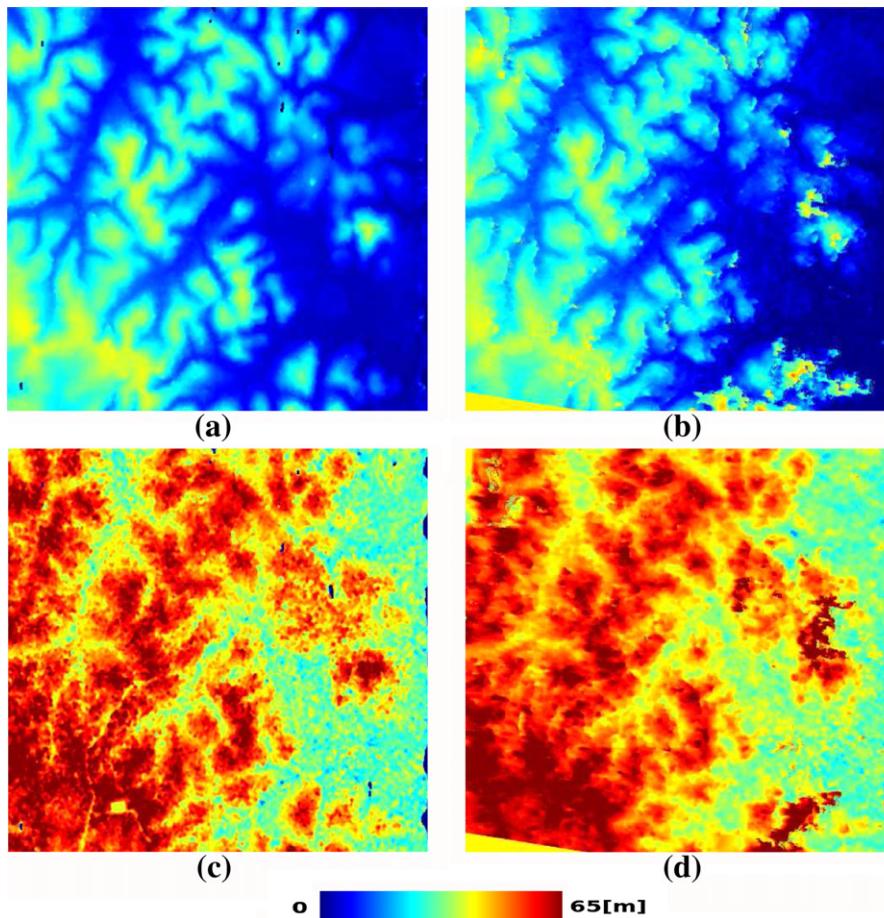
3-D reconstruction over an urban zone of Dresden using FP-NSF estimator ($n_s = 2$). (a) Google map, (b) height map, and (c) α map.

**FIGURE 21.83**

Tree top heights and underlying ground topography by TomSAR approaches. Black line: Lidar profiles. Gray line: estimated profiles by TomSAR. Vertical: height (m). Horizontal: azimuth (bin) (a) HH, (b) VV, (c) HV, and (d) FP (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.).

estimator suffers spurious sidelobes (blue¹ circle), whereas FP-NSF estimator illustrates precisely the vertical profile of buildings without any sidelobe. The 3-D reconstructions of an urban zone are depicted in Figure 21.82.

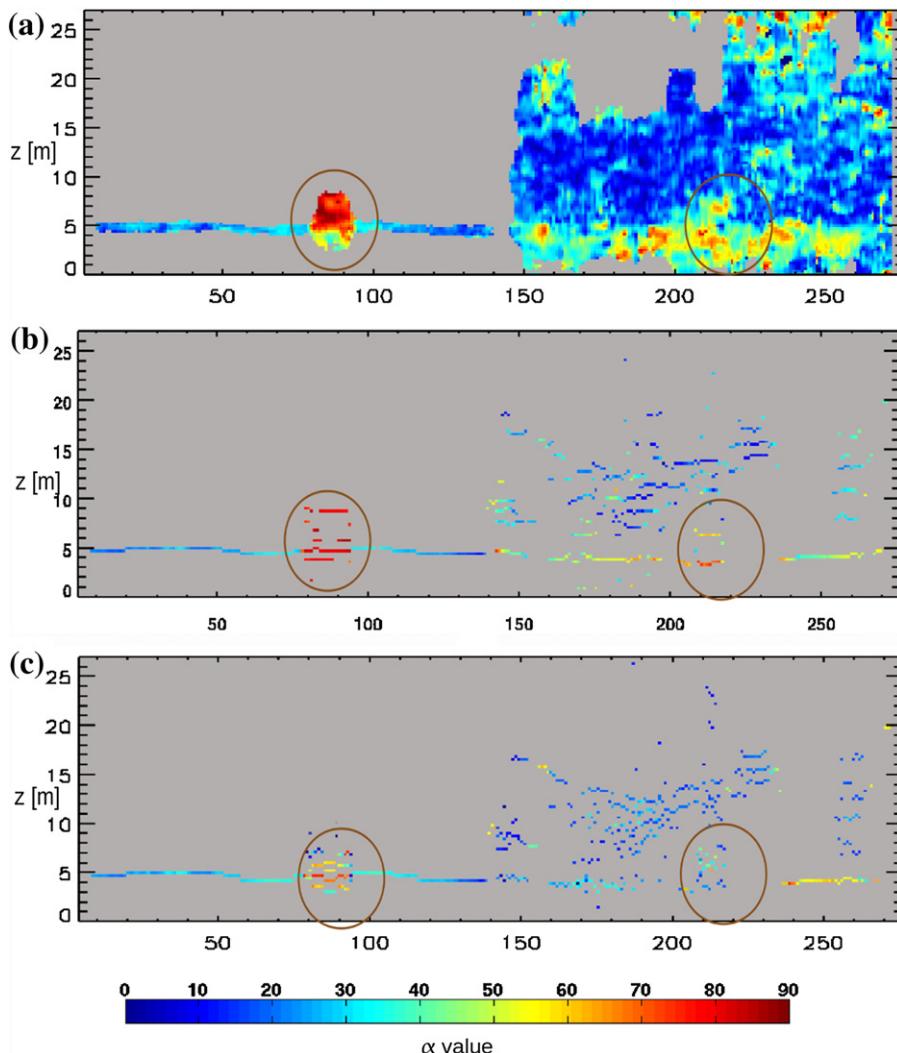
¹For interpretation of color in Figures 21.66, 21.82, 21.83 the reader is referred to the web version of this book.

**FIGURE 21.84**

Estimated ground topography and tree top heights (HH) in ground range (a) lidar z_g , (b) Estimated z_g , (c) lidar z_{top} , and (d) Estimated z_{top} .

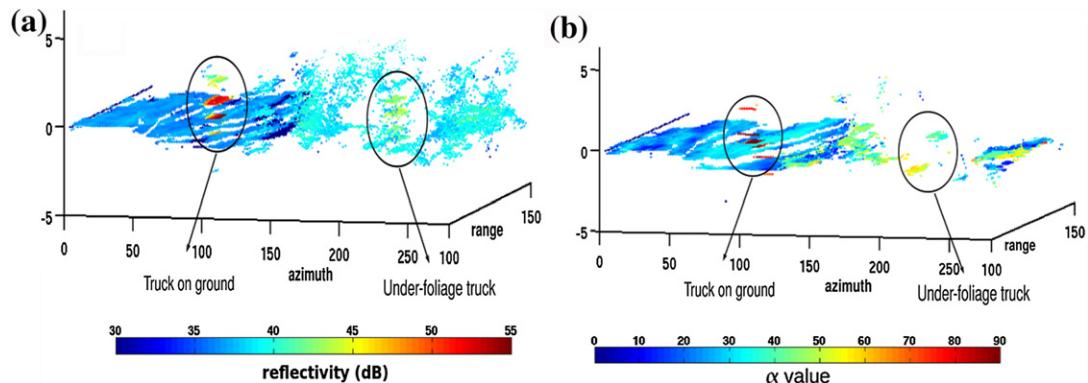
2.21.4.3.4.2 Forestry remote sensing

In the forest scenario, the canopy consists of a large number of elementary scatterers distributed in the vertical direction and is characterized by a continuous spectrum, whereas the ground is an impenetrable medium that possesses an isolated localized phase center and is characterized by a discrete spectrum. So the vertical distribution of backscattered power from forested areas leads to a mixed spectrum. The performance of conventional tomographic estimators may degrade due to their lack of adaptation to the type of spectrum. A novel hybrid tomographic approach is proposed to deal with the mixed spectrum and provides robust estimates for both tree top heights z_{top} and the underlying ground topography z_g [77].

**FIGURE 21.85**

α Tomograms. Circle: truck. Vertical axis: height (m). Horizontal axis: azimuth (bin). (a) FP-Capon, (b) FP-MUSIC, and (c) FP-NSF.

Using P-band data sets (including 6 tracks) acquired over a test area of Paracou, French Guiana, by the ONERA SETHI sensor, the tree top and the ground elevation over an azimuth cut are estimated by the hybrid spectral approach (in gray) and the corresponding Lidar profiles (in black) are drawn over the normalized Capon tomograms in Figure 21.83. It can be observed that the estimated profiles coincide well with the Lidar ones. The ground elevation is precisely estimated using *HH* data sets, but still

**FIGURE 21.86**

3-D visualization of under-foliage objects. (a) SSF and (b) FP-NSF.

overestimated in *VV* and *HV* polarization. The forest top heights are well-estimated over all polarization channels. Using FP data sets, the results are similar to those obtained by *HH* data. Using *HH* data sets, Figure 21.84 depicts the estimated z_{top} and z_g over a test zone as well as the corresponding Lidar data in ground range. Globally, the estimated z_{top} and z_g match well with the Lidar ones in terms of texture and heights, except for some overestimated areas due to topographic effects that may influence scattering responses.

2.21.4.3.4.3 Underfoliage imaging

For the study of under-foliage imaging, Capon's method has been applied to extract the shape of objects and canopy profile [78], but considering the complex structure of such objects, HR approaches are required to discriminate their closely spaced scattering features in the vertical direction. The performance of these spectral analysis techniques is conditioned by the statistical nature of the scattering response of the observed objects, as it has been shown in [79,80]. Under-foliage objects may be described by a deterministic response embedded within a speckle affected environment and hence associated to a series of complex scattering centers. To distinguish under-foliage objects, Figure 21.85 illustrates the α tomograms over a test line with a truck beneath the canopy [81], using L-band PolInSAR data sets acquired over Dornstetten, Germany. The 3-D tomogram of the area, performed over azimuth bins and for varying range positions using the SP-SSF approach. Limiting the reconstruction height to 4 m above the terrain permits to isolate the under-foliage truck response (Figure 21.86a). Its reconstructed shape is similar to the uncovered one, with a reflectivity slightly higher (about 5 dB more) than the surrounding environment one. Complementary to distinct the objects in reflectivity, a 3-D reconstruction using FP-NSF estimator, permits to visualize the corresponding scattering pattern by α value. To better visualize the underlying truck response, the height is limited to 4 m above the ground and the reflectivity is limited to values inferior to 43 dB (Figure 21.86b). It shows that the truck outside the forest has a strong double bounce reflection at the ground-truck interaction, and the shape of the truck beneath the

canopy is a precisely reconstructed with an α value about 45° due to the sidelobe effect from the canopy. Using PolTomSAR techniques, the under-foliage truck can be well described both in terms of shape and scattering pattern.

2.21.5 Conclusion and trends

As it has been mentioned in the introduction of this chapter, SAR polarimetry is a multi-disciplinary subject, with well-established theoretical foundations and tight connections with applied electromagnetism, multivariate signal processing, . . .

Polarization diversity is particularly well adapted to SAR remote sensing since it provides both, a dimension over which one may filter out undesirable responses (or select responses of interest), and a very good way to characterize physical parameters of objects or media. Specific estimation techniques, adapted to the random nature of radar signals, have been developed to provide in an unsupervised way features that correspond to canonical scattering behaviors which can be efficiently used to determine a wide range of physical properties.

A comparison of a polarimetric color-coded SAR image with a single polarization one permits to realize that fully polarimetric measurements bring a lot more information and hence shall reveal much more useful. The most famous and widely used applications of SAR polarimetry mainly concern qualitative approaches, i.e., whose output parameter is not the estimate of the value of a physical parameter. Indeed, numerous polarization-based methods have been developed for object, or natural matter (like snow, water, . . .), or structure detection, environment classification and mapping in general, hazard monitoring, vegetation observation, . . . For such applications, radar polarimetry provides very efficient tools that require among the smallest levels of supervision, i.e., that can be implemented in a quasi automatic way within badly known environments.

Concerning quantitative approaches, the situation is slightly different. For some applications related to parameters that can be well separated by polarimetry, e.g., azimuthal orientation, or whose level of robustness may be estimated from the signal itself, radar polarimetry is a much valuable measuring approach. For some others, a reliable estimation may reveal difficult depending on the environment of the measure. For known conditions, a solution among the wide range offered by polarimetry can generally be found. In the opposite case, radar polarimetry may become a victim of the qualities of polarization. Polarization is widely used as diversity mode due to its good sensitivity to many of the electromagnetic properties of observed media. Estimating one particular parameter among the large number of other factors that modify the polarization of a wave and for a wide range of configurations can be problematic. Polarimetric measurements present a low dimensionality, 9 at most, while they do not possess any intrinsic structure, i.e., a target vector may be represented by an amplitude and a unitary vector. For applications, like those based on coherent spatial diversity, clutter echoes have low rank, whereas in the polarimetric case the clutter may occupy an important part of the eigen-space if the observed echos do not have sufficiently similar characteristics.

Current trends in SAR polarimetry correspond to different possible solutions considered to maintain polarimetry robustness to a high level even in particularly unfavorable configurations. Some solutions use external information provided by other types of sensors, like hyperspectral or other types of measurements.

Other particularly efficient solutions rely on incoherent diversity modes, based on the use of several spectral bands (carriers frequencies) or incidence angles or acquisition dates . . . This solution is strongly encouraged by the development of space-borne sensors delivering time series of data sets.

Finally coherent diversity modes, mainly based on spatial diversity using two (interferometry) or more (tomography) images have shown spectacular results. In particular, forest structure estimation using single or multi-baseline POL-inSAR acquisitions can give very good results when polarimetric only measurements are quasi-useless.

Anyway, one should avoid doing predictions about the future of radar polarimetry, since the historical review given in the introduction clearly mentions that radar polarimetry is regularly deeply modified by a surprising discovery!

This is why we really enjoy this research field.

References

- [1] W.M. Boerner, Inverse methods in electromagnetic imaging, in: Wolfgang M. Boerner et al. (Ed.), Nato Advanced Research Workshop on Inverse Methods in Electromagnetic Imaging, vol. 1–2, Reidel, Bad windsheim, Germany, Holland, Dordrecht, 1983.
- [2] G. Sinclair, Modification of the radar target equation for arbitrary targets and arbitrary polarization, Technical Report Report 302–19, The Ohio State University Research Foundation, Antenna Laboratory, 1948.
- [3] S. Sinclair, The transmission and reception of elliptically polarized waves, Proc. IRE 38 (2) (1950) 148–151.
- [4] E.M. Kennaugh, Effects of the type of polarization on echo characteristics, Technical Report 381–1 394–24, Antenna Laboratory, The Ohio State University Research Foundation, 1949.
- [5] G.A. Deschamps, Geometrical representation of the polarization of a plane electromagnetic wave, Proc. IRE 39 (5) (1951) 540–544. *****
- [6] J. Huynen, Phenomenological theory of radar targets, PhD Thesis, Drukkerij Brondor-Offset NV, 1970.
- [7] S.R. Cloude, Group theory and polarisation algebra, OPTIK 75 (1986) 26–36.
- [8] A. Freeman, Sar calibration: an overview, IEEE Trans. Geosci. Remote Sens. 30 (6) (1992) 1107–1121.
- [9] A. Freeman, S. Durden, A three-component scattering model for polarimetric sar data, IEEE Trans. Geosci. Remote Sens. 36 (3) (1998) 963–973.
- [10] J. van Zyl, Unsupervised classification of scattering behavior using radar polarimetry data, IEEE Trans. Geosci. Remote Sens. 27 (1) (1989) 36–45.
- [11] J.S. Lee, L. Jurkevich, P. Dewaele, P. Wambacq, A. Oosterlinck, Speckle filtering of sar images: a review, Remote Sens. Rev. 8 (1994) 313–340.
- [12] S. Cloude, E. Pottier, A review of target decomposition theorems in radar polarimetry, IEEE Trans. Geosci. Remote Sens. 34 (2) (1996) 498–518.
- [13] S. Cloude, E. Pottier, An entropy based classification scheme for land applications of polarimetric sar, IEEE Trans. Geosci. Remote Sens. 35 (1) (1997) 68–78.
- [14] Y. Yamaguchi, W.-M. Boerner, H. Eom, M. Sengoku, S. Motooka, T. Abe, On characteristic polarization states in the cross-polarized radar channel, IEEE Trans. Geosci. Remote Sens. 30 (5) (1992) 1078–1080.
- [15] S. Cloude, K. Papathanassiou, Polarimetric sar interferometry, IEEE Trans. Geosci. Remote Sens. 36 (5) (1998) 1551–1565.
- [16] E. Lueneburg, Direct and Inverse Electromagnetic Scattering, Pittman Research Notes in Mathematic Series 361, Addison Wesley Longman, Harlow, UK, 1996 (Chapter Radar polarimetry: A revision of basic concepts).
- [17] A. Kostinski, W. Boerner, On foundations of radar polarimetry, IEEE Trans. Antennas Propag 34 (12) (1986) 1395–1404.

- [18] J.S. Lee, E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*, CRC Press, 2008.
- [19] E. Lueneburg, Principles of radar polarimetry, *Proc. IEICE Trans. Electron. Theory* 10 (1995) 1339–1345.
- [20] S. Cloude, Lie groups in electromagnetic wave propagation and scattering, *J. Electromagnet. Waves Appl.* 6 (8) (1992) 947–974.
- [21] R. Bamler, P. Hartl, Synthetic aperture radar interferometry, *Inverse probl.* 14 (4) (1998).
- [22] J.-S. Lee, K. Hoppel, S. Mango, A. Miller, Intensity and phase statistics of multilook polarimetric and interferometric sar imagery, *IEEE Trans. Geosci. Remote Sens.* 32 (5) (1994) 1017–1028.
- [23] R. Touzi, A review of speckle filtering in the context of estimation theory, *IEEE Trans. Geosci. Remote Sens.* 40 (11) (2002) 2392–2404.
- [24] S.R. Cloude, *Polarimetry: the characterization of polarization effects in EM scattering*, PhD Thesis, University of Birmingham, England, 1986.
- [25] Y. Yamaguchi, T. Moriyama, M. Ishido, H. Yamada, Four-component scattering model for polarimetric sar image decomposition, *IEEE Trans. Geosci. Remote Sens.* 43 (8) (2005) 1699–1706.
- [26] A. Freeman, Fitting a two-component scattering model to polarimetric sar data from forests, *IEEE Trans. Geosci. Remote Sens.* 45 (8) (2007) 2583–2592.
- [27] M. Neumann, L. Ferro-Famil, A. Reigber, Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric sar data, *IEEE Trans. Geosci. Remote Sens.* 48 (3) (2009) 1086–1104.
- [28] J. van Zyl, M. Arii, Y. Kim, Model-based decomposition of polarimetric sar covariance matrices constrained for nonnegative eigenvalues, *IEEE Trans. Geosci. Remote Sens.* 49 (9) (2011) 3452–3459.
- [29] M. Arii, J. van Zyl, Y. Kim, Adaptive model-based decomposition of polarimetric sar covariance matrices, *IEEE Trans. Geosci. Remote Sens.* 49 (3) (2011) 1104–1113.
- [30] M. Arii, J. van Zyl, Y. Kim, A general characterization for polarimetric scattering from vegetation canopies, *IEEE Trans. Geosci. Remote Sens.* 48 (9) (2010) 3349–3357.
- [31] R. Paladini, L. Ferro-Famil, E. Pottier, M. Martorella, F. Berizzi, E. Dalle Mese, Lossless and sufficient ψ -invariant decomposition of random reciprocal target, *IEEE Trans. Geosci. Remote Sens.* 99 (2012) 1–15.
- [32] L. Ferro-Famil, E. Pottier, J.S. Lee, *Unsupervised Classification of Natural Scenes form Polarimetric Interferometric SAR data*, World Scientific Publishing, Frontiers of Remote Sensing Information Processing, 2003, pp. 105–137.
- [33] J.-S. Lee, R. Jansen, D. Schuler, T. Ainsworth, G. Marmorino, S. Chubb, Polarimetric analysis and modeling of multifrequency sar signatures from gulf stream fronts, *IEEE J. Oceanic Eng.* 23 (4) (1998) 322–333.
- [34] J.-S. Lee, D. Schuler, T. Ainsworth, E. Krogager, D. Kasilingam, W.-M. Boerner, On the estimation of radar polarization orientation shifts induced by terrain slopes, *IEEE Trans. Geosci. Remote Sens.* 40 (1) (2002) 30–41.
- [35] D. Schuler, J.-S. Lee, G. De Grandi, Measurement of topography using polarimetric sar images, *IEEE Trans. Geosci. Remote Sens.* 34 (5) (1996) 1266–1277.
- [36] J.S. Lee, M.R. Grunes, E. Pottier, L. Ferro-Famil, Unsupervised terrain classification preserving polarimetric scattering characteristics, *IEEE Trans. Geosci. Remote Sens.* 42 (4) (2004) 722–731.
- [37] I. Hajnsek, E. Pottier, S. Cloude, Inversion of surface parameters from polarimetric sar, *IEEE Trans. Geosci. Remote Sens.* 41 (4) (2003) 727–744.
- [38] A. Fung, *Microwave Scattering and Emission Models and Their Applications*, Artech House Remote Sensing Library, Artech House, 1994. <<http://books.google.fr/book?id=MHxtQgAACAAJ>>.
- [39] S. Allain, L. Ferro-Famil, E. Pottier, Relevant polarimetric parameters for surface parameter retrieval using multi-frequency sar data, in: *Proceedings of the Fourth International Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, 2004.

- [40] S. Allain, C.L. Martinez, L. Ferro-Famil, E. Pottier, New eigenvalue-based parameters for natural media characterization, in: Proceedings of the IGARSS, 2005.
- [41] K. Chen, T.-D. Wu, L. Tsang, Q. Li, J. Shi, A. Fung, Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations, IEEE Trans. Geosci. Remote Sens. 41 (1) (2003) 90–101.
- [42] P. Leducq, L. Ferro-Famil, E. Pottier, Matching-pursuit-based analysis of moving objects in polarimetric sar images, IEEE Geosci. Remote Sens. Lett. 5 (2) (2008) 123–127.
- [43] F. Gatelli, A. Guamieri, F. Parizzi, P. Pasquali, C. Prati, F. Rocca, The wavenumber shift in sar interferometry, IEEE Trans. Geosci. Remote Sens. 32 (4) (1994) 855–865.
- [44] P. Flandrin, Time-Frequency/Time-Scale Analysis, Academic Press, London, 1999, (Wavelet Anlaysis Appl.).
- [45] L. Ferro-Famil, W.M. Boerner, E. Pottier, A. Reigber, Analysis of sar response anisotropic behavior using sub-aperture polarimetric data, in: Proceedings of the POLINSAR, 2003.
- [46] L. Ferro-Famil, A. Reigber, E. Pottier, W.M. Boerner, Scene characterization using subaperture polarimetric sar data, IEEE Trans. Geosci. Remote Sens. 41 (10) (2003) 2264–2276.
- [47] A. Yueh, R.T. Shin, J.A. Kong, Scattering from randomly perturbed periodic and quasiperiodic surfaces, Prog. Electromagnet. Res. 1 (1988) 297–358.
- [48] L. Ferro-Famil, A. Reigber, E. Pottier, Nonstationary natural media analysis from polarimetric sar data using a two-dimensional time-frequency decomposition approach, Can. J. Remote Sens. 31 (1) (2005) 20–29.
- [49] J.-C. Souyris, C. Henry, F. Adragna, On the use of complex sar image spectral analysis for target detection: assessment of polarimetry, IEEE Trans. Geosci. Remote Sens. 41 (12) (2003) 2725–2734.
- [50] A. Ferretti, C. Prati, F. Rocca, Permanent scatterers in sar interferometry, IEEE Trans. Geosci. Remote Sens. 39 (1) (2001) 8–20.
- [51] R. Schneider, K. Papathanassiou, I. Hajnsek, A. Moreira, Polarimetric and interferometric characterization of coherent scatterers in urban areas, IEEE Trans. Geosci. Remote Sens. 44 (4) (2006) 971–984.
- [52] M. Spigai, C. Tison, J.-C. Souyris, Time-frequency analysis in high-resolution sar imagery, IEEE Trans. Geosci. Remote Sens. 49 (7) (2011) 2699–2711.
- [53] G. Franceschetti, A. Iodice, D. Riccio, G. Ruello, Sar raw signal simulation for urban structures, IEEE Trans. Geosci. Remote Sens. 41 (9) (2003) 1986–1995.
- [54] L. Ferro-Famil, E. Pottier, Urban area remote sensing from l- band polsar data using time-frequency techniques, in: Proceedings Urban Conference 2007.
- [55] S. Cloude, Polarization coherence tomography, Radio Sci. 41 (4) (2006) RS4017.
- [56] S. Cloude, K. Papathanassiou, Three-stage inversion process for polarimetric sar interferometry, IEE Proc. Radar Sonar Navig. 150 (3) (2003) 125–134.
- [57] S.R. Cloude, Polarisation Applications in Remote Sensing, Oxford University Press, 2009.
- [58] K. Papathanassiou, S. Cloude, Single-baseline polarimetric sar interferometry, IEEE Trans. Geosci. Remote Sens. 39 (11) (2001) 2352–2363.
- [59] L. Ferro-Famil, M. Neumann, Y. Huang, Robust estimation of multi-baseline pol-insar parameters for the analysis of natural environments, in: Proceedings of the EUSAR, 2010.
- [60] F. Gini, F. Lombardini, Multibaseline cross-track sar interferometry: a signal processing perspective, IEEE Aerosp. Electron. Syst. Mag. 20 (8) (2005) 71–93.
- [61] P. Stoica, R. Moses, Spectral Analysis of Signals, Prentice Hall, Upper Saddle River, NJ, 2005.
- [62] A. Reigber, A. Moreira, First demonstration of airborne sar tomography using multibaseline l-band data, IEEE Trans. Geosci. Remote Sens. 38 (5) (2000) 2142–2152.
- [63] S. Tebaldini, Algebraic synthesis of forest scenarios from multibaseline polinsar data, IEEE Trans. Geosci. Remote Sens. 47 (12) (2009) 4132–4142.
- [64] S. Tebaldini, Single and multipolarimetric sar tomography of forested areas: a parametric approach, IEEE Trans. Geosci. Remote Sens. 48 (5) (2010) 2375–2387.

- [65] S. Tebaldini, F. Rocca, Multibaseline polarimetric sar tomography of a boreal forest at p- and l-bands, *IEEE Trans. Geosci. Remote Sens.* 50 (1) (2012) 232–246.
- [66] Y. Huang, L. Ferro-Famil, A. Reigber, Under-foliage object imaging using sar tomography and polarimetric spectral estimators, *IEEE Trans. Geosci. Remote Sens.* 99 (2011) 1–13.
- [67] F. Gini, F. Lombardini, M. Montanari, Layover solution in multibaseline sar interferometry, *IEEE Trans. Aerosp. Electron. Syst.* 38 (4) (2002) 1344–1356.
- [68] S. Sauer, L. Ferro-Famil, E. Pottier, A. Reigber, Physical parameter extraction over urban areas using l-band polsar data and interferometric baseline diversity, in: *IEEE International Geoscience and Remote Sensing Symposium IGARSS 2007*, 2007, pp. 1–5.
- [69] M. Wax, T. Kailath, Detection of signals by information theoretic criteria, *IEEE Trans. Acoust. Speech Signal Process.* 33 (2) (1985).
- [70] M. Wax, I. Ziskind, Detection of the number of coherent signals by the mdl principle, *IEEE Trans. Acoust. Speech Signal Process.* 37 (8) (1989) 1190–1196.
- [71] M. Wax, I. Ziskind, Detection and localization of multiple sources via the stochastic signal model, *IEEE Trans. Signal Process.* 39 (11) (1991) 2450–2456.
- [72] P. Stoica, K. Sharman, Maximum likelihood methods for direction-of-arrival estimation, *IEEE Trans. Acoust. Speech Signal Process.* 38 (7) (1990) 1132–1143.
- [73] B. Ottersten, M. Viberg, P. Stoica, A. Nehorai, Radar array processing, in: *Exact and Large Sample Approximations of Maximum Likelihood Techniques for Parameter Estimation and Detection in Array Processing*, Springer-Verlag, 1993, pp. 99–151.
- [74] P. Stoica, B. Ottersten, M. Viberg, R. Moses, Maximum likelihood array processing for stochastic coherent sources, *IEEE Trans. Signal Process.* 44 (1) (1996) 96–105.
- [75] M. Viberg, B. Ottersten, Sensor array processing based on subspace fitting, *IEEE Trans. Signal Process.* 39 (5) (1991) 1110–1121.
- [76] Y. Huang, L. Ferro-Famil, 3d characterization of buildings in a dense urban environment using l-band polinsar data with irregular baselines, in: *Proceedings of the IGARSS*, 2009.
- [77] Y. Huang, L. Ferro-Famil, C. Lardeux, Polarimetric sar tomography of tropical forest at p-band, in: *Proceedings of the IGARSS*, 2011.
- [78] M. Nannini, R. Scheiber, R. Horn, Imaging of targets beneath foliage with sar tomography, in: *Proceedings of the EUSAR*, 2008.
- [79] Y. Huang, L. Ferro-Famil, Building height estimation using multibaseline l-band sar data and polarimetric subspace fitting methods, in: *Proceedings of the POLINSAR*, 2009.
- [80] P. Stoica, A. Nehorai, Performance study of conditional and unconditional direction-of-arrival estimation, *IEEE Trans. Acoust. Speech Signal Process.* 38 (10) (1990) 1783–1795.
- [81] Y. Huang, L. Ferro-Famil, A. Reigber, Under foliage object imaging using sar tomography and polarimetric spectral estimators, in: *Proceedings of the EUSAR*, 2010.

Integrated Sensor Systems and Data Fusion for Homeland Protection

22

Alfonso Farina^{*}, Luciana Ortenzi^{*}, Branko Ristic[†], and Alex Skvortsov[†]

^{*}SELEX Electronic Systems, Rome, Italy

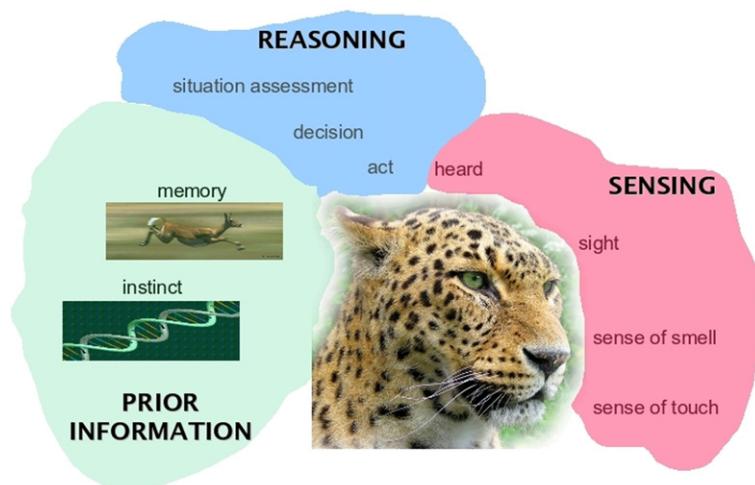
[†]DSTO, Melbourne, Vic. 3207, Australia

2.22.1 Introduction

As stated by John Naisbitt in his bestseller “Megatrends” [1], published in 1982, about the new trends and directions transforming our lives: “We are drowning in information but starved for knowledge. This level of information is clearly impossible to be handled by present means. Uncontrolled and unorganized information is no longer a resource in an information society, instead it becomes the enemy.” This successful sentence can be taken as a statement of the problem of information fusion: how can knowledge, awareness and decision making capability be achieved starting from the available information?

This chapter is intended as an attempt to technical and mathematical answer to the previous question; in particular it addresses the application of data and information fusion to the design of integrated systems in the Homeland Protection (HP) domain. HP refers to the broad civilian and military effort produced by a Country to protect its territory—including citizens, assets and activities which are vital and fundamental for its growth and prosperity—against internal and external hazards and to reduce its vulnerability to attacks, whatever their origin, as well as natural disasters. HP is therefore a wide and complex domain: systems in this domain are large, to mean that size and scope of such systems are conspicuous and that system boundaries may not be easy to identify; systems are integrated, to mean that it is generally not sufficient to study each subsystem in isolation; systems are different in purpose and require a multidisciplinary approach for their design and analysis.

The design and analysis of such systems devoted to operate in such scenarios are necessarily required to provide data and information fusion in the most general sense. Information fusion is about combining, or fusing, information from different sources to provide knowledge that is not evident from individual sources. Numerous real world problems benefit from the combination of heterogeneous information sources, for instance, as depicted in Figure 22.1, multi-sensor data fusion is naturally performed by animals and humans to access more accurately the surrounding environment and to identify threats or food, thereby improving their chances of survival. The field of information fusion is commonly characterized as multidisciplinary research area and includes and/or overlaps with a number of other areas. The information fusion at sensor level includes signal processing; at data level data processing; at meta-data level it overlaps with knowledge representation and, finally, at the decision level it involves the decision making capability. Data fusion has been defined in [2] as “the process of combining evidence to support intelligence generation.” Mainly the methods

**FIGURE 22.1**

Why information fusion. (Kindly provided by Dr. A. Benavoli—IDSIA, Istituto Dalle Molle di Studi sull’Intelligenza Artificiale, Switzerland.)

employed to achieve this scope can be divided into two general classes: quantitative and qualitative. The former are based on numerical techniques, the latter ones are based on symbolic representation of information.

Examples of quantitative methods can be found in stochastic estimation theory, that aim to estimate the state of a system, using all available information, and to characterize the fusion uncertainty in the framework of probability theory. Most of the algorithms developed for quantitative fusion are based on Bayes filter [3], as Kalman filter [4], information filter [5] and neural networks. The application of these algorithms is employed usually to perform multi-target and multi-sensor tracking. The new generation methods, applying a qualitative approach, are based on a symbolic representation of information. They are, of course, based on mathematical models and their output is numeric, however they can be employed to model qualitative information (e.g., fuzzy). They include expert systems, heuristic, behavioral and structural modeling. Qualitative methods are based on artificial intelligence techniques, such fuzzy logic [6], Dempster-Shafer theory [7,8] Dezert-Smarandache theory [9,10] and rules based-methods.

The first data fusion algorithms employed in real systems in the radar field go back to the early seventies, when they had been developed for multi-radar tracking (MRT) for netted sensors. It was late 1970s, beginning 1980s when these new algorithms and the corresponding means to mitigate unavoidable sources of errors due to practical world (e.g., the time synchronization, the radar alignment to the North, the inaccurate knowledge of coordinates of radar sites) were provided. Probably one of the first MRT system for Air Traffic Control (ATC) ever installed was the system operating in the center and South of Italy [11]. When the competence on tracking was so mature, it was collected in a brand new book on radar data processing [12,13], translated also in Russian and Chinese.

Later in 1990s further advancements were done in the field of multi-sensor fusion for Airborne Early Warning (AEW) systems, setting up an algorithm suite to track targets on the basis of the data provided by surveillance radar, an Identification Friend or Foe (IFF), an Electronic Support Measurement (ESM)

and data links. After conceiving also algorithms to track targets on the basis of the angle and identification measurements provided by ESM on a moving platform, data fusion of active and passive tracks was provided [14–17].

Nowadays concepts have been developed and spread to be applied to very complex systems with the aim to achieve the highest level of intelligence as possible and hopefully to support decision. Data fusion is aimed to enhance situation awareness and decision making through the combination of information/data obtained by networks of homogeneous and/or heterogeneous sensors. A sensor network presents advantages over a single sensor under different points of views, as it supplies both redundant and complementary information. Redundant information is exploited to make the system robust to the failure in order that a malfunction of an entity of the system means only a degradation of the performances, rather than the complete failure of the system, since information about the same environment can be obtained from different sources. More robustness can be achieved also with respect to interferences, both intentional and unintentional, due to frequency and spatial diversity of the sensors. Complementary information build up a more complete picture of the observed system; for example sensors are dislocated over large regions providing diverse viewing angles of observed phenomenon and different technologies can be employed in the same application to provide improved system performance.

A large number of different applications, algorithms and architectures have been developed exploiting these advantages. Several examples can be found in robotics, military applications, Homeland Protection and management of large and complex critical infrastructures. Although the specific nature of each problem is different, the final goal, from the point of view of the sensed information, is always the same: using all the available data to better understand the investigated phenomena. The aim of this chapter is to give an overview of the several approaches that can be followed to design and analyze systems for Homeland Protection. Different fusion architectures can be drawn on the basis of the employed algorithms; according to this approach, three general categories can be identified in the literature [18, 19]: centralized, hierarchical, and decentralized/netcentric.

The traditional architecture is centralized: in this framework several sensing devices are connected to a central component, the fusion node. For example, in the case of a sensor network employed for the surveillance of an area, usually the information traffic goes from the sensor nodes to a single sink node called information fusion center. According to the information received from the sensors, the fusion center monitors the area where the sensors are deployed and decides the actions to take. Conceptually, the algorithms employed in this case are relatively simple and the resource allocation is straightforward because the central component has an overall view of the whole system. This kind of architecture presents several drawbacks: high computational load, the possibility of catastrophic failure when the fusion node goes down and the lack of flexibility to changes of the system and sensor entities. Therefore this approach is still valid if the number of sensors, whose information is fused, independently of the width of the area to be monitored, is limited and also the relationship and interconnections among sensors are limited too.

In hierarchical architectures, there are several fusion nodes, where intermediate fusion processes are performed, and an ending central fusion node. The principle of a hierarchy is to reduce the communications and computational loads of centralized systems by distributing data fusion tasks among a hierarchy of sensor entities. However in a hierarchy there is still a central component acting as a fusion center. Entities constituting local fusion center, locally process information and send it to the central fusion node. This approach is commonly used in robotics and surveillance applications. Although this architecture reduces the computational and communication loads, there are still some drawbacks

connected to the centralized model. In addition to these problems, there are some disadvantages related to the resource allocation balancing and the vulnerability to communication bottlenecks.

In certain cases the traditional data fusion algorithms may still be valid; however, in some cases the great variety of sub-systems and the complexity of interconnections may require new approaches. Most of the drawbacks of centralized and hierarchical architectures can be overcome by decentralized architectures. The trend in surveillance today is towards Network Centric Operation (NCO) [20]. The vision for NCO is to provide seamless access to timely information to all operators (e.g., soldier, officer) and decision-makers at every echelon in the military hierarchy. The goal is to enable all elements, including individual infantry soldiers, ground vehicles, command centers, aircraft and naval vessels, to share the collected information and to combine it into a coherent, accurate picture of the battlefield.

The same approach can be followed in the organization of a sensor network. In recent years the decreasing sensor cost and the development of telecommunication technology have made possible the deployment of networks with a huge number of sensors; in this case the use of information fusion centers is unpractical. Consequently a new class of sensor networks, whose way of functioning is called network centric, has emerged. These networks do not have a fusion center and their functioning is based on the information exchange between near-by sensors. Under this approach the information can be considered as a property of the network rather than of the own sensor. This solution is strongly advocated for its robustness and ease of implementation, but it might suffer when the number of sensors grows very much. It has a broad range of potential applications in the field of Homeland Protection: surveillance of habitat and environmental monitoring, structural monitoring (e.g., bridges), contaminants, smart roads, intruder detection, battlefield. It is complementary to the classical surveillance with few large-costly sensors hierarchically organized.

The network of numerous sensors and communication nodes (for instance: peer to peer networks) may have link topology varying with time due to natural interferences, electromagnetic propagation masked by the terrain surface, meteorological conditions, dust and smoke which might be present in the environment, allowing therefore modularity, robustness and flexibility. These networks should be designed to be resilient to Electronic Counter Measures (ECM), cyber attacks and should be able to manage increasing and highly variable flow of data. The satisfaction of such demanding requirements, maintaining however the limitation of resources such as energy, bandwidth and node complexity, can be achieved borrowing from biological systems several mechanism. For example bio-inspired sensor networks employ decentralized decisions through the self-synchronization mechanism observed in nature that allows forcing every single node of the network to reach the globally optimal decision, without the need of any fusion center. However there are also drawbacks associated to these architectures: in fully decentralized systems, communication issues are more complex and depend on the topology of the network; generally, communication overheads are higher than in centralized systems.

In this chapter these aspects will be investigated in depth for networks respectively of homogeneous and heterogeneous sensors with the description of real study cases applied to real world problems of Homeland Protection. In particular the possibility of netting different sensors operating with different characteristics of domain, coverage, frequency and resolution allows a multi-scale¹ approach.

¹In engineering, mathematics, physics, meteorology and computer science, multi scale modeling is the field of solving physical problems which have important features at multiple scales, particularly multiple spatial and (or) temporal scales. (http://en.wikipedia.org/wiki/Multiscale_modeling).

This approach is particularly suitable for the surveillance of wide areas such as national borders or critical strategic regions.

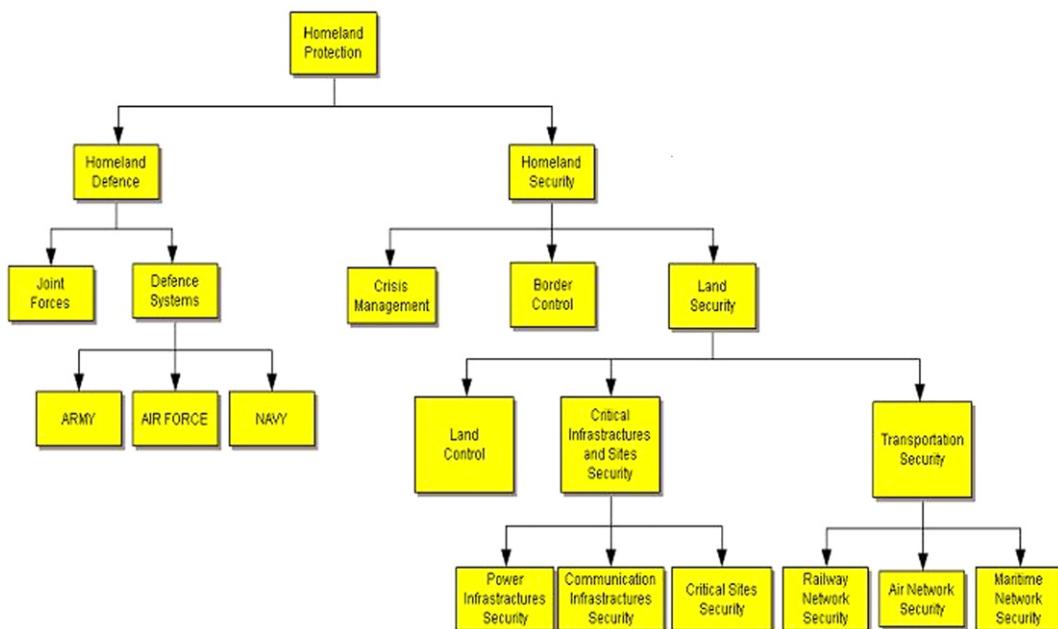
The chapter is organized as follows: Section 2.22.2 illustrates the Homeland Protection domain and highlights some of the characteristics of systems in this specific domain; Section 2.22.3 briefly reviews the development of data fusion and gives references to new emerging trends in the domain of high level data fusion. Section 2.22.4 gives a broad and very general description of the basic categories of intelligence that are the source of data and information employed to perform the fusion process. The Sections 2.22.5 and 2.22.6 tackle different aspects related to homogeneous sensor networks. The former proposes several issues from a theoretical point of view, illustrating, next to traditional approaches, the new trends of Collaborative Signal and Information Processing (CSIP), self-organizing and self-synchronizing sensor network; Section 2.22.5 proposes also some remarks about real applications and the need to rethink some mathematical algorithms to overcome the network centric approach. The latter, Section 2.22.6, proposes three real study cases where the novel approaches give significant results. Likewise Section 2.22.7 tackles the aspects related to heterogeneous sensor networks, dealing with the problems of deployment, behavior assignment and coordination of the different sensors. Also in this case a special attention is focused on the mathematical issues related to these new approaches. Real applications of this kind of sensor networks are described in Sections 2.22.8 and 2.22.9, respectively for the border control problem and the forecasting and estimation of an epidemics. Finally Section 2.22.9, with the concluding remarks, follows.

2.22.2 The Problem of homeland protection

The diagram of Figure 22.2 provides a decomposition of the Homeland Protection domain: the two main sub-domains are Homeland Defense (HD) and Homeland Security (HS) [21].

HD includes the typical duties and support systems of military joint forces and single armed forces. Usually HD systems are strictly military, are employed by military personnel only, satisfy specific technical requirements, operational needs and environmental scenarios, and in most cases are designed to face only military threats. The new trend aims to employ military surveillance systems in combined military and civil operations, especially to face terrorism [22]. The military domain has also been swept in recent years by the NCO paradigm; NCO predicates a tighter coupling among forces, especially in the cognitive domain, to achieve synchronization, agility and decision superiority and it is a strong driver in the transformation from a platform-centric force to a network-centric force [20].

HS is a very broad and complex domain that requires coordinated action among national and local governments, private sector and concerned citizens across a country; it covers issues such as crisis management, border control, critical infrastructure protection and transportation security [23,24]. Crisis management is the ability of identifying and assessing a crisis, planning a response, and acting to resolve the crisis situation. Border control aims to build a smart protection belt all around a country to counter terrorism and illegal activities; yet it is not resolute due to the difficulty of controlling the country boundaries along their full and variegated extension, the non necessarily physical nature of attacks in the current information age, and the threats which often arise internally to the country itself. HS includes also land security that is particularly critical because of its complexity and strategic importance; the security of critical assets, such as electric power plants, communication infrastructures, strategic areas and railway networks, must be ensured continuously in space and time [25–27]. The most recent terrorist

**FIGURE 22.2**

Homeland Protection domain. (From [21], reprinted with permission.)

attacks have shown the vulnerability of national critical infrastructures [28] and have made the world aware of the possibility of large-scale terrorist offensive actions against civil society: the September 11th, 2001 attack on the World Trade center in New York City is the most dramatic example of this new terrorism. The main emphasis has been put on the terrorist threat, but what emerges is the fragility and vulnerability of modern society to both deliberate threats and natural disasters. Figure 22.3 shows a Synthetic Aperture Radar (SAR) image collected by a satellite of the Italian CosmoSkyMed constellation of the area of the Fukushima nuclear plant hit in 2011 by the tsunami.

The HP domain includes also the protection from deliberate attacks against the commercial activities of a Country led also out of the national territory, comprehensive also of the territorial waters and Exclusive Economic Zone (EEZ). Seaborne piracy against transport vessels remains a significant issue (with estimated worldwide losses of US\$13–16 billion per year), particularly in the waters between the Red Sea and Indian Ocean, off the Somali coast, and also in the Strait of Malacca and Singapore, which are navigated by over 50,000 commercial ships a year [29,30].

The globalization, the pervasiveness of information technologies and the transformation of the industrial sector and civil society have created new vulnerabilities in the system as a whole, but all this has happened without a corresponding effort to increase its robustness and security. As an example, single infrastructure networks have grown over the years independently, creating autonomous “vertical” systems with limited points of contact; around year 2000, as a consequence of the change of trend in the

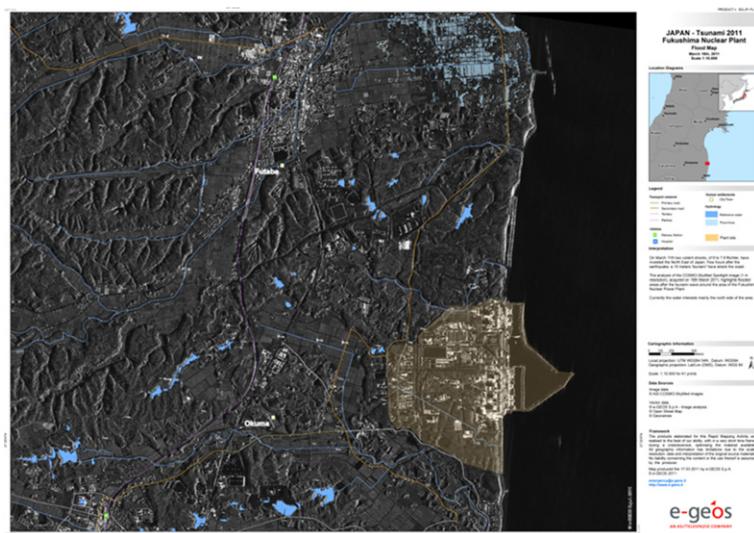


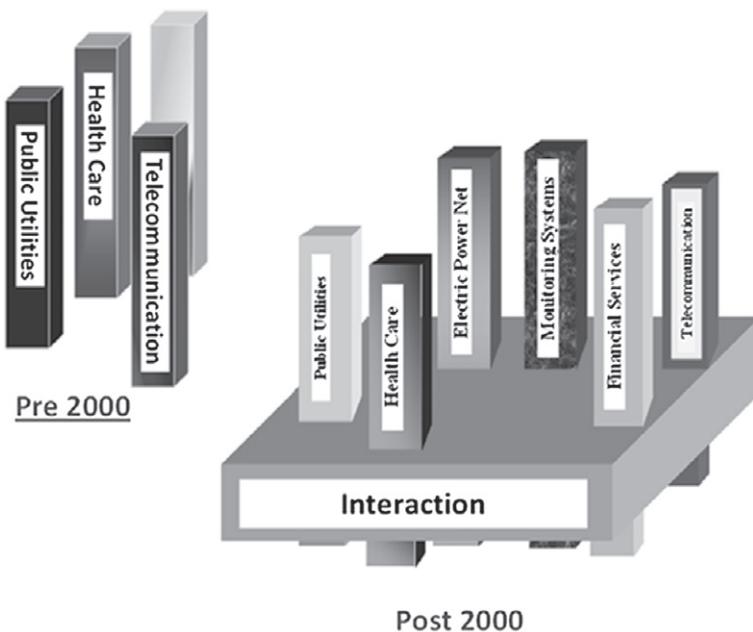
FIGURE 22.3

A CosmoSkyMed SAR image of Fukushima nuclear plant zone after the tsunami 2011 showing the flooded areas. (Courtesy of E-geos, a Telespazio Company.)

socio-techno scenario, the infrastructures have begun to share services and thus to create interconnected and interdependent systems. Nowadays infrastructures are interconnected and mutually dependent in a complex way: a phenomenon that affects one infrastructure can have a direct or indirect impact on other infrastructures, spreading on a wide geographical area and affecting several sectors of the citizen life. This is schematically represented in Figure 22.4 [31,32].

Beside the physical protection of territory, citizens, critical assets and activities, the security of information and computer systems is one the greatest challenges for a Country. Information and communication technologies have enhanced the efficiency and the comfort of the civil society on one hand, but added complexity and vulnerability on the other hand. The cyber security consists in ensuring the protection of information and property from hackers, corruption, or natural disaster, maintaining however the information and property accessible and productive to its intended users. This problem is pervasive in nearly all the systems supporting a nation: financial, energy, healthcare and transportation. The new trend toward the mobile communications is revealing a new cyber vulnerability, for instance the sheer mass of mobile endpoints gives more protection to hackers leading a cyber attack starting from a mobile. Therefore, the mobile infrastructure is becoming a critical infrastructure as well [33].

Nowadays the challenge is to understand this new scenario and to address the use of new and efficient algorithms for the information fusion in the domain of large integrated systems [34]. To integrate such heterogeneous information the necessity emerges to develop new algorithms of data fusion and information fusion to achieve an operational picture. In such scenario, where the attack can be lead with unconventional manners, information of heterogeneous sources, despite appearing uncorrelated,

**FIGURE 22.4**

Interdependencies between present infrastructures. (From [31], reprinted with permission.)

can be related and hence exploited by its fusion. Therefore particular attention is due to the information sources; Section 2.22.4 is devoted to this aspect of the problem, giving an overview of the sensors and the systems that traditionally provide information.

2.22.3 Definitions and background

Before addressing in more detail the topic of data fusion applied to the domain of Homeland Protection, it is useful to briefly review the evolution of data fusion and, more recently, the definition of the new paradigms and the introduction to high-level data fusion and information fusion.

A definition of data fusion is provided in [35]: “Data fusion is a process that combines data and knowledge from different sources with the aim of maximizing the useful information content, for improved reliability or discriminant capability, while minimizing the quantity of data ultimately retained.” Another definition is provided by the Joint Directors of Laboratories (JDL) Data Fusion Subpanel (DFS) which, in its latest revision of its data fusion model, Steinberg and Bowman [36] settle with the following short definition: “Data fusion is the process of combining data or information to estimate or predict entity states.” Due to its generality, the definition of JDL encompasses the previous one. One aspect of the data fusion process, which is not included in the first definition and is implicit in the second, is process refinement, i.e., the improving of data fusion process and data acquisition. Many authors,

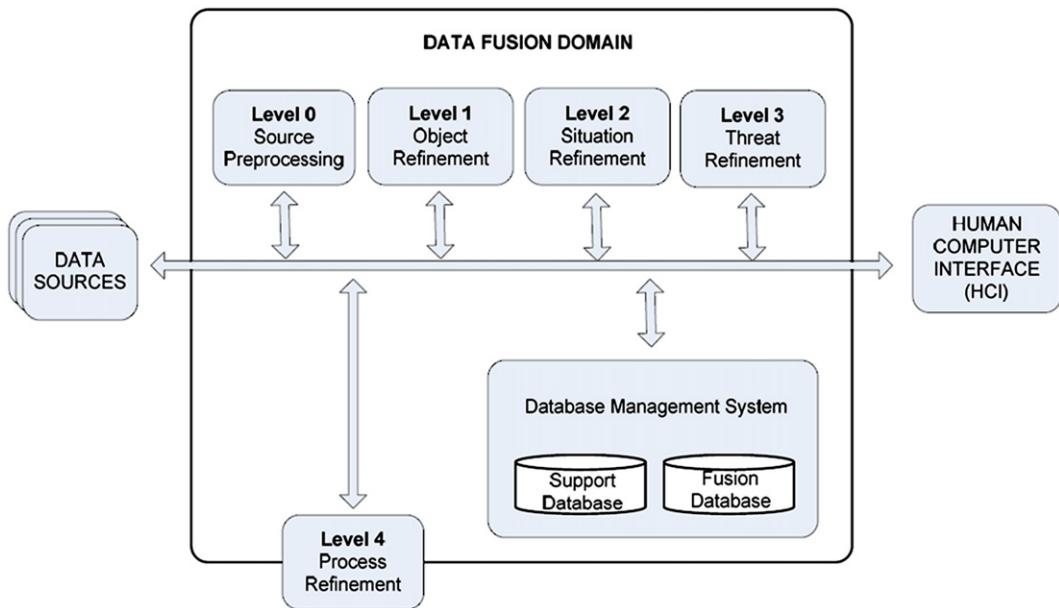
recognize process refinement and data fusion to be so closely coupled that process refinement should be considered to be a part of the data fusion process. This is not a new technique in itself, rather a framework for incorporating reasoning and learning with perceived information into systems, utilizing both traditional and new areas of research. These areas include decision theory, management of uncertainty, digital signal processing, and computer science. The data fusion process comprises techniques for data reduction, data association, resource management, and fusion of uncertain, incomplete, and contradictory information.

In 1986, an effort to standardize the terminology related to data fusion began and the JDL data fusion working group was established. The result of that effort was the conception of a process model for data fusion and a data fusion lexicon. The so-called JDL fusion model [37] is a functional model, developed to overcome potential confusion in the community and to improve communications among military researchers and system developers. The model provides a common frame of reference for fusion discussions and to facilitate understanding and recognizing the problems where data fusion is applicable. The first issue of the model, dated 1988, provided four fusion levels:

- level 1: Object refinement,
- level 2: Situation refinement,
- level 3: Threat refinement,
- level 4: Process refinement.

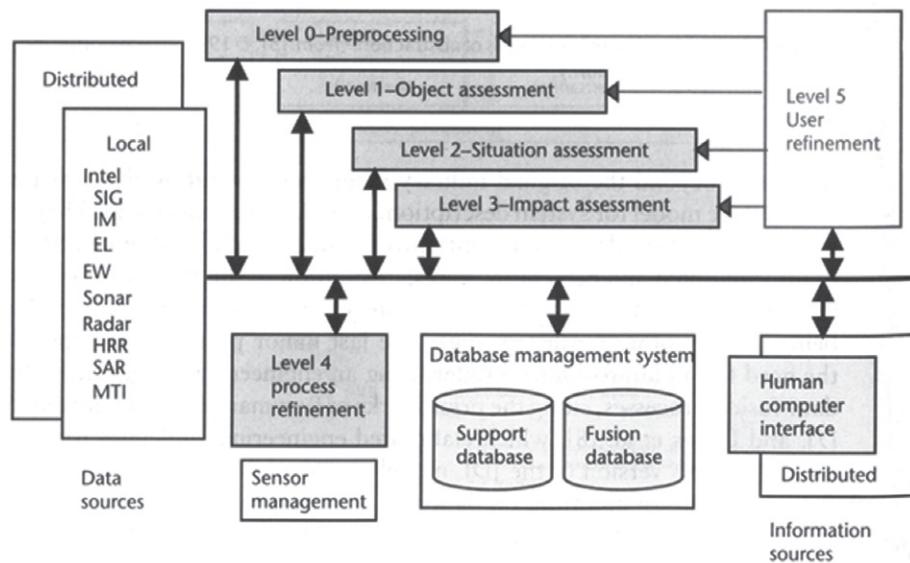
In 1998 Steinberg et al. [38] revised and expanded the JDL model to broaden the functional model and related taxonomy beyond the original military focus. They introduced a level 0 to the model for estimation and prediction of signal/object observable states on the basis of pixel/signal-level data association and characterization. They also suggested renaming and re-interpretation of level 2 and level 3 to focus on understanding the external world beyond military situation and threat focus. Figure 22.5 reports a block diagram representing this functional model. Although originally developed for military applications, the model is generally applicable. Furthermore, the model does not assume its functions to be automated, they could equally well be maintained by human labor. Hence, the model is both general and flexible. The revised JDL model levels specify logical separations in the data fusion process and divide information into different levels of abstraction depending on the kind of information they produce, where the lower levels yield more specific, and the higher more general, information. The model is divided into the following five levels [18]:

- *Level 0—sub-object assessment*: the pre-detection activities such as pixel or signal processing, spatial or temporal registration is present. Level 0 deals with the estimation and prediction of signal/object observable states on the basis of pixel/signal level data association and characterization.
- *Level 1—object assessment*: is concerned with estimation and prediction of target locations, behavior or identity. In this level, which is sometimes referred to as *multi-sensor data fusion* or *multi-sensor integration*, data is combined to assign dynamic features (e.g., velocity) as well as static (e.g., identity) to objects, hence adding semantic labels to data. This level includes techniques for data association and management of objects (including creation and deletion of hypothesized objects, and state updates of the same). Level 1 addresses the following functions: data alignment, data/object correlation, object positional/kinematic/attribute estimation, object identity estimation.

**FIGURE 22.5**

JDL model. (From [40], reprinted with permission.)

- *Level 2—situation assessment:* investigates the relations among entities such as force structure and communication roles. This level involves aggregation of level 1 entities into high-level, more abstract entities, and relations between entities. An entity in this level might be a pattern of connected objects of level 1 entities. Input data are assessed with respect to the environment, relationship among level 1 entities, and entity patterns in space and time. Level 2 addresses the following functions: object aggregation, contextual interpretation/fusion, event/activity aggregation, multi-perspective assessment.
- *Level 3—impact assessment:* outlines sets of possible courses of action and the effect on the current situation. The impact assessment, which is sometimes called *significance estimation* or *threat refinement*, estimates and predicts the combined effects of system control plans and the entities of level 2 (possibly including estimated or predicted plans of other environment agents) on system objectives. Level 3 addresses the following functions: estimate/aggregate force capabilities, predict enemy intent, identify threat opportunities, estimate implications, multi perspective assessment.
- *Level 4—process refinement:* is an element of Resource Management used to close the loop by re-tasking resources to support the objectives of the mission. Process refinement evaluates the performance of the data fusion process during its operation and encompasses everything that refines it, e.g., acquisition of more relevant data, selection of more suitable fusion algorithms, optimization of resource usage with respect to, for instance, electrical power consumption. Process refinement is sometimes called process adaption to emphasize that it is dynamic and should be able to evolve with respect both its internal properties and the surrounding environment. The function of this level

**FIGURE 22.6**

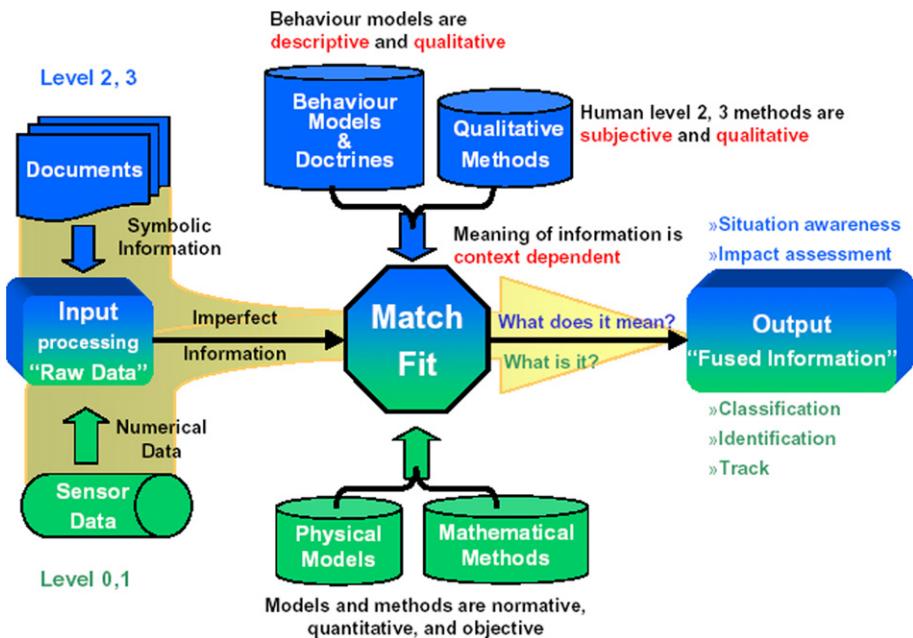
JDL model including level 5. (From [40], reprinted with permission.)

is in some literature handled by a so called meta-manager or meta-controller. It is also rewarding to compare level 4 fusion to the concept of *covert attention* in biological vision which involves, e.g., sifting through an abundance of visual information and selecting properties to extract. Level 4 addresses the following functions: evaluation (real-time control/long term improvement), fusion control, source requirements, mission management.

The 1998 revised JDL fusion model recognized the original Process Refinement level 4 function as a Resource Management function. In 2002, a level 5 was added [39,40], named User Refinement, into the JDL model to support a user's trust, workload, attention, and situation awareness. Mainly the level 5 was added to distinguish between machine-process refinement and user refinement of either human control action or the user's cognitive model. In many cases the data fusion process is focused on the machine point of view, however a full advantage can be taken by considering also the human factor, not only as a qualified expert to refine the fusion process, but also as a customer for whom the fusion system is designed. Figure 22.6, taken from [40], shows the JDL fusion model including also the level 5.

Later in [41] also a level 6, Mission Management, was added; this level tackles the adaptive determination of spatial-temporal control of assets (e.g., airspace operations) and route planning and goal determination to support team decision making and actions (e.g., theater operations) over social, economic, and political constraints.

Figure 22.7 shows a multi-sensor data fusion architecture with a representation of the levels involved into each process of data fusion. Level 0 and level 1 concern the combination of data from different

**FIGURE 22.7**

Data fusion architecture.

sensors, level 2 and level 3 are often referred to as *information fusion*. Under the proposed partitioning scheme, the same entity can simultaneously be the subject of level 0, 1, 2, and 3 fusion processes. Entity features can be estimated from one or more entity signal observations (e.g., pixel intensities, emitter pulse streams) via a level 0 data preparation/association/estimation process. The identity, location, track and activity state of an entity (whether it be a man, a vehicle, or a military formation) can be estimated on the basis of attributes inferred from one or more observations; i.e., via a level 1 data preparation/association/estimation process. The same entity's compositional or relational state (e.g., its role within a larger structure and its relations with other elements of that structure) can be inferred via level 2 processes. Thus, a single entity—anything with internal structure, whether man, machine, or mechanized infantry brigade—can be treated either as an individual, subject to level 1 observation and state estimation—or as a “situation,” subject to compositional analysis via level 2 entity/entity association and aggregate state estimation. The impact of a signal, entity, or situation on the user goal or mission can then be predicted based upon an association of these to alternative courses of action for each entity via a level 3 process.

There are also other fusion models developed on the basis of different perspectives, including a purely computational and a human information processing. In the following an overview of different models [42].

The DIKW (Data Information Knowledge and Wisdom) [43] hierarchy organizes data, information, knowledge, and wisdom in layers with an increasing level of abstraction and addition of knowledge,

starting from the bottommost data layer. The hierarchy can be considered alike the JDL data fusion model because both start from raw transactional data to yield knowledge at an increasing level of abstraction.

The JDL model and many other computational models do not simulate the complex human cognitive process that leads to “become aware,” because they do not model the fusion process from a human perspective. In 1988, Endsley defined the situation awareness as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their state in the near future” [44]. In [45,46] he identified three levels of situation awareness, namely perception, comprehension, and projection, parallel to the corresponding levels in the JDL model. Therefore the levels in the JDL model can be considered as processes producing results to help a human operator became aware of the situation. In [47] in addition to this three different aspects identified by Endsley, the model included also “intention” (i.e., the understanding of own options and courses of action relative to own goals) and “metacognition” (i.e., accounting for how reliable own situation awareness is likely to be). These levels summarize the fact that situation awareness requires the understanding of information, events, and the impact of own actions on own goals and objectives. This process involves several capabilities as learning, detection of anomalies, prediction of future behaviors, managing uncertainty, and analysis of heterogeneous sources.

The OODA (Observe-Orient-Decide-Act) loop, developed by Boyd in 1987 [48], is one of the first C4I (Command, Control, Communications, Computers, and Intelligence) architectures and it represents the classic decision-support mechanism in military information operations. Because decision-support systems for situational awareness are tightly coupled with fusion systems, the OODA loop has also been used for sensor fusion [49]. Observations in OODA refer to scanning the environment and gathering information from it; orientation is the use of the information to form a mental image of the circumstances; decision is considering options and selecting a subsequent course of action; and action refers to carrying out the conceived decision. Bedworth and O’Brien [50] report a comparison of the OODA loop to the levels of the JDL model.

The human information processing can be modeled by the Rasmussen model [51,52]. It is composed of three layers, namely skill-based, rule-based, and knowledge-based processing. The input of the process is a perception (e.g., the detection of a target by a sensor) and the output is an action. An example of result at the first level may be represented by the automatic identification of a tank by processing of raw sensors data; at the next level an enemy unit composition can be identified on the basis of its number and relative locations. Knowledge-based behavior represents the most complex cognitive processing used to handle novel, complex, situations where no routine or rule is available to manage situations. An example of this type of processing may be the interpretation of unusual behavior, and the consequent generation of a course of actions based on enemy unit size and behavior.

The Generic Error modeling System (GEMS) [53] is an extension of Rasmussen’s approach, which describes the competencies needed by workers to perform their roles in complex systems. GEMS describes three major categories of errors: skill-based slips and lapses, rule-based mistakes, and knowledge-based mistakes.

Table 22.1, from [42], shows a correspondence, and not a comparison, among levels and layers of various models presented before. This table is intended as a guide to identify the components of a data fusion architecture, where the separation between the columns is not so sharp. Notice that the JDL model does not explicitly model into a level the action consequent to the threat assessment. The action

Table 22.1 Comparison Among Fusion Models (From [42], Reprinted with Permission)

JDL Data Fusion Model	Level 1 = Object Assessment	Level 2 = Situation Assessment	Level 3 = Threat Assessment	Action (Not Explicit In The JDL Model)	Level 4 = Process Refinement
Knowledge pyramid	Data	Information and knowledge	Knowledge	Wisdom	n/a
Situation awareness	Perception	Comprehension	Projection	Decision	n/a
OODA loop	Observe	Orient	Decide	Act	n/a
Information processing hierarchy	Skill/rule-based processing	Rule/knowledge-based processing	Rule/knowledge-based processing	n/a	n/a

level, with the sense of a reaction is only in part included in the process refinement level 4, for this reason the column “action” has been inserted in the table, to allow a more clear correspondence with the other models that explicitly account for the reaction. The JDL model is the one that allows the most global view of the data fusion process from an operative perspective: there is not any correspondence of the other models with JDL level 4.

2.22.4 The information sources

This section gives a broad and very general description of the basic categories of intelligence that are the source of data/information employed to perform the fusion process. The USAF (United States Air Force) in 1998 first and the ODNI (Office of Directors of National Intelligence) later in 2008 described in their studies that there are six basic intelligence categories [54,55]:

- Signals Intelligence (SIGINT),
- Imagery Intelligence (IMINT),
- Measurement and Signature Intelligence (MASINT),
- Human Intelligence (HUMINT),
- Open-Source Intelligence (OSINT),
- Geospatial Intelligence (GEOINT).

In addition, there is also Scientific and Technical (S&T) Intelligence resulting from the analysis of foreign scientific and technical information. In the following is an overview of the categories.

SIGINT is achieved by the interception/detection of electromagnetic (em) emissions. SIGINT includes Electronic Intelligence (ELINT) and Communications Intelligence (COMINT). The former derives from the processing and analysis of em radiation emitted from emitters, in most of cases radars, not employed for communications, other than nuclear detonations or radioactive sources. An emitter may be related closely to a specific threat. The information that can be achieved by a typical ESM (Electronic Support Measures) device consists of an estimate of the emitter category, location, with a certain accuracy, and various electronic attributes, such as frequency and pulse duration. This information can

be employed in a high-level fusion process. COMINT derives from the processing and analysis of intercepted communications from emitters. The communications may be encrypted and they may be of several forms such as voice, e-mail, fax and the like.

IMINT is obtained by sensors working in several bandwidths which are able to produce a view of the scenario or of the specific target: electro/optical sensors, infrared, radar (e.g., Synthetic Aperture Radar (SAR) and Inverse SAR (ISAR), and Moving Target Indicator (MTI)), laser, laser radar (LADAR), and multi-spectral sensors. Each sensor has a unique capability. Some work in all weather conditions, some may work also in night conditions, and some produce high-quality images with detectable signatures.

MASINT is obtained by the collection and the analysis of several and heterogeneous sensors and instruments usually working in different regions or domains of the em spectrum, such as infrared or magnetic fields. MASINT includes Radar Intelligence (RADINT), Nuclear Intelligence (NUCINT), Laser Intelligence (LASINT), and Chemical and Biological Intelligence (CBINT). RADINT, for example, is a specialized form of ELINT, which categorizes and locates as active or passive collection of energy reflected from a target.

HUMINT is the collection of information derived by the human contact. Information of interest might include target name, size, location, time, movement, and intent. HUMINT typically includes structured text (e.g., tables, lists), annotated imagery, and free text (e.g., sentences, paragraphs). HUMINT provides comprehension of adversary actions, capability and capacity, plans and intentions, decisions, research goals and strategies.

OSINT is publicly available information appearing either in print or in electronic form including radio, television, newspapers, journals, the Internet, commercial databases, videos, graphics, and drawings. OSINT can be considered as a complement to the other intelligence categories and can be used to fill gaps and improve accuracy and confidence in classified information. A special mentioning is for the Internet, that, with its blogs, e-mails, videos, messages and mobile systems, favors an ever greater interaction between users. Moreover notice that there is a little overall planning in the development of the World Wide Web, but rather a myriad of initiatives by individuals of small groups. Government have always tried to use telephone tapping, surveillance, files, i.e., intelligence. Now this is possible on a different scale given the technical possibilities offered by satellites, mobile, phones, credit cards management systems, information storage, etc. From the topological point of view, Internet is a scale-free complex network with a power-law of the distribution of the nodes [56]; this technical remark should be considered in the data exploitation analysis.

GEOINT is the analysis and the visual representation of the activities on the earth related to the security achieved by the sensors (radar, optical, IR, multispectral) deployed in the space. The information related to GEOINT is obtained through an integration of imagery, imagery intelligence, and geospatial information.

2.22.5 Homogeneous sensor networks

Stand-alone sensors usually provide a fragmentary view of a complex situation of interest. A significant enhancement of performance can therefore be accomplished by a combination of networked sensors in the close vicinity to the region of interest. Using efficient methods of centralized or decentralized

multiple sensor fusion, the quality of the produced situation picture can significantly be improved. In practice, improvements with respect to the following aspects are of interest:

- production of accurate and continuous tracks (e.g., objects, persons, single vehicles, group objects),
- system reaction rates (e.g., track extraction, detection of target maneuvers, track monitoring),
- sustainment of reconnaissance capabilities in case of either system or network failures (e.g., graceful degradation),
- system robustness against jamming and deception,
- compensation of degradation effects (e.g., sensor misalignment, limited sensor resolution),
- robustness against sub-optimal real-time realizations of sensor data fusion algorithms,
- processing of eventually delayed sensor data (e.g., out-of sequence measurements).

In the following, several sections tackle different aspects related to homogeneous sensor networks.

2.22.5.1 Sensor configuration

Sensor fusion networks can be categorized according to the type of sensor configuration. Durrant-Whyte distinguishes three types of sensor configuration as schematized in Figure 22.8 [57,58].

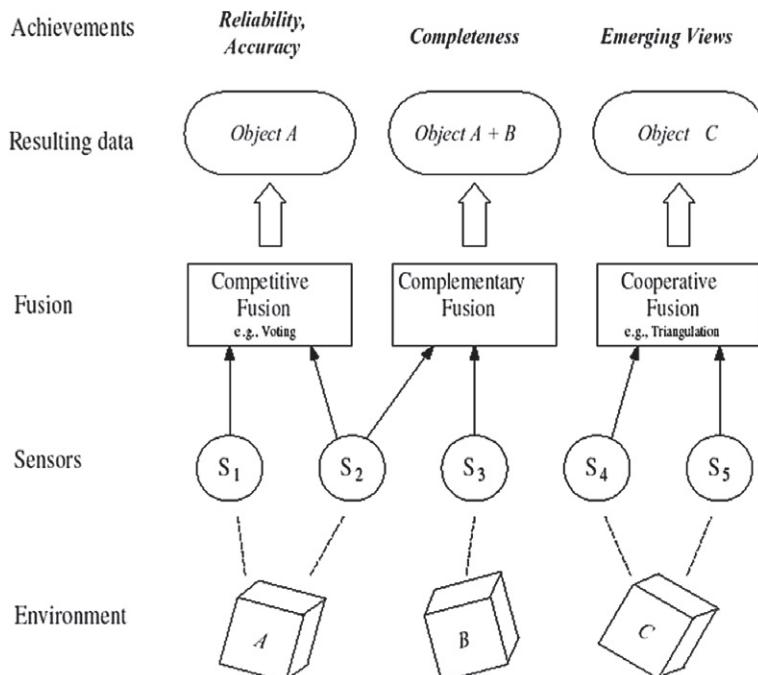


FIGURE 22.8

Sensors configuration (from [57], reprinted with permission).

Competitive sensor data fusion: Sensors are configured competitive if each sensor delivers independent measurements of the same property. Sensor data represent the same attribute, and the fusion is to reduce uncertainty and resolve conflicts. Competitive sensor configuration is also called a redundant configuration. Sensors S1 and S2 in Figure 22.8 represent a competitive configuration, where both sensors redundantly observe the same property of an object in the environment space.

Complementary sensor data fusion: A sensor configuration is called complementary if the sensors do not directly depend on each other, but can be combined to give a more complete image of the phenomenon under observation. Fusion of the sensor data provides an overall and complete model. Examples for a complementary configuration is the employment of multiple cameras each observing disjoint parts of a room, or using multiple spectrum signatures to identify a land cover type, or using different waveform to identify an aircraft type. Sensor S2 and S3 in Figure 22.8 represent a complementary configuration, since each sensor observes a different part of the environment space.

In both competitive and complementary sensor configurations, there is an improvement of the accuracy of the target characteristics estimation consequent to the data fusion. In their seminal work H. Cramer and C.R. Rao found how to compute the best theoretical accuracy that can be achieved by an estimator. The lower bound of accuracy, i.e., the mean square error of any unbiased estimator, is given by the inverse of the so-called Fisher Information Matrix (FIM). The computation of the CRLB (Cramer-Rao Lower Bound) applies to problems involving the maximum likelihood estimation of unknown constant parameters from noisy measurements [59]. The best achievable improvement of target location and track accuracy can be quantified by the reduction of the CRLB consequent to the track fusion. In [60] this computation is reported in case of fusion of data from two sensors with an ideal unitary detection probability. In [61,62] the same computation has been proposed in case of detection probability less than one and false alarm probability higher than zero.

Cooperative sensor data fusion: A cooperative sensor network uses the information provided by two independent sensors to derive information that would not be available from the single sensors. An example for a cooperative sensor configuration is stereoscopic vision: by combining two-dimensional images from two cameras at slightly different viewpoints a three-dimensional image of the observed scene is derived. Cooperative sensor fusion is the most difficult to design, because the resulting data are sensitive to inaccuracies in all individual participating sensors. Thus, in contrast to competitive fusion, cooperative sensor fusion generally decreases accuracy and reliability. Sensor S4 and S5 in Figure 22.8 represent a cooperative configuration. Both sensors observe the same object, but the measurements are used to form an emerging view on object C that could not have been derived from the measurements of S4 or S5 alone.

These three categories of sensor configuration are not mutually exclusive. Many applications implement aspects of more than one of the three types. An example for such a hybrid architecture is the application of multiple cameras that monitor a given area. In regions covered by two or more cameras the sensor configuration can be competitive or cooperative. For regions observed by only one camera the sensor configuration is complementary.

2.22.5.2 Classical approach to surveillance

Sensor networks have countless applications, for example, we mention the sensor networks used in computer science and telecommunications, in biology, where they can be used to monitor the behavior of animal species such as birds or fishes, and in habitat monitoring, where they can be used to provide

real-time rainfall and water level information used to evaluate the possibility of flooding. In the field of Homeland Protection one of the main tasks to be assigned to a sensor network is the surveillance with its most general significance. Automatic surveillance is a process of monitoring the behavior of selected objects (targets and/or anomalies) inside a specific area by means of sensors. A target generally consists of an object (e.g., a tank close to a land border or a rubber approaching to the coast) whose presence and characteristics can be detected and estimated by the sensor; an anomaly consists in a non usual behavior (e.g., a jeep moving off-road, the increasing of the radioactivity level within an area) that can be revealed by the sensor. Sensors typically provide the following functions:

- detection of a targets or anomalies inside the surveillance area,
- estimation of target position or the anomaly localization and extension,
- monitoring of the target kinematic (tracking) or of the anomaly behaviors,
- classification and/or recognition of the targets.

To perform the previous functions, the sensors can be organized on the bases of several approaches. The classical approach to surveillance of wide areas is based on the use of a single or few sensors with long range capabilities. The signal received by the single sensor is processed by means of suitable digital signal processing subsystems. In this case the sensors are costly, with adequate computation and communication capabilities. Sensors are normally located in properly selected sites, to mitigate terrain masking problems; nevertheless, they provide different performance depending on the location of target inside the surveillance area. Typical sensors are radars (ground-based, air-borne, ship-borne or space-based), infrared or TV cameras, seismic, acoustical, radioactive sensors. Usually in this kind of networks, as represented in Figure 22.9, the information traffic goes from the sensor nodes to a single

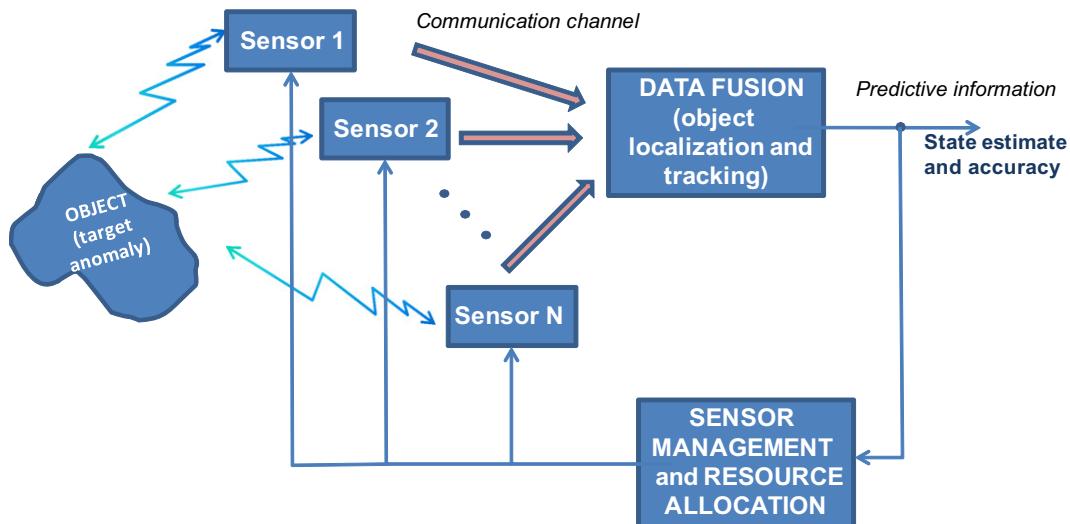


FIGURE 22.9

Block-diagram for optimal system resource management in a sensor network.

sink node called information fusion center that performs the target localization and tracking. According to the information received from the sensors the fusion center monitors the area where the sensors are deployed and decides, on the basis of the state estimates and their accuracy (e.g., a covariance matrix for a Kalman filter or a particle cloud for a particle filter) the actions to take.

In [63] an example of high-performance radar netted for Homeland Security application with a centralized data fusion process is described. The same classical approach is presented in [64] where this kind of sensor network is employed for natural resource management and bird air strike hazard (BASH) applications.

However if an intruder reaches and neutralizes the fusion center, the communication between the network nodes are interrupted and the whole network is exposed to the risk of becoming useless as a network even if the individual sensors may still be all working.

2.22.5.3 Collaborative signal and information processing (CSIP)

Nowadays, a novel approach to the automatic surveillance has been adopted; it is based on the use of many sensors with short range capabilities, low costs, and limited computation and communication capabilities. In case of a huge number of sensors, the use of information fusion centers is unpractical and their functioning is based on the information exchange between “near-by” sensors. The sensors can be distributed in fixed positions of the territory, but they could also be deployed adaptively to the change of the scenario. There are several approaches: they can be randomly distributed inside the surveillance area and if the number of sensors is high, the performance of the surveillance system can be considered independent of the location of the targets; then the signal received by each sensor is processed using the computational capabilities of a sub-portion of the sensor system and employed to re-organize dynamically the network. Sensors may be agile in a variety of ways, e.g., the ability to reposition, point an antenna, choose sensing mode, or waveform. Notice that the number of potential tasking of the network grows exponentially with the number of sensors. The goal of sensor management in a large network is to choose actions for individual sensors dynamically so as to maximize overall network utility. This process is called Collaborative Signal and Information Processing (CSIP) [65]. One of the central issues for CSIP to address is energy-constrained dynamic sensor collaboration: how to dynamically determine who should sense, what needs to be sensed, and who the information must be passed onto. This kind of processing system allows a limitation in the consumption of power. Applying a surveillance strategy which accounts for the target tracking accuracy and the sensor random location, only a limited number of sensors are awake and follow/anticipate the target movement; thus, the network self-organizes to detect and track the target, allowing an efficient performance from the energetic point of view with limited sensor prime power and with a reduced number of sensors working in the whole network. For example in [66], instead of requesting data from all the sensors, the fusion center iteratively selects sensors for the target localization: first a small number of anchor sensors send their data to the fusion center to obtain a coarse location estimate, then, at each step a few non-anchor sensors are activated to send their data to the fusion center to refine the location estimate iteratively. Moreover the possibility to actively probe certain nodes allows to disambiguate multiple interpretations of an event.

In [67] the techniques of information-driven dynamic sensor collaboration is introduced. In this case an information utility measurement is defined as the statistical entropy and it is exploited to evaluate the

benefits in employing part of the network that consequently is re-organized. Other cost/utility functions can be employed as criteria to dynamically re-organize the sensor network as described in [68,69].

Several analytical efforts have been done to evaluate the performance of such networks in terms of tracking accuracy. As usual the CRLB has been taken as reference of the best achievable accuracy; in particular a new concept of conditional PCRLB (Posterior Cramer Rao Lower Bound) is proposed and derived in [70]. This quantity is dependent on the actual observation data up to the current time, and is implicitly dependent on the underlying system state. Therefore, it is adaptive to the particular realization of the underlying system state and provides a more accurate and effective online indication of the estimation performance than the unconditional PCRLB. In [71,72] the PCRLB is proposed as a criterion to dynamically select a subset of sensors over time within the network to optimize the tracking performance in terms of mean square error. In [73] the same criterion is proposed as a framework for the systematic management of multiple sensors in presence of clutter.

2.22.5.4 Self-Organizing Sensor Networks

Self-organization can be defined as the spontaneous set-up of a globally coherent pattern out of local interactions among initially independent components. Sensors are randomly spread out over a two dimensional surveillance area. In a self-organized system, its elements affect only close elements; distant parts of the system are basically unaffected. The control is distributed, i.e., all the elements contribute to the fulfillment of the task. The system is relatively insensitive to perturbations or errors, and have a strong capacity to restore itself. Initially independent components form a coherent whole able to efficiently fulfill a particular function [74]. Flocks of birds, shoals of fish, swarms of bees are examples of self-organizing systems; they move together in an elegantly synchronized manner without a leader which coordinates them and decides their movement. It has been shown that flocks of birds self-organize into V-formations when they need to travel long distances to save energy, by taking advantage of the upwash generated by the neighboring birds. Cattivelli and Sayed [75] propose a model for the upwash generated by a flying bird, and shows that a flock of birds is able to self-organize into a V-formation as if every bird processes spatial and network information by means of an adaptive diffusive process. This result has interesting implications. First, a simple diffusion algorithm is able to account for self-organization of birds. Second, according to the model, that birds can self-organize on the basis of the upwash generated by the other birds. Third, some information is necessarily shared among birds to reach the optimal flight formation. The paper also proposes a modification to the algorithm that allows birds to organize, starting from a V-formation, into a U-formation, leading to an equalization effect, where every bird in the flock observes approximately the same upwash. The same algorithm based on birds flight is extended in [76] to the problem of distributed detection, where a set of sensors/nodes is required to decide between two hypotheses on the basis of the collected measurements. Each node makes individual real-time decisions and communicates only with its immediate neighbors, in order that any fusion center is not necessary. The proposed distributed detection algorithms are based on diffusion strategies described in [77–79] and their performance is evaluated by means of classical probabilities of detection and false alarms.

These diffusion detection schemes are attractive in the context of wireless and sensor networks thanks to their intrinsic adaptability, scalability, improved robustness to node and link failure as compared to centralized schemes, and their potential to save energy and communication resources.

2.22.5.5 Self-synchronization mechanism applied to sensor network

Several studies have shown how a simple self-synchronization mechanism, borrowed from biological systems, can form the basic tool for achieving globally optimal distribution decisions in a wireless sensor network with no need for a fusion center. Self-synchronization is a phenomenon first observed between pendulum clocks (hooked to the same wooden beam) by Christian Huygens in 1658. Since then, self-synchronization has been observed in a myriad of natural phenomena, from flashing fireflies in South East Asia to singing crickets, from cardiac pacemaker or neuron cells to menstrual cycles of women living in strict contact with each other [80]. The goal of these studies is to find a strategy of interaction among the sensors/nodes that could allow them to reach globally optimal decisions in terms of a “consensus” value in a totally decentralized manner. Distributed consensus algorithms are indeed techniques largely studied in distributed computing [81, 82]. The approaches suggested in [83, 84] give a form of consensus achieved through self-synchronization that may result critical in wide-area networks, where propagation delays might induce an ambiguity problem. This problem is overcome in [85–87] where also a model of the network and of the sensors is proposed. Each of the N nodes composing the network is equipped with four basic components: (1) a transducer that senses the physical parameter of interest y_i (e.g., temperature, concentration of contaminants, radiation, etc.); (2) a local processing unit that provides a function $g_i(y_i)$ of the measurements; (3) a dynamical system, initialized with the local measurements, whose state $x_i(t)$ evolves as a function of its own measurement $g_i(y_i)$ and of the state of nearby sensors; (4) a radio interface that makes possible the interactions among the sensors. The criterion to reach a consensus value is the asymptotical convergence toward a common value of all the derivatives of the state, for any set of initial conditions and for any set of bounded. This condition makes the convergence to the final consensus independent of the network graph topology. However the topology has an impact on several aspects: the overall energy necessary to achieve the consensus and the convergence time. In general there exists a trade-off between the local power transmitted by each sensor and the converge time depending on the algebraic connectivity of the network graph, as shown in [88]. In the practical applications these aspects cannot be neglected; for instance, the design of a network should account for the precision to achieve, and the time to get the consensus value at the given precision, versus such constraints as the energy limitations of the sensors. A global overview of the problem is given in [89].

2.22.5.6 From theory to real application problems

Moving from the functional model to a working implementation in a real environment involves a number of design considerations: including what information sources to use and what fusion architecture to employ, communication protocols, etc.

Admittedly, the fusion of data is decoupled from the actual number of information sources and, hence, does not require necessarily multiple sensors: the fusion, in fact, may be performed also on a temporal sequence of data that was generated by a single information source (e.g., a fusion algorithm may be applied to a sequence of images produced by a single camera sensor). However, employing a number of sensors provides many advantages as well explained in the previous Sections. Unsurprisingly, there are also difficulties associated with the use of multiple sensors.

A missed sensor registration may cause a failure in the correct association between signals or features of different measurements. This problem and the similar data association problem are very important

and apply also to single sensor data processing. To perform data registration, the relative locations of the sensors, the relationship between their coordinate systems, and any timing errors need to be known, or estimated, and accounted for otherwise a mismatch between the compiled picture and the truth may result. An overstated confidence in the accuracy of the fused output, and inconsistencies between track databases, such as multiple tracks that correspond to a single target may appear. A missed registration can result from location and orientation errors of the sensor relative to the supporting platform, or of the platform relative to the Earth, such as a bearing measurement with an incorrect North alignment. Errors may be present in data time stamping, and numerical errors may occur in transforming data from one coordinate system to another. Automatic sensor registration can correct for these problems by estimating the bias in the measurements along with the kinematics of the target. However, the errors in sensor registration need to be known and accounted for [90]. In [91] a maximum likelihood (ML) algorithm for registration is presented using a recursive two-step optimization that involves a modified Gauss-Newton procedure to ensure fast convergence. In [92] a novel joint sensor association, registration, and fusion is performed exploiting the expectation–maximization algorithm incorporated with the linear Kalman filter (KF) to give simultaneous state and parameter estimates. The same approach can be followed also with non linear filtering techniques as the Extended KF (EKF) and the Unscented KF (UKF) as proposed in [93], where also the performance is evaluated by means of the PCRLB.

Next to the spatial sensor registration also the temporal alignment cannot be neglected. For instance, a critical aspect of a sensor network is its vulnerability to temporary node sleeping, due to duty-cycling for battery recharge, permanent failures, or even intentional attacks.

Other realistic problems, such as conflicting information and noise model assumptions, may enable the use of some fusion techniques. Noisy input data sometimes yield conflicting observations, a problem that has to be addressed and which does not arise in single sensor data processing. The administration of multiple sensors have to be coordinated and information must be shared between them.

2.22.5.7 Rethinking mathematical algorithms for net-centric approaches

Most of the optimization algorithms have been developed in a centralized framework, i.e., they have been conceived to perform centralized data fusion process. In the last years the trend is to employ network centric approaches, and the mathematical optimization algorithms must be able to support this approach. In the following an example of the adaptation of a “centralized-conceived” algorithm to the new trend is presented.

Consider the following minimization problem to solve:

$$\min_{(x,y,z) \in \mathbb{R}^3} f(x, y, z) = \alpha x^2 + \beta y^2 + \gamma z^2 + 2\delta xy + 2\varepsilon xz + 2\eta yz, \quad (22.1)$$

where $\alpha, \beta, \gamma, \delta, \varepsilon, \eta$ are real positive values and the function $f(x, y, z) = w \geq 0$ represents an ellipsoid function, whose axes do not coincide with the reference frame axes if $\delta \neq 0, \varepsilon \neq 0, \eta \neq 0$. The problem of Eq. (22.1) can be solved by the steepest descent method in a centralized fusion process frame, hence it will be named “centralized steepest descent.” The centralized steepest descent method when used to solve minimization problems is an iterative procedure that, beginning from an initial guess, updates at every iteration the current approximation of the solution of the function to minimize with a step in the direction of the gradient of the own function. In a network centric approach it may

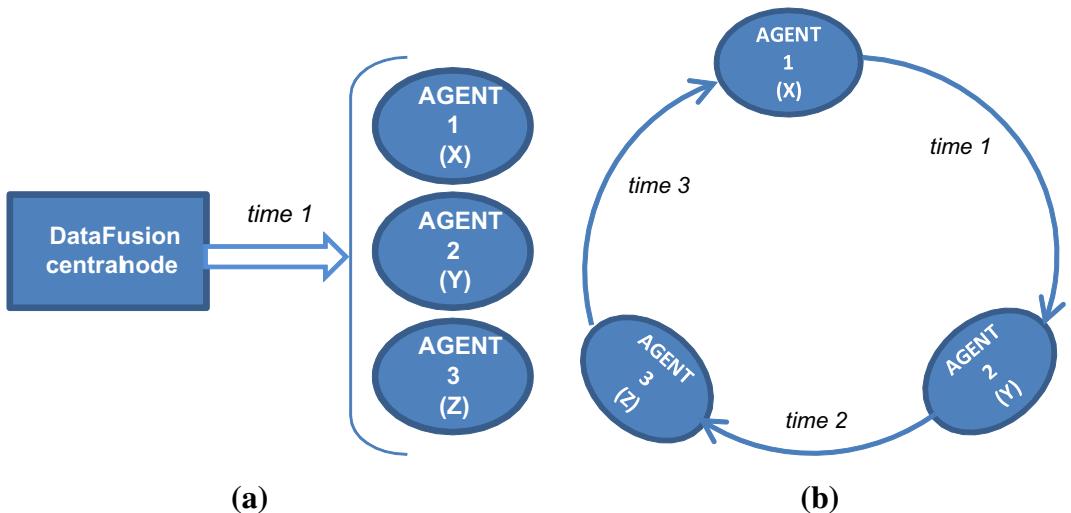


FIGURE 22.10

(a) centralized data fusion process model; (b) network centric data fusion process model.

be solved by the application of the Jacobi method² usually employed for the iterative solution of linear system equation.

Consider three agents (namely agent 1, 2, and 3) controlling the three variables x , y , and z . In the centralized data fusion process, represented in Figure 22.10a, the communication between the three agents is completely performed at the same instant of time; in the network centric case this does not happen. Consider the model of Figure 22.10b with the following communication scheme:

- agent 1 communicates to agent 2,
- agent 2 communicates agent 3,
- agent 3 communicates to agent 1;

moreover the communications among agents is not instantaneous, but they succeeds in time.

The method of the centralized steepest descent applied to the function $f(x, y, z)$, given a starting point (x^0, y^0, z^0) , is based on the following iterations:

$$\begin{aligned} x^{k+1} &= x^k - h(2\alpha x^k + 2\delta y^k + 2\varepsilon z^k), \\ y^{k+1} &= y^k - h(2\beta y^k + 2\delta x^k + 2\eta z^k), \\ z^{k+1} &= z^k - h(2\gamma z^k + 2\epsilon x^k + 2\eta y^k), \end{aligned} \quad (22.2)$$

where $k = 0, 1, \dots$, and $h \geq 0$ represents the step employed in the steepest descent method.

²The Jacobi method is an algorithm for determining the solutions of a system of linear equations with largest absolute values in each row and column dominated by the diagonal element. Each diagonal element is solved for, and an approximate value plugged in. The process is then iterated until it converges. This algorithm is a stripped-down version of the Jacobi transformation method of matrix diagonalization. The method is named after German mathematician Carl Gustav Jakob Jacobi [96].

A network centric steepest descent method can be derived by the communication scheme represented in Figure 22.10b and described below. Given the starting point (x^0, y^0, z^0) , the following iterations can be done:

$$\begin{aligned} x^{k+1} &= x^k - h(2\alpha x^k + 2\delta y^k + 2\varepsilon z^k), \\ y^{k+1} &= y^k, \\ z^{k+1} &= z^k, \\ x^{k+1} &= x^{k+1}, \\ y^{k+2} &= y^{k+1} - h(2\beta y^{k+1} + 2\delta x^{k+1} + 2\eta z^{k+1}), \\ z^{k+1} &= z^{k+1}, \\ x^{k+3} &= x^{k+2}, \\ y^{k+3} &= y^{k+2}, \\ z^{k+3} &= z^{k+2} - h(2\gamma z^{k+2} + 2\varepsilon x^{k+2} + 2\eta y^{k+2}), \end{aligned} \quad (22.3)$$

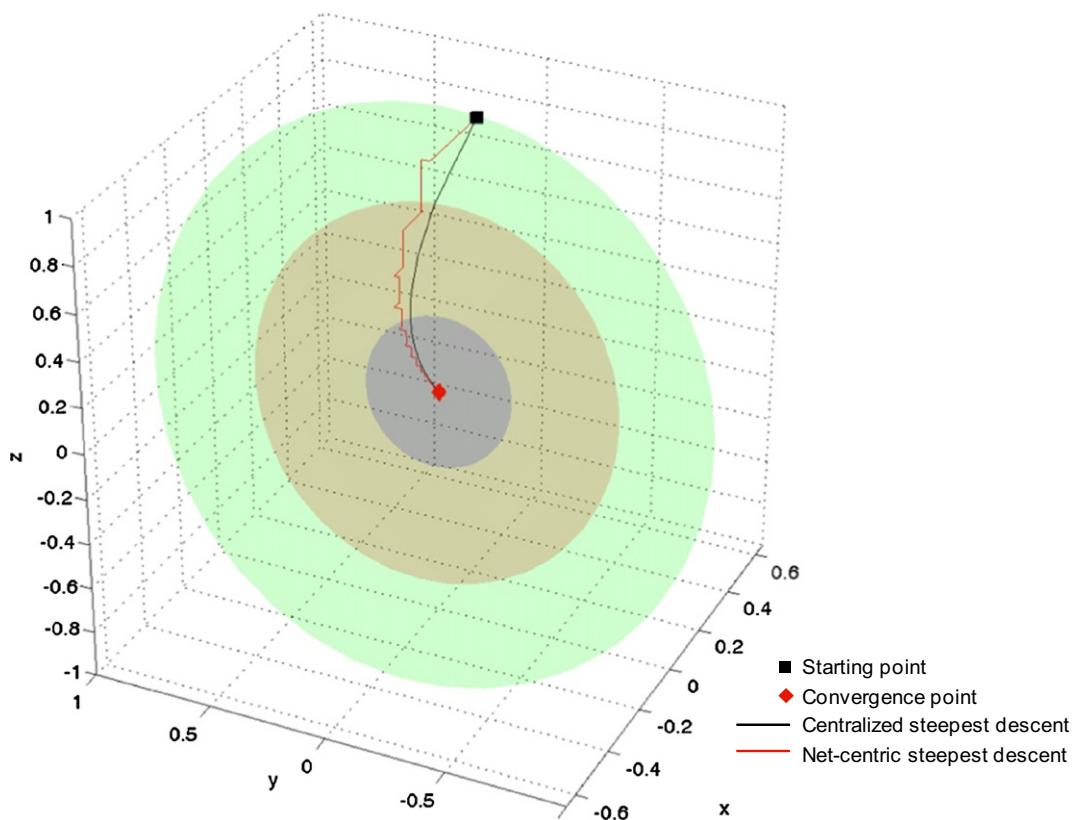
where $k = 0, 3, 6, \dots$ and $h \geq 0$ represents the step employed in the steepest descent method. Figure 22.11 shows the comparison of the two methods for the previous model. Note that the three agents in the net-centric approach are those looking at the function to be minimized along the x , y , and z axes respectively. The black square and the red diamond in the curves represent respectively the starting point of the iteration and the final position. The black solid line shows the trajectory described by the variables (x, y, z) obtained by the application of the centralized steepest descent method; the red solid line shows the behavior of the variables obtained by the net-centric steepest descent method. Note that the red line approaches the minimum by moving along the x , y , and z axes separately. The ellipsoids of Figure 22.11 represent the iso-level surfaces of the objective function. Notice that the telecommunication network modeled for the net-centric steepest descent determines the usual Jacobi iteration employed for the solution of linear systems associated to minimization problems [94–96]. In the following Section 2.22.6.1 this approach is applied to reach the optimal deployment of a sensor network.

2.22.6 Real study cases: novel approaches to sensor networks

This section proposes several study cases of sensor networks employing novel approaches. Section 2.22.6.1 proposes an optimization method, projected in the network centric frame, to obtain the optimal deployment of a cooperative sensor network; Section 2.22.6.2 describes how to employ the so-called bio-inspired models of dynamic sensor collaboration in a chemical sensor network to detect a chemical pollutant; finally Section 2.22.6.3 gives a description of the typical problem of detection of radioactive sources.

2.22.6.1 A cooperative sensor network: optimal deployment and functioning

This section presents a mathematical model for the deployment of a sensor network, for the creation of consensus values from the noisy data measured and a statistical methodology to detect local anomalies

**FIGURE 22.11**

Comparison between the trajectories computed by the centralized and the network centric steepest descent.

in these data. A local anomaly in the data is associated to the presence of an intruder. The model of sensor network presented here is characterized by the absence of a fusion center. In other words the deployment, the construction of the consensus values, and the detection of local anomalies in the data are the result of local interactions between sensors. Nevertheless the local interactions will lead to global solution of the considered problem. This is an example of model of a *network centric* sensor network. The sensors are assumed to be identical and they measure a quantity pertinent to the properties of the area to survey able to reveal the presence of an intruder. In the proposed study case the sensors are able to measure the temperature of the territory in the position or in the “area” where they are located; in absence of anomalies there is a uniform temperature on the territory where the sensors are deployed. The sensor measures are noisy and can be considered synchronous. This measurement process is repeated periodically in time with a given frequency. From these measures a “consensus” temperature is deduced, pertinently to the territory where the sensors are deployed and an estimate of the magnitude of the noise contained in the data. Finally using these consensus values as reference values local anomalies are



FIGURE 22.12

Territory of the city of Urbino (Italy) selected for the study case.

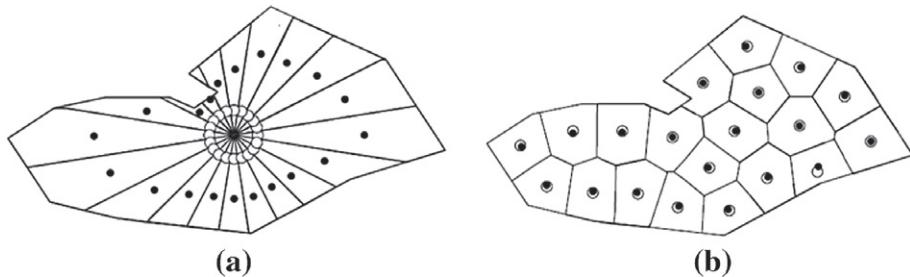
detected by the individual sensors. In the following we give some analytical details of the consensus method [97].

Let Ω be a bounded connected polygonal domain in two dimensional real Euclidean space R^2 . The domain Ω represents the territory where the sensor network must be deployed; in our case the downtown part of the Italian city of Urbino, shown in Figure 22.12. Let $\|\cdot\|$ denote the Euclidean norm of in R^2 . Consider N sensors s_1, s_2, \dots, s_N , located respectively, in the points $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N \in \Omega$, assumed to be distinct. To the sensor network deployed in the points $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$ corresponds a graph whose nodes are the sensors location and whose edges join the sensors able to communicate between themselves. This graph is assumed to be connected and can be imagined as laid on the territory. The assumption that the graph is connected is equivalent to assuming that the sensors constitute a network. For $i = 1, 2, \dots, N$, a polygonal region $\Omega_i \subset \Omega$ is associated to each sensor s_i ; this region is defined by the condition that the points belonging to Ω_i are closest to the sensor s_i , that is they are closest to $\underline{\xi}_i$, than to any other of the remaining sensors s_j located in $\underline{\xi}_j$, $j \neq i$, $j = 1, 2, \dots, N$. It follows:

$$\Omega_i = \left\{ \underline{x} \in \Omega : i = \arg \min \left(f(j) = \left\| \underline{x} - \underline{\xi}_j \right\| \right), \ j = 1, 2, \dots, N \right\}, \quad i = 1, 2, \dots, N. \quad (22.4)$$

When for a given $\underline{x} \in \Omega$ the minimizer of the function $f(j) = \|\underline{x} - \xi_j\|$, $j = 1, 2, \dots, N$ is not unique we attribute \underline{x} to Ω_i , where i is the smallest index between the indices that are minimizers of the function f .

The collection of subsets $\{\Omega_1, \Omega_2, \dots, \Omega_N\}$ defined in Eq. (22.4) and further specified by the condition above is a partition of Ω and it is a Voronoi partition of Ω associated to the Voronoi centers $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$, as represented in Figure 22.13 [98], where the sets $\Omega_1, \Omega_2, \dots, \Omega_N$ are the Voronoi cells. The sensor s_i is located in $\underline{\xi}_i$, with $\underline{\xi}_i \in \Omega_i$, $i = 1, 2, \dots, N$, and monitors the sub-region Ω_i of Ω . Note that there is a Voronoi partition of Ω associated to each choice of the Voronoi centers

**FIGURE 22.13**(a) Voronoi partition of Ω ; (b) optimal Voronoi partition of Ω .

$\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$, that, completed with the graph that defines the communication between the sensors, constitute a deployment of the sensors s_1, s_2, \dots, s_N on the territory Ω .

After the definition of a Voronoi partition of Ω , we want to determine the optimal one with respect to a pre-specified criterion, that in this study case is the fact that the Voronoi centers $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$ should coincide (as much as possible) with the centers of mass of the corresponding Voronoi cells $\Omega_1, \Omega_2, \dots, \Omega_N$. This property translates in mathematical terms the request that the sensors are well distributed on the territory. That is what is called optimal Voronoi partition, i.e., the Voronoi partition associated to the Voronoi centers whose coordinates $\underline{\xi}_1^*, \underline{\xi}_2^*, \dots, \underline{\xi}_N^*$ are the solution of the following problem:

$$\min_{\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N} F(\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N) = \sum_{j=1}^N \left\| \underline{B}_j - \underline{\xi}_j \right\|^2 \quad (22.5)$$

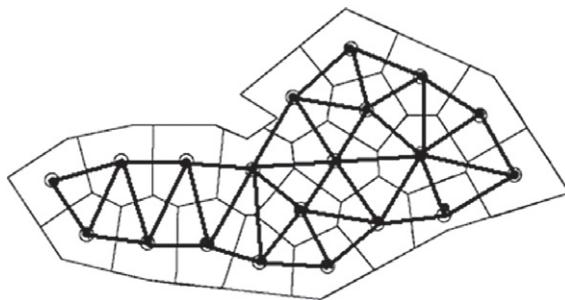
subject to the constraints:

$$\underline{\xi}_i \neq \underline{\xi}_j, \quad i \neq j, \quad i, j = 1, 2, \dots, N, \quad (22.6)$$

where \underline{B}_j is the center of mass of the Voronoi cell Ω_j , $j = 1, 2, \dots, N$. Moreover we require:

$$F(\underline{\xi}_1^*, \underline{\xi}_2^*, \dots, \underline{\xi}_N^*) = 0. \quad (22.7)$$

That is the Voronoi centers and the centers of mass of the Voronoi cells coincide. Note that in general \underline{B}_j depends on $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$ and that the function $F(\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N)$ is in general a non linear function of $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N$. The solution of the problem expressed in Eqs. (22.5)–(22.7) after having specified the communications between the sensors is the optimal deployment, represented in Figures 22.13b and 22.14. When the condition of Eq. (22.7) cannot be satisfied, we may accept the available solution of Eqs. (22.5) and (22.6) as location of the Voronoi centers corresponding to the optimal deployment. Note that in general the solution of problem expressed in Eqs. (22.5)–(22.7) is not unique and it can be solved by the application of the steepest descent concept, revised in a network centric frame as shown conceptually in Section 2.22.5.7 [94]. This method can be used to solve the problem of Eq. (22.5) with an iterative procedure, that beginning from an initial guess, updates at every iteration the current approximation of the solution with a step in the direction of the gradient of the

**FIGURE 22.14**

Graph associated to the optimal Voronoi partition of Ω shown in Figure 22.13b.

function $F(\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N)$. Moreover the steepest descent method must be adapted to the presence of the constraints of Eq. (22.6), of the condition of Eq. (22.7) and to the requirement that its implementation must lead to a network centric solution of the deployment problem. For sake of brevity, how to impose Eq. (22.6) will not be discussed here, however a treatment of constraints in the continuous analog of the steepest descent algorithm can be found in [99]. Note also that the solutions of Eqs. (22.5) and (22.6) that are of interest are usually interior points of the constraints (6). That is the constraint issue usually is not relevant in the solution of Eqs. (22.5) and (22.6). Similarly we will not pay attention to condition of Eq. (22.7). In fact with respect to Eq. (22.7), we will simply verify if the solution of the optimization problem determined by the steepest descent method satisfies Eq. (22.7). Let us concentrate our attention on the issue of building a network centric implementation of the continuous analog of the steepest descent method to solve Eq. (22.5). Assume that the sensor s_i knows only the position of its neighbor sensors, that is of the sensors that belong to a disk with center $\underline{\xi}_j$ and radius $r > 0$, $i = 1, 2, \dots, N$. Later we will show how to choose r . The solution of the optimization problem of Eq. (22.5) is found approximating the solution of the system of differential equations:

$$\dot{\underline{\xi}}_i = -\nabla_{\underline{\xi}_i} F(\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N), \quad \lambda_1 > 0, \quad i = 1, 2, \dots, N, \quad (22.8)$$

where λ_1 denotes a real parameter, with the solution of the “network centric” system differential equations:

$$\dot{\underline{\xi}}_i = -\nabla_{\underline{\xi}_i} F_i(\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_N), \quad \lambda_1 > 0, \quad i = 1, 2, \dots, N \quad (22.9)$$

with

$$F_i = \sum_{j \in L_i} \left\| \tilde{B}_{i,j} - \underline{\xi}_j \right\|^2, \quad i = 1, 2, \dots, N, \quad (22.10)$$

where

$$L_i = \left\{ \underline{\xi}_j, \quad j = 1, 2, \dots, N : \underline{\xi}_j \text{ is neighbor to } \underline{\xi}_i \right\}, \quad i = 1, 2, \dots, N \quad (22.11)$$

and $\tilde{B}_{i,j}$ being the center of mass of the Voronoi cell $\tilde{\Omega}_{i,j}$ obtained computing the Voronoi partition of Ω associated to the Voronoi centers $\underline{\xi}_j$, $j \in L_i$, $i = 1, 2, \dots, N$. Assume that $r > 0$ is large enough to

guarantee that $\underline{\xi}_j$ is neighbor of $\underline{\xi}_i$ when the distance between Ω_j and Ω_i is zero, $i, j = 1, 2, \dots, N$.

Note that with this assumption we have $\dot{\underline{B}}_{i,i} = \underline{B}_i$, $i = 1, 2, \dots, N$. In Eqs. (22.8) and (22.9) the dot denotes the differentiation with respect to λ_1 . We observe that Eq. (22.8) is known as the steepest descent differential equation. The continuous analog of the steepest descent method consists in obtaining the solution of the optimization problem of Eq. (22.5) computing the asymptotic value as λ_1 goes to infinity of a solution of Eq. (22.8) equipped when $\lambda_1 = 0$ with a suitable initial condition. This asymptotic value hopefully is a point that solves Eq. (22.5) and satisfies Eqs. (22.6) and (22.7).

Note that the function F_i depends only on $\underline{\xi}_j$, $j \in L_i$, that is can be computed in the location $\underline{\xi}_i$ using only information available in $\underline{\xi}_j$, $i = 1, 2, \dots, N$. Approximating the gradient of F with the appropriate pieces of the gradients of the function F_i , $i = 1, 2, \dots, N$, and using Eq. (22.9) instead than Eq. (22.8) we can find an approximation of the solution of Eq. (22.5) integrating numerically the initial value problem for Eq. (22.9). Note that the solution for the i th differential equation of Eq. (22.9) is computed in the location $\underline{\xi}_i$, $i = 1, 2, \dots, N$. This approximation of the solution of Eq. (22.5) is obtained using only local information so that it is “network centric.” When the asymptotic value as λ_1 goes to infinity, the solution of Eq. (22.9) coincides with an asymptotic value of a solution of Eq. (22.8), solving numerically Eq. (22.9), we can obtain in a network centric manner a solution of Eq. (22.5). The choices of the optimal Voronoi partition and of the steepest descent method to determine it, are only one of the many other legitimate choices. In Figure 22.13 the polygonal region shown represents Ω , for $N = 20$ and for $i = 1, 2, \dots, 20$, denoting with the full circle the position of the center of mass \underline{B}_i of the subset Ω_i and with the empty circle the position of the sensors $\underline{\xi}_i$. The Figure 22.13a shows the Voronoi partition of the domain Ω , associated to 20 Voronoi centers $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_{20}$ and the corresponding centers of mass $\underline{B}_1, \underline{B}_2, \dots, \underline{B}_{20}$ of the associated Voronoi cells $\Omega_1, \Omega_2, \dots, \Omega_{20}$. Note that in Figure 22.13a we have $\underline{\xi}_i \neq \underline{B}_i$, $i = 1, 2, \dots, 20$. The Figure 22.13b shows an optimal Voronoi partition. Note that in Figure 22.13b we have $\underline{\xi}_i = \underline{B}_i$, $i = 1, 2, \dots, 20$. The Voronoi partition shown in Figure 22.13b satisfies Eqs. (22.5)–(22.7). The centers of Figure 22.13b have been obtained integrating numerically, using the explicit Euler method in Eq. (22.9), equipped with the initial condition given by the centers shown in Figure 22.13a. In Figure 22.14 we show the graph associated to the optimal Voronoi partition of Ω shown in Figure 22.13b. The graph is obtained joining with branches the Voronoi centers that are (distinct) neighbors. In Figure 22.13 and Figure 22.14 we have chosen $r = k (\text{area}(\Omega))^{1/2}$, where k is a parameter that can be changed during the optimization procedure used to solve Eqs. (22.5)–(22.7).

Remind that we have assumed that the graph G associated to the optimal deployment is connected (see Figures 22.14 and 22.13b). Moreover we remind that, since there is not a fusion center, each node of the graph G does not know the positions of all the remaining nodes of the graph, in fact it knows only the positions of its neighbor nodes. Let L be the Laplacian matrix associated to G [100]. The matrix L is a symmetric positive semi-definite $N \times N$ matrix. Let $\underline{x}(\lambda_2) = (x_1(\lambda_2), x_2(\lambda_2), \dots, x_N(\lambda_2))^T$, $\lambda_2 > 0$, be a real N dimensional vector depending on the real parameter λ_2 . The superscript $(\cdot)^T$ means transposed. We consider the system of ordinary differential equations:

$$\dot{\underline{x}}(\lambda_2) = -L\underline{x}(\lambda_2), \quad \lambda_2 > 0 \quad (22.12)$$

equipped with the initial conditions:

$$\underline{x}(0) = \underline{\alpha}, \quad (22.13)$$

where $L\underline{x}$ denotes the usual matrix vector multiplication, $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$ is a known initial condition and the dot denotes differentiation with respect to λ_2 . Since G is connected we have:

$$\lim_{\lambda_2 \rightarrow +\infty} x_i(\lambda_2) = \frac{1}{N} \sum_{j=1}^N \alpha_j, \quad i = 1, 2, \dots, N, \quad (22.14)$$

where $\underline{x}(\lambda_2)$, $\lambda_2 > 0$, is the solution of Eqs. (22.12) and (22.13). This result follows easily from the spectral properties of L [100]. Note also that the right hand side of Eq. (22.14) is the “average” of the initial condition $\underline{\alpha}$. Note that Eq. (22.12) can be interpreted as the “heat equation” on the graph G , that the problem of Eqs. (22.12) and (22.13) can be seen as an initial value problem for the heat equation on G and that Eq. (22.14) can be understood as the approach to an asymptotic equilibrium “temperature” in an “heat transfer” problem. We assume that during the monitoring phase the sensor measures a physical quantity, such as, for example, the temperature, of the region Ω_i where it is located. The sensors are identical, the measures made by the sensors are synchronous, repeated periodically in time and of course they are noisy. Moreover they are assumed to be independent. A first set of measures is taken by the sensors at time $t = t_0$ and is collected in the vector $\underline{\beta}_0 = (\beta_{0,1}, \beta_{0,2}, \dots, \beta_{0,N})^T$, where $\beta_{0,i}$ is the measure done by the sensor s_i . The set of measure $\underline{\beta}_0$ will be used to obtain the “consensus” value $\bar{\beta}_0$ of the quantity monitored in Ω at time $t = t_0$. We choose:

$$\bar{\beta}_0 = \frac{1}{N} \sum_{i=1}^N \beta_{0,i}. \quad (22.15)$$

Remind that the sensor s_i located in ξ_i knows $\beta_{0,i}$ and communicates with the sensors s_j located in ξ_j , $j \in L_i$, $i = 1, 2, \dots, N$. In order to provide to the sensor s_i , the consensus value $\bar{\beta}_0$ in a network centric manner we proceed as follow: we choose $\underline{\alpha} = \underline{\beta}_0$ in Eq. (22.13) and we integrate numerically the initial value problem of Eqs. (22.12) and (22.13) using the explicit Euler method to obtain a numerical approximation of $\lim_{\lambda_2 \rightarrow +\infty} \underline{x}(\lambda_2)$. Note that the i th differential equation of Eq. (22.12) is integrated in the location ξ_i , and that using the explicit Euler method this can be done using only information available in the location ξ_i . Note that the analytic solution of Eqs. (22.12) and (22.13) is not “network centric” but its approximation with the explicit Euler method is “network centric.” In the former case to achieve the solution each node should know the whole graph, i.e., all the nodes. The i th node is not able to achieve the solution exploiting only the information in its posses: in this sense the solution is not “network centric.” Otherwise, exploiting the Euler approximation of the exponential of a matrix, the whole knowledge of the graph is not necessary: in this sense a “network centric” solution is achieved.

Once obtained $\bar{\beta}_0$ we consider the following vector:

$$\underline{\gamma}_0 = \frac{N}{N-1} \left((\beta_{0,1} - \bar{\beta}_0)^2, (\beta_{0,2} - \bar{\beta}_0)^2, \dots, (\beta_{0,N} - \bar{\beta}_0)^2 \right)^T. \quad (22.16)$$

Then we choose $\underline{\alpha} = \underline{\gamma}_0$ in Eq. (22.13) and we integrate Eqs. (22.12) and (22.13) with the explicit Euler method as done above. In this way we obtain asymptotically a numerical approximation of $\bar{\gamma}_0$ where:

$$\bar{\gamma}_0 = \frac{N}{N-1} \sum_{i=1}^N (\beta_{0,i} - \bar{\beta}_0)^2. \quad (22.17)$$

This approximation of $\bar{\gamma}_0$ is provided to each sensor in a network centric manner. Note that $\bar{\gamma}_0$ is an estimate of the magnitude of the noise contained in the data; in fact $\bar{\gamma}_0$ is the “sample” variance of the measures $\beta_{0,i}$ made by the sensor at time $t = t_0$. The approximation of $\bar{\beta}_0$ and $\bar{\gamma}_0$ obtained integrating numerically Eqs. (22.12) and (22.13) are the consensus values. These values are “global” values (that is they depend on all the measures made by the sensor network at time $t = t_0$) and have been provided to each sensor in a network centric manner (that is using only “local” interactions between sensors).

The sensor s_i repeats periodically in time the measure of the quantity of interest and after a given time interval has as its disposal a set of measures that can be compared with the consensus values $\bar{\beta}_0$ and $\bar{\gamma}_0$ to detect (local) anomalies. Let us assume that the set of measures made by the sensor s_i is a sample taken from a set of independent identically distributed Gaussian random variables of mean μ_i and variance σ_i^2 . In these hypotheses the Student t -test and the Chi-square test [101] are the elementary statistical tools that must be used to compare μ_i and σ_i^2 (that are unknown) to $\bar{\beta}_0$ and $\bar{\gamma}_0$. The result of this comparison is the detection of local anomalies. A (statistical) significance is associated to the detected anomalies. The statistical tests used are based on the assumption that the measures come from a set of independent identically distributed Gaussian random variables. Note that the estimators $\bar{\beta}_0$ and $\bar{\gamma}_0$ can be used in more general circumstances.

2.22.6.2 Modeling and performance analysis of a network of chemical sensors with dynamic collaboration

Typically the challenge in the deployment of an operational wireless sensor network (WSN) resides in establishing the balance between its operational requirements (e.g., minimal detection threshold, the size of surveillance region, detection time, the rate of false negatives, etc.) and the available resources (e.g., energy supply, number of sensors, communication range, fixed detection threshold of individual sensors, limited budget for the cost of hardware, maintenance, etc.) [102]. The issue of resource constraints is particularly important for a network of chemical sensors, because modern chemical sensors are equipped with air-sampling units (fans), which turn on when the sensor is active. Operating a fan requires a significant amount of energy as well as a frequent replacement of some consumable items (i.e., cartridges, filters). This leads to the critical requirement in the design of a WSN to reduce the active (air-sampling) time of its individual sensors.

One attractive way to achieve the described balance between the requirements and the constraints of WSN is to exploit the idea of *dynamic sensor collaboration* (DSC) [103, 104]. The DSC implies that a sensor in the network should be invoked (or activated) only when the network will gain information by its activation [104]. For each individual sensor this information gain can be evaluated against other performance criteria of the sensor system, such as the detection delay or the detection threshold, to find an optimal solution in given circumstances. However, the DSC-based algorithms involve continuous estimation of the state of each sensor in the network and usually require extensive computer simulations [103, 104]. These simulations may become unpractical as the number of sensors in the network increases. Furthermore, the simulations can provide the numerical values for optimal network parameters only for a specific scenario.

This motivates the development of another simple and analytic approach to the problem of network analysis and design. The main idea is to phenomenologically employ the so-called bio-inspired

(epidemiology, population dynamics) or physics inspired (percolation and graph theory) models of DSC in the sensor network in order to describe the dynamics of collaboration as a single entity [105–110]. From a formal point of view, the equations of bio inspired models of DSC are the ones of the “mean-field” theory, meaning that instead of working with dynamic equations for each individual sensor we use only a small number of equations for the “averaged” sensor state (i.e., passive, active, faulty, etc.), regardless of the actual number of sensors in the system.

The analytic approach can lead to the valuable insights into the performance of the proposed sensor network system by providing simple analytical expressions to calculate the vital network parameters, such as the detection threshold, robustness, responsiveness and stability and their functional relationships.

The fluctuations in concentration C of the pollutant are modeled by the probability density function (pdf) with the mean C_0 as a parameter [111]:

$$\rho(C|C_0) = (1 - \omega)\delta(C) + \frac{\omega^2(\gamma - 1)}{C_0(\gamma - 2)} \left(1 + \frac{\omega C}{(\gamma - 2)C_0}\right)^{-\gamma}. \quad (22.18)$$

Here the value $\gamma = 26/3$ can be chosen to make it compliant with the theory of tracer dispersion in Kolmogorov turbulence [111], but it may vary with meteorological conditions. The parameter ω , which models the tracer intermittency in the turbulent flow, can be in the range $[0, 1]$, with $\omega = 1$ corresponding to the non-intermittent case. In general it also depends on the sensor position within a chemical plume, thus $0.95 < \omega < 0.98$ near the plume centroid and may drop to $0.3 < \omega < 0.5$ near the plume edge. For $\omega \neq 0$, the pdf ρ of Eq. (22.18) has a delta impulse in zero, meaning that the measured concentration in the presence of intermittency can be zero on some occasions. It can be easily shown that the pdf of Eq. (22.18) integrates to unity, so it is appropriately normalized.

Depending on the values of parameters γ , ω , Eq. (22.18) allows simulation of pollutant distributions with the different correlation structure (e.g., intermittent and strongly non-Gaussian) corresponding to the rich variety of possible regimes of turbulent mixing occurring in the ambient environment; Figure 22.15 shows two examples of the same WSN operating in two different correlation structures of the chemical tracer.

We adopt a binary model of a chemical sensor, with reading V specified as:

$$V = \begin{cases} 1, & C \geq C_*, \\ 0, & C < C_*, \end{cases} \quad (22.19)$$

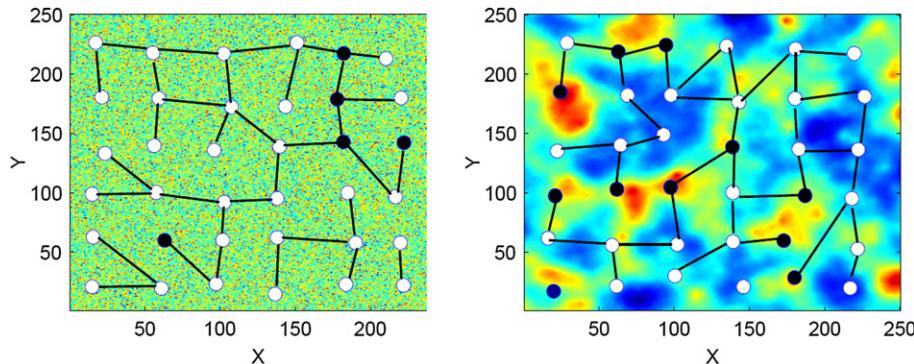
where C_* is the threshold (an internal characteristic of the sensor). It can be shown [112] that the probability of detection of an individual sensor embedded in the environmental model described by Eq. (22.18) is given by:

$$p = 1 - F(C_*|C), \quad (22.20)$$

where

$$F(C|C_0) = 1 - \omega \left[1 + \left(\frac{1}{\gamma - 2}\right) \frac{C}{C_0}\right]^{1-\gamma} \quad (22.21)$$

is the cumulative distribution function corresponding to pdf of Eq. (22.18), see [113].

**FIGURE 22.15**

Examples of WSN network operating in the tracer filed with different correlation structure [117]. (reprinted with permission.)

Suppose that N chemical sensors are uniformly distributed over the surveillance domain of area S and adopt the following network protocol for dynamic collaboration. Each sensor can be only in one of the two states: active and passive. The sensor can be activated only by a message it receives from another sensor. Once activated, the sensor remains in the active state during an interval of time τ_* ; then it “dies out” (becomes passive). While being in the active state, the sensor senses the environment and if the chemical tracer is detected, it broadcasts a (single) message. The broadcast capability of the sensor is characterized by its communication range r_* . This network with the described dynamic collaboration can be modeled using the epidemic SIS model (susceptible-infected-susceptible) [114]:

$$\begin{aligned} \frac{dN_+}{dt} &= \alpha N_+ N_- - \frac{N_+}{\tau_*}, \\ \frac{dN_-}{dt} &= -\alpha N_+ N_- + \frac{N_+}{\tau_*}, \end{aligned} \quad (22.22)$$

where N_+ , N_- denote the number of active and passive sensors, respectively. The nonlinear terms on the right hand side of Eq. (22.22) are responsible for the interaction between the sensors; parameter α is a measure of this interaction. The number of sensor is assumed constant, hence we have an additional equation: $N_+ + N_- = N$. Since the parameter alpha describes the intensity of social interaction in a community [114] we can propose that:

$$\alpha \propto \frac{m \cdot p}{N \tau_*}, \quad (22.23)$$

where m is the number of contacts made by the activated (“infected”) sensor during its infectious period τ_* (i.e., the number of sensors that received the wake-up message from an alerting sensor). In our case $m = \pi \cdot r_*^2 \cdot N/S$. Then we have:

$$\alpha = G \frac{\pi \cdot r_*^2 \cdot p}{S \tau_*}, \quad (22.24)$$

where G is a calibration constant. In order to simplify notation we will further assume that G is absorbed in the definition of r_* . Equation (22.22) combined with $N_+ + N_- = N$ can be reduced to one equation for $y = N_+$:

$$\frac{dy}{dt} = \alpha \cdot y \cdot (N - y) + \frac{y}{\tau_*} = y \cdot (b - \alpha \cdot y), \quad (22.25)$$

where $b = \alpha N - 1/\tau_*$. By simple change of variables $z = \alpha y/b$, this equation can be reduced to the standard logistic equation [115, 116]:

$$\frac{dz}{dt} = bz(1 - z). \quad (22.26)$$

The solution of the logistic equation is well-known:

$$z(t) = \frac{z_0}{(1 - z_0) \exp(-bt) + z_0}, \quad (22.27)$$

where $z_0 = z(0)$. Observe that the WSN will be able to detect the presence of a pollutant only if $b > 0$, because then $z \rightarrow 1$ as $t \rightarrow \infty$ independent of z_0 . In this case, after a certain transition interval, the WSN will reach a new steady state with:

$$\frac{N_+}{N} = 1 - \theta, \quad \frac{N_-}{N} = \theta, \quad \theta = \frac{1}{\alpha \tau_* N}. \quad (22.28)$$

From (22.27) and using the expression for b stated above, the activation time (transition interval) is given by:

$$\tau \approx \frac{1}{b} = \frac{\tau_*}{\alpha \tau_* N - 1}. \quad (22.29)$$

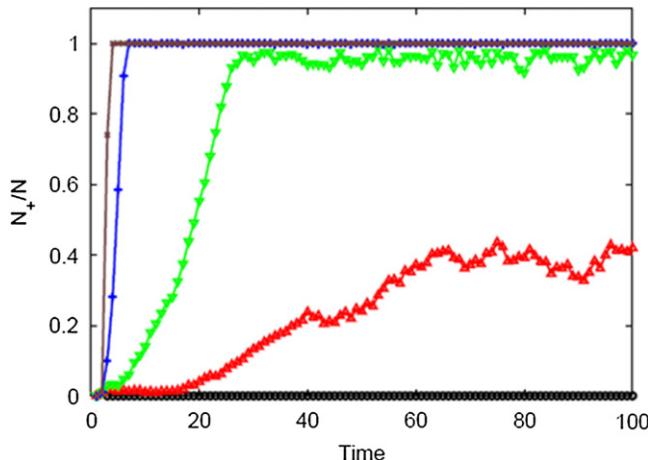
From Eq. (22.29) it follows that the key requirement for the network to be operational $b > 0$ is that $\alpha \tau_* N > 1$, that is:

$$R_0 = p \cdot N \cdot \pi \cdot r_*^2 / S > 1, \quad (22.30)$$

where R_0 is a well-known parameter in epidemiology, referred to as the *basic reproductive number* [114]. Observe that R_0 is independent of τ_* ; however, according to Eq. (22.29) the response time of the WSN is strongly dependent on τ_* .

It remains to specify q , the number of sensors that should initially be active for the described WSN with dynamic collaboration to be effective. The initial condition is simply $q \cdot p > 1$, that is on average $q > 1/p$. Eqs. (22.28)–(22.30) are important analytic results. For a given level of mean pollutant concentration C_0 and meteorological conditions (γ, ω) , these expressions provide a simple yet rigorous way to estimate how a change in network and sensor parameters (i.e., N , C_* , τ_*) will affect the network performance (i.e., N_+ , τ).

The examples of agent-based simulation of “information epidemic” in WSN, which satisfies the threshold condition of Eq. (22.30) is presented in Figure 22.16. We can observe that by change of the configuration parameters of WSN we can vary the activation time and the saturation limit of the detection system. Further development of the theoretical framework presented in this section can be found in [117–120].

**FIGURE 22.16**

Examples of Information Epidemic in WSN of chemical sensors for different values of parameters ω in Eq. (22.18). (Black, blue, green and red lines correspond to $\omega = 1, 0.95, 0.5, 0.03$, respectively.) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

2.22.6.3 Detection and localization of radioactive point sources with experimental verification

Recently there has been an increased interest in detection and localization of radioactive material [121–125]. Radioactive waste material is relatively easy to obtain with numerous accidents involving its loss or theft reported. The danger is that a terrorist group may acquire some radiological material and use it to build a dirty-bomb. The dirty bomb would consist of waste by products from nuclear reactors wrapped in conventional explosives, which upon detonation would expel deadly radioactive particles into the environment. The ability to rapidly detect and localize radioactive sources is important in order to disable and isolate the potential threat in emergency situations.

This section is concerned with radiological materials that emit gamma rays. The probability that a gamma radiation detector registers $z \in N$ counts (N being the set of natural numbers including zero) in τ seconds, from a source that emits on average μ counts per second is [126]:

$$\mathcal{P}(z; \lambda) = \frac{\lambda^z}{z!} e^{-\lambda}, \quad (22.31)$$

where $\lambda = \mu\tau$ is the mean and variance of the Poisson distribution. The measurements of radiation field are assumed to be made using a network of low-cost Geiger-Müller (GM) counters as sensors. In general, the problem of detection and localization of point sources or radioactive sources can be solved using either controllable or uncontrollable sensors. Controllable sensors can move and vary the radiation exposure time [127, 128]. In this Section we will focus on uncontrollable sensors, placed at known locations with constant and known exposure times.

Assume that $r \geq 0$ sources (r is unknown) are present in the area of interest. Furthermore, the assumption is that the area is flat without obstacles (“open field”). Each source $i = 1, 2, \dots, r$ is parameterized by its 2D location (x_i, y_i) and its equivalent strength α_i (a single parameter which takes into account the activity of the source, the value of gamma energy per integration and scaling factors involved, see [129]). Thus the parameter vector of source i is $\vartheta_i = [x_i \ y_i \ \alpha_i]^T$, while the total parameter vector is a stacked vector: $\vartheta = [\vartheta_1^T \ \dots \ \vartheta_r^T]^T$. Suppose a network of GM counters is deployed in the field of interest. Let GM counter $j = 1, \dots, m$, located at $(\xi_j \ \zeta_j)$, reports its count z_j every τ seconds. Assuming that each GM counter has a uniform directional response and that attenuation of gamma radiation due to air can be neglected, the joint density of the measurement vector $\mathbf{z} = [z_1 \ \dots \ z_m]^T$, conditional on the parameter vector ϑ and the knowledge that r sources are present, can be modeled as [129]:

$$l(\mathbf{z}|\vartheta) = \prod_{j=1}^m P(z_j; \lambda_j(\vartheta)). \quad (22.32)$$

Here $\lambda_j(\vartheta)$ is the mean radiation count at sensor j :

$$\lambda_j(\vartheta) = \lambda_b + \sum_{i=1}^r \frac{\alpha_i}{d_{ij}^2} \quad (22.33)$$

with

$$d_{ij}^2 = \sqrt{(x_i - \xi_j)^2 + (y_i - \zeta_j)^2} \quad (22.34)$$

being the distance between the source i and sensor j , and λ_b the average count due to the background radiation (assumed known). The problem for the network of GM counters is to estimate the number of sources r and the parameter vector for each source ϑ_i , $i = 1, \dots, r$. In this section we will present the experimental results obtained using real data and a Bayesian estimation algorithm combined with the minimum description length (MDL) for source number estimation.

A radiological field trial was conducted on a large, flat, and open area without any obstacles at the Puckapunyal airfield site in Victoria, Australia. The measurements were collected using the DSTOs³ Low Cost Advanced Airborne Radiological Survey (LCAARS) survey system which consists of an AN/PDR-77 radiation survey meter equipped with an RS232 interface module, a gamma probe and software written in Visual Basic running on a laptop computer. The gamma probe contains two GM tubes to cover both low and high ranges of dose rates. It was capable of measuring gamma radiation dose rates from background to 9.99 Sv/h⁴ without saturating [130] with a fairly flat response [131]. Three radiation sources were used in the field trial: source 1 was a cesium source (¹³⁷Cs) with $\vartheta_1 = [11 \text{ m } 10 \text{ m } 9105 \mu\text{Sv}/\text{h}]^T$, source 2 was also a cesium source with $\vartheta_2 = [3 \text{ m } 50 \text{ m } 1868 \mu\text{Sv}/\text{h}]^T$, and source 3 was a cobalt

³The Defense Science and Technology Organisation (DSTO) is part of Australia’s Department of Defense. DSTO is the Australian Government’s lead agency charged with applying science and technology to protect and defend Australia and its national interests (<http://www.dsto.defence.gov.au/>).

⁴Sievert/Plank constant: the Sievert (symbol: Sv) is the International System of Units derived unit of dose equivalent radiation. It attempts to quantitatively evaluate the biological effects of ionizing radiation as opposed to just the absorbed dose of radiation energy, which is measured in gray.

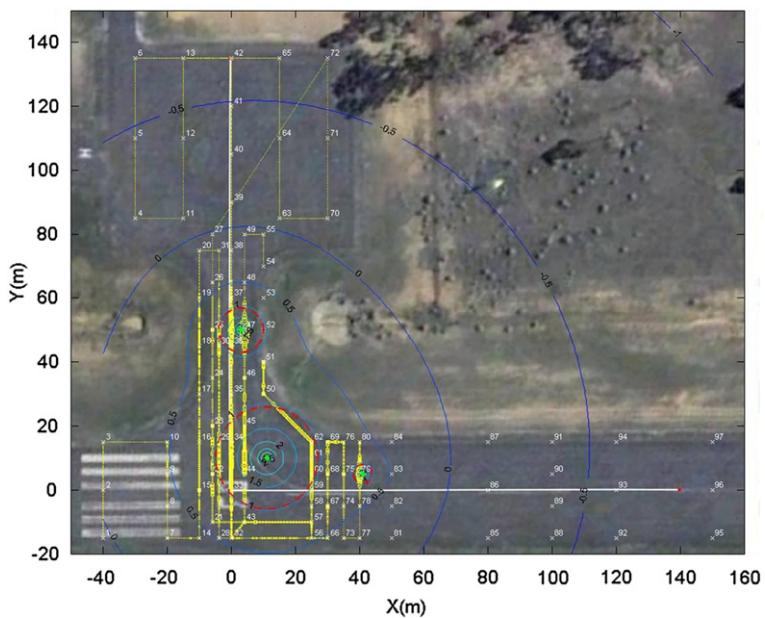


FIGURE 22.17

Aerial image of the experimental site with the local coordinate system. (green points indicate the locations of three sources; red circles indicate the zones with dangerous levels of radiation.) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

source (^{60}Co) with $\vartheta_3 = [41\text{m } 5\text{m } 467 \mu\text{Sv/h}]^T$. The aerial image of the experimental site with the location of sources and the local Cartesian coordinate system is shown in Figure 22.17. Four data sets were collected during the field trials in the presence of r sources, with respectively $r = 0, 1, 2, 3$ [132]. Data sets with $r > 0$ sources contains 50 count measurements in each measurement point.

Estimation of parameter vector ϑ , under the assumption that r is known, was carried out using the Bayesian importance sampling technique known as the progressive correction [125, 133]. This technique assumes that prior distribution of ϑ , denoted $p_0(\vartheta)$, is available. The information contained in the measurement vector \mathbf{z} is combined with the prior to give the posterior pdf: $p(\vartheta|\mathbf{z}) \propto l(\mathbf{z}|\vartheta) \cdot p_0(\vartheta)$. The minimum mean squared error estimate of ϑ is then the posterior expectation:

$$\hat{\vartheta} = E(\vartheta|\mathbf{z}) = \int \vartheta \cdot p(\vartheta|\mathbf{z}) \cdot d\vartheta. \quad (22.35)$$

The problem is that the posterior pdf and hence the posterior expectation of Eq. (22.35) cannot be found analytically for the described problem. Instead, an approximation of Eq. (22.35) is computed via the importance sampling: it involves drawing N_p samples of the parameter vector from an importance density and approximating the integral by a weighted sum of the samples. This is carried out in a few stages, each stage drawing samples from a “target distribution” which is gradually approaching the true

posterior. The “target distribution” at stage $s = 1, \dots, S$ is constructed as:

$$p_s(\vartheta | \mathbf{z}) \propto l(\mathbf{z} | \vartheta)^{G_s} \cdot p_0(\vartheta), \quad (22.36)$$

where $G_s = \sum_{l=1}^s \gamma_l$ with $\gamma \in [0, 1]$ and $G_S = \sum_{l=1}^S \gamma_l = 1$. An adaptive scheme for the computation of S and factors $\gamma_1, \gamma_2, \dots, \gamma_S$ is given in [125, 133]. Assume that a random sample $\{\vartheta_{s-1}^n\}_{n=1}^{N_p}$ from $p_{s-1}(\vartheta | \mathbf{z})$ is available and one wants to generate the samples or particles from $p_s(\vartheta | \mathbf{z})$. The progressive correction algorithm steps are then as follows [125]:

1. compute $G_s = G_{s-1} + \gamma_s$;
2. compute not-normalized weight of each sample as: $w_s^n = l(\mathbf{z} | \vartheta)^{\gamma_s}$, for $n = 1, \dots, N_p$;
3. normalize weights;
4. perform re-sampling of particles [134];
5. carry out Markov chain Monte Carlo (MCMC) move step for each particle [134].

The procedure is repeated for every stage $s < S$ until $G_s < 1$. The initial set of particles is drawn from the prior density $p_0(\vartheta)$. The final estimate in Eq. (22.35) is approximated as

$$\hat{\vartheta} = \frac{1}{N_p} \sum_{n=1}^{N_p} \vartheta_s^n. \quad (22.37)$$

The number of sources was estimated using the MDL algorithm [59], which will choose $r \in \{1, 2, \dots, r_{\max}\}$ that will maximize the following quantity:

$$\beta_r = \log l(\mathbf{z} | \hat{\vartheta}(r)) - \frac{1}{2} \log |\mathbf{J}(\hat{\vartheta}(r))|, \quad (22.38)$$

where $\hat{\vartheta}(r)$ is the estimate obtained under the assumption that r sources are present and

$$\mathbf{J}(\vartheta) = -E \left\{ \nabla_{\vartheta} \nabla_{\vartheta}^T \log l(\mathbf{z} | \vartheta) \right\} \quad (22.39)$$

is the Fisher Information Matrix. It can be shown that

$$\mathbf{J}(\vartheta) = \sum_{j=1}^m \frac{[\nabla_{\vartheta} \lambda_j(\vartheta)] [\nabla_{\vartheta} \lambda_j(\vartheta)]^T}{\lambda_j(\vartheta)} \quad (22.40)$$

with

$$\begin{aligned} \frac{\partial \lambda_j(\vartheta)}{\partial x_i} &= 2\alpha_i(\xi_j - x_i)/d_{ij}^4, \\ \frac{\partial \lambda_j(\vartheta)}{\partial y_i} &= 2\alpha_i(\zeta_j - y_i)/d_{ij}^4, \\ \frac{\partial \lambda_j(\vartheta)}{\partial \alpha_i} &= d_{ij}^{-2}. \end{aligned} \quad (22.41)$$

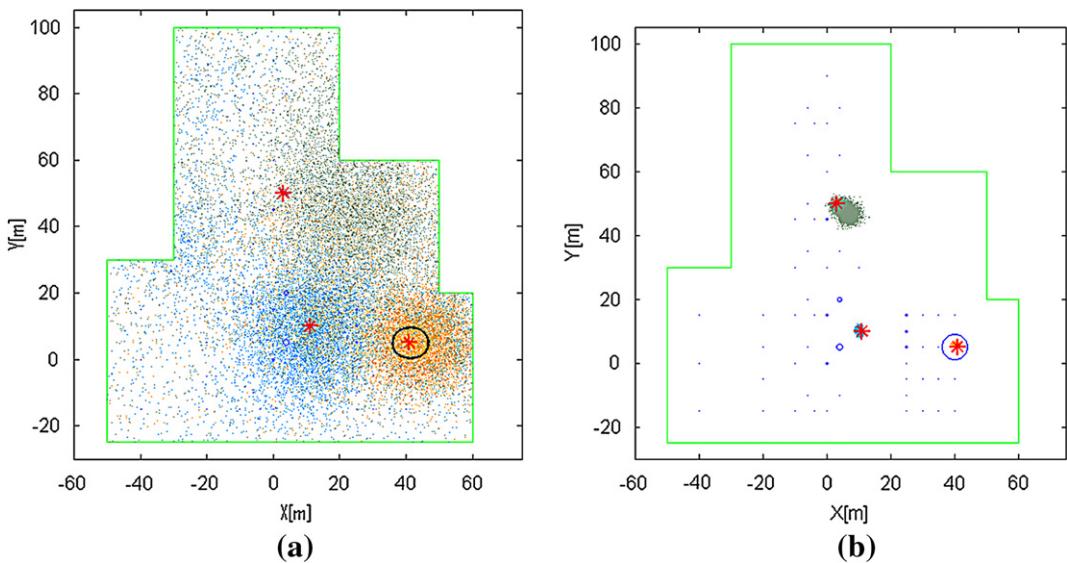


FIGURE 22.18

The output of progressive correction algorithm after (a) $s = 2$ and (b) $s = 11$ stages: data set 3 with $r = 3$ sources present (indicated by red stars). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

The inverse of the FIM gives us the CRLB, which represents the theoretical lower bound for estimation error covariance [135]. Figure 22.18 shows the output of the progressive correction algorithm for data set 3 (with three sources present) after (a) $s = 2$ and (b) $s = 11$ stages of processing. The red stars indicate the locations of three sources. The green line shows the initial polygon A for the location of sources. The prior density for sampling the initial set of particles for source $i = 1, \dots, r$ is:

$$p_0(\vartheta_i) = U_A(x_i, y_i) \cdot \Gamma_{\kappa, v}(\alpha_i), \quad (22.42)$$

where $U_A(x_i, y_i)$ stands for uniform distribution over the polygon A and $\Gamma_{\kappa, v}(\alpha_i)$ is the gamma distribution with parameters $\kappa = 1.5$ and $v = 8000$. From Figure 22.18 we observe how the progressive correction algorithm localizes the three sources fairly accurately.

As we mentioned earlier, 50 count measurements have been collected by each sensor. This allows us to find the root mean square (rms) estimation error using each snapshot of measurement data from all sensors. Table 22.2 shows the resulting rms errors versus the theoretical CRLB.

The theoretical CRLB was computed using the idealized measurement model as stated by Eqs. (22.32)–(22.34). Considering that this measurement model was very crude with a number of factors neglected (e.g., uniform directional response, neglected air attenuation, perfect knowledge of sensor locations, known and constant average background radiation, etc.), the agreement between the theoretical bound and the RMS estimation errors in Table 22.2 is remarkable. The experimental results in this table effectively verify the measurement model as well as the estimation algorithm.

Table 22.2 RMS Error of Progressive Correction Algorithm Versus the Theoretical Bound for Data Set 3 ($r = 3$)

	$\sqrt{\text{CRLB}}$	RMS Error
x_1 (m)	0.35	0.62
y_1 (m)	0.46	0.57
α_1 ($\mu\text{Sv/h}$)	85	160
x_2 (m)	1.12	2.6
y_2 (m)	0.57	1.46
α_2 ($\mu\text{Sv/h}$)	63	195
x_3 (m)	0.20	1.27
y_3 (m)	2.66	1.22
α_3 ($\mu\text{Sv/h}$)	39	95

Table 22.3 Estimation of the Number of Sources Using the Progressive Correction in the MDL Algorithm

Test	True			Est. \hat{r}		
	Data Set	r	0	1	2	3
Backgr.	0	50	0	0	0	0
Test 1	1	0	48	2	0	0
Test 2	2	0	0	49	1	
Test 3	3	0	0	6	44	

Results for estimation of r are shown in Table 22.3. The table lists the number of runs (out of 50) that resulted in $r \in \{0, 1, 2, 3\}$. It can be observed that the number of sources is estimated correctly in the majority of cases.

More results of experimental data processing can be found in [131, 132]. In a recent study [136] it was found that by using all 50 snapshots of measurement data for estimation by progressive correction, results in a posterior pdf which is very narrow but does not include the true source positions. This indicates that the measurement model is not perfect, which is not surprising considering that it is based on many approximations. In situations where the measurement likelihood is not exact, it is necessary to introduce a degree of caution to make the estimation more robust. In the framework of progressive correction this can be achieved by $G_S = \sum_{l=1}^S \gamma_l < 1$. In this way the measurement likelihood is effectively approximated by a fuzzy membership function which has a theoretical justification in random set theory [137, Chapter 7].

If one wants to relax the assumption that radioactive sources are point sources, the problem becomes the one of radiation field estimation. This is an inverse problem, difficult to solve in general. By modeling the radiation field by a Gaussian mixture, however, the problem becomes tractable and some recent results are reported in [138].

2.22.7 Heterogeneous multi-sensor network management

Multi-sensor management concerns with the control of environment perception activities by managing or coordinating the usage of multiple heterogeneous sensor resources. Multi-sensor systems are becoming increasingly important in a variety of military and civilian applications. Since a single sensor generally can only perceive limited partial information about the environment, multiple similar and/or dissimilar sensors are required to provide sufficient local pictures with different focus and from different viewpoints in an integrated manner. As viewed, information from heterogeneous sensors can be combined using data fusion algorithms to obtain synergistic observation effects. Thus the benefit of multi-sensors system are to broaden perception and enhance awareness of the state of the world compared to what could be acquired by a single sensor system. The increased sophistication of sensor assets along with the large amounts of data to be processed has pushed the information acquisition problem far beyond what can be handled by human operator. This motivates the emerging interest in research into automatic and semi-automatic management of sensor resources for improving overall perception performance beyond basic fusion of data.

Multi-sensor management is formally described as a system or process that seeks to manage or coordinate the usage of a suite of sensors or measurement devices in a dynamic, uncertain environment, to improve the performance of data fusion and ultimately that of perception.

The basic objective of sensor management is to select the right sensors to do the right service on the right object at the right time. Sensor management, aiming at improving data fusion performance by controlling sensor behavior, plays the role of level 4 functions in JDL model presented in Section 2.22.3. Mainly the same considerations made for homogeneous sensor networks are still valid: the criteria followed to manage the network remains the same, however there is an increasing of complexity due to the diversity of the sensors. In the following Sections the problems related to multi-sensor management are divided into three main categories i.e., sensor deployment, sensor behavior assignment, and sensor coordination.

2.22.7.1 Sensor deployment

Sensor deployment is a critical issue for intelligence collection in an uncertain dynamic environment. It concerns with making decisions about when, where, and how many sensing resources need to be deployed in reaction to the state of the environment and its changes.

Sensor placement needs special attention in sensor deployment. It consists of positioning multiple sensors simultaneously in optimal or near optimal locations to support surveillance tasks when necessary. Typically it is desired to locate sensors within a particular region determined by tactical situations to optimize a certain criterion usually expressed in terms of global detection probability, quality of tracks, etc. This problem can be formulated as one of constrained optimization of a set of parameters. It is subject to constraints due to the following factors:

- sensors are usually restricted to specified regions due to tactical considerations;
- critical restrictions may be imposed on relative positions of adjacent sensors to enable their mutual communication when sensors are arranged as distributed assets in a decentralized network (e.g., net-centric approach);

- the amount of sensing resources that can be positioned in a given period is limited due to logistical restrictions.

In simple cases, decisions on sensor placement are to be made with respect to a well-prescribed and stationary environment. An example of a stationary problem is the placing of radars to minimize the terrain screening effect in detection of an aircraft approaching a fixed site. Another example is the arrangement of a network of intelligence gathering assets in a specified region to target another well-defined area. In the above scenarios, mathematical or physical models such as terrain models, propagation models, etc. are commonly available and they are used as the basis for evaluation of sensor placement decisions. Paper [139] presents a study for finding a solution to the placement of territorial resources for multi-purpose telecommunication services considering also the restrictions imposed by the orography of the territory itself. To solve this problem genetic algorithms⁵ are used to identify sites to place the resources for the optimal coverage of a given area. The used algorithm has demonstrated to be able to find optimal solutions in a variety of considered situations.

More challenging are those situations in which the environment is dynamic and sensors must repeatedly be repositioned to be able to refine and update the state estimation of moving targets in real time. Typical situations where reactive sensor placement is required are, for instance, submarine tracking by means of passive sonobuoys in an anti-submarine warfare scenario; locating moving transmitters using ESM (Electronic Support Measures) receivers; tracking of tanks on land by dropping passive acoustic sensors.

2.22.7.2 Sensor behavior assignment

The basic purpose of sensor management is to adapt sensor behavior to dynamic environments. By sensor behavior assignment is meant efficient determination and planning of sensor functions and usage according to changing situation awareness or mission requirements. Two crucial points are involved. Firstly the decisions about the set of observation tasks (referred to as system-level tasks) that the sensor system is supposed to accomplish currently or in the near future, on the basis of the current/predicted situation as well as the given mission goal. Secondly the planning and scheduling of actions of the deployed sensors to best accomplish the proposed observation tasks and their objectives.

Owing to limited sensing resources, it is prevalent in real applications that available sensors are not able to serve all desired tasks and achieve all their associated objectives simultaneously. Therefore a reasonable compromise between conflicting demands is sought. Intuitively, more urgent or important tasks should be given higher priority in their competition for resources. Thus a scheme is required to prioritize observation tasks. Information about task priority can be very useful in scheduling of sensor actions and for negotiation between sensors in a decentralized paradigm.

To focus on this class of problems, let us consider a scenario including a number of targets as well as multiple sensors, which are capable of focusing on different objects with different modes for target tracking and/or classification. The first step for the sensor management system should be to utilize evidences gathered to decide objects of interest and to prioritize which objects to look at in the time

⁵Genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. GA belongs to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover [140].

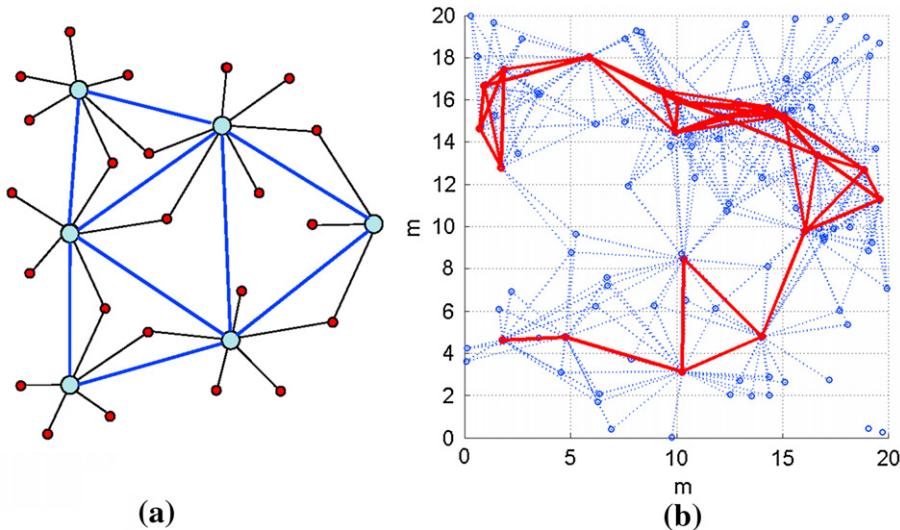
following. Subsequently, in the second step, different sensors together with their modes are allocated across the interesting objects to achieve best situation awareness. In fact, owing to the constraints on sensors and computational resources, it is in general not possible to measure all targets of interest with all sensors in a single time interval. Also, improvement of the accuracy on one object may lead to degradation of performance on another object. What is required is a suitable compromise among different targets.

2.22.7.3 Sensor coordination in a decentralized sensor network

As stated in the previous Sections, there are two general ways to integrate a set of sensors into a sensor network. One is the centralized paradigm, where all actions of all sensors are decided by a central mechanism. The other alternative is to treat sensors in the network as distributed intelligent agents with some degree of autonomy. In such a decentralized architecture, bi-directional communication between sensors is enabled, so that communication bottlenecks possibly existing in a centralized network can be avoided. A major research objective of decentralized sensor management is to establish cooperative behavior between sensors with no or little external supervision. In a decentralized sensor network scenario a local view perceived from a sensor can be shared by some members of the sensor community. Intuitively, a local picture from one sensor can be used to direct the attention of other sensors or transfer tasks such as target tracking from one sensor to another. An interesting question is how participating sensors can autonomously coordinate their movements and sensing actions, on grounds of shared information, to develop an optimal global awareness of the environment with parsimonious consumption of time and resources.

As for homogeneous sensor network, the CSIP approach can be exploited [141, 142]: the network consists of different kinds of sensors, randomly distributed inside the surveillance area and if the number of sensors is high, the performance of the surveillance system can be considered independent of the location of the targets. Each sensor has a different functioning level. A first level sensor, with small sensing and communication capabilities may provide only detection information; a second level sensor may provide detection and localization information, with medium sensing and communication capabilities. Finally a third level sensor may provide tracking information and may be able to perform target recognition and classification. Usually the number of low level sensors exceeds the number of higher level sensors and only close sensors exchange data.

In [143] the network consists of two types of sensors: simple and complex as represented in Figure 22.19a. The simple ones have only the capability of sensing their coverage area with a reduced computation capabilities and they transmit data to complex sensors. The information they provide may be encoded, for example, by a “1” if sensor detects something crossing its coverage area and by a “0” otherwise. Complex sensors, instead, have computation capabilities; they are able to locate the target by applying sophisticated algorithms (e.g., in [143] the maximum likelihood estimation algorithm is applied). The topology simulated in [143], constituted by 80 simple sensors and 20 complex sensors, is represented in Figure 22.19b: the sensors are indicated by circles; the complex sensors are connected by the solid lines, simple and complex sensor by dashed lines. Figure 22.20 shows the number of active sensors during the target tracking: the theoretical value and the simulated value are compared. It is evident that in a self-organizing configuration the number of active sensors is optimized with the consequent advantage of saving of power.

**FIGURE 22.19**

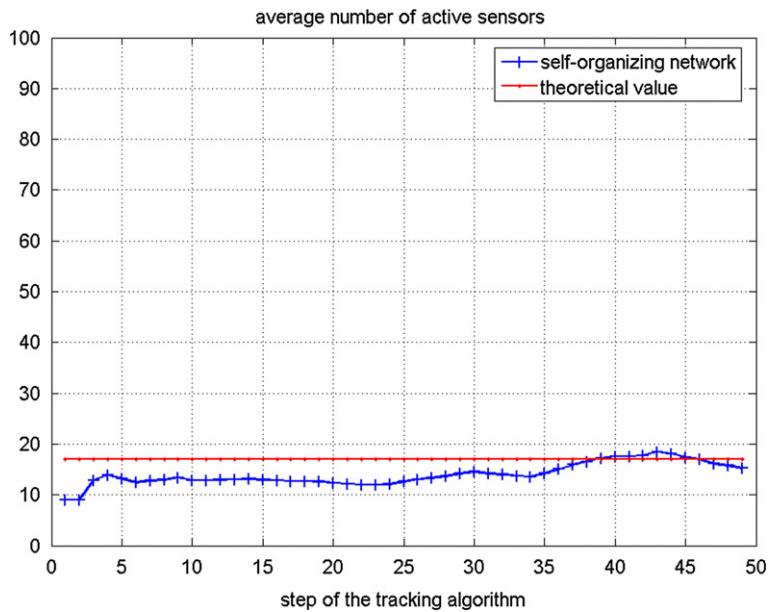
(a) Network architecture scheme; (b) deployment of simple and complex nodes simulated in [143]. (Reprinted with permission.)

An adaptive self-configuring system consists of a collection of independent randomly located sensors that, carrying ahead local interactions, estimate the position of the target without a centralized control unit that coordinates their communication. It is fault tolerant and adapts to changing conditions. Furthermore, it is able to self-configuring, i.e., there is not an external entity that configures the network. Finally, the task is performed efficiently, i.e., it guarantees both a reasonably long network life and good target tracking performances. From local interactions, sensors form an efficient system that follows the target, i.e., local communication leads to a self-organizing network that exploits the features of the theories of random graphs and of self-organizing systems. The most natural way to approach random network topology is by means of the theory of random graphs [144, 145]. The theory of random graphs allows, for instance, to compute an upper bound to the estimated number of active sensors at each time step.

2.22.7.4 A mathematical issue for multi-sensor networks

When the fusion of heterogeneous signals is performed, there is a formal problem to solve. The signal received by the different sensors may be statistically dependent because of the complex intermodal interactions; usually the statistical dependence is either ignored or not adequately considered. Usually the multiple hypotheses testing theory is based on the statistical independence of the received signals, in our case this condition is not maintained, therefore techniques as the “copula probability theory” may be useful.

In probability theory and statistics, a copula can be used to describe the dependence between random variables [146]. The cumulative distribution function of a random vector can be written in terms of

**FIGURE 22.20**

Average number of active sensors as function of the step number of the algorithm; simulation and theoretical values. (From [143], reprinted with permission.)

marginal distribution functions and a copula. The marginal distribution functions describe the marginal distribution of each component of the random vector and the copula describes the dependence structure between the components. Copulas are popular in statistical applications as they allow one to easily model and estimate the distribution of random vectors by estimating marginal distributions and copula separately. The Sklar's theorem ensures that the joint cumulative distribution function (cdf) $F_Z(z_1, z_2, \dots, z_N)$ of random variables Z_1, Z_2, \dots, Z_N are joined by a copula function $C(\cdot)$ to the respective marginal distributions $F_{Z_1}(z_1), F_{Z_2}(z_2), \dots, F_{Z_N}(z_N)$ as [147]:

$$F_Z(z_1, z_2, \dots, z_N) = C(F_{Z_1}(z_1), F_{Z_2}(z_2), \dots, F_{Z_N}(z_N)). \quad (22.43)$$

Further, if the marginals are continuous, $C(\cdot)$ is unique. By the differentiation of the joint cdf, the joint pdf is obtained:

$$f(z_1, z_2, \dots, z_N) = \left(\prod_{i=1}^N f(z_i) \right) c(F_{Z_1}(z_1), F_{Z_2}(z_2), \dots, F_{Z_N}(z_N)). \quad (22.44)$$

The copula density $c(\cdot)$, function of the N marginals from the N sensors, represents a correction term of the independent product of densities of Eq. (22.44).

Processing heterogeneous data set is not straightforward as they may not be commensurate. In addition, the signals may also exhibit statistical dependence due to overlapping fields of view. In [148] the authors propose a copula-based solution to incorporate statistical dependence between disparate sources of information. The important problem of identifying the best copula for binary classification problems is also addressed and a copula based test-statistic, able to decouples marginals and dependency information, is developed.

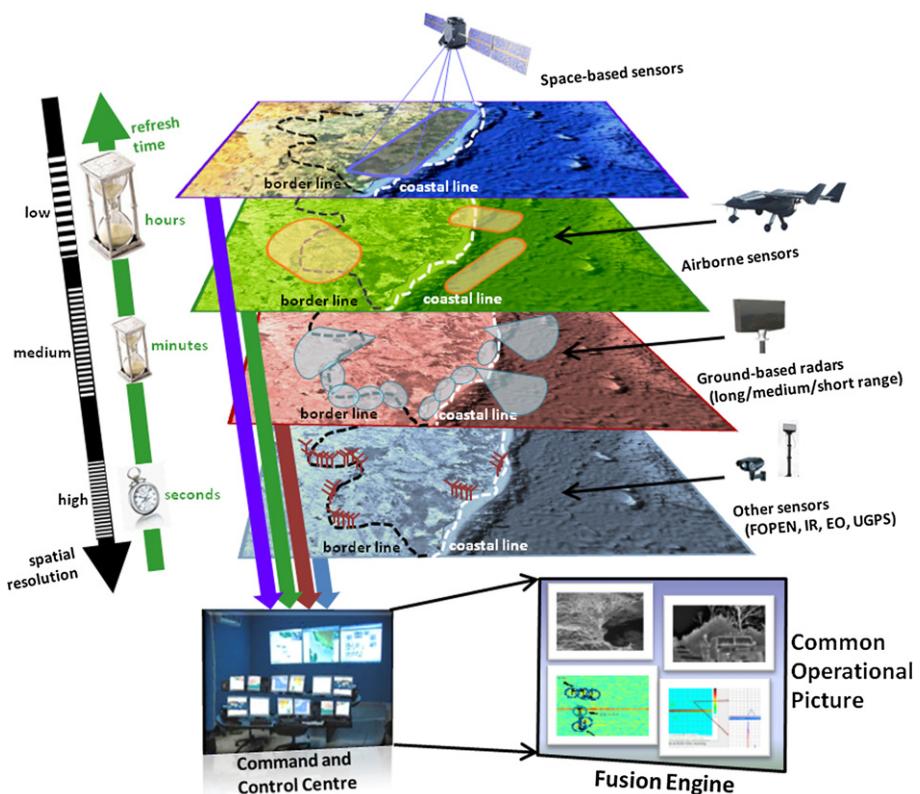
2.22.8 Border control problem via electronic fence

This section tackles the problem of the surveillance of the borders of a nation. The region of interest, in general, may be very wide consisting even of thousands of kilometers of coastline and land border line, and millions of square kilometers. Such a system must face threats such as drug trafficking, intrusions (man, vehicles and airplanes), illegal immigration, smuggling, human trafficking, arms smuggling, unauthorized deforestation, terrorist activities over the military defense of the borders in order to ensure the territorial defense and the national sovereignty in the areas close to the border line. In the following Sections an overview of the range of possibilities and solutions in the design of the surveillance asset and data fusion process of such systems devoted to border control is given.

2.22.8.1 Multi-scale approach

The size of the region, the nature of the border and the complexity of the scenario require the provision of different pictures of the region with different field of view at different resolution and time scales, suggesting a multi-sensor/multi-scale approach integrated in a hierarchical architecture of the whole system. Typically a global field of view of the whole region is necessary at the higher Command and Control (C2) level to capture the overall situation. A higher level of resolution and refresh rate is necessary at the lower and local level to analyze and control in depth each single zone of a region. Therefore the surveillance segment may be structured according to a multilayer architecture where layers realize different trade-offs in terms of field of view and granularity and refresh time. The surveillance segment comprises several types of sensors, each one characterized by different achievable resolution, field of view, and revisiting time. A pictorial sketch of the surveillance architecture is depicted in Figure 22.21 for a notional country: sensors on board of satellites are expected to provide a global coverage of the monitored area at medium resolution with a low refresh rate, typically in the order of several hours or days; a higher resolution data and a higher refresh rate, in the order of seconds or tens of seconds, is provided by ground sensors on limited areas; airborne sensors (e.g., Unmanned Air Vehicle, UAV) will provide data on remote areas with good resolution data and short deployment time.

All data collected by the sensors are exploited by the fusion engine, highlighted in the figure. It is responsible to track and classify relevant entities present in the scenario and to provide a high quality representation of the situation. Also the data fusion process supports this multi-scale approach performing a distributed and network-centric processing at the various levels of the architecture, in accordance with available communication bandwidth and latency.

**FIGURE 22.21**

Pictorial of the surveillance architecture.

2.22.8.2 The electronic fence

The surveillance of critical perimeters is one of the most important issues in Homeland defense and Homeland Protection systems. The ground surveillance needs are relevant to border protection applications, but include also local area protection, such as critical infrastructure, military/civilian posts.

During the last 10 years special attention has been focused on the realization of so-called “electronic fence” for perimeter/border control and several developments have been carried out to demonstrate the efficiency of such systems. However several problems occurred when the electronic fences became operational, showing lacks in the practical use by the operators (i.e., high number of false alarms, loss of/slow communication links) together with the problem of the high funding required for the whole system. One example is described in [149], that requires now a total different approach for the surveillance of a wide national border (>500 km).

In the following an overview of the problems and solutions related to the implementation of an electronic fence is presented. The major components are:

- *Sensors*: they may be either active or passive, radar networks or heterogeneous sensor networks, (e.g., passive IR—infrared, seismic, acoustic, electro-optic—E/O, etc.).
- *Communication network*: necessary to data exchange, may be subdivided into sub-networks if necessary.
- *Fusion engines*: they perform data collection, data fusion and classification; this capability can be spread across the layers that compose the electronic fence (i.e., in the master stations, but also in the C2 centers).

Depending on the geographical deployment of the protection system, the data are then exchanged with C2 centers, both at local level and wide area (i.e., national) level. In Figure 22.22 an example of an electronic fence architecture is depicted. In this case a wide area to be controlled, such as a border of a nation, has been considered; the subnets are geographically distributed along the boundaries. The architecture has the advantage to be modular and scalable and it can be organized with different level C2 centers (local, regional, national), depending also on the size of the considered boundaries. Each subnet is able to ensure the data exchange among the sensors. An overview of the sensors that can be employed in an electronic fence is presented.

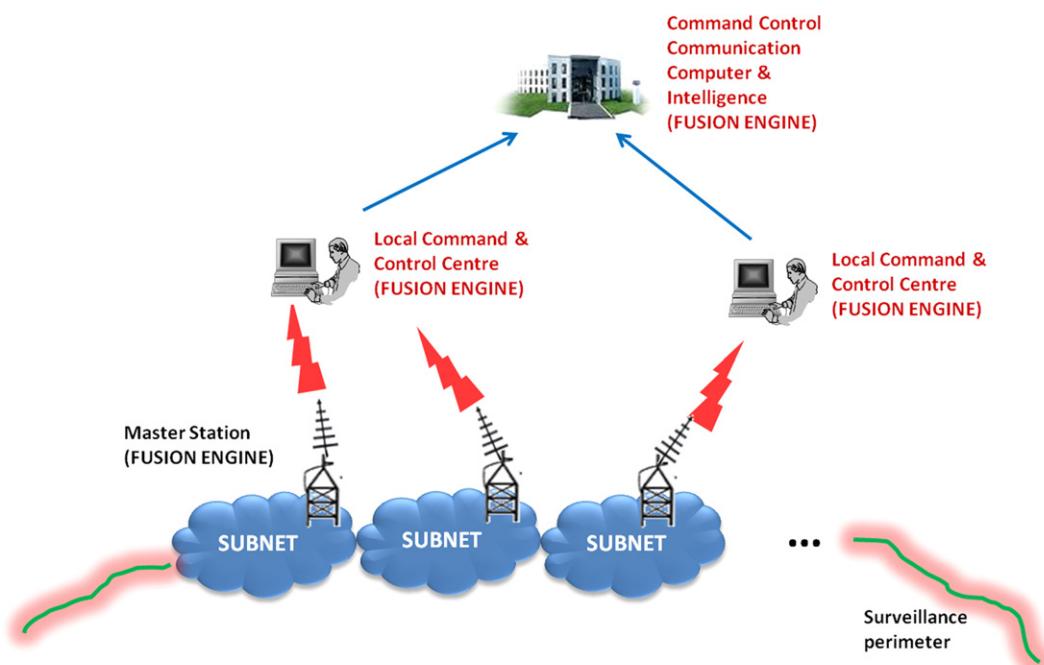


FIGURE 22.22

Example of electronic fence architecture.

2.22.8.2.1 Sensors

Ground-based sensors

Microwave (X, Ku, Ka band) ground based radars are widely used to perform the monitoring of open wide areas. The monitoring of walking people and vehicles for ground applications, and of small sized boat for sea and river applications are relevant. The detection ranges varies from 2 km to 10 km for people, and from 5 km to 20 km for vehicles. Aerial targets (e.g., helicopters, low level aircraft) are also detected. Depending on the technology used these radars can be subdivided into the following two categories:

- *Incoherent*: they are low cost devices, FMCW (Frequency Modulated Continuous Wave) or pulsed (often a magnetron is used as most of the navigational radars), where the detection of the moving targets is based on inter-clutter visibility. Resolutions are typically of few meters or tenths of meters both in range and cross-range.
- *Coherent*: they are solid state transmitter based, FMCW or pulse compressed, where the detection of the moving targets is based on sub-clutter visibility. The MTD (Moving Target Detection) filtering, even if the radar is working at X-band, requires low scan rates (in the order of 1–3 RPM—Round Per Minute) to allow high Doppler frequency resolution (0.2–0.5 m/s) to resolve slow moving target also in presence of strong clutter [150].

Airborne sensors

The attention is for sensors able to operate in critical environments and many studies have been performed, in this direction, mainly using aerial platforms equipped with SAR. The aircraft equipped with sensors are used for wide areas where ground based sensors are not suitable or cannot be installed, such as in forest or jungle. However the use of airborne platforms to perform surveillance, are limited to missions “on spot” because it is not practical or cost/effective for continuous surveillance. The radar sensor can be mounted on manned or unmanned aircraft, usually equipped with electro-optic devices, and they can be used to monitor areas of several tenths of kilometer length. Other solutions take into account the installation of the radar either on a tethered aerostat or on a hovering helicopter. GMTI (Ground Moving Target Indication) from a stationary platform has been demonstrated.

FOLiage PENetrating (FOPEN) radar

Fixed radars for border control are usually in X and Ku band, but, because of the attenuation they suffer from foliage, they cannot be used for FOPEN applications. The ability of traditional microwave radars in operating in an environment with dense foliage is severely limited by foliage backscatter and attenuation of microwave frequencies through foliage [151]. As attenuation falls with increasing wavelength, lower frequencies such as those in the VHF and UHF bands (30–1000 MHz) may be suitable for FOPEN radar applications [152–155]. FOPEN SAR (Synthetic Aperture Radar) systems started to be used in the early 1990s. They are usually mounted on manned or unmanned aircraft and mainly address illegal activity control and search-and-rescue operations. The focus is now for ground based systems and/or sensors with capabilities to detect walking personnel and moving vehicles [156]. Logistic constraints drive the technology to very low power devices, that are able to operate for several months or years, without maintenance. Another important issue is, together with a good probability of detection, the low

false alarm probability, that is requested to be lowered up to 1 false alarm per day, or lower, even in presence of specific weather conditions (rain, wind) and/or local seasonal fauna.

A special attention is due to the effect of environment. In dense foliage environments the main clutter effects are the backscatter and the attenuation.

Backscatter: The fixed clutter returns can have a zero Doppler component raising up to 60–70 dB above the noise level with spectra amplitude and shape without large variations with frequency, but depending mainly on the wind strength [150]. Considering the measurements reported in [150] of the backscatter Doppler spectra, in order to perform efficient clutter rejection, two values of thresholds can be used: i.e., 1 m/s in case of light air, 2 m/s in case of windy/gale.

Attenuation: The attenuation depends mainly on the frequency used and the radar beam grazing angle, even if small variations are reported with different polarizations [153]. Many studies have been carried out for SAR application and several studies report data for attenuation measured directly at ground level [151, 153, 154, 157]. The total attenuation, taking in account the major effects of the environment for a ground radar, can be summarized as follows:

$$L_t = L_r L_h L_f, \quad (22.45)$$

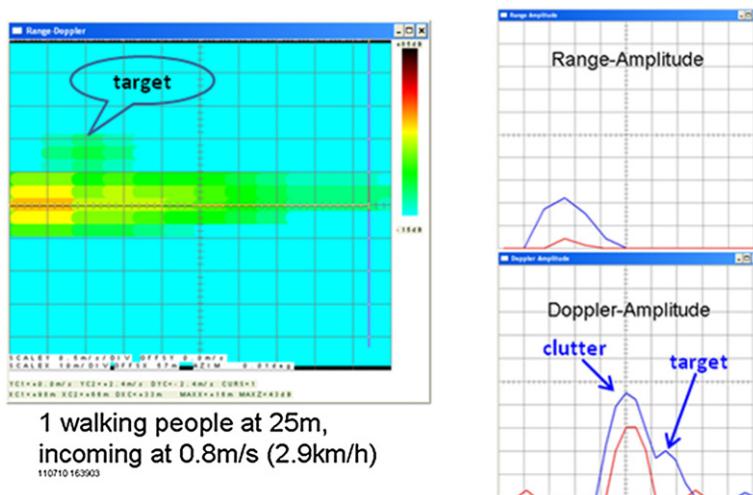
where:

- $L_r = \frac{1}{(4\pi R^2)^2}$ is the attenuation due to distance R ,
- $L_h = (4\pi \frac{h_t h_r}{\lambda R})^4$ is the attenuation due to the ground reflection at the heights of the antenna (h_r) and the target (h_t), for the wavelength λ ,
- L_f is the attenuation due to the foliage: it depends on the distance, the polarization and the forest type. It depends also on the distribution of the trees and the diameter of the masts, that can limit the line of sight, together with the height and density of lower canopy level.

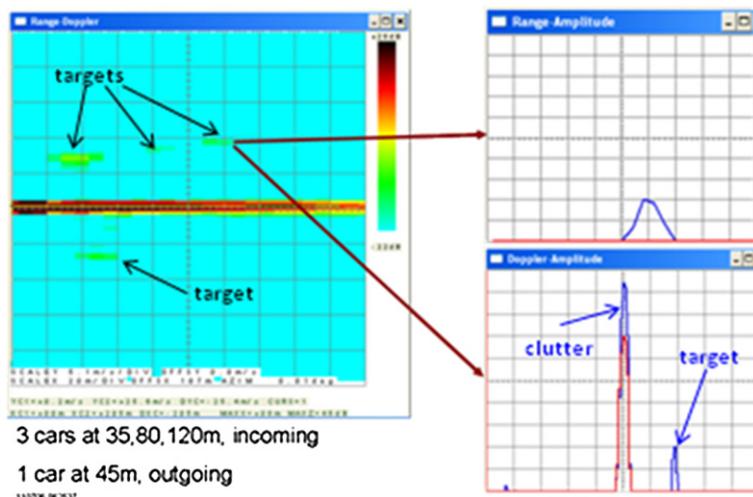
The main requirements/constraints addressed are the range of the detections, which is reduced by the attenuation due to foliage and the low antenna height, that is usually limited to 1–2 m for logistic purposes. Also the power consumption must be kept at minimum level, also considering that photovoltaic cells are not suitable for installation on the ground in the forest. As a consequence the emitted power must be kept at a level of several mW. Camouflage and anti-tamper are often required. Very low cost is a mandatory requirement. Low Probability of Intercept (LPI) capabilities are necessary. Walking personnel and moving vehicles should be detected.

Even if the FOPEN radars are referred to the forest environment, the sensor described above is suitable to operate also in different installations, considering, for example, riverside or sea harbor protection applications. In these cases the different environmental conditions allow to achieve better radar performances. In addition, several other constraints (for example the management of transmitted power) can be mitigated by the use of photovoltaic cells and/or different antenna installations.

In Figures 22.23 and 22.24 some outputs of the target detected by the UHF radar are shown. The information are displayed on range-Doppler maps, that are suitable to be read by a trained operator, giving information on the radial speed and, with a medium–high resolution in range, it helps the operator in the targets discrimination and alarm recognition.

**FIGURE 22.23**

UHF radar range-Doppler map: walking people.

**FIGURE 22.24**

UHF radar range-Doppler map: vehicles.

**FIGURE 22.25**

An example of positioning of UGPS in operative field.

Other sensors

In this section we consider the Unattended Ground Passive Sensors (UGPS) and Electro-Optic (EO) to detect moving people or vehicles.

UGPS. They are used in case of small areas or critical infrastructure perimeter surveillance. They give alarms in presence of target in the operational range and, in some cases, can give a pre-classification of the target detected. The range of each sensor is usually limited to 10 m, but the latest technologies promise to reach detection ranges up to 50 m. They have very small dimension (less than 11 volume) and low weight (less than 1 kg); they can be rapidly installed on rough ground or roads. Figure 22.25 gives an example of positioning of UGPS in an operative field. They are of following basic types:

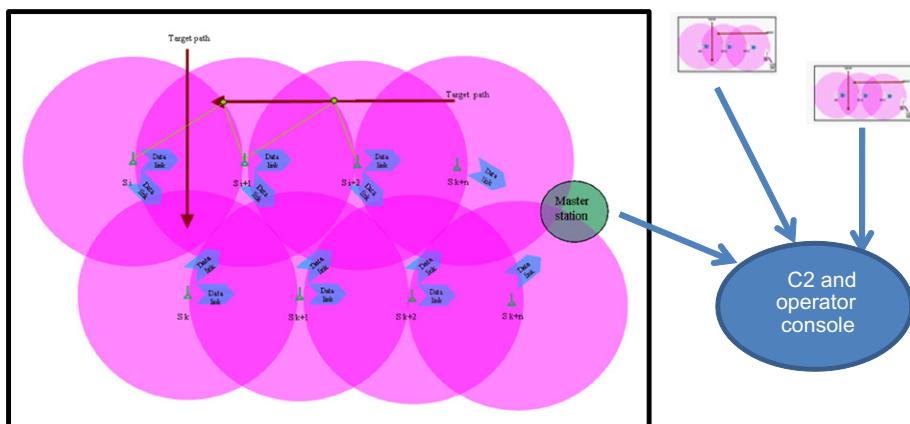
- seismic: to detect seismic movement produced by vehicles wheels or people walking,
- acoustic: to detect vehicle engine noise,
- infrared: to detect differences in thermal data from the environment due to the infrared signature of people and vehicles,
- magnetic: to detect magnetic filed variation produced by vehicles.

Electro-Optic: They are widely used for surveillance, and many signal processing techniques assist the operator for target detection alerts.

They can be fixed or rotating covering up to 360° in azimuth. For the night vision infrared EO are used, either passive or active, and they can reach a visibility of several kilometers in range. The EO are normally used stand alone or connected with radar sensor to help the operator for classification and identification of the detected targets. For example, with active infrared the operator can read (up to 2 km far from the camera) the license plate of a vehicle previously detected and tracked by the radar.

2.22.8.2.2 *Sensor network*

The sensors operate in cluster, and they are connected via a low power RF link, operating at UHF or L/S bands. The data of the unmanned radars can be combined with the data of other UGPS sensors

**FIGURE 22.26**

Unattended ground radar network.

(infrared, acoustic, seismic), or connected to an existing network, to perform a more reliable detection system.

In Figure 22.26 an example of sensor network is reported. As shown, adjacent sensor nodes are connected together and the information are sent, to the master station, via the short range radio link; the master station performs data fusion and medium range connection with the other master stations, or the C2 center. In case of long range connection the master stations are connected via radio link repeaters or satellite connections.

Special care must be taken to avoid interactions among the sensors, where two or more sensors share the same visibility area. Mutual interferences can be avoided using different frequencies and/or different timing for the transmitted waveform and also orthogonally coded waveforms.

The data transfer among the nodes is performed using the radio link between adjacent nodes. In case of linear geometric distribution the data grow up linearly with the number of nodes in the subnet; as a consequence the number of nodes in the subnet is limited by the maximum data rate of the single connection link.

The linear electronic fence can be composed of two or more parallel sections to allow redundancy in case of failure or loss of visibility of one or more sensors.

An example of electronic fence is shown in Figure 22.27. In this case different environment conditions have been considered (riverside, forest, manmade buildings, and obstacles) and a network of FOPEN unattended ground radar sensors is used.

2.22.8.2.3 Fusion engine

The fusion engine allows to fuse heterogeneous sensor data at multiple levels to perform tracking and classification of relevant entities present in the scenario and to provide a high quality representation of the situation together with cartographic layers and sensed images of the terrain. Figure 22.28 provides an example of architecture for the fusion engine.

**FIGURE 22.27**

Border surveillance: a notional case.

**FIGURE 22.28**

Fusion engine architecture.

The tracking function processes the raw data provided by sensors and generates a set of tracks, representative of the real entities present in the scenario. A track typically carries the following information: a timestamp, position coordinates, velocity components, uncertainty on the kinematic components as expressed through the covariance matrix and additional attributes such as class/type and identity. In consideration of the potentially huge geographic extension of the system and of the importance

to optimize the deployment of sensors as well as communication and processing power resources, a distributed tracking architecture is necessary. At the first level of the tracking architecture each sensor produces its own “local” tracks, in order to make available to the fusion engine a filtered information. Then a second level tracking combines local tracks originating from different sources into system tracks. This solution distributes the computational load on the peripheral nodes and reduces considerably the communication traffic which must be transmitted from the local level to the higher echelons; this is extremely important in consideration of the reduced bandwidth generally available between the peripheral elements and the center of the system.

In this step of the process, information of different nature can be fused producing a unique high quality information. Radar tracks can be fused with multiple images acquired by SAR and optical sensors, even if acquired at different resolutions, to achieve an improved representation of the scene with respect to the one achievable by processing data sets separately, in particular in terms of detection and false alarm probabilities when dealing with small targets (i.e., targets that occupy only few pixels of the image) [158–162]. The cartographic layers, superimposed with SAR or optical images, allow to put into context all the available information and support the fusion process (e.g., target tracking for ground vehicles especially during maneuvers).

Another output of the fusion engine is the classification of the tracked targets and entities of the scenario, i.e., the attribution of a class to the track under examination, hence supporting the capability to achieve a situation awareness.

From an operational point of view, the fusion engine can be considered as the responsible of producing a multi-resolution and multi-layer COP (Common Operating Picture), whose definition, as provided by [163], is the following: “A single identical display of relevant information shared by more than one command that facilitates collaborative planning and assists all echelons to achieve situational awareness.” The COP therefore provides to the operators at the different levels the capability to view each time a well-suited map, both in terms of proper scale (with respect to the scale of the observed situation) as well as in terms of number and type of information, according to the situation under analysis. This characteristic allows the system to properly support the operator without overloading him with unimportant information and keep him focused on events and information that might be related with his goal in terms of spatial, temporal, and logic correlation.

In the following the main constituents of the fusion engine are described.

Local tracking

The local tracking function processes the measurements provided by the sensor and produces a local track for each of the observed targets present in the surveillance region. The task of the tracking function at the local level is therefore of using the measurements made available by the sensor to estimate the number of targets and their kinematic components [164–166]. Local tracks provide position and velocity estimates at a given time, together with an indication of track quality; the track may also include other attributes relative to track classification, derived directly from radar measurements, from other sensors (EO/IR, UGPS, UAV) or assigned by a human operator.

In the scenario of a generic land border may be necessary to form low altitude tracks, surface tracks and ground tracks. Tracking of ground targets is especially critical due to the characteristics of the ground environment and of ground targets. The main criticality may be the masking effect due to terrain orography and vegetation. Another interesting feature of the ground environment is the presence of areas,

mainly roads, where the probability of finding targets is higher, and areas such as off-road where the presence of targets is less probable. Distinguishing features of ground targets are high maneuverability and move-stop-move dynamics.

Even a well trained operator would be unable to select the correct hypothesis when a ground target is maneuvering since available information is insufficient. In these situations the best strategy is to defer the final decision until more data is available. To take into account these difficulties, the tracking function must be designed so as to handle several concurrent hypotheses and to make final decisions with a deferred logic [167–169], i.e., when more data is available which allows to make a final decision with sufficient confidence. The choice of hypotheses is also dependent on the environment and on the target type. The management of multiple hypotheses is then the capability of the function to consider at each time instant a set of hypotheses, such as:

- the target is proceeding regularly/is maneuvering on road;
- the target is moving/maneuvering off-road;
- the target has stopped, etc.

The tracking function assigns a score to each hypothesis and identifies the most probable; the function keeps alive for some time not only the most likely hypothesis but also a set of alternative hypotheses which represent different kinematic evolutions of the target. Figure 22.29 shows an example of the set of hypotheses generated by the function: each hypothesis is relative to a path in the tree from time t_0 to time t_3 and the single branches may be relative to the choice of a specific dynamic model and/or a specific correlation hypothesis with a measurement in the set. For example in the path highlighted in red it is assumed that the target trajectory in the interval t_0-t_3 is described by the dynamic model m_1 ; the other branches are relative to alternative hypotheses where it is assumed for example that the target has maneuvered (m_2) or stopped (m_3), etc. As new information is acquired, the probability of each hypothesis is updated according to new information; hypothesis which initially have a low score may gain credibility and vice versa. This characteristic, i.e., defer the decision until the available information is considered sufficient, allows to resolve most critical situations.

To take into account terrain and geographic information, the tracking solution leverages also context information provided by the GIS (Geographic Information System) in accordance with logics of terrain

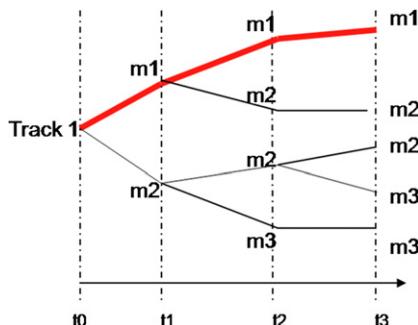
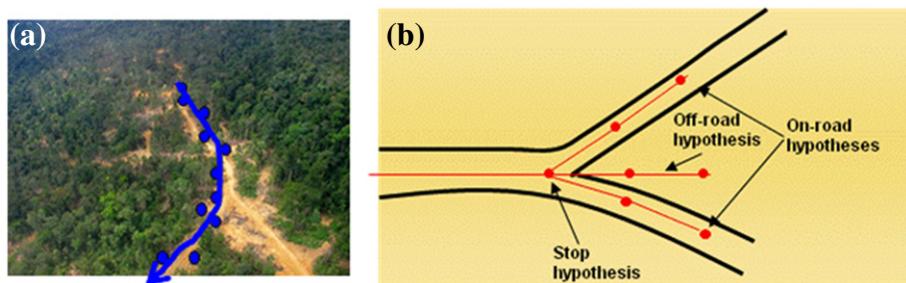


FIGURE 22.29

Set of hypotheses generated for a track.

**FIGURE 22.30**

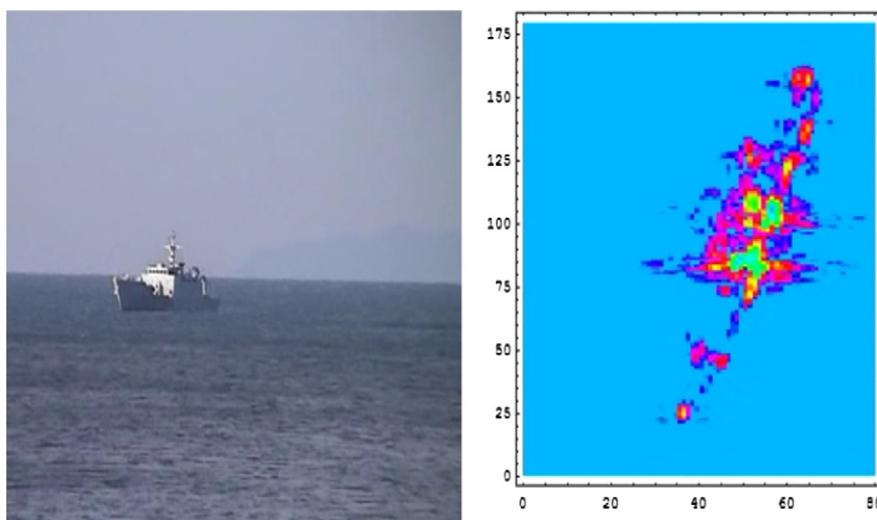
(a) Terrain aided tracking, (b) road aided tracking.

and road aided tracking. Digital Terrain Elevation Data (DTED) are also used to perform accurate projections of the tracks on the terrain and to identify zones where the target trajectory will be masked by obstacles and thus improve track continuity and the estimate of track kinematic parameters (e.g., maximum target velocity given the terrain type). The following Figure 22.30 shows, for instance, how environmental knowledge can be exploited to improve the tracking function [170,171]. Figure 22.30a shows a landscape covered by forests and crossed by a network of paths; due to the nature of the environment, targets especially if motorized, will preferentially move along the track, avoiding off-road areas more difficult to traverse. The blue line represents the trajectory of a track which moves along a winding path in the forest. Figure 22.30b on the other side shows how information relative to roads and viability in general can be exploited to improve the tracking performance. When the track approaches a bifurcation or a crossing, different hypotheses are generated to take into account possible target trajectories, such as on-road, off-road and also move-stop motion. More specifically the adoption of techniques such as road aided tracking is specifically important since it allows to improve the accuracy in the estimation of target kinematic parameters and therefore to make longer term projections. Finally weather information is exploited to further improve the tracking processing by feeding in information about areas where target detection is less probable (e.g., flooded areas) and expected target velocity is low given the past days weather conditions (e.g., heavy rain is expected to result in limited target velocity).

Classification

The classification function allows attributing a class to the track under examination, i.e., to determine its belonging to a class of targets. Target classification is extremely important since it helps to determine target identity and its threat level. Part of the classification process is the non-cooperative target recognition (NCTR), in order to avoid fratricide and to allow proper allocation of defensive means against the threat. In a coastal scenario NCTR capabilities are needed against ships, potentially involved in terrorism, illegal immigration or contraband operations, in order to assess and prioritize threats and to provide the appropriate response.

Sensors such as radar, EO/IR, may provide useful information for classification. In the radar case, the NCTR technology facilitates the identification of non-co-operative targets by transmitting wide band signals and by processing the radar echoes in a suitable multidimensional domain; e.g., time-frequency and range-angle. In the former case the target is discriminated on the basis of the jet engine or the

**FIGURE 22.31**

A snapshot of ISAR (Inverse SAR) signal processing, a profile of a ship along range and cross-range.

helicopter rotor modulations of the echo [172–175]; in the latter case the target is discriminated on the basis of the measured two-dimensional radar image obtained by ISAR techniques [176–178] (Figure 22.31 shows a snapshot of the radar image of a ship).

The automatic classification, that the radar is capable of providing by means of these processing techniques, is used directly within the tracking function, to support the plot-track correlation process and to attribute a class to the track. The classification process allows therefore determining the class to which the track belongs (such as pedestrians, vehicles, convoys, helicopters, and small low altitude aircrafts) and performing cueing to other sensors (e.g., EO/IR sensors, high resolution radars) or demanding a patrolling mission (e.g., a mission with UAV).

While data provided by sensors are needed to perform the classification processing, once the target has been assigned to a class, this information can be exploited at sensor level to achieve better accuracy in the performed processing (e.g., target classification can be used to refine kinematic target parameters used in the tracking processing).

The range-Doppler information can be furthermore employed to produce a confusion matrix useful for target classification. The confusion matrix expresses the a posteriori probability that a target has been classified correctly among a finite number of classes that have been a priori established. References [21, 179] give an example of the use of confusion matrix in the classification issue.

2.22.9 Estimation and forecasting of an epidemic

Epidemics can impose serious challenges on societies in modern times. The poor health of general population due to a disease causes hardship and pain but also negative trends in the economy through

absenteeism from work, missed business opportunities, etc. The ongoing epidemics of AIDS (Acquired Immune Deficiency Syndrome), tuberculosis and the recent outbreaks of SARS (Severe Acute Respiratory Syndrome) and H1N1 (swine flu) provide some revealing examples.

In the absence of an effective cure against an infectious disease, the best approach to mitigate its malicious or natural epidemic outbreak resides in the development of a capability for its early detection and prediction of its further development [180]. This enables typical countermeasures, such as the quarantine, vaccination, medical treatment, to be much more effective and less costly [181, 182]. Therefore this issue can be approached as a surveillance problem in the context of Homeland Protection.

Syndromic surveillance is referred to as a systematic collection, analysis, and interpretation of public health data for the purpose of early detection of an epidemic outbreak and the mobilization of a rapid response [180, 182]. The key idea is to detect an epidemic outbreak using early symptoms, well before the clinical or laboratory data result in a definite diagnosis. The rationale is that a spread of an infectious disease is usually associated with the measurable changes in the social behavior, which can be measured by non-medical means. Recent studies [183–185] have demonstrated that these non-medical sources of syndromic data streams, such as the absenteeism from work/school, the pharmaceutical sales, internet queries, twitter messages, and alike, can enable one to draw important conclusions regarding the epidemic state in the community. The “Google Flu” project [186] (flu-related searches in Google) is a well publicized example of this approach.

The algorithms for syndromic surveillance have recently attracted significant attention by scientists and practitioners; there is a vast amount of literature devoted to this topic (for more comprehensive review see [180, 182] and references therein). In general, all algorithms applied in this area can be divided into two main groups, the *data mining* methods and the *information fusion* (also known as data assimilation) methods. Data mining is primarily concerned with the extraction of patterns from massive amounts of raw data without using dynamic models of the underlying process (i.e., epidemic spread) [183, 185]. Information fusion algorithms, on the contrary, strongly rely of mathematical models: in this case, the *dynamic model* of an epidemic outbreak and the *measurement model* of a particular syndromic data stream [187, 188]. Naturally, the accuracy of information fusion algorithms is significantly determined by the fidelity of the underlying models.

This section presents a study of a recursive information fusion algorithm for syndromic surveillance, formulated in the Bayesian context of stochastic nonlinear filtering and solved using a particle filter [134]. While a similar work has been considered earlier, see [189–192], this section introduces two novelties. First, in order to overcome the limitations of the standard “compartment” model of epidemic spread (the “well-mixed” approximation) we employ a more flexible alternative, see [193, 194]. The adopted epidemic model has the explicit parameter of “mixing efficiency” (or level of social interaction) and is therefore more appropriate to represent a variety of social interactions in a small community (e.g., self-isolation and panic). An advantage of the adopted epidemiological model is also that it enables to estimate the scaling law of the noise level with respect to the population size of a community. Second, a more flexible model of syndromic measurements, validated with data sets available in the literature [183, 186], is adopted in the section. This measurement model is robust in the sense that some of its parameters are specified imprecisely, as interval values. The optimal sequential estimator (filter) and predictor are then formulated in the Bayesian framework and solved using a particle filter.

2.22.9.1 Modeling

To describe the dynamics of an epidemic outbreak we employ the generalized SIR (Susceptible, Infectious and Recovered) epidemic model with stochastic fluctuations [195–197]. According to this model, the population of a community can be divided into three interacting groups: susceptible, infectious and recovered. Let the number of susceptible, infectious and recovered be denoted by S , I , and R , respectively, so that $S + I + R = P$, where P is the total population size. The dynamic model of epidemic progression in time can be then expressed by two stochastic differential equations subject to the “conservation” law for the population:

$$\begin{aligned}\frac{ds}{dt} &= -\alpha \cdot i \cdot s^\nu + \sigma_q \xi, \\ \frac{di}{dt} &= \alpha \cdot i \cdot s^\nu - \beta \cdot i - \sigma_q \xi + \sigma_\beta \zeta, \\ r &= 1 - s - i,\end{aligned}\tag{22.46}$$

where $s = S/P$, $i = I/P$, $r = R/P$, and ξ , ζ are two uncorrelated white Gaussian noise processes, both with zero mean and unit variance. The terms $\sigma_q \xi$ and $\sigma_\beta \zeta$ are introduced into Eq. (22.47) to capture the demographic noise (random variations in the contact rate α and in the recovery time β) [197, 198]. Parameter ν in Eq. (22.47) is the population mixing parameter, which for a homogeneous population equals 1. In the presence of an epidemic, however, ν may vary as people change their daily habits to reduce the risk of infection (e.g., panic, self-isolation). In general, model parameters α , β , ν can be assumed to be partially known as interval values. In order to insure $P\{s, i, r \in [0, 1]\} \approx 1$, standard deviations σ_q , σ_β need to satisfy [199]:

$$\sigma_q \approx \frac{\sqrt{\alpha}}{P}, \quad \sigma_\beta \approx \frac{\sqrt{\beta}}{P}.\tag{22.47}$$

Assuming that non-medical syndromic data are available for estimation and forecasting of the epidemic, we adopt a measurement model verified by [185, 186], where a power law relationship holds for the odds-ratio between the observable syndrome z_j and the (normalized) number of infected people i :

$$\frac{z_j}{1 - z_j} \propto \left(\frac{i}{1 - i} \right)^{\varsigma_j}, \quad j = 1, \dots, N_z.\tag{22.48}$$

The power law exponent ς_j in Eq. (22.48) is in general syndrome specific. Since at the initial stages of an epidemic (which is of main interest for early detection and forecasting) we have: $i \ll 1$ and $z_j \ll 1$, Eq. (22.48) can be reduced to a simple power-law model:

$$z_j = b_j \cdot i^{\varsigma_j} + \tau_j,\tag{22.49}$$

where b_j is a constant and τ_j is introduced to model the random nature of measurement noise. It is assumed that τ_j is uncorrelated to other syndromes and dynamic noises ξ, ζ . Since $z_j \geq 0$ (e.g., number of Google searches), the noise term τ_j associated with syndrome j should be modeled by a random variable that provides strictly non-negative realizations. For this purpose we adopt the log-normal distribution, that is $\tau_j = \sigma_j \eta_j$, with $\eta_j \sim \ln N(0, 1)$ and $N(0, 1)$ being the standard Gaussian distribution.

Parameters b_j , σ_j , ς_j typically are not known, but with a representative data set of observations the model of Eq. (22.49) can be easily calibrated (see for example the results of the linear regression fits in [186]). The data fit reported in [183] suggests that ς_j may be close to unity, although it is difficult to precisely specify its value because of significant scattering of data points). To cater for this uncertainty, we assume that ς_j can take any value in an interval, $\varsigma \in [\varsigma_1, \varsigma_2]$ around $\varsigma = 1$. Unfortunately [185, 186] do not report any specific values of fitting parameters, so we use in this study some heuristic values for b_j , σ_j in our simulations.

The problem now is to estimate the (normalized) number of infected i , and susceptible s at time t , using syndromic observations z_j of Eq. (22.49), collected up to time t . Let \mathbf{x} denote the state vector to be estimated; it includes i and s , but also the imprecisely known epidemic model parameters α , β and v . The formal Bayesian solution is given in the form of the posterior pdf $p(\mathbf{x}_t | z_{1:t})$, where \mathbf{x}_t is the state vector at time t and $z_{1:t}$ denotes all observations up to time t . Using the posterior $p(\mathbf{x}_t | z_{1:t})$, one can predict the progress of the epidemic using the dynamic model of Eq. (22.47).

2.22.9.2 Sequential Bayesian solution

For the purpose of computer implementation, first we need a discrete-time approximation of dynamic model of Eq. (22.47). The state vector is adopted as: $\mathbf{x} = [i \ s \ \alpha \ \beta \ v]^T$, where T is the matrix transpose. Using Euler's method with small integration interval δ , the nonlinear differential equations in Eq. (22.47) can be approximated as

$$\mathbf{x}_{k+1} \approx f_k(\mathbf{x}_k) + \mathbf{w}_k, \quad (22.50)$$

where $k = t_k/\delta$ is the discrete-time index and

$$f_k(\mathbf{x}) = \begin{bmatrix} \mathbf{x}[1] + \delta \cdot \mathbf{x}[1] \cdot (\mathbf{x}[3] \cdot \mathbf{x}[2]^{\mathbf{x}[5]} - \mathbf{x}[4]) \\ \mathbf{x}[2] - \delta \cdot \mathbf{x}[3] \cdot \mathbf{x}[1] \cdot \mathbf{x}[2]^{\mathbf{x}[5]} \\ \mathbf{x}[3] \\ \mathbf{x}[4] \\ \mathbf{x}[5] \end{bmatrix} \quad (22.51)$$

is the transition function; here $\mathbf{x}[i]$ denotes the i th component of vector \mathbf{x} . Discrete-time process noise \mathbf{w}_k in Eq. (22.50) is assumed to be zero-mean white Gaussian with diagonal covariance matrix \mathbf{Q} , which according to Eq. (22.47) can be expressed as $\mathbf{Q} = \text{diag}[(\alpha + \beta)\delta^2/P^2, \alpha\delta^2/P^2, 0, 0, 0]$.

The optimal Bayes filter is typically presented in two steps, *prediction* and *update*. Suppose the posterior pdf at time t_k is given by $p(\mathbf{x}_k | Z|_{1:k})$. Then the prediction step computes the pdf predicted to time $t_m = t_k + \delta$ as [194]:

$$p(\mathbf{x}_m | z_{1:k}) = \int \pi(\mathbf{x}_m | \mathbf{x}_k) p(\mathbf{x}_k | z_{1:k}) d\mathbf{x}_k, \quad (22.52)$$

where $\pi(\mathbf{x}_m | \mathbf{x}_k)$ is the transitional density. According to Eq. (22.50), we can write $\pi(\mathbf{x}_m | \mathbf{x}_k) = N(f_k(\mathbf{x}_k), \mathbf{Q})$. The prediction step is carried out many times with tiny sampling intervals δ until observation $z_{j,k+1}$ about syndrome j becomes available at t_{k+1} . The predicted pdf at t_{k+1} is denoted $p(\mathbf{x}_{k+1} | z_{1:k})$.

In the standard Bayesian estimation framework, the predicted pdf is updated using measurement $z_{j,k+1}$ by multiplication with the measurement likelihood function [200]. According to Eq. (22.49), the likelihood function in this case is $g(z_{j,k+1} | \mathbf{x}_k) = \ln N(h(\mathbf{x}_{k+1}; \varsigma_j), \sigma_j^2)$, where $h(\mathbf{x}; \varsigma) = b_j \cdot \mathbf{x}[1]^\varsigma$.

The standard Bayesian approach, however, cannot be applied because $h(\mathbf{x}; \underline{\varsigma})$ defined in this way is not a function: $\underline{\varsigma}$ is effectively an infinite set (an interval) and therefore $h(\mathbf{x}; \underline{\varsigma})$ is one-to-many mapping.

An elegant solution to the imprecise measurement transformation is available in the framework of random set theory [137]. In this approach $h(\mathbf{x}; \underline{\varsigma}) + \tau$ is modeled by a random set $\Sigma_{\mathbf{x}}$ and the likelihood function represents the *probability*: $\tilde{g}(z|\mathbf{x}) = \Pr\{z \in \Sigma_{\mathbf{x}}\}$, and is referred to as the *generalized likelihood*. More details and a theoretical justification of this approach can be found in [201]. The Bayes update using syndromic measurement $z_{j,k+1}$ is now defined as [137]:

$$p(\mathbf{x}_{k+1}|z_{1:k+1}) = \frac{\tilde{g}(z|\mathbf{x}_{k+1}) \cdot p(\mathbf{x}_{k+1}|z_{1:k})}{\int \tilde{g}(z|\mathbf{x}_{k+1}) \cdot p(\mathbf{x}_{k+1}|z_{1:k}) d\mathbf{x}_{k+1}}. \quad (22.53)$$

For the measurement model Eq. (22.49) with additive Gaussian noise, the generalized likelihood has an analytic expression [201]:

$$\tilde{g}(z_j|\mathbf{x}) = \varphi(z; \underline{\Sigma}_{\mathbf{x}}, \sigma_j^2) - \varphi(z; \bar{\Sigma}_{\mathbf{x}}, \sigma_j^2), \quad (22.54)$$

where $\underline{\Sigma}_{\mathbf{x}} = \min\{h(\mathbf{x}; \underline{\varsigma}), h(\mathbf{x}; \bar{\varsigma})\}$, $\bar{\Sigma}_{\mathbf{x}} = \max\{h(\mathbf{x}; \underline{\varsigma}), h(\mathbf{x}; \bar{\varsigma})\}$ define the limits of the set and $\varphi(u; \mu, P) = \int_{-\infty}^u \ln N(y; \mu, P) dy$ is the cumulative log-normal distribution. The recursions of the Bayes filter start with an initial pdf (at time $t_k = 0$), denoted $p(\mathbf{x}_0)$, which is assumed known.

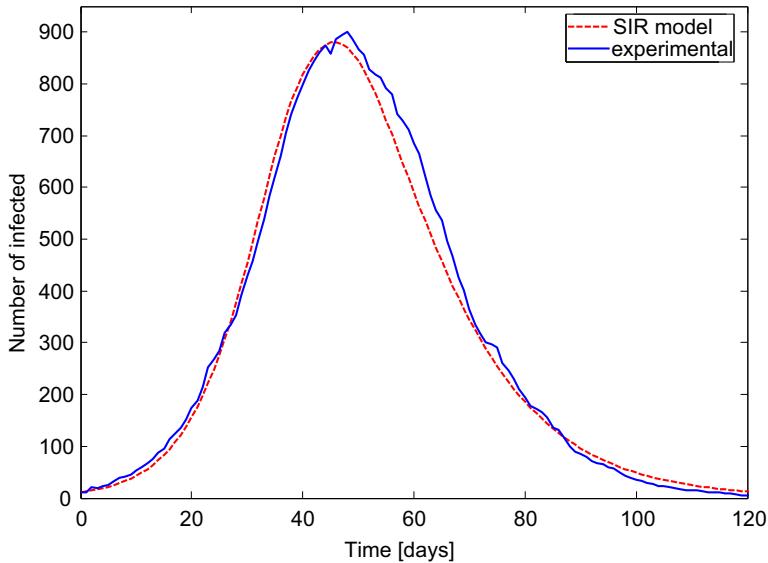


FIGURE 22.32

Experimental data set: the solid blue line represents the number of infected people over time (obtained by agent based simulation); the dashed red line is the fitted non-homogeneous mixing SIR model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

The proposed Bayesian estimator cannot be solved in the closed form. Instead we developed an approximate solution based on the particle filter (PF) [134, 202]. The PF approximates the posterior pdf $p(\mathbf{x}_k | z_{1:k})$ by a weighted random samples; details can be found in [134, 202]. The only difference here is that importance weight computation is based on the generalized likelihood function.

2.22.9.3 Numerical results

Epidemic forecasting will be demonstrated using an experimental data set obtained using a large-scale agent based simulation model [203, 204] of a virtual town of $P = 5000$ inhabitants, created in accordance with the Australian Census Bureau data. The agent based model is rather complex (takes a long time to run) and incorporates a typical age/gender breakdown, family-household-workplace habits, including the realistic day-to-day people contacts for a disease spread. The blue line in Figure 22.32 shows the number of people of this town infected by a fictitious disease, reported once per day during a period of 154 days (only first 120 days shown). The dashed red line represents the adopted SIR model fit, using the entire batch of 154 data points and integration interval $\delta = 0.0052$ days, with no process noise, i.e., $\mathbf{w}_k = 0$ in Eq. (22.50). The estimated model parameters are: $\hat{\alpha} = 0.2399$, $\hat{\beta} = 0.1066$, $\hat{\nu} = 1.2042$.

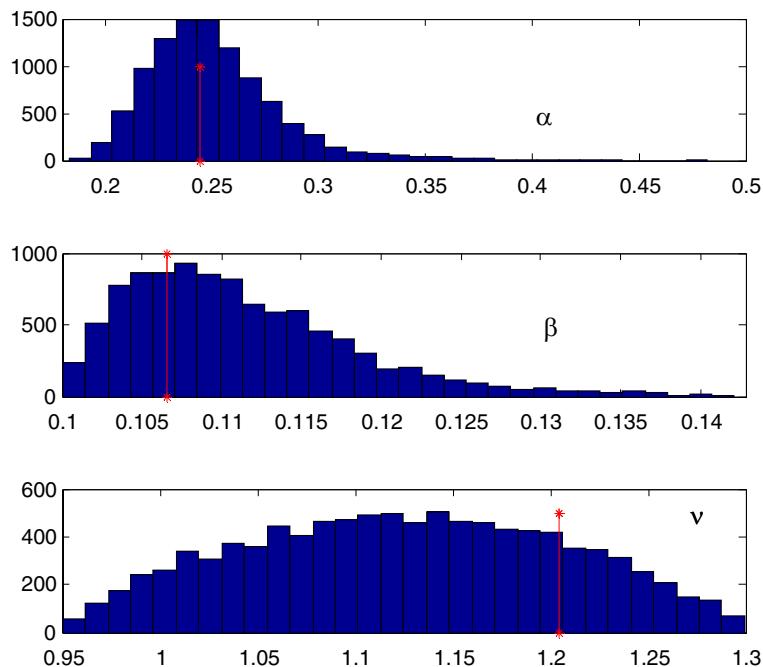
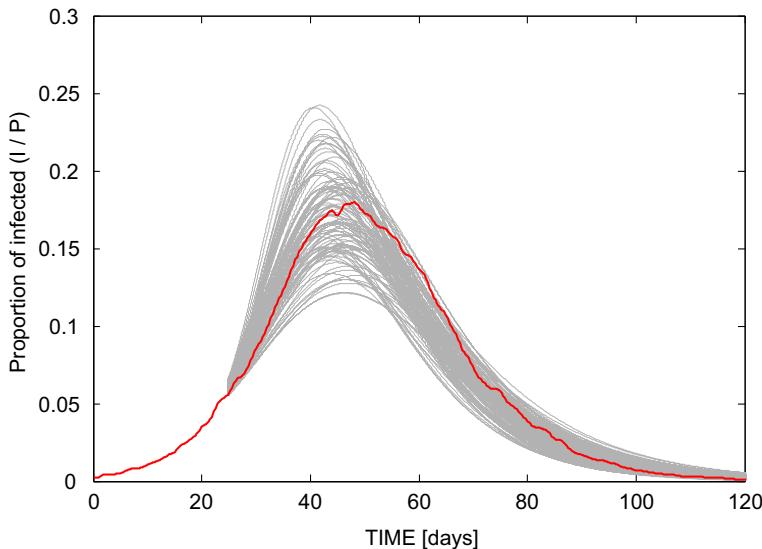


FIGURE 22.33

Histograms of particle filter estimated values of epidemic model parameters after processing 25 days of syndromic measurements. Red vertical lines indicate the true values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

**FIGURE 22.34**

Prediction results for a random sample of 100 particles (gray lines); the red line is the experimental curve from Figure 22.32. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

These estimates were obtained using the importance sampling technique of progressive correction [202]. Figure 22.32 serves to verify that the adopted non-homogeneous mixing SIR model, although very simple and fast to run, is remarkably accurate in explaining the data obtained from a very complex simulation system.

The *true* number of infected people in forecasting simulations is chosen to be the output of the agent based population model, shown by the solid blue line in Figure 22.32. The measurements are generated synthetically in accordance with Eq. (22.49) and discussions above, using the following parameters: $\varsigma = 1.05$, $b_j = 0.25$, $\sigma_j = 0.01$, for all $j = 1, 2, 3, 4$ monitored syndromes. Independent measurements concerning all $N_z = 4$ syndromes are assumed available on a daily basis during the first 25 days. The problem is to perform the estimation sequentially as the measurements become available until the day number 25, and at that point of time to forecast the number of infected people as a function of time.

The initial pdf for the state vector was chosen as $p(\mathbf{x}_0) = p(i_0)p(s_0)p(\alpha_0)p(\beta_0)p(v_0)$ with $p(i_0) = N_{[0,1]}(z_{j,1}/b_j, \sigma_j^2)$, $p(s_0) = N_{[0,1]}(1 - z_{j,1}/b_j, \sigma_j^2)$, $p(\alpha_0) = U[0.18, 0.5]$, $p(\beta_0) = U[0.1, 0.143]$, $p(v_0) = U[0.95, 1.3]$, where $N_{[a,b]}(\mu, P)$ and $U[a, b]$ denote the truncated Gaussian distribution, restricted to $[a, b]$, and uniform distribution, respectively. The imprecise measurement parameter is adopted as $\varsigma \in [1.03, 1.07]$, while its true value is 1.05. The number of particles is set to 10,000.

Figure 22.33 shows the histograms of particle filter estimated values of α , β , and v , after processing 25 days of syndromic data (i.e., in total 100 measurements). The histograms in this figure reveals that the

uncertainty in parameters α and β has been substantially reduced after processing the data (compared with the initial $p(\alpha_0)$ and $p(\beta_0)$). The uncertainty in v , on the other hand, has not been reduced, indicating that this parameter cannot be estimated from syndromic data. While this is unfortunate, it does not appear to be a serious problem in forecasting the epidemic mainly because the prior on v in practice is fairly tight ($v \approx 1$). This is confirmed in Figure 22.34 which shows a sample of 100 overlaid predicted epidemic curves (gray lines) based on the estimate of i, s, α, β, v obtained after 25 days. Figure 22.34 indicates that the forecast of the peak of the epidemic is fairly accurate, while the forecast of the size of the peak is more uncertain. Most importantly, however, the *true* epidemic curve (solid red line) appears to be always enveloped by the prediction curves. More experimental results can be found in [199].

2.22.10 Conclusions

Integrated sensor systems and data fusion have been the main focus of this chapter. The discussed matter has been subdivided in nine sections which have covered a *long trip* starting from the description of the Homeland Protection problem, to the illustration of a wide spectrum of information sources (sensors and the like), to the netting of such sensors (both homogeneous and heterogeneous), with a broad range of practical applications: cooperative sensing to defend a urban territory, network of cooperative chemical sensors, detection and localization of radioactive point sources, use of so-called electronic fence to protect long borderlines of a territory, up to the estimation and forecasting of an epidemic. This work, an unofficial collaboration between experts from industry, research centers and academia, has brought together a wide spectrum of competences scientific, technical/technological/systemic and on the field.

List of Acronyms

AEW	Airborne Early Warning
AIDS	Acquired Immune Deficiency Syndrome
AJP	Allied Joint Publication
ATC	Air Traffic Control
BASH	Bird Air Strike Hazard
C2	Command and Control
C4I	Command, Control, Communications, Computers, and Intelligence
CBINT	Chemical and Biological Intelligence
cdf	Cumulative distribution function
COMINT	Communications Intelligence
COP	Common Operating Picture
CRLB	Cramer-Rao Lower Bound
CSIP	Collaborative Signal and Information Processing
CTR	Cooperative Target Recognition
DFS	Data Fusion Subpanel

DIKW	Data Information Knowledge and Wisdom)
DSC	Dynamic Sensor Collaboration
DSTO	Defense Science and Technology Organisation
DTED	Digital Terrain Elevation Data
EA	Evolutionary Algorithms
ECM	Electronic Counter Measures
EEZ	Exclusive Economic Zone
EKF	Extended KF
ELINT	Electronic Intelligence
EML	Exact Maximum Likelihood
EO	Electro-Optical
ESM	Electronic Support Measurement
ESM	Electronic Support Measures
FIM	Fisher Information Matrix
FMCW	Frequency Modulated Continuous Wave
FOPEN	Foliage PENetration
GA	Genetic Algorithm
GEMS	Generic Error modeling System
GEOINT	Geospatial Intelligence
GIS	Geographic Information System
GM	Geiger-Müller
GMTI	Ground Moving Target Indicator
HAP	High Altitude Platform
HD	Homeland Defense
HP	Homeland Protection
HS	Homeland Security
HUMINT	Human Intelligence
IFF	Identification Friend or Foe
IMINT	Imagery Intelligence
IR	Infra Red
ISAR	Inverse SAR
JDL	Joint Directors of Laboratories
KF	Kalman Filter
LADAR	Laser Radar
LASINT	Laser Intelligence
LPI	Low Probability of Intercept
MASINT	Measurement and Signature Intelligence
MCMC	Markov Chain Monte Carlo
MDL	Minimum Description Length
MRT	Multi-Radar Tracking
MTD	Moving Target Detector
MTI	Moving Target Indicator
NCO	Network Centric Operation

NCTR	Non-Cooperative Target Recognition
NUCINT	Nuclear Intelligence
ODNA	Office of Directors of National Intelligence
OODA	Observe-Orient-Decide-Act)
OSINT	Open-Source Intelligence
PCRLB	Posterior cramer rao lower bound
pdf	Probability density function
PF	Particle Filter
RADINT	Radar Intelligence
RPM	Route Per Minute
SAR	Synthetic Aperture Radar
SARS	Severe Acute Respiratory Syndrome
SIR	Susceptible Infectious Recovered
SIGINT	Signals Intelligence
TV	Television
UAV	Unmanned Air Vehicle
UGPS	Unattended Ground Passive Sensor
UHF	Ultra High Frequency
UKF	Unscented KF
USAF	United States Air Force
VHF	Very High Frequency
WSN	Wireless Sensor Network

Acknowledgments

The first two authors wish to warmly thank Prof F. Zirilli (Univ. of Rome) for his contribution to Section 2.22.6.1, Dr. S. Gallone (SELEX Sistemi Integrati) for contributing to Section 2.22.8.2 and Dr. A. Graziano (SELEX Sistemi Integrati) for the continuous and fruitful cooperation on the many topics described in the Chapter for years.

References

- [1] J. Naisbitt, *Megatrends*, GCP, 1982.
- [2] R.T. Antony, *Principles of Data Fusion Automation*, Artech House, 1995.
- [3] A. Doucet, N. de Freitas, N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, Springer Verlag, New York, 2001.
- [4] R.E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME – J. Basic Eng.* 82 (1960) 35–45.
- [5] A.G.O. Mutambara, *Decentralized Estimation and Control for Multi-Sensor Systems*, CRC Press, 1998.
- [6] G.J. Klir, Y. Bo, *Fuzzy Logic: Theory and Applications*, Prentice Hall PTR, 1995.
- [7] A.P. Dempster, A generalization of Bayesian inference, *J. Roy. Stat. Soc.* 30 (1968) 205–247.
- [8] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

- [9] F. Smarandache, J. Dezert (Eds.), *Advances and Applications of DSmT for Information Fusion*, vol. 1, American Research Press, 2004.
- [10] F. Smarandache, J. Dezert (Eds.), *Advances and Applications of DSmT for Information Fusion*, vol. 2, American Research Press, 2004.
- [11] A. Farina, S. Pardini, Introduction to multiradar tracking system, *Rivista Tecnica Selenia* 8 (1) (1982) 14–26.
- [12] A. Farina, F.A. Studer, Radar data processing, in: P. Bowron (Ed.), *Introduction and Tracking*, vol. I, Researches Studies Press, John Wiley, England, 1985 (Translated in Russian—Radio I Sviaz-, Moscow, 1993—and in Chinese—China Defense Publishing House in 1988).
- [13] A. Farina, F.A. Studer, Data Processing, in: P. Bowron (Ed.), *Advanced Topics and Applications*, vol. 2, Researches Studies Press, John Wiley, England, 1986 (Translated in Chinese—China Defence Publishing House in 1992).
- [14] A. Farina, Target tracking with bearings-only measurements, *Signal Process.* 78 (1) (1999) 81–78.
- [15] A. Farina, R. Miglioli, Association between active and passive tracks for airborne sensors, *Signal Process.* 69 (3) (1998) 209–217.
- [16] A. Farina, B. La Scala, Methods for association of active and passive tracks for airborne sensors, in: Int. Symp. Radar, IRS98, Munich, Germany, September 15–17, 1998, pp. 735–744.
- [17] B. La Scala, A. Farina, Choosing a track association method, *Inform. Fusion* 3 (2) (2002) 119–133.
- [18] E. Waltz, J. Llinas, *Multisensor Data Fusion*, Artech House, Norwood, MA, 1990.
- [19] D. Hall, *Mathematical Techniques in Multisensory Data Fusion*, Artech House, Norwood, MA, 1992.
- [20] D.S. Alberts, J. Garstka, F.P. Stein, *Network Centric Warfare: Developing and Leveraging Information Superiority*, National Defence Press, USA, 1999.
- [21] A. Farina, A. Graziano, L. Ortenzi, E. Spinogatti, The role of modelling and simulation (M and S) in the analysis of integrated systems for homeland protection, in: G. Franceschetti, M. Grossi (Eds.), *Homeland Security Technology Challenges, From Sensing and Encrypting to Mining and Modelling*, Artech House, 2008 (Chapter 9).
- [22] C.J. Skinner, S. Cochrane, M. Field, R. Johnston, *Defence Against Terrorism: The Evolution of military Surveillance Systems into Effective Counter Terrorism Systems Suitable for Use in Combined Military Civil Environments, Dream or Reality?* NATO Panel Systems, Concepts and Integration (SCI) Methods and Technologies for Defence Against Terrorism, London, UK, October 25–27, 2004.
- [23] US Office of Homeland Security: National Strategy for Homeland Security, Washington, DC, July 2002. <http://www.whitehouse.gov/homeland/book/nat_strat_hls.pdf>.
- [24] U.S. Environmental Protection Agency: Strategic Plan for Homeland Security, Washington, DC, September 2002. <http://www.epa.gov/epahome/downloads/epa_homeland_security_strategic_plan.pdf>.
- [25] J. Motteff, C. Copeland, J. Fischer, Critical Infrastructure: What Makes an Infrastructure Critical? Report for Congress RL31556, The Library of Congress, August 2002. <www.fas.org/irp/crs/RL31556.pdf>.
- [26] US Government, The National Strategy for the Physical Protection of Critical Infrastructure and Key Assets, The White House, Washington, DC, February 2003.
- [27] S. Bologna, R. Setola, The need to improve local self-awareness in CIP/CIIP, in: Proceeding of 1st IEEE Int. Workshop on CIP (IWCIP 2005), Darmstadt, Germany, November 3–4, 2005, pp. 84–89.
- [28] S. Rinaldi, J. Peerenboom, T. Kelly, Identifying, understanding and analyzing critical infrastructure interdependencies, *IEEE Control Syst. Mag.* 21 (6) (2001) 11–25.
- [29] <<http://www.icc-ccs.org/piracy-reporting-centre>>.
- [30] Archive of International Maritime Bureau Piracy and Armed Robbery Reports, IMB Publications, 2008.
- [31] A. Farina, S. Giompapa, A. Graziano, Analysis of emerging phenomena in large complex systems, in: Franceschetti G, Grossi M (Eds.), *Homeland Security—Facets: Threats, Countermeasures, and the Privacy Issue*, Artech House, 2011 (Chapter 4).

- [32] R. Zimmerman, Decision making and the vulnerability of interdependent critical infrastructure, in: Proceeding of the IEEE Conference on Systems, Man and Cybernetics 2004, The Hague, The Netherlands, October 10–13, 2004, pp. 4059–4063.
- [33] R. Setola, S. Geretschuber, Critical information infrastructure security, in: Third International Workshop CRITIS, Frascati (Rome), Italy, October 2008, pp. 13–15.
- [34] S.T. Smith, A. Silberfarb, S. Philips, E.K. Kao, C. Anderson, Network discovery using wide-area surveillance data, in: Proceedings of IEEE Conference on Information Fusion 2011, Lexington, MA, USA, July 5–8, 2011, pp. 1–8.
- [35] A. Starr, M. Desforges, Strategies in data fusion—sorting through the tool box, in: Bedworth, O'Brien (Eds.), Proceedings of EuroFusion98 International Conference on Data Fusion, 1998, pp. 85–90.
- [36] A.N. Steinberg, C.L. Bowman, A systems engineering approach for implementing data fusion systems, in: D.L. Hall, J. Llinas (Eds.), Handbook of Multisensor Data Fusion, CRC Press, London, 2001.
- [37] F.E. White, A model for data fusion, in: Proceedings of the 1st National Symposium on Sensor Fusion, vol. 2, Chicago, USA, 1988, pp. 143–158.
- [38] A.N. Steinberg, C.L. Bowman, F.E. White, Revisions to the JDL data fusions models, in: Proceedings of the 3rd NATO/IRIS Conference, Quebec City, Canada, 1998.
- [39] E. Blasch, S. Plano, JDL Level 5 fusion model: user refinement issues and applications in group tracking, in: Proceedings of the SPIE Signal Processing Sensor Fusion Target Recognition XI, vol. 4729, 2002, pp. 270–279.
- [40] E. Blasch, S. Plano, Level 5: user refinement to aid the fusion process, in: Proceedings of the SPIE, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, vol. 5099, 2003, pp. 288–297.
- [41] E. Blasch, S. Plano, DFIG Level 5: user refinement issues supporting situational assessment reasoning, in: Proceedings of 8th International Conference on Information Fusion, Philadelphia, USA, July, 2005, pp. 25–29.
- [42] Subrata Das, High-Level Data Fusion, Artech House, 2008.
- [43] J. Rowley, The wisdom hierarchy: representations of the DIKW hierarchy, *J. Inform. Sci.* 33 (2) (2007) 163–180.
- [44] M.R. Endsley, Design and evaluation for situation awareness enhancement, in: Proceedings of the 32nd Annual Meeting of the Human Factors Society, 1988, pp. 97–101.
- [45] M.R. Endsley, Toward a theory of situation awareness in dynamic systems, *Human Factors* 37 (1) (1995) 32–64.
- [46] M.R. Endsley, Measurement of situation awareness in dynamic systems, *Human Factors* 37 (1) (1995) 65–84.
- [47] P.P. Perla, M. Markowitz, A.A. Nofi, C. Weuve, J. Loughran, M. Stahl, Gaming and shared situation awareness, Alexandria, Virginia, Centre for Naval Analyses, 2000.
- [48] J.R. Boyd, A Discourse on Winning and Losing, Unpublished Set of Briefing Slides, Air University Library, Maxwell AFB, AL, USA, May, 1987.
- [49] T. Bass, Intrusion detection systems and multisensor data fusion: creating cyberspace situational awareness, *Commun. ACM* 43 (4) (2000) 99–105.
- [50] M.D. Bedworth, J. O'Brien, The omnibus model: a new architecture for data fusion? *IEEE Trans. Aerosp. Electron. Syst.* 15 (4) (1999) 30–36.
- [51] J. Rasmussen, Skills, rules and knowledge: signals, signs and symbolism, and other distinctions in human performance models, *IEEE Trans. Syst. Man Cybern.* 12 (1983) 257–266.
- [52] J. Rasmussen, Information Processing and Human Machine Interaction: An Approach to Cognitive Engineering, North Holland, NY, 1986.
- [53] J. Reason, Human Error, Cambridge University Press, Cambridge, UK, 1990.

- [54] USAF (United States Air Force): Air Force Pamphlet 14–210, Intelligence, USAF Intelligence Targeting Guide, Department of Defense, USA, 1998.
- [55] ODNI (Office of Director of National Intelligence): How Do We Collect Intelligence? USA, 2008. <http://www.dni.gov/who_what/what_collection.htm>.
- [56] D. Krioukov, F. Papadopoulos, F. Vahdat, M. Boguna, Curvature and temperature of complex networks, *Phys. Rev. E* 80 (2009) 035101(R).
- [57] H.F. Durrant-Whyte, Sensor models and multi-sensor integration, *Int. J. Robot. Res.* 7 (6) (1988) 97–113.
- [58] W. Elmenreich, Sensor fusion in time-triggered systems, PhD Thesis, Institut für Technische Informatik, Vienna University of Technology, 2002.
- [59] S.M. Kay, *Statistical Signal Processing: Detection Theory*, Prentice Hall, 1998.
- [60] A. Farina, B. Ristic, L. Timmoneri, Cramer-Rao bound for nonlinear filtering with $P_d < 1$ and its application to target tracking, *IEEE Trans. Signal Process.* 50 (8) (2002) 1916–1924.
- [61] M. Hernandez, B. Ristic, A. Farina, L. Timmoneri, A comparison of two Cramer-Rao bounds for non-linear filtering with $P_d < 1$, *IEEE Trans. Signal Process.* 52 (9) (2004) 2361–2370.
- [62] A. Farina, B. Ristic, S. Immediata, L. Timmoneri, CRLB with $P_d < 1$ fused tracks, *Fusion 2005*, in: Proceedings of the 7th International Conference on Information Fusion, Philadelphia, USA, July 25–28, 2005, pp. 191–196.
- [63] T.J. Nohara, P. Weber, G. Jones, A. Ukrainec, A. Premji, Affordable high-performance radar networks for homeland security applications, *RADAR '08*, in: Proceedings of IEEE International Radar Conference, Rome, Italy, May 26–30, 2008, pp. 1–6.
- [64] T.J. Nohara, P. Weber, A. Premji, C. Krasnor, S. Gauthereaux, M. Brand, G. Key, Affordable avian radar surveillance systems for natural resource management and BASH applications, *RADAR 2005*, in: Proceedings of IEEE International Radar Conference, Arlington, Virginia, USA, May 9–12, 2005, pp. 10–15.
- [65] Proceedings of 2nd Workshop on Cognitive Information Processing, Elba Island, Tuscany, Italy, June 14–16, 2010.
- [66] Masazade, R. Niu, P.K. Varshney, M. Keskinoz, Energy aware iterative source localization for wireless sensor networks, *IEEE Trans. Signal Process.* 58(9) (2010) 4824–4835.
- [67] Feng Zhao, Jaewon Shin, J. Reich, Information-driven dynamic sensor collaboration, *IEEE Signal Process. Mag.* 19 (2) (2002) 61–72.
- [68] C.M. Kreucher, A.O. Hero, K.D. Kastella, M.R. Morelande, An information-based approach to sensor management in large dynamic networks, *Proc. IEEE* 95 (5) (2007) 978–999.
- [69] J.L. Williams, J.W. Fisher, A.S. Willsky, Approximate dynamic programming for communication-constrained sensor network management, *IEEE Trans. Signal Process.* 55 (8) (2007) 4300–4311.
- [70] L. Zuo, R. Niu, P.K. Varshney, Conditional posterior Cramér–Rao lower bounds for nonlinear sequential Bayesian estimation, *IEEE Trans. Signal Process.* 59 (1) (2001) 1–14.
- [71] L. Zuo, R. Niu, P.K. Varshney, Posterior CRLB based sensor selection for target tracking in sensor networks, *ICASSP 2007*, in: IEEE International Conference on Acoustic Speech and Signal Processing, vol. 2, Honolulu, Hawaii, April 15–20, 2007, pp. II-1041–II-1044.
- [72] L. Zuo, R. Niu, P.K. Varshney, A sensor selection approach for target tracking in sensor networks with quantized measurements, in: IEEE International Conference on Acoustic Speech and Signal Processing, *ICASSP 2008*, Las Vegas, Nevada, USA, March 30–April 4, 2008, pp. 2521–2524.
- [73] M.L. Hernandez, T. Kirubarajan, Y. Bar-Shalom, Multi-sensor resource deployment using posterior Cramer–Rao bounds, *IEEE Trans. Aerosp. Electron. Syst.* 40 (2) (2004) 399–416.
- [74] T.C. Collier, C. Taylor, Self-organization in sensor networks, *J. Parall. Distr. Comput.* 64 (7) (2004) 866–873.
- [75] F.S. Cattivelli, A.H. Sayed, Modelling bird flight formations using diffusion adaptation, *IEEE Trans. Signal Process.* 59 (5) (2011) 2038–2051.

- [76] F.S. Cattivelli, A.H. Sayed, Distributed detection over adaptive networks using diffusion adaptation, *IEEE Trans. Signal Process.* 59 (5) (2011) 1917–1932.
- [77] C.G. Lopes, A.H. Sayed, Diffusion least-mean squares over adaptive networks: formulation and performance analysis, *IEEE Trans. Signal Process.* 56 (7) (2008) 3122–3136.
- [78] F.S. Cattivelli, A.H. Sayed, Diffusion LMS strategies for distributed estimation, *IEEE Trans. Signal Process.* 58 (3) (2010) 1035–1048.
- [79] F.S. Cattivelli, C.G. Lopes, A.H. Sayed, Diffusion recursive least-squares for distributed estimation over adaptive networks, *IEEE Trans. Signal Process.* 56 (5) (2008) 1865–1877.
- [80] A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization – A Universal Concept in Non Linear Sciences*, Cambridge University Press, Cambridge, UK, 2001.
- [81] N.A. Lynch, *Distributed Algorithms*, Morgan-Kaufmann, San Mateo, CA, 1997.
- [82] R. Olfati-Saber, J.S. Shamma, Consensus filters for sensor networks and distributed sensor fusion, in: Proceeding of the Joint IEEE Conference On Decision and Control and the European Control Conf., Seville, Spain, December 15, 2005, pp. 6698–6703.
- [83] Y.W. Hong, L.F. Cheow, A. Scaglione, A simple method to reach detection consensus in massively distributed sensor networks, in: Proceedings of International Symposium on Information Technology, ISIT 2004, Chicago, USA, July 2004, p. 250.
- [84] D. Lucarelli, I.J. Wang, Decentralized synchronic protocols with nearest neighbor communication, in: Proceedings of the 2nd International Conference on Embedded Networked Sensor System, SenSys 2004, Baltimore, USA, November 3–5, 2004, pp. 62–68.
- [85] S. Barbarossa, Self-organizing sensor networks with information propagation based on mutual coupling of dynamic systems, in: Proceedings of International Workshop on Ad-hoc Wireless Network 2005, IWAN 2005, London, UK, May 23–26, 2005.
- [86] S. Barbarossa, G. Scutari, Decentralized maximum likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems, *IEEE Trans. Signal Process.* 55 (7-I) (2007) 3456–3470.
- [87] G. Scutari, S. Barbarossa, L. Pescosolido, Distributed decision through self-synchronizing sensor networks in presence of propagation delays and asymmetric channels, *IEEE Trans. Signal Process.* 56 (4) (2008) 1667–1684.
- [88] S. Barbarossa, G. Scutari, A. Swami, Achieving consensus in self-organizing wireless sensor networks, the impact of network topology on energy consumption, in: Proceedings of ICASSP 2007, vol. 2, Honolulu, Hawaii, April 15–20, 2007, pp. 841–844.
- [89] S. Barbarossa, G. Scutari, Bio-inspired sensor network design, *IEEE Signal Process. Mag.* 24 (3) (2007) 26–35.
- [90] S. Fortunati, F. Gini, M.S. Greco, A. Farina, A. Graziano, S. Giomppapa, Least squares estimation and Cramér-Rao type lower bounds for relative sensor registration process, *IEEE Trans. Signal Process.* 59 (3) (2011) 1075–1087.
- [91] Yifeng Zhou, H. Leung, P.C. Yip, An exact maximum likelihood registration algorithm for data fusion, *IEEE Trans. Signal Process.* 45 (6) (1997) 1560–1573.
- [92] Zhenhua Li, Siyue Chen, H. Leung, E. Bosse, Joint data association, registration, and fusion using EM-KF, *IEEE Trans. Aerosp. Electron. Syst.* 46 (2) (2010) 496–507.
- [93] Siyue Chen, H. Leung, E. Bosse, A maximum likelihood approach to joint registration, association and fusion for multi-sensor multi-target tracking, in: Proceedings of 12th International Conference on Information Fusion, FUSION 2009, Seattle, USA, July 6–9, 2009, pp. 686–693.
- [94] S. Herzel, M.C. Recchioni, F. Zirilli, A quadratically convergent method for linear programming, *Linear Algebra Appl.* 152 (1991) 255–289.
- [95] N. Karmarkar, A new polynomial time algorithm in linear programming, *Combinatorica* 4 (1984) 373–395.
- [96] L. Hageman, D. Young, *Applied Iterative Methods*, Academic Press, New York, 1981.

- [97] A. Farina, A. Graziano, F. Mariani, F. Zirilli, A cooperative sensor network: optimal deployment and functioning, *RAIRO Oper. Res.* 44 (4) (2009) 379–388 (special issue COGIS).
- [98] F. Aurenhammer, Voronoi diagrams—a survey of a fundamental geometric data structure, *ACM Comput. Surv.* 23 (3) (1991) 345–405.
- [99] J.A. Snyman, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, Springer, Cambridge, Massachusetts, USA, 2005.
- [100] F.R. Chung, *Spectral Graph Theory*, American Mathematical Society, Providence, USA, CBMS 92, 1997.
- [101] M.H. DeGroot, M.J. Schervish, *Probability and Statistics*, Pearson Addison Wesley, New York, 2002.
- [102] C.S. Raghavendra, K.M. Sivalingam, Taieb Znati, *Wireless Sensor Networks*, Springer, USA, 2005.
- [103] E. Ertin, J.W. Fisher, L.C. Potter, Maximum mutual information principle for dynamic sensor query problems, *Lect. Notes Comput. Sci. Inform. Process. Sensor Networks* 2634 (2003) 91–104.
- [104] F. Zhao, J. Shin, J. Reich, Information-driven dynamic sensor collaboration for tracking applications, *IEEE Signal Process. Mag.* 19 (2) (2002) 61–72.
- [105] J. Mathieu, G. Hwang, J. Dunyak, *The State of the Art and the State of the Practice: Transferring Insights from Complex Biological Systems to the Exploitation of Netted Sensors in Command and Control Enterprises*, MITRE Technical Papers, MITRE Corporation, USA, July 2006.
- [106] A. Khelil, C. Becker, J. Tian, K. Rothermel, An epidemic model for information diffusion in MANETs, in: *Proceedings of the 5th ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems*, MSWiM 2002, Atlanta, Georgia, USA, September 28, 2002, pp. 54–60.
- [107] P. De, Y. Liu, S.K. Das, An epidemic theoretic framework for evaluating broadcast protocols in wireless sensor networks, , in: *Proceedings of the 4th IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, MASS 2007, Pisa, Italy, October 8–11, 2007, pp. 1–9.
- [108] B. Ristic, A. Skvortsov, M. Morelande, Predicting the progress and the peak of an epidemics, in: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 2009*, ICASSP 2009, Taiwan, April 19–24, 2009, pp. 513–516.
- [109] A. Dekker, A. Skvortsov, Topological issues in sensor networks. in: *MODSIM 2009: 2009 MSSANZ International Congress on Modelling and Simulation*, Cairns, Australia, July 13–17, 2009, pp. 952–958.
- [110] S. Eubank, Anil Kumar VS, Marathe M: Epidemiology and wireless communication: tight analogy or loose metaphor? *Lect. Notes Comput. Sci. Bio-Inspir. Comput. Commun.* 5151 (2008) 91–104.
- [111] V. Bisignanesi, M.S. Borgas, Models for integrated pest management with chemicals in atmospheric surface layers, *Ecol. Modell.* 201 (1) (2007) 2–10.
- [112] A. Skvortsov, B. Ristic, M. Morelande, Networks of chemical sensors: a simple mathematical model for optimisation study, in: *5th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, ISSNIP 2009, Melbourne, Australia, December 7–10, 2009, pp. 385–390.
- [113] A. Gunatilaka, B. Ristic, A. Skvortsov, M. Morelande, Parameter estimation of a continuous chemical plume source, in: *IEEE 11th International Conference on Information Fusion*, FUSION 2008, Cologne, Germany, June 30–July 3, 2008, pp. 1–8.
- [114] J.D. Murray, *Mathematical Biology*, vols. 8–12, Springer, USA, 2002.
- [115] P.F. Verhulst, Recherches mathématiques sur la loi d'accroissement de la population, *Nouv. mém. de l'Academie Royale des Sci. et Belles-Lettres de Bruxelles* 18 (1845) 1–41.
- [116] Weisstein EW, Logistic Equation, *MathWorld-A Wolfram Web Resource*. <<http://mathworld.wolfram.com/LogisticEquation.html>>.
- [117] C. Mendis, A. Gunatilaka, A. Skvortsov, S. Karunasekera, The effect of correlation of chemical tracers on chemical sensor network performance, in: *Proceedings of 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, ISSNIP 2010, Brisbane, Australia, December 7–10, 2010, pp. 103–108.

- [118] A. Skvortsov, B. Ristic, Modelling and performance analysis of a network of chemical sensors with dynamic collaboration, *Int. J. Distr. Sensor Networks* 2012 (2012) (article ID65231).
- [119] S. Karunasekera, C. Mendis, A. Skvortsov, A. Gunatilaka, A decentralized dynamic sensor activation protocol for chemical sensor networks, in: Proceedings of the 9th IEEE Symposium on Network Computing and Applications, Cambridge, MA, USA, July 15–17, 2010, pp. 218–223.
- [120] S. Karunasekera, J. Beaton, A. Dimech, A. Skvortsov, A. Gunatilaka, A distributed e-research tool for evaluating source backtracking algorithms, in: Proceedings of IEEE 6th International Conference on e-Science, Brisbane, Australia, December 7–10, 2010, pp. 17–24.
- [121] R.J. Nemzek, J.S. Dreicer, D.C. Torney, T.T. Warnock, Distributed sensor networks for detection of mobile radioactive sources, *IEEE Trans. Nucl. Sci.* 51 (4) (2004) 1693–1700.
- [122] S.M. Brennan, A.M. Mielke, D.C. Torney, Radioactive source detection by sensor networks, *IEEE Trans. Nucl. Sci.* 52 (3) (2005) 813–819.
- [123] A.V. Klimenko, W.C. Priedhorsky, N.W. Hengartner, K.N. Borozin, Efficient strategies for low-level nuclear searches, *IEEE Trans. Nucl. Sci.* 53 (3) (2006) 1435–1442.
- [124] A. Sundaresan, P.K. Varshney, N.S.V. Rao, Distributed detection of a nuclear radioactive source using fusion of correlated decisions, in: Proceedings of IEEE International Conference on Information Fusion, FUSION 2007, Quebec, Canada, July 9–12, 2007, pp. 1–8.
- [125] M.R. Morelande, B. Ristic, Radiological source detection and localisation using Bayesian techniques, *IEEE Trans. Signal Process.* 57 (11) (2009) 4220–4231.
- [126] N. Tsoulfanidis, Measurement and Detection of Radiation, Taylor and Francis, Washington, DC, 1995.
- [127] R.A. Cortez, X. Papageorgiou, H.G. Tanner, A.V. Klimenko, K.N. Borozdin, R. Limia, W.C. Priedhorsky, Smart radiation sensor management, *IEEE Robot. Autom. Mag.* 15 (3) (2008) 85–93.
- [128] B. Ristic, M.R. Morelande, A. Gunatilaka, Information driven search for point sources of gamma radiation, *Signal Process.* 90 (4) (2010) 1225–1239.
- [129] A. Martin, S.A. Harbison, An Introduction to Radiation Protection, Chapman and Hall, 1987.
- [130] AN/PDR-77 User's Guide, Canberra Industries Inc., CT, USA.
- [131] A. Gunatilaka, B. Ristic, LCAARS radiological field trial and validation of source localisation algorithms, DSTO Tech. Report, DSTO-TR-1988, 2009.
- [132] A. Gunatilaka, B. Ristic, M.R. Morelande, Experimental verification of algorithms for detection and estimation of radioactive sources, in: Proceedings of 13th International Conference on Information Fusion, FUSION 2010, Edinburgh, UK, July 26–29, 2010, pp. 1–8.
- [133] C. Musso, N. Oudjane, F. LeGland, Improving regularised particle filters, in: N. Doucet, N. DeFreitas, N.J. Gordon (Eds.), Sequential Monte Carlo methods in Practice, Springer-Verlag, New York, 2001.
- [134] B. Ristic, S. Arulampalam, Beyond the Kalman Filter, Artech House, Gordon N, 2004.
- [135] H.L. Van Trees, Detection, Estimation, and Modulation Theory (Part I), John Wiley and Sons, 1968.
- [136] E. Yee, A. Gunatilaka, Comparison of two approaches for detection and estimation of radioactive sources, *Applied Mathematics*, Ristic B, 2011.
- [137] R. Mahler, Statistical Multi-Source Multi-Target Information Fusion, Artech House, 2007.
- [138] M.R. Morelande, A. Skvortsov, Radiation field estimation using a Gaussian mixture, in: Proceedings of 12th International Conference on Information Fusion, FUSION 2009, Seattle, USA, July 6–9, 2009, pp. 2247–2254.
- [139] D. Cacciani, F. Garzia, A. Neri, R. Cusani, Optimal territorial placement for multi-purpose wireless service using genetic algorithms, *Wireless Eng. Technol.* 2 (3) (2011) 184–195.
- [140] D.E. Goldberg, Genetic Algorithms, Addison-Wesley, 1989.
- [141] IEEE Signal Processing Magazine, Special Issue on Collaborative Processing, IEEE Press, March 2002.

- [142] R. Niu, P. Varshney, M. Moore, D. Klamer, Decision fusion in a wireless sensor network with a large number of sensors, in: Proceedings of 7th IEEE International Conference on Information Fusion, FUSION 2004, Stockholm, Sweden, June 28–July 1, 2004.
- [143] A. Farina, G. Golino, A. Capponi, C. Pilotto, Surveillance by means of a random sensor network: a heterogeneous sensor approach, in: Proceedings of the 8th IEEE International Conference on Information Fusion, FUSION 2005, vol. 2, Philadelphia, USA, July, 2005, pp. 25–28.
- [144] R. Diestel, Graph Theory, Graduate Texts in Mathematics, Springer, 1997.
- [145] B. Bollobas, Modern Graph Theory, Graduate Texts in Mathematics, vol. 184, Springer, 1998.
- [146] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 1999.
- [147] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publications de l’Institut de Statistique de l’Université de Paris 8 (1959) 229–231.
- [148] S.G. Iyengar, Varshney PK, T. Darmala, A parametric copula-based framework for hypothesis testing using heterogeneous data, IEEE Trans. Signal Process. 59 (5) (2001) 2308–2319.
- [149] R.N. Charette, The virual fence’s long good-bye, IEEE Spectr. (2011).
- [150] J.B. Billingsley, Low-Angle Radar Land Clutter, SciTech. Pub. Inc., New York, 2001.
- [151] B.T. Binder, M.F. Toups, S. Ayasli, E.M. Adams, SAR Foliage Penetration Phenomenology of Tropical Rain Forest and Northern US Forest, in: Proceeding of IEEE International Radar Conference 1995, Washington DC, USA, May 8–11, 1995, pp. 158–163.
- [152] Y. Lu, Y. Cheng, W. Liu, H.W. Seah, H.L. Chan, L.C. Tai’, M. Lesturgie, P. Borderies, R. Guern, Low frequency radar phenomenology study in equatorial vegetation – preliminary results, in: Proceeding of IEEE Radar Conference 2002, October 15–17, 2002, pp. 70.
- [153] M.E. Davis, Developments in foliage penetration radar, in: Proceeding of IEEE Radar Conference 2010, Washington DC, USA, May 10–14, 2010, pp. 1233.
- [154] M.N. Roy, S. Swamp, S.K. Tewari, Radio wave propagation through rain forests of India, IEEE Trans. Antennas Propag. 38 (4) (April 1990) 433–449.
- [155] M.E. Davis, Foliage Penetration Radar: Detection and Characterization of Objects Under Trees, SciTech. Pub. Inc., New York, 2011.
- [156] S. Gallone, FOPEN radar for UGS applications, in: Proceeding of CIE International Conference on Radar, Chengdu, China, October 24–27, 2011.
- [157] <<http://www.cosmo-skymed.it/>>.
- [158] M. Costantini, A. Farina, F. Zirilli, The fusion of different resolution radar images: a new formalism invited paper, Proc. IEEE 85 (1) (1997) 139–146 (special issue on Data Fusion).
- [159] A.M. Signorini, A. Farina, G. Zappa, Application of multi-scale estimation algorithm to SAR images fusion, in: Proceedings of International Symposium on Radar, IRS98, Munich, Germany, September 15–17, 1998, pp. 1341–1352.
- [160] G. Simone, C. Morabito, A. Farina, Radar image fusion by multiscale Kalman filter, in: Proceedings of 3rd IEEE International Conference on Fusion, FUSION 2000, Paris, France, July 10–13, 2000, pp. WeD3. 10–WeD3.17.
- [161] G. Simone, C. Morabito, A. Farina, Multifrequency and multiresolution fusion of SAR images for remote sensing applications, Fusion 2001, Montreal, Canada, August 2001, pp. 7–10.
- [162] G. Simone, A. Farina, F.C. Morabito, S.B. Serpico, L. Bruzzone, Image fusion techniques for remote sensing applications, Inform. Fusion 3 (1) (2002) 3–15.
- [163] US Department of Defence Dictionary of Military Terms, USA, 2010. <http://www.dtic.mil/doctrine/dod_dictionary/>
- [164] Y. Bar-Shalom, X.R. Li, Estimation and Tracking: Principles, Techniques, and Software, Artech House, 1993.
- [165] S.S. Blackman, R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House, 1999.

- [166] Y. Bar-Shalom, X.R. Li, Kirubarajan T, *Estimation with Applications to Tracking and Navigation*, John Wiley and Sons, 2004.
- [167] A. Graziano, A. Farina, R. Miglioli, deFeo M: IMMJPDA vs., MHT and Kalman filter with NN correlation performance comparison, *IEE Proc. Radar Sonar Naviga.* 144 (2) (1997) 49–56.
- [168] W.G. Bath, Use of multiple hypothesis in radar tracking, Presented at Radar '92, John Hopkins Applied Physics Laboratory, USA, pp. 90–93.
- [169] M. Hernandez, A. Benavoli, A. Graziano, A. Farina, F. Morelande, Performance measure and MHT for tracking move-stop-move targets with MTI sensors, *IEEE Trans. Aerosp. Electron. Syst.* 47 (1) (2011) 996–1025.
- [170] A. Farina, G. Golino, L. Ferrante, Constrained tracking filters for A-SMGCS, in: Proceedings of 6th International Conference on Information Fusion, FUSION 2003, Cairns, Queensland, Australia, July 8–11, 2003, pp. 414–421.
- [171] G. Golino, A. Farina, Track-plot correlation in A-SMGCS using the target images by a surface movement radar, in: Proceedings of 7th International Conference on Information Fusion, FUSION 2004, Stockholm, Sweden, June 28–July 1, 2004, pp. 999–1005.
- [172] A. Di Lallo, A. Farina, L. Timmoneri, T. Volpi, Bi-dimensional analysis of simulated herm (helicopter rotor modulation) and jem (jet engine modulation) radar signals for target recognition, in: Proceeding of 1st International Conference on Waveform Diversity and Design, WDDC 2004, Edinburgh, UK, November 8–10, 2004.
- [173] S. Palumbo, S. Barbarossa, A. Farina, M.R. Toma, Classification techniques of radar signals backscattered by helicopter blades, in: Proceeding of International Symposium on Digital Signal Processing, ISDSP96, London, UK, July 23–24, 1996, pp. 44–50.
- [174] (a) A. Farina, F. Gini, A matched subspace approach to CFAR detection of hovering helicopters, in: International Symposium on Radar, IRS98, Munich, Germany, September 15–17, 1998, pp. 597–606; (b) F. Gini, A. Farina, Matched subspace CFAR detection of hovering helicopters, *IEEE Trans. Aerosp. Electron. Syst.* 35 (4) 1293–1305.
- [175] M.N. Saidi, K. Daoudi, A. Khenchaf, B. Hoeltzener, D. Aboutajdine, Automatic target recognition of aircraft models based on ISAR images, in: Proceeding of IEEE International, IGARSS 2009, Cape Town, South Africa, July 12–17, 2009, pp. IV-685–IV-688.
- [176] D. Pastina, C. Spina, Multi-feature based automatic recognition of ship targets in ISAR images, in: Proceeding of IEEE International Radar Conference, Radar08, Rome, Italy, May 26–30, 2008, pp. 1–6.
- [177] D. Pastina, C. Spina, Multi-frame data fusion techniques for ATR of ship targets from multiple ISAR images, Proceeding of European Radar Conference, EuRAD 2009, Rome, Italy, September 30–October 2, 2009, pp. 409–412.
- [178] S. Giompapa, A. Farina, F. Gini, A. Graziano, R. Croci, R. Di Stefano, Naval target classification based on the confusion matrix, in: Proceeding of IEEE Aerospace Conference 2008, Big Sky, Montana, USA, March 1–8, 2008, pp. 1–9.
- [179] M. Wagner, A. Moore, Aryel R, *Handbook of Biosurveillance*, Elsevier, 2006.
- [180] J. Walden, E.H. Kaplan, Estimating time and size of bioterror attack, *Emerg. Infect. Dis.* 10 (7) (2004) 1202–1205.
- [181] A.G. Wilson, G.D. Wilson, D.H. Olwell, *Statistical Methods in Counterterrorism: Aame Theory, Modelling, Syndromic Surveillance, and Biometric Authentication*, Springer, 2006.
- [182] Eysenbach, Infodemiology: tracking flu-related searches on the web for syndromic surveillance, in: Proceedings of Symposium of American College Medical Informatics Association 2006, AMIA 2006, Washington, DC, USA, November 11–15, 2006, pp. 244–248.
- [183] N.M. Schuster, M.A. Rogers, Using search engine query data to track pharmaceutical utilization: a study of statin, *Am. J. Manag. Care* 16 (8) (2010) e215–e219.

- [184] A. Culotta, Detecting influenza outbreaks by analyzing twitter messages, in: Proceedings of Conference on Knowledge Discovery and Data Mining 2010, Washington, DC, USA, July 2010, pp. 25–28.
- [185] J. Ginsberg, M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, L. Brilliant, Detecting influenza epidemics using search engine query data, *Nature* 457 (2009) 1012–1015.
- [186] B. Cazelles, N.P. Chau, Using the Kalman filter and dynamic models to assess the changing HIV/AIDS epidemic, *Math. Biosci.* 140 (2) (1997) 131–154.
- [187] J. Mandela, J.D. Beezleya, L. Cobba, A. Krishnamurthy, Data driven computing by the morphing fast Fourier transform ensemble Kalman filter in epidemic spread simulations, *Proc. Comput. Sci.* 1 (1) (2010) 1221–1229.
- [188] C. Jégat, F. Carrat, C. Lajaunie, H. Wackernagel, Early detection and assessment of epidemics by particle filtering, in: A. Soares, M.J. Pereira, R. Dimitrakopoulos (Eds.), geoENV VI Geostatistics for Environmental Applications, vol. 15, Springer, 2008, pp. 23–35.
- [189] J.B.S. Ong, M.I.C. Chen, A.R. Cook, H.C. Lee, V.J. Lee, R.T.P. Lin, P.A. Tambyah, L.G. Goh, Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in Singapore, *PLoS ONE* 5 (4) (2010) e10036.
- [190] B. Ristic, A. Skvortsov, M. Morelande, Predicting the progress and the peak of an epidemic, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009, Taipei, Taiwan, April 19–24, 2009, pp. 513–516.
- [191] A. Skvortsov, B. Ristic, C. Woodruff, Predicting an epidemic based on syndromic surveillance, in: Proceedings of 13th International Conference on Information Fusion, FUSION 2010, Edinburgh, UK, July 26–29, 2010, pp. 1–8.
- [192] P.D. Stroud, S.J. Sydoriak, J.M. Riese, J.P. Smith, S.M. Mniszewski, P.R. Romero, Semiempirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing, *Math. Biosci.* 203 (2006) 301–318.
- [193] A.S. Novozhilov, On the spread of epidemics in a closed heterogeneous population, *Math. Biosci.* 215 (2) (2008) 177–185.
- [194] R.M. Anderson, R.M. May, Population biology of infectious diseases: Part 1, *Nature* 280 (1979) 361–367.
- [195] D.J. Daley, J. Gani, *Epidemic Modelling*, Cambridge University Press, 1996.
- [196] C.E. Dangerfield, J.V. Ross, M.J. Keeling, Integrating stochasticity and network structure into an epidemic model, *J. Roy. Soc. Interf.* 6 (38) (2009) 761–774.
- [197] O.A. van Herwaarden, J. Grasman, Stochastic epidemics: major outbreaks and the duration of the endemic period, *J. Math. Biol.* 33 (4) (1995) 581–601.
- [198] A. Skvortsov, B. Ristic, Monitoring and prediction of an epidemic outbreak using syndromic observations, *Math. Biosci.* 240 (2012) 12–19.
- [199] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [200] B. Ristic, Bayesian estimation with imprecise likelihoods: random set approach, *IEEE Signal Process. Lett.* 18 (7) (2011) 395–398.
- [201] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Process.* 50 (2) (2002) 174–188.
- [202] A. Skvortsov, R. Connell, P. Dawson, R. Gailis, Epidemic modelling: validation of agent based simulation by using simple mathematical models, in: International Congress on Modelling and Simulation, MODSIM 2007, Christchurch, New Zealand, December 10–13, 2007, pp. 657–662.
- [203] A. Skvortsov, R. Connell, P. Dawson, R. Gailis, Epidemic spread modelling: alignment of agent-based simulation with a simple mathematical model, in: Proceedings of International Conference on Bioinformatics and Computational Biology, BIOCOMP 2007, Las Vegas, USA, June 25–28, 2007, pp. 487–890.
- [204] N. Oudjane, C. Musso, Progressive correction for regularized particle filters, in: Proceedings of 3rd International Conference on Information Fusion, FUSION 2000, Paris, France, vol. 2, July 10–13, 2000, pp. WEB5.26–WEB5.33.

Index

A

ACMA. *See* Algebraic Constant Modulus Algorithms (ACMA)
Across-track interferometry, 1044
Across-track SAR interferometry
statistical characterization of, 1055
surface topography, measuring, 1054
Across-track velocity
parameter estimation, 957
SAR imaging, 936
Active constellation extension (ACE), 242–243
Adaptive beamformer pattern synthesis, 624
Adaptive channel estimation, 118–119
Adaptive filter implementations
reduced-dimension STAP (RD-STAP), 630
post-Doppler STAP, 632
pre-Doppler STAP, 634
reduced-rank STAP, 622
adaptive beamformer pattern synthesis, 624
cross spectral metric, 627
diagonal loading, 626
eigencanceler, 625
Hung Turner Projection (HTP), 626
Multi-Stage Wiener Filter (MWF), 629
principle components inverse (PCI), 624
Adaptive precoding, 314
Adaptive SINR loss, 613
Adaptive SISO nonlinear Kalman equalizer, 122
ADC. *See* Analog to digital conversion (ADC)
Additive colored noise case, 730
Additive Gaussian colored noise (AGCN), 732
Additive noise, 182
Additive white Gaussian noise (AWGN), 190–191, 414–415
Advanced polarimetric decomposition techniques, 1162
adaptive model-based decomposition, 1168
2-component decomposition, improved volume
description, 1165
efficient eigenvector based decomposition, 1171
Yamaguchi 4-component decomposition, 1162
Airborne Early Warning (AEW) systems, 1246–1247
Algebraic Constant Modulus Algorithms (ACMA), 138–140
Algebraic graph theory, 85, 343–344, 400
Algorithms, for optimal resource allocation, 425
Along-track acceleration
SAR imaging, 940
Along-track interferometry (ATI), 943–948, 1044–1045
Along-track velocity
parameter estimation, 958–959
SAR imaging, 934–936

Alternating direction method of multipliers (ADMM), 354
Alternating projection timing estimator (APTE), 76–77
Ambiguity function (AF), 668, 675
Analog to digital conversion (ADC), 9
APC algorithm. *See* Autonomous power control (APC) algorithm
a-priori distribution, 12
Arbitrary networks, optimal topology and power
allocation for, 388
Architectures, multitarget multisensor tracking, 783
AR models. *See* Autoregressive (AR) models
ASKF. *See* Augmented state Kalman filter (ASKF)
Attenuation, 846, 1294
Augmented state Kalman filter (ASKF), 786
Autonomous power control (APC) algorithm, 426
Autoregressive (AR) models, 98, 140–141, 564
AWGN channel. *See* Additive white Gaussian noise (AWGN)

B

Backscatter, 1294
Band occupancy prediction, 487
Band occupancy scheduling, 488
Barankin Bound (BB), 28–29
Basis expansion models (BEMs), 99, 126
Bayesian and Neyman-Pearson (NP) formulations, 380–381
Bayesian approach, 358–359, 761
BB. *See* Barankin Bound (BB)
Beaufort Wind Scale, 528
BEMs. *See* Basis expansion models (BEMs)
Bennett Equality, 144
BER. *See* Bit-error-rate (BER)
BICM. *See* Bit-Interleaved Coded Modulation (BICM)
Bio-inspired approaches, 85
Bistatic ambiguity functions, 825
Bistatic challenges, 649
Bistatic ISAR, 1031
cross-range image formation, 1035
PSF of, 1033
range compression, 1034
signal modeling, 1032
Bistatic radar, 816
geometry, 816
radar equation, 817
Bistatic radar clutter, 821, 848
characterization of, 848
geometry, 822
Bistatic radar imaging, 1030
bistatic ISAR, 1031–1032

- Bit-error-rate (BER), 14–15, 192
 Bit-Interleaved Coded Modulation (BICM), 121, 192, 449–450
 Blind channel estimation, 104
 combined channel and symbol estimation, 104
 methods of moments, 106
 Blind signal separation, 5
 Blind Source Separation (BSS)
 additive noise, 182
 for communication signals, 136
 convulsive mixtures, 171, 173
 generalities on, 135
 instantaneous mixtures, 148
 non-circular sources, case of, 182
 non-stationarity, exploiting the, 182
 organisation of paper, 138
 signals, 141
 simulation, 180
 Block-adaptive channel estimation, 118
 Border control problem via electronic fence, 1290
 electronic fence, 1291–1292
 multi-scale approach, 1290
 Bragg scatterers, 522
 Bragg scattering and long waves, 556
 Broadband multipath interference, 734
 Broadcast consensus based phase synchronization, 703
 BSS. *See* Blind Source Separation (BSS)
- C**
- Canonical scatterers, 1131–1132
 Canonical scattering mechanisms, 1131
 Carrier frequency offset (CFO), 25, 38, 57, 250, 252
 Cartesian polarization basis, canonical scattering
 mechanisms, 1131
 Cartography-enabled route optimization, 493
 CAS. *See* Subband carrier allocation scheme (CAS)
 Cauchy-Schwarz theorem, 733
 CCRB derivations, 45
 CE-BEM. *See* Complex exponential basis expansion
 model (CE-BEM)
 Cellular networks, 487
 Central fusion center (CFC), 767
 Centralized data fusion process model, 1267
 Centralized tracking, 784
 Central limit theorem (CLT), 234, 518, 1137
 CFC. *See* Central fusion center (CFC)
 CFO. *See* Carrier frequency offset (CFO)
 CG map construction, 482
 Change detection based MTI approach, 888
 change detection, 891
 experimental results, 893
 signal model, 890
 Channel estimation
 blind channel estimation, 104
 MIMO channel estimation, 108
 semi-blind approaches, 107
 superimposed training-based approaches, 107
 training-based channel estimation, 103
 Channel gain cartography, 480
 CG map construction, 482
 coverage region estimation, 485
 spatio-temporal channel correlation, 482
 Channel gain map, 481
 Channel impulse response (CIR), 53, 200–201
 Channel matrix structure, 301
 Channel models
 MIMO channels, 102
 time-invariant channels, 100
 time-variant (doubly selective) channels, 96
 Channel state information (CSI), 116, 278, 442
 Chemical sensors networks, dynamic collaboration, 1275
 Circular polarization, 1124
 Circular polarization basis, canonical scattering
 mechanisms, 1131
 Circular-polarization scattering matrix, 525–526
 Clairvoyant SINR loss, 612
 Classical dual-channel techniques, 943
 along-track interferometry (ATI), 943–944
 displaced phase center antenna (DPCA) technique, 948, 950
 Clock characteristics, comparions of, 82
 Clock drift, 81–82
 Clutter
 discrete clutter, 617
 distributed clutter, 615
 Clutter diversity, 823
 Clutter modeling, 505–506, 764
 amplitude statistics, 517
 clutter volume reflectivity, 517
 discrete scatterers, 526
 Doppler spectrum, 522
 normalized clutter reflectivity, 515
 polarization characteristics, 524
 spatial correlation, 526
 Clutter suppression, GMTI, 952
 adaptive clutter suppression, 953
 DPCA technique, 952
 Clutter volume reflectivity, 517
 Coarse synchronization methods, 58
 Code-aided synchronization, 46
 Code-division multiple-access (CDMA) receiver, 195–196
 Coded-OFDM, 217–218
 Coding gain, 221–222
 Cognitive radio applications, 276
 Cognitive radio networks (CRN), 454–456, 474

- Cognitive radio sensing and cross-layer design, 489
 cartography-enabled route optimization, 493
 joint sensing and resource optimization, 490
 real-time traffic, sequential sensing for, 492
 throughput-aware sequential sensing, 490
- Coherency and covariance matrices, 1139
- Coherent ambiguity function, 675
- Coherent and non stationary pixels (yellow class), 1211
- Coherent and stationary pixels (white class), 1211
- Coherent decomposition techniques, 1143
- Coherent MIMO radar, 669
- Coherent processing, 672–673
- Coherent time-frequency characterization, 1187
 SAR image spectral content, 1188–1190
 time frequency decomposition, 1190–1191
- Collaborative signal and information processing (CSIP), 1263
- Column-wise diagonally dominant (CWDD), 301–302
- Communication network, 1292
- Communication networks, resource allocation, 409
- Competitive sensor data fusion, 1261
- Complementary sensor data fusion, 1261
- Complex exponential basis expansion model (CE-BEM), 99
- Complex scattering coefficients, 1129
- Compound-Gaussian model, 505–506, 519
- Compressed sensing approach, 476
- Compressive sensing approach, 896
 experimental results, 900–901
 signal model, 897–898
- Computational complexity theory, 424
- Conditional-CCRB, 43
- Conformal array, 650
- Consensus algorithm, 343–344, 348
 over realistic channels, 348
- Consensus-type approaches, 85
- Constant acceleration, 762
- Constant velocity, 762
- Control Time Protocol (CTP), 82
- Conventional SAR processing, 952
- Convolutional mixtures
 case of sparse channels, 171
 frequency-domain approaches, 176
 global BSS approaches, 175
 identifying the symbols, algebraic methods, 173
 iterative BSS, general cyclo-stationary, 178
 iterative BSS, stationary case, 176
 MA/AR structures (stationary data), 174
 subspace methods (cyclo-stationary data), 174
 temporal approaches, 175
- Cooperative sensor data fusion, 1261
- Cooperative sensor network, optimal deployment and functioning, 1268
- Cost function, 22, 27, 34
- Covariance Intersection (CI) algorithm, 764
- Coverage region estimation, 485
- CP. *See* Cyclic prefix (CP)
- CP-OFDM. *See* Cyclic-prefix OFDM (CP-OFDM)
- CP philosophy, 211
- Cramer-Rao bound (CRB), 10, 18, 42, 673–674
- CRB. *See* Cramer-Rao bound (CRB)
- CRN. *See* Cognitive radio networks (CRN)
- Cross-covariance, of estimation errors, 761
- Cross-layer design
 cellular networks, 487
 channel gain cartography, 480
 CR sensing and cross-layer design, 489
 MAC layer, 486
 physical layer, 473
 sparsity-aware power spectrum cartography, 475
 wireless regional area networks (WRAN), 486
- CSI. *See* Channel state information (CSI)
- CSI information, 276
- CSIP. *See* Collaborative signal and information processing
- CTP. *See* Control Time Protocol (CTP)
- CWDD. *See* Column-wise diagonally dominant (CWDD)
- Cyclic prefix (CP), 50
- Cyclic-prefix OFDM (CP-OFDM), 204
 vector-matrix representation of, 209
- Cyclo-correlation function, 143–144
- Cyclostationarity detection, 473
- Cyclostationarity tool, 37

D

- DA approach. *See* Data-aided approach
- 2-D assignment, measurement-to-track association, 776
- Data-aided (DA) approach, 17–18
- Data collection programs, 654
 knowledge-aided sensor signal processing and expert reasoning (KASSPER) data, 657
- mountaintop database, 656
- multichannel airborne radar measurements (MCARM), 654
- naval research laboratory (NRL) database, 656
- Data fusion, 510, 1245–1246, 1253
- Data Information Knowledge and Wisdom (DIKW), 1256–1257
- Data uploading, 333–336
- Decentralized observations with centralized estimation, 367
- Decentralized observations with decentralized estimation, 353
 conditionally independent observations, 353
- distributed Bayesian estimation under Gaussian noise and Laplacian prior, 358
- distributed ML estimation under Gaussian noise, 357
- distributed recursive least square estimation with sparsity constraint, 361
- spatially correlated observations, 364

- Decentralized sensor network, sensor coordination in, 1287–1288
 Decentralized tracking, 785
 Decision center
 conditionally independent observations, 379
 nodes send data to, 376
 signal embedded in additive noise, 376
 Decision-Directed (DD) mode, 24–25
 Decision feedback equalization, 111
 Decision feedback equalizer, 323
 Degree of polarization (DoP), 1139
 Degrees of Freedom (DoF), 415, 458, 1143
 Degrees of freedom region, 415
 Delay-Doppler spread function, 96
 Deployment of simple and complex nodes, 1288
 DERD. *See* Double bounce Eigenvalue Relative Difference
 Deterministic maximum likelihood estimation, 105–106
 Deterministic scatterers, 1227
 DFT. *See* Discrete Fourier transform (DFT)
 DFT-based implementation, OFDM, 198
 Difference of convex (DC) problem, 395
 Differential SAR Interferometry (DInSAR), 1061
 Digital audio broadcasting (DAB), 199
 Digital Terrain Elevation Data (DTED), 1300–1301
 Dihedral scattering, 1144
 Dihedral, scattering matrices, 1133
 DIKW. *See* Data Information Knowledge and Wisdom
 DInSAR. *See* Differential SAR Interferometry
 Directly achievable rate regions, 417
 Discrete cosine transform (DCT) matrix, 220–221
 Discrete Fourier transform (DFT), 176, 187–188, 299, 1053
 Discrete Karhunen-Lo  e (DKL), 255–257
 Discrete prolate spheroidal BEM (DPS-BEM), 99–100
 Discrete scatterers, 526
 Displaced phase center antenna (DPCA) technique, 948, 950
 Distributed Bayesian estimation under Gaussian noise and
 Laplacian prior, 358
 Distributed coherency matrix, 1142
 Distributed ML estimation under Gaussian noise, 357
 Distributed projection algorithms, 381
 Distributed resource allocation, interference channel, 445
 Distributed scatterers, 1226–1227
 Distributed tracking, 785
 Dithered quantization, 349
 Dolph-Chebyshev weighting function, 523–524
 Dominant scatterer autofocus, ISAR, 1009
 Doppler effects, 12, 99, 262–263
 Doppler parameter estimation, 962
 fractional Fourier transform (FrFT), 966
 matched filter bank, 963
 radon transform, 971
 Doppler spectrum, 522, 534, 536
 Double bounce Eigenvalue Relative Difference (DERD), 1183
 Douglas Sea State, 528
 DPS-BEM. *See* Discrete prolate spheroidal BEM (DPS-BEM)
 4D SAR imaging, 1103
 DSC. *See* Dynamic sensor collaboration
 DTED. *See* Digital Terrain Elevation Data
 Dual polarization SAR, 1138
 Dynamic sensor collaboration (DSC), 1275
- E**
- EDPCA. *See* Extended displaced phase center antenna technique
 Efficient eigenvector based decomposition, 1171–1173
 parametrization using number of degrees of
 freedom, 1174–1175
 target vector, circular basis, 1173
 Eigenvector based decomposition technique, 1144
 cloud dominant scattering mechanism decomposition,
 1144–1145
 H/A/ φ decomposition, 1145
 Eigenvector decomposition, 174–175, 1143
 EKF. *See* Extended Kalman filter (EKF)
 Ellipse amplitude, 1123
 Ellipse aperture, 1123
 Ellipse orientation, 1123
 Ellipticity, 1123
 Elliptic polarization, 1124
 EM algorithm. *See* Expectation-Maximization (EM) algorithm
 Energy detection, 473
 Epidemics, estimation and forecasting, 1302
 modeling, 1304
 numerical results, 1307–1309
 sequential Bayesian solution, 1305
 Equal-gain combining (EGC), 230–231
 Equalization
 decision feedback equalization, 111
 linear equalization, 108
 maximum likelihood sequence detection, 113
 turbo equalization, 114
 Equation error, 106
 ESPRIT algorithm. *See* Estimation of signal parameters via
 rotational invariance technique (ESPRIT) algorithm
 Estimation of signal parameters via rotational invariance
 technique (ESPRIT) algorithm, 72–73
 Euclidean Random Matrix (ERM), 403
 Euler's totient function, 221–222
 Expectation-Maximization (EM) algorithm, 33
 Extended displaced phase center antenna technique (EDPCA), 977
 Extended Kalman filter (EKF), 767, 834
- F**
- Far end crosstalk (FEXT), 298
 FCs. *See* Fusion centers (FCs)

- FECs. *See* Forward Error Correcting codes (FECs)
- FEXT. *See* Far end crosstalk (FEXT)
- Filter-bank multicarrier (FBMC), 277
- FIM. *See* Fisher information matrix (FIM)
- Fine synchronization methods, 65
- Finite Set Statistics (FISST), 782
- First-order statistics based methods, 63
- First scheme, DA phase and CFO, 20
- Fisher information matrix (FIM), 43–44, 673–674
 - for coherent MIMO radar joint estimation, 718
 - for localization with phase errors, 719
- FISST. *See* Finite Set Statistics (FISST)
- Fitting error, 106
- Flat fading channel, synchronization in
 - code-aided synchronization, 46
 - CRB, 42
 - DA case, 18
 - DA phase and CFO estimation, 25
 - DA timing estimation, 21
 - NDA case, 31
 - NDA ML phase and CFO estimator, 33
 - NDA ML timing estimator, 32
 - sub-optimal estimators, 36
- Flat plate, scattering matrices, 1131
- FOLiage PENetrating (FOPEN) radar, 1293
- Forestry remote sensing, 1237–1239
- Forward Error Correcting codes (FECs), 219
- Fourier series expansion, 37
- Fourier transform, 55–56, 144, 189
- Fractional Fourier transform (FrFT), 966
- Freeman-Durden decomposition, 1149–1150
- Free-space propagation, 858
- Frequency division multiplexing, 193
 - concept of, 194
- Frequency-selective multipath channels, 200
- Frequency spread signals, 714
- FrFT. *See* Fractional Fourier transform
- Fully polarimetric SAR, 1139
- Fully polarimetric tomography
 - mono-dimensional estimators, 1231–1233
 - multi-dimensional estimators, 1233
 - numerical examples, 1233–1235
- Fundamental information-theoretical issues, 341
- Fusion centers (FCs), 785
 - nodes send local decisions to, 380
- Fusion engines, 1297–1298
 - classification function, 1301
 - local tracking function, 1299
- ## G
- Game theoretical formulations, 447
- Game theoretic approaches, 474
- Gamma texture (GK), 520
- Gaussian amplitude statistics, 526
- Gaussian-GCRB, 43
- Gaussian Markov Random Field (GMRF), 339
- Gaussian noise, 13, 412–413
- Gaussian process, 11, 234, 550–551
- Gaussian statistics, 518–519
- Gauss-Markov Random field (GMRF), 371–373
- GCRB derivations, 45
- GDFE. *See* Generalized decision feedback equalizer (GDFE)
- GEMS. *See* Generic Error modeling System
- General GMTI processing chain, 950
 - clutter suppression, 952
 - detection, 953
 - parameter estimation, 957
 - preprocessing, 951
 - signal extraction, 955
 - visualization, 962
- Generalizations, BSS methods, 162
- Generalized decision feedback equalizer (GDFE), 323
- Generalized eigen-analysis, 609
- Generalized sidelobe canceler, 606
- Generic Error modeling System (GEMS), 1257
- GEOINT. *See* Geospatial Intelligence
- Geometric Programming (GP) technique, 432
- Geospatial Intelligence (GEOINT), 1258–1259
- Global BSS, stationary mixture, 158
 - first result, 159
 - generalization: notion of contrast function, 160
 - JADE, 161
- GMRF. *See* Gaussian Markov Random Field (GMRF)
- GP technique. *See* Geometric Programming (GP) technique
- Graphical models and consensus algorithm, 343
- Ground clutter
 - Doppler spectra for land clutter, 536
 - land clutter, 535
- Ground moving target indication (GMTI), 508–509
- Group and coefficient levels, sparsity, 477
- Grouped linear precoding (GLP), 222
- ## H
- $H/A/\bar{\alpha}$ decomposition
 - decomposition parameters, 1146–1149
 - mean scattering mechanism, 1145–1146
- Han-Kobayashi region, 415
- Heterogeneous clutter, 645
- Heterogeneous multi-sensor network management, 1285
- Hidden Markov Model (HMM), 104–105
- High entropy multiple scattering, 1153
- High entropy vegetation scattering, 1152
- High resolution image formation, 1046

- High resolution radar and radar imaging, 988
 high cross-range resolution, 992
 high range resolution, 990
 resolution, 989
- Homeland Protection (HP), 1249
- HOMHT. *See* Hypothesis-oriented MHT (HOMHT)
- Homogeneous sensor networks, 1259–1260
 sensor configuration, 1260
- Horizontal dipole, scattering matrices, 1132
- HP. *See* Homeland Protection
- HTP. *See* Hung Turner Projection (HTP)
- Human Intelligence (HUMINT), 1258–1259
- HUMINT. *See* Human Intelligence
- Hung Turner Projection (HTP), 626
- Huynen parameters, 1140–1141
- Hybrid scatterers, 1227
- Hypothesis-oriented MHT (HOMHT), 780
- I**
- IBI. *See* Inter-block interference (IBI)
- ICBF. *See* Iterative Coordinated BeamForming (ICBF)
- ICI. *See* Inter-carrier interference (ICI)
- IF. *See* Improvement factor (IF)
- IHP. *See* Internal Hermitian Product
- ILSP. *See* Iterative Least Square with Projection (ILSP)
- Image autofocus, ISAR, 1007
 dominant scatterer autofocus (DSA), 1009
 IEBA, 1009
 image contrast based autofocus (ICBA) algorithm, 1007
 PGA, 1010–1011
- Image contrast based autofocus (ICBA) algorithm, 1007
- Image formation, ISAR, 998
 radial motion compensation, 998
 range-doppler image formation, 999
- Imagery Intelligence (IMINT), 1258–1259
- Image scaling, ISAR, 1015
- Imaging space-time adaptive processing (ISTAP), 978
- IMINT. *See* Imagery Intelligence
- Impact assessment, 1254
- Improvement factor (IF), 614, 870
- Impulse response function (IRF), 1046–1047
- Incoherent and non-stationary pixels (red class), 1211
- Incoherent and stationarity pixels (green class), 1211
- Incoherent decompositions, 1143
- Information fusion, 1255–1256
- Information sources, 1258
- Information-theoretic results, 414
- Instantaneous mixtures, 146–147
 algebraic methods, 149
 deflation step, 157
 extensions, 158
- global BSS, 158
- improving the deflation, 157
- indeterminacies, 149
- iterative BSS, 151
 practical aspects, 155
 second-order based identification, 150
 theoretical considerations, 152
- Integrated sensor systems, 510
- Integration time, 1049
- Interacting multiple-model estimator, 770
- Inter-block interference (IBI), 50, 202
- Inter-carrier interference (ICI), 55–56, 197–198
- Interference alignment, 458
- Interference channel (IC), 6, 410
 capacity results for, 414
 distributed resource allocation in, 445, 449
 optimal resource allocation in, 420
 algorithms for, 425
 comparisons of, 445
 complexity of, 424
 hybrid formulations, algorithms for, 441
 lagrangian dual decomposition, 428
 min-rate maximization, algorithms for, 426
 problem formulations, 420
 QoS constrained power minimization, 439
 robust resource allocation, algorithms for, 442
 weighted sum-utility maximization, 428
 system model, 412
- Interfering broadcast channel (IBC), 411–412
- Interfering multiple access channel (IMAC), 411
- Internal Hermitian Product (IHP), 1198–1199
 polarimetric, 1200
 single polarization, 1199–1200
- Inter-Symbol Interference (ISI), 11
- Inverse DFT (IDFT), 199
- Inverse synthetic aperture radar (ISAR), 509
 geometry, 995–997
 high resolution radar and radar imaging, 988
 historical overview, 987–988
 image formation, 998
 SAR to ISAR, 995
 signal modeling, 996
- IRF. *See* Impulse response function
- ISAR. *See* Inverse synthetic aperture radar
- ISAR image evaluation, 1001
 image coordinates, 1001
 image projection plane, 1002
 image resolution, 1002
 images, examples of, 1004–1007
 point spread function, 1002
- ISB algorithm. *See* Iterative spectrum balancing (ISB)
 algorithm
- ISTAP. *See* Imaging space-time adaptive processing

Iterative BSS, general cyclo-stationary case

global BSS, 168

semi-analytical considerations, 169

Iterative Coordinated BeamForming (ICBF), 430

Iterative Least Square with Projection (ILSP), 138–140, 149

Iterative spectrum balancing (ISB) algorithm, 429

J

JADE algorithm. *See* Joint Approximate Diagonalization of Eigenmatrices (JADE) algorithm

Joint Approximate Diagonalization of Eigenmatrices (JADE) algorithm, 138–140

Joint Directors of Laboratories (JDL) fusion model, 1253–1255, 1257–1258

Joint domain space-time adaptive processing, 972

Joint location and velocity estimation, 670, 680

Joint probabilistic data association (JPDA), 778

Joint sensing and resource optimization, 490

Joint separating methods, 138–140

Jones vector, 1125

JPDA. *See* Joint probabilistic data association (JPDA)

K

Kalman detector, 109

Kalman filter, 109–110, 767

Karush-Kuhn-Tucker optimal solution, 495

KASSPER data. *See* Knowledge-aided sensor signal processing and expert reasoning (KASSPER) data

k-mean segmentation technique, 1155, 1158

Knowledge-aided (KA) methods, 753

Knowledge-aided sensor signal processing and expert reasoning (KASSPER) data, 657

Kriged Kalman filtering (KKF), 483

Kronecker delta function, 103

Kronecker index, 13–14

L

Lagrangian dual function, 429

Laplacian matrix, 344

Lasso algorithm, 394

Least-absolute shrinkage and selection operator (lasso) method, 359

Left helix, scattering matrices, 1133

LFC. *See* Local fusion center (LFC)

Lightweight Tree-based Synchronization (LTS), 82

Likelihood function, 102

Linear constraints, 747

pre-nulling on transmit, 748

Linear equalization, 108

Linearly precoded-OFDM (LP-OFDM), 219

Linear polarization, 1124

Linear precoding

partial ZF precoding, 312

reduced complexity ZF, 309

word length for ZF, 311

zero forcing precoder, 307

Linear precoding (LP) matrix, 219–220

Linear receiver, 322

Linear rotated basis, canonical scattering mechanisms, 1131

Linear time-varying (LTV)

channel equalization, 253

channel estimation, 255

Line-of-sight (LOS), 1046–1047

Local fusion center (LFC), 767, 785, 1247–1248

Lorentzian function, 522

Low-cost scientific measurements, 836

bistatic adjunct to weather radar, 837

Manastash Ridge Radar (MRR), 836

ocean remote sensing using gps signals, 838

Low entropy dipole scattering, 1152

Low entropy multiple scattering events, 1152

Low entropy surface scatter, 1152

LP-OFDM. *See* Linearly precoded-OFDM (LP-OFDM)

LTI channels for OFDM, 243

LTS. *See* Lightweight Tree-based Synchronization (LTS)

LTV. *See* Linear time-varying (LTV)

Lyapunov theorem, 429

M

MAI. *See* Multiple-access interference (MAI)

Manastash Ridge Radar (MRR), 836

MANET. *See* Mobile ad-hoc networks (MANET)

Markov observations, 337–338

MASINT. *See* Measurement and Signature Intelligence

Master-slave closed-loop phase synchronization, 700

Matching communication network topology, 393

numerical results, 396

precision matrix, 396

preserving total transmit power, encouraging sparsity by, 394

Maximum likelihood sequence detection (MLSD), 113

Maximum ratio combining (MRC), 230–231

Maximum SINR filter, 603

MCARM. *See* Multichannel airborne radar measurements (MCARM)

Measurement and Signature Intelligence (MASINT), 1258–1259

Medium access control (MAC) protocol, 332

Medium entropy multiple scattering, 1152

Medium entropy surface scatter, 1152

- Medium entropy vegetation scattering, 1152
MHT. *See* Multiple hypothesis tracker (MHT)
 Microwave high resolution imaging, 1043
MIMO. *See* Multiple-input multiple-output (MIMO)
 Minimum energy consensus, 385
 analytic approach, 391
 arbitrary networks, optimal topology and power allocation for, 388
 numerical examples, 389
 optimization criterion, 387
 random geometric graphs, 391
 Minimum variance beamformer, 605
MLSD. *See* Maximum likelihood sequence detection (MLSD)
 Mobile ad-hoc networks (MANET), 10
 Mode-conditioned filtering, 772
 Mode jump process, 771
 Modified-MCRB, 43–44
 Monochromatic plane wave, polarization of, 1121
 Motion adapted SAR processing, 952
 Mountaintop database, 656
 Moving point target signal model, 925–926
 multi-channel signal model, 929
 single-channel signal model, 926
 Moving target indication (MTI), 914–915, 925
 Moving target indication (MTI) radar, 611–612
 adaptive SINR loss, 613
 Clairvoyant SINR loss, 612
 improvement factor (IF), 614
 optimal and adaptive filter patterns, 614
 MTI radar. *See* Moving target indication (MTI) radar
 Multi-band testing, 473
 Multibaseline SAR interferometry, 1071
 Multicarrier CDMA (MC-CDMA), 225
 Multicarrier communications in nonlinear systems, 233
 Multicarrier (MC) communication system, 193
 Multichannel airborne radar measurements (MCARM), 654
 Multichannel SAR, 508
 Multi-channel signal model, 929
 Multidimensional polarimetric SAR signal processing, 1187
 Multipass interferometry, 1080
 multipass phase unwrapping, 1088
 Multipath exploitation, 874
 experimental results, 886
 image formation algorithm, 876
 multipath exploitation algorithm, 885
 multipath model, 877
 Multipath exploitation principle, 884
 Multipath model, 877
 multipath focusing analysis, 882–884
 multipath locations, 880
 Multiple-access interference (MAI), 49, 230–231
 Multiple hypothesis tracker (MHT), 779
 hypothesis-oriented MHT (HOMHT), 780
 track oriented MHT (TOMHT), 781
 Multiple-input multiple-output (MIMO)
 channel estimation, 108
 channels, 102
 coherent MIMO radar, 669
 coherent ambiguity function, 675
 joint location and velocity estimation, 670
 phase errors, localization with, 676
 distinguishing coherent and noncoherent processing, 716
 equalization for, 276
 FIM, for coherent MIMO radar joint estimation, 718
 FIM, for localization with phase errors, 719
 frequency spread signals, 714
 IC model, 413, 423–424
 Neyman-Pearson signal detection, 692
 signal model, 692
 signal space dimension, 693
 noncoherent MIMO radar, 679
 joint location and velocity estimation, 680
 noncoherent ambiguity function, 685
 performance and complexity analysis for, 688
 normalized mean square error difference, 688
 numerical examples, 690
 phase synchronization for, 699
 broadcast consensus based phase synchronization, 703
 master-slave closed-loop phase synchronization, 700
 round-trip phase synchronization, 701
 radars, 506–507
 technologies, 5
 waveform design for, 705
 optimal waveform design for, 709
 signal model and design criteria, 706
 Multiple input single output (MISO), 414
 Multiple Mode (MM) approach, 770–771
 Multi-sensor data fusion, 1253
 Multisensor integration, 1253
 Multi-sensor networks, mathematical issue for, 1288
 Multisensor tracking, 507
 Multi-StageWiener Filter (MWF), 629
 Multistatic radar, 506–507
 Multitarget multisensor tracking
 architectures of, 783
 centralized tracking, 786
 parallel updating, 787
 sequential updating, 786
 data association
 measurement-to-measurement association, 778
 measurement-to-track association, 774
 distributed tracking
 covariance intersection, 764
 cross-covariance of the estimation errors, 761
 with dependent errors, 762
 tracklet fusion, 763

- filter initialization, 773
 issues related to track initialization, 773
 single-point track initialization, 773
 two-point difference track initialization, 773
- filters, 766
 extended Kalman filter (EKF), 767
 interacting multiple-model estimator, 770
 Kalman filter, 767
 particle filter, 769
 unscented Kalman filter (UKF), 768
- formulation of, 761
 clutter model, 764
 sensor models, 763
 spatial clutter intensity estimation, 765
 target dynamic models, 762
- multitarget tracking algorithms, 778
 joint probabilistic data association (JPDA), 778
 multiple hypothesis tracker (MHT), 779
 probabilistic data association (PDA), 778
 probability hypothesis density (PHD) method, 782
- performance evaluation, 794
- Posterior Cramér-Rao Lower Bound (PCRLB) of tracking, 794
- simulations, 799
 tracker-related measures, 798
- Multi-terminal source/channel *coding* problem, 399
- Multicarrier systems
versus CDMA, 232
 equalization and data recovery for, 228
- Multiuser synchronization
 with generalized CAS, 74
 with interleaved CAS, 71
 subband CAS, 70
 uplink OFDMA systems, 77
 uplink signal model and synchronization policy, 68
- MWF. *See* Multi-StageWiener Filter (MWF)
- N**
- Naval research laboratory (NRL) database, 656
 NDA approach. *See* Non-data-aided (NDA) approach
 Near end crosstalk (NEXT), 298
 Network architecture scheme, 1288
 Network centric data fusion process model, 1267
 Network Centric Operation (NCO), 1248
 Network synchronization
 clock models, 81
 distributed clock sync, 85
 pairwise synchronization, 83
 protocols, 82
 unilateral synchronization, 83
 Network Time Protocol (NTP), 83
 NEXT. *See* Near end crosstalk (NEXT)
- Neyman-Pearson (NP) hypothesis testing, 473
 Neyman-Pearson (NP) signal detection, 692
 Node positions, 348
 Noise, achievable rate regions, 416
 Non-Circularity tool, 37
 Non-circular sources, case of, 182
 Noncoherent ambiguity function, 685
 Noncoherent MIMO radar, 679
 Non-data-aided (NDA) approach, 18
 Non-flat fading channels, synchronization for
 carrier frequency offsets, 57
 coarse synchronization methods, 58
 downlink OFDMA, 57
 fine synchronization methods, 65
 OFDM transmission, 52
 signal model and preliminaries, 50
 simulation results for downlink OFDMA systems, 66
 SIR *versus* fractional carrier frequency offset, 58
 subcarrier allocation strategies, 51
 timing errors, 55
- Nonlinear constraints
 constant modulus and the method of stationary phase, 752
 NLFM to achieve constant modulus, 753
 relaxed projection approach, 750
 relaxed projection example, 751
- Nonlinear frequencymodulated (NLFM) pulse, 752
 to achieve constant modulus, 753
- Nonlinear precoding, 304
- Nonparametric basis pursuit method, 478
- Non-stationarity, exploiting the, 182
 Nonstationary clutter, 647
 Non-stationary polarimetric SAR responses, 1192
 Bragg resonant scattering, 1196–1198
 discrete time-frequency decomposition, range and azimuth, 1192–1194
 TF behaviors detection, 1194–1196
- NRL database. *See* Naval research laboratory (NRL) database
- NTP. *See* Network Time Protocol (NTP)
- Nyquist filter, 13–14
 Nyquist raised-cosine filter, 136
 Nyquist rate, 100–101
 Nyquist-Shannon sampling theorem, 196, 198–199
- O**
- Observation model, impact of, 336
 Observe-Orient-Decide-Act (OODA) loop, 1257
 Ocean remote sensing, GPS signals, 838
 OFDM. *See* Orthogonal frequency-division multiplexing (OFDM)
- OFDMA. *See* Orthogonal frequency-division multiple access (OFDMA)

- OHGR. *See* Osborne Head Gunnery Range (OHGR)
- Open-Source Intelligence (OSINT), 1258–1259
- Operational modes, SAR system, 1050
- scanSAR mode, 1051–1052
 - spotlight mode, 1051
 - staring spotlight mode, 1051
 - stripmap mode, 1050
- Optical Network Unit (ONU), 297–298
- Optimal and adaptive filter patterns, 614
- Optimal radar waveform design
- additive colored noise case, 730
 - broadband multipath interference, 734
 - clutter case, 737
 - optimum pulse shaping for maximizing SCR, 740
 - sidelobe clutter discrete suppression, 739
 - constrained optimum transmit-receiver radar, 747
 - linear constraints, 747
 - nonlinear constraints, 750
 - open issues and problems, 753
 - transmit-receive functions, for target identification, 742
 - multitarget ID, 746
 - multi-target identification case, 745
 - two-target identification example, 744
- Optimal receiver structure, 14
- Optimal resource allocation problems, 424
- Optimal Voronoi partition, 1272
- Optimal waveform design, 709
- ORC. *See* Orthogonal restoring combining (ORC)
- Oriented dipole, scattering matrices, 1132
- Orthogonal frequency-division multiple access (OFDMA), 49, 410
- downlink OFDMA representation, 59
 - downlink OFDMA transmission, 53, 57
 - simulation results for, 66, 77
 - uplink representation, 69
- Orthogonal frequency-division multiplexing (OFDM), 5, 49
- bit and power-loading, 269
 - coded, 218
 - cognitive radio applications, 276
 - CSI information, 276
 - filter-bank multicarrier (FBMC), 277
 - in frequency-selective multipath channels, 200
 - linearly precoded-OFDM, 219
 - LTI channels for, 243
 - mathematical background, 189
 - MIMO, 276
 - and multicarrier communications in nonlinear systems, 233
 - multiuser multicarrier systems, 224
 - operational principle, 51
 - path to, 193
 - practical guidelines and commercial standards, 260
 - principle, 196
 - rapidly time-varying channels, 248
 - relay communications, 277
- single carrier background, 189
- symbol error rate (SER) performance analysis, 214
- time and frequency synchronization, 263
- time-varying channel estimation, 276
- transmission, 52
- underwater communications, 276
- vector-matrix representation for, 207
- Orthogonal multiplexing principle, 194
- Orthogonal restoring combining (ORC), 230–231
- Osborne Head Gunnery Range (OHGR), 559
- OSINT. *See* Open-Source Intelligence
- Outlook, 6
- Overlap-and-add (OLA) approach, 217–218
- P**
- Pairwise-error probability (PEP), 215
- PAMF. *See* Parametric adaptive matched filtering (PAMF)
- PAPRs. *See* Peak-to-average power ratios (PAPRs)
- Parallel IC model, 413, 424
- Parameter estimation, 957
- accelerations, 959–960
 - across-track velocity, 957–958
 - along-track velocity, 958–959
 - true target position, 960–961
- Parametric adaptive matched filtering (PAMF), 635
- Parseval's identity, 22–23
- Partial transmit sequences (PTS), 242
- Particle filter, 769
- Particle orientation, 1167
- Passive bistatic radar (PBR)
- airborne PBR, 839–840
 - applications, 836
 - border/perimeter surveillance, 838
 - HF skywave signals, 846
 - imaging, 839
 - issues and problems, 847
 - countermeasures and detection of, 847
 - identifying applications, 847
- Low-cost scientific measurements, 836
- short-range surveillance, WiFi and WiMAX, 840
- waveforms, 823
- See also* Bistatic radar
- Passive radar, 506–507
- Pauli target vector, 1130
- PBR. *See* Passive bistatic radar
- PCI. *See* Principle components inverse (PCI)
- PCRLB. *See* Posterior Cramér-Rao Lower Bound (PCRLB)
- PDA. *See* Probabilistic data association (PDA)
- Peak-to-average power ratio (PAPR)
- reduction methods, 241
- Peak-to-average power ratios (PAPRs), 50, 241

- Perron-Frobenius theorem, 346
 Perron matrices, 346
 Phase synchronization for MIMO, 699
 Phase unwrapping, 1066
 Bayesian statistical solution to, 1076
 graph cuts solution to, 1078
 minimum cost flow solution, 1070
 residue cut algorithms for, 1067
 PHD. *See* Probability hypothesis density (PHD)
 PLL scheme, 31
 Point Spread Function (PRF), 865–866, 1002, 1046–1047
 Poisson point processes, 765–766
 Polarimetric anisotropy, 1149
 Polarimetric decomposition techniques, 1142
 eigenvector based decomposition technique, 1144
 necessity of polarimetric decompositions, 1142
 types of, 1143
 Polarimetric decomposition techniques, soil
 characterization, 1181
 H, A, and $\bar{\alpha}$ and the Small Perturbation surface scattering
 Model (SPM) model, 1183
 reflection symmetry and Integral Equation Model (IEM)
 model, 1183
 Polarimetric diversity, 1127–1128
 Polarimetric ISAR (Pol-ISAR), 1021, 1023
 image formation, 1025
 initialisation, 1027
 most focussed ISAR image, 1028
 optimization, 1028
 polarimetric autofocus, 1026
 polarimetric ISAR imaging, 1023–1024
 polarimetric SAR imaging, 1022
 signal model, 1024
 Polarimetric SAR imaging, 1022
 Polarimetric SAR interferometry, 1045
 Polarimetric SAR interferometry (PolinSAR), 1120
 Polarimetric SAR tomography, 1046, 1222
 basics of, 1222
 tomographic focusing techniques, 1228
 tomographic signal models, 1225
 Polarimetric scattering, 1121
 Polarimetric scattering descriptors, 1127
 Polarimetric time-frequency characterization, 1204
 Polarimetry, 1119
 Polarization (SAR), 1119
 Polarization basis change, 1126, 1130
 Polarization characteristics, 524
 Polarization ellipse, 1121, 1123
 Polarizations, types of, 1124
 PolSAR TF analysis, 1210
 classification, 1210–1211
 TF cleaning of PolSAR data, 1211
 PolSAR TF signal modeling, 1204, 1207
 coherent pixel discrimination, 1209–1210
 non-stationary pixel discrimination, 1208–1209
 second order statistics, 1207–1208
 Possible architectures, 342
 Post-doppler STAP, 974
 Posterior Cramér-Rao Lower Bound (PCRLB), 760–761, 1264
 tracking of, 794
 Practical vegetation height estimation, 1220
 Precision matrix, 396
 Precoding
 CSI at transmitter, 117
 for MIMO channels, 117
 for SISO channels, 116
 Preprocessing, GMTI Kernel
 conventional SAR processing, 952
 motion adapted SAR processing, 952
 range compression with RCMC, 952
 Preserving total transmit power, encouraging sparsity by, 394
 PRF. *See* Point Spread Function
 Principle components inverse (PCI), 624
 Probabilistic data association (PDA), 778
 Probability density function (pdf), 1137
 Probability evaluation, 772
 Probability hypothesis density (PHD)
 filter, 766
 method, 782
 Probability hypothesis density (PHD) filtering and
 tracking, 835
 Process refinement, 1254
 Progressive correction algorithm, 1283
 Propagation model, 477
 Pure coherency matrix, 1143
- Q**
- QoS. *See* Quality of service (QoS)
 Quality of service (QoS), 49, 409, 439
- R**
- Radar, 507–508
 Radar clutter, 505–506
 Radar clutter analysis
 land clutter analysis, farmland area, 567
 sea clutter, 540
 Radar clutter modeling and analysis
 clutter modeling, 515
 radar clutter analysis, 539
 radar design and analysis, 585
 simulation methods, 581
 Radar design and analysis, 585
 performance acceptance and trials analysis, 587

- Radar design and analysis (*Continued*)
 performance prediction, 585
 radar performance, specification and measurement of, 587
 signal processing and design of detectors, 586
- Radar polarimetry, 510
 organization of work, 1120
 overview of, 1119
 polarimetric SAR interferometry (PolinSAR), 1120
- Radar signal processing, 505
- Radioactive point sources, 1279
- Radio frequency interference, 619
- Rain clutter
 atmospheric attenuation, 537
 rain attenuation, 537
 rain Doppler spectrum, 538
 theoretical and empirical model for, 537
- Random Geometric Graphs (RGG)
 adjacency matrix, 403
 spectrum of, 403
- Random link failuresmodel, 349
- Range compression with RCMC, 952
- Rapidly time-varying channels, 248
 carrier frequency offset (CFO), 252
 LTV channel equalization, 253
 LTV channel estimation, 255
- Rate region, definition of, 416
- Rayleigh flat-fading channel, 218
- Rayleigh quotient, 608
- RBS. *See* Reference Broadcast Systems (RBS)
- RD-STAP. *See* Reduced-dimension STAP (RD-STAP)
- Received signals
 assumptions on channels, 148
 models of sampled data, 146
- Receiver noise, 620
- Receiver structure, 121–122
- Reduced-dimension STAP (RD-STAP), 630
- Reduced-rank STAP, 622
- Reference Broadcast Systems (RBS), 82
- Reflection symmetry and Integral Equation Model (IEM) model
 roughness and moisture retrieval, 1185–1186
 SERD and the DERD parameters, 1183–1185
- Relaxed projection approach, 750
- Reliable coherent digital communication systems, 9–10
- Residual range cell migration
 SAR imaging, 931
- Resource allocation
 in communication networks, 409
 directly achievable rate regions, 417
 information-theoretic results, 414
 interference channel (IC), 410, 420, 445
 MIMO IC model, 413
 multiple input multiple output (MIMO), 411
 notation, 410
- parallel IC model, 413
 rate region, definition of, 416
 scalar IC model, 412
 via interference alignment, 458
- Rethinking mathematical algorithms, net-centric approaches, 1266
- RF cartography, 472, 490
- Right helix, scattering matrices, 1133
- Roll-off factor, 143
- Round-trip phase synchronization, 701
- Round trip time (RTT), 82–83
- Row-wise diagonally dominant (RWDD), 304
- RTT. *See* Round trip time (RTT)
- RWDD. *See* Row-wise diagonally dominant (RWDD)

S

- Sample matrix inversion, 609, 974
- SAR imaging
 across-track velocity, 936
 along-track acceleration, 940
 along-track velocity, 934
 basics concepts, 1046
 effects on, 931
 effects, summary of, 941
 residual range cell migration, 931
- SAR interferometry, 1054, 1212
 basics concepts, 1046
 interferometric coherence, 1214
 interferometric phase and height of scatterer, 1212
 polarimetric RVOG model, 1218–1219
 Pol-inSAR representations, 1218
 single polarization channel, random volume over ground model, 1214
- SAR interferometry (InSAR), 509–510
- SAR response, 1135–1136
- SAR tomography, 1093
 compressive sensing inversion for, 1098
 linear adaptive inversion for, 1096
 linear non-adaptive inversion for, 1095
 methods, performance comparison of, 1101
- SAR tomography (Tomo-SAR), 509–510
- SCA. *See* Successive convex approximation (SCA)
- Scalar IC model, 412, 423–424
- SCALE. *See* Successive Convex Approximation for Low complexity (SCALE)
- S&C and S&S estimators, 62–63
- Scattering matrix, 1129
- Scattering mechanisms characterization, 1175
 azimuthal orientation, 1175–1178
 statistical Wishart segmentation, 1178–1179
- Schmidl and Cox (S&C) algorithm, 59–61

- SCNR. *See* Signal-to-clutter-plus-noise ratio
 Sea clutter, 527
 - amplitude statistics, 532
 - AR model, 564
 - Bragg scattering and long waves, 556
 - clutter amplitude, statistical models of, 541
 - correlation analysis and power spectrum estimation, 552
 - cumulant domain analysis, 550
 - Doppler spectrum, 534
 - mean range texture autocovariance sequence, 556
 - speckle autocorrelation and cross-correlation sequences, 552
 - texture autocorrelation sequence, 553
 - theoretical and empirical models for, 529
 Second-Order Blind Identification (SOBI) algorithm, 138–140
 Second order cone programs (SOCP), 426
 Second order incoherent representations, 1133
 Second-order statistics based methods, 59
 Second scheme, DA phase and CFO, 20
 Self-organizing sensor networks, 1264
 Self-synchronization mechanism, 1265
 Semi-blind approaches, 107
 Sensor behavior assignment, 1286–1287
 Sensor deployment, 1285–1286
 Sensor models, 763
 Sensor network, 1296–1297
 Sensor networks, study cases of, 1268
 Sensor registration, 786, 1265–1266
 Sensors, 1292–1293
 - airborne sensors, 1293
 - electro-optic (EO), 1296
 - FOliage PENetrating (FOPEN) radar, 1293
 - ground-based sensors, 1293
 - and scenario parameters, 800
 - unattended ground passive sensors (UGPS), 1296
 Sequential alternatives, 473
 SERD. *See* Single bounce Eigenvalue Relative Difference
 SER performance analysis. *See* Symbol error rate
 - performance analysis
 Shi and Serpedin (S&S) timing estimator, 61–62
 SIGINT. *See* Signals Intelligence
 Signal and interference environment, 827, 829
 - multilateration and tracking, 834
 - performance prediction, 831
 - suppression techniques, 830
 Signal extraction
 - azimuth lines, extraction of, 956
 - range history tracking, 955
 Signal models
 - and design criteria, 706
 - receiver noise, 620
 Signal processing and communications
 - blind signal separation, 5
 - channel estimation, equalization, precoding, and tracking, 4
 contents and contributors, 4
 history, 3
 OFDM and multicarrier signal processing, 5
 and optimal resource allocation, 6
 outlook, 6
 spectrum sensing and cross-layer design for cognitive radio networks, 6
 synchronization, 4
 vectored multichannel VDSL, 5
 wireless sensor networks, 5
 Signals
 - cyclo-stationarity of source, 143
 - received signals, 145
 - source signals, 142
 Signals Intelligence (SIGINT), 1258–1259
 Signal-to-clutter-plus-noise ratio (SCNR), 975–976
 Signal-to-clutter ratio (SCR), 737–738
 Signal-to-noise ratio (SNR), 13, 414–415, 595, 889–890
 SIMO Channel Estimation, 107
 Simple polarimetric classification scheme, 1150
 - $H\alpha$ classification, 1150–1152
 - polarimetric anisotropy, 1153
 Simulation methods, 180
 - correlated non-Gaussian random numbers, 582
 - generating correlated Gaussian random numbers, 582
 - generating uncorrelated random numbers, 581
 Single bounce Eigenvalue Relative Difference (SERD), 1183
 Single carrier background, 189
 Single-channel signal model, 926
 Single-input multi-output (SIMO) systems, 109
 Single-input single-output (SISO) systems, 100–101, 109
 Single polarization SAR, 1133
 Single polarization tomography, 1228
 - mono-dimensional estimators, 1229
 - multi-dimensional estimators, 1229–1230
 - numerical examples, 1230–1231
 SISO channel estimation, 106
 Situation assessment, 1254
 SNR. *See* Signal-to-noise ratio (SNR)
 SOBI algorithm. *See* Second-Order Blind Identification (SOBI)
 - algorithm
 SOCP. *See* Second order cone programs (SOCP)
 Space-time adaptive processing (STAP), 971
 - adaptive filter implementations, 622
 - application, 639
 - algorithm performance, 642
 - interference characteristics, 639
 - with nonadaptive solution, 644
 - summary, 644
 - basic concepts, 602
 - challenges
 - bistatic, 649
 - conformal array, 650

- Space-time adaptive processing (*Continued*)
 - heterogeneous clutter, 645
 - nonstationary clutter, 647
 - data collection programs, 654
 - detection, 602
 - extended displaced phase center antenna technique (EDPCA), 977
 - imaging space-time adaptive processing (ISTAP), 978
 - implementation, 651
 - joint domain, 972
 - metrics, 611
 - post-doppler, 974
 - sample matrix inversion, 609
 - signal models, 615
 - clutter, 615
 - radio frequency interference, 619
 - receiver noise, 620
 - space-time snapshot, 621
 - target, 620
 - space-time filter formulations, 603
- Space-time filter formulations, 603
 - generalized eigen-analysis, 609
 - generalized sidelobe canceler, 606
 - maximum SINR filter, 603
 - minimum variance beamformer, 605
 - Rayleigh quotient, 608
- Space-time snapshot, 621
- Sparsity-aware power spectrum cartography, 475
 - compressed sensing approach, 476
 - group and coefficient levels, sparsity, 477
 - nonparametric basis pursuit, 478
 - propagation model, 477
 - splines, group-lasso on, 479
- Spatial clutter intensity estimation, 765
- Spatial filtering, 862
 - imaging results, 869–870
 - moving average wall removal, 865
 - notch filtering, 868
 - target spatial signature, 864
 - wall reflection, characteristic of, 862
- Spatio-temporal channel correlation, 482
- Speckle effect, 1120–1121, 1133–1134
- Speckle filtering, 1137
- Spectrum sensing, 486
- Sphere, scattering matrices, 1131
- Splines, group-lasso on, 479
- STAP. *See* Space-time adaptive processing; Space-time adaptive processing (STAP)
- Stationary and mutually independent, 151
- Stationary point target signal model, 921
- Statistically independent observations, 336
- Stochastic maximum likelihood estimation, 104–105
- Stochastic-SCRB, 42–43
- Straight-iso-range approximation, 998
- Strongly connected component (SCC), 401
- Structure of adaptive soft-input soft-output equalizer, 123–124
- Subband carrier allocation scheme (CAS), 51–52
 - generalized CAS, 74
 - interleaved CAS, 71
 - synchronization with, 70
- Subcarrier allocation strategies, 51–52
- Sub-object assessment, 1253
- Sub-optimal estimators, 36
- Suboptimal linear equalization, 218
- Successive convex approximation (SCA), 431–432
- Successive Convex Approximation for Low complexity (SCALE), 431–432
- Superimposed training-based approaches, 107
- Surface scattering, 1144
- Surveillance architecture, 1291
- Surveillance, classical approach to, 1261
- Symbol error rate (SER) performance analysis, 214
- Synchronization, 4
 - flat fading channel, 11
 - multiuser synchronization, 68
 - network synchronization, 79
 - non-flat fading channels, 49
- Synthetic aperture radar (SAR) imaging, 1133–1134
- Synthetic aperture radar (SAR) polarimetry, 1119–1120
 - basics, 1121
 - organization of work, 1120
 - polarimetric decomposition techniques, 1142
 - polarimetric scattering, 1121
- Synthetic aperture radar principle, 918
 - acquisition geometry and operation, 918
 - pulse compression and image formation, 923
 - stationary point target signal model, 921
- T**
- Tapped delay line model, 96
- Target dynamic models, 762
- Target parameters, 801
- Target signatures, 818
 - forward scatter, 818
- Target vector, 1129
- TCRB derivations, 44
- Telecommunication signals. *See* Signals
- TF polarimetric characterization, 1198
 - CS detection based on multiple criteria, 1203
 - CS detection based on TF entropy, 1200
 - point-like scatterers, Internal Hermitian Product (IHP), 1198
- Theory to real application problems, 1265
- Threat refinement, 1254
- Through-the-wall radar imaging

change detection based MTI approach, 888
 compressive sensing approach, 896
 data sets, 903
 issues and problems, 902–903
 multipath exploitation, 874
 wall clutter mitigation, 861
 wall parameter estimation/modeling, and subtraction, 871
 Through-the-wall radar imaging (TWRI), 508, 889–890
 Through-the-wall sensing, 857–858
 Through-the-wall synthetic aperture radar model, 876
 Time and frequency synchronization, OFDM, 263
 Time-division multiple-access (TDMA)-fashion, 409, 482
 Time-frequency image formation, 1019
 Time-variant (doubly selective) channels
 autoregressive (AR) models, 98
 basis expansion models (BEMs), 99
 tapped delay line model, 96
 Time-varying channel estimation, 276
 Time-windowing, 1013
 Timing estimation, 37
 Timing-sync Protocol for Sensor Networks (TPSN), 82
 Tiny/Mini-Sync (TMS) protocol, 82
 TMS protocol. *See* Tiny/Mini-Sync (TMS) protocol
 Toeplitz convolution matrix, 207–208
 TOMHT. *See* Track oriented MHT (TOMHT)
 Tomographic focusing techniques, 1228
 fully polarimetric tomography, 1231
 single polarization tomography, 1228
 Tomographic signal models
 PolTomSAR signal model, 1227
 TomSAR signal models, 1253–1255
 Tone injection (TI), 242–243
 Tone reservation (TR), 242
 TPSN. *See* Timing-sync Protocol for Sensor Networks (TPSN)
 Trackers, 801
 Tracking
 adaptive channel estimation
 for slowly varying channels, 118
 via subblock tracking, 119
 block-adaptive channel estimation using CE-BEM, 118
 symbol-adaptive joint channel estimation and
 data detection, 120
 Tracklet fusion, 763
 Track oriented MHT (TOMHT), 781
 Training-based channel estimation, 103
 Trihedral, scattering matrices, 1131
 True target position, 960
 True-TCRB, 42–43
 Turbo equalization, 125–128
 for doubly-selective channels, 115
 principle of, 114
 using EKF and CE-BEM, 121
 Turbo-synchronization, 47

TWRI. *See* Through-the-wall radar imaging (TWRI)
 Typical telecommunication (TLC) network, 332

U

UGPS. *See* Unattended ground passive sensors
 UKF. *See* Unscented Kalman filter (UKF)
 Unattended ground passive sensors (UGPS), 1296
 Unconditional-UCRB, 42–43
 Underfoliage imaging, 1239
 Underwater communications, 276
 Uniform linear array (ULA), 622, 739
 Unilateral synchronization, 83
 Unscented Kalman filter (UKF), 768
 recursion, 768
 sigma point generation, 768
 Unsupervised $H/\Lambda/\alpha$ segmentation, 1159–1163
 Urban remote sensing, 1235

V

Vectored multichannel VDSL
 channel matrix structure, 301
 downstream transmission, 304
 linear precoding, 307
 nonlinear precoding, 304
 system model, 297
 upstream transmission, 322
 decision feedback equalizer, 323
 linear receiver, 322
 Vector-matrix representation, OFDM
 of CP-OFDM, 209
 CP philosophy, 211
 of ZP-OFDM, 212
 Vegetation bias removal, 1219
 Volume scattering, 1144

W

Wall clutter mitigation, 861
 spatial filtering, 862
 wall parameter estimation/modeling, and
 subtraction, 871
 Wall parameter estimation/modeling, and subtraction, 871
 approach, 871
 imaging results, 873
 Water-filling (WF) principle, 270–271
 Wave covariance matrix, 1138–1139
 Waveform design, 506–507, 705
 Waveforms, passive bistatic radar (PBR), 823
 Wave polarization, 1125
 Wave temporal trajectory, 1122

- Werner channel model, 312
 Wide-sense stationary uncorrelated scattering (WSSUS) channel, 96, 98
 Wireless local area networks (WLANS), 102
 Wireless personal area networks (WPANs), 102
 Wireless regional area networks (WRAN), 486
 Wireless sensor networks (WSN), 5, 10, 1275
 computing while communicating, 332
 distributed detection, 371
 distributed estimation, 351
 distributed projection algorithms, 381
 fundamental information-theoretical issues, 341
 general framework, 332
 graphical models and consensus algorithm, 343
 Markov observations, 337
 matching communication network topology, 393
 minimum energy consensus, 385
 observation model, impact of, 336
 possible architectures, 342
 statistically independent observations, 336
 Wishart statistics, unsupervised statistical segmentation, 1155
 WLANs. *See* Wireless local area networks (WLANS)
 WPANs. *See* Wireless personal area networks (WPANs)
 WRAN. *See* Wireless regional area networks (WRAN)
 WSN. *See* Wireless sensor networks (WSN)
 WSSUS channel. *See* Wide-sense stationary uncorrelated scattering (WSSUS) channel

Y

Yule-Walker equation, 98

Z

Zero forcing precoder, 307
 Zero-mean Gaussian vector, 207
 Zero-padded OFDM (ZP-OFDM), 202
 detection for, 218
 vector-matrix representation of, 212
 Ziv-Zakai Bound (ZZB), 28–29
 ZP-OFDM. *See* Zero-padded OFDM (ZP-OFDM)
 ZZB. *See* Ziv-Zakai Bound (ZZB)