



J. R. Guerci

Space-Time
Adaptive Processing
for Radar

Space-Time Adaptive Processing for Radar

For a listing of recent titles in the *Artech House Radar Library*,
turn to the back of this book.

Space-Time Adaptive Processing for Radar

J. R. Guerci



Artech House
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data

Guerci, J. R.

Space-time adaptive processing for radar / J. R. Guerci.

p. cm. — (Artech House radar library)

Includes bibliographical references and index.

ISBN 1-58053-377-9 (alk. paper)

1. Radar. 2. Adaptive signal processing. 3. Adaptive antennas. 4. Space and time. I. Title. II. Series.

TK6580.G84 2003

621.3848—dc21

2003052296

British Library Cataloguing in Publication Data

Guerci, J. R.

Space-time adaptive processing for radar. — (Artech House radar library)

1. Adaptive signal processing 2. Radar I. Title

621.3'848

ISBN 1-58053-377-9

Cover design by Yekaterina Ratner

© 2003 ARTECH HOUSE, INC.

685 Canton Street

Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

International Standard Book Number: 1-58053-377-9

Library of Congress Catalog Card Number: 2003052296

10 9 8 7 6 5 4 3 2 1

Contents

Preface	<i>ix</i>
References	<i>x</i>
Acknowledgments	<i>xiii</i>
References	<i>xiv</i>
1 Introduction	<u>1</u>
1.1 The Need for Space-Time Adaptive Processing in Moving Target Indicator Radar	1
1.2 STAP for MTI Radar	3
1.3 Book Organization	7
References	10
2 Adaptive Array Processing	<u>11</u>
2.1 Introduction	11
2.2 Optimum Spatial (Angle) Beamforming	11
2.2.1 Derivation of the Optimum Beamformer	18
2.2.2 Case I: Additive White Noise	19
2.2.3 Case II: Additive Colored Noise	23

2.3	Optimum Temporal (Doppler/Pulse) Processing	29
2.4	Adaptive 1-D Processing	31
2.5	Adaptivity in Nonstationary Environments	39
2.6	Summary References	42 42
	Appendix 2A: ULA Antenna Pattern Response	44
	Appendix 2B: Derivation of the Maximum Likelihood Sample Covariance Matrix	46
3	Space-Time Adaptive Processing	51
3.1	Introduction	51
3.2	Need for Joint Space and Time Processing	52
3.2.1	Joint Clutter and Jamming Characteristics	63
3.3	Optimum Space-Time Processing for MTI Radar	65
3.4	STAP	71
3.5	Summary References	71 72
4	Other Important Factors Affecting STAP Performance	75
4.1	Introduction	75
4.2	Channel Mismatch	77
4.2.1	Angle-Independent Channel Mismatch	78
4.2.2	Angle-Dependent Channel Mismatch	92

4.3	Other Interference Subspace Leakage Effects	95
4.4	Antenna Array Misalignment	99
4.5	Nonlinear Arrays	103
4.6	Interference Nonstationarity and the Iceberg Effect	103
4.7	Summary	106
	References	108
5	STAP for Radar: Methods, Algorithms, and Performance	111
5.1	Introduction	111
5.2	Data-Independent Reduced-Rank STAP	114
5.2.1	Pre-Doppler (Signal-Independent) Reduced-Rank STAP: DPCA and Adaptive DPCA	114
5.2.2	Post-Doppler (Signal-Dependent) Reduced-Rank STAP	122
5.2.3	Other Rank-Reducing Linear Transformations	125
5.3	Data-Dependent Reduced-Rank STAP	129
5.3.1	Signal-Independent Methods	130
5.3.2	Signal-Dependent Methods	136
5.3.3	Comparison of Data-Dependent Rank-Reduction Methods	147
5.4	Structured-Covariance and Model-Based Methods	149
5.4.1	Covariance Matrix Tapers	151
5.4.2	Other Structured-Covariance Methods	156
5.5	Illustrative Design Examples	158
5.5.1	Signal-Independent Approach	159
5.5.2	Signal-Dependent Approach	161

5.6	Summary	163
	References	165
6	<u>Other Topics</u>	169
6.1	Introduction	169
6.2	Statistical Basis for STAP	169
6.3	STAP Implementation	173
6.4	Knowledge-Aided STAP	178
6.5	Summary	179
	References	179
	<u>About the Author</u>	181
	<u>Index</u>	183

Preface

The burgeoning popularity of *space-time adaptive processing* (STAP) is easily demonstrated with a quick keyword search. Although originally coined for airborne multichannel *moving target indicator* (MTI) radar [1, 2], the acronym has been adopted in many disciplines in which joint adaptive sensor temporal and spatial processing are performed (e.g., multidimensional adaptive filtering). Although a widely published topic, there is a void in book-form coverage at the introductory to intermediate level—a niche which this book is designed to address.

Multichannel space-time array processing is an extremely rich topic area in and of itself (see, for example, the recent authoritative text by H. L. Van Trees [3]). When coupled with the modern marvel of a radar system, it is doubtful that any single source could come close to providing comprehensive coverage. In selecting both the scope and treatment for this book, I had the benefit of having taught both introductory and advanced courses on STAP for radar for several years in industry, academia, and the IEEE Radar Conference series. The audience for such courses is extremely diverse, ranging from pure theoretical academicians to working engineers trying to implement STAP on time and on budget (and most occupation categories in between). Moreover, such forums were generally quite time constrained, placing a premium on efficient teaching methods.

Space-Time Adaptive Processing for Radar is my first attempt to distill the many lessons learned from not only my teaching experiences, but from my interactions with many esteemed colleagues and my own research. My overarching goal is to take the reader equipped with a basic foundation

in radar, electromagnetics, and signal processing from basic adaptive array principles to the state of the art in STAP algorithm design.

Echoing in my psyche during the creation of this treatment was the often-cited student goal, how do I make STAP work in the real world? The source of such pleas was often the disappointing results achieved when STAP methods were blindly applied to challenging, albeit realistic interference environments. A forensic analysis of such results invariably revealed that the root cause was an extreme oversimplification of the nature of the underlying interference.

For example, in the derivation of the optimum space-time beamformer, stationarity assumptions are imposed that are generally not met in practice (except, of course, in highly contrived scenarios—e.g., a flat uninhabited desert). Thus, after establishing basic first-order space-time models for clutter and jamming in Chapter 3, many important second-order and higher effects are detailed in Chapter 4 before modern STAP algorithms are introduced in Chapter 5. The perhaps obvious reason for this is simply that many algorithms perform very well on first-order models, but can produce disappointing results when higher-fidelity models or real data are employed. This effect is exacerbated as the strength of the interference increases (see the iceberg effect in Chapter 4)—precisely the situation one wants to remedy with STAP.

At the time this book was written, the only other books available devoted to STAP for radar were by Klemm [2, 4]. I have copies of both and refer to them often (as evidenced throughout this book). However, these are advanced treatments and, thus, not readily accessible to the newly initiated. Moreover, due to their enormous breadth of coverage, particular topics could not always be explored in great depth. Thus, to come even close to accomplishing my aforementioned objective in a single source, sacrifices in breadth had to be made. For example, I have chosen not to cover explicitly bistatic STAP or newly emerging application areas in *synthetic aperture radar* (SAR)—topics which are covered in [2, 4]. However, many of the robust STAP algorithms developed in Chapter 5 are applicable to these more exotic applications. Indeed, it is my hope that after completing Chapter 5, the motivated reader could simply learn the particulars of another application and then readily adopt a suitable STAP algorithm from his or her toolkit.

References

- [1] Ward, J., *Space-Time Adaptive Processing for Airborne Radar*, MIT Technical Report 1015, MIT Lincoln Laboratory, December 1994.

- [2] Klemm, R., *Principles of Space-Time Adaptive Processing*, London, England: IEEE Press, 2002.
- [3] Van Trees, H. L., *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, New York: Wiley Interscience, 2002.
- [4] Klemm, R., *Space-Time Adaptive Processing: Principles and Applications*, London, England: IEEE Press, 1998.

Acknowledgments

I have been extremely fortunate to have had first-hand interactions and collaborations with some of the founding “fathers” of both the theory of STAP (e.g., Professors Irving S. Reed and Dante C. Youla, who derived the colored-noise multichannel matched filter in continuous form in 1959—the year I was born! [1]) and its practical implementation (e.g., Marshall Greenspan). Their alacrity and clarity are truly inspirational. Other colleagues with whom it has been my privilege to collaborate include E. H. Feria, S. U. Pillai, A. O. Steinhardt, J. S. Goldstein, P. M. Techau, and J. S. Bergin (the latter two of whom were students of Harry Van Trees and were acknowledged in his recent array processing book [2]).

Lee Moyer (RPN rules!), Jamie Bergin, Paul Techau, Allen Adler, Michael Zatman, and Amy Alving provided substantial feedback on earlier drafts of this book, and I gratefully acknowledge their support. Also, the anonymous reviewer for Artech House provided many excellent suggestions for improving the technical presentation. I’d also like to acknowledge the gentle yet firm encouragement and assistance of my editor Ms. Barbara Lovenvirth of Artech House.

STAP is a vast field and I have benefited over the years from so many generous colleagues that have also directly and indirectly influenced the material in this book. I cannot possibly mention them all, but feel especially compelled to acknowledge M. Zatman, M. Wicks, P. Zulch, E. Baranoski, J. Ward, A. Farina, W. Melvin, and J. Carlini.

Last, but by no means least, I would like to acknowledge the support and understanding of my wife, Evelyn, and four children, Sophia, Raymond,

John, and Diana. I can never truly make up for the time together lost while I was writing this book, but I will spend the rest of my life trying.

References

- [1] Youla, D. C., *Theory and Design of Multiple-Channel “Matched” Filters*, Polytechnic University Report prepared for the Atlantic Research Corporation, June 1959.
- [2] Van Trees, H. L., *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, New York: Wiley Interscience, 2002.

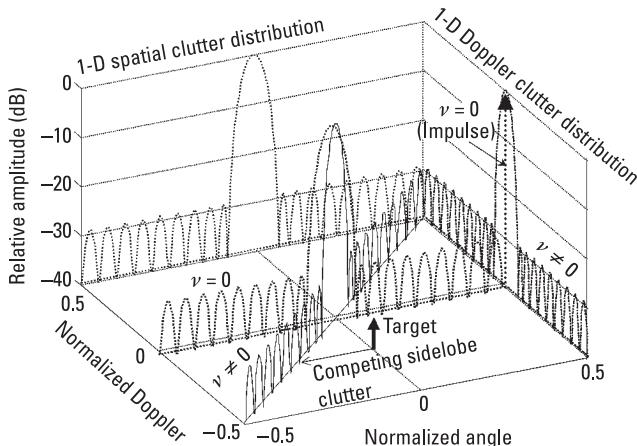
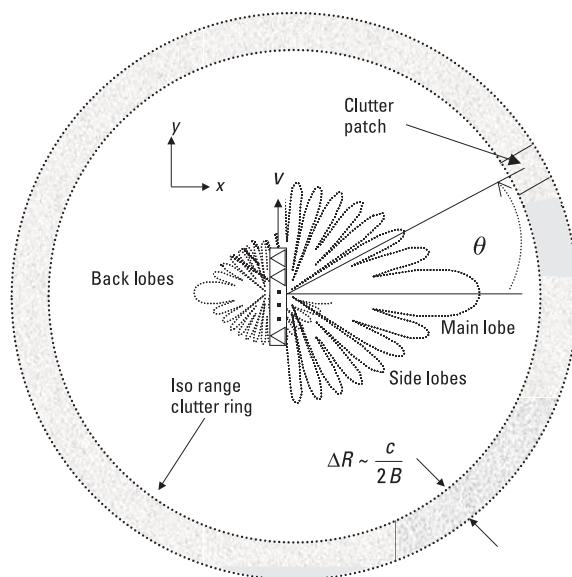
1

Introduction

1.1 The Need for Space-Time Adaptive Processing in Moving Target Indicator Radar

A quick literature search on the keyword “STAP” will quickly serve to illustrate the ubiquitous nature of this rich topic area. Although originally coined for airborne multichannel *moving target indicator* (MTI) radar [1, 2], *space-time adaptive processing* (STAP) has been adopted in many disciplines in which joint adaptive sensor temporal and spatial processing are performed (e.g., multidimensional adaptive filtering).

The need for joint space and time processing in either airborne (or spaceborne) MTI radar arises from the inherent two-dimensional (2-D) nature of ground clutter. Figure 1.1 illustrates the distribution of ground clutter (power spectral density arising from the two-way antenna pattern) for a side-looking radar employing a *uniform linear array* (ULA; see Chapter 2) as a function of normalized angle (i.e., the spatial dimension) and normalized Doppler (temporal dimension)—as defined in Chapter 2. Notice that in the absence of both ownship platform motion ($v = 0$ case in Figure 1.1) and *internal clutter motion* (ICM; see Chapter 4), the ground-clutter returns have no Doppler shift and can thus be characterized by a simple *one-dimensional* (1-D) process; that is, filtering of ground clutter can be achieved by simply rejecting signals from the 0-Hz Doppler filter [3, 4]. Moving targets are thus detected by looking for the presence of signals (of sufficient strength) in the remaining Doppler filters.



(b)

Figure 1.1 Illustration of (a) clutter iso-range ring and angle-Doppler dependence for a ULA radar in constant-velocity motion, and (b) corresponding angle-Doppler clutter “ridge” arising from main lobe and sidelobe antenna patterns. Note that a target of interest generally must compete with either sidelobe or main lobe clutter leakage.

However, when ownship motion is present, ground returns are Doppler shifted according to the relationship [1, 2]

$$\bar{f}_d = \beta \bar{\theta} \quad (1.1)$$

where $\bar{\theta}$ and \bar{f}_d are the normalized angle and Doppler respectively (see Chapter 3) and β is a proportionality constant that depends on ownship speed, *pulse repetition frequency* (PRF), and array interelement spacing (see Chapter 3).

As illustrated in Figure 1.1, ownship velocity gives rise to 2-D clutter “ridges” whose slope is governed by (1.1). Moreover, due to the Nyquist frequency of $\pm \text{PRF}/2$ [3], there may be more than one distinct ridge due to Doppler aliasing [1, 2]. As a consequence, it is likely that for every target of interest, there is a region of ground clutter with the same Doppler frequency—but offset in angle—that potentially could mask the target through either sidelobe or main lobe leakage [1, 2]. The objective of space-time processing is to place a null in the angle-Doppler beampattern where clutter may compete. Due to the aforementioned 2-D nature of clutter, this null will depend on the target angle-Doppler of interest [1, 2].

While (1.1) establishes the basic need for space-time (angle-Doppler) processing, it in no way implies the need for adaptivity. Since it is conceivable to know the relative orientation of the radar and ground surface through ownship *inertial navigation systems* (INS) and *global positioning systems* (GPS), as well as basic ground topology and clutter prediction maps [5], one could simply steer a deterministic null in the predicted clutter direction [see (1.2)]. Unfortunately, a multitude of practical real-world considerations severely limits the efficacy of this technique. Indeed, many of these considerations constitute both the impetus and subject matter of this book. Examples of such considerations include channel mismatch, clutter heterogeneity/nonstationarity, ICM, and other *interference subspace leakage* (ISL) mechanisms (see Chapters 2 through 4). It is these effects that give rise to the inherent need for adaptivity—hence STAP.

1.2 STAP for MTI Radar

Figure 1.2 shows the basic schematic architecture of a STAP beamformer. For ease of analysis we will nominally consider an N -element ULA and a *coherent processing interval* (CPI) consisting of M pulses with a fixed *pulse repetition interval* (PRI) (as notionally depicted in Figure 1.2)—although

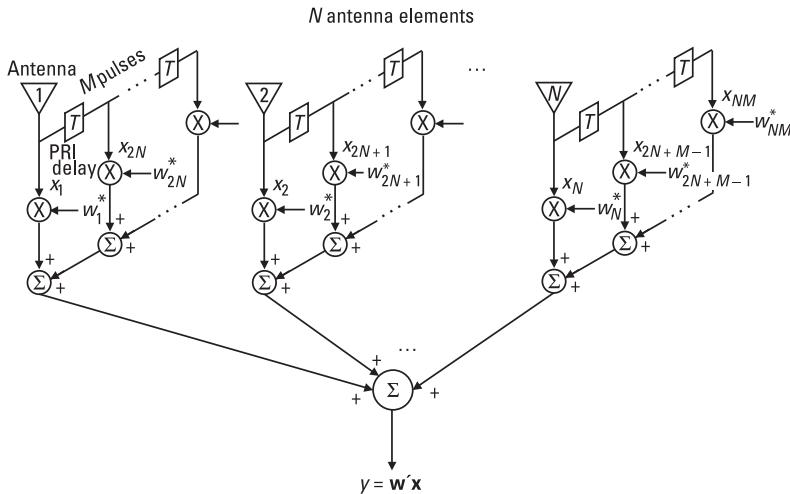


Figure 1.2 Space-time (angle-Doppler) beamformer consisting of N independent antenna channels and M pulses comprising the CPI. A specific angle-Doppler pattern is obtained by judicious selection of the complex linear combiner weights.

many of the concepts and techniques discussed are readily extensible to nonuniform/nonlinear [6] arrays and variable PRFs (see Chapters 4 and 5). As with ordinary one-dimensional (1-D) spatial-only beamforming, a 2-D angle-Doppler (space-time) beampattern can be formed by a judicious selection of the complex linear combiner weights $\{w_i\}$ [1, 2] (see Chapter 3 for further details).

As discussed in Chapter 2 (and elsewhere [7]), to maximize the response [and/or *signal-to-noise ratio* (SNR)] to a uniform narrowband plane wave corresponding to a given angle and Doppler, the linear combiner weight vector $\mathbf{w} = \text{vec}(w_1, \dots, w_{NM})$ (where $\text{vec}(\cdot)$ is the vector operator that simply forms an NM -dimensional column vector from the NM elements w_1, \dots, w_{NM}) should be set equal to the anticipated structure of the desired signal \mathbf{s} ; that is, $\mathbf{w} = \mathbf{s}$ (see Chapter 2 for further details). For the ideal ULA example under consideration, the result is the 2-D sinc pattern shown in Figure 1.3. Note the presence of a peak main lobe and both angle and Doppler sidelobes, which for a sinc pattern may be unacceptably high. Although tapering, or “windowing,” can be used to reduce these sidelobes, they come at the expense of an increase in main lobe width and a decrease in SNR (see Chapter 2) [7]. Moreover, channel mismatch places further limitations on the amount of tapering possible (see Chapter 2).

As detailed in Chapter 2, a linear combiner of this type has NM degrees of freedom (DoFs). Thus, for a desired angle-Doppler look direction \mathbf{s} , we

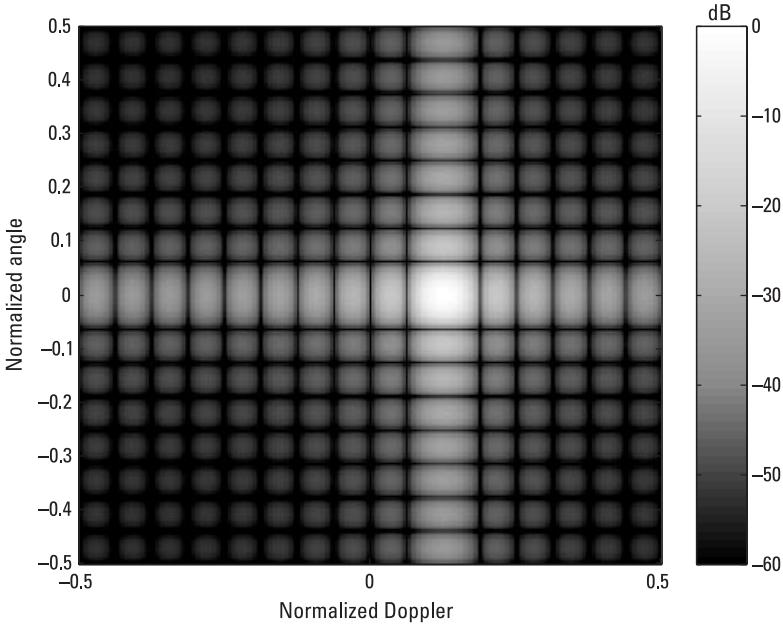


Figure 1.3 Example space-time (angle-Doppler) beamformer response for $N = M = 16$ sidelooking ULA. Note the presence of a finite-width main lobe and significant angle and Doppler sidelobes.

can also specify up to $NM - 1$ “nulls.” Specifically, if the set of NM angle-Doppler vectors $\{\mathbf{s}, \mathbf{j}_1, \dots, \mathbf{j}_{NM-1}\}$ is linearly independent [8], then a deterministic set of beamformer weights can be obtained from the linear set of equations (see Chapter 2)

$$\begin{aligned} \mathbf{w}'\mathbf{s} &= 1 \\ \mathbf{w}'\mathbf{j}_1 &= 0 \\ &\vdots \\ \mathbf{w}'\mathbf{j}_{NM} &= 0 \end{aligned} \tag{1.2}$$

Unfortunately, due to the relative strength of potential interference (clutter and/or jamming), and the aforementioned practical limitations on deterministic nulling, this approach is untenable in practice [1, 2]. Instead, the radar community has looked towards statistical signal processing for a means of addressing the limitations of deterministic nulling. In particular, clutter and jamming are treated as stochastic processes and the optimum space-time beamformer is derived via a statistical optimization procedure

[1, 2]. Although there are many variations on the theme (many of which are described in this text), the basic underlying optimum beamformer result is given by

$$\mathbf{w} = R^{-1}\mathbf{s} \quad (1.3)$$

where as before, \mathbf{s} is the desired signal, and \mathbf{w} is the optimal set of combiner weights. R is the positive-definite $NM \times NM$ dimensional covariance matrix [1, 2] associated with the total interference (clutter plus jamming plus receiver noise) for the given range cell under test [1, 2]. Since we have already stipulated that it is not possible to characterize the space-time interference a priori with sufficient precision to effect adequate nulling, it may at first appear that (1.3) is a mere tautology. However, under conditions discussed in Chapters 3 and 4, and elsewhere [9], it may be possible to estimate R from sample data obtained in the normal course of radar operation. Indeed, it is the estimation of R on the fly that is the true basis for the inclusion of the word “adaptive” in STAP [1, 2].

However, as will be made clear in Chapters 2 through 5, great care may be required in practice to glean the essential information from the environment since the amount of data available for estimating R is fundamentally limited not only by the resolution properties of the radar, but by the environment itself [10]. This latter consideration is particularly acute in highly heterogeneous and nonstationary terrain that can be expected to occur in a variety of potential real-world operating conditions [5, 10, 11]. It is the intention of this book to introduce the reader to an important cross section of these issues and to introduce STAP methods which have the potential of meeting those challenges.

A clear illustration of the conspicuous differences between real-world and synthetic, or mathematically idealized, ground clutter is afforded by examination of the returns from the *Defense Advanced Research Projects Agency* (DARPA) mountaintop radar [12], shown in Figure 1.4. Although ground-based, airborne MTI radar operation was emulated by a so-called *inverse displaced phase center array* (IDPCA) technique in which successive pulses were transmitted from a different phase center along the ULA [12]. Thus, as far as ground clutter was concerned, the illumination appeared to be moving.

Figure 1.5 shows the processed returns as a function of range and Doppler for both the actual radar and a simulation based on a homogenous clutter assumption [5]. The radar was situated at the White Sands Missile Range (WSMR). It is clearly not hard to imagine the mismatch pitfalls that

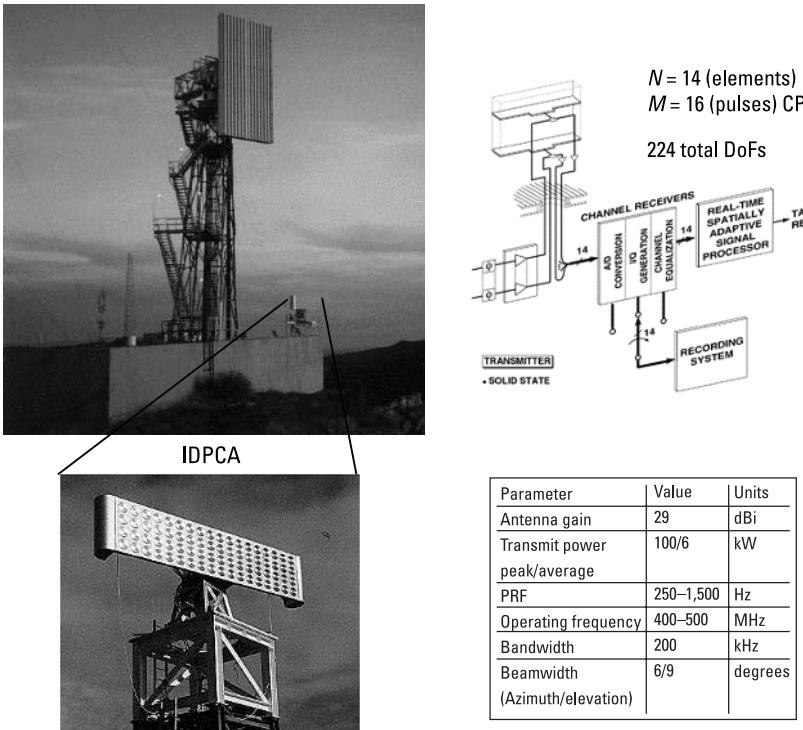


Figure 1.4 Photograph and specifications of the DARPA mountaintop radar. The IDPCA array was used on transmit to emulate platform motion.

await the unsuspecting radar engineer if care is not taken to properly segment the returns for both training (estimating the covariance—see Chapter 5) as well as establishing a *constant false-alarm rate* (CFAR) detection scheme [13].

1.3 Book Organization

Chapter 2 introduces the basic concepts of adaptive array processing by first considering the 1-D ULA case. Surprisingly, many of the issues that are at the forefront in advanced STAP research, such as convergence and robust adaptivity in heterogeneous and nonstationary environments [10], can be introduced with the 1-D case. After introducing the basic operation of a 1-D phased array, the optimum beamformer is derived that maximizes the total signal-to-interference-plus-noise ratio (SINR). This formulation introduces the concept of the interference covariance matrix R . Next, since R is

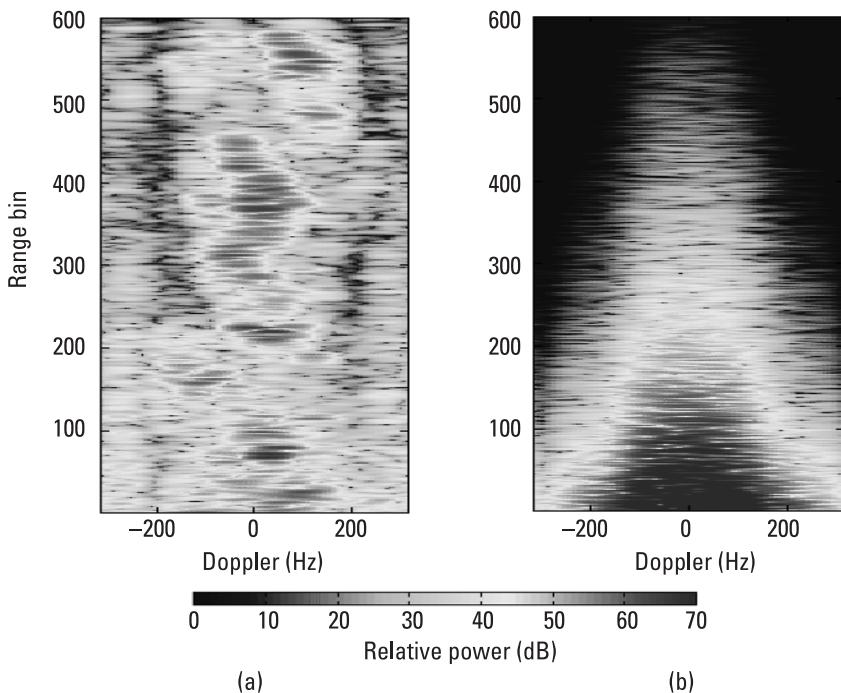


Figure 1.5 (a) Actual processed clutter returns for the DARPA mountaintop radar; and (b) simulated returns assuming homogenous clutter. (Courtesy of Information Systems Laboratories (ISL), Vienna, Virginia.)

not known a priori in practice, the concept of an adaptive array is introduced as a practical attempt to estimate R on the fly from the assumed-to-be-available sensor data [1, 2]. Properties of the sample covariance-based adaptive beamformer are examined via an eigen analysis of the adaptive beamformer. This leads to the introduction of the method of principal components to help alleviate the potentially deleterious effects of finite sample estimates [14]. Finally, a covariance augmentation technique [7], referred to as *covariance matrix tapers* (CMTs) [15], is introduced to illustrate a potential solution to the so-called stale weights problem.

Chapter 3 introduces the basic underlying signals and interference models for STAP for airborne MTI radar. Again, beginning with a ULA radar, the basic 2-D properties of ground clutter are developed along with a derivation of Brennan's rule that provides an analytical estimate of the underlying rank of the clutter covariance matrix [1, 2]. The optimum space-time beamformer is then derived and applied to a notional $N = M = 16$

sidelooking ULA radar. As with the 1-D case, the optimum beamformer requires knowledge of the total interference covariance matrix. However, the dimensionality of the space-time covariance is dramatically greater (due to the Kronecker relationship of angle and Doppler)—thus exacerbating the convergence/implementation issues introduced with 1-D adaptive array processing.

Next, in Chapter 4, important refinements to the signal models introduced in Chapter 3 are detailed. These factors can have a profound deleterious impact on STAP performance in practice and must therefore be properly accounted for in the beginning design stages. Factors considered include channel mismatch (narrow and finite bandwidth) [16], ICM [17], antenna crabbing, other ISL effects [11], as well as clutter heterogeneity [10].

In Chapter 5, a broad class of STAP algorithms is presented and taxonomized based on their fundamental formulation. Emphasis is placed on those techniques that make the most efficient use of the adaptive DoFs of the STAP processor. The primary objective of Chapter 5 is to provide a pallet of techniques that can be combined and tailored to a specific application in order to maximize STAP efficiency—loosely defined as maximizing convergence rate (minimizing requisite sample support), yet closely preserving near-optimal performance. The chapter culminates with a design example in which the requisite sample support for a sample matrix-based STAP processor is reduced by an order of magnitude with commensurate computational reductions in the adaptation process.

Finally in Chapter 6, some important additional topics are discussed. In Section 6.2, the statistical basis for STAP is introduced. For the *additive Gaussian noise* (AGN) case, it is shown that the max SINR [or minimum *mean-squared error* (MSE)] space-time beamformer based on the total interference covariance matrix is indeed optimal in the statistical sense (e.g., maximum likelihood sense [7]). References are provided for the reader interested in exploring the latest research trends related to non-Gaussian extensions of relevance to STAP. In Section 6.3, basic real-time implementation issues are discussed with an emphasis on the popular data-domain QR factorization approach [18]. This formulation leads to an extremely efficient parallel-processing implementation of the *sample matrix inverse* (SMI). Due to its ubiquitous adoption, we have extended it to include both diagonal loading and CMT—the latter covariance augmentation being of value in implementing a class of robust STAP algorithms introduced in Chapter 5. Lastly, we introduce the concept of *knowledge-aided STAP* (KA-STAP), which has recently emerged to address the difficulties encountered when attempting to adapt to extremely complex clutter environments [5].

References

- [1] Ward, J., *Space-Time Adaptive Processing for Airborne Radar*, MIT Technical Report 1015, MIT Lincoln Laboratory, December 1994.
- [2] Klemm, R., *Principles of Space-Time Adaptive Processing*, London, England: IEEE Press, 2002.
- [3] Barton, D. K., *Radar Systems Analysis*, Dedham, MA: Artech House, 1976.
- [4] Schleher, D. C., *MTI and Pulsed Doppler Radar*, Norwood, MA: Artech House, 1991.
- [5] Guerci, J. R., "Knowledge-Aided Sensor Signal Processing and Expert Reasoning," *Proc. of 2002 Workshop on Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER)*, Washington, D.C., April 3, 2002 (CD-ROM).
- [6] Zatman, M., "Circular Array STAP," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 510–517.
- [7] Van Trees, H. L., *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, New York: Wiley Interscience, 2002.
- [8] Strang, G., *Introduction to Linear Algebra*, Wellesley, MA: Wellesley-Cambridge Press, 1998.
- [9] Reed, I. S., J. D. Mallet, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 10, No. 6, November 1974, pp. 853–863.
- [10] Melvin, W., (ed.), "Space-Time Adaptive Processing and Adaptive Arrays: Special Collection of Papers," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 508–509.
- [11] Guerci, J. R., and J. S. Bergin, "Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January, 2002, pp. 152–162.
- [12] Titi, G. W., and D. F. Marshall, "The ARPA/NAVY Mountaintop Program: Adaptive Signal Processing for Airborne Early Warning Radar," *Proc. of 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 2, Atlanta, GA, May 7–10, 1996, pp. 1165–1168.
- [13] Nitzberg, R., *Radar Signal Processing and Adaptive Systems*, 2nd ed., Norwood, MA: Artech House, 1999.
- [14] Tufts, D. W., R. Kumaresan, and I. Kirssteins, "Data Adaptive Signal Estimation by Singular Value Decomposition of a Data Matrix," *Proc. of the IEEE*, Vol. 70, No. 6, June 1982, pp. 684–685.
- [15] Guerci, J. R., "Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming," *IEEE Transactions on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–985.
- [16] Billingsley, J. B., *Radar Clutter*, Norwood, MA: Artech House, 2001.
- [17] Zatman, M., "The Effect of Bandwidth on STAP," *Proc. of IEEE Antennas and Propagation International Symposium*, Vol. 2, Baltimore, MD, July 21–26, 1996, pp. 1188–1191.
- [18] Farina, A., and L. Timmoneri, "Real-Time STAP Techniques," *Electronics and Communication Engineering Journal*, Vol. 11, Issue 1, February 1999, pp. 13–22.

2

Adaptive Array Processing

2.1 Introduction

In this chapter we introduce both optimum and adaptive, spatial (angle) and temporal (Doppler) processing. Many of the key concepts and issues in STAP can be readily grasped by first successively examining the 1-D constituent angle and Doppler processing chains. In Chapter 3, we consider the joint spatiotemporal (2-D) process required for advanced clutter suppression in airborne radars.

2.2 Optimum Spatial (Angle) Beamforming

Consider the effect of a unit-amplitude, narrowband *electromagnetic* (EM) plane wave impinging on an N -element ULA with interelement spacing d , as depicted in Figure 2.1. In this context, the term *narrowband* refers to a signal whose modulation bandwidth, B , is such that $c/B \gg Nd$ [1]. This condition insures that propagation delay across the array is manifested as a simple phase shift [see (2.1)]. The impact of finite bandwidth will be addressed in Chapter 4.

If we define the plane wave *angle of arrival* (AoA), θ_0 , relative to boresight as shown in Figure 2.1, the complex envelope phasor at baseband observed at the n th antenna element as a function of θ_0 is

$$s_n = e^{j2\pi(n-1)\frac{d}{\lambda} \sin \theta_0}, \quad n = 1, \dots, N \quad (2.1)$$

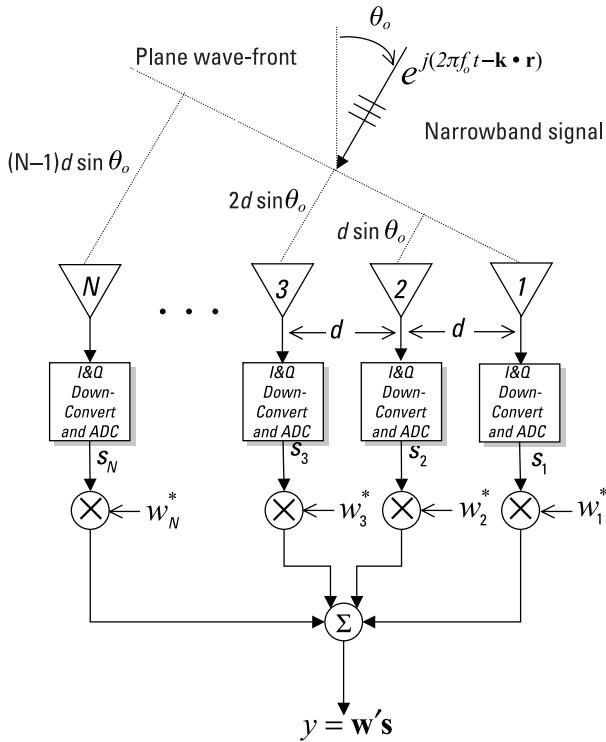


Figure 2.1 A ULA adaptive beamformer. The spatial (angle) response is controlled by the complex weighting factors.

where λ is the operating wavelength (units consistent with d), and θ_0 is the AoA in radians [1]. Note that the phase progression of a plane wave is linear across a ULA. The impact of a nonlinear array geometry for airborne MTI radar is discussed in Chapter 4.

By introducing multiplicative complex weighting factors, w_n , in each receive channel of the array (as shown in Figure 2.1), the output response can be maximized for any desired AoA. More specifically, let y denote the scalar beamformer output defined as

$$y = \sum_{n=1}^N w_n^* s_n = \mathbf{w}' \mathbf{s} \quad (2.2)$$

where an asterisk (*) denotes complex conjugation, the prime ('') denotes vector complex conjugate transposition (i.e., Hermitian transpose [2]), and

the vectors $\mathbf{s} \in \mathbb{C}^N$ and $\mathbf{w} \in \mathbb{C}^N$ (\mathbb{C}^N denotes the space of N -dimensional complex vectors) are defined as

$$\mathbf{s}(\theta_0) \triangleq \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_N \end{bmatrix} = \begin{bmatrix} e^{j0} \\ e^{j2\pi\frac{d}{\lambda}\sin\theta_0} \\ e^{j2\pi(2)\frac{d}{\lambda}\sin\theta_0} \\ \vdots \\ e^{j2\pi(N-1)\frac{d}{\lambda}\sin\theta_0} \end{bmatrix} \quad (2.3)$$

and

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_N \end{bmatrix} \quad (2.4)$$

To maximize the response of the beamformer to a plane wave arriving at an AoA of θ_0 , we have the following elementary optimization problem:

$$\max_{\{\mathbf{w}\}} |y|^2 = \max_{\{\mathbf{w}\}} |\mathbf{w}'\mathbf{s}|^2 \quad (2.5)$$

$$\text{subject to } \|\mathbf{w}\|^2 = \text{constant} < \infty$$

where the constant gain constraint is imposed to insure a finite solution. Since $\mathbf{w}'\mathbf{s}$ is simply an inner (or dot) product of the two nonzero norm vectors \mathbf{w} and \mathbf{s} [2], Schwarz's inequality can be applied [2, 3]; that is, $|\mathbf{w}'\mathbf{s}|^2 \leq \|\mathbf{w}\|^2 \|\mathbf{s}\|^2$, with equality if (and only if) the vectors are colinear. This yields the result

$$\mathbf{w} = \kappa \mathbf{s} \quad (2.6)$$

where κ is a scalar chosen to satisfy the normalization constraint.

This result is intuitive inasmuch as it states that the optimum beamformer applies phase corrections to each channel to compensate for the

time delays associated with the plane wave traveling across the array. More specifically, at the n th channel, the beamformer forms the product $w_n^* s_n \propto e^{-j\alpha_n} e^{j\alpha_n} = 1$, thereby canceling the phase term. The beamformer thus coherently integrates the signal outputs from each channel [1]. Without this compensation, destructive interference would occur with a commensurate decrease in output signal strength.

An important and fundamental limitation of linear beamformers is that they will also, in general, respond to signals arriving from other angles. This can lead to many practical problems, as strong unwanted signals from other directions can interfere with the signal of interest. To visualize this effect, consider the response of the above beamformer to plane waves arriving from -90° to $+90^\circ$ with $\mathbf{w} = \kappa \mathbf{s}$, where \mathbf{s} is chosen to be a plane wave with AoA of θ_0 . If we let x_n denote the output of the n th receive channel (and $\mathbf{x} \in \mathbb{C}^N$ the corresponding vector of received values), then the total beamformer output, steered to angle θ_0 , is given by

$$y = \mathbf{w}' \mathbf{x} = \kappa \sum_{n=1}^N x_n e^{-j2\pi(n-1)\frac{d}{\lambda} \sin \theta_0} \quad (2.7)$$

which has the form of a *discrete Fourier transform* (DFT) [1]. If we set $d/\lambda = 0.5$ (half-wavelength element spacing), $N = 16$, $\theta_0 = 30^\circ$, and

$$x_n = e^{j2\pi(n-1)\frac{d}{\lambda} \sin \theta}, \quad n = 1, \dots, N \quad (2.8)$$

while varying θ from -90° to $+90^\circ$, the beamformer response of Figure 2.2 results (with $\kappa = 1$). Note that for a ULA, the beamformer response can be obtained via a *fast Fourier transform* (FFT) [1]. For this particular example, an analytical expression for the normalized beamformer response ($|y| \leq 1$) exists and is given by [1] (see Appendix 2A)

$$|y| = \frac{1}{N} \left| \frac{\sin \left[N\pi \frac{d}{\lambda} (\sin(\theta) - \sin(\theta_0)) \right]}{\sin \left[\pi \frac{d}{\lambda} (\sin(\theta) - \sin(\theta_0)) \right]} \right| \quad (2.9)$$

The beamformer response of Figure 2.2 has several interesting features. First, note that signals close to 30° also produce a significant response. This region is generally referred to as the *main lobe*. The lobing structures outside

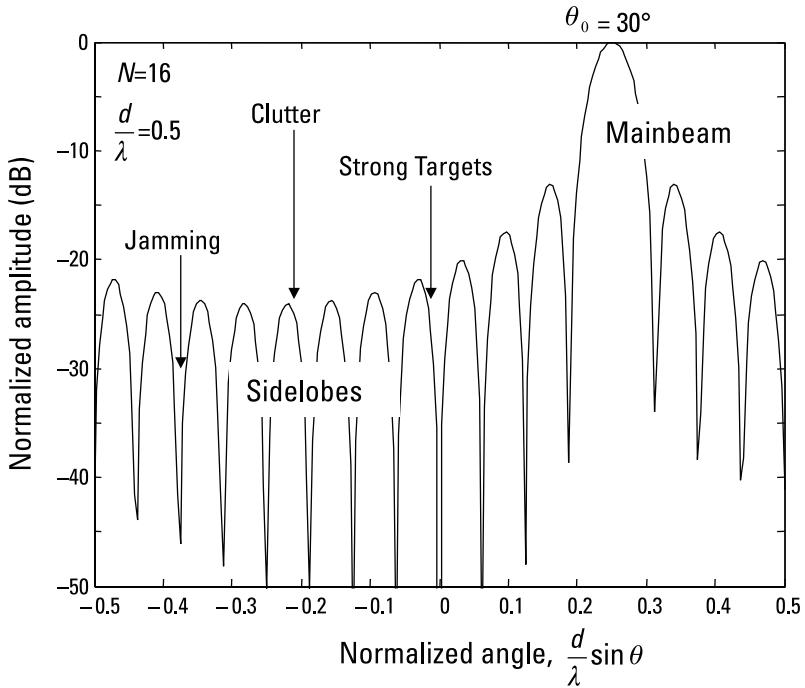


Figure 2.2 ULA beamformer response steered to $\theta_0 = 30^\circ$ (with respect to boresight).

of the main lobe region are referred to as *sidelobes*. For the ULA considered, the first near-in sidelobe is approximately 13 dB down from the peak of the main lobe [1]. The null-to-null width of the main lobe for a ULA depends on the number of elements, N (i.e., the antenna length), and the scan angle, θ_0 . It is easily obtained by setting $N\pi \frac{d}{\lambda} (\sin(\theta) - \sin(\theta_0)) = \pi$, and solving for θ_{MB} , which is the first null. The null-to-null width is thus given by

$$2\theta_{MB} = 2 \sin^{-1} \left[\frac{\lambda}{(Nd)} - \sin(\theta_0) \right] \quad (2.10)$$

which, for modest scan angles [1], can be approximated by

$$2\theta_{MB} = 2 \sin^{-1} \frac{\left[\frac{\lambda}{(Nd)} \right]}{\cos(\theta_0)} \quad (2.11)$$

In modern radar systems, the presence of strong targets, clutter, and jamming in the main lobe or sidelobes can mask the presence of a potentially weaker target of interest. There are basically two approaches to combat these effects: array tapering [1] and adaptive beamforming [4].

Tapering is accomplished by applying an amplitude (positive and real) weighting vector $\mathbf{t} \in \mathbb{R}_+^N$ across the array, such as the 40-dB Chebyshev taper illustrated in Figure 2.3 (for more details on antenna tapering, see [1]). Mathematically, this is tantamount to replacing the untapered beamformer weight, $\mathbf{w} = \mathbf{s}$, with $\mathbf{w} = \mathbf{s} \circ \mathbf{t}$, where \circ denotes the Hadamard product or elementwise matrix multiplication [2], which in general is defined as

$$\begin{aligned} A \circ B &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nm} \end{bmatrix} \circ \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \\ b_{n1} & & & b_{nm} \end{bmatrix} \\ &= \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & & \\ \vdots & & \ddots & \\ c_{n1} & & & c_{nm} \end{bmatrix} \\ &= C \end{aligned} \tag{2.12}$$

where $A, B, C \in \mathbb{C}^{n \times m}$, and the ij th element of C is given by $c_{ij} = a_{ij} b_{ij}$. In words, given two conformal matrices of dimension $n \times m$, the Hadamard product $A \circ B = C$ is an $n \times m$ matrix whose elements are the pairwise products of the corresponding elements of A and B . Note that the output of the beamformer y in (2.7) can be expressed as a Hadamard product of the form

$$y = [1 \ 1 \ \dots \ 1] (\mathbf{w}^* \circ \mathbf{s}) \tag{2.13}$$

where $*$ denotes conjugation without transposition, and $[1 \ 1 \ \dots \ 1]$ is a $1 \times N$ row vector (Hadamard identity row vector) that acts as the summer operator. Note that the Hadamard product in parenthesis must be performed first for the expression to have meaning.

Figure 2.3(b) shows the resulting tapered patterns, normalized to a peak gain of unity, for two Chebyshev weighting schemes. Although sidelobe levels are significantly reduced, there is an increase in the main lobe width and a corresponding loss in gain [1] (see discussion on impact of tapering

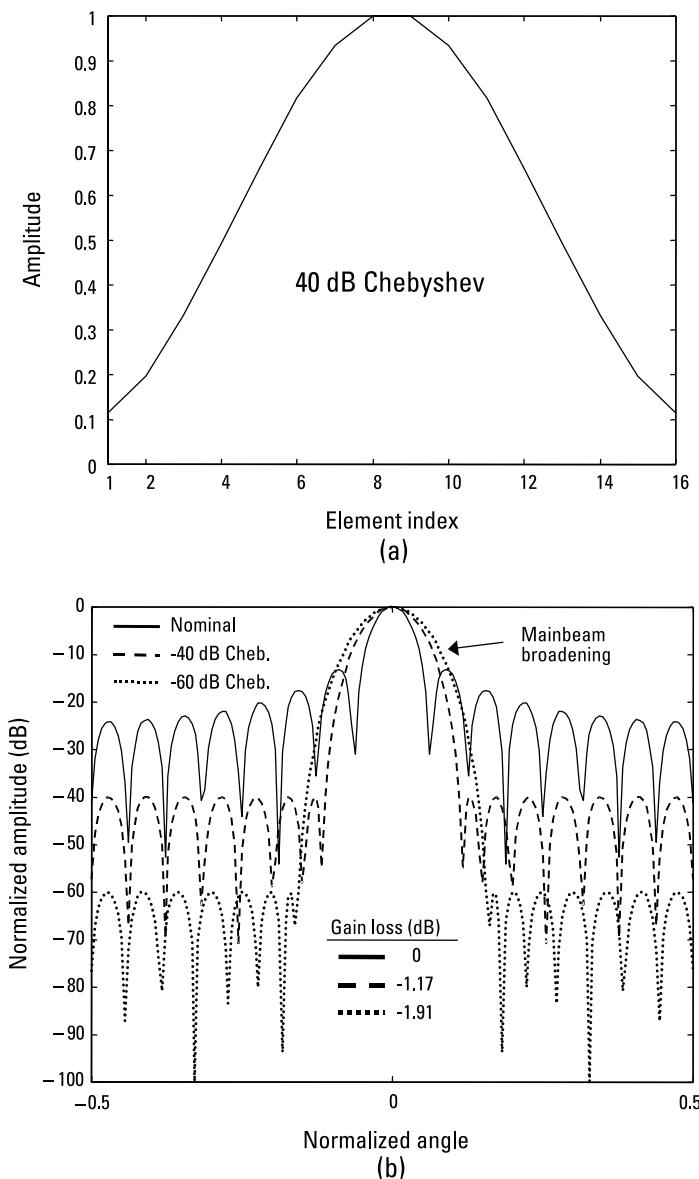


Figure 2.3 (a) Example of a 40-dB Chebyshev taper function for a 16-element ULA, and (b) resulting beam patterns for two different Chebyshev tapers. Note the broadening of the main lobe.

on SNR). Also, there are practical limitations to the amount of tapering possible due to ever-present channel-mismatch errors [1]. Further details on the impact of various forms of channel mismatch are discussed in Chapter 4. As we will see, adaptive beamforming has the potential to overcome some of these disadvantages.

Adaptive beamforming attempts to take advantage of the spatial DoFs afforded by a phased array by adapting the pattern (i.e., the weights, w_n) in real time. The adaptive beamformer thus attempts to steer pattern nulls in the direction of the interference. The nulls must be determined adaptively as, in general, the direction of the interference is not known *a priori*. In practice, adaptive beamforming techniques are the practical attempt to implement optimum beamforming methods. We thus begin by examining the following optimization problem.

2.2.1 Derivation of the Optimum Beamformer

Consider the response of a ULA beamformer to signal $\mathbf{s} \in \mathbb{C}^N$ plus additive interfering signals which we will simply refer to as noise, $\mathbf{n} \in \mathbb{C}^N$. There are many potential sources of noise: clutter, jamming (intentional or not), and ever-present thermal or receiver noise [5]. However, for the moment, we will simply lump these signals into a single term \mathbf{n} . Since the array response is linear, the output is given by

$$y = \mathbf{w}' \mathbf{s} + \mathbf{w}' \mathbf{n} = y_s + y_n \quad (2.14)$$

where \mathbf{n} is a zero-mean, finite-variance, vector *random variable* (RV) [6].

Our objective is to maximize, in some appropriate sense, the signal response while simultaneously minimizing the response due to noise. A ubiquitously employed metric in radar (and many other sensor) applications is SNR [7], which, for the known signal-in-noise case, is defined as

$$\text{SNR} \triangleq \frac{|y_s|^2}{\mathbb{E}\{|y_n|^2\}} \quad (2.15)$$

where $\mathbb{E}\{|y_n|^2\}$ is the expected value of the magnitude squared noise response [6], which formally, is given by [6]

$$\mathbb{E}\{|y_n|^2\} \triangleq \int_{-\infty}^{\infty} |y|^2 f_{y_n}(y) dy \quad (2.16)$$

where $f_{y_n}(y)$ is the *probability density function* (pdf) associated with the RV y_n . SNR is thus a measure of the signal energy to the average noise energy. Since \mathbf{w} is not random, the expectation operation $E\{\cdot\}$ only acts on \mathbf{n} as follows [6]:

$$\begin{aligned} E\{|y_n|^2\} &= E\{|\mathbf{w}'\mathbf{n}|^2\} \\ &= \mathbf{w}'E\{\mathbf{n}\mathbf{n}'\}\mathbf{w} \\ &= \mathbf{w}'R\mathbf{w} \end{aligned} \quad (2.17)$$

where

$$R \triangleq E\{\mathbf{n}\mathbf{n}'\} \in \mathbb{C}^{N \times N} \quad (2.18)$$

is the covariance matrix associated with \mathbf{n} (assumed to be zero-mean). In Chapter 6, we will see that maximizing SNR is also optimum in a rigorous statistical setting for the additive Gaussian noise case.

2.2.2 Case I: Additive White Noise

We first consider the case when the only competing interference is so-called white noise, which is typical of thermal receiver noise [5]. For this case, $R = \sigma^2 I$, where I is the $N \times N$ identity matrix, and the scalar σ^2 is the receiver noise variance, which is assumed to be the same in each channel. For thermal receiver noise, $\sigma^2 = kT_{\text{eff}}B$, where k is Boltzmann's constant $(1.38 \times 10^{-23} \frac{\text{W}}{\text{Hz} \cdot \text{K}^\circ})$, T_{eff} is the effective temperature in absolute degrees Kelvin, and B is the receiver bandwidth in hertz [5]. For this case, we have

$$\max_{\{\mathbf{w}\}} \frac{|\mathbf{w}'\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} = \max_{\{\mathbf{w}\}} \frac{|\mathbf{w}'\mathbf{s}|^2}{\sigma^2 \|\mathbf{w}\|^2} \quad (2.19)$$

$$\text{subject to: } \mathbf{w}'\mathbf{w} = 1$$

The unity norm constraint is chosen for both convenience and to insure a finite solution. Substitution of the constraint into the primary objective function yields the equivalent optimization problem

$$\begin{aligned} \max_{\{\mathbf{w}\}} \frac{|\mathbf{w}'\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} &= \max_{\{\mathbf{w}\}} \frac{|\mathbf{w}'\mathbf{s}|^2}{\sigma^2} \\ &= \max_{\{\mathbf{w}\}} |\mathbf{w}'\mathbf{s}|^2 \end{aligned} \quad (2.20)$$

which is mathematically identical to the maximum-output deterministic beamformer of (2.5); that is,

$$\mathbf{w} = \kappa \mathbf{s} \quad (2.21)$$

where to meet the unity gain constraint in (2.19), $\kappa = 1/\sqrt{N}$ (although this scalar does not effect the SNR as shown below).

This is also known as the matched filter solution for the white-noise case [7]. It is certainly interesting to note that the solution to (2.9), which accounts for noise, is identical to the solution obtained for simply maximizing the beamformer response. Intuitively this makes sense since white noise has no spatial structure and is completely uncorrelated. Consequently, the DoFs are best utilized maximizing the signal response and allowing the summation operation to average out the noise. Moreover, from (2.7), we see that for the white-noise ULA case, the optimum matched filter can be implemented with an FFT. This is certainly an important benefit of the ULA since the FFT greatly facilitates real-time implementation.

The corresponding optimum SNR is obtained by substituting the optimum weight solution into (2.19), and is given by

$$\begin{aligned} \text{SNR}_{\text{opt}} &= \frac{|\mathbf{w}'\mathbf{s}|^2}{\sigma^2 \|\mathbf{w}\|^2} \\ &= \frac{|\kappa \mathbf{s}'\mathbf{s}|^2}{\sigma^2 \kappa^2 \|\mathbf{s}\|^2} \\ &= \frac{\|\mathbf{s}\|^2}{\sigma^2} \end{aligned} \quad (2.22)$$

which is independent of the constant κ . The gain in output SNR achieved by the optimum beamformer is the ratio of output to input SNR

$$\begin{aligned}
 \text{SNR}_{\text{gain}} &= \frac{\text{SNR}_{\text{opt}}}{\text{SNR}_{\text{in}}} \\
 &= \frac{\frac{\|\mathbf{s}\|^2}{\sigma^2}}{\frac{|s_n|^2}{\sigma^2}} \\
 &= \frac{\|\mathbf{s}\|^2}{|s_n|^2}
 \end{aligned} \tag{2.23}$$

where SNR_{in} is the single channel SNR (assumed to be equal for all channels, $n = 1, \dots, N$). Note that the expression is independent of σ^2 , since the gain in SNR is a relative measure and is based on the white-noise assumption (not its strength).

In our previous ULA example, $N = 16$, \mathbf{s} and s_n are given by (2.3), and yield

$$\begin{aligned}
 |s_n|^2 &= \left| e^{j2\pi(n-1)\frac{d}{\lambda} \sin \theta_0} \right|^2 = 1 \\
 \text{and} \\
 |\mathbf{s}|^2 &= \mathbf{s}'\mathbf{s} = 16
 \end{aligned} \tag{2.24}$$

which corresponds to $\text{SNR}_{\text{gain}} = 16 = N$ (or, in decibels, $10 \log 16 = 12.04$ dB). Thus, in general, for an N -element ULA, the white-noise coherent integration gain of a matched filter is given by

$$\text{SNR}_{\text{gain_dB}} = 10 \log N \text{ (dB)} \tag{2.25}$$

If a taper or window function is employed [1], there is a loss in SNR since the weight vector \mathbf{w} is no longer perfectly matched to the signal \mathbf{s} of interest. It is straightforward to calculate this loss in SNR as follows. First, define the taper efficiency ratio, η_{eff} , as the ratio of the tapered output SNR (SNR_{tap}) to optimal SNR (SNR_{opt}); that is,

$$\begin{aligned}\eta_{\text{eff}} &= \frac{\text{SNR}_{\text{tap}}}{\text{SNR}_{\text{opt}}} \\ &= \frac{|\mathbf{w}'\mathbf{s}|^2}{N}\end{aligned}\quad (2.26)$$

where (2.24) was invoked. For the tapered case, $\mathbf{w} = \mathbf{s} \odot \mathbf{t}$, where to insure the unity gain constraint $\mathbf{w}'\mathbf{w} = 1$, \mathbf{t} must satisfy

$$\begin{aligned}\mathbf{w}'\mathbf{w} &= (\mathbf{s} \odot \mathbf{t})'(\mathbf{s} \odot \mathbf{t}) \\ &= (\mathbf{s}^* \odot \mathbf{s})'(\mathbf{t}^* \odot \mathbf{t}) \\ &= [1 \ \dots \ 1](\mathbf{t}^* \odot \mathbf{t}) \\ &= \mathbf{t}'\mathbf{t} \\ &= \|\mathbf{t}\|_2^2 = 1\end{aligned}\quad (2.27)$$

where $\|\mathbf{t}\|_2$ denotes the usual Euclidean (L_2 -norm) vector norm $\|\mathbf{t}\|_2 \triangleq \sqrt{\mathbf{t}'\mathbf{t}}$ [2]. Substituting $\mathbf{w} = \mathbf{s} \odot \mathbf{t}$ into (2.26) yields

$$\begin{aligned}\eta_{\text{eff}} &= \frac{|(\mathbf{s} \odot \mathbf{t})'\mathbf{s}|^2}{N} = \frac{|\mathbf{s}'(\mathbf{s} \odot \mathbf{t})|^2}{N} \\ &= \frac{|(\mathbf{s}^* \odot \mathbf{s})'\mathbf{t}|^2}{N} \\ &= \frac{|[1 \ \dots \ 1]\mathbf{t}|^2}{N} \\ &= \frac{\|\mathbf{t}\|_1^2}{N}\end{aligned}\quad (2.28)$$

where $\|\mathbf{t}\|_1$ denotes the L_1 -norm (sum of the absolute values of \mathbf{t} [2]—recall that $[1 \ \dots \ 1]\mathbf{t}$ is the summation operator, which for a positive taper is also the L_1 -norm). Note that an elementary property of Hadamard products was also invoked, namely, $\mathbf{s}'(\mathbf{s} \odot \mathbf{t}) = (\mathbf{s}^* \odot \mathbf{s})'\mathbf{t}$ (which is readily verified by an elementwise comparison of both sides of the equality [8]).

For the 16-element 40-dB Chebyshev taper shown in Figure 2.3(a), $\|\mathbf{t}\|_1 = 3.5$, which when substituted into (2.28) yields $\eta_{\text{eff}} = 0.7642$ (-1.17 dB). Note that for the uniform taper case (i.e., no taper), $\|\mathbf{t}\|_1 = \sqrt{N}$ (recall

normalization constraint) which yields, as expected, $\eta_{\text{eff}} = 1$ (0 dB). It is interesting that for the uniform case, $\|\mathbf{t}\|_1 = \|\mathbf{t}\|_2$.

2.2.3 Case II: Additive Colored Noise

In general the additive noise term, \mathbf{n} , consists of both white noise and colored noise (has correlation and/or spatial structure) due to the presence of structured interference sources such as clutter and jamming [7]. Thus, in general the total noise (interference plus white noise) covariance matrix R will not typically be diagonal, but will have a more complex structure. However, assuming only that R is positive-definite, a property guaranteed in practice since receiver noise is always present [7], a closed form solution for the optimum beamformer can still be obtained. Again, our objective is to maximize the SINR:

$$\max_{\{\mathbf{w}\}} \frac{|\mathbf{w}'\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} \quad (2.29)$$

subject to: $\mathbf{w}'\mathbf{w} = 1$

Noting that $R^{1/2}R^{-1/2} = I$, and again applying Schwarz's inequality, we have the following inequality:

$$\begin{aligned} \frac{|\mathbf{w}'\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} &= \frac{|\mathbf{w}'R^{1/2}R^{-1/2}\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} \\ &\leq \frac{(\mathbf{w}'R\mathbf{w})(\mathbf{s}'R^{-1}\mathbf{s})}{\mathbf{w}'R\mathbf{w}} \\ &= \mathbf{s}'R^{-1}\mathbf{s} \end{aligned} \quad (2.30)$$

or more simply

$$\frac{|\mathbf{w}'R^{1/2}R^{-1/2}\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} \leq \mathbf{s}'R^{-1}\mathbf{s} \quad (2.31)$$

Equality is achieved when $R^{1/2}\mathbf{w} = \kappa R^{-1/2}\mathbf{s}$ (colinearity requirement), or equivalently,

$$\mathbf{w} = \kappa R^{-1} \mathbf{s} \quad (2.32)$$

where the Hermitian property (complex conjugate transpose symmetry [2]) of R has been invoked. Note that since the upper bound in (2.31) is independent of \mathbf{w} , (2.32) is strictly optimum, and achieves the upper bound; that is,

$$\text{SINR}_{\text{opt}} = \mathbf{s}' R^{-1} \mathbf{s} \quad (2.33)$$

which reduces to (2.22) for the white-noise case ($R = \sigma^2 I$).

Since R is of the form

$$R = R_I + \sigma^2 I \quad (2.34)$$

where R_I is a generally positive-semidefinite covariance matrix [7] associated with the colored-noise source (i.e., clutter and/or jamming) that is uncorrelated with the white noise, SINR_{opt} is bounded by the white-noise SNR_{opt} result (2.22) (i.e., thermal noise limited performance).

To illustrate the colored-noise case, consider the effect that a persistent interfering point source (i.e., a jammer) has on our example ULA considered earlier. For this case,

$$\mathbf{n} = \mathbf{n}_J + \mathbf{n}_{Rx} \quad (2.35)$$

where $\mathbf{n}_J \in \mathbb{C}^N$ and $\mathbf{n}_{Rx} \in \mathbb{C}^N$ are the jammer and receiver noise vector RV signals, respectively [4]. The simplest model for a jammer (or strong cochannel interferer) is $\mathbf{n}_J = \tilde{\gamma} \mathbf{s}_J$, where $\tilde{\gamma}$ is a zero-mean, complex scalar RV, assumed to be uncorrelated with the receiver noise (i.e., $E\{\mathbf{n}_J \mathbf{n}_{Rx}'\} = \emptyset$), with a variance (jammer power)

$$E\{|\bar{\gamma}|^2\} = \sigma_J^2 \quad (2.36)$$

at the receiver input (per element), and \mathbf{s}_J is the jammer steering vector associated with a plane wave AoA of θ_J [see (2.3)] [4]. Therefore, for the single jammer case,

$$\begin{aligned} R &= E\{\mathbf{n}\mathbf{n}'\} \\ &= E\{\mathbf{n}_J \mathbf{n}_J'\} + E\{\mathbf{n}_{Rx} \mathbf{n}_{Rx}'\} \\ &= \sigma_J^2 \mathbf{s}_J \mathbf{s}_J' + \sigma^2 I \end{aligned} \quad (2.37)$$

Note that the matrix $\sigma_j^2 \mathbf{s}_j \mathbf{s}'_j$ is rank one and is thus singular (but positive-semidefinite since $\sigma_j^2 > 0$). In general, if there are N_J uncorrelated jammers, R has the form [4]

$$R = \sum_{i=1}^{N_J} \sigma_{J_i}^2 \mathbf{s}_i \mathbf{s}'_i + \sigma^2 I \quad (2.38)$$

If, furthermore, the jammer steering vectors are linearly independent [2], then

$$\text{rank}\left(\sum_{i=1}^{N_J} \sigma_{J_i}^2 \mathbf{s}_i \mathbf{s}'_i\right) = N_J \quad \text{for } N_J \leq N \quad (2.39)$$

As an example, consider the effect of six uncorrelated jammers impinging on the aforementioned 16-element ULA, shown in Figure 2.4. Assuming each interferer produces an element *jammer-to-noise ratio* (JNR) of 50 dB (i.e., $10 \log_{10} (\sigma_j^2 / \sigma^2) = 50$, for $j = 1, \dots, N_J$), and that the single element

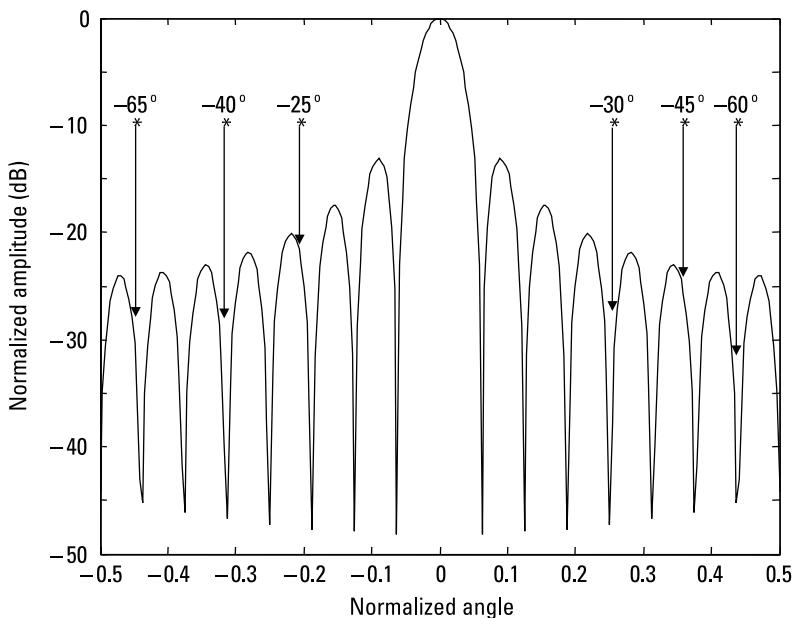


Figure 2.4 Illustration of the angular distribution of the six-jammer example.

SINR is 0 dB (or equivalently, $\mathbf{s}'\mathbf{s} = 16$), the resulting quiescent beamformer output SINR_q with $\theta_0 = 0^\circ$ (i.e., $\mathbf{s} = [1 \ 1 \ \dots \ 1]^T$) is

$$\begin{aligned}\text{SINR}_q &= \frac{|\mathbf{w}'\mathbf{s}|^2}{\mathbf{w}'R\mathbf{w}} \\ &= \frac{|\mathbf{s}'\mathbf{s}|^2}{\mathbf{s}'R\mathbf{s}} \\ &= 6.3 \times 10^{-4} \ (-32 \text{ dB})\end{aligned}\tag{2.40}$$

This implies that any signal of interest at $\theta_0 = 0^\circ$ would be 32 dB lower than the output response due to interference—a generally intolerable condition. If, on the other hand, we implement optimum beamforming, there is a dramatic improvement. Specifically,

$$\begin{aligned}\text{SINR}_{\text{opt}} &= \mathbf{s}'R^{-1}\mathbf{s} \\ &= 15.7\end{aligned}\tag{2.41}$$

or 12.0 dB, a loss of only ~ 0.1 dB relative to the white-noise case. Even this simple example highlights the significant potential of optimum beamforming. Of course, if even one of the jammers had an AoA of $\theta_0 = 0^\circ$, the optimum beamformer would be unable to null out the jammer and maximize the signal response simultaneously. There would thus be no net gain in SINR relative to that jammer.

Figure 2.5 shows the corresponding optimum antenna response for the six-jammer case. Note that nulls or notches have been introduced into the pattern at precisely the angles the jammers are present. Additionally, the pattern still has a distinct and well-formed main lobe to insure maximum response to desired signals at boresight.

As mentioned previously, it is possible to combine aperture tapering (e.g., Chebyshev windowing) with optimum beamforming by including in (2.32) the desired aperture taper \mathbf{t} as follows

$$\mathbf{w} = \kappa R^{-1}(\mathbf{s} \odot \mathbf{t})\tag{2.42}$$

Figure 2.6 illustrates an example of (2.42) and corresponds to the optimum pattern of Figure 2.5 modified by the presence of a 30-dB Chebyshev taper. Of course, as previously discussed, the inclusion of a taper mismatch

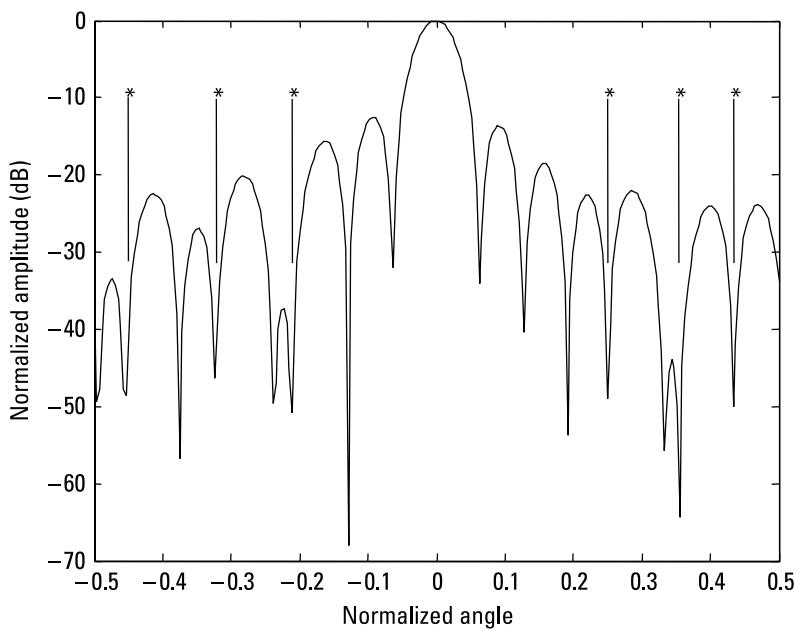


Figure 2.5 Optimum beam pattern for the six-jammer case of Figure 2.4.

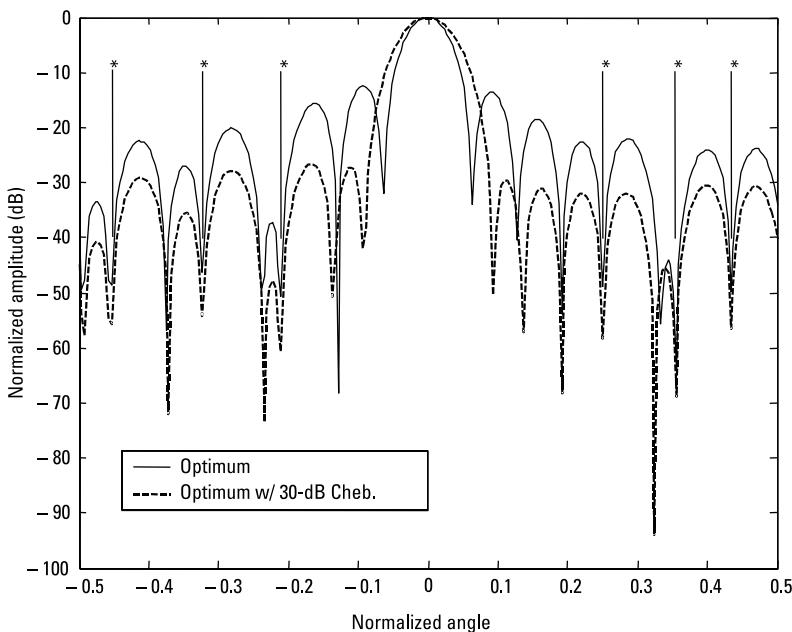


Figure 2.6 Example of combining tapering with optimal beamforming.

introduces additional SINR loss (in this case resulting output SINR is 11.3 dB, a loss of 0.7 dB relative to the optimum case). However, in practice, the practical benefits of maintaining low sidelobes is often well worth the modest loss.

An important fundamental question is, how many jammers can be nulled with an N -element array? If we let \mathbf{s}_{J_i}' denote the steering vector associated with the i th jammer, then a null is achieved if $\mathbf{s}_{J_i}'\mathbf{w} = 0$. To insure that a target of interest is not inadvertently nulled, we need to also invoke a so-called mainbeam constraint of the form $\mathbf{s}'\mathbf{w} = 1$ (sometimes referred to as a unity gain constraint). Assuming that there are N_J linearly independent jammer steering vectors (independent of \mathbf{s} as well), the following linear system of equations results:

$$\begin{aligned}\mathbf{s}'\mathbf{w} &= 1 \\ \mathbf{s}_{J_1}'\mathbf{w} &= 0 \\ \mathbf{s}_{J_2}'\mathbf{w} &= 0 \\ &\vdots \\ \mathbf{s}_{J_{N_J}}'\mathbf{w} &= 0\end{aligned}\tag{2.43}$$

which is a linear system of $N_J + 1$ equations in N unknowns. From elementary linear algebra [3], we know that there exists a \mathbf{w} (not necessarily unique) that satisfies the above set of equations provided that $N_J + 1 \leq N$. This, in turn, immediately implies that the number of jammers that can be nulled is bounded by

$$N_J \leq N - 1\tag{2.44}$$

The above optimum beamformer calculations were highly idealized as they assumed exact knowledge of the total interference covariance matrix, as well as perfectly matched channels and idealized jammer characteristics—conditions not met in practice. Much of this book is devoted to strategies designed to preserve these desirable gains in more realistic operating environments. We next turn our attention to the other significant measurement dimension, namely time. More specifically, we will consider pulse-to-pulse time processing manifested as Doppler filtering, which plays a fundamental role in MTI radar.

2.3 Optimum Temporal (Doppler/Pulse) Processing

Besides AoA, another key physical observable for separating moving targets from noise is Doppler frequency [5, 9, 10]. Fortunately, the mathematical framework employed in modern Doppler radar systems is essentially identical to that already introduced in our discussion of the ULA; specifically, a spatial point target produces a linear phase progression across a ULA. We will see that a Doppler-shifted point target similarly produces a linear phase progression in time and can thus be processed in a manner directly analogous to the beamformer case.

Consider the effect of a Doppler-shifted return propagating through a single channel, M -tap delay-line filter as shown in Figure 2.7. For a pulsed Doppler radar [5], the delay T is chosen to match the PRI. The output of the m th tap is then given by $s_m = e^{j2\pi(m-1)\bar{f}_d t}$, $m = 1, \dots, M$, where \bar{f}_d is the normalized Doppler frequency given by [5]

$$\begin{aligned}\bar{f}_d &= \frac{f_d}{\text{PRF}} \\ &= f_d T \\ &= \frac{2T}{\lambda} (\mathbf{v}_{tgt} - \mathbf{v}_{Rx}) \cdot \hat{\mathbf{i}}_{Rx}\end{aligned}\quad (2.45)$$

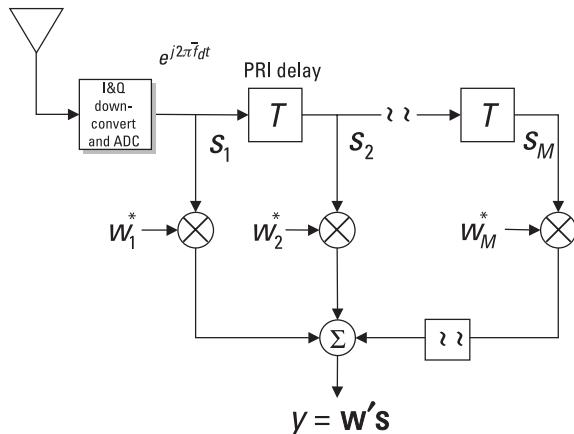


Figure 2.7 Uniform tapped delay-line linear combiner for processing Doppler-shifted returns. Note that the combiner output has the same mathematical form as the ULA beamformer of Figure 2.1.

where PRF is the pulse repetition frequency ($\text{PRF} = 1/\text{PRI} = 1/T$), f_d is the Doppler frequency of the point target in units of 1/time (e.g., hertz), \mathbf{v}_{tgt} and \mathbf{v}_{Rx} are the target and receiver velocity vectors (Cartesian coordinates), $\hat{\mathbf{i}}_{Rx}$ is a unit direction vector pointing from the receiver to the target ($|\hat{\mathbf{i}}_{Rx}| = 1$), λ is the operating wavelength, and \cdot denotes the vector “dot” product; that is, $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$, where θ is the angle between \mathbf{a} and \mathbf{b} . Note for a fixed PRF, the unambiguous Doppler region for complex sampling based on the Nyquist criterion is from $-\text{PRF}/2$ to $+\text{PRF}/2$ [5]. Thus, the unambiguous normalized Doppler region is

$$-0.5 \leq \bar{f}_d \leq +0.5 \quad (2.46)$$

Note that (2.45) is only valid (in general) for the monostatic case (transmitter and receiver collocated).

As with the optimum ULA beamformer, an optimized Doppler filter response can also be constructed by a judicious selection of the complex weighting factors w_m , $m = 1, \dots, M$, in Figure 2.7. Let

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_M \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j2\pi\bar{f}_d} \\ \vdots \\ e^{j2\pi(M-1)\bar{f}_d} \end{bmatrix} \quad (2.47)$$

denote the M -tap Doppler steering vector for a signal (Doppler) of interest and $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_M]^T$ denote the vector of weights. Then, as with the ULA case, we desire to choose \mathbf{w} to maximize SINR. Since the math is identical, the optimum SINR solution is also given by (2.32); namely,

$$\mathbf{w} = \kappa R^{-1} \mathbf{s} \quad (2.48)$$

where R is the $M \times M$ total noise covariance matrix, and κ is a scalar constant that does not affect SINR.

We will defer examples and rationale of optimum Doppler processing until Chapter 3, where ground clutter will be discussed. In this context, we will also introduce the notion of 2-D angle-Doppler (or space-time) processing. Next, the notion of adaptive, vice-optimum processing is discussed for the previously introduced 1-D beamformer.

2.4 Adaptive 1-D Processing

Unfortunately, in much of the radar and signal processing literature the expression for the optimum linear combiner (beamformer) is referred to as an adaptive linear combiner (or adaptive beamformer, and so forth). This equivalence is not only potentially misleading and confusing, it is in general incorrect—and this is not just pedantry. In real-world systems, adaptivity is used to adjust radar processing on the fly. The reasons for this are obvious. In the case of jamming, for example, one cannot know *a priori* both the strength and exact angular location. Similar arguments are true for airborne radar clutter.

This lack of prior information implies that the essential total interference covariance matrix R of (2.32) must be estimated concurrently with the beamforming operation. To accomplish this, some form of sample estimate, derived from actual radar measurements, must be either explicitly or implicitly formed. Referred to somewhat generically as *training* (the term can be traced to Bernard Widrow's pioneering work on adaptive neural networks [11]), it is precisely this adaptation process that is a major impetus for this book, many of the techniques described herein, and much of the research literature [12].

Historically, adaptation to interference was accomplished by the explicit or implicit estimation of the local noise statistics. Beginning with *cell averaging constant false-alarm rate* (CACFAR) techniques, developed concurrently with the very first airborne MTI radars [13], adaptation has been accomplished by identifying a subset of the radar measurements from which noise statistics are estimated. Figure 2.8 presents a pictorial representation of this process. Let \mathbf{x}_i denote the N -dimensional output vector of a ULA corresponding to the i th range bin. To process \mathbf{x}_i according to (2.32), that is, to look for a target in a given direction, \mathbf{s} , while simultaneously minimizing the influence of all other interfering signals, the covariance matrix for that i th range bin R_i is required.

A reasonable approach is to compute a sample estimate of R_i (which we will denote as \hat{R}_i) using surrounding range bins (without the cell under test to prevent so-called self-nulling) that are likely to contain similar interference; that is,

$$\hat{R}_i = \frac{1}{L} \sum_{l \in \Omega_i} \mathbf{x}_l \mathbf{x}'_l \quad (2.49)$$

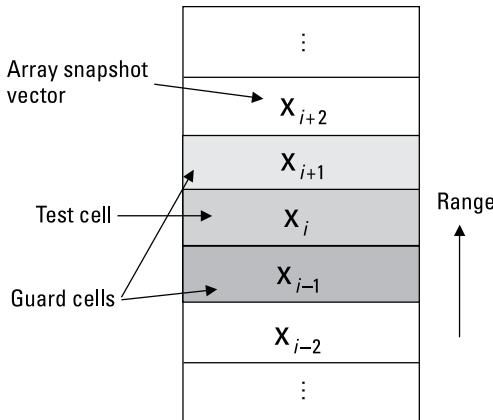


Figure 2.8 Conventional training strategy for estimating interference covariance matrix.

where L is the total number of samples used in the summation, and Ω_i denotes the set of training samples used for the i th range bin. It is easy to show that if the data samples $\{\mathbf{x}_l\}$, $l = 1, \dots, L$, are uncorrelated and have identical covariance R , (2.49) is an unbiased estimate of R , since

$$\begin{aligned} E\{\hat{R}_i\} &= E\left\{\frac{1}{L} \sum_{l \in \Omega_i} \mathbf{x}_l \mathbf{x}_l'\right\} \\ &= \frac{1}{L} \sum_{l \in \Omega_i} E\{\mathbf{x}_l \mathbf{x}_l'\} \\ &= \frac{1}{L} \sum_{l \in \Omega_i} R \\ &= R \end{aligned} \quad (2.50)$$

If, additionally, $\{\mathbf{x}_l\}$, $l = 1, \dots, L$, are Gaussian and *independent identically distributed* (i.i.d.), than (2.49) corresponds to the maximum likelihood estimate of R_i [7] as shown in Appendix 2B.

Of course in practice, there is no guarantee that the interference in the i th range bin will have statistics identical with the training data. Nonetheless, it is a reasonable first choice assuming that only range bins from surrounding, and thus hopefully similar, regions are utilized. Another practical constraint on the training data selection process is the presence of a buffer region around the cell under test comprised of so-called guard cells, as shown

in Figure 2.8. This is sometimes included to prevent so-called self-nulling, since a target response may straddle more than one range bin or have temporal sidelobes present in other range bins, a common problem in pulse compression radars [14].

It is important to realize that if (2.49) is used in (2.32), then a minimum of N i.i.d. samples are required to insure that \hat{R}_i is nonsingular (almost surely [6]). However, under certain circumstances, as we will see later in this and subsequent chapters, very useful adaptive solutions can be obtained with significantly fewer than N samples. This can be a very desirable practical feature since, in the case of clutter, statistics can vary significantly with range, thereby invalidating the necessary stationarity assumption tacitly employed in (2.49).

Reed, Mallet, and Brennan (RMB) [15], in a seminal paper, were able to rigorously characterize the impact of replacing the actual covariance, R_i , with its sample estimate, \hat{R}_i , under certain conditions. Specifically, if the L training samples $\Omega_i = \{\mathbf{x}_j\}_j$ are free of target signal contamination and are independent i.i.d. Gaussian vector RVs with statistics identical to the i th test cell (not included in the training data), then the following are true: (1) the elements of \hat{R}_i are complex Wishart-distributed [15], and (2) a closed form expression for the ratio ρ of the expected SINR using (2.23) to optimum SINR using (2.32) is beta-distributed and is given by [15]

$$\rho = \frac{L - N + 2}{L + 1}, \quad L \geq N \quad (2.51)$$

This is the important RMB result, which states that the SMI method produces an SINR loss that is ~ 3 dB if $L \sim 2N$. Note that (2.51) is also extremely useful for estimating the amount of loss due to finite L . Referred to colloquially as RMB loss, this equation has become an integral part of a radar systems engineer's SINR budget spreadsheets. As expected, $\rho \rightarrow 1$ as $L \rightarrow \infty$. An elegant and accessible derivation of this result can be found in the chapter by Steinhardt in [16].

Although $2N$ i.i.d. Gaussian samples yield an SINR that is within ~ 3 dB of optimum, the corresponding adapted pattern may not be suitable for most situations due to pattern distortions. For example, Figure 2.9 displays the adapted pattern based on $2N$ samples for the $N = 16$ element ULA example of Figure 2.6, along with the optimum tapered pattern. Note the presence of extremely high sidelobes and shape-distortion of the main lobe. Although the adapted SMI pattern can be improved by increasing L (note the response in Figure 2.9 corresponding to $5N$ samples), there are

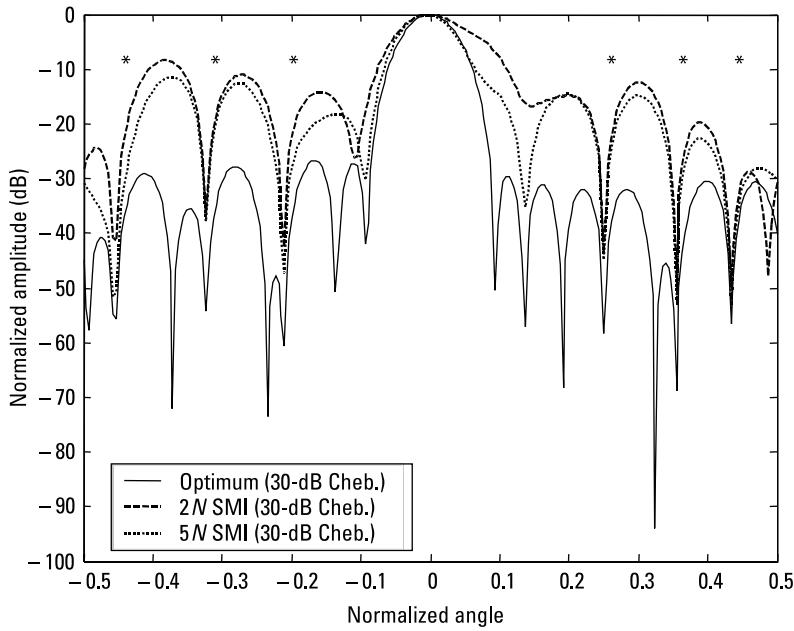


Figure 2.9 Effect of covariance estimation error due to a finite data sample on adapted beam patterns.

many practical limits to doing so due to interference nonstationarity and/or computational burden. Instead, it is worth examining the cause of the adapted pattern distortion at this stage by conducting an eigen analysis of the underlying adaptation process. In so doing, we will also lay the groundwork for an important class of robust adaptive signal-processing techniques based on optimum subspace estimation and principal components [17, 18].

We begin by first considering an eigendecomposition of a general Hermitian, positive-definite, N -dimensional covariance matrix R ; that is,

$$R = U \Lambda U' = \sum_{n=1}^N \lambda_n \mathbf{u}_n \mathbf{u}_n' \quad (2.52)$$

where

$$U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_N] \in \mathbb{C}^{N \times N} \quad (2.53)$$

is the $N \times N$ matrix whose columns are the orthonormal eigenvectors of R , and

$$\Lambda = \text{diag}\{\lambda_1 \ \lambda_2 \ \dots \ \lambda_N\} \in \mathbb{R}_+^{N \times N} \quad (2.54)$$

is the $N \times N$ diagonal matrix of corresponding strictly positive eigenvalues [8].

Substituting (2.52) into (2.32) and recalling that

$$\begin{aligned} R^{-1} &= (U\Lambda U')^{-1} \\ &= U\Lambda^{-1}U' \\ &= \sum_{n=1}^N \frac{\mathbf{u}_n \mathbf{u}_n'}{\lambda_n} \end{aligned} \quad (2.55)$$

(U is a unitary matrix; that is, $U^{-1} = U'$ [8]), and rearranging terms yields

$$\begin{aligned} \mathbf{w} &= \kappa R^{-1} \mathbf{s} \\ &= \tilde{\kappa} \lambda_{\min} U \Lambda^{-1} U' \mathbf{s} \\ &= \tilde{\kappa} (\mathbf{s} - \mathbf{s} + U \tilde{\Lambda}^{-1} U' \mathbf{s}) \\ &= \tilde{\kappa} [\mathbf{s} - (I - U \tilde{\Lambda}^{-1} U') \mathbf{s}] \\ &= \tilde{\kappa} [\mathbf{s} - U(I - \tilde{\Lambda}^{-1}) U' \mathbf{s}] \end{aligned} \quad (2.56)$$

where

$$\tilde{\kappa} = \kappa / \lambda_{\min} \quad \text{and} \quad \tilde{\Lambda} = (1 / \lambda_{\min}) \Lambda \quad (2.57)$$

and λ_{\min} denotes the minimum eigenvalue of R . Thus,

$$\tilde{\Lambda} = \text{diag}\{\tilde{\lambda}_1 \ \tilde{\lambda}_2 \ \dots \ \tilde{\lambda}_N\} \quad (2.58)$$

where $\tilde{\lambda}_n = \lambda_n / \lambda_{\min}$ is the n th normalized eigenvalue.

Finally, expressing (2.56) as a summation yields a very useful equivalent representation of the optimum beamformer result expressed in eigen form, namely,

$$\begin{aligned}\mathbf{w} &= \tilde{\kappa} \left[\mathbf{s} - \sum_{n=1}^N \left(1 - \frac{1}{\bar{\lambda}_n} \right) (\mathbf{u}'_n \mathbf{s}) \mathbf{u}_n \right] \\ &= \tilde{\kappa} \left[\mathbf{s} - \sum_{n=1}^N \left(\frac{\lambda_n - \lambda_{\min}}{\lambda_n} \right) (\mathbf{u}'_n \mathbf{s}) \mathbf{u}_n \right]\end{aligned}\quad (2.59)$$

Equation (2.59) shows that the optimum beamformer weight vector is formed by subtracting a weighted sum of the interference eigenvectors from the quiescent steering vector \mathbf{s} . Note that only the eigenvectors with corresponding eigenvalues greater than λ_{\min} may contribute to the sum. The exact amount they contribute is proportional to the difference between their eigenvalues and the minimum eigenvalue term, as well as the amount of colinearity [i.e., $(\mathbf{u}'_n \mathbf{s})$].

To explain the distorted pattern results of Figure 2.9, consider first the eigendecomposition of the exact covariance for the six-jammer, 16-element ULA example. Figure 2.10 shows the eigenvalue distribution; there are exactly six large eigenvalues (due to the jammers). The remaining 10 eigenvalues are often referred to as the noise-floor components. To see why

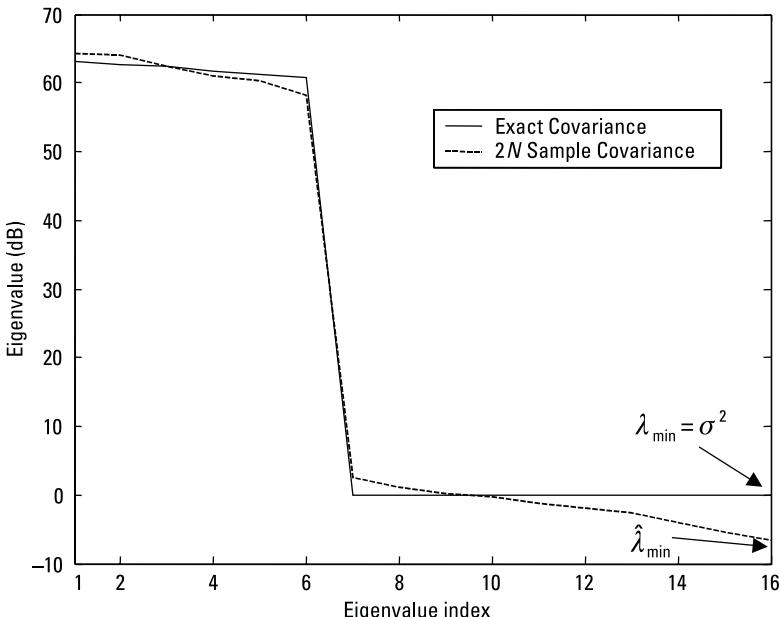


Figure 2.10 Comparison of exact and estimated eigenvalues for the six-jammer example.

this is so, consider the structure of the exact covariance described by (2.38); that is,

$$R = \sum_{i=1}^{N_J} \sigma_{J_i}^2 \mathbf{s}_i \mathbf{s}'_i + \sigma^2 I = R_J + \sigma^2 I \quad (2.60)$$

where

$$R_J \triangleq \sum_{i=1}^{N_J} \sigma_{J_i}^2 \mathbf{s}_i \mathbf{s}'_i \quad (2.61)$$

For the current example $N_J = 6$. Thus, the eigendecomposition of R_J is of the form

$$\begin{aligned} R_J &= U_J \Lambda_J U'_J \\ &= U_J \text{diag}\{\lambda_{J_1} \ \lambda_{J_2} \ \dots \ \lambda_{J_6} \ 0 \ \dots \ 0\} U'_J \end{aligned} \quad (2.62)$$

where there are six nonzero eigenvalues (the rank of R_J is exactly six since the jammers are linearly independent). Although the remaining 10 eigenvalues are zero, we can still use the Gram-Schmidt method to construct a complete 16-dimensional orthogonal basis. Thus, we can assume that U_J is full rank and, therefore, that $U_J^{-1} = U'_J$.

Since any unitary matrix diagonalizes a diagonal matrix [8], the eigendecomposition of the total covariance matrix, R , is given by

$$R = U \Lambda U' = U_J \Lambda U'_J \quad (2.63)$$

The eigenvalues of R are thus given by

$$\begin{aligned} \Lambda &= U'_J R U_J \\ &= U'_J (R_J + \sigma^2 I) U_J \\ &= U'_J R_J U_J + \sigma^2 I \\ &= \text{diag}\{\lambda_{J_1} + \sigma^2 \ \dots \ \lambda_{J_6} + \sigma^2 \ \sigma^2 \ \dots \ \sigma^2\} \end{aligned} \quad (2.64)$$

From (2.64) we see that $\lambda_{\min} = \sigma^2$ and that only the first six eigenvalues are greater than λ_{\min} . Thus, there are only six terms that contribute to the sum in (2.59) for the ideal case (known-covariance).

Now consider the eigenvalues of the estimated covariance \hat{R} , shown in Figure 2.10, for an example in which $L = 2N$ i.i.d. Gaussian samples are used in (2.49). Note that while there are certainly six large distinct eigenvalues, there is no definitive noise floor. More specifically, since the first 15 estimated eigenvalues $\{\hat{\lambda}_1, \dots, \hat{\lambda}_{15}\}$ are greater than the minimum eigenvalue $\hat{\lambda}_{16}$, they may all contribute in (2.59). Unfortunately, except for the very strong jammer eigenvalues, the eigenvalues/vector pairs associated with the noise floor are not well estimated with only $2N$ samples. The biggest impact, as illustrated in Figure 2.9, is significant corruption of the adapted pattern sidelobes.

Although we will defer an in-depth discussion of principal components methods until Chapter 4, it is clear from the above discussion that if we only used the six dominant eigenvalues/vector pairs in (2.59) and set $\hat{\lambda}_{\min} = \sigma^2$ (since the receiver noise floor is reasonably well-known in a calibrated radar), we should have a much better pattern. Indeed, as shown in Figure 2.11, the adapted pattern using only the six principal components in (2.59) is extremely close to the optimum. The corresponding SINR loss relative to the optimum tapered response is only 0.8 dB (for $L = 2N$).

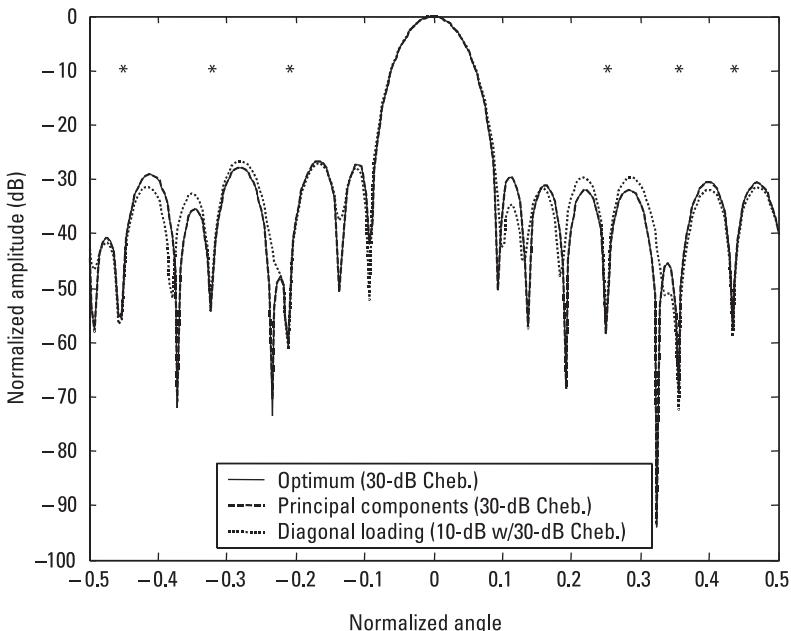


Figure 2.11 Adapted patterns using principal components and diagonal loading. Note that the principal components pattern is virtually identical to the optimum pattern.

Figure 2.11 also displays the adapted pattern obtained by adding a diagonal loading term [19] to \hat{R} ; that is,

$$\hat{R}_{DL} = \hat{R} + \delta^2 I \quad (2.65)$$

where $\delta^2 = 10\sigma^2$ was chosen (i.e., 10 dB of diagonal loading). Again the adapted pattern is quite good compared with SMI. The reason diagonal loading works somewhat similarly to the principal components method is clear from an examination of the eigenvalues of \hat{R}_{DL} , which are

$$\hat{\Lambda}_{DL} = \text{diag}\{\hat{\lambda}_1 + \delta^2, \hat{\lambda}_2 + \delta^2, \dots, \hat{\lambda}_{16} + \delta^2\} \quad (2.66)$$

Since, in the example considered, $\hat{\lambda}_n \ll \delta^2$, for $n = 7, \dots, 16$, we again see that the 10 smallest eigenvalues/vectors will not appreciably contribute to the sum of (2.59). However, the depth of the nulls is not the same for diagonal loading as it is for principal components because the weighting factors in (2.59) have been decreased (uniformly). This resulting decrease in null depth, which can be an issue in some applications, highlights some of the limitations of diagonal loading (not to mention the issue of choosing a correct value of loading factor [20]). Note also that the corresponding SINR loss relative to the optimum tapered response is slightly worse (0.9 dB) than that for principal components, but is still quite good compared to SMI.

2.5 Adaptivity in Nonstationary Environments

The availability of a sufficiently large training set of i.i.d. samples is highly suspect in stressing applications involving large numbers of adaptive DoFs. Variability in clutter terrain and other interference nonstationarities place practical limits on the size and quality of the training data available [19]. Thus, an extremely important issue is how many adaptive DoFs can be utilized or supported in a given interference environment.

The answer, as we have already seen, is highly dependent on the specific adaptation algorithm employed. For example, in the principal components method, we saw that instead of supporting 16 adaptive DoFs, we only needed to estimate the subspace associated with the interference, which was six-dimensional. Of course, we somehow needed to know that only six eigenvalues were required. For this example, a reasonably simple criterion could

be used based on the size of the estimated eigenvalue relative to the known noise floor. When clutter interference is introduced in the next chapter, we will see that there is often a continuum of eigenvalues with monotonic decreasing amplitudes down to the noise floor, thus complicating the determination of a simple threshold approach. Nonetheless, by a judicious incorporation of prior structural information, an order of magnitude or greater reduction in sample support requirements can be achieved while preserving near optimal performance.

While minimizing sample support requirements is an indirect way of addressing nonstationarity, there are other methods that attempt to directly model or parameterize the nonstationarity. Examples of this approach include the so-called time varying weights technique [21], which has been proposed for both bistatic and circular array STAP applications, and CMT, which models the subdominant rank inflation mechanisms that give rise to nonstationarity in interference rank [22].

Although we will discuss CMTs at much greater length in the context of clutter cancellation and STAP in Chapters 4 and 5, they have also proved straightforward to apply in spatial-only adaptivity for the so-called stale weights problem [23–27]—a form of nonstationarity. For a multitude of practical reasons, the adaptive weight vector derived from the necessarily finite training set (secondary data) may need to be applied to data outside of that training region [24]. If a jammer is present, it may be at a slightly different angle if the training data is old relative to the dynamics of the jammer [24]. Although derivative constraints have been proposed to remedy this [24] by producing wider jammer nulls, a much simpler and more straightforward approach that essentially produces the same effect (that is, readily extensible to nonlinear arrays) is via the application of a CMT to the sample covariance matrix (also referred to as covariance augmentation in [28]); that is,

$$\hat{R} \rightarrow \hat{R} \odot T \quad (2.67)$$

where \hat{R} is the sample covariance matrix (which may also include diagonal loading [23]), \odot denotes the Hadamard matrix multiplication operator (simply conformal elementwise multiplication analogous to matrix addition and subtraction [2]), and $T \in \mathbb{C}^{N \times N}$ is a CMT matrix [23]. Mailloux [26] and Zatman [27] independently showed that by setting T equal to a correlation matrix of the form [23]

$$[T]_{i,j} = \text{sinc}(\Delta|i - j|/\pi) \quad (2.68)$$

where $\text{sinc}(x) \stackrel{\Delta}{=} \sin(\pi x)/\pi x$, $[T]_{i,j}$ denotes the (i, j) th element of T , and Δ is a positive scalar used to set the amount of notch widening (dimensions of radians [23]) and can also be interpreted as the amount of uniform random dither [23] or bandwidth [26]. The specific CMT of (2.68) is referred to as a Mailloux-Zatman CMT in [23]. Figure 2.12 shows an example of the

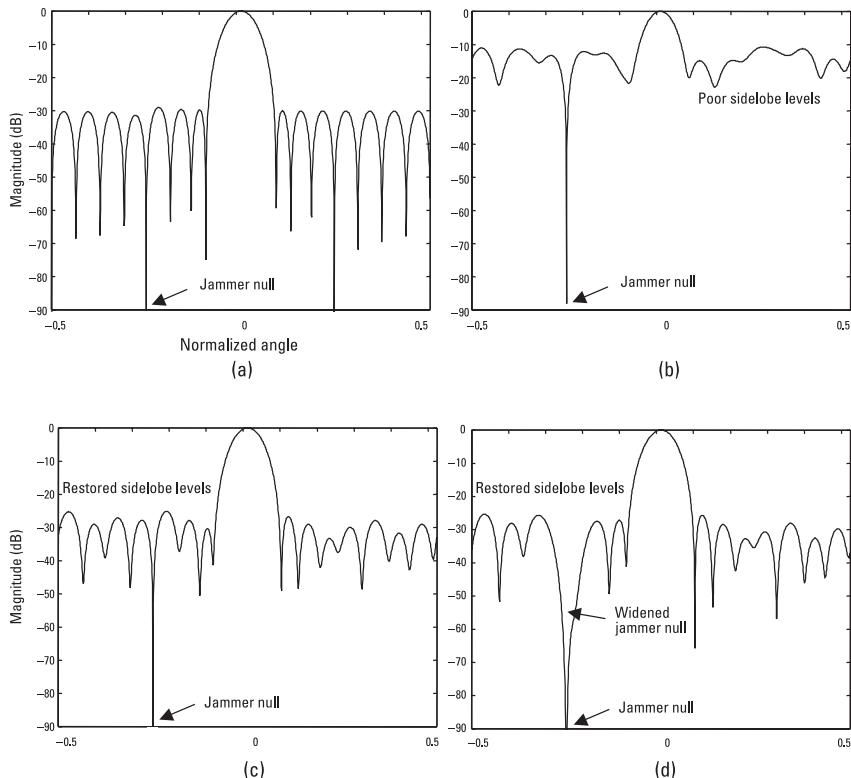


Figure 2.12 Example of the application of a CMT to produce a wider jammer null to help address the stale weights problem. Adapted patterns for a 16-element half-wavelength ULA in the presence of a 60-dB per element JNR at a normalized angle of -0.25 (-30° off boresight): (a) optimum adapted pattern derived from exact (ideal) covariance with a -35 -dB sidelobe level Chebyshev taper, (b) pattern obtained with a $2N$ sample covariance matrix (i.e., SMI method), (c) pattern obtained if 10 dB of diagonal loading is added to the sample covariance matrix, and (d) pattern obtained if a diagonally loaded (10 dB) Mailloux-Zatman CMT is employed with a $\Delta = 0.01$. Note the preservation of good sidelobes associated with diagonal loading as well as a substantially widened (and more robust) jammer null.

application of a Mailloux-Zatman CMT; the jammer notch widening is quite readily apparent, as are the preservation of the main lobe and average sidelobe levels [23].

We will defer a detailed taxonomical discussion of general methods for addressing other forms of nonstationarity that arise in space-time clutter cancellation, including additional techniques for minimizing sample support requirements, to Chapter 5.

2.6 Summary

In this chapter we began our journey into STAP with a sojourn into its origins in 1-D adaptive beamforming and Doppler processing. After deriving the equations for optimal beamforming and illustrating its application to a multijammer scenario, we introduced the concept of adaptive beamforming. An important distinction was drawn between optimal and adaptive beamforming, and some fundamental differences were highlighted, particularly for the finite sample covariance estimation case. For this case, an eigen analysis was conducted to gain key insights into the origins of the degradation from optimality for the finite sample case. In particular, the occurrence of erratic adapted sidelobes and main lobe distortion were shown to be the result of estimation errors associated with the so-called noise-floor eigenvalues/vectors. Both principal components and diagonal loading were introduced to remedy this problem.

Finally, a brief discussion of issues arising when applying adaptive processing in nonstationary environments was introduced, along with an outline of strategies that can be brought to bear, including minimal sample support methods (indirect approach) and nonstationary modeling (direct approach). The concept of CMT covariance augmentation was also introduced to simply address the so-called stale weights problem. A detailed survey of these methods will be presented in Chapter 5.

References

- [1] Mailloux, R. J., *Phased Array Antenna Handbook*, Norwood, MA: Artech House, 1994.
- [2] Horn, R. A., and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, England: Cambridge University Press, 1991.
- [3] Strang, G., *Introduction to Linear Algebra*, 2nd ed., Wellesley, MA: Wellesley-Cambridge Press, 1998.

- [4] Farina, A., *Antenna-Based Signal Processing Techniques for Radar Systems*, Norwood, MA: Artech House, 1992.
- [5] Barton, D. K., *Radar System Analysis*, Dedham, MA: Artech House, 1976.
- [6] Papoulis, A., and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 2001.
- [7] Van Trees, H. L., *Detection, Estimation, and Modulation Theory*, Part I, New York: John Wiley & Sons, 1968.
- [8] Horn, R. A., and C. R. Johnson, *Matrix Analysis*, Cambridge, England: Cambridge University Press, 1985.
- [9] Schleher, D. C., *MTI and Pulsed Doppler Radar*, Norwood, MA: Artech House, 1991.
- [10] Morris, G. V., and L. Harkness, *Airborne Pulsed Doppler Radar*, 2nd ed., Norwood, MA: Artech House, 1996.
- [11] Widrow, B., and M. A. Lehr, "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation," *Proc. of IEEE*, Vol. 78, No. 9, 1990, pp. 1415–1442.
- [12] Melvin, W., (ed.), "Space-Time Adaptive Processing and Adaptive Arrays: Special Collection of Papers," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 508–509.
- [13] Nitzberg, R., *Radar Signal Processing and Adaptive Systems*, 2nd ed., Norwood, MA: Artech House, 1999.
- [14] Cook, C. E., and M. Bernfeld, *Radar Signals: An Introduction to Theory and Application*, Norwood, MA: Artech House, 1993.
- [15] Reed, I. S., J. D. Mallet, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 10, No. 6, November 1974, pp. 853–863.
- [16] Steinhardt, A., "Adaptive Multisensor Detection," in *Adaptive Radar Detection and Estimation*, S. Haykin and A. Steinhardt, (eds.), New York: John Wiley & Sons, 1992.
- [17] Guerci, J. R., J. S. Goldstein, and I. S. Reed, "Optimal and Adaptive Reduced-Rank STAP," *IEEE Trans. on Aerospace and Electronic Systems, Special Section on STAP*, Vol. 36, No. 2, April 2000, pp. 647–663.
- [18] Tufts, D. W., R. Kumaresan, and I. Kirssteins, "Data Adaptive Signal Estimation by Singular Value Decomposition of a Data Matrix," *Proc. of IEEE*, Vol. 70, No. 6, June 1982, pp. 684–685.
- [19] Carlson, B. D., "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 24, No. 3, July 1988, pp. 397–401.
- [20] Guerci, J. R., Y. L. Kim, and S. U. Pillai, "Optimal Loading Factor for Minimal Sample Support Space-Time Adaptive Radar," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 4, Seattle, WA, May 12–15, 1998, pp. 2505–2508.
- [21] Zatman, M., "Circular Array STAP," *IEEE Trans. on Aerospace and Electronic Systems, Special Section on STAP*, Vol. 36, No. 2, April 2000, pp. 510–517.

- [22] Guerci, J. R., and J. S. Bergin, ‘‘Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem,’’ *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002.
- [23] Guerci, J. R., ‘‘Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming,’’ *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–985.
- [24] Gershman, A. B., U. Nickel, and J. F. Bohme, ‘‘Adaptive Beamforming Algorithms with Robustness Against Jammer Motion,’’ *IEEE Trans. on Signal Processing*, Vol. 45, No. 7, July 1997, pp. 1878–1886.
- [25] Zatman, M., and J. R. Guerci, ‘‘Comments on ‘Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming’ [and Reply],’’ *IEEE Trans. on Signal Processing*, Vol. 48, No. 6, June, 2000, pp. 1796–1800.
- [26] Mailloux, R. J., ‘‘Covariance Matrix Augmentation to Produce Adaptive Array Pattern Troughs,’’ *Electronics Letters*, Vol. 31, No. 10, 1995, pp. 771–772.
- [27] Zatman, M., ‘‘Production of Adaptive Array Troughs Through Dispersion Synthesis,’’ *Electronics Letters*, Vol. 31, No. 25, December 1995, pp. 2141.
- [28] Van Trees, H. L., *Optimum Array Processing: Part IV of Detection Estimation and Modulation Theory*, New York: Wiley Interscience, 2002.
- [29] Swokoski, E. W., *Calculus with Analytic Geometry*, Boston, MA: Prindle, Weber, and Schmidt, 1975.
- [30] Pillai, S. U., *Array Signal Processing*, New York: Springer-Verlag, 1989.
- [31] Muirhead, M., *Introduction to Multivariate Analysis*, Englewood Cliffs, NJ: Prentice Hall, 1985.
- [32] Pierre, D. A., *Optimization Theory with Applications*, New York: Dover, 1986.

Appendix 2A: ULA Antenna Pattern Response

In this appendix, we derive the expression of (2.9) for the normalized antenna pattern of a ULA, which is tantamount to evaluating the output response in (2.7); that is,

$$y = \mathbf{w}' \mathbf{x} = \sum_{n=1}^N x_n e^{-j2\pi(n-1)\frac{d}{\lambda} \sin \theta_0} \quad (2A.1)$$

where for convenience, we have set $\kappa = 1$. For a plane wave impinging on the array from an angle θ with respect to boresight, we have

$$x_n = e^{j2\pi(n-1)\frac{d}{\lambda} \sin \theta}, \quad n = 1, \dots, N \quad (2A.2)$$

Substituting (2A.2) into (2A.1) yields

$$\begin{aligned}
y &= \sum_{n=1}^N e^{j2\pi(n-1)\frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \\
&= \sum_{n=0}^{N-1} e^{j2\pi n \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \\
&= \sum_{n=0}^{N-1} r^n
\end{aligned} \tag{2A.3}$$

where

$$r = e^{j2\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \tag{2A.4}$$

From the Geometric Sum Formula [29],

$$\sum_{n=0}^{N-1} r^n = \frac{1 - r^N}{1 - r} = \frac{r^N - 1}{r - 1} \tag{2A.5}$$

we thus have

$$\begin{aligned}
y &= \frac{e^{j2\pi N \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} - 1}{e^{j2\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} - 1} \\
&= \frac{e^{j\pi N \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \left(e^{j\pi N \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} - e^{-j\pi N \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \right)}{e^{j\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \left(e^{j\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} - e^{-j\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \right)} \\
&= \frac{e^{j\pi N \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \sin \left(\pi N \frac{d}{\lambda} (\sin \theta - \sin \theta_0) \right)}{e^{j\pi \frac{d}{\lambda}(\sin \theta - \sin \theta_0)} \sin \left(\pi \frac{d}{\lambda} (\sin \theta - \sin \theta_0) \right)}
\end{aligned} \tag{2A.6}$$

where we have used Euler's Identity, $2j \sin \alpha = e^{j\alpha} - e^{-j\alpha}$, to arrive at the final form. Taking the absolute value of (2A.6) and dividing by N (for normalization) yields (2.9).

Appendix 2B: Derivation of the Maximum Likelihood Sample Covariance Matrix

The derivation in this appendix of the maximum likelihood estimate (MLE) of the sample covariance matrix follows that of Steinhardt [16] and Pillai [30]. The reader is encouraged to consult [16, 30] for further statistical details and properties of the adaptive beamformer including an elegant derivation of the important RMB result of (2.51) in [16].

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L \in \mathbb{C}^N$ denote L observations (array snapshots) that are assumed to be complex zero-mean i.i.d. and Gaussian, each with a marginal pdf given by [28]

$$f_{x_i}(\mathbf{x}) = \frac{1}{\pi^N |R|} e^{-\mathbf{x}' R^{-1} \mathbf{x}}, \quad i = 1, \dots, L \quad (2B.1)$$

where $R \in \mathbb{C}^{N \times N}$ is an unknown-covariance matrix (assumed to be positive-definite) with nonzero determinant $|R|$. The associated likelihood function $f(\mathbf{x}_1, \dots, \mathbf{x}_L | R)$ (joint density of $\mathbf{x}_1, \dots, \mathbf{x}_L$ conditioned on R [7]) is thus given by

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_L | R) &= \prod_{i=1}^L f_{x_i}(\mathbf{x}_i | R) \\ &= \frac{1}{\pi^{NL} |R|^L} e^{-\sum_{i=1}^L \mathbf{x}_i' R^{-1} \mathbf{x}_i} \end{aligned} \quad (2B.2)$$

Letting

$$X \triangleq [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_L] \in \mathbb{C}^{N \times L} \quad (2B.3)$$

(2B.2) can be rewritten as

$$\begin{aligned}
f(X|R) &= \frac{1}{\pi^{NL} |R|^L} e^{-\sum_{i=1}^L \mathbf{x}_i' R^{-1} \mathbf{x}_i} \\
&= \frac{1}{\pi^{NL} |R|^L} e^{-\text{Tr}(X' R^{-1} X)} \\
&= \frac{1}{\pi^{NL} |R|^L} e^{-\text{Tr}(XX' R^{-1})} \\
&= \frac{1}{\pi^{NL} |R|^L} e^{-\text{Tr}(R^{-1} XX')} \\
&= \left[\pi^{-N} |R|^{-1} e^{-\text{Tr}(R^{-1} \hat{R})} \right]^L
\end{aligned} \tag{2B.4}$$

where $\text{Tr}(\cdot)$ denotes the trace operator (sum of the diagonal elements of a square matrix [2]) and

$$\hat{R} \triangleq \frac{1}{L} XX' = \frac{1}{L} \sum_{i=1}^L \mathbf{x}_i \mathbf{x}_i' \tag{2B.5}$$

which is recognized as the $N \times N$ *sample covariance matrix* (SCM). Use of the identity $\text{Tr}(AB) = \text{Tr}(BA)$ was utilized in arriving at the final form in (2B.4).

The MLE of the unknown-covariance matrix R is obtained by maximizing the likelihood function with respect to R or minimizing the negative log-likelihood [7]; specifically,

$$\begin{aligned}
\hat{R}_{\text{ML}} &\triangleq \arg \max_{\{R\}} (f(X|R)) \\
&= \arg \min_{\{R\}} (-\ln f(X|R)) \\
&= \arg \min_{\{R\}} (\ln |R| + \text{Tr}(R^{-1} \hat{R}))
\end{aligned} \tag{2B.6}$$

where use was made of the monotonicity of the natural logarithm, and only the terms involving R were retained. Since it is assumed that $L \geq N$, \hat{R} is

positive-definite (almost surely [31]) and therefore possesses a matrix square root $\hat{R}^{1/2}$ [2] (such that $\hat{R} = \hat{R}^{1/2}\hat{R}^{1/2}$). This allows us to rewrite (2B.6) as

$$\begin{aligned}
\hat{R}_{\text{ML}} &= \arg \min_{\{R\}} (-\ln |R^{-1}| + \text{Tr}(R^{-1}\hat{R})) \\
&= \arg \min_{\{R\}} (-\ln |R^{-1}\hat{R}^{-1}\hat{R}| + \text{Tr}(R^{-1}\hat{R})) \\
&= \arg \min_{\{R\}} (\ln |\hat{R}^{-1}| - \ln |R^{-1}\hat{R}| + \text{Tr}(R^{-1}\hat{R})) \\
&= \arg \min_{\{R\}} (-\ln |R^{-1}\hat{R}^{1/2}\hat{R}^{1/2}| + \text{Tr}(R^{-1}\hat{R}^{1/2}\hat{R}^{1/2})) \\
&= \arg \min_{\{R\}} (-\ln |\hat{R}^{1/2}R^{-1}\hat{R}^{1/2}| + \text{Tr}(\hat{R}^{1/2}R^{-1}\hat{R}^{1/2})) \\
&= \arg \min_{\{R\}} (-\ln |\Psi| + \text{Tr}(\Psi))
\end{aligned} \tag{2B.7}$$

where

$$\Psi \triangleq \hat{R}^{1/2}R^{-1}\hat{R}^{1/2} \tag{2B.8}$$

and use was made of the following identities at arriving at the final form in (2B.7):

$$\det(AB) = \det(BA), \det(A^{-1}) = \frac{1}{\det(A)}, \text{Tr}(AB) = \text{Tr}(BA)
\tag{2B.9}$$

Note also that additive terms not involving R were dropped in (2B.7) since they do not affect the result.

Since there is a one-to-one mapping between R and Ψ [i.e., (2B.8)] and both are positive-definite Hermitian matrices, we can first minimize with respect to Ψ , then solve for R to obtain the final result. This process is further facilitated by expressing Ψ in diagonal form; that is,

$$\Psi = U\Lambda U' \tag{2B.10}$$

where $U \in \mathbb{C}^{N \times N}$ is a unitary matrix [2] (i.e., $UU' = I$) and $\Lambda \in \mathbb{R}_+^{N \times N}$ is a diagonal matrix of strictly positive eigenvalues of Ψ ; that is, $\Lambda =$

$\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$. Substituting (2B.10) into (2B.7) yields the equivalent minimization problem

$$\begin{aligned}
 \Psi_{\text{ML}} &= \arg \min_{\{\Psi\}} (-\ln(|\Psi|) + \text{Tr}(\Psi)) \\
 &= \arg \min_{\{U\Lambda U'\}} (-\ln(|U\Lambda U'|) + \text{Tr}(U\Lambda U')) \\
 &= \arg \min_{\{\Lambda\}} (-\ln(|\Lambda|) + \text{Tr}(\Lambda)) \\
 &= \arg \min_{\{\lambda_1, \lambda_2, \dots, \lambda_N\}} \left(-\sum_{n=1}^N \ln(\lambda_n) + \sum_{n=1}^N \lambda_n \right)
 \end{aligned} \tag{2B.11}$$

Due to the convexity of the original likelihood function [32], a necessary and sufficient condition for the minimization of (2B.11) is that each of the derivatives of the argument with respect to each eigenvalue vanish; that is,

$$\frac{d \left(-\sum_{n=1}^N \ln(\lambda_n) + \sum_{n=1}^N \lambda_n \right)}{d \lambda_n} = 0 \tag{2B.12}$$

$$\forall n : n = 1, \dots, N$$

which yields

$$-\frac{1}{\lambda_n} + 1 = 0 \rightarrow \lambda_n = 1 \tag{2B.13}$$

$$\forall n : n = 1, \dots, N$$

which in turn implies that $\Lambda_{\text{ML}} = I \rightarrow U_{\text{ML}} = I \rightarrow \Psi_{\text{ML}} = I$, and thus

$$I = \hat{R}^{1/2} R_{\text{ML}}^{-1} \hat{R}^{1/2} \tag{2B.14}$$

or

$$R_{\text{ML}} = \hat{R} = \frac{1}{L} XX' = \frac{1}{L} \sum_{i=1}^L \mathbf{x}_i \mathbf{x}'_i \tag{2B.15}$$

which is the desired result.

3

Space-Time Adaptive Processing

3.1 Introduction

The term *STAP* was first applied to multidimensional adaptive filtering of clutter and jamming in airborne MTI radars [1, 2]. Unlike ground-based (or ground-stationary) MTI radars, clutter returns manifest themselves as fully 2-D (nonfactorable) structures due to the motion-induced Doppler-spreading effect described in Section 3.2. As a consequence, the traditional factored, or decoupled, approaches of beamforming followed by Doppler filtering (or vice versa) are not optimal [1]. Instead, as we will see, a better approach is to perform multidimensional filtering that accounts for angle-Doppler coupling.

As we will shortly see, space-time clutter is generally colored noise, that is, a nondiagonal covariance matrix. This is a good news/bad news situation: On the one hand, the fact that it has structure suggests that there may be an opportunity for separating the clutter subspace from the signal subspace via space-time filtering—as described later. On the other hand, this can only be accomplished if we have an accurate model for the clutter structure—therein lies the rub.

The requisite accuracy of this model is, of course, highly dependent on the particular application. For example, in a high PRF (Doppler unambiguous [3]) *airborne MTI* (AMTI) radar with good two-way antenna sidelobes, the targets of interest will be well removed (in Doppler) from the strongest main lobe clutter interference [3]. This situation only requires a sidelobe notch that is well removed in angle-Doppler from the target main lobe.

Since widening and deepening this notch width (e.g., with CMTs [4] or *power selected training* (PST) [5]) will have virtually no effect on the target main lobe response, the clutter model need not be very complex.

On the other hand—in stark contrast to the AMTI case—*ground MTI* (GMTI) radars can often encounter slow targets that are very close (in angle-Doppler) to mainbeam clutter. In this case, a much higher-fidelity clutter model and multidimensional filtering scheme are required [6]. Since ground clutter can be exquisitely complex—comprised of all sorts of terrain, surface reflectivities, and internal motion—this modeling problem can be extremely challenging [6]. To make matters worse, this model must be implementable in real time. Indeed, much of the current STAP research is aimed at ever more refined and robust modeling (explicit or implicit) of ground clutter for just such applications [6–11].

In this chapter, we introduce a basic model of clutter combining both basic Doppler physics and stochastic modeling. This first-order 2-D clutter model will be useful for establishing basic performance characteristics and properties of both optimum and adaptive STAP filtering algorithms. While adequate for some nonstressing applications, many necessary real-world refinements to this model will be described in Chapter 4. Only then will a comprehensive survey of many current STAP techniques be introduced and applied to higher-fidelity simulations.

3.2 Need for Joint Space and Time Processing

When a radar emits a pulse, it propagates outward at the speed of light with a directional intensity distribution governed by the transmit antenna pattern [12]. In addition to potentially illuminating a target of interest, everything else in the radar’s field of regard will likewise be radiated. Of particular concern are the reflections from ground clutter whose aggregate return can be many orders of magnitude greater than a potential target of interest.

In addition to angle and range separation, for MTI radars, the Doppler shift induced by a moving target is pivotal in separating it from ground clutter (which is assumed to be stationary [3]). For a nonmoving radar (e.g., a fixed ground-based or statically deployed airborne system such as an aerostat [13]), ground-clutter returns should not have appreciable Doppler shifts and can thus be eliminated by notching out the zero-Doppler filter [3]. Unfortunately, this straightforward 1-D processing strategy is insufficient when the radar is moving—such as in airborne or even spaceborne MTI systems [14].

To see why this is the case, consider Figure 3.1, which depicts the N -element ULA radar of Chapter 2 in uniform motion. Neglecting for the moment any small elevation angle and curved Earth effects [1, 2], we see that ground clutter contained in an iso-range ring centered at the radar will have a Doppler-shift distribution due to the motion of the radar. The width of the ring, which is governed by the bandwidth of the radar pulse, is approximately $\Delta R \sim c/(2B)$ [2], where c is the speed of light, and B is the pulse bandwidth (consistent units assumed). For the side-looking case illustrated in Figure 3.1, we see that the normalized Doppler shift induced on a differential clutter patch located at an angle θ relative to the array boresight is given by [1]

$$\bar{f}_d = \frac{2vT}{\lambda} \sin \theta \quad (3.1)$$

where v is the speed of the radar in units consistent with λ and T . Note that (3.1) is consistent with (2.23) if we recognize that for the geometry of

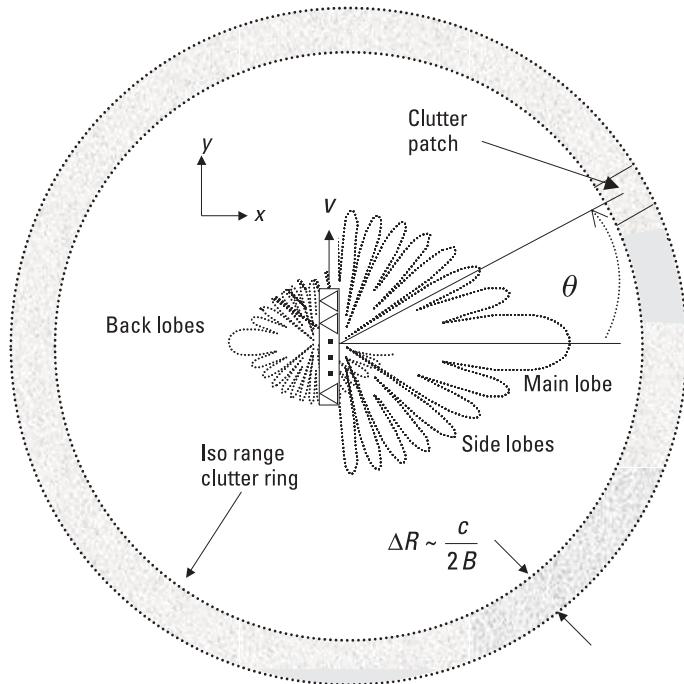


Figure 3.1 Illustration of a clutter iso-range ring for a ULA radar in uniform constant-velocity motion relative to the ground.

Figure 3.1, $\mathbf{v}_{tgt} = 0$ (i.e., the clutter patch is stationary), $\mathbf{v}_{Rx} = v\hat{\mathbf{i}}_y$, where $\hat{\mathbf{i}}_y$ is a unit vector in the y direction, and $\hat{\mathbf{i}}_y \cdot \hat{\mathbf{i}}_{Rx} = -\sin \theta$.

In addition to a spatially dependent Doppler shift, each clutter patch will have a spatially dependent reradiating intensity that depends on the transmit antenna pattern and the intrinsic reflectivity of the clutter [15]. In general, this pattern consists of both front lobe (main lobe and sidelobes) and back lobe radiation (see Figure 3.1). Most airborne MTI radars are designed to insure that the back lobe radiation is significantly attenuated and can often be ignored [16]. However, since this may not always be the case (especially at low frequencies), we will explicitly account for it in Chapter 4.

Figure 3.2 illustrates the impact of uniform radar motion. In the absence of platform motion ($v = 0$), the stationary clutter is concentrated along the zero-Doppler contour for all $\bar{\theta}$. However, when $v \neq 0$, there is a linear relationship between Doppler and $\sin \theta$ (or $\bar{\theta}$) governed by (3.1). As a result, the clutter energy is distributed along a line, or “clutter ridge,” as shown in Figure 3.2. Note that for the case illustrated, the antenna is aligned with direction of motion. The effect of antenna crabbing (array misalignment with the ground-referenced velocity vector) will be considered in Chapter 4.

The slope β of the clutter ridge is easily obtained by introducing the normalized angle, $\bar{\theta} = \frac{d}{\lambda} \sin \theta$, in (3.1), which yields

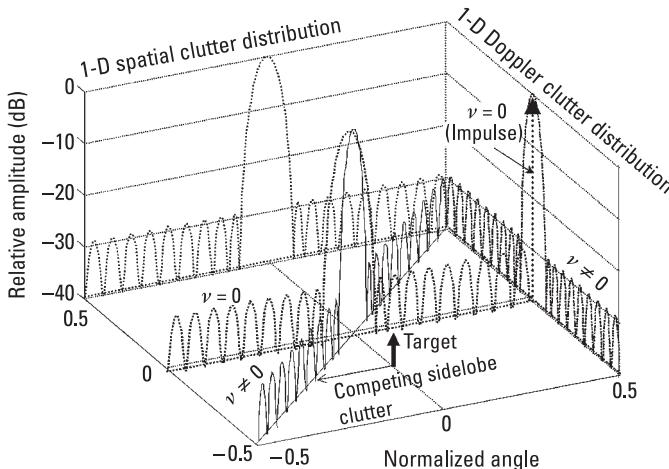


Figure 3.2 Angle-Doppler (space-time) structure of airborne clutter due to ownship platform motion.

$$\begin{aligned}
\bar{f}_d &= \frac{2vT}{\lambda} \left(\frac{\lambda}{d} \right) \left(\frac{d}{\lambda} \right) \sin \theta \\
&= \frac{2vT}{d} \left(\frac{d}{\lambda} \sin \theta \right) \\
&= \beta \bar{\theta}
\end{aligned} \tag{3.2}$$

from which we see that [1]

$$\beta = \frac{2vT}{d} = \frac{4vT}{\lambda} \Big|_{d=\frac{\lambda}{2}} \tag{3.3}$$

Thus, the case illustrated in Figure 3.2 corresponds to $\beta = 1$. Figure 3.2 also clearly illustrates that a simple 1-D filtering scheme is not optimum for separating target from clutter since competing clutter can potentially come from several possible directions (depending on β) [1], hence the need for joint space and time (i.e., 2-D) processing.

We are now in a position to quantify the total clutter return from a given iso-range ring. Assuming that the i th clutter patch at a normalized angle, $\bar{\theta}_i$, in the iso-range ring is far enough away so that elevation angle is negligible [1], its spatial N -dimensional steering vector \mathbf{a}_i is given by

$$\mathbf{a}_i = \left[1 \ e^{j2\pi\bar{\theta}_i} \ e^{j2\pi(2)\bar{\theta}_i} \ \dots \ e^{j2\pi(N-1)\bar{\theta}_i} \right]^T \tag{3.4}$$

Since this i th clutter patch also has a normalized Doppler shift \bar{f}_{d_i} , each vector of array outputs (i.e., a “snapshot”) from successive pulses due to the i th clutter patch will have a temporal linear phase progression; that is, at the m th PRI, the clutter patch snapshot is given by

$$e^{j2\pi(m-1)\bar{f}_{d_i}} \mathbf{a}_i \tag{3.5}$$

If M pulses are to be processed in a CPI [1], the total $N \times M$ matrix of space-time steering vectors for the i th clutter patch is of the form

$$\left[\mathbf{a}_i \ e^{j2\pi\bar{f}_{d_i}} \mathbf{a}_i \ \dots \ e^{j2\pi(M-1)\bar{f}_{d_i}} \mathbf{a}_i \right] \tag{3.6}$$

This corresponds to an observation vector consisting of a total of NM measurements that can conveniently be represented by the Kronecker product [17] of the Doppler and spatial steering vectors, namely,

$$\mathbf{v}_i = \mathbf{b}_i \otimes \mathbf{a}_i \quad (3.7)$$

where \mathbf{v}_i is the NM -dimensional space-time steering vector for the i th clutter patch, and \mathbf{b}_i is its corresponding M -dimensional Doppler steering vector given by

$$\mathbf{b}_i = [1 \ e^{j2\pi\bar{f}_{d_i}} \ \dots \ e^{j2\pi(M-1)\bar{f}_{d_i}}]^T \quad (3.8)$$

\otimes denotes the Kronecker (or tensor) product, which is defined as follows [17]:

Let A and B denote matrices of dimensions $m \times n$ and $p \times q$, respectively. Then $A \otimes B$ is an $(mp) \times (nq)$ -dimensional matrix of the form

$$\begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & & a_{2n}B \\ \vdots & & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \quad (3.9)$$

where a_{ij} denotes the (i, j) th element of A . Note that (3.7) is obtained from the concatenation of the columns of (3.6).

Figure 3.3 displays a sample normalized angle-Doppler signal of the form (3.7); as was the case with the 1-D angle response, main lobe and sidelobe features are evident. The pattern in Figure 3.3 can be obtained via a 2-D FFT since both the spatial and temporal sampling are uniform [1].

The total space-time clutter return from a given iso-range is thus an NM -dimensional random vector $\mathbf{X}_c \in \mathbb{C}^{NM}$ of the form

$$\mathbf{X}_c = \sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i \quad (3.10)$$

where $\tilde{\gamma}_i$ is a complex scalar RV that accounts for the amplitude and phase of the i th clutter patch, and N_c is the total number of clutter patches in the iso-range ring. Note that (3.10) is essentially a Riemann sum approximation to the actual continuous clutter integral [1, 15]. However, due to the finite spatial and temporal resolution (or bandwidth) of the radar, this approximation is accurate provided that N_c and the corresponding clutter patches are chosen properly. For a radar-centric treatise of the validity of integral approximations like (3.8), the reader is referred to Brennan and Mallett [18].

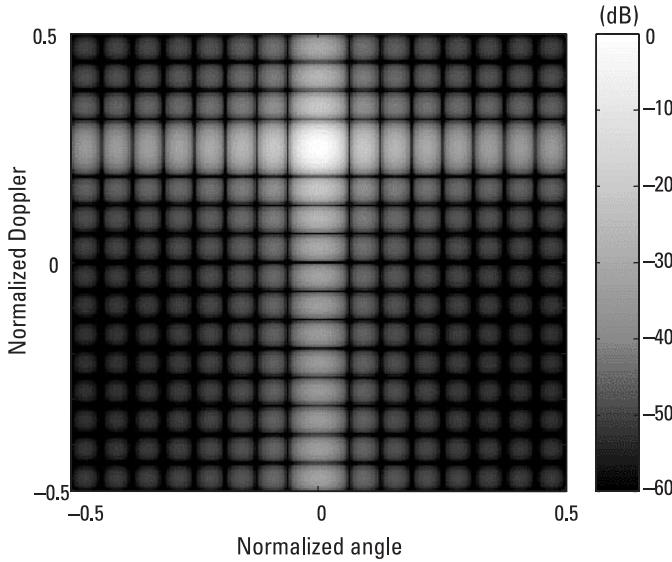


Figure 3.3 Example of a normalized 2-D angle-Doppler steering-vector pattern achieved via a 2-D DFT (side-looking case with a normalized Doppler of +0.25). Note the presence of both angle and Doppler sidelobes.

In general, $\tilde{\gamma}_i$ will depend not only on the transmit antenna pattern, but also on the intrinsic nature of the clutter and can be quite complex in its scattering mechanisms [15, 19–21]. Several of these complicating factors will be considered in later chapters (e.g., inhomogeneities, internal clutter motion). At present, we will assume that the underlying clutter is homogenous and can thus be modeled as a *wide sense stationary* (wss) process [22] with uncorrelated complex reflectivity factors; that is,

$$\mathbb{E}\{\tilde{\gamma}_i \tilde{\gamma}_j^*\} = 0, \forall i, j: i \neq j \quad (3.11)$$

Moreover, the average (or expected) strength of the i th scatterer is assumed to be proportional to the transmit antenna gain [1]; that is,

$$\mathbb{E}\{|\tilde{\gamma}_i|^2\} = G_i, \text{ for } i = 1, \dots, N_c \quad (3.12)$$

where G_i is a strictly positive real-valued constant that is proportional to the transmit antenna gain in the direction θ_i . For this case, the resulting NM -dimensional space-time clutter covariance matrix $R_c \in \mathbb{C}^{NM \times NM}$ is given by

$$\begin{aligned}
R_c &= \mathbb{E}\{\mathbf{X}_c \mathbf{X}_c'\} \\
&= \mathbb{E}\left\{\left(\sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i\right) \left(\sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i\right)'\right\} \\
&= \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} \mathbb{E}\{\tilde{\gamma}_i \tilde{\gamma}_j^*\} \mathbf{v}_i \mathbf{v}_j' \\
&= \sum_{i=1}^{N_c} G_i \mathbf{v}_i \mathbf{v}_i'
\end{aligned} \tag{3.13}$$

The total space-time covariance matrix $R \in \mathbb{C}^{NM \times NM}$, due to both clutter and receiver noise, is of the form

$$R = R_c + \sigma^2 I \tag{3.14}$$

where it is assumed that clutter is uncorrelated with receiver noise. The *clutter-to-noise ratio* (CNR) [1] is defined as the ratio of the clutter power to the receiver noise power on a single element and a single pulse; that is,

$$\text{CNR} = \frac{R_c(1, 1)}{\sigma^2} \tag{3.15}$$

where the first diagonal element of R_c was arbitrarily chosen since each channel is assumed to have the same receiver noise level and elemental antenna pattern. Throughout this text, we will set the noise floor to 0 dB ($\sigma^2 = 1$) for notational and graphical convenience and, thus, adjust the clutter (and possibly jammer) level to achieve a specified CNR (and/or JNR).

We are now in a position to examine the space-time, or equivalently, the angle-Doppler structure of R . One method for characterizing the clutter is by examining the 2-D (angle-Doppler) power spectrum corresponding to R [2]. The so-called Fourier-based *power spectral density* (psd) is the 2-D positive function defined as [2]

$$P_F(\bar{f}_d, \bar{\theta}) = \mathbf{s}'(\bar{f}_d, \bar{\theta}) R \mathbf{s}(\bar{f}_d, \bar{\theta}) \tag{3.16}$$

where $\mathbf{s}(\bar{f}_d, \bar{\theta})$ is the NM -dimensional target steering vector specifying the normalized Doppler and angle of interest, and R is the total interference covariance matrix. Equation (3.16) can be interpreted as the expected value

of the 2-D discrete Fourier transform of the total clutter plus noise signal out of the space-time beamformer [2, 23]; that is,

$$\begin{aligned} P_F(\bar{f}_d, \bar{\theta}) &= E(|\mathbf{s}'(\bar{f}_d, \bar{\theta})\mathbf{x}|^2) \\ &= \mathbf{s}'R\mathbf{s} \end{aligned} \quad (3.17)$$

where $\mathbf{x} \in \mathbb{C}^{NM}$ is the received NM -dimensional total clutter-plus-noise vector RV. It is assumed in (3.16) that the receiver noise floor is set to 0 dB (i.e., $\sigma^2 = 1$) and that the steering vectors are normalized to unity (i.e., $\mathbf{s}'\mathbf{s} = 1$), which insures that $P_F(\bar{f}_d, \bar{\theta}) = 1$ if only receiver noise is present. Later in this chapter, a much higher resolution estimate will be obtained that is far more representative of the information exploited by optimum space-time processors.

Figure 3.4 shows the psds for four different Doppler spreads ($\beta = 0, 0.5, 1$, and 2) for a 16-element, 16-pulse example (i.e., $N = M = 16$) with a CNR set to 40 dB. Note the presence of Doppler aliasing when $\beta > 1$ [3]. In this case, for a given target Doppler of interest, competing sidelobe clutter can exist in multiple directions. In certain *low Earth orbit* (LEO) *space-based radar* (SBR) scenarios, Doppler aliasing can occur in the mainbeam due to the extremely high ground speeds (~ 7 km/sec) [16, 24], unless a very large antenna is employed.

It is apparent from Figures 3.2 and 3.4 that there can be a substantial reduction in the percentage of the unambiguous angle-Doppler observation space occupied by clutter when 2-D processing is employed. This colored-noise subspace compression is best understood and quantified by conducting an eigen analysis of the total interference covariance matrix to determine the extent of the clutter subspace.

Figure 3.5 shows the eigenvalue distributions corresponding to the four different β values considered in Figure 3.4. Note the increase in the effective rank of the clutter portion of the total interference eigenspectrum with increasing β , where the effective rank is defined as

$$\text{rank}_{\text{eff}}\{R_c\} = \sum_{i=1}^{NM} U^+(\lambda_{c_i} - \sigma^2) \leq NM \quad (3.18)$$

where

$$U^+(z) = \begin{cases} 1, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

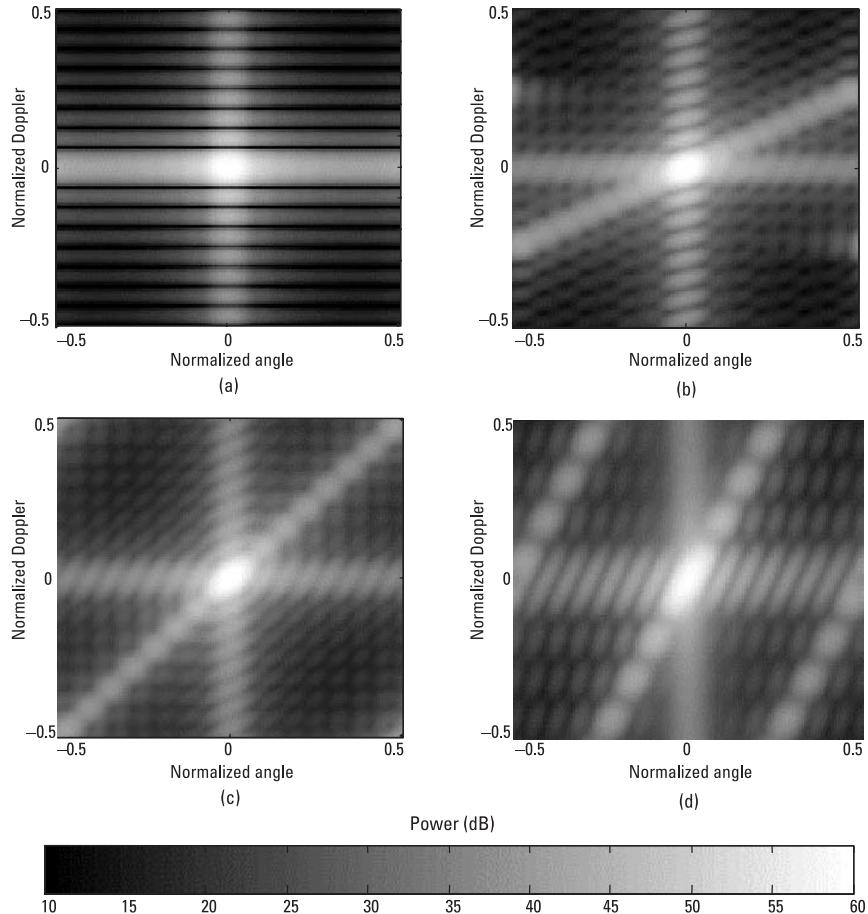


Figure 3.4 Power spectral densities (Fourier-based) for four different values of β : (a) $\beta = 0$, (b) $\beta = 0.5$, (c) $\beta = 1$, and (d) $\beta = 2$.

and R_c is the clutter-only covariance matrix. Equation (3.18) defines the effective rank as the number of eigenvalues that exceed the receiver noise floor [25]. This is often a more meaningful definition of rank for radar engineers since eigenvalues below the noise floor can be ignored (recall from Chapter 2 that the eigencanceller form of the optimum weight vector only contains contributions from eigenvectors whose corresponding eigenvalues are greater than $\lambda_{\min} = \sigma^2$ [25]).

A useful expression for estimating the rank (under ideal conditions) of the clutter-only covariance matrix, R_c , has been obtained by Brennan for the case of a side-looking ULA with fixed PRF, constant velocity, and no crab [1, 26]. Specifically

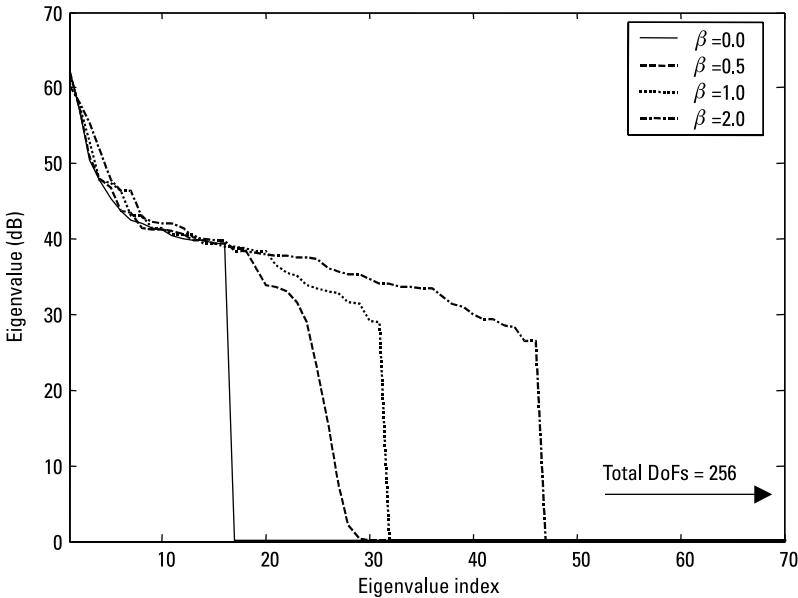


Figure 3.5 Eigenvalue distributions corresponding to the four different β values of Figure 3.4.

$$\text{rank} \{R_c\} \approx N + (M - 1)\beta \quad (3.19)$$

Equation (3.19) is an exact equality for integer values of β . Since the effective rank depends on CNR according to (3.18), in general $\text{rank}_{\text{eff}}\{R_c\} \leq \text{rank}\{R_c\}$, especially for weak clutter cases.

It is worth examining the origins of (3.19) as it sheds light on the issues associated with space-time clutter cancellation, as well as on an historically important space-time clutter-rejection scheme known as *displaced phase center array* (DPCA) processing [27]. To accomplish this, we recall from (3.13) that R_c is of the form

$$R_c = \sum_{i=1}^{N_c} G_i \mathbf{v}_i \mathbf{v}_i' = VGV' \quad (3.20)$$

where

$$V \triangleq [\mathbf{v}_1 \ \dots \ \mathbf{v}_{N_c}] \quad (3.21)$$

that is, an $NM \times N_c$ matrix whose columns are the N_c clutter steering vectors, and

$$G = \text{diag} \{ G_1 \dots G_{N_c} \} \quad (3.22)$$

that is, an $N_c \times N_c$ diagonal matrix of the relative clutter patch strengths (including transmit antenna pattern effects). Since G_i represents the average power from each clutter patch, we may assume that $G_i > 0, \forall i$ [if a particular $G_i = 0$, simply remove its corresponding steering vector from the summation in (3.20)]. Thus, the rank $\{R_c\}$ is given by

$$\begin{aligned} \text{rank } \{R_c\} &= \min (\text{rank } \{\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_{N_c}\}), NM \\ &= \min (\text{rank } \{V\}, NM) \end{aligned} \quad (3.23)$$

where it is assumed that $N_c \geq NM$ and $\mathbf{v}_i \neq \xi \mathbf{v}_j$ for $i \neq j$ (where ξ is an arbitrary complex scalar). In other words, the rank of R_c is determined by the number of linearly independent space-time clutter steering vectors, \mathbf{v}_i , $i = 1, \dots, N_c$: $N_c \geq NM$, but may be less than NM (indeed, this is desirable).

For the conditions stated, and assuming that β is an integer, the i th clutter steering vector has a Vandermonde structure [1] of the form

$$\begin{aligned} \mathbf{v}_i = & \left[1 \ w_i \ \dots \ w_i^{N-1} \mid w_i^\beta \ w_i^{\beta+1} \ \dots \ w_i^{\beta+N-1} \mid \dots \right. \\ & \left. \dots \mid w_i^{(M-1)\beta} \ w_i^{(M-1)\beta+1} \ \dots \ w_i^{(M-1)\beta+N-1} \right]^T \end{aligned} \quad (3.24)$$

where $w_i = e^{j2\pi\bar{\theta}_i}$. Inspection of (3.24) (with β an integer) reveals that there are only $N + (M - 1)\beta$ unique elements; the rest are redundant. For example, if $\beta = 1$, (3.24) becomes

$$\begin{aligned} \mathbf{v}_i = & \left[1 \ w_i \ \dots \ w_i^{N-1} \mid w_i^1 \ w_i^2 \ \dots \ w_i^N \mid \dots \right. \\ & \left. \dots \mid w_i^{(M-1)} \ w_i^{(M-1)+1} \ \dots \ w_i^{(M-1)+N-1} \right]^T \end{aligned} \quad (3.25)$$

Starting with the first N entries from the first PRI (first block), we see that only one element from the second PRI (second block) is distinct from the first N , namely w_i^N . This trend continues for the third and subsequent pulses, with only one distinct new element being added for each new PRI. Thus, after M pulses, the total number of distinct elements is $N + (M - 1)$, which agrees with (3.19) exactly for $\beta = 1$.

Since the rank of a matrix is unaffected by an interchange of rows, the above observations imply that, in general, V can be rearranged as follows:

$$\left[\begin{array}{ccc} 1 & \dots & 1 \\ w_1 & & w_{N_c} \\ w_1^2 & & w_{N_c}^2 \\ \vdots & & \vdots \\ w_1^{(N-1)+\beta(M-1)} & \dots & w_{N_c}^{(N-1)+\beta(M-1)} \\ \hline NM - [N + \beta(M - 1)] \text{ Repeated Rows} \end{array} \right] = \left[\begin{array}{c} \Xi \\ \text{Repeated Rows} \end{array} \right] \quad (3.26)$$

where Ξ is an $[N + \beta(M - 1)] \times N_c$, matrix. Since the rank of a matrix is equal to the number of linearly independent rows or columns (which ever is smaller) [28], we see that

$$\text{rank } \{V\} = \text{rank } \{\Xi\} = \min \{N + \beta(M - 1), NM\} \quad (3.27)$$

since $w_i \neq w_j$, for $i \neq j$, which is the desired result.

It is particularly illuminating to examine the $\beta = 1$ case, which corresponds to the situation when the antenna moves precisely one interelement spacing per PRI. If only the first $N - 1$ elements are used for reception on the first pulse, and only the last $N - 1$ elements are likewise used for the next (second) pulse, then the ground appears stationary to the antenna (assuming perfect channel match) [27]. This is the so-called DPCA case [27]. Clutter cancellation can then be “simply” accomplished by coherently subtracting the returns from successive pulses on adjacent antennas (i.e., the two-pulse MTI canceller [2]). Apart from the challenges of matching antenna channels and eliminating crab angle of the array relative to the true ground-track flight path, DPCA requires that the PRI (or PRF) be chosen in (3.3) to satisfy $\beta = 1$ —an often demanding restriction in practice. DPCA will be studied in more detail in Chapter 5.

3.2.1 Joint Clutter and Jamming Characteristics

If clutter and jamming are both present, the total space-time interference covariance matrix has the form

$$R = R_c + R_J + \sigma^2 I \quad (3.28)$$

where it is assumed that the jamming is uncorrelated with clutter and receiver noise [22]. The single pulse, spatial-only jammer covariance matrix was

derived in Chapter 2. Assuming that the jammer signal is uncorrelated from pulse-to-pulse, the off-diagonal block terms (corresponding to temporal cross correlation) are null matrices. Thus, the NM -dimensional jammer-only covariance, R_J , has the following block-diagonal form:

$$R_J = \begin{bmatrix} \sum_{j=1}^{N_J} \sigma_j^2 \mathbf{s}_j \mathbf{s}'_j & & \emptyset \\ & \sum_{j=1}^{N_J} \sigma_j^2 \mathbf{s}_j \mathbf{s}'_j & \\ & & \ddots \\ \emptyset & & \sum_{j=1}^{N_J} \sigma_j^2 \mathbf{s}_j \mathbf{s}'_j \end{bmatrix} \quad (3.29)$$

where \emptyset is used to indicate that the off-block-diagonal entries are zero. Since the rank of a block-diagonal matrix is equal to the sum of the ranks of the constituent blocks [28], we see that the rank of R_J is given by

$$\begin{aligned} \text{rank } \{R_J\} &= \sum_{m=1}^M \text{rank} \left\{ \sum_{j=1}^{N_J} \sigma_j^2 \mathbf{s}_j \mathbf{s}'_j \right\} \\ &= M \text{rank} \left\{ \sum_{j=1}^{N_J} \sigma_j^2 \mathbf{s}_j \mathbf{s}'_j \right\} \end{aligned} \quad (3.30)$$

For example, if there are three uncorrelated noise jammers, and $M = 16$, the rank of R_J is $3 \times 16 = 48$. If $N = 16$, the fraction of the total available observation space occupied by the jammers is $48 \div 256 = 3/16$, which is precisely the same ratio if spatial-only processing is performed ($N_J = 3, N = 16$). In other words, there is no advantage to using space-time processing against conventional noise jammers—spatial-only processing will suffice. The advantage of space-time processing is realized against clutter (although simultaneous jammer nulling is readily accommodated—as will be shown later).

Figure 3.6 displays a space-time eigenspectrum with and without jamming. Note that the presence of jamming generally increases the effective rank of the combined clutter and jamming subspace.

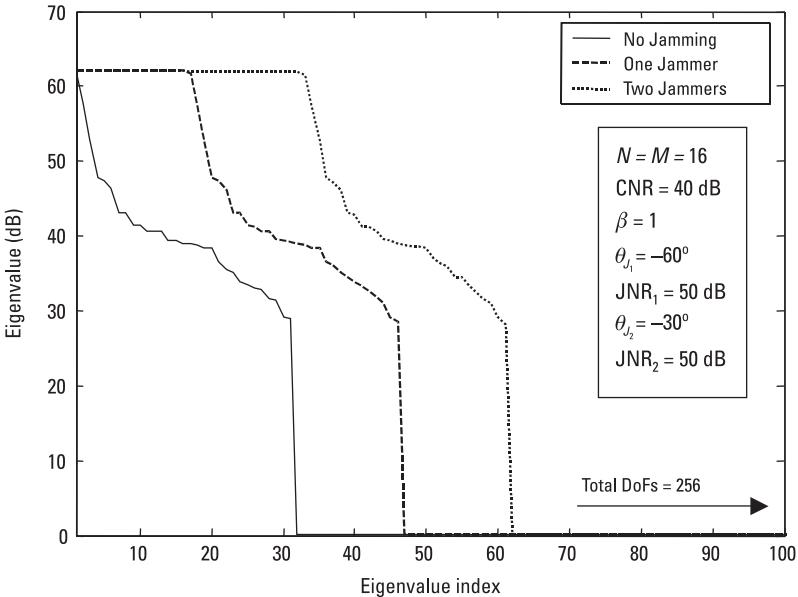


Figure 3.6 Impact of jamming on the total interference eigenspectrum.

3.3 Optimum Space-Time Processing for MTI Radar

We next turn our attention to deriving the optimal space-time beamformer for rejecting both clutter and jamming. Consider the space-time linear beamformer of Figure 3.7, consisting of N identical antenna elements (spatial DoFs) and M PRI time taps (temporal DoFs). Analogous to the 1-D beamformer of Figure 2.1, the output y is in general the linear superposition of the desired signal response y_s and an undesired noise/interference response y_n . Our objective will thus be to choose an optimal set of complex space-time weights, \mathbf{w} , so as to maximize SINR.

By adopting the vector RV notation of Chapter 2, we have an optimization problem identical to that already solved in Chapter 2. Specifically, the weight vector that maximizes SINR is given by

$$\mathbf{w} = \kappa R^{-1} \mathbf{s} \quad (3.31)$$

where $R \in \mathbb{C}^{NM \times NM}$ is the $NM \times NM$ total interference (clutter and/or jamming) plus receiver noise covariance matrix, κ is a scalar that does not affect the SINR, and $\mathbf{s} \in \mathbb{C}^{NM}$ is the NM -dimensional space-time (angle-Doppler) steering vector of the desired signal which is given by

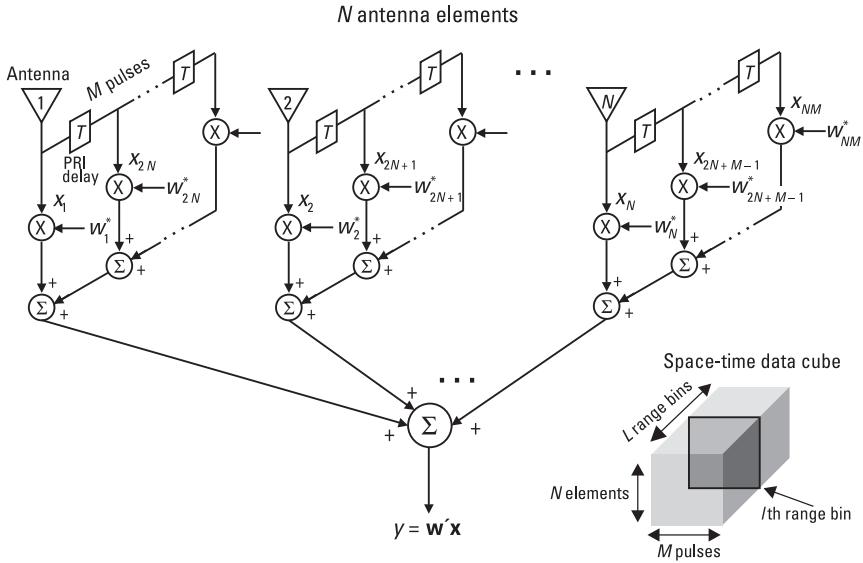


Figure 3.7 Space-time (angle-Doppler) beamformer and “data cube” illustration.

$$\mathbf{s} = \mathbf{b} \otimes \mathbf{a} \quad (3.32)$$

where \mathbf{b} denotes the desired Doppler steering vector and \mathbf{a} the spatial component; that is,

$$\mathbf{b} = [1 \ e^{j2\pi\bar{f}_d} \ \dots \ e^{j2\pi(M-1)\bar{f}_d}]^T \quad (3.33)$$

and

$$\mathbf{a} = [1 \ e^{j2\pi\bar{\theta}_0} \ \dots \ e^{j2\pi(N-1)\bar{\theta}_0}]^T \quad (3.34)$$

where $\bar{\theta}_0$ and \bar{f}_d denote the desired angle and Doppler, respectively, and the superscript T denotes transposition without conjugation. The corresponding optimal SINR is given by (see Chapter 2),

$$\text{SINR}_{\text{opt}} = \mathbf{s}' R^{-1} \mathbf{s} \quad (3.35)$$

As with the 1-D optimal beamformer, tapering can be incorporated into (3.31) to reduce both angle and Doppler sidelobes at the expense of

mainbeam broadening and additional SINR loss (see Section 2.2). For 2-D beamforming, tapering is incorporated as follows:

$$\mathbf{w} = \kappa R^{-1}(\mathbf{s} \odot \mathbf{t}) \quad (3.36)$$

where

$$\mathbf{t} = \mathbf{t}_d \otimes \mathbf{t}_\theta \quad (3.37)$$

$\mathbf{t}_d \in \mathbb{R}_+^M$ is the M -dimensional Doppler taper, and $\mathbf{t}_\theta \in \mathbb{R}_+^N$ is the N -dimensional angle taper (thus $\mathbf{t} \in \mathbb{R}_+^{NM}$ is NM -dimensional as expected).

Figure 3.8 provides an example optimum space-time beam pattern (with and without tapering) for an $N = M = 16$ example, with look angle and Doppler given by $\theta_0 = 0$ (side-looking case), $f_d = +0.25$. The β is set to 1, the CNR is 40 dB, and all JNRs are 50 dB (at the element level). Note the simultaneous presence of both clutter and jamming nulls. Since for a given radar pointing angle there may be a range of Doppler frequencies of interest, it is often convenient to plot SINR versus Doppler for a fixed angle. Figure 3.9 shows an example of this for the side-looking case ($\theta_0 = 0^\circ$); the SINR dropout due to main lobe clutter is clearly evident.

It is evident from an examination of Figure 3.8 that the optimum space-time beamformer is utilizing very high-resolution information regarding the interference, which is manifested as sharp nulls in the angle-Doppler patterns.

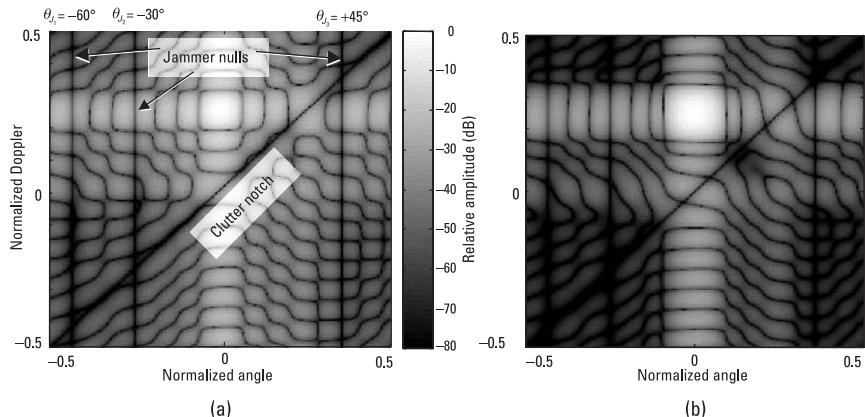


Figure 3.8 Optimum space-time (angle-Doppler) beam pattern illustrating the simultaneous nulling of both clutter and jamming: (a) without tapering, and (b) with 30-dB Chebyshev angle and Doppler tapers.

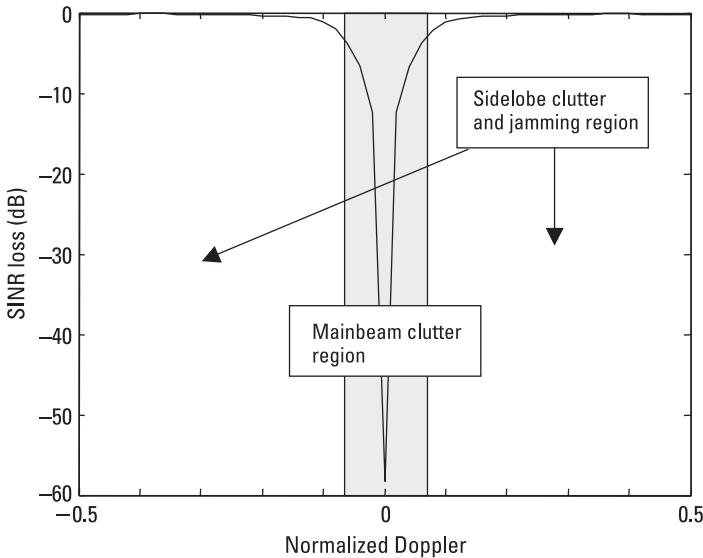


Figure 3.9 SINR versus Doppler for $\theta_0 = 0^\circ$ (side-looking case) of Figure 3.7. Note the SINR dropout due to mainbeam clutter.

Indeed, a very high-resolution (i.e., a so-called super-resolution [29]) spectral estimator underlies the beamformer and is given by

$$P_{MF}(\bar{f}_d, \bar{\theta}_0) = \frac{1}{\mathbf{s}'(\bar{f}_d, \bar{\theta}_0) R^{-1} \mathbf{s}(\bar{f}_d, \bar{\theta}_0)} \quad (3.38)$$

which is simply the reciprocal of the optimal SINR given by (3.35). Note that consistent with (3.16), it is assumed that $\sigma^2 = 1$ and $\mathbf{s}'\mathbf{s} = 1$, which insures that $P_{MF}(\bar{f}_d, \bar{\theta}_0) = 1$ when only receiver noise is present. Equation (3.38) is generally referred to as the minimum variance or Capon super-resolution spectral estimator [23, 29, 30] and can be interpreted as the expected value of the output power of the optimum colored-noise matched filter (optimum space-time beamformer), when a unity gain on signal constraint is invoked.

The reason (3.38) achieves a much higher resolution than the Fourier-based estimator of (3.16) is that the weight vector minimizes leakage from all other angle-Doppler, while maintaining a mainbeam constraint [23]. To see this explicitly, consider the Capon estimator (weight vector), which satisfies the following optimization problem:

$$\min_{\{\mathbf{w}\}} E(|\mathbf{w}' \mathbf{x}|^2) = \mathbf{w}' R \mathbf{w} \quad (3.39)$$

$$\text{subject to: } \mathbf{w}' \mathbf{s} = 1$$

Since R is assumed to be positive-definite, the objective function is a convex quadratic form subject to a linear equality constraint [31]. Thus, a necessary and sufficient condition for an extremum is that the first (vector) derivative of the following augmented objective function vanish:

$$\frac{d(\mathbf{w}' R \mathbf{w} + \mu(1 - \mathbf{w}' \mathbf{s}))}{d\mathbf{w}} = 0 \quad (3.40)$$

where μ is the Lagrange multiplier associated with the unity gain constraint [31]. From the vector derivative identities [31]

$$\frac{d\mathbf{w}' R \mathbf{w}}{d\mathbf{w}} = 2R\mathbf{w} \quad (3.41)$$

and

$$\frac{d\mathbf{w}' \mathbf{s}}{d\mathbf{w}} = 2\mathbf{s} \quad (3.42)$$

we have, upon substitution into (3.40), the following optimum Capon spectral estimator:

$$\mathbf{w} = \kappa R^{-1} \mathbf{s} \quad (3.43)$$

where to satisfy the unity gain constraint $\mathbf{w}' \mathbf{s} = 1$,

$$\kappa = \frac{1}{\mathbf{s}' R^{-1} \mathbf{s}} \quad (3.44)$$

Note that (3.43) is the same expression as the max SINR beamformer. The only difference is the choice of normalization parameter. The beamformer output power corresponding to (3.43) is thus given by

$$\begin{aligned}
E\{|w'x|^2\} &= w'E\{xx'\}w \\
&= w'Rw \\
&= (\kappa^* s'R^{-1})R(\kappa R^{-1}s) \\
&= |\kappa|^2 s'R^{-1}s \\
&= \frac{s'R^{-1}s}{|s'R^{-1}s|^2} \\
&= \frac{1}{s'R^{-1}s}
\end{aligned} \tag{3.45}$$

which is the expression in (3.38).

Figure 3.10 contains a comparison of the Fourier-based [i.e., (3.16)] and Capon power-spectral density plots; the super-resolution properties of the space-time matched-filter-based method are clearly evident. Finer resolution is often very useful for better understanding the underlying properties of the interference, as well as the performance of optimum space-time beamformers. This will be made evident in Chapter 4, when clutter subspace leakage phenomena (e.g., internal clutter motion) are introduced.

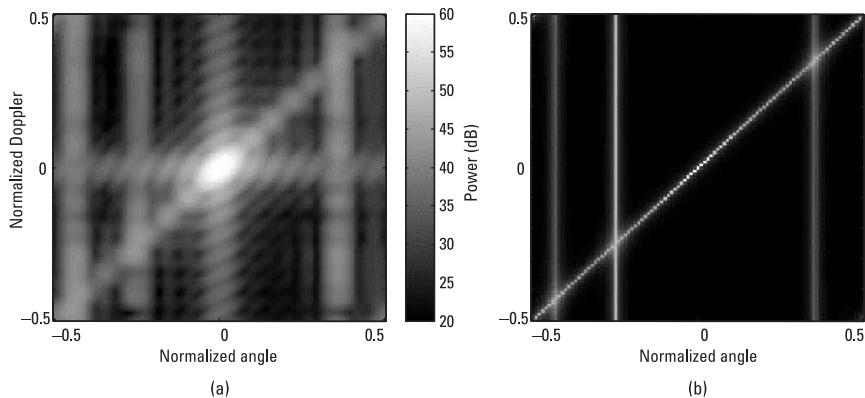


Figure 3.10 Comparison of (a) Fourier-based and (b) minimum variance (MV) power spectral estimators. Note the high resolution of the *minimum variance* (MV) technique implicitly utilized by the optimal space-time beamformer.

3.4 STAP

With Chapter 2 as background, it is clear that STAP is simply the practical attempt to implement optimum space-time processing [i.e., (3.31)]. Since, in practice, the interference environment is generally not known *a priori*, it is necessary to approximate the “exact” or “ideal” covariance required in (3.31) with a finite sample estimate. Again, this is assumed to be accomplished by the availability of a suitable training region in which the interference is present (without any target-like signals)—as was the assumption in Chapter 2.

Unfortunately, the sample support requirements are much greater for STAP than for the relatively simple 1-D adaptive beamforming examples considered in Chapter 2. For example, from the RMB result [see (2.24)], we see that at least $2NM = 512$ samples are required for the $N = M = 16$ case to insure that the resulting SINR is within about 3 dB of optimum, compared with only 32 samples for the $N = 16$ 1-D case. Assuming that the interference is stationary over hundreds of range bins is simply not realistic in many applications. Thus, from the very onset, we are faced with the inherent requirement to develop techniques that retain the benefits of 2-D filtering, but reduce sample support requirements. Whatever methods are considered must also be robust to other real-world effects (internal clutter motion, channel match, and so forth) described in Chapter 4. Thus, we will defer a discussion of potential candidate algorithms to Chapter 5.

Figure 3.11(a) shows the impact on the adapted pattern of replacing the ideal covariance with a finite sample estimate for the case considered in Figure 3.8. Note the presence of very high sidelobes and a distorted main lobe (even though 30-dB Chebyshev angle and Doppler tapers were applied). However, as previously analyzed in Section 2.4, a potential remedy is the use of principal components or diagonal loading. Figure 3.11(b) shows the adapted pattern resulting from the addition of 10 dB of diagonal loading to the sample estimate (along with 30-dB Chebyshev angle and Doppler tapers). It is important to note that these results still required a significant number of i.i.d. training data and, thus, may still not be practical.

3.5 Summary

In this chapter, the rationale for joint 2-D space and time (angle-Doppler) processing was established. The detailed nature of the angle-Doppler structure of clutter was thoroughly examined from a variety of perspectives. In particu-

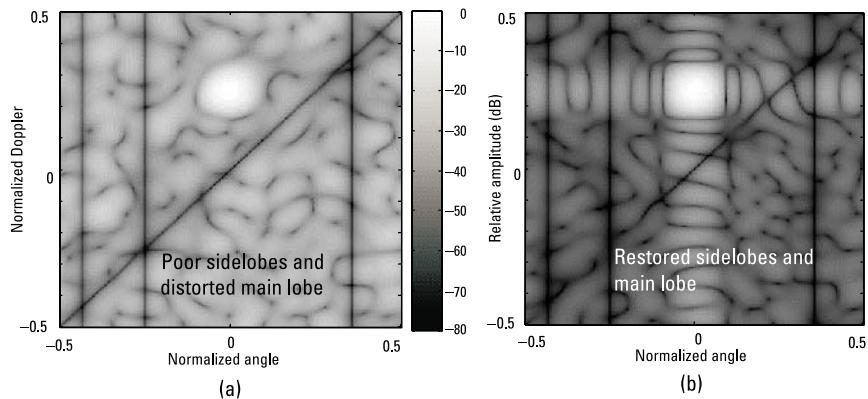


Figure 3.11 Impact of finite sample support on 2-D adapted patterns: (a) adapted pattern resulting from $L = 2NM = 512$ samples, and (b) result if 10 dB of diagonal loading is added. Both patterns include 30-dB Chebyshev angle and Doppler tapers.

lar, a detailed eigen analysis of the space-time clutter covariance matrix was performed and an expression (i.e., Brennan's rule) for its rank as a function of β was derived. Next, an expression for the total space-time clutter-plus-jamming-plus-noise covariance was derived for the case of uncorrelated noise jamming. Utilizing the optimization framework established in Chapter 2, an expression for the optimal SINR space-time beamformer was derived and illustrated with a multiple jammer-plus-clutter scenario. Finally, STAP was introduced via the substitution of the ideal covariance matrix (unknown a priori) with an estimate obtained from sample data. Some adapted pattern-distortion issues due to finite sample support (estimation errors) were also illustrated along with some rudimentary basic fixes (i.e., diagonal loading).

References

- [1] Ward, J., *Space-Time Adaptive Processing for Airborne Radar*, MIT Technical Report 1015, MIT Lincoln Laboratory, December 1994.
- [2] Klemm, R., *Space-Time Adaptive Processing: Principles and Applications*, London, England: IEEE Press, 1998.
- [3] Schleher, D. C., *MTI and Pulsed Doppler Radar*, Norwood, MA: Artech House, 1991.
- [4] Guerci, J. R., "Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming," *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–986.

- [5] Rabideau, D. J., and A. O. Steinhardt, "Improved Adaptive Clutter Cancellation Through Data Adaptive Training," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 35, No. 3, July 1999, pp. 879–891.
- [6] Guerci, J. R., "Knowledge-Aided Sensor Signal Processing and Expert Reasoning," *Proc. of 2002 Workshop on Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER)*, Washington, D.C., April 3, 2002 (CD-ROM).
- [7] Melvin, W. L., "Space-Time Adaptive Radar Performance in Heterogeneous Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 621–633.
- [8] Bergin, J. S., et al., "GMTI STAP in Target-Rich Environments: Site-Specific Analysis," *Proc. of IEEE 2002 Radar Conference*, Long Beach, CA, April 22–25, 2002, pp. 391–396.
- [9] Melvin, W. L., and J. R. Guerci, "Adaptive Detection in Dense Target Environments," *Proceedings of the IEEE 2002 Radar Conference*, Atlanta, GA, May 1–3, 2001, pp. 187–192.
- [10] Melvin, W., (ed.), "Space-Time Adaptive Processing and Adaptive Arrays: Special Collection of Papers," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 508–509.
- [11] Klemm, R., (ed.), Special Issue on Space-Time Adaptive Processing, *Electronics and Communication Engineering Journal*, Vol. 11, February 1999.
- [12] Barton, D. K., *Radar Systems Analysis*, Dedham, MA: Artech House, 1976.
- [13] "Tethered Aerostat Radar System (TARS)," <http://www2.acc.af.mil/library/factsheets/tars.html>.
- [14] Davis, M. E., "Technology Challenges in Affordable Space Based Radar," *Record of the 2000 IEEE International Radar Conference*, Alexandria, VA, May 7–12, 2000, pp. 18–23.
- [15] Long, M. W., *Radar Reflectivity of Land and Sea*, 3rd ed., Norwood, MA: Artech House, 2001.
- [16] Mailloux, R. J., *Phased Array Antenna Handbook*, Norwood, MA: Artech House, 1994.
- [17] Horn, R. A., and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, England: Cambridge University Press, 1991.
- [18] Brennan, L. E., and J. D. Mallett, "Efficient Simulation of External Noise Incident on an Array," *IEEE Trans. on Antennas and Propagation*, Vol. 24, 1976, pp. 740–741.
- [19] Knott, F. K., J. F. Shaeffer, and M. T. Tuley, *Radar Cross Section*, 2nd ed., Norwood, MA: Artech House, 1993.
- [20] Billingsley, J. B., *Exponential Decay in Windblown Radar Ground Clutter Doppler Spectra: Multifrequency Measurements and Model*, Technical Report 997, MIT Lincoln Laboratory, Lexington, MA, July 29, 1996.
- [21] Billingsley, J. B., *Low Angle Radar Land Clutter: Measurements and Empirical Models*, Raleigh, NC: SciTech Publishing, 2002.
- [22] Papoulis, A., and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 2001.

- [23] Pillai, S. U., *Array Signal Processing*. New York: Springer-Verlag, 1989.
- [24] Rabideau, D., and S. Kogon, "A Signal Processing Architecture for Space-Based GMTI Radar," *Proc. of IEEE Radar Conference*, Waltham, MA, 1999, pp. 96–101.
- [25] Guerci, J. R., and J. S. Bergin, "Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002.
- [26] Brennan, L. E., and F. M. Staudaher, *Subclutter Visibility Demonstration*, Technical Report RL-TR-92-21, Adaptive Sensors Inc., 1992.
- [27] Skolnik, M., *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
- [28] Horn R. A., and C. R. Johnson, *Matrix Analysis*, Cambridge, England: Cambridge University Press, 1985.
- [29] Gabriel, W. F., "Spectral Analysis and Adaptive Array Superresolution Techniques," *Proc. of IEEE*, Vol. 68, No. 6, 1980, pp. 654–666.
- [30] Capon, J., "High-Resolution, Frequency-Wavenumber Spectral Analysis," *Proc. of IEEE*, Vol. 57, No. 8, 1969.
- [31] Pierre, D., *Optimization Theory with Applications*, New York: Dover, 1986.

4

Other Important Factors Affecting STAP Performance

4.1 Introduction

Due to a multitude of practical considerations, the idealized performance presented in Chapter 3 represents a generally unachievable upper bound on performance—for both the known and iid sample estimate covariance cases. In this chapter, we will examine some of the salient factors that contribute to this degradation and must be considered when both designing a STAP system, as well as predicting realized performance. Most, if not all, of the factors described in this chapter are present to a greater or lesser degree in any STAP radar. Thus, when comparing various STAP algorithms (such as those described in Chapter 5 and elsewhere), careful attention should be paid to how they perform when one or more of the generally deleterious effects described herein is present.

This chapter is not meant to serve as a comprehensive treatise on factors that can affect STAP performance. It is, however, designed to illustrate an important phenomenon of real-world STAP systems: The effective clutter/interference rank is almost certainly greater than would be the case in idealized, simplified signal and systems models, and the rank increases with increasing *interference-to-noise ratio* (INR). This latter dependency on INR is referred to as the *iceberg effect* (see Section 4.6), since increasing INR tends to “lift” the colored interference eigenvalues further above the “ocean” noise floor—thereby exposing more of the “iceberg” (effective interference rank).

In varying clutter terrain situations, this can be a significant source of nonstationarity and often will result in wider Doppler clutter notches (and poorer minimum detectable velocities). Advanced methods for addressing these issues will be discussed and analyzed in Chapter 5.

Although a multitude of factors contribute to increasing the effective rank (as defined in Chapter 3) of the interference subspace, the net effect is essentially a decorrelation between spatial and temporal measurements (channels and/or pulses). The phenomena that give rise to this rank-increasing effect are referred to as interference subspace leakage, or ISL, since the decorrelation tends to cause a smearing or leakage of the interference subspace [1]. For example, in the derivation of Brennan's rule (Chapter 3), it was assumed that if the clutter was sampled at a later time (e.g., one PRI later) with an identical antenna channel observing from exactly the same spot, the total clutter return would be identical (i.e., perfectly correlated). If, however, ICM is present, the signals will not be the same, and the net effect is a decorrelation between samples. Moreover, any mismatches between antenna channels could likewise cause differences in the observed signals, again potentially resulting in decorrelation.

As we will see later in this chapter, a convenient mathematical representation of decorrelating ISL (and nondecorrelating channel mismatch) is via CMT [1–4]. Specifically, if $R \in \mathbb{C}^{NM \times NM}$ denotes the total interference covariance matrix, then the resultant covariance including all ISL effects has the form [1]

$$R \xrightarrow{\text{ISL}} R \circ T_1 \circ T_2 \circ \dots \circ T_k \stackrel{\Delta}{=} \tilde{R} \in \mathbb{C}^{NM \times NM} \quad (4.1)$$

where $T_1, T_2, \dots, T_k \in \mathbb{C}^{NM \times NM}$ are positive- (semi-) definite Hermitian matrices associated with uncorrelated ISL phenomenon (such as ICM and decorrelating channel mismatch—see below), and \circ denotes Hadamard (elementwise) matrix multiplication (see Chapter 2).

This mathematical form turns out to be more than just a convenient compact representation of realistic covariance structure. For example in Chapter 5, the CMT structure is explicitly incorporated into two STAP algorithms (PC-CMT and SMI-CMT) that retain the desirable minimal sample support properties of principal-component and diagonal-loading SMI techniques (i.e., non-ISL environments) respectively, yet account for ISL effects with an increase in sample support requirements [1].

The remainder of this chapter is organized as follows: In Section 4.2 we examine channel-mismatch effects, which are delineated into two major

categories: bandwidth and angle-dependent effects. For example, in the narrowband signal model case, channel mismatch results in an amplitude and phase offset between channels that can be modeled as an unknown (but fixed over the CPI) multiplicative complex scalar for each channel. This form of mismatch is nondecorrelating and is shown not to result in an increase in the effective interference rank, whereas in the finite (nonzero) bandwidth case, both bandwidth dispersion effects [5] and transfer function differences [6] result in decorrelating mismatch.

Angle-dependent channel-mismatch effects arise when the amount of channel mismatch depends on the AoA. For example, element position errors or mutual coupling [7] tend to have an angle dependency. However, as discussed in Section 4.2, since these errors are small to begin with, the dominant effects are associated with the look direction of the radar. Thus, the angle-dependent CMT $T(\theta)$ can be approximated by the fixed CMT $T(\theta) \rightarrow T(\theta_0) = T_0$ and then incorporated into (4.1).

Section 4.3 discusses other ISL effects such as ICM and range-walk. Again, it is shown that these effects result in a spatial and/or temporal decorrelation that can be conveniently modeled by (4.1) (under certain general conditions).

Section 4.4 examines the impact of antenna misalignment with the ground-track velocity vector (i.e., crabbing [8]) on the effective clutter rank and derives an expression for the angle-Doppler clutter locus. A brief discussion of nonlinear array geometries is contained in Section 4.5, which also includes a circular array example.

Section 4.6 contains a discussion and illustration of the mechanism by which the aforementioned ISL phenomenon couple into CNR nonstationarity (which occurs in nature due to terrain variability [9]) to produce a more complex nonstationary vector stochastic process whose rank is likewise nonstationary ([1, 10]). This iceberg effect is fundamental, omnipresent, and must be accounted for in the STAP design stage (as illustrated in Chapter 5)—otherwise clutter under- or overnulling will result [1]. More exotic radar applications such as bistatic space-time adaptive radar [11] generally only exacerbate this effect.

4.2 Channel Mismatch

An important and ever-present factor affecting STAP performance is so-called channel mismatch (referred to colloquially as *channel match*). The N separate antenna channels employed in our ULA model have been assumed

up to this point to be ideal; that is, the output response of a ULA was assumed to have a perfect linear phase (Vandermonde) response and no mismatch-induced decorrelation between channels.

In reality, it is impossible to have each channel pathway perfectly matched from an input-output perspective for all AoAs. For analysis purposes, it is convenient to taxonomize each contributing channel-mismatch source based on whether or not it is dependent on AoA and whether it is narrowband (or not). Angle-dependent channel-mismatch errors are often due to array manifold and possibly radome effects (e.g., element position errors and pattern differences, near field scattering and multipath) and are often quite complex and difficult to model and/or predict accurately [1, 6, 12]. Angle-independent sources are generally due to mismatches between the channels after the antenna elements [6, 12].

To facilitate the analysis of channel mismatch, we will adopt the CMT framework that provides a convenient and unifying mathematical framework for a broad class of channel mismatch [1, 2]. Moreover, this framework will also prove useful when examining other subspace leakage phenomenon such as ICM, as well as when considering model-based minimal sample support STAP methods in Chapter 5.

4.2.1 Angle-Independent Channel Mismatch

Angle-independent channel mismatch is due to variations between the transfer functions, $H_n(\omega)$, $n = 1, \dots, N$, associated with each receive channel (see Figure 4.1) without angle-dependent effects (see following discussion). We will first consider the ideal narrowband case in which each transfer function simplifies to a complex gain multiplier.

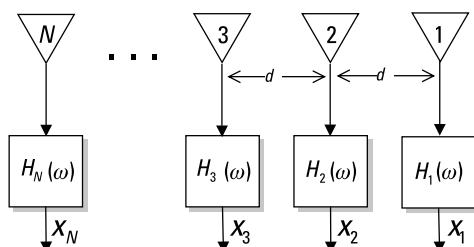


Figure 4.1 Angle-independent channel mismatch for narrowband arrays arises from variations in the signal paths in each receive channel generally after the array manifold.

4.2.1.1 Narrowband Case

The impact of angle-independent narrowband channel mismatch that is stable over a CPI (normally the case) is the introduction of unknown but constant gain and phase differences (i.e., mismatch) between channels. This is conveniently modeled as a constant tapering of all signal vectors impinging on the array. Specifically, if \mathbf{a}_i denotes the steering vector associated with the i th incoming signal, then the resulting actual array response vector $\hat{\mathbf{a}}_i$ is given by

$$\hat{\mathbf{a}}_i = \mathbf{a}_i \odot \mathbf{t} \quad (4.2)$$

where \mathbf{t} is an N -dimensional vector (or taper) of the form

$$\mathbf{t} = [\epsilon_1 e^{j\varphi_1} \dots \epsilon_N e^{j\varphi_N}]^T \quad (4.3)$$

In (4.3), $\epsilon_1, \dots, \epsilon_N$ and $\varphi_1, \dots, \varphi_N$ denote the amplitude and phase errors, respectively. The amplitude errors represent the differences in gain through each channel, and thus $0 \leq \epsilon_i \leq 1$ for $i = 1, \dots, N$. Typically, the variation in gain is quite small and is usually less than 0.1 dB (i.e., $\sim 0.99 \leq \epsilon_i \leq 1$). Typical values for phase errors are 5° or less. Note that in general $\|\mathbf{t}\| \leq 1$. Also, it has been tacitly assumed that the noise floor in each channel is identical.

As discussed in Chapter 2, these errors introduce a signal mismatch that limits achievable tapered sidelobe levels and signal integration gain. For a given mismatch taper \mathbf{t} , the SNR loss ratio, defined as the ratio of the realized SNR to optimum SNR [see (2.10)] is given by

$$\begin{aligned} \text{SNR}_{\text{loss}} &= \frac{\text{SNR}_{\text{act}}}{\text{SNR}_{\text{opt}}} = \frac{\frac{|\mathbf{s}'(\mathbf{s} \odot \mathbf{t})|^2}{\sigma^2}}{\frac{|\mathbf{s}'\mathbf{s}|^2}{\sigma^2}} \\ &= \frac{\mathbf{s}'(\mathbf{s} \odot \mathbf{t})(\mathbf{s} \odot \mathbf{t})'\mathbf{s}}{|\mathbf{s}'\mathbf{s}|^2} \\ &= \frac{\mathbf{s}'[(\mathbf{s}\mathbf{s}') \odot (\mathbf{t}\mathbf{t})']\mathbf{s}}{|\mathbf{s}'\mathbf{s}|^2} \\ &= \frac{\mathbf{s}'[(\mathbf{s}\mathbf{s}') \odot T]\mathbf{s}}{|\mathbf{s}'\mathbf{s}|^2} \end{aligned} \quad (4.4)$$

where

$$T = \mathbf{t}\mathbf{t}' \in \mathbb{C}^{N \times N} \quad (4.5)$$

is a rank-one matrix referred to as a CMT [2]. The final factored form made use of the Hadamard matrix multiplication identity

$$(A \circ B)(C \circ D)' = (AC') \circ (BD') \quad (4.6)$$

where A, B are $m \times n$ and C, D are $n \times p$ matrices [13]. Note that in the absence of channel mismatch, T is the Hadamard identity matrix (i.e., matrix with all unity entries) [13], and (4.4) reduces to unity (i.e., no loss).

In a well-calibrated radar, the SNR loss due to signal mismatch is quite small (~ 1 dB or less). Of greater concern in general is the impact of channel mismatch on interference suppression. For the angle-independent narrow-band case, each signal impinging on the array is tapered according to (4.2). Thus, the total received space-time interference signal is of the form

$$\begin{aligned} \mathbf{X} &= \mathbf{X}_c + \mathbf{X}_J + \mathbf{n} \\ &= \sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i \odot \mathbf{t} + \sum_{j=1}^{N_J} \tilde{\zeta}_j \mathbf{s}_j \odot \mathbf{t} + \mathbf{n} \end{aligned} \quad (4.7)$$

where $\mathbf{X}_c, \mathbf{X}_J, \mathbf{n} \in \mathbb{C}^{NM}$ denote the space-time clutter, jamming, and receiver noise RVs, respectively. Note that the mismatch taper only affects the structured (colored) noise signals and not the thermal or white noise. This is a consequence of assuming that each channel has the same noise floor (differences in receiver noise levels can be viewed as differences in SNR, which ultimately can be represented as differences in signal gain for a fixed noise-floor level).

If only spatial channel mismatch is present, that is, (4.2), then the corresponding space-time mismatch taper, \mathbf{t} , in (4.7) is of the form

$$\mathbf{t} = [\underbrace{1 \dots 1}_M]^T \otimes [\epsilon_1 e^{j\varphi_1} \dots \epsilon_N e^{j\varphi_N}]^T \quad (4.8)$$

Examining (4.7) in further detail, we see that the first sum is due to clutter (see Chapter 3) with the inclusion of mismatches, while the second sum is due to N_J uncorrelated jammers (see Chapters 2 and 3). Assuming that

the clutter, jamming, and receiver noise obey the aforementioned statistical models described in Chapters 2 and 3, and that they are mutually uncorrelated, the resulting total interference covariance in the presence of simple channel mismatch is given by

$$\begin{aligned}
 \text{cov}(\mathbf{X}) &\stackrel{\Delta}{=} R \\
 &= \sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) \odot (\mathbf{t} \mathbf{t}') + \sum_{j=1}^{N_c} \sigma_j^2 (\mathbf{v}_j \mathbf{v}'_j) \odot (\mathbf{t} \mathbf{t}') + \sigma^2 I \\
 &= \left(\sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) + \sum_{j=1}^{N_c} \sigma_j^2 (\mathbf{v}_j \mathbf{v}'_j) \right) \odot (\mathbf{t} \mathbf{t}') + \sigma^2 I \quad (4.9) \\
 &= R_I \odot T + \sigma^2 I \\
 &= \tilde{R}_I + \sigma^2 I
 \end{aligned}$$

where $\tilde{R}_I = R_I \odot T$, and R_I is the total interference covariance matrix (clutter plus jamming) in the absence of mismatch. T is the rank-one CMT associated with the mismatch; that is, $T = \mathbf{t} \mathbf{t}'$.

An important result regarding the impact of simple channel mismatch on STAP performance can be immediately inferred from (4.9) based on the properties of Hadamard products. Specifically, since $\text{rank}(A \odot B) \leq \text{rank}(A) \text{rank}(B)$, we have [13]

$$\begin{aligned}
 \text{rank}(\tilde{R}_I) &= \text{rank}(R_I \odot T) \\
 &= \text{rank}(R_I \odot (\mathbf{t} \mathbf{t}')) \quad (4.10) \\
 &\leq \text{rank}(R_I) \text{rank}(T) \\
 &= \text{rank}(R_I) \text{rank}(\mathbf{t} \mathbf{t}'')
 \end{aligned}$$

However, since $\text{rank}(T) = \text{rank}(\mathbf{t} \mathbf{t}') = 1$, we immediately see that

$$\text{rank}(\tilde{R}_I) \leq \text{rank}(R_I) \quad (4.11)$$

That is, the rank of the total interference covariance matrix cannot be increased by simple angle-independent narrowband channel mismatch. The practical consequence of this result is that optimum nulling performance is

essentially unaffected (there is slight degradation when amplitude mismatch is present—see following discussion). However, there will still be a small amount of SNR loss due to steering-vector mismatch as described by (4.4). Also, limits on nominal tapered sidelobe levels will also be present as described in Chapter 2.

Stronger statements regarding the exact impact of simple channel mismatch can be derived for certain special cases. For example, for phase only mismatch, the eigenvalues of the total interference covariance matrix can be shown to be invariant to the phase errors. Coupled with the rank property of (4.11), this implies that no loss in nulling performance is incurred (there is still a small signal-mismatch loss). To see this explicitly, we begin by assuming that \mathbf{t} is of the form

$$\mathbf{t} = \overbrace{[1 \ \dots \ 1]^T}^M \otimes [e^{j\varphi_1} \dots e^{j\varphi_N}]^T \quad (4.12)$$

that is, phase-only mismatch. Expressing the resulting covariance in terms of the eigen-based representation of the original covariance (no mismatch), we have

$$\begin{aligned} R &= R_I \odot T + \sigma^2 I \\ &= \left(\sum_{i=1}^{NM} \lambda_i \mathbf{u}_i \mathbf{u}'_i \right) \odot T + \sigma^2 I \\ &= \sum_{i=1}^{NM} \lambda_i \mathbf{u}_i \mathbf{u}'_i \odot T + \sigma^2 I \\ &= \sum_{i=1}^{NM} \lambda_i (\mathbf{u}_i \mathbf{u}'_i) \odot (\mathbf{t} \mathbf{t}') + \sigma^2 I \\ &= \sum_{i=1}^{NM} \lambda_i (\mathbf{u}_i \odot \mathbf{t}) (\mathbf{u}_i \odot \mathbf{t})' + \sigma^2 I \end{aligned} \quad (4.13)$$

where R_I is the total colored-noise interference covariance matrix (clutter plus jamming), and $\sigma^2 I$ is the receiver noise (white noise). Correspondingly, $\{\lambda_i, \mathbf{u}_i\}$ denotes the i th eigenvalue–eigenvector pair for R_I .

To prove the eigenvalue invariance property, we need to show that the set of modified vectors $\{\mathbf{u}_i \odot \mathbf{t}\}$ retains the orthonormality property (orthogonal and unit norm). Orthonormality is proved as follows

$$\begin{aligned}
(\mathbf{u}_i \circ \mathbf{t})' (\mathbf{u}_j \circ \mathbf{t}) &= (\mathbf{u}_i^* \circ \mathbf{u}_j)^T (\mathbf{t}^* \circ \mathbf{t}) \\
&= (\mathbf{u}_i^* \circ \mathbf{u}_j)^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \mathbf{u}_i' \mathbf{u}_j = \delta_{ij}
\end{aligned} \tag{4.14}$$

where δ_{ij} denotes the Kronecker delta function. Note that $\mathbf{t}^* \circ \mathbf{t} = [1 \ 1 \ \dots \ 1]^T$ since each element of \mathbf{t} is of the form $e^{j\varphi_i}$ (unity modulus). Thus, the eigenvalues of $R_J \circ T + \sigma^2 I$ are the same as $R_I + \sigma^2 I$, and the eigenvectors of $R_J \circ T + \sigma^2 I$ are those of $R_I + \sigma^2 I$ modified by the phase-only mismatch taper; that is, $\mathbf{u}_i \rightarrow \mathbf{u}_i \circ \mathbf{t}$. Note, however, that if amplitude errors are present, the eigenvalues are not invariant (which impacts SINR—although very modestly in practice).

Figure 4.2 shows the optimum beam pattern (with a 30-dB Chebyshev taper) with channel mismatch for the six-jammer case described in Section 2.2. The phase-only channel-mismatch taper, \mathbf{t} , was selected from a 5° (one-sigma) Gaussian random-number generator. In generating the adapted pattern, modified DFT vectors are required since plane waves impinging on the array no longer obey a perfect linear phase response. If $[\mathbf{s}_0 \ \dots \ \mathbf{s}_j \ \dots]$ denotes the DFT vectors used to generate the optimum response in the absence of mismatch, then modified DFT vectors of the form $[\mathbf{s}_0 \circ \mathbf{t} \ \dots \ \mathbf{s}_j \circ \mathbf{t} \ \dots]$ are required to generate the correct response when mismatch is present.

As expected, the adapted pattern has nulls placed at the jammer locations and an average sidelobe level of approximately 30 dB (variations due to channel mismatch). The corresponding SNR loss computed from (4.4) is 0.6 dB—an extremely modest loss. Indeed, as emphasized in Chapter 2, the ability of adaptive beamforming to effect deep nulls even in the presence of certain channel mismatches is a major practical benefit.

4.2.1.2 Finite (Nonzero) Bandwidth Case

For the finite (nonzero) bandwidth case, the N receive channel transfer functions, $H_n(\omega)$, $n = 1, \dots, N$, can no longer be simply modeled as unknown scalar complex gain multipliers (i.e., a rank-one CMT). Instead, the instantaneous frequency content of a PRI (pulse spectrum) is subtly altered between channels. The net effect for a stochastic signal (i.e., clutter and jamming) is a decorrelation of the signals between channels. To see why this is the case, we first recall the form of the received total interference

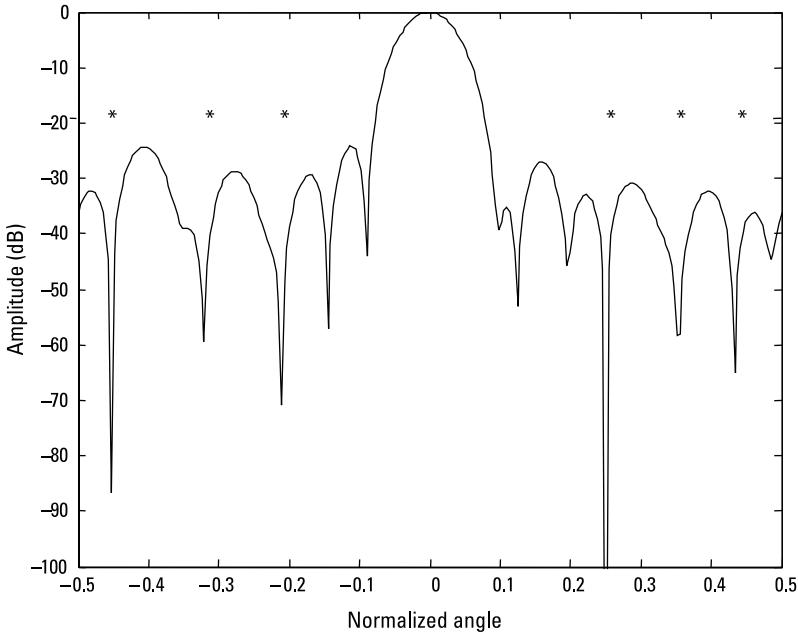


Figure 4.2 Optimum pattern (with 30-dB Chebyshev taper) with 5° (one-sigma) phase-only channel mismatch for the six-jammer case of Chapter 2. Although there is modest 0.6-dB SNR loss due to signal mismatch, the adapted pattern places sufficiently deep nulls at the jammer locations to effectively cancel the interference.

signal with channel mismatch given by (4.7), but with \mathbf{t} a vector RV (as opposed to a fixed, but unknown, constant vector). Since we are focusing on spatial decorrelation effects resulting from channel-to-channel random mismatch [6, 12], it suffices to analyze the pairwise cross correlation between channels. Specifically,

$$\mathbf{E}\{x_i x_k^*\} = \mathbf{E}\{((x_{c_i} + x_{J_i}) \cdot t_i + n_i)((x_{c_k} + x_{J_k}) \cdot t_k + n_k)^*\} \quad (4.15)$$

where (x_i, x_k) denotes the complex (in-phase and quadrature) interference-only outputs of the i th and k th channels for a single PRI, respectively. Under the assumption that the clutter, jamming, thermal (white) noise, and channel amplitude/phase mismatch gains are all mutually uncorrelated (physically justifiable assumptions), (4.15) can be written as

$$\mathbf{E}\{x_i x_k^*\} = (\mathbf{E}\{x_{c_i} x_{c_k}^*\} + \mathbf{E}\{x_{J_i} x_{J_k}^*\}) \cdot \mathbf{E}\{t_i t_k^*\} + \sigma^2 \delta_{ik} \quad (4.16)$$

$$R_{i,k} = R_{I_{i,k}} \cdot T_{i,k} + \sigma^2 \delta_{ik}$$

where $R_{i,k}$ and $R_{I_{i,k}}$ denote the (i, k) th elements of the total and colored-noise-only (clutter plus jamming) spatial covariance matrices, respectively. $T_{i,k} = \mathbf{E}\{t_i t_k^*\}$ is the cross-correlation coefficient due to channel mismatch and has the property $0 \leq |T_{i,k}| \leq 1$, although in practice it is much closer to unity than to zero (see the following discussion).

In matrix form, (4.16) becomes

$$R = R_I \odot T + \sigma^2 I \quad (4.17)$$

where R is the spatial-only covariance matrix, and $T = \mathbf{E}\{\mathbf{t}\mathbf{t}'\}$ is the CMT associated with the channel mismatch. Equation (4.17) is of the same mathematical form as (4.9) with one major exception: T is generally full rank. Thus, unlike the simple rank-one narrowband CMT, the presence of decorrelating channel mismatch will alter the eigenvalues and eigenvectors of the interference covariance matrix and may increase the effective rank of the colored-noise subspace (see the following discussion) [1]. This increased effective rank, in turn, will often result in a widening of the angle-Doppler nulls—thereby reducing the amount of observation space available for target detection.

To illustrate the impact of decorrelating random amplitude and phase channel mismatch on the interference signal, consider a simple uncorrelated amplitude and phase mismatch model:

$$1 - \delta_{\epsilon_i} \leq \epsilon_i \leq 1, \quad (4.18)$$

$$\text{where } p(\delta_{\epsilon_i}) = \begin{cases} \frac{1}{\Delta_\epsilon}, & \text{for } 0 \leq \delta_{\epsilon_i} \leq \Delta_\epsilon, \forall i \\ 0, & \text{elsewhere} \end{cases}$$

and

$$\frac{-\Delta_\varphi}{2} \leq \varphi_i \leq \frac{\Delta_\varphi}{2} \quad (4.19)$$

$$\text{where } p(\varphi_i) = \begin{cases} \frac{1}{\Delta_\varphi}, & \text{for } \frac{-\Delta_\varphi}{2} \leq \varphi_i \leq \frac{\Delta_\varphi}{2}, \forall i \\ 0, & \text{elsewhere} \end{cases}$$

where $p(\delta_{\epsilon_i})$ and $p(\varphi_i)$ are the pdfs (uniform) associated with the amplitude and phase errors respectively. Note that other authors have used different stochastic representations to model decorrelating channel mismatch (e.g., [6, 12, 14, 15])—all of which are but approximations to the actually underlying mismatch mechanisms. In practice, the correlation coefficients required by (4.16) would be measured experimentally as part of the calibration process. Once known, a CMT comprised of those coefficients completely characterizes the effects from an SINR standpoint. That said, we continue with the analytical illustration: With (4.18) and (4.19), we can now calculate the elements of the spatial-only CMT as follows:

$$\begin{aligned} [T]_{i,k} &= E\{t_i t_k^*\} \\ &= E\{\epsilon_i e^{j\varphi_i} \epsilon_k e^{-j\varphi_k}\} \\ &= E\{\epsilon_i\} E\{\epsilon_k\} E\{e^{j\varphi_i}\} E\{e^{-j\varphi_k}\} \\ &= \left(1 - \frac{\Delta_\epsilon}{2}\right)^2 \text{sinc}^2\left(\frac{\Delta_\varphi}{2}\right) \end{aligned} \quad (4.20)$$

for $i \neq k$

with diagonal elements

$$\begin{aligned} [T]_{i,i} &= E\{\epsilon_i^2 e^{j\varphi_i} e^{-j\varphi_i}\} \\ &= E\{\epsilon_i^2\} \\ &= 1 - \Delta_\epsilon + \frac{1}{3}\Delta_\epsilon^3 \end{aligned} \quad (4.21)$$

Thus, T is of the following linear matrix form:

$$T = \rho_1 \mathbf{1} + \rho_2 I \quad (4.22)$$

where $\mathbf{1}$ is the rank-one Hadamard identity matrix [13] (all elements equal unity), I is the usual identity matrix, and

$$\rho_1 = [T]_{i,k} \quad (4.23)$$

$$\rho_2 = [T]_{i,i} - [T]_{i,k} = [T]_{i,i} - \rho_1 \quad (4.24)$$

Applying (4.22) to R_I yields

$$\begin{aligned} R_I \odot T &= R_I \odot (\rho_1 \mathbf{1} + \rho_2 I) \\ &= \rho_1 R_I \odot \mathbf{1} + \rho_2 R_I \odot I \\ &= \rho_1 R_I + \rho_2 \operatorname{diag}\{R_I\} \end{aligned} \quad (4.25)$$

where the Hadamard product identities $R_I \odot \mathbf{1} = R_I$ and $R_I \odot I = \operatorname{diag}\{R_I\}$ were invoked ($\operatorname{diag}\{R_I\}$ is a diagonal matrix whose entries are the diagonal elements of R_I). From (4.25) we see that the net effect of uniform decorrelating channel mismatch is to effectively raise the white-noise floor (increased diagonal entries). Thus, for this particular case, the spatial colored-noise-only covariance rank is unaffected (although the effective rank (Chapter 3) may change since the INR has changed). However, this is not generally the case for the space-time clutter rank as illustrated below. Nor is it generally the case when the mismatch is nonuniform, that is, when the off-diagonal terms of T are not identical—which, in other words, means the pairwise correlation between channels is nonuniform.

Note also that in the absence of amplitude modulation, that is, $\Delta_\epsilon = 0$, $\rho_2 = 1 - \rho_1$, and thus (4.22) is a convex sum of the Hadamard and ordinary identity matrices. The resulting covariance is of the form

$$\begin{aligned} R &= R_I \odot T + \sigma^2 I \\ &= R_I \odot (\rho_1 \mathbf{1} + \rho_2 I) + \sigma^2 I \\ &= R_I \odot (\rho_1 \mathbf{1} + (1 - \rho_1)I) + \sigma^2 I \\ &= R_I \odot \begin{bmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & & \\ \vdots & & \ddots & \\ \rho_1 & & & 1 \end{bmatrix} + \sigma^2 I \end{aligned} \quad (4.26)$$

The expression in (4.22) is for the spatial-only CMT. The corresponding full-dimensional space-time CMT is simply given by

$$T = \mathbf{1}_{M \times M} \otimes T_{\text{space}} \quad (4.27)$$

where $\mathbf{1}_{M \times M}$ is the $M \times M$ Hadamard identity matrix, and T_{space} denotes the spatial-only CMT.

Equation (4.27) is a special case of a more general expression

$$T = T_{\text{time}} \otimes T_{\text{space}} \quad (4.28)$$

when there are also temporal mismatches and/or decorrelating effects that are uncorrelated with the spatial factors (e.g., ICM). A proof of (4.28) is relatively straightforward: Let $\mathbf{t}_t \in \mathbb{C}^M$, $\mathbf{t}_s \in \mathbb{C}^N$ denote the uncorrelated temporal and spatial vector RVs, respectively, with corresponding correlation matrices $\text{cov}(\mathbf{t}_t) = T_{\text{time}} \in \mathbb{C}^{M \times M}$, $\text{cov}(\mathbf{t}_s) = T_{\text{space}} \in \mathbb{C}^{N \times N}$, then $\text{cov}(\mathbf{t}_t \otimes \mathbf{t}_s) \stackrel{\Delta}{=} T$ is given by

$$\begin{aligned} T &\stackrel{\Delta}{=} E((\mathbf{t}_t \otimes \mathbf{t}_s)(\mathbf{t}_t \otimes \mathbf{t}_s)') \\ &= E((\mathbf{t}_t \mathbf{t}_t') \otimes (\mathbf{t}_s \mathbf{t}_s')) \\ &= E(\mathbf{t}_t \mathbf{t}_t') \otimes E(\mathbf{t}_s \mathbf{t}_s') \\ &\stackrel{\Delta}{=} T_{\text{time}} \otimes T_{\text{space}} \end{aligned} \quad (4.29)$$

where use was made of the Kronecker identity $(A \otimes B)(C \otimes D)' = (AC') \otimes (BD')$ [13] (where conformality of the constituent matrix dimensions is assumed). We will make use of (4.28) later in this book when a multitude of spatial and temporal effects are present simultaneously.

To gain insight into the impact decorrelating channel mismatch can have on the space-time clutter covariance matrix, recall from Chapter 3 the Riemann sum approximation for the total space-time clutter RV $\mathbf{X}_c \in \mathbb{C}^{NM}$,

$$\begin{aligned} \mathbf{X}_c &= \sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i \\ &= \sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{b}_i \otimes \mathbf{a}_i \end{aligned} \quad (4.30)$$

where $\{\tilde{\gamma}_i\}$ represents mutually uncorrelated zero-mean complex-valued scalar RVs with $E\{|\tilde{\gamma}_i|^2\} = G_i$, and $\mathbf{v}_i = \mathbf{b}_i \otimes \mathbf{a}_i \in \mathbb{C}^{NM}$ is the space-time steering vector corresponding to the i th clutter patch (see Chapter 3) with corresponding temporal $\mathbf{b}_i \in \mathbb{C}^M$ and spatial $\mathbf{a}_i \in \mathbb{C}^N$ steering vectors. Uniform spatial-only decorrelating channel mismatch alters (4.30) as follows:

$$\sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{b}_i \otimes \mathbf{a}_i \rightarrow \sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{b}_i \otimes (\mathbf{a}_i \odot \mathbf{t}_s) \quad (4.31)$$

where $\mathbf{t}_s \in \mathbb{C}^N$ is the vector RV associated with the decorrelating channel mismatch. The corresponding resultant covariance matrix, assuming $\{\tilde{\gamma}_i\}$ and \mathbf{t}_s are uncorrelated, is thus given by

$$\begin{aligned}
 \text{cov}\left(\sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{b}_i \otimes (\mathbf{a}_i \odot \mathbf{t}_s)\right) &= \sum_{i=1}^{N_c} \text{cov}(\tilde{\gamma}_i \mathbf{b}_i \otimes (\mathbf{a}_i \odot \mathbf{t}_s)) \\
 &= \sum_{i=1}^{N_c} \text{cov}(\tilde{\gamma}_i (\mathbf{b}_i \otimes \mathbf{a}_i) \odot (\mathbf{1}_M \otimes \mathbf{t}_s)) \\
 &= \sum_{i=1}^{N_c} \text{cov}(\tilde{\gamma}_i) \text{cov}((\mathbf{b}_i \otimes \mathbf{a}_i) \odot (\mathbf{1}_M \otimes \mathbf{t}_s)) \\
 &= \sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) \odot \text{cov}((\mathbf{1}_M \otimes \mathbf{t}_s)) \quad (4.32) \\
 &= \sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) \odot (\mathbf{1}_{M \times M} \otimes T_{\text{space}}) \\
 &= \left(\sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) \right) \odot (\mathbf{1}_{M \times M} \otimes T_{\text{space}}) \\
 &= R_c \odot (\mathbf{1}_{M \times M} \otimes T_{\text{space}})
 \end{aligned}$$

which is of the factored form in (4.17) with a CMT of the form given by (4.27). Note that the above derivation depends on the angle-independence of the decorrelating channel mismatch. However, as shown below, even for angle-dependent mismatch, it is still possible to approximate the resultant covariance with the factored form.

To illustrate the impact of decorrelating channel-mismatch errors on the space-time clutter eigenspectrum, consider the following three cases: (1) $\Delta_\epsilon = 0$, $\Delta_\varphi = 0^\circ$, that is, no errors; (2) $\Delta_\epsilon = 0.01$, $\Delta_\varphi = 2^\circ$; and (3) $\Delta_\epsilon = 0.02$, $\Delta_\varphi = 5^\circ$. To make a fair comparison, we will hold the original CNR constant regardless of the amount of mismatch. This is tantamount to adjusting the original white-noise floor in (4.17) to preserve CNR as defined in Chapter 3.

Figure 4.3(a) shows the total interference eigenvalues for the three cases. Note that the effective rank of the clutter is increased due to the

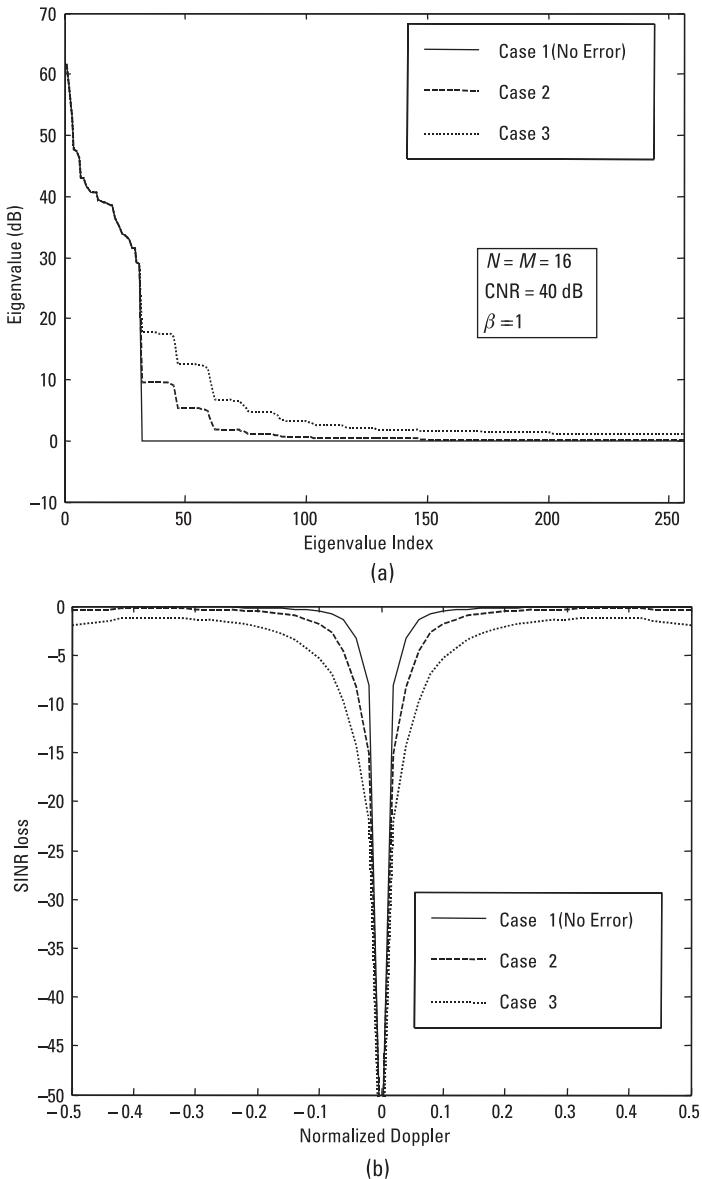


Figure 4.3 Impact of decorrelating channel mismatch on (a) the clutter eigenspectrum and (b) SINR loss for (1) $\Delta_\epsilon = 0$, $\Delta_\varphi = 0^\circ$, that is, no errors; (2) $\Delta_\epsilon = 0.01$, $\Delta_\varphi = 2^\circ$; and (3) $\Delta_\epsilon = 0.02$, $\Delta_\varphi = 5^\circ$.

presence of decorrelating channel mismatches, which will have a deleterious impact on SINR performance. Figure 4.3(b) shows the impact on SINR loss. Note the increased width in the mainbeam clutter notch—resulting in poorer low-Doppler-shift signal detection. Also evident is the effective increase in the noise floor as a consequence of the structure of (4.22) and (4.26).

In general, both the simple rank-one CMT of (4.9), denoted as T_1 , and the (generally) full-rank CMT of (4.22), denoted as T_2 , will be present. The net effect, assuming that the errors are uncorrelated, is a composite CMT of the form

$$T = T_1 \circ T_2 \quad (4.33)$$

A proof of the Hadamard multiplicative property can be found in [1] and is based on the fact that the expected value of the product of two uncorrelated RVs is simply the product of their expected values [16]. Specifically, consider the (i, k) th element of (4.33); that is,

$$\begin{aligned} [T]_{i,k} &= [\mathbb{E}((\mathbf{t}_1 \circ \mathbf{t}_2)(\mathbf{t}_1 \circ \mathbf{t}_2)')]_{i,k} \\ &= \mathbb{E}(t_{1_i} t_{2_i} t_{1_k}^* t_{2_k}^*) \\ &= \mathbb{E}(t_{1_i} t_{1_k}^*) \mathbb{E}(t_{2_i} t_{2_k}^*) \\ &= [T_1 \circ T_2]_{i,k} \end{aligned} \quad (4.34)$$

where $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{C}^{NM}$ are the vector RV modulations with associated correlation matrices $T_1, T_2 \in \mathbb{C}^{NM \times NM}$, respectively. Note that the above argument can be repeated to accommodate any number of uncorrelated random effects, resulting in the Hadamard factored form of (4.1).

An interesting fact associated with the Hadamard product of two matrices is that if T_1 and T_2 are both positive-definite, then $T = T_1 \circ T_2$ is also guaranteed to be positive-definite as a direct consequence of the Schur Product theorem [13]. While the reader can consult [13] for a general proof (and extensions), it is obvious from (4.34) that this should be the case since $T_1 \circ T_2$ can be associated with a finite covariance stochastic process $\mathbf{t}_1 \circ \mathbf{t}_2$ [1].

Before we leave this section, mention should be made of how the colloquial term *cancellation ratio*, or *CR* [6, 12], usually specified in decibels, relates to the preceding discussion. CR is defined as the expected value of the normalized energy residue (in decibels) resulting from subtracting signals

from two different channels [6]. Thus, if the correlation coefficient between two adjacent channels is ρ , then the cancellation ratio (in decibels) is defined as

$$\text{CR (dB)} \triangleq 10 \log_{10} (1 - \rho)^{-1} \quad (4.35)$$

For example, if $\rho = 0$ (no correlation between channels), then the CR is 0 dB (no cancellation). If $\rho = 0.9$, the CR is 10 dB. The CR is simply a measure of the pairwise correlation between channels. The CMT framework of this section shows explicitly how to relate the correlation between channels (and thus the CR between channels) to its impact on interference cancellation. Namely, if the pairwise correlation between channels is uniform and equal to ρ , (which could be experimentally determined), then the corresponding spatial-only CMT is simply given by

$$T = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & & \\ \vdots & & \ddots & \\ \rho & & & 1 \end{bmatrix} \quad (4.36)$$

4.2.2 Angle-Dependent Channel Mismatch

Due to the potential presence of a multitude of array manifold effects (mutual coupling, near-field multipath, element position errors, bandwidth dispersion [3], and so forth), T of (4.33) may be a function of AoA; that is,

$$T = T(\theta) \quad (4.37)$$

which can significantly complicate the analysis of the impact on the interference covariance matrix with no mismatch errors. Specifically, when present, the total (clutter plus jamming plus noise) covariance matrix R is of the form

$$R = \text{cov}(\mathbf{X}_c) + \text{cov}(\mathbf{X}_{J_1}) + \dots + \text{cov}(\mathbf{X}_{J_{N_J}}) + \sigma^2 I \quad (4.38)$$

where

$$\begin{aligned} \text{cov}(\mathbf{X}_{J_i}) &= R_{J_i} \odot T(\theta_{J_i}) \\ \forall i : i &= 1, \dots, N_J \end{aligned} \quad (4.39)$$

where R_{J_1} is the covariance associated with the i th jammer in the absence of channel mismatch, and $T(\theta_{J_1})$ is the total CMT for jammer AoA θ_{J_1} . The expression for $\text{cov}(\mathbf{X}_c)$ is more complicated due to the fact that clutter originates from all angles (in general). Specifically, from Chapter 3 we have

$$\begin{aligned}\text{cov}(\mathbf{X}_c) &= \text{cov}\left(\sum_{i=1}^{N_c} \tilde{\gamma}_i \mathbf{v}_i \odot \mathbf{t}(\theta_i)\right) \\ &= \sum_{i=1}^{N_c} G_i(\mathbf{v}_i \mathbf{v}'_i) \odot T(\theta_i)\end{aligned}\quad (4.40)$$

where we have assumed that the clutter and random RV mismatch error are uncorrelated. Notice that since $T(\theta_i)$ is inside the summation, $\text{cov}(\mathbf{X}_C)$ is not of the previously encountered factored form $R \odot T$. However, in a normally operating radar, the transmit and receive antenna patterns are pointing in the same direction. This implies that G_i will have a significant maximum in the look direction. Since the effect of a CMT is the introduction of sideband energy that is generally several (if not many) orders of magnitude down from the primary (mainbeam) unmodulated eigenvectors (see [1] and the discussion on the PC-CMT technique in Chapter 5), only the dominant eigenvectors will generally play a role in introducing subspace leakage.

The practical consequence of these observations is that the unfactored result of (4.40) can reasonably be approximated by the factored result

$$\text{cov}(\mathbf{X}_c) \approx R_c \odot T(\theta_0) \quad (4.41)$$

where R_c is the covariance in the absence of angle-dependent channel mismatch, and $T(\theta_0)$ is the CMT associated with the radar transmit-receive look direction θ_0 .

Combining (4.41) with (4.38) and (4.39) yields a simplified, yet useful, result for the impact of angle-dependent channel mismatch when both clutter and jamming are present; that is,

$$\begin{aligned}R &= \text{cov}(\mathbf{X}_c) + \text{cov}(\mathbf{X}_{J_1}) + \dots + \text{cov}(\mathbf{X}_{J_{N_J}}) + \sigma^2 I \\ &\approx R_c \odot T(\theta_0) + R_{J_1} \odot T(\theta_{J_1}) + \dots + R_{J_{N_J}} \odot T(\theta_{J_{N_J}}) + \sigma^2 I\end{aligned}\quad (4.42)$$

One potential source of angle-dependent channel-to-channel spatial decorrelation is antenna dispersion due to finite bandwidth [7, 17]. Referring

back to Figure 2.1, we see that the absolute time delay between the k th and n th channels in a ULA is

$$\tau_{|n-k|} = |n - k| \frac{d}{c} \sin \theta \quad (4.43)$$

where c is the speed of light, d is the element separation (consistent units assumed), and θ is the angle off boresight (as defined in Chapter 2). For the finite (nonzero) bandwidth case, the receiver (pulse) will have an autocorrelation function associated with it [18]. For an ideal bandpass, with bandwidth B (hertz), the corresponding autocorrelation function, $r(\tau)$, is given by [16]:

$$r(\tau) = \text{sinc}(B\tau) \quad (4.44)$$

where to be consistent, τ is expressed in seconds. Other correlation functions are of course possible, such as that corresponding to a *linear frequency modulation* (LFM) with a temporal taper for range sidelobe reduction.

The corresponding angle-dependent CMT is thus given by

$$T(\theta) = \begin{bmatrix} 1 & r(\tau_1) & r(\tau_2) & \dots & r(\tau_{N-1}) \\ r(\tau_1) & 1 & r(\tau_1) & & \\ r(\tau_2) & r(\tau_1) & 1 & & \\ \vdots & & & \ddots & r(\tau_1) \\ r(\tau_{N-1}) & & & r(\tau_1) & 1 \end{bmatrix} \quad (4.45)$$

where

$$\tau_m = m \frac{d}{c} \sin \theta \quad (4.46)$$

Although for a side-looking radar the peak of the mainbeam is nominally pointing at $\theta = 0^\circ$, for which $\tau_m = 0$, it is not advisable to ignore the above dispersion since mainbeam clutter is still spread in angle. A reasonable approximate angle-independent CMT for this case [i.e., (4.41)] is $T(\Delta\theta)$, where $\Delta\theta$ is some suitable measure of mainbeam width. Figure 4.4 illustrates an example of the impact of antenna dispersion on the clutter rank.

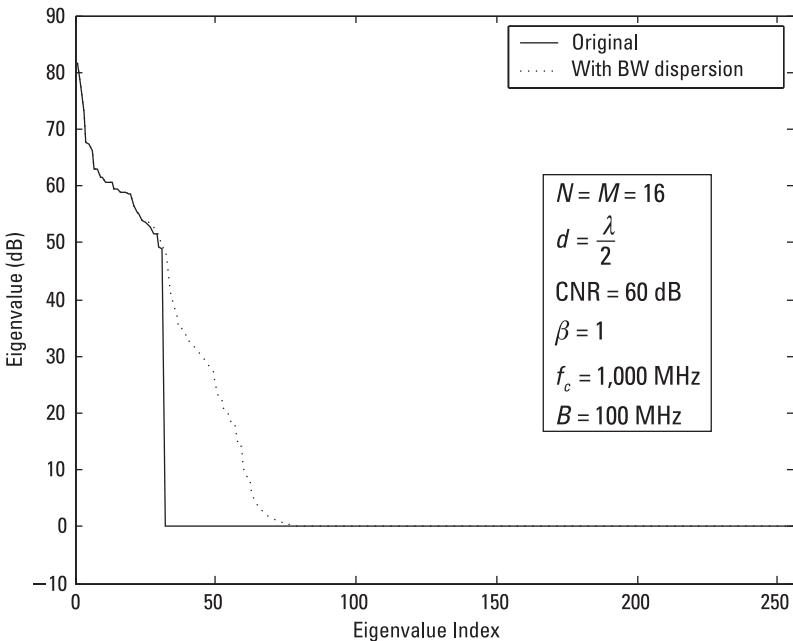


Figure 4.4 Impact of antenna array dispersion on the clutter eigenspectrum for the finite bandwidth case (10% bandwidth example).

4.3 Other Interference Subspace Leakage Effects

ISL [1] refers generally to situations in which an increase in the rank of the colored noise (clutter and/or jamming) is present. As we saw in the previous section, decorrelating channel mismatch can cause an increase in rank and is thus a special case of ISL. There are, in general, many other ISL mechanisms. For example, random amplitude and/or phase modulation of the clutter signal will generally cause an increase in its rank.

One common example is ICM [4, 9, 19] (e.g., wind-blown foliage—an extensive research and measurement study was recently conducted by Billingsley of MIT Lincoln Laboratory [19]). A product of this investigation is a very useful and accurate empirical exponential model that captures the salient characteristics of ICM and is colloquially referred to as the Billingsley model. The only parameters required to specify the clutter Doppler power spectrum are essentially the operating wavelength and wind speed. Specifically,

$$P_c(f) = \frac{r}{r+1} \delta(f) + \frac{1}{r+1} \frac{b\lambda}{4} e^{-\frac{b\lambda}{2}|f|} \quad (4.47)$$

where $P_c(f)$ is the clutter Doppler psd, λ is the operating wavelength, f is the Doppler frequency (hertz), and $\delta(\cdot)$ is the Dirac delta function. The shape parameter b depends primarily on wind conditions and has been tabulated [19]. The ratio r , $0 \leq r \leq \infty$, between the dc [first term in (4.47)] and ac components (second term) was found to be a function of both the carrier frequency and wind speed [19], namely,

$$10 \log r = -15.5 \log w - 12.1 \log f_c + 63.2 \quad (4.48)$$

where w is the wind speed in *miles per hour* (mph), f_c is the carrier frequency in megahertz, and $\log(\cdot)$ is the base 10 logarithm.

The corresponding correlation (temporal) function is given by the inverse Fourier transform of (4.48) [4], that is,

$$\begin{aligned} r_c(\tau) &= \int_{-\infty}^{\infty} P_c(f) e^{j2\pi f\tau} \\ &= \frac{r}{r+1} + \frac{1}{r+1} \frac{(b\lambda)^2}{(b\lambda)^2 + (4\pi\tau)^2} \end{aligned} \quad (4.49)$$

The CMT corresponding to this ICM model is easily obtained by sampling $r_c(\tau)$ at multiples of the PRI; that is, $r_c(0)$ (first pulse), $r_c(\text{PRI})$ (correlation between first and second pulses), $\dots, r_c((M-1)\text{PRI})$ (correlation between first and last (M th) pulses [4]. The full-up space-time CMT is thus given by

$$T = T_{\text{ICM}} \otimes \mathbf{1}_{N \times N} \quad (4.50)$$

where

$$[T_{\text{ICM}}]_{i,j} = r_c(|i-j|\text{PRI}) \quad (4.51)$$

Note that ICM is a temporal modulation effect, as evidenced by (4.50), which is a special case of (4.28).

Figure 4.5 shows an example of the impact of ICM on both the eigenspectrum and SINR for a modest average wind speed of 10 mph. Note that similar to the decorrelating channel-mismatch case, the most significant impact is an increase in the mainbeam clutter notch—with a commensurate decrease in MDV.

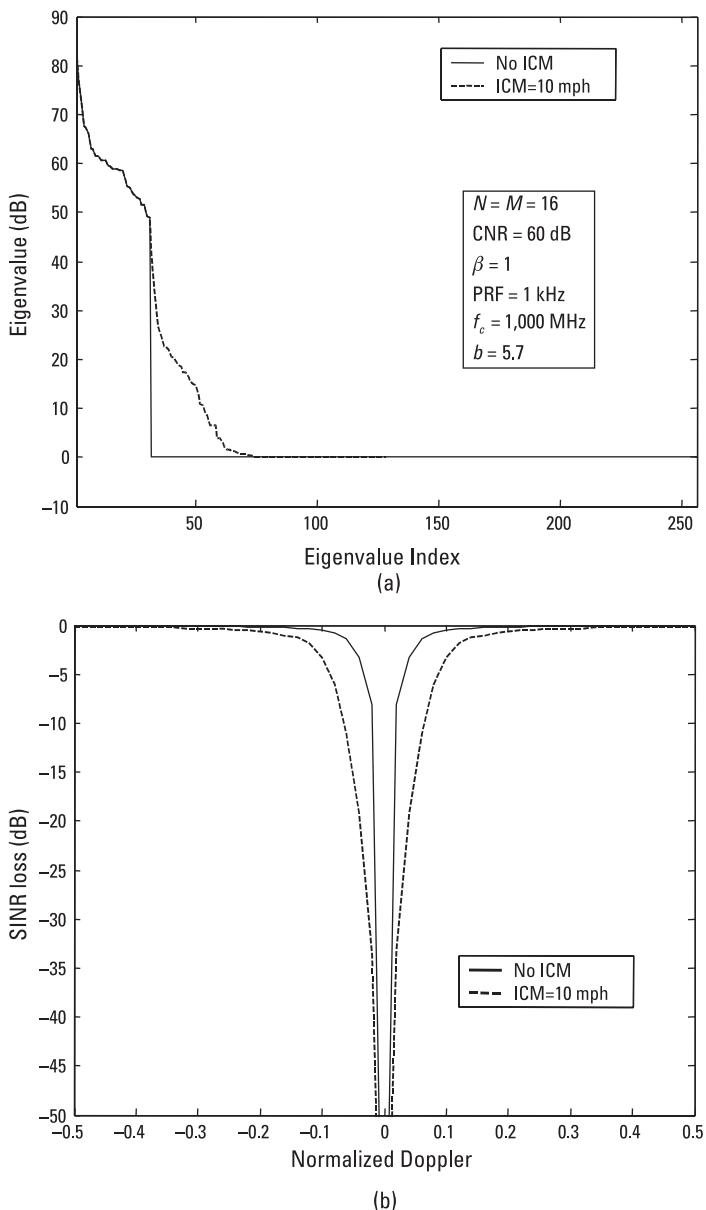


Figure 4.5 Impact of ICM on (a) eigenspectrum and (b) SINR loss. Note that even a modest average wind speed can have an impact on performance in high CNR environments.

Another potential source of ISL is so-called range walk, which can occur due to range-bin migration of the clutter during a CPI (see [20] and references cited therein). Figure 4.6 illustrates this effect for the general case when the look direction is not broadside. Since, to a reasonable approximation, the amount of decorrelation is proportional to the amount of area overlap ΔA (see Figure 4.6), we see that this effect is in general angle-dependent—with a minimum at broadside and a maximum in the forward-looking direction. However, for the same reason cited above, a useful approximation can be obtained by considering the amount of decorrelation in the look direction thereby resulting in an approximate, but angle-independent, CMT. Assuming a relatively short CPI with a constant velocity, we see from Figure 4.6 that the correlation ρ between successive pulses is approximately given by

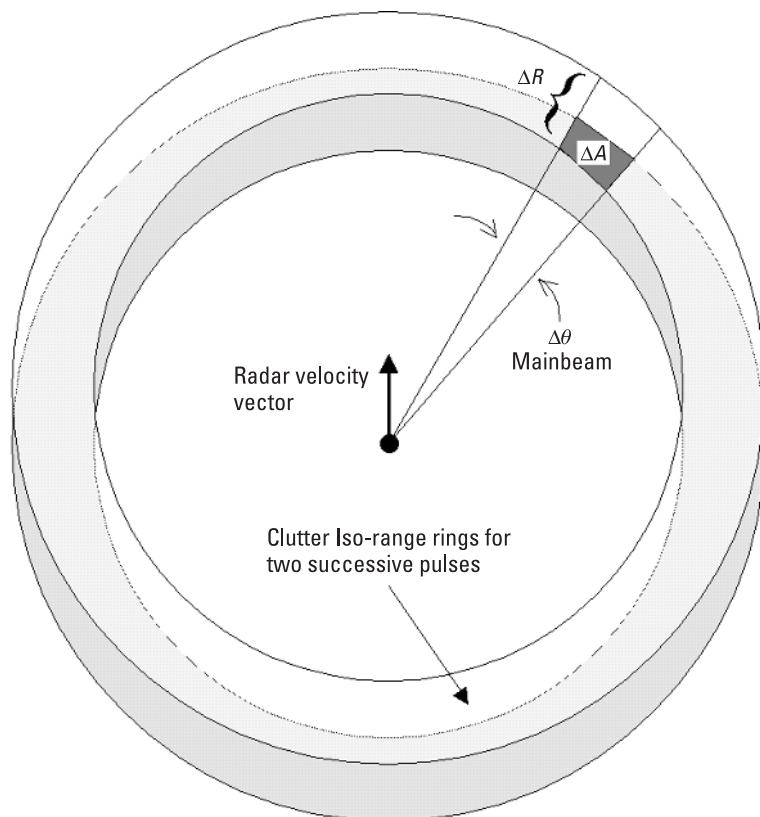


Figure 4.6 Illustration of the cause of clutter decorrelation due to range-bin migration.

$$\begin{aligned}
\rho &\approx \frac{\Delta A}{A} \\
&\approx \frac{\Delta A}{\Delta R \cdot \Delta \theta} \\
&= \frac{\Delta A}{\left(\frac{c}{B}\right) \cdot \Delta \theta}
\end{aligned} \tag{4.52}$$

where $\Delta\theta$ is the mainbeam width, and ΔR is the range-bin size—which is taken to be c/B , where c is the speed of light and B is the bandwidth of the compressed pulse (consistent units assumed).

For a short CPI, it is reasonable to assume that the change in area overlap is approximately constant between successive pulses. Thus, if the decorrelation between successive pulses is ρ , the temporal decorrelation between the m th and n th pulses, denoted by $\rho_{m,n}$, is then simply given by

$$\rho_{m,n} = \rho^{|m-n|} \tag{4.53}$$

The corresponding temporal CMT is therefore of the form

$$T = \begin{bmatrix} 1 & \rho & \dots & \rho^{M-1} \\ \rho & 1 & & \\ \vdots & & \ddots & \\ \rho^{M-1} & & & 1 \end{bmatrix} \tag{4.54}$$

Note that this has a similar (not identical) structure to the ICM case inasmuch as there is a monotonic decrease in correlation with increasing PRI separation. Thus, the resulting clutter eigenspectrum will have a similar appearance to that of the ICM case.

4.4 Antenna Array Misalignment

Brennan's rule, as derived in Chapter 3, assumed that the ULA antenna was aligned with the ground velocity vector. In fact, this was crucial for showing that the clutter rank was generally significantly less than the total spatiotemporal DoFs [i.e., the redundancy argument demonstrated in (3.18)]. Moreover,

when aligned, the back lobe clutter exactly coincides with the front lobe clutter as a function of angle-Doppler as evidenced by an examination of Figure 3.1 and (3.1) [8]. In practice, it is not possible to maintain antenna alignment to ground track (e.g., aircraft crabbing and implementation constraints). Consequently, the clutter rank (and angle-Doppler spectrum) is altered.

The impact of velocity misalignment on the angle-Doppler relationship of ground clutter (at long ranges) is the introduction of a so-called crab angle θ_c in (3.1) as follows [8]:

$$\bar{f}_d = \frac{2vT}{\lambda} \sin(\theta + \theta_c) \quad (4.55)$$

Note that the symmetry between back lobe and front lobe has generally been lost. Specifically, with $\theta_c = 0$ (i.e., no crab) there is even Doppler frequency symmetry about the nose of the aircraft; that is, $\sin\left(\frac{\pi}{2} + \alpha\right) = \sin\left(\frac{\pi}{2} - \alpha\right)$, for $|\alpha| \leq \pi$, where α is the angle of a given clutter patch with respect to the nose of the aircraft. Front and back lobe clutter will thus lie along the same angle-Doppler contour. With $\theta_c \neq 0$, $\sin\left(\frac{\pi}{2} + \alpha\right) \neq \sin\left(\frac{\pi}{2} - \alpha\right)$, for $|\alpha| \leq \pi$. This will result in a so-called back lobe ridge, which may be of concern in strong clutter with poorly attenuated back lobes [8].

Another observation gleaned from (4.55) is that there is no longer a strictly linear relationship between normalized angle and Doppler [i.e., (3.2)] no longer holds. Instead, the clutter locus forms an ellipse [8]. To see this explicitly, we manipulate (4.55) as follows:

$$\begin{aligned} \bar{f}_d &= \frac{2vT}{\lambda} \sin(\theta + \theta_c) \\ &= \frac{2vT}{\lambda} [\sin(\theta) \cos(\theta_c) + \cos(\theta) \sin(\theta_c)] \\ &= \beta\eta [\sin(\theta) \cos(\theta_c) + \cos(\theta) \sin(\theta_c)] \end{aligned} \quad (4.56)$$

where

$$\beta = \frac{2vT}{d}, \quad \eta = \frac{d}{\lambda} \quad (4.57)$$

Rearranging (4.56) and employing a trigonometric identity, we have

$$\begin{aligned} \bar{f}_d - \beta \bar{\theta}_d \cos \theta_c &= \beta \eta \cos \theta \sin \theta_c \\ &= \pm \beta \eta \sqrt{1 - \sin^2 \theta} \sin \theta_c \end{aligned} \quad (4.58)$$

where the expression for normalized angle $\bar{\theta}_d = \eta \sin \theta$ is inserted and the trigonometric identity $\cos \theta = \pm \sqrt{1 - \sin^2 \theta}$ is employed. Squaring both sides of (4.58) and rearranging terms yields

$$\bar{f}_d^2 + \beta^2 \bar{\theta}_d^2 - 2\beta \cos \theta_c \bar{f}_d \bar{\theta}_d - \beta^2 \eta^2 \sin^2 \theta_c = 0 \quad (4.59)$$

Equation (4.59) yields a rotated ellipse in the normalized angle-Doppler coordinates. The standard form for a rotated (and generally translated) ellipse is given by [21]

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (4.60)$$

Comparing this with (4.59) yields the following assignments

$$\begin{aligned} x &= \bar{\theta}_d \\ y &= \bar{f}_d \\ A &= \beta^2 \\ B &= -2\beta \cos \theta_c \\ C &= 1 \\ D &= E = 0 \\ F &= -\beta^2 \eta^2 \sin^2 \theta_c \end{aligned} \quad (4.61)$$

from which it is evident that the ellipse is centered at the origin, but is rotated from the $\bar{\theta}_d$ axis by an angle ϕ given by [21]

$$\begin{aligned} \cot 2\phi &= \frac{A - C}{B} \\ &= \frac{1 - \beta^2}{2\beta \cos \theta_c} \\ B &\neq 0 \end{aligned} \quad (4.62)$$

Note that for $\beta^2 = 1$, $\cot 2\phi = 0$, which implies that $\phi = \frac{\pi}{4}$, that is, an ellipse oriented at 45° .

Figure 4.7 displays angle-Doppler contours for different values of crab angle, along with corresponding clutter-plus-noise eigenspectrum and Capon

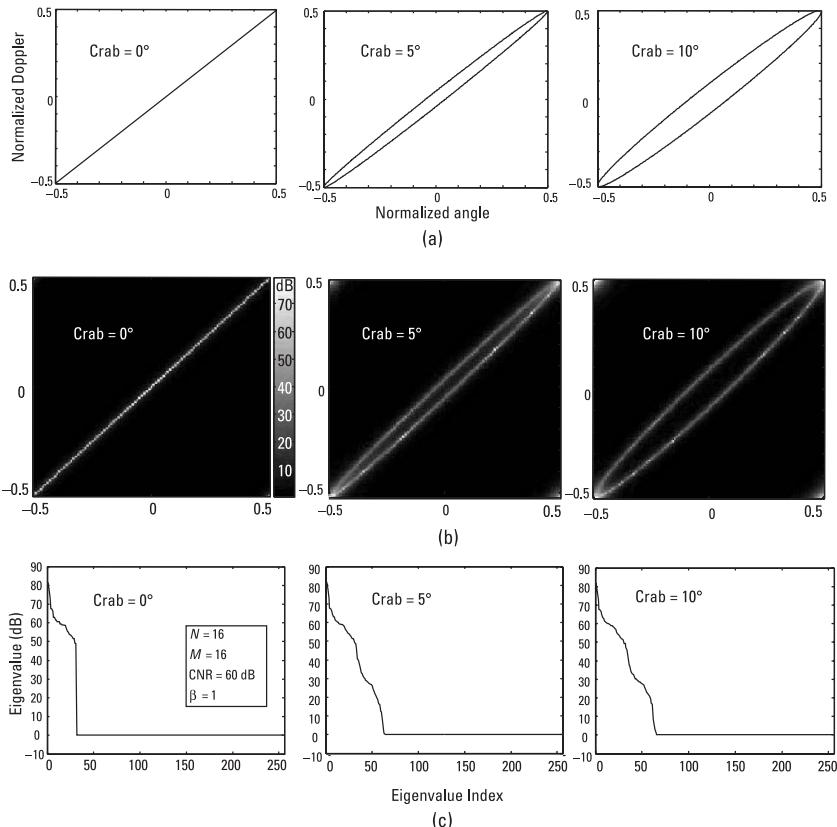


Figure 4.7 Effect of crabbing and finite back lobe rejection ratio on (a) angle-Doppler clutter loci, (b) power spectra, and (c) eigenspectra, respectively. Back lobe rejection ratio is fixed to 30 dB.

power-spectral densities. Note the appearance of a second clutter ridge whose strength is determined by the back lobe rejection level of the array.

4.5 Nonlinear Arrays¹

Brennan's rule, as derived in Chapter 3, assumed that the side-looking ULA antenna was aligned with the ground-referenced velocity vector. Indeed, it was because of this assumption (and several others) that a generally lower rank for the clutter eigenspectrum was realized. When this assumption is violated—as is the case in Section 4.4 due to crabbing—the rank of the clutter is generally increased. Unfortunately, this rank inflation also arises when nonlinear array geometries are employed (even if the geometrical axis of the array is aligned with the ground-referenced velocity vector) [8]. As with the crabbing case, the violation of Brennan's rule is easily established since the requisite linear phase relationship and redundancy in (3.16) no longer hold.

There are obviously too many possible nonlinear array geometries to contemplate them all in this venue. However, an increasingly popular configuration is the so-called circular array [22, 23]. A static (nonrotating), electronically scanned, circular array can eliminate the need for bulky, heavy, and energy-consuming mechanically scanned servos, as well as provide enhanced *track-while-scan* (TWS) capabilities [22].

Figure 4.8 illustrates an example UHF linear and circular array configuration. To insure the formation of proper transmit and receive patterns, only a subsection of the total circular array is active [23]. Note that to make a fair comparison with the linear case, a few extra elements are included in the circular subsection to maintain an approximately equivalent aperture size.

Figure 4.9 shows the impact of circular-array geometry on the clutter-plus-noise eigenspectrum or the case of 30-dB and 100-dB back lobe rejection. Note that as with the crabbing array, low back lobe rejection can have a major impact on clutter spread.

4.6 Interference Nonstationarity and the Iceberg Effect

Due to both man-made and natural variations in terrain, real-world clutter is heterogeneous [24]. Thus, the statistics (e.g., pdf, CNR, and correlation

1. Simulation material in this section was provided by Mr. Jamie Bergin, Information Systems Laboratories, Inc.

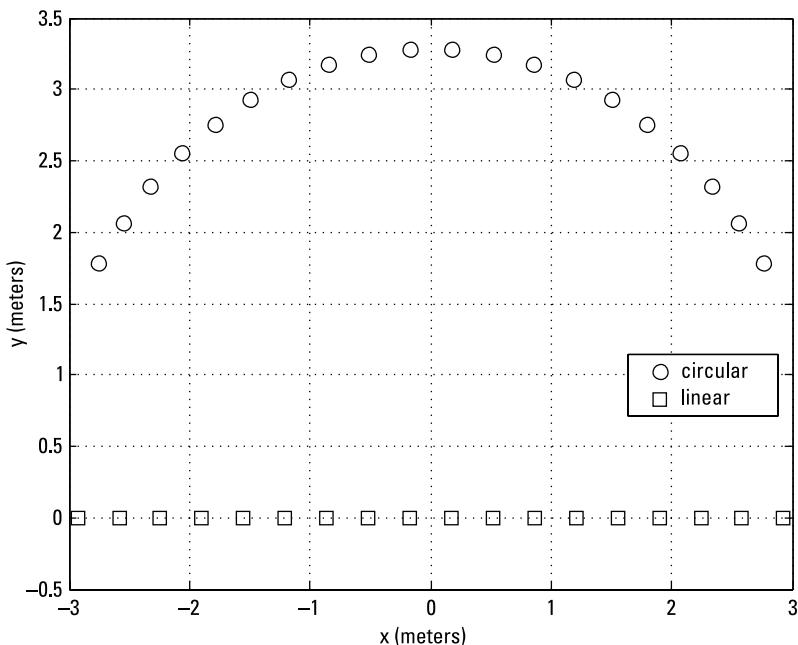


Figure 4.8 UHF linear and circular arrays illustrating the impact of nonlinear array geometry. (Data provided courtesy of Information Systems Laboratories, Vienna, Virginia.)

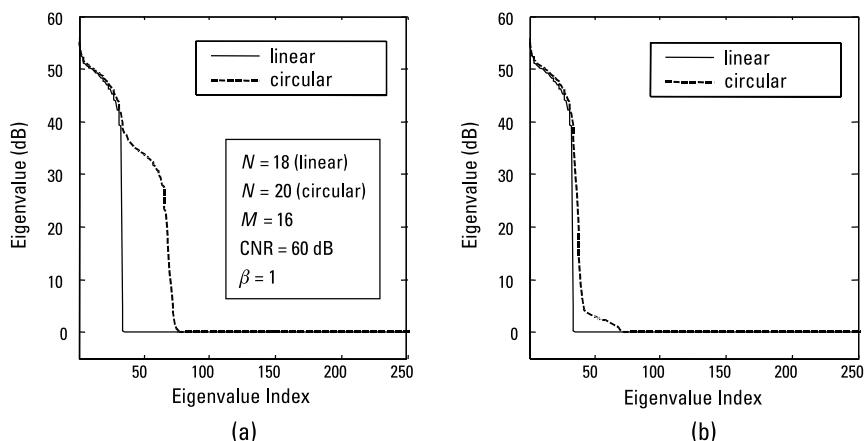


Figure 4.9 Impact of nonlinear-array geometry on clutter eigenspectrum for back lobe rejection ratios of (a) 30 dB and (b) 100 dB. First 250 eigenvalues displayed. (Data provided courtesy of Information Systems Laboratories, Vienna, Virginia.)

properties) will vary with location. This fact obviously presents a severe challenge to any adaptive clutter-cancellation scheme that must estimate requisite statistical information on the fly.

To begin to appreciate the impact nonstationarity can have on STAP performance, we first consider perhaps the simplest form of heterogeneity, namely CNR variation. Figure 4.10 shows the impact of a change in CNR for the ideal ULA case (for which Brennan's rule holds) and for the more realistic case in which subspace leakage is present (10-mph ICM). Note that for the low CNR case, the effective ranks of the clutter eigenspectra are approximately equal. However, for high CNR, there is a significant increase in the rank when subspace leakage is present. If one thinks of the white-noise floor (0 dB on graph) as the surface of a perfectly calm “ocean,” and the clutter-only eigenspectrum as an “iceberg,” one can see that the increase in effective clutter rank with increasing CNR is analogous to an iceberg rising above the ocean surface. Although ICM was used to provide the additional subspace leakage, any of the decorrelating effects described in this chapter will produce the same qualitative result.

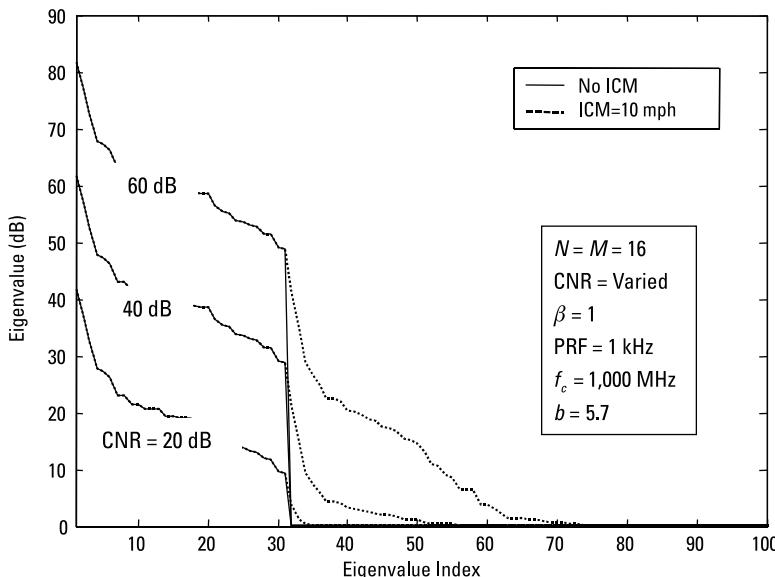


Figure 4.10 Illustration of the iceberg effect, which describes the increase in interference rank (eigenvalues above noise floor) with increasing CNR when realistic eigenspectra are considered (i.e., Brennan's rule no longer holds). Note that for low CNR the effect is negligible, while for large CNR it can be quite significant.

This iceberg effect can have a significant impact on sample covariance-(adaptive-) based beamforming. As a simple illustration, consider an ideal side-looking ULA (no crab) where the CNR decreases from 70 to 20 dB over a range support of 6,000 samples (bins), where $N = M = 16$, and there is 10-mph ICM. This power variation can simply be the result of decreasing CNR with increasing range from the radar [24], or it can be due to gradual variations in terrain elevation or coverage [24].

Figure 4.11 shows SINR versus Doppler and range for both the ideal optimum (known-covariance) and estimated covariance ($2NM$ SMI centered at range bin 3,000) cases. The additional loss in the near-in ranges is due to clutter undernulling (i.e., the notch width and depth are inadequate), while the loss in the far range bins is due to clutter overnulling (i.e., notch width and depth are too large—resulting in unnecessary target signal loss). In both cases, the biggest impact is on the *minimum detectable velocity* (MDV), which is defined as the velocity at which SINR loss is 10 dB (see [11] for further discussion of MDV and alternate definitions). For GMTI radar, this severe loss in MDV (mainbeam clutter notch width) may be undesirable [25].

4.7 Summary

In this chapter, we touched upon a few of the many real-world effects that can have a significant impact on STAP performance, as compared with so-called ideal performance. Although one cannot reduce the effective rank of the colored-interference subspace when ISL and other rank-inflation mechanisms are present, judicious algorithm design can minimize its impact. This is a primary consideration in the next chapter.

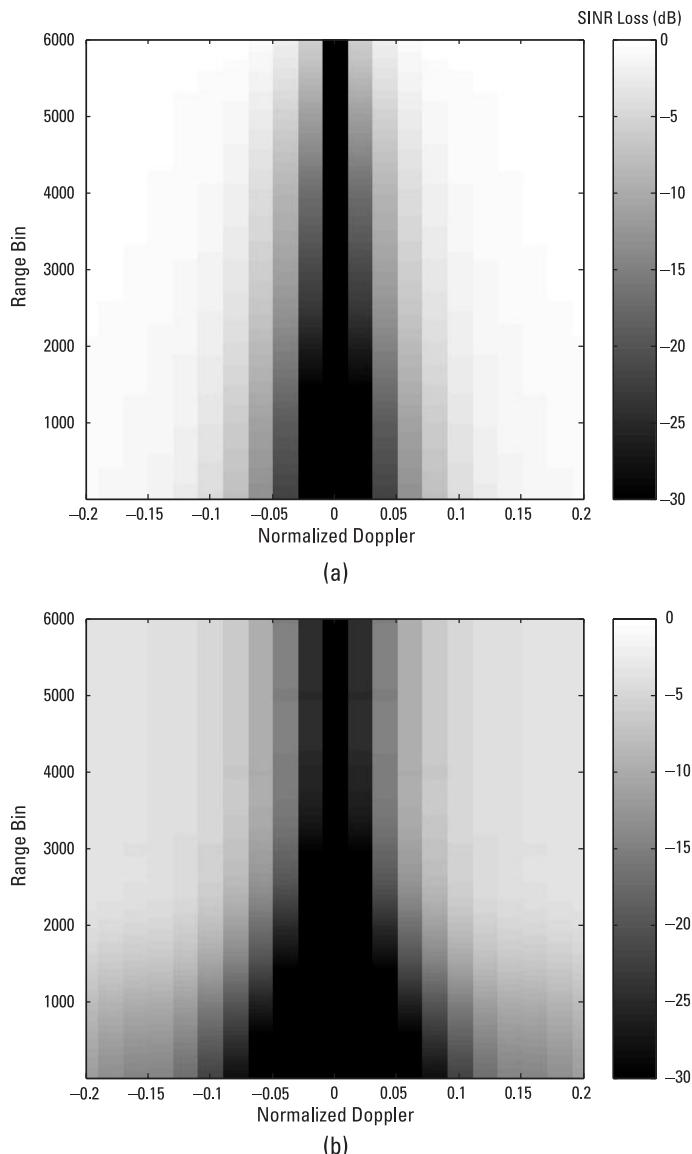


Figure 4.11 Impact of clutter nonstationarity on SINR: (a) ideal SINR performance achieved when exact covariance is used for each range bin, and (b) actual SINR performance when a $2NM$ sample (centered at range bin 3,000) covariance is used. The additional SINR loss at near-in range bins is due to clutter undernulling, while the additional loss at far ranges is due to overnulling. This example illustrates the importance of tracking clutter statistics for optimal performance.

References

- [1] Guerci, J. R., and J. S. Bergin, "Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002.
- [2] Guerci, J. R., "Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming," *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–986.
- [3] Zatman, M., "Production of Adaptive Array Troughs by Dispersion Synthesis," *Electronics Letters*, Vol. 31, No. 25, 1995, pp. 2141–2142.
- [4] Techau, P. M., J. S. Bergin, and J. R. Guerci, "Effects of Internal Clutter Motion on STAP in a Heterogeneous Environment," *Proc. of 2001 IEEE Radar Conference*, May 2001.
- [5] Zatman, M., "How Narrow Is Narrowband?" *IEE Proc. of Radar, Sonar and Navigation*, Vol. 145, April 1998, pp. 85–91.
- [6] Monzingo, R. A., and T. W. Miller, *Introduction to Adaptive Arrays*, New York: John Wiley & Sons, 1980, pp. 56–64.
- [7] Mailloux, R. J., *Phased Array Antenna Handbook*, Norwood, MA: Artech House, 1994.
- [8] Ward, J., *Space-Time Adaptive Processing for Airborne Radar*, MIT Technical Report 1015, MIT Lincoln Laboratory, December 1994.
- [9] Billingsley, J. B., *Radar Clutter*, Norwood, MA: Artech House, 2001.
- [10] Melvin, W. L., "Space-Time Adaptive Radar Performance in Heterogeneous Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 36, No. 2, pp. 621–633.
- [11] Melvin, W. L., M. J. Callahan, and M. C. Wicks, "Bistatic STAP: Application to Airborne Radar," *Proc. of 2002 IEEE Radar Conference*, Long Beach, CA, April 22–25, 2002, pp. 1–7.
- [12] Compton, R. T., Jr., *Adaptive Antennas: Concepts and Performance*, Englewood Cliffs, NJ: Prentice Hall, 1988.
- [13] Horn, R. A., and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, England: Cambridge University Press, 1991.
- [14] Hudson, J. E., *Adaptive Array Principles*, London, England: Peter Peregrinus Ltd. on behalf of the IEE: (reprint) 1991.
- [15] Farina, A., *Antenna-Based Signal Processing Techniques for Radar Systems*, Norwood, MA: Artech House 1992.
- [16] Papoulis, A., and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 2001.
- [17] Zatman, M., "Production of Adaptive Array Troughs Through Dispersion Synthesis," *Electronics Letters*, Vol. 31, No. 25, December 1995, pp. 2141.
- [18] Techau, P. M., "Effects of Receiver Filtering on Hot Clutter Mitigation," *Proc. of the 1999 IEEE Radar Conference*, Waltham, MA, April 20–22, 1999, pp. 84–89.
- [19] Billingsley, J. B., *Exponential Decay in Windblown Radar Ground Clutter Doppler Spectra: Multifrequency Measurements and Model*, Technical Report 997, MIT Lincoln Laboratory, Lexington, MA, July 29, 1996.

- [20] Klemm, R., *Principles of Space-Time Adaptive Processing*, London, England: IEEE Press, 2002.
- [21] Swokoski, E. W., *Calculus with Analytic Geometry*, Boston, MA: Prindle, Weber and Schmidt, 1975.
- [22] Zatman, M., "Circular Array STAP," *Proc. of IEEE National Radar Conference*, Waltham, MA, 1999.
- [23] Guerci, J. R., et al., "Optimal Reduced-Rank STAP for Circular Adaptive Arrays," *Proc. of Space-Time Processing Methods for Circular Ring Arrays with Application to Navy Airborne Surveillance Radar*, George Mason University, Fairfax, VA, February 22–23, 1999.
- [24] Long, M. W., *Radar Reflectivity of Land and Sea*, 3rd ed., Norwood, MA: Artech House, 2001.
- [25] Guerci, J. R., "Knowledge-Aided Sensor Signal Processing and Expert Reasoning," *Proc. of the 2002 Workshop on Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER)*, Washington, D.C., April 3, 2002 (CD-ROM).

5

STAP for Radar: Methods, Algorithms, and Performance

5.1 Introduction

Although in theory we have already specified the theoretically optimum strategy for space-time filtering of interference—that is, the ubiquitous Wiener-Hopf equation $\mathbf{w} = R^{-1}\mathbf{s}$ —many practical considerations (described in previous chapters and elsewhere [1–4]) preclude its direct implementation. Most notable is the fact that the interference covariance matrix R is unknown *a priori* and must be estimated on the fly, then utilized in an appropriate real-time space-time filtering architecture. Even in so-called stationary Gaussian environments (which, in fact, never truly exist) when a sample covariance estimate is indicated (see Section 2.4), care must be taken to avoid erratic sidelobe levels introduced due to estimation errors and potentially limited sample support. Even greater care is required when confronted with nonstationary environments—particularly when ISL and other rank-inflating phenomena (crabbing, nonlinear array geometries, and so forth) are present.

In this chapter, we introduce a broad range of STAP algorithms that have been developed (many in recent years) to combat one or more of the aforementioned performance degrading effects. Although it is not practical to consider every current STAP algorithm that has been proposed (a keyword search on “STAP” should suffice to make the point), a representative sampling of algorithms drawn from a STAP taxonomy are analyzed in this chapter. It should be mentioned that our emphasis here is on performance and not

specifically on implementation complexity. Certain key aspects of real-time STAP will be explored in Chapter 6.

Figure 5.1 illustrates a STAP algorithm taxonomy. All of these algorithms either explicitly or implicitly attempt to reduce the dimensionality of the unknown interference parameters via a rank-reducing transformation or an explicit model-based method (or potentially both), while still preserving acceptable performance [1]. Thus, although the STAP CPI dimension NM might be quite large, the dimension of the unknown parameters vector or matrix could be substantially less—since clutter is generally restricted in angle-Doppler space [1, 2, 5]. This reduced-dimension adaptivity is only

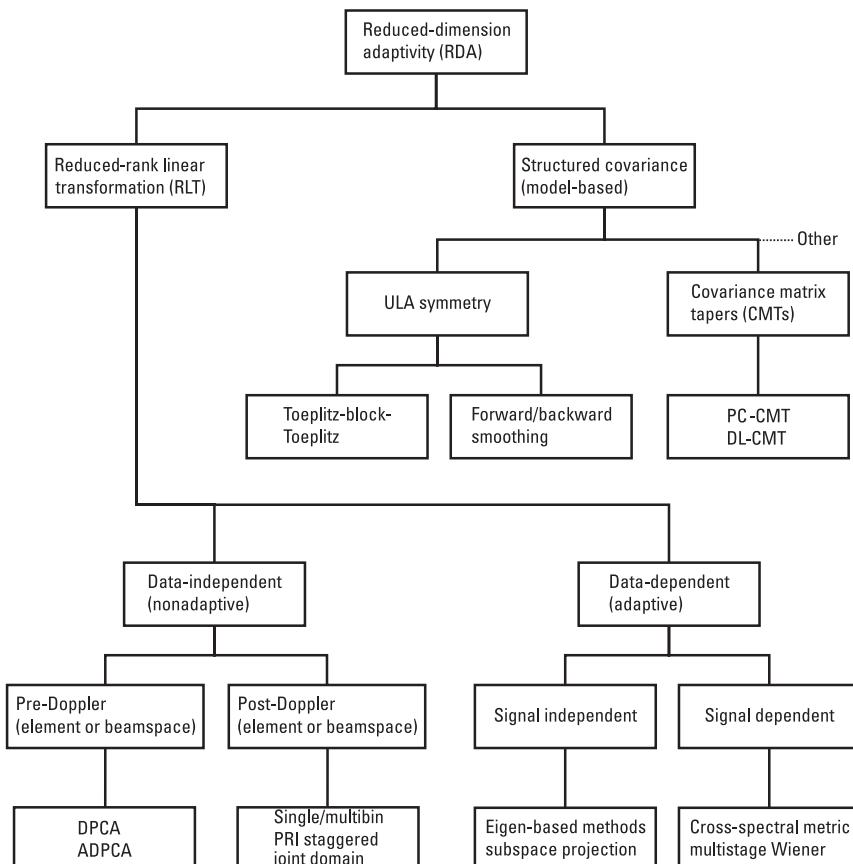


Figure 5.1 Taxonomy of reduced-dimension adaptivity STAP algorithms. It is important to note that many of the methods can be combined, thus greatly increasing the palette of techniques from which to choose.

possible if additional prior knowledge of the interference is assumed. Indeed, it is critical to know what prior knowledge is explicitly or implicitly assumed in order to select an algorithm appropriate to a given radar scenario.

For example, in post-Doppler methods [for an example of a reduced-rank *linear transformation* (RLT) data-independent method, see Figure 5.1], an assumption on the Doppler extent of ground clutter is exploited to subband the rejection problem [1, 2, 5]. Thus, a different, but substantially reduced-rank (or reduced-dimension) adaptive problem is solved for each Doppler bin. Since a much smaller covariance matrix is utilized in these methods, it can provide significant advantages in nonstationary environments and also be substantially easier to implement. Under certain conditions, which may or may not be sufficiently satisfied in any given application, post-Doppler methods can closely approximate the performance of optimum full-DoF STAP [1, 2].

An example of an explicit model-based algorithm that can reduce the effective adaptive DoFs is Toeplitz matrix fitting [6, 7]. For an ideal sidelooking ULA (no crab or channel mismatch) and uniform PRF in the presence of uniform uncorrelated clutter scatterers (such as considered in Chapter 3), the ideal STAP covariance matrix has a *Toeplitz-block-Toeplitz* (TBT) structure [1, 2, 6, 7]. The redundancy in the covariance structure effectively reduces the dimension of the unknown parameter vector (unknown elements of covariance matrix) which, in turn, reduces requirements on sample support for estimation—a highly desirable property in nonstationary environments or where sample support is limited by other factors (e.g., radar horizon). However, the Toeplitz assumption is an approximation at best and thus must be applied judiciously.

The remainder of Chapter 5 will be spent exploring the STAP taxonomy tree of Figure 5.1 and comparing the different methods in a variety of simulated, but nonetheless realistic scenarios. Though this treatise is not exhaustive, the reader should gain a foundation from which to craft a custom STAP solution well suited to his or her application.

A unique pedagogical feature of this chapter is the explicit mathematical derivation of an *effective covariance inverse* (ECI), denoted by R_{eff}^I , for each STAP algorithm. That is, each reduced-dimension STAP algorithm can be shown to result in a mathematical form

$$\mathbf{w} = R_{\text{eff}}^I \mathbf{s} \quad (5.1)$$

where R_{eff}^I is always an $NM \times NM$ matrix. Adopting this approach not only significantly aids in the mathematical analysis of various methods, but also

provides insight into precisely how a particular algorithm works—an invaluable aid when attempting to choose among the many different methods. It is important to note that the ECI may be singular and thus may not have an inverse. Obviously, for the full-DoF known-covariance case, $R_{\text{eff}}^I = R$.

5.2 Data-Independent Reduced-Rank STAP

The first branch point in the STAP taxonomy of Figure 5.1 differentiates between RLT and structured-covariance (model-based) methods. An RLT is either data-independent (nonadaptive) or data-dependent (adaptive) and can be thought of as a nonsquare matrix transformation (with more columns than rows) applied to the concatenated space-time snapshot data vectors (column vectors).

The most extreme example of a data-independent RLT is the so-called factored approach that consists of nonadaptive Doppler filtering and beamforming. Since there is no joint space-time (or joint angle-Doppler) processing, this technique is also referred to as 1-D factored space-time processing [1, 2]. This method is extremely suboptimal for airborne radar in the presence of ground clutter since it makes no provision for the Doppler dependency on angle as described in Chapter 3. Succinctly stated, this method is mathematically tantamount to simply $\mathbf{w} = \mathbf{s}$, which clearly shows the complete lack of interference structure (i.e., interference covariance). Figure 5.2 shows the extreme suboptimality of this approach and serves as a reminder of why STAP for clutter cancellation was invented in the first place.

5.2.1 Pre-Doppler (Signal-Independent) Reduced-Rank STAP: DPCA and Adaptive DPCA

The first practical step, both historically and pedagogically, towards joint space-time processing is the DPCA method [1]. Taxonomically, it is a pre-Doppler data-independent RLT. The DPCA method attempts to operate a ULA in such a manner that (1) the elements (or subarrays) of the ULA are extremely well matched (just how well will be discussed momentarily); (2) the axis of the ULA is aligned with the ground-track velocity vector (not necessarily the aircraft orientation due to crabbing); (3) the PRI is precisely chosen to insure that the array advances one interelement spacing per pulse; and (4) the trailing and leading elements are blanked on receive, as shown in Figure 5.3, which makes the indicated subarray appear stationary relative to the ground. If all four of these steps are performed to requisite tolerances,

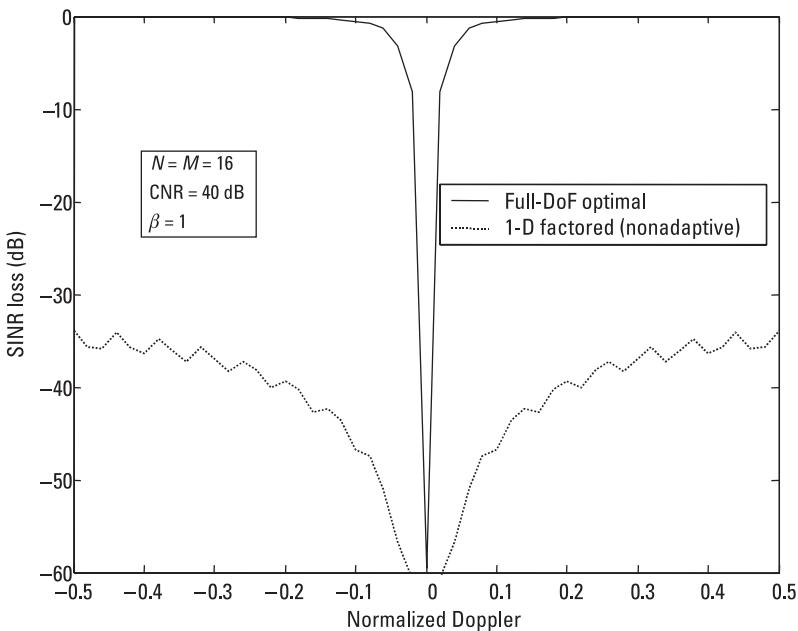


Figure 5.2 Comparison of full-DoF optimal space-time processing with a nonadaptive 1-D factored approach.

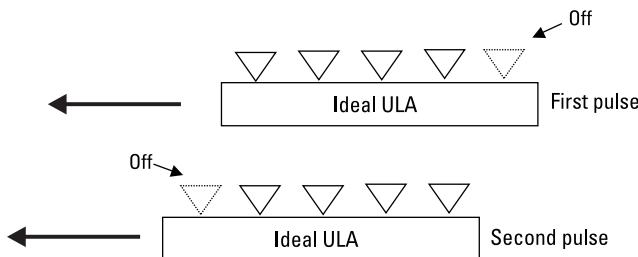


Figure 5.3 Illustration of the DPCA concept for an ideal ULA perfectly aligned with the ground-track velocity vector. The PRI is chosen so that the array advances exactly one interelement spacing per pulse (i.e., $\beta = 1$). By using only the subarrays indicated above, the array appears to be stationary relative to the ground, vastly simplifying the clutter-cancellation process.

the ground clutter will appear stationary (without ICM, and so forth), thus vastly simplifying the clutter-cancellation process—at the cost of aperture efficiency.

More precisely, consider the N -dimensional DPCA ULA “snapshot” returns from two successive pulses, \mathbf{x}_1 and \mathbf{x}_2 , with a corresponding concate-

nated two-pulse space-time data vector $\mathbf{x}' = [\mathbf{x}_1' \ \mathbf{x}_2']$. Since the clutter will appear to be stationary (i.e., no Doppler shift), simply subtracting the returns from the two pulses should suffice in cancelling ground clutter [1, 2]. If we assume that we are interested in returns emanating from boresight (0° , that is, side-looking), then mathematically the entire two-pulse MTI DPCA canceller is equivalent to

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 \\ -\mathbf{s}_2 \end{bmatrix} \quad (5.2)$$

where $\mathbf{s}_1, \mathbf{s}_2$ are the N -dimensional spatial-only steering vectors for the first and second pulses, respectively. Thus, to achieve the DPCA effect we must have

$$\mathbf{s}_1 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \text{(1st element off)}, \quad \mathbf{s}_2 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} \text{(Nth element off)} \quad (5.3)$$

Equation (5.2) can also be cast into the ECI framework as follows

$$\begin{aligned} \mathbf{w} &= R_{\text{eff}}^I \mathbf{s} \\ &= (I - e_1 e_1' - e_{2N} e_{2N}') \mathbf{s} \\ &= (I - e_1 e_1' - e_{2N} e_{2N}') \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right) \\ &= \left(I - \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & 0 \end{bmatrix} - \begin{bmatrix} 0 & & \dots & 0 \\ & \ddots & & \vdots \\ & & 0 & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \right) \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right) \end{aligned} \quad (5.4)$$

where I is the $2N \times 2N$ identity matrix, and e_1, e_{2N} are the first and $2N$ th Euclidean basis vectors (e_i is a vector of zeros except for a unity entry in the i th position).

Equation (5.4) is worthy of further examination. It is left as an exercise for the reader to prove equivalence between (5.4) and (5.2) (i.e., that the weight vectors are identical, which can be accomplished by comparing both

vectors on an element basis). The implied space-time steering vector for the two-pulse DPCA system is

$$\mathbf{s} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (5.5)$$

This corresponds to a sidelooking (boresight-aligned) spatial steering vector [second vector in (5.5)] and a Doppler steering vector of $[1 \ -1]'$ —which in turn corresponds to a simple two-pulse MTI canceller [8]. Since there are only two pulses being processed for clutter cancellation and since clutter in a perfect DPCA system occupies the Doppler space corresponding to dc (0 Hz), the only Doppler space left for signal detection is that which is orthogonal to $[1 \ 1]',$ namely $[1 \ -1]'$. Further inspection of (5.4) reveals the effective matrix inverse for two-pulse DPCA

$$R_{\text{eff}}^I = (I - e_1 e_1' - e_{2N} e_{2N}') \quad (5.6)$$

which is recognized to be a projection operator [9] that eliminates the 2-D subspace spanned by the Euclidean basis vectors $e_1, e_{2N}.$ This is the operator responsible for alternately shutting off the leading and trailing elements on the first and second pulse, respectively.

Comparing (5.6) with the eigen-based form for the inverse of the covariance [see (2.59)] implies that for an ideal two-pulse DPCA system, the clutter rank after the DPCA operation is only two—regardless of the number of antenna elements. If the antenna were not operated in this DPCA fashion, Brennan's rule would set the clutter rank at $N + \beta(M - 1) = N + 1,$ since $\beta = 1$ and $M = 2.$ This apparent violation of Brennan's rule can be explained by the fact that it is only applicable to a *linear time-invariant* (LTI) space-time ULA. A DPCA antenna is time-varying (since the array manifold is different on each pulse) and is thus not covered by Brennan's rule.

Before we analyze the performance of DPCA, we need to extend it to the processing of an entire STAP CPI, that is, more than just two pulses. This is easily accomplished by simply Doppler filtering the scalar (but complex) outputs from the two-pulse DPCA MTI filter as shown in Figure 5.4. The scalar outputs from the MTI canceller, $y_i,$ that are the inputs to the Doppler filter bank are related to the original NM -dimensional STAP data by a rank-reducing (nonsquare) linear transformation Ω of the form

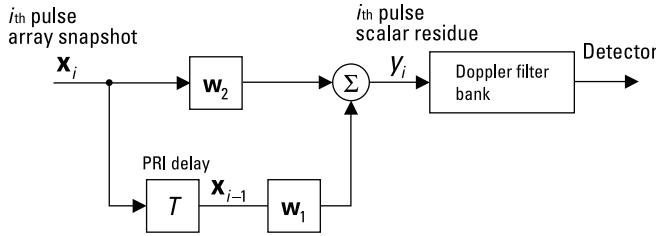


Figure 5.4 A two-pulse DPCA MTI canceller extensible to an arbitrarily long CPI. The scalar residues from the two-pulse DPCA MTI canceller (first stage) are sequentially fed into a Doppler filter bank for further temporal integration gain.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{M-1} \end{bmatrix} = \begin{bmatrix} \mathbf{w}'_1 & \mathbf{w}'_2 & \mathbf{0}' & \dots & \mathbf{0}' \\ \mathbf{0}' & \mathbf{w}'_1 & \mathbf{w}'_2 & \dots & \mathbf{0}' \\ \vdots & & \ddots & & \vdots \\ \mathbf{0}' & \mathbf{0}' & \dots & \mathbf{w}'_1 & \mathbf{w}'_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} = \boldsymbol{\Omega} \mathbf{x} \quad (5.7)$$

where $\mathbf{0}'$ is an N -dimensional row vector, and the DPCA weight vectors \mathbf{w}_1 and \mathbf{w}_2 are given by (5.2). Thus, the original space-time data has been transformed into a clutter-whitened space spanned by \mathbf{y} . For two-pulse DPCA, the dimension of \mathbf{y} is $M - 1$, as indicated by the index subscript in (5.7). This corresponds roughly to a factor of N rank reduction.

The final stage of filtering is thus accomplished by matched filtering in this transformed space to a steering vector of interest. If we are interested in detecting a signal with original NM -dimensional space-time steering vector \mathbf{s} , the corresponding steering vector in the transformed space is $\mathbf{s}_y = \boldsymbol{\Omega} \mathbf{s}$. We are now in a position to state explicitly what the equivalent NM -dimensional space-time weight vector is (along with its ECI) for a two-pulse DPCA followed by Doppler filtering. This is accomplished by working backwards from the final scalar output of the DPCA filtered CPI; that is,

$$\begin{aligned} z &= \mathbf{s}'_y \mathbf{y} \\ &= \mathbf{s}' \boldsymbol{\Omega}' \boldsymbol{\Omega} \mathbf{x} \\ &= \mathbf{w}'_{\text{DPCA}} \mathbf{x} \end{aligned} \quad (5.8)$$

from which we readily ascertain

$$\begin{aligned} \mathbf{w}_{\text{DPCA}} &= \boldsymbol{\Omega}' \boldsymbol{\Omega} \mathbf{s} \\ &= R_{\text{eff}}^I \mathbf{s} \end{aligned} \quad (5.9)$$

and thus

$$R_{\text{eff}}^I = \Omega' \Omega \quad (5.10)$$

Figure 5.5 displays the ideal performance for both full-DoF STAP and DPCA for an $N = M = 16$ sidelooking ULA example. Also shown is the performance of DPCA when errors are present, namely, simple channel-mismatch errors (5° rms phase) and/or PRI mismatch (which is equivalent to $\beta \neq 1$). The loss can be as much as 20 dB or more, especially for lower Doppler shifts.

As we demonstrated in Chapters 2 and 4, adaptive processing can be robust to certain channel-mismatch errors. Thus, it is natural to introduce adaptivity into the DPCA process and hopefully restore performance in the presence of channel mismatch and allow for operation when $\beta \neq 1$. *Adaptive DPCA* (ADPCA) replaces the nonadaptive DPCA MTI canceller weight vector in (5.2) with an adaptive one, namely,

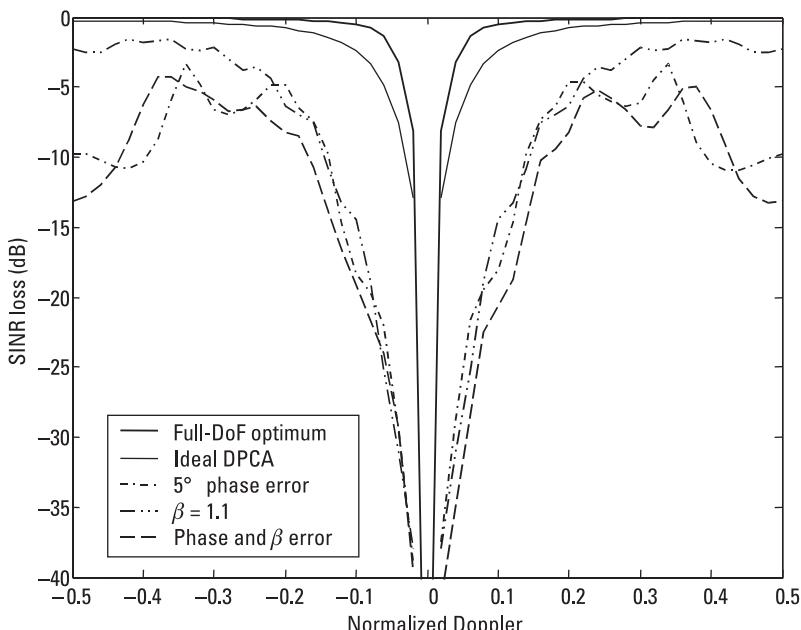


Figure 5.5 Performance of two-pulse DPCA for both ideal and realistic (with errors) conditions. The extreme sensitivity of DPCA to simple channel and β mismatches is one of the prime reasons for considering STAP.

$$\begin{aligned}
\mathbf{w} &= \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} \\
&= R^{-1} \mathbf{s} \\
&= R^{-1} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix}
\end{aligned} \tag{5.11}$$

where R is the $2N \times 2N$ interference covariance matrix associated with two successive array snapshots (i.e., two pulses). This two-pulse ADPCA canceller is then followed by Doppler filtering in precisely the same manner as DPCA (see Figure 5.4). Indeed, ADPCA has a similar mathematical structure to DPCA, namely,

$$\mathbf{w}_{\text{ADPCA}} = \Omega' \Omega \mathbf{s} = R_{\text{eff}}^I \mathbf{s} \tag{5.12}$$

where \mathbf{s} is the full-DoF steering vector, and, in contrast to DPCA, Ω is comprised of the adaptive weights from (5.11).

More insight into the rank-reducing mechanism employed by ADPCA can be gained by examining the final scalar ADPCA output for the entire CPI as was done in (5.8). Specifically,

$$\begin{aligned}
z &= \mathbf{w}'_{\text{ADPCA}} \mathbf{x} \\
&= \mathbf{s}' R_{\text{eff}}^{-1} \mathbf{x} \\
&= \mathbf{s}' \Gamma' \left[\begin{array}{cccccc} 1 R_{2N \times 2N}^{-1} & \mathbf{0}_{2N \times 2N} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} \\ \mathbf{0}_{2N \times 2N} & 2 R_{2N \times 2N}^{-1} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} \\ \vdots & & & \ddots & \\ \mathbf{0}_{2N \times 2N} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} & M-1 R_{2N \times 2N}^{-1} \end{array} \right] \Gamma \mathbf{x} \\
&= \mathbf{s}' \Gamma' \left[\begin{array}{cccccc} 1 R_{2N \times 2N}^{-1} & \mathbf{0}_{2N \times 2N} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} \\ \mathbf{0}_{2N \times 2N} & 2 R_{2N \times 2N}^{-1} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} \\ \vdots & & & \ddots & \\ \mathbf{0}_{2N \times 2N} & \mathbf{0}_{2N \times 2N} & \dots & \mathbf{0}_{2N \times 2N} & M-1 R_{2N \times 2N}^{-1} \end{array} \right]^{-1} \Gamma \mathbf{x} \\
&= \mathbf{s}' \Gamma' (\Gamma R \Gamma')^{-1} \Gamma \mathbf{x}
\end{aligned} \tag{5.13}$$

where R is the full-DoF $NM \times NM$ covariance matrix, ${}_m R_{2N \times 2N}$ is the m th $2N \times 2N$ covariance submatrix associated with the m th and $(m + 1)$ th pulses, and the $2N(M - 1) \times NM$ sub-CPI selection transformation matrix Γ is given by

$$\Gamma = \begin{bmatrix} {}_1 I_{2N \times 2N} & \mathbf{0}_{2N \times N} & \mathbf{0}_{2N \times N} & \dots & \mathbf{0}_{2N \times N} \\ \mathbf{0}_{2N \times N} & {}_2 I_{2N \times 2N} & \mathbf{0}_{2N \times N} & \dots & \mathbf{0}_{2N \times N} \\ \vdots & & & \ddots & \\ \mathbf{0}_{2N \times N} & \mathbf{0}_{2N \times N} & \dots & \mathbf{0} & {}_{M-1} I_{2N \times 2N} \end{bmatrix} \quad (5.14)$$

where $\mathbf{0}_{2N \times N}$ is a $2N \times N$ matrix of zeros and $I_{2N \times 2N}$ is the $2N \times 2N$ identity matrix. Equation (5.14) explicitly shows the rank-reduction mechanism employed by ADPCA: Only consider the space-time covariance matrix of successive pulses. Thus, an alternative equivalent (yet more revealing) expression for the ECI of two-pulse ADPCA is

$$\begin{aligned} \mathbf{w}_{\text{ADPCA}} &= R_{\text{eff}}^I \mathbf{s} \\ &= \Gamma' (\Gamma R \Gamma')^{-1} \Gamma \mathbf{s} \end{aligned} \quad (5.15)$$

from which

$$R_{\text{eff}}^I = \Gamma' (\Gamma R \Gamma')^{-1} \Gamma \quad (5.16)$$

Although we have only considered two pulses in the clutter-rejection filter, the form of (5.16) is the same for an arbitrary number of pulses in the first MTI stage with a suitable modification of (5.14). For example, three-pulse ADPCA has a transformation matrix Γ given by

$$\Gamma = \begin{bmatrix} {}_1 I_{3N \times 3N} & \mathbf{0}_{3N \times N} & \mathbf{0}_{3N \times N} & \dots & \mathbf{0}_{3N \times N} \\ \mathbf{0}_{3N \times N} & {}_2 I_{3N \times 3N} & \mathbf{0}_{3N \times N} & \dots & \mathbf{0}_{3N \times N} \\ \vdots & & & \ddots & \\ \mathbf{0}_{3N \times N} & \mathbf{0}_{3N \times N} & \dots & \mathbf{0} & {}_{M-2} I_{3N \times 3N} \end{bmatrix} \quad (5.17)$$

However, when the number of pulses equals the entire CPI (i.e., M), ADPCA becomes equivalent to full-DoF STAP and can thus no longer be considered a rank-reducing linear transformation.

Figure 5.6 demonstrates the potential of ADPCA to retain performance in the presence of both channel-mismatch and nonunity β conditions. Of course, a sample covariance is required in practice, and this can reduce the effectiveness of ADPCA from what is shown (although all STAP techniques must work with sample statistics). However, we will defer these considerations for the moment and will revisit them later in this chapter when a comparison can be drawn with other STAP techniques.

5.2.2 Post-Doppler (Signal-Dependent) Reduced-Rank STAP

Post-Doppler rank-reducing methods apply adaptivity after Doppler filtering has first been applied. The primary rationale for this is the spatial (angle) dependency of clutter Doppler as described in Chapter 3. This Doppler subbanding approach typically allows for significantly smaller (reduced-rank) adaptivity, with a commensurate reduction in sample support and implementation requirements [1, 2, 5].

The simplest form of post-Doppler rank reduction is the single-bin method, which transforms an NM -dimensional space-time filtering problem

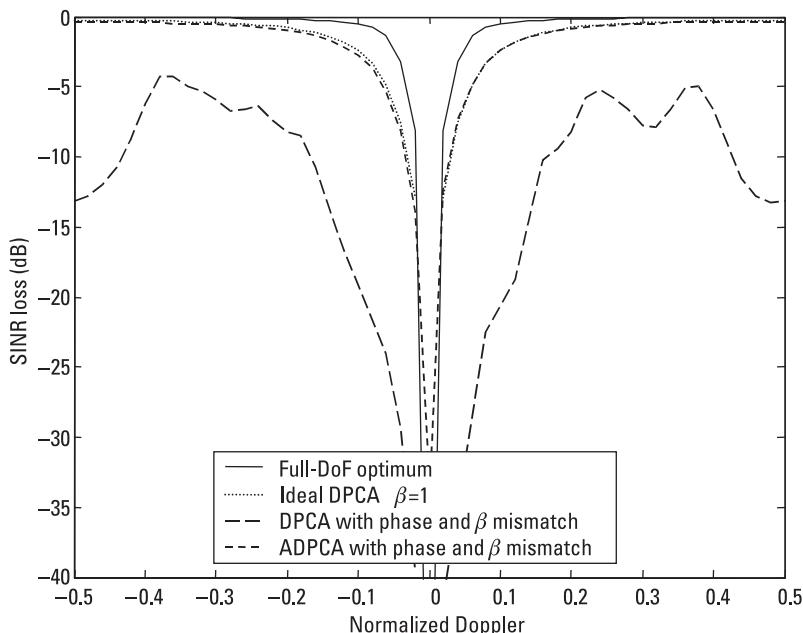


Figure 5.6 Theoretical performance of ADPCA in the presence of both channel and β mismatch.

into (typically) M separate N -dimensional spatial-only adaptive beamforming problems as shown in Figure 5.7. Mathematically, the first 1-D Doppler filtering stage can be represented by

$$\begin{aligned}
 \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} &= \begin{bmatrix} I & *w_0^1 I & *w_0^2 I & \dots & *w_0^{M-1} I \\ I & *w_1^1 I & *w_1^2 I & \dots & *w_1^{M-1} I \\ \vdots & & \ddots & & \vdots \\ I & *w_{M-1}^1 I & *w_{M-1}^2 I & \dots & *w_{M-1}^{M-1} I \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{f}'_0 \otimes I \\ \mathbf{f}'_1 \otimes I \\ \vdots \\ \mathbf{f}'_{M-1} \otimes I \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} \\
 &= \begin{bmatrix} \Omega_0 \\ \Omega_2 \\ \vdots \\ \Omega_{M-1} \end{bmatrix} \mathbf{x}
 \end{aligned} \tag{5.18}$$

where the $N \times NM$ linear rank-reducing transformation is given by $\Omega_m = \mathbf{f}'_m \otimes I$, and I is the $N \times N$ identity matrix. The $M \times 1$ vector \mathbf{f}_m is readily recognized as the m th Doppler DFT steering vector given by

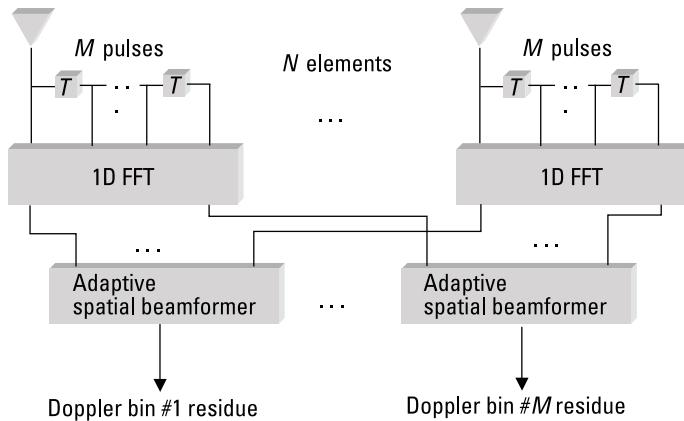


Figure 5.7 Single-bin post-Doppler STAP beamformer.

$$\mathbf{f}_m = \begin{bmatrix} w_m^0 \\ w_m^1 \\ \vdots \\ w_m^{M-1} \end{bmatrix} \quad (5.19)$$

where $w_m^k = e^{j2\pi k \bar{f}_m}$ and \bar{f}_m is the m th normalized Doppler frequency defined in Chapter 2 (Doppler frequency divided by PRF). Also, $*w_m^k$ is the scalar complex conjugate of w_m^k .

Rank reduction is achieved by solving a different spatial-only adaptive beamforming problem for each of the M Doppler bins. Thus, M different $N \times N$ adaptive problems are solved in contrast to M different $NM \times NM$ full-DoF STAP problems (assuming a single look direction and M Doppler bins), which can be a dramatic computational savings.

The m th Doppler bin, \mathbf{y}_m , is thus processed as follows:

$$\begin{aligned} z_m &= \mathbf{w}'_m \mathbf{y}_m \\ &= (R_{y_m y_m}^{-1} \mathbf{s}_{y_m})^{-1} \mathbf{y}_m \\ &= \mathbf{s}' \Omega'_m R_{y_m y_m}^{-1} \Omega_m \mathbf{x} \\ &= \mathbf{s}' \Omega'_m (\Omega_m R \Omega'_m)^{-1} \Omega_m \mathbf{x} \\ &= \mathbf{s}' R_{m_{\text{eff}}}^I \mathbf{x} \end{aligned} \quad (5.20)$$

where $R_{y_m y_m}$ is the $N \times N$ (spatial-only) covariance matrix for the m th Doppler bin, and $\mathbf{s}_{y_m} = \Omega_m \mathbf{s}$, that is, the steering vector corresponding to the full dimensional \mathbf{s} in the reduced rank subspace (m th Doppler filter). Note that the Doppler frequency selected in the full dimensional \mathbf{s} does not necessarily have to agree exactly with that implied by Ω_m . For example, if the cost associated with estimating $R_{y_m y_m}$ is high, one may use a coarse Doppler grid (i.e., M subbands) for $R_{y_m y_m}$, even though many more Doppler bins may be processed. From (5.20), we readily infer that the effective covariance inverse for single-bin post-Doppler STAP (m th Doppler bin) is

$$R_{m_{\text{eff}}}^I = \Omega'_m (\Omega_m R \Omega'_m)^{-1} \Omega_m \quad (5.21)$$

The reader is encouraged to contrast (5.21) with (5.16). In particular, note that the effective inverse in (5.21) is dependent on the desired Doppler

(signal), in contrast to the signal-independent effective inverse for ADPCA in (5.16).

As will shortly be made apparent, so-called single-bin post-Doppler STAP performance is usually poor for most applications. A much better approach is to apply space-time adaptivity to two or more Doppler bins. For example, the effective matrix inverse for three-bin (adjacent bins on either side of the bin of interest) post-Doppler STAP is

$$\begin{aligned} R_{\text{eff}}^I &= \left[\Omega'_{m-1} \quad \Omega'_m \quad \Omega'_{m+1} \right] \left(\begin{bmatrix} \Omega_{m-1} \\ \Omega_m \\ \Omega_{m+1} \end{bmatrix} R \begin{bmatrix} \Omega'_{m-1} & \Omega'_m & \Omega'_{m+1} \end{bmatrix} \right)^{-1} \\ &= {}_3\Omega'_m ({}_3\Omega_m R {}_3\Omega'_m)^{-1} {}_3\Omega_m \end{aligned} \quad (5.22)$$

where

$${}_3\Omega_m = \begin{bmatrix} \Omega_{m-1} \\ \Omega_m \\ \Omega_{m+1} \end{bmatrix} \quad (5.23)$$

Note that, similarly to multipulse ADPCA, in the limit, as the number of bins processed equals M , full-DoF STAP is realized—and thus no rank reduction is achieved.

Figure 5.8 demonstrates the performance of single-bin post-Doppler STAP for several different tapering levels. The need for a heavy Doppler taper is particularly acute in the single-bin approach since mainbeam clutter can leak through the Doppler sidelobes (nominally a sinc response). This problem is significantly alleviated by the multibin approach (see Figure 5.9) at the expense of a substantially increased computational burden.

5.2.3 Other Rank-Reducing Linear Transformations

As will become apparent throughout this chapter, there are many combinations and variations possible in designing a STAP algorithm—indeed too many to address adequately in any one source. Nonetheless, our objective

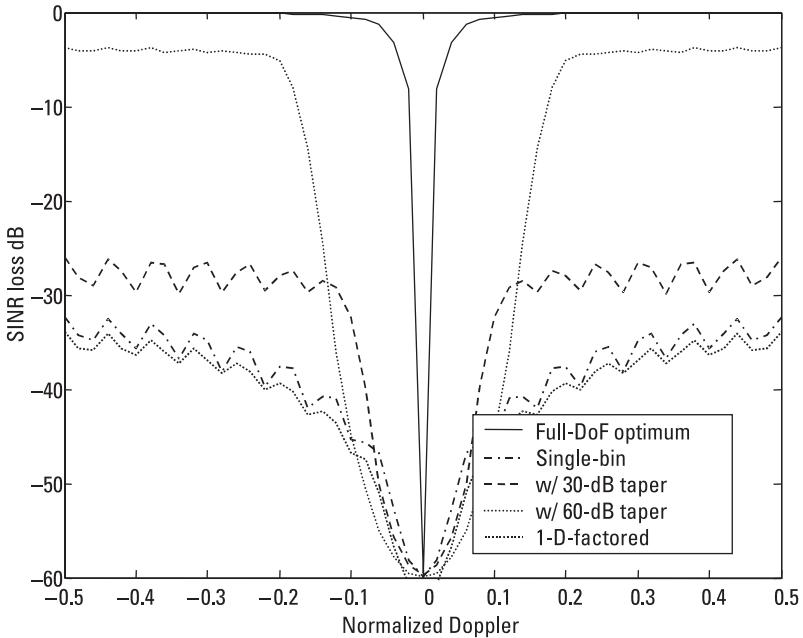


Figure 5.8 Theoretical performance of single-bin post-Doppler STAP. Note the need for heavy Doppler tapering and the presence of significant loss in the mainbeam clutter region.

in this chapter is to “span the space” of salient methods in sufficient detail as to allow the reader to design his or her own tailored solution.

An example of a natural extension to the above pre- and post-Doppler techniques is the so-called beamspace method [1, 5]. In this approach N_B beams are first formed from the N antenna elements (either with phase shifts or time delays [10]), such that $N_B < N$. The beams are then substituted for the ULA outputs. Beamspace rank reduction can be represented by a linear $N_B \times N$ matrix operator, B , applied to each spatial-only array snapshot \mathbf{x}_i , whose output, \mathbf{x}_{B_i} , is given by

$$\mathbf{x}_{B_i} = B\mathbf{x}_i \quad (5.24)$$

where to be of value as an RLT, $\dim(\mathbf{x}_{B_i}) < \dim(\mathbf{x}_i) = N$. The corresponding full-DoF space-time beamspace RLT is thus given by

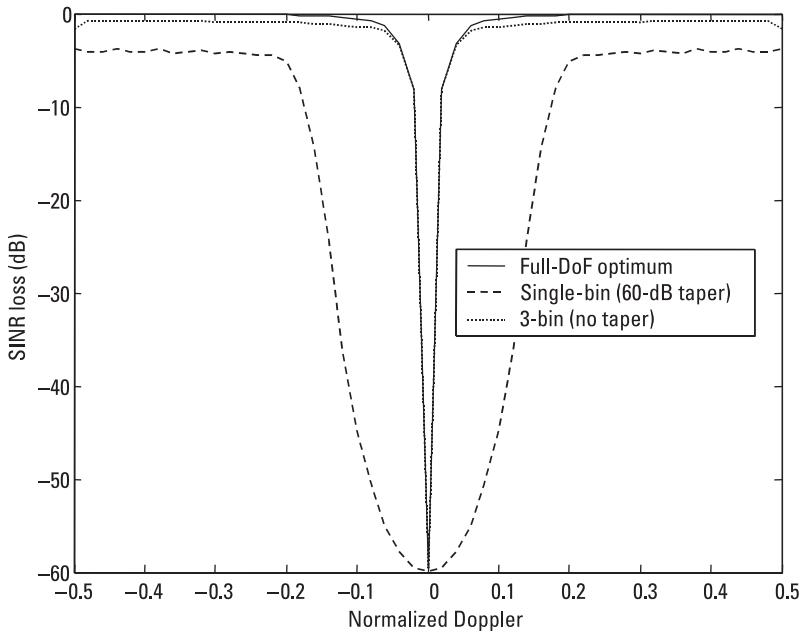


Figure 5.9 Theoretical performance of multibin (three-bin) post-Doppler STAP.

$$\mathbf{x}_B = \begin{bmatrix} B & \dots & \mathbf{0} \\ & B & \\ \vdots & & \ddots \\ \mathbf{0} & & B \end{bmatrix} \mathbf{x} \quad (5.25)$$

$$= B_{ST} \mathbf{x}$$

where B_{ST} is an $N_B \times N_B M$ block diagonal matrix comprised of M repetitions of B .

Due to the linear nature of beamspace rank reduction, other rank-reducing methods (e.g., DPCA, post-Doppler) can be extended to include beamspace via a composite linear operator. For example, the effective covariance inverse for single-bin post-Doppler modified to include beamspace rank reduction, has the general form

$$R_{m_{\text{eff}}}^I = B_{ST}' \Omega_m' (\Omega_m B_{ST} R B_{ST}' \Omega_m')^{-1} \Omega_m B_{ST} \quad (5.26)$$

where it is understood that since the beamspace rank-reduction operator is applied first, the dimensions of Ω_m are now given by $N_B \times N_B M$ instead

of $N \times NM$, as was the case for the element space formulation. It is also tacitly assumed in (5.26) that the beams are linearly independent (likely orthogonal in practice) [10], thus insuring that the indicated inverse exists. The reader is encouraged to derive similar expressions for the ADPCA expression in (5.16) and the multibin post-Doppler approach, along with a dimensional analysis of the constituent matrices.

If the competing clutter of concern is of limited spatial extent and its AoA is generally known a priori, a substantially reduced number of beams could be substituted for the N omnibeams. An example of such an approach for ground clutter cancellation is the *point beams at clutter* (PBC) approach [11]. In one potential manifestation of this approach, separate beams are formed for each Doppler bin (based on the angle-Doppler relationship described in Chapter 3) and then adaptively combined in either a single or multibin approach. However, this approach requires precise knowledge of ownership orientation and velocity along with terrain characteristics to maximize beamspace rank reduction.

Although an indispensable tool for rank reduction in practice (particularly at higher frequencies), we will not explicitly include beamspace methods in the examples considered in this chapter because their efficacy is extremely dependent on specific implementation details. For example, poor sidelobe control can lead to interference leakage. Also, beam-pointing errors can have a significant deleterious effect. Nonetheless, it is easily accommodated as described above and can thus be readily incorporated into any analyses or examples considered in this book.

One last reduced-rank method worthy of discussion is the so-called PRI-staggered [1, 2] method of rank reduction, which is in some sense a cross between time-domain (e.g., ADPCA) and Doppler-domain approaches. Similar to post-Doppler methods, a rank-reducing linear transformation is performed that subbands the full-DoF space-time CPI into distinct Doppler bins. However, unlike the previously discussed single or multibin post-Doppler methods, a single Doppler bin is formed from multiple-PRI-staggered windows or sub-CPI segments. For example, the rank-reducing transformation, ${}_2\Psi_m$, for the m th Doppler bin, two-PRI-staggered case, is given by

$${}_2\Psi_m = \begin{bmatrix} \mathbf{f}'_{m_1} \otimes I_{N \times N} & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \mathbf{f}'_{m_2} \otimes I_{N \times N} \end{bmatrix} \quad (5.27)$$

where ${}_2\Psi_m$ is a $2N \times NM$ matrix, and $\mathbf{f}_{m_1}, \mathbf{f}_{m_2}$ are both $M - 1$ dimensional vectors comprised of the first and last $M - 1$ elements, respectively, of \mathbf{f}_m defined in (5.19). Thus, the effective inverse for the two-PRI-staggered case is given by

$$R_{\text{eff}}^I = {}_2\Psi'_m ({}_2\Psi_m R {}_2\Psi'_m)^{-1} {}_2\Psi_m \quad (5.28)$$

Again the reader is encouraged to extend (5.27) to the three-PRI-staggered (or higher) case.

The performance of the PRI-staggered approach is similar to that achieved by the multibin post-Doppler approach for like adaptive dimensionality (e.g., the three-bin post-Doppler and the three-PRI-staggered approaches have the same size reduced-dimension covariance matrices, and thus to first order, also have similar complexity).

5.3 Data-Dependent Reduced-Rank STAP

Recently, to achieve even more efficient use of the available adaptive DoFs in a STAP beamformer, advanced methods of rank reduction have been developed whose exact behavior is dependent on the received data (generally sample statistics associated with the incoming data)—in other words adaptive rank reduction [12]. Indeed, we have already encountered one of the more popular methods in Chapter 2, namely *principal components* (PCs). Analogous to the deterministic or data-independent case, these methods also can be delineated into signal-dependent or -independent categories, which is to say, the rank-reduction method may or may not (respectively) choose its reduced-rank subspace based on the particular steering vector of interest (see the taxonomy of Figure 5.1).

Another subtle, yet important, discriminant is whether the space-time snapshot that is being interrogated for the presence of targets is included in the data-dependent rank reduction. For example, the PC method performs (directly or indirectly) an eigendecomposition of a sample covariance matrix (as was the case in Chapter 2). If the sample covariance was computed without the current resolution-cell-under-test data, then the PC method is still a linear transformation, because it is tantamount to a linear matrix transformation (nonsquare) whose elements are independent of the data being transformed (see below).

If, on the other hand, the sample covariance was computed with the current resolution-cell-under-test data (along with other cells in general),

then technically this is a nonlinear rank-reduction method. A nonlinear operation can be a significant complicating factor if not carefully controlled and understood. Nonlinearities can introduce spurious false targets (violation of linear superposition) and can transform the statistics of the interference data. For example, the application of a nonlinear transformation to a Gaussian stochastic process results in a generally non-Gaussian output, which can significantly impact the design of the detector (e.g., CFAR detector) [13]. However, reasons to include the cell under test, despite these concerns, might be to detect the presence of discretes (i.e., strong interferers that only exist in a single resolution cell [11]), or if there is a scarcity of representative training data due to extreme nonstationarity [14]. Methods have been developed that allow for the inclusion of the test cell in the training data without significantly altering the linear behavior of the filter. One such method is PC and its variants [15].

Section 5.3.1 will examine the signal-independent variants of the data-dependent rank-reduction methods. These methods choose a linear subspace solely on the basis of the data and not on what particular steering vector will be applied for filtering. In contrast, Section 5.3.2 will introduce signal-dependent methods. Again, we emphasize that these methods can be combined with the deterministic rank-reducing methods previously described. This fact will be illustrated later on in this chapter.

5.3.1 Signal-Independent Methods

The most fundamental and important method in the signal-independent (data-dependent) category is principal components. Here, the *rank-ordering metric* (ROM) is simply the expected energy, or variance, along an eigenvector dimension of the interference covariance matrix (sample or ideal). Rank reduction is achieved by truncating the eigen-based expansion of the covariance. For an illustration of this process, consider Figure 5.10, which shows a block diagram for the *Karhunen-Loeve transform* (KLT) analysis filter bank [16, 17].

The input NM -dimensional random-vector process (space-time array output) is projected onto an orthonormal basis in which the ROM is the variance along an eigenvector dimension. Thus, for example, at the first stage \mathbf{u}_1 (an NM -dimensional vector) is chosen to maximize the objective function J given by

$$J = E(|\mathbf{u}_1' \mathbf{x}|^2) = \mathbf{u}_1' R \mathbf{u}_1 \quad (5.29)$$

Subject to: $\mathbf{u}_1' \mathbf{u}_1 = 1$

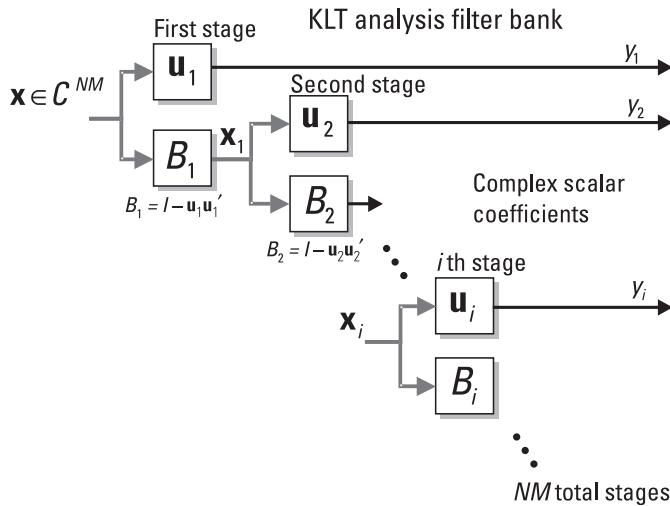


Figure 5.10 KLT analysis filter bank illustrating the decomposition employed in the PC methods. The ROM is simply the expected value of the energy (or variance) at each stage. Rank reduction is achieved by truncating the analysis tree.

Since R is positive-semidefinite, J is a quadratic form and convex functional [18]. Thus, a necessary and sufficient condition for a maximum is that the vector derivative of the Lagrange multiplier augmented objective function vanish; that is,

$$\frac{d}{d\mathbf{u}_1} (\mathbf{u}_1' R \mathbf{u}_1 - \lambda_1 (\mathbf{u}_1' \mathbf{u}_1 - 1)) = 2(R\mathbf{u}_1 - \lambda_1 \mathbf{u}_1) = \mathbf{0}_{NM \times 1} \quad (5.30)$$

where λ_1 is the Lagrange multiplier associated with the unit norm constraint in (5.29) [18]. The result is an eigenvalue/vector problem of the form

$$R\mathbf{u}_1 = \lambda_1 \mathbf{u}_1 \quad (5.31)$$

which for a positive-definite R is known to have exactly NM solutions (not all necessarily corresponding to distinct eigenvalues [16]). Thus, at the first stage, we choose that solution \mathbf{u}_1 which corresponds to a maximum eigenvalue λ_1 . Note that λ_1 is also the variance of y_1 , that is, the variance of the output of the first stage of the KLT analysis tree:

$$\begin{aligned}
E(|y_1|^2) &= E(|\mathbf{u}_1' \mathbf{x}|^2) \\
&= \mathbf{u}_1' R \mathbf{u}_1 \\
&= \lambda_1 \mathbf{u}_1' \mathbf{u}_1 \\
&= \lambda_1
\end{aligned} \tag{5.32}$$

where (5.31) and the unity norm constraint were invoked in the second and third steps, respectively, of (5.32).

Except for the case when no clutter is present (and at most only one rank-one jammer), a 1-D subspace is inadequate for representing the colored-noise process. Consequently, the above procedure is repeated for k stages, where for rank-reduction $k < NM$ (typically $k \ll NM$).

Referring to Figure 5.10, we see that the second stage of the KLT consists of selecting that second orthonormal basis vector (eigenvector) \mathbf{u}_2 such that

$$\max_{\mathbf{u}_2} \mathbf{u}_2' R_1 \mathbf{u}_2 \tag{5.33}$$

$$\text{Subject to: } \mathbf{u}_2' \mathbf{u}_2 = 1$$

where

$$\begin{aligned}
R_1 &= E(\mathbf{x}_1 \mathbf{x}_1') \\
&= E(B_1 \mathbf{x} \mathbf{x}' B_1') \\
&= B_1 R B_1' \\
&= (I - \mathbf{u}_1 \mathbf{u}_1') R (I - \mathbf{u}_1 \mathbf{u}_1')
\end{aligned} \tag{5.34}$$

Thus, R_1 is the covariance (positive-semidefinite, even if R was full-rank) associated with the subspace that is orthogonal to \mathbf{u}_1 since B_1 is a projection operator that effectively removes the \mathbf{u}_1 subspace [16].

The mathematical procedure for finding \mathbf{u}_2 is identical to that for finding \mathbf{u}_1 . The solution is again that eigenvector that maximizes the eigenvalue λ_2 associated with

$$R_1 \mathbf{u}_2 = \lambda_2 \mathbf{u}_2 \tag{5.35}$$

Principal components rank reduction is achieved by truncating the above procedure at k stages where $k < NM$. To be useful, however, k should

be $\ll NM$. Recall that for the jammer-only case considered in Chapter 2, the number of independent jammers, N_J , was shown to be equal to the rank of the colored-noise covariance matrix. Moreover, $N_J < N$, and thus only a few eigenvectors were required to form the adaptive weights. Resulting in a well-behaved adapted pattern and commensurately reduced sample support.

For the clutter case (or clutter-plus-jamming case), Brennan's rule establishes a lower bound for the clutter rank and can thus be used to aid in the selection of a practical stopping rule. However, as was detailed in Chapter 4, Brennan's rule may significantly underestimate the true clutter rank due to a multitude of important factors (e.g., crabbing, nonlinear arrays, ISL, and other decorrelating effects). This can lead to significantly suboptimal performance and/or a loss of the desirable minimal sample support property. However, a method for extending the PC method in the presence of ISL is considered in Section 5.4.

As derived in Chapter 2, the effective covariance inverse for the rank- k PC method is

$$R_{\text{eff}}^I = I - \sum_{i=1}^k \frac{\lambda_i - \lambda_{\min}}{\lambda_i} \mathbf{u}_i \mathbf{u}'_i \quad (5.36)$$

Note that if $\lambda_k \gg \lambda_{\min}$, then (5.36) can be approximated by

$$R_{\text{eff}}^I = I - \sum_{i=1}^k \mathbf{u}_i \mathbf{u}'_i \quad (5.37)$$

which is a projection matrix [16]; that is, $(R_{\text{eff}}^I)^2 = R_{\text{eff}}^I$ (which is left as an exercise for the reader to verify). In this case, the PC method projects the data out of the subspace spanned by the k dominant eigenvectors.

Figure 5.11 shows the performance of the PC method relative to both the straight SMI and diagonally loaded SMI. In this example, Brennan's rule holds (ideal case) and the PC method can produce results with only 62 training samples that SMI can only achieve with $2NM$ or 512 samples. However, as detailed in Chapter 2, the proper choice of diagonal loading can approximate the PC method.

As discussed in Chapter 4 (and elsewhere [1, 2, 4]), Brennan's rule never strictly holds in practice due to a multitude of potential rank-increasing effects. As a consequence, PC methods lose their appeal in practical applications as they require the inclusion of more subdominant eigenvalues to

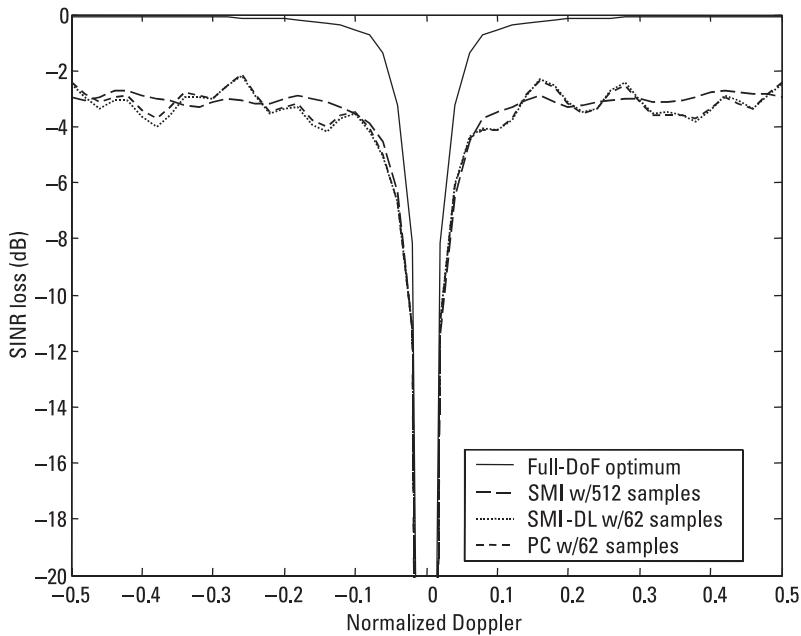


Figure 5.11 Performance of the PC method for an ideal ULA for which Brennan’s rule holds. The efficiency of the PC method results in a substantial reduction in sample support requirements compared with the SMI method. Note that a properly loaded SMI (10 dB in this case) achieves similar performance since it mathematically approximates the PC method for this case (see Chapter 2 for an explanation of this effect).

account for the additional interference subspaces. For example, in Figure 5.12, the same example is repeated, but with the inclusion of ICM. If Brennan’s rule is employed as a PC rank-selection criterion, it will result in significant SINR loss. This can be alleviated by including more components, but at the expense of a commensurately increased sample support (and computational burden). A method of extending the PC approach to more realistic environments—without increasing the requisite sample support—will be discussed in Section 5.4.

Another less-known, signal-independent, rank-reduction method somewhat related to principal components, is the Hung-Turner projection [19]. This method has the form

$$R_{\text{eff}}^I = I - YY' \quad (5.38)$$

where

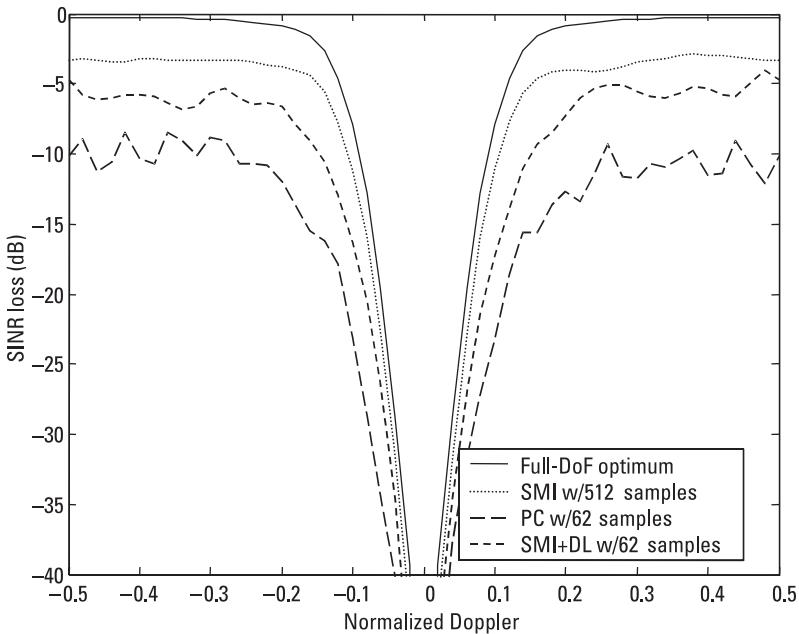


Figure 5.12 Impact when interference subspace leakage (10 mph ICM) is present. Note that for this case, Brennan's rule is not a good choice for rank selection. A method for addressing this problem is described in Section 5.4.

$$Y = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_L] \quad (5.39)$$

is a $NM \times L$, $L \ll NM$, matrix comprised of space-time array snapshots (i.i.d. training data) generally selected from surrounding range-bins. If the colored-noise interference is both strong (relative to the noise floor) and low rank (i.e., $k \ll NM$), then the PC method can be approximated by (5.38) if $L \approx k$ [19]. This is because k i.i.d. samples will, in the limit as $\text{CNR} \rightarrow \infty$, exactly span the k -dimensional colored-noise subspace (almost surely [20]). Hence, projecting orthogonal to that subspace [i.e., (5.38)] will result in an optimum filter. Zatman has analyzed the optimum choice of L for finite CNR [19].

Although not generally considered a robust practical technique, it is relatively simple to implement (no need to perform an eigendecomposition) and, for the dominant low-rank colored-noise case, can yield essentially equivalent performance as PC (i.e., same SINR for properly selected training data)—hence its appeal. The reader is encouraged to experiment with this

technique and compare it to PC. However, for the reasons previously cited, we will not consider it further in this book.

5.3.2 Signal-Dependent Methods

As was the case for the data-independent rank-reducing methods, the main reason for considering signal-dependency in the rank-reduction process is to make more efficient use of the available adaptive DoFs. Post-Doppler methods (an example of a signal-dependent data-independent rank-reduction method) solved a different adaptive problem for each Doppler bin of interest, in contrast to ADPCA, which solved only one adaptive clutter-cancellation problem for all Doppler. If the number of signals (i.e., steering vectors) is large, signal-dependent methods can have dramatically greater overall computational burdens. However, they also enjoy a high degree of parallelism since in most cases the signal-dependent solutions can be derived independently of each other. However, this feature can only be exploited if a large number of parallel CPUs is available in the real-time signal processor.

Perhaps the most straightforward method for introducing signal dependency (i.e., steering-vector-specific dependency) into the subspace selection process is to modify the ROM of principal components. For example, the standard signal-independent ROM for PC is simply the strength of the eigenvalue. There are several relatively straightforward methods for modifying the ROM to include a desired steering vector [12].

5.3.2.1 Relative Importance of Eigenbeam Method

One method is based on examining the amount of influence a particular eigenvector has on the weight vector associated with a particular steering-vector solution. For example, recall from Chapter 2 the expression for the optimum weight vector expressed in terms of the eigenvalues/vectors of the interference covariance matrix; that is,

$$\begin{aligned} \mathbf{w} &= \left(I - \sum_{i=1}^{NM} \frac{\lambda_i - \lambda_{\min}}{\lambda_i} \mathbf{u}_i \mathbf{u}'_i \right) \mathbf{s} \\ &= \mathbf{s} - \sum_{i=1}^{NM} \frac{\lambda_i - \lambda_{\min}}{\lambda_i} (\mathbf{u}'_i \mathbf{s}) \mathbf{u}_i \end{aligned} \quad (5.40)$$

An examination of (5.40) reveals a natural choice of ROM; namely

$$\max_{\{i\}} \frac{\lambda_i - \lambda_{\min}}{\lambda_i} |\mathbf{u}'_i \mathbf{s}| \quad (5.41)$$

Thus, for example, $\{\tilde{\lambda}_1, \tilde{\mathbf{u}}_1\}$ is the eigenvalue-eigenvector pair that maximizes (5.41), which may or may not correspond to the PC ROM; that is, $\{\lambda_1, \mathbf{u}_1\}$. Assuming, without loss of generality, that $\mathbf{s}'\mathbf{s} = \mathbf{u}'_i \mathbf{u}_i = 1$, (5.41) represents the normalized amplitude of the eigenbeam (i.e., \mathbf{u}_i) subtracted from \mathbf{s} . For convenience, we will refer to (5.41) as the relative importance of eigenbeam, or RIE, metric of rank ordering.

There are three ways in which the metric in (5.41) could be zero: (1) $\lambda_i = \lambda_{\min}$, which implies (in practice) that the eigenvalue is equal to the white-noise floor and thus should not be subtracted (recall that the optimum filter for the white-noise case is $\mathbf{w} = \mathbf{s}$); (2) $|\mathbf{u}'_i \mathbf{s}| = 0$, which implies that \mathbf{u}_i is orthogonal to \mathbf{s} and can thus be ignored (regardless of the strength of the eigenvalue); and (3) both conditions are met. In practice, one should be very careful about condition 2, since errors in the assumed steering vector can cause erroneous results. A remedy is to place a lower bound on $|\mathbf{u}'_i \mathbf{s}|$ to insure that it is not completely ignored (a natural choice is to set the lower bound equal to the expected steering-vector mismatch residue error). Indeed, sensitivity to steering-vector mismatches is a serious concern in all signal-dependent reduced-rank methods.

From (5.40) and (5.41), it follows that the ECI for the *relative importance of eigenbeam* (RIE) method is

$$R_{\text{eff}}^I = \left(I - \sum_{i=1}^k \frac{\tilde{\lambda}_i - \lambda_{\min}}{\tilde{\lambda}_i} \tilde{\mathbf{u}}_i \tilde{\mathbf{u}}'_i \right) \quad (5.42)$$

where $\{\tilde{\lambda}_i, \tilde{\mathbf{u}}_i\}$ denotes the eigenvalue-eigenvector pair that ranked in the i th position based on the RIE rank-ordering metric. Thus, in general, $\tilde{\lambda}_i \neq \lambda_i$ (although this is entirely possible). Again for meaningful rank reduction, it is tacitly assumed that $k \ll NM$.

5.3.2.2 Cross Spectral Metric

Another technique for modifying the PC approach to include signal dependency is based on viewing the STAP problem from a *sidelobe canceller* (SLC) perspective [21]. From Figure 5.13, we see that the space-time filtering problem can reformulated from a max SINR problem (which is the framework adopted so far in this book) to a minimum variance estimation problem [21]. Specifically, the objective is to choose the auxiliary space-time weight

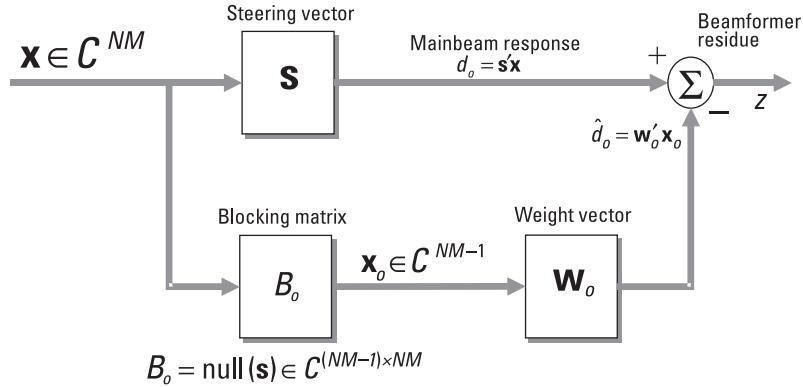


Figure 5.13 Space-time sidelobe canceller structure.

vector \mathbf{w}_o [which is of dimension $(NM - 1) \times 1$] so as to minimize the scalar MSE; that is,

$$\min_{\mathbf{w}_o} E(|d_o - \hat{d}_o|^2) \quad (5.43)$$

Since \hat{d}_o is a linear estimator, a necessary and sufficient condition for minimizing the quadratic convex functional (5.43) is setting its vector derivative equal to zero (on an element basis):

$$\begin{aligned} \frac{d}{d\mathbf{w}_o} E(|d_o - \hat{d}_o|^2) &= \frac{d}{d\mathbf{w}_o} E((d_o - \mathbf{w}_o' \mathbf{x}_o)' (d_o - \mathbf{w}_o' \mathbf{x}_o)) \\ &= -2E(\mathbf{x}_o d_o^*) + 2E(\mathbf{x}_o \mathbf{x}_o') \mathbf{w}_o \\ &= -2R_{\mathbf{x}_o d_o} + 2R_{\mathbf{x}_o \mathbf{x}_o} \mathbf{w}_o \\ &= \mathbf{0} \end{aligned} \quad (5.44)$$

which yields

$$\mathbf{w}_o = R_{\mathbf{x}_o \mathbf{x}_o}^{-1} \mathbf{r}_{\mathbf{x}_o d_o} \quad (5.45)$$

where

$$R_{\mathbf{x}_o \mathbf{x}_o} = B_o R B_o' \quad (5.46)$$

is the $(NM - 1) \times (NM - 1)$ covariance matrix associated with the auxiliary data \mathbf{x}_o and

$$\mathbf{r}_{\mathbf{x}_o d_o} = B_o R \mathbf{s} \quad (5.47)$$

is the $(NM - 1) \times 1$ cross-correlation matrix (actually a vector). B is a $(NM - 1) \times (NM)$ blocking [21] or null [10] matrix whose $(NM - 1)$ rows are orthonormal to \mathbf{s} , the desired steering vector (also referred to as the mainbeam response [21]).

To obtain the form of the effective covariance inverse for the SLC formulation, we again examine the output and work back as follows:

$$\begin{aligned} z &= \mathbf{s}' \mathbf{x} - \mathbf{w}'_o B_o \mathbf{x} \\ &= (\mathbf{s}' \mathbf{x} - \mathbf{w}'_o B_o) \mathbf{x} \end{aligned} \quad (5.48)$$

from which the NM -dimensional space-time weight vector \mathbf{w} is readily extracted and further reduced

$$\begin{aligned} \mathbf{w} &= \mathbf{s} - B_o \mathbf{w}'_o \\ &= \mathbf{s} - B_o R_{\mathbf{x}_o \mathbf{x}_o}^{-1} \mathbf{r}'_{\mathbf{x}_o d_o} \\ &= \mathbf{s} - B_o R_{\mathbf{x}_o \mathbf{x}_o}^{-1} B_o R \mathbf{s} \\ &= (I - B_o R_{\mathbf{x}_o \mathbf{x}_o}^{-1} B_o R) \mathbf{s} \end{aligned} \quad (5.49)$$

and thus

$$R_{\text{eff}}^I = I - B_o R_{\mathbf{x}_o \mathbf{x}_o}^{-1} B_o R \quad (5.50)$$

While seemingly a different approach than the max SINR method, the above SLC formulation is equivalent and yields identical output SINR. One way to see this is first to observe that the output from the mainbeam channel (scalar) concatenated with the $(NM - 1)$ outputs from the auxiliary channels formed by the $(NM - 1) \times 1$ blocking matrix is the result of applying an $NM \times NM$ unitary (and thus invertible) matrix Q given by

$$Q = \begin{bmatrix} \mathbf{s}' \\ B_o \end{bmatrix} \quad (5.51)$$

The unitary property of Q is easily established since, by definition, the rows of B are pairwise orthonormal and orthonormal to \mathbf{s} (and, without loss of generality, we can assume $\mathbf{s}'\mathbf{s} = 1$). Since the application of a unitary matrix does not alter the SINR [21], a completely equivalent (and optimal from a max SINR perspective) filter is given by

$$\begin{aligned}\mathbf{w}_{\text{eq}} &= (QRQ')^{-1}\mathbf{y}_s \\ &= (QRQ')^{-1}Q\mathbf{s} \\ &= (QR^{-1}Q')Q\mathbf{s}\end{aligned}\quad (5.52)$$

where \mathbf{y}_s is the steering vector corresponding to \mathbf{s} in the transformed space, and the unitary property of Q was invoked in obtaining the final inverse form. Note, as expected from Figure 5.13, $\mathbf{y}_s = [1 \ 0 \ \dots \ 0]'$ since the first channel is the mainbeam response. The corresponding SINR is given by

$$\begin{aligned}\text{SINR} &= \mathbf{y}_s' QR^{-1} Q' \mathbf{y}_s' \\ &= [1 \ 0 \ \dots \ 0] \begin{bmatrix} \sigma_{d_o}^2 & \mathbf{r}_{\mathbf{x}_o d_o}' \\ \mathbf{r}_{\mathbf{x}_o d_o} & R_{\mathbf{x}_o \mathbf{x}_o} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \frac{1}{\sigma_{d_o}^2 - \mathbf{r}_{\mathbf{x}_o d_o}' R_{\mathbf{x}_o \mathbf{x}_o}^{-1} \mathbf{r}_{\mathbf{x}_o d_o}}\end{aligned}\quad (5.53)$$

where

$$\begin{aligned}\sigma_{d_o}^2 &\triangleq \mathbb{E}(|d_o|^2) \\ &= \mathbb{E}(|\mathbf{s}'\mathbf{x}|^2) \\ &= \mathbf{s}'R\mathbf{s}\end{aligned}\quad (5.54)$$

and use was made of the partitioned matrix inverse lemma [9], that is,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{12} \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & -A_{11}^{-1}A_{12}B_2 \\ -B_2A_{21}A_{11}^{-1} & B_2 \end{bmatrix}\quad (5.55)$$

where

$$B_1 \triangleq (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \quad (5.56)$$

and

$$B_2 = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \quad (5.57)$$

Note that only B_1 needed to be computed in (5.53) to obtain the SINR.

Again, (5.53) is the max SINR for the transformed (via Q) colored-noise matched filter, which is also equal to the original optimum SINR, $\mathbf{s}'R^{-1}\mathbf{s}$, since a unitary transformation cannot alter the SINR. Showing that this is also the SINR corresponding to the minimum MSE SLC is straightforward. From an examination of the signal paths in Figure 5.13 we have

$$\begin{aligned} \text{SINR}_{\text{SLC}} &= \frac{|\mathbf{s}'\mathbf{s}|^2}{\mathbf{E}(|d_o - \hat{d}_o|^2)} \\ &= \frac{1}{\mathbf{E}(|\mathbf{s}'\mathbf{x} - \mathbf{w}'_o\mathbf{x}_o|^2)} \\ &= \frac{1}{\mathbf{E}((\mathbf{s}'\mathbf{x} - \mathbf{w}'_o\mathbf{x}_o)(\mathbf{s}'\mathbf{x} - \mathbf{w}'_o\mathbf{x}_o)')} \\ &= \frac{1}{\mathbf{E}(\mathbf{s}'\mathbf{x}\mathbf{x}'\mathbf{s}) - \mathbf{w}'_o\mathbf{E}(\mathbf{x}_o\mathbf{x}'_o)\mathbf{w}_o} \\ &= \frac{1}{\sigma_{d_o}^2 - \mathbf{w}'_o R_{\mathbf{x}_o\mathbf{x}_o} \mathbf{w}_o} \\ &= \frac{1}{\sigma_{d_o}^2 - \mathbf{r}'_{\mathbf{x}_o d_o} R_{\mathbf{x}_o\mathbf{x}_o}^{-1} \mathbf{r}_{\mathbf{x}_o d_o}} \end{aligned} \quad (5.58)$$

which is identical to (5.53), thus establishing the equivalence between the SLC and max SINR (colored-noise match filter) approaches. Note that the orthogonality between \mathbf{s} and the rows of B_o was invoked to eliminate the cross terms in line four of the above proof.

The *cross-spectral metric* (CSM) rank-ordering metric as developed in [22] is a method for prioritizing the eigenvalues/vectors of the covariance matrix associated with the auxiliary channels. This is first accomplished by expressing the SINR in terms of these eigenvalues; that is,

$$\begin{aligned} \text{SINR} &= \frac{1}{\sigma_{d_o}^2 - \mathbf{r}_{\mathbf{x}_o d_o}' R_{\mathbf{x}_o \mathbf{x}_o}^{-1} \mathbf{r}_{\mathbf{x}_o d_o}} \\ &= \frac{1}{\sigma_{d_o}^2 - \sum_{i=1}^{NM-1} \frac{|\mathbf{u}_i' \mathbf{r}_{\mathbf{x}_o d_o}|^2}{\lambda_i}} \end{aligned} \quad (5.59)$$

where $(\lambda_i, \mathbf{u}_i)$ denotes the i th eigenvalue/vector pair associated with $R_{\mathbf{x}_o \mathbf{x}_o}$, and use was made of the eigen-based form of the inverse of $R_{\mathbf{x}_o \mathbf{x}_o}$; that is,

$$R_{\mathbf{x}_o \mathbf{x}_o}^{-1} = \sum_{i=1}^{NM-1} \frac{\mathbf{u}_i \mathbf{u}_i'}{\lambda_i} \quad (5.60)$$

From (5.59) it is apparent that the eigenvectors that contribute the most to maximizing SINR (thus minimizing MSE) are those for which the following quantity is relatively large:

$$\frac{|\mathbf{u}_i' \mathbf{r}_{\mathbf{x}_o d_o}|^2}{\lambda_i} \quad (5.61)$$

Using this CSM as the rank-ordering metric yields the following reduced-rank SLC beamformer:

$$\begin{aligned} \mathbf{w}_o &= R_{\mathbf{x}_o \mathbf{x}_o}^{-1} \mathbf{r}_{\mathbf{x}_o d_o} \\ &\approx \sum_{i=1}^k \frac{(\tilde{\mathbf{u}}_i' \mathbf{r}_{\mathbf{x}_o d_o})}{\tilde{\lambda}_i} \tilde{\mathbf{u}}_i' \end{aligned} \quad (5.62)$$

where $(\tilde{\lambda}_i, \tilde{\mathbf{u}}_i)$ denotes the eigenvalue/vector pair that ranked in the i th position based on the CSM. For effective rank reduction, it is again assumed that $k \ll NM - 1$.

Finally, the corresponding effective covariance inverse for the CSM method is obtained by examining the effective full-DoF space-time weight vector; specifically,

$$\begin{aligned}
 \mathbf{w}_{\text{CSM}} &= \mathbf{s} - B'_o \mathbf{w}_{o_{\text{CSM}}} \\
 &= \mathbf{s} - B'_o \mathbf{w}_{o_{\text{CSM}}} \\
 &= \mathbf{s} - B'_o \sum_{i=1}^k \frac{(\tilde{\mathbf{u}}'_i \mathbf{s})}{\tilde{\lambda}_i} \tilde{\mathbf{u}}_i \\
 &= \left(I - B'_o \sum_{i=1}^k \frac{\tilde{\mathbf{u}}_i \tilde{\mathbf{u}}'_i}{\tilde{\lambda}_i} \right) \mathbf{s}
 \end{aligned} \tag{5.63}$$

from which we readily ascertain

$$R_{\text{eff}}^I = I - B' \sum_{i=1}^k \frac{\tilde{\mathbf{u}}_i \tilde{\mathbf{u}}'_i}{\tilde{\lambda}_i} \tag{5.64}$$

Equation (5.64) has a similar form as the PC methods previously described, but with a different rank-ordering metric and unitary transformation Q .

5.3.2.3 Multistage Wiener Filter

The above signal-dependent methods (RIE and CSM) for modifying the data-dependent PC method are relatively straightforward methods for modifying the basis selected for rank reduction. However, both methods utilize an existing basis (eigenvectors of R or $R_{\mathbf{x}_o \mathbf{x}_o}$) that was developed without any knowledge of \mathbf{s} . An entirely different method for selecting a useful basis that explicitly takes \mathbf{s} into account was developed by Goldstein et al. and is referred to as the *multistage Wiener filter* (MWF) [23].

To see how the MWF selects suitable basis vectors, consider the first stage depicted in Figure 5.14. The objective is to choose a rank-one $NM - 1$ dimensional basis vector, \mathbf{h}_1 , that is useful for minimizing the MSE. The optimal choice for \mathbf{h}_1 is, of course, the minimum MSE weight vector \mathbf{w}_o given by (5.45). But this solution requires knowledge of the full covariance matrix $R_{\mathbf{x}_o \mathbf{x}_o}$, and is thus counter to the objective of rank reduction. We thus seek a basis vector that is useful for estimating d_o , but does not require full knowledge of $R_{\mathbf{x}_o \mathbf{x}_o}$.

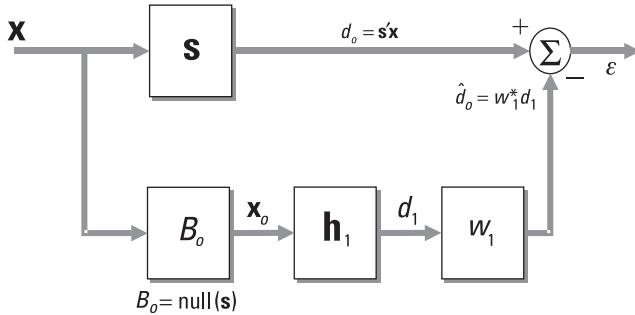


Figure 5.14 First stage of the MWF.

One candidate, which is the rationale of the MWF, is to choose that rank-one basis vector \mathbf{h}_1 such that the cross-correlation energy between $d_1 = \mathbf{h}_1' \mathbf{x}_o$ and d_o is maximized [12]; that is,

$$\max_{\mathbf{h}_1} E(|\mathbf{h}_1' \mathbf{x}_o d_o^*|^2) = |\mathbf{h}_1' \mathbf{r}_{\mathbf{x}_o d_o}|^2 \quad (5.65)$$

Since (5.65) is the magnitude squared of an inner product, a direct application of Schwarz's inequality yields the optimum \mathbf{h}_1 , namely,

$$\mathbf{h}_1 = \frac{\mathbf{r}_{\mathbf{x}_o d_o}}{\|\mathbf{r}_{\mathbf{x}_o d_o}\|} \quad (5.66)$$

where a unity norm constraint was imposed on the basis vector \mathbf{h}_1 ; that is, $\|\mathbf{h}_1\| = (\mathbf{h}_1' \mathbf{h}_1)^{1/2} = 1$. Note that as desired, (5.66) does not depend directly on $R_{\mathbf{x}_o \mathbf{x}_o}$. Equation (5.66) is an intuitively appealing result: Set \mathbf{h}_1 equal to the cross correlation between \mathbf{x}_o and d_o . With this choice of \mathbf{h}_1 , the optimal rank-one MWF (scalar) weight vector w_1 is given by

$$w_1 = R_{d_1 d_1}^{-1} \mathbf{r}_{d_1 d_o} \quad (5.67)$$

$$= \frac{\sigma_{d_1 d_o}}{\sigma_{d_1}^2}$$

where $\sigma_{d_1 d_o}$ is the cross correlation between d_1 and d_o , and $\sigma_{d_1}^2$ is the variance of d_1 .

Since \mathbf{h}_1 is not in general colinear with the optimal Wiener vector \mathbf{w}_o (as this requires the full covariance and cross-correlation information), additional basis vectors will be required to adequately span the colored-noise subspace that is interfering with the mainbeam response d_o . To remedy this, we can add more stages to the above process—hence the meaning of “multistage” in MWF.

Figure 5.15 shows a two-stage (rank-two) MWF. The space spanned by \mathbf{x}_1 can be seen to be orthogonal to $\text{span}(\mathbf{s}, \mathbf{h}_1)$ due to the blocking matrices B_o and B_1 . Our objective at this stage is to choose a basis vector \mathbf{h}_2 , which, reapplying the criterion from the first stage, is maximally correlated with the coefficient from the first stage d_1 , [12, 23], that is,

$$\max_{\mathbf{h}_2} : E(|\mathbf{h}_2' \mathbf{x}_1 d_1^*|^2) = |\mathbf{h}_2' \mathbf{r}_{x_1 d_1}|^2 \quad (5.68)$$

subject to the usual unit norm constraint $\mathbf{h}_2' \mathbf{h}_2 = 1$. Again from Schwarz’s inequality we have

$$\mathbf{h}_2 = \frac{\mathbf{r}_{x_1 d_1}}{\|\mathbf{r}_{x_1 d_1}\|} \quad (5.69)$$

The corresponding optimal SLC weight for the two-stage MWF is thus

$$\mathbf{w}_2 = R_{\mathbf{d}_2 \mathbf{d}_2}^{-1} \mathbf{r}_{\mathbf{d}_2 d_o} \quad (5.70)$$

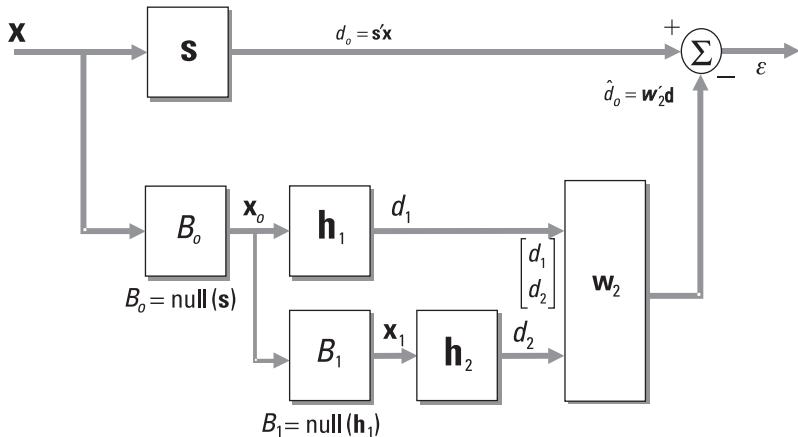


Figure 5.15 Two-stage MWF.

where

$$\begin{aligned} R_{\mathbf{d}_2 \mathbf{d}_2} &= E(\mathbf{d}_2 \mathbf{d}'_2) \\ &= E\left(\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \begin{bmatrix} d_1^* & d_2^* \end{bmatrix}\right) \end{aligned} \quad (5.71)$$

and

$$\mathbf{r}_{\mathbf{d}_2 d_o} = E(\mathbf{d}_2 d_o^*) \quad (5.72)$$

\mathbf{w}_2 can be related back to the original covariance and cross-correlation quantities of (5.46) and (5.47) as follows:

$$\begin{aligned} R_{\mathbf{d}_2 \mathbf{d}_2} &= E(\mathbf{d}_2 \mathbf{d}'_2) \\ &= E(L'_2 \mathbf{x} \mathbf{x}' L_2) \\ &= L'_2 R L_2 \end{aligned} \quad (5.73)$$

where

$$L'_2 = \begin{bmatrix} \mathbf{h}'_1 B_o \\ \mathbf{h}'_2 B_1 B_o \end{bmatrix} \quad (5.74)$$

with

$$\begin{aligned} \mathbf{r}_{\mathbf{d}_2 d_o} &= E(\mathbf{d}_2 d_o^*) \\ &= E(L'_2 \mathbf{x} \mathbf{x}' \mathbf{s}) \\ &= L'_2 R \mathbf{s} \end{aligned} \quad (5.75)$$

From (5.50) and (5.70), the corresponding NM -dimensional space-time weight vector \mathbf{w} for the two-stage MWF is

$$\begin{aligned} \mathbf{w} &= \mathbf{s} - L_2 \mathbf{w}_2 \\ &= \mathbf{s} - L_2 R_{\mathbf{d}_2 \mathbf{d}_2}^{-1} \mathbf{r}_{\mathbf{d}_2 d_o} \\ &= \mathbf{s} - L_2 (L'_2 R L_2)^{-1} L'_2 R \mathbf{s} \\ &= (I - L_2 (L'_2 R L_2)^{-1} L'_2 R) \mathbf{s} \end{aligned} \quad (5.76)$$

Thus, the ECI for the two-stage MWF is readily seen to be

$$R_{\text{eff}}^I = I - L_2(L'_2 RL_2)^{-1}L'_2 R \quad (5.77)$$

In general, for the k -stage case, we have

$$R_{\text{eff}}^I = I - L_k(L'_k RL_k)^{-1}L'_k R \quad (5.78)$$

where

$$L'_k = \begin{bmatrix} \mathbf{h}'_1 B_o \\ \mathbf{h}'_2 B_1 B_o \\ \vdots \\ \mathbf{h}'_k B_{k-1} \dots B_1 B_o \end{bmatrix} \quad (5.79)$$

Recently, Weippert et al. [24] have shown the relationship between the above sequential decomposition and the method of conjugate gradients. This observation is of more than just theoretical interest since many previously established convergence analyses can be readily adapted to the special MWF case [24].

5.3.3 Comparison of Data-Dependent Rank-Reduction Methods

Before applying these methods to the STAP problem, we will examine the relative rank-compression performance of the data-dependent methods for the 1-D beamforming example first considered in Chapter 2. This simple illustration will aid in understanding the differences between the various methods without extraneous complications. Section 5.5 will illustrate application of these methods to STAP.

Figure 5.16 shows the SINR performance for the six-jammer example of Figure 2.4 as a function of the rank k of the adaptive beamformer (number of principal components or MWF stages) for each of the data-dependent reduced-rank methods. Figure 5.16(a) is for the ideal case (exact covariance); Figure 5.16(a) shows the finite sample case ($2k$ samples, where k is the jammer rank). Note that in the ideal case, the MWF enjoys a significant rank-compression advantage over the other methods. However, this advantage is essentially removed when a finite amount of training data is available. In fact, all six basis vectors are required to achieve good nulling performance—exactly the rank required by the PC methods. This last behavior is due to

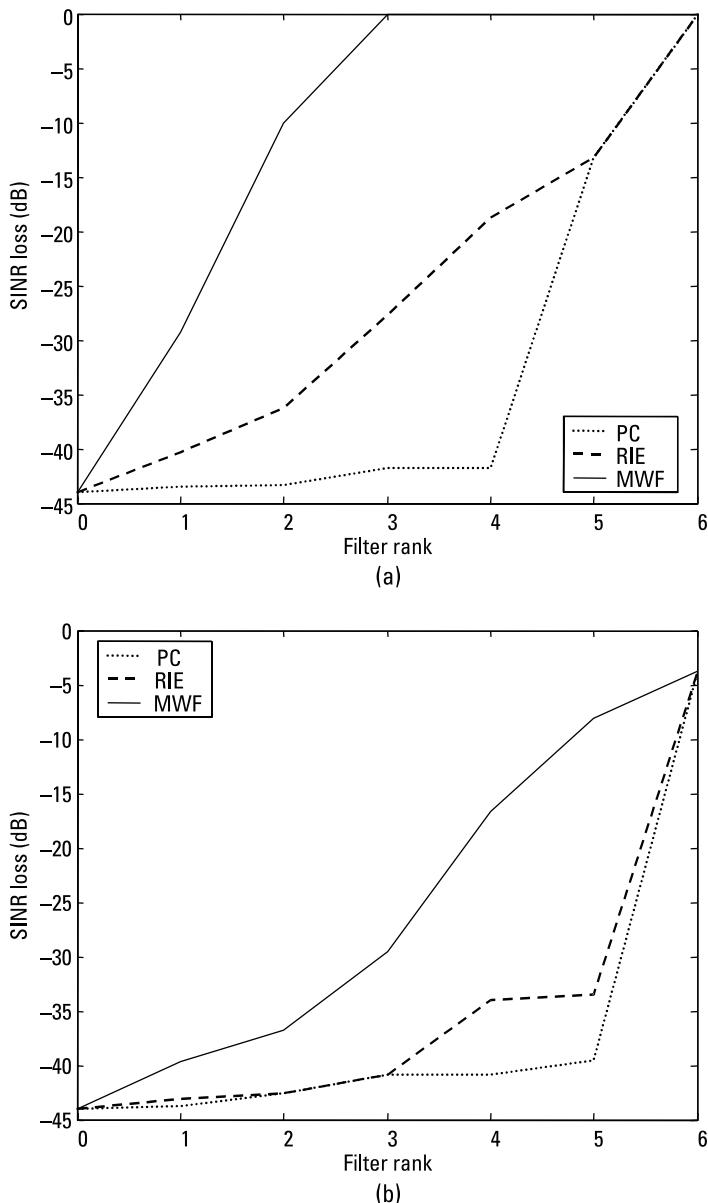


Figure 5.16 Rank-compression performance as a function of filter rank: (a) known-covariance case, and (b) finite sample covariance case.

the need to fully characterize the colored-noise subspace before an optimal choice of basis can be determined. This characterization is fundamentally limited by the amount of sample support (as well as the estimation procedure). The interested reader is referred to [12] for further details regarding SINR performance as a function of filter rank and sample support.

To gain further insight into how the MWF can achieve higher rank-compression (for the known or large sample-support covariance case), consider the eigenbeam responses of Figure 5.17. Depicted is the PC response, along with the principal MWF response. Since PC is based strictly on energy and not the steering vector of interest (in this case broadside), it simply selects an eigenvector that is essentially aligned with the maximum jammer signal. In contrast, the MWF uses the shaping of the quiescent steering vector to select responses corresponding to jammers that have a greater influence, that is, stronger JNR after beamforming. Of course, if there are steering-vector mismatches, the efficiency of the MWF will be reduced. However, this effect is highly scenario-dependent, and it is difficult to make general statements regarding MWF robustness.

5.4 Structured-Covariance and Model-Based Methods

Structured-covariance methods, also referred to sometimes as model-based methods, refer to techniques that explicitly incorporate specific mathematical constraints or structures into the covariance estimation process. Implicit constraints, such as Doppler or beamspace binning, are almost always present in any MTI radar and are generally not considered explicit model-based methods. Instead, they can be considered data-independent/signal-dependent methods in accordance with the taxonomy of Figure 5.1.

One of the primary reasons for considering structured-covariance methods is their potential for significantly reducing sample-support requirements. This can be of considerable practical value in nonstationary environments where there may be a dearth of auxiliary data available for covariance estimation (or weight training) [12].

Section 5.4.1 will consider a broadly applicable class of structured covariance methods based on the observation that many realistic clutter/interference scenarios involve a multitude of ISL phenomena, such as ICM, dispersion, and so on, as described in Chapter 4. In particular, a method that combines the CMT structure with either PC or diagonal loading will be explored. Due to the ubiquitous nature of ISL, this class of methods enjoys particularly broad applicability [12].

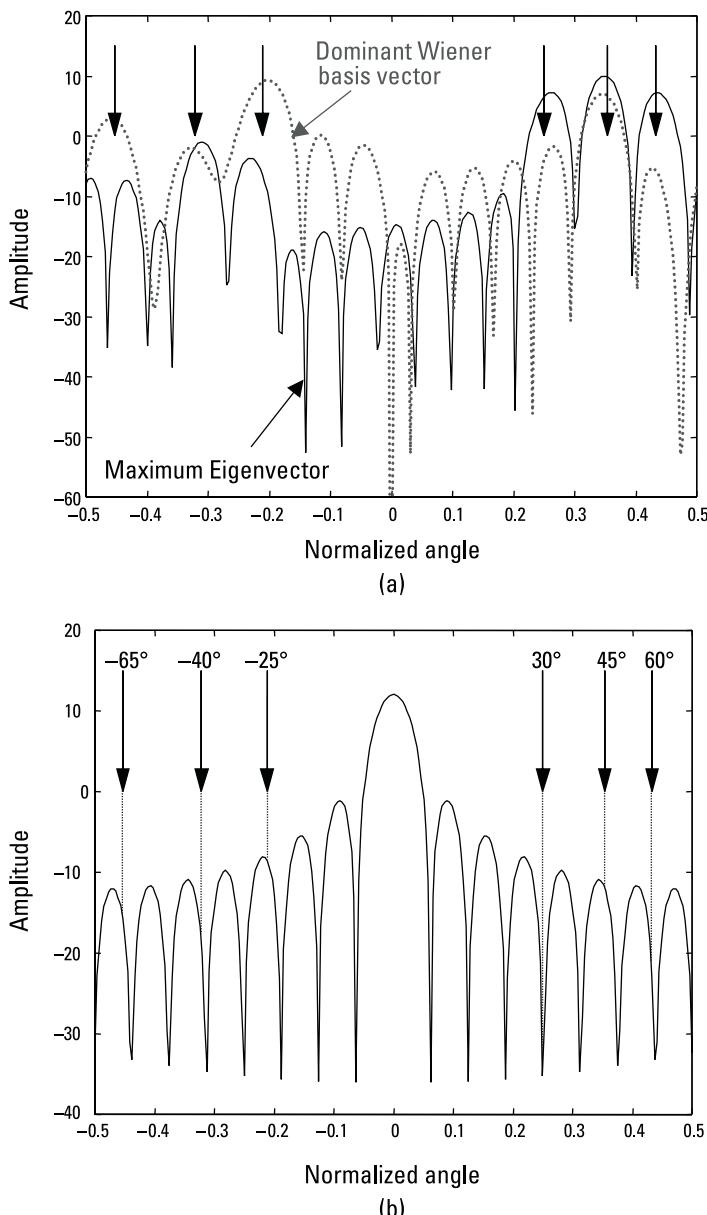


Figure 5.17 Eigenbeams corresponding to the primary (first stage) basis functions associated with the PC and MWF methods: (a) eigenbeam patterns, and (b) quiescent pattern with incident jammer locations.

Other model-based methods with far less broad applicability include Toeplitz and TBT [6, 7], *forward-backward smoothing* (FBS) [25, 26], and parametric multichannel signal modeling [27–29]. Due to their narrow scope, we will only briefly discuss their basic modeling assumptions in Section 5.4.2 and provide references for the interested reader.

5.4.1 Covariance Matrix Tapers

As discussed in Chapter 4, many real-world phenomena give rise to a specific covariance structure involving a Hadamard-multiplicative matrix operation referred to as a CMT [30]. Specifically, the structure is of the form $R \circ T$, where R is the underlying space-time covariance matrix in the absence of any ISL or random modulation effects, and T is the generally composite CMT that accounts for all ISL effects.

To see how this CMT structure can be exploited to provide a more efficient covariance estimation process, consider the space-time eigenspectrum of our standard $N = M = 16$ ULA with and without ICM (see Figure 5.18). Notice that the spectrum of the dominant eigenvalues ($k = 31$ for this example) appears essentially unaffected by the ICM. As discussed in Chapter 4, ICM is well modeled as a CMT and only requires a few simple parameters to characterize (e.g., Billingsley’s model [31]). This suggests a hybrid estimation approach in which the dominant eigenvalues/vectors are estimated in the usual manner using the PC method (with Brennan’s rule serving as a nominal stopping criterion), and then a final estimated covariance is obtained by applying a CMT and reestablishing the noise floor [4]; that is,

$$\hat{R}_{PC-CMT} = \left(\sum_{i=1}^k \hat{\lambda}_i \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i' \right) \circ \hat{T} + \hat{\sigma}_n^2 \quad (5.80)$$

where the set of k eigenvalues/vectors $\{\hat{\lambda}_i, \hat{\mathbf{u}}_i\}$ is estimated in the normal manner (e.g., an SVD [9]) with the usual (e.g., $2k$) sample support, \hat{T} is the estimated CMT, and $\hat{\sigma}_n^2$ is the estimated noise floor. Before commenting on the methods for obtaining \hat{T} and $\hat{\sigma}_n^2$, a bit more justification for (5.80) is in order, other than just a simple visual examination of an eigenspectrum invariance. Note that the corresponding ECI for the PC-CMT of (5.80) is simply its inverse.

Consider the Karhunen-Loeve [21] representation of the unmodulated space-time clutter signal, which is based on a stochastic expansion of the clutter-only eigenvectors; that is,

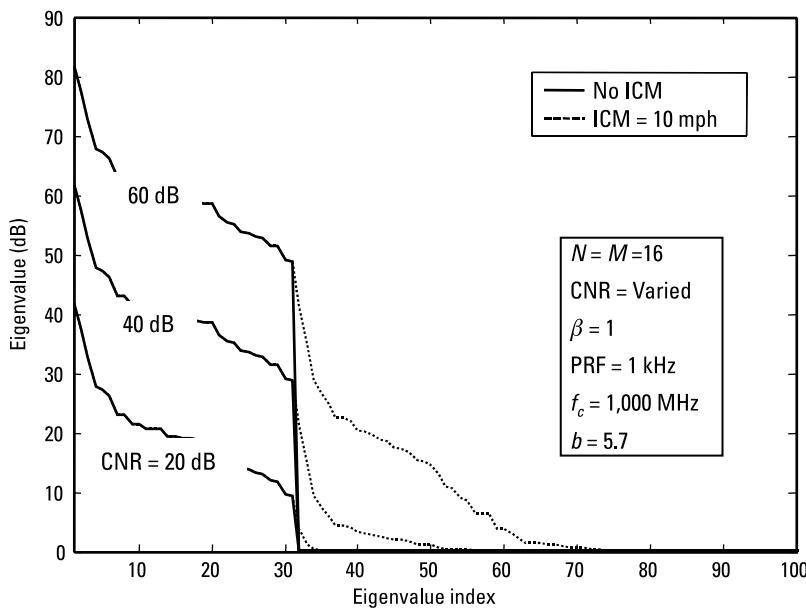


Figure 5.18 Illustration of the relative invariance of the dominant eigenvalues to small amounts of ISL. This property can be exploited by a structured-covariance estimation procedure that models the subdominant eigenvalues/vectors by a CMT.

$$\mathbf{x}_c = \sum_{i=1}^k \gamma_{c_i} \mathbf{u}_{c_i} \quad (5.81)$$

where \mathbf{x}_c is an NM -dimensional vector RV corresponding to the clutter-only signal in the absence of random modulation, \mathbf{u}_{c_i} is the i th eigenvector associated with the clutter-only covariance matrix (see Chapter 3), and the k complex scalars and mutually uncorrelated RVs $\{\gamma_{c_i}\}$ have the property

$$E(\gamma_{c_i} \gamma_{c_q}^*) = \lambda_{c_i} \delta_{iq} \quad (5.82)$$

where δ_{iq} is the Kronecker delta function.

As described in Chapter 4, the presence of a CMT implies that a vector random modulation \mathbf{t} is present; that is,

$$\begin{aligned}\mathbf{x}_c \odot \mathbf{t} &= \left(\sum_{i=1}^k \gamma_{c_i} \mathbf{u}_{c_i} \right) \odot \mathbf{t} \\ &= \sum_{i=1}^k \gamma_{c_i} (\mathbf{u}_{c_i} \odot \mathbf{t})\end{aligned}\quad (5.83)$$

where the distributive property of Hadamard products has been invoked [9].

In practice, the amount of modulation is small; otherwise, the colored noise (clutter/jamming) would be effectively whitened, precluding the need for structured space-time filtering as a simple white-noise matched filter would suffice. Indeed, it is this characteristic of ISL that results in the seemingly unperturbed eigenvalues of Figure 5.18. To see this, we first explicitly invoke the small modulation approximation; namely

$$\mathbf{t} \approx \mathbf{1} + \boldsymbol{\epsilon} \quad (5.84)$$

where $\mathbf{1}$ is an NM -dimensional vector with unity entries, and $\boldsymbol{\epsilon}$ is a zero-mean, NM -dimensional complex-valued vector RV such that $E(\|\boldsymbol{\epsilon}\|) \ll 1$. Substituting this approximation into (5.83) yields

$$\begin{aligned}\mathbf{x}_c \odot \mathbf{t} &\approx \sum_{i=1}^k \gamma_{c_i} \mathbf{u}_{c_i} \odot (\mathbf{1} + \boldsymbol{\epsilon}) \\ &= \sum_{i=1}^k \gamma_{c_i} \mathbf{u}_{c_i} + \sum_{i=1}^k \gamma_{c_i} \mathbf{u}_{c_i} \odot \boldsymbol{\epsilon}\end{aligned}\quad (5.85)$$

where the vector Hadamard identity property was invoked; that is, $\mathbf{u}_{c_i} \odot \mathbf{1} = \mathbf{u}_{c_i}$. Equation (5.85) shows that the introduction of a small ISL modulation can be viewed as a small complex modulation of a large carrier signal (the eigenvectors). From elementary communication systems theory [32], it is known that this will result in a large carrier (original eigenvectors) with small sideband signals (subdominant eigenvectors) due to the modulation.

The corresponding covariance matrix is given by [9]

$$\begin{aligned}\text{cov}(\mathbf{x}_c \odot \mathbf{t}) &\approx \text{cov}(\mathbf{x}_c \odot (\mathbf{1} + \boldsymbol{\epsilon})) \\ &= R_c \odot (\mathbf{1}_{NM \times NM} + R_{\boldsymbol{\epsilon}}) \\ &= R_c + \delta R\end{aligned}\quad (5.86)$$

where $R_c = \text{cov}(\mathbf{x}_c)$, $R_\epsilon = \text{cov}(\boldsymbol{\epsilon})$, and $\delta R = R_c \odot R_\epsilon$. While by definition $\|\delta R\| \ll 1$, it is in general of full rank [9]. Thus, it can be decomposed as follows

$$\delta R = P_c \delta R P_c' + (I - P_c) \delta R (I - P_c)' \quad (5.87)$$

where P_c is an $NM \times NM$ projection matrix onto the subspace spanned by the k clutter eigenvectors, and, thus, $I - P_c$ is the projection onto the orthogonal compliment [16]. Since, again by definition, $\|P_c \delta R P_c'\| \ll 1$, we see that the subspace spanned by R_c is essentially undisturbed—which explains our observations in Figure 5.18. The subdominant eigenvalues/vectors introduced into the subspace orthogonal to the unmodulated clutter spectrum may or may not be of concern. As discussed in Chapter 4, these subdominant components only have an impact if their eigenvalues exceed the elemental noise floor (see the discussion on the iceberg effect in Chapter 4).

The preceding argument justifies the use of (5.80) under the conditions (normally encountered) of modest ISL. Thus, the indicated eigenvalues/vectors in (5.80) can be estimated in the usual manner with—and this is the important part—a sample support based on k , the nominal clutter rank (for which Brennan's rule can be used as a guide). Of course, it is tacitly assumed that the CNR is high; otherwise, the effective rank of the clutter may be even less than that predicted by Brennan's rule (and the subdominant eigenvalues due to ISL and so on are below the noise floor—and thus of no concern). The noise-floor estimate $\hat{\sigma}_n^2$ is easily obtained from routine calibration procedures and thus will not be further discussed. However, the choice of CMT \hat{T} is a different matter.

There are a number of different ways in which \hat{T} may be estimated, such as using knowledge of average channel-mismatch specifications, bandwidth dispersion, array manifold calibration tables, and real-time equalization procedures. Also, knowledge of the general clutter type and wind state can be used. However, since the presence of some form of space and/or time decorrelating mechanism is a virtual certainty, having some reasonable estimate of \hat{T} is far better than none at all. To illustrate this point, Figure 5.19 compares a nominal PC method (rank = k) and several PC-CMT methods with varying degrees of CMT accuracy for a fixed sample support of $2k = 62$ (two times the Brennan rank). Even with imperfect knowledge of \hat{T} , the PC-CMT method consistently outperforms direct PC, as evidenced by the ± 5 -mph mismatch cases in Figure 5.19.

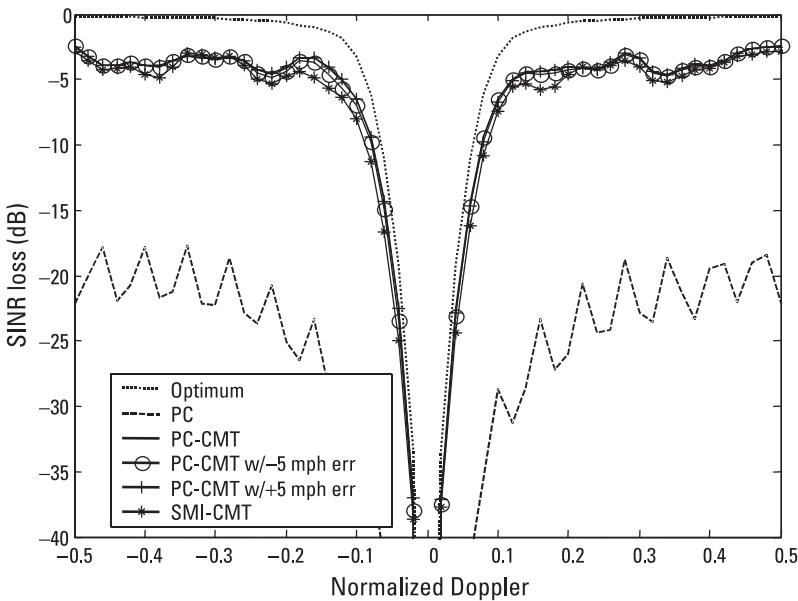


Figure 5.19 Performance of the PC-CMT and SMI-CMT approaches in the presence of ICM. Note the robustness of mismatches in the CMT.

Figure 5.19 also shows a diagonally loaded sample matrix method augmented with a CMT. As described in Chapter 2, diagonal loading approximates PC methods and can thus be used as a surrogate for the eigendecomposition indicated in (5.80). This can be of considerable practical implementation value since this SMI-CMT algorithm is amenable to conventional systolic array implementation (see Chapter 6). Specifically, the SMI-CMT estimator is of the form

$$\hat{R}_{\text{SMI-CMT}} = (\hat{R}_{\text{SMI}}) \odot \hat{T} + \delta I \quad (5.88)$$

where \hat{R}_{SMI} is the usual sample matrix estimate (based on $2k$ samples, and thus generally positive-semidefinite), and δ is the loading factor [33]. Note that the effective covariance inverse is simply the inverse of (5.88).

Since ISL in its many forms is a ubiquitous real-world phenomenon, the CMT structure should likewise be widely adopted to account for subdominant interference. The above methods are merely two simple illustrations of how one might incorporate the CMT structured covariance. Many others will undoubtedly be developed. For example, a more formal estimation procedure might entail a composite joint MLE approach. However these

approaches are generally only efficient asymptotically [34] and thus might require significantly larger sample supports than the divide-and-conquer approach considered here.

5.4.2 Other Structured-Covariance Methods

Due to the ubiquitous presence of spatial and/or temporal decorrelating effects and channel mismatch, the above CMT structured-covariance methods are fairly general and are thus not overly constraining. Other structured-covariance methods have been suggested that are far more restrictive and, thus, less generally applicable. For example, from Section 3.2, under the assumptions of a perfectly matched ULA in the presence of mutually uncorrelated clutter point scatterers in the absence of antenna crabbing and any ISL, it can be deduced that the space-time covariance matrix of an ideal ULA with constant PRF has a TBT structure of the form [1]

$$R = \begin{bmatrix} R_o & R_1 & R_2 & \dots & R_{M-1} \\ R_1^* & R_o & R_1 & \dots & R_{M-2} \\ R_2^* & R_1^* & R_o & \dots & R_{M-3} \\ \vdots & & & \ddots & \vdots \\ R_{M-1}^* & R_{M-2}^* & R_{M-3}^* & \dots & R_o \end{bmatrix} \quad (5.89)$$

which is an $NM \times NM$ block-Toeplitz matrix comprised of $M, N \times N$ block matrices, R_o, R_1, \dots, R_{M-1} , which in turn are Toeplitz (hence the term *Toeplitz-block-Toeplitz*); that is,

$$R_m = \begin{bmatrix} r_{m_o} & r_{m_1} & r_{m_2} & \dots & r_{m_{N-1}} \\ r_{m_{-1}} & r_{m_o} & r_{m_1} & \dots & r_{m_{N-2}} \\ r_{m_{-2}} & r_{m_{-1}} & r_{m_o} & \dots & r_{m_{N-3}} \\ \vdots & & & \ddots & \vdots \\ r_{m_{-N+1}} & r_{m_{-N+2}} & r_{m_{-N+3}} & \dots & r_{m_o} \end{bmatrix} \quad (5.90)$$

where the entries are complex scalars. Thus, the space-time matrix is comprised of $O(NM)$ distinct elements, dramatically reducing the number of unknowns from the general case that involves $O(N^2 M^2)$ unknowns. This implies that there exists significant potential for reducing sample-support requirements [6, 7].

Unfortunately, fitting a TBT matrix to sample data is a highly nonlinear process (see, for example, [6, 7], which discuss the MLE and EM approaches) that involves iterative processes that are not guaranteed to converge—a significant drawback for practical applications. Since the efficacy of this technique is tightly coupled to the specifics of any given application, it is difficult to draw any general conclusions regarding its utility. The interested reader is thus referred to the research literature for further details.

Another very closely related technique that applies under the same general ULA constraints cited for Toeplitz is the method of FBS [25, 26, 35]. Consider the linear flip-up-down-then-conjugate operator $J(\cdot)$: $\mathbb{C}^{NM} \rightarrow \mathbb{C}^{NM}$, given by

$$J(\mathbf{x}) = \text{conj} \left(\begin{bmatrix} \mathbf{0} & & 1 \\ & \ddots & \\ 1 & & \mathbf{0} \end{bmatrix} \mathbf{x} \right) \quad (5.91)$$

It is straightforward to show that a linear phase Vandermonde steering vector, \mathbf{v} , is, to within a complex scalar phase term, invariant to $J(\cdot)$, that is, $J(\mathbf{v}) = e^{j\varphi} \mathbf{v}$ [25, 26]. However, if \mathbf{n} is an NM -dimensional vector RV comprised of complex, zero-mean, and uncorrelated entries, then $\mathbf{E}(J(\mathbf{n})\mathbf{n}') = \emptyset$. This suggests that if the colored-noise portion of the total interference signal is a linear superposition of Vandermonde steering vectors, then additional data samples can be constructed whose white-noise components are uncorrelated. For the Gaussian case, this further implies that the extended samples have additive white noise that is independent of the original samples—thus effectively increasing the available sample support. Again, the efficacy of this technique in practice has been found to be extremely scenario-dependent, and it is thus not considered further in this text.

Another class of structured-covariance methods effectively reduces the number of independent dimensions of the full space-time stochastic process by imposing a vector autoregressive multichannel signal model [27–29]. For example, the output of the ULA can be thought of as a vector AR time series of the form

$$\mathbf{x}(m + 1) = A(m)\mathbf{x}(m) + \mathbf{v}(m) \quad (5.92)$$

where $\mathbf{x}(m) \in \mathbb{C}^N$ is the array snapshot vector associated with the m th pulse, and $\mathbf{v}(m)$ is assumed to be a white-noise vector RV, that is, $\mathbf{E}(\mathbf{v}(m)\mathbf{v}'(q)) = \delta_{mq} D$, where D is a diagonal matrix. Thus, all of the correlation structure

is contained in the state transition matrix $A(m) \in \mathbb{C}^{N \times N}$ (and the implied autoregressive structure). Although certainly a simplifying structural constraint, its effectiveness is also highly scenario-dependent and thus not amenable to generalizations. The interested reader should consult [27–29] and the references cited therein for further details.

5.5 Illustrative Design Examples

In this section, STAP design examples are presented that illustrate ways in which various reduced-rank STAP methods can be combined to yield good SINR performance, but with dramatically reduced sample-support (and computational) requirements. Specifically, we consider both signal-independent (ADPCA with data-dependent CMT rank reduction) and signal-dependent (multibin post-Doppler with MWF rank reduction). Our motivation is simple: to minimize sample support to insure reliable performance and/or graceful degradation in nonstationary heterogeneous clutter environments. While this design process does not explicitly include real-time embedded computing constraints, they are indirectly, but significantly, alleviated due to the presence of a deterministic rank-reducing prestage transformation [1, 2]. Further discussions of real-time implementation issues and methods can be found in Chapter 6 (and references cited therein).

A somewhat stressing clutter scenario is chosen to magnify the impact of each of the design stages and to provide a better gauge of algorithm robustness. Figure 5.20 shows the full-DoF eigenspectrum for an $N = M = 16$ ULA with a CNR of 60 dB, $\beta = 1$, ICM of 15 mph, 5° rms of simple channel mismatch (rank-one CMT), 2° of crab with 40-dB back lobe rejection, and decorrelating channel mismatch (full-rank CMT) corresponding to $\Delta_\epsilon = 0.01$, $\Delta_\phi = 1^\circ$ (see Chapter 4). Note that the effective rank of the clutter (~ 100 ; see Chapter 4) is substantially greater than that predicted by Brennan's rule, which for this example is 31.

At this stage, the reader might be wondering why we have chosen not to illustrate the design methods on real data. There are actually several reasons: First, there is a paucity of public-domain multichannel STAP data that is truly representative of diverse and complex clutter environments. Although some data does exist (e.g., [24, 36–39]), it is by no means enough to draw definitive conclusions regarding the best algorithm for a given operating condition. Second, the reader may very well be interested in selecting a STAP algorithm for a radar configuration and operation that is dissimilar to the publicly available data. So rather than potentially bias the

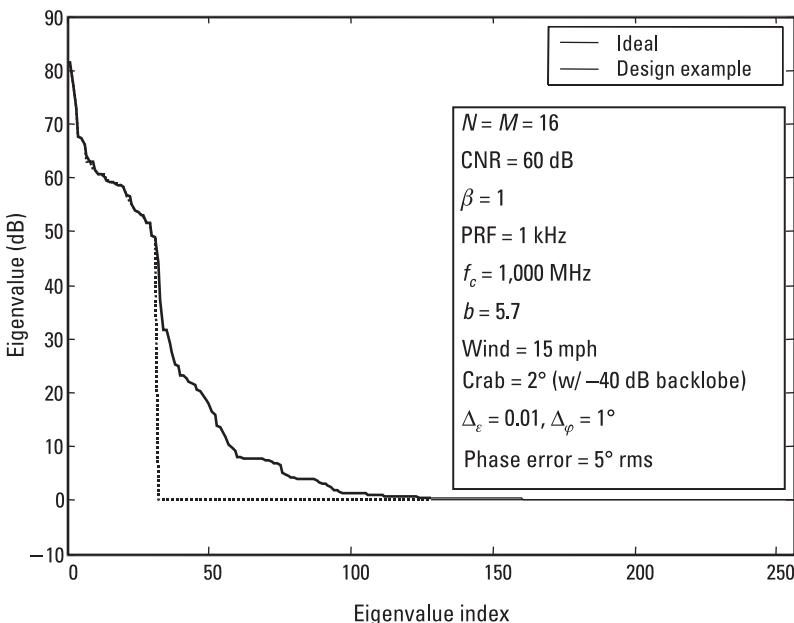


Figure 5.20 Eigenspectrum for the design example involving both strong clutter and a multitude of real-world deleterious effects.

reader with information based on sparse experimentation, this book has opted to develop the salient constitutive signal and system models that allow for the exploration of countless scenarios—thereby affording the reader the ability to build a comprehensive experience base.

5.5.1 Signal-Independent Approach

Our first design step is the choice of a data-independent (deterministic) rank-reducing transformation. Even at this stage, there is a significant number of options available, including beamspace versus element space, and pre-Doppler (PRI space) versus post-Doppler. As beamspace is usually indicated when there is a large number of independent receive elements or when severe implementation constraints exist, we will not consider it in this particular design example. We are therefore left with a choice of pre- or post-Doppler.

Figure 5.21 shows the ideal SINR loss curves for the full-DoF optimum, three-pulse ADPCA, and three-bin post-Doppler cases. While both reduced-rank methods perform well, the post-Doppler method involves substantially greater overall computations since a different adaptive problem must be

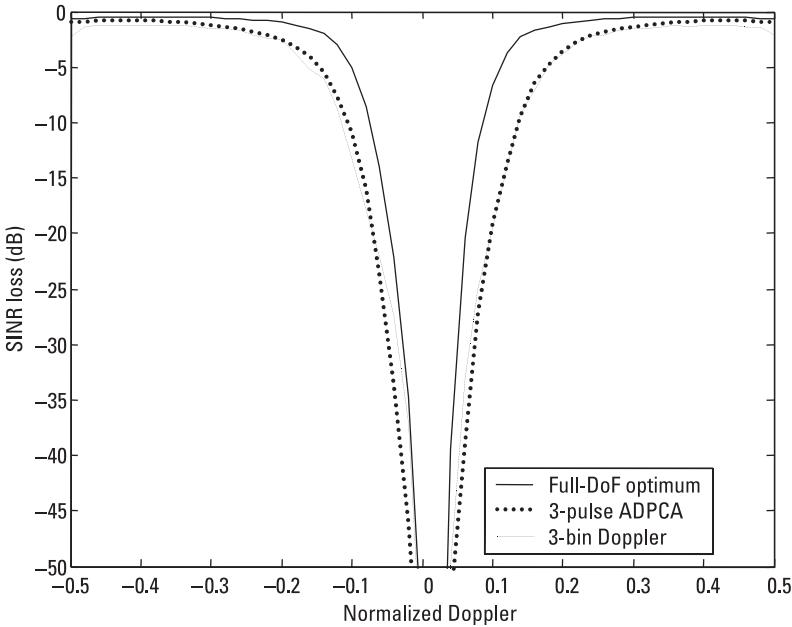


Figure 5.21 Comparison between three-pulse ADPCA and three-bin post-Doppler reduced-rank linear transformations relative to full-DoF optimum performance.

solved for each Doppler bin (see Section 5.5.2). In this example, we select three-pulse ADPCA. The three-bin post-Doppler case is considered in Section 5.5.2.

Although we have reduced the adaptive DoFs from 256 to $3 \times 16 = 48$, with a commensurate reduction in sample-support requirements, we would still like to achieve further reductions to boost robustness. To accomplish this, we will next consider data-dependent methods.

Figure 5.22 shows the eigenspectrum for the three-pulse ADPCA covariance matrix. Brennan's rule predicts a nominal clutter rank of $N + \beta(M - 1) = N + \beta 2 = 18$, but the actual effective rank is again substantially higher due to the composite ISL effects and crabbing. However, by modeling the additional clutter subspace by a CMT, we can theoretically reduce the sample support to that associated with only 18 DoFs using PC-CMT—which is more than an order-of-magnitude reduction from the original full-DoF case. In other words, the requisite $3N \times 3N$ ADPCA covariance matrix R is replaced with the corresponding PC-CMT estimate

$$\hat{R}_{PC-CMT} = \left(\sum_{i=1}^{18} \hat{\lambda}_i \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i' \right) \circ \hat{T} + \hat{\sigma}_n^2 \quad (5.93)$$

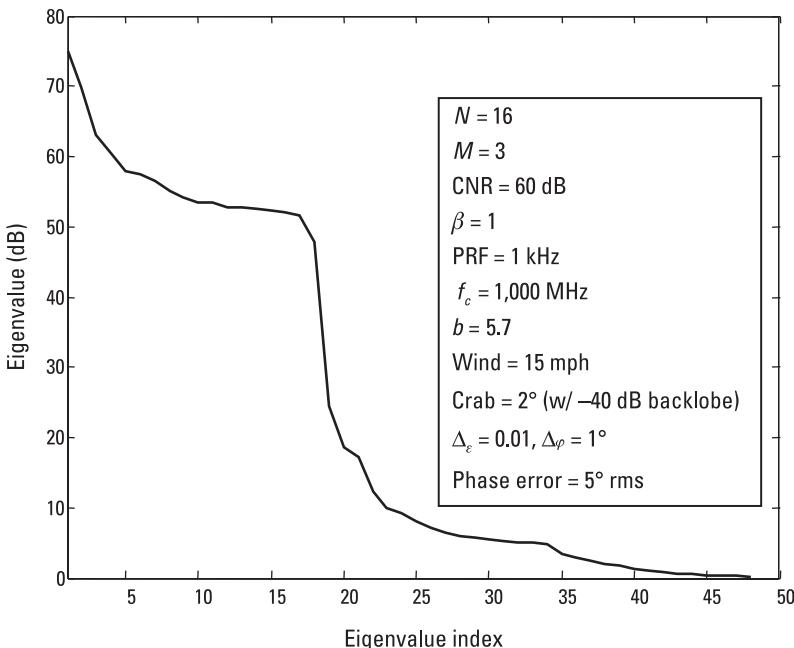


Figure 5.22 Eigenspectrum of the three-pulse ADPCA covariance.

Shown in Figure 5.23 is the performance of the hybrid ADPCA/PC-CMT approach and a alternate method that substitutes a 10-dB diagonally loaded SMI-CMT for the PC-CMT (easier to implement since no eigendecomposition is required). The performance was achieved with only $2k = 36$ training samples—as compared with $2NM = 512$. Figure 5.24, which shows wind-speed errors of ± 10 mph, demonstrates the robustness of this approach to errors in the estimated CMT. Notice that even in the presence of mismatch, the hybrid approach still provides performance commensurate with that associated with a far greater sample support. In general, it has been observed that it is better to apply a more pessimistic CMT (greater ISL) to impart robustness and prevent undernulling [4]. The reader is encouraged to create his or her own recipe for achieving similar or perhaps better results.

5.5.2 Signal-Dependent Approach

As described previously in this chapter, a popular and effective first-stage signal-dependent and deterministic rank-reduction transformation is Doppler filtering. Since the spatial distribution of clutter is highly dependent on Doppler, there exists an opportunity to dramatically reduce the number of

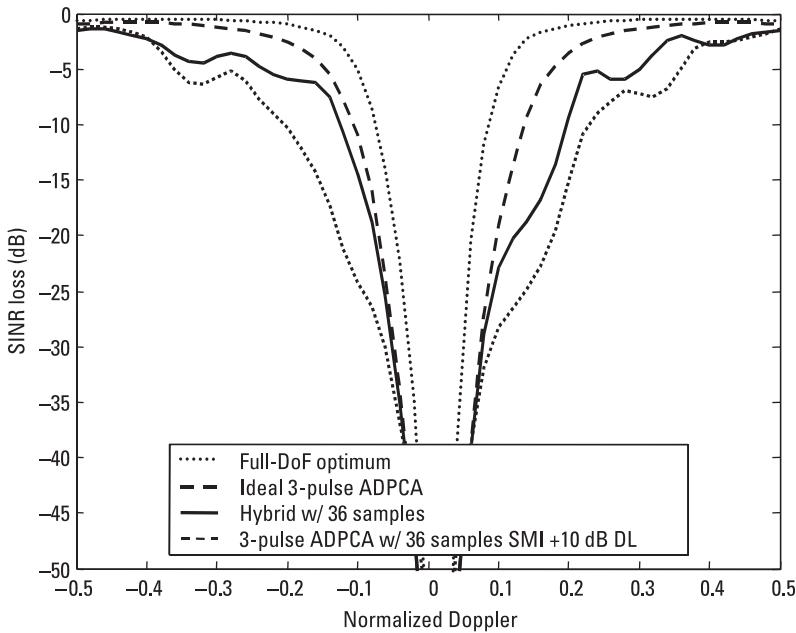


Figure 5.23 Relative performance of the hybrid three-pulse ADPCA + PC-CMT approach. Also shown is a similar approach with a 10-dB diagonally loaded SMI-CMT substituted for PC-CMT.

adaptive DoFs without significant loss [1, 2, 5]. For example, Figure 5.25 shows the eigenspectrum corresponding to the three-bin post-Doppler covariance ($3N \times 3N$) for a normalized Doppler of 0.125 (a relatively low Doppler, thus a stressing case). Note that the effective rank of the clutter has been reduced from ~ 100 (see Section 5.5.1) to ~ 25 .

At this stage there remain $(M - 2, 3N)$ -dimensional adaptive filtering problems which could be solved with either signal-independent adaptivity (e.g., SMI, PC, or PC-CMT) or signal-dependent methods—which is the course pursued in this example. If SMI is used, then $6N \sim 96$ or more samples are required to reduce SINR losses to within 3 dB or so (as discussed in Chapter 2). However, as we have already seen in the previous example, this sample support can be dramatically reduced if we exploit the fact that the effective rank of the clutter can be substantially less than the total number of DoFs (in this case $3N$).

As previously described in this chapter, a highly efficient signal-dependent adaptive rank-reduction method is the *multistage Wiener filter* (MWF). The primary design variables in the MWF implementation are the number

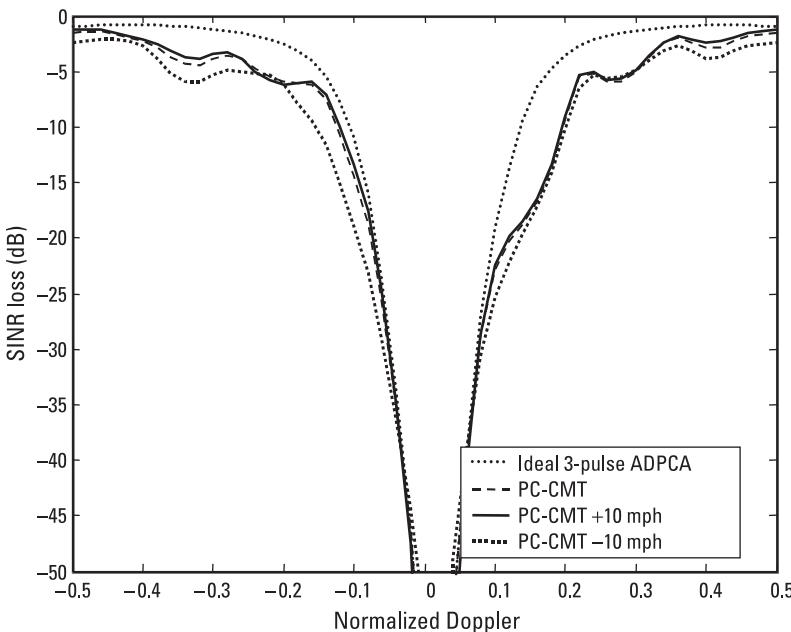


Figure 5.24 Illustration of the CMT's robustness to errors.

of stages and the sample support [12]. To illustrate the interplay of these two design parameters, Figure 5.26 plots the results of a Monte Carlo simulation of SINR loss versus the number of MWF stages and sample support for the selected normalized Doppler of 0.125. Notice that as sample support increases, the number of stages can be reduced and approaches a theoretical optimum [12]. Since our objective in this design example is to minimize sample support while preserving good performance, we might select the (*stages, samples*) combination of (20, 40) with a corresponding SINR loss of approximately 4 dB relative to the ideal (known-covariance) three-bin post-Doppler case (see Figure 5.21). Thus, again we see that it is possible to significantly reduce sample-support requirements, in this case from 512 samples to 40.

5.6 Summary

In this chapter, we surveyed key algorithms of a STAP taxonomy with an eye towards their application to stressing real-world scenarios. In particular, both data-independent (nonadaptive or deterministic) and data-dependent

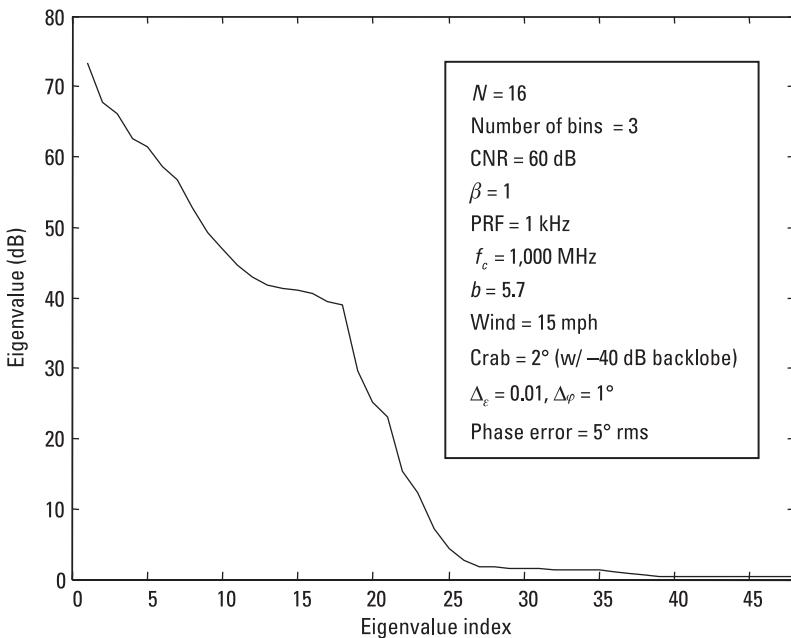


Figure 5.25 Eigenspectrum for three-bin post-Doppler rank reduction.

(or adaptive) reduced-rank STAP techniques were discussed and contrasted. The need for either deterministic or adaptive reduced-rank methods is inherent in real-world applications due to a multitude of potential sources of nonstationarity and deviation from idealized conditions—which, therefore, places a premium on minimizing sample support. Rank reduction (especially deterministic methods) can also dramatically reduce real-time computational burdens [1, 2]. Deterministic and adaptive reduced-rank methods were further delineated into either signal-dependent or -independent methods; that is, whether or not they depended explicitly on the desired target steering vector.

Illustrative design examples were then presented that highlighted the methods of creating hybrid STAP algorithms that combine various rank-reduction mechanisms. Specifically considered were both a purely signal-independent example (ADPCA with CMT-augmented PC or diagonally loaded SMI) and purely signal-dependent example (three-bin post-Doppler with MWF). The results illustrated that robust solutions were possible that preserve good performance with dramatically reduced sample-support (and computational) requirements compared with a full-DoF STAP implementa-

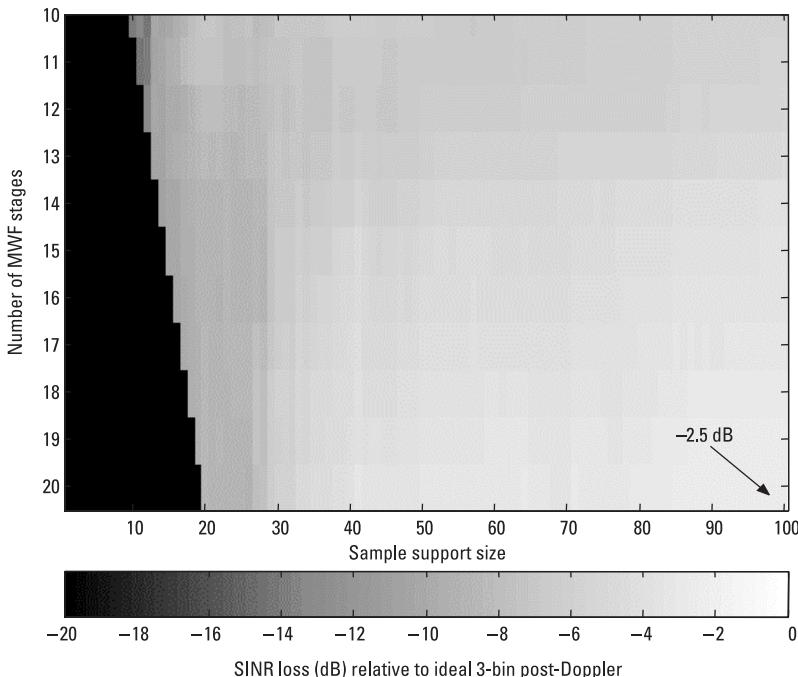


Figure 5.26 SINR loss relative to ideal (known-covariance) three-bin post-Doppler case versus the number of MWF stages and sample-support size.

tion. The statistical basis for STAP, further computational considerations, and emerging areas of research are addressed in the next chapter.

References

- [1] Ward, J., *Space-Time Adaptive Processing for Airborne Radar*, MIT Technical Report 1015, MIT Lincoln Laboratory, December 1994.
- [2] Klemm, R., *Space-Time Adaptive Processing: Principles and Applications*, London, England: IEEE Press, 1998.
- [3] Melvin, W. L., et al., “Design of Adaptive Detection Algorithms for Surveillance Radar,” *Proc. of IEEE International Radar Conference*, Alexandria, VA, May 8–12, 2000.
- [4] Guerci, J. R., and J. S. Bergin, “Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem,” *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002, pp. 152–162.
- [5] Wicks, M., H. Wang, and L. Cai, “Adaptive Array Processing for Airborne Radar,” *Proc. of International Radar Conference (Radar 92)*, Brighton, England, October 12–13, 1992, pp. 159–162.

- [6] Fuhrmann, D. R., "Application of Toeplitz Covariance Estimation to Adaptive Beamforming and Detection," *IEEE Trans. on Signal Processing*, Vol. 39, No. 10, October 1991, pp. 2194–2198.
- [7] Barton, T. A., and S. T. Smith, "Structured Covariance Estimation for Space-Time Adaptive Processing," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 5, April 2–24, 1997, pp. 3493–3496.
- [8] Barton, D., *Radar System Analysis*, Dedham, MA: Artech House, 1976.
- [9] Horn, R. A., and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, England: Cambridge University Press, 1991.
- [10] Mailloux, R. J., *Phased Array Antenna Handbook*, Norwood, MA: Artech House, 1994.
- [11] Kreithen, D., and A. O. Steinhardt, "Target Detection in Post-STAP Undernulled Clutter," *Proc. of IEEE 29th ASILOMAR Conference on Signals, Systems and Computers*, Pacific Grove, CA, Vol. 2, October 30–November 2, 1995, pp. 1203–1207.
- [12] Guerci, J. R., J. S. Goldstein, and I. S. Reed, "Optimal and Adaptive Reduced-Rank STAP," Special Section on Space-Time Adaptive Processing, *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 647–663.
- [13] Gibson, J. D., and J. L. Melsa, *Introduction to Nonparametric Detection with Applications*, Piscataway, NJ: IEEE Press, 1996.
- [14] Melvin, W. L., "Space-Time Adaptive Radar Performance in Heterogeneous Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Special Section on Space-Time Adaptive Processing, Vol. 36, No. 2, April 2000, pp. 608–613.
- [15] Haimovich, A. M., and M. Berin, "Eigenanalysis-Based Space-Time Adaptive Radar: Performance Analysis," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 33, No. 4, October 1997, pp. 1170–1179.
- [16] Strang, G., *Introduction to Linear Algebra*, 2nd ed., Wellesley, MA: Wellesley-Cambridge Press, 1998.
- [17] Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, Wellesley, MA: Wellesley-Cambridge Press, 1996.
- [18] Pierre, D. A., *Optimization Theory with Applications*, New York: Dover, 1986.
- [19] Zatman, M., "Properties of Hung-Turner Projections and Their Relationship to the Eigencanceller," *Proc. of IEEE 30th ASILOMAR Conference on Signals, Systems and Computers*, Vol. 2, Pacific Grove, CA, November 3–6, 1996, pp. 1176–1180.
- [20] Doob, J. L., *Stochastic Processes*, New York: John Wiley & Sons, 1953.
- [21] Van Trees, H. L., *Optimum Array Processing: Part IV of Detection Estimation and Modulation Theory*, New York: Wiley Interscience, 2002.
- [22] Goldstein, J. S., and I. S. Reed, "Reduced-Rank Adaptive Filtering," *IEEE Trans. on Signal Processing*, Vol. 45, No. 2, February 1997, pp. 492–496.
- [23] Goldstein, J. S., I. S. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Trans. on Information Theory*, Vol. 44, No. 7, November 1998, pp. 2943–2959.

- [24] Weippert, M., et al., "Insights from the Relationship Between the Multistage Wiener Filter and the Method of Conjugate Gradients," *Proc. of Second IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2002)*, Rosslyn, VA, August 5–6, 2002.
- [25] Pillai, S. U., and B. H. Kwon, "Forward/Backward Spatial Smoothing for Coherent Signal Identification," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 37, No. 1, January 1989, pp. 8–15.
- [26] Zatman, M., and D. Marshall, "Forward–Backward Averaging in the Presence of Array Manifold Errors," *IEEE Trans. on Antennas and Propagation*, Vol. 46, No. 11, November 1998, pp. 1700–1704.
- [27] Michels, J. H., P. Varshney, and D. Weiner, "Multichannel Detection Using a Model-Based Approach," *Proc. of IEEE ICASSP*, Vol. 5, April 14–17, 1991, pp. 3553–3556.
- [28] Guerci, J. R., and E. H. Feria, "On a Least Squares Predictive-Transform Modeling Methodology," *IEEE Trans. on Signal Processing*, Vol. 44, No. 7, July 1996, pp. 1752–1764.
- [29] Guerci, J. R., and E. H. Feria, "Application of a Least Squares Predictive-Transform Modeling Methodology to Space-Time Adaptive Array Processing," *IEEE Trans. on Signal Processing*, Vol. 44, No. 7, July 1996, pp. 1825–1833.
- [30] Guerci, J. R., "Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming," *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–986.
- [31] Billingsley, J. B., *Exponential Decay in Windblown Radar Ground Clutter Doppler Spectra: Multifrequency Measurements and Model*, Technical Report 997, MIT Lincoln Laboratory, Lexington, MA, July 29, 1996.
- [32] Haykin, S., *Communication Systems*, 3rd ed., New York: John Wiley & Sons, 1994.
- [33] Carlson, B. D., "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 24, No. 3, July 1988, pp. 397–401.
- [34] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, 2nd ed., New York: John Wiley & Sons, 1984.
- [35] Fante, R. L., E. C. Barile, and T. P. Guella, "Clutter Covariance Smoothing by Subaperture Averaging," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 30, No. 4, July 1994, pp. 941–945.
- [36] Staudaher, F. M., "An Experimental Evaluation of Adaptive RADAR," *1974 S-MTT International Microwave Symposium Digest*, Vol. 74, June 1974, pp. 21.
- [37] Titi, G. W., and D. F. Marshall, "The ARPA/NAVY Mountaintop Program: Adaptive Signal Processing for Airborne Early Warning Radar," *Proc. of 1996 IEEE ICASSP*, Vol. 2, Atlanta, GA, May 7–10, 1996, pp. 1165–1168.
- [38] Brennan, L. E., D. J. Piwinski, and F. M. Staudaher, "Space-Time Adaptive Processing Algorithm Performance Evaluation," *Proc. of IEEE Long Island Section Adaptive Antenna Systems Symposium*, November 1992, pp. 105–110.

- [39] Melvin, W. L., M. C. Wicks, and R. D. Brown, "Assessment of Multichannel Airborne Radar Measurements for Analysis and Design of Space-Time Processing Architectures and Algorithms," *Proc. of 1996 IEEE National Radar Conference*, Ann Arbor, MI, May 13–16, 1996, pp. 130–135.

6

Other Topics

6.1 Introduction

In this chapter, we briefly discuss several important topics in STAP applications and current research. The statistical basis for STAP is presented in Section 6.2, along with a brief discussion of current research trends in this area. Section 6.3 considers implementation issues. An important class of real-time QR factorization computing architectures is extended to include the CMT approach introduced in Chapter 5—thereby expanding the class of robust STAP algorithms amenable to current real-time implementation architectures. Finally, Section 6.4 introduces some concepts in the newly emerging area of research known as *knowledge-aided STAP* (KA-STAP).

6.2 Statistical Basis for STAP

The space-time beamformer derived in Chapters 2 and 3 was optimum in both a max SINR sense and, if viewed from an SLC perspective, a minimum MSE sense as well. However, nothing was said concerning its statistical (i.e., probabilistic) optimality [1].

For the additive Gaussian interference case, the max SINR beamformer can be shown to be statistically optimal. To see this, consider the basic target detection binary hypothesis testing problem [1, 2]:

$$\begin{aligned} H_0: \mathbf{x} &= \mathbf{n} && \text{(target absent)} \\ H_1: \mathbf{x} &= \mathbf{s} + \mathbf{n} && \text{(target present)} \end{aligned} \tag{6.1}$$

where $\mathbf{x} \in \mathbb{C}^{NM}$ is the received space-time array snapshot corresponding to a given range bin, $\mathbf{s} \in \mathbb{C}^{NM}$ is the target steering vector (assumed to be known, or as is done in practice, hypothesized), and $\mathbf{n} \in \mathbb{C}^{NM}$ is the total interference RV (receiver noise, clutter, jamming, and so on). Our objective is to construct a decision function that is, in some sense, optimal in terms of maximizing the probability of detection, while simultaneously maintaining the probability of false alarms.

For many sensor surveillance systems, the universally adopted framework for constructing an optimum detection test is the Neyman-Pearson criterion [1, 2], which maximizes the probability of detection for a fixed false-alarm rate. The corresponding decision criterion is the intuitively appealing likelihood ratio test given by

$$\Lambda(\mathbf{x}) \triangleq \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)} \begin{cases} > \gamma, & \text{say } H_1 \\ < \gamma, & \text{say } H_0 \end{cases} \quad (6.2)$$

where $p(\mathbf{x}|H_0)$ and $p(\mathbf{x}|H_1)$ are the conditional probabilities for the observation RV \mathbf{x} under the target-absent and target-present conditions, respectively [1, 2]. The strictly positive threshold γ is chosen to fix the probability of false alarm [1]. Note that a completely equivalent test based on the log-likelihood (due to the monotonicity of the log function for positive-valued arguments) is given by

$$\ln \Lambda(\mathbf{x}) \triangleq \ln p(\mathbf{x}|H_1) - \ln p(\mathbf{x}|H_0) \begin{cases} > \ln \gamma, & \text{say } H_1 \\ < \ln \gamma, & \text{say } H_0 \end{cases} \quad (6.3)$$

For the complex Gaussian case (see Chapter 2), we can make further progress in specifying the exact nature of the test in (6.3). Based on (6.1), $p(\mathbf{x}|H_0)$ is given by

$$p(\mathbf{x}|H_0) = \frac{1}{\pi^N |R|} e^{-\mathbf{x}' R^{-1} \mathbf{x}} \quad (6.4)$$

where R is the total $NM \times NM$ -dimensional interference covariance matrix, which is assumed to be nonsingular, and $|R|$ is its corresponding determinant. The corresponding target-present conditional probability, $p(\mathbf{x}|H_1)$, is given by

$$p(\mathbf{x}|H_0) = \frac{1}{\pi^N |R|} e^{-(\mathbf{x}-\mathbf{s})' R^{-1} (\mathbf{x}-\mathbf{s})} \quad (6.5)$$

Substituting (6.5) and (6.4) into (6.3) yields

$$\begin{aligned} \ln \Lambda(\mathbf{x}) &= -[(\mathbf{x}-\mathbf{s})' R^{-1} (\mathbf{x}-\mathbf{s}) + \mathbf{x}' R^{-1} \mathbf{x}] \\ &= 2 \operatorname{Re}\{\mathbf{s}' R^{-1} \mathbf{x}\} - \mathbf{s}' R^{-1} \mathbf{s} \end{aligned} \quad (6.6)$$

The only term in (6.6) that depends on the observed data (and thus cannot be precomputed) is the quantity

$$z = \mathbf{s}' R^{-1} \mathbf{x} \quad (6.7)$$

which we can write as

$$z = \mathbf{w}' \mathbf{x} \quad (6.8)$$

where

$$\mathbf{w} = R^{-1} \mathbf{s} \quad (6.9)$$

Thus, an equivalent optimum detector for the additive (known statistics) Gaussian case is precisely the max SINR space-time beamformer, whose residue is compared to a threshold. In the statistical literature, (6.7) is known as a sufficient statistic [1] since it can be substituted for all of the data so far as detection is concerned.

Of course in practice, as discussed in Chapter 4, the total interference covariance matrix is generally not known a priori (apart from the additional Gaussian assumption). Thus, in reality, the simple binary detection problem is a composite hypothesis testing problem [1].

A well-known approach in this case is the *generalized likelihood ratio test* (GLRT), which consists of a modified likelihood ratio test in which the unknown quantities are replaced with their MLEs under each of the assumed conditions [1, 2], that is,

$$\hat{\Lambda}(\mathbf{x}) \triangleq \frac{\max_{\hat{R}_1} p(\mathbf{x}|H_1, \hat{R}_1)}{\max_{\hat{R}_0} p(\mathbf{x}|H_0, \hat{R}_0)} \begin{cases} > \gamma, & \text{say } H_1 \\ < \gamma, & \text{say } H_0 \end{cases} \quad (6.10)$$

As mentioned in Chapter 2 for the Gaussian case under the target-absent hypothesis H_0 , under the assumption that there are $L \geq NM$ i.i.d. target-free samples available, the MLE of \hat{R}_0 is given by the SCM [3], that is,

$$\hat{R}_0 = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l' \quad (6.11)$$

where $L \geq NM$. If the presence of a target is assumed not to affect the interference statistics (assumed identical for each test cell), then \hat{R}_0 can be substituted for \hat{R}_1 .

The above highly restrictive derivations should motivate the reader to appreciate the theoretical difficulties associated with extending optimal statistical procedures to the generally nonstationary, non-Gaussian case in the presence of unknown signal-mismatch statistics [4–6], which of course is the real world. Some progress in developing detectors that preserve the CFAR property has been made for the class of *elliptically contoured distributions* (ECDs), which is a generalization of the Gaussian case [5, 6]. Also, a more effective detection scheme than the GLRT for the finite sample case has been developed by Kelly [7] (popularly known as the Kelly test).

At the time of this writing, the most common practical approach to the detection problem is a two-stage method in which the scalar residue from a robust STAP beamformer is then fed to some form of a cell-averaging CACFAR detector [8], as shown in Figure 6.1. Due to any or all of the aforementioned practical STAP shortcomings, the noise-only residue from the STAP beamformer is generally greater than the theoretical optimum (i.e., the white-noise floor) and is also nonstationary [8]. Thus, the job of

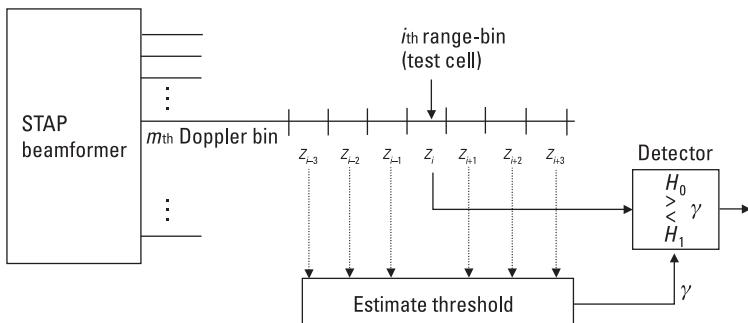


Figure 6.1 Illustration of conventional detector for a space-time beamformer. The insertion of the CACFAR circuit after the STAP processor is due to inevitable statistical mismatches leading to improperly nulled interference.

the CACFAR circuit is to estimate the local noise-floor residue (generally accomplished via a sliding window mean or median estimator [8]) and set a detection threshold based on this estimate so as to maintain a prescribed false-alarm rate. Designing a robust CACFAR algorithm is a challenging problem in and of itself. Indeed, one of the major motivations for considering the advanced minimal-sample-support STAP techniques of Chapter 5 is to lessen the burden on the CACFAR designer—since the only DoF is essentially a scalar threshold. The reader is referred to the authoritative text by Nitzberg [8] for further details on CACFAR design.

6.3 STAP Implementation

Consider a hypothetical radar with a 1-kHz PRF and a bandwidth of 10 MHz. It has approximately 10,000 unambiguous range bins to process. If each range bin has N spatial and M temporal DoFs, then a single CPI could involve as many as $10,000 NM \times NM$ covariance matrix estimation/inversion STAP operations, easily resulting in a compute throughput measured in the hundreds of GFLOPS (hundreds of billions of floating-point operations). Thus, from the very onset, a radar designer must look for very efficient parallel-processing implementations.

Many such architectures have been developed for SMI-based techniques [9, 10]. One of the more popular is the systolic array implementation of the QR-factorized SMI. It results from recasting the SMI problem in terms of an equivalent QR decomposition as follows.

First, recall the SMI formulation for the STAP weight vector \mathbf{w} :

$$\left(\sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l' \right) \mathbf{w} = (YY') \mathbf{w} = \mathbf{s} \quad (6.12)$$

where by definition, $Y = [\mathbf{x}_1, \dots, \mathbf{x}_L]$ is the $NM \times L$ data matrix (matrix of space-time array snapshots), and the $1/L$ factor has been dropped from the sample covariance for convenience since it does not affect the SINR (it could easily be absorbed into Y). A QR factorization of Y' is of the form

$$QY' = \begin{bmatrix} r \\ \emptyset \end{bmatrix} \quad (6.13)$$

where Q is an $L \times L$ unitary matrix (i.e., $Q'Q = I$), r is an $NM \times NM$ upper triangular positive-definite matrix, and \emptyset' is an $(L - NM) \times NM$ null matrix. Substituting $Q'Q$ into (6.12) yields

$$\begin{aligned} YY' \mathbf{w} &= YQ'QY' \mathbf{w} \\ &= [r' \ \emptyset'] \begin{bmatrix} r \\ \emptyset \end{bmatrix} \mathbf{w} \\ &= r'r \mathbf{w} = \mathbf{s} \end{aligned} \quad (6.14)$$

which can be solved in two steps involving straightforward backsubstitutions (due to the triangular nature of r) [10] as follows:

$$\begin{aligned} r' \mathbf{a} &= \mathbf{s} \\ r \mathbf{w} &= \mathbf{a} \end{aligned} \quad (6.15)$$

An important benefit of this formulation is that the numerical condition number (ratio of largest to smallest eigenvalue [10]) of r is the square root of the sample covariance. Additionally, the QR decomposition itself can be efficiently implemented via a sequence of Givens rotations [10].

Diagonal loading can easily be accommodated in this framework by augmenting the data matrix Y with a diagonal matrix whose entries are the square root of the desired loading factor; that is,

$$\begin{aligned} Y_{DL} &= \left[Y \mid \sqrt{\delta} \mathbf{e}_1 \dots \sqrt{\delta} \mathbf{e}_{NM} \right] \\ &= \left[Y \mid \sqrt{\delta} I \right] \end{aligned} \quad (6.16)$$

where $\mathbf{e}_1, \dots, \mathbf{e}_{NM}$ are the NM Euclidean basis vectors, and I is the NM -dimensional identity matrix. The corresponding diagonally loaded sample covariance matrix is thus given by

$$Y_{DL} Y_{DL}' = YY' + \delta I \quad (6.17)$$

where δ is the desired loading factor. From (6.17) we see that the inclusion of diagonal loading in the data domain [10] has the effect of creating more data samples—albeit artificially. This is a reasonable result since diagonal loading is a form of structured-covariance estimation and has been shown to produce results (under conditions that agree with the model) equivalent to those associated with far greater sample supports (see Chapters 2 and 5).

Interestingly, this data-augmentation approach can also be extended to include CMTs. This is of important practical value since a diagonally loaded SMI augmented with a CMT can closely approximate the highly effective PC-CMT algorithm—without the need for a difficult-to-implement eigendecomposition. To see how this is accomplished, we first need to introduce an interesting Hadamard matrix multiplication identity discovered by Khatri and Rao [11].

Theorem

Let A and B be matrices of size $n \times n$ with corresponding factorizations $A = \Psi' \Psi$ and $B = \Omega' \Omega$, where Ψ is of order $q \times n$ and Ω is of order $p \times n$. Then the Hadamard product of A and B expressed in terms of its factors is given by

$$A \odot B = (\Psi \odot \Omega)'(\Psi \odot \Omega) \quad (6.18)$$

where \odot is the Khatri-Rao product, which consists of the pairwise Kronecker products of the columns of Ψ and Ω ; that is,

$$\Psi \odot \Omega = [\psi_1 \otimes \omega_1 \quad \psi_2 \otimes \omega_2 \dots \quad \psi_n \otimes \omega_n] \quad (6.19)$$

where ψ_i and ω_i are the i th columns of Ψ and Ω , respectively.

Proof

Although algebraically straightforward, the interested reader is referred to the book by Rao for a proof [11].

Now consider the application of a CMT to the SCM

$$\hat{R} \odot T = (YY') \odot T \quad (6.20)$$

As discussed in Chapter 5, the structure of T is quite simple in practice and can be essentially fixed for a given operating environment (channel mismatch and more). Being positive-semidefinite [12], a Gram or square-root factorization of T is guaranteed [13]. Thus, (6.20) can be rewritten as

$$\begin{aligned} (YY') \odot T &= (YY') \odot (\Gamma'\Gamma) \\ &= (Z'Z) \odot (\Gamma'\Gamma) \\ &= (Z \odot \Gamma)'(Z \odot \Gamma) \end{aligned} \quad (6.21)$$

where $T = \Gamma' \Gamma$ and, by definition, $Z = Y'$. The advantage of (6.21) is that the same QR factorization and implementation architecture can be applied to the augmented data matrix $(Z \odot \Gamma)'$. A simple example will illustrate the basic procedure.

Consider a simple two-channel adaptive beamformer implementing a diagonally loaded SMI-CMT (obviously this is of no practical interest, but it is useful for illustrating the process). Assume that there are four i.i.d. samples available for estimating the SCM; that is,

$$Y = \begin{bmatrix} 2 + j2.5 & 0.9 + j4.5 & 2.3 - j3.6 & -0.6 - j1.7 \\ -1.2 + j2.6 & -1.5 - j4.7 & -3 + j2.9 & 2.3 + j1.9 \end{bmatrix} \quad (6.22)$$

from which

$$YY' = \begin{bmatrix} 52.8 & -40.4 + j9.4 \\ -40.4 + j9.4 & 58.9 \end{bmatrix} \quad (6.23)$$

To add 0-dB of diagonal loading, we augment Y with a simple 2×2 diagonal matrix whose diagonal entry is 2 [the 1/4 factor omitted from (6.23) is absorbed into the loading factor]; that is,

$$Y_{DL} = \begin{bmatrix} 2 + j2.5 & 0.9 + j4.5 & 2.3 - j3.6 & -0.6 - j1.7 & 2 & 0 \\ -1.2 + j2.6 & -1.5 - j4.7 & -3 + j2.9 & 2.3 + j1.9 & 0 & 2 \end{bmatrix} \quad (6.24)$$

Assuming a CMT of the form

$$T = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix} \quad (6.25)$$

the corresponding diagonally loaded SMI-CMT covariance matrix is given by

$$\begin{aligned} \hat{R} &= (Y_{DL} Y_{DL}') \odot T \\ &= \begin{bmatrix} 56.8 & -36.3 - j8.4 \\ -36.3 + j8.4 & 62.9 \end{bmatrix} \end{aligned} \quad (6.26)$$

In factored form, $(Y_{DL} Y'_{DL}) \odot T$ can be rewritten as

$$(Y_{DL} Y'_{DL}) \odot T = (Z'Z) \odot (\Gamma'\Gamma) \quad (6.27)$$

where $Z = Y'_{DL}$ and

$$\Gamma = \begin{bmatrix} 0.85 & 0.53 \\ 0.53 & 0.85 \end{bmatrix} \quad (6.28)$$

The corresponding 12×2 Khatri-Rao factor, $Z \odot \Gamma$, is given by

$$Z \odot \Gamma = [\mathbf{z}_1 \otimes \gamma_1 \quad \mathbf{z}_2 \otimes \gamma_2] \quad (6.29)$$

where

$$\mathbf{z}_1 = \begin{bmatrix} 2 - j2.5 \\ 0.9 - j4.5 \\ 2.3 + j3.6 \\ -0.6 + j1.7 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{z}_2 = \begin{bmatrix} -1.2 - j2.6 \\ -1.5 + j4.7 \\ -3 - j2.9 \\ -2.3 - j1.9 \\ 0 \\ 2 \end{bmatrix} \quad (6.30)$$

and γ_1 and γ_2 are the first and second columns of Γ . The diagonally loaded SMI-CMT covariance matrix corresponding to the Khatri-Rao factored form is given by

$$\begin{aligned} \hat{R} &= (Z \odot \Gamma)'(Z \odot \Gamma) \\ &= \begin{bmatrix} 56.8 & -36.3 - j8.4 \\ -36.3 + j8.4 & 62.9 \end{bmatrix} \end{aligned} \quad (6.31)$$

which is identical to (6.26).

Although highly contrived and simple, the foregoing example nonetheless illustrates that the final STAP algorithm, which includes data-domain formulations and techniques for improving its robustness (e.g., diagonal loading and/or CMTs), can be more computationally demanding than one might at first expect. This is certainly true for the far more complex structured-covariance estimation methods described in Chapter 5 and elsewhere. For

further details on real-time implementation, the reader is referred to the literature ([9, 10]).

6.4 Knowledge-Aided STAP

Much of the emphasis in this book has been on developing STAP techniques that can work in real-world environments. A moment's reflection will reveal that real-world ground clutter can be exquisitely complex—especially when man-made objects are factored in (buildings, power lines, roadways, and so forth). The presence of these large clutter discrete and abrupt discontinuities can have a significantly deleterious effect on STAP performance if not somehow accounted and/or compensated for [14–16].

One of the most exciting and far-reaching areas of research beginning to address these problems is so-called *knowledge-aided STAP* (KA-STAP) [14, 16]. The idea behind KA-STAP is simple (its implementation, however, is very challenging [16]): Databases derived from a multitude of sources [e.g., *digital terrain and elevation data* (DTED) and SAR [16]] contain information about the environment that can be extremely useful in the STAP adaptation process. For example, as described by Melvin et al. [14], if one knows where the roadways are, resolution cells corresponding to potentially competing ground traffic can be excised from the training data, which can result in dramatically improved detection performance of targets with Doppler frequencies similar to the competing background traffic [14].

While certainly a new area of research, KA-STAP is tantamount to the inclusion of priors in the statistical-estimation process [1], and it is thus a well-established statistical concept from a classical Bayesian perspective [1]. For example, if the underlying interference has a quiescent space-time pdf R_q , prior knowledge of the presence of a discrete, for example, would modify the total covariance R , resulting in $R = R_q + R_d$, where R_d is the a priori covariance, associated with the clutter discrete only (that is, assumed to be independent of the nominal background clutter). If the clutter discrete is a single-point scatterer with associated space-time steering vector \mathbf{d} and power (variance) σ_d^2 , then for the Gaussian case, the composite total interference pdf is given by

$$p(\mathbf{x}) = \frac{1}{\pi^N |R|} e^{-\mathbf{x}' R^{-1} \mathbf{x}} \quad (6.32)$$

where the total covariance R is the sum of the quiescent covariance R_q and the factor due to the discrete; that is,

$$R = R_q + \sigma_d^2 \mathbf{d} \mathbf{d}' \quad (6.33)$$

In this example, the discrete was treated as a rank-one signal. In practice, due to the generally space-time-distributed nature of the scatterer and uncertainties associated with channel mismatch, the actual covariance will generally be of full rank, that is, $\mathbf{d} \mathbf{d}' \rightarrow (\mathbf{d} \mathbf{d}') \odot T$, where T is a CMT. The interested reader is referred to the literature for further details [14, 16].

6.5 Summary

In this chapter, several important topics were surveyed. Section 6.2 introduced the statistical basis for STAP establishing the optimality of the max SINR beamformer for the additive Gaussian case and discussed current research trends in this area. Section 6.3 considered implementation issues with an emphasis on an important class of real-time QR factorization computing architectures. This popular implementation method was extended to the CMT approach introduced in Chapter 5, thereby expanding the class of robust STAP algorithms amenable to real-time implementation. Finally, Section 6.4 introduced some concepts in the newly emerging area of research known as KA-STAP.

References

- [1] Van Trees, H. L., *Detection, Estimation, and Modulation Theory: Part I*, New York: John Wiley & Sons, 1968.
- [2] DiFranco, J. V., and W. L. Rubin, *Radar Detection*, Dedham, MA: Artech House, 1980.
- [3] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, 2nd ed., New York: John Wiley & Sons, 1984.
- [4] Wegman, E. J., S. C. Schwartz, and J. B. Thomas, (eds.), *Topics in Non-Gaussian Signal Processing*, New York: Springer-Verlag, 1989.
- [5] Fang, K. T., and Y. T. Zhang, *Generalized Multivariate Analysis*, New York: Springer-Verlag, 1990.
- [6] Richmond, C. D., “PDF’s, Confidence Regions, and Relevant Statistics for a Class of Sample Covariance-Based Array Processors,” *IEEE Trans. on Signal Processing*, Vol. 44, No. 7, July 1996, pp. 1779–1793.
- [7] Kelly, E. J., *Adaptive Detection in Nonstationary Interference: Parts I and II*, Technical Report 724, MIT Lincoln Laboratory, 1985.

- [8] Nitzberg, R., *Adaptive Signal Processing for Radar*, Norwood, MA: Artech House, 1992.
- [9] Ward, C. R., P. J. Hargrave, and J. G. McWirther, "A Novel Algorithm and Architecture for Adaptive Digital Beamforming," *IEEE Trans. on Antennas and Propagation*, Vol. 34, No. 3, March 1986, pp. 338–346.
- [10] Farina, A., and L. Timmoneri, "Real-Time STAP Techniques," *Electronics and Communication Engineering Journal*, Vol. 11, No. 1, February 1999, pp. 13–22.
- [11] Rao, C. R., and M. B. Rao, *Matrix Algebra and Its Applications to Statistics and Econometrics*, Singapore: World Scientific, 1998.
- [12] Guerci, J. R., "Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming," *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, April 1999, pp. 977–985.
- [13] Horn, R. A., and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, England: Cambridge University Press, 1991.
- [14] Melvin, W. L., "Space-Time Adaptive Radar Performance in Heterogeneous Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 621–633.
- [15] Guerci, J. R., and J. S. Bergin, "Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problem," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002, pp. 152–162.
- [16] Guerci, J. R., "Knowledge-Aided Sensor Signal Processing and Expert Reasoning," *Proc. of 2002 Workshop on Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER)*, Washington, D.C., April 3, 2002 (CD ROM).

About the Author

J. R. Guerci is the deputy director of the Special Projects Office for the Defense Advanced Research Projects Agency (DARPA), where he is responsible for the development of next generation radar and sensor systems, including space-based radar. He has more than 15 years of industrial and academic experience, and has published numerous technical papers including a book chapter he coauthored in *Intelligent Signal Processing* (IEEE Press, 2001). A member of the IEEE Radar Systems Panel, he received his Ph.D. in system engineering from Polytechnic University in New York.

Index

- Adaptive 1-D processing, 31–39
Adaptive array processing, 11–42
 1-D, 31–39
 introduction, 11
 optimum spatial (angle) beamforming, 11–28
 optimum temporal (Doppler/pulse), 29–30
 summary, 42
Adaptive DPCA (ADPCA), 119
 effectiveness reduction, 122
 rank-reduction mechanism, 121
 signal-independent effective inverse, 125
 theoretical performance, 122
 three-pulse, 121, 160, 161, 162
 two-pulse, 121
 See also Displaced phase center array (DPCA)
- Adaptivity
 in nonstationary environments, 39–42
 reduced-dimension, 112
 spatial-only, 40
- Additive colored-noise, 23–28
Additive Gaussian noise (AGN), 9
Additive white noise, 19–23
Airborne MTI (AMTI) radar
 design, 54
 with two-way antenna sidelobes, 51
- See also* Moving target indicator (MTI)
- Angle-dependent channel mismatch, 92–95
 defined, 77
 errors, 78
 source, 93–94
 See also Channel mismatch
- Angle-Doppler
 airborne clutter structure, 54
 contours, 101, 102
 dependence for ULA radar, 2
 steering vectors, 65
- Angle-independent channel mismatch, 78–92
 cause, 78
 finite (nonzero) bandwidth case, 83–92
 narrowband, impact, 79
 narrowband array illustration, 78
 narrowband case, 79–83
 See also Channel mismatch
- Angle of arrival (AoA), 11, 92
- Antenna array dispersion, 95
- Antenna array misalignment, 99–103
 clutter rank and, 100
 velocity, 100
- Arrays
 manifold effects, 92
 nonlinear, 103, 104
 UHF linear/circular, 104

- Beamformers
 - 1-D optimal, 66
 - max SINR, 69, 169
 - optimum, derivation of, 18–19
 - optimum ULA, 30
 - response, 14
 - single-bin post-Doppler STAP, 123
 - space-time, 4, 5, 66, 72
 - ULA adaptive, 12
- Beamforming
 - 1-D, data-dependent methods for, 147
 - adaptive, 18
 - optimal, 27
 - optimum spatial (angle), 11–28
- Beamspace rank reduction, 127
- Brennan’s rule, 99, 103, 133
 - true clutter rank and, 100
 - violation of, 103
- Cancellation ratio (CR), 91–92
- Capon estimator, 68, 69
- Cell averaging constant false-alarm rate (CACFAR), 31
 - cell-averaging detector, 172
 - circuit insertion, 172
 - robust algorithm, 173
- Channel mismatch, 77–95
 - analysis, 78
 - angle-dependent, 77, 92–95
 - angle-independent, 78–92
 - bandwidth, 77
 - categories, 76–77
 - decorrelating, 85, 86, 88
 - decorrelating random amplitude and phase, 85
 - defined, 77
 - i*th jammer in absence of, 93
 - phase-only, 82
 - spatial, 80
 - total interference covariance in presence of, 81
- Chebyshev
 - taper function, 17
 - weighting schemes, 16
- Clutter
 - airborne, angle-Doppler structure, 54
 - cancellation, 63, 114
 - decorrelation, 98
 - Doppler power, 95, 96
 - eigenspectrum, 99
 - nonstationarity impact on SINR, 107
 - in perfect DPCA system, 117
 - point beams at (PBC) approach, 128
 - true rank, 133
- Clutter-plus-noise eigenspectrum, 103
- Clutter-to-noise ratio (CNR)
 - defined, 58
 - nonstationarity, 77
 - specified, 58
- Coherent processing interval (CPI), 3
 - DPCA filtered, 118
 - short, 99
- Constant false-alarm rate (CFAR), 7
- Covariance augmentation, 40
- Covariance estimation error, 34
- Covariance matrix
 - clutter/jamming and, 63
 - clutter-only, 60
 - CMT application to, 40
 - diagonally-loaded SMI-CMT, 176–77
 - estimation/inversion STAP operations, 173
 - jammer, spatial-only, 63
 - NM*-dimensional space-time clutter, 57–58
 - sample, 40, 46–49
 - total space-time, 58, 63
 - unknown, 46
 - weight vector formulation, 173
- Covariance matrix tapers (CMTs)
 - application, to sample covariance matrix, 40
 - application example, 41
 - application to SCM, 40, 175
 - composite, 151
 - defined, 8
 - ICM model, 96
 - Mailloux-Zatman, 41–42
 - matrix, 40
 - rank-one, 91
 - robustness to errors, 163
 - spatial-only, 40, 86, 87, 92
 - temporal, 99
- Crab angle, 100

- Crabbing, effect of, 102
Cross spectral metric (CSM), 137–43
 defined, 142
 effective covariance inverse for, 143
 using, as rank-ordering metric, 142
 See also Signal-dependent methods
- Data-dependent reduced-rank STAP, 129–49
 Hung-Turner projection, 134–36
 methods comparison, 147–49
 PC method, 129, 130–34
 signal-dependent methods, 136–47
 signal-independent methods, 130–36
- Data-independent reduced-rank STAP, 114–29
 linear transformations, 125–29
 post-Doppler, 122–25
 pre-Doppler, 114–22
- Decorrelating channel mismatch, 85
 clutter eigenspectrum impact, 90
 errors, impact of, 89
 impact on space-time clutter covariance matrix, 88
 random amplitude and phase, 85
 SINR loss impact, 90
 spatial-only, 88
 See also Channel mismatch
- Defense Advanced Research Projects Agency (DARPA) radar, 6
photograph, 7
processed clutter returns, 8
specifications, 7
- Degrees of freedom (DoFs)
 adaptive, 39
 NM , 4
 optimal space-time processing, 115
 spatiotemporal, 99
- Diagonal loading, 39, 174
 data matrix augmentation and, 174
 limitations, 39
 in PC method, 133
 SMI, 76
 SMI-CMT, 176
- Digital terrain and elevation data (DTED), 178
- Dirac delta function, 96
- Discrete Fourier transform (DFT), 14
Displaced phase center array (DPCA)
 adaptive (ADPCA), 119, 120, 121, 122
 case, 63
 clutter, 117
 concept illustration, 115
 performance, 119
 PRI requirement, 63
 processing, 61
 two-pulse, performance, 119
 two-pulse canceller, 116, 118
 two-pulse system, 117
 ULA, 115
- Doppler
 aliasing, 59
 filtering, 28, 114
 processing, 30
- Doppler shift
 distribution, 53
 induced by moving targets, 52
 normalized, 53, 55
 spatially dependent, 54
- Doppler-shifted return, 29
- Doppler steering vectors, 30, 66
 Kronecker product, 55
 M -dimensional, 56
 normalized 2-D angle pattern, 57
- Effective covariance inverse (ECI)
 mathematical derivation, 113
 RIE method, 137
 singular, 114
 for two-stage MWF, 147
- Eigendecomposition, 37
- Eigenvalues, 37, 49
 distributions, 61
 of estimated covariance, 38
 exact and estimated comparison, 36
 invariance property, 82
 as noise-floor components, 36
 total interference, 89
- Elliptically contoured distributions (ECDs), 172
- Euler's Identity, 45
- Fast Fourier transform (FFT), 14
- Forward-backward smoothing (FBS), 151, 157

- Fourier transform
 - 2-D discrete, 59
 - inverse, 96
- Generalized likelihood ratio test (GLRT),
 - 171–72
 - defined, 171
 - for finite sample case, 172
- Geometric Sum Formula, 45
- Global positioning systems (GPS), 3
- Gram-Schmidt method, 37
- Ground MTI (GMTI) radars, 52
- Hadamard
 - identity matrix, 86, 87
 - identity property, 153
 - identity row vector, 16
 - matrix multiplication identity, 80, 175
 - multiplicative property, 91
- Hadamard products, 16, 22
 - identities, 87
 - performing, 16
- Hung-Turner projection, 134–36
 - form, 134–35
 - PC method comparison, 135–36

See also Signal-independent methods
- Iceberg effect
 - characteristics, 77
 - defined, 75
 - illustrated, 105
 - interference nonstationarity and, 103–6
- Inertial navigation systems (INS), 3
- Interference covariance matrix
 - colored-noise, 82
 - estimating, 32
 - total, 59, 81

See also Covariance matrix
- Interference subspace leakage (ISL), 3
 - decorrelating, 76
 - defined, 95
 - effects, 95–99
 - mechanisms, 95
 - range-walk source, 98
 - uncorrelated phenomenon, 76
- Interference-to-noise ratio (INR), 75
- Internal clutter motion (ICM), 1
 - CMT corresponding to, 96
 - defined, 95
- impact on eigenspectrum, 96, 97
- modeling, 151
- SINR loss impact, 96, 97
- for subspace leakage, 105
- as temporal modulation effect, 96
- Inverse displaced phase center array (IDPCA), 6, 7
- Inverse Fourier transform, 96
- Jammers
 - impact on total interference eigenspectrum, 65
 - nulls, 40
 - power, 24
 - spatial only, covariance matrix, 63–64
 - steering vector, 24
 - uncorrelated, 25, 80
 - uncorrelated noise, 64
- Jammer-to-noise ratio (JNR), 25, 149
- Karhunen-Loeve representation, 151
- Karhunen-Loeve transform (KLT), 130
 - analysis filter, 131
 - analysis tree, 131–32
- Khatri-Rao
 - factor, 177
 - product, 175
- Knowledge-aided STAP (KA-STAP), 9, 178–79
 - defined, 178
 - idea behind, 178

See also STAP
- Kronecker
 - delta function, 152
 - identity, 88
 - products, 55, 56, 175
- Lagrange multiplier, 69, 131
- Likelihood function
 - convexity of, 49
 - maximizing, 47
- Linear frequency modulation (LFM), 94
- Low Earth orbit (LEO) SBR, 59
- Mailloux-Zatman CMT, 41–42
 - defined, 41
 - illustrated, 41
 - jammer notch, 42

See also Covariance matrix tapers (CMTs)

- Main lob, 14
- Maximum likelihood estimate (MLE), 46–49
derivation, 46–49
of unknown-covariance matrix, 47
- Mean-squared error (MSE), 9
scalar, minimizing, 138
space-time beamformer, 9
- Minimum detectable velocity (MDV), 106
- Minimum variance (MV), 70
- Model-based methods. *See* Structured-covariance methods
- Moving target indicator (MTI) radar, 1
airborne (AMTI), 51, 54
ground (GMTI), 52
optimum space-time processing for, 65–70
spaceborn, 52
STAP for, 3–7
two-pulse canceller, 63
- Multistage Wiener filter (MWF), 143–47, 162
advantage, 147–49
bias vector selection, 143
comparison, 147–49
defined, 143
design variables, 162–63
first stage, 144
optimal rank-one weight vector, 144
shaping of quiescent steering vector, 149
two-stage, 145–47
See also Signal-dependent methods
- Neyman-Pearson criterion, 170
- Noise
additive colored, 23–28
additive Gaussian, 9
additive white, 19–23
total, 23
white, 19–23, 24
- Noise-floor components, 36
- Nonlinear arrays, 103, 104
- Optimal beamforming, 27
- Optimum spatial (angle) beamforming, 11–28
- Optimum temporal (Doppler/pulse) processing, 29–30
- Organization, this book, 7–9
- Orthonormality, 82–83
- Parametric multichannel signal modeling, 157–58
- PC-CMT method, 76, 154
PC method comparison, 154
performance, 155
- Point beams at clutter (PBC) approach, 128
- Post-Doppler reduced-rank STAP, 122–25
multibin, 127
single-bin, 126
three-bin, 164–65
- Power selected training (PST), 52
- Power spectral densities (psd), 58
clutter Doppler, 96
illustrated, 60
- Principal components (PC) method, 130–34
comparison, 147–49
defined, 129
performance, 133, 134
in practical applications, 133–34
rank- k , 133
rank reduction, achieving, 132–33
straight SMI/diagonally loaded SMI and, 133, 134
See also Signal-independent methods
- Pulse repetition frequency (PRF), 3
- Pulse repetition interval (PRI), 3
DPCA requirement, 63
staggered approach, performance, 129
- Range walk, 98
- Rank linear transformation (RLT), 113
data-dependent, 114
data-independent, 114
pre-Doppler data-independent, 114
- Rank-ordering metric (ROM), 130, 136
- Rank reduction
beamspace, 127
data dependent, 129–49
data independent, 114–29
post-Doppler, 122–25
pre-Doppler, 114–22

- Rank reduction (continued)
 - principal components, 129, 130–34
 - signal-dependent methods, 136–47
 - signal-independent methods, 130–36
- Relative importance of eigenbeam (RIE)
 - method, 136–37
 - comparison, 147–49
 - ECI, 137
- Riemann sum approximation, 56, 88
- Sample covariance matrix (SCM)
 - CMT application to, 40, 175
 - estimating, 176
 - maximum likelihood, derivation of, 46–49
 - $N \times N$, 47
- Sample matrix inverse (SMI), 9
 - diagonal-loading, 76
 - STAP weight vector formulation, 173
 - See also* SMI-CMT
- Schur Product theorem, 91
- Schwartz’s inequality, 13, 23, 144
- Self-nulling, 33
- Sidelobe canceller (SLC), 137
 - effective covariance inverse for, 139
 - space-time, structure, 138
- Sidelobes
 - defined, 15
 - Doppler, 66
 - range, reduction, 94
- Signal-dependent methods, 136–47
 - cross spectral metric (CSM), 137–43
 - illustrative design example, 161–63
 - multistage Wiener filter (MWF), 143–47
 - relative importance of eigenbeam (RIE) method, 136–37
 - See also* Data-dependent reduced-rank STAP
- Signal-independent methods, 130–36
 - Hung-Turner projection, 134–35
 - illustrative design example, 159–61
 - PC, 129, 130–34
 - rank-ordering metric (ROM), 130
 - See also* Data-dependent reduced-rank STAP
- Signal-interference-to-noise ratio (SINR)
 - clutter nonstationarity impact on, 107
 - Doppler vs., 68
 - dropout, 67
 - loss, 28, 39, 67, 80, 163
 - loss curves, 159
 - loss relative to ideal three-bin post-Doppler, 165
 - max, 141
 - max beamformer, 69, 169
 - optimal, 68
 - performance impact, 91
 - for six-jammer example, 147, 148
 - space-time beamformer, 72
 - unitary matrix and, 140
- Signal-to-noise ratio (SNR), 4
 - defined, 19
 - loss, 21, 83
 - loss ratio, 79
 - maximizing, 19
 - optimum, 20, 21–22, 79
 - output, 20–21
 - realized, 79
 - single-channel, 21
 - tapered output, 21–22
 - white-noise, 24
- Six jammer case
 - angular distribution, 25
 - covariance, 36
 - optimum beam pattern, 27
 - SINR performance for, 147, 148
- SMI-CMT, 76, 155
 - covariance matrix, 176, 177
 - diagonally-loaded, 176, 177
- Space-based radar (SBR), 59
- Space-time adaptive processing. *See* STAP
- Space-time beamformers
 - illustrated, 66
 - optimal SINR, 72
 - response for sidelooking ULA, 5
- Space-time mismatch taper, 80
- Spatial-only CMT, 87, 92
 - adaptivity, 40
 - elements, calculating, 86
 - See also* Covariance matrix tapers (CMTs)

- STAP, 51–72
beamformer, 3
bistatic, x , 77
circular array applications, 40
for clutter cancellation, 114
CPI dimension NM , 112
data-dependent reduced-rank, 129–49
data-independent reduced-rank,
 114–29
efficiency, 9
full-DoF, 119
implementation, 173–78
introduction, 51–52
knowledge-aided (KA-STAP), 9,
 178–79
for MTI radar, 3–7
multibin post-Doppler, 127
need in moving target indicator radar,
 1–3
nonlinearity impact on, 105
optimum, for MTI radar, 65–70
performance factors, 75–107
post-Doppler reduced-rank, 122–25
pre-Doppler reduced-rank, 114–22
real-world systems, 75
single-bin post-Doppler, 126
statistical basis for, 169–73
taxonomy, 111
- STAP algorithms
comparison, 75
PC-CMT, 76, 154, 155
SMI-CMT, 76, 155
taxonomy, 112
- Steering vectors
angle-Doppler, 65
Doppler, 30, 55, 56, 57, 66
space-time, 55, 117
spatial, 55
Vandermonde, 157
- Structured-covariance methods, 149–58
applicability, 149
CMT, 151–56
defined, 149
FBS, 157
parametric multichannel signal
 modeling, 157–58
TBT, 156–57
- Synthetic aperture radar (SAR), x
- Three-bin post-Doppler rank reduction,
 164–65
 eigenspectrum for, 164
 SINR loss relative to, 165
- Three-pulse ADPCA, 121, 160
 eigenspectrum, 161
 relative performance, 162
 See also Adaptive DPCA (ADPCA)
- Toeplitz-block-Toeplitz (TBT), 151,
 156–57
 defined, 156
 matrix, fitting, 157
 structure, 113, 156
- Track-while-scan (TWS), 103
- Two-stage MWF, 145–47
 ECI for, 147
 illustrated, 145
 NM-dimensional space-time weight
 vector for, 146
 optimal SLC weight for, 145–46
 See also Multistage Wiener filter
 (MWF)
- Uniform linear array (ULA)
adaptive beamformer, 12
antenna array response, 44–45
DPCA, 115
 N -element, 3
optimum beamformer, 30
side-looking, 60
white-noise case, 20
- Uniform linear array (ULA) radar, 1
angle-Doppler dependence for, 2
clutter iso-range ring for, 2
 N -element, 53
- Vandermonde steering vectors, 157
- Vandermonde structure, 62
- White noise
additive, 19–23
defined, 19
matched filter solution, 20
SNR, 24
ULA, 20
- Wide sense stationary (wss) process, 57
- Zero-Doppler filter, 52