

Second Edition

Boualem Boashash

Editor

Time-Frequency Signal Analysis and Processing

A Comprehensive Reference



Time-Frequency Signal Analysis and Processing

EURASIP and Academic Press Series in Signal and Image Processing

Series Advisor

Sergios Theodoridis

Titles include:

Duhamel and Kieffer

Joint Source-Channel Decoding, ISBN 9780123744494

Thiran, Marqués and Bourlard

Multimodal Signal Processing, ISBN 9780123748256

Boashash

Time-Frequency Signal Analysis and Processing, ISBN 9780123984999

Time-Frequency Signal Analysis and Processing

A Comprehensive Reference

Edited by

Boualem Boashash



AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
125 London Wall, London, EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016, 2003 Elsevier Ltd. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

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

British Library Cataloguing in Publication Data

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

ISBN:978-0-12-398499-9

For information on all Academic Press publications
visit our website at <http://store.elsevier.com/>

Printed and bound in the United Kingdom



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Todd Green

Acquisitions Editor: Tim Pitts

Editorial Project Manager: Charlie Kent

Production Project Manager: Lisa Jones

Designer: Greg Harris

*To my sons
Jamil Mehdi and Issam Alam,
my wife Nadia,
my late spouse Theresa,
and my whole family.*

This page intentionally left blank

Contents

| | |
|------------------------------------|--------|
| Preface to the Second Edition..... | xxxi |
| Preface to the First Edition | xxxiii |
| List of Contributors..... | xxxv |
| Abbreviations | xli |
| Book Standard Notations | xlvi |

PART I A TIME-FREQUENCY TUTORIAL

| | |
|---|-----------|
| CHAPTER I The Time-Frequency Approach: Essence and Terminology | 3 |
| <i>B. Boashash</i> | |
| Introduction and Overview | 3 |
| I.1 Traditional Signal Representations | 4 |
| I.1.1 Origin and Types of Signals | 4 |
| I.1.2 Time-Domain Representation of Typical Signals | 5 |
| I.1.3 Frequency Domain Representation of Typical Signals..... | 8 |
| I.1.4 Positive and Negative Frequencies | 10 |
| I.1.5 Relationship Between t Domain and f Domain Representations | 11 |
| I.1.6 Limitation of t Domain or f Domain Representations..... | 13 |
| I.2 Joint Time-Frequency Representations | 14 |
| I.2.1 Uncovering Hidden Information Using (t,f) Representations..... | 14 |
| I.2.2 Concept and Meaning of a (t,f) Representation | 17 |
| I.2.3 Signal Characteristics in (t,f) Representations | 18 |
| I.2.4 Historical Note: The Editor's view | 22 |
| I.3 An Overview of Remaining Chapters and Key (t,f) Issues..... | 23 |
| I.4 Conclusions and Advice to the Reader..... | 27 |
| I.5 Additional Exercises..... | 27 |
| References | 28 |
| CHAPTER 1 Time-Frequency and Instantaneous Frequency Concepts | 31 |
| <i>B. Boashash</i> | |
| Introduction and Overview..... | 31 |
| 1.1 The Benefits of Time-Frequency Distributions (TFDs)..... | 32 |
| 1.1.1 Representation of Three Real-Life Signals..... | 32 |
| 1.1.2 Time-Domain Representation | 34 |
| 1.1.3 Spectral Characteristics of Typical signals | 35 |
| 1.1.4 Joint Time-Frequency Representation..... | 39 |
| 1.1.5 Desirable Characteristics of a TFD..... | 39 |

| | | |
|------------|--|----|
| 1.2 | Signal Formulations and Characteristics in the (t,f) Domain..... | 43 |
| 1.2.1 | Examples of Signal Models..... | 43 |
| 1.2.2 | Analytic Signals | 43 |
| 1.2.3 | Hilbert Transform (HT) and Analytic Associate..... | 44 |
| 1.2.4 | Duration, Bandwidth, and BT Product..... | 45 |
| 1.2.5 | Asymptotic Signals | 48 |
| 1.2.6 | Monocomponent vs Multicomponent Signals..... | 49 |
| 1.3 | Instantaneous Frequency (IF) and Spectral Delay (SD)..... | 49 |
| 1.3.1 | Instantaneous Frequency..... | 49 |
| 1.3.2 | Spectral Delay | 51 |
| 1.3.3 | Mean IF and Group delay | 54 |
| 1.3.4 | Relaxation Time, Dynamic Bandwidth..... | 56 |
| 1.4 | Defining Amplitude, Phase, and IF Using the Analytic Signal..... | 58 |
| 1.4.1 | Amplitude and Phase Formulation for the AM/FM Model..... | 58 |
| 1.4.2 | Analytic Signal, HT, and AM/FM Model..... | 58 |
| 1.4.3 | Characterization of AM/FM Canonical Pairs $[a(t), \phi(t)]$ | 59 |
| 1.5 | Summary and Discussion..... | 60 |
| 1.6 | Additional Exercises..... | 60 |
| | References | 61 |

CHAPTER 2 Heuristic Formulation of Time-Frequency Distributions 65

B. Boashash

| | | |
|------------|--|----|
| | Introduction and Overview..... | 65 |
| 2.1 | Method 1: The Wigner-Ville Distribution | 66 |
| 2.1.1 | Knife-Edge IF Indication | 66 |
| 2.1.2 | Formulation of the Signal Kernel..... | 66 |
| 2.1.3 | The Wigner Distribution..... | 67 |
| 2.1.4 | The Wigner-Ville Distribution (WVD)..... | 69 |
| 2.2 | Method 2: Time-Varying Power Spectral Density..... | 74 |
| 2.2.1 | Spectra of Nonstationary Random Processes..... | 74 |
| 2.2.2 | Estimating the Wigner-Ville Spectrum..... | 75 |
| 2.3 | Method 3: Windowed FT and Spectrogram..... | 76 |
| 2.3.1 | STFT, Inverse STFT, and Spectrogram..... | 76 |
| 2.3.2 | Optimal Window Length of the Spectrogram..... | 78 |
| 2.3.3 | STFT vs Gabor Transform | 79 |
| 2.4 | Method 4: Filtered Function of Time | 81 |
| 2.4.1 | Filter Banks and the Sonograph..... | 81 |
| 2.4.2 | Equivalence to Spectrogram | 81 |
| 2.5 | Method 5: Instantaneous Power Spectra..... | 82 |
| 2.5.1 | Time-Dependent Spectrum and Page Distribution..... | 82 |
| 2.6 | Method 6: Energy Density | 83 |

| | | |
|---|--|------------|
| 2.6.1 | The Complex (t, f) Energy Density | 83 |
| 2.6.2 | The Real Energy (t, f) Density | 84 |
| 2.6.3 | Windowed Rihaczek and Levin Distributions | 85 |
| 2.7 | Relationship Between TFDs Using Time-Lag Kernels..... | 85 |
| 2.7.1 | Spectrogram and WVD | 86 |
| 2.7.2 | Rihaczek and Levin Distributions and Their Time-Lag Kernels..... | 87 |
| 2.7.3 | Page Distribution and Its Time-Lag Kernel | 89 |
| 2.7.4 | Relationship Between the WVD and Other TFDs..... | 90 |
| 2.7.5 | MBD, EMBD, and Other Quadratic TFDs of Interest..... | 90 |
| 2.7.6 | Time-Scale and Wavelet Transforms..... | 92 |
| 2.7.7 | Links Among Linear, Quadratic TFDs, and Time-Scale Methods..... | 97 |
| 2.8 | Summary and Discussion..... | 99 |
| 2.9 | Additional Exercises..... | 99 |
| | References..... | 100 |
| CHAPTER 3 Theory and Design of High-Resolution Quadratic TFDs..... | | 103 |
| <i>B. Boashash</i> | | |
| | Introduction and Overview | 103 |
| 3.1 | The Wigner-Ville Distribution..... | 104 |
| 3.1.1 | Properties of the WVD | 104 |
| 3.1.2 | Limitations of the WVD | 106 |
| 3.2 | Formulations of Quadratic TFDs (QTFDs)..... | 110 |
| 3.2.1 | Time-Lag, Doppler-Frequency, and Other Formulations | 110 |
| 3.2.2 | Time-Frequency Formulation..... | 112 |
| 3.2.3 | Doppler-Lag Formulation and TFD Design | 114 |
| 3.2.4 | Doppler-Frequency Formulation and the SCF..... | 115 |
| 3.2.5 | Examples of Simple TFD Formulations..... | 117 |
| 3.3 | Properties of Quadratic TFDs (QTFDs)..... | 119 |
| 3.3.1 | Desirable Properties | 119 |
| 3.3.2 | TFD Properties and Equivalent Kernel Constraints..... | 121 |
| 3.3.3 | Basic Design Methods for High-Resolution TFDs | 121 |
| 3.3.4 | Advanced Design of High-Resolution TFDs | 123 |
| 3.4 | Positivity of QTFDs: Examples and Conditions | 128 |
| 3.4.1 | The Spectrogram as an Example of Positive TFD | 128 |
| 3.4.2 | The Sonogram Filter Bank..... | 129 |
| 3.5 | QTFDs, Ambiguity Function, and Radar | 129 |
| 3.5.1 | Radar AF: Where Is the Ambiguity?..... | 129 |
| 3.5.2 | WVDs, AFs, and Inner-Product Invariance..... | 130 |
| 3.6 | Concluding the Tutorial..... | 132 |

| | | |
|------------|---|------------|
| 3.6.1 | Summary, Discussion, Conclusions, and Perspectives..... | 132 |
| 3.6.2 | Relationship to Parts II, III, IV, and V | 133 |
| 3.7 | Additional Exercises | 134 |
| | References..... | 135 |

PART II FUNDAMENTAL PRINCIPLES OF TFSAP

CHAPTER 4 Advanced Time-Frequency Signal and System Analysis 141

| | | |
|------------|--|------------|
| | Introduction and Overview | 141 |
| 4.1 | Relationships Between Quadratic TFDs and Time-Scale Representations | 142 |
| | <i>B. Boashash and S. Touati</i> | |
| 4.1.1 | Introduction | 142 |
| 4.1.2 | Linear or Atomic Decompositions..... | 142 |
| 4.1.3 | Energy Quadratic TFDs and TSEDs | 145 |
| 4.1.4 | QTFDs, Wavelet Transform, and Scalogram: Relationships | 148 |
| 4.1.5 | Summary and Conclusions | 149 |
| 4.1.6 | Proofs | 149 |
| 4.2 | Cross-terms and Localization in Quadratic Time-Frequency Distributions..... | 151 |
| | <i>P. Flandrin</i> | |
| 4.2.1 | Identifying Cross-Terms | 151 |
| 4.2.2 | Reducing Cross-Terms | 153 |
| 4.2.3 | Cross-Terms and Localization..... | 155 |
| 4.2.4 | Summary and Conclusions | 157 |
| 4.3 | The Covariance Theory of Time-Frequency Analysis..... | 158 |
| | <i>F. Hlawatsch and G. Tauböck</i> | |
| 4.3.1 | The Covariance Principle | 158 |
| 4.3.2 | Time-Frequency Displacement Operators | 159 |
| 4.3.3 | Covariant Signal Representations: Group Domain | 160 |
| 4.3.4 | The Displacement Function | 161 |
| 4.3.5 | Covariant Signal Representations: Time-Frequency Domain | 164 |
| 4.3.6 | Example: Hyperbolic Wavelet Transform and Hyperbolic Class..... | 166 |
| 4.3.7 | Summary and Conclusions | 168 |
| 4.4 | Uncertainty in Time-Frequency Analysis..... | 168 |
| | <i>P.M. Oliveira and V. Barroso</i> | |
| 4.4.1 | The Time-Frequency Plane..... | 170 |
| 4.4.2 | Information and Spectral Estimation..... | 170 |
| 4.4.3 | Summary and Conclusions | 175 |
| 4.5 | Generalized TFRs via Unitary Transforms | 175 |
| | <i>R. Baraniuk</i> | |
| 4.5.1 | Three Approaches to Joint Distributions | 176 |
| 4.5.2 | Linking Signal and Axis Transformations | 177 |

| | | |
|-------------|--|------------|
| 4.5.3 | Examples of Linked Signal/Axis Transformations | 178 |
| 4.5.4 | Summary and Conclusions | 179 |
| 4.6 | Measurement Concepts in the Time-Frequency Plane..... | 179 |
| | <i>G. Jones and B. Boashash</i> | |
| 4.6.1 | Quadratic TFDs and Spectrogram | 180 |
| 4.6.2 | Signal Measurements in the (t,f) Domain..... | 180 |
| 4.6.3 | Properties and Interpretation of Local (t,f) Measurements..... | 182 |
| 4.6.4 | Application of Local (t,f) Measurements to Spectrograms | 183 |
| 4.6.5 | Example: Result for an Adaptive Spectrogram..... | 184 |
| 4.6.6 | Summary and Conclusions | 184 |
| 4.7 | Time-Frequency Transfer Function Calculus of Linear Time-Varying Systems | 185 |
| | <i>G. Matz and F. Hlawatsch</i> | |
| 4.7.1 | Linear Time-Varying Systems | 185 |
| 4.7.2 | The Generalized Weyl Symbol..... | 186 |
| 4.7.3 | The Generalized Spreading Function | 187 |
| 4.7.4 | Underspread LTV Systems | 188 |
| 4.7.5 | Time-Frequency Transfer Function Calculus | 190 |
| 4.7.6 | Summary and Conclusions | 192 |
| 4.8 | Wigner Distribution and Fractional Fourier Transform..... | 193 |
| | <i>T. Alieva and M.J. Bastiaans</i> | |
| 4.8.1 | Time-Frequency Representations | 193 |
| 4.8.2 | Wigner Distribution and Ambiguity Function | 194 |
| 4.8.3 | Fractional Fourier Transform | 194 |
| 4.8.4 | Fractional Power Spectrum and Radon-Wigner Transform..... | 196 |
| 4.8.5 | Fractional Fourier Transform Moments | 197 |
| 4.8.6 | Applications | 199 |
| 4.8.7 | Summary and Conclusions | 200 |
| 4.9 | A Time-Frequency Perspective on Systems: From SISO to MIMO..... | 200 |
| | <i>Lorenzo Galleani</i> | |
| 4.9.1 | SISO Systems..... | 201 |
| 4.9.2 | MIMO Systems..... | 205 |
| 4.9.3 | Summary and Conclusions | 209 |
| 4.10 | Teager-Kaiser Energy Operators in Time-Frequency Analysis..... | 209 |
| | <i>A.-O. Boudraa and F. Salzenstein</i> | |
| 4.10.1 | Concept of Energy | 209 |
| 4.10.2 | Teager-Kaiser Energy Operator | 210 |
| 4.10.3 | Cross Teager-Kaiser Energy Operators..... | 213 |
| 4.10.4 | Link of Ψ_B to Dynamic Signal | 215 |
| 4.10.5 | Summary and Conclusions | 217 |
| 4.11 | Gabor Spectrogram, WVD, and Energy Atoms | 217 |
| | <i>B. Boashash, Md A. Awal, and S. Dong</i> | |

| | | |
|--------|--|------------|
| 4.11.1 | WVD Properties and Gaussian Functions..... | 217 |
| 4.11.2 | Gabor Decomposition of the WVD..... | 218 |
| 4.11.3 | Gabor Spectrogram..... | 219 |
| 4.11.4 | Tests on Simulated and Real Signals..... | 220 |
| 4.11.5 | Relationship Between Gabor Spectrogram and Matching Pursuit..... | 222 |
| 4.11.6 | EEG Burst-Suppression Detection Using Gabor Spectrogram | 223 |
| 4.11.7 | Summary and Conclusions | 224 |
| 4.11.8 | Proof of Eq. (4.11.11)..... | 225 |
| 4.12 | Emperical Mode Decomposition and Hilbert Spectrum..... | 225 |
| | <i>B. Boashash and A. Akan</i> | |
| 4.12.1 | Empirical Mode Decomposition | 225 |
| 4.12.2 | Hilbert Spectral Analysis | 227 |
| 4.12.3 | Relationship Between the EMD and QTDFs..... | 229 |
| 4.12.4 | Relationship Between EMD and Other (t,f) /Time-Scale Representations..... | 230 |
| 4.12.5 | Summary and Conclusions | 230 |
| | References..... | 231 |
| | CHAPTER 5 Advanced Design and Specifications of TFDs | 237 |
| | Introduction and Overview | 237 |
| 5.1 | Ambiguity Functions | 238 |
| | <i>P. Flandrin</i> | |
| 5.1.1 | The Radar/Sonar Problem | 238 |
| 5.1.2 | Definitions of Ambiguity Functions..... | 238 |
| 5.1.3 | Properties of Narrowband Ambiguity Functions | 240 |
| 5.1.4 | Remarks on Wideband Ambiguity Functions..... | 243 |
| 5.1.5 | Summary and Conclusions | 243 |
| 5.2 | Reduced Interference Time-Frequency Distributions..... | 244 |
| | <i>W.J. Williams</i> | |
| 5.2.1 | Nonstationarity, Resolution, and Interference..... | 244 |
| 5.2.2 | The Reduced Interference Distribution..... | 244 |
| 5.2.3 | Kernel Selection for Reduced Interference Distribution (RID)..... | 246 |
| 5.2.4 | Comparisons of TFD Results | 250 |
| 5.2.5 | Summary and Conclusions | 250 |
| 5.3 | Adaptive Time-Frequency Analysis..... | 252 |
| | <i>R.G. Baraniuk and D.L. Jones</i> | |
| 5.3.1 | Adaptive Short-Time Fourier Transforms..... | 252 |
| 5.3.2 | Adaptive Quadratic (t,f) Representations..... | 254 |
| 5.3.3 | Summary and Conclusions | 256 |

| | | |
|---|---|-----|
| 5.4 | Polynomial Wigner-Ville Distributions..... | 257 |
| <i>B. Boashash and G.R. Putland</i> | | |
| 5.4.1 | Polynomial FM Signals | 257 |
| 5.4.2 | Principles of Formulation of PWVDs..... | 257 |
| 5.4.3 | IF Estimates With Zero Deterministic Bias..... | 259 |
| 5.4.4 | Calculation of Coefficients | 260 |
| 5.4.5 | Examples..... | 261 |
| 5.4.6 | Summary and Conclusions | 263 |
| 5.5 | Design of Polynomial TFDs, Time-Varying Higher Order Spectra, and Complex-Lag TFDs | 263 |
| <i>B. Boashash</i> | | |
| 5.5.1 | Introduction | 263 |
| 5.5.2 | Foundations of PTFDs..... | 263 |
| 5.5.3 | Multicomponent Signals and PTFDs..... | 264 |
| 5.5.4 | IF Estimation of PFM Signals Using PTFDs and TVHOS | 264 |
| 5.5.5 | Complex Argument TFDs for Polynomial FM Signals Analysis | 264 |
| 5.5.6 | Multicomponent Signal Analysis and Implementation Issues | 265 |
| 5.5.7 | Summary and Conclusions | 266 |
| 5.6 | Time-Frequency Representations Covariant to Spectral Delay Shifts | 266 |
| <i>A. Papandreou-Suppappola</i> | | |
| 5.6.1 | Spectral Delay Shift Covariance Property | 266 |
| 5.6.2 | Classes of SDS-Covariant QTFRs | 270 |
| 5.6.3 | Simulation Example..... | 273 |
| 5.6.4 | Summary and Conclusions | 273 |
| 5.7 | Design of High-Resolution Quadratic TFDs with Separable Kernels..... | 274 |
| <i>B. Boashash and G.R. Putland</i> | | |
| 5.7.1 | RIDs and Quadratic TFDs | 274 |
| 5.7.2 | Separable Kernel Formulations..... | 275 |
| 5.7.3 | Properties | 277 |
| 5.7.4 | Design Examples of Separable-Kernel TFDs | 279 |
| 5.7.5 | Results and Discussion | 280 |
| 5.7.6 | Summary and Conclusions | 283 |
| 5.8 | Fractional Fourier Transform and Generalized-Marginal TFDs..... | 284 |
| <i>X.-G. Xia</i> | | |
| 5.8.1 | Fractional Fourier Transform | 284 |
| 5.8.2 | Generalized-Marginal Time-Frequency Distribution..... | 285 |
| 5.8.3 | Summary and Conclusions | 288 |
| 5.9 | Design of High-Resolution Quadratic TFDs With Multidirectional Kernels..... | 288 |
| <i>B. Boashash, B.K. Jawad, and N.A. Khan</i> | | |
| 5.9.1 | Limitations of Separable Kernels for TFD Resolution..... | 288 |

| | | |
|-------------|---|------------|
| 5.9.2 | Location of Auto-Terms and Cross-Terms..... | 289 |
| 5.9.3 | Defining a High-Resolution Multidirectional TFD..... | 295 |
| 5.9.4 | MDD Performance: Results and Discussions..... | 296 |
| 5.9.5 | Summary and Conclusions | 298 |
| 5.10 | Adaptive Directional Time-Frequency Distributions..... | 299 |
| | <i>N.A. Khan, B. Boashash, and E. Sejdić</i> | |
| 5.10.1 | The Need for Adaptive Directional Kernels..... | 299 |
| 5.10.2 | Adaptive Directional (t,f) Distributions..... | 301 |
| 5.10.3 | Illustrative Examples..... | 304 |
| 5.10.4 | Summary and Conclusions | 305 |
| 5.11 | The Modified S-Transform Applied to Time-Frequency Phase Synchrony | 307 |
| | <i>S. Assous, B. Boashash, and S. Ventosa</i> | |
| 5.11.1 | Phase Synchrony Measures..... | 307 |
| 5.11.2 | Phase Synchrony in Time-Frequency Representations | 308 |
| 5.11.3 | Cross-MST Applied to Phase Synchrony | 312 |
| 5.11.4 | Summary and Conclusions | 317 |
| 5.12 | Kernel Conditions for the Positivity of Quadratic TFDs | 317 |
| | <i>B. Boashash</i> | |
| 5.12.1 | Introduction | 317 |
| 5.12.2 | Necessary and Sufficient Condition | 317 |
| 5.12.3 | IF Implications, Positivity, and (t,f) Resolution | 321 |
| 5.12.4 | Summary and Conclusions | 321 |
| 5.12.5 | Proofs | 322 |
| | References..... | 323 |

PART III TIME-FREQUENCY METHODS

| | | |
|------------------|--|------------|
| CHAPTER 6 | Advanced Implementation and Realization of TFDs | 331 |
| | Introduction and Overview | 331 |
| 6.1 | Discrete Time-Frequency Distributions | 332 |
| | <i>B. Boashash and G.R. Putland</i> | |
| 6.1.1 | The Discrete Wigner-Ville Distribution | 332 |
| 6.1.2 | The Windowed DWVD..... | 334 |
| 6.1.3 | The Discrete Quadratic TFD | 335 |
| 6.1.4 | Desirable Properties: Kernel Constraints..... | 339 |
| 6.1.5 | Examples..... | 340 |
| 6.1.6 | Summary and Conclusions | 342 |
| 6.2 | Quadratic and Higher Order Time-Frequency Analysis Based on the STFT..... | 342 |
| | <i>L.J. Stanković</i> | |
| 6.2.1 | STFT-Based Realization of the Quadratic Representations..... | 343 |
| 6.2.2 | Discrete Realization of the Basic S-Method Form..... | 345 |

| | | |
|------------|--|-----|
| 6.2.3 | STFT-Based Realization of Higher Order Representations | 348 |
| 6.2.4 | Summary and Conclusions | 351 |
| 6.3 | Gabor's Signal Expansion for a Nonorthogonal Sampling Geometry | 351 |
| | <i>M.J. Bastiaans and A.J. van Leest</i> | |
| 6.3.1 | Historical Perspective | 351 |
| 6.3.2 | Gabor's Signal Expansion on a Rectangular Lattice | 352 |
| 6.3.3 | Fourier Transform and Zak Transform | 352 |
| 6.3.4 | Rational Oversampling | 354 |
| 6.3.5 | Nonorthogonal Sampling | 355 |
| 6.3.6 | Gabor's Signal Expansion on a Nonorthogonal Lattice | 356 |
| 6.3.7 | Summary and Conclusions | 357 |
| 6.4 | Spectrogram Decompositions of Time-Frequency Distributions | 358 |
| | <i>W.J. Williams and S. Aviyente</i> | |
| 6.4.1 | Decomposition-Based Approaches | 358 |
| 6.4.2 | Decomposition of Time-Frequency Kernels | 359 |
| 6.4.3 | Development of the Method | 359 |
| 6.4.4 | Wigner Example | 360 |
| 6.4.5 | Optimum Orthogonal Windows | 361 |
| 6.4.6 | Kernel Decomposition Results | 362 |
| 6.4.7 | Summary and Conclusions | 364 |
| 6.5 | Computation of Discrete Quadratic TFDs | 364 |
| | <i>B. Boashash and G.R. Putland</i> | |
| 6.5.1 | General Computational Procedure | 364 |
| 6.5.2 | Computation of the Analytic Signal | 365 |
| 6.5.3 | Real-Time Computation of TFDs | 366 |
| 6.5.4 | Computational Approximations for Discrete-Time Kernels | 366 |
| 6.5.5 | Special Case: Direct Form of the Discrete Spectrogram | 369 |
| 6.5.6 | Sample Code Fragments | 369 |
| 6.5.7 | The <i>TFSAP</i> Toolbox | 373 |
| 6.5.8 | Summary and Conclusions | 373 |
| 6.6 | Memory-Efficient Algorithms for Quadratic TFDs | 374 |
| | <i>J.M. O'Toole and B. Boashash</i> | |
| 6.6.1 | Algorithms to Reduce Computation and Memory | 374 |
| 6.6.2 | TFD Algorithms for Specific Kernel Types | 375 |
| 6.6.3 | Further Memory Reduction: Decimated TFDs | 378 |
| 6.6.4 | Example | 380 |
| 6.6.5 | Precision of Decimated TFDs: Numerical Examples | 380 |
| 6.6.6 | Summary and Conclusions | 383 |
| | References | 383 |

| | |
|---|------------|
| CHAPTER 7 Measures, Performance Assessment, and Enhancement of TFDs..... | 387 |
| Introduction and Overview | 387 |
| 7.1 TFD Design Based on the Affine Group..... | 388 |
| <i>B. Boashash and S. Touati</i> | |
| 7.1.1 Signal Scale Transformations to Construct TFDs..... | 388 |
| 7.1.2 Tomographic Methods and the Affine WVD..... | 389 |
| 7.1.3 Comparison of the Affine and Usual Wigner-Ville Functions | 391 |
| 7.1.4 Hyperbolic Chirps and Affine Group Extension..... | 392 |
| 7.1.5 Unitarity Property and Some of Its Consequences..... | 394 |
| 7.1.6 Summary and Conclusions | 395 |
| 7.2 Time-Frequency Reassignment..... | 395 |
| <i>F. Auger, P. Flandrin, and E. Chassande-Mottin</i> | |
| 7.2.1 Basic Principle..... | 396 |
| 7.2.2 Variations and Related Approaches | 400 |
| 7.2.3 Summary and Conclusions | 400 |
| 7.3 Measuring Time-Frequency Distributions Concentration..... | 401 |
| <i>L.J. Stanković</i> | |
| 7.3.1 Concentration Measurement | 401 |
| 7.3.2 Numerical Examples..... | 405 |
| 7.3.3 Parameter Optimization | 406 |
| 7.3.4 Summary and Conclusions | 408 |
| 7.4 Resolution Performance Assessment for Comparing and Selecting Quadratic TFDs | 408 |
| <i>B. Boashash and V. Sucic</i> | |
| 7.4.1 Selecting and Comparing TFDs..... | 408 |
| 7.4.2 Performance Criteria for TFDs | 410 |
| 7.4.3 Resolution Performance Measure for TFDs..... | 412 |
| 7.4.4 Selecting the Optimal TFD for a Multicomponent Signal | 413 |
| 7.4.5 Applying the Performance Measure in Real-Life Situations | 414 |
| 7.4.6 Summary and Conclusions | 417 |
| 7.5 Postprocessing in the (t,f) Domain: Methods and Performance Comparison..... | 418 |
| <i>B. Boashash and N.A. Khan</i> | |
| 7.5.1 Overview of Postprocessing Methods | 419 |
| 7.5.2 Results..... | 421 |
| 7.5.3 Discussion and Interpretation of Results | 421 |
| 7.5.4 Summary and Conclusions | 423 |
| 7.6 Time-Frequency Compressive Sensing..... | 424 |
| <i>E. Sejdić</i> | |
| 7.6.1 Sampling With Compression..... | 425 |

| | | |
|------------------|---|------------|
| 7.6.2 | Compressive Sensing and Time-Frequency Analysis | 425 |
| 7.6.3 | Illustrating Examples | 427 |
| 7.6.4 | Summary and Conclusions | 429 |
| 7.7 | Signal Complexity Estimation Using Time-Frequency Short-Term Entropy | 430 |
| | <i>V. Sucic, N. Saulig, and B. Boashash</i> | |
| 7.7.1 | Entropy Measures and Time-Frequency Distributions..... | 431 |
| 7.7.2 | The Entropy Counting Property | 432 |
| 7.7.3 | Definition and Properties of the Short-Term Entropy | 433 |
| 7.7.4 | Summary and Conclusions | 436 |
| 7.8 | Time-Frequency Analysis Using Neural Networks..... | 438 |
| | <i>I. Shafit, J. Ahmad, and S.I. Shah</i> | |
| 7.8.1 | Requirements and Motivation..... | 438 |
| 7.8.2 | Computing De-Blurred Time-Frequency Distributions..... | 439 |
| 7.8.3 | Results and Discussion | 440 |
| 7.8.4 | Summary and Conclusions | 444 |
| 7.9 | Defining Joint-Domain Representations via Discrete-Domain Frames..... | 445 |
| | <i>B. Boashash</i> | |
| 7.9.1 | Frames and Reconstruction Collections | 445 |
| 7.9.2 | Product-Function Frames | 446 |
| 7.9.3 | Cascaded Frames..... | 447 |
| 7.9.4 | Summary and Conclusions | 448 |
| | References..... | 448 |
| CHAPTER 8 | Multisensor, Multichannel, and Time-Space Processing | 453 |
| | Introduction and Overview | 453 |
| 8.1 | Multisensor Time-Frequency Analysis and Processing: Methods for Multichannel Nonstationary Data..... | 454 |
| | <i>B. Boashash and S. Ali</i> | |
| 8.1.1 | Extension of Single-Sensor QTFDs to Multisensor QTFDs..... | 454 |
| 8.1.2 | Fundamentals of Array Processing | 456 |
| 8.1.3 | Foundations of Source Separation and Localization | 459 |
| 8.1.4 | Extension of Array Signal Model to the Nonstationary Case | 462 |
| 8.1.5 | Summary and Conclusions | 467 |
| 8.2 | Spatial Time-Frequency Distributions | 467 |
| | <i>M.G. Amin and Y.D. Zhang</i> | |
| 8.2.1 | Spatial Time-Frequency Distributions | 467 |
| 8.2.2 | Fundamental Properties | 468 |
| 8.2.3 | Examples..... | 470 |
| 8.2.4 | Cross-Term Issues in STFD | 474 |
| 8.2.5 | Summary and Conclusions | 476 |

| | | |
|------------|--|-----|
| 8.3 | Quadratic Detection in Arrays Using TFDs..... | 476 |
| | <i>B. Boashash and S. Ali</i> | |
| 8.3.1 | The Detection Problem..... | 477 |
| 8.3.2 | Quadratic Detection in a Linear Array..... | 477 |
| 8.3.3 | TFD-Based Array Detection..... | 479 |
| 8.3.4 | Application: Multisensor Fetal Movement Detection System | 481 |
| 8.3.5 | Summary and Conclusions | 482 |
| 8.4 | Blind Source Separation Using Time-Frequency Distributions | 483 |
| | <i>K. Abed-Merai, A. Belouchrani, and A.R. Leyman</i> | |
| 8.4.1 | Separation of Instantaneous Mixtures..... | 483 |
| 8.4.2 | Separation of Convulsive Mixtures..... | 487 |
| 8.4.3 | Illustrative Examples..... | 489 |
| 8.4.4 | Summary and Conclusions | 490 |
| 8.5 | Underdetermined Blind Source Separation for FM-Like Signals..... | 491 |
| | <i>N. Linh-Trung, K. Abed-Merai, A. Belouchrani, and B. Boashash</i> | |
| 8.5.1 | Data Model and Assumptions..... | 492 |
| 8.5.2 | Separation Using Vector Clustering | 493 |
| 8.5.3 | Separation Using Monocomponent Extraction | 496 |
| 8.5.4 | Summary and Conclusions | 501 |
| 8.6 | Audio Source Localization and Separation Using Time-Frequency Representations | 502 |
| | <i>A. Aïssa-El-Bey and B. Boashash</i> | |
| 8.6.1 | Data Model and Assumptions..... | 503 |
| 8.6.2 | Audio Source Localization | 504 |
| 8.6.3 | Underdetermined Audio Source Separation of Convulsive Mixtures Based on Localization Information..... | 504 |
| 8.6.4 | Illustrative Example | 508 |
| 8.6.5 | Summary and Conclusions | 508 |
| 8.7 | Implementation and Code of STFDs-Based Source Separation Algorithms | 508 |
| | <i>A. Belouchrani</i> | |
| 8.7.1 | Implementation of the STFD..... | 508 |
| 8.7.2 | Implementation of STFDs-Based Source Separation | 511 |
| 8.7.3 | Implementation of the Whitening..... | 512 |
| 8.7.4 | Selection of Auto-Terms and Cross-Terms | 513 |
| 8.7.5 | Implementation of JD and JOD | 514 |
| 8.7.6 | Summary and Conclusions | 516 |
| | References..... | 516 |

PART IV TF STATISTICAL TECHNIQUES

| | | |
|------------------|---|------------|
| CHAPTER 9 | Noise Analysis and Random Processes in the (t,f) Domain | 521 |
| | Introduction and Overview | 521 |
| 9.1 | Analysis of Noise in Time-Frequency Distributions | 522 |
| | <i>LJ. Stanković</i> | |
| 9.1.1 | Wigner Distribution | 522 |
| 9.1.2 | Noise in Quadratic Time-Frequency Distributions..... | 524 |
| 9.1.3 | Noisy Signals | 526 |
| 9.1.4 | Numerical Example | 529 |
| 9.1.5 | Summary and Conclusions | 529 |
| 9.2 | Statistical Processing of Dispersive Systems and Signals | 531 |
| | <i>A. Papandreou-Suppappola, B.-G. Iem, and G.F. Boudreaux-Bartels</i> | |
| 9.2.1 | Processing Tools for Time-Varying Systems and Signals..... | 531 |
| 9.2.2 | Dispersive Time-Frequency Symbols..... | 533 |
| 9.2.3 | Special Cases of Dispersive Time-Frequency Symbols | 535 |
| 9.2.4 | Analysis Application Examples..... | 537 |
| 9.2.5 | Summary and Conclusions | 538 |
| 9.3 | Robust Time-Frequency Distributions..... | 539 |
| | <i>V. Katkovnik, I. Djurović, and LJ. Stanković</i> | |
| 9.3.1 | Robust Spectrogram..... | 539 |
| 9.3.2 | Realization of the Robust STFT | 541 |
| 9.3.3 | Robust Wigner-Ville Distribution..... | 543 |
| 9.3.4 | Example | 544 |
| 9.3.5 | Summary and Conclusions | 546 |
| 9.4 | Time-Varying Power Spectra of Nonstationary Random Processes | 546 |
| | <i>G. Matz and F. Hlawatsch</i> | |
| 9.4.1 | Nonstationary Random Processes..... | 546 |
| 9.4.2 | The Generalized Wigner-Ville Spectrum..... | 546 |
| 9.4.3 | The Generalized Evolutionary Spectrum..... | 547 |
| 9.4.4 | The Generalized Expected Ambiguity Function | 549 |
| 9.4.5 | Underspread Processes | 550 |
| 9.4.6 | Time-Varying Spectral Analysis of Underspread Processes..... | 551 |
| 9.4.7 | Summary and Conclusions | 553 |
| 9.5 | Time-Frequency Characterization of Random Time-Varying Channels..... | 554 |
| | <i>G. Matz and F. Hlawatsch</i> | |
| 9.5.1 | Time-Varying Channels | 554 |
| 9.5.2 | WSSUS Channels..... | 555 |
| 9.5.3 | Underspread WSSUS Channels..... | 558 |
| 9.5.4 | Summary and Conclusions | 561 |

| | | |
|---------------------------------------|--|------------|
| 9.6 | Characterization of Cyclostationary Signals and Their Generalizations | 562 |
| <i>A. Napolitano</i> | | |
| 9.6.1 | Cyclostationary Signals | 562 |
| 9.6.2 | Generalized Almost-Cyclostationary Signals | 566 |
| 9.6.3 | Spectrally Correlated Signals | 567 |
| 9.6.4 | Summary and Conclusions | 569 |
| | References..... | 569 |
| CHAPTER 10 | Instantaneous Frequency Estimation and Localization | 575 |
| Introduction and Overview | | |
| 10.1 | Iterative Instantaneous Frequency Estimation for Random Signals | 576 |
| <i>A. El-Jaroudi and M.K. Emresoy</i> | | |
| 10.1.1 | IF Estimation: Introduction and Background | 576 |
| 10.1.2 | Iterative Algorithm for IF Estimation | 577 |
| 10.1.3 | Convergence of the Estimation Algorithm..... | 578 |
| 10.1.4 | Summary and Conclusions | 580 |
| 10.2 | Adaptive Instantaneous Frequency Estimation Using TFDs | 580 |
| <i>LJ. Stanković</i> | | |
| 10.2.1 | Optimal Window Width..... | 581 |
| 10.2.2 | Adaptive Algorithm | 581 |
| 10.2.3 | Numerical Example | 585 |
| 10.2.4 | Summary and Conclusions | 586 |
| 10.3 | IF Estimation for Multicomponent Signals..... | 588 |
| <i>Z.M. Hussain and B. Boashash</i> | | |
| 10.3.1 | Time-Frequency Peak IF Estimation..... | 588 |
| 10.3.2 | Properties of IF Estimates Based on Quadratic TFDs..... | 590 |
| 10.3.3 | Design of Quadratic TFDs for Multicomponent IF Estimation..... | 592 |
| 10.3.4 | An Adaptive Algorithm for Multicomponent IF Estimation | 594 |
| 10.3.5 | Summary and Conclusions | 596 |
| 10.4 | Parameter Estimation for Polynomial FM Signals in Additive Gaussian Noise..... | 596 |
| <i>B. Boashash and P.J. O'Shea</i> | | |
| 10.4.1 | Problem Definition | 596 |
| 10.4.2 | The Polynomial Wigner-Ville Distributions..... | 598 |
| 10.4.3 | Higher Order Ambiguity Functions | 602 |
| 10.4.4 | Comparative Performance Evaluation of PWVDs and HAFs..... | 603 |
| 10.4.5 | Summary and Conclusions | 604 |
| 10.5 | IF Estimation of FM Signals in Multiplicative Noise..... | 604 |
| <i>B. Boashash and B. Barkat</i> | | |
| 10.5.1 | Random Amplitude Modulation..... | 604 |
| 10.5.2 | Linear FM Signal | 604 |
| 10.5.3 | Polynomial FM Signals | 608 |

| | |
|--|------------|
| 10.5.4 Time-Varying Higher-Order Spectra | 609 |
| 10.5.5 Summary and Conclusions | 611 |
| 10.6 Component Extraction from TFDs for Multicomponent Signal IF Estimation..... | 611 |
| <i>V. Sucic, J. Lerga, L.J. Rankine, and B. Boashash</i> | |
| 10.6.1 A Blind Source Separation-Based Method | 612 |
| 10.6.2 An Image Processing-Based Method | 615 |
| 10.6.3 A Real-Life Signal Example | 617 |
| 10.6.4 Summary and Conclusions | 619 |
| 10.7 Instantaneous Frequency and Polynomial-Phase Parameter Estimation Using Linear Time-Frequency Representations..... | 620 |
| <i>I. Djurović, M. Simeunović, and S. Djukanović</i> | |
| 10.7.1 Polynomial Phase Signals..... | 620 |
| 10.7.2 PD-Based Estimators | 621 |
| 10.7.3 QML Method | 623 |
| 10.7.4 Simulations | 624 |
| 10.7.5 Summary and Conclusions | 626 |
| 10.8 Sequential Bayesian Estimation of Instantaneous Frequencies..... | 626 |
| <i>B. Ristic</i> | |
| 10.8.1 Introduction | 626 |
| 10.8.2 The Sequential Bayesian Estimation Framework | 627 |
| 10.8.3 Summary and Conclusions | 630 |
| 10.9 Instantaneous Frequency Estimation Using the Viterbi Algorithm..... | 631 |
| References..... | 631 |
| CHAPTER 11 Time-Frequency Synthesis and Filtering | 637 |
| Introduction and Overview | 637 |
| 11.1 Linear Time-Frequency Filters..... | 638 |
| <i>F. Hlawatsch and G. Matz</i> | |
| 11.1.1 Time-Frequency Design of Linear, Time-Varying Filters..... | 638 |
| 11.1.2 Explicit Design—The Generalized Weyl Filter..... | 638 |
| 11.1.3 Implicit Design I—The STFT Filter | 639 |
| 11.1.4 Implicit Design II—The Gabor Filter..... | 641 |
| 11.1.5 The Discrete-Time Case..... | 642 |
| 11.1.6 Simulation Results..... | 643 |
| 11.1.7 Summary and Conclusions | 645 |
| 11.2 Time-Varying Filtering Using the STFT and Gabor Expansion..... | 645 |
| <i>B. Boashash and S. Ouelha</i> | |
| 11.2.1 The Short-Time Fourier Transform Approach to (t,f) Filtering | 645 |
| 11.2.2 Vector Formulation of the STFT | 647 |
| 11.2.3 Analysis, Synthesis, and (t,f) Filtering Using the STFT | 649 |
| 11.2.4 The Gabor Expansion Approach to (t,f) Filtering | 652 |
| 11.2.5 Summary and Conclusions | 656 |

| | |
|---|------------|
| 11.3 Time-Frequency Filtering of Speech Signals in Hands-Free Telephone Systems..... | 656 |
| <i>S. Stanković</i> | |
| 11.3.1 Limitations of the Quasi-Stationary Approach..... | 657 |
| 11.3.2 Time-Variant Filtering of Speech Signals..... | 658 |
| 11.3.3 Time-Varying Filtering Based on the Spectrogram..... | 659 |
| 11.3.4 Results..... | 660 |
| 11.3.5 Using RIDs for Time-Frequency Filtering..... | 660 |
| 11.3.6 Summary and Conclusions | 663 |
| 11.4 Signal Denoising by Time-Frequency Peak Filtering (TFPF) | 663 |
| <i>B. Boashash</i> | |
| 11.4.1 Denoising for SNR Enhancement..... | 663 |
| 11.4.2 Time-Frequency Peak Filtering (TFPF) | 663 |
| 11.4.3 Accuracy of TFPF Estimates..... | 666 |
| 11.4.4 Discrete-Time Algorithm for TFPF | 667 |
| 11.4.5 TFPF of Multicomponent Signals in Noise..... | 668 |
| 11.4.6 Summary and Conclusions | 670 |
| 11.5 Subspace Noise Filtering Using Time-Frequency Distributions | 670 |
| <i>H. Hassanpour and B. Boashash</i> | |
| 11.5.1 Introduction | 670 |
| 11.5.2 The Noise-Reduction Technique..... | 671 |
| 11.5.3 Performance Evaluation..... | 676 |
| 11.5.4 Summary and Conclusions | 677 |
| 11.6 Evaluation of Time-Frequency Denoising Algorithms for Speech Enhancement..... | 678 |
| <i>S. Ouelha and B. Boashash</i> | |
| 11.6.1 Signal Enhancement by Time-Frequency Denoising..... | 678 |
| 11.6.2 Estimation of Nonstationary Noise Power Spectrum..... | 679 |
| 11.6.3 General Process for Speech Enhancement Algorithms | 679 |
| 11.6.4 An Advanced Sonogram Using Mel Filters for Speech Enhancement: The Denoised Hearingogram..... | 681 |
| 11.6.5 Directional Noise Reduction | 683 |
| 11.6.6 Findings, Discussions, and Interpretations..... | 685 |
| 11.6.7 Summary and Conclusions | 686 |
| References..... | 688 |
| CHAPTER 12 Detection, Classification, and Estimation in the (t,f) Domain..... | 693 |
| Introduction and Overview | 693 |
| 12.1 Optimal Time-Frequency Detectors..... | 694 |
| <i>A.M. Sayeed</i> | |

| | |
|---|-----|
| 12.1.1 Time-Frequency Detection | 694 |
| 12.1.2 Time-Frequency Representations | 694 |
| 12.1.3 Time-Frequency Detection Framework..... | 695 |
| 12.1.4 Extensions | 700 |
| 12.1.5 Summary and Conclusions | 701 |
| 12.2 Time-Frequency Signal Analysis and Classification Using Matching Pursuits..... | 701 |
| <i>A. Papandreou-Suppappola and S.B. Suppappola</i> | |
| 12.2.1 Signal Time-Frequency Structures..... | 701 |
| 12.2.2 Matching Pursuits for Analysis and Classification..... | 702 |
| 12.2.3 Simulation Example Using a Synthetic Multicomponent Signal..... | 706 |
| 12.2.4 Summary and Conclusions | 708 |
| 12.3 System Identification Using Time-Frequency Filtering | 708 |
| <i>X.-G. Xia</i> | |
| 12.3.1 Problem Description | 708 |
| 12.3.2 Time-Frequency Filtering..... | 709 |
| 12.3.3 Denoising for Received Signals..... | 711 |
| 12.3.4 System Identification..... | 714 |
| 12.3.5 Summary and Conclusions | 716 |
| 12.4 Time-Frequency Methods for Signal Estimation and Detection..... | 716 |
| <i>F. Hlawatsch and G. Matz</i> | |
| 12.4.1 Nonstationary Signal Estimation..... | 716 |
| 12.4.2 Nonstationary Signal Detection | 719 |
| 12.4.3 Summary and Conclusions | 724 |
| 12.5 A General Approach to Time-Frequency-Matched Filtering..... | 724 |
| <i>B. Boashash and G. Azemi</i> | |
| 12.5.1 Classical Detection Problem and Its Solution..... | 724 |
| 12.5.2 General Formulation of Time-Frequency-Matched Filters | 725 |
| 12.5.3 Time-Frequency Kernel Selection for Optimal Detection..... | 726 |
| 12.5.4 Illustration | 727 |
| 12.5.5 Summary and Conclusions | 730 |
| 12.5.6 Appendix: Pseudocode for the Implemented Algorithm | 730 |
| 12.6 Defining Time-Frequency Image Features by Extension from Frequency Domain or Time Domain | 732 |
| <i>B. Boashash and N.A. Khan</i> | |
| 12.6.1 Extension of f -Domain Features to the Joint (t,f) Domain..... | 733 |
| 12.6.2 Extension of t -Domain Features to Joint (t,f) Domain | 736 |
| 12.6.3 Time-Frequency Image Features for Pattern Recognition | 737 |
| 12.6.4 Results and Discussion | 738 |
| 12.6.5 Summary and Conclusions | 739 |
| References | 740 |

PART V ENGINEERING APPLICATIONS

| | |
|---|------------|
| CHAPTER 13 Time-Frequency Methods in Communications | 747 |
| Introduction and Overview | 747 |
| 13.1 Time-Frequency Interference Mitigation in Spread Spectrum | |
| Communication Systems | 748 |
| <i>M.G. Amin and A.R. Lindsey</i> | |
| 13.1.1 Spread-Spectrum Systems and Interference | 748 |
| 13.1.2 Typical Signal Model | 750 |
| 13.1.3 A Time-Frequency Distribution Perspective | 750 |
| 13.1.4 Example | 752 |
| 13.1.5 Summary and Conclusions | 753 |
| 13.2 Communication Over Linear Dispersive Channels: | |
| A Time-Frequency Perspective | 753 |
| <i>A.M. Sayeed</i> | |
| 13.2.1 Linear Dispersive Channels | 753 |
| 13.2.2 Time-Frequency Model for Dispersive Channels | 754 |
| 13.2.3 Communication over Dispersive Channels | 755 |
| 13.2.4 Summary and Conclusions | 760 |
| 13.3 Eigenfunctions of Underspread Linear Time-Varying Systems | 760 |
| <i>S. Barbarossa and M. Tsitsvero</i> | |
| 13.3.1 Channel Characterizations | 761 |
| 13.3.2 Channel Decomposition | 761 |
| 13.3.3 Exact and Approximate Solutions | 762 |
| 13.3.4 Numerical Validation | 764 |
| 13.3.5 Optimal Waveforms for LTV Digital Communication Channels | 765 |
| 13.3.6 Summary and Conclusions | 768 |
| 13.3.7 Appendix | 768 |
| 13.4 Fractional Autocorrelation for Detection in Communications | 770 |
| <i>O. Akay and G.F. Boudreaux-Bartels</i> | |
| 13.4.1 Fractional Fourier Transform (FrFT) | 770 |
| 13.4.2 Fractional Convolution and Correlation | 771 |
| 13.4.3 Fractional Autocorrelation and Ambiguity Function | 774 |
| 13.4.4 Detection and Chirp Rate Parameter Estimation of Chirps | 774 |
| 13.4.5 Summary and Conclusions | 777 |
| 13.5 Time-Frequency Estimation of Radio-Signal Modulation Parameters | 777 |
| <i>A.Z. Sha'ameri and B. Boashash</i> | |
| 13.5.1 Radio-Monitoring Problem | 777 |
| 13.5.2 High-Resolution (t, f) Design | 779 |
| 13.5.3 Illustration of Two Examples with Results and Discussions | 786 |
| 13.5.4 Time-Frequency Analysis for Cognitive Radio Applications | 788 |

| | |
|--|------------|
| 13.5.5 Summary and Conclusions | 788 |
| References..... | 789 |
| CHAPTER 14 Time-Frequency Methods in Radar, Sonar, and Acoustics | 793 |
| Introduction and Overview | 793 |
| 14.1 Time-Frequency Analysis of Helicopter Doppler Radar Data..... | 794 |
| <i>S.L. Marple Jr.</i> | |
| 14.1.1 Dynamic Range Considerations in TF Analysis..... | 794 |
| 14.1.2 Classical Linear and Quadratic TFDs..... | 794 |
| 14.1.3 Alternative High-Resolution Linear TFD..... | 796 |
| 14.1.4 Application to Simulated and Actual Data..... | 799 |
| 14.1.5 Summary and Conclusions | 804 |
| 14.2 Time-Frequency Motion Compensation Algorithms for ISAR Imaging..... | 804 |
| <i>S. Barbarossa</i> | |
| 14.2.1 Echo from a Rotating Rigid Body | 806 |
| 14.2.2 Signal Analysis based on Time-Frequency Representations | 807 |
| 14.2.3 Parametric Estimation of Instantaneous Phases | 807 |
| 14.2.4 Summary and Conclusions | 810 |
| 14.3 Flight Parameter Estimation Using Doppler and Lloyd's Mirror Effects | 810 |
| <i>B.G. Ferguson and K.W. Lo</i> | |
| 14.3.1 Acoustical Doppler Effect | 810 |
| 14.3.2 Acoustical Lloyd's Mirror Effect | 812 |
| 14.3.3 Time-Frequency Signal Analysis | 815 |
| 14.3.4 Source Parameter Estimation: An Inverse Time-Frequency Problem..... | 815 |
| 14.3.5 Summary and Conclusions | 817 |
| 14.4 High-Frequency Radar Measurements of a Ballistic Missile Using the WVD | 818 |
| <i>G.J. Frazer and B. Boashash.</i> | |
| 14.4.1 Experiment Description | 818 |
| 14.4.2 Signal Description..... | 818 |
| 14.4.3 Signal Model..... | 820 |
| 14.4.4 Instantaneous Doppler Estimation | 821 |
| 14.4.5 Results..... | 821 |
| 14.4.6 Summary and Conclusions | 825 |
| 14.5 Time-Frequency Sonar Processing | 826 |
| <i>V. Chandran and B. Boashash</i> | |
| 14.5.1 Principles of Sonar Processing | 827 |
| 14.5.2 Standard Methods Used in Sonar | 830 |
| 14.5.3 Time-Frequency Approach to Sonar | 832 |
| 14.5.4 Prony and Higher-Order Spectral Methods in Sonar | 834 |
| 14.5.5 Dispersion and Angle Frequency Representation | 835 |
| 14.5.6 Summary and Conclusions | 836 |

| | | |
|---|--|------------|
| 14.6 | Sparse Time-Frequency Distributions Applied to Geophysics..... | 837 |
| <i>A. Gholami</i> | | |
| 14.6.1 | Linear Discrete Time-Frequency Distributions..... | 837 |
| 14.6.2 | Sparsity Induced Time-Frequency Distributions..... | 839 |
| 14.6.3 | Geophysical Applications..... | 841 |
| 14.6.4 | Summary and Conclusions | 843 |
| 14.7 | Audio Speech and Underwater Signals Time-Frequency Characteristic Enhancement..... | 843 |
| <i>B. Boashash and S. Ouelha</i> | | |
| 14.7.1 | Examples of Audio Signals and Their (t,f) Characteristics | 843 |
| 14.7.2 | Preprocessing of Audio Signals for SNR Enhancement..... | 846 |
| 14.7.3 | Results for Speech and Underwater Acoustic Signals | 847 |
| 14.7.4 | Summary and Conclusion..... | 852 |
| | References..... | 853 |
| CHAPTER 15 | Time-Frequency Diagnosis, Condition Monitoring, and Fault Detection | 857 |
| | Introduction and Overview | 857 |
| 15.1 | Time-Frequency Analysis of Electric Power Disturbances..... | 858 |
| <i>E.J. Powers, Y.-J. Shin, and W. Mack Grady</i> | | |
| 15.1.1 | Time-Frequency Analysis: Reduced Interference Distribution | 858 |
| 15.1.2 | Power Quality Assessment via Time-Frequency Analysis..... | 859 |
| 15.1.3 | Application of Instantaneous Frequency for Disturbance Propagation | 860 |
| 15.1.4 | Summary and Conclusions | 865 |
| 15.2 | Combustion Diagnosis by TF Analysis of Car Engine Signals..... | 865 |
| <i>J.F. Böhme and S. Carstens-Behrens</i> | | |
| 15.2.1 | Knocking Combustions..... | 865 |
| 15.2.2 | Signal Models..... | 865 |
| 15.2.3 | Signal Analysis Using Wigner-Ville Spectrum..... | 867 |
| 15.2.4 | Signal Analysis Using S-Method | 869 |
| 15.2.5 | Summary and Conclusions | 871 |
| 15.3 | Power Class Time-Frequency Representations and Their Applications..... | 871 |
| <i>A. Papandreou-Suppappola, F. Hlawatsch, and G.F. Boudreaux-Bartels</i> | | |
| 15.3.1 | Power Class Quadratic Time-Frequency Representations | 871 |
| 15.3.2 | Power Class Applications..... | 875 |
| 15.3.3 | Summary and Conclusions | 877 |
| 15.4 | The 2D Wigner-Ville Distribution Applied to Image Distortion Estimation | 878 |
| <i>A. Beghdadi, R. Iordache, and B. Boashash</i> | | |
| 15.4.1 | Image Quality and Spatial/Spatial-Frequency Representations..... | 878 |
| 15.4.2 | Continuous 2D Wigner-Ville Distribution | 878 |
| 15.4.3 | Discrete 2D Wigner-Ville Distribution | 881 |

| | |
|---|------------|
| 15.4.4 An Image Quality Metric Based on the 2D WVD | 883 |
| 15.4.5 Summary and Conclusions | 886 |
| 15.5 Extracting Time-Frequency Features from PCG Signals for Medical Diagnosis ... | 886 |
| <i>B. Boashash and A. Djebbari</i> | |
| 15.5.1 Time-Frequency Change Detection in Biomedical Signals..... | 886 |
| 15.5.2 Principles of Diagnosis in Biomedical and PCG Signals..... | 887 |
| 15.5.3 Time-Frequency Analysis of PCG Signals | 888 |
| 15.5.4 Summary and Conclusions | 893 |
| 15.6 Diagnosis and Condition Monitoring Using Time-Frequency Pattern | |
| Recognition and Machine Learning..... | 894 |
| <i>B. Boashash, S. Ouelha</i> | |
| 15.6.1 Condition Monitoring vs. Fault Diagnosis..... | 894 |
| 15.6.2 Machine-Condition Monitoring | 894 |
| 15.6.3 Three Stages of Machine-Condition Monitoring and Diagnosis | 895 |
| 15.6.4 Standard DSP Methods for Feature Extraction | 898 |
| 15.6.5 Nonstationary Signals in Machines..... | 899 |
| 15.6.6 Example 1: Gearbox Fault Detection Using the WVD | 900 |
| 15.6.7 Example 2: Detection of EEG Abnormalities in the Brain | 901 |
| 15.6.8 Summary and Conclusions | 904 |
| 15.7 Condition Monitoring of Assets Using Time-Frequency Methods | 904 |
| <i>D.M. Onchis</i> | |
| 15.7.1 Time-Frequency Methods for Condition Monitoring | 904 |
| 15.7.2 Improving the Readability of Time-Frequency Spectra | 905 |
| 15.7.3 Assessment of the Methods on Cantilever Beams | 906 |
| 15.7.4 Condition Monitoring Using the EMD Metric | 907 |
| 15.7.5 Summary and Conclusions | 909 |
| References | 910 |
| CHAPTER 16 Time-Frequency Methodologies in Neurosciences..... | 915 |
| Introduction and Overview | 915 |
| 16.1 Time-Frequency Diagnosis of Abnormalities in Newborn Physiological Signals..... | 916 |
| <i>B. Boashash and N.J. Stevenson</i> | |
| 16.1.1 The Clinical Context of the Newborn | 916 |
| 16.1.2 Newborn Signal Analysis: The (t, f) Approach | 917 |
| 16.1.3 Detection of Seizures in Newborn EEG | 920 |
| 16.1.4 Estimating the Heart Rate from ECG | 920 |
| 16.1.5 Classifying Abnormality in Newborn EEG..... | 923 |
| 16.1.6 Summary and Conclusions | 923 |
| 16.2 Time-Frequency Modeling of Nonstationary Signals for Newborn EEGs..... | 924 |
| <i>L.J. Rankine, N.J. Stevenson, and B. Boashash</i> | |

| | | |
|-------------|--|------------|
| 16.2.1 | The Time-Frequency Modeling Process..... | 925 |
| 16.2.2 | Example of Time-Frequency Modeling on Synthetic Data..... | 926 |
| 16.2.3 | Application to Newborn EEG Modeling..... | 928 |
| 16.2.4 | Summary and Conclusions | 930 |
| 16.3 | Time-Frequency Features for Nonstationary Signal Classification with Illustration on Newborn EEG Burst-Suppression Detection..... | 931 |
| | <i>G. Azemi and B. Boashash</i> | |
| 16.3.1 | Classification of Nonstationary Signals Using Information Extracted from TFDs..... | 931 |
| 16.3.2 | Time-Frequency Features | 931 |
| 16.3.3 | Illustration on Real Newborn EEG Data | 933 |
| 16.3.4 | Summary and Conclusions | 936 |
| 16.3.5 | Appendix: Pseudocode for the Implemented Algorithm | 936 |
| 16.4 | Time-Varying Analysis of Brain Networks..... | 937 |
| | <i>E. Sejdic, S. Aviyente, and B. Boashash</i> | |
| 16.4.1 | Brain Networks..... | 937 |
| 16.4.2 | Time-Frequency Analysis for Brain Networks | 938 |
| 16.4.3 | Illustrative Example..... | 940 |
| 16.4.4 | Summary and Conclusions | 943 |
| 16.5 | Time-Frequency Analysis and EEG Noise Reduction Using Empirical Mode Decomposition | 944 |
| | <i>A. Akan and A. Mert</i> | |
| 16.5.1 | Empirical Mode Decomposition | 944 |
| 16.5.2 | Hilbert-Huang Transform..... | 946 |
| 16.5.3 | EMD-Based EEG Denoising..... | 947 |
| 16.5.4 | Classification of Normal and Epileptic EEGs Using Time-Domain Properties by DFA and EMD | 951 |
| 16.5.5 | Summary and Conclusions | 952 |
| 16.6 | Time-Frequency Methodologies for Assessment of Biosignals in Neurosciences | 953 |
| | <i>B. Boashash, S. Dong, A. Omidvarnia, P. Zarjam, J.M. O'Toole, and P. Colditz</i> | |
| 16.6.1 | Time-Frequency Analysis Assessment of the Nervous System..... | 953 |
| 16.6.2 | Autonomic Nervous System Response to Perinatal Hypoxia | 954 |
| 16.6.3 | Seizure Detection Using Combined EEG and HRV | 957 |
| 16.6.4 | Measuring Phase Synchrony for EEG Abnormality Assessment..... | 958 |
| 16.6.5 | Measuring Cognitive Workload by Assessing EEG Signals..... | 960 |
| 16.6.6 | Assessing Brain Injury by IF Estimation of PLEDs..... | 961 |

| | |
|---|------------|
| 16.6.7 Summary and Conclusions | 963 |
| References..... | 963 |
| CHAPTER 17 Getting Started with a Practical and Efficient Time-Frequency Toolbox TFSAP-7.0 | 967 |
| <i>B. Boashash</i> | |
| Introduction and Overview..... | 967 |
| 17.1 Introduction and (t, f) Toolbox Description..... | 968 |
| 17.1.1 Features and Contents | 968 |
| 17.1.2 Test Signals Update | 969 |
| 17.1.3 Standard Formulations and Their Implementations | 969 |
| 17.2 Technical Content of the (t, f) Toolbox TFSAP 7.0..... | 970 |
| 17.2.1 List of Updated Functions in Version 7.0..... | 970 |
| 17.2.2 TFSAP Development Process..... | 971 |
| 17.3 User's Guide for (t, f) Toolbox TFSAP-7.0..... | 971 |
| 17.3.1 Installation..... | 971 |
| 17.3.2 Running <i>TFSAP</i> 7.0..... | 972 |
| 17.3.3 Data Accesses and Signal Generation GUI..... | 973 |
| 17.3.4 IF Estimation GUI..... | 974 |
| 17.3.5 Quadratic TFD Analysis GUI..... | 975 |
| 17.3.6 Signal Synthesis GUI..... | 977 |
| 17.3.7 TFSAP-MATLAB Flexible Interface | 978 |
| 17.4 A Brief (t, f) Tutorial Using the (t, f) Toolbox..... | 978 |
| 17.4.1 Results for Simulated Signals..... | 978 |
| 17.4.2 Results for Real Signals | 978 |
| 17.4.3 Results for Signal Synthesis and (t, f) Filtering | 983 |
| 17.4.4 Dual Use of (t, f) Toolbox (GUI or Command Line)..... | 986 |
| 17.5 Expanding and Personalizing the <i>TFSAP</i> Toolbox | 987 |
| 17.5.1 Principle and Illustration | 987 |
| 17.5.2 Exercise | 987 |
| 17.6 Summary and Conclusions | 987 |
| References..... | 987 |
| Additional Notes and Further Reading | 989 |
| Index | 995 |

This page intentionally left blank

Preface to the Second Edition

In the decade after the publication of the first edition of this book, advances in the field of time-frequency signal analysis and processing accumulated to an extent that called for an expanded and updated edition. This second edition contains:

1. over 30 new sections by leading experts in their respective fields; covering new time-frequency (t,f) methods, techniques, algorithms, and applications;
2. new tutorial sections on AM/FM decomposition of signals ([Section 1.4](#)), time-scale and wavelet transforms ([Section 2.7.6](#)), time-frequency distributions (TFDs) vs. time-scale methods ([Section 2.7.7](#)), high-resolution TFDs ([Section 3.3.4](#)), positivity of (t,f) distributions ([Section 3.4](#)), Ambiguity function and Radar ([Section 3.5](#)); thus making [Chapter 3](#) the most important;
3. a new chapter on “Getting Started” in (t,f) computation using the original (t,f) toolbox updated for this second Edition ([Chapter 17](#));
4. a new introductory chapter (numbered \mathcal{I} to minimize numbering changes), written at a more intuitive level, making the book accessible to a wider readership reflecting the widening use of the (t,f) approach; and
5. a new chapter on (t,f) methods in neuroscience. This emerging field of (t,f) application is the subject of the new [Chapter 16](#), containing six new sections; these concern the assessment of brain EEG and ECG abnormalities ([Section 16.1](#)), (t,f) modeling of EEG signals ([Section 16.2](#)), EEG (t,f) features for classification ([Section 16.3](#)), brain networks analysis ([Section 16.4](#)), EEG (t,f) analysis using empirical mode decomposition ([Section 16.5](#)), and perspectives for further advances ([Section 16.6](#)).

Among the new sections, four relate (t,f) methods to time-scale methods ([Section 4.1](#)), systems ([Section 4.9](#)), energy ([Section 4.10](#)) and the EMD with the Hilbert spectrum ([Section 4.12](#)); four concern directional kernels ([Section 5.9](#)), adaptive directional kernels ([Section 5.10](#)), the modified S-transform ([Section 5.11](#)) and TFDs positivity ([Section 5.12](#)). Other new topics include memory-efficient computation of TFDs ([Section 6.6](#)), compressive sensing ([Section 7.6](#)), signal complexity ([Section 7.7](#)), neural networks for (t,f) localization ([Section 7.8](#)), the improvement of (t,f) representations using (t,f) post-processing ([Section 7.5](#)), introducing multichannel/multisensor (t,f) processing ([Section 8.1](#)), audio source separation ([Section 8.6](#)), cyclostationary signals ([Section 9.6](#)), extraction of signal components ([Section 10.6](#)), polynomial-phase parameter estimation ([Section 10.7](#)), IF sequential Bayesian estimation ([Section 10.8](#)), subspace noise filtering ([Section 11.5](#)), speech enhancement ([Section 11.6](#)), generalizing one-dimensional concepts in matched filtering ([Section 12.5](#)), (t,f) feature definition ([Section 12.6](#)), modulation parameter estimation ([Section 13.5](#)), sparse TFDs in geophysics ([Section 14.6](#)), audio signals’ (t,f) characteristics ([Section 14.7](#)), using (t,f) features to improve medical diagnosis ([Section 15.5](#)), and structural condition monitoring ([Section 15.7](#)).

More information on the scope of new and existing chapter/sections appears in the table of contents and chapter overviews.

The (t,f) toolbox package (available at the Elsevier website) complements the other chapters of this book. It describes in a simple way the most important algorithms presented in the book and makes them ready for use via an easy practical graphical user interface (GUI). With the aid of [Chapter 17](#), new users can start using the algorithms on simulated and real examples, compare with the results presented in the

book, and insert the algorithms in their own applications. Source code will be made available. Readers can use the package as an introduction to the field (in combination with the tutorial in [Chapter 17](#)), or as a set of MATLAB™-compatible commands that can be accessed from the command line or called from other programs, or as source code that can be adapted for particular applications.

The (t,f) toolbox TFSAP is a recent upgrade to the package formerly known as TFSA, the first package in its field, which was released publicly in 1987 at the inaugural ISSPA conference in Brisbane, Australia, and frequently upgraded thereafter. It was originally purchased by several hundred universities and research institutions throughout the world, and also made available free of charge as part of evaluation agreements.

In addition to new material, this second edition also updates the text and references of existing sections, ensures more consistent use of common terminology and notations, and enriches the cross-referencing between the various parts of the book. The result is a comprehensive and connected presentation of standard and novel topics in (t,f) signal analysis and processing. My hope is that it achieves the best possible outcome, within the constraints of a single volume, in equipping modern researchers and engineers with the information and tools necessary to undertake any study, research, or development project in this important scientific field.

ACKNOWLEDGMENTS

I thank several colleagues, postdocs, and research assistants who provided much needed help to complete this project. First, Dr Gavin Putland and Ms Ayat Salim assisted with L^AT_EX editing of the sections in the book and all related tasks. In addition, several other colleagues contributed to the technical reviews of individual chapters or sections; they include: S. Ouelha, N. A. Khan, E. Sejdic, S. Touati, S. Ali, A. Aïssa-El-Bey, G. Azemi, J. O'Toole, S. Aviyente, B. Jawad, N. Stevenson, H. Hassanpour, A. Saleem, S. Dong, A. Zuri-Sha'ameri, M. A. Awal, A. Ouahabi and M. Alsdad. In addition, all individual sections of this book were technically reviewed by at least two experts in the field, as shown in the first page of each section.

Publication of this second edition was facilitated by several grants from Qatar Foundation. In particular, the work in the contributions in [Sections 4.1, 11.6, 13.5, 14.7, and 15.6](#) and the updates to [Sections 7.1, 7.9, 14.4, and 14.5](#), was funded by Qatar National Research Fund, grant number NPRP 6-680-2-282. The work in [Sections 5.11, 6.6, 7.7, 10.6, 11.5, 12.5, 15.5, 16.2, 16.3, 16.4, and 16.6](#), and the updates to [Chapters 1 and 2](#) and [Sections 4.6, 4.12, 5.4, 5.5, 5.7, 5.12, 6.1, 6.5, 7.4, 8.1, 10.3, 10.4, 10.5, 11.2, 11.4, 15.4 and 16.1](#) was funded by Qatar National Research Fund, grant number NPRP 6-885-2-364. The work in [Chapter 17, Sections 5.9, 5.10, 7.5, and 12.6](#), and the updates to [Chapter 3](#), was funded by Qatar National Research Fund, grant number NPRP 4-1303-2-517. The work reported in Section 16.6 was funded by the Australian Research Council and the National Health & Medical Research Council, Australia.

Boualem Boashash, Editor

Preface to the First Edition

Time-Frequency Signal Analysis and Processing (TFSAP) is a collection of theory and algorithms used for analysis and processing of nonstationary signals, as found in a wide range of applications including telecommunications, radar, and biomedical engineering. This book brings together the main knowledge of TFSAP, from theory to applications, in a user-friendly reference suitable for both expert and nonexpert readers.

The **contents** of the book include:

1. a *comprehensive tutorial introduction* to TFSAP, accessible to anyone who has taken a first course in signals and systems;
2. more specialized theory and algorithms, concisely presented by some of the leading authorities on the respective topics; and
3. studies of *key applications*, written by leading researchers, showing how to use TFSAP methods to solve practical problems.

The **motivation** for producing this book was twofold:

- My original and widely used decade-old tutorial on TFSAP [1] needed updating in two respects. First, some of the advances of the last decade are sufficiently fundamental to warrant inclusion in an introductory treatment, while others are sufficiently important to demand coverage in any comprehensive review of TFSAP. Second, new applications have widened the range of disciplines interested in TFSAP, and thus reduced the common background knowledge that may be expected of readers. **Part I** of this book addresses these needs.
- The need for a standard language of discourse became apparent in 1990 while I was editing the 23 contributions to the first comprehensive book in the field [2]. These seminal contributions to TFSAP led to further developments throughout the 1990s, including some significant advances in practical methods suitable for nonstationary signals including instantaneous frequency estimation [3]. These efforts continued apace as this book was being written. Such rapid progress produced a variety of new terminologies and notations that were in need of standardization and inclusion in an updated reference book.

The **organization** of this book uses five parts, each part including several chapters, and each chapter comprising several articles. **Part I** introduces the basic concepts while **Parts II–V** cover more advanced or specialized areas.

Part I defines and explains the basic concepts of TFSAP, intuitively derives a variety of well-known time-frequency distributions (TFDs), and then reduces them to a common form. This leads to the general treatment of quadratic TFDs in [Chapter 3](#), which should be regarded as the core of the book and as a prerequisite for the later chapters.

Part II gives more details on some fundamental topics of TFSAP, such as TFD design and signal analysis in the (t,f) plane.

Part III describes specialized techniques used in implementation, measurement, and enhancement of TFDs.

Part IV presents the key statistical techniques for TFSAP of noisy signals, including a full treatment of detection and classification methods.

Part V describes a representative selection of TFSAP applications, encompassing telecommunications, radar, sonar, power generation, image quality, automotive applications, machine condition monitoring, and biomedical engineering.

Usability is enhanced by an updated *consolidated bibliography* (alphabetical by author) and a *two-level index* (which also serves as a dictionary of abbreviations).

Under the standard review procedure used for this book, each article had two (usually external) reviewers concentrating on scientific rigor and accuracy, plus two anonymous internal reviewers concentrating on clarity and consistency.

ACKNOWLEDGMENTS

Acknowledgments are due to a number of people who made possible the completion of this book. Foremost among them are my two sons, who aided me to continue this work during and after my wife's final illness, thus contributing to my sense of balance and purpose during this difficult period. I thank all authors and reviewers, and the organizers of the Special Sessions on TFSAP at ISSPA conferences, for their expertise, timely effort, and professionalism, and for facilitating the exchange of ideas between contributors to this book. I thank my research students and the SPRC staff for valuable assistance. In particular, Gavin Putland assisted with the technical editing of portions of Part I and was responsible for the final mix-down of the authors' L^AT_EX™ and PostScript™ files.

Boualem Boashash, Editor

REFERENCES

- [1] B. Boashash, "Time-Frequency Signal Analysis", in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [2] B. Boashash (Ed.), "Time-Frequency Signal Analysis: Methods and Applications", Longman-Cheshire/Wiley, Melbourne/New York, 1992.
- [3] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal – Part 1: Fundamentals; Part 2: Algorithms and applications", Proc. IEEE, vol. 80, pp. 519-568, April 1992.

List of Contributors

K. Abed-Meraim

Université d'Orléans, Orléans, France

J. Ahmad

Abasyn University, Islamabad, Pakistan

A. Aïssa-El-Bey

Institut Mines-Télécom, Télécom Bretagne, Brest Cedex 3, France

A. Akan

Istanbul University, Avcılar, İstanbul, Turkey

O. Akay

Dokuz Eylül University, Buca/Izmir, Turkey

S. Ali

Qatar University, Doha, Qatar

T. Alieva

Universidad Complutense de Madrid, Madrid, Spain

M.G. Amin

Villanova University, Villanova, PA, USA

S. Assous

Geoscience Development, Weatherford Laboratories Ltd, Gotham Road, East Leake LE12 6JX, UK

F. Auger

L'Université Nantes Angers Le Mans (LUNAM), Saint Nazaire Cedex, France

S. Aviyente

Michigan State University, East Lansing, MI, USA

M.A. Awal

University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia

G. Azemi

Razi University, Kermanshah, Iran (formerly University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia)

M.J. Bastiaans

Eindhoven University of Technology, Eindhoven, The Netherlands

R.G. Baraniuk

Rice University, Houston, TX, USA

S. Barbarossa

Sapienza University of Rome, Rome, Italy

B. Barkat

Petroleum Institute, Abu Dhabi, United Arab Emirates

V. Barroso

Instituto Superior Técnico, Lisboa, Portugal

A. Beghdadi

Université Paris 13, Villetaneuse, France

A. Belouchrani

Ecole Nationale Polytechnique, Algiers, Algeria

B. Boashash

Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia

J.F. Böhme

Ruhr-Universität Bochum, Bochum, Germany

A.-O. Boudraa

Institut de Recherche de l'École Navale, Almada, France

G.F. Boudreaux-Bartels

University of Rhode Island, Kingston, RI, USA

S. Carstens- Behrens

University of Applied Sciences Koblenz, Remagen, Germany

V. Chandran

Queensland University of Technology, Brisbane, QLD, Australia

E. Chassande -Mottin

Université Paris Diderot, Paris Cedex 13, France

P. Colditz

University of Queensland Centre for Clinical Research, Brisbane, Australia

A. Djebbari

Université Abou Bekr Belkaid, Tlemcen, Algeria

S. Djukanović

University of Montenegro, Podgorica, Montenegro

I. Djurović

University of Montenegro, Podgorica, Montenegro

S. Dong

University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia

A. El-Jaroudi

University of Pittsburgh, Pittsburgh, PA, USA

M.K. Emresoy

University of Pittsburgh, Pittsburgh, PA, USA

B.G. Ferguson

Defence Science and Technology Organisation, Pyrmont, NSW, Australia

P. Flandrin

École Normale Supérieure de Lyon, Lyon Cedex 07, France

G.J. Frazer

Defence Science and Technology Organisation, Edinburgh, SA, Australia

D. Frunzaverde

Eftimie Murgu University of Resita, Resita, Romania

L. Galleani

Politecnico di Torino, Torino, Italy

A. Gholami

Institute of Geophysics, University of Tehran, Iran

H. Hassanpour

Shahrood University of Technology, Shahrood, Iran

F. Hlawatsch

Vienna University of Technology, Vienna, Austria

Z.M. Hussain

University of Kufa and, Edith Cowan University, Australia

B.-G. Iem

Gangneung-Wonju National University, Gangwon-do, Republic of Korea

R. Jordache

GE Healthcare, 78530 Buc, France

B.K. Jawad

Qatar University, Doha, Qatar

G. Jones

Accipiter Radar Technologies Inc., Fenwick, ON, Canada

D.L. Jones

University of Illinois at Urbana-Champaign, Urbana, IL, USA

V. Katkovnik

Tampere University of Technology, Hervanta, Tampere, Finland

N.A. Khan

Qatar University, Doha, Qatar

J. Lerga

University of Rijeka, Rijeka, Croatia

A.R. Leyman

National University of Singapore, Singapore

N. Linh-Trung

Vietnam National University, Hanoi, Vietnam

A.R. Lindsey

Austral Engineering and Software, Inc., Athens, OH, USA

K.W. Lo

Defence Science and Technology Organisation, Pyrmont, NSW, Australia

xxxviii List of Contributors

W. Mack Grady

Baylor University, Waco, TX, USA

S.L. Marple Jr.

Georgia Tech Research Institute, Atlanta, GA, USA

G. Matz

Vienna University of Technology, Vienna, Austria

A. Mert

Piri Reis University, Tuzla, Istanbul, Turkey

A. Napolitano

University of Napoli “Parthenope”, Napoli, Italy

P.M. Oliveira

Escola Naval, Almada, Portugal

A. Omidvarnia

The University of Melbourne, Melbourne, Australia

D.M. Onchis

Eftimie Murgu University of Resita, Resita, and West University of Timisoara, Timisoara, Romania

P.J. O’Shea

Qatar University, Doha, Qatar; University of Queensland, Brisbane, QLD, Australia

S. Ouelha

Qatar University, Doha, Qatar

J.M. O’Toole

University College Cork, Corcaigh, Ireland

A. Papandreou-Suppappola

Arizona State University, Tempe, AZ, USA

E.J. Powers

The University of Texas at Austin, Austin, TX, USA

G.R. Putland

Lund Values Research Group, Prosper Australia, Melbourne, VIC, Australia

L.J. Rankine

Business Engineers Australia, QLD (formerly University of Queensland, Brisbane, QLD, Australia)

B. Ristic

RMIT University, Swanston Street, Melbourne, VIC 3000 Australia

F. Salzenstein

University of Strasbourg, Strasbourg, France

N. Saulig

University of Rijeka, Rijeka, Croatia

A.M. Sayeed

University of Wisconsin–Madison, Madison, WI, USA

E. Sejdić

University of Pittsburgh, Pittsburgh, PA, USA

A.Z. Sha'ameri

Universiti Teknologi Malaysia, Johor, Malaysia

I. Shafi

Abasyn University, Islamabad, Pakistan

S.I. Shah

Abasyn University, Islamabad, Pakistan

Y.-J. Shin

Yonsei University, Seoul, Korea

M. Simeunović

University of Montenegro, Podgorica, Montenegro

S. Stanković

University of Montenegro, Podgorica, Montenegro

LJ. Stanković

University of Montenegro, Podgorica, Montenegro

N.J. Stevenson

Helsinki University Central Hospital, P.O.Box 280, 00029 HUS, Finland

V. Sucic

University of Rijeka, Rijeka, Croatia

S.B. Suppappola

Cirrus Logic Inc., Mesa, AZ, USA

G. Tauböck

Acoustics Research Institute, The Austrian Academy of Sciences, Vienna, Austria

S. Touati

King Saud University, Riyadh, Kingdom of Saudi Arabia

M. Tsitsvero

Sapienza University of Rome, Rome, Italy

A.J. van Leest

ASML, Veldhoven, The Netherlands

S. Ventosa

Institut de Physique du Globe de Paris, Paris, France

W.J. Williams

The University of Michigan, Ann Arbor, MI, USA

X.-G. Xia

University of Delaware, Newark, DE, USA

P. Zarjam

Kermanshah Azad University, Iran, and University of New South Wales, Sydney, Australia

Y.D. Zhang

Villanova University, Villanova, PA, USA

Abbreviations

The index of this book also serves as the main dictionary of abbreviations; the index entry for each abbreviation is a cross-reference to the corresponding unabbreviated entry. The most common abbreviations are listed below.

| | |
|-------------------|-------------------------------------|
| FSK | frequency-shift keying |
| AOK | adaptive optimal kernel |
| AUC | area-under-the-curve |
| AWVD | adaptive Wigner-Ville distribution |
| B-S | burst-suppression |
| BD | B-distribution |
| BJD | Born-Jordan distribution |
| BSS | blind source separation |
| <i>BT</i> product | bandwidth-duration product |
| CDMA | code-division multiple-access |
| CFD | central finite-difference |
| CKD | compact-support-Kernel distribution |
| CRLB | Cramer-Rao lower bound |
| CSK | compact-support-kernel |
| CWD | Choi-Williams distribution |
| DCT | discrete cosine transform |
| DGF | directional Gaussian filter |
| DGT | discrete Gabor transform |
| DI | Doppler-independent |
| DOA | direction-of-arrival |
| DWT | discrete wavelet transform |
| DWVD | discrete WVD |
| EEG | electroencephalogram |
| EMBD | extended modified B-distribution |
| EMD | empirical mode decomposition |
| FFT | fast Fourier transform |
| FM | frequency modulation |
| FrFT | fractional Fourier transform |
| FT | Fourier transform |

| | |
|------|--|
| GAF | generalized ambiguity function |
| GLRT | generalized likelihood ratio test |
| GPS | global positioning system |
| GWD | generalized Wigner distribution |
| HAF | higher-order ambiguity function |
| HHT | Hilbert-Huang transform |
| HVS | human visual system |
| IA | instantaneous amplitude |
| IAF | instantaneous autocorrelation function |
| IDGT | discrete Gabor transform (inverse) |
| IF | instantaneous frequency |
| IFT | inverse Fourier transform |
| IMF | intrinsic mode function |
| IP | instantaneous phase |
| ISAR | inverse synthetic aperture radar |
| KLT | Karhunen-Loëve transform |
| LD | Levin distribution |
| LFM | linear FM |
| LI | lag-independent |
| LTV | linear time-varying |
| MBD | modified B-distribution |
| MDD | Multi-directional distribution |
| MDK | Multi-directional kernel |
| MIMO | multiple-input multiple-output |
| ML | maximum likelihood |
| MSE | mean squared error |
| MST | modified S-transform |
| NMF | nonnegative matrix factorization |
| PCA | principal component analysis |
| PD | phase differentiation |
| PLV | phase locking value |
| PPS | polynomial-phase signal |
| PSD | power spectral density |
| PSK | phase-shift keying |

| | |
|--------|--|
| PWVDs | polynomial WVDs |
| QMF | quadrature matched filter |
| QTFRs | quadratic time-frequency representations |
| RD | Rihaczek distribution |
| RID | reduced-interference distribution |
| ROC | receiver operating characteristic |
| RWT | Radon-Wigner transform |
| SAR | synthetic aperture radar |
| SDS | spectral delay shift |
| SIMO | single-input multiple-output |
| SISO | single-input single-output |
| SM | S-method |
| SNR | signal-to-noise ratio |
| SS | spread-spectrum |
| s.t. | such that |
| ST | S-transform |
| STFD | spatial TFD |
| STFT | short-time Fourier transform |
| SVD | singular-value decomposition |
| TF | time-frequency |
| TFD | time-frequency distribution |
| TFEC | time-frequency energy concentration |
| TFMF | time-frequency matched filter |
| TFPF | time-frequency peak filtering |
| TFRE | time-frequency Rényi entropy |
| TFR | time-frequency representation |
| TFSP | time-frequency signal processing |
| TK | Teager-Kaiser |
| TVIR | time-varying impulse response |
| TV-HOS | time-varying higher-order spectra |
| TVTS | time-varying transfer function |
| WPT | wavelet packet transform |
| WSS | wide-sense stationary |
| WT | wavelet transform |
| XWVD | cross-Wigner-Ville distribution |

This page intentionally left blank

Book Standard Notations

Symbols frequently used in this book are listed below in alphabetical order, with Latin (English) letters before Greek letters. The meaning in the list below should be assumed unless the symbol is otherwise defined in context.

Latin

| | |
|--|--|
| $A(f)$ | amplitude spectrum |
| $a(t)$ | instantaneous amplitude (real envelope), e.g., of $s(t) = a(t) \cos \phi(t)$ or $z(t) = a(t)e^{j\phi(t)}$ |
| $A_z(v, \tau)$ | symmetrical ambiguity function (SAF) of $z(t)$ $= \int_{-\infty}^{\infty} z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) e^{-j2\pi vt} dt$ |
| $\mathcal{A}_z(v, \tau)$ | filtered ambiguity function of $z(t)$ $= g(v, \tau) A_z(v, \tau)$ |
| B | bandwidth of signal |
| $\text{DFT}_{n \rightarrow k}\{x[n]\}$ | discrete Fourier transform of $x[n]$, (from n to k domain) |
| $\text{DTFT}_{n \rightarrow f}\{x[n]\}$ | discrete-time Fourier transform of $x[n]$, to f domain |
| f | frequency |
| $f_i(t)$ | instantaneous frequency |
| $\mathcal{F}_{t \rightarrow f}\{x(t)\}$ | Fourier transform (FT) of $x(t)$, to f domain |
| $\mathcal{F}_{f \rightarrow t}^{-1}\{X(f)\}$ | inverse Fourier transform (IFT) of $X(f)$, back to t domain |
| $F_z^w(t, f)$ | short-time Fourier transform (STFT) of $z(t)$ with window $w(t)$ |
| \mathbb{F}^α | fractional FT with parameter α |
| F_v | frequency-shift operator: $(F_v s)(t) = e^{j2\pi vt} s(t)$ |
| $\mathcal{G}[l, k]$ | discrete Doppler-frequency kernel |
| $g[l, m]$ | discrete Doppler-lag kernel |
| $G[n, m]$ | discrete time-lag kernel |
| $G(t, \tau)$ | time-lag kernel $= \int_{-\infty}^{\infty} g(v, \tau) e^{j2\pi vt} dv$ |
| $\mathcal{G}(v, f)$ | Doppler-frequency kernel $= \int_{-\infty}^{\infty} g(v, \tau) e^{-j2\pi f\tau} d\tau$ |
| $g(v, \tau)$ | Doppler-lag kernel |
| $g_\theta(v, \tau)$ | multi-directional Doppler-lag kernel with set of signal direction angles θ |
| H | operator of system or channel |
| $\mathcal{H}\{s(t)\}$ | Hilbert transform of real signal $s(t)$ |
| $\text{Im}\{\cdot\}$ or $\Im\{\cdot\}$ | imaginary part of $\{\cdot\}$ |

| | |
|--|--|
| k | discrete frequency (in bins) |
| $K_z(t, \tau)$ | instantaneous auto-correlation function (IAF) of $z(t)$ $= z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$ |
| $k_z(v, f)$ | spectral correlation function (SCF) of $z(t)$ $= Z(f + \frac{v}{2})Z^*(f - \frac{v}{2})$, where $Z(f) = \mathcal{F}_{t \rightarrow f}\{z(t)\}$ |
| $\mathbf{K}_{zz}(t, f)$ | spatial instantaneous autocorrelation function (SIAF) |
| l | discrete Doppler (frequency shift) |
| m | discrete lag (delay, time shift) |
| n | discrete time (in samples) |
| $\text{Re}\{\cdot\}$ or $\Re\{\cdot\}$ | real part of $\{\cdot\}$ |
| $\text{rect}(x)$ | function equal to 1 for $-\frac{1}{2} \leq x \leq \frac{1}{2}$, and 0 elsewhere |
| $R_z(t, \tau)$ | smoothed instantaneous autocorrelation function (IAF) of $z(t)$ $= G(t, \tau) * K_z(t, \tau)$ |
| $\mathcal{R}_z(t, \tau)$ | time-varying autocorrelation of random process $z(t)$ |
| $\mathcal{R}_z(\tau)$ | autocorrelation of wide-sense stationary random process $z(t)$ |
| $\text{sinc } x$ | $\frac{1}{\pi x} \sin(\pi x)$ |
| $\mathcal{S}_z(t, f)$ | evolutive (Wigner-Ville) spectrum of random process $z(t)$ $= \mathcal{F}_{\tau \rightarrow f}\{\mathcal{R}_z(t, \tau)\}$ |
| $\mathcal{S}_z(f)$ | power spectral density of wide-sense stationary process $z(t)$ $= \mathcal{F}_{\tau \rightarrow f}\{\mathcal{R}_z(\tau)\}$ |
| $S_z^{(H)}(t, f)$ | sonograph (sonogram) of $z(t)$ with filter $H(f)$ |
| $S_z^w(t, f)$ | spectrogram (squared STFT) of $z(t)$ with window $w(t)$ |
| T | duration of signal |
| t | time |
| \mathbf{T}_τ | time-shift operator: $(\mathbf{T}_\tau s)(t) = s(t - \tau)$ |
| $w(t)$ | window function, <i>or</i> white noise |
| $W_z(t, f)$ | Wigner-Ville distribution (WVD) of $z(t)$ $= \int_{-\infty}^{\infty} z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})e^{-j2\pi f\tau} d\tau$ |
| $\text{WT}_x(t, a)$ | Wavelet transform of signal $x(t)$ $= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(u)h^*\left(\frac{u-t}{a}\right) du$ |

Greek

| | |
|-------------------|---|
| $\gamma[n, k]$ | discrete time-frequency kernel |
| $\gamma(t, f)$ | time-frequency kernel = $\int_{-\infty}^{\infty} G(t, \tau) e^{-j2\pi f \tau} d\tau$ |
| $\epsilon(t)$ | additive noise with mean μ_ϵ and variance σ_ϵ^2 |
| $\theta(f)$ | phase spectrum |
| ν | Doppler (frequency shift) |
| $\Pi_T(t)$ | rectangular or box function, also written as $\text{rect}(t/T)$ defined as: |
| $\rho_z(t, f)$ | $\text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1 & \text{if } t \leq T/2 \\ 0 & \text{if } t > T/2 \end{cases}$ <p>quadratic time-frequency distribution (TFD) of signal $z(t)$ $= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j2\pi \nu(t-u)} g(\nu, \tau) z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) e^{-j2\pi f \tau} d\nu du d\tau$ where $g(\nu, \tau)$ is the Doppler-lag kernel</p> |
| $\rho_{zz}(t, f)$ | spatial time-frequency distributions |
| τ | lag (delay, time shift) |
| $\tau_d(f)$ | spectral delay (time delay) |
| $\tau_g(f)$ | group delay |
| $\phi(t)$ | instantaneous phase |
| $\psi_{d,c}(u)$ | discrete wavelet |
| ∇ | partial differential vector operator |

Math symbols

| | |
|------------------------------|--|
| $*_t$ | convolution in time |
| $*_{(t,f)}$ | 2D convolution in both time and frequency |
| \equiv, \triangleq or $:=$ | equal by definition, or defined as |
| \propto | proportional to |
| $\int_{\mathbb{R}}$ | $\equiv \int_{-\infty}^{\infty}$ |
| $\int_{\mathbb{R}^2}$ | $\equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty}$ |
| \exists | there exists |

This page intentionally left blank

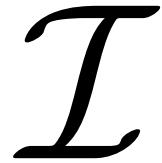
PART

A TIME-FREQUENCY TUTORIAL

I

This page intentionally left blank

THE TIME-FREQUENCY APPROACH: ESSENCE AND TERMINOLOGY⁰



INTRODUCTION AND OVERVIEW

This chapter starts the introductory tutorial that constitutes [Part I](#) of the book. It provides the basic insights needed to understand time-frequency (t,f) fundamental principles, techniques, methods, and algorithms; it also provides the background material required to do the (t,f) toolbox tutorial in [Chapter 17](#).

Time-frequency signal processing (TFSP) is a set of signal processing (SP) methods, techniques, and algorithms in which the two natural variables time (t) and frequency (f) are used *concurrently*. This contrasts with traditional SP methods in which either time t or frequency f is used exclusively and independently of the other.

TFSP has become a standard digital signal processing (DSP) technology with applications found in all traditional areas of DSP and beyond. This chapter presents the essential concepts on which the field of TFSP is built. It uses simple language, describes the most important concepts in qualitative terms and illustrates them with examples. An understanding of these basic (t,f) concepts and terminology is a prerequisite for comprehending, mastering, and using the more specialized methods reported in the other chapters.

The choice of the contents, style, and organization of this chapter is motivated by the desire to provide a nonspecialist with:

- a bridge or stepping stone to the key fundamental (t,f) concepts presented in [Chapter 1](#);
- a step-by-step approach to conceptually grasp the complexities of the field and be able to represent a signal in a 2D (t,f) space;
- a framework that relates general and generic TFSP principles to those more specific methods and techniques applicable in typical situations;
- a presentation of TFSP using examples that are familiar to students, engineers, and researchers in related fields;
- an insight into key characteristic features that are (t,f) specific, such as instantaneous frequency (IF), spectral delay (SD), and instantaneous bandwidth.

Finally, a brief overview of the following chapters is provided at the end of this chapter.

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa).

I.1 TRADITIONAL SIGNAL REPRESENTATIONS

Nature shows us in our daily experiences that these two variables, t and f , are usually interrelated in signals, such as human speech, recorded music, or whale song (Fig. I.1.1). Such signals are described as *time-varying* or *nonstationary* because their frequency content varies with time. In Fig. I.1.1, the time-domain representation (a) does not clearly show what frequencies are present, and the frequency-domain representation (b) does not show *when* they are present, but the time-frequency representation (TFR) (c) shows both [1–4].

I.1.1 ORIGIN AND TYPES OF SIGNALS

An objective of DSP is to extract as much information as possible from a signal and use it for effective decision making. It follows that precisely defining the characteristics of the signal under investigation is often the most important step in the process of solving practical engineering problems.

A step-by-step problem-solving procedure for a real-life engineering application is shown in Fig. I.1.2. It usually involves several tasks such as:

1. stating the physical problem and determining the physical laws at work;
2. deriving the mathematical relationship relating signals to the system;
3. determining the corresponding type of signals (signal characterization);
4. deriving the relationship between the characteristics of the signal type and the characteristics of the physical model (system analysis); and
5. developing algorithms to estimate the characteristics of the signal and/or system from the observations (implementation phase).

The two classical representations of a signal s are the temporal or time-domain representation $s(t)$, and the spectral or frequency-domain representation $S(f)$. We illustrate these representations using six examples of signal types that may be identified in the “signal characterization” step of the above

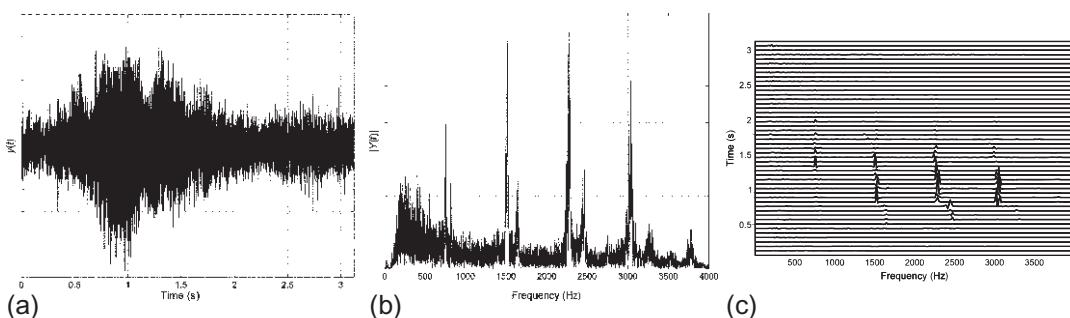
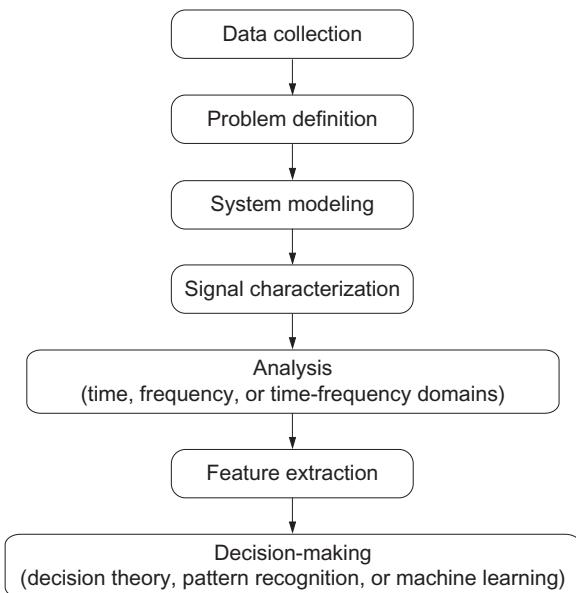


FIGURE I.1.1

Three representations of a humpback whale song: (a) time domain, (b) frequency domain, and (c) joint time-frequency domain (see tutorial in Chapter 17 to reproduce this plot).

**FIGURE I.1.2**

Typical flowchart for a basic DSP problem-solving and decision-making procedure (other procedures can include preprocessing stages, filtering, and postprocessing stages).

procedure. The simplest example is a pure sine wave (like the 50 or 60 Hz voltage signal in a domestic power supply). The second simplest example is a sum of several different sine waves. The successive examples become more complex and more general, as encountered in practical applications such as studying brain malfunction [3].

I.1.2 TIME-DOMAIN REPRESENTATION OF TYPICAL SIGNALS

This section illustrates in a simple conceptual way how to characterize a signal in the time domain before the next chapter goes into more depth and presents the detailed mathematical formulations for the six typical signals.¹ The choice of signal characteristics depends on the degree of complexity of the signal, which can serve as a criterion for classifying the signal. The algorithms that are effective for extracting such characteristics from observations of the signal will depend on the type of signal.

The time-domain representation $s(t)$ reveals information about the actual presence of the signal, its start and end times, its strength and temporal evolution, and it indicates how the signal energy is distributed along the t axis.

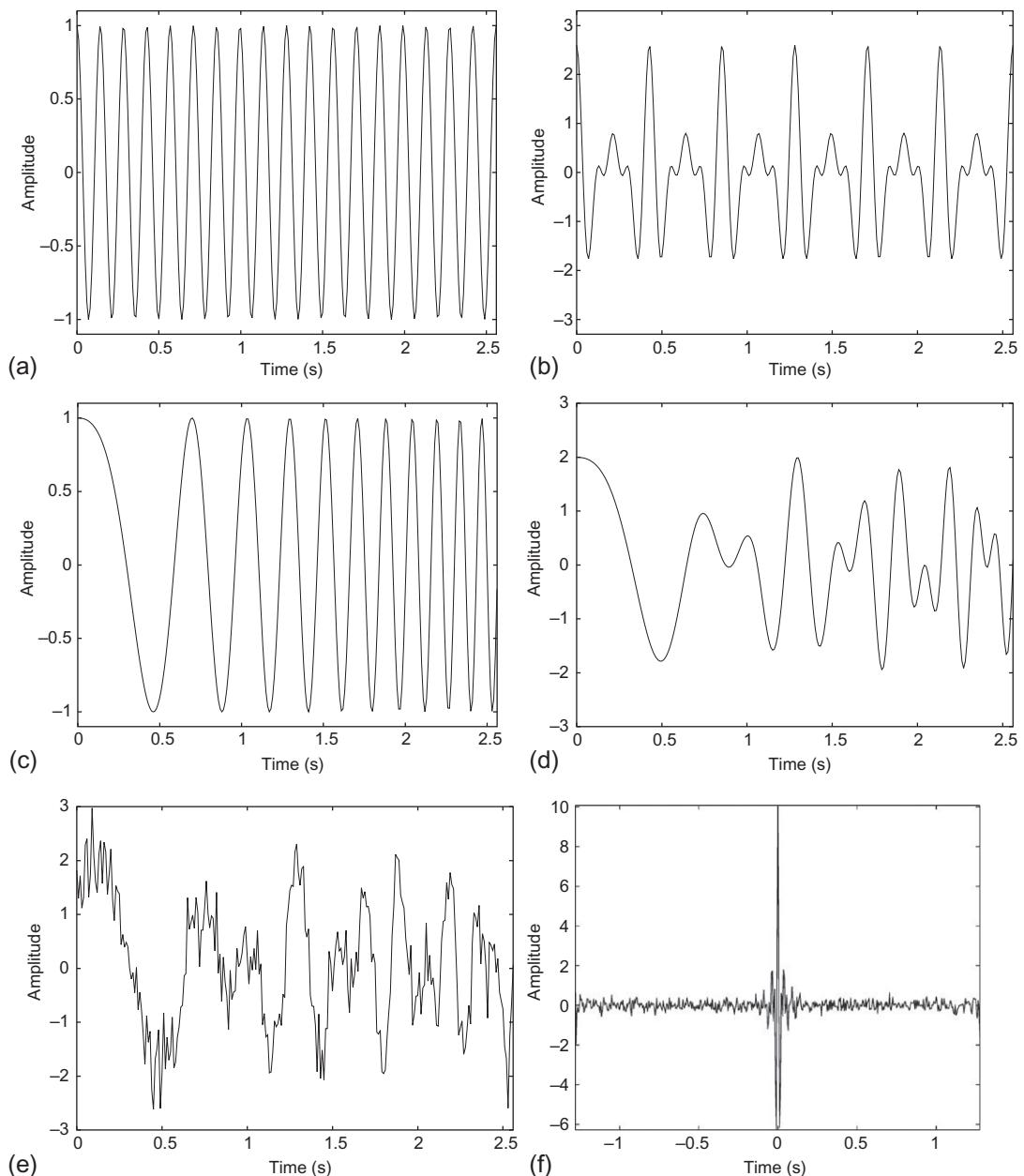
¹A common goal in SP is to extract information about the physical system generating the signal, from measurements, recordings, or observations.

Figure I.1.3 shows the time-domain representations of six typical signals. The first four cases may be classified as *noiseless* and *deterministic*.² The last two cases may be classified as *noisy* or *random* signals.³ The six signals are described below heuristically first in this chapter as a brief introduction with full analytical expressions given in Chapter 1.

- $s_a(t)$, which we shall call a type 1 signal, shown in Fig. I.1.3(a), represents a simple sinusoidal oscillation with a constant characteristic frequency f_c . For this type of signal, in some applications, a single characteristic f_c is sufficient to specify the whole signal. If the amplitude and phase are significant in the application, two more characteristics are needed to describe them.
- $s_b(t)$, which we shall call a type 2 signal, shown in Fig. I.1.3(b), is a “multicomponent,” signal comprising several “components” each of which is a sinusoid of type 1. Specification of a linear mixture of M sinusoids would require specification of the amplitudes a_k and frequencies f_k for $k = 1, \dots, M$, giving a total of $2M$ characteristics. For some applications, the initial phase ψ_k of each component may also be significant, giving a total of $3M$ characteristics. A variant of this signal can be obtained by defining a modified form of $s_b(t)$, where frequencies can have arbitrary start and stop times. Such signals are used in communications (e.g., frequency-hopping spread-spectrum communication).
- $s_c(t)$, which we shall call type 3, shown in Fig. I.1.3(c), has a single “component” of unit amplitude, whose local frequency $f_i(t)$ changes as time evolves. Examples include frequency modulated (FM) signals used in telecommunications and “vibroseis” signals used in geological exploration. Such a signal is described as monocomponent and nonstationary. It is characterized by the frequency-variation law $f_i(t)$, which can be approximated in some cases (e.g., linear FM) by half the rate of zero crossings in the time-domain representation of the signal. In this case, the dominant frequency at a certain time t is a linear function of time t , which we call “instantaneous frequency” $f_i(t)$.
- $s_d(t)$, which we shall call type 4, shown in Fig. I.1.3(d), is a weighted sum of M components, each of which is of type 3. Thus each component has an amplitude a_k and a time-varying frequency $f_k(t)$, and the signal is characterized by M coefficients a_k and M functions $f_k(t)$. Such a signal is described as multicomponent and nonstationary. This signal $s_d(t)$ is shown in the plot for two components, each of which has unit amplitude and a linear IF law.
- $s_e(t)$, which we shall call type 5, shown in Fig. I.1.3(e), is a type 4 signal ($s_d(t)$ in this case) to which noise has been added to the sum of the components; and, in addition, the amplitude a_k of each component has become random. The former type of noise is called *additive noise*. The latter is called *multiplicative noise*: the multiplier a_k becomes a random process $m_k(t)$, produced, for example, by adding noise to a_k . A simple model of such a random process affected by both additive and multiplicative noise is, for example: $x_k(t) = a_k(t) \cdot s(t) + \text{noise}_1(t)$ with $a_k(t) = \text{constant}_k + \text{noise}_2(t)$. To characterize the signal, we need not only the amplitudes and instantaneous frequency laws of the components, but also the statistics of the noise terms. In other words, we must be able not only to distinguish the components from each other in spite of their

²A signal is considered deterministic if its values are completely specified for any given time; it is not affected by any kind of random noise, whether additive or multiplicative. A noiseless signal can also be random.

³A random signal takes random values at any given time; it can be obtained simply by considering a deterministic signal and adding noise to it or replacing one of its characteristic constants (e.g., amplitude or initial phase) by a random variable.

**FIGURE I.1.3**

Time-domain representations of the six signals described in Section I.1.2. (a) $s_a(t)$: Sinusoidal signal. (b) $s_b(t)$: Sum of sinusoids. (c) $s_c(t)$: Monocomponent, nonstationary signal. (d) $s_d(t)$: Multicomponent, nonstationary signal (for this particular example, $M = 2$). (e) $s_e(t)$: Same as (d), but with additive and multiplicative noise. (f) $s_g(t)$: Sinc pulse with additive noise.

drifting amplitudes and frequencies but also to separate them from the noise. Multiplicative noise causes a convolution of the signal with the noise in the frequency domain and is therefore harder to remove than additive noise, which is merely added to the signal in both time and frequency. An example of additive noise is the distortion created by the internal instrumental noise of microphones (which is uncorrelated to the speech signal). An example of multiplicative noise is the ambient or reverberation noise caused by delayed replicas of speech captured by microphones (which is correlated to the speech signal).

- $s_g(t)$, which we shall call type 6, shown in Fig. I.1.3(f), is a sinc function with additive white Gaussian noise. This type of signal is useful as a basic model of impulsive noise or electroencephalogram (EEG) spikes.

As mentioned above, knowing the signal class (or type) is important in algorithm development. From the knowledge of the signal type, we can select the signal characteristics that we wish to estimate, and design DSP tools appropriate for the signal analysis. An example of specific signal is the frequency shift keying (FSK) signal used in digital communications (see Section 13.5 for more details).

I.1.3 FREQUENCY DOMAIN REPRESENTATION OF TYPICAL SIGNALS

The Fourier theorem indicates that, under some conditions,⁴ any bounded signal can be written as a weighted sum of sines and cosines (or complex exponentials) at different frequencies f . The weights are given by the complex coefficient $S(f)$. If $s(t)$ is periodic, we obtain a line spectrum. If it is nonperiodic, then these complex exponentials occur at a continuum of frequencies f weighted by “ $S(f)$ ” according to the following integral expression:

$$s(t) = \int_{-\infty}^{\infty} S(f) e^{j2\pi ft} df. \quad (\text{I.1.1})$$

This equation shows that the value of $s(t)$ at time t is determined from the values of $S(f)$ at *all* f from $-\infty$ to $+\infty$. Equation (I.1.1) is commonly known as the inverse Fourier transform (IFT) of $S(f)$ [1,2].

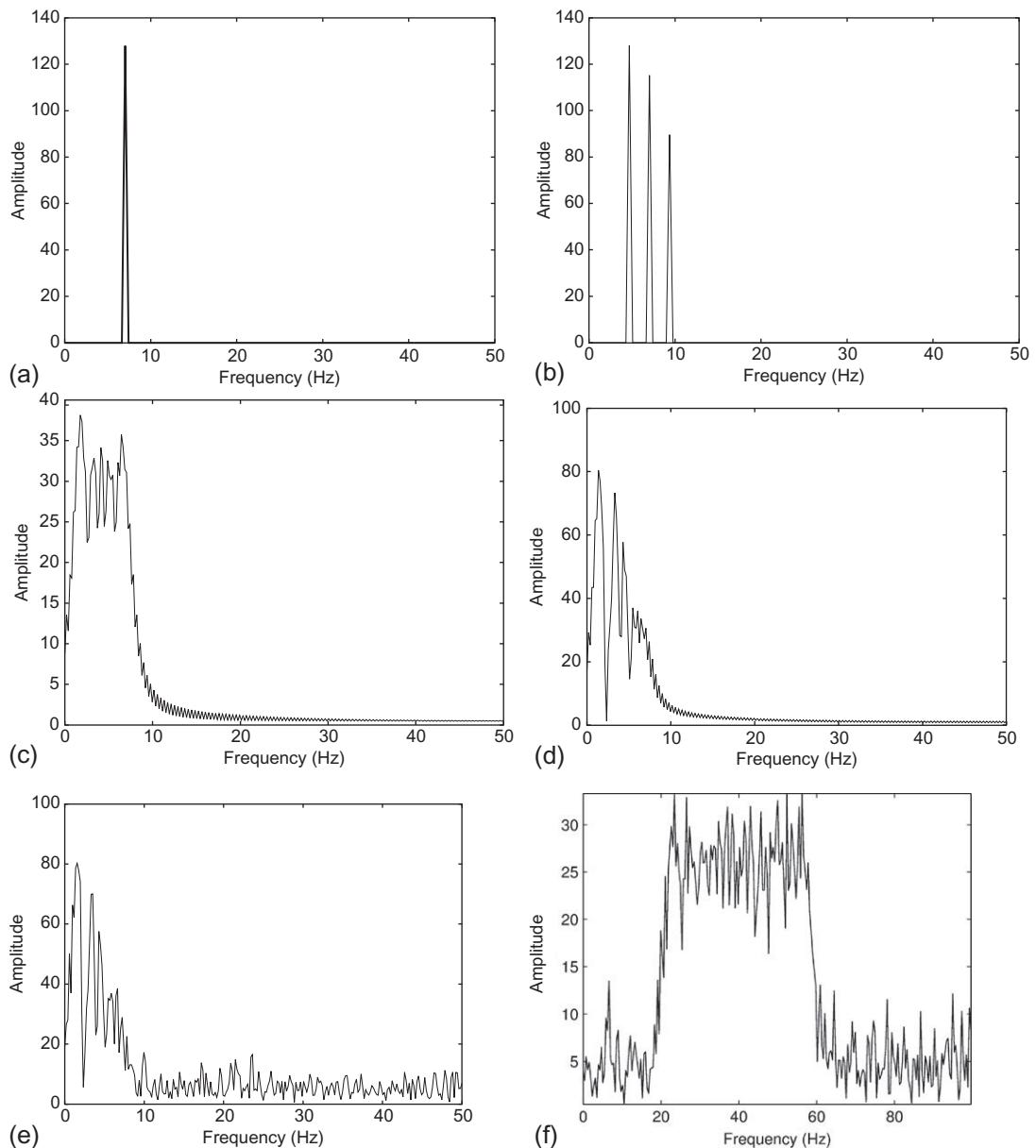
The weights $S(f)$ in Eq. (I.1.1) define a frequency domain representation for the deterministic signal $s(t)$; it allows the evaluation of periodicities in the signal (the unit is the number of oscillations per second, expressed in Hz). This may be obtained from $s(t)$ by taking the Fourier transform (FT):

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt. \quad (\text{I.1.2})$$

The spectral representation, $S(f)$, indicates which frequencies are present in the signal, their relative magnitudes, minimum frequency, maximum frequency, and the bandwidth formed by their difference.

The six signal types presented earlier in the t domain can be represented in the f domain as shown in Fig. I.1.4; (a) and (b) show that sinusoids in the t domain give rise to impulse (or delta) functions in the f domain (“line spectra”).

⁴More precisely, any signal that satisfies the Dirichlet conditions has an FT.

**FIGURE I.1.4**

Frequency-domain representations of the six signals in Fig. I.1.3. Only magnitudes (not the phases) are shown. In plot (f), the magnitude is squared, showing how the noise energy is distributed over the frequency axis.
 (a) $|S_a(f)|$, (b) $|S_b(f)|$, (c) $|S_c(f)|$, (d) $|S_d(f)|$, (e) $|S_e(f)|$, and (f) $|S_f(f)|^2$.

For $s_a(t)$ or $s_b(t)$, the FT is the most appropriate tool for a sparse representation of the signal. There is no point using more complex tools, since the FT will perform the decomposition of the signal into its harmonic components adequately.⁵

For $s_c(t)$ and $s_d(t)$, however, the FT is not adequate, because it assumes the sinusoidal components exist at all t , when, here, in fact the frequency is continuously shifting (for the single component of $s_c(t)$ and both components of $s_d(t)$).

Neither is the FT adequate for $s_e(t)$; this raises the same issues as $s_d(t)$, but it is further complicated by noise. The FT-based power spectral density (PSD) adequately describes the additive noise, whose characteristics do not change with time. But it does not fully describe the multiplicative noise, whose frequency content varies with time because each noise term modulates a component whose frequency varies with time (such random modulation may also be caused by a movement such as helicopter blades).

For $s_g(t)$, the FT fails to show the *transient* nature of the signal—that is, the precise localization of the signal in time.

Figure I.1.5 shows a real-life example of a type 4 signal—a sound emitted by a bat. This signal is more complex than that in Fig. I.1.1 as the different components of the signal observed in the (t, f) domain vary, indicating that the (t, f) relationships for each component need to be known in order to identify the signal precisely.

I.1.4 POSITIVE AND NEGATIVE FREQUENCIES

Mathematically, the frequency f ranges from $-\infty$ to $+\infty$; but in Fig. I.1.4, the plots show only the positive frequencies. Why? The intuitive explanation is that f represents the number of oscillations per second, which is expected to be positive. A more mathematical explanation relies on an important property of the FT, namely that if $s(t)$ is real, its FT $S(f)$ has *Hermitian symmetry*; that is, $S(f) = S^*(-f)$. This means that the real part is even and the imaginary part is odd, or equivalently that the magnitude is even and the phase is odd. Thus, the information contained in the negative frequencies is a duplication

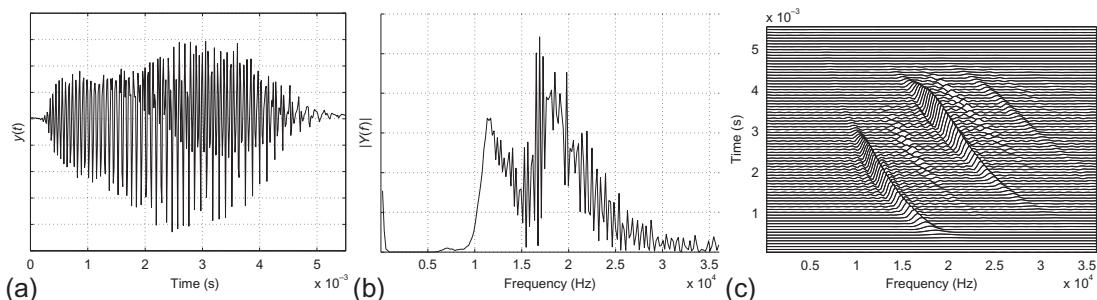


FIGURE I.1.5

Echolocation pulse emitted by a bat and (a) its time-domain representation, (b) its magnitude spectrum, and (c) its time-frequency representation.

⁵The FT of a signal $s(t)$ is the inner product of $s(t)$ and the complex exponential $e^{j2\pi f_k t} = \cos 2\pi f_k t + j \sin 2\pi f_k t$. This results in a harmonic decomposition of the signal into its components at frequencies f_k .

of the information contained in the positive frequencies. Therefore, in practice, one prefers to deal only with positive frequencies. This is achieved by introducing a particular complex signal $z(t)$ called the “analytic signal,” which contains only positive frequencies.

Definition I.1.1. A signal $z(t)$ is said to be *analytic* iff

$$Z(f) = 0 \quad \text{for } f < 0, \quad (\text{I.1.3})$$

where $Z(f)$ is the FT of $z(t)$, written symbolically as $Z(f) = \mathcal{F}\{z(t)\}$.

A *real* signal $s(t)$ cannot have only positive frequencies, because the Hermitian symmetry ($S(f) = S^*(-f)$) would fail Eq. (I.1.3); so, $z(t)$ cannot be real.

In the frequency domain, Eq. (I.1.3) expresses the fact that the spectra of the real and imaginary parts of $z(t)$ are of opposite sign for negative f (nullifying each other when added to obtain the spectrum $Z(f)$), but identical for positive f (see Chapter 1 for more details). An analytic signal is typically expressed as $z(t) = a(t)e^{j\phi(t)}$, with its real part $s(t) = \text{Re}\{z(t)\}$ [1,5]. Thus, there is a relation of equivalence between the real signal $s(t)$ and the associated analytic signal $z(t)$. The reason the analytic associate, $z(t)$, is used in practice rather than the real signal $s(t)$ is discussed in Chapter 1 (where we will see that the imaginary part of the analytic signal $z(t)$ is related to the real part $s(t)$ by the Hilbert Transform) [5].

I.1.5 RELATIONSHIP BETWEEN t DOMAIN AND f DOMAIN REPRESENTATIONS

I.1.5.1 Fourier transform pairs

The FT (Eq. I.1.2) and the IFT (Eq. I.1.1) are said to form an FT pair, symbolically represented as $s(t) \xleftrightarrow{tf} S(f)$, which links the t domain to the f domain. For convenience, a list of basic common signals and their FT is given in Table I.1.1. Note that the first signal $\delta(t)$ is the Dirac delta function, whose defining features are $\delta(t) = 0$ for $t \neq 0$, and $\int_{-\infty}^{\infty} \delta(t) dt = 1$, or equivalently

$$\int_{-\infty}^{\infty} \delta(\tau)x(t - \tau) d\tau = x(t), \quad \forall t. \quad (\text{I.1.4})$$

I.1.5.2 Energy distribution of deterministic signals

The energy of a deterministic signal $s(t)$ can be distributed over t using the instantaneous power $|s(t)|^2$; it can also be distributed over f using its equivalent, the energy spectral density $|S(f)|^2$ (which is the squared modulus of the FT given by Eq. (I.1.2)). These two quadratic representations are related by Parseval’s theorem, which states that

$$\int_{-\infty}^{\infty} |s(t)|^2 dt = E_s = \int_{-\infty}^{\infty} |S(f)|^2 df, \quad (\text{I.1.5})$$

where E_s denotes the energy of the signal.

Equation (I.1.5) reflects the fact that the FT is a unitary transform and as a consequence, the norm ($|\cdot|^2$) is preserved when shifting from the t domain to the f domain and vice versa. These equations show also that $|s(t)|^2$ is an energy time density, while $|S(f)|^2$ is an energy frequency (or spectral) density. Thus, the classical representations of a signal can show how the signal energy is distributed over time t (left-hand equality) or over frequency f (right-hand equality). But they cannot show how the energy

Table I.1.1: Fourier Transform Pairs of Selected Basic Signals

| Signal Name | Time Domain $s(t)$ | Frequency Domain $S(f)$ |
|-----------------------|---|--|
| Delta or Dirac | $\delta(t)$ | 1 |
| Heaviside step | $u(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$ | $\frac{1}{2} \left[\delta(f) - \frac{j}{\pi f} \right]$ |
| Sign | $\text{sgn}(t) = \begin{cases} -1 & \text{if } t < 0 \\ 0 & \text{if } t = 0 \\ +1 & \text{if } t > 0 \end{cases}$ | $-\frac{j}{\pi f}$ |
| Exponential | $e^{j2\pi f_0 t}$ | $\delta(f - f_0)$ |
| Sine | $\sin(2\pi f_0 t)$ | $\frac{\delta(f-f_0)-\delta(f+f_0)}{2j}$ |
| Cosine | $\cos(2\pi f_0 t)$ | $\frac{\delta(f-f_0)+\delta(f+f_0)}{2}$ |
| Rectangular or box | $\text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1 & \text{if } t \leq T/2 \\ 0 & \text{if } t > T/2 \end{cases}$ | $T \text{sinc}(Tf) = \frac{\sin(\pi Tf)}{\pi f}$ |
| Sinc or cardinal sine | $\text{sinc}(Bt) = \frac{\sin(\pi Bt)}{\pi Bt}$ | $\frac{1}{ B } \text{rect}\left(\frac{f}{B}\right)$ |

of a nonstationary signal like $s_c(t)$ is distributed simultaneously over t and f , because they do not show the instantaneous frequency $f_i(t)$ that links the two variables—as discussed next in [Section I.2.2](#) and in more detail in [Chapter 1](#).

The question that then arises is: “How can we define such a joint (t,f) density?” Before answering this, let us consider, for completeness, the case of random signals.

I.1.5.3 Power distributions of random signals

For a complex random signal $z(t)$, the symmetrical *autocorrelation* function is

$$R_z(t, \tau) = \mathcal{E} \left\{ z \left(t + \frac{\tau}{2} \right) z^* \left(t - \frac{\tau}{2} \right) \right\}, \quad (\text{I.1.6})$$

where $\mathcal{E}\{\cdot\}$ denotes the expected value, also known as the *ensemble* average taken over all possible realizations of the random process. For random signals, the pair of FTs $z(t) \longleftrightarrow Z(f)$ is replaced by $R_z(\tau) \longleftrightarrow S_z(f)$, where $R_z(\tau)$ is the autocorrelation function and $S_z(f)$ is the PSD. If $z(t)$ is *stationary* in the wide sense, its autocorrelation function is independent of t , and its *PSD* is defined as [6]

$$S_z(f) \triangleq \lim_{T \rightarrow \infty} \mathcal{E} \left\{ \frac{1}{T} \left| \int_{-T/2}^{T/2} z(t) \text{rect}(t/T) dt \right|^2 \right\}, \quad (\text{I.1.7})$$

where $\text{rect}(t/T)$ is defined as 1 for $-T/2 \leq t \leq T/2$, and 0 elsewhere, so that $z(t)$ is time-limited to $-T/2 \leq t \leq T/2$ before its FT is taken. According to the *Wiener-Khintchine theorem*, the PSD $S_z(f)$ equals then the FT ($\tau \rightarrow f$) of the autocorrelation function $R_z(\tau)$ (this relationship is pursued further in [Section 2.2.1](#)).

If $z(t)$ is *ergodic* in the wide sense,⁶ the ensemble average in Eq. (I.1.6) is equal to the time average, so that

$$R_z(\tau) = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) dt \right]. \quad (\text{I.1.8})$$

In taking the time average, time t is “integrated out,” so that the time average is independent of t . Its FT gives the PSD and thus, *ergodicity implies stationarity* (but not vice versa).

The total signal power may be expressed as

$$P_z = R_z(0) = \int_{-\infty}^{\infty} S_z(f) df. \quad (\text{I.1.9})$$

These analysis tools are suitable for many problems related to stationary signals.

I.1.6 LIMITATION OF t DOMAIN OR f DOMAIN REPRESENTATIONS

The amplitude or power spectra of the signals (deterministic or random) shown in Fig. I.1.4 gives no indication of how the frequency content of a signal changes with time, information that is important when one deals with FM signals, or other kind of nonstationary signals. This frequency variation often contains critical information about the signal and process studied in applications. The magnitude spectrum cannot therefore suitably represent signals such as $s_c(t)$ to $s_g(t)$, above.

The FT is inadequate for *nonstationary* signals because the spectral representation $S(f)$ requires the variables of time, t , and frequency, f , to be mutually exclusive. The evaluation of one value of the spectrum $S(f)$ at frequency f requires the knowledge of the values of $s(t)$ at all t (i.e., from $-\infty$ to $+\infty$); however, the nonstationarity of a signal means that there is time-dependency of the signal frequency spectrum or equivalently time-dependency of the signal autocorrelation function.

This limitation of “classical” spectral representations is well illustrated by the fact that we can find totally different signals related to different physical phenomena, for example, a linear-FM signal and a limited impulse, which have similar “spectra” (i.e., magnitude spectra), as examined in detail with the following example.

Example I.1.2. Consider a type 3 signal defined as a finite-duration linear-FM (LFM) signal expressed as:

$$s_A(t) = \text{rect}\left(\frac{t-T/2}{T}\right) \cos\left[2\pi\left(f_0 t + \alpha \frac{t^2}{2}\right)\right], \quad (\text{I.1.10})$$

which equals $\cos\left[2\pi\left(f_0 t + \alpha \frac{t^2}{2}\right)\right]$ for $0 < t < T$, and 0 elsewhere, and the finite-duration modulated sinc function

$$s_B(t) = \text{rect}\left(\frac{t-T/2}{T}\right) \frac{\sin(\pi Bt)}{\pi t} \cos(2\pi f_{ct}); \quad (\text{I.1.11})$$

$s_A(t)$ is an LFM signal, commonly used in radar and seismic applications, and analogous to a musical note with a steadily rising pitch. Here, f_0 is its starting frequency, T its duration, and $\alpha = B/T$ its frequency sweep rate (rate of change of frequency), so that its bandwidth is $B = \alpha T$ and its center

⁶Other authors use the equivalent terminology of mean ergodic processes and auto correlation ergodic processes [7].

frequency is $f_c = f_0 + B/2$. The parameters are chosen in such a way that the bandwidth-duration product⁷ is large ($BT \gg 1$). The second signal $s_B(t)$ is defined in terms of the same B and f_c .

Let us consider two signal recordings $x_A(t)$ and $x_B(t)$, which are realizations of the two random processes $X_A(t) = s_A(t) + N_A(t)$ and $X_B(t) = s_B(t) + N_B(t)$, respectively, where $N_A(t)$ and $N_B(t)$ are zero-mean white Gaussian noise processes with standard deviations σ_A and σ_B , respectively. Thus $x_A(t)$ where $x_B(t)$ are noisy versions of $s_A(t)$ and $s_B(t)$, and $x_B(t)$ is of type 6.⁸ The signal-to-noise ratio (SNR) is chosen high enough so that the influence of $N_A(t)$ and $N_B(t)$ is small. Figure I.1.6 shows the two signals and their spectra.

Apart from the noise, the sinc function and the LFM chirp have similar magnitude spectra. To differentiate between them, we would need to know their respective phases. The sinc function has all frequencies arriving at once (zero phase), while the chirp signal has the frequencies arriving in a certain order: the time of arrival is a linear function of frequency. The SD (also known as *time delay*⁹) is a measure of the “time of arrival” of the frequencies, which can be read directly from a joint TFR (in short a (t,f) representation).

Exercise I.1.3. (i) Write the code to generate a chirp signal $x_A(t) = s_A(t) + n_A(t)$ for values of B and T such that: $BT = 2.5, 10$, and 100 . (ii) Plot their magnitude and phase spectra, and compare with Fig. I.1.6. What is the effect of the parameter BT on the representation?

To deal with signal types such as $s_c(t)$ to $s_g(t)$, we need a representation that can reveal not only which frequencies are present, but *when* they are present.

I.2 JOINT TIME-FREQUENCY REPRESENTATIONS

I.2.1 UNCOVERING HIDDEN INFORMATION USING (t,f) REPRESENTATIONS

Revealing the time and frequency dependence of a signal, such as $s_A(t)$ in Example I.1.2, can be achieved by plotting a joint TFR using a time-frequency distribution (TFD),¹⁰ as shown in Fig. I.2.1. This is another example of two different signals with similar magnitude spectra. Each signal contains threeLFMs whose start and stop times are different. The differences are shown only roughly in the t domain (left-hand traces) and are hardly visible in the f domain (bottom traces), but appear clearly in the (t,f) representation, therefore allowing precise measurements of actual frequencies *and* their epochs (the times at which they appear).

The start and end times of each component are clearly identifiable in the (t,f) domain,¹¹ as is the time variation of their frequency content (this information cannot be easily retrieved directly from $s(t)$ or $S(f)$; it is lost when the FT is squared in modulus and the phase spectrum is thereby discarded). This internal organization of the signal features includes such details as the times at which the signal power

⁷The various definitions and interpretations of the bandwidth-duration product BT are developed in more detail in Chapter 1.

⁸Indeed, $x_B(t)$ in Fig. I.1.6 is identical to $s_g(t)$ in Figs. I.1.3(f) and I.1.4(f).

⁹We will use the following terminology: “spectral delay” for signals such as the chirp and “group delay” when the signal can be conceived as a sum of several time-varying signals such as the sum of two or more chirps. For each chirp, we can calculate an SD. For the overall signal, the group delay represents the overall combined group effect.

¹⁰The representation shown here uses the B-distribution described in Section 2.7.5 and Table 3.3.2 [3].

¹¹Except that there is a poor resolution at the beginning and end in this example, which can be improved with methods presented in later chapters.

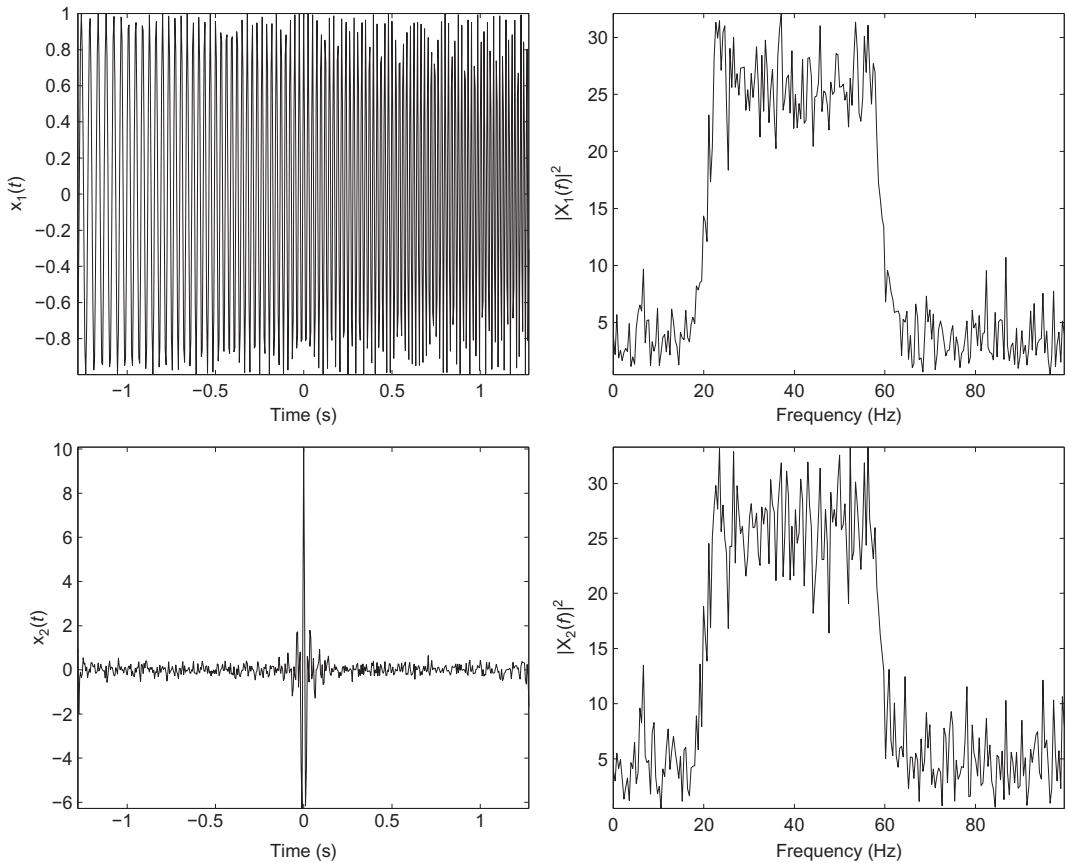


FIGURE I.1.6

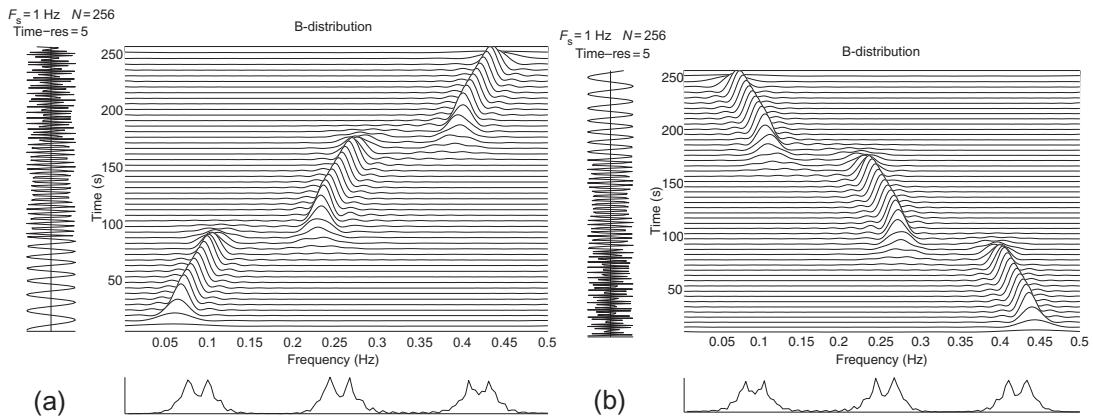
A noisy LFM signal with an SNR of 16dB (upper left) and its energy spectrum (upper right), and a noisy modulated sinc function with the same SNR (lower left) and its energy spectrum (lower right).

is above or below a particular threshold, in addition to the order of appearance in time of the different frequencies present in the signal. The visual appearance of this information makes the concept of a joint (t,f) signal representation useful for practical applications.

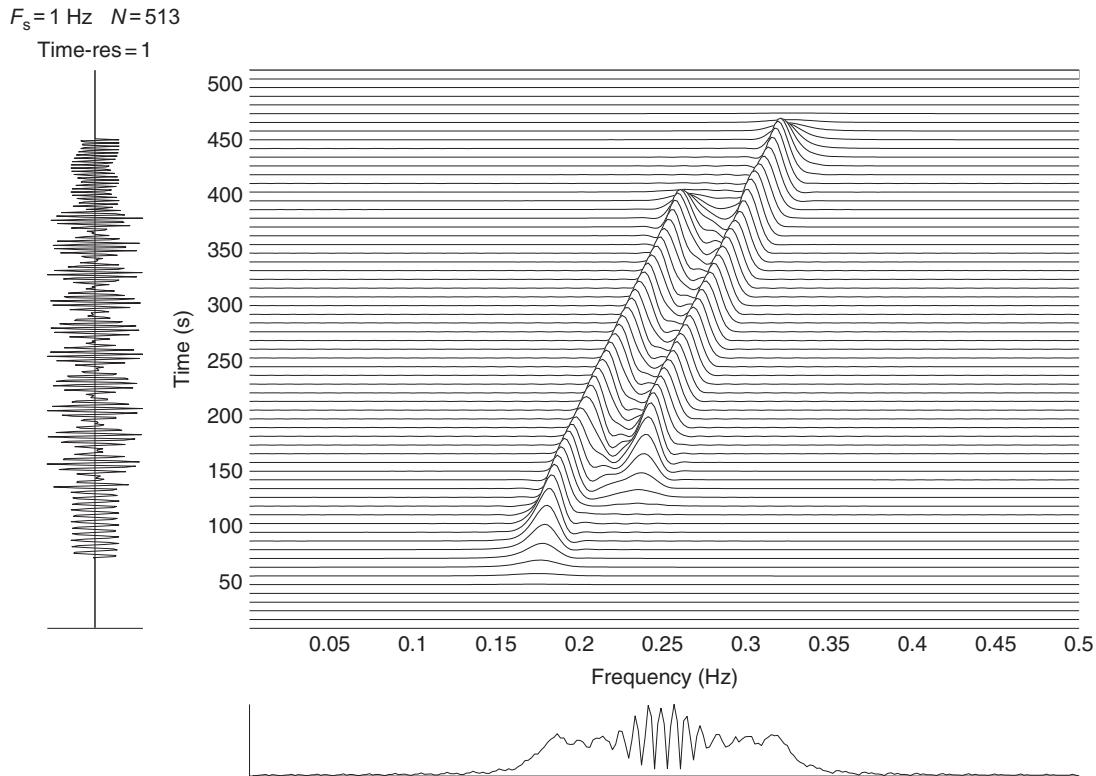
Another significant reason to use TFRs and TFDs of signals is that they also reveal whether the signal is monocomponent or multicomponent, a fact that cannot easily be revealed by conventional spectral analysis, especially when individual components are also time-varying, as in Fig. I.2.2.¹²

In brief, for nonstationary signals, the (t,f) representation is more adequate and easily interpretable. It is essentially a time-varying spectrum (or time-dependent PSD) [1,8].

¹²The terms *monocomponent* and *multicomponent* can be best understood by observing Figs. I.1.1, I.1.5(c), I.2.1, and I.2.2, and noting that all show several (t,f) areas with significant energy. In practice, we will describe a nondiscontinuous pattern in the (t,f) domain as a component. A desired property of (t,f) methods is to allow a precise estimation of such patterns and their (t,f) boundaries. In this context, we can observe that the signal in Fig. I.2.2 has two clear components with a (t,f) shift between them.

**FIGURE I.2.1**

Time-frequency representation of two different three-component signals.

**FIGURE I.2.2**

Time-frequency representation of a type 4 multicomponent signal (sum of two LFM signals with close parallel FM laws).

Exercise I.2.1. Compute the spectrum of an LFM signal; form its amplitude spectrum and phase spectrum; then replace its phase spectrum by zero, and invert the new spectrum to obtain a new signal. Explain what you observe.

I.2.2 CONCEPT AND MEANING OF A (t, f) REPRESENTATION

TFSP is a “natural” extension of both time domain and frequency domain processing, using representations of signals in a complete space that can display “all” the signal information received from one sensor in a more accessible way [9]. Such a representation is intended to provide a *distribution* of signal energy over both time t and frequency f simultaneously. For this reason, the representation is commonly called a TFD.

A concept intimately related to joint TFR is that of IF and SD. The IF, the local maximum at a given time, corresponds to the frequency of a sine wave that locally (at a given time) fits best the signal under analysis. The SD, as defined in the previous section, is a measure of the “order of arrival” of the frequencies.

The TFD is expected to exhibit the (t, f) law of each signal component in the (t, f) domain, thereby making the estimation of the IF $f_i(t)$ and SD $\tau_d(f)$ easier. It could also provide information about the amplitude and duration of each component and its instantaneous bandwidth $B_i(t)$, that is, its spectral spread around the IF, as shown in Fig. I.2.3.

The TFD is often expected to satisfy a certain number of properties that are intuitively desirable for a practical analysis. Let us denote by $\rho_z(t, f)$ such a TFD. In order to generate a TFR of a signal $z(t)$, it is expected to satisfy the following desirable properties [3,4]:

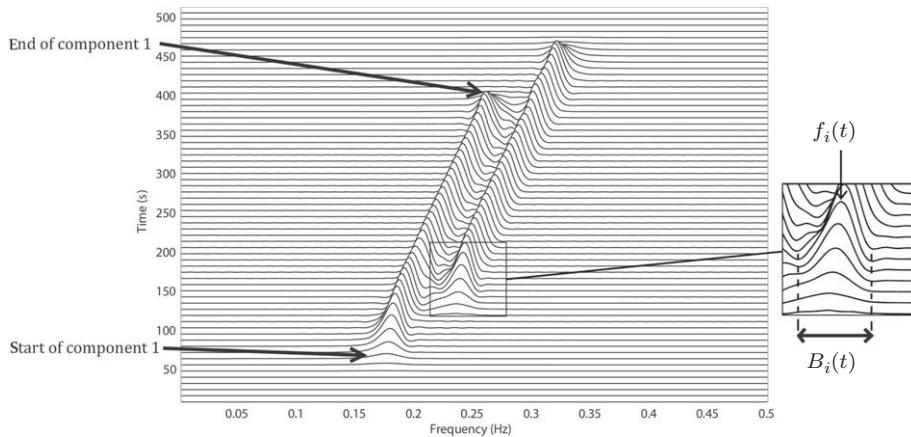


FIGURE I.2.3

Illustration of how a time-frequency distribution is expected to show the beginning, the end, the instantaneous frequency $f_i(t)$, and the instantaneous bandwidth $B_i(t)$ of each component of a multicomponent signal.

- P1:** The TFD should be real and its integration over the whole (t,f) domain should give the total signal energy E_z :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_z(t,f) dt df = E_z. \quad (\text{I.2.1})$$

This equation effectively merges the two parts of Eq. (I.1.5), giving a distribution of energy over both t and f .

- P2:** The energy in a certain region R in the (t,f) plane, E_{zR} , should be expressible as in Eq. (I.2.1), except that the integral is over the region R .
- P3:** At any time, the frequency corresponding to the peak of the (t,f) representation of a monocomponent signal should be equal to the IF of the signal, thus, showing clearly the law that relates t and f in terms of energy localization; $f_i(t) = \arg[\max \rho_z(t,f)]$;

Naturally, the following questions arise: Can we design a TFD that meets the (t,f) specifications listed above? If yes, how can we do it? What are the significant signal characteristics and parameters that will impact the construction of a joint TFR? How do these relate to the TFD? How do we obtain them from the TFR? The next section details some of the most relevant (t,f) characteristics and features necessary to fully understand the meaning of a TFD, and use such (t,f) tools effectively.

I.2.3 SIGNAL CHARACTERISTICS IN (t,f) REPRESENTATIONS

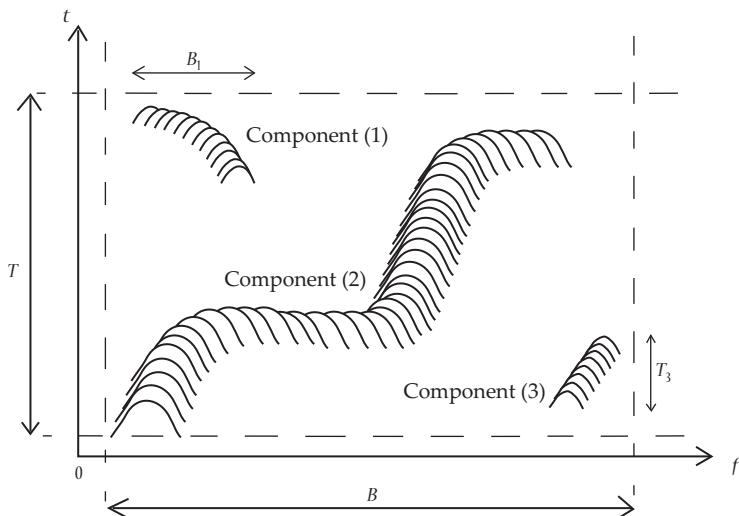
In essence, the important signal characteristics that can be represented in the two-dimensional (t,f) domain include the following [3,4]:

1. frequency-domain:
 - bandwidth, $B = f_{\max} - f_{\min}$ for each component;
 - ranges and relative amplitudes of frequencies within the instantaneous bandwidth $B_i(t)$;
 - the instantaneous frequency $f_i(t)$, for each component;
 - the instantaneous bandwidth $B_i(t)$, for each component (see Fig. I.2.3);
 - the total signal instantaneous bandwidth (see Fig. I.2.4).
2. time-domain:
 - duration, $T = t_{\text{end}} - t_{\text{start}}$ for each component;
 - time-ranges and relative strengths of transients between t_{start} and t_{end} ;
 - SD, $\tau_d(f)$, which is a frequency-dependent time-delay.¹³
3. combined time and frequency domain:
 - bandwidth-duration (BT) product;
 - number of components in the (t,f) domain.

Figure I.2.4 shows how these characteristics are represented by a hypothetical TFD in the case of a three-component signal.

In the case of a monocomponent signal, the IF $f_i(t)$ assigns one frequency to each time, and one or multiple times to each frequency [1,2]. This complements the information provided by the time-domain representation $s(t)$. In the audio frequency range, it describes the signal as it would be perceived by a

¹³We will see in Chapter 1 that the parameters IF and SD are related.

**FIGURE I.2.4**

Illustrative representation of a multicomponent signal in the (t, f) domain, showing the signal bandwidth B and duration T , the bandwidth B_1 of component (1), and the duration T_3 of component (3). The IF of one component at time t can be read off the f axis, and the SD of one component at frequency f can be read off the t -axis.

listener; for example, as a sound sweeping from low frequency to high frequency, or vice versa, or alternating in frequency, noting that the frequency scale may not necessarily be linear over the full range to reflect the characteristics of human perception.

Additional (t, f) features that can be important for signal identification, detection, and classification include, for example [10]:

1. time-frequency concentration (see [Section 12.6](#)) [11];
2. time-frequency complexity using, for example, entropy as a measure (see [Section 7.7](#)); and
3. time-frequency flux (see [Section 12.6](#)).

I.2.3.1 Finite-duration, finite-bandwidth signals

In real life, the signal analyst expects to deal with signals that have both a finite duration T and a finite bandwidth B . It can be shown mathematically that these expectations are incompatible; however, a class of “practical” signals can be defined to ensure that the error made in assuming both band- and time-limited functions is small [1]. This is achieved by imposing the following four conditions on the signal: (1) finite energy; (2) finite duration; (3) BT large (>10); and (4) bounded amplitude (i.e., no infinite amplitudes).

As B is the Nyquist sampling rate required to represent the signal (e.g., if the spectrum is confined to $-B/2 < f < B/2$, then B is twice the maximum absolute value of frequency), BT is the required

number of samples. Thus, BT is the number of “degrees of freedom” in the signal and, for a given SNR, is proportional to the information content [1,9].

I.2.3.2 Physical interpretation of TFDs

Most TFDs used in practical applications can take negative values as they are not necessarily positive-definite, so they do not represent an instantaneous energy spectral density at any arbitrary (t,f) location. For example, one particular TFD discussed in Section 2.5.1 is defined as the time derivative of the squared magnitude of the “running spectrum,” that is, spectrum of the signal from $-\infty$ to time t :

$$\rho_z(t,f) = \frac{\partial}{\partial t} \left| \int_{-\infty}^t z(u) e^{-j2\pi fu} du \right|^2. \quad (\text{I.2.2})$$

The truncation of the signal at time t causes oscillations in the f domain, with the result that the distribution takes both positive and negative values.

To relate a distribution $\rho_z(t,f)$ to physical quantities, we interpret $\rho_z(t,f)$ as a measure of energy flow through the spectral window $(f - \Delta f/2, f + \Delta f/2)$ during the time interval $(t - \Delta t/2, t + \Delta t/2)$ [12]. The signal energy localized in this (t,f) window of dimensions Δt by Δf around a point (t_0, f_0) is thus given by

$$E_{\Delta t, \Delta f} = \int_{t_0 - \Delta t/2}^{t_0 + \Delta t/2} \int_{f_0 - \Delta f/2}^{f_0 + \Delta f/2} \rho_z(t,f) dt df. \quad (\text{I.2.3})$$

The larger the window, the more likely $E_{\Delta t, \Delta f}$ will correspond to a true measure of physical energy. The window should be chosen large enough to satisfy the Heisenberg/Gabor uncertainty relation of standard deviations for basic atoms:¹⁴

$$\Delta t \Delta f \geq 1/(4\pi), \quad (\text{I.2.4})$$

as proven below and as discussed in Example 1.2.5, p. 48 [9].

I.2.3.3 The uncertainty principle in the time-frequency domain

Let $x(t)$ be a zero-mean deterministic signal and $X(f)$ its FT. Let us try to localize simultaneously $x(t)$ and $X(f)$. Let us assume that $x(t)$ is a finite energy signal centered around $t = 0$; and $X(f)$ is centered around $f = 0$. Let $p(t)$ be the energy density versus time t : $p(t) = \frac{|x(t)|^2}{\|x\|^2}$, where, $p(t)$ is real ≥ 0 and has the dimension of a probability density and where $\|x\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 dt$. The “time interval where the signal is concentrated” is given by Δt , which defines an equivalent duration given by: $(\Delta t)^2 = \frac{\int_{-\infty}^{\infty} t^2 |x(t)|^2 dt}{\int_{-\infty}^{\infty} |x(t)|^2 dt}$. We have the same relationship in the frequency domain with Δf equivalent bandwidth, given by: $(\Delta f)^2 = \frac{\int_{-\infty}^{\infty} f^2 |X(f)|^2 df}{\int_{-\infty}^{\infty} |X(f)|^2 df}$.

¹⁴Equation (I.2.4) is perhaps best known for its corollary in quantum mechanics. If we write x (position) for t , and k (wave number) for $2\pi f$, we obtain $\Delta x \Delta k \geq 1/2$. De Broglie’s “matter wave” relation may be written $k = 2\pi p/h$, where p is the momentum and h is Planck’s constant. Making this substitution for k in the above equation, we obtain $\Delta x \Delta p \geq \frac{h}{4\pi}$, which, for some readers, may be more familiar than Eq. (I.2.4).

Let us relate these two quantities. We have: $\frac{dx(t)}{dt} = \dot{x}(t) \Leftrightarrow 2j\pi f X(f)$, which leads to: $\int_{-\infty}^{\infty} |\dot{x}|^2 dt = \int_{-\infty}^{\infty} 4\pi^2 f^2 |X(f)|^2 df$, and then: $\Delta t \Delta f = [\int_{-\infty}^{\infty} |tx(t)|^2 dt]^{1/2} \cdot \left[\frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} |\dot{x}(t)|^2 dt \right]^{1/2} \cdot \|x\|^{-2}$. Note that this derivation shows that the product $\Delta t \cdot \Delta f$ can be computed without using a Fourier transformation.

Using Schwartz' inequality and integrating by parts yields Eq. (I.2.4). If the signal $x(t)$ is centered around t_0 , Δt is defined about t_0 , and the same applies if $X(f)$ is centered around $f = f_0$. Gabor's relation remains the same.

I.2.3.3.1 Consequences of the uncertainty principle for (t, f) relationships

- $\Delta t \Delta f \geq \frac{1}{4\pi}$ implies that Δt and Δf cannot both approach zero.
- In communications, the bandwidth required to transmit a message increases when the speed of information transmission is increased.
- There is an analogy with quantum mechanics where the position and velocity of a particle obey: $\Delta p \Delta x \geq \frac{\hbar}{2}$, which is the Heisenberg's uncertain principle.
- If we use another definition of bandwidth and duration, we will still obtain an inequality but with a different limit.
- Given a linear filter $h(t) \Leftrightarrow H(f)$, if Δt represents the memory of the system, and Δf is its bandwidth, then we have: $\Delta t \Delta f \geq \frac{1}{4\pi}$.
- In the case of stationary random process: $R_x(\tau) \Leftrightarrow S_x(f)$. Here, Δt represents the statistical memory or “correlation time” and Δf is the spectral bandwidth; this again results in $\Delta t \Delta f \geq \frac{1}{4\pi}$.
- The equality $\Delta t \Delta f = \frac{1}{4\pi}$ is obtained for the Gaussian signal only, so that:

$$x(t) = k_1 e^{-\frac{t^2}{\tau^2}} \Leftrightarrow k_2 e^{-\pi^2 \tau^2 f^2}. \quad (\text{I.2.5})$$

This signal is in essence its own FT. A consequence is that the Gaussian signal has the minimum required bandwidth for a given duration of message.

I.2.3.4 Wavelets and time-scale methods

There are also a number of methods that are often called time-frequency methods but are in fact time-scale methods, which can be used for the analysis of nonstationary signals and especially transient signals. These methods are related to the quadratic class of TFDs presented in Chapter 3, and the link is described in detail in Section 4.1, as well as other subsequent chapter sections. These time-scale methods are often derived by translating, modulating, and scaling a basis function (called mother wavelet) that has a definite time and frequency localization, as will be described in Section 2.7.6 [6].

These methods are based on a linear decomposition of a given signal by translating, modulating, and scaling a basis function. These linear (t, f) decompositions preserve the phase information, so they allow estimation of signal parameters such as local phase and local frequency from the derivative of the phase [13]. The resolution performance of these methods for a given signal depends on the choice of parameters selected for the basis function [14].

I.2.4 HISTORICAL NOTE: THE EDITOR'S VIEW

At the request of readers who asked about the history of the field, the editor/author decided to give the most general description of these developments as they appeared from his perspective.

The editor/author was introduced to the field while an undergraduate student in Lyon, France in the year 1977, in the context of an electronics engineering final-year project; the topic was specifically defined in recognition of the editor/author's interests and achievements at the time as a student in related subjects. The first findings were found to be significant and partially published in French in local conferences [15] and in [16].

That work was done in parallel with another team in the Netherlands, unaware of each other (the World Wide Web and Google did not exist then), and we reached similar conclusions. The key literature that formed the basis of our research work included the works of Gabor [9], Wigner [17], Ville [18], Rihaczek [19], and Flanagan [20]. The total number of directly relevant papers then available to a typical student researcher was about 20, and it took great time and effort to get them through library exchanges or by sending a handwritten letter to the authors.

In those years, the work was mostly theoretical and analytical, focusing on mathematical derivations, and the Wigner-Ville distribution (WVD) was seen as a kind of correlation function. One of the editor/author's contributions was then to see it as the FT of a signal kernel. His other contribution was to write a program in A Programming Language (APL) to implement the WVD, which was seen to be the first program at the time. To do this, in 1978, he benefitted from the access to an IBM supercomputer (of the time).

The editor/author then continued the project as part of a Masters and then Doctorate in the context of a geophysical application within the elf-Aquitaine oil company. He was then invited to take an academic position at the INSA, Lyon (1982-1983), at which time he became aware of another research group in the United States that produced several students who in turn established other labs elsewhere in the United States focusing on the (t,f) field mostly remaining in North America.

At that time, the main focus was the study of the WVD, smoothed versions of it, its implementation, use for (t,f) filtering and validation on applications.

In December 1983, the editor/author moved to Australia and took an academic position at the University of Queensland, Brisbane.

From that point on, while in Queensland University, the editor/author significantly promoted the (t,f) field by setting up a new research group in the (t,f) field, a new approach to teach DSP [21], established the series of Information Science Signal Processing and their Applications (ISSPA) international conference (first organized in 1987 in Brisbane), the series of Workshop on Signal Processing and its Applications (WoSPA) (later renamed Workshop on Systems, Signal Processing and their Applications [WoSSPA]) and a continuous Special Session at the SPIE conference in San Diego, CA (1985-1995). He also organized in 1990 in Brisbane the first international conference in (t,f) analysis; subsequently he published the main findings in an edited book that helped further develop the field [1]. He then became the first Associate Editor of the Institute of Electrical and Electronics Engineers (IEEE) transactions on SP with focus on the (t,f) field, and subsequently the technical chairman of the first International Conference on Acoustics, Speech, and Signal Processing (ICASSP) held in Australia in 1994, providing additional focus to the (t,f) field.

Some of the most important developments in the field occurred during this period and in the following years (1995-2000). During this time, a number of other significant groups emerged in

the United States, Montenegro, and other parts of the world. The original French group continued contributing, while a part of the Dutch group transferred to Vienna, Austria.

During that period, a main focus was on defining advanced methods such as quadratic TFDs that reduce the cross-terms or linear methods that eliminate them altogether while preserving the properties of localization, resolution and robustness to noise. Another complementary focus was on defining IF estimation methods that are efficient for nonstationary signals.

After the 1995-2000 period, the focus became more about real applications involving monocomponent and multicomponent signals, refining implementations and simplifying the concepts, methods, and techniques to make them available to a wider audience (the main aim of this book).

In the mid-1990s, the editor/author moved on and took several positions of leadership and management (Center Director, Department Head, Dean, etc.) in several universities in Australia, the United Arab Emirates, and Qatar, but remained active in the field by continuing his research with his students, assistants as well as postdoctoral fellows.

The (t,f) toolbox presented in [Chapter 17](#) reflects the advances made in the field since the late 1970s. It was made available to hundreds of other labs and groups since 1987, both as executable and source; the main code is described in [\[22\]](#). That code allowed a major boost to the field as all (t,f) methods could then be evaluated numerically, a major challenge at the time. Current at the time of publication, the editor/author continues to receive a continual stream of requests for the (t,f) toolbox, indicating that the Time-Frequency Signal Processing and Analysis (TFSAP) toolbox remains an important aid to students, engineers, and researchers.

At the time of publication of this book, from the editor/author's point of view, the main areas of focus include:

1. simplifying the presentation of the (t,f) concepts, terminology, and notations so as to make it easier for new entrants in the field, whether students, engineers, or researchers;
2. providing clear methodologies for the efficient use of (t,f) signal analysis and processing techniques in engineering and science;
3. improving the resolution and robustness of (t,f) methods to noise in applications such as accurate estimation of signal parameters, for example, IF estimation; and
4. widening the domains of applications to include, for example, the important area of multisensory/multichannel SP.

The editor/author believes that (t,f) methods will become integrated with most future DSP software and image processing/vision techniques as new computer technologies allow easy processing of the full TFD as a (t,f) image. This will in turn define new areas for research with renewed efforts on improving the precision, resolution and robustness of (t,f) methods as well as their memory and computational efficiency for real-time implementations. Specific future hot areas include the refined use of (t,f) features in automatic classification for change detection, fault identification, and condition monitoring.

I.3 AN OVERVIEW OF REMAINING CHAPTERS AND KEY (t, f) ISSUES

So far, this chapter has briefly described the essential (t,f) terminology, concepts, and characteristics that must be understood for a correct and effective application of TFSP methods and techniques. For completeness, we present in a concise form, an overview of the remaining material covered in the

book with focus on the objectives, motivation, and rationale for the presented concepts, methods, and techniques. As TFDs reveal the important characteristics of a signal, they are able to produce useful time-frequency *signatures* that uniquely characterize the signal by some key features in the (t, f) domain. These signatures have been found useful in a wide range of applications such as the detection of machine faults [2, Chap. 18] and electrocardiogram (ECG) abnormalities [23], telecommunications, biomedical engineering, seismic surveying, radar, sonar, oceanography, and many others. The key issues arising from such applications are briefly described chapter by chapter below, with further details provided within the corresponding chapters.

Part I: Foundations of TFSP: Spectrogram, WVD, and ambiguity functions

[Chapter 1](#): Defining (t, f) signal models

[Chapter 1](#) gives more formal definitions for the concepts of IF, SD, time duration, instantaneous bandwidth, and group delay. The notion of IF has been fundamental to the development of TFSP, and it remains fundamental to a good understanding and use of TFSP methods. As the modern definition of IF is based on the analytic signal and the Hilbert transform, these are studied in detail in terms of their practical application to real-life signals.

[Chapter 2](#): Formulating TFDs

[Chapter 2](#) introduces heuristic formulations of various TFDs, including those that have played key roles in the development of TFSP—such as the WVD, the spectrogram and filter banks, and the Page distribution, as well as more advanced formulations such as the extended modified B-distribution (EMBD). The spectrogram is still a very popular tool despite its low resolution and high sensitivity to the choice of a window. Its main advantage is its simplicity. In order to improve upon it, one can use it adaptively by iteratively refining the choice of the window. Another approach is to use a relevant method from a class of Quadratic TFDs formed by smoothing WVDs.

[Chapter 3](#): Multicomponent signal analysis using high-resolution TFDs

Many TFDs are nonlinear in the sense that they are FTs of intermediate expressions that are *quadratic* in the signal $z(t)$. These expressions are said to “dechirp” the signal; if the signal has a linear IF law, the intermediate expression has *constant* IF. But for multicomponent signals, the quadratic expression creates spurious “cross-terms” that occur *between* individual components in the (t, f) space. Various “reduced-interference” distributions, such as the EMBD, were developed to remedy such problems while maintaining good (t, f) resolution. These are based on 2D filtering in the ambiguity domain, as described in [Chapter 3](#) and further detailed in, for example, [Sections 5.2](#) and [5.7](#).

Part II: Fundamental methods

[Chapter 4](#): (t, f) Signal and system analysis

[Chapter 4](#) presents a number of additional techniques and insights that are essential for an advanced and thorough understanding of (t, f) methods. This includes a presentation of the relationships between quadratic time-frequency distributions (QTFDs), time-scale methods and the empirical mode decomposition, the localization of cross-terms in the ambiguity and (t, f) domain, uncertainty relationships in (t, f) signal analysis that establish a lower bound on achievable (t, f) resolution. Then formulations of measurements from the (t, f) plane, such as spread, are provided. Finally, a link between the spectrogram and the WVD is discussed.

Chapter 5: Key issues in TFD design

Chapter 5 discusses constraints involved in TFD design as well a number of alternate approaches for obtaining TFDs with desirable properties. It starts with the discussion of localization of cross-terms in the ambiguity domain. Then some classical as well as recent methods for defining high-energy concentration and high-resolution TFDs using separable kernel, multidirectional kernel, and adaptive kernel approaches are discussed. In addition, the chapter also discusses polynomial and higher-order time-frequency methods as well as complex-argument TFDs and the covariance property of TFDs.

Part III: TF methods

Chapter 6: TFSP algorithms and implementation issues

The computational properties of TFDs are further studied, and this leads to an efficient real-time implementation, which exploits the symmetry properties of TFDs [24]. Most current algorithms for TFSP are based on these original algorithms. Finite-duration signals require taking a windowed FT of the quadratic kernel (see [24, 2, Chap. 6]). The choice of the windows results in a similar effect to the case of windowing the FT for stationary signals but extended to the 2D (t, f) domain.

Chapter 7: Performance measure, selection, and implementation of TFDs

A practical application of TFSP with real-life signals requires a calibration of the TFSP tools, so that we can predict the behavior and performance of a TFSP TFD in a given situation. This chapter addresses this issue and provides objective criteria to measure the resolution performance of TFDs in some specific situations.

Chapter 8: Multisensor SP and multichannel data

Chapter 8 presents several applications of TFDs in the multisensory case, including the design of spatial TFDs. It extends conventional (t, f) signal analysis techniques to deal with multichannel/multisensor data. Spatial TFDs are defined to solve two important array SP problems: direction finding and blind-source separation. One advantage of multichannel (t, f) algorithms is that they can extract sources even when there are less sensors than sources.

Part IV: Statistical techniques

Chapter 9: (t, f) Noise analysis and random processes

Just as conventional spectral techniques can be extended to higher-order spectra so that measurements like skewness and kurtosis can be used for improved characterization of processes, so too can time-varying spectra be extended. This results in time-varying higher-order spectra (TVHOS). Particular slices of these TVHOS have some interesting properties. The *polynomial spectra* are slices that ideally concentrate energy in the (t, f) plane (see [2, Chap. 12, 25]). *Cyclostationary* analysis—that is, the spectral analysis of processes whose autocorrelation functions are pseudoperiodic [26]—can be represented in a simple manner within a time-frequency perspective.

Chapter 10: Estimating the IF

As the IF is a key characteristic of nonstationary signals, its estimation is important in many applications, in the same way that the topic of frequency estimation is important in conventional SP. The peaks of TFDs are often used as an estimate of the IF of nonstationary signals, and several algorithms are presented for the difficult case of multicomponent signals.

[Chapter 11: \$\(t, f\)\$ Synthesis and filtering](#)

TFDs are used in the design of signals that have some predefined properties. The ambiguity function has long been used for the design of radar signals; its 2D FT (composed by one FT and one IFT) results in a (t, f) representation that can therefore be naturally used in the design of signals exhibiting particular (t, f) features [2, Chap. 17], with the advantage that such TFDs are real, while the ambiguity function is complex.

The input-output convolution relationships of filters are preserved when one uses basic TFDs to represent the signals. Therefore, they can be used to describe a time-varying filter [2, Chap. 17]. A simple filter can then be applied to mask regions of the (t, f) plane of the input signal and then perform an inversion of the TFD to recover the filtered signal (see [1] and Chap. 17 of [2]).

[Chapter 12: Detection and estimation using TFDs](#)

The optimal solutions to many detection problems are expressed as inner products formed by time-domain correlations, as in the matched filter [27]. [Section 12.5](#) shows that these time-domain correlations are equivalent to (t, f) correlations [28].

We can therefore formulate many important detection problems using TFDs [28,29]. The advantage of this is that one can highlight certain regions of the (t, f) plane and perform detection based on selective time-frequency features. Therefore, many of the classical detection and estimation problem solutions have alternate implementations based on TFDs. The 2D time-frequency nature of the implementation allows greater flexibility than the classical one [2, Chap. 12]. Such (t, f) -based optimal detection schemes are described in [Chapter 12](#).

Part V: Engineering applications

[Chapter 13: TFSP in telecommunications](#)

This application, along with seismic prospecting and radar, are the areas in which SP initially developed and grew to become a separate field. With biomedicine, they are still major application areas of DSP, and it is therefore expected that TFSP would have important applications in all these areas and beyond.

[Chapter 14: \$\(t, f\)\$ Methods in radar, sonar, and acoustics](#)

[Chapter 14](#) presents several applications in the areas of radar, sonar, speech, and acoustics. The study of five concrete cases is realized. Each one illustrates the usefulness of (t, f) methods for resolving problems within this field.

[Chapter 15: \$\(t, f\)\$ Diagnosis and monitoring](#)

[Chapter 15](#) presents several applications in the important area of diagnosis and monitoring. In these applications, the given signal is transformed in the (t, f) domain using TFDs. These are then processed to extract key signal features, such as the IF. The extracted features are then used by a classifier to detect abnormalities or changes for use in diagnosis. Several applications illustrate the finding that (t, f) approaches often outperform conventional signal classification approaches.

[Chapter 16: Biomedical TFSP](#)

TFSP has been applied to a wide range of biomedical problems. [Chapter 16](#) describes the methodologies used in some selected applications in an important subfield—the neurosciences—for analysis and

processing of EEGs, ECGs, heart rate variabilities (HRVs), and other associated signals recorded during the experiments.

Studies of these signals have revealed that newborn EEGs are highly nonstationary, making TFSP a suitable method for developing seizure detection techniques. In addition, TFD plots can be used to represent the ECG nonstationarities, and are helpful for isolating particular heart disorders [23]. In addition, information obtained from TFDs can complement EEG analysis for seizure detection, given the link between heart and brain functions.

[Chapter 17](#): Getting started with the TFSAP software package

A version of the TFSAP software package accompanies this book. The manual provided in [Chapter 17](#) includes several tutorials. We recommend that Tutorial 1 be completed in conjunction with the reading of this first chapter.

I.4 CONCLUSIONS AND ADVICE TO THE READER

This introductory chapter was intended to take the reader to a stage where the essence of the (t,f) approach is revealed clearly. Next, [Chapter 1](#) consolidates the knowledge gained so far by revisiting in more depth the fundamentals of DSP from a (t,f) perspective. [Chapter 2](#) then reviews the most basic formulations of TFDS from several perspectives. [Chapter 3](#) is the most important chapter of this book, as it introduces a modern approach to the design of high-resolution TFDs and it forms the foundation for all subsequent chapters. After studying these chapters, the reader will have all the necessary background to study any other chapter that is relevant to their specific interests. To further assist the reader, this book includes a table of standard (t,f) notations, commencing on page [xlv](#), as well as a detailed table of contents and a two-level index.

The newcomer to the field is advised to read this chapter and [Chapters 1 to 5](#), and then do the tutorial given in [Chapter 17](#). The reader with more advanced knowledge can skip this chapter. Note that [Chapters 3, 2, and 1](#) (in this order) are considered the most important in this book, as they establish the foundations of the field.

I.5 ADDITIONAL EXERCISES

This section provides additional exercises that cover important topics from this chapter and prepares the reader for the next chapters in the book. It is recommended to use the (t,f) toolbox described in [Chapter 17](#) as an aid for doing these exercises. In a formal class setting, the instructor may wish to guide the student by adding intermediate questions with focus on one or two assessment criteria.

Exercise I.5.1. Synthesize a type 1 signal, that is, a signal with constant frequency, and analyze it using time-domain, frequency domain and TFRs using the *TFSAP* package described in [Chapter 17](#). The analysis is intended to result in the definition of features that characterize the signal and that can be used for identification or classification.

1. Comment on the list of features that were observed from the obtained results
2. Explain the advantages of using TFRs

Exercise I.5.2. Synthesize a type 2 signal and analyze it using time-domain, frequency domain, and TFRs.

1. Comment on the list of features that were observed from the obtained results
2. Explain the advantages of using TFRs

Exercise I.5.3. Synthesize a type 3 signal and analyze it using time-domain, frequency domain, and TFRs.

1. Comment on the list of features that were observed from the obtained results
2. Explain the advantages of using TFRs

Exercise I.5.4. We wish to analyze a bat signal using a range of representations including a time domain, frequency domain, and time-frequency domain representations (a demo bat signal is available in the *TFSAP* package described in [Chapter 17](#) and available in the supplementary material). Using TFRs, estimate the following:

1. Total number of components
2. The starting and stopping times of each component
3. Maximum amplitude of each component
4. Maximum instantaneous bandwidth of the signal
5. Maximum instantaneous bandwidth of each component

Exercise I.5.5. A bird signal is available in the supplementary material, for example, in the (t, f) toolbox described in [Chapter 17](#). Do the same for the bird signal as in [Exercise I.5.4](#).

REFERENCES

- [1] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [2] B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992.
- [3] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [4] B. Boashash, N.A. Khan, T. Ben-Jabeur, “Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review”, *Digital Signal Process.* 40 (2015) 1-30, doi:10.1016/j.dsp.2014.12.015.
- [5] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 519-568.
- [6] S. Kay, *Intuitive Probability and Random Processes Using MATLAB®*, Springer Science & Business Media, Berlin, 2006.
- [7] A. Papoulis, S.U. Pillai, *Probability, Random Variables, and Stochastic Processes*, fourth ed., Tata McGraw-Hill Education, New York, 2002.
- [8] L.B. White, B. Boashash, “Cross spectral analysis of nonstationary processes”, *IEEE Trans. Inf. Theory* 36 (4) (1990) 830-835.
- [9] D. Gabor, “Theory of communication”, *J. IEE* 93(III) (26) (1946) 429-457.

- [10] B. Boashash, G. Azemi, N.A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: applications to newborn EEG abnormality detection", *Pattern Recognit.* 48 (3) (2015) 616-627.
- [11] E. Sejdić, I. Djurović, J. Jiang, "Time-frequency feature representation using energy concentration: an overview of recent advances", *Digital Signal Process.* 19 (1) (2009) 153-183.
- [12] B. Bouachache, "Representation temps-frequence", Thesis for diplome de docteur ingenieur, Institut National Polytechnique de Grenoble, France, 1982, 164 pp (in French).
- [13] P. Guillemain, R. Kronland-Martinet, "Characterization of acoustic signals through continuous linear time-frequency representations", *Proc. IEEE* 84 (4) (1996) 561-585.
- [14] D. Iatsenko, P.V.E. McClintock, A. Stefanovska, "Linear and synchrosqueezed time-frequency representations revisited: overview, standards of use, resolution, reconstruction, concentration, and algorithms", *Digital Signal Process.* (2015) doi:10.1016/j.dsp.2015.03.004, URL <http://www.sciencedirect.com/science/article/pii/S1051200415000792>.
- [15] B. Bouachache, B. Escudie, J.M. Komatitsch, "Sur la possibilité d'utiliser la représentation conjointe en temps et fréquence dans l'analyse des signaux modulés en fréquence émis en vibroismiques", in: Seventh GRETSI Symp. on Signal Processing and Its Applications, 1979, pp. 121/1-121/6.
- [16] B. Bouachache, B. Escudé, P. Flandrin, J. Gréa, "Sur une condition nécessaire et suffisante de positivité de la représentation conjointe en temps et fréquence des signaux d'énergie finie", in: Comptes rendus de l'Académie des Sciences, Paris, vol. 288, Ser. A, 1979, pp. 307-309, in French.
- [17] E.P. Wigner, "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.* 40 (1932) 749-759.
- [18] J. Ville, "Théorie et applications de la notion de signal analytique", *Cables et Transmissions* 2A (1) (1948) 61-74, in French. English translation: I. Selin, Theory and applications of the notion of complex signal, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).
- [19] A.W. Rihaczek, "Signal energy distribution in time and frequency", *IEEE Trans. Inf. Theory* 14 (3) (1968) 369-374.
- [20] J.L. Flanagan, Speech Analysis; Synthesis and Perception, second ed., Springer-Verlag, Berlin, New York, 1972, ISBN 0387055614 3540055614, x, 444 pp.
- [21] B. Boashash, S. Sridharan, V. Chandran, "The development of a new signal processing program at the Queensland University of Technology", *IEEE Trans. Educ.* 39 (2) (1996) 186-191, ISSN 0018-9359, doi: 10.1109/13.502065.
- [22] B. Boashash, A.P. Reilly, "Algorithms for time-frequency signal analysis", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, chap. 7, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 163-181 (Chapter 7).
- [23] S.S. Abeysekera, B. Boashash, "Methods of signal classification using the images produced by the Wigner-Ville distribution", *Pattern Recogn. Lett.* 12 (11) (1991) 717-729.
- [24] B. Boashash, P.J. Black, "An efficient real-time implementation of the Wigner-Ville distribution", *IEEE Trans. Acoust. Speech Signal Process.* 35 (1987) 1611-1618.
- [25] B. Boashash, P.J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.* 42 (1) (1994) 216-220.
- [26] W.A. Gardner, Statistical Spectral Analysis: A Nonprobabilistic Theory, Prentice-Hall, Upper Saddle River, NJ, 1988.
- [27] H.L.L. Van Trees, Detection, Estimation, and Modulation Theory, vol. I, Wiley, New York, 1968, reprinted 2001.
- [28] B. Boashash, G. Azemi, "A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG", *Digital Signal Process.* 28 (2014) 28-38.
- [29] B. Boashash, P.J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques", *IEEE Trans. Acoust. Speech, Signal Process.* 38 (11) (1990) 1829-1841.

This page intentionally left blank

TIME-FREQUENCY AND INSTANTANEOUS FREQUENCY CONCEPTS⁰

1

INTRODUCTION AND OVERVIEW

The previous chapter introduced the field of time-frequency (t,f) signal analysis and processing (TFSAP), which concerns the analysis and processing of signals with time-varying frequency content. Such signals are best represented by a *time-frequency distribution (TFD)*, which is intended to show how the energy of the signal is distributed over the two-dimensional (t,f) space. Processing of the signal may then exploit the features produced by the concentration of signal energy in two dimensions (time *and* frequency) instead of only one (time *or* frequency).

This chapter begins with the more formal, in-depth part of the tutorial which occupies the rest of **Part I (Chapters 1–3)**. The present tutorial updates the one given in Ref. [1] and the first edition of this book by including more recent advances, refining terminology, and simplifying both the presentations of concepts and formulations of methodologies. This core material needs to be understood by anybody seeking to specialize in this field. Precise definitions and formulations of the key concepts needed to formulate (t,f) methods are provided.

Chapter 1 includes four sections. Section 1.1 describes in detail the justifications why time-frequency methods are preferred for a wide range of applications in which the signals have time-varying spectral characteristics or multiple components for which the variables t and f are related. Section 1.2 provides the signal models and mathematical formulations needed to describe temporal and spectral characteristics of nonstationary signals in the time-frequency domain. It defines such basic concepts as analytic signals, Hilbert transform (HT), bandwidth-duration product, and asymptotic signals. Section 1.3 defines the key quantities related to time-frequency methods, including the instantaneous frequency (IF), spectral delay (SD), and group delay (GD). Section 1.4 reinforces the material in the previous sections with a simple tutorial on defining the AM/FM characteristics of signals using the concepts of analytic signal and IF; and relating to the definition of the amplitude and phase of a nonstationary signal.

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa).

1.1 THE BENEFITS OF TIME-FREQUENCY DISTRIBUTIONS (TFDs)

As introduced in [Chapter I](#), the two classical representations of a signal are the time-domain representation $s(t)$ and the frequency-domain representation $S(f)$. In both cases, the variables t and f are treated as mutually exclusive: to obtain a representation in terms of one variable, the other variable is “integrated out.” Consequently, each classical representation of the signal is *nonlocalized* w.r.t. the excluded variable; that is, the frequency representation is essentially averaged over the values of the time representation at *all* times, and the time representation is essentially averaged over the values of the frequency representation at *all* frequencies.

As the conventional representations in the time or frequency domains are inadequate in situations such as those described above, an obvious solution is to seek a representation of the signal as a *two-variable* function or distribution whose domain is the two-dimensional (t,f) space. Such a representation is called a *time-frequency representation (TFR)* and is obtained using a *TFD*.

In the TFD, denoted by $\rho(t,f)$, the variables t and f are *not* mutually exclusive, but present together. The TFD representation is *localized* in both t and f subject to the interpretation and within the limits of the uncertainty principle, as mentioned in [Section I.2.3.3](#) of [Chapter I](#).

Before proceeding further with the (t,f) approach, for the sake of completeness, the next sections first use practical examples to present a systematic step-by-step approach linking the time domain and the frequency domain with the (t,f) domain.

1.1.1 REPRESENTATION OF THREE REAL-LIFE SIGNALS

The usefulness of representing a signal as a function of both time *and* frequency is illustrated by considering three signals of practical importance:

1. *Sinusoidal FM signal*: The sound in monophonic analog TV, and in monophonic FM radio, is transmitted on a frequency-modulated carrier. If the audio signal is a pure tone of frequency f_m (the modulating frequency), then the frequency of the carrier is of the form

$$f_i(t) = f_c + f_d \cos[2\pi f_m t + \phi], \quad (1.1.1)$$

where t is time, $f_i(t)$ is the frequency modulation law (FM law), f_c is the mean (or “center”) carrier frequency, f_d is the peak frequency deviation, and ϕ allows for the phase of the modulating signal. The amplitude of the carrier is considered constant.

2. *Linear-FM signal*: Consider a “sinusoidal” signal of total duration T , with constant amplitude, whose frequency increases from f_0 to $f_0 + B$ at a constant rate $\alpha = B/T$. If the signal begins at $t = t_0$, the FM law may be written

$$f_i(t) = f_0 + \alpha(t - t_0); \quad t_0 \leq t \leq t_0 + T. \quad (1.1.2)$$

In an electronics laboratory, such a signal is called a *linear frequency sweep*, and might be used in an experiment to measure the frequency response of an amplifier or filter because of the wide-band characteristics of this signal. In mineral exploration, a linear-FM signal is called a *chirp* or *vibroseis* signal, and is used as an acoustic “ping” for detecting underground formations [2].

- 3. Musical performance:** A musical note consists of a number of “components” of different frequencies, of which the lowest is called the *fundamental* and the remainder are called *overtones* [3, p. 56]; overtones at whole-number multiples of the fundamental frequency are also called *harmonics*. These components are present during a specific time interval and may vary in amplitude during that interval. In modern musical notation, each note is represented by a “head.” The position of the head in the vertical dimension (together with other information such as the clef and key signature) indicates the *pitch*, that is, the frequency of the most prominent component (usually the fundamental). The horizontal position of the head in relation to other symbols indicates the starting time, and the duration is specified by the shading of the head, attached bars, dots, stem and flags, and tempo markings such as *Allegro*. The choice of instrument—each one being characterized by its overtones and relationships with the fundamental—is indicated by using a separate stave (group of five horizontal lines) for each instrument or group of instruments, or a pair of staves for a keyboard instrument. Variations in amplitude are indicated by dynamic markings such as *f* or *p*, and *crescendo*. Figure 1.1.1 illustrates the system. By scanning a set of staves vertically, one can see which fundamentals are present on which instruments at any given time. By scanning the staves horizontally, one can see the times at which a given fundamental is present on a given instrument. This is clearly a (t, f) representation.

The above three signals are *nonstationary*; that is, each of them has a time-varying frequency or time-varying “frequency content.” Such signals are comprehensible partly because our sense of hearing readily interprets sounds in terms of variations of frequency or “frequency content” with time. However, conventional representations of a signal in the time domain or frequency domain do not facilitate such interpretation, as shown below.

The following two sections revisit in more depth and more details the concepts first introduced in Sections I.1.2 and I.1.3.

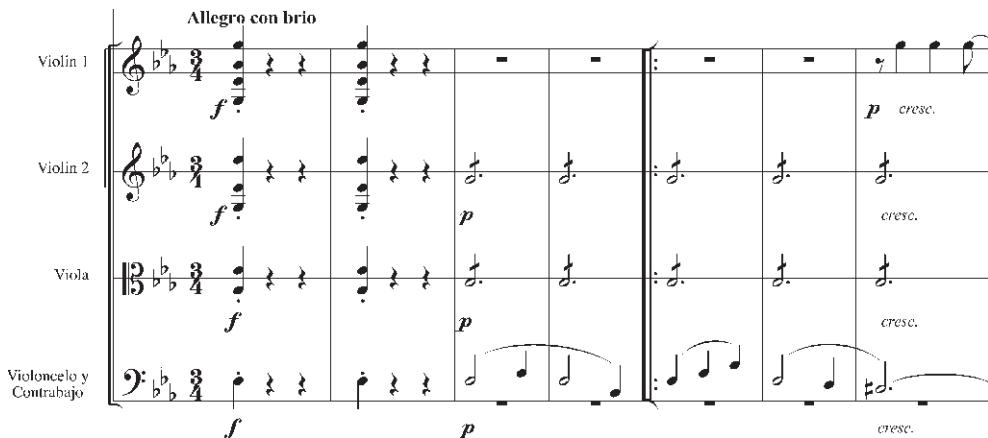


FIGURE 1.1.1

Musical notation, showing the string parts at the beginning of Beethoven's *Third Symphony* [4]. Roughly speaking, the horizontal dimension is time and the vertical dimension is frequency.

1.1.2 TIME-DOMAIN REPRESENTATION

A signal can be described naturally as a function of time, which we may write $s(t)$. This representation leads immediately to the *instantaneous power*, given by $|s(t)|^2$, which shows how the energy of the signal is distributed over time; the total signal energy is expressed as: $E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt$. But the time-domain description has limitations, as may be seen by applying it to the above three examples:

1. The *sinusoidal FM signal* whose frequency satisfies Eq. (1.1.1) may be written as:

$$s_1(t) = A \cos \left(2\pi f_c t + \frac{f_d}{f_m} \sin[2\pi f_m t + \phi] + \psi \right), \quad (1.1.3)$$

where A is the amplitude and ψ a phase offset; the fraction f_d/f_m is called the *modulation index* and is equal to the peak phase deviation (in radians) from $2\pi f_c t$. This equation, by itself, does not clearly indicate how the frequency varies with time. Figure 1.1.2 is a graph of $s_1(t)$ against time for $A = 1$, $f_c = 0.25$ Hz, $f_d = 0.1$ Hz, $f_m = (3/64)$ Hz, and $\phi = \psi = 0$. The graph gives an impression of an oscillating frequency, but extracting that frequency as a function of time from the graph would be a tedious and imprecise exercise.

2. The *linear-FM signal* whose frequency satisfies Eq. (1.1.2) may be written

$$s_2(t) = A \text{rect} \left[\frac{t - t_0 - T/2}{T} \right] \cos \left(2\pi \left[f_0(t - t_0) + \frac{\alpha}{2}(t - t_0)^2 \right] + \psi \right), \quad (1.1.4)$$

where, again, A is the amplitude and ψ a phase offset. The rect function is a rectangular pulse of unit height and unit duration, centered on the origin of time, as defined in Table I.1.1; that is, the

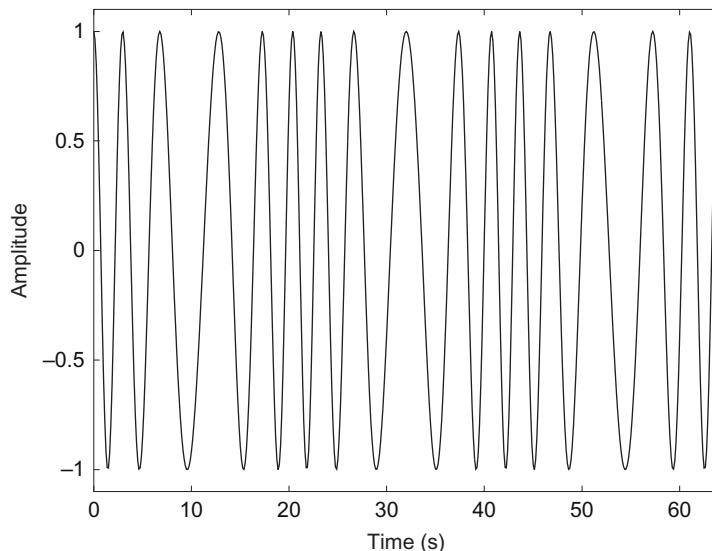


FIGURE 1.1.2

Time-domain representation of a sinusoidal FM signal (Eq. 1.1.3).

rect [...] factor in Eq. (1.1.4) is equal to unity for $t_0 \leq t \leq t_0 + T$, and 0 elsewhere. But again it is not immediately apparent why Eq. (1.1.4) has the required FM law.

The graph of $s_2(t)$ versus t is shown on the left side of Fig. 1.1.4(a), for $A = 1$, $t_0 = 18$ s, $T = 64$ s, $f_0 = 0.1$ Hz, $\alpha = (3/640)$ Hz s $^{-1}$, and $\psi = -\pi/2$. Although the graph gives a clear impression of a steadily increasing frequency, the exact FM law is not directly represented, and would be even further obscured if the signal were corrupted by noise (as in Fig. I.1.6 in the preceding chapter).

3. A *musical performance* can be represented in the time domain by variations in time of the air pressure at a particular point in space. The time-varying pressure may be converted by a microphone and amplifier into an electrical signal, say $s_3(t)$. Indeed, music is routinely recorded and broadcast in this way. However, this signal $s_3(t)$ is nothing like the form in which a composer would write music, or the form in which most musicians read music for the purpose of performance. Neither is it much help, by itself, to an engineer who wants to remove noise and distortion from an old “vintage” recording. The problem is that musical waveforms are far more complex than the underlying artistic ideas. As an example, Fig. 1.1.3 shows the waveform of the opening chord of a performance of Beethoven’s *Third Symphony*—the single *staccato* chord of which the string parts are shown at the left of Fig. 1.1.1.

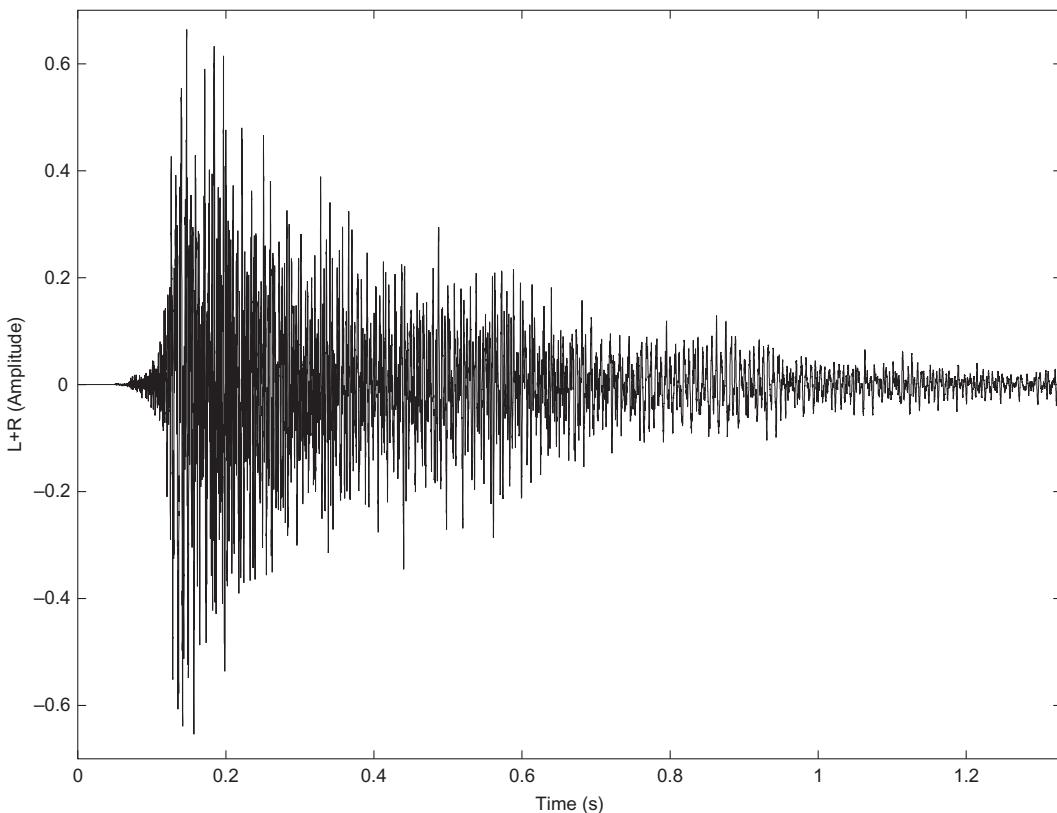
These three examples show that the time-domain representation tends to obscure information about frequency because it assumes that the two variables t and f are mutually exclusive.

1.1.3 SPECTRAL CHARACTERISTICS OF TYPICAL SIGNALS

Let us consider in more detail three cases of a signal $s(t)$ and its FT $S(f)$, given in Eq. (I.1.2) by $S(f) = \mathcal{F}_{t \rightarrow f} \{s(t)\} \triangleq \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt$. For convenience, the relation between $s(t)$ and $S(f)$ may be written “ $s(t) \xleftrightarrow{t \rightarrow f} S(f)$ ” or simply “ $s(t) \leftrightarrow S(f)$.” The FT is in general complex; its magnitude is called the *magnitude spectrum* and its phase is called the *phase spectrum*. The square of the magnitude spectrum is the *energy spectrum*, which describes how the energy of the signal is distributed over the frequency domain; the total energy of the signal is defined in Eq. (I.1.5) as $E_s = \int_{-\infty}^{\infty} |S(f)|^2 df = \int_{-\infty}^{\infty} S(f) S^*(f) df$, where the superscripted asterisk (*) denotes the complex conjugate. Although the representation $S(f)$ is a function of frequency only—time having been “integrated out”—the FT is a complete representation of the signal because the signal can be recovered from the FT by taking the inverse FT (IFT): $s(t) = \mathcal{F}_{t \leftarrow f}^{-1} \{S(f)\} = \int_{-\infty}^{\infty} S(f) e^{j2\pi ft} df$ defined in Eq. (I.1.1). But the “completeness” of the FT representation does not make it convenient for all purposes, as may be seen by considering the same three examples:

1. *Sinusoidal FM signal*: If $\phi = \psi = 0$ in Eq. (1.1.3), the expression for $s_1(t)$ can be expanded into an infinite series (using Jacobi-Anger identity) as

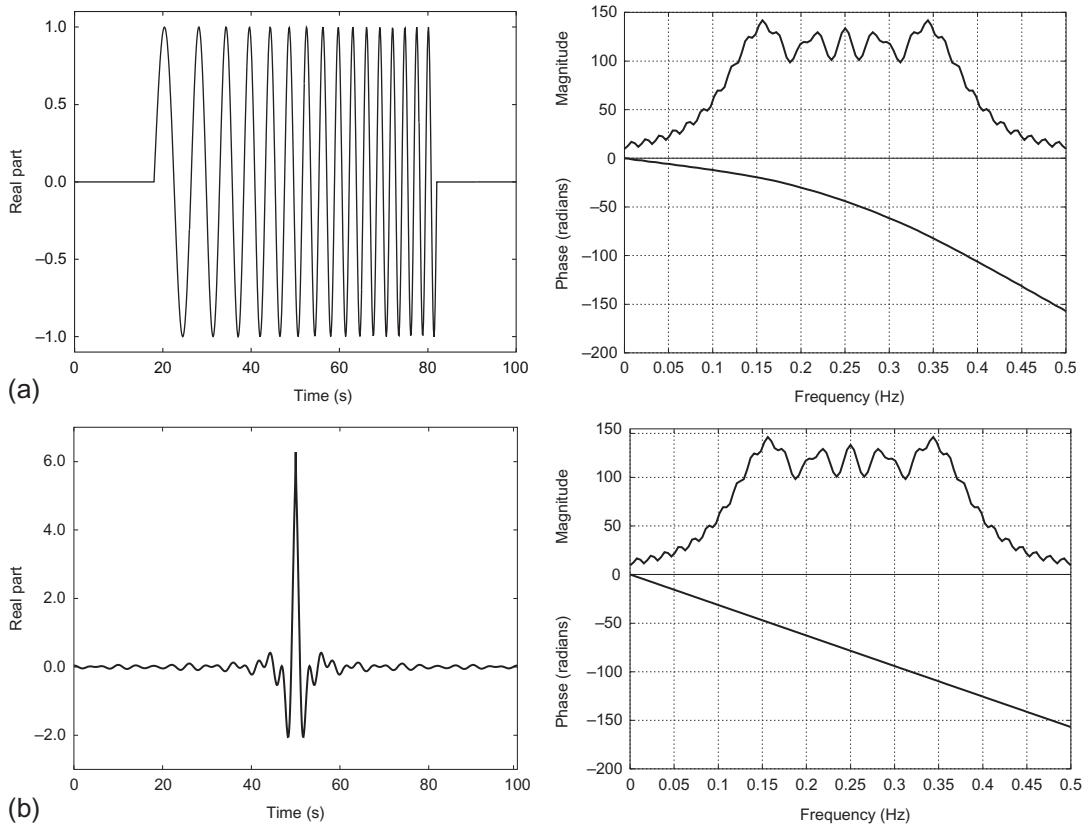
$$s_1(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos 2\pi(f_c + nf_m)t, \quad (1.1.5)$$

**FIGURE 1.1.3**

Time-domain representation of a single short E_b -major chord played by a full orchestra [5]. The left and right stereo channels are averaged.

where β is the modulation index ($\beta = f_d/f_m$) and J_n denotes the Bessel function of the first kind, of order n [6, p. 226]. In the frequency domain, this becomes an infinite series of delta functions; one of these (the *carrier*) is at the mean frequency f_c , and the remainder (the *sideband* tones) are separated from the carrier by multiples of the modulating frequency f_m , as shown in Fig. 1.1.5(a). Although the number of sideband tones is theoretically infinite, the *significant* ones¹ may be assumed to lie between the frequencies $f_c \pm (f_d + f_m)$ or, more conservatively, $f_c \pm (f_d + 2f_m)$. This information is essential if one is designing a tuning filter to isolate the TV audio carrier or separate

¹For more detail, see Carlson [6, pp. 220-237]. The above description considers only positive frequencies; similar comments apply to the negative frequencies. Theoretically, the lower sideband tones belonging to the positive-frequency carrier extend into the negative frequencies, while the corresponding sideband tones belonging to the negative-frequency carrier extend into the positive frequencies; but such “aliased” components are negligible if f_c and f_d are appropriately chosen.

**FIGURE 1.1.4**

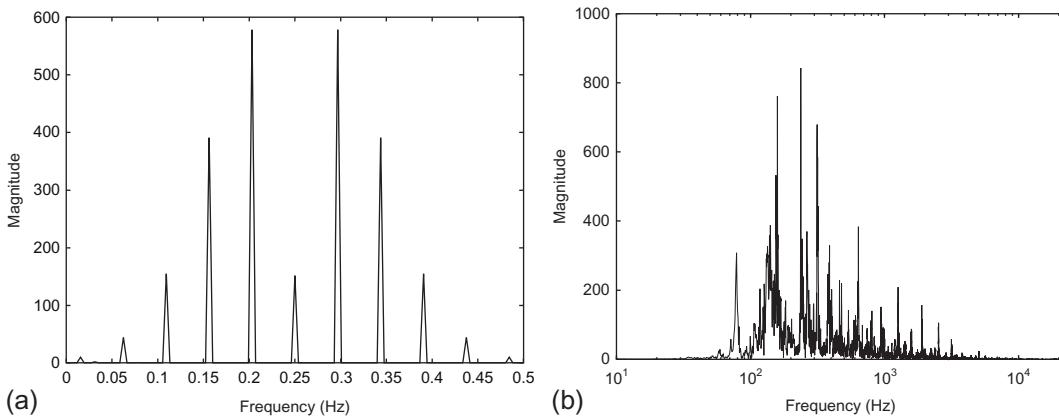
The importance of phase: (a) time-domain representation (left) and magnitude and phase spectra (right) of a linear-FM signal (Eq. 1.1.6) with duration 64 s, starting frequency 0.1 Hz, and finishing frequency 0.4 Hz; (b) corresponding representations of another signal with the same magnitude spectrum as that in part (a). The sampling rate is 8 Hz (see also Fig. I.1.6).

one FM channel from adjacent channels. However, it is inadequate if one is designing a modulator or demodulator because its connection with the FM law is even more obscure than that of Eq. (1.1.3).

2. The *linear-FM signal* in Eq. (1.1.4) is the real part of the complex signal

$$z_2(t) = A \operatorname{rect} \left[\frac{t - t_0 - T/2}{T} \right] e^{j2\pi[f_0(t-t_0) + \frac{\alpha}{2}(t-t_0)^2] + \psi}. \quad (1.1.6)$$

The magnitude spectrum of $z_2(t)$ is shown by the upper trace in the right-hand graph of Fig. 1.1.4(a), for $A = 1$, $t_0 = 18$ s, $T = 64$ s, $f_0 = 0.1$ Hz, $\alpha = (3/640)$ Hz s^{-1} , and $\psi = -\pi/2$. It shows that the magnitude is significant in the band corresponding to the frequency-sweep range (0.1 Hz $< f < 0.4$ Hz) and that the energy is mostly confined to that band. But the magnitude

**FIGURE 1.1.5**

Magnitude spectra of (a) the sinusoidal FM signal in Fig. 1.1.2, and (b) the E_5 -major chord in Fig. 1.1.3. The tall narrow spikes in (a) are computed approximations to delta functions.

spectrum fails to show that the frequency is increasing with time; it tells us what frequencies are present in the signal, but not the “times of arrival” of those frequencies. The latter information, which we call spectral delay (SD), is encoded in the phase spectrum, shown by the *lower* trace in the same graph.

The shortcomings of the magnitude spectrum are further emphasized in Fig. 1.1.4(b), which shows a signal whose magnitude spectrum (right, upper trace) is identical to that of the linear-FM signal in part (a), but whose appearance in the time domain (left) is very different. The explanation is to be found in the phase spectrum (right, lower trace): the phase is nonlinear in part (a) but linear in part (b).²

3. Similarly, the *musical performance* has a magnitude spectrum, which tells us what frequencies are present, but not *when* they are present; the latter information is again encoded in the phase.

Figure 1.1.5(b) shows the magnitude spectrum of the single chord in Fig. 1.1.3. The magnitude spectrum of a longer and more complex musical signal could have up to 120 peaks, corresponding to the notes of the chromatic scale in the audible frequency range. The relative heights of those peaks tell us something about the tonality of the music (or whether it is tonal at all), but the timing of the notes is represented in the magnitude spectrum and will not be obvious from the phase spectrum.

These examples show that the frequency-domain representation “hides” information about timing, as $S(f)$ does not mention the variable t .

We shall see in Section 1.3.2, p. 51 that the SD, which is the localization in time of the frequency components, is proportional to the derivative of the phase spectrum w.r.t. frequency [1,7]. This is similar

²The signal in part (b) of Fig. 1.1.4 was obtained from that in part (a) by padding the signal with zeros in the time domain, taking the fast Fourier transform (FFT), setting the phase to zero, taking the inverse FFT, and delaying the result by 50 s. The magnitude and phase spectra were then recomputed. Due to the padding and the generous over-sampling rate (8 Hz) (i.e., much greater than Shannon limit), the FFT and its inverse closely approximate the FT and its inverse.

to the relationship between group delay (GD) and phase in filter design (cf. [Section 1.3.3, p. 54](#)). Thus, the linear phase spectrum in [Fig. 1.1.4\(b\)](#) means constant delay (all frequencies arrive at once), whereas the increasingly negative slope of the phase spectrum in [Fig. 1.1.4\(a\)](#) means that the delay increases with frequency. We shall also see in [Section 1.3.1, p. 49](#) that the IF, which is the localization in frequency of the time components, is proportional to the derivative of the instantaneous phase of the complex signal w.r.t. time. The (t,f) approach postulates that it would be more convenient if we could infer the SD and IF from a *single* representation of a signal, without the need to interpret additional representations such as the phase spectrum. This is part of the motivation for devising a “sophisticated” and “practical” tool for analysis of nonstationary signals, preserving “all” the information needed for characterization and identification as this can be critical in applications requiring condition monitoring and fault detection such as medical engineering [8].

1.1.4 JOINT TIME-FREQUENCY REPRESENTATION

In a TFR, a constant- t cross-section should show the frequency or frequencies present at time t , and a constant- f cross-section should show the time or times at which the frequency f is present. Key questions are: how to define a TFD to obtain a desirable TFR? Are there any required constraints? [Figure 1.1.6](#) illustrates what is desired. The sinusoidal FM law is clear in part (a), as is the linear-FM law in part (b). Part (c) shows the appearance and decay of the components of the E_\flat -major chord. The signals in parts (b) and (d) cannot be distinguished by their magnitude spectra (cf. [Fig. 1.1.4](#)), but their (t,f) representations show that in the former case the higher frequencies appear after the lower frequencies, whereas in the latter case all frequencies appear at once.

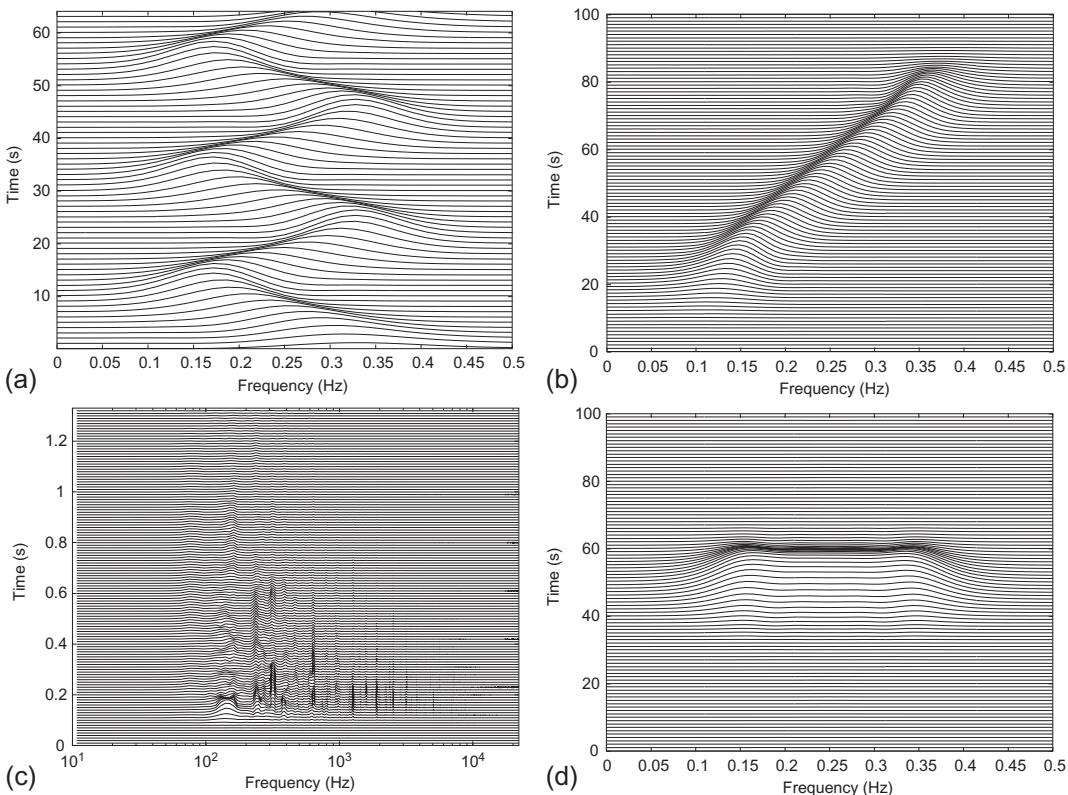
[Figure 1.1.7](#) shows a different TFD of a linear-FM signal (the same signal as in [Fig. 1.1.4](#), without the delayed onset). This particular TFD, known as the Wigner-Ville distribution (WVD), is noted for its sharp resolution of linear-FM laws; note the clear indication of the frequency range and the variation in frequency with time. Such a variation may be described by the IF $f_i(t)$. A signal may have more than one component, each with its own IF; for example, [Fig. 1.1.8](#) shows a TFD of a sum of two linear-FM signals, each of which has its own IF.³ These IF features are not apparent in conventional signal representations.

Nonstationary signals for which a TFD representation may be useful occur not only in broadcasting, seismic exploration, and audio, from which our three examples are taken, but also in numerous other engineering and interdisciplinary fields such as telecommunications, radar, sonar, vibration analysis, speech processing, and medical diagnosis. *Time-frequency signal processing (TFSP)* is the processing of such signals in the (t,f) domain by means of TFDs.

1.1.5 DESIRABLE CHARACTERISTICS OF A TFD

The use of a TFD for a particular purpose is inevitably based on particular assumptions concerning the properties of TFDs. To ensure that the conclusions of the analysis are sound, these assumptions must be identified and verified. In general, we may say that the desirable properties of TFDs are those on

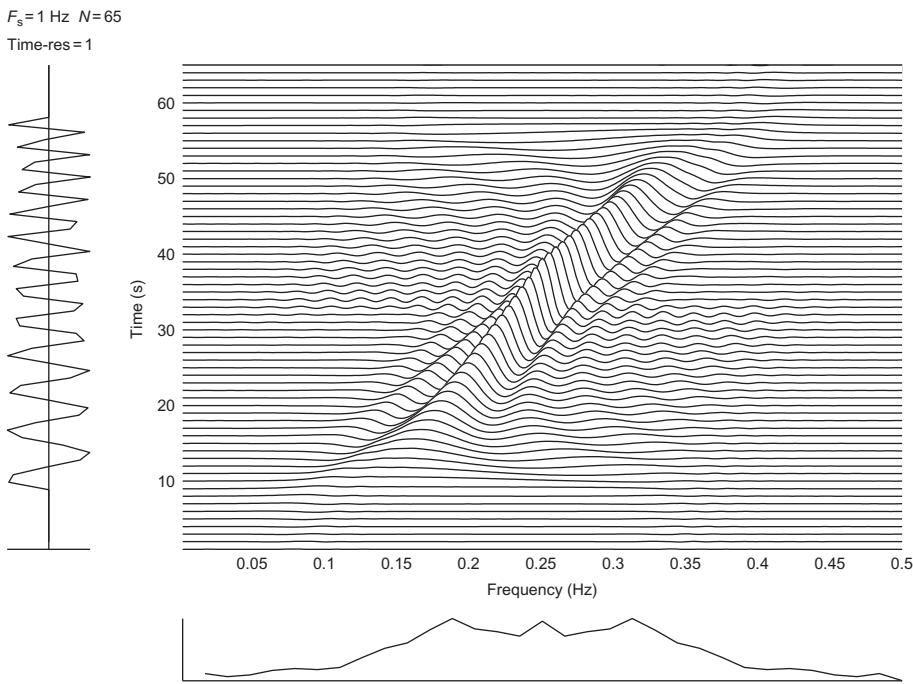
³N.B.: In subsequent graphs of TFDs in [Part I](#), the labels on axes are similar to those in [Fig. 1.1.7](#), but may be smaller due to space constraints.

**FIGURE 1.1.6**

Time-frequency representations of (a) the sinusoidal-FM signal in Figs. 1.1.2 and 1.1.5; (b) the linear-FM signal in Fig. 1.1.4; (c) the E_b -major chord in Figs. 1.1.3 and 1.1.5; and (d) the linear-phase signal in Fig. 1.1.4(b). (The TFRs used here are called spectrograms, except in part (c), where we use the square root of the spectrogram in order to compress the dynamic range. These TFRs will be defined in the next chapter; to reproduce these plots, see toolbox in Chapter 17.)

which their most likely applications depend. The following generic applications illustrate some typical uses of TFDs, applications, and corresponding required properties [9–11]:

- Analyze the raw signal in the (t,f) domain to identify characteristics such as number of components, relative amplitudes, IF laws, (t,f) complexity, (t,f) flatness, (t,f) energy concentration areas, etc.
- Separate the components or (t,f) regions from each other and from the background noise by filtering in the (t,f) domain (see Sections 10.6 and 11.4).
- Synthesize the filtered TFD to reconstruct $s(t)$ in the time domain.
- Estimate and analyze specific components separately, and in particular:
 - track the instantaneous amplitude;

**FIGURE 1.1.7**

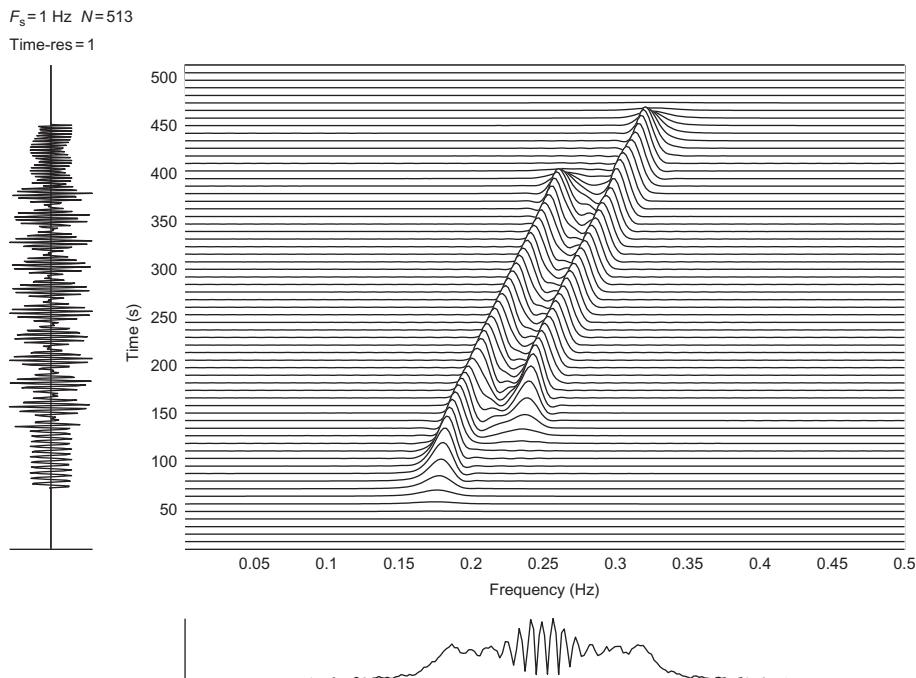
A time-frequency representation of a linear-FM signal (Eq. 1.1.4) with duration 65 samples, starting frequency 0.1 Hz, and finishing frequency 0.4 Hz (sampling rate 1 Hz). The time-domain representation appears on the left, and the magnitude spectrum at the bottom; the same pattern is followed in most TFD graphs in Part I of this book. (The TFD used in this case is the WVD option of the (t, f) toolbox described in Chapter 17.)

- track the IF; and
- track the instantaneous bandwidth (spread of energy about the IF).
- Choose a mathematical model of the signal, showing clearly the significant characteristics, such as the IF $f_i(t)$ or SD $\tau_d(f)$.

Specific applications include, for example, watermarking, that is, hiding a watermark in a particular predefined secret (t, f) location of a signal such as an audio recording [12].

These applications can be carried out using a TFD with the following properties:

1. The TFD is preferably an energy density so as to allow an energetic physical interpretation. That is,
 - (a) the TFD is real;
 - (b) the integral of the TFD over the entire (t, f) plane is the total energy of the signal;
 - (c) the integral over a rectangular region of the (t, f) plane, corresponding to a finite bandwidth and finite time interval, is approximately the energy of the signal in that bandwidth over that interval, provided that the bandwidth and interval are sufficiently large (see Section 1.2.5).

**FIGURE 1.1.8**

A time-frequency representation of two linear-FM signals with close parallel FM laws, duration 512 samples, frequency separation 0.04 Hz (sampling rate 1 Hz).

2. For a monocomponent FM signal, the peaks of the constant-time cross-sections of the TFD should give the IF law which describes the signal FM law.
3. For a multicomponent FM signal, the dominant peaks of the TFD should reflect the components' respective FM laws; and the TFD should resolve any close components, as in Fig. 1.1.8.

A simple measure of the resolution of a TFD is the concentration of energy about the IF law(s) of a signal (Section 7.4; Chapter 10).

The linear-FM signal, widely used in oil exploration [2] and elsewhere because of its simplicity, is a convenient test signal for verifying properties 1 and 2. Property 3 may be tested by a sum of linear-FM signals.

The above properties have further implications. For example, property 3 helps to ensure TFD robustness in the presence of noise, while properties 2 and 3 together allow discernment and discrimination of the multiple IFs of multicomponent signals. Note, however, that in some cases, such as to perform time-frequency filtering, linear time-frequency methods are useful, although they cannot be interpreted directly as energy distributions (see Chapter 11 for more details).

1.2 SIGNAL FORMULATIONS AND CHARACTERISTICS IN THE (t,f) DOMAIN

1.2.1 EXAMPLES OF SIGNAL MODELS

In the terminology of Chapter I, a “type-5” signal (p. 6) may have the form

$$s_e(t) = \left(\sum_{k=1}^M m_k(t) e^{j2\pi \int_0^t f_k(\tau) d\tau} \right) + w(t), \quad (1.2.1)$$

where $m_k(t)$ is the noisy instantaneous amplitude of the k th component, implying the presence of multiplicative noise, while $f_k(t)$ is the IF of the k th component, and $w(t)$ is additive noise. To analyze such a general signal, we must distinguish the *time-varying* components from each other and from the noise. The form of Eq. (1.2.1) strongly suggests a time-frequency method of analysis.

If each $m_k(t)$ is replaced by a deterministic but time-varying amplitude $a_k(t)$, the noise becomes merely additive. If $w(t)$ is also eliminated, the entire signal becomes deterministic—and, with appropriate choices of $a_k(t)$ and $f_k(t)$, might be a passable imitation of speech or music. If the amplitudes $a_k(t)$ are then replaced by constants, the signal becomes what we have called “type 4.” If we then reduce the number of components to 1, we obtain what we have called a “type-3” signal, of which the sinusoidal-FM and linear-FM signals are examples. Even these can be advantageously represented in time-frequency form (Fig. 1.1.6). Only if the amplitudes and frequencies of all the components are constant (as in a “type-2” or “type-1” signal) does the FT give an immediately clear representation.

1.2.2 ANALYTIC SIGNALS

As discussed in Section I.1.4, a signal $s(t)$ is real if and only if (iff)

$$S(-f) = S^*(f), \quad (1.2.2)$$

where $S(f)$ is the FT of $s(t)$. In other words, a real signal is one that exhibits *Hermitian symmetry* between the positive-frequency and negative-frequency components, allowing the latter to be deduced from the former. Hence the negative-frequency components of a real signal *may be eliminated from the signal representation without losing information by forming the analytic signal*. In the case of a real low-pass signal, removal of negative frequencies has two beneficial effects. First, for narrow-band signals, it halves the total bandwidth, allowing the signal to be sampled at half the usual Nyquist rate without aliasing [13,14]. Second, it avoids the appearance of some interference terms generated by the interaction of positive and negative components in quadratic TFDs (to be treated in detail in Section 3.1.2). (Moreover, there is an argument that negative frequencies have no clear physical meaning that directly reflects the real world).

Note that, although an analytic signal contains *no negative frequencies*, it may have a spectral component at zero frequency (DC).

Theorem 1.2.1. *Let us consider two real signals $s(t)$ and $y(t)$; the signal*

$$z(t) = s(t) + jy(t) \quad (1.2.3)$$

is analytic with a real DC component, iff

$$Y(f) = (-j \operatorname{sgn} f) S(f), \quad (1.2.4)$$

where $S(f)$ and $Y(f)$ are the FTs of $s(t)$ and $y(t)$, respectively, and where

$$\operatorname{sgn} \xi \stackrel{\Delta}{=} \begin{cases} -1 & \text{if } \xi < 0; \\ 0 & \text{if } \xi = 0; \\ +1 & \text{if } \xi > 0. \end{cases} \quad (1.2.5)$$

Proof and (Exercise). Take the FT of Eq. (1.2.3) and use Eq. (I.1.3). \square

1.2.3 HILBERT TRANSFORM (HT) AND ANALYTIC ASSOCIATE

If the FTs of $s(t)$ and $y(t)$ are related according to Eq. (1.2.4), we say that $y(t)$ is the Hilbert transform (HT) of $s(t)$, and we write

$$y(t) = \mathcal{H}\{s(t)\}. \quad (1.2.6)$$

Hence, we may restate [Theorem 1.2.1](#) as follows: *A signal is analytic with a real DC component iff its imaginary part is the HT of its real part.*

By invoking the “if” form of [Theorem 1.2.1](#) and restating the sufficient condition in terms of Eq. (1.2.3), we may now see the practical significance of the theorem and the HT: Given a real signal $s(t)$, we can construct the complex signal

$$z(t) = s(t) + j\mathcal{H}\{s(t)\} \quad (1.2.7)$$

and know that $z(t)$ is analytic. This $z(t)$ is called the analytic signal “corresponding to” or “associated with” the real signal $s(t)$. In this book, for convenience, we shall usually call $z(t)$ the *analytic associate* of $s(t)$. Other advantages of using $z(t)$ instead of $s(t)$ are described in [Chapters 2 and 6](#) [12,15].

By taking the IFT of Eq. (1.2.4) and applying Eq. (1.2.6), we arrive at the following concise definition of the HT:

Definition 1.2.2. The Hilbert transform of a signal $s(t)$, denoted by $\mathcal{H}\{s(t)\}$, may be expressed using the FT $\mathcal{F}\{\cdot\}$ of $s(t)$ as:

$$\mathcal{H}\{s(t)\} = \mathcal{F}_{t \leftarrow f}^{-1} \left\{ (-j \operatorname{sgn} f) \mathcal{F}_{t \rightarrow f} \{s(t)\} \right\}. \quad (1.2.8)$$

In other words, the HT of $s(t)$ can be evaluated as follows:

1. take the FT $S(f)$ of $s(t)$;
2. multiply $S(f)$ by $-j$ for $f > 0$, by $+j$ for $f < 0$, and by 0 for $f = 0$; and
3. take the IFT.

According to Step 2 of the above procedure, a Hilbert transformer introduces a phase lag of 90° (as $-j = e^{-j\pi/2}$), producing a signal in *quadrature* to the input signal. The effect is well illustrated by the following result, which is easily verified using [Definition 1.2.2](#) and a table of transforms:

Example 1.2.3. If f_0 is a positive constant, then

$$\mathcal{H}\{\cos(2\pi f_0 t)\} = \sin(2\pi f_0 t) \quad (1.2.9)$$

$$\mathcal{H} \{ \sin(2\pi f_0 t) \} = -\cos(2\pi f_0 t). \quad (1.2.10)$$

It would be convenient for constructing the analytic signal if the pattern of [Example 1.2.3](#) were applicable to modulated signals so that, for example, we could have

$$\mathcal{H} \{ a(t) \cos \phi(t) \} = a(t) \sin \phi(t), \quad (1.2.11)$$

which would imply that the analytic associate of a real signal $s(t) = a(t) \cos \phi(t)$ is

$$\begin{aligned} z(t) &= a(t) \cos \phi(t) + j\mathcal{H} \{ a(t) \cos \phi(t) \} = a(t) \cos \phi(t) + ja(t) \sin \phi(t) \\ &= a(t) e^{j\phi(t)}. \end{aligned} \quad (1.2.12)$$

The condition for Eqs. [\(1.2.11\)](#) and [\(1.2.12\)](#) to hold is that the variation of $a(t)$ is sufficiently slow to ensure “spectral disjointness” [\[15\]](#), that is, to avoid overlap between the spectrum of $a(t)$ and the spectrum of $\cos \phi(t)$. More details appear in [Section 1.4.3](#) and Ref. [\[16\]](#).

Equation [\(1.2.4\)](#) indicates that the transfer function of a Hilbert transformer is $-j \operatorname{sgn} f$. The corresponding impulse response is

$$\mathcal{F}_{t \leftarrow f}^{-1} \{ -j \operatorname{sgn} f \} = \frac{1}{\pi t}. \quad (1.2.13)$$

Definition 1.2.4. Using the above result and applying the convolution property to Eq. [\(1.2.8\)](#), we obtain a definition of the Hilbert transform in the time domain:

$$\mathcal{H} \{ s(t) \} = s(t) * \frac{1}{\pi t} \quad (1.2.14)$$

$$= \frac{1}{\pi} \operatorname{p.v.} \left\{ \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau \right\}, \quad (1.2.15)$$

where $\operatorname{p.v.} \{ \cdot \}$ denotes the Cauchy *principal value* of the improper integral (i.e., to account for the $t = \tau$ situation) [\[17\]](#); the p.v. is given in this case by¹

$$\lim_{\varepsilon \rightarrow 0} \left[\int_{-\infty}^{t-\varepsilon} \frac{s(\tau)}{t - \tau} d\tau + \int_{t+\varepsilon}^{\infty} \frac{s(\tau)}{t - \tau} d\tau \right]. \quad (1.2.16)$$

1.2.4 DURATION, BANDWIDTH, AND BT PRODUCT

If a signal $s(t)$ has the FT $S(f)$, the *duration* of the signal is the range of time outside which $s(t) = 0$, while the *bandwidth* of the signal is the range of frequencies outside which $S(f) = 0$. These definitions, as we shall see, lead to the conclusion that a finite duration implies infinite bandwidth and vice versa.

¹In practice, Eq. [\(1.2.14\)](#) is rarely used in actual computations and implementations because the FT properties make it easier to work in the frequency domain.

In practice, however, signals are observed for finite periods of time, and measured or processed by devices with finite usable bandwidths.² In practice, therefore, the literal definitions need to be relaxed in some way.

1.2.4.1 Finite duration vs finite bandwidth

A *time-limited signal*, of duration T centered at $t = 0$, can be expressed as

$$s_T(t) = s(t) \operatorname{rect}[t/T], \quad (1.2.17)$$

where the subscript T indicates the duration (see Table I.1.1).

The FT of $s_T(t)$ is

$$S_T(f) = S(f) * T \operatorname{sinc} fT, \quad (1.2.18)$$

where $*$ denotes convolution in frequency. Thus, the bandwidth of $S_T(f)$ is infinite.

If, in order to avoid the effects of discontinuities, we replace $\operatorname{rect}[t/T]$ with a smoother window $w(t)$ of the same duration T , we can write

$$s_T(t) = s(t) \operatorname{rect}[t/T] w(t), \quad (1.2.19)$$

whose FT still involves a convolution with $\operatorname{sinc} fT$ and therefore still gives an infinite bandwidth.

In analogy to the time-limited case, a *band-limited signal*, of bandwidth B centered at the origin, can be expressed in the frequency domain as

$$S_B(f) = S(f) \operatorname{rect}[f/B]. \quad (1.2.20)$$

In the time domain, the signal is given by

$$s_B(t) = s(t) * B \operatorname{sinc} Bt, \quad (1.2.21)$$

which has an infinite duration.

Thus, under the “literal” definitions, finite T implies infinite B and vice versa.

1.2.4.2 Practical measures of bandwidth and duration

If there is no finite bandwidth containing *all* the energy of the signal, there may still be a finite bandwidth containing *most* of the energy. Hence, for example, the minimum bandwidth containing 99% of the signal energy might be accepted as a useful measure of the signal bandwidth. If the nominated fraction of the signal energy were confined between the frequencies f_{\min} and f_{\max} , the bandwidth would be $B = f_{\max} - f_{\min}$.

Similarly, the minimum duration containing a nominated fraction of the signal energy might be accepted as a useful measure of the signal duration.

A less arbitrary but less conservative measure of bandwidth, published by Gabor³ in 1946 [18], is the so-called *effective bandwidth* B_e , defined by

²All signal acquisition and measuring systems are low-pass as their $|f|$ must be bounded.

³Dennis Gabor (1900-1979), a Hungarian-born electrical engineer who settled in Britain, is best known for the invention of holography (1947-1948), for which he was awarded the Nobel Prize for Physics in 1971.

$$B_e^2 = \frac{1}{E_s} \int_{-\infty}^{\infty} f^2 |S(f)|^2 df, \quad (1.2.22)$$

where $S(f)$ is the FT of the signal, and E_s is the total energy of the signal:

$$E_s = \int_{-\infty}^{\infty} |S(f)|^2 df. \quad (1.2.23)$$

Thus B_e^2 is the normalized *second moment* of $|S(f)|^2$ w.r.t. frequency, about the origin ($f = 0$). For brevity, we call B_e^2 the “2nd moment of the signal w.r.t. f .”

Similarly, the so-called *effective duration* T_e is defined as

$$T_e^2 = \frac{1}{E_s} \int_{-\infty}^{\infty} t^2 |s(t)|^2 dt, \quad (1.2.24)$$

that is, the normalized 2nd moment of $|s(t)|^2$ w.r.t. time, about the origin ($t = 0$). For brevity, we refer to T_e^2 as “the 2nd moment of the signal w.r.t. time.”

As an aid to remembering these definitions, note that if f were a random variable and $|S(f)|^2$ were its probability density function (p.d.f.), then we would have $E_s = 1$ so that B_e^2 would be the variance of f if the mean of f were 0. Thus the effective bandwidth B_e is analogous to the standard deviation of f . Similarly, the effective duration T_e is analogous to the standard deviation of t . When we consider how little of a typical probability distribution falls within an interval of one standard deviation, we realize that the “effective” bandwidth or duration is only a mathematical construction, *not* an estimate of the actual bandwidth or duration required to measure or process the signal.

The bandwidth or duration needed for sufficiently accurate processing of a signal is known as the *significant* or *essential* bandwidth or duration. For a given signal $s(t)$ whose FT is $S(f)$, let us define a “time-truncation” signal $\hat{s}(t)$ that satisfies

$$\hat{s}(t) = \begin{cases} 0 & \text{if } t < t_1 \\ s(t) & \text{if } t_1 \leq t \leq t_2 \\ 0 & \text{if } t > t_2, \end{cases} \quad (1.2.25)$$

where $t_2 > t_1$, so that the duration of $\hat{s}(t)$ is $t_2 - t_1$. Let a “frequency-truncation” be defined similarly in the frequency domain. Then the “essential” duration of $s(t)$ is the duration of the shortest time-truncation $\hat{s}(t)$ that is an acceptable substitute for $s(t)$, while the “essential” bandwidth of $s(t)$ is the bandwidth of the most narrow-band frequency-truncation $\hat{S}(f)$ that is an acceptable substitute for $S(f)$. Notice that the previous definitions of effective duration/bandwidth do not rely on the idea of “truncation.” What is “acceptable” depends on the application. According to Slepian’s definition [19], an acceptable time- or frequency-truncation is one that the detecting or measuring apparatus cannot distinguish from the original signal.

Given B and T , the *bandwidth-duration product* BT is self-explanatory.⁴ If we use conservative (essential) values of B and T , the BT product has a simple practical significance: because a signal of total bandwidth B can be reconstructed from samples at the sampling rate B , the product BT is the

⁴If the bandwidth is written $|f| < W$, then $B = 2W$. Hence, “ BT ” is also known as “ $2WT$.”

total number of samples required to represent the signal.⁵ If, due to noise or other sources of error, each sample has an accuracy of n bits, then the signal can encode nBT bits of information. Thus BT is proportional to the *information richness* of the signal: one cannot extract large amounts of information from a signal with a small BT product [16].

Example 1.2.5. For the Gaussian signal $s(t) = e^{-\alpha^2 t^2}$, the effective duration is

$$T_e = \frac{1}{2\alpha}. \quad (1.2.26)$$

The FT of $s(t)$ is $S(f) = \frac{\sqrt{\pi}}{\alpha} e^{-\pi^2 f^2 / \alpha^2}$ so that the effective bandwidth is

$$B_e = \frac{\alpha}{2\pi}. \quad (1.2.27)$$

From Eqs. (1.2.26) and (1.2.27), $B_e T_e = \frac{1}{4\pi}$.

The Gaussian signal is the only signal for which this equality holds, and that for all other signals $B_e T_e > \frac{1}{4\pi}$ as per Eq. (T.2.4) [18] (see also p. 20).

1.2.5 ASYMPTOTIC SIGNALS

For signals with significant information content, it is desired to know not only the overall bandwidth B but also the distribution of energy through the bandwidth, for example, the frequencies present, their relative amplitudes, and the times during which they are significant. Similarly we may want to know not only the overall duration T but also the distribution of energy throughout the duration, for example, the times during which the signal is present, the relative amplitudes at those times, and the significant frequencies present during those times. Such signals may be modeled by the class of asymptotic signals.

Definition 1.2.6. A signal $s(t)$ is *asymptotic* iff it has the following properties:

- (a) The essential duration T , as defined previously, is finite.
- (b) The essential bandwidth B , as defined previously, is finite.
- (c) The product BT is large (e.g., > 10).
- (d) The amplitude is bounded, and the energy $E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt$ is finite.

As an example, the two-component linear-FM signal in Fig. 1.1.8 (p. 42) is asymptotic because:

- (a) from the time-domain graph (left) or the TFD, we see that the duration is just under 400 s;
- (b) from the magnitude spectrum (bottom) or the TFD, we see that the signal is almost entirely between 0.15 and 0.35 Hz, giving an essential bandwidth of about 0.2 Hz if negative frequencies are not considered significant (or 0.7 Hz if they *are* considered significant);
- (c) if we accept the lower estimate of the essential bandwidth, the BT product is almost 80; and
- (d) from the time-domain graph or the TFD, we get a clear impression that the amplitude is bounded.

Asymptotic signals allow useful approximations for deriving simple signal models (e.g., to express analytic signals).

⁵In the language of statistics, BT is the number of *degrees of freedom* in the signal. In the language of linear algebra, BT is the *dimension* of the signal, considered as a vector.

1.2.6 MONOCOMPONENT VS MULTICOMPONENT SIGNALS

A *monocomponent* signal is described in the (t, f) domain by one single “ridge,” corresponding to an elongated region of energy concentration. Furthermore, interpreting the crest of the “ridge” as a graph of IF versus time, we require the IF of a monocomponent signal to be a single-valued function of time [16, p. 527]. [Figure 1.1.7](#) shows an example.

Such a monocomponent signal $s(t)$ has an analytic associate of the form

$$z(t) = a(t) e^{j\phi(t)}, \quad (1.2.28)$$

where

- $a(t)$, known as the instantaneous amplitude, is real and positive;
- $\phi(t)$, known as the instantaneous phase, is differentiable; and
- $a(t)$ and $e^{j\phi(t)}$ are spectrally disjoint.

Without the requirement of spectral disjointness, *any* complex signal would qualify as monocomponent because any complex signal can be written in the form of Eq. (1.2.28). If, in addition, $s(t)$ itself is real and asymptotic with amplitude modulation $a(t)$, it can also be expressed approximately as [16]

$$s(t) = a(t) \cos \phi(t). \quad (1.2.29)$$

A *multicomponent* signal may be described as the sum of two or more monocomponent signals (having analytic associates of the form Eq. (1.2.28)) such that:

$$z(t) = \sum_{i=1}^M a_i(t) e^{j\phi_i(t)}. \quad (1.2.30)$$

The decomposition into components is not necessarily unique [16, p. 535]. This model allows the extraction and separation of components from a given multi-component signal using (t, f) filtering methods [20]. Further in-depth discussion of monocomponent and multicomponent signals can be found in Refs. [1, 10]. [Figure 1.1.8](#) (p. 42) shows an example of a multicomponent signal with two components.

1.3 INSTANTANEOUS FREQUENCY (IF) AND SPECTRAL DELAY (SD)

1.3.1 INSTANTANEOUS FREQUENCY

Definition 1.3.1. The *instantaneous frequency* of a monocomponent signal is

$$f_i(t) = \frac{1}{2\pi} \phi'(t) \equiv \frac{1}{2\pi} \frac{d\phi}{dt}(t), \quad (1.3.1)$$

where $\phi(t)$ is the instantaneous phase of the signal, and the prime ('') indicates differentiation.

This formulation is justified below by considering first a constant-frequency signal, then a variable-frequency signal.

Consider the amplitude-modulated signal

$$x(t) = a(t) \cos(2\pi f_c t + \psi), \quad (1.3.2)$$

where f_c and ψ are constant. As t increases by the increment $1/f_c$, the argument of the cosine function increases by 2π and the signal passes through one cycle. So the period of the signal is $1/f_c$, and the frequency, being the reciprocal of the period, is f_c . The same signal can be written as

$$x(t) = a(t) \cos \phi(t), \quad (1.3.3)$$

where

$$\phi(t) = 2\pi f_c t + \psi, \quad (1.3.4)$$

from which we obtain

$$f_c = \frac{1}{2\pi} \phi'(t) \equiv \frac{1}{2\pi} \frac{d\phi}{dt}(t). \quad (1.3.5)$$

Although the left-hand side of this equation (the frequency) has been assumed constant, the right-hand side would be variable if $\phi(t)$ were a nonlinear function. So, let us check whether this result can be extended to a *variable frequency*.

Consider a monocomponent signal $s(t)$ whose analytic associate is

$$z(t) = a(t) e^{j\phi(t)}, \quad (1.3.6)$$

$s(t)$ is then given by Eq. (1.2.29), where $a(t)$ and $\phi(t)$ are real and $a(t)$ is positive; then $a(t)$ is called the *instantaneous amplitude* and $\phi(t)$ is called the *instantaneous phase*. Let $\phi(t)$ be evaluated at $t = t_1$ and $t = t_2$, where $t_2 > t_1$. By the mean value theorem of elementary calculus, if $\phi(t)$ is differentiable, there exists a time instant t between t_1 and t_2 such that

$$\phi(t_2) - \phi(t_1) = (t_2 - t_1) \phi'(t). \quad (1.3.7)$$

Let p_i be the period of one particular oscillation of $s(t)$, and let $f_i = 1/p_i$. If we let $t_2 = t_1 + p_i$, then $\phi(t_2) = \phi(t_1) + 2\pi$ so that Eq. (1.3.7) becomes

$$2\pi = p_i \phi'(t); \quad (1.3.8)$$

that is,

$$f_i = \frac{\phi'(t)}{2\pi} \equiv \frac{1}{2\pi} \frac{d\phi}{dt}(t). \quad (1.3.9)$$

Now t is an instant during a cycle of oscillation and f_i is the frequency (inverse of the period) of that oscillation, suggesting that the right-hand side be defined as the *IF* at time t , as in Definition 1.3.1 above.

Comparing Eqs. (1.2.29) and (1.3.6), we see that

$$s(t) = \operatorname{Re}\{z(t)\}, \quad (1.3.10)$$

where $\operatorname{Re}\{\cdot\}$ denotes the real part. Now let us define

$$y(t) \stackrel{\Delta}{=} \operatorname{Im}\{z(t)\} = a(t) \sin \phi(t), \quad (1.3.11)$$

where $\operatorname{Im}\{\cdot\}$ denotes the imaginary part.

Using Eq. (1.3.1), we can easily confirm that the signals described in the time domain by Eqs. (1.1.3) and (1.1.4) have the IFs given in Eqs. (1.1.1) and (1.1.2), respectively (sinusoidal FM and linear FM).

Definition 1.3.1 is strictly meaningful only for a monocomponent signal; a multicomponent signal ought to have a separate IF for each component. In general, the IF of the sum of two signals is *not* the sum of their two IFs. Further details regarding the IF can be found in Ref. [21].

The IF is a detailed description of the frequency characteristics of a signal. This contrasts with the notion of “average frequency” defined next.

Definition 1.3.2. The *average frequency* of a signal is

$$f_0 = \frac{\int_0^\infty f |S(f)|^2 df}{\int_0^\infty |S(f)|^2 df}, \quad (1.3.12)$$

where $S(f)$ is the FT of the signal.

In other words, f_0 is the *first moment* of $|S(f)|^2$ w.r.t. frequency; that is, we define the “average” frequency as if $|S(f)|^2$ were the p.d.f. of the frequency.

Notice that if $|S(f)|^2$ is replaced by a TFD, f_0 becomes a function of time, suggesting that perhaps the first moment of a TFD w.r.t. frequency is a measure of IF. The conditions under which this is true will be explained in due course (and summarized in the table on p. 122). For now we simply note that any reasonable (t, f) representation of a signal should contain information about the IF laws of the components. Ideally, a TFD should show the signal energy concentrated along the IF. This criterion is often used for adapting the parameters of the analysis window of the Spectrogram [22]. The IF estimate obtained as derivative of phase can be related to the first-order moment of a quadratic TFD (QTFD) [23]. Specifically, the IF estimate obtained using the first-order moment of the TFD is equivalent to applying a low-pass filtering function to the direct estimator based on the central finite differencing (CFD) of the phase of the analytic signal [23]. In particular, it would be most convenient if the *crests* of the ridges in the (t, f) domain represented the IF laws.

1.3.2 SPECTRAL DELAY

The IF of a signal indicates the dominant frequency of the signal at a given time. Let us now seek a *dual* or “inverse” of the IF, indicating the dominant time when a given frequency occurs.

If $z(t)$ is an analytic signal with FT

$$Z(f) = A \delta(f - f_i), \quad (1.3.13)$$

where A is a real constant, the dominant frequency is f_i . Taking the IFT of $Z(f)$ gives

$$z(t) = A e^{j2\pi f_i t}. \quad (1.3.14)$$

The instantaneous phase of $z(t)$, denoted by $\phi(t)$, is

$$\phi(t) = \arg z(t) = 2\pi f_i t \quad (1.3.15)$$

so that

$$f_i = \frac{1}{2\pi} \phi'(t) \equiv \frac{1}{2\pi} \frac{d\phi}{dt}(t). \quad (1.3.16)$$

Although this result has been obtained for constant frequency, it is also valid for a variable-frequency signal, as explained earlier. More specifically, for a variable frequency $f_i(t)$, the phase can be written as $\phi(t) = 2\pi \int_0^t f_i(\tau) d\tau + \psi$, where ψ is the initial phase (often considered to be zero). The real signal then becomes:

$$s(t) = A \cos \phi(t). \quad (1.3.17)$$

Now let us repeat the argument with the time and frequency variables interchanged. The signal

$$z(t) = a \delta(t - \tau_d), \quad (1.3.18)$$

is an impulse at time τ_d , and a is in general complex. If we ask what is the “delay” of this signal, the only sensible answer is τ_d . The FT of this signal is

$$Z(f) = a e^{-j2\pi f \tau_d} \quad (1.3.19)$$

and its phase, denoted by $\theta(f)$, is

$$\theta(f) = \arg Z(f) = -2\pi f \tau_d + \arg a \quad (1.3.20)$$

so that

$$\tau_d = -\frac{1}{2\pi} \theta'(f) \equiv -\frac{1}{2\pi} \frac{d\theta}{df}(f). \quad (1.3.21)$$

Again, although this result has been obtained for a constant τ_d , the right-hand side of Eq. (1.3.21) is well defined even if $\theta'(f)$ varies with f .

Definition 1.3.3. If $z(t)$ is an analytic signal with FT $Z(f)$, and $\theta(f) = \arg Z(f)$ then the spectral delay of $z(t)$, denoted by $\tau_d(f)$, is

$$\tau_d(f) = -\frac{1}{2\pi} \theta'(f) \equiv -\frac{1}{2\pi} \frac{d\theta}{df}(f). \quad (1.3.22)$$

Notice that the definitions of τ_d and f_i are similar, except that time and frequency are interchanged and Eq. (1.3.22) has an extra minus sign; hence, we say that SD is the *dual* of IF.¹ Figure 1.3.1 illustrates these concepts.

Seeing that the IF $f_i(t)$ is a function assigning a frequency to a given time, whereas the SD $\tau_d(f)$ is a function assigning a time to a given frequency, we may well ask whether the two functions are inverses of each other. In fact, they are not *always* inverses because the IF function may not be invertible. So, let us restrict the question to *invertible* signals, that is, monocomponent signals whose IFs are monotonic functions of time.

One example of an invertible signal is the generalized Gaussian signal, that is, a linear-FM signal with a Gaussian envelope. Let such a signal peak at time t_0 , with peak amplitude A , center frequency f_c ,

¹The SD τ_d is often called the “time delay.” This term, although well established, is tautological in that “delay” implies “time.” The term *instantaneous frequency* is a quantity with the dimension of frequency modified by an adjective indicating localization in time. The dual of this term should therefore be a quantity with the dimension of time modified by an adjective indicating localization in frequency, for example, “frequency-localized delay” or, more succinctly, *spectral delay*. This term is used as a standard terminology in this book.

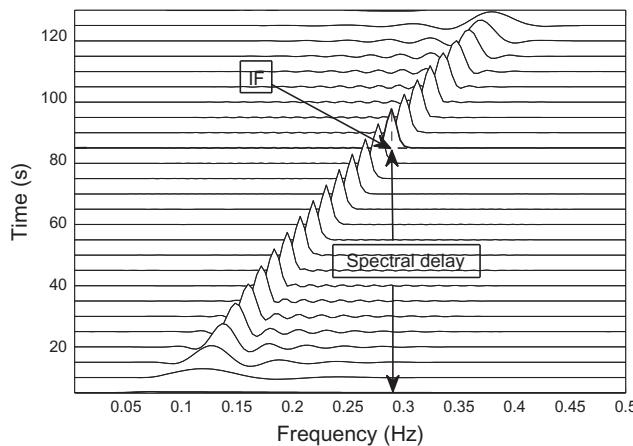
**FIGURE 1.3.1**

Illustration of instantaneous frequency (IF) and spectral delay (SD): LFM signal of length 128 samples, starting frequency = 0.1 Hz, finishing frequency = 0.4 Hz with sampling rate = 1; at time 80, the value of the IF is about 0.3; in other words, for the frequency value 0.3, the SD is 80.

sweep rate α , and decay constant β , and suppose that the IF is positive while the envelope is significant. Then the analytic associate of the signal is

$$z(t) = A \exp \left(j2\pi \left[f_c [t - t_0] + \frac{\alpha + j\beta}{2} [t - t_0]^2 \right] \right) \quad (1.3.23)$$

and, from Eq. (1.3.16) with $a(t) = A \exp(-\pi\beta(t - t_0)^2)$, its IF is

$$f_i(t) = f_c + \alpha[t - t_0]. \quad (1.3.24)$$

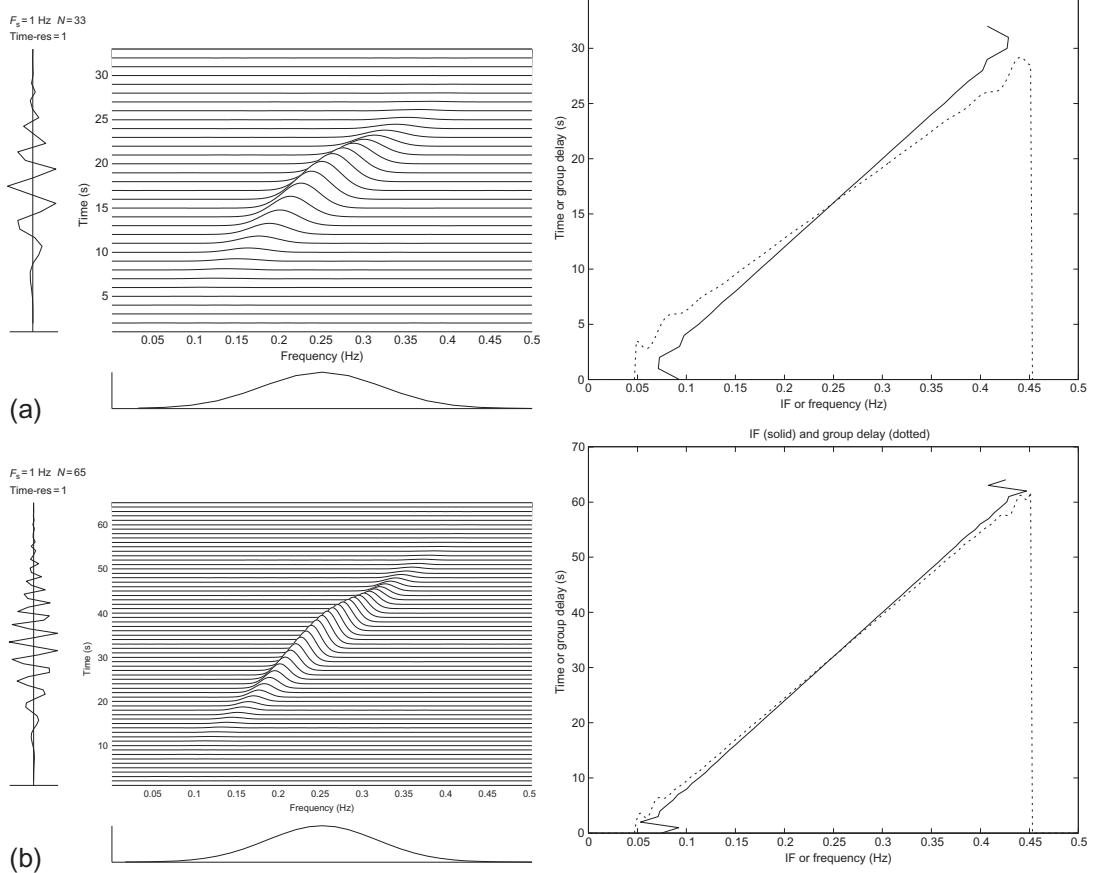
To find the inverse function of $f_i(t)$, we simply solve for t , obtaining

$$t = t_0 + \frac{f_i(t) - f_c}{\alpha}, \quad (1.3.25)$$

which suggests that the SD $\tau_d(f)$ of $z(t)$ can be estimated by

$$\hat{\tau}_d(f) = t_0 + \frac{f - f_c}{\alpha}. \quad (1.3.26)$$

In general, τ_d and $\hat{\tau}_d$ are not equal, but converge to each other as the BT product increases. As an example, Fig. 1.3.2 shows the IF (solid line) and SD (dotted line) of the generalized Gaussian signal $z(t)$ for two values of $B_e T_e$, where the subscript “e” means “effective.” Note that the two curves are closer for a larger value of $B_e T_e$.

**FIGURE 1.3.2**

Instantaneous frequency and spectral delay for a linear-FM signal with a Gaussian envelope: (a) total duration $T = 33$ s, $B_e T_e = 0.1806$; (b) $T = 65$ s, $B_e T_e = 0.3338$. For each signal, the left-hand graph shows the time trace, spectrum, and TFD (WVD, defined later in Eq. (2.1.17)), while the right-hand graph shows the IF (solid line) and SD (dotted line). The vertical dotted segments are caused by truncation of the frequency range (to avoid finite-precision effects).

1.3.3 MEAN IF AND GROUP DELAY

Let $z(t)$ be a band-pass analytic signal with center frequency f_c and FT:

$$Z(f) = M(f - f_c) e^{j\theta(f)}, \quad (1.3.27)$$

where the magnitude $M(f - f_c)$ and phase $\theta(f)$ are real. If the signal has *linear phase* in support of $Z(f)$, that is, if $\theta(f)$ is a linear function of f wherever $Z(f)$ is nonzero, we can let

$$\theta(f) = -2\pi \tau_p f_c - 2\pi \tau_g [f - f_c], \quad (1.3.28)$$

where τ_p and τ_g are real-time constants. Equation (1.3.27) then becomes

$$Z(f) = M(f - f_c) e^{-j(2\pi\tau_p f_c + 2\pi\tau_g [f - f_c])} \quad (1.3.29)$$

$$= e^{-j2\pi f_c \tau_p} M(f - f_c) e^{-j2\pi \tau_g [f - f_c]}. \quad (1.3.30)$$

Taking the IFT of $Z(f)$, we find

$$z(t) = m(t - \tau_g) e^{j2\pi f_c [t - \tau_p]}, \quad (1.3.31)$$

where $m(t)$ is the IFT of $M(f)$. Since $M(f)$ is real, $m(t)$ is Hermitian (i.e., $m(-t) = m^*(t)$) so that $|m(t)|$ is even. Hence τ_g is the time about which the envelope function is symmetrical; for this reason, τ_g is called the *GD* [24, pp. 123-124]. The phase of the oscillatory factor is $-2\pi f_c \tau_p$, wherefore τ_p is called the *phase delay*. These observations and Eq. (1.3.28) lead to the definition below:

Definition 1.3.4. If an analytic signal has the FT $Z(f) = |Z(f)| e^{j\theta(f)}$, then the group delay of the signal reduces to the spectral delay, that is,

$$\tau_g(f) = -\frac{1}{2\pi} \theta'(f) \quad (1.3.32)$$

and the *phase delay* of the signal is

$$\tau_p(f) = -\frac{1}{2\pi f} \theta(f). \quad (1.3.33)$$

Equation (1.3.32) is found by differentiating Eq. (1.3.28) w.r.t. f , and Eq. (1.3.33) is found by putting $f_c = f$ in Eq. (1.3.28). Whereas Eq. (1.3.28) assumes linear phase, the above definition is meaningful whether the phase is linear or not.

Thus, the GD (Eq. 1.3.32) equals the SD (Eq. 1.3.22), although the physical interpretations are different; the SD applies to an impulse, whereas the GD applies to the envelope of a narrowband signal.

Now consider the dual of the above argument. Let $z(t)$ be a time-limited signal centered on $t = t_c$, and let

$$z(t) = a(t - t_c) e^{j\phi(t)}, \quad (1.3.34)$$

where $a(t)$ and $\phi(t)$ are real. If the signal has *constant IF* in support of $z(t)$, that is, if $\phi(t)$ is a linear function of t wherever $z(t)$ is nonzero, we can let

$$\phi(t) = 2\pi f_0 t_c + 2\pi f_m [t - t_c], \quad (1.3.35)$$

where f_0 and f_m are real frequency constants. Equation (1.3.34) then becomes

$$z(t) = a(t - t_c) e^{j(2\pi f_0 t_c + 2\pi f_m [t - t_c])} \quad (1.3.36)$$

$$= e^{j2\pi f_0 t_c} a(t - t_c) e^{j2\pi f_m [t - t_c]}. \quad (1.3.37)$$

Taking the FT of $z(t)$, we find

$$Z(f) = A(f - f_m) e^{-j2\pi [f - f_0] t_c}, \quad (1.3.38)$$

where $A(f)$ is the FT of $a(t)$. Now because $a(t)$ is real, $A(f)$ is Hermitian so that $|A(f)|$ is even. Hence, f_m is the frequency about which the amplitude spectrum is symmetrical; for this reason, f_m is called the *mean IF*. Differentiating Eq. (1.3.35) w.r.t. t leads to the following definition:

Definition 1.3.5. For the signal

$$z(t) = |z(t)| e^{j\phi(t)}, \quad (1.3.39)$$

the *mean IF* is

$$f_m(t) = \frac{1}{2\pi} \phi'(t) \equiv \frac{1}{2\pi} \frac{d\phi}{dt}(t). \quad (1.3.40)$$

Thus the mean IF is the same as the IF defined earlier (Eq. 1.3.1), but the physical interpretations are different. The IF has been derived for a tone (and earlier for a modulated sinusoid), whereas the mean IF applies to the spectrum of a short-duration signal.

1.3.4 RELAXATION TIME, DYNAMIC BANDWIDTH

For a linear-FM signal, the instantaneous phase $\phi(t)$ is quadratic. So $\phi(t)$ can be expanded in a Taylor series about $t = t_0$:

$$\phi(t) = \phi(t_0) + \phi'(t_0)[t - t_0] + \frac{1}{2}\phi''(t_0)[t - t_0]^2 \quad (1.3.41)$$

$$= \phi(t_0) + 2\pi f_i(t_0)[t - t_0] + \frac{1}{2}2\pi f_i'(t_0)[t - t_0]^2. \quad (1.3.42)$$

The *relaxation time* T_r , as defined by Rihaczek [25, p. 374], is the duration over which the instantaneous phase deviates no more than $\pi/4$ from linearity. That is,

$$\left| \frac{1}{2}2\pi f_i'(t_0)T_r^2/4 \right| = \pi/4. \quad (1.3.43)$$

Solving this equation leads to the following definition:

Definition 1.3.6. The *relaxation time* of a signal is

$$T_r(t) = \left| \frac{df_i(t)}{dt} \right|^{-1/2}, \quad (1.3.44)$$

where $f_i(t)$ is the IF.

The dual of relaxation time, known as *dynamic bandwidth*, is the bandwidth over which the phase spectrum, assumed to be a quadratic function of frequency, deviates no more than $\pi/4$ from linearity. The result is as follows:

Definition 1.3.7. The *dynamic bandwidth* of a signal is

$$B_d(f) = \left| \frac{d\tau_d(f)}{df} \right|^{-1/2}, \quad (1.3.45)$$

where $\tau_d(f)$ is the SD.

1.3.4.1 Interpretation

As the relaxation time is a measure of the time needed to observe significant variations in IF, so the dynamic bandwidth is a measure of the bandwidth needed to observe significant variations in SD.

Definition 1.3.8. The *instantaneous bandwidth (IB)* of a signal $s(t)$ is a measure of frequency spread at a given time, and it can be expressed as [26,27]:

$$B_i(t) = \sigma_f^2(t) = \frac{\int_{-\infty}^{\infty} (f - f_i(t))^2 \rho(t, f) df}{\int_{-\infty}^{\infty} \rho(t, f) df}. \quad (1.3.46)$$

Heuristic Proof. By analogy with the PDF, the classical bandwidth of a stationary signal may be defined as the deviation of frequency about the mean frequency of the signal and it can be expressed as:

$$B = \sigma_{BW}^2(t) = \frac{\int_{-\infty}^{\infty} (f - f_c)^2 |S(f)|^2 df}{\int_{-\infty}^{\infty} |S(f)|^2 df}. \quad (1.3.47)$$

For nonstationary signals, the square magnitude spectrum representing an energy density is replaced by the TFD and the concept of bandwidth can be extended by replacing the mean frequency f_c by the IF so that the IB of a monocomponent signal represents the deviations about the IF for a given time, resulting in Eq. (1.3.46). For a multicomponent signal, at a certain time, the IB represents the difference between the highest value of subcomponent IF and the lowest value of subcomponent IF, assuming the subcomponents IB are negligible. Otherwise, a corrective factor needs to be introduced. \square

1.3.4.2 Illustration

Consider a newborn EEG signal which is multicomponent and nonstationary in nature with subcomponents IFs shown in Fig. 1.3.3 [9]. One can observe that the multicomponent signal IB at the 4th second

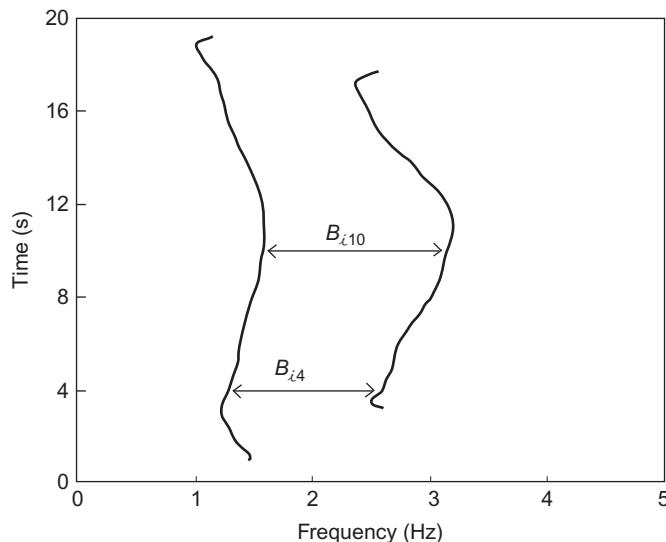


FIGURE 1.3.3

IF laws and IB of an epoch of a multicomponent newborn EEG signal. (The signal IB has two distinct values at $t = 4$ and $t = 10$; not to be confused with the subcomponents IB.)

is different from the multicomponent signal IB at the 10th second. As the IFs change over time, the IB will also vary. This IB measure is another important characteristic of a nonstationary multicomponent signal, useful in some applications [26].

1.4 DEFINING AMPLITUDE, PHASE, AND IF USING THE ANALYTIC SIGNAL

This section revisits material discussed earlier and further elaborates on the equivalence between a signal $x(t)$ and the two characteristics of amplitude and phase.

1.4.1 AMPLITUDE AND PHASE FORMULATION FOR THE AM/FM MODEL

Following the approach taken in [Section 1.3](#), let us consider a monochromatic signal $a \cos(2\pi ft + \psi)$ with amplitude a , frequency f , and initial phase ψ . The argument of the cosine function is the linear function $\phi(t) = 2\pi ft + \psi$ so that $f = \phi'(t)/(2\pi)$. Hence, for a more general amplitude- and frequency-modulated signal

$$x(t) = a(t) \cos \phi(t), \quad (1.4.1)$$

where $a(t) \geq 0$ and $\phi(t)$ is *not* necessarily linear, it is natural to define the *instantaneous amplitude* as $a(t)$, the *instantaneous phase* as $\phi(t)$, and the IF as $f_i(t) = \phi'(t)/(2\pi)$.

Unfortunately, these widely used “natural” definitions are sometimes ambiguous. Consider, for example, the problem of defining the amplitude and phase of the signal $y(t) = \cos \phi_1(t) \cos \phi_2(t)$. One solution could be $a(t) = \cos \phi_1(t)$ and $\phi(t) = \phi_2(t)$. Another could be $a(t) = \cos \phi_2(t)$ and $\phi(t) = \phi_1(t)$. By specifying $a(t)$ and $\phi(t)$, we uniquely specify $x(t)$ in Eq. (1.4.1). But the *inverse* problem—finding $a(t)$ and $\phi(t)$, given $x(t)$ —has infinitely many solutions [28]: if $\phi_2(t)$ is *any* function such that $\cos \phi_2(t)$ has the same sign as $x(t)$, we can find $a_2(t) \geq 0$ such that

$$x(t) = a_2(t) \cos \phi_2(t). \quad (1.4.2)$$

To obtain an inverse problem with a unique solution, we must impose further conditions. There are numerous ways to do this. The most classical approach uses the HT and the *analytic signal* (AS), as presented in [Section 1.3](#). Let us now further explore these concepts and some of their implications concerning the definitions and properties of the IF.

1.4.2 ANALYTIC SIGNAL, HT, AND AM/FM MODEL

Let us seek to define a one-to-one correspondence between the real signal $x(t)$, and the pair of real functions $[a(t), \phi(t)]$ comprising the instantaneous amplitude and phase, with $f_i(t) = \phi'(t)/(2\pi)$; in other words, find a one-to-one correspondence between $x(t)$ and a complex signal $z(t)$, whose modulus and argument are $[a(t), \phi(t)]$.

Let $X(f)$ be the FT of the real signal $x(t)$. As in Eq. (1.2.2), $X(f)$ satisfies the Hermitian symmetry $X(-f) = X^*(f)$. The analytic signal (AS) $z(t)$ is then obtained from $x(t)$ by applying $Z(f) = 0$ for $f < 0$ to find the FT of $z(t)$ as

$$Z(f) = (1 + \text{sgn } f) X(f). \quad (1.4.3)$$

Hence the FT of $z^*(t)$ is $Z^*(-f) = (1 - \text{sgn} f)X^*(-f) = (1 - \text{sgn} f)X(f)$. Adding this result to Eq. (1.4.3) and taking IFTs returns $x(t) = (1/2)[z(t) + z^*(t)]$ as the *real part* of the AS. Thus for any $x(t)$, we have specified $z(t)$ and vice versa: a one-to-one correspondence, as previously defined in Eq. (1.2.3).

As per Eq. (1.2.6), the HT of the real signal $x(t)$, $\mathcal{H}\{x(t)\}$, defines the *imaginary part* $y(t)$ of the AS. Let its FT be $Y(f)$; then $z(t) = x(t) + jy(t)$, so that $Z(f) = X(f) + jY(f)$, which satisfies Eq. (1.4.3) iff $Y(f) = -jX(f) \text{ sgn } f$ (cf. Eq. 1.2.4 and Example 1.2.3). Then, taking the IFT yields $y(t)$.

Equation (1.4.3) indicates that $z(t)$ and $y(t)$ are deduced from $x(t)$ through linear filters with frequency responses $1 + \text{sgn } f$ and $-j \text{sgn } f$, respectively.

As $Z(f)$ does not have Hermitian symmetry, $z(t)$ cannot be real. Let the modulus (non-negative) of $z(t)$ be $a(t)$ and the argument (modulo 2π) be $\phi(t)$ so that

$$z(t) = a(t) \exp[j\phi(t)]. \quad (1.4.4)$$

Then the real signal $x(t)$ is associated with a unique pair $[a(t), \phi(t)]$, which may be called its *canonical pair* [29]. This is a consequence of Eq. (1.4.3).

There are questions on the physical meaning of the AS, some of which are discussed in Ref. [28]. It is shown in Refs. [15,30] that if we make some *a priori* physical assumptions, the only possible definition of the instantaneous amplitude and phase is the one given above; and it allows easy calculations and interpretations.

1.4.3 CHARACTERIZATION OF AM/FM CANONICAL PAIRS $[a(t), \phi(t)]$

An arbitrary pair $[a(t), \phi(t)]$ is not necessarily canonical in the sense defined above, given the requirement that the complex function $z(t)$ defined by Eq. (1.4.4) is an AS (i.e., its FT is zero for negative frequencies). Let us discuss the consequences.

As the characterization of an AS is deduced from the spectral condition (zero FT for negative frequencies), it is tempting to translate the same idea for $a(t)$ and $\phi(t)$. Noting that most applications of AM/FM concern narrow-band signals, it is appropriate in such cases to separate the contribution of the carrier frequency by using, instead of Eq. (1.4.1), a signal of the form

$$x(t) = a(t) \cos[2\pi f_0 t + \psi(t)]. \quad (1.4.5)$$

This is the real part of $w(t) \exp(j2\pi f_0 t)$, where $w(t) = a(t) \exp[j\psi(t)]$ is called the *complex envelope* of $x(t)$. We therefore seek conditions on $[a(t), \psi(t)]$ which ensure that $w(t) \exp(j2\pi f_0 t)$ is an AS, as a generalization of Eq. (1.2.11).

An elementary answer, introducing the idea of an asymptotic AS (see below), is that any physical signal $w(t)$ is approximately band-limited and may therefore be made approximately an AS by shifting it far enough in the positive frequency direction—that is, by multiplying it by $\exp(j2\pi f_0 t)$ for a sufficiently high f_0 [16,31]. This reasoning applies rigorously if the signal is strictly band-limited.

For amplitude modulation, the instantaneous amplitude-phase pair is $[a(t), 2\pi f_0 t + \psi]$, where ψ is constant. It is easy to verify that this pair is canonical iff $a(t)$ is a positive band-limited signal whose FT is zero for $|f| > f_0$ [28].

Let us now try to use spectral methods for the characterization of more general pairs of functions $[a(t), \phi(t)]$.

Saying that $a(t) \exp[j\phi(t)]$ is an AS implies (or is equivalent to saying) that the HT of $a(t) \cos \phi(t)$ equals $a(t) \sin \phi(t)$, which is ensured by Bedrosian theorem [32] on the HT of a product of two real functions $x_1(t)$ and $x_2(t)$. A simple derivation of this theorem and some extensions can be found in Ref. [28]. The main result is as follows: let $X_1(f)$ and $X_2(f)$ be the FTs of $x_1(t)$ and $x_2(t)$, respectively. If $x_1(t)$ and $x_2(t)$ are low- and high-frequency signals such that $X_1(f) = 0$ for $f > B$ and $X_2(f) = 0$ for $f < B$, then

$$\mathcal{H}\{x_1(t)x_2(t)\} = x_1(t)\mathcal{H}\{x_2(t)\}. \quad (1.4.6)$$

It follows that if $a(t)$ is a low-frequency signal and $\cos \phi(t)$ a high-frequency signal as defined above, then

$$\mathcal{H}\{a(t)\cos \phi(t)\} = a(t)\mathcal{H}\{\cos \phi(t)\} = a(t)\sin \phi(t). \quad (1.4.7)$$

The estimation and computational aspects of the HT, analytic signal, and IF are discussed in Refs. [15,33] and Section 6.5, with applications presented in Ref. [16].

1.5 SUMMARY AND DISCUSSION

Many practical signals are characterized by a variation of frequency content with time. A TFR of a signal shows this variation directly. The magnitude spectrum (magnitude of the FT) does not show this variation at all. In principle, the signal can be completely represented as a function of time or as a *complex* FT; but in practice, neither of these is as effective as a TFR in showing how the frequency content varies with time. The bandwidth-duration (*BT*) product of a signal is proportional to its information content. A large *BT* product is among the characteristics of an *asymptotic* signal.

The use of the *analytic associate* of a given real signal, rather than the signal itself, is essential for obtaining an unambiguous IF, and useful for reducing the required sampling rate. The analytic associate contains only non-negative frequencies. It is related to the HT, which removes the zero-frequency component, changes the signs of the negative-frequency components, and introduces a phase lag of $\pi/2$. The analytic associate of a *monocomponent* asymptotic signal is fully characterized by its instantaneous amplitude and instantaneous phase, from which we can determine its IF $f_i(t)$. IF estimation methods can be found in Refs. [15,34] and Chapter 10. An understanding of IF and its dual, SD, is necessary for the interpretation of TFDs. The concept of the analytic signal $z(t)$ is related to the definition of the canonical pairs $[a(t), \phi(t)]$ for a given real signal $s(t)$. Other applications of the HT appear in [35–37].

The next chapter introduces various formulations of TFDs and demonstrates the importance of using the analytic associate.

1.6 ADDITIONAL EXERCISES

This section provides additional exercises that cover important topics from this chapter and prepares the reader for the next chapters in the book. It is recommended to use the (t,f) toolbox described in Chapter 17 as an aid for doing these exercises. In a formal class setting, the instructor may wish to guide the student by adding intermediate questions with focus on one or two assessment criteria.

Exercise 1.6.1. Synthesize an LFM signal with starting frequency 10 Hz, end frequency 90 Hz, and duration 10 s.

1. Plot the signal in the time domain.
2. Plot the signal in the frequency domain.

Exercise 1.6.2. Compute the analytic signal associated with the one synthesized in the previous exercise.

1. Show the frequency domain representation of the analytic signal.
2. Write a code to estimate the IF and GD from the analytic signal and its frequency domain representation. *Note that before implementing a derivative operation in Matlab, you need to unwrap the phase of the signal.*

Exercise 1.6.3. Illustrate the (t, f) representation of the above analytic signal using the WVD.

1. Compute the WVD of this analytic signal and estimate the IF and GD from the location of peaks in the (t, f) domain; then relate IF to GD and SD.
2. Compare your results with the results obtained using the phase derivative algorithm using the TFSAP toolbox described in [Chapter 17](#).

Exercise 1.6.4. Repeat [Exercises 1.6.2](#) and [1.6.3](#) for the same signal corrupted by noise with SNR-0, 5, and 10 dB. Find which algorithm is more robust to noise, in terms of mean square error.

Exercise 1.6.5. Generate and plot the following signals:

1.

$$s(t) = \begin{cases} \cos(0.1\pi t) + \cos(2\pi(0.35t - 0.005t^2)), & 0 \leq t \leq 255 \\ 0, & \text{otherwise.} \end{cases}$$

2. The test signal given in Eq. (1.1.3), p. 34 with the same values given in the text below the equation.

Exercise 1.6.6. Using time-frequency representations, find out whether the signals given in [Exercise 1.6.5](#) are multi-component or mono-component. Justify your answer.

REFERENCES

- [1] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [2] B. Boashash, B. Escudie, “Wigner-Ville analysis of asymptotic signals and applications”, *Signal Process.* 8 (3) (1985) 315-327.
- [3] T.D. Rossing, *The Science of Sound*, second ed., Addison-Wesley, Reading, MA, 1990.
- [4] L.V. Beethoven (1770-1827), Symphony No. 3 in E Flat, Op. 55, *Eroica* (Full Score), IMSLP Petrucci Music Library, www.imslp.org, 2009.
- [5] L.V. Beethoven (1770-1827), Symphony No. 3 in E Flat, Op. 55, *Eroica* (audio recording), in *Beethoven, Symphonies 1-9*, North German Radio Symphony Orchestra (Günther Wand, cond.), RCA/BMG 74321-20277-2 (5-CD set); 1994, recorded 1985.
- [6] A.B. Carlson, *Communication Systems*, second ed., McGraw-Hill, Tokyo, 1975.

- [7] B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992.
- [8] L.J. Rankine, N.J. Stevenson, M. Mesbah, B. Boashash, “A nonstationary model of newborn EEG”, *IEEE Trans. Biomed. Eng.* 54 (1) (2007) 19-28.
- [9] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [10] B. Boashash, N.A. Khan, T. Ben-Jabeur, “Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review”, *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [11] B. Boashash, G. Azemi, N.A. Khan, “Principles of time-frequency feature extraction for change detection in non-stationary signals: applications to newborn EEG abnormality detection”, *Pattern Recog.* 48 (3) (2015) 616-627.
- [12] S. Stanković, I. Orović, N. Žarić, “An application of multidimensional time-frequency analysis as a base for the unified watermarking approach”, *IEEE Trans. Image Process.* 19 (3) (2010) 736-745.
- [13] B. Boashash, “Note on the use of the Wigner distribution for time-frequency signal analysis”, *IEEE Trans. Acoust. Speech Signal Process.* 36 (9) (1988) 1518-1521.
- [14] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, “The Wigner distribution—a tool for time-frequency signal analysis”, *Philips J. Res.* 35 (1980) 217-250 (Part 1), 276-300 (Part 2), and 372-389 (Part 3).
- [15] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 540-568.
- [16] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 519-568.
- [17] F.B. Hildebrand, Advanced Calculus for Engineers, Prentice-Hall, New York, 1949.
- [18] D. Gabor, “Theory of communication”, *J. IEE* 93(III) (26) (1946) 429-457.
- [19] D. Slepian, “On bandwidth”, *Proc. IEEE* 64 (3) (1976) 292-300.
- [20] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, B. Torrésani, “Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies”, *IEEE Trans. Inf. Theory* 38 (2) (1992) 644-673.
- [21] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, M. Daković, “Instantaneous frequency in time-frequency analysis: enhanced concepts and performance of estimation algorithms”, *Digital Signal Process.* 35 (0) (2014) 1-13, ISSN 1051-2004, <http://dx.doi.org/10.1016/j.dsp.2014.09.008>, URL <http://www.sciencedirect.com/science/article/pii/S1051200414002838>.
- [22] G. Jones, B. Boashash, “Generalized instantaneous parameters and window matching in the time-frequency plane”, *IEEE Trans. Signal Process.* 45 (5) (1997) 1264-1275.
- [23] P.J. Kootsookos, B.C. Lovell, B. Boashash, “A unified approach to the STFT, TFDs, and instantaneous frequency”, *IEEE Trans. Signal Process.* 40 (8) (1992) 1971-1982.
- [24] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.
- [25] A.W. Rihaczek, “Signal energy distribution in time and frequency”, *IEEE Trans. Inf. Theory* 14 (3) (1968) 369-374.
- [26] J. Imberger, B. Boashash, “Application of the Wigner-Ville distribution to temperature gradient microstructure: a new technique to study small-scale variations”, *J. Phys. Oceanogr.* 16 (12) (1986) 1997-2012.
- [27] G.S. Jones, B. Boashash, “Instantaneous frequency, instantaneous bandwidth and the analysis of multicomponent signals”, in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP’90), vol. 5, 1990, pp. 2467-2470.
- [28] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals”, *Proc. IEEE* 80 (4) (1992) 520-538.

- [29] B. Picinbono, “On instantaneous amplitude and phase of signals”, *IEEE Trans. Signal Process.* 45 (3) (1997) 552-560.
- [30] D. Vakman, “On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency”, *IEEE Trans. Signal Process.* 44 (4) (1996) 791-797.
- [31] A.H. Nutall, “On the quadrature approximation to the Hilbert transform of modulated signals”, *Proc. IEEE* 54 (1966) 1458-1459.
- [32] E. Bedrosian, “A product theorem for Hilbert transforms”, *Proc. IEEE* 51 (1963) 686-689.
- [33] A.P. Reilly, G.J. Frazer, B. Boashash, “Analytic signal generation—tips and traps”, *IEEE Trans. Signal Process.* 42 (11) (1994) 3241-3245.
- [34] L.B. White, B. Boashash, “On estimating the instantaneous frequency of a Gaussian random signal by the use of the Wigner-Ville distribution”, *IEEE Trans. Acoust. Speech Signal Process.* 36 (3) (1988) 417-420.
- [35] Z. Hussain, B. Boashash, “Hilbert transformer and time delay: statistical comparison in the presence of Gaussian noise”, *IEEE Trans. Signal Process.* 50 (3) (2002) 501-508, <http://dx.doi.org/10.1109/78.984723>.
- [36] Z.M. Hussain, B. Boashash, “The time-delay digital tanlock loop: performance analysis in additive Gaussian noise”, *J. Frankl. Inst.* 339 (1) (2002) 43-60, [http://dx.doi.org/10.1016/S0016-0032\(01\)00059-X](http://dx.doi.org/10.1016/S0016-0032(01)00059-X), URL <http://www.sciencedirect.com/science/article/pii/S001600320100059X>.
- [37] Z. Hussain, B. Boashash, M. Hassan-Ali, S. Al-Araji, “A time-delay digital tanlock loop”, *IEEE Trans. Signal Process.* 49 (8) (2001) 1808-1815, <http://dx.doi.org/10.1109/78.934151>.

This page intentionally left blank

HEURISTIC FORMULATION OF TIME-FREQUENCY DISTRIBUTIONS⁰

2

INTRODUCTION AND OVERVIEW

Having established the basic signal formulations in the previous chapter, we now consider the problem of formulating how to represent signals in a joint time-frequency (t,f) domain using a (t,f) distribution (TFD). The major finding of this study is that there is no unique solution to this problem; or in other words, there is a class of solutions.

The chapter then describes a variety of *ad hoc* approaches to providing a solution to the above problem resulting in six different methods and related formulations, namely the Wigner-Ville distribution (WVD) (Section 2.1), a time-varying spectrum or power spectral density (PSD) (Section 2.2), localized forms of the Fourier transform (FT) (Section 2.3), filter banks (Section 2.4), the spectrum gradient (Section 2.5), related energy densities (Section 2.6), and others. Finally, Section 2.7 shows how all these apparently different TFDs are related mathematically to the WVD, thus setting the scene for the more systematic treatment in the next chapter using a general and generic formula.

The behavior and performance of the various TFDs are illustrated using a linear-frequency modulated (LFM) asymptotic signal, which is regarded as a basic test signal for TFDs because it is the simplest example of a signal whose frequency content varies with time. It is monocomponent, and is asymptotic if its *BT* product is large. The minimum requirement for a useful TFD is that it clearly shows the IF law for such an asymptotic LFM signal, by following the peaks in the (t,f) domain, giving a reasonable concentration of energy about the IF law (which, for an asymptotic monocomponent FM signal, is in essence equivalent to the spectral delay [SD] law that describes the same linear segment in a (t,f) plot).

Other related (t,f) approaches are also discussed for completeness; these include the ambiguity function, cross-TFDs, time-scale, and wavelet methods.

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa).

2.1 METHOD 1: THE WIGNER-VILLE DISTRIBUTION

Let us consider a real signal $s(t)$ from which we construct the analytic signal $z(t)$ for reasons discussed in [Section 1.1](#); we seek to construct a TFD $\rho_z(t,f)$ that expresses how the signal energy is distributed versus time and frequency and how it is related to the standard energy, temporal and spectral characteristics of the signal. Let us choose the symbol ρ_z in the expectation that the TFD will represent an “energy density of z ” in the (t,f) plane, with the understanding that the constant- t cross-section of $\rho_z(t,f)$ will be some sort of “instantaneous spectrum” at time t .

For a monocomponent signal, it is reasonable to expect that the TFD should take the form of a knife-edge ridge whose crest is a graph of the IF law in the (t,f) plane. Mathematically the “knife edge” is idealized as a delta function w.r.t. frequency (e.g., [Fig. 1.1.7](#)).

2.1.1 KNIFE-EDGE IF INDICATION

Note that ρ_z is a function of frequency f and represents a kind of time-varying spectrum, we can define its inverse FT which we denote K_z . Let us call this function the *signal kernel*, which can be written as $K_z(t, \tau)$, such that:

$$\rho_z(t,f) = \mathcal{F}_{\tau \rightarrow f} \{K_z(t, \tau)\}. \quad (2.1.1)$$

Note that all TFDs in this chapter, unless otherwise stated, are defined from the complex analytic associate of the signal, not from the real or “raw” signal.

2.1.2 FORMULATION OF THE SIGNAL KERNEL

To find a suitable form for $K_z(t, \tau)$, for simplicity, let us first consider the case of the unit-amplitude monocomponent FM signal

$$z(t) = e^{j\phi(t)}, \quad (2.1.2)$$

whose instantaneous frequency is

$$f_i(t) = \frac{\phi'(t)}{2\pi}. \quad (2.1.3)$$

We would like the TFD of $z(t)$ at any given time to be a unit delta function at the instantaneous frequency, so that the “instantaneous spectrum” reduces to the ordinary FT in the constant-frequency case; that is, we want

$$\rho_z(t,f) = \delta(f - f_i(t)). \quad (2.1.4)$$

Substituting this into Eq. (2.1.1) and taking the inverse FT (IFT), we obtain

$$K_z(t, \tau) = \mathcal{F}_{\tau \leftarrow f}^{-1} \{\delta(f - f_i(t))\} = e^{j2\pi f_i(t)\tau} = e^{j\phi'(t)\tau}. \quad (2.1.5)$$

By definition $\phi'(t)$ in Eq. (2.1.5) can be expressed as:

$$\phi'(t) = \lim_{\tau \rightarrow 0} \frac{\phi(t + \frac{\tau}{2}) - \phi(t - \frac{\tau}{2})}{\tau}, \quad (2.1.6)$$

which can be estimated by removing the limit and using the central finite difference (CFD) approximation [1,2]

$$\phi'(t) \approx \frac{1}{\tau} \left[\phi\left(t + \frac{\tau}{2}\right) - \phi\left(t - \frac{\tau}{2}\right) \right]. \quad (2.1.7)$$

Substituting Eq. (2.1.7) into Eq. (2.1.5) and using Eq. (2.1.2) gives the signal kernel

$$\begin{aligned} K_z(t, \tau) &= e^{j\phi(t + \frac{\tau}{2})} e^{-j\phi(t - \frac{\tau}{2})} \\ &= z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right). \end{aligned} \quad (2.1.8)$$

Substituting Eq. (2.1.8) into Eq. (2.1.1), we obtain

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \left\{ z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) \right\} = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f \tau} d\tau. \quad (2.1.9)$$

The approximation in Eq. (2.1.7) is exact if $\phi(t)$ is quadratic (or linear) [3, p. 298], that is, if $\phi'(t)$ is linear (or constant). Thus, the WVD gives an “unbiased” estimate of the IF for a complex LFM signal. For nonlinear FM signals the central difference formula fails to give an exact estimate of the phase derivative. For such cases, higher-order approximations can be used as discussed in Sections 5.4 and 5.5.

2.1.3 THE WIGNER DISTRIBUTION

Applying the above expression to the case of a real signal yields:

$$\rho_s(t, f) = \int_{-\infty}^{\infty} s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f \tau} d\tau. \quad (2.1.10)$$

Equation (2.1.10) is given the symbol $\mathcal{W}_s(t, f)$ and is often called the *Wigner distribution* (WD) in honor of its first publisher,¹ who derived it in 1932 in a quantum-mechanical context [4]. Figure 2.1.1 shows the steps in the formulation of the WD.

The constant-frequency real signal

$$s(t) = \cos 2\pi f_c t \quad (2.1.11)$$

leads to the signal kernel

$$\begin{aligned} K_s(t, \tau) &= s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) = \cos 2\pi f_c \left(t + \frac{\tau}{2}\right) \cos 2\pi f_c \left(t - \frac{\tau}{2}\right) \\ &= \frac{1}{2} \cos 2\pi f_c \tau + \frac{1}{2} \cos 2\pi 2f_c t. \end{aligned} \quad (2.1.12)$$

¹E. P. Wigner (1902-1995) was born in Hungary, studied chemical engineering in Germany and eventually settled in the United States, where he specialized in mathematical physics. He was a joint winner of the 1963 Nobel Prize for Physics for his many contributions to particle physics, including his law of conservation of parity and his work on the strong nuclear force.

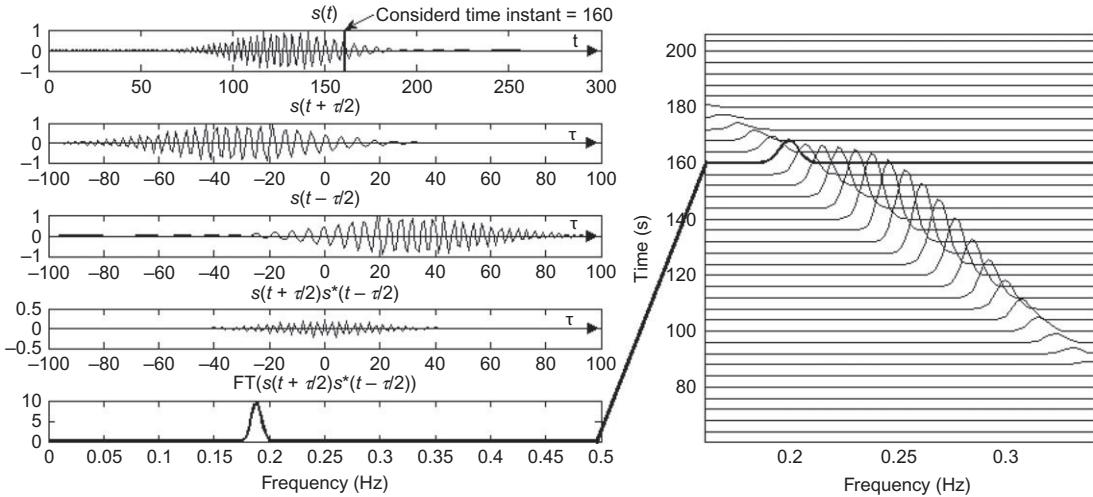


FIGURE 2.1.1

WD calculation of a linear frequency modulated (LFM) signal at time instant $t = 160$ s (length 256 samples: starting frequency = 0.4 Hz, finishing frequency 0.1 Hz (sampling rate 1 Hz), Gaussian window length = 256 and $\alpha = 3$).

Taking the FT w.r.t. τ gives the WD

$$\mathcal{W}_s(t, f) = \frac{1}{4} \delta(f - f_c) + \frac{1}{4} \delta(f + f_c) + \frac{1}{2} [\cos 2\pi 2f_c t] \delta(f). \quad (2.1.13)$$

The terms in $\delta(f \mp f_c)$ are naturally expected and arise because $s(t)$ may be expressed as a sum of complex sinusoids at frequencies $\pm f_c$. However, the term in $\delta(f)$, called a *cross-term*, is an artifact arising because the *nonlinearity* of the WD causes interaction between the positive-frequency terms and the negative-frequency terms. (Cross-terms are discussed further in [Section 3.1.2](#).)

By a similar argument, we find that the nonwindowed LFM signal

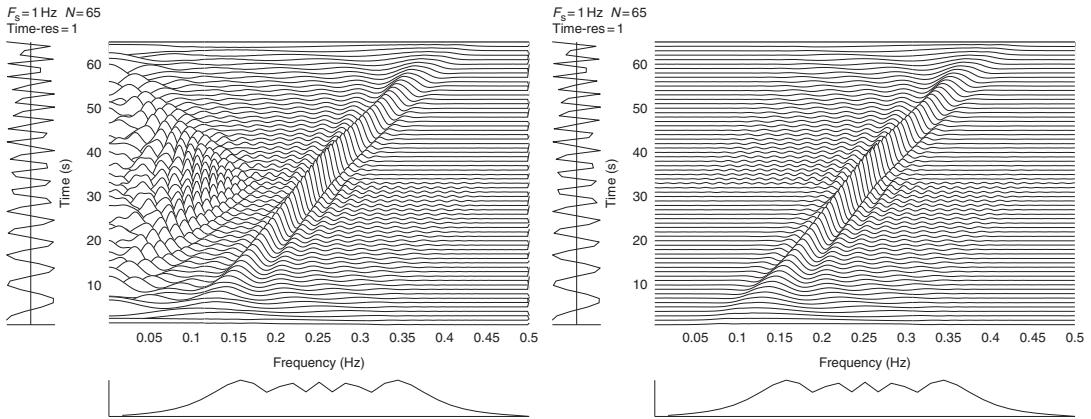
$$s(t) = A \cos \left(2\pi \left[f_0 t + \frac{\alpha}{2} t^2 \right] \right) \quad (2.1.14)$$

leads to the signal kernel

$$K_s(t, \tau) = \frac{1}{2} A^2 \cos 2\pi f_i(t) \tau + \frac{1}{2} A^2 \cos 2\pi \left[\frac{\alpha \tau^2}{4} + 2f_0 t + \alpha t^2 \right], \quad (2.1.15)$$

where $f_i(t) = f_0 + \alpha t$. Taking the FT w.r.t. τ gives the WD

$$\mathcal{W}_s(t, f) = \frac{1}{4} A^2 \delta(f - f_i(t)) + \frac{1}{4} A^2 \delta(f + f_i(t)) + \frac{1}{2} A^2 \underset{\tau \rightarrow f}{\mathcal{F}} \left\{ \cos 2\pi \left[\frac{\alpha \tau^2}{4} + 2f_0 t + \alpha t^2 \right] \right\}. \quad (2.1.16)$$

**FIGURE 2.1.2**

Wigner distribution (left) and Wigner-Ville distribution (right) of a real finite duration LFM signal (Eq. (1.1.4)) with duration 65 samples, starting frequency 0.1 Hz and finishing frequency 0.4 Hz (sampling rate 1 Hz). Note that the low-frequency artifacts in the Wigner distribution, caused by interaction between positive and negative frequencies, have disappeared in the right-hand plot.

The terms in $\delta(f \mp f_i(t))$ are naturally expected, while the last term in the signal kernel gives rise to a continuum of artifacts in the WD (see Fig. 2.1.2 (left)). These artifacts, which greatly diminish the usefulness of the WD for real signals, are removed by using the WVD, that is, modifying the WD using the analytic signal as originally intended at the beginning of Section 2.1.1, as described in the next section.

2.1.4 THE WIGNER-VILLE DISTRIBUTION (WVD)

Definition 2.1.1. The WVD of a signal $s(t)$, denoted by $W_z(t,f)$, is defined as the WD of its analytic associate, that is

$$\rho_z(t,f) \equiv W_z(t,f) = \mathcal{F}_{\tau \rightarrow f} \left\{ z \left(t + \frac{\tau}{2} \right) z^* \left(t - \frac{\tau}{2} \right) \right\}, \quad (2.1.17)$$

where $z(t)$ is complex and defined as the analytic associate of $s(t)$.

The name “Wigner-Ville distribution,” as opposed to “Wigner distribution,” emphasizes the use of the analytic signal [5] and recognizes the contribution of Ville [6], who derived the TFD in a signal-processing context in 1948. Note that a signal can have a time-dependent frequency content,

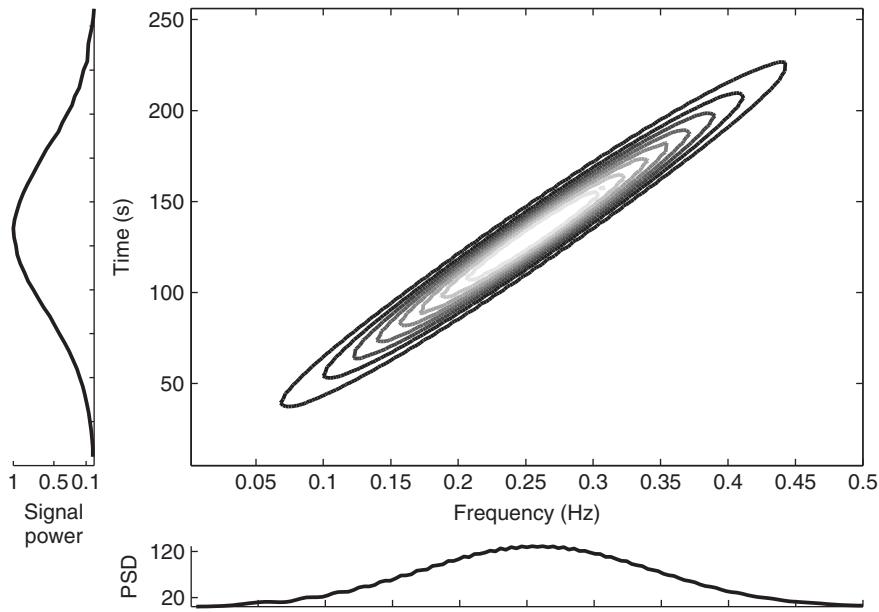
**FIGURE 2.1.3**

Illustration of the marginal properties of typical TFD: LFM signal (length 256 samples: starting frequency = 0.01 Hz, finishing frequency = 0.5 Hz [sampling rate 1 Hz], Gaussian window length = 256 and $\alpha = 6$); to reproduce the plots, see toolbox in [Chapter 17](#).

Ville sought an “instantaneous spectrum” having the attributes of an energy density (property 1 in [Section 1.1.5](#)) and satisfying the so-called *marginal conditions*² that is, so that:

- the integral of the TFD $\rho_z(t, f)$ w.r.t. f is the instantaneous power $|z(t)|^2$ and
- the integral of the TFD $\rho_z(t, f)$ w.r.t. t is the energy spectrum $|Z(f)|^2$.

The two marginal conditions are shown in [Fig. 2.1.3](#). By analogy with the conditional moments of a p.d.f., and using known relationships between the moments of a p.d.f. and its characteristic function, we can show that the TFD (now known as WVD) had the desired properties [6] (see [Fig. 2.1.3](#)). Using Eq. (2.1.8), Eq. (2.1.1) is immediately verified when $\rho_z(t, f) = W_z(t, f)$, that is,

$$W_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{K_z(t, \tau)\}. \quad (2.1.18)$$

The signal kernel $K_z(t, \tau)$ is also called the *instantaneous autocorrelation function* (IAF) of $z(t)$.

²The name “marginal” can be explained with reference to the discrete-time, discrete-frequency case: if the TFD were written as a two-dimensional array of discrete “energies,” each “energy” value corresponding to a discrete time (vertical axis) and a discrete frequency (horizontal axis), then the sum over time for each frequency could be written in the horizontal “margin” of the array, and the sum over frequency for each time could be written in the vertical “margin.”

2.1.4.1 The WVD of an LFM signal

Equation (2.1.2) describes a constant-amplitude, infinite-duration signal. We can allow for nonconstant amplitude and finite duration using the form

$$z(t) = a(t) e^{j\phi(t)}, \quad (2.1.19)$$

where $a(t)$ is real. For this signal we find

$$K_z(t, \tau) = K_a(t, \tau) e^{j\psi(t, \tau)}, \quad (2.1.20)$$

where

$$K_a(t, \tau) = a\left(t + \frac{\tau}{2}\right) a\left(t - \frac{\tau}{2}\right), \quad (2.1.21)$$

$$\psi(t, \tau) = \phi\left(t + \frac{\tau}{2}\right) - \phi\left(t - \frac{\tau}{2}\right). \quad (2.1.22)$$

If $\phi(t)$ is quadratic (i.e., if $f_i(t)$ is linear), then the CFD approximation given in Eq. (2.1.7) is exact and gives $\psi(t, \tau) = \phi'(t)\tau = 2\pi f_i(t)\tau$, so that

$$K_z(t, \tau) = K_a(t, \tau) e^{j2\pi f_i(t)\tau}. \quad (2.1.23)$$

So $K_z(t, \tau)$, considered as a function of τ , has a constant frequency equal to $f_i(t)$; the reduction to constant frequency is sometimes called *dechirping*. If we let $\mathcal{W}_a(t, f) = \underset{\tau \rightarrow f}{\mathcal{F}} \{K_a(t, \tau)\}$, so that $\mathcal{W}_a(t, f)$ is the WD of $a(t)$, and take the FT of Eq. (2.1.23) w.r.t. τ , we obtain

$$W_z(t, f) = \mathcal{W}_a(t, f) *_{\tau} \delta(f - f_i(t)) \quad (2.1.24)$$

$$= \mathcal{W}_a(t, f - f_i(t)). \quad (2.1.25)$$

Now $K_a(t, \tau)$ is real and even in τ . Hence, $\mathcal{W}_a(t, f)$ is real and even in f , so that $W_z(t, f)$ is real and symmetrical about $f = f_i(t)$. Further, $\mathcal{W}_a(t, f)$ is maximum at $\mathcal{W}_a(t, 0)$, which results in $W_z(t, f)$ having a maximum along $f_i(t)$.

Thus, even for a *finite*-duration LFM signal, such as the one analyzed in Fig. 2.1.2, we expect the WVD to give a clear indication of the IF law. Indeed, the TFD shown in Fig. 2.1.2 is the WVD.

Note: research in the early 1980s focused on the WVD due to its good frequency localization for chirp signals. For example, in a seismic context, the WVD was shown to be a very effective tool to represent Vibroseis chirp signals emitted in seismic processing [7]. When the signal emitted is a pure LFM, the WVD exhibits a sharp peak along the IF representing the FM law. This property was lost if the signal was contaminated by harmonic coupling effects and other distortions [8]. Because the LFM is the simplest way of making f vary as a function of t , the chirp signal is the most used nonstationary signal as a model or approximation; and the WVD, being an optimal (t, f) representation for such a signal, then became then the most important TFD to be studied in detail.

2.1.4.2 The WVD in terms of the spectrum

The variables t , f , and τ are called time, frequency, and lag, respectively. We now introduce the variable v , which represents frequency *shift* just as τ represents time shift; accordingly, v will be called *Doppler*.³ Let us now define

$$k_z(v, f) = \mathcal{F}_{t \rightarrow v} \{W_z(t, f)\}. \quad (2.1.26)$$

Writing out the definitions of the FT and the WVD and taking invariant factors inside the integral signs, we obtain

$$k_z(v, f) = \iint z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi(f\tau + vt)} dt d\tau, \quad (2.1.27)$$

where the integrals are from $-\infty$ to ∞ . If we write

$$u = t + \frac{\tau}{2}; \quad v = t - \frac{\tau}{2} \quad (2.1.28)$$

and solve for t and τ , obtaining

$$t = \frac{1}{2}(u + v); \quad \tau = u - v, \quad (2.1.29)$$

then the use of the Jacobian⁴ yields

$$dt d\tau = du dv. \quad (2.1.30)$$

With these substitutions, we find that Eq. (2.1.27) can be factored into

$$\begin{aligned} k_z(v, f) &= \int_{-\infty}^{\infty} z(u) e^{-j2\pi[f+v/2]u} du \cdot \int_{-\infty}^{\infty} z^*(v) e^{j2\pi[f-v/2]v} dv \\ &= Z\left(f + \frac{v}{2}\right) Z^*\left(f - \frac{v}{2}\right), \end{aligned} \quad (2.1.31)$$

where $Z(f) = \mathcal{F}\{z(t)\}$. Noting that $k_z(v, f)$ has a similar form to $K_z(t, \tau)$ in the Doppler-frequency domain; we denote $k_z(v, f)$ as the *spectral correlation function (SCF)*.⁵

Substituting Eq. (2.1.31) into Eq. (2.1.26) and taking the IFT yields an expression for the WVD in terms of $Z(f)$:

$$W_z(t, f) = \int_{-\infty}^{\infty} Z\left(f + \frac{v}{2}\right) Z^*\left(f - \frac{v}{2}\right) e^{j2\pi vt} dv. \quad (2.1.32)$$

The above expression of the WVD may be particularly useful in the case of narrow-band signals as it may reduce the computation costs by taking into account the shorter lengths of the signal spectrum

³The well-known “Doppler effect” is actually a frequency *scaling*. However, when the effect is used to measure velocity, the scaling factor is usually close to unity and the bandwidth of interest is usually narrow. Under these conditions, the frequency scaling is well approximated by a frequency *shift* proportional to the velocity.

⁴More precisely, the absolute value of the determinant (equal to 1 here) of the Jacobian matrix associated to this change variable.

⁵Note that normally SAF should be used, but we use in this case SCF to avoid confusion with the SAF abbreviation used for the symmetrical ambiguity function.

$Z(f)$. The reason is that for narrowband signals, we can reduce the limit of integration from $(-\infty, \infty)$ to $(-B/2, B/2)$, where B is the signal bandwidth. By inverting Eq. (2.1.26) and using Eq. (2.1.27), one can recognize that the WVD is the 2D FT, w.r.t. the lag and Doppler variables τ and ν , of the “ambiguity function,” commonly used by specialists in radar and sonar. This important property is discussed in detail later, for example, in Section 5.1 and Chapter 14.

2.1.4.3 Effects of time- and frequency-limiting

As discussed in Sections 1.2.3.1 and 1.2.4.2, a practical signal $z(t)$ is often expected to be both time-limited and band-limited, despite theoretical constraints. Let us assume that $z(t)$ is windowed in both time and frequency. For the time windowing around $t = t_0$ with window $w(t)$, we can replace $z(t)$ by

$$z_w(t) = z(t) w(t - t_0). \quad (2.1.33)$$

The WVD of $z_w(t)$ is then:

$$W_{z_w}(t, f) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) w\left(t - t_0 + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) w^*\left(t - t_0 - \frac{\tau}{2}\right) e^{-j2\pi f \tau} d\tau. \quad (2.1.34)$$

Estimating the above at $t = t_0$ results in [9,10]:

$$W_{z_w}(t_0, f) = \int_{-\infty}^{\infty} g_2(\tau) K_z(t_0, \tau) e^{-j2\pi f \tau} d\tau, \quad (2.1.35)$$

where

$$g_2(\tau) = w\left(\frac{\tau}{2}\right) w^*\left(-\frac{\tau}{2}\right). \quad (2.1.36)$$

If g_2 has a finite duration, that is, $g_2(\tau) = 0$ for $|\tau| > T/2$, then the limits of integration in Eq. (2.1.35) reduce to $\pm T/2$. Note that $W_{z_w}(t, f)$ differs from the WVD of $z(t)$ in that the IAF is multiplied by $g_2(\tau)$ before being Fourier-transformed w.r.t. τ . This $g_2(\tau)$ is thus the effective lag window corresponding to the sliding time window $w(t - t_0)$. For the frequency windowing, we can replace $Z(f)$ by

$$Z_H(f) = Z(f) H(f - f_0). \quad (2.1.37)$$

Using Eq. (2.1.32), we can take the WVD corresponding to $Z_H(f)$ and evaluate the WVD at $f = f_0$, obtaining in a similar manner to Eq. (2.1.35):

$$W_{Z_H}(t, f_0) = \int_{-\infty}^{\infty} G_1(\nu) Z\left(f_0 + \frac{\nu}{2}\right) Z^*\left(f_0 - \frac{\nu}{2}\right) e^{j2\pi \nu t} d\nu, \quad (2.1.38)$$

where

$$G_1(\nu) = H\left(\frac{\nu}{2}\right) H^*\left(-\frac{\nu}{2}\right). \quad (2.1.39)$$

If $G_1(\nu) = 0$ for $|\nu| > B/2$, then the limits of integration in Eq. (2.1.38) may be changed to $\pm B/2$. Note that $W_{Z_H}(t, f)$ differs from the WVD of $z(t)$ in that the SCF is multiplied by $G_1(\nu)$ before being inverse-Fourier-transformed w.r.t. ν . This $G_1(\nu)$ is the effective Doppler window corresponding to the sliding frequency window $H(f - f_0)$.

Note that the effects of time windowing and frequency windowing may be combined; this then results in the (ν, τ) domain as a multiplication by the factor

$$g(\nu, \tau) = G_1(\nu) g_2(\tau). \quad (2.1.40)$$

If we consider time-limiting alone, the resulting TFD is given by Eq. (2.1.35) and is called the *windowed WVD* [10–12]. If we consider band-limiting alone, the resulting TFD is given by Eq. (2.1.38). We shall call this the *filtered WVD*. The window shapes of G_1 and g_2 should be selected to ensure the properties 1 to 3 in Section 1.1.5 are verified. We can therefore conclude that a general TFD may be obtained by constructing a filtered and windowed WVD, or equivalently by a (t, f) filtered WVD, with filters designed in the (ν, τ) domain, known as the “ambiguity” domain.

2.2 METHOD 2: TIME-VARYING POWER SPECTRAL DENSITY

2.2.1 SPECTRA OF NONSTATIONARY RANDOM PROCESSES

In this approach, let us assume $z(t)$ is a complex random signal, its symmetrical *autocorrelation* function is

$$\mathcal{R}_z(t, \tau) = \mathcal{E} \left\{ z \left(t + \frac{\tau}{2} \right) z^* \left(t - \frac{\tau}{2} \right) \right\}, \quad (2.2.1)$$

where $\mathcal{E}\{\cdot\}$ denotes the expected value. In this approach, let us assume $z(t)$ is wide-sense stationary, then $\mathcal{R}_z(t, \tau)$ is independent of t , and the limit

$$\mathcal{S}_z(f) \triangleq \lim_{T \rightarrow \infty} \mathcal{E} \left\{ \frac{1}{T} \left| \mathcal{F}_{t \rightarrow f} \{z(t) \text{rect}(t/T)\} \right|^2 \right\} \quad (2.2.2)$$

is called the PSD of the random process $z(t)$, and describes the distribution of signal power over the frequencies, where $\text{rect}(t/T)$ is the rectangular function defined in Table I.1, p. 12. The PSD is related to the autocorrelation function by the *Wiener-Khintchine theorem*, which states that, for $R_z(t, \tau)$ independent of t , that is, $\mathcal{R}_z(t, \tau) = \mathcal{R}_z(\tau)$, we have

$$\mathcal{S}_z(f) = \mathcal{F}_{\tau \rightarrow f} \{ \mathcal{R}_z(t, \tau) \} = \mathcal{F}_{\tau \rightarrow f} \{ \mathcal{R}_z(\tau) \}. \quad (2.2.3)$$

If $z(t)$ is not wide-sense stationary, the right-hand side of this equation depends on t , so that the left-hand side also depends on t , suggesting the generalization

$$\mathcal{S}_z(t, f) \triangleq \mathcal{F}_{\tau \rightarrow f} \{ \mathcal{R}_z(t, \tau) \}. \quad (2.2.4)$$

This $\mathcal{S}(t, f)$ may be called the time-varying spectrum or *evolutive spectrum* (ES), and is interpreted as the “time-dependent PSD” of the nonstationary signal. Using Fubini’s theorem, the right-hand side of Eq. (2.2.4) may be expanded as

$$\mathcal{F}_{\tau \rightarrow f} \{ \mathcal{R}_z(t, \tau) \} = \int_{-\infty}^{\infty} \mathcal{E} \left\{ z \left(t + \frac{\tau}{2} \right) z^* \left(t - \frac{\tau}{2} \right) \right\} e^{-j2\pi f \tau} d\tau \quad (2.2.5)$$

$$= \mathcal{E} \left\{ \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \right\} \quad (2.2.6)$$

$$\mathcal{S}_z(t,f) = \mathcal{E} \{ W_z(t,f) \}. \quad (2.2.7)$$

Equation (2.2.7) shows that the expected value of the WVD is the FT of the time-dependent autocorrelation function [13]; that is, *the ES is the expected value of the WVD*. For this reason, the ES is also called the *Wigner-Ville spectrum*.

If $z(t)$ is deterministic, Eq. (2.2.7) reduces to

$$\mathcal{S}_z(t,f) = W_z(t,f). \quad (2.2.8)$$

2.2.2 ESTIMATING THE WIGNER-VILLE SPECTRUM

If the signal $z(t)$ is random, then the (t,f) transformation also results in a random TFD, here $W_z(t,f)$, which can be characterized by averages over all possible realizations of the random process. Equation (2.2.7) refers to an ensemble average of the random process $W_z(t,f)$. If we have only one realization of this process, we may be able to assume that $W_z(\theta,f)$ is locally ergodic on the interval $t - \Delta/2 < \theta < t + \Delta/2$, where Δ is positive and independent of t . An estimate of $W_z(t,f)$ for the local values of $z(t)$, yields

$$\hat{W}_z(t,f) = \mathcal{F}_{\tau \rightarrow f} \{ g_2(\tau) K_z(t,\tau) \} \quad (2.2.9)$$

$$= G_2(f) *_{\! f} W_z(t,f), \quad (2.2.10)$$

where $g_2(\tau)$, the effective analysis window, is real, even and time-limited, and $G_2(f) = \mathcal{F} \{ g_2(\tau) \}$.¹ The notation $*$ stands for convolution in frequency.

Then, to obtain an estimate of $\mathcal{S}_z(t,f)$, we replace the ensemble average over all realizations of $W_z(t,f)$ by a time average of $\hat{W}_z(t,f)$ over the interval Δ . We can calculate this average using a real, even weighting function $g_1(t)$, such that

$$g_1(t) \begin{cases} > 0 & \text{if } |t| \leq \Delta/2, \\ = 0 & \text{otherwise,} \end{cases} \quad (2.2.11)$$

$$\int_{-\infty}^{\infty} g_1(t) dt = 1. \quad (2.2.12)$$

The resulting estimate of $\mathcal{S}_z(t,f)$ is

$$\hat{\mathcal{S}}_z(t,f) = \int_{-\infty}^{\infty} g_1(\theta - t) \hat{W}_z(\theta,f) d\theta. \quad (2.2.13)$$

¹This lag window $g_2(\tau)$ may be used to reduce the computations required for estimating the WVD. It has the effect of causing smoothing along the frequency axis, whether desired or undesired.

Using the assumed evenness of g_1 and substituting from Eq. (2.2.10), we have

$$\hat{S}_z(t,f) = g_1(t) * \hat{W}_z(t,f) = g_1(t) * \frac{G_1(t)}{t} * G_2(f) * \frac{W_z(f)}{f}, \quad (2.2.14)$$

$$\hat{S}_z(t,f) = \gamma(t,f) \underset{(t,f)}{**} W_z(t,f), \quad (2.2.15)$$

where

$$\gamma(t,f) = g_1(t) G_2(f) \quad (2.2.16)$$

and the double asterisk denotes *double convolution*.

Equation (2.2.15) defines a class of estimates for time-varying spectra obtained by a double convolution of the WVD with a 2D filter. This will be used in the next chapter to define quadratic TFDs with specific properties (see also [14]).

2.3 METHOD 3: WINDOWED FT AND SPECTROGRAM

2.3.1 STFT, INVERSE STFT, AND SPECTROGRAM

Consider a signal $s(\tau)$ and a real, even window $w(\tau)$, whose FTs are $S(f)$ and $W(f)$, respectively. To obtain a localized spectrum of $s(\tau)$ at time $\tau = t$, multiply the signal by the window $w(\tau)$ centered at time $\tau = t$, obtaining

$$s_w(t, \tau) = s(\tau) w(\tau - t), \quad (2.3.1)$$

and then take the FT w.r.t. τ , obtaining

$$F_s^w(t,f) = \mathcal{F}_{\tau \rightarrow f} \{s(\tau) w(\tau - t)\}. \quad (2.3.2)$$

$F_s^w(t,f)$ is called the *short-time Fourier transform (STFT)*.

2.3.1.1 Inverse STFT

The window $w(t)$ used in the definition of the STFT is selected to verify the relation:

$$\int_{-\infty}^{\infty} w(\tau) d\tau = 1, \quad (2.3.3)$$

where τ represents here the running time for ease of notation (while t is the actual time). It follows from the previous equation that:

$$s(t) = s(t) \int_{-\infty}^{\infty} w(t - \tau) d\tau = \int_{-\infty}^{\infty} s(t) w(t - \tau) d\tau. \quad (2.3.4)$$

The FT of $s(t)$ is:

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-2j\pi ft} dt. \quad (2.3.5)$$

Substituting $s(t)$ from Eq. (2.3.4) in $S(f)$ yields:

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} s(t)w(t-\tau) d\tau \right] e^{-2j\pi ft} dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t)w(t-\tau) e^{-2j\pi ft} d\tau dt. \end{aligned} \quad (2.3.6)$$

Now, changing the order of integration yields:

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t)w(t-\tau) e^{-2j\pi ft} dt d\tau \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [s(t)w(t-\tau) e^{-2j\pi ft}] dt d\tau \\ &= \int_{-\infty}^{\infty} F_s^w(\tau, f) d\tau. \end{aligned} \quad (2.3.7)$$

Applying the inverse FT on the previous equation, $s(t)$ can be recovered from $F_s^w(t, f)$ as:

$$s(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_s^w(\tau, f) e^{2j\pi ft} d\tau df, \quad (2.3.8)$$

which defines the inverse STFT. This formulation is the basis for several application areas such as (t, f) signal filtering and (t, f) signal synthesis [Section 11.2] [15].

2.3.1.2 The spectrogram

The squared magnitude of the STFT, denoted by $S_s^w(t, f)$, is called the *spectrogram* and is expressed as follows:

$$S_s^w(t, f) = |F_s^w(t, f)|^2 \quad (2.3.9)$$

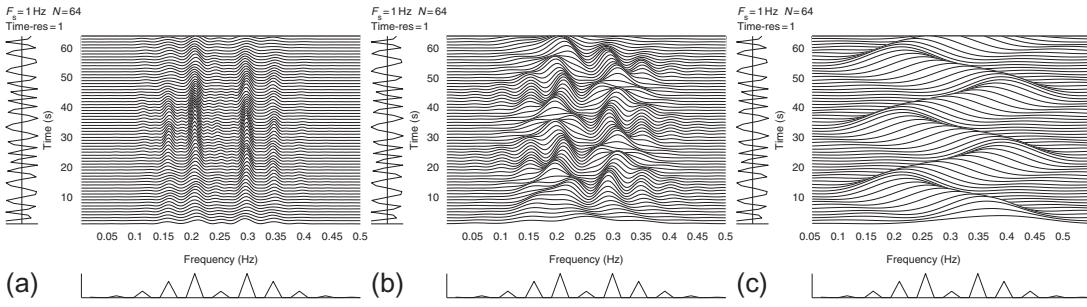
$$= \left| \mathcal{F}_{\tau \rightarrow f} \{ s(\tau) w(\tau - t) \} \right|^2 \quad (2.3.10)$$

$$= \left| \int_{-\infty}^{\infty} s(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right|^2, \quad (2.3.11)$$

where, in the notation $S_s^w(t, f)$, the upper-case S stands for “spectrogram,” while the subscript stands for the signal and the superscript for the filtering window.

The observation window $w(\tau)$ allows localization of the spectrum in time, but also smears the spectrum in frequency¹ in accordance with the “uncertainty relationship” described in Section 4.4 [16], leading to a trade-off between time resolution and frequency resolution. The problem is illustrated in Fig. 2.3.1, which shows the spectrogram of a sinusoidal FM signal for a rectangular window of three different lengths. If the window is long compared with the modulating signal, the frequency resolution

¹Smearing is caused by the convolution operation. If an image (a function of two coordinates) is convolved with a confusion pattern (another function of the same two coordinates), the result is a blurred image. If the confusion pattern is a line, we tend to describe the blurring as a “smearing.”

**FIGURE 2.3.1**

Spectrogram of a sinusoidal FM signal (Eq. (1.1.3)) with 65 samples (sampling rate 1 Hz), $A = 1$, $f_c = 1/4$, $f_m = 3/64$, $f_d = 6/64$, rectangular window of length Δ samples: (a) $\Delta = 63$; (b) $\Delta = 23$; and (c) $\Delta = 7$; to reproduce the plots, see toolbox in Chapter 17.

is sufficient to show the sideband tones (the “multicomponent aspect” [17] of the signal), but the time resolution is insufficient to show the FM law (the “monocomponent aspect”). If the window is short compared with the modulating signal, the time resolution is sufficient to show the FM law but the frequency resolution is insufficient to show the sideband tones.

The spectrogram is nonlinear; but the nonlinearity is introduced only in the final step (taking the squared magnitude) and therefore does not lead to undesirable artifacts (i.e., cross-terms) present in the WVD. This freedom from artifacts, together with simplicity, robustness, and ease of interpretation, has made the spectrogram a popular tool in many applications including speech analysis (resolution of speech into phonemes and formants) since its publication in 1946 [18]. However, its sensitivity to the window size, shown in Fig. 2.3.1, is a major disadvantage, which requires optimization of the window size via adaptation or the use of other TFDs for improved precision and resolution.

2.3.2 OPTIMAL WINDOW LENGTH OF THE SPECTROGRAM

As mentioned above, the spectrogram involves a compromise between time resolution and frequency resolution: a longer window provides less localization in time and more discrimination in frequency. The purpose of the window is to obtain a time-slice of the signal during which the spectral characteristics are nearly constant. If the window is too long, it fails to capture the most rapid variations of spectral content. If it is too short, it smears the TFD in the frequency dimension without a commensurate improvement in detail in the time dimension. The more rapidly the spectral content changes, the shorter the window must be.

Hence, for a monocomponent signal of constant amplitude, the optimal window duration is inversely related to the rate of change of the IF. If the window is rectangular and has total duration Δ , it can be shown [19,20] that the value

$$\Delta = \sqrt{2} \left| \frac{df_i(t)}{dt} \right|^{-1/2} \quad (2.3.12)$$

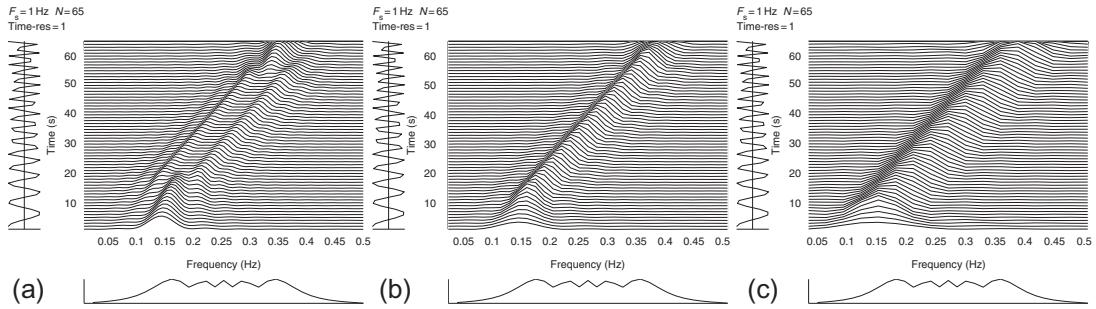


FIGURE 2.3.2

Spectrogram of an LFM signal (Eq. (1.1.4)) with duration 65, starting frequency 0.1, and finishing frequency 0.4, for a rectangular window of length Δ : (a) $\Delta = 33$; (b) $\Delta = 21$; and (c) $\Delta = 11$. The optimal window length according to Eq. (2.3.12) is 20.7.

is optimal in the sense that it minimizes the half-height width of the resulting ridge in the (t, f) plane. This optimal Δ is proportional to the relaxation time T_r ; compare it with Eq. (1.3.44). For an LFM signal, the optimal window duration simplifies to $\Delta = \sqrt{2T/B}$, where T is the signal duration and B is the signal bandwidth [9]. Figure 2.3.2 shows the spectrogram of a chirp signal (the same signal as in Fig. 1.1.7) for a rectangular window of three different lengths, one of which ($\Delta = 21$) is optimal according to Eq. (2.3.12).

Even for the optimal window length, the spectrogram is not a delta function describing the IF law. The use of this optimal window is inconvenient because it requires knowledge of the IF, and this knowledge might be obtainable only by a prior exploratory (t, f) analysis and an adaptive procedure. Moreover, if the IF law is nonlinear, the optimal window duration varies with time. In the case of the sinusoid FM signal of Fig. 2.3.1, the optimal window length is time-varying if the signal is considered as a modulated carrier. Although it is possible to vary the spectrogram window length with time and even with frequency, such procedures have a cost in computational efficiency (see Section 5.3) [21]. Section 10.1 described and iterative algorithm for matching the spectrogram window to the estimated IF, so that the spectrogram of a monocomponent signal is concentrated along the IF law [22]. We will also see in Chapter 3 that the spectrogram can also be processed by a 2D filtering to obtain a TFD with concentration close to the WVD (see Section 3.3.4 for more details).

2.3.3 STFT VS GABOR TRANSFORM

In the context of telecommunication applications, the requirements for efficient signal transmission [16] indicate that the (t, f) plane can be divided into an array of rectangles using a bank of filters, each of which is switched on for a single interval of time and passes a single band of frequencies. Each rectangle was called a *logon*, and its dimensions were called the *decay time* and the *tuning width* whose dimensions must satisfy the uncertainty relation

$$\Delta t \Delta f \geq \frac{1}{4\pi}, \quad (2.3.13)$$

where Δt and Δf are the effective duration and bandwidth of the logon [16]. This relationship is “at the root of the fundamental principle of communication” [16], in that it puts a lower limit on the spread of a signal in time and frequency. For the minimum elemental area, which is obtained in the case of a complex Gaussian signal, Eq. (2.2.4) becomes an equality as in Example 1.2.5.

In Gabor’s representation, each logon is assigned a complex coefficient $c_{n,k}$, where n is the time index and k the frequency index. The signal $s(t)$ is expanded in a doubly infinite series, where the sum is over all integers n and k , as [23]:

$$s(t) = \sum_{n,k} c_{n,k} \psi_{n,k}(t), \quad (2.3.14)$$

where $\psi_{n,k}(t)$ is a function centered about time $n\Delta t$ and frequency $k\Delta f$. To find the coefficients $c_{n,k}$, let $h_{n,k}(t)$ and $\psi_{n,k}(t)$ be related by

$$\int_{-\infty}^{\infty} h_{m,l}^*(t) \psi_{n,k}(t) dt = \begin{cases} 1 & \text{if } m = n \text{ and } l = k; \\ 0 & \text{otherwise.} \end{cases} \quad (2.3.15)$$

In other words, let $h_{m,l}(t)$ be orthogonal to every $\psi_{n,k}(t)$ except $\psi_{m,l}(t)$, or, equivalently, let $\psi_{m,l}(t)$ be orthogonal to every $h_{n,k}(t)$ except $h_{m,l}(t)$; functions related in this way are called *dual functions*. Multiplying Eq. (2.3.14) by $h_{n,k}^*(t)$ and integrating w.r.t. time, we obtain an expression for $c_{n,k}$:

$$c_{n,k} = \int_{-\infty}^{\infty} s(\tau) h_{n,k}^*(\tau) d\tau, \quad (2.3.16)$$

where τ is a running variable of dimension t . If we choose

$$h_{n,k}(\tau) = w(\tau - n\Delta t) e^{j2\pi k\Delta f \tau}, \quad (2.3.17)$$

where w denotes a real Gaussian function, then Eq. (2.3.16) becomes

$$c_{n,k} = \int_{-\infty}^{\infty} s(\tau) w(\tau - n\Delta t) e^{-j2\pi k\Delta f \tau} d\tau \quad (2.3.18)$$

$$= \mathcal{F}_{\tau \rightarrow k\Delta f} \{s(\tau) w(\tau - n\Delta t)\}. \quad (2.3.19)$$

Expression (2.3.19) is known as the *Gabor transform*. When $\psi_{n,k}(t)$ is chosen as the dual function of $h_{n,k}(t)$, Eq. (2.3.14) gives the *inverse Gabor transform*. Equation (2.3.19) has the same form as Eq. (2.3.2) except that t and f are discretized, which is why the Gabor transform can be described as a sampled STFT with Gaussian window. Suitable selections of ψ and h need to be made for the TFD to verify the properties 1 to 3 in Section 1.1.5, at least approximately.

Lerner [24] extended Gabor’s work by allowing the elements of the (t,f) plane to be nonrectangular. Helstrom [25] generalized the expansion by replacing the discrete elementary cell weighting with a continuous function $\xi(\tau, t, f)$ (see Section 6.3 for details). Wavelet theory [26] was later developed as a further extension of Gabor’s work, but choosing the partitions of the (t,f) plane so as to give constant- Q filtering [27,28]. The relationships between the spectrogram, Gabor transform, time-scale representations, and other TFDs are further discussed in Section 4.1.

2.4 METHOD 4: FILTERED FUNCTION OF TIME

2.4.1 FILTER BANKS AND THE SONOGRAPH

Whereas the spectrogram is conceived as a function of frequency with dependence on the timing of a window, the sonograph is conceived as a function of time with dependence on the tuning of a filter. Consider a signal $s(t)$ with spectrum $S(v)$, and a low-pass filter with real impulse response $h(t)$ and transfer function $H(v)$. To extract the band-pass “component” of $s(t)$ at frequency $v = f$, we shift the filter function in frequency so that it is centered at $v = f$, multiply the signal spectrum by the shifted filter transfer function, and take the inverse FT w.r.t. v , that is,

$$B_s^H(t,f) = \mathcal{F}_{t \leftarrow v}^{-1} \{S(v) H(v - f)\}, \quad (2.4.1)$$

where the “ B ” stands for “bandpass.” This signal may be understood as the complex output of a band-pass filter with center frequency $+f$; the input $s(t)$ may be reconstructed as the sum of the outputs of such *filter bank* whose transfer functions add up to unity within the bandwidth of $s(t)$.

The squared magnitude of $B_s^H(t,f)$ is denoted by $S_s^{(H)}(t,f)$ and is called the *sonograph* or *sonogram* so that:

$$S_s^{(H)}(t,f) = |B_s^H(t,f)|^2 \quad (2.4.2)$$

$$= \left| \mathcal{F}_{t \leftarrow v}^{-1} \{S(v) H(v - f)\} \right|^2 \quad (2.4.3)$$

$$= \left| \int_{-\infty}^{\infty} S(v) H(v - f) e^{j2\pi vt} dv \right|^2. \quad (2.4.4)$$

For the sonograph, the optimal bandwidth of the band-pass filter is related to the SD and dynamic bandwidth B_d in the same way that the optimal window length for the spectrogram is related to the IF and relaxation time T_r (see the duality relationship between these variables in [Chapter 1, Section 1.3](#)).

Like the spectrogram, the sonograph is nonlinear, but the nonlinearity is introduced only in the final step and does not lead to artifacts. Filter banks—the practical realization of the sonograph—have long been used in many applications including music broadcasting, recording, and even home entertainment, although the frequency division is not necessarily linear; for example, graphic equalizers and analyzers usually have logarithmic frequency division.

2.4.2 EQUIVALENCE TO SPECTROGRAM

Theorem 2.4.1. *The spectrogram and sonograph are equal if the window function $w(t)$ of the spectrogram is real and even and equal to the impulse response $h(t)$ of the sonograph filter for $f = 0$.*

Proof: Applying the inverse convolution property to Eq. (2.4.1), we obtain

$$B_s^H(t,f) = s(t) *_t \left[h(t) e^{j2\pi ft} \right] \quad (2.4.5)$$

$$= \int_{-\infty}^{\infty} s(\tau) h(t - \tau) e^{j2\pi f(t-\tau)} d\tau \quad (2.4.6)$$

$$= e^{j2\pi ft} \mathcal{F}_{\tau \rightarrow f} \{s(\tau) h(t - \tau)\}, \quad (2.4.7)$$

which yields

$$S_s^{(H)}(t, f) = \left| \mathcal{F}_{\tau \rightarrow f} \{s(\tau) h(t - \tau)\} \right|^2. \quad (2.4.8)$$

Comparing Eqs. (2.3.10) and (2.4.8), we see the spectrogram equals the sonograph if

$$h(t) = w(-t), \quad (2.4.9)$$

which is the case if $w(t)$ is even and equal to $h(t)$. The condition that $w(t)$ be real is redundant in the proof, but is assumed in the standard definition of the spectrogram. \square

2.5 METHOD 5: INSTANTANEOUS POWER SPECTRA

2.5.1 TIME-DEPENDENT SPECTRUM AND PAGE DISTRIBUTION

This approach attempts to define a kind of “running spectrum” by considering the variations of the signal spectrum as time t increases.

Given a signal $s(t)$, let us define the *running transform* $S_-(t, f)$ as the FT of the signal s up to time t . To do this, we first define the “auxiliary signal” $s_t(\theta)$ as

$$s_t(\theta) = \begin{cases} s(\theta) & \text{if } \theta \leq t, \\ 0 & \text{if } \theta > t. \end{cases} \quad (2.5.1)$$

Then the running transform is simply

$$S_-(t, f) = \mathcal{F}_{\theta \rightarrow f} \{s_t(\theta)\} = \int_{-\infty}^t s(\theta) e^{-j2\pi f \theta} d\theta. \quad (2.5.2)$$

As the normal energy spectrum is the squared magnitude of the FT, so the *running energy spectrum* up to time t , denoted by $e_s(t, f)$, is the squared magnitude of the running transform:

$$e_s(t, f) = |S_-(t, f)|^2 = S_-(t, f) S_-^*(t, f). \quad (2.5.3)$$

Differentiating each expression in this equation w.r.t. time, and denoting the time-derivative of $e_s(t, f)$ by $p_s(t, f)$, we obtain

$$p_s(t, f) = \frac{\partial}{\partial t} [|S_-(t, f)|^2] \quad (2.5.4)$$

$$= S_-(t, f) \frac{\partial}{\partial t} [S_-^*(t, f)] + S_-^*(t, f) \frac{\partial}{\partial t} [S_-(t, f)]. \quad (2.5.5)$$

This $p_s(t, f)$, being the time-derivative of a time-dependent energy spectrum, may be understood as a kind of *time-dependent power spectrum*. It is now known as the *Page distribution* in honor of its publisher [29].

Substituting the right-hand expression of Eq. (2.5.2) into Eq. (2.5.4), yields

$$P_s(t,f) = \frac{\partial}{\partial t} \left[\left| \int_{-\infty}^t s(\theta) e^{-j2\pi f\theta} d\theta \right|^2 \right], \quad (2.5.6)$$

which is the usual definition of the Page distribution showing a clear interpretation as a “spectral gradient”; the term between square brackets represents the spectral “accumulation” as t varies and its derivative represents the variation of this spectral “accumulation” at time t . Using Eq. (2.5.2) to evaluate the partial derivatives in Eq. (2.5.5), we obtain the alternative expression

$$P_s(t,f) = 2 \operatorname{Re} \left\{ s^*(t) S_-(t,f) e^{j2\pi ft} \right\}. \quad (2.5.7)$$

Substituting from Eq. (2.5.2) and writing $\tau = t - \theta$,

$$P_s(t,f) = 2 \operatorname{Re} \left\{ \int_0^\infty s^*(t) s(t - \tau) e^{j2\pi f\tau} d\tau \right\}. \quad (2.5.8)$$

If $s(t)$ is real as Page assumed [29], Eq. (2.5.8) becomes

$$P_s(t,f) = 2 \int_0^\infty s(t) s(t - \tau) \cos(2\pi f\tau) d\tau. \quad (2.5.9)$$

The Page distribution of a time-limited LFM signal is shown in Fig. 2.7.1(c). Note that the distribution can take *negative values*, which are inconsistent with the notion of an energy *distribution*, but perfectly consistent with the notion of an energy *gradient* as defined in Eq. (2.5.6). As negative values compensate for earlier spurious positive values caused by the truncation of $s(t)$ to produce $s_t(\theta)$, the presence of negative values implies that energy is delocalized in the (t,f) domain.

2.6 METHOD 6: ENERGY DENSITY

2.6.1 THE COMPLEX (t, f) ENERGY DENSITY

In search of a TFD localized in both time and frequency, Rihaczek [30] considered the energy of a complex deterministic signal over finite ranges of t and f , and allowed those ranges to become infinitesimal, obtaining what he called a *complex energy density*. Here, we offer a simpler derivation than that given by Rihaczek. The energy of a complex signal $z(t)$, with FT $Z(f)$, is

$$\begin{aligned} E &= \int_{-\infty}^{\infty} |z(t)|^2 dt = \int_{-\infty}^{\infty} z(t) z^*(t) dt = \int_{-\infty}^{\infty} z(t) \int_{-\infty}^{\infty} Z^*(f) e^{-j2\pi ft} df dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Re_z(t,f) dt df, \end{aligned} \quad (2.6.1)$$

where

$$\Re_z(t,f) = z(t) Z^*(f) e^{-j2\pi ft}. \quad (2.6.2)$$

This $\Re_z(t,f)$, known as the *Rihaczek distribution* (RD), is a (t,f) energy density in the sense that its integral over the (t,f) plane is the signal energy (Eq. (2.6.1)). If we express $Z(f)$ and hence $Z^*(f)$ in terms of $z(\lambda)$ and use the substitution $\tau = t - \lambda$, we obtain the alternative form

$$\mathfrak{R}_z(t, f) = \int_{-\infty}^{\infty} z(t) z^*(t - \tau) e^{-j2\pi f \tau} d\tau, \quad (2.6.3)$$

which looks like a nonsymmetric form of the WVD that shows more clearly the relationship between $z(t)$ and the TFD (respectively, $Z(f)$ and the TFD) by comparing Eq. (2.6.3) with Eq. (2.6.2). A TFD equivalent to Rihaczek's was derived earlier, in the context of quantum mechanics [31].

From Eq. (2.6.2), it is easily verified that

$$\int_{-\infty}^{\infty} \mathfrak{R}_z(t, f) df = |z(t)|^2, \quad (2.6.4)$$

$$\int_{-\infty}^{\infty} \mathfrak{R}_z(t, f) dt = |Z(f)|^2. \quad (2.6.5)$$

Thus, the RD satisfies the *marginal conditions* (see discussion after Eq. (2.1.17)).

Integrating Eq. (2.6.4) w.r.t. t and Eq. (2.6.5) w.r.t. f , we obtain, respectively,

$$\int_{t_1}^{t_2} \int_{-\infty}^{\infty} \mathfrak{R}_z(t, f) df dt = \int_{t_1}^{t_2} |z(t)|^2 dt, \quad (2.6.6)$$

$$\int_{f_1}^{f_2} \int_{-\infty}^{\infty} \mathfrak{R}_z(t, f) dt df = \int_{f_1}^{f_2} |Z(f)|^2 df. \quad (2.6.7)$$

The right-hand side of Eq. (2.6.6) is the energy in the time interval between t_1 and t_2 , while the right-hand side of Eq. (2.6.7) is the energy in the frequency band between f_1 and f_2 . Together, the two equations indicate that $\mathfrak{R}_z(t, f)$ can be interpreted as an energy density over an arbitrary time interval δt and an arbitrary frequency band δf .

In Eq. (2.6.2), the RD has the signal $z(t)$ as a factor. It follows that the RD is 0 at those times when $z(t)$ is 0; this property is called *strong time support*. Similarly, we see that the RD is 0 at those frequencies for which the spectrum $Z(f)$ is 0; this property is called *strong frequency support*.

2.6.2 THE REAL ENERGY (t, f) DENSITY

By simply taking the *real part* of the RD, we obtain a TFD that, like the RD, has strong time support and strong frequency support. Taking the real parts of Eqs. (2.6.4) and (2.6.5), we further conclude that this TFD, known as the *Levin distribution* (LD) [32], satisfies the marginal conditions and their corollaries.

Let the LD of the complex signal $z(t)$ be denoted by $L_z(t, f)$. Taking the real part of Eq. (2.6.2), we obtain the definition

$$L_z(t, f) = \operatorname{Re} \left\{ z(t) Z^*(f) e^{-j2\pi f t} \right\}. \quad (2.6.8)$$

Taking the real part of Eq. (2.6.3) yields the alternative expression

$$L_z(t, f) = \operatorname{Re} \left\{ \int_{-\infty}^{\infty} z(t) z^*(t - \tau) e^{-j2\pi f \tau} d\tau \right\}. \quad (2.6.9)$$

If $z(t)$ is replaced by a real signal $s(t)$, Eq. (2.6.9) reduces to the cosine form obtained by Levin [32]. Historically, the TFD obtained by Levin was a modification of the Page distribution and a precursor of the RD. It was first defined in a quantum-mechanical context [33], it is also called the *Margenau-Hill distribution* [34, p. 26].

2.6.3 WINDOWED RIHACZEK AND LEVIN DISTRIBUTIONS

Because energy is a real quantity, the real part of the RD is more significant than the imaginary part or the magnitude. Hence, a “plot of the RD” is usually a plot of the real part, that is, the LD. Figure 2.7.1(d) shows such a plot for a time-limited LFM signal. Although Eqs. (2.6.6) and (2.6.7) indicate that the energy of the TFD is well localized in the time and frequency dimensions separately, that is, in strips parallel to the f and t axes, it does not follow that the energy is well localized in both dimensions at once. Indeed, Fig. 2.7.1(d) shows that the peaks of the TFD are not confined to the IF law, but show many spurious features. The WVD of this signal (Fig. 2.7.1(a)) is much cleaner. Because the conventional RD/LD performs so poorly on such a simple signal, its relevance to practical applications is not straightforward.

By comparison with the RD, the spectrogram is remarkably free of artifacts. Recall that the spectrogram is the squared magnitude of the STFT. So one way to reduce artifacts in the RD is to introduce the STFT as a factor instead of the spectrum $Z(f)$, in Eq. (2.6.2). The resulting distribution, which might be called the *windowed RD*, is

$$\mathfrak{R}_z^w(t,f) = z(t) \left[\mathcal{F}_{\tau \rightarrow f} \{z(\tau) w(\tau - t)\} \right]^* e^{-j2\pi ft} \quad (2.6.10)$$

$$= z(t) \left[\int_{-\infty}^{\infty} z(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right]^* e^{-j2\pi ft}, \quad (2.6.11)$$

where w is the window function of the STFT. As the real part of the RD is the LD, we might as well designate the real part of the windowed RD as the *windowed LD*:

$$L_z^w(t,f) = \operatorname{Re} \left\{ z(t) \left[\int_{-\infty}^{\infty} z(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right]^* e^{-j2\pi ft} \right\}. \quad (2.6.12)$$

A windowed LD of a time-limited LFM signal is shown in Fig. 2.7.1(e); note the reduction in artifacts compared with the conventional LD.

From the form of Eq. (2.6.10), we see that the windowed RD and its real part, the windowed LD, have strong time support. Other properties of these distributions will emerge in later sections.

2.7 RELATIONSHIP BETWEEN TFDs USING TIME-LAG KERNELS

So far, we have considered six different approaches to defining a TFD. All the approaches seem natural and reasonable, and yet lead paradoxically to at least five different definitions of a TFD, all of them *quadratic* in the signal. The physical interpretation is that power is proportional to the product of an *effort* signal (such as voltage, force, or pressure) and the corresponding *flow* signal (current, linear velocity, or volume velocity, respectively). If the effort and flow signals are proportional to each

other, power then is proportional to the *square* of one or the other. Using the “signal kernel” approach (Eq. (2.1.1)), in general, we may write for any TFD $\rho_z(t, f)$:

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{R_z(t, \tau)\}, \quad (2.7.1)$$

where $R_z(t, \tau)$ is found simply by taking the inverse FT of $\rho_z(t, f)$ w.r.t. f . In the special case of the WVD, $R_z(t, \tau)$ is the IAF, denoted by $K_z(t, \tau)$. In other cases, we shall call $R_z(t, \tau)$ the *smoothed IAF*. The reason for this term will become apparent as we will show that $R_z(t, \tau)$ is a smoothed version of $K_z(t, \tau)$ for each of the TFDs that we have defined.

2.7.1 SPECTROGRAM AND WVD

2.7.1.1 Spectrogram and its time-lag kernel

From Eq. (2.3.10), the spectrogram with window function w can be rewritten as

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{z(\tau) w(\tau - t)\} \left[\mathcal{F}_{\tau \rightarrow f} \{z(\tau) w(\tau - t)\} \right]^*. \quad (2.7.2)$$

Taking the IFT ($f \rightarrow \tau$) of both sides gives

$$\begin{aligned} R_z(t, \tau) &= [z(\tau) w(\tau - t)] * [z^*(-\tau) w^*(-\tau - t)] \\ &= \int_{-\infty}^{\infty} z(\lambda) w(\lambda - t) z^*(\lambda - \tau) w^*(\lambda - \tau - t) d\lambda \\ &= \int_{-\infty}^{\infty} z\left(u + \frac{\tau}{2}\right) w\left(u - t + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) w^*\left(u - t - \frac{\tau}{2}\right) du, \end{aligned} \quad (2.7.3)$$

where $\lambda = u + \frac{\tau}{2}$ is the dummy variable in the convolution. Exploiting the evenness of w , this can be written

$$R_z(t, \tau) = \int_{-\infty}^{\infty} w^*\left(t - u + \frac{\tau}{2}\right) w\left(t - u - \frac{\tau}{2}\right) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) du. \quad (2.7.4)$$

If we let

$$G(t, \tau) = w^*\left(t + \frac{\tau}{2}\right) w\left(t - \frac{\tau}{2}\right), \quad (2.7.5)$$

we have

$$R_z(t, \tau) = \int_{-\infty}^{\infty} G(t - u, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) du \quad (2.7.6)$$

$$= G(t, \tau) *_t K_z(t, \tau), \quad (2.7.7)$$

where $K_z(t, \tau)$ is given by Eq. (2.1.8). $G(t, \tau)$ is called the *time-lag kernel*.¹ Equation (2.7.7) defines the time-lag kernel as that which must be convolved in time with the IAF to obtain the smoothed IAF;

¹The term “kernel” was used in this sense in [35]. To avoid confusion between “time-lag kernel” $G(t, \tau)$ and “signal kernel” $K_z(t, \tau)$, the latter is usually called the IAF in this book.

the word “smoothed” refers to the convolution. We shall adopt this definition for *all* TFDs given by Eq. (2.7.1). Substituting Eq. (2.7.6) into Eq. (2.7.1) and writing out the FT, we obtain one expression of the general form of quadratic TFDs.

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t-u, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau. \quad (2.7.8)$$

For the spectrogram, $G(t, \tau)$ is given by Eq. (2.7.5). As w is real, we may interchange w and w^* , leading to the conclusion that the time-lag kernel for the spectrogram is simply the IAF of the observation window function.

2.7.1.2 Wigner-Ville distribution and its time-lag kernel

The “smoothed IAF” for the WVD reduces to the trivial case as it yields

$$R_z(t, \tau) = K_z(t, \tau) = \delta(t) * K_z(t, \tau); \quad G(t, \tau) = \delta(t). \quad (2.7.9)$$

In this trivial case, the “smoothing” makes no difference, as $\delta(t)$ is the identity element for the convolution operation. In essence, there is no actual smoothing in this case. There would be, however, if we considered another TFD as the core one, for example, the spectrogram.

2.7.2 RIHACZEK AND LEVIN DISTRIBUTIONS AND THEIR TIME-LAG KERNELS

2.7.2.1 Rihaczek distribution and its time-lag kernel

Equation (2.6.3) may be written

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{z(t) z^*(t - \tau)\}. \quad (2.7.10)$$

Taking the IFT gives

$$R_z(t, \tau) = z(t) z^*(t - \tau) = \delta\left(t - \frac{\tau}{2}\right) * \left[z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right)\right] \quad (2.7.11)$$

$$= \delta\left(t - \frac{\tau}{2}\right) * K_z(t, \tau) \quad (2.7.12)$$

so that

$$G(t, \tau) = \delta\left(t - \frac{\tau}{2}\right). \quad (2.7.13)$$

2.7.2.2 Levin distribution and its time-lag kernel

Using Eq. (2.7.10) as the definition of the RD, and taking the real part, we obtain for the LD

$$\rho_z(t, f) = \operatorname{Re} \left\{ \mathcal{F}_{\tau \rightarrow f} \{z(t) z^*(t - \tau)\} \right\} \quad (2.7.14)$$

$$= \frac{1}{2} \mathcal{F}_{\tau \rightarrow f} \{z(t) z^*(t - \tau)\} + \frac{1}{2} \left[\mathcal{F}_{\tau \rightarrow f} \{z(t) z^*(t - \tau)\} \right]^*. \quad (2.7.15)$$

Taking the IFT, we obtain

$$R_z(t, \tau) = \frac{1}{2} z(t) z^*(t - \tau) + \frac{1}{2} z^*(t) z(t + \tau) \quad (2.7.16)$$

$$\begin{aligned} &= \frac{1}{2} \delta\left(t - \frac{\tau}{2}\right) * K_z(t, \tau) + \frac{1}{2} \delta\left(t + \frac{\tau}{2}\right) * K_z(t, \tau) \\ &= \frac{1}{2} \left[\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right) \right] * K_z(t, \tau) \end{aligned} \quad (2.7.17)$$

so that

$$G(t, \tau) = \frac{1}{2} \left[\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right) \right]; \quad (2.7.18)$$

yielding a TFD expressed as a cosine function that is the real part of the RD.

2.7.2.3 Windowed Rihaczek distribution and its time-lag kernel

Equation (2.6.10) can be written

$$\rho_z(t, f) = z(t) \mathcal{F}_{\tau \rightarrow f} \{z^*(-\tau) w^*(-\tau - t)\} e^{-j2\pi ft} \quad (2.7.19)$$

$$= z(t) \mathcal{F}_{\tau \rightarrow f} \{z^*(t - \tau) w^*(-\tau)\}. \quad (2.7.20)$$

Taking the IFT gives

$$R_z(t, \tau) = z(t) z^*(t - \tau) w^*(-\tau) = \left[\delta\left(t - \frac{\tau}{2}\right) * K_z(t, \tau) \right] w^*(-\tau) \quad (2.7.21)$$

$$= \left[w^*(-\tau) \delta\left(t - \frac{\tau}{2}\right) \right] * K_z(t, \tau) \quad (2.7.22)$$

so that

$$G(t, \tau) = w^*(-\tau) \delta\left(t - \frac{\tau}{2}\right). \quad (2.7.23)$$

Because w is real and even, this reduces to

$$G(t, \tau) = w(\tau) \delta\left(t - \frac{\tau}{2}\right). \quad (2.7.24)$$

A comparison of Eqs. (2.7.12) and (2.7.22) indicates that the smoothed IAFs of the RD and the windowed RD differ by the factor $w^*(-\tau)$ (i.e., by a windowing operation in the lag domain prior to Fourier transformation from τ to f).

2.7.2.4 Windowed Levin distribution and its time-lag kernel

From Eq. (2.7.20), the real part of the windowed RD is

$$\frac{1}{2} z(t) \mathcal{F}_{\tau \rightarrow f} \{z^*(t - \tau) w^*(-\tau)\} + \frac{1}{2} z^*(t) \left[\mathcal{F}_{\tau \rightarrow f} \{z^*(t - \tau) w^*(-\tau)\} \right]^*. \quad (2.7.25)$$

Taking the IFT, we find that the smoothed IAF of the windowed LD is

$$\begin{aligned} R_z(t, \tau) &= \frac{1}{2} z(t) z^*(t - \tau) w^*(-\tau) + \frac{1}{2} z^*(t) z(t + \tau) w(\tau) \\ &= \frac{1}{2} \left[\delta\left(t - \frac{\tau}{2}\right) *_t K_z(t, \tau) \right] w^*(-\tau) + \frac{1}{2} \left[\delta\left(t + \frac{\tau}{2}\right) *_t K_z(t, \tau) \right] w(\tau) \\ &= \frac{1}{2} \left[w(\tau) \delta\left(t + \frac{\tau}{2}\right) + w^*(-\tau) \delta\left(t - \frac{\tau}{2}\right) \right] *_t K_z(t, \tau) \end{aligned} \quad (2.7.26)$$

so that

$$G(t, \tau) = \frac{1}{2} \left[w(\tau) \delta\left(t + \frac{\tau}{2}\right) + w^*(-\tau) \delta\left(t - \frac{\tau}{2}\right) \right] \quad (2.7.27)$$

or, because w is real and even,

$$G(t, \tau) = \frac{1}{2} w(\tau) \left[\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right) \right]. \quad (2.7.28)$$

A comparison of Eqs. (2.7.17) and (2.7.26) indicates that the smoothed IAFs of the LD and the windowed LD differ by the factor $w(\tau)$ (i.e., by a windowing operation in the lag domain before taking the FT from τ to f).

2.7.3 PAGE DISTRIBUTION AND ITS TIME-LAG KERNEL

Rewriting Eq. (2.5.8) using the unit step function $u(t)$, we obtain for the Page distribution

$$\rho_z(t, f) = 2 \operatorname{Re} \left\{ \int_{-\infty}^{\infty} z^*(t) z(t - \lambda) u(\lambda) e^{j2\pi f \lambda} d\lambda \right\}. \quad (2.7.29)$$

With the substitution $\tau = -\lambda$, this becomes

$$\rho_z(t, f) = 2 \operatorname{Re} \left\{ \mathcal{F}_{\tau \rightarrow f} \{z^*(t) z(t + \tau) u(-\tau)\} \right\} \quad (2.7.30)$$

$$= \mathcal{F}_{\tau \rightarrow f} \{z^*(t) z(t + \tau) u(-\tau)\} + \left[\mathcal{F}_{\tau \rightarrow f} \{z^*(t) z(t + \tau) u(-\tau)\} \right]^*. \quad (2.7.31)$$

Taking the IFT, we obtain

$$R_z(t, \tau) = z^*(t) z(t + \tau) u(-\tau) + z(t) z^*(t - \tau) u(\tau) \quad (2.7.32)$$

$$= u(-\tau) \left[\delta\left(t + \frac{\tau}{2}\right) *_t K_z(t, \tau) \right] + u(\tau) \left[\delta\left(t - \frac{\tau}{2}\right) *_t K_z(t, \tau) \right] \quad (2.7.33)$$

$$= \left[u(-\tau) \delta\left(t + \frac{\tau}{2}\right) + u(\tau) \delta\left(t - \frac{\tau}{2}\right) \right] *_t K_z(t, \tau) \quad (2.7.34)$$

so that

$$G(t, \tau) = u(-\tau) \delta\left(t + \frac{\tau}{2}\right) + u(\tau) \delta\left(t - \frac{\tau}{2}\right). \quad (2.7.35)$$

2.7.4 RELATIONSHIP BETWEEN THE WVD AND OTHER TFDs

2.7.4.1 TFDs as filtered WVDs

By taking the FT of Eq. (2.7.7) w.r.t. τ and using specific forms for $G(t, \tau)$, all the considered TFDs can be written in the same form as Eq. (2.2.15):

$$\rho_z(t, f) = \gamma(t, f) \ast\ast_{(t, f)} W_z(t, f), \quad (2.7.36)$$

where

$$\gamma(t, f) = \mathcal{F}_{\tau \rightarrow f} \{G(t, \tau)\} \quad (2.7.37)$$

is called the time-frequency kernel. This then suggests that all the TFDs naturally introduced so far can be considered to be smoothed WVDs. Full insights and interpretations are provided in [Chapter 3](#).

2.7.4.2 Relationship between the WVD and the ambiguity function

If we take the 2D-FT of Eq. (2.2.15) (more precisely one FT operation and one inverse FT operation for consistency of dimensions and meaning), the double (t, f) convolution is replaced by a product in the Doppler-lag (v, τ) domain, also called the *ambiguity domain*. The symmetrical *ambiguity function* (AF) is defined as

$$A_z(v, \tau) = \mathcal{F}_{t \rightarrow v} \{K_z(t, \tau)\} = \mathcal{F}_{t \rightarrow v} \left\{ \mathcal{F}_{\tau \leftarrow f}^{-1} \{W_z(t, f)\} \right\}, \quad (2.7.38)$$

where $\mathcal{F}\{\cdot\}$ is the forward (direct) FT, and $\mathcal{F}^{-1}\{\cdot\}$ its inverse. This 2D-FT is written

$$W_z(t, f) \stackrel{t \rightarrow v}{\stackrel{f \rightarrow \tau}{\equiv}} A_z(v, \tau). \quad (2.7.39)$$

Because convolution with $W_z(t, f)$ corresponds to multiplication by $A_z(v, \tau)$, design of a TFD is likely to be simpler in the (v, τ) domain.² The above observation and relationships are important precursors of advanced procedures that led to the design or rediscovery of several other TFDs (see discussion of the AF in [Section 5.1](#)).

2.7.5 MBD, EMBD, AND OTHER QUADRATIC TFDs OF INTEREST

Having derived some TFDs by intuitive methods and then determined their time-lag kernels, let us now define a few more TFDs directly in terms of their time-lag kernels. Although for now we specify each TFD in terms of $G(t, \tau)$, the actual design is often more conveniently done in the ambiguity (v, τ)

²*Note:* The symmetric ambiguity function, which plays a fundamental role in radar, is the two-dimensional FT of the WVD. As a consequence, knowledge gained in the field of radar contributed to the development of (t, f) methods. A key property of the ambiguity function is that it separates significantly cross-terms from auto-terms (for details, see [Sections 3.3.3, 5.2](#), and [6.1](#)). Another important property is that convolutions in (t, f) become products in the ambiguity domain, so efficient that filtering procedures can be introduced in the ambiguity domain to improve or correct some features of the TFD—for example, by reducing cross-terms.

domain, as summarized in [Table 3.3.3](#) in the next chapter. Implementation of the kernel could be in the (t, τ) domain as in [Section 6.5](#), or the (ν, τ) domain as in [Section 6.6](#), and need not be done in the same domain as the design.

2.7.5.1 The modified B-distribution

In [Table 2.7.1](#), the parameters α, σ, a , and β are real and positive, and $w(\tau)$ is a window function. The exponential distribution (also known as the Choi-Williams distribution) was defined in [36] and the ZAM distribution in [37]. The BJ distribution, called the “sinc distribution” in [34, p. 26], was defined in [34] using an operational rule of Born and Jordan [38, p. 873]. Observe that the BJ distribution is a special case of the ZAM with $a = 1/\alpha$ and $w(\tau) = a/|2\tau|$. These TFDs are illustrated for an LFM signal in [Fig. 2.7.1\(f\)-\(i\)](#). The kernels of the BD, w -WVD, modified B-distribution (MBD), and extended modified B-distribution (EMBD) [39,40] are “separable”; the w -WVD kernel is also “Doppler-independent” and similarly, the MBD kernel is also “lag-independent”; these concepts are discussed later, especially in [Section 5.7](#).

2.7.5.2 The extended modified B-distribution

The EMBD is an improvement of the MBD obtained by adding an identical kernel function in the lag direction in the ambiguity domain. This modification allows the EMBD to include smoothing over both time and frequency axes (whereas the MBD performs smoothing only over the time axis). As a consequence, the EMBD can better characterize signals whose components appear along either time or frequency directions or elsewhere, by tuning two parameters, each of which control the kernel shape and size in one axis. Applications of the EMBD include the analysis of newborn EEG seizure signals which often manifest as quasi-stationary components or a train of impulses [39]. In [Fig. 2.7.1\(j\)-\(l\)](#), we have, respectively, plotted a Doppler-independent, a lag-independent, and a separable-kernel TFD of an LFM signal. Note the undesirable wave pattern shown in parts (c) and (d) are caused by the cos function in TFDs related to Rihaczek’s. We shall see (e.g., in [Table 3.3.3](#) on p. 124) that the separable, Doppler-independent, and lag-independent kernels have been introduced in order to obtain certain desired TFD properties.

Table 2.7.1: A Selection of TFDs Designed Using Filter-Based Specifications

| Name | $G(t, \tau)$ |
|---|---|
| Windowed WVD (w -WVD) | $\delta(t) w(\tau)$ |
| Sinc or Born-Jordan (BJ) | $\frac{1}{ 2\alpha\tau } \text{rect} \frac{t}{2\alpha\tau}$ |
| Exponential (ED) or Choi-Williams | $\frac{\sqrt{\pi\sigma}}{ \tau } e^{-\pi^2\sigma t^2/\tau^2}$ |
| Windowed sinc or Zhao-Atlas-Marks (ZAM) | $w(\tau) \text{rect} \frac{t}{2\tau/a}$ |
| B-distribution (BD) | $ \tau ^\beta \cosh^{-2\beta} t$ |
| Modified B-distribution (MBD) | $\frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi}$ |
| Extended modified B-distribution (EMBD) | $\frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi} \cdot \frac{[\Gamma(\alpha+j\pi\tau)]^2}{\Gamma^2(\alpha)}$ |

Table 2.7.2: Special Forms of Selected Quadratic TFDs (Third Column), Together with Their Time-Lag Kernels (Second Column)

| TFD Name | $G(t, \tau)$ | $\rho_z(t, f)$ |
|--------------|--|---|
| Wigner-Ville | $\delta(t)$ | $\int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau$ |
| Levin | $\frac{1}{2} [\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right)]$ | $\text{Re} \{z(t) Z^*(f) e^{-j2\pi f\tau}\}$ |
| Born-Jordan | $\frac{1}{ 2\alpha\tau } \text{rect} \frac{t}{2\alpha\tau}$ | $\int_{-\infty}^{\infty} \int_{t- \alpha\tau }^{t+ \alpha\tau } \frac{1}{ 2\alpha\tau } z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| Modified BD | $\frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi}$ | $\iint \frac{\cosh^{-2\beta}(t-u)}{\int \cosh^{-2\beta} \xi d\xi} z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| EMBD | $\frac{ \Gamma(\alpha+j\pi\tau) ^2 \cosh^{-2\beta} t}{\Gamma^2(\alpha) \int \cosh^{-2\beta} \xi d\xi}$ | $\iint \frac{ \Gamma(\alpha+j\pi\tau) ^2 \cosh^{-2\beta}(t-u)}{\Gamma^2(\alpha) \int \cosh^{-2\beta} \xi d\xi} z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| w-WVD | $\delta(t) w(\tau)$ | $\int_{-\infty}^{\infty} w(\tau) z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau$ |
| w-Levin | $\frac{w(\tau)}{2} [\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right)]$ | $\text{Re} \left\{ z(t) \left[\int_{-\infty}^{\infty} z(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right]^* e^{-j2\pi f\tau} \right\}$ |
| ZAM | $w(\tau) \text{rect} \frac{t}{2\tau/a}$ | $\int_{-\infty}^{\infty} \int_{t- \frac{\tau}{a} }^{t+ \frac{\tau}{a} } w(\tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| Rihaczek | $\delta\left(t - \frac{\tau}{2}\right)$ | $z(t) Z^*(f) e^{-j2\pi f\tau}$ |
| w-Rihaczek | $w(\tau) \delta\left(t - \frac{\tau}{2}\right)$ | $z(t) \left[\int_{-\infty}^{\infty} z(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right]^* e^{-j2\pi f\tau}$ |
| Page | $u(-\tau) \delta\left(t + \frac{\tau}{2}\right) + u(\tau) \delta\left(t - \frac{\tau}{2}\right)$ | $\frac{\partial}{\partial t} \left[\left \int_{-\infty}^t z(\tau) e^{-j2\pi f\tau} d\tau \right ^2 \right]$ |
| ED | $\frac{\sqrt{\pi\sigma}}{ \tau } e^{-\pi^2 \sigma t^2 / \tau^2}$ | $\iint \frac{\sqrt{\pi\sigma}}{ \tau } e^{-\frac{\pi^2 \sigma(t-u)^2}{\tau^2}} z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| BD | $ \tau ^\beta \cosh^{-2\beta} t$ | $\iint \frac{ \tau ^\beta}{\cosh^{2\beta}(t-u)} z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau$ |
| Spectrogram | $w\left(t + \frac{\tau}{2}\right) w\left(t - \frac{\tau}{2}\right)$ | $\left \int_{-\infty}^{\infty} z(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \right ^2$ |

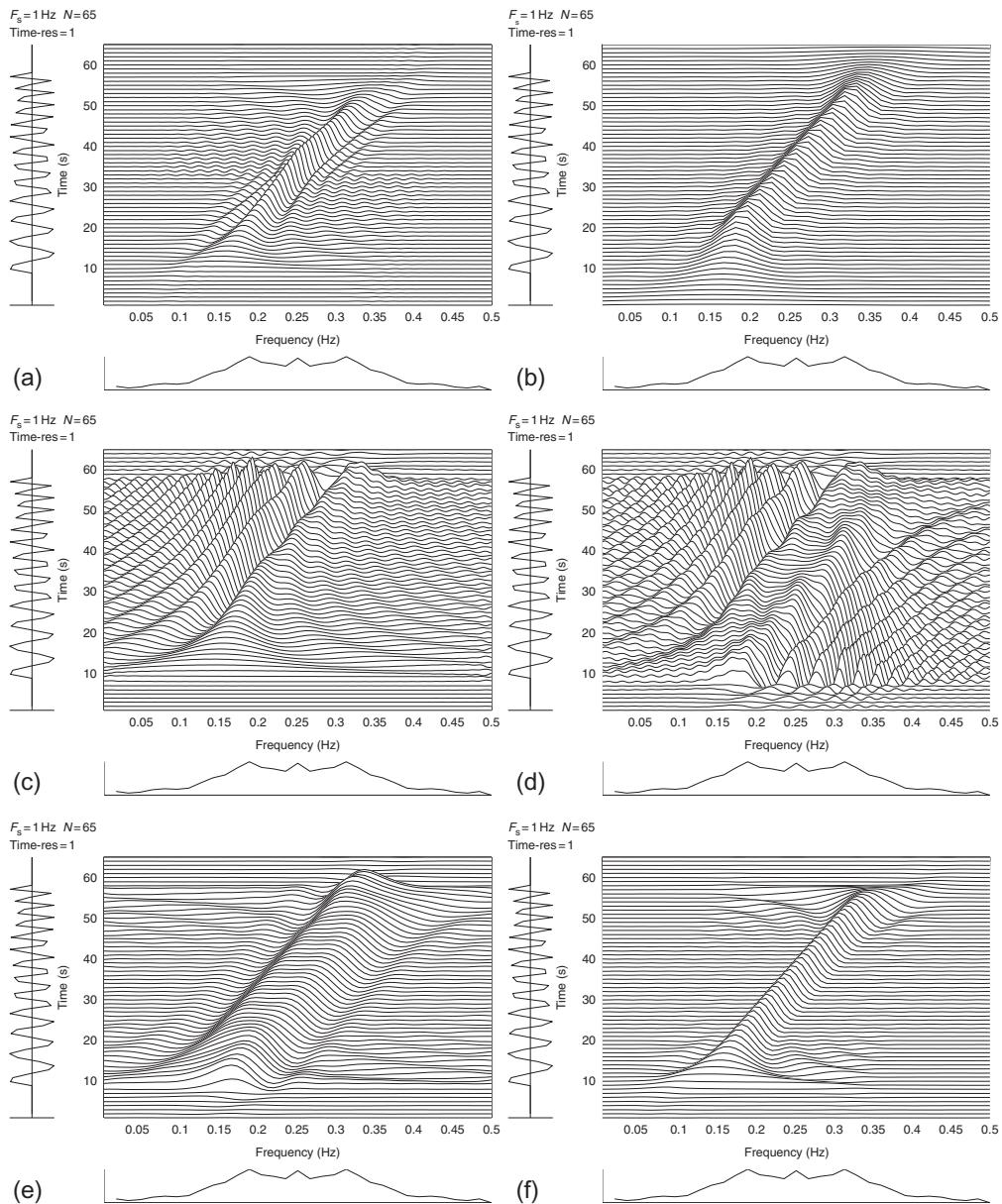
Notes: The window $w(\tau)$ is assumed to be real and even. Integrals, unless otherwise noted, are from $-\infty$ to ∞ . The forms involving double integrals are obtained by direct substitution of the time-lag kernel into Eq. (2.7.8). The w-WVD may also be so obtained. Other forms are quoted from this chapter.

For convenience, Table 2.7.2 collects and tabulates the definitions of the TFDs that have been derived heuristically in this chapter, and of the seven TFDs that have just been defined in terms of their time-lag kernels. Chapter 3 reviews these methods in more depth within a general approach and Section 6.1 extends these definitions to the discrete-time case. Other related methods such as polynomial TFDs and complex-lag TFDs are presented in Chapter 5.

2.7.6 TIME-SCALE AND WAVELET TRANSFORMS

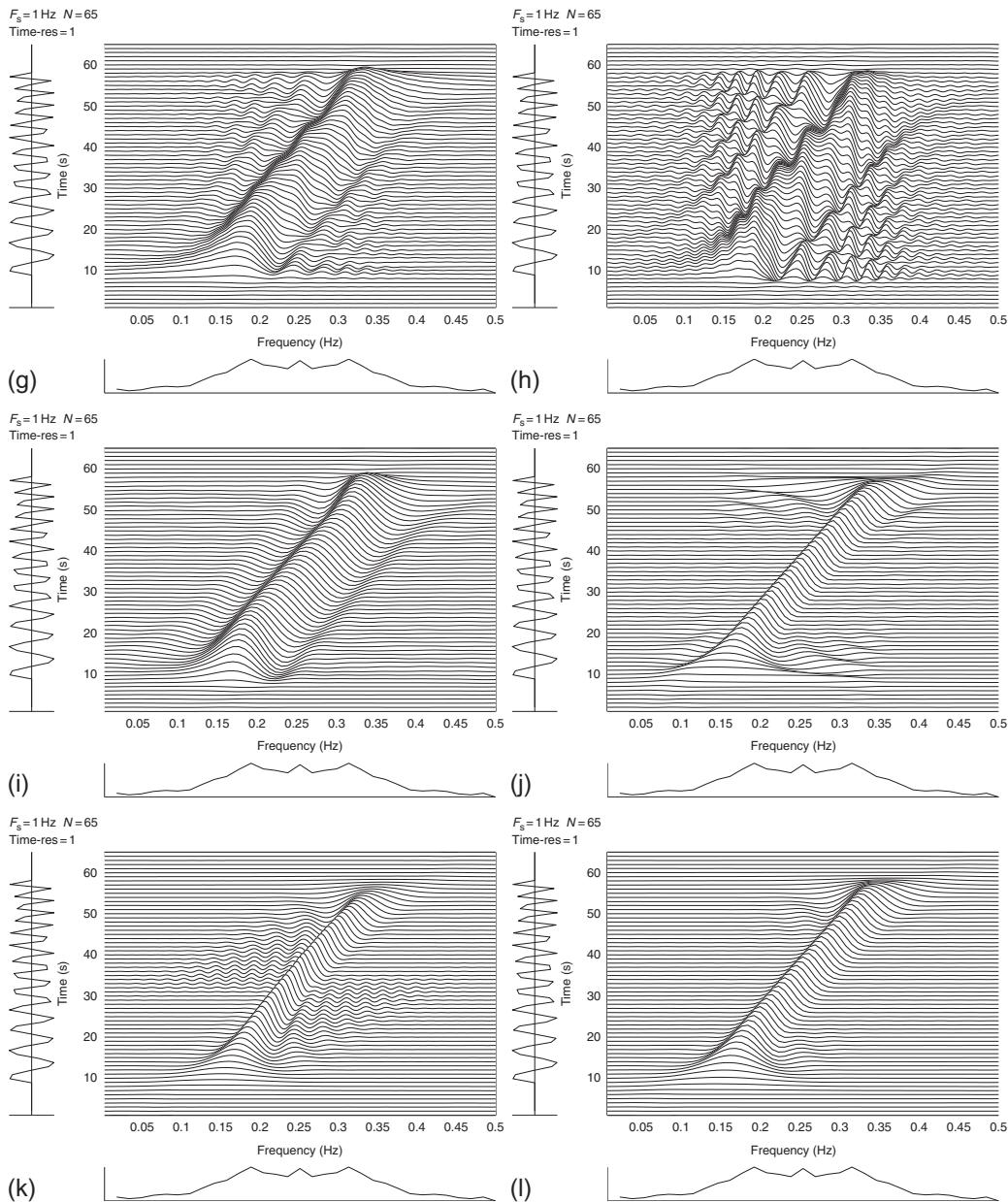
2.7.6.1 Signal decomposition-based methods

The previous sections have explained how TFDs can be obtained by taking the FTs of various smoothed autocorrelation functions of the signal. For completeness, let us present another approach to (t, f) localization, which is to expand a signal using a basis of orthogonal signals, called *wavelets*, which are already localized in time and frequency. Wavelets are wave-like transients that can be interpreted

**FIGURE 2.7.1**

TFDs of an LFM signal with displayed duration 65 s or samples (sampling rate 1 Hz), unit amplitude from sample 9 to sample 57 and zero amplitude elsewhere, frequency range 0.1 (at sample 1) to 0.4 (at sample 65): (a) Wigner-Ville; (b) Spectrogram, 21-point rectangular window; (c) Page; (d) Levin; (e) Windowed Levin, 19-point Hamming window; (f) exponential, $\sigma = 4$;

Continued

**FIGURE 2.7.1, CONT'D**

(g) Born-Jordan, $\alpha = 1/2$; (h) Zhao-Atlas-Marks, $a = 2$, $g_2(\tau) = 1$; (i) Zhao-Atlas-Marks, $a = 2$, $g_2(\tau) = 19\text{-point Hamming window}$; (j) Doppler-independent, $g_2(\tau) = 29\text{-point Hamming window}$; (k) Lag-independent, $g_1(t) = 9\text{-point Hamming}$; and (l) Separable kernel, $g_1(t) = 9\text{-point Hamming}$, $g_2(\tau) = 29\text{-point Hamming}$.

as sinusoids of short duration (e.g., if they are in the audible frequency range, they can be perceived by the ear as pure tones of short duration). They are defined as [28]

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right), \quad (2.7.40)$$

where τ and s are the translation and scale (dilation) parameters, respectively. The function ψ is the basis function called the mother wavelet. Decompositions of a signal on such a basis are called wavelet transforms (WTs) and are localized equivalents of FTs. They provide an alternative tool for representing local features of a signal.

Recall that a finite-energy signal $x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df$ can be represented by its FT $X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$. The FT decomposes the time-domain signal into a linear combination of harmonics $e^{j2\pi ft}$, which are considered as the “elementary messages” of which the signal is made. The wavelet transform (WT) is defined in a similar manner except that the chosen “elementary messages” are wavelets. Using wavelets as a basis, the time-domain signal can be represented as (Section 4.1)

$$x(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} \Psi_x^\psi(\tau, s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \frac{ds}{s^2} d\tau, \quad (2.7.41)$$

where the coefficients in the expansion are given by [41]

$$\Psi_x^\psi(\tau, s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-\tau}{s}\right) dt. \quad (2.7.42)$$

The coefficient c_ψ is a constant that depends on the wavelet used, and $\Psi_x^\psi(\tau, s)$ is the continuous WT of the signal $x(t)$. A number of mother wavelets have been proposed, such as the Mexican hat wavelet and the Morlet wavelet [42].

The discrete version of the WT is called discrete wavelet transform (DWT). It is realized by first discretizing the parameter scale s on a logarithmic grade. The time parameter is then discretized with respect to the scale parameter; that is, a different sampling rate is used for every scale. Usually, the sampling is performed on a dyadic sampling grid. With this sampling, a signal $x(t)$ can be decomposed into a set of orthogonal basis functions (scaled and shifted versions of the mother wavelet ψ); that is [28] (Section 4.1):

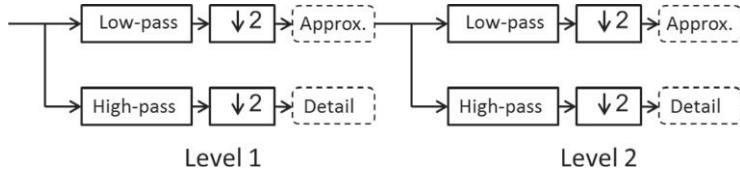
$$x(t) = c_\psi \sum_l \sum_k a_{lk} s_0^{-l/2} \psi\left(s_0^{-l/2} t - k\tau_0\right), \quad (2.7.43)$$

where

$$a_{lk} = \int_{-\infty}^{\infty} x(t) s_0^{-l/2} \psi\left(s_0^{-l/2} t - k\tau_0\right) dt \quad (2.7.44)$$

with τ_0 and s_0 being positive constants usually taken as 1 and 2, respectively [28]. The integer l describes the different levels of resolution, and k covers the number of time locations in each level.

The WT allows localization in both the time domain via translations of the mother wavelet, and in the scale (frequency) domain via dilations. The wavelet is irregular in shape and compactly supported, thus making it a good tool for analyzing signals of a transient nature. Irregularity of the wavelet basis makes it suitable for the analysis of signals with discontinuities or sharp changes, while the compactly supported nature of wavelets enables temporal localization of a signal features.

**FIGURE 2.7.2**

Multiresolution analysis using the DWT as a tree of cascading filter banks.

The multiresolution analysis of the DWT can be represented as a tree of low- and high-pass filters, with each step transforming the low-pass filter as shown in Fig. 2.7.2. The original signal is successively decomposed into components of lower resolution, while the high-frequency components are not analyzed any further, unlike wavelet packet analysis.

In contrast with the regular DWT, discrete wavelet packet analysis (DWPA) can significantly increase the versatility and power of the DWT. Unlike the DWT, DWPA utilizes both the low-frequency components (approximations), and the high-frequency components (details). From this family of bases, methods for choosing the optimal scheme (i.e., best basis) for a particular signal can be developed such as that based on energy compaction (entropy) [43]. Section 7.9 illustrates the applications of the DWT and DWPA (see also [44,45]). For further details about the theory, implementation of WTs and wavelet packet transforms, the reader can consult [41,46–48].

2.7.6.2 Relationship of S-transform with WT and STFT

Due to the fixed width of the window function, the STFT has poor time frequency resolution [49]. The WT on the other hand uses a basis function that dilates and contracts with frequency; it does not retain the absolute phase information. A (t,f) representation that combines the good features of the STFT and the WT is the S-transform [49,50]; it retains the phase information of a signal as in the STFT and provides the variable resolution as in the case of WT [51]. It is based on a scalable localizing Gaussian window and provides the frequency-dependent resolution [49,52]. The continuous wavelet transform (CWT) of a signal $x(t)$ is defined as [49,52]:

$$\Psi_x^\psi(t,f) = \int_{-\infty}^{\infty} x(\tau)\psi(\tau-t,f) d\tau, \quad (2.7.45)$$

where $\psi(t,f)$ is a scaled replica of the fundamental mother wavelet. The S-transform $S_T(t,f)$ of the signal $x(t)$ is defined as [49]:

$$S_T(t,f) = e^{j2\pi ft} \Psi_x^\psi(t,f), \quad (2.7.46)$$

which is the CWT $\Psi_x^\psi(t,f)$ with a specific mother wavelet multiplied by the phase correction factor $e^{j2\pi ft}$. The mother wavelet for Eq. (2.7.46) is defined as [49]:

$$\psi(t,f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{-t^2 f^2}{2}} e^{-j2\pi ft}. \quad (2.7.47)$$

The final expression of the S-transform $S_T(t, f)$ becomes

$$S_T(t, f) = \int_{-\infty}^{\infty} x(\tau) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-(\tau-t)^2 f^2}{2}} e^{-j2\pi f\tau} d\tau. \quad (2.7.48)$$

On the other hand, the STFT of the signal $x(t)$ is defined as:

$$F_s^w(t, f) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau. \quad (2.7.49)$$

The S-transform can be derived from Eq. (2.7.49) by replacing the window function $w(t)$ with the Gaussian function shown as:

$$w_f(t) = \frac{|f|}{\sqrt{2\pi}} e^{\frac{-t^2 f^2}{2}} = \frac{1}{\sigma(f) \sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2(f)}}, \quad (2.7.50)$$

where $\sigma(f) = \frac{1}{|f|}$. Thus, the S-transform is a special case of the STFT with frequency-dependent Gaussian window function given in Eq. (2.7.50). The advantage of the S-transform over the STFT is that the standard deviation $\sigma(f)$ is a function of frequency and the window function $w_f(t)$ is a function of both time and frequency. As the time-domain width of the analysis window is controlled by the frequency, the window is wider in the time domain for the lower frequencies, and narrower for the higher frequencies. In other words, the window provides good localization in the frequency domain for the low frequencies, while it provides good localization in time domain for higher frequencies [53] (see Section 5.11 for details).

It follows that the generalized S-transform can be expressed as:

$$S_T(t, f, \sigma) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t, \sigma(f)) e^{-j2\pi f\tau} d\tau, \quad (2.7.51)$$

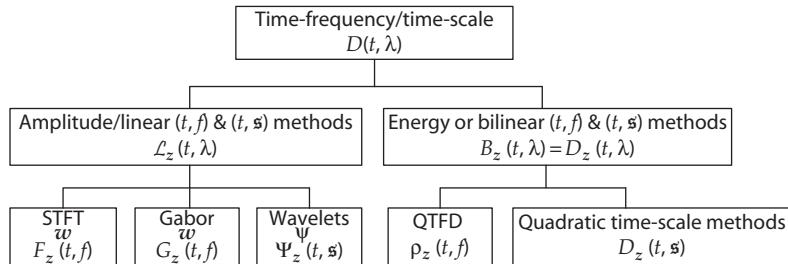
where $w(t, \sigma(f))$ is the frequency-controlled signal analysis window. Observing the definition of the multiresolution Fourier transform (MFT) given in [54,55], the generalized S-transform $S_T(t, f, \sigma)$ falls within the definition of the MFT [56]. The S-transform performs a multiresolution analysis of the signal with a frequency-controlled window. Section 4.1 elaborates on the relationships between such time-scale methods and quadratic TFDs.

2.7.7 LINKS AMONG LINEAR, QUADRATIC TFDs, AND TIME-SCALE METHODS

2.7.7.1 Relationships between nonstationary signal analysis methods

The aim of this section is to summarize schematically all the linear and quadratic (t, f) and (t, s) methods³ discussed in this chapter. Figure 2.7.3 shows a classification of these methods in term of whether they are linear versus quadratic, and (t, f) or (t, s) . The linear (t, f) representation discussed in this chapter is the STFT and wavelets (Sections 2.3 and 2.7.6). In essence, one can summarize that the STFT (and therefore the spectrogram) is a “constant window” method and the WT $((t, s)$ method) is “constant Q” transform, where in the context of filterbanks (sonogram) Q is the ratio of a filter center

³The symbol s is used here to represent the scale parameter and avoid confusion with other uses of the like, for example, $s(t)$.

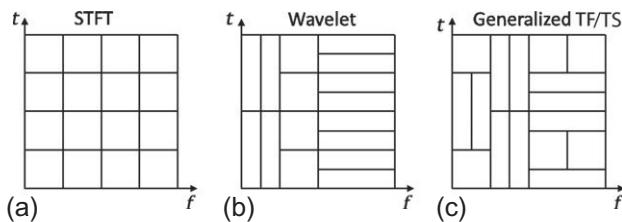
**FIGURE 2.7.3**

Linear and quadratic (t, f) methods for nonstationary signals analysis (S-transform combines STFT and wavelets).

frequency f_c to the filter bandwidth Δf , that is, $Q = \frac{f_c}{\Delta f}$. On the other hand, the WVD-based quadratic TFDs discussed in this chapter are (t, f) varying window methods.

2.7.7.2 Tiling of the (t, f) plane

Figure 2.7.4 shows the tiling of the (t, f) plane with rectangular tiles for the STFT, WT, and a generalized (t, f) method [57]. Each rectangle within the tiling diagram represents a filter, whose pass-band and temporal resolution are presented by the width and the height of the rectangle, respectively. Figure 2.7.4(a) shows the STFT, which is obtained by using fixed-signal-analysis window for all time instants [58]. Figure 2.7.4(b) shows the WT, which has narrow time windows that yield better time resolution at higher frequencies and wider time windows that yield better frequency resolution at low frequencies. Hence, the STFT gives fixed resolution at all times, while the WT gives variable resolutions. Figure 2.7.4(c) indicates that the generalized (t, f) method adapts the signal analysis window in time as well as frequency; STFT and WT are particular cases of the generalized (t, f) representation [59]. Note that the area of each rectangular box within the tiling diagram is the same, which implies that achieving arbitrarily precise resolution of both timing and frequency is not possible as per the Heisenberg principle (Eq. (I.2.4) in Chapter I). The way boxes within the tiling (t, f) diagram are shaped reflects the chosen trade-off between the time and frequency resolutions.

**FIGURE 2.7.4**

Tilings of the (t, f) plane: (a) short-time Fourier transform (STFT); (b) wavelet transform; and (c) generalized (t, f) methods.

2.8 SUMMARY AND DISCUSSION

Constructing a quadratic TFD from the *analytic associate* of a given real signal, rather than from the signal itself, avoids spurious terms caused by interference between positive-frequency and negative-frequency components.

Every TFD that we have derived heuristically from basic principles is *quadratic in the signal*; that is, if the signal is scaled by a factor k , the TFD is scaled by a factor k^2 . This is to be expected because

- (a) Each TFD considered so far is related to some sort of *energy density*. For example, the Page distribution is defined as a gradient of energy, hence accommodating for the negative values that occur in the TFD.
- (b) We have seen in [Section 2.7](#) that every TFD considered in this chapter can be written as the FT of a smoothed IAF (Eq. (2.7.1)), which is the convolution of an auxiliary function (the time-lag kernel filter $G(t, \tau)$) and the ordinary IAF $K_z(t, \tau)$, which in turn is *quadratic* in the signal.

Equation (2.7.7) effectively defines a TFD in terms of its time-lag kernel and results in a general expression of a TFD which is the key to the theory and design of quadratic TFDs, as will be detailed in the next chapter. [Section 2.7.6](#) will also be expanded in the next chapter to show the relationship between time-scale methods and quadratic TFDs.

2.9 ADDITIONAL EXERCISES

This section provides additional exercises that cover important topics from this chapter and prepares the reader for the next chapters in the book. It is recommended to use the (t, f) toolbox described in [Chapter 17](#) as an aid for doing these exercises. In a formal class setting, the instructor may wish to guide the student by adding intermediate questions with focus on one or two assessment criteria.

Exercise 2.9.1. Using [Fig. 2.1.1](#), p. 68, and Eq. (2.1.17), p. 69, implement the WVD.

1. Show the flowchart.
2. Show the code.
3. Validate the code on the signal defined as:

$$s(t) = \begin{cases} \cos(0.1\pi t) + \cos(2\pi(0.35t - 0.005t^2)), & 0 \leq t \leq 255 \\ 0, & \text{otherwise} \end{cases}$$

by comparing the theoretical and estimated WVD.

4. Calculate the IF of this signal using an IF estimation method of your choice.
5. Compare with the results given by the (t, f) toolbox described in [Chapter 17](#).

Exercise 2.9.2. Using Eq. (2.3.11), p. 77, implement the spectrogram $S_s^w(t, f)$.

1. Show the code.
2. Implement the spectrogram using Eq. (2.7.7), p. 86.
3. Validate the two methods on the signal given in [Exercise 2.9.1](#) and the test signal given in Eq. (1.1.3), p. 34.
4. Compare with the results given by the (t, f) toolbox described in [Chapter 17](#).

Exercise 2.9.3. Using Table 2.7.2, p. 92, do the same procedure as in Exercise 2.9.2 for all quadratics TFDs presented, and compare their performance using the same test signal given in Exercise 2.9.1 and the test signal given in Eq. (1.1.3), p. 34.

REFERENCES

- [1] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals”, *Proc. IEEE* 80 (4) (1992) 520-538.
- [2] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 540-568.
- [3] S.D. Conte, C. de Boor, Elementary Numerical Analysis: An Algorithmic Approach, third ed., McGraw-Hill, Tokyo, 1980.
- [4] E.P. Wigner, “On the quantum correction for thermodynamic equilibrium”, *Phys. Rev.* 40 (1932) 749-759.
- [5] B. Boashash, “Note on the use of the Wigner distribution for time-frequency signal analysis”, *IEEE Trans. Acoust. Speech Signal Process.* 36 (9) (1988) 1518-1521.
- [6] J. Ville, “Théorie et applications de la notion de signal analytique”, *Cables et Transmissions* 2A (1) (1948) 61-74, in French. English translation: I. Selin, Theory and applications of the notion of complex signal, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).
- [7] B. Bouachache, “Representation temps-frequence”, Thesis for diplome de docteur ingenieur, Institut National Polytechnique de Grenoble, France, 1982, 164 pp. (in French).
- [8] B. Bouachache, “Note d’information sur la representation des signaux dans le domaine temps-frequence”, Tech. Rep. 135/81, ELF-Aquitaine; 1981.
- [9] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), Advances in Spectrum Analysis and Array Processing, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [10] B. Boashash, P.J. Black, “An efficient real-time implementation of the Wigner-Ville distribution”, *IEEE Trans. Acoust. Speech Signal Process.* 35 (1987) 1611-1618.
- [11] B. Boashash, H.J. Whitehouse, “High resolution Wigner-Ville analysis”, in: Eleventh GRETSI Symp. on Signal Processing and Its Applications, 1987, pp. 205-208.
- [12] H.J. Whitehouse, B. Boashash, J.M. Speiser, “High-resolution processing techniques for temporal and spatial signals”, in: M. Bouvet, G. Bienvenu (Eds.), High-Resolution Methods in Underwater Acoustics, Springer, Berlin, 1991, pp. 127-176 (Chapter 4).
- [13] W. Martin, “Time-frequency analysis of random signals”, in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP’82), vol. 3, 1982, pp. 1325-1328.
- [14] L.B. White, B. Boashash, “Cross spectral analysis of nonstationary processes”, *IEEE Trans. Inf. Theory* 36 (4) (1990) 830-835.
- [15] S. Nawab, T. Quatieri, J. Lim, “Signal reconstruction from short-time Fourier transform magnitude”, *IEEE Trans. Acoust. Speech Signal Process.* 31 (4) (1983) 986-998, ISSN 0096-3518, <http://dx.doi.org/10.1109/TASSP.1983.1164162>.
- [16] D. Gabor, “Theory of communication”, *J. IEE* 93(III) (26) (1946) 429-457.
- [17] G.R. Putland, B. Boashash, “Can a signal be both monocomponent and multicomponent?”, in: Third Australasian Workshop on Signal Processing Applications (WoSPA 2000), 2000, Paper No. 32.
- [18] W. Koenig, H.K. Dunn, L.Y. Lacy, “The sound spectrograph”, *J. Acoust. Soc. Am.* 18 (1) (1946) 19-49.
- [19] J. Imberger, B. Boashash, “Application of the Wigner-Ville distribution to temperature gradient microstructure: a new technique to study small-scale variations”, *J. Phys. Oceanogr.* 16 (12) (1986) 1997-2012.
- [20] L.R.O. Storey, “An investigation of whistling atmospherics”, *Phil. Trans. R. Soc.* A246 (1953) 113-141.

- [21] D.L. Jones, T.W. Parks, "A high-resolution data-adaptive time-frequency representation", *IEEE Trans. Acoust. Speech Signal Process.* 38 (12) (1990) 2127-2135.
- [22] M.K. Emresoy, A. El-Jaroudi, "Iterative instantaneous frequency estimation and adaptive matched spectrogram", *Signal Process.* 64 (2) (1998) 157-165.
- [23] S. Qian, D. Chen, "Joint time-frequency analysis", *IEEE Signal Process. Mag.* 16 (2) (1999) 52-65.
- [24] J. Kay, R. Lerner, *Lectures in Communications Theory*, McGraw-Hill, New York, 1961.
- [25] C.W. Helstrom, "An expansion of a signal in Gaussian elementary signals", *IEEE Trans. Inf. Theory* 12 (1) (1966) 81-82.
- [26] I. Daubechies, "The wavelet transform: a method for time-frequency localization", in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 366-417 (Chapter 8).
- [27] Y. Meyer, *Wavelets: Algorithms and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1993 (translated and revised by Robert D. Ryan).
- [28] S.G. Mallat, *A Wavelet Tour of Signal Processing*, second ed., Academic Press, San Diego/London, 1999.
- [29] C.H. Page, "Instantaneous power spectra", *J. Appl. Phys.* 23 (1) (1952) 103-106.
- [30] A.W. Rihaczek, "Signal energy distribution in time and frequency", *IEEE Trans. Inf. Theory* 14 (3) (1968) 369-374.
- [31] J.G. Kirkwood, "Quantum statistics of almost classical ensembles", *Phys. Rev.* 44 (1933) 31-37.
- [32] M.J. Levin, "Instantaneous spectra and ambiguity functions", *IEEE Trans. Inf. Theory* 10 (1) (1964) 95-97.
- [33] H. Margenau, R.N. Hill, "Correlation between measurements in quantum theory", *Prog. Theor. Phys.* 26 (1961) 722-738.
- [34] L. Cohen, "Introduction: a primer on time-frequency analysis", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 3-42 (Chapter 1).
- [35] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, "The Wigner distribution—a tool for time-frequency signal analysis; part 3: relations with other time-frequency signal transformations", *Philips J. Res.* 35 (6) (1980) 372-389.
- [36] H.-I. Choi, W.J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE Trans. Acoust. Speech Signal Process.* 37 (6) (1989) 862-871.
- [37] Y. Zhao, L.E. Atlas, R.J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals", *IEEE Trans. Acoust. Speech Signal Process.* 38 (7) (1990) 1084-1091.
- [38] M. Born, P. Jordan, "Zur Quantenmechanik", *Zeitschrift für Physik*. 34 (1925) 858-888.
- [39] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [40] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [41] O. Rioul, M. Vetterli, "Wavelets and signal processing", *IEEE Signal Process. Mag.* 8 (1991) 14-38.
- [42] I. Daubechies, *Ten Lectures on Wavelets*, second ed., Society for Industrial and Applied Mathematics, Philadelphia, 1992, <http://dx.doi.org/10.1137/1.9781611970104>, <http://epubs.siam.org/doi/pdf/10.1137/1.9781611970104>, URL <http://epubs.siam.org/doi/abs/10.1137/1.9781611970104>.
- [43] R. Coifman, M. Wickerhauser, "Entropy-based algorithms for best basis selection", *IEEE Trans. Inf. Theory* 38 (2) (1992) 713-718, ISSN 0018-9448, <http://dx.doi.org/10.1109/18.119732>.
- [44] S. Kasaei, M. Deriche, B. Boashash, "A novel fingerprint image compression technique using wavelets packets and pyramid lattice vector quantization", *IEEE Trans. Image Process.* 11 (12) (2002) 1365-1378, ISSN 1057-7149, <http://dx.doi.org/10.1109/TIP.2002.802534>.

- [45] S. Kasaei, M. Deriche, B. Boashash, “An efficient quantization technique for wavelet coefficients of finger-print images”, *Signal Process.* 62 (3) (1997) 361-366, ISSN 0165-1684, [http://dx.doi.org/10.1016/S0165-1684\(97\)00169-2](http://dx.doi.org/10.1016/S0165-1684(97)00169-2), URL <http://www.sciencedirect.com/science/article/pii/S0165168497001692>.
- [46] M. Farge, “Wavelet transforms and their applications to turbulence”, *Ann. Rev. Fluid Mech.* 24 (1) (1992) 395-458.
- [47] P. Bentley, J. McDonnell, “Wavelet transforms: an introduction”, *Electron. Commun. Eng. J.* 6 (4) (1994) 175-186.
- [48] A. Cohen, J. Kovacevic, “Wavelets: the mathematical background”, *Proc. IEEE* 84 (4) (1996) 514-522, ISSN 0018-9219, <http://dx.doi.org/10.1109/5.488697>.
- [49] R.G. Stockwell, L. Mansinha, R.P. Lowe, “Localization of the complex spectrum: the S transform”, *IEEE Trans. Signal Process.* 44 (4) (1996) 998-1001.
- [50] S. Stanković, “Time-frequency analysis and its application in digital watermarking”, *EURASIP J. Adv. Signal Process.* 2010 (2010) 5:1-5:20, ISSN 1110-8657, <http://dx.doi.org/10.1155/2010/579295>.
- [51] I. Djurović, E. Sejdić, J. Jiang, “Frequency-based window width optimization for S-transform”, *Int. J. Electron. Commun.* 62 (4) (2008) 245-250.
- [52] S. Ventosa, C. Simon, M. Schimmel, J.J. Da nobeitia, A. Mànuel, “The S-transform from a wavelet point of view”, *IEEE Trans. Signal Process.* 56 (7) (2008) 2771-2780.
- [53] S. Assous, B. Boashash, “Evaluation of the modified S-transform for time-frequency synchrony analysis and source localisation”, *EURASIP J. Adv. Signal Process.* 2012 (2012) <http://dx.doi.org/10.1186/1687-6180-2012-49>, Article ID 49, 18 pp.
- [54] R. Wilson, A.D. Calway, E.R.S. Pearson, “A generalized wavelet transform for Fourier analysis: the multiresolution Fourier transform and its application to image and audio signal analysis”, *IEEE Trans. Inf. Theory* 38 (2) (1992) 674-690, ISSN 0018-9448, <http://dx.doi.org/10.1109/18.119730>.
- [55] E. Sejdić, I. Djurović, J. Jiang, “Time-frequency feature representation using energy concentration: an overview of recent advances”, *Digital Signal Process.* 19 (1) (2009) 153-183.
- [56] P.K. Dash, M.V. Chilukuri, “Hybrid S-transform and Kalman filtering approach for detection and measurement of short duration disturbances in power networks”, *IEEE Trans. Instrum. Meas.* 53 (2) (2004) 588-596, ISSN 0018-9456.
- [57] C. Herley, J. Kovacevic, K. Ramchandran, M. Vetterli, “Tilings of the time-frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms”, *IEEE Trans. Signal Process.* 41 (12) (1993) 3341-3359, ISSN 1053-587X.
- [58] L. Durak, O. Arikan, “Short-time Fourier transform: two fundamental properties and an optimal implementation”, *IEEE Trans. Signal Process.* 51 (5) (2003) 1231-1242, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2003.810293>.
- [59] N. Hess-Nielsen, M.V. Wickerhauser, “Wavelets and time-frequency analysis”, *Proc. IEEE* 84 (4) (1996) 523-540.

THEORY AND DESIGN OF HIGH-RESOLUTION QUADRATIC TFDs⁰

3

INTRODUCTION AND OVERVIEW

The quadratic time-frequency distributions (QTFDs) introduced in the previous chapter represent the majority of traditional time-frequency $((t,f))$ methods used in practical applications that deal with nonstationary signals. This key chapter completes the introductory tutorial and prepares the reader for the more advanced material presented in the following chapters. The core material shows that the particular QTFDs introduced in Chapter 2 actually belong to a general quadratic class of TFDs whose design follows a common procedure, and whose properties are governed by common laws. This quadratic (or bilinear) class may be considered as the class of “smoothed” or filtered Wigner-Ville distributions (WVDs). The “smoothing” (or “filtering”) is described in the (t,f) domain by a 2D convolution with a “ (t,f) kernel” $\gamma(t,f)$ and in other domains by multiplication and/or 1D convolution with various transforms of $\gamma(t,f)$ such as kernel filter $g(\nu,\tau)$. The generalized approach allows the definition of new TFDs that are better adapted to particular signal types, using a simple and systematic procedure [1].

Section 3.1 reviews in detail the key properties and limitations of the WVD to motivate the introduction of general QTFDs (Section 3.2) and discuss their properties (Section 3.3). In Section 3.2, using Fourier transforms (FTs) from lag to frequency and from time to Doppler (frequency shift), the QTFDs and their kernels are formulated in four different but related 2D domains. This leads to the definition of the “ambiguity function” and allows the smoothing of the WVD to be understood as a filtering operation. Section 3.3 lists some desirable TFD properties not shared by the WVD. The various TFD properties are then expressed in terms of constraints on the kernel, so that TFD design is reduced to kernel design. Three tables are provided showing the kernel properties equivalent to various TFD properties in the various 2D domains and the properties of those same TFDs. Three state-of-the-art high-resolution TFDs are then discussed. Section 3.4 deals with the question of TFD positivity, the corresponding condition on the kernel filter, and its implications. Section 3.5 concludes the tutorial.

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa).

3.1 THE WIGNER-VILLE DISTRIBUTION

3.1.1 PROPERTIES OF THE WVD

We have seen that the WVD has the simplest time-lag kernel (see Eq. 2.7.9), and that each of the other TFDs can be written as a filtered WVD using a specific time-lag kernel filter (see Eq. 2.7.36). In this sense, the WVD is the basic or prototype TFD and the other TFDs are variations thereon [2]. Moreover, we see in Fig. 2.7.1 that the WVD gives the sharpest indication of the IF law of a linear-frequency modulation (LFM) signal. These are some of the reasons why the WVD is the most widely studied TFD and deserves a further detailed description of its properties, as listed below.

- *Realness (RE)*: $W_z(t,f)$ is real for all z , t , and f .
- *Time-shift invariance*, also called *time covariance*: A time shift in the signal causes the same time shift in the WVD; that is, if

$$z_r(t) = z(t - t_0), \quad (3.1.1)$$

then

$$W_{z_r}(t,f) = W_z(t - t_0, f). \quad (3.1.2)$$

- *Frequency-shift invariance*, also called *frequency covariance*: A frequency shift in the signal causes the same frequency shift in the WVD; that is, if

$$z_r(t) = z(t) e^{j2\pi f_0 t}, \quad (3.1.3)$$

then

$$W_{z_r}(t,f) = W_z(t, f - f_0). \quad (3.1.4)$$

- *Time marginal (TM)*: Integration of the WVD over frequency gives the instantaneous power:

$$\int_{-\infty}^{\infty} W_z(t,f) df = |z(t)|^2. \quad (3.1.5)$$

- *Frequency marginal (FM)*: Integration of the WVD over time gives the energy spectrum:

$$\int_{-\infty}^{\infty} W_z(t,f) dt = |Z(f)|^2. \quad (3.1.6)$$

- *Global energy*: Integration of the WVD over the entire (t,f) plane yields the signal energy E_z :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t,f) dt df = E_z. \quad (3.1.7)$$

- *Instantaneous frequency (IF)*: For a monocomponent analytic signal, the first moment (i.e., the mean) of the WVD w.r.t. frequency is the IF:

$$\frac{\int_{-\infty}^{\infty} f W_z(t,f) df}{\int_{-\infty}^{\infty} W_z(t,f) df} = \frac{1}{2\pi} \frac{d}{dt} [\arg z(t)]. \quad (3.1.8)$$

- *Spectral delay (SD)*: The first moment of the WVD w.r.t. time is the SD:

$$\frac{\int_{-\infty}^{\infty} tW_z(t,f) dt}{\int_{-\infty}^{\infty} W_z(t,f) dt} = -\frac{1}{2\pi} \frac{d}{df} [\arg Z(f)], \quad (3.1.9)$$

where $Z(f)$ is the FT of $z(t)$.

- *Time support (TS)*: The time support of $W_z(t,f)$ is limited by the duration of $z(t)$; that is, if $z(t) = 0$ for $t < t_1$ and for $t > t_2$, then $W_z(t,f) = 0$ for $t < t_1$ and for $t > t_2$.
- *Frequency support (FS)*: The frequency support of $W_z(t,f)$ is limited by the bandwidth of $z(t)$; that is, if $\mathcal{F}\{z(t)\} = Z(f) = 0$ for $f < f_1$ and for $f > f_2$, then $W_z(t,f) = 0$ for $f < f_1$ and for $f > f_2$.
- *Convolution invariance*: The WVD of the time-convolution of two signals is the time-convolution of the WVDs of the two signals; that is, if

$$z_3(t) = \underset{t}{\ast} z_1(t) z_2(t), \quad (3.1.10)$$

where \ast is the time convolution; then

$$W_{z_3}(t,f) = W_{z_1}(t,f) \underset{t}{\ast} W_{z_2}(t,f). \quad (3.1.11)$$

- *Modulation invariance*: The WVD of the frequency-convolution of two signals is the frequency-convolution of the WVDs of the two signals; that is, if $z_3(t) = z_1(t) z_2(t)$, then $Z_3(f) = Z_1(f) \underset{f}{\ast} Z_2(f)$, and it follows that

$$W_{z_3}(t,f) = W_{z_1}(t,f) \underset{f}{\ast} W_{z_2}(t,f), \quad (3.1.12)$$

where \ast is the frequency convolution.

- *Invertibility*: If $W_z(t,f)$ is the known WVD of an unknown signal $z(t)$, we may need to find $z(t)$ explicitly; it may be shown [3, pp. 223-224] that

$$\int_{-\infty}^{\infty} W_z(t/2,f) e^{j2\pi ft} df = z(t) z^*(0). \quad (3.1.13)$$

Putting $t = 0$ in this result yields the magnitude, but not the phase, of $z(0)$. Hence, a signal may be recovered from its WVD up to a constant phase factor.

- *Inner-product invariance*: The WVD is a unitary transformation; that is, it preserves inner products (also called scalar products) [4,5]

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{z_1}(t,f) W_{z_2}(t,f) dt df = \left| \int_{-\infty}^{\infty} z_1(t) z_2^*(t) dt \right|^2, \quad (3.1.14)$$

which also equals the inner product of the FTs (by Parseval's theorem).

- *Ambiguity function*: The WVD is the 2D FT of the ambiguity function (AF), denoted by $A_z(\nu, \tau)$. More precisely,

$$W_z(t,f) = \mathcal{F}_{\tau \rightarrow f} \left\{ \mathcal{F}_{t \leftarrow \nu}^{-1} \{ A_z(\nu, \tau) \} \right\}. \quad (3.1.15)$$

This follows directly from Eqs. (2.7.38) and (2.7.39).

The above properties are not independent; for example, it is a trivial exercise to show that either of the marginals (TM or FM) implies the global energy condition.

A more comprehensive list of the WVD properties is given in [3,6]. These either follow directly from the definition or are proven in [3,7].

A notable omission from the above list is *positivity*; the WVD can assume negative values for almost every signal. This is easily explained when related to the Page distribution, which is an energy spectrum gradient (see Sections 2.5.1 and 3.4).

Applicability: Many properties of the WVD are desirable in applications; for example, realness is convenient for simplification and interpretation as energy gradient, while convolution invariance and modulation invariance make the WVD partly compatible with linear filtering theory. Nevertheless, the WVD was not applied to real-life problems until the late 1970s, when it was implemented for the purpose of processing LFM signals used in seismic prospecting [8,9], as well as the analysis of loudspeakers signals [2,10].

Exercise 3.1.1. Derive the proofs for the above properties.

3.1.2 LIMITATIONS OF THE WVD

Despite its many desirable properties, the WVD has some drawbacks. It may assume large negative values. Further, it is *bilinear* (i.e., quadratic) in the signal rather than linear; hence, it exhibits some spurious features called *artifacts* or *cross-terms*, which appear midway between true signal components in the case of multicomponent signals as well as nonlinear mono- and multicomponent FM signals (see Section 4.2). Such cross-terms may be useful in some applications like classification [11] but are undesirable in other applications, including analysis and interpretation as well as multicomponent IF estimation.

3.1.2.1 Nonlinear monocomponent FM signals and inner artifacts

Consider a monocomponent signal $z(t) = a(t) e^{j\phi(t)}$. In the case of a *linear-FM* signal, the WVD gives an accurate representation of the IF law (Fig. 1.1.7) because the CFD approximation to $\phi'(t)$ is exact (see Eq. 2.1.7), so that the signal kernel is a dechirped function of τ (see Eqs. 2.1.19–2.1.25). In the case of a *nonlinear* FM signal, the CFD approximation is *not* exact and the signal kernel (IAF) is *not* dechirped. This results in the formation of *inner artifacts* “within” a single component [12].

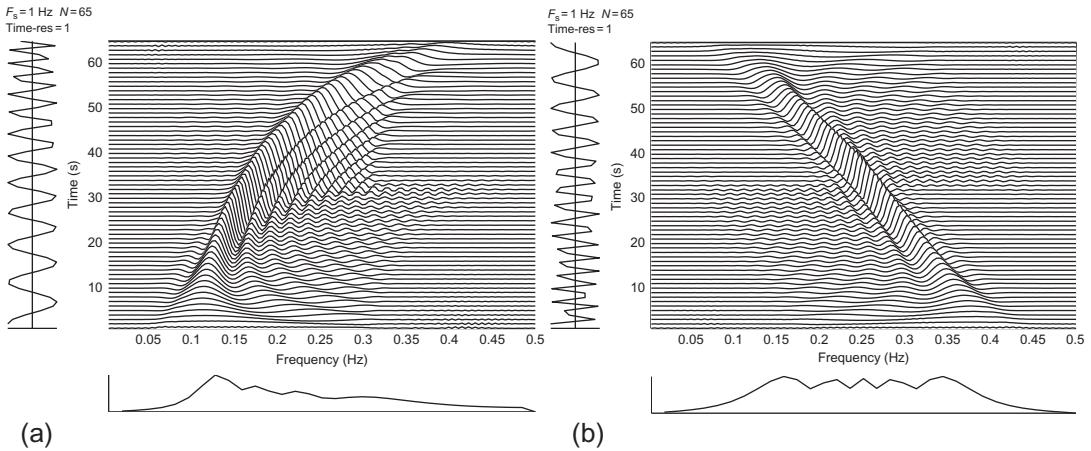
An example of a nonlinear FM signal is the *hyperbolic FM* signal with rectangular amplitude $a(t)$ and the phase given by

$$\phi(t) = \frac{2\pi f_0}{\alpha} \ln(1 + \alpha t). \quad (3.1.16)$$

The IF is

$$f_i(t) = \frac{\phi'(t)}{2\pi} = \frac{f_0}{1 + \alpha t}. \quad (3.1.17)$$

For the time interval $0 \leq t \leq T$, the starting frequency is $f_0 = f_i(0)$. The finishing frequency is $f_{\max} = f_i(T) = f_0/(1 + \alpha T)$.

**FIGURE 3.1.1**

WVD of (a) hyperbolic FM signal with starting frequency 0.1 Hz and finishing frequency 0.4 Hz; (b) LFM signal with starting frequency 0.4 Hz and finishing frequency 0.1 Hz. Both plots are for a duration of 65 samples (sampling rate 1 Hz); see [Chapter 17](#) to reproduce these plots.

[Figure 3.1.1\(a\)](#) shows the WVD of a hyperbolic FM signal with duration 65 samples, $f_0 = 0.1$ and $f_{\max} = 0.4$. While the crest of the WVD seems to be a reasonable approximation to the IF law, the energy concentration is poorer than expected. The many spurious ridges are the inner artifacts. They oscillate and alternate in sign as we move *normal* to the IF law in the (t, f) plane; this is a characteristic feature of inner artifacts (see [Sections 4.2](#) and [5.9](#)). For comparison, [Fig. 3.1.1\(b\)](#) shows the WVD of the LFM signal with the same duration and the same frequency limits (with falling frequency). Note the superior energy concentration for the LFM case and the more attenuated artifacts distributed on both sides of the main ridge.

Artifacts caused by nonlinear FM laws can be reduced by windowing the IAF in τ before taking the FT, leading to the *windowed WVD*. This procedure, however, causes a loss of frequency resolution [6].

3.1.2.2 Multicomponent signals and outer artifacts

If $z(t)$ is a multicomponent signal, the algebraic expansion of $K_z(t, \tau)$ contains cross-product terms which, when Fourier-transformed, give rise to spurious features in the WVD that oscillate and alternate in sign as shown in Eq. (2.1.16). These are the *outer artifacts* or *cross-terms*. (The word “cross-terms,” when used without qualification, usually refer to *outer* artifacts). To explain these, consider the signal

$$z(t) = z_1(t) + z_2(t), \quad (3.1.18)$$

where $z(t)$, $z_1(t)$, and $z_2(t)$ are analytic. Expanding the IAF, we obtain

$$K_z(t, \tau) = K_{z_1}(t, \tau) + K_{z_2}(t, \tau) + K_{z_1 z_2}(t, \tau) + K_{z_2 z_1}(t, \tau), \quad (3.1.19)$$

where $K_{z_1 z_2}(t, \tau)$ and $K_{z_2 z_1}(t, \tau)$ are the “signal cross-kernels” or instantaneous cross-correlation functions (e.g., $K_{z_1 z_2}(t, \tau) = z_1(t + \frac{\tau}{2}) z_2^*(t - \frac{\tau}{2})$). Taking FTs of Eq. (3.1.19) w.r.t. τ (using Eq. (2.1.18)), we find

$$W_z(t, f) = W_{z_1}(t, f) + W_{z_2}(t, f) + 2\operatorname{Re}\{W_{z_1 z_2}(t, f)\}, \quad (3.1.20)$$

where $W_{z_1}(t, f)$ and $W_{z_2}(t, f)$ are the WVDs of $z_1(t)$ and $z_2(t)$, and $W_{z_1 z_2}(t, f)$ is the *cross-Wigner-Ville distribution* (XWVD) of $z_1(t)$ and $z_2(t)$, defined by

$$W_{z_1 z_2}(t, f) = \int_{-\infty}^{\infty} z_1\left(t + \frac{\tau}{2}\right) z_2^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau. \quad (3.1.21)$$

So the WVD of the sum of two signals is not the sum of their WVDs but the sum of their WVDs and XWVDs. If $z_1(t)$ and $z_2(t)$ are monocomponent signals, $W_{z_1}(t, f)$ and $W_{z_2}(t, f)$ are auto-terms, and $2\operatorname{Re}\{W_{z_1 z_2}(t, f)\}$ is a cross-term.

[Figure 3.1.2](#) shows the WVD of the sum of two parallel LFM signals. There seem to be three components rather than two; the “extra” component at the mean frequency of the expected components has a large oscillating amplitude and occurs in a region of the (t, f) plane where we expect no energy at all. [Figure 3.1.3\(a\)](#) shows the WVD of the sum of two FM signals crossing over in frequency. A large

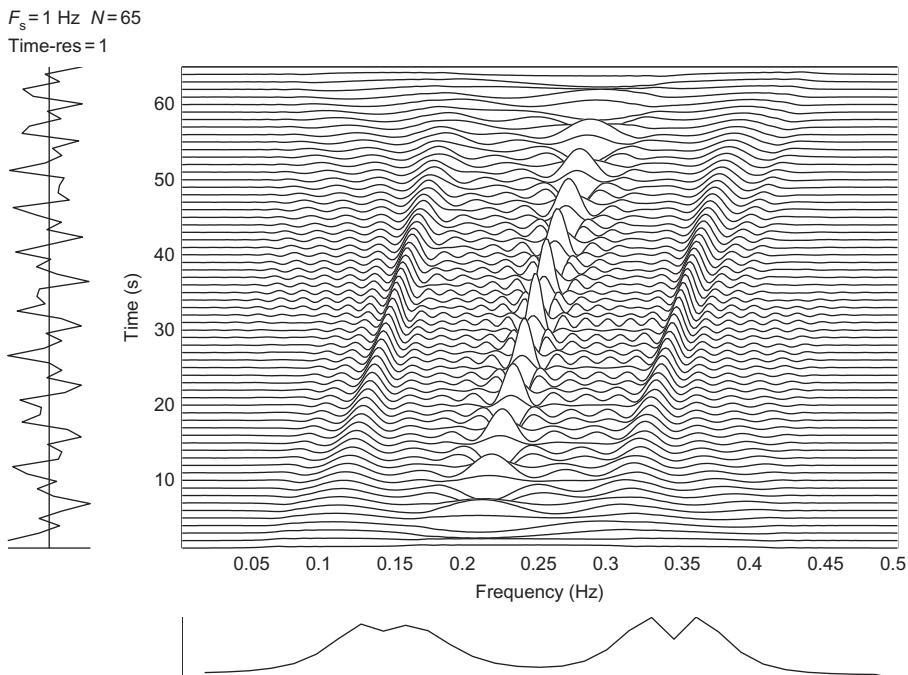
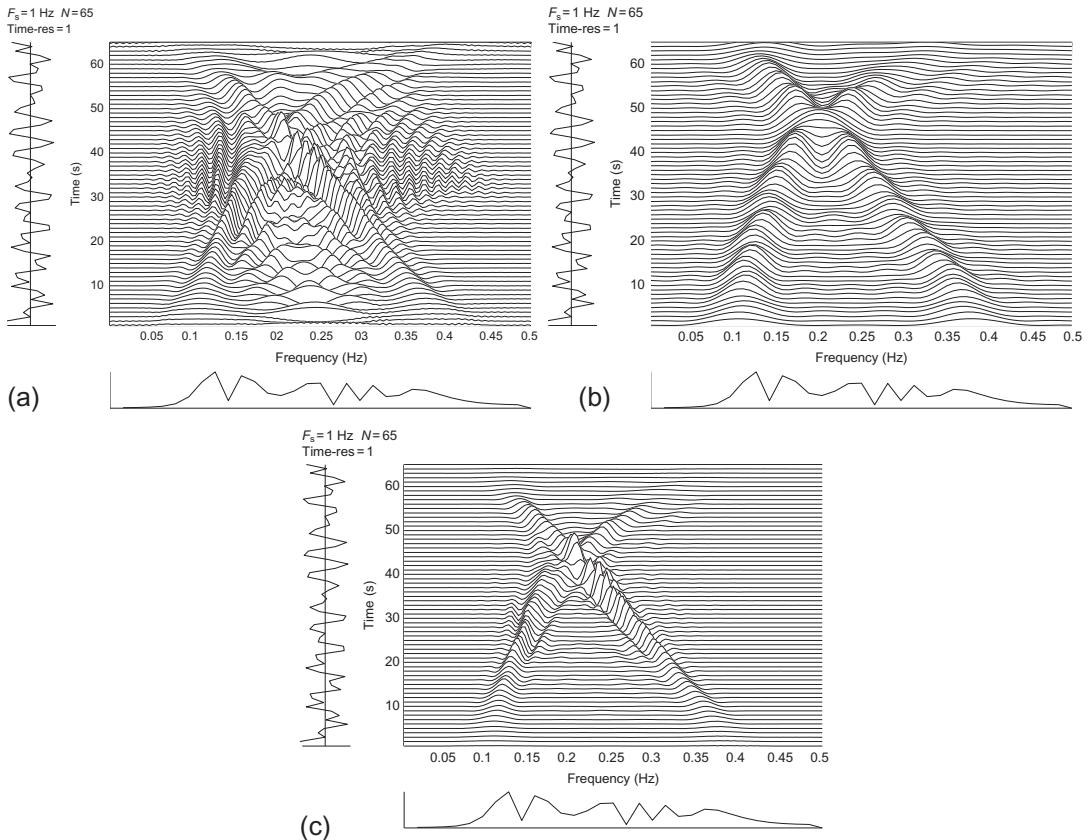


FIGURE 3.1.2

The WVD of the sum of two LFM signals with frequency ranges 0.1–0.2 and 0.3–0.4 Hz, unit amplitudes, and a duration of 65 samples (sampling rate 1 Hz).

**FIGURE 3.1.3**

Suppression of cross-terms in the sum of a rising hyperbolic FM signal (frequency range 0.1-0.4 Hz) and a falling LFM signal (frequency range 0.4-0.1 Hz), with unit amplitudes and a duration of 65 samples (sampling rate 1 Hz): (a) WVD; (b) spectrogram with 21-point rectangular window; (c) masked WVD, being the product of (a) and (b).

number of undulations appear in addition to the two main ridges. In both examples, the cross-terms oscillate and alternate in sign as we move *parallel* to the expected features (e.g., FM law) in the (t,f) plane; this is a characteristic feature of cross-terms (see more details in [Sections 4.2](#) and [5.9](#)).

Note that following the above approach summarized by Eqs. [\(3.1.20\)](#) and [\(3.1.21\)](#), one can also define the cross-TFD $\rho_{z_1 z_2}(t,f)$ by extending Eq. [\(2.7.36\)](#) and then replacing the WVD by the XWVD in these expressions, that is, as follows:

$$\rho_{z_1 z_2}(t,f) = \gamma(t,f) *_{(t,f)}^* W_{z_1 z_2}(t,f); \quad \gamma(t,f) = \mathcal{F}_{\tau \rightarrow f} \{G(t,\tau)\}. \quad (3.1.22)$$

This concept is used in later chapters (e.g., [Sections 6.1, 6.5, 8.4](#), and [12.5](#)).

3.1.2.3 Suppression of cross-terms

Cross-terms can make the WVD difficult to analyze and interpret, especially if the components are numerous, close to each other or when cross-terms overlap with auto-terms. This is especially so in the presence of noise, as cross-terms between signal components and noise exaggerate the effects of noise and cause rapid degradation of performance as the SNR decreases. For such reasons, cross-terms are often regarded as the fundamental limitation on the applicability of QTFDs, and the desire to suppress them has led to several strategies. For example:

1. If we multiply the WVD by the spectrogram, we obtain the so-called *masked WVD*, which combines the cross-term suppression of the spectrogram with the high resolution of the WVD. To illustrate the effect, Fig. 3.1.3 shows the WVD, spectrogram, and masked WVD of the sum of two crossed FM signals; observe that the masked WVD is “cleaner” than the WVD and has better resolution than the spectrogram.
2. Equation (3.1.21) indicates that the XWVD of $z_1(t)$ and $z_2(t)$ is linear in $z_2(t)$. If $z_1(t)$ is a reference signal and $z_2(t)$ is the signal under analysis, there will be no cross-terms between components of $z_2(t)$. This observation inspired efforts to displace the WVD with the XWVD in relevant areas of application. A reference signal is readily available in some detection applications [13], including sonar, radar [14], and seismic exploration [15]. In applications where reference signals are not available, a filtered version of the observed signal can be used as a reference. The filtering procedure may use the IF as a critical feature of the signal; for example, the data-dependent TFDs defined in [16] use, as a reference signal, the signal component that maximizes the energy concentration around the IF law in the TFD.
3. All QTFDs in Chapter 2, including the spectrogram which suppresses cross-terms, can be identified by a time-lag kernel (Eqs. 2.7.36 and 2.7.37). By analyzing the properties of time-lag kernels and/or Doppler-lag kernels (see below), we can define and design TFDs that attenuate cross-terms in the desired way. A QTFD in which the cross-terms are attenuated relative to the auto-terms is often called a *reduced-interference distribution (RID)*.
4. In addition to the above points, there are other strategies to remove cross-terms. Some (t,f) methods attenuate cross-terms by first extracting the signal individual components using methods such as time-varying filtering or the empirical mode decomposition (see Sections 4.1 and 16.5); other methods include matching pursuit algorithms (see Section 12.2) or blind source separation (BSS) algorithms (see Chapter 8). Then each component may be analyzed separately using the WVD or a suitable high-resolution TFD.

This chapter concentrates on the RID approach for defining QTFDs with the property of cross-terms suppression or reduction.

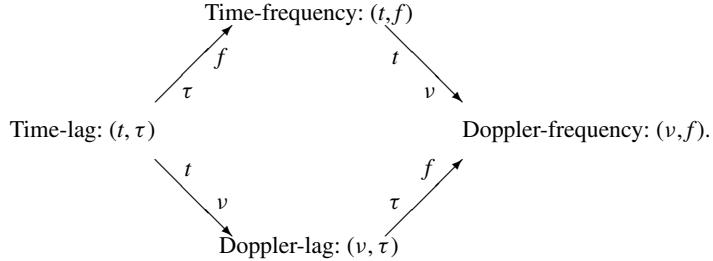
3.2 FORMULATIONS OF QUADRATIC TFDs (QTFDs)

3.2.1 TIME-LAG, DOPPLER-FREQUENCY, AND OTHER FORMULATIONS

Given an analytic signal $z(t)$, the IAF is defined in Eq. (2.1.8) as

$$K_z(t, \tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right). \quad (3.2.1)$$

The IAF $K_z(t, \tau)$ is a function of two time variables: the actual time t and the time lag τ . By taking the dual domains of t and τ in frequency, we obtain two frequency variables: the frequency shift (Doppler) v and the actual frequency f . This allows for four possible domains of representation, as illustrated below [17].



Starting with the IAF, we define the WVD by taking the FT ($\tau \rightarrow f$):

$$W_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{K_z(t, \tau)\}. \quad (3.2.2)$$

The FT ($t \rightarrow v$) of the WVD defines the spectral correlation function (SCF) (Eq. 2.1.26) as

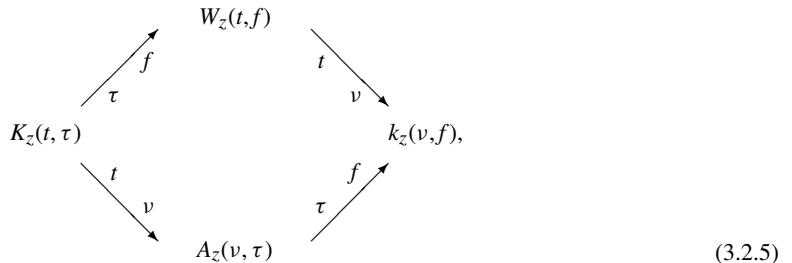
$$k_z(v, f) = \mathcal{F}_{t \rightarrow v} \{W_z(t, f)\}. \quad (3.2.3)$$

The FT ($t \rightarrow v$) of the IAF $K_z(t, \tau)$ equals the IFT ($\tau \leftarrow f$) of the SCF $k_z(v, f)$ and defines the symmetrical *ambiguity function* (AF) as

$$A_z(v, \tau) = \mathcal{F}_{t \rightarrow v} \{K_z(t, \tau)\} = \mathcal{F}_{\tau \leftarrow f}^{-1} \{k_z(v, f)\}, \quad (3.2.4)$$

where $\mathcal{F}\{\cdot\}$ represents the forward (direct) FT and $\mathcal{F}^{-1}\{\cdot\}$ its inverse.

If we represent Fourier transformations by arrows labeled with the participating variables, Eqs. (3.2.2)–(3.2.4) may be combined into the single graphical equation



which we shall call a *diamond diagram*.¹ This single graphical representation of several equations is very useful in that it links several known methods by simple FT operations. The knowledge of FT properties then allows us to use characteristics of a method in one domain and transfer them to another domain. This also relates to radar methods used in the (ν, τ) Doppler-lag domain (see [Section 5.1](#)) to (t, f) and cyclostationary methods used in the Doppler-frequency (ν, f) domain (see [Section 9.6](#)).

Now let us represent similarly the TFD kernel in these four domains by taking various FTs of the *time-lag kernel* $G(t, \tau)$:

```

    graph TD
      G[G(t, tau)] -->|f| gamma[gamma(t, f)]
      G -->|tau| g[g(v, tau)]
      gamma -->|t| Gv[G(v, f)]
      gamma -->|nu| g
      g -->|nu| Gv
      g -->|tau| Gt[G(t, f)]
      Gt -->|f| gamma
  
```

(3.2.6)

We refer to $g(\nu, \tau)$ as the *Doppler-lag kernel*, to $\mathcal{G}(\nu, f)$ as the *Doppler-frequency kernel*, and to $\gamma(t, f)$ as the *time-frequency kernel*.

3.2.2 TIME-FREQUENCY FORMULATION

We have defined the smoothed IAF as

$$R_z(t, \tau) = G(t, \tau) *_{\tau} K_z(t, \tau), \quad (3.2.7)$$

where $G(t, \tau)$ is the time-lag kernel. Following Eq. (2.7.1), we define a class of *QTFDs* (which are smoothed WVDs) as

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{R_z(t, \tau)\}. \quad (3.2.8)$$

Substituting Eq. (3.2.7) into Eq. (3.2.8), writing out the convolution and the transform, and substituting from Eq. (3.2.1), we obtain

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t - u, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f \tau} du d\tau, \quad (3.2.9)$$

¹An alternative definition of $A_z(\nu, \tau)$ uses an *inverse* FT from t to ν in Eq. (3.2.4), so that the lower-left arrow in Eq. (3.2.5) and in the other “diamond diagrams” is reversed. If we adopt the “alternative” definition of $A_z(\nu, \tau)$ but continue to define $k_z(\nu, f)$ as $Z(f + \nu/2) Z^*(f - \nu/2)$ (see [Eq. 2.1.31](#), p. 72), then we no longer obtain a simple FT relationship between $A_z(\nu, \tau)$ and $k_z(\nu, f)$; that is, we no longer have a lower-right arrow. The convention adopted in this book is that transforms from a time variable (t or τ) to a frequency variable (ν or f) are always *forward*, with the result that all arrows in the diamond diagrams are present and point to the *right*.

which defines the general class of QTDFs in terms of the signal and the time-lag kernel. The set of TFDs of this form is called the *quadratic class*.²

Equation (3.2.8) is included in the graphical equation

$$\begin{array}{ccc} & \rho_z(t,f) & \\ \nearrow f & & \searrow t \\ R_z(t,\tau) & & r_z(v,f) \\ \downarrow t & & \uparrow v \\ & A_z(v,\tau) & \end{array} \quad (3.2.10)$$

which assigns symbols to the various FTs of the smoothed IAF [6, p. 436]. By analogy with Eqs. (3.2.3) and (3.2.4), we call $r_z(v,f)$ the “generalized SCF” and $A_z(v,\tau)$ the “generalized ambiguity function” (GAF)³ or more precisely *filtered ambiguity function*; note the distinction between the normal A in Eq. (3.2.5) and the calligraphic \mathcal{A} in Eq. (3.2.10).

Using Eq. (3.2.7) and the lower two arrows in Eq. (3.2.10), together with the convolution properties of the FT, we obtain in sequence

$$A_z(v,\tau) = g(v,\tau) A_z(v,\tau); \quad (3.2.11)$$

$$r_z(v,f) = \mathcal{G}(v,f) * k_z(v,f). \quad (3.2.12)$$

Then, by using Eq. (3.2.7) and the upper left arrow in Eq. (3.2.10) (or using Eq. (3.2.12) and the upper right arrow in Eq. (3.2.10)), together with the convolution properties, we obtain

$$\rho_z(t,f) = \gamma(t,f) \ast\ast_{(t,f)} W_z(t,f), \quad (3.2.13)$$

where the double asterisk denotes double convolution in t and f (cf. [6], pp. 437, 475).

In Eq. (3.2.11), multiplication by $g(v,\tau)$ may reduce the range of v or τ (or both) over which $A_z(v,\tau)$ is nonzero. In computation (with discrete time and frequency), this may reduce the fast Fourier transform length, hence the (t,f) resolution and the number of values of t or f for which $\rho_z(t,f)$ is computed, as detailed in Sections 6.1 and 6.6. But, even for continuous time and frequency, Eq. (3.2.13) implies that $\rho_z(t,f)$ will have lower (t,f) resolution than the WVD.

²The *quadratic class* as defined here satisfies the time-shift- and frequency-shift-invariance properties. Some authors seem to equate “*Cohen’s class*” with the quadratic class. In addition, some authors apply the term “quadratic time-frequency representation” or “QTFR” to all time-frequency representations that are quadratic in the signal (whether they have the form of Eq. (3.2.9) or not) so that “Cohen’s class” becomes a subset of the “QTFRs.” However, the class originally defined by Cohen in a quantum-mechanical context [18] differs from the quadratic class (Eq. 3.2.9) in that the marginal conditions must be satisfied, while the kernel may depend on the signal (in which case the TFD is *not* quadratic in the signal). We shall see (e.g., in Table 3.3.3) that many useful QTDFs do not satisfy the marginals, so that they are not members of “Cohen’s class” as originally defined. For these reasons, the terminology “quadratic class” is more appropriate.

³The term *generalized ambiguity function* is also used with a different meaning in connection with polynomial TFDs; see Section 5.5 and references therein.

Equation (3.2.13) defines the quadratic class of TFDs in terms of the WVD and the (t,f) kernel. For this reason and as mentioned in Section 3.1.1, we regard the WVD as the core QTFD, and all other QTFDs as filtered versions of the WVD. Equation (3.2.13) also suggests a method of suppressing cross-terms in the WVD. Because inner artifacts alternate in general (e.g., for signals with parabolic IF law) as we move in the frequency direction, they can be attenuated by choosing a $\gamma(t,f)$ with a sufficient spread in the frequency direction. Similarly, because outer artifacts (cross-terms) alternate in general (e.g., pure tones or components with IF parallel to the time axis) as we move in the time direction, they can be attenuated by choosing a $\gamma(t,f)$ with a sufficient spread in the time direction. More specifically, Eq. (3.2.11) is directly used in the design of the filter $\gamma(t,f)$ shown in Eq. (3.2.13), as the convolution in Eq. (3.2.13) is replaced by a product in Eq. (3.2.11), thereby simplifying filter design.

Equation (3.2.11) shows how one QTFD (described by A in the Doppler-lag domain) can be transformed to another (described by \mathcal{A}) by 2D filtering product; this is expressed in the (t,f) domain by a linear convolution operation.

3.2.3 DOPPLER-LAG FORMULATION AND TFD DESIGN

Substituting for $K_z(t, \tau)$ in Eq. (3.2.4) and writing out the transform, we obtain

$$A_z(v, \tau) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi v t} dt \quad (3.2.14)$$

$$= \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) \left[z\left(t - \frac{\tau}{2}\right) e^{j2\pi v t} \right]^* dt. \quad (3.2.15)$$

The expression in square brackets can be obtained by delaying $z(t + \frac{\tau}{2})$ in time by τ and shifting it in frequency by v , indicating that $A_z(v, \tau)$ is the correlation of a signal $z(t)$ with a time-delayed and frequency-shifted version of itself. This correlation is well known in radar theory as the Sussman ambiguity function [19]; the name ‘‘ambiguity’’ arises from the equivalence between time-shifting and frequency-shifting for LFM signals, which are frequently used in radar [20,21]. Hence, the Doppler-lag (v, τ) domain is also called the *ambiguity domain*. This can also be extended by defining the cross ambiguity function of two signals $z_1(t)$ and $z_2(t)$ as

$$A_{z_1 z_2}(v, \tau) = \int_{-\infty}^{\infty} z_1\left(t + \frac{\tau}{2}\right) z_2^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi v t} dt. \quad (3.2.16)$$

Equations (3.2.10) and (3.2.11) indicate that the QTFD is a 2D FT (half inverse, half forward) of the filtered ambiguity function; this is expressed as

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(v, \tau) A_z(v, \tau) e^{j2\pi(v t - f \tau)} dv d\tau. \quad (3.2.17)$$

Then writing the dummy u for t in Eq. (3.2.14) and substituting the result into Eq. (3.2.17) gives

$$\rho_z(t, f) = \iiint g(v, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{j2\pi(v t - v u - f \tau)} du dv d\tau, \quad (3.2.18)$$

where the integrals are from $-\infty$ to ∞ ; this defines the QTFD in terms of the Doppler-lag kernel $g(v, \tau)$.

From the lower left side of Eq. (3.2.6), the relationship between the time-lag kernel and the Doppler-lag kernel is

$$g(v, \tau) = \mathcal{F}_{t \rightarrow v} \{G(t, \tau)\}. \quad (3.2.19)$$

Using this equation, the time-lag kernels determined in Section 2.7 may be converted to Doppler-lag form. For example, using Eq. (2.7.9), we find that the Doppler-lag kernel for the WVD is $g(v, \tau) = \mathcal{F}_{t \rightarrow v} \{\delta(t)\} = 1$.

Equation (3.2.13) and its 2D FT, Eq. (3.2.11) express that a QTFD can be designed using basic filter design principles. As in 1D filter design, the filter specifications are best given in the domain where the filtering operation is expressed as a multiplication as in Eq. (3.2.11), rather than in the dual domain with a convolution as in Eq. (3.2.13). If the Doppler-lag kernel $g(v, \tau)$ has the separable form $G_1(v)g_2(\tau)$, multiplication by this kernel may include the combined effect of time windowing and frequency windowing (see the discussion of Eq. (2.1.40)).

We can use the same process to define a filter which attenuates cross-terms in QTFDs. The cross-terms in the (t, f) domain tend to be highly oscillatory, so that the corresponding terms in the dual (v, τ) ambiguity domain tend to be away from the origin (high-pass). The auto-terms in the (t, f) domain tend to be smooth and well delineated, so that the corresponding terms in the dual (v, τ) ambiguity domain are mostly concentrated about the origin or “pass through” the origin (low-pass) as shown in Fig. 3.2.1. This behavior is well known in the field of radar (see Section 5.1). Hence, the cross-terms in the ambiguity domain can be “filtered out” by selecting a kernel filter⁴ $g(v, \tau)$ that deemphasizes information far from the origin in the Doppler-lag domain (see Sections 4.2, 5.2, 5.9 and [1] for more details).

Various authors [1, 22] have shown that other desirable TFD properties are equivalent to constraints on the kernel filter, and we shall see that most of these constraints are conveniently expressed in the Doppler-lag (ambiguity) domain. For all the above reasons, *the design of the TFD kernel filter is usually performed in the Doppler-lag domain*. Often, the resulting kernel is then described in the time-lag domain for ease of TFD implementation, as given by Eq. (3.2.9).

3.2.4 DOPPLER-FREQUENCY FORMULATION AND THE SCF

Equations (3.2.10) and (3.2.12) relate QTFDs and SCF; we have

$$\rho_z(t, f) = \mathcal{F}_{t \leftarrow v}^{-1} \left\{ \mathcal{G}(v, f) *_{\! f} k_z(v, f) \right\} \quad (3.2.20)$$

or, writing out the transform and convolution,

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{G}(v, f - \eta) k_z(v, \eta) e^{j2\pi vt} d\eta dv. \quad (3.2.21)$$

⁴The term “kernel filter” is used to reinforce the idea that designing QTFDs essentially reduces to filter design with specific constraints.

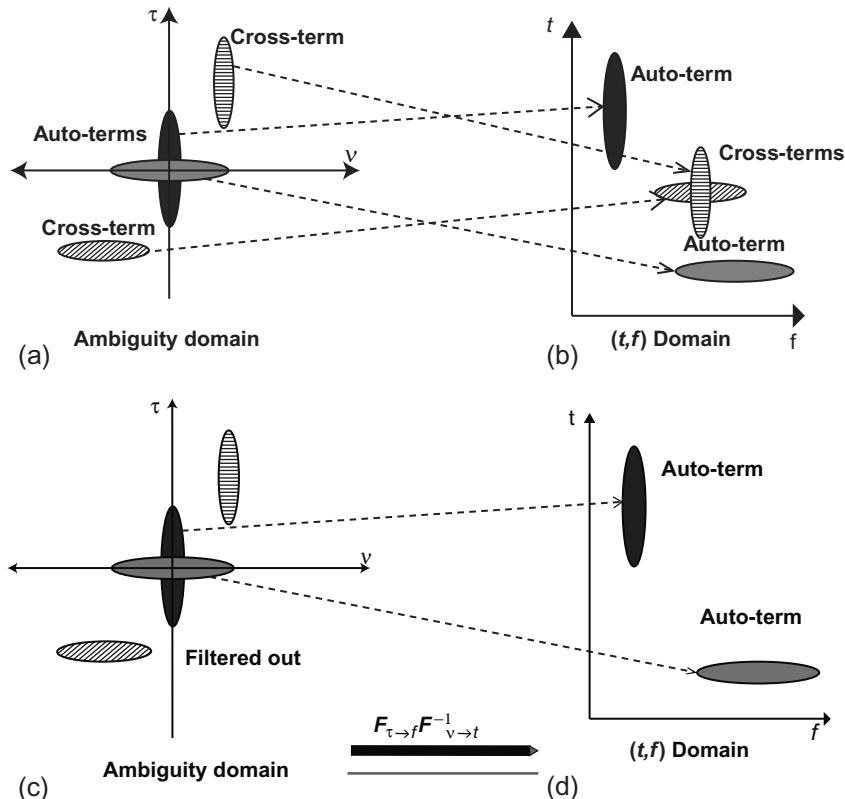
**FIGURE 3.2.1**

Illustration of the design process for high-resolution QTFD using 2D masking in the ambiguity domain:
 (a) ambiguity function; (b) WVD; (c) ambiguity domain filtering; and (d) the resultant high-resolution TFD with reduced cross-terms.

Writing η for f in Eq. (2.1.31) and substituting the result into Eq. (3.2.21), we obtain

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{G}(v, f - \eta) Z\left(\eta + \frac{v}{2}\right) Z^*\left(\eta - \frac{v}{2}\right) e^{j2\pi vt} d\eta dv. \quad (3.2.22)$$

This defines QTFDs in terms of the Doppler-frequency kernel $\mathcal{G}(v, f)$ and the signal spectrum. The interest of this Doppler-frequency formulation is that TFDs of narrow-band signals expressed by their spectra may be more efficiently computed in this form as the computation of the SCF requires less calculations as compared to IAF for such signals. In addition, cyclostationary spectral methods are often defined in this domain (see Section 9.6.1.2).

3.2.5 EXAMPLES OF SIMPLE TFD FORMULATIONS

The approach to QTFD design is illustrated by the diagram in Fig. 3.2.2. The next sections provide the detail.

3.2.5.1 Separable kernels

A simple way to design kernel filters for QTFDs is to consider the special case of a separable kernel (see Section 5.7), for which

$$g(v, \tau) = G_1(v) g_2(\tau), \quad (3.2.23)$$

so that Eq. (3.2.11) becomes

$$\mathcal{A}_z(v, \tau) = G_1(v) g_2(\tau) A_z(v, \tau). \quad (3.2.24)$$

If we let

$$G_1(v) = \mathcal{F}_{t \rightarrow v} \{g_1(t)\} \quad (3.2.25)$$

and

$$G_2(f) = \mathcal{F}_{\tau \rightarrow f} \{g_2(\tau)\}, \quad (3.2.26)$$

then, from Eqs. (3.2.10) and (3.2.24) and the convolution property, we obtain

$$\rho_z(t, f) = g_1(t) *_{\tau} W_z(t, f) *_{f} G_2(f). \quad (3.2.27)$$

Equation (3.2.24) shows that the design of the kernel filter $G_1(v) g_2(\tau)$ is greatly simplified as the 2D filtering operation is replaced by two successive 1D filtering operations as shown in Fig. 3.2.3. Equivalently, in Eq. (3.2.27), the two convolutions can be evaluated in either order, indicating that the Doppler-dependent and lag-dependent factors in the separable kernel $g(v, \tau)$ lead to separate convolutions in t and f , respectively. A separable kernel is not to be confused with a *product kernel*, which is a function of the product $v\tau$, for example, $g(v, \tau) = g_3(v\tau)$.

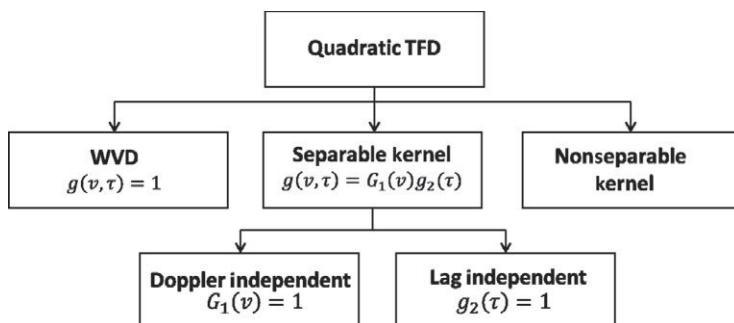


FIGURE 3.2.2

Different approaches for QTFD design.

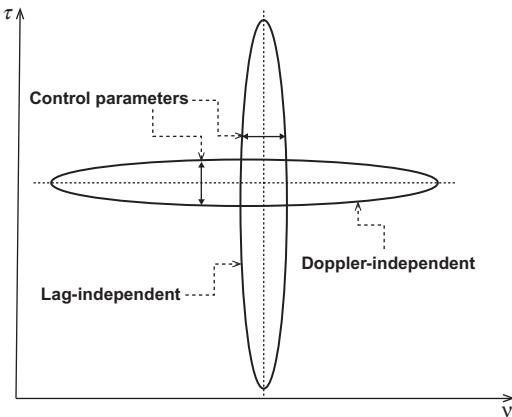
**FIGURE 3.2.3**

Illustration of the lag-independent (LI) vs Doppler-independent (DI) kernel (this diagram explains graphically; Eqs. (3.2.28) and (3.2.33)).

3.2.5.2 Doppler-independent kernels

A *Doppler-independent* (DI) kernel is a special case of a separable kernel obtained by putting $G_1(v)$ constant or in normalized form by selecting

$$G_1(v) = 1 \quad (3.2.28)$$

in Eqs. (3.2.23) and (3.2.25), which then become

$$g(v, \tau) = g_2(\tau), \quad (3.2.29)$$

$$g_1(t) = \delta(t). \quad (3.2.30)$$

Making these substitutions in Eqs. (3.2.24) and (3.2.27), we obtain

$$\mathcal{A}_z(v, \tau) = g_2(\tau) A_z(v, \tau), \quad (3.2.31)$$

$$\rho_z(t, f) = G_2(f) * W_z(t, f). \quad (3.2.32)$$

The last result shows that in the (t, f) domain, a DI kernel causes smoothing of the WVD in the frequency direction only. (Doppler and frequency have the same dimensions which serves as a mnemonic.) But in all four domains in Eq. (3.2.6), a “DI” kernel remains independent of Doppler. TFDs defined by DI kernels are best suited for signals whose energy is concentrated in the (t, f) domain along a direction parallel to the frequency axis; examples include EEG seizure spikes [1].

3.2.5.3 Lag-independent kernels

A *lag-independent* (LI) kernel is another special case of the separable kernel, obtained by putting $g_2(\tau)$ constant or in normalized form by selecting

$$g_2(\tau) = 1 \quad (3.2.33)$$

in Eqs. (3.2.23) and (3.2.26), which then become

$$g(\nu, \tau) = G_1(\nu), \quad (3.2.34)$$

$$G_2(f) = \delta(f). \quad (3.2.35)$$

Making these substitutions in Eqs. (3.2.24) and (3.2.27), we obtain

$$\mathcal{A}_z(\nu, \tau) = G_1(\nu) A_z(\nu, \tau), \quad (3.2.36)$$

$$\rho_z(t, f) = g_1(t) * \underset{t}{W}_z(t, f). \quad (3.2.37)$$

The last result shows that in the (t, f) domain, an LI causes smoothing of the WVD in the *time* direction only. (Lag and time have the same dimensions which serves as a mnemonic.) In the (t, τ) domain, an LI kernel is a function of time only; but in all four domains in Eq. (3.2.6), it remains independent of lag. TFDs defined by LI kernels are best suited for signals whose energy is concentrated in the (t, f) domain along a direction parallel to the time axis; examples include some FM like newborn EEG seizure signals [1]. Graphically, this LI example is “orthogonal” to the DI example given earlier.

The kernel of the WVD is $g(\nu, \tau) = 1$, which is both DI and LI. The kernel of the windowed WVD (Eq. 2.1.35) is DI, and the kernel of the filtered WVD (Eq. 2.1.38) is LI.

Section 5.7 provides an in-depth treatment of such separable kernels, including examples of QTFDs obtained by this simple kernel filter design procedure. More advanced methods with higher-resolution TFDs are briefly presented next, so as to introduce the top performing methods including directional kernel-based TFDs that are discussed in more detail in subsequent chapters.

3.3 PROPERTIES OF QUADRATIC TFDs (QTFDs)

3.3.1 DESIRABLE PROPERTIES

3.3.1.1 Core properties

Some QTFDs verify desirable properties that are not shared by the WVD, and vice versa. Later we shall relate the properties of a TFD to the constraints of its kernel and tabulate properties for selected TFDs. But first we discuss some properties that are considered fundamental [23–25] for many applications.

1. *Concentration of local energy:* The energy in a certain region R in the (t, f) plane, denoted by E_{z_R} , is given by the integral of the TFD over the region R; for example, if R is the region within a time interval Δt and a frequency band Δf , the energy within R is

$$E_{z_R} = \int_{\Delta t} \int_{\Delta f} \rho_z(t, f) \, df \, dt. \quad (3.3.1)$$

2. *IF/SD visualization:* The TFD of a monocomponent signal directly depicts the instantaneous frequency $f_i(t)$ and spectral delay $\tau_d(f)$ as a range of peaks along the curve representing the FM law in the (t, f) plane. That is, if $z(t)$ is a monocomponent signal and $\rho_z(t, f)$ is its TFD, then

$$\max_f \rho_z(t, f) = f_i(t), \quad (3.3.2)$$

$$\max_t \rho_z(t, f) = \tau_d(f). \quad (3.3.3)$$

- 3. Reduced interference (RI):** The TFD attenuates or suppresses inner and outer artifacts (cross-terms) relative to the signal components (auto-terms).

There are other conditions considered by some authors as fundamental, for example, *positivity*, or more precisely *nonnegativity* (NN), defined as

$$\rho_z(t, f) \geq 0 \quad \forall z, t, f. \quad (3.3.4)$$

Among known QTFDs, only *sums of spectrograms* and *filterbanks* possess NN. This means that the Doppler-lag kernel of a “positive” TFD is a sum of ambiguity functions, which makes the NN property incompatible with both the IF property (see [Table 3.3.1](#)) and IF visualization [26,27]. Hence, NN is usually considered nonessential because its cost is excessive, and it contradicts the meaning of gradient provided by the Page distribution (see [Section 2.5.1](#)). In practice, some specialists discard the spectrogram not just for this reason but for the additional reasons of resolution and sensitivity of the window as shown in [Fig. 2.3.1](#). Other specialists still prefer to continue to use it as it is still a very popular and widely used method due to its simplicity of use and the fact that it is nonnegative. Its main limitation of poor resolution and sensitivity to window length limits its scope for many applications, for example, accurate IF estimation of close signal components.

3.3.1.2 The subclasses \mathcal{P} , \mathcal{P}' , and reduced interference distributions

A subclass \mathcal{P} of QTFDs comprises those TFDs which satisfy the realness, time marginal, frequency marginal, instantaneous frequency, time support, and frequency support properties. Researchers have shown much interest in TFDs of this class and especially RIDs of this class. To some extent the design of RIDs is the art of “improving” on the spectrogram by

- sacrificing NN,
- improving resolution, and
- retaining sufficient reduced-interference ability for the application.

The outcome is usually a compromise between the spectrogram and the WVD, involving a time-frequency kernel filter less extensive than that of the spectrogram and resulting in a TFD with higher resolution than the spectrogram and better suppression of cross-terms than the WVD. The compromise may involve sacrificing one or both of the marginal properties (time or frequency). Although the marginals are critical in the field of quantum physics [18], they seem less important in signal processing. As evidence of this we note that

- The spectrogram does not satisfy the marginals and yet has always been regarded as a very useful TFD.
- Attempts to improve on the spectrogram are usually motivated by the need for higher resolution (e.g., [28]) rather than satisfying the marginals.
- If the marginal conditions do not hold, it is still possible that the integral forms thereof (Eqs. [\(2.6.6\)](#), [\(2.6.7\)](#)) are approximately true over sufficiently broad time intervals or frequency bands.

For these reasons it is suggested that the class \mathcal{P} is not necessarily the most appropriate for general nonstationary signal processing applications, and a more relevant class \mathcal{P}' can be defined, which comprises reduced-interference distributions satisfying the realness, global and local energy, IF

visualization, and components resolution properties. Note that the properties of class \mathcal{P}' are those listed on p. 41 at the end of [Section 1.1.5](#).

3.3.2 TFD PROPERTIES AND EQUIVALENT KERNEL CONSTRAINTS

Using some properties listed in [6,29], the relevant kernel constraints are adapted for separable, DI and LI kernels. The results are collected in [Table 3.3.1](#).

The proofs of these kernel constraints use Eq. (3.2.13), which states that the QTFD is the WVD convolved in t and f with the (t,f) kernel $\gamma(t,f)$, that is, $\rho_z(t,f) = \gamma(t,f) \ast\ast_{(t,f)} W_z(t,f)$. Using this relation, we can find *sufficient* conditions under which certain properties of the WVD carry over to $\rho_z(t,f)$. For example:

- Realness holds if $\gamma(t,f)$ is real.
- Because 2D convolution is shift-invariant, time- and frequency-shift invariance hold for any fixed $\gamma(t,f)$.
- Time support holds for a DI kernel, which does not redistribute the WVD in time, whereas frequency support holds for an LI kernel, which does not redistribute the WVD in frequency.
- The IF moment property holds for a DI kernel if $G_2(f)$ has a first moment (mean frequency) of zero, so that convolution w.r.t. f does not change the first moment of the WVD w.r.t. f . Similarly, the GD property holds for an LI kernel if $g_1(t)$ has a first moment (mean time) of zero, so that convolution w.r.t. t does not change the first moment of the WVD w.r.t. t .

Some further proofs of kernel constraints can be found in [3,6].

3.3.3 BASIC DESIGN METHODS FOR HIGH-RESOLUTION TFDs

3.3.3.1 Examples of TFDs with specific properties

For the TFDs defined so far, [Table 3.3.2](#) lists their kernels in various domains, and [Table 3.3.3](#) shows their properties.

Note: a time-lag kernel satisfying the time-support constraint is described as a *butterfly function* [6] or *cone-shaped kernel* [30]; that is, the nonzero values of the kernel are confined to the interior of a 2D cone in the (t, τ) plane (see the entry in the “General” column of [Table 3.3.1](#)). The Born-Jordan and ZAM distributions have kernels of this type.

3.3.3.2 Standard design kernel procedure

The reduced-interference capabilities of separable kernels warrant special attention. Cross-terms oscillate in the direction orthogonal to the line joining two signal components ([Section 4.2](#)). For small duration spikes, the direction of oscillation of cross-terms thus becomes parallel to the frequency axis ([Section 5.9](#)). Such cross-terms can be suppressed by smoothing along the frequency axis by using a DI kernel with a sufficiently long $G_2(f)$ (see [Eq. 3.2.37](#)), which corresponds to sufficiently short $g_2(\tau)$. For pure tones, the direction of oscillation of cross-terms becomes parallel to the time axis. Such

Table 3.3.1: TFD Properties and Associated Constraints on the Doppler-Lag Kernel for General, Separable, DI, and LI Kernels

| Property | Kernel Constraints | | | |
|----------|---|--|------------------------------------|---------------------------------|
| | General – | Separable $g(v, \tau) = G_1(v)g_2(\tau)$ | DI $g(v, \tau) = g_2(\tau)$ | LI $g(v, \tau) = G_1(v)$ |
| EC | $g(0, 0) = 1$ | $G_1(0)g_2(0) = 1$ | $g_2(0) = 1$ | $G_1(0) = 1$ |
| RE | $g(v, \tau) = g^*(-v, -\tau)$ | $G_1(v)g_2(\tau) = G_1^*(-v)g_2^*(-\tau)$ | $G_2(f)$ is real | $g_1(t)$ is real |
| TI | $g(v, \tau)$ does not depend on t | $G_1(v)g_2(\tau)$ does not depend on t | $g_2(\tau)$ does not depend on t | $G_1(v)$ does not depend on t |
| FI | $g(v, \tau)$ does not depend on f | $G_1(v)g_2(\tau)$ does not depend on f | $g_2(\tau)$ does not depend on f | $G_1(v)$ does not depend on f |
| TM | $g(v, 0) = 1 \forall v$ | $G_1(v)g_2(0) = 1 \forall v$ | $g_2(0) = 1$ | WVD only |
| FM | $g(0, \tau) = 1 \forall \tau$ | $G_1(0)g_2(\tau) = 1 \forall \tau$ | WVD only | $G_1(0) = 1$ |
| IF | $g(v, 0) = \text{const.}$ $\frac{\partial g}{\partial \tau} _{\tau=0} = 0 \forall v$ | $G_1(v)g_2(0) = \text{const.}$ $g'_2(0) = 0$ | $g'_2(0) = 0$ | WVD only* |
| SD | $g(0, \tau) = \text{const.}$ $\frac{\partial g}{\partial v} _{v=0} = 0 \forall \tau$ | $G_1(0)g_2(\tau) = \text{const.}$ $G'_1(0) = 0$ | WVD only* | $G'_1(0) = 0$ |
| TS | $G(t, \tau) = 0$ if $ \tau < 2 t $ | DI only | Always | WVD only* |
| FS | $G(f, v) = 0$ if $ v < 2 f $ | LI only | WVD only* | Always |
| RI | Unrestricted | Unrestricted | x-terms parallel to frequency axis | x-terms parallel to time axis |
| NN | $g(v, \tau)$ is an AF (or sum of) | $g(v, \tau)$ is an AF (or sum of) | Never | Never |

Notes: TI and FI are automatically satisfied provided that $g(v, \tau)$ has no hidden dependence on t or f . The asterisk in “WVD only*” means that the WVD may be multiplied by a scale factor. Abbreviations for properties: EC, energy conservation; FI, frequency-shift invariance (or frequency covariance); FM, frequency marginal; FS, frequency support; IF, instantaneous frequency; NN, nonnegativity; RE, realness; RI, reduced interference; TI, time-shift invariance (or time covariance); TM, time marginal; TS, time support; SD, spectral delay.

cross-terms can be easily removed by smoothing along the time axis, that is, using an LI kernel with a sufficiently long $g_1(t)$, which corresponds to sufficiently short $G_1(v)$. A general separable kernel with sufficiently short $G_1(v)$ and $g_2(\tau)$ can therefore attenuate both kinds of artifacts; a TFD with such a kernel is shown in Fig. 2.7.1 part (l) for an LFM signal.

Properties of separable kernels and their special cases, such as the B-distribution and modified B-distribution, are studied further and illustrated on examples in Section 5.7 (see also discussion in [31,32]). Design of RIDs is discussed more generally in Section 5.2. Section 6.1 extends these definitions to the discrete-time case. The next section presents the state of the art in advanced design of TFDs.

Table 3.3.2: A Kernels of Selected TFDs in the Time-Lag, Doppler-Lag, and (Where Possible) Doppler-Frequency Domains

| TFD Name | $G(t, \tau)$ | $g(v, \tau)$ | Kernel |
|--------------|--|--|---|
| Wigner-Ville | $\delta(t)$ | 1 | $\delta(f)$ |
| Levin | $\frac{1}{2} \left[\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right) \right]$ | $\cos(\pi v \tau)$ | $\frac{1}{2} \left[\delta\left(f + \frac{v}{2}\right) + \delta\left(f - \frac{v}{2}\right) \right]$ |
| Born-Jordan | $\frac{1}{ 2\alpha\tau } \operatorname{rect} \frac{t}{2\alpha\tau}$ | $\operatorname{sinc}(2\alpha v \tau)$ | $\frac{1}{ 2\alpha v } \operatorname{rect} \frac{f}{2\alpha v}$ |
| Modified BD | $\frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi}$ | $\frac{ \Gamma(\beta + j\pi v) ^2}{\Gamma^2(\beta)}$ | $\frac{ \Gamma(\beta + j\pi v) ^2}{\Gamma^2(\beta)} \delta(f)$ |
| EMBD | $\frac{ \Gamma(\alpha + j\pi \tau) ^2 \cosh^{-2\beta} t}{\Gamma^2(\alpha) \int \cosh^{-2\beta} \xi d\xi}$ | $\frac{ \Gamma(\beta + j\pi v) ^2}{\Gamma^2(\beta)} \frac{ \Gamma(\alpha + j\pi \tau) ^2}{\Gamma^2(\alpha)}$ | $\frac{\cosh^{-2\alpha} f}{\int \cosh^{-2\alpha} \xi d\xi} \frac{ \Gamma(\beta + j\pi v) ^2}{\Gamma^2(\beta)}$ |
| w-WVD | $\delta(t) w(\tau)$ | $w(\tau)$ | $W(f)$ |
| w-Levin | $\frac{w(\tau)}{2} \left[\delta\left(t + \frac{\tau}{2}\right) + \delta\left(t - \frac{\tau}{2}\right) \right]$ | $w(\tau) \cos(\pi v \tau)$ | $\frac{1}{2} \left[W\left(f + \frac{v}{2}\right) + W\left(f - \frac{v}{2}\right) \right]$ |
| ZAM | $w(\tau) \operatorname{rect} \frac{t}{2\tau/a}$ | $w(\tau) \frac{2 \tau }{a} \operatorname{sinc} \frac{2v\tau}{a}$ | $\frac{-1}{\pi^2 a} W(f) * \frac{1}{f^2 - v^2/a^2}$ |
| Rihaczek | $\delta\left(t - \frac{\tau}{2}\right)$ | $e^{-j\pi v \tau}$ | $\delta\left(f + \frac{v}{2}\right)$ |
| w-Rihaczek | $w(\tau) \delta\left(t - \frac{\tau}{2}\right)$ | $w(\tau) e^{-j\pi v \tau}$ | $W\left(f + \frac{v}{2}\right)$ |
| Page | $u(-\tau) \delta\left(t + \frac{\tau}{2}\right) + u(\tau) \delta\left(t - \frac{\tau}{2}\right)$ | $e^{-j\pi v \tau }$ | $\frac{1}{2} \left[\delta\left(f + \frac{v}{2}\right) + \delta\left(f - \frac{v}{2}\right) \right] + jv/[2\pi(f^2 - v^2/4)]$ |
| Gaussian | $\frac{\sqrt{\pi\sigma}}{ \tau } e^{-\pi^2 \sigma t^2 / \tau^2}$ | $e^{-v^2 \tau^2 / \sigma}$ | $\frac{\sqrt{\pi\sigma}}{ v } e^{-\pi^2 \sigma f^2 / v^2}$ |
| BD | $ \tau ^\beta \cosh^{-2\beta} t$ | $\frac{ \tau ^\beta \Gamma(\beta + j\pi v) ^2}{2^{1-2\beta} \Gamma(2\beta)}$ | $\frac{-\sin(\pi\beta/2) \Gamma(\beta + 1)}{2^{1-\beta} \Gamma(2\beta)} \frac{ \Gamma(\beta + j\pi v) ^2}{ \pi f ^{\beta+1}}$ |
| Spectrogram | $w\left(t + \frac{\tau}{2}\right) w\left(t - \frac{\tau}{2}\right)$ | $A_w(v, \tau)$ | $W\left(f + \frac{v}{2}\right) W\left(f - \frac{v}{2}\right)$ |

Notes: The window $w(t)$ is assumed to be real and even. Its FT and AF are $W(f)$ and $A_w(v, \tau)$, respectively. The prefix “w” means “windowed.”

3.3.4 ADVANCED DESIGN OF HIGH-RESOLUTION TFDs

Previous sections focused on QTFDs because of their potential for high resolution and simple interpretation as a distribution of signal energy in a (t, f) plane. A remaining question is: Which TFD should be used to get started in a particular application? The choice is not obvious given that the WVD has a good resolution but with a problem of cross-terms; the spectrogram is easy to use with no apparent cross-terms but has a problem of resolution and window sensitivity; and the EMBD offers a good compromise, and other QTFDs appear more complex to design and calibrate and still remain data-dependent. To answer the above question, let us first define an enhanced spectrogram that can be used as the basic reference for an objective comparison.

Table 3.3.3: A Properties of the TFDs Whose Kernels Are Listed in Table 3.3.2

| TFD Name | Property | | | | | | |
|------------------------------|----------|----|----|----|----|----|----|
| | RE | TM | FM | IF | TD | TS | FS |
| Wigner-Ville | ! | ! | ! | ! | ! | ! | ! |
| Levin (product kernel) | ! | ! | ! | ! | ! | ! | ! |
| Born-Jordan (product kernel) | ! | ! | ! | ! | ! | * | * |
| Modified B (LI kernel) | ! | | ! | | ! | | ! |
| w-WVD (DI kernel) | ! | * | | ! | | ! | |
| w-Levin | ! | * | | ! | | ! | |
| ZAM | ! | * | | ! | | * | |
| Rihaczek (product kernel) | | ! | ! | | | ! | ! |
| w-Rihaczek | | * | | | | ! | |
| Page | ! | ! | ! | ! | | ! | |
| Exponential (product kernel) | ! | ! | ! | ! | ! | | |
| B (separable kernel) | ! | | | | | | |
| Spectrogram | ! | | | | | | |

Notes: The window $w(t)$ is assumed to be real and even. An exclamation (!) means that the property is always satisfied, while an asterisk (*) means that the property is satisfied subject to normalization of the window (for TM) or the value of the parameter (for TS and FS), while blank spaces mean that the property is not satisfied by the given TFD. Comments on the kernel type are in parentheses (thus).

3.3.4.1 Enhanced spectrogram

Previous sections indicated that the spectrogram is a popular tool in applications that do not require high resolution. Improvements to the spectrogram can be made in a number of ways. One approach, called the S-method (SM) [33], is defined as

$$SM_s(t, f) = 2 \int_{-\infty}^{\infty} P(v) F_s^w(t, f + v) F_s^{w*}(t, f - v) dv, \quad (3.3.5)$$

where $F_s^w(t, f)$ is the short-time FT of $s(t)$, obtained with a window w and $P(v)$ is a narrow window that can have shapes like Hamming, Gaussian, or rectangular. The length of $P(v)$ controls the cross-term suppression and auto-term resolution properties of the enhanced spectrogram. Previous studies have shown that the appropriate selection of the length of $P(v)$ can combine the advantages of both the spectrogram and WVD so that the resultant distribution can have auto-term resolution close to the WVD with significant suppression of cross-terms [34]. This enhanced spectrogram is a QTDF with kernel defined as

$$g(v, \tau) = P\left(\frac{v}{2}\right)_v * \int_{-\infty}^{\infty} w\left(u + \frac{\tau}{2}\right) w^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi uv} du = P\left(\frac{v}{2}\right)_v * A_w(v, \tau). \quad (3.3.6)$$

The shape of this SM kernel in the ambiguity domain depends on the shapes and sizes of $w(t)$ and $p(t)$. This implies that just like separable kernel methods, in addition to shape parameters, the SM kernel has two parameters (that are lengths of $w(t)$ and $p(t)$) to adjust smoothing along both lag and Doppler

axes. Note that the standard spectrogram lacks such flexibility [33]. See Section 3.3.2 for standard kernel constraints. An application of the SM is presented in Section 11.3.5, p. 660.

3.3.4.2 Compact kernel TFD

Compact support kernels are designed to vanish outside a given range in the ambiguity domain; unlike Gaussian windows, they do not have infinite length, so there is no need to truncate them using rectangular windows that may cause loss of information. These TFDs have been shown to outperform some other fixed kernel-based methods in terms of their ability to suppress cross-terms while retaining the resolution of auto-terms in some cases [35]. Such high-resolution performance is achieved by these kernels by combining their compact support with a flexibility to adjust both shape and size of the kernel independently, as shown below.

$$g(v, \tau) = G_1(v)g_2(\tau) = \begin{cases} e^{2c} e^{\frac{cD^2}{v^2-D^2} + \frac{cE^2}{\tau^2-E^2}} & |v| < D, \quad |\tau| < E, \\ 0 & \text{otherwise.} \end{cases} \quad (3.3.7)$$

The above equations show that both Doppler v and lag τ windows are determined, respectively, by the parameters D and E . While parameter c controls their shape [33]. Moreover, the kernel width in the ambiguity domain (e.g., the parameters D and E in Eq. (3.3.7)) can be explicitly determined by prior knowledge of the signal components. The cut-off value of the kernel in each axis equals the minimum distance between any two components of a signal in that axis, that is, for a signal with N components, the cut-off value of the kernel function in the Doppler axis equals to $\min |f_p - f_q|$ while the cut-off value in the lag axis is $\min |t_p - t_q|$, where $p, q = 1, \dots, N, p \neq q$ [36]. As a result, these compact support parameters can be optimally defined rather than arbitrarily selected. The prior knowledge of the locations of signal components can be obtained using an initial exploratory reference TFD. Separable kernel TFDs such as the compact kernel distribution (CKD) give best performance for signals whose auto-terms are nearly parallel to either time or frequency axis; but its performance degrades for signals whose auto-terms have a specific direction away from the time axis or frequency axis in the (t, f) domain. In such cases, the optimum can be reached using multidirectional kernels as defined below.

3.3.4.2.1 Example of data-dependent kernel design

The MBD was shown to be adapted to signals with IFs quasi parallel to the time axis. For such signals, the MBD and compact support principle can be combined to yield the compact support modified B-distribution (CS-MBD) [36] expressed as

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \left\{ z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) \mathcal{F}_{t \leftarrow v}^{-1} \{ g(v, \tau) \} \right\}, \quad (3.3.8)$$

$$g(v, \tau) = g(v) = \begin{cases} \frac{|\Gamma(\beta + j\pi v)|^2}{\Gamma^2(\beta)} & \text{if } v^2 < v_0^2, \\ 0 & \text{otherwise,} \end{cases} \quad (3.3.9)$$

where $g(v, \tau)$ represents the LI kernel of the CS-MBD in the ambiguity domain, Γ represents the gamma function in which β is a parameter controlling the trade-off between the resolution enhancement and cross-term elimination; and v_0 is defined as the minimal frequency distance between any two

components in the signal [36]. An application of the CS-MBD to heart-rate variability (HRV) signals is presented in [Section 16.6.2](#), p. 954.

3.3.4.3 Design of a multidirectional kernel TFD

Separable kernel TFDs have high energy concentration for signals whose energy is either concentrated along the time axis or frequency axis in the (t, f) domain, but these TFDs fail to offer high resolution for signals whose energy is concentrated along several directions away from both time-axis and frequency-axis in the (t, f) domain, for example, multicomponent LFM signals with different nonzero chirp rates. In order to obtain high-resolution TFDs for such signals, a rotation parameter can be included in the formulation of smoothing kernels [33]. Such TFDs are denoted multidirectional kernel TFDs (in short MDD). Considering first the case of one direction (e.g., one LFM), one can include an additional parameter θ to rotate the CKD kernel as per the following expression:

$$g_\theta(v, \tau) = \begin{cases} e^{\frac{c}{\left(\frac{v \cos(\theta) - \tau \sin(\theta)}{D}\right)^2 - 1}} e^{\frac{c}{\left(\frac{\sin(\theta)v + \cos(\theta)\tau}{E}\right)^2 - 1}} & \text{for } |\cos(\theta)v - \sin(\theta)\tau| < D \\ 0 & \text{for } |\sin(\theta)v + \cos(\theta)\tau| < E \\ & \text{otherwise,} \end{cases} \quad (3.3.10)$$

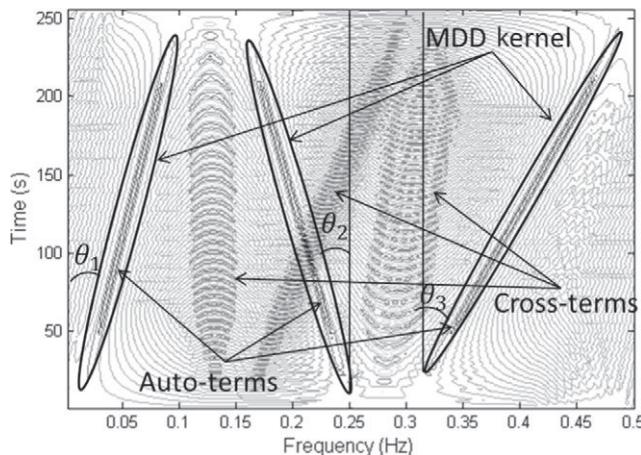
where θ is the angle of the kernel with the Doppler axis in the ambiguity domain or the time axis in the (t, f) domain⁵; E is the half-length of the kernel along its principal direction in the ambiguity domain, and D is the half-support of $g_\theta(v, \tau)$ along the direction perpendicular to the angle of the kernel. The kernel $g(v, \tau)$ in Eq. (3.3.7) is a decreasing function that approaches zero as v approaches D or τ approaches E , unlike the Gaussian function that never goes to zero. The expression $e^{2c} e^{\frac{cD^2}{v^2-D^2} + \frac{cE^2}{\tau^2-E^2}}$ is defined for $v < D$ and $\tau < E$ because of discontinuity at $v = D$ or $\tau = E$. For signals with multiple directions of energy concentration in the (t, f) domain, smoothing is performed along multiple directions, resulting in a summation of a predetermined number N_D of directional kernels that defines the multidirectional kernel (MDK)

$$g(v, \tau) = \frac{e^c}{N_D} \sum_{i=1}^{N_D} g_{\theta_i}(v, \tau), \quad (3.3.11)$$

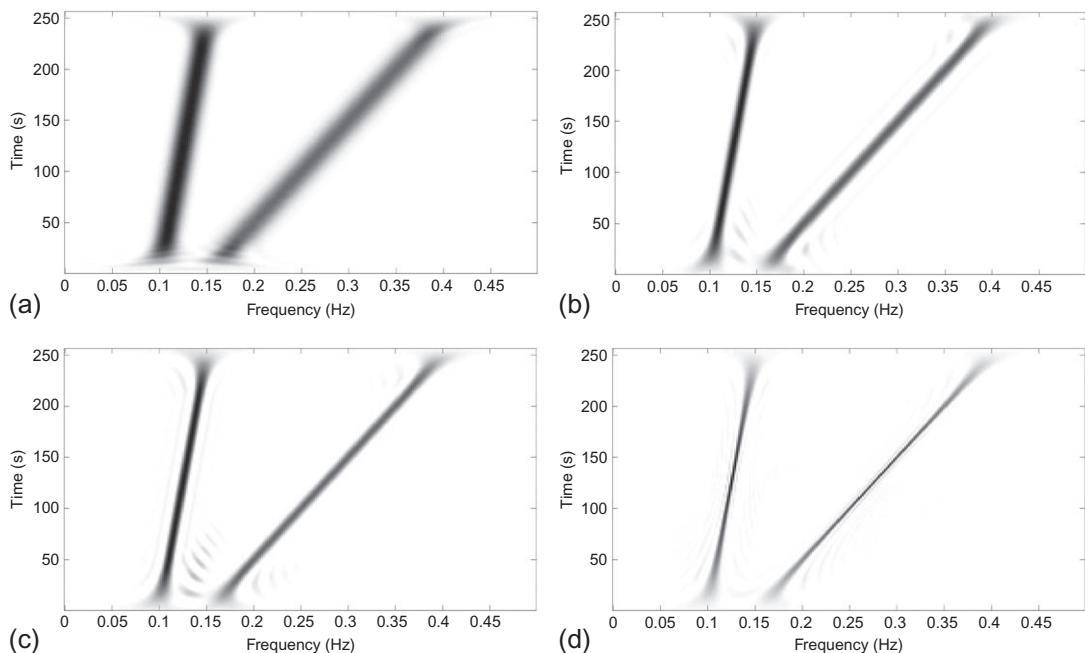
where θ_i is the angle of the i th kernel as shown in [Fig. 3.3.1](#) (see [Section 5.9](#) for more details). *Performance evaluation of the above advanced high-resolution TFDs:* To do this comparison, let us consider a two-component LFM signal having two different directions of energy concentration in the (t, f) domain. The signal is analyzed using the spectrogram, S-method, MDD, and CKD. [Figure 3.3.2](#) shows that the spectrogram gives the poorest energy concentration. High-resolution methods such as the CKD and enhanced spectrogram achieve better results than the spectrogram. Their improved performance is due to the flexibility to independently control smoothing along the time and frequency axes (see [Section 5.7](#) for details). The MDD can achieve the highest energy concentration, as it takes into account the direction of signal components in the (t, f) domain (see [Section 5.9](#) for details).

The study in [33] indicated that an accurate estimation of the signal components IF is directly related to the cross-term suppression and auto-term resolution properties of the (t, f) methods employed. There is therefore a need to develop better methods for improving the resolution of TFDs, including (t, f)

⁵The angle θ can be easily estimated by measuring the direction of energy concentration in the (t, f) or (v, τ) domain. It is the same in both domains due to the unitary property of the 2D FT, which preserves angles.

**FIGURE 3.3.1**

Selection of the angle parameter for the MDD based on the orientation of auto-terms in the (t, f) domain.

**FIGURE 3.3.2**

TFDs of a two-component signal: (a) spectrogram (the Hamming window with its length equal to 71 points); (b) S-method (Hamming window of length 151 samples, and a rectangular window $P(v)$ of length 11 frequency samples is used to compute the STFT); (c) compact kernel TFD ($c = 0.5$, $D = 0.2$, $E = 0.1$ in Eq. (3.3.7)); (d) MDD ($c = 0.1$, $D = 0.05$, $E = 1$, $\theta_1 = 5^\circ$, $\theta_2 = 23^\circ$ in Eq. (3.3.10)).

postprocessing methods. In particular, image-processing-based techniques can be used for such desired TFD enhancement in clarity, which is essential for TFDs, so that their relevant features can be read, selected, and extracted. To improve the readability, let us consider the TFD as a (t,f) image. Some of the state-of-the-art (t,f) image enhancement techniques discussed in [33] and Section 7.5 can then be employed as postprocessing operations. By including such image enhancement techniques within the overall methodology of change detection, improvements in performance can be obtained in applications such as the classification of physiological signal abnormalities. For example, a study in [33] indicated that the two best (t,f) features (i.e., time-frequency flatness and time-frequency flux) for the detection of seizure activity in EEG signals can be obtained from enhanced (t,f) images (see Section 12.6). The expense for this result is an additional computational cost, whereas among the nonadaptive methods (i.e., without enhancement), both the EMBD and CKD give the best performing features and this method is computationally less expensive [33].

3.4 POSITIVITY OF QTDFs: EXAMPLES AND CONDITIONS

3.4.1 THE SPECTROGRAM AS AN EXAMPLE OF POSITIVE TFD

For physical interpretation, it is intuitively desirable for TFDs to be nonnegative valued so they can be interpreted as an energy density, and TFD parameters have clear meaning. A particular positive QTDF is the standard spectrogram of a given real signal $s(t)$ defined in Eq. (2.3.10) as

$$\rho_s^w(t,f) = |F_s^w(t,f)|^2 = |\mathcal{F}_{\tau \rightarrow f}\{s(\tau) w(\tau - t)\}|^2,$$

where w is a window function and the subscript s represents the signal $s(t)$. By construction, the spectrogram is clearly positive. Let us now study its characteristics to understand the constraint on the kernel that is needed to obtain a positive TFD.

Equation (3.2.9) indicates that the spectrogram of an analytic signal $z(t)$ can be expressed as

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t-u, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f \tau} du d\tau, \quad (3.4.1)$$

where $G(t, \tau)$ is the time-lag kernel of the spectrogram (see Table 2.7.2 for more details). If the spectrogram window $w(t)$ is a real function, $G(t, \tau)$ can be considered as its IAF and is defined as follows (see also Eq. 2.7.5 for more details):

$$G(t, \tau) = w(t + \frac{\tau}{2}) w^*(t - \frac{\tau}{2}) = K_w(t, \tau). \quad (3.4.2)$$

By taking the FT of $G(t, \tau)$ from the time-lag domain (t, τ) to the Doppler-lag domain (v, τ) according to Eq. (3.2.6), we find

$$g(v, \tau) = \mathcal{F}_{t \rightarrow v}\{G(t, \tau)\} = \mathcal{F}_{t \rightarrow v}\{K_w(t, \tau)\} \quad (3.4.3)$$

$$= \int_{-\infty}^{\infty} w(t + \frac{\tau}{2}) w^*(t - \frac{\tau}{2}) e^{-j2\pi t v} dt. \quad (3.4.4)$$

By comparing Eq. (3.4.3) with Eq. (2.7.5) and Eq. (3.4.4) with Eq. (3.2.4) (pp. 86 and 111), we can recognize that $g(v, \tau)$ is the ambiguity function (AF) of $w(t)$, that is, $A_w(v, \tau)$. That is, the Doppler-lag kernel of the spectrogram is the AF of the window function used in the spectrogram. This result is generalized in Section 5.12 to any positive-valued TFD.

3.4.2 THE SONOGRAM FILTER BANK

Recalling [Theorem 2.4.1](#) (p. 81) and the second part of Eq. (5.12.23) in [Section 5.12](#), we get an insight in the other practical way to define positive TFDs using filter banks, also called the sonogram which is defined in Eq. (2.4.3) as [6]

$$S_s^{(H)}(t,f) = \left| B_s^H(t,f) \right|^2 = \left| \mathcal{F}_{t \leftarrow v}^{-1} \{ S(v) H(v-f) \} \right|^2,$$

where $B_s^H(t,f)$ is a band-pass filter, and $H(v-f)$ is redefined as a low-pass filter with a central frequency $v = f$. By following the same steps in [Sections 2.7.1](#) and [3.4.1](#), we can find that the sonogram can be expressed as a QTFD as follows:

$$\rho_s(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{G}(v,f-v) S(v + \frac{v}{2}) S^*(v - \frac{v}{2}) e^{j2\pi vt} dv dv, \quad (3.4.5)$$

where $\mathcal{G}(v,f)$ is the Doppler-frequency kernel of the sonogram and is defined, exploiting the evenness of H , as follows:

$$\mathcal{G}(v,f) = H(f - \frac{v}{2}) H^*(f + \frac{v}{2}) = k_H(v,f). \quad (3.4.6)$$

Let us relate $\mathcal{G}(v,f)$ to the AF of the low-pass filter. By transferring $\mathcal{G}(v,f)$ from the Doppler-frequency domain to the Doppler-lag domain via inverse FT, we find

$$g(v,\tau) = \mathcal{F}_{\tau \leftarrow f}^{-1} \{ \mathcal{G}(v,f) \} = \mathcal{F}_{\tau \leftarrow f}^{-1} \{ k_H(v,f) \} \quad (3.4.7)$$

$$= \int_{-\infty}^{\infty} H(f - \frac{v}{2}) H^*(f + \frac{v}{2}) e^{j2\pi \tau f} df. \quad (3.4.8)$$

By comparing Eqs. (3.2.5), (3.4.6), and (3.4.8), we recognize that the above kernel $g(v,\tau)$ is the ambiguity function of the low-pass filter $H(f)$, that is, $A_H(v,\tau)$, which confirms that the sonogram is a positive TFD. [Section 5.12](#) shows the general proof for a QTFD to be positive and elaborates on the implications of this requirement in terms of IF estimation.

3.5 QTFDs, AMBIGUITY FUNCTION, AND RADAR

Several equations such as Eqs. (3.2.5) and (3.2.10) indicated the importance of the relationship between QTFDs and the AF. The latter originates from the fields of Radar and Sonar as an extension of the correlation function to account for Doppler (frequency shift) effect. This section explores this link in more detail and illustrates the meaning of these relationships and AF in the context of a simple radar experiment.

3.5.1 RADAR AF: WHERE IS THE AMBIGUITY?

Let us assume a target is traveling at a constant radial velocity v ; the reflected radar signal is then delayed in time and frequency. Let the transmitted signal be

$$z(t) = e^{j2\pi ft^2}, \quad (3.5.1)$$

where $z(t)$ is an LFM [37]. The received signal $r(t)$ is then

$$\begin{aligned} r(t) &= z(t - \tau) e^{j2\pi\nu t} = e^{j2\pi f(t-\tau)^2} e^{j2\pi\nu t} \\ &= e^{j2\pi ft^2} e^{j2\pi f\tau^2} e^{-j4\pi f\tau} e^{j2\pi\nu t}, \end{aligned} \quad (3.5.2)$$

where τ is the round trip delay such that $\tau \triangleq \frac{2d}{c}$, ν is the Doppler shift such that $\nu = \frac{2\pi f v}{c}$, and c is the speed of light. To determine the location and velocity of the object, one simply correlates the received signal with shifts of the transmitted signal. The maximum correlation value corresponds to the location and velocity of the target. The nonsymmetric ambiguity function is represented by the correlation (i.e., inner product) between the transmitted and received signals, that is

$$A_z(\nu, \tau) = \int_{-\infty}^{\infty} z(t)r^*(t) dt = \int_{-\infty}^{\infty} z(t)z^*(t - \tau) e^{-j2\pi\nu t} dt, \quad (3.5.3)$$

which equals the symmetric definition by up to a multiplicative phase factor $\exp(j\pi\nu\tau)$ (see [Section 3.2.3](#)):

$$A_z(\nu, \tau) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\nu t} dt. \quad (3.5.4)$$

Putting values from Eqs. [\(3.5.1\)](#) and [\(3.5.2\)](#) in Eq. [\(3.5.3\)](#), we get

$$\begin{aligned} A_z(\nu, \tau) &= \int_{-\infty}^{\infty} e^{j2\pi ft^2} e^{-j2\pi f\tau^2} e^{-j2\pi f\tau^2} e^{j4\pi f\tau} e^{-j2\pi\nu t} dt \\ &= \int_{-\infty}^{\infty} e^{-j2\pi f\tau^2} e^{j4\pi f\tau} e^{-j2\pi\nu t} dt = e^{-j2\pi f\tau^2} \int_{-\infty}^{\infty} e^{j2\pi(2f\tau - \nu)t} dt. \end{aligned} \quad (3.5.5)$$

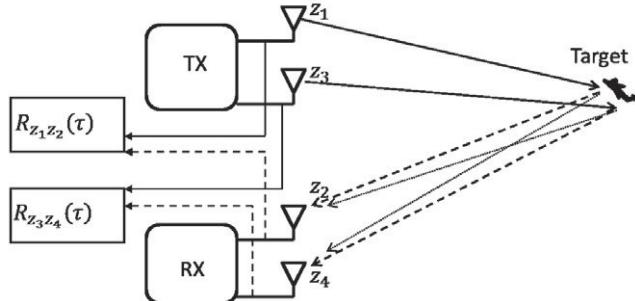
[Equation \(3.5.5\)](#) indicates that the ambiguity arises because the change in the expression $(2f\tau - \nu)$ is due to two unknown parameters τ and ν causing a range ambiguity and a Doppler ambiguity. This is also called the LFM effect [[20](#), Chap. 3].

3.5.2 WVDs, AFs, AND INNER-PRODUCT INVARIANCE

Consider the multiple-input multiple-output (MIMO) radar system shown in [Fig. 3.5.1](#) where two antennas of the transmitter (TX) send two LFM signals $\{z_1(t), z_3(t)\}$ and two antennas of the receiver (RX) collect reflected and noisy copies of them, $z_2(t)$ reflected copy of $z_1(t)$ and $z_4(t)$ reflected copy of $z_3(t)$. In the case of a moving target (Doppler shift is present), the reflected signals ($z_2 = z_1(t - \tau) e^{j2\pi\nu t}$ and $z_4 = z_3(t - \tau) e^{j2\pi\nu t}$) are correlated with the transmitted signals (z_1 and z_3) to find $R_{z_1, z_2}(\tau) = \int_{-\infty}^{\infty} z_1(t)z_2^*(t) dt = \int_{-\infty}^{\infty} z_1(t)z_1^*(t - \tau) e^{-j2\pi\nu t} dt$ and $R_{z_3, z_4}(\tau) = \int_{-\infty}^{\infty} z_3(t)z_4^*(t) dt = \int_{-\infty}^{\infty} z_3(t)z_3^*(t - \tau) e^{-j2\pi\nu t} dt$, respectively. The cross-correlation functions $R_{z_1, z_2}(\tau)$ (and resp. $R_{z_3, z_4}(\tau)$) can be seen to reduce to the ambiguity functions $A_{z_1}(\nu, \tau)$ (and resp. $A_{z_3}(\nu, \tau)$). Using [Eq. \(3.1.14\)](#), for the case of two auto-WWDs $W_{z_1}(t, f) = \mathcal{F}_{v \rightarrow t}^{-1} \mathcal{F}_{\tau \rightarrow f}(A_{z_1}(\nu, \tau))$ and

$W_{z_3}(t, f) = \mathcal{F}_{v \rightarrow t}^{-1} \mathcal{F}_{\tau \rightarrow f}(A_{z_3}(\nu, \tau))$, the inner-product invariance property becomes

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{z_1}(t, f) W_{z_3}^*(t, f) dt df = \left| \int_{-\infty}^{\infty} z_1(t) z_3^*(t) dt \right|^2. \quad (3.5.6)$$

**FIGURE 3.5.1**

MIMO radar system.

Apart from the two auto-WVDs, there may be the case to have two cross-WVDs: cross-WVD between the two transmitted signals $\{z_1(t), z_3(t)\}$ and cross-WVD between the two received signals $\{z_2(t), z_4(t)\}$. Using Eq. (3.1.14) (known as inner-product invariance or Moyal's formula [38]), the generalized inner-product invariance (GIPI) property of the WVD is given as [39]

$$\int_{\mathbb{R}^2} W_{z_1, z_3}(t, f) W_{z_2, z_4}^*(t, f) dt df = \left(\int_{-\infty}^{\infty} z_1(t) z_3^*(t) dt \right) \left(\int_{-\infty}^{\infty} z_2(t) z_4^*(t) dt \right)^*, \quad (3.5.7)$$

where $W_{z_1, z_3}(t, f)$ and $W_{z_2, z_4}(t, f)$ are the cross-WVDs of $\{z_1(t), z_3(t)\}$ and $\{z_2(t), z_4(t)\}$, respectively.

The diamond shape graphical equation (3.5.9) indicates that the GIPI property of the cross-WVDs is equally valid for the cross-ambiguity functions [19], where $A_{z_1, z_2}(\nu, \tau)$ and $A_{z_3, z_4}(\nu, \tau)$ are cross-ambiguity functions of the analytic signals $\{z_1(t), z_2(t)\}$ and $\{z_3(t), z_4(t)\}$, respectively. The correspondence between the GIPI properties of cross-WVDs and cross-ambiguity functions can also be described as

$$A_{z_1, z_2}(\nu, \tau) A_{z_3, z_4}^*(\nu, \tau) = \mathcal{F} \mathcal{F}^{-1} (W_{z_1, z_3}(t, f) W_{z_2, z_4}^*(t, f)), \quad (3.5.8)$$

which is also known as the Sussman identity [19].

$$\begin{array}{ccc}
 & W_{z_1 z_3}(t, f) W_{z_2 z_4}^*(t, f) & \\
 & \swarrow f \quad \searrow t & \\
 K_{z_1 z_2}(t, \tau) * K_{z_3 z_4}(t, \tau) & & k_{z_1 z_2}(\nu, f) * k_{z_3 z_4}(\nu, f). \\
 & \swarrow t \quad \searrow \nu & \\
 & A_{z_1 z_2}(\nu, \tau) A_{z_3 z_4}^*(\nu, \tau) &
 \end{array} \quad (3.5.9)$$

3.6 CONCLUDING THE TUTORIAL

3.6.1 SUMMARY, DISCUSSION, CONCLUSIONS, AND PERSPECTIVES

This key foundation chapter ends the tutorial introduction to (t,f) principles constituted by Part I of this book. In essence, the main message and findings of this chapter are as follows. Characteristic features of signals in the (t,f) domain can be enhanced by using data-dependent TFDs. In particular, for a monocomponent LFM signal, the WVD is optimal for energy concentration about the IF and for unbiased estimation of the IF (see Chapter 10 for details). If a signal has nonlinear frequency modulation and/or multiple components, the WVD suffers from inner artifacts and/or outer artifacts (cross-terms), respectively; in either case, reduced interference QTFDs (RIDs) should be preferred over the WVD in most applications. The design of RIDs is best undertaken by designing the desired kernel filter in the ambiguity domain and using FTs to see the effects in the time-lag (t, τ) and (t,f) domains. To be a useful tool for practical applications, QTFDs are expected to be real, to satisfy the global and local energy requirements, and to resolve signal components while reflecting the components IF laws through the peaks of their dominant ridges in the (t,f) plane. Several RIDs were designed using simple separable kernels (including, e.g., the EMBD), demonstrating the procedure to construct QTFDs that meet the desired requirements. Special-purpose QTFDs can be easily designed to meet the specifications of particular applications (as we shall also see in several sections in the following chapters such as multidirectional TFDs). In general, therefore, QTFDs such as the WVD and various RIDs are the most useful TFDs; the spectrogram, which has been widely used, is at best subsumed by quadratic RIDs (of which it is a special case), and at worst made obsolete by them, due mostly to its sensitivities to window size as shown in Fig. 2.3.1. Even optimized or enhanced spectrograms (such as the S-method) is also a particular QTFD. Another conclusion is that the findings in Chapter 3 imply that among all common QTFDs, only the spectrogram and sonogram offer “positivity,” but these cannot be used for the analysis of close signal components and signals which have highly variable IF because of poor resolution (see Section 5.12). RIDs still suffer from lower amplitude cross-terms but they can resolve close signal components. Once these signal components are resolved in the (t,f) domain, if needed, the cross-terms can be subsequently suppressed by postprocessing methods as discussed in Sections 5.10 and 7.5, respectively. The resultant TFDs would have higher resolution with suppression of cross-terms and can be used for applications such as IF estimation. An alternative to the design of RIDs is to introduce a preprocessing stage to decompose the multicomponent signal into several additive monocomponent signals using methods such as BSS (see Sections 8.4 and 8.5).

In addition, advanced procedures for the design of high-resolution TFDs can allow extraction of more precise information defining discriminatory features such as the IF. Other specific (t,f) features can then be obtained, for example, by extending traditional time domain or frequency domain features to (t,f) features such as (t,f) flatness (see details in Section 12.6 and [33]). Current state of the art in terms of TFD design indicates that

1. The enhanced spectrogram⁰ is useful for initial exploratory (t,f) analysis, but its use should be complemented by higher resolution TFDs such as the EMBD or CKD; or S-method; or multiview TFDs [40], or adaptive TFDs.

⁰Note that the enhanced spectrogram is the result of a postprocessing operation while the optimized spectrogram is obtained using an optimized window.

2. Separable kernel TFDs such as the EMBD or CKD are easy to implement and offer better resolution as compared to the spectrogram as they have more control parameters to independently adjust smoothing along the frequency and time axes.
3. The multidirectional distribution (MDD) is an extra refinement that is especially suitable for signals that have more than one direction of energy distribution in the (t,f) plane (see details in [Section 5.9](#)).
4. SNR considerations affect the choice of TFD in a particular application.

So far, we have considered only 1D single channel signals. The methods presented so far can be easily extended to the multichannel case (when dealing with recordings made by multisensors); this case is dealt with in [Chapter 8](#). The case of 2D signals (i.e., images) is discussed in [Section 15.4](#). Other advanced methods that are outside the scope of this introductory tutorial can be found in subsequent chapters, including polynomial TFDs, complex lag TFDs, and (t,f) postprocessing methods. See also the companion tutorials in [6,33]. In addition, QTFDs are generally not invariant to changes in scale; so for applications where scale is important, a class of affine and hyperbolic TFDs can be defined that are scale invariant as discussed in [41] and later chapters.

3.6.2 RELATIONSHIP TO PARTS II, III, IV, AND V

The remaining four parts of the book elaborate on the material presented so far with examples and applications, discuss advanced design methods for TFDs, and present a wide selection of methodologies, algorithms, and applications that demonstrate the use of (t,f) methods in practice. In particular, questions and issues such as how to select a TFD for a particular application, how to implement it, and how best to apply it are covered. These chapters constituting the remaining four parts include a number of sections, which tend to cover more advanced, specific, and detailed material, complementing and supplementing the tutorial introduction of [Part I](#).

[Part II](#) of the book ([Chapters 4 and 5](#)) gives more details on some fundamental topics of TFSAP such as TFD design and signal analysis in the (t,f) plane as well as related topics such as time-scale methods. [Chapter 4](#) presents some advanced concepts for (t,f) signal analysis, (t,f) signal processing, (t,f) system analysis, uncertainty principle, and Teager-Kaiser operators. [Chapter 5](#) presents a number of methods for designing TFDs, including the ambiguity function, EMD, S-transform, and adaptive TFDs.

[Part III](#) ([Chapters 6–8](#)) describes specialized techniques used in implementation, measurement, and enhancement of TFDs. [Chapter 6](#) deals with the implementation and realization of TFDs; in particular, the formulation of discrete-time QTFDs is presented and computation efficiencies are described to take care of precision, speed, and memory requirements. [Chapter 7](#) presents quality measures for TFDs and methods for performance enhancement, including entropy measures, neural networks, and preprocessing as well as postprocessing methods. [Chapter 8](#) describes methods and algorithms for multisensor and time-space processing used in applications such as sonar and telecommunications, including (t,f) multicomponent BSS, as well as direction of arrival estimation. [Part IV](#) ([Chapters 9–12](#)) presents key statistical techniques for TFSA of random signals. [Chapter 9](#) presents time-frequency methods for random processes and noise analysis; [Chapter 10](#) describes methods for instantaneous-frequency estimation in the two cases of monocomponent and multicomponent signals; [Chapter 11](#) deals with the field of time-frequency synthesis and filtering, including time-varying filter design

and (t,f) noise reduction; and Chapter 12 presents time-frequency methods for signal detection, classification, and estimation, including (t,f) matched filtering and feature attraction.

Part V (Chapters 13–16) describes a representative selection of TFSAP applications encompassing a wide range of fields and industries. Chapter 13 presents time-frequency applications in telecommunications; Chapter 14 describes time-frequency methods in radar, sonar, speech, and acoustics; Chapter 15 details a number of time-frequency methods for diagnosis and monitoring used in a wide range of diverse applications such as asset management and condition monitoring; and Chapter 16 is dedicated to (t,f) applications to the important field of neurosciences with focus on enhancing newborn health outcomes using (t,f) analysis and detection and classification of EEG abnormalities in the newborn.

To assist the reader in locating topics, the book includes a three-level table of contents and a two-level alphabetical index.

Supplementary material is provided on the Elsevier book site, including a PowerPoint presentation and a (t,f) toolbox to reproduce most plots.

3.7 ADDITIONAL EXERCISES

This section provides additional exercises that cover important topics from this chapter and prepares the reader for the next chapters in the book. It is recommended to use the (t,f) toolbox described in Chapter 17 as an aid for doing these exercises. In a formal class setting, the instructor may wish to guide the student by adding intermediate questions with focus on one or two assessment criteria.

Exercise 3.7.1. Derive analytically the following properties of the WVD.

1. Signal energy preservation (see Eq. 3.1.7, p. 104)
2. Time-marginal (see Eq. 3.1.5, p. 104)
3. Frequency-marginal (see Eq. 3.1.6, p. 104)
4. Time-shift invariance (see Eq. 3.1.2, p. 104)
5. Frequency shift invariance (see Eq. 3.1.4, p. 104)

Compare with the analytic results and comment on the results from the point of view of accuracy.

Exercise 3.7.2. We consider the following signal for analysis:

$$s(t) = \begin{cases} \cos(0.1\pi t) + \cos\left(2\pi\left(0.1t - 0.0005t^2\right)\right) + \cos\left(2\pi\left(0.35t - 0.0005t^2\right)\right), & \text{for } 0 \leq t \leq 255; \\ 0, & \text{otherwise.} \end{cases}$$

Analyze the signal using the spectrogram, optimized spectrogram, enhanced spectrogram, compact kernel TFD, windowed WVD, MBD, EMBD, and multidirectional distribution.

1. Compare the energy concentration performance of the WVD, the spectrogram, and a high-resolution TFD of your choice using the normalized Rényi entropy, which is defined as

$$\text{RV}_3 = -\frac{1}{2} \log_2 \iint \left[\rho_z(t,f) / \iint |\rho_z(t,f)| dt df \right]^3 dt df.$$

Note that a low value of the normalized Rényi entropy implies high energy concentration for the TFD while a high value of the normalized Rényi entropy implies low energy concentration for the TFD. (More details about the normalized Rényi entropy can be found in [Section 7.3.](#))

2. Optimize the EMBD and another QTDF of your choice using the following procedure.
 - a. Analyze the signal in the ambiguity domain.
 - b. Locate separately the auto-terms and then the cross-terms. (Note that the auto-terms pass through the origin in the ambiguity domain.)
 - c. Choose the parameters for the selected kernel that retains most of the auto-term energy while suppressing most of the cross-terms.
 - d. Apply the mask in the ambiguity domain and transform the ambiguity domain to the (t, f) domain by taking a suitable 2D FT.

Show your results and compare the TFDs using the normalized Rényi entropy.

REFERENCES

- [1] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [2] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, "The Wigner Distribution—a tool for time-frequency signal analysis; part 3: relations with other time-frequency signal transformations", *Philips J. Res.* 35 (6) (1980) 372-389.
- [3] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, "The Wigner Distribution—a tool for time-frequency signal analysis; part 1: continuous-time signals", *Philips J. Res.* 35 (3) (1980) 217-250.
- [4] B. Boashash, P.J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 38 (11) (1990) 1829-1841.
- [5] B. Boashash, G. Azemi, "A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG", *Digital Signal Process.* 28 (2014) 28-38.
- [6] B. Boashash, "Time-frequency signal analysis", in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [7] N.G. de Bruijn, "A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence", *Nieuw Archief voor Wiskunde* 21 (1973) 205-280.
- [8] B. Bouachache, "Representation temps-fréquence", Tech. Rep. 373/78, Soc. Nat. ELF-Aquitaine, Pau, France, 1978, 56 pp.
- [9] B. Bouachache, B. Escudie, J.M. Komatitsch, "Sur la possibilité d'utiliser la représentation conjointe en temps et fréquence dans l'analyse des signaux modulés en fréquence émis en vibroismiques", in: Seventh GRETSI Symp. on Signal Processing and its Applications, 1979, pp. 121/1-121/6.
- [10] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, "The Wigner Distribution—a tool for time-frequency signal analysis; part 2: discrete-time signals", *Philips J. Res.* 35 (4/5) (1980) 276-300.
- [11] B.D. Forrester, "Time-frequency analysis in machine fault detection", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 406-423 (Chapter 18).
- [12] F. Hlawatsch, G.F. Boudreux-Bartels, "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.* 9 (2) (1992) 21-67.

- [13] B.V.K. Vijaya Kumar, C.W. Carroll, "Performance of Wigner distribution function based detection methods", *Opt. Eng.* 23 (6) (1984) 732-737.
- [14] H.H. Szu, "Two-dimensional optical processing of one-dimensional acoustic data", *Opt. Eng.* 21 (5) (1982) 804-813.
- [15] P.J. Boles, B. Boashash, "Application of the cross-Wigner-Ville distribution to seismic data processing", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 445-466 (Chapter 20).
- [16] D.L. Jones, T.W. Parks, "A high-resolution data-adaptive time-frequency representation", *IEEE Trans. Acoust. Speech Signal Process.* 38 (12) (1990) 2127-2135.
- [17] P. Bello, "Time-frequency duality", *IEEE Trans. Inf. Theory* 10 (1) (1964) 18-33.
- [18] L. Cohen, "Introduction: a primer on time-frequency analysis", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 3-42 (Chapter 1).
- [19] S.M. Sussman, "Least-squares synthesis of radar ambiguity functions", *IRE Trans. Inf. Theory* 8 (1962) 246-254.
- [20] N. Levanon, E. Mozeson, *Radar Signals*, Wiley-IEEE Press, Hoboken, New Jersey 2004.
- [21] M.H. Kahaei, B. Boashash, M. Deriche, "Improving detection of linear FM signals in noise using adaptive lattice filters", *Electron. Lett.* 33 (12) (1997) 1032-1034, ISSN 0013-5194, <http://dx.doi.org/10.1049/el:19970695>.
- [22] J. Jeong, W.J. Williams, "Kernel design for reduced interference distributions", *IEEE Trans. Signal Process.* 40 (2) (1992) 402-412.
- [23] B. Boashash, "Wigner analysis of time-varying signals—its application in seismic prospecting", in: Proc. European Signal Processing Conf. (EUSIPCO-83), 1983, pp. 703-706.
- [24] B. Boashash, B. Escudie, "Wigner-Ville analysis of asymptotic signals and applications", *Signal Process.* 8 (3) (1985) 315-327.
- [25] B. Bouachache, "Representation temps-frequence", Thesis for diplome de docteur ingenieur, Institut National Polytechnique de Grenoble, France, 1982, 164 pp (in French).
- [26] B. Bouachache, B. Escudie, P. Flandrin, J. Gréa, "Sur une condition nécessaire et suffisante de positivité de la représentation conjointe en temps et fréquence des signaux d'énergie finie", *Comptes rendus de l'Académie des Sciences, Paris* 288, Ser. A (1979) 307-309 (in French).
- [27] B. Boashash, L.B. White, J. Imberger, "Wigner-Ville analysis of non-stationary random signals (with application to turbulent microstructure signals)", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'86), vol. 4, 1986, pp. 2323-2326.
- [28] E.F. Velez, H. Garudadri, "Speech analysis based on smoothed Wigner-Ville distribution", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 351-374 (Chapter 15).
- [29] W.J. Williams, J. Jeong, "Reduced interference time-frequency distributions", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 74-97 (Chapter 3).
- [30] Y. Zhao, L.E. Atlas, R.J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals", *IEEE Trans. Acoust. Speech Signal Process.* 38 (7) (1990) 1084-1091.
- [31] B. Barkat, B. Boashash, "Reply to ‘Comments on ‘A high-resolution quadratic time-frequency distribution for multicomponent signals analysis’’", *IEEE Trans. Signal Process.* 52 (1) (2004) 290-291, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2003.820059>.
- [32] A.Z. Sha'ameri, B. Boashash, "The lag windowed Wigner-Ville distribution: an analysis method for HF data communication signals", *J. Teknol. UTM* 30 (D) (1999) 33-54.

- [33] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [34] L. Stanković, "A method for time-frequency analysis", *IEEE Trans. Signal Process.* 42 (1) (1994) 225-229.
- [35] M. Abed, A. Belouchrani, M. Cheriet, B. Boashash, "Time-frequency distributions based on compact support kernels: properties and performance evaluation", *IEEE Trans. Signal Process.* 60 (6) (2012) 2814-2827.
- [36] S. Dong, G. Azemi, B. Boashash, "Improved characterization of HRV signals based on instantaneous frequency features estimated from quadratic time-frequency distributions with data-adapted kernels", *Biomed. Signal Process. Control* 10 (2014) 153-165, ISSN 1746-8094, <http://dx.doi.org/10.1016/j.bspc.2013.11.008>, URL <http://www.sciencedirect.com/science/article/pii/S1746809413001717>.
- [37] P.M. Woodward, "Radar ambiguity analysis", Tech. Rep., Royal Radar Establishment, Malvern, U.K 1967, 731 pp.
- [38] J.E. Moyal, "Quantum mechanics as a statistical theory", in: Mathematical Proc. Cambridge Philosophical Soc., vol. 45, 1949, pp. 99-124.
- [39] D. Farden, L. Scharf, "Estimating time-frequency distributions and scattering functions using the Rihaczek distribution", in: Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2004, pp. 470-474, <http://dx.doi.org/10.1109/SAM.2004.1502992>.
- [40] G.J. Frazer, B. Boashash, "Multiple window spectrogram and time-frequency distributions", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94), vol. IV, 1994, pp. 293-296.
- [41] A. Papandreou-Suppappola, F. Hlawatsch, G.F. Boudreux-Bartels, "Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: a unified framework for the affine, hyperbolic, and power classes", *Digital Signal Process.* 8 (1) (1998) 3-48.

This page intentionally left blank

PART

FUNDAMENTAL PRINCIPLES OF TFSAP

II

This page intentionally left blank

ADVANCED TIME-FREQUENCY SIGNAL AND SYSTEM ANALYSIS

4

INTRODUCTION AND OVERVIEW

This chapter extends the material described in [Chapters 1–3](#). It presents additional advanced key principles underlying and justifying the use of time-frequency (t,f) methods. The topic is covered in 11 focused sections with appropriate internal cross-referencing to this and other chapters. The issues and techniques described in this chapter were selected to provide different facets and perspectives to allow for a deeper insight into the foundations of the field and its methodologies.

The chapter begins by relating the class of quadratic TFDs to time-scale methods such as wavelets, as an extension of [Section 2.3.3](#); it further motivates the use of quadratic TFDs in most situations and not just in exploratory signal analysis ([Section 4.1](#)). Then, the key issue of cross-terms generation and their localization are described in detail. The quadratic superposition principle is used to explain the mechanism generating the cross-terms and the subsequent trade-off between cross-term reduction and increased localization as well as (t,f) resolution ([Section 4.2](#)). This is followed by an examination of the covariance property of TFDs for important signal transformations like (t,f) shifts or scaling, a key characteristic in some applications ([Section 4.3](#)). Another significant aspect of (t,f) methods is that the (t,f) uncertainty relations determine the issue of lower bounds in achievable (t,f) resolution ([Section 4.4](#)). Using methods such as coordinate change methods, we can also define joint distributions of other variables than t and f that may be better suited for specific applications ([Section 4.5](#)). Having calculated and plotted the (t,f) representation of a given signal, we wish to make precise measurements and estimate signal parameters directly from the (t,f) plane. Formulations of measures such as spread measures are then provided ([Section 4.6](#)). (t,f) methods can also be used to describe linear time-varying input-output relationships ([Section 4.7](#)). The relationships between (t,f) methods such as the WVD and the fractional FT is described using the Radon-Wigner transform (RWT) ([Section 4.8](#)). Then, the next two sections focus on a (t,f) perspective of MIMO dynamical systems ([Section 4.9](#)) and Teager-Kaiser (TK) operators in (t,f) analysis ([Section 4.10](#)). This is followed by the presentation of the Gabor spectrogram ([Section 4.11](#)), which relates the properties of the WVD to energy atoms. The chapter ends with a presentation of the empirical mode decomposition and Hilbert spectrum.

4.1 RELATIONSHIPS BETWEEN QUADRATIC TFDs AND TIME-SCALE REPRESENTATIONS⁰

4.1.1 INTRODUCTION

Section 2.7.7 indicated that time-frequency (t,f) methods can be grouped into two large sets: “amplitude” (linear) “atomic” decompositions and “energy” quadratic (or bilinear) distributions (see Fig. 2.7.3). The linear class is concerned with representing a signal as a linear combination of a family of elementary signals (called atoms) that are well localized both in time and in frequency. The linear class includes the short-time Fourier transform (STFT) and the wavelet transform (WT). Figure 2.7.3 illustrates the relationship between both. The bilinear class includes quadratic TFDs (QTFDs) that distribute the energy of the signal between two variables, namely time and frequency (or time and scale). The STFT forms a link between these two classes. Each energy QTFD can be expressed as a filtered version of the Wigner-Ville distribution (WVD). An energy QTFD can be, hence, expressed using the WVD or equivalently the STFT. In particular, the spectrogram is the squared modulus of the STFT and also a special smoothed (or filtered) version of the WVD as shown in Eqs. (2.7.7) and (2.7.8) and Table 3.3.2. Based on the principles of covariance, two large subclasses of QTFDs can be identified: the shift covariant (also known as quadratic) subclass and the scale covariant (also known as the affine) subclass. The next sections establish a link between all these classes, for example, between time-scale methods such as wavelets and energy distributions such as QTFDs formed by (t,f) smoothed WVDs.

4.1.2 LINEAR OR ATOMIC DECOMPOSITIONS

4.1.2.1 Short-time Fourier transform

The linear STFT introduces the temporal dependence to the spectral representation by using a window function h and replacing the pure waves of the Fourier transform (FT) with localized wave packets given by

$$h_{t,f}(u) = h(u - t) e^{j2\pi fu}. \quad (4.1.1)$$

Using these wave packets, we then obtain the following joint (t,f) representation:

$$F_x^h(t,f) = \langle x, h_{t,f} \rangle = \int_{-\infty}^{\infty} x(u) h^*(u - t) e^{-j2\pi fu} du, \quad (4.1.2)$$

where $\langle ., . \rangle$ stands for the inner product. The STFT is then just the projection of the analyzed signal $x(t)$ onto a family of atoms $\{h_{t,f}, t, f \in \mathbb{R}\}$ derived from one unique 1D element $h(t)$ by (t,f) -shifts. The invertibility (closure) condition indicates that the STFT is an admissible representation if $h(t)$ has unity energy; that is

$$\int_{-\infty}^{\infty} |h(t)|^2 dt = 1 \Rightarrow x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_x^h(u,f) h_{u,f}(t) du df. \quad (4.1.3)$$

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Touati**, King Saud University, Riyadh, Kingdom of Saudi Arabia (stouati@ksu.edu.sa). Reviewers: A. Akan, L. Rankine, B.K. Jawad, S. Ouelha, and N. Stevenson.

Due to its continuous nature, the STFT is a redundant representation. This redundancy can be reduced by (t, f) discretization. This leads to the formulation

$$F_x^h[n, k] = F_x^h[nt_0, kf_0] = \int_{-\infty}^{\infty} x(u)h^*(u - nt_0) e^{-j2\pi kf_0 u} du, \quad (4.1.4)$$

where $n, k \in \mathbb{Z}$ and t_0 and f_0 represent the mesh size in time and frequency, respectively. The Balian-Low theorem shows that, assuming h is well localized in both time and frequency (i.e., the second moments of h in time and frequency are finite), the family of atoms $\{h_{n,k}(t) = h(t - nt_0) e^{j2\pi kf_0 t}\}$ used in the discrete STFT cannot constitute an orthogonal basis. So, this family cannot easily be used to recover the time domain signal from its (t, f) coefficients $F_x^h[n, k]$ by simple projection.

The Gabor transform is a special case of discrete STFT with $h(t)$ given by [1] (see Section 2.3.3)

$$h(t) = (\pi t_0)^{-1/4} \exp\left(-\frac{t^2}{t_0^2}\right). \quad (4.1.5)$$

When the Gabor filter is modulated by a complex exponential, it becomes a 1D Gabor wavelet, which is a special case of continuous wavelets [2].

4.1.2.2 Continuous wavelets

A time-scale representation is obtained using atoms generated from a function $h(u)$ (called a wavelet) through the following mapping:

$$h(u) \rightarrow h_{t,a}(u) = |a|^{-1/2} h\left(\frac{u-t}{a}\right). \quad (4.1.6)$$

This transformation corresponds to a shift in time and scaling (stretching $|a| > 1$ or compression $|a| < 1$). The term $|a|^{-1/2}$ is added for normalization. A special case of wavelet functions, closely related to human perception, both hearing and vision [3], is the Gabor wavelet (or Morlet wavelet) described as follows:

$$h_{t,a}(u) = |a|^{-1/2} \exp\left(-\frac{(u-t)^2}{a^2}\right) \exp\left(-\frac{jk_0(u-t)}{a}\right), \quad (4.1.7)$$

where k_0 controls the rate of modulation.

The continuous wavelet transform (WT) of a signal $x(t)$ is given by [2]:

$$\text{WT}_x(t, a) = \langle x, h_{t,a} \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(u)h^*\left(\frac{u-t}{a}\right) du. \quad (4.1.8)$$

For an orthogonal wavelet h [2] (i.e., h and its dual function \tilde{h} are equal¹), an admissibility condition (sufficient but not necessary) that ensures the invertibility of Eq. (4.1.8) is [4]:

$$\int_{-\infty}^{\infty} |H(f)|^2 \frac{df}{|f|} = 1, \quad (4.1.9)$$

¹This condition is specific to the case of orthogonal wavelet: where $h = \tilde{h}$. Otherwise, $|H(f)|^2$ should be replaced by $H(f)^* \tilde{H}(f)$ in Eq. (4.1.9), and h by \tilde{h} in Eq. (4.1.10).

where $H(f)$ is the FT of $h(t)$. This condition implies that: (1) $H(f)$ decreases at least as fast as $|f|^{-1/2}$ and (2) $H(0) = \int_{-\infty}^{\infty} h(t) dt = 0$ ($h(t)$ is zero mean). Inversely, when $h(t)$ is zero mean and $\int_{-\infty}^{\infty} (1 + |t|^{\alpha}) |h(t)| dt$ is finite for some $\alpha > 0$, then the WT is invertible and the relation known as resolution of identity or Calderon's reproducing identity, reads:

$$x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} \text{WT}_x(u, a) h\left(\frac{t-u}{a}\right) du \frac{da}{a^2}. \quad (4.1.10)$$

The WT can be interpreted as a (t, f) representation. Assume that $H(f)$, is unimodal and localized to a neighborhood of a frequency f_0 , used as a reference for the scale $a = 1$ (i.e., if $a = 1 \rightarrow f = f_0$). This allows the interpretation of the WT as a function of time and frequency using $f = f_0/a$. In other words, one gets the wavelet transform $\widehat{\text{WT}}_x$ as a function of time and frequency:

$$\widehat{\text{WT}}_x(t, f) \equiv \text{WT}_x\left(t, \frac{f_0}{f}\right). \quad (4.1.11)$$

4.1.2.3 Relationship between STFT and WT

The WT contains the same information as the STFT and we can go from one to the other without loss of information. These two representations are related as follows:

$$\text{WT}_x(t, a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_x^h(u, f; h^{\text{ST}}) \langle h_{u,f}^{\text{ST}}, h_{t,a}^W \rangle du df, \quad (4.1.12)$$

$$F_x^h(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WT}_x(u, a; h^W) \langle h_{u,a}^W, h_{t,f}^{\text{ST}} \rangle \frac{du da}{a^2}, \quad (4.1.13)$$

where h^{ST} is the window used in the STFT and h^W is the window (wavelet) used in the WT. Equation (4.1.12) is obtained by combining Eqs. (4.1.8) and (4.1.3); and Eq. (4.1.13) is obtained from Eqs. (4.1.2) and (4.1.10) (see Section 4.1.6 for proofs). The signal energy is given by:

$$E_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F_x^h(t, f)|^2 dt df = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\text{WT}_x(t, a)|^2 \frac{dt da}{a^2}.$$

4.1.2.4 The discrete wavelet transform

A natural way to sample the parameters a and t is to use a logarithmic discretization of the a scale and link this, in turn, to the size of steps taken between t consecutive locations. To link t to a we move in discrete steps to each location t which is proportional to the a scale. This “discretization” of the mother wavelet h leads to:

$$\psi_{d,c}(u) = \frac{1}{\sqrt{a_0^d}} h\left(\frac{u - cb_0 a_0^d}{a_0^d}\right) = a_0^{-d/2} h(a_0^{-d} u - cb_0), \quad (4.1.14)$$

where the integers d and c are related to the wavelet dilation and translation, respectively. The constants a_0 and b_0 are the fixed dilation step and the (temporal) location parameter, respectively. The size of the translation steps, $\Delta t = b_0 a_0^d$, is directly proportional to the wavelet scale a_0^d .

A common choice for discrete wavelet parameters is: $a_0 = 2$ and $b_0 = 1$ [2]. This power-of-two logarithmic scaling is known as the dyadic grid arrangement. This dyadic grid is the most computationally efficient discretization that results in an orthonormal wavelet basis (orthonormal across both scale and time dilatation). The dyadic grid wavelet can be written as [2]

$$\psi_{d,c}(u) = \frac{1}{\sqrt{2^d}} \psi\left(\frac{u - 2^d c}{2^d}\right) = 2^{-d/2} \psi(2^{-d} u - c). \quad (4.1.15)$$

For the case of discrete wavelet transform (DWT) for discrete-time, to convert from scale to frequency, the same procedure discussed above Eq. (4.1.11) can be used. The reference frequency f_0 can be chosen as a characteristic frequency of the mother wavelet, such as the spectral peak or band-pass center frequency [5]. The scale depending frequency f_d becomes²

$$f_d = \frac{f_0}{2^d}. \quad (4.1.16)$$

4.1.3 ENERGY QUADRATIC TFDs AND TSEDs

Two classes of energy distributions are considered: the time-frequency distributions (TFDs) (t, f) and time-scale energy distributions (TSEDs). As energy is quadratic by nature, an energy distribution in a 2D domain can be written as:

$$D_x(t, \lambda) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(u, v, t, \lambda) x(u) x^*(v) du dv, \quad (4.1.17)$$

where D is a generic form that stands for “distribution” and the variable λ can be either f for the case of TFDs or a for the case of TSEDs. To determine the kernel $K(u, v, t, \lambda)$, constraints need to be imposed on the distribution $D_x(t, \lambda)$. To define the two classes mentioned above, two different constraints are used: (t, f) shift covariance and time shift and dilation (affine) covariance.

4.1.3.1 Bilinear (t, f) distributions

Let us consider the following (t, f) shift operation

$$x(t) \rightarrow x_{t', f'}(t) = x(t - t') e^{j2\pi f' t}. \quad (4.1.18)$$

Then the covariance condition reads:

$$D_{x_{t', f'}}(t, f) = D_x(t - t', f - f'). \quad (4.1.19)$$

Applying Eq. (4.1.19) to Eq. (4.1.17) and using the notation “ $\int_{\mathbb{R}^2}$ ” yields (see Section 4.1.6):

$$D_x(t, f) = \int_{\mathbb{R}^2} K\left(u - t + \frac{\tau}{2}, u - t - \frac{\tau}{2}, 0, 0\right) x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau du. \quad (4.1.20)$$

²Because of the nonunique way of choosing f_0 , the frequencies f_d have been referred to as pseudo-frequencies [6].

Let us consider the following definitions:

$$\gamma(t, f) = \int_{-\infty}^{\infty} K\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}, 0, 0\right) e^{-j2\pi f\tau} d\tau, \quad (4.1.21)$$

$$W_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau, \quad (4.1.22)$$

where W_x is the WVD. Using Parseval's formula yields (see [Section 4.1.6](#)):

$$D_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(u - t, \xi - f) W_x(u, \xi) du d\xi := \rho_x(t, f), \quad (4.1.23)$$

which is the class of QTFDs defined as smoothed WVD as it is the same as Eqs. [\(2.2.15\)](#) and [\(3.2.13\)](#), which can be rewritten as:

$$\rho_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(v, \tau) x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) e^{j2\pi(vt - vu - f\tau)} dv du d\tau, \quad (4.1.24)$$

where $g(v, \tau)$ is the Doppler-lag kernel defined in Eq. [\(3.2.6\)](#) as the 2D FT of $\gamma(t, f)$.

Note: this characterization of (t, f) shift covariant class assumes that the kernel $g(v, \tau)$ or $\gamma(t, f)$ does not depend on the signal; that is, it is signal independent.

The energy constraint: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x(t, f) dt df = E_x$ requires that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(t, f) dt df = 1 \quad (4.1.25)$$

or equivalently, as per [Section 4.1.6](#):

$$g(0, 0) = 1. \quad (4.1.26)$$

4.1.3.2 Bilinear time-scale distributions

The formulation of bilinear time-scale representations follows the same steps used for QTFDs apart from replacing the covariance relative to the (t, f) shifts with the covariance relative to “shift and dilation.” The time-shift and dilation operation applied on an arbitrary signal $x(t)$ is given by

$$x(t) \rightarrow x_{t', a'}(t) = \frac{1}{\sqrt{|a'|}} x\left(\frac{t - t'}{a'}\right). \quad (4.1.27)$$

The constraint of covariance becomes

$$D_{x_{t', a'}}(t, a) = D_x\left(\frac{t - t'}{a'}, \frac{a}{a'}\right). \quad (4.1.28)$$

Following the same steps as in the case of (t, f) distributions, we get

$$D_x(t, a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma\left(\frac{u - t}{a}, a\xi\right) W_x(u, \xi) du d\xi, \quad (4.1.29)$$

where

$$\gamma(u, f) = \int_{-\infty}^{\infty} K\left(u + \frac{\tau}{2}, u - \frac{\tau}{2}, 0, 1\right) e^{-j2\pi f\tau} d\tau, \quad (4.1.30)$$

where K is the generic kernel defined in Eq. (4.1.17). By analogy to the TFDs case (Eq. 4.1.18), the scale parameter is replaced here by 1 because the constraint of covariance Eq. (4.1.28) reads as scale normalization, while Eq. (4.1.19) reads as frequency shifting. This parameterization constitutes the “affine class” of bilinear representations, which are covariant relative to translations-dilations. As $D_x(t, a)$ is an energy distribution, the following equation needs to be verified (see Section 4.1.6):

$$\int_{-\infty}^{\infty} \mathcal{G}(0, f) \frac{df}{|f|} = 1, \quad (4.1.31)$$

where $\mathcal{G}(\nu, f)$ denotes the FT of $\gamma(t, f)$ along the time axis, as defined in Eq. (3.2.6)

$$\mathcal{G}(\nu, f) = \int_{-\infty}^{\infty} \gamma(t, f) e^{-j2\pi \nu t} dt. \quad (4.1.32)$$

Like the (t, f) shift covariant class, the affine class has different equivalent forms depending on the representational space chosen to express the kernel, for example:

$$D_x(t, a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma\left(\frac{u-t}{a}, a\xi\right) W_x(u, \xi) du d\xi, \quad (4.1.33)$$

$$D_x(t, a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g\left(av, \frac{\tau}{a}\right) A_x(v, \tau) e^{+j2\pi(vt)} dv d\tau, \quad (4.1.34)$$

where A_x is the symmetric ambiguity function (AF) (see Eqs. (3.2.6) and (3.2.10)).

Note: the WVD belongs to both bilinear classes (i.e., quadratic and affine) and the other TFDs within these classes can be expressed as filtered WVDs.

By applying Moyal’s formula and the scaling and translation properties of WVD, the scalogram, which is the square of the WT, can be related to the WVD as [7]

$$|\text{WT}_x(t, a)|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_h\left(\frac{u-t}{a}, a\xi\right) W_x(u, \xi) du d\xi, \quad (4.1.35)$$

where W_h is the WVD of the wavelet (i.e., window) h . This result is similar to that obtained by the spectrogram (square of the STFT) when expressed as a filtered version of the WVD. Equation (4.1.35) can be generalized to TFDs other than WVD if the ambiguity domain kernel, g , of these TFDs satisfies two conditions: (1) $g(v, \tau) = g_0(v\tau)$ and (2) $|g_0(v\tau)| = 1$ (e.g., with the Gaussian kernel). This yields³

$$|\text{WT}_x(t, a)|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_h^*\left(\frac{u-t}{a}, a\xi\right) \rho_x(u, \xi) du d\xi, \quad (4.1.36)$$

where ρ_h is the TFD of the wavelet/window h whose ambiguity domain kernel satisfies the above two conditions [7,8].

³A similar relation exists for the case of spectrogram when the ambiguity domain kernel has a modulus of one; that is, $|g(v, \tau)| = 1$.

4.1.3.3 Relationship between the two bilinear classes

As mentioned above, by introducing an arbitrary nonzero reference frequency f_0 , needed for the formal identification “scale = inverse of frequency,” we can have the following relationship [9]:

$$f = f_0/a. \quad (4.1.37)$$

Using Eq. (4.1.37), we obtain a (t, f) representation by rewriting Eq. (4.1.33) as

$$\hat{\rho}_x(t, f) = D_x \left(t, \frac{f_0}{f} \right). \quad (4.1.38)$$

In practice, to convert from scale to frequency, a characteristic frequency of the wavelet/window is used. One of the most commonly used characteristic frequencies is the band-pass center f_c of the wavelet’s power spectrum defined as:

$$f_c = \sqrt{\frac{\int_0^\infty f^2 |H(f)|^2 df}{\int_0^\infty |H(f)|^2 df}}, \quad (4.1.39)$$

where $H(f)$ is the FT of the wavelet $h(t)$ [9]. Other representative frequencies of the wavelet (e.g., spectral peak frequency or central frequency) could be used.

Special cases. Suppose that the kernel g in Eq. (4.1.34) has the following form:

$$g(v, \tau) = g_0(v\tau) e^{-j2\pi f_0 \tau}. \quad (4.1.40)$$

This leads to the following TSED:

$$\varrho_x(t, a; g) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_0(v\tau) A_x(v, \tau) e^{j2\pi(vt - f_0\tau/a)} dv d\tau = \rho_x \left(t, \frac{f_0}{a}; g_0 \right). \quad (4.1.41)$$

This last expression is the QTFD in the (t, f) shift covariant class associated with the parameterization of $g_0(v\tau)$ and the identification $f = f_0/a$.

Another particular class of time-scale representations is the affine Wigner distributions. Their parameter function admits the generic form

$$g(v, \tau) = M(v) e^{-j2\pi N(v)\tau}, \quad (4.1.42)$$

where $M(v)$ and $N(v)$ are two real, positive, and continuous functions. As previously, by substituting $f = f_0/a$, every TSED ϱ_x of this type (see Eq. (4.1.41)) can be associated with a (t, f) distribution $\hat{\rho}_x$ as follows

$$\hat{\rho}_x(t, f; \eta) \equiv \varrho_x(t, f_0/a; g). \quad (4.1.43)$$

Members of the affine Wigner class include other distributions, as reported in [8], including those discussed in [Section 7.1](#).

4.1.4 QTFDs, WAVELET TRANSFORM, AND SCALOGRAM: RELATIONSHIPS

The relationship between QTFDs given in Eq. (3.2.13) and the WT (time-scale method), and scalogram can be further clarified by noting that Eq. (4.1.11) gives a (t, f) interpretation of wavelet. [Section 4.1.2.3](#) established the different relationships between the STFT and WT and consequently between the spectrogram and the scalogram; the relationship between QTFDs and TSED (e.g., WTs) results, as

the spectrogram is a smoothed (or filtered) WVD. Section 4.1.3.3 discussed the relationships between the shift-invariant (quadratic) class and the time-dilation invariant (affine) class of (t,f) representations. In the 2D case, both QTFDs and wavelets can be extended to deal with images (see, e.g., Section 15.4 for 2D-WVD and [10] for the WT applied to human recognition using the iris).

4.1.5 SUMMARY AND CONCLUSIONS

Regarding the link between frequency and scale, the inversely proportional relationship between the two is well-established, but the exact formula is still arbitrary. The constant of proportionality is usually chosen as one of the parameters of the wavelet function. This indicates that the knowledge of the wavelet generating the (t,f) representation is essential for a frequency interpretation of the results.

The strong relationship between the STFT and the WT is clearly illustrated by Eqs. (4.1.12) and (4.1.13).

The common origin of the two bilinear classes of (t,f) representations is clearly given by Eq. (4.1.17). What differentiates the two classes is the form of the kernel obtained when imposing the different covariance constraints.

The WVD has a central role in both shift-invariant and affine classes of bilinear (t,f) representations. Every other member of these two classes can be generated by smoothing the WVD in different ways.

Taking everything into account, one can conclude that QTFDs appear intuitively to be more appealing given their direct physical meaning and interpretation. For these reasons, this book has more focus on QTFDs.

4.1.6 PROOFS

Proof of Eq. (4.1.12): Substituting Eq. (4.1.3) $x(t) = \iint_{-\infty}^{\infty} F_x^h(u,f) h_{uf}(t) du df$ into Eq. (4.1.8), and changing slightly the notations for h^{ST} , the window used in the STFT, and h^W used in the WT, gives:

$$\text{WT}_x(t,a) = \iint_{-\infty}^{\infty} F_x^h(u,f; h^{\text{ST}}) h_{uf}^{\text{ST}}(v) \underbrace{\frac{1}{\sqrt{|a|}} h^* \left(\frac{v-t}{a} \right)}_{\langle h_{uf}^{\text{ST}}, h_{t,a}^W \rangle} du df dv \quad (4.1.44)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_x^h(u,f; h^{\text{ST}}) \langle h_{uf}^{\text{ST}}, h_{t,a}^W \rangle du df, \quad (4.1.45)$$

which equals Eq. (4.1.12). \square

Proof of Eq. (4.1.13): Following the same method, by substituting Eq. (4.1.10) $x(t) = \iint_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} \text{WT}_x(u,a) h \left(\frac{t-u}{a} \right) du \frac{da}{a^2}$ into Eq. (4.1.2) we get:

$$F_x^h(t,f) = \iint_{-\infty}^{\infty} \text{WT}_x(u,a) \underbrace{\frac{1}{\sqrt{|a|}} h \left(\frac{v-u}{a} \right) h^*(v-t,f)}_{\langle h_{u,a}^W, h_{t,f}^{\text{ST}} \rangle} du \frac{da}{a^2} dv \quad (4.1.46)$$

$$= \int_{-\infty}^{\infty} \text{WT}_x(u,a; h^W) \langle h_{u,a}^W, h_{t,f}^{\text{ST}} \rangle du \frac{da}{a^2}, \quad (4.1.47)$$

which equals Eq. (4.1.13). \square

Proof of Eq. (4.1.20): By applying Eq. (4.1.19) to Eq. (4.1.17) and choosing $t' = t$ and $f' = f$, we get the following constraint on the kernel K :

$$K(u, v, t, f) = K(u - t, v - t, 0, 0) \exp(-j2\pi f(u - v)).$$

Then, by usual change of variables we get:

$$D_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K\left(u - t + \frac{\tau}{2}, u - t - \frac{\tau}{2}, 0, 0\right) x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau du, \quad (4.1.48)$$

which equals Eq. (4.1.20). \square

Proof of Eq. (4.1.23): As γ and W_x are FTs w.r.t. τ , the Parseval formula leads to the equality between the inner integrals in Eq. (4.1.20) (w.r.t. τ) and the inner integral in Eq. (4.1.23) (w.r.t. ξ), which is the convolution of γ and W_x . Then, we obtain Eq. (4.1.23). \square

Proof of Eq. (4.1.26): By integrating Eq. (4.1.23) w.r.t t and f , and using the energy property $\iint_{-\infty}^{\infty} W_x(t, f) dt df = E_x$ satisfied by the WVD, we obtain

$$\iint_{-\infty}^{\infty} \rho_x(t, f) dt df = E_x \cdot \iint_{-\infty}^{\infty} \gamma(t, f) dt df, \quad (4.1.49)$$

which equals Eq. (4.1.26) (see also the next proof of Eq. (4.1.31)). \square

Proof of Eq. (4.1.31): By integrating Eq. (4.1.29) w.r.t. t and a , one get

$$E_x = \iiint_{-\infty}^{\infty} \mathcal{G}(0, f) W_x(u, \xi) du d\xi \frac{df}{|f|}, \quad (4.1.50)$$

where $\mathcal{G}(v, f)$ denotes the FT of $\gamma(t, f)$ along the time axis; that is

$$\mathcal{G}(v, f) = \int_{-\infty}^{\infty} \gamma(t, f) e^{-j2\pi vt} dt. \quad (4.1.51)$$

Using the property $\iint_{-\infty}^{\infty} W_x(t, f) dt df = E_x$, the energy constraint equation (4.1.50) can be written

$$\int_{-\infty}^{\infty} \mathcal{G}(0, f) \frac{df}{|f|} = 1, \quad (4.1.52)$$

which equals Eq. (4.1.31). \square

4.2 CROSS-TERMS AND LOCALIZATION IN QUADRATIC TIME-FREQUENCY DISTRIBUTIONS⁰

4.2.1 IDENTIFYING CROSS-TERMS

4.2.1.1 Cross-terms in spectrum analysis

Let $x(t)$ and $y(t)$ be any two signals, and a and b any two complex numbers. If $X(f) := \mathcal{F}_{t \rightarrow f}\{x(t)\}$ stands for the FT of $x(t)$ —with a similar notation for $y(t)$ —we then have

$$|\mathcal{F}_{t \rightarrow f}\{ax(t) + by(t)\}|^2 = |a|^2|X(f)|^2 + |b|^2|Y(f)|^2 + 2\operatorname{Re}\{ab^*X(f)Y^*(f)\}. \quad (4.2.1)$$

This means that the spectrum energy density of the sum of two signals does not reduce to the sum of the individual densities (unless the signals are spectrally disjoint), but satisfies a quadratic superposition principle involving a third term, referred to as a *cross-term*. While this situation is, mathematically speaking, a necessary by-product of the quadratic nature of the transformation and of the elementary identity $(a + b)^2 = a^2 + b^2 + 2ab$, it also corresponds to a physical reality in terms of interfering waves. This can be illustrated by the simple example where a delayed replica is superimposed to a given waveform. Assuming the model $y(t) = x(t - \tau)$, with $a = b = 1$, we get:

$$|\mathcal{F}_{t \rightarrow f}\{x(t) + x(t - \tau)\}|^2 = 4(\cos \pi \tau f)^2|X(f)|^2, \quad (4.2.2)$$

and the existence of a cross-term is instrumental in the creation, in the composite spectrum, of “fringes” whose periodicity in f is controlled by the delay τ : it is in fact the basis of interferometry.

4.2.1.2 Cross-terms, from spectrograms to the WVD

Switching to TFDs, similar considerations apply naturally to spectrograms defined by $S_x(t, f) := |F_x(t, f)|^2$, with

$$F_x(t, f) := \mathcal{F}_{s \rightarrow f}\{x(s)h^*(s - t)\}, \quad (4.2.3)$$

where $h(t)$ is a short-time analysis window. We get in this case

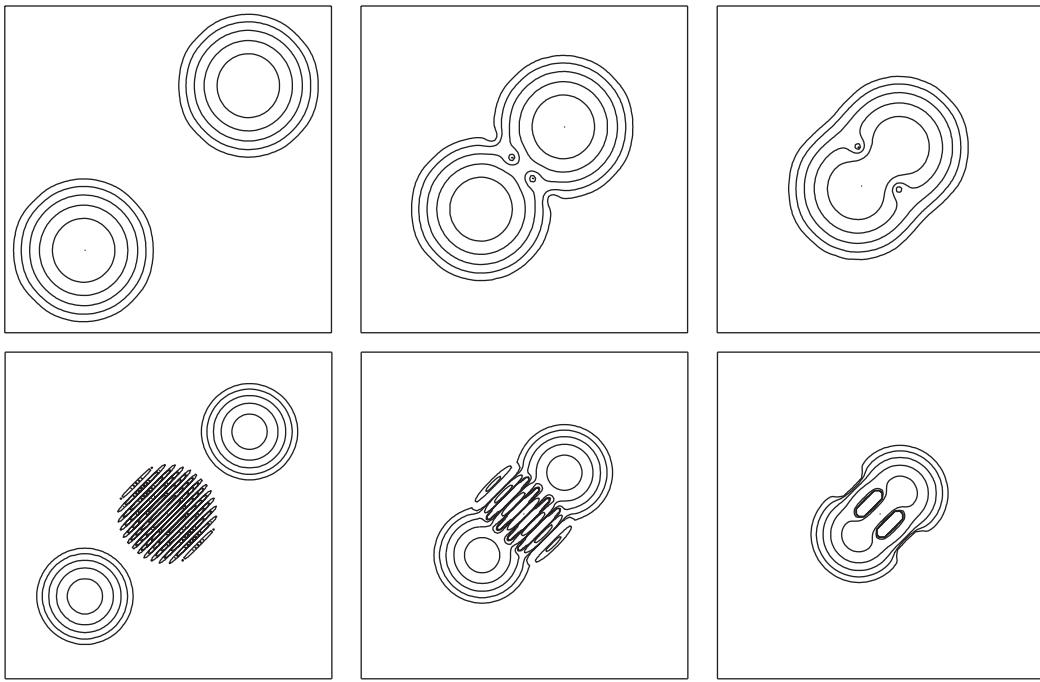
$$S_{ax+by}(t, f) = |a|^2S_x(t, f) + |b|^2S_y(t, f) + 2\operatorname{Re}\{ab^*F_x(t, f)F_y^*(t, f)\}, \quad (4.2.4)$$

and cross-terms show up as long as the (essential) time-frequency (t, f) supports of the STFTs $F_x(t, f)$ and $F_y(t, f)$ overlap (see Fig. 4.2.1). A similar situation is observed with the scalograms (i.e., the squared modulus of the wavelet transform).

As squared linear transforms, spectrograms happen, however, to be a restricted version of quadratic transforms. In fact, they only correspond to a special case of the more general quadratic class, which consists of all quadratic (t, f) energy distributions covariant to time and frequency shifts, and whose most prominent member is the WVD [11]:

$$W_x(t, f) := \mathcal{F}_{\tau \rightarrow f}\{x(t + \tau/2)x^*(t - \tau/2)\}. \quad (4.2.5)$$

⁰Author: **P. Flandrin**, École Normale Supérieure de Lyon, Lyon Cedex 07, France (flandrin@ens-lyon.fr). Reviewers: F. Hlawatsch and W.J. Williams.

**FIGURE 4.2.1**

Cross-terms in the (t, f) analysis of two shifted Gaussian wave packets. In the spectrogram case (top row), cross-terms are only present when the effective supports of the individual contributions overlap, and they fade out when the distance between the components increases. In the Wigner-Ville case (bottom row), cross-terms persistently exist as an oscillating contribution located midway between the two components, with an increasing oscillation rate when the inter-components distance increases. (In all diagrams, time is horizontal, frequency is vertical, and isocontours of the (positive parts of the) distributions are displayed.)

The WVD itself satisfies a quadratic superposition principle according to which

$$W_{ax+by}(t,f) = |a|^2 W_x(t,f) + |b|^2 W_y(t,f) + 2\operatorname{Re}\{ab^* W_{x,y}(t,f)\} \quad (4.2.6)$$

with

$$W_{x,y}(t,f) := \mathcal{F}_{\tau \rightarrow f}\{x(t + \tau/2)y^*(t - \tau/2)\}. \quad (4.2.7)$$

As compared to spectrograms, the situation is, however, drastically changed in the sense that increasing the (time-frequency) distance between the components no longer implies a vanishing of the cross-term controlled by Eq. (4.2.7). More precisely, if we let

$$x_{\pm}(t) := x_0(t \mp \Delta t/2) \exp\{\pm i2\pi(\Delta f/2)t\}, \quad (4.2.8)$$

we get (see Fig. 4.2.1):

$$W_{x_+ + x_-}(t, f) = W_{x_0}(t - \Delta t/2, f - \Delta f/2) + W_{x_0}(t + \Delta t/2, f + \Delta f/2) + I(t, f) \quad (4.2.9)$$

with

$$I(t, f) = 2W_{x_0}(t, f) \cos[2\pi(t\Delta f - f\Delta t)]. \quad (4.2.10)$$

This simplified situation exemplifies the general rules, which control the *interference geometry* of the WVD [12,13], whose cross-terms

1. are located *midway* between the interacting components;
2. *oscillate* proportionally to the inter-components' (t, f) distance;
3. have a direction of oscillation *orthogonal* to the straight line connecting the components.

4.2.2 REDUCING CROSS-TERMS

In the case of multicomponent signals of the form

$$x(t) = \sum_{n=1}^N x_n(t), \quad (4.2.11)$$

WVD cross-terms are created in between any two components, leading to a combinatorial proliferation that quickly hampers readability. Cross-terms' reduction is, therefore, an important issue in many applications. Since cross-terms appear between any two contributions, at positive as well as negative frequencies, a first recipe is to compute WVDs on analytic signals rather than on real-valued ones, as described in Section 2.1.4 (see also Section 1.2.2). The reason is that the recourse to the analytic signal forces spectrum contributions to be zero on the real half-line of negative frequencies, thus suppressing, *de facto*, cross-terms stemming from interactions between negative and positive frequencies.

4.2.2.1 The ambiguity domain interpretation

A further way of reducing WVD cross-terms' amounts to identifying them as oscillating contributions (as opposed to the smoother variations attached to "signal terms"), thus suggesting to make use of a low-pass smoothing. The spectrogram case discussed above is in fact a first instance of such a procedure, since it is well-known (e.g., [11]) that:

$$S_x(t, f) = \iint_{-\infty}^{+\infty} W_x(s, \xi) W_h(s - t, \xi - f) ds d\xi, \quad (4.2.12)$$

with the low-pass nature of the window $h(t)$ carrying over to its WVD $W_h(t, f)$. More generally, replacing in Eq. (4.2.12) $W_h(t, f)$ by some (almost) arbitrary function $\gamma(t, f)$ defines the quadratic class [11]:

$$\rho_x(t, f; \gamma) := \iint_{-\infty}^{+\infty} W_x(s, \xi) \gamma(s - t, \xi - f) ds d\xi, \quad (4.2.13)$$

and offers a versatile framework for achieving cross-terms reduction.

Among the many possible parameterizations of the quadratic class, one that proves most powerful for the considered problem operates in the so-called *ambiguity plane*, obtained by Fourier duality from the (t,f) plane (cf. Sections 3.2.1 and 5.1). Starting from the general form equation (4.2.13) of the quadratic class with (time-frequency) kernel $\gamma(t,f)$, we obtain by a 2D Fourier transformation that

$$\rho_x(t,f;\gamma) = \mathcal{A}_x(v,\tau;g) := g(v,\tau)A_x(v,\tau), \quad (4.2.14)$$

with $g(v,\tau)$ the 2D inverse FT of $\gamma(t,f)$ and¹

$$A_x(v,\tau) := \mathcal{F}_{t \rightarrow v}\{x(t + \tau/2)x^*(t - \tau/2)\} \quad (4.2.15)$$

the AF of $x(t)$: in the ambiguity plane, the quadratic class is nothing but a weighted AF.

In the case of the two-component signal equation (4.2.8), we have²

$$A_{x_+ + x_-}(v,\tau) = 2A_{x_0}(v,\tau) \cos[\pi(v\Delta t + \tau\Delta f)] + J(v,\tau) \quad (4.2.16)$$

with

$$J(v,\tau) = A_{x_0}(v + \Delta f, \tau - \Delta t) + A_{x_0}(v - \Delta f, \tau + \Delta t). \quad (4.2.17)$$

In accordance with the interpretation of the AF as a (t,f) correlation function (cf. Section 5.1), it thus appears that signal terms are mainly concentrated in the vicinity of the origin of the ambiguity plane, while cross-terms are (symmetrically) located at a distance from the origin, which is equal to the (t,f) distance between the interacting components. A detailed analysis of the shape and orientation of cross-terms for piecewise linear frequency modulated (LFM) signals is given in Section 5.9.

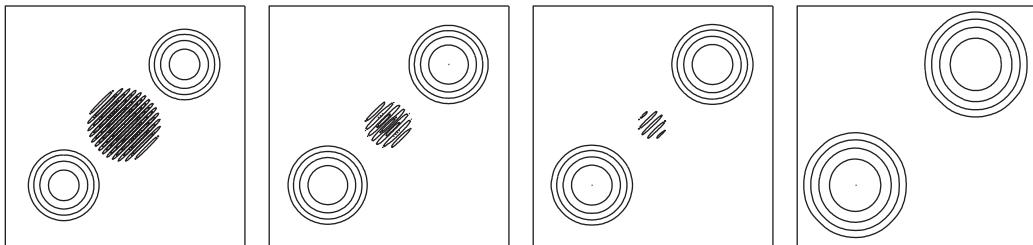
4.2.2.2 Kernel design for reduced interference

The above AF interpretation [14] offers an alternative distinction between “signal terms” and “cross-terms,” and is the clue for a possible reduction of the latter: according to Eq. (4.2.14), it suffices to choose a kernel function $g(v,\tau)$ so as to keep the AF as unchanged as possible in the vicinity of the origin, while suppressing as much as possible contributions off the origin. In a first approximation, the kernel function has, therefore, to be of *low-pass* type, thus guaranteeing that it equivalently operates a smoothing in the (t,f) plane. In this respect, a versatile solution is to make use of a *separable* kernel function of the form $\gamma(t,f) = g_1(t)G_2(f)$. The associated distribution is referred to as a “smoothed pseudo-WVD” [11], and it allows qualitatively for a smooth transition between the WVD (no smoothing, i.e., $\gamma(t,f) = \delta(t)\delta(f)$) and spectrograms (lower bounded smoothing, $\gamma(t,f) = W_h(t,f)$), see Fig. 4.2.2.

More elaborate approaches have been developed, either by incorporating specific constraints to be satisfied by “reduced interference distributions (RID)” (see [15] and Section 5.9) or by making the kernel signal-dependent (see [16] and Section 5.3), so as to reject at best cross-terms’ contributions

¹With the adopted definitions of ρ , g , and A_x , which are slightly different from usual ones, we get in Eq. (4.2.14) $g(v,\tau)A_x^*(v,\tau)$. Otherwise, from Eq. (4.2.13), we can apply $\mathcal{F}_{t \rightarrow v} \mathcal{F}_{f \rightarrow \tau}^{-1}$, we obtain, then, Eq. (4.2.14) as it is, but with $g(v,\tau) = \mathcal{F}_{f \rightarrow \tau}^{-1} \mathcal{F}_f(\gamma(t,f))$ and $A_x(v,\tau) = \mathcal{F}_t(x(t + \tau/2)x^*(t - \tau/2))$.

²These are true when the second definition set is adopted in the previous comment in the footnote.

**FIGURE 4.2.2**

Cross-terms reduction by (t, f) smoothing. Using a smoothed pseudo-WVD with a separable kernel allows for a continuous transition between a WVD (left) and a spectrogram (right).

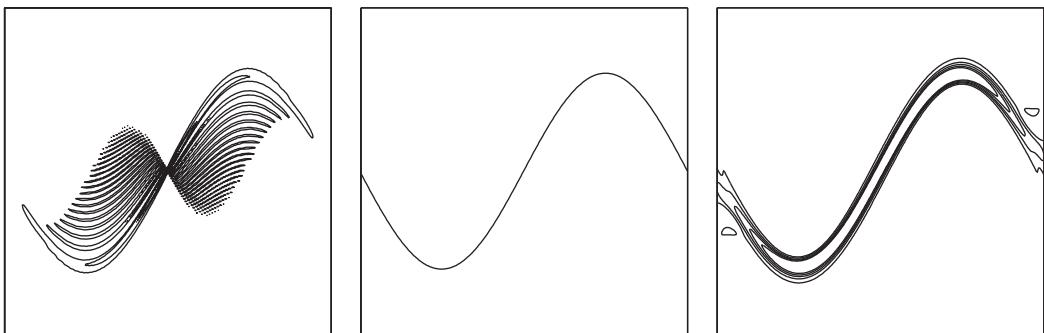
located off the origin in the ambiguity plane. Those different techniques being based on some FT inversion after weighting, they all face a trade-off between interference rejection and localization in the (t, f) plane. An improvement upon this limitation can come from the observation that, in the common situation where the number of amplitude and FM components is limited, the idealized representation (which consists of weighted trajectories on the (t, f) plane) is intrinsically sparse. Methods of optimal recovery from sparsity constraints thus suggested revisiting the issue of (t, f) localization by exploiting sparsity, as adapted to the specific context of (quadratic) distributions. Based on the AF interpretation sketched above, it is known that the relevant information is mostly concentrated in a restricted subset of Fourier coefficients of the WVD neighboring the origin of the AF plane. Using this incomplete information as the primary constraint, it has been shown that a well-localized distribution with reduced interference can follow as the minimum ℓ_1 -norm solution in the transformed domain [17].

4.2.3 CROSS-TERMS AND LOCALIZATION

As presented so far, cross-terms could be thought of as being specific of “multicomponent” situations. However, the notion of a multicomponent signal—as it is modeled in Eq. (4.2.11)—is somehow artificial, in the sense that any signal can always be, at will, split into arbitrary parts. This is especially the case for *FM* signals that intuition (and physics) would like to identify as monocomponent, even if any arbitrary chopping is expected to create cross-terms between the so-created “components.” In such a situation, cross-terms are referred to as “inner interference terms,” in contrast to “outer interference terms” resulting from the interaction between contributions attached to (t, f) domains that are essentially disconnected [12,13].

4.2.3.1 Localization from interference

Iterating *ad infinitum* the procedure of chopping a signal into more and more elementary parts leads to a pointwise application of the WVD cross-term construction rules enounced before. In a nutshell, if we admit to describe a chirp signal by means of its instantaneous frequency (IF) trajectory in the (t, f) plane, inner cross-terms are created midway of the chord between any two points located on this trajectory, with a possible reduction when applying suitable smoothing (see Fig. 4.2.3).

**FIGURE 4.2.3**

Inner interference terms in the (t, f) analysis of a chirp signal. In the WVD case (left), cross-terms appear midway between any two points belonging to the IF trajectory (middle). As for the “outer” interference terms of Fig. 4.2.2, these “inner” interference terms can be reduced by using a smoothed pseudo-WVD (right).

While a more refined description of this mechanism can be obtained on the basis of stationary phase arguments, a companion viewpoint is given by Janssen’s interference formula:

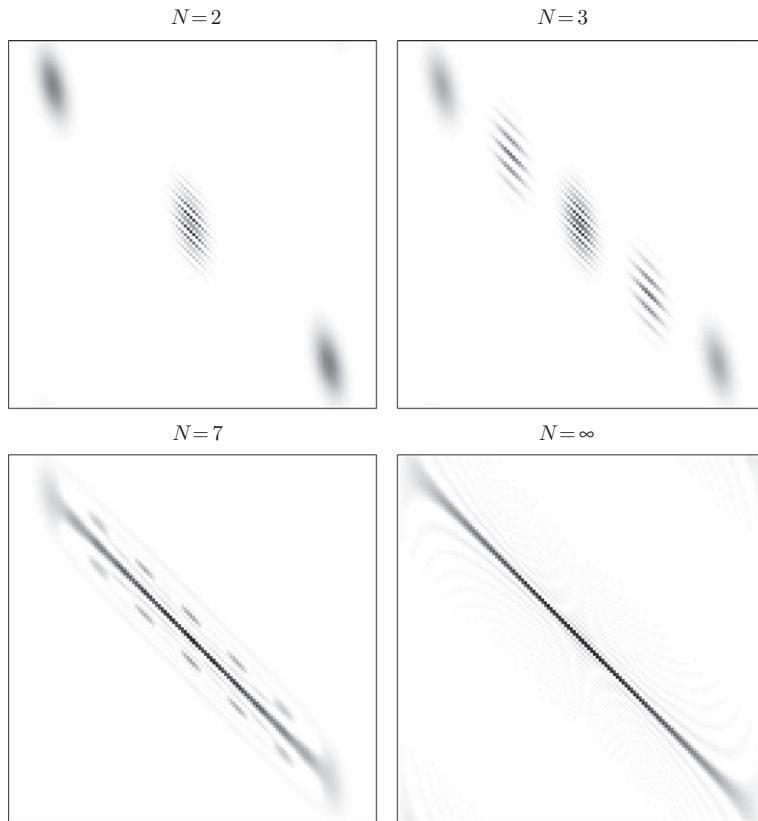
$$W_x^2(t, f) = \iint_{-\infty}^{+\infty} W_x(t + \tau/2, f + \nu/2) W_x(t - \tau/2, f - \nu/2) d\tau d\nu, \quad (4.2.18)$$

according to which a nonzero value of the WVD at a given (t, f) point results from the superposition of all values of the very same distribution that are symmetrically located with respect to the considered point. This means that the WVD is indeed the result of some “holographic” construction that, in some sense, makes impossible a clear-cut distinction between “signal terms” and “cross-terms.”

As a further consequence, *localization* of the WVD for linear chirps [11] can also be viewed as a by-product of interference. The geometric explanation of this well-known property is as follows: if a signal is to have a WVD which is perfectly localized on a given (t, f) curve of the plane, it is then necessary—by application of the interference geometry principle—that all of the mid-points between any two points located on the curve belong themselves to the same curve. It follows that *linear* chirps are the only admissible solutions in the WVD case, since straight lines are the only curves of the plane that are the geometric locus of all of their mid-points (see Fig. 4.2.4). This principle (“a signal term is nothing but the emergence of coherent cross-terms”) is very general and can be extended to other types of quadratic (t, f) distributions (e.g., affine) equipped with suitably modified geometries, thus justifying geometrically their localization on nonlinear chirps [18].

4.2.3.2 The interference/localization trade-off

In ordinary Fourier analysis, and therefore in spectrogram analysis, there exists a necessary trade-off between the resolution in time and the resolution in frequency. In the more general framework of the quadratic class, the situation has to be interpreted differently, with another trade-off between the *joint* resolution (in both time and frequency) on the one hand and the level of cross-terms on the other hand. This clearly appears in the examples of Figs. 4.2.2 and 4.2.3, where a fading out of cross-terms (be they “inner” or “outer”) is achieved only at the expense of spreading out signal terms, that is, a

**FIGURE 4.2.4**

Localization from interference. The localization property of the WVD on straight lines of the (t, f) plane can be seen as resulting from a constructive interference process. Cross-terms of the WVD being located midway between any two interacting components (see Fig. 4.2.1), it is here shown how an increasing number N of aligned wave packets creates an increasing number of cross-terms that are aligned too. In the limit where $N \rightarrow \infty$, this leads to a perfect localization of the distribution along the line, which is the geometric locus of all of its midpoints.

resolution loss. It is worth noting that overcoming such a trade-off is possible to some extent, by using reassignment techniques (cf. Section 7.2). In addition, the loss of resolution can be sufficiently reduced by using methods that take into account the direction of signal energy concentration in the (t, f) domain, for example, multidirection TFD (see Section 5.9).

4.2.4 SUMMARY AND CONCLUSIONS

Quadratic (t, f) distributions satisfy a quadratic superposition principle, which necessarily results in the existence of cross-terms. On the one hand, such cross-terms may be considered as troublesome

since, in the case of a signal composed of many $((t,f)$ disjoint) components, their combinatorial proliferation hampers readability; on the other hand, in the case of individual components like chirps, the interferential generation of cross-terms may guarantee an improved localization as compared to linear transforms. Understanding the mechanism underlying cross-terms is therefore a key for trading-off interference reduction and increased localization in quadratic (t,f) distributions.

4.3 THE COVARIANCE THEORY OF TIME-FREQUENCY ANALYSIS⁰

4.3.1 THE COVARIANCE PRINCIPLE

Many important classes of linear and bilinear/quadratic time-frequency representations (TFRs) (such as those described in [Chapter 3](#), Eq. (3.2.13)) can be defined by a *covariance property*. For example, let us consider the family of STFTs [11, Sect. 2.2.2], [19, Sect. 3.2]

$$F_x^h(t,f) = \int_{-\infty}^{\infty} x(t') h^*(t' - t) e^{-j2\pi ft'} dt', \quad (4.3.1)$$

where $x(t)$ is the signal and $h(t)$ is a function independent of $x(t)$. Equation (4.3.1) defines all linear TFRs L that are *covariant* to time-frequency (t,f) shifts as

$$L_{S_{\tau,v}x}(t,f) = e^{-j2\pi(f-v)\tau} L_x(t-\tau, f-v). \quad (4.3.2)$$

Here, $S_{\tau,v}$ is the (t,f) shift operator defined as $(S_{\tau,v}x)(t) = x(t-\tau) e^{j2\pi vt}$. Thus, among all linear TFRs, the STFT is axiomatically defined by the (t,f) shift covariance property ([Eq. 4.3.2](#)). Similarly, the class of TFRs defined by the expression [11, Sect. 2.3.1], [19, Chap. 5] (see [Section 3.2.2](#), p. 112)

$$\rho_{x,y}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1) y^*(t_2) h^*(t_1 - t, t_2 - t) e^{-j2\pi f(t_1-t_2)} dt_1 dt_2, \quad (4.3.3)$$

where $h(t_1, t_2)$ is a “kernel” function independent of $x(t)$ and $y(t)$, comprises all bilinear/quadratic TFRs B that are covariant to (t,f) shifts according to

$$B_{S_{\tau,v}x, S_{\tau,v}y}(t,f) = B_{x,y}(t-\tau, f-v). \quad (4.3.4)$$

Similar covariance-based interpretations and definitions can be given for many other important classes of linear TFRs (e.g., wavelet transform, hyperbolic wavelet transform, and power wavelet transform [11, 19–21]) as well as bilinear/quadratic TFRs (e.g., affine, hyperbolic, power classes [19–22]; see also [Sections 4.1, 5.6, 7.1, and 15.3](#)).

This section presents a unified *covariance theory of (t,f) analysis* that allows the systematic construction of covariant TFRs [23–26]. Covariance properties are important in (t,f) analysis because unitary signal transformations often occur in practice—for example, time or frequency shifts as described by the (t,f) shift operator $S_{\tau,v}$ represent delays and Doppler shifts, which are encountered in radar and mobile communications.

⁰Authors: **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at) and **G. Tauböck**, Acoustics Research Institute, The Austrian Academy of Sciences, Vienna, Austria (georg.tauboeck@oeaw.ac.at). Reviewers: J.-P. Ovarlez and A. Papandreou-Suppappola.

4.3.2 TIME-FREQUENCY DISPLACEMENT OPERATORS

The proposed covariance theory is based on (t,f) displacement operators (DOs), which are operators that displace signals in the (t,f) plane. These operators generalize the (t,f) shift operator $\mathbf{S}_{\tau,v}$, for which the (t,f) displacements are simple shifts (translations), to other types or geometries of (t,f) displacements.

4.3.2.1 Group fundamentals

A set \mathcal{G} together with a binary operation \star that maps $\mathcal{G} \times \mathcal{G}$ to \mathcal{G} is called a *group* if it satisfies the following properties: (1) There exists an *identity element* $g_0 \in \mathcal{G}$ such that $g \star g_0 = g_0 \star g = g$ for all $g \in \mathcal{G}$. (2) To every $g \in \mathcal{G}$, there exists an *inverse element* $g^{-1} \in \mathcal{G}$ such that $g \star g^{-1} = g^{-1} \star g = g_0$. (3) Associative law: $g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3$ for all $g_1, g_2, g_3 \in \mathcal{G}$. If, in addition, $g_1 \star g_2 = g_2 \star g_1$ for all $g_1, g_2 \in \mathcal{G}$, the group is called *commutative* or *Abelian*. An elementary example of a commutative group is $(\mathbb{R}, +)$ for which $g_1 \star g_2 = g_1 + g_2$, $g_0 = 0$, and $g^{-1} = -g$. Two groups (\mathcal{G}, \star) and (\mathcal{H}, \diamond) are said to be *isomorphic* if there exists an invertible mapping $\psi : \mathcal{G} \rightarrow \mathcal{H}$ such that $\psi(g_1 \star g_2) = \psi(g_1) \diamond \psi(g_2)$ for all $g_1, g_2 \in \mathcal{G}$.

4.3.2.2 Formal definition and examples of DOs

A DO is a family of “unitary” operators $\mathbf{D}_{\alpha,\beta}$ indexed by a 2D “displacement parameter” (α, β) that belongs to some group (\mathcal{D}, \circ) . This operator family $\{\mathbf{D}_{\alpha,\beta}\}_{(\alpha,\beta) \in \mathcal{D}}$ satisfies the following two properties:

1. A displacement by the group identity parameter (α_0, β_0) is no displacement, that is, if we define \mathbf{I} as the identity operator, then $\mathbf{D}_{\alpha_0, \beta_0} = \mathbf{I}$.
2. A displacement by (α_1, β_1) followed by a displacement by (α_2, β_2) is equivalent (up to a phase factor) to a single displacement by $(\alpha_1, \beta_1) \circ (\alpha_2, \beta_2)$, that is,

$$\mathbf{D}_{\alpha_2, \beta_2} \mathbf{D}_{\alpha_1, \beta_1} = e^{j\phi(\alpha_1, \beta_1; \alpha_2, \beta_2)} \mathbf{D}_{(\alpha_1, \beta_1) \circ (\alpha_2, \beta_2)} \quad \forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathcal{D} \quad (4.3.5)$$

with $\phi(\alpha_1, \beta_1; \alpha_2, \beta_2)$ being a continuous function.

More precisely, a DO $\{\mathbf{D}_{\alpha,\beta}\}_{(\alpha,\beta) \in \mathcal{D}}$ is an *irreducible and faithful projective representation* of a group (\mathcal{D}, \circ) ; the function $e^{j\phi(\alpha_1, \beta_1; \alpha_2, \beta_2)}$ is known as the *cocycle*. For $e^{j\phi(\alpha_1, \beta_1; \alpha_2, \beta_2)} \equiv 1$, the DO $\{\mathbf{D}_{\alpha,\beta}\}_{(\alpha,\beta) \in \mathcal{D}}$ is a *unitary group representation* [26]. Two examples of a DO are the following:

- The (t,f) shift operator $\mathbf{S}_{\tau,v}$. Here, $(\alpha, \beta) = (\tau, v)$ and (\mathcal{D}, \circ) is the commutative group $(\mathbb{R}^2, +)$ with operation $(\tau_1, v_1) \circ (\tau_2, v_2) = (\tau_1 + \tau_2, v_1 + v_2)$; furthermore $(\tau_0, v_0) = (0, 0)$, $(\tau, v)^{-1} = (-\tau, -v)$, and $\phi(\tau_1, v_1; \tau_2, v_2) = -2\pi v_1 \tau_2$.
- The (t,f) scaling/time shift operator $\mathbf{R}_{\sigma,\tau}$ defined as $(\mathbf{R}_{\sigma,\tau}x)(t) = \frac{1}{\sqrt{|\sigma|}} x\left(\frac{t-\tau}{\sigma}\right)$. Here, $(\alpha, \beta) = (\sigma, \tau)$, (\mathcal{D}, \circ) is the noncommutative *affine group* with $\mathcal{D} = \mathbb{R}^+ \times \mathbb{R}$ and group operation $(\sigma_1, \tau_1) \circ (\sigma_2, \tau_2) = (\sigma_1 \sigma_2, \tau_1 \sigma_2 + \tau_2)$; furthermore $(\sigma_0, \tau_0) = (1, 0)$, $(\sigma, \tau)^{-1} = (1/\sigma, -\tau/\sigma)$, and $\phi(\sigma_1, \tau_1; \sigma_2, \tau_2) \equiv 0$.

4.3.2.3 Additional structure of DOs

The interpretation that a DO $\mathbf{D}_{\alpha,\beta}$ performs (t,f) displacements and motivates certain topological assumptions which imply that (\mathcal{D}, \circ) is a *simply connected 2D Lie group*, resulting in two important consequences [26]:

1. The group (\mathcal{D}, \circ) underlying $\mathbf{D}_{\alpha,\beta}$ is isomorphic to the group $(\mathbb{R}^2, +)$ underlying $\mathbf{S}_{\tau,v}$ or to the affine group underlying $\mathbf{R}_{\sigma,\tau}$ (see examples above).
2. The DO $\mathbf{D}_{\alpha,\beta}$ is *separable* (possibly up to a phase factor) in the following sense: there exists an isomorphism $\psi : (\alpha, \beta) \rightarrow (\alpha', \beta')$ such that the parameter-transformed DO $\mathbf{D}_{\psi^{-1}(\alpha', \beta')}$ (briefly written as $\mathbf{D}_{\alpha', \beta'}$ hereafter) can be factored as [26]

$$\mathbf{D}_{\alpha', \beta'} = e^{j\mu(\alpha', \beta')} \mathbf{B}_{\beta'} \mathbf{A}_{\alpha'}. \quad (4.3.6)$$

Here, $\mathbf{A}_{\alpha'}$ and $\mathbf{B}_{\beta'}$ (termed *partial DOs*) are indexed by 1D displacement parameters $\alpha' \in (\mathcal{A}, \bullet)$ and $\beta' \in (\mathcal{B}, *)$, respectively, where (\mathcal{A}, \bullet) and $(\mathcal{B}, *)$ are *commutative groups* that are isomorphic to $(\mathbb{R}, +)$. For example, $\mathbf{S}_{\tau,v} = \mathbf{F}_v \mathbf{T}_\tau$ and $\mathbf{R}_{\sigma,\tau} = \mathbf{T}_\tau \mathbf{C}_\sigma$, with the time-shift operator \mathbf{T}_τ , frequency-shift operator \mathbf{F}_v , and (t,f) scaling operator \mathbf{C}_σ defined as $(\mathbf{T}_\tau x)(t) = x(t - \tau)$, $(\mathbf{F}_v x)(t) = x(t) e^{j2\pi v t}$, and $(\mathbf{C}_\sigma x)(t) = \frac{1}{\sqrt{|\sigma|}} x\left(\frac{t}{\sigma}\right)$, respectively.

4.3.3 COVARIANT SIGNAL REPRESENTATIONS: GROUP DOMAIN

The construction of (t,f) representations that are covariant to a given DO $\mathbf{D}_{\alpha,\beta}$ is a two-stage process: first, we construct covariant signal representations that are functions of the displacement parameter (i.e., the group variables) (α, β) . Subsequently (in Sections 4.3.4 and 4.3.5), we convert these covariant (α, β) representations into covariant (t,f) representations.

4.3.3.1 Covariance in the group domain

A linear (α, β) representation $L_x(\alpha, \beta)$ is called *covariant to a DO* $\mathbf{D}_{\alpha,\beta}$ if

$$L_{\mathbf{D}_{\alpha', \beta'} x}(\alpha, \beta) = e^{j\phi((\alpha, \beta) \circ (\alpha', \beta')^{-1}; \alpha', \beta')} L_x\left((\alpha, \beta) \circ (\alpha', \beta')^{-1}\right), \quad (4.3.7)$$

for all signals $x(t)$ and for all $(\alpha, \beta), (\alpha', \beta') \in \mathcal{D}$ [26]. Similarly, a bilinear/quadratic (α, β) representation $B_{x,y}(\alpha, \beta)$ is called covariant to a DO $\mathbf{D}_{\alpha,\beta}$ if

$$B_{\mathbf{D}_{\alpha', \beta'} x, \mathbf{D}_{\alpha', \beta'} y}(\alpha, \beta) = B_{x,y}\left((\alpha, \beta) \circ (\alpha', \beta')^{-1}\right), \quad (4.3.8)$$

for all signal pairs $x(t), y(t)$ and for all $(\alpha, \beta), (\alpha', \beta') \in \mathcal{D}$ [25,26]. The “linear” covariance property (Eq. 4.3.7) differs from the “bilinear” covariance property (Eq. 4.3.8) in that it contains a phase factor. For example, for $\mathbf{S}_{\tau,v}$, Eqs. (4.3.7) and (4.3.8) become

$$L_{\mathbf{S}_{\tau', v'} x}(\tau, v) = e^{-j2\pi(v-v')\tau'} L_x(\tau - \tau', v - v'),$$

$$B_{\mathbf{S}_{\tau', v'} x, \mathbf{S}_{\tau', v'} y}(\tau, v) = B_{x,y}(\tau - \tau', v - v'),$$

which are identical to Eqs. (4.3.2) and (4.3.4), respectively. For $\mathbf{R}_{\sigma,\tau}$, we obtain

$$L_{\mathbf{R}_{\sigma',\tau'}x}(\sigma, \tau) = L_x\left(\frac{\sigma}{\sigma'}, \frac{\tau - \tau'}{\sigma'}\right), \quad B_{\mathbf{R}_{\sigma',\tau'}x, \mathbf{R}_{\sigma',\tau'}y}(\sigma, \tau) = B_{x,y}\left(\frac{\sigma}{\sigma'}, \frac{\tau - \tau'}{\sigma'}\right).$$

4.3.3.2 Construction of covariant (α, β) representations

It can be shown [26] that all linear (α, β) representations covariant to a DO $\mathbf{D}_{\alpha,\beta}$ as defined in Eq. (4.3.7) are given by

$$L_x(\alpha, \beta) = \langle x, \mathbf{D}_{\alpha,\beta}h \rangle = \int_{\mathcal{I}} x(t)(\mathbf{D}_{\alpha,\beta}h)^*(t) dt, \quad (4.3.9)$$

where $h(t)$ is an arbitrary function and \mathcal{I} is the time interval on which $\mathbf{D}_{\alpha,\beta}$ is defined. Similarly, all bilinear/quadratic (α, β) representations covariant to a DO $\mathbf{D}_{\alpha,\beta}$ as defined in Eq. (4.3.8) are given by [25,26]

$$B_{x,y}(\alpha, \beta) = \langle x, \mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}y \rangle = \int_{\mathcal{I}} \int_{\mathcal{I}} x(t_1)y^*(t_2) [\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}]^*(t_1, t_2) dt_1 dt_2, \quad (4.3.10)$$

where \mathbf{H} is an arbitrary linear operator and $[\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}](t_1, t_2)$ denotes the kernel of the composed operator $\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}$. Equations (4.3.9) and (4.3.10) provide canonical expressions for all covariant linear and bilinear/quadratic (α, β) representations. For example, for $\mathbf{S}_{\tau,\nu}$, these expressions yield

$$\begin{aligned} L_x(\tau, \nu) &= \int_{-\infty}^{\infty} x(t)h^*(t - \tau) e^{-j2\pi\nu t} dt, \\ B_{x,y}(\tau, \nu) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1)y^*(t_2) h^*(t_1 - \tau, t_2 - \tau) e^{-j2\pi\nu(t_1-t_2)} dt_1 dt_2, \end{aligned}$$

which are seen to be the STFT in Eq. (4.3.1) and the quadratic class in Eq. (4.3.3), respectively. Similarly, for $\mathbf{R}_{\sigma,\tau}$ we obtain time-scale versions of the wavelet transform and the affine class [11]:

$$\begin{aligned} L_x(\sigma, \tau) &= \frac{1}{\sqrt{|\sigma|}} \int_{-\infty}^{\infty} x(t)h^*\left(\frac{t - \tau}{\sigma}\right) dt, \\ B_{x,y}(\sigma, \tau) &= \frac{1}{|\sigma|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1)y^*(t_2) h^*\left(\frac{t_1 - \tau}{\sigma}, \frac{t_2 - \tau}{\sigma}\right) dt_1 dt_2. \end{aligned}$$

4.3.4 THE DISPLACEMENT FUNCTION

The covariant (α, β) representations constructed above can be converted into covariant (t,f) representations (ultimately, we are interested in (t,f) representations and not in (α, β) representations). This conversion uses a mapping $(\alpha, \beta) \rightarrow (t,f)$ that is termed the *displacement function* (DF) since it describes the (t,f) displacements performed by a DO $\mathbf{D}_{\alpha,\beta}$ in terms of (t,f) coordinates [26]. This DF concept is based on the following reasoning. If a signal $x(t)$ is (t,f) localized about some (t,f) point

(t_1, f_1) , then the transformed (“displaced”) signal $(\mathbf{D}_{\alpha, \beta}x)(t)$ will be localized about some other (t, f) point (t_2, f_2) that depends on (t_1, f_1) and α, β . We can thus write

$$(t_2, f_2) = e_{\mathbf{D}}(t_1, f_1; \alpha, \beta),$$

with some function $e_{\mathbf{D}}(t, f; \alpha, \beta)$ that is called the *extended DF* of the DO $\mathbf{D}_{\alpha, \beta}$. For example, for $\mathbf{S}_{\tau, v}$ and $\mathbf{R}_{\sigma, \tau}$, the extended DF is given by

$$e_{\mathbf{S}}(t, f; \tau, v) = (t + \tau, f + v), \quad e_{\mathbf{R}}(t, f; \sigma, \tau) = \left(\sigma t + \tau, \frac{f}{\sigma} \right). \quad (4.3.11)$$

4.3.4.1 Construction of the extended DF

We now present a systematic construction of the extended DF of a given DO $\mathbf{D}_{\alpha, \beta}$ [26]. The expression¹ $\mathbf{D}_{\alpha, \beta} = e^{j\mu(\alpha, \beta)} \mathbf{B}_{\beta} \mathbf{A}_{\alpha}$ (see Eq. (4.3.6)) states that $\mathbf{D}_{\alpha, \beta}$ is, up to a phase factor, the composition of \mathbf{A}_{α} and \mathbf{B}_{β} . Hence, $e_{\mathbf{D}}(t, f; \alpha, \beta)$ can be obtained by composing the extended DF of the partial DO \mathbf{A}_{α} , $e_{\mathbf{A}}(t, f; \alpha)$, and the extended DF of the partial DO \mathbf{B}_{β} , $e_{\mathbf{B}}(t, f; \beta)$, as

$$e_{\mathbf{D}}(t, f; \alpha, \beta) = e_{\mathbf{B}}(e_{\mathbf{A}}(t, f; \alpha); \beta). \quad (4.3.12)$$

Using Eq. (4.3.12), the task of constructing $e_{\mathbf{D}}(t, f; \alpha, \beta)$ reduces to constructing $e_{\mathbf{A}}(t, f; \alpha)$ and $e_{\mathbf{B}}(t, f; \beta)$. We explain the construction of $e_{\mathbf{A}}(t, f; \alpha)$ [26]; the construction of $e_{\mathbf{B}}(t, f; \beta)$ is of course analogous.

We first recall the “definition” of $e_{\mathbf{A}}(t, f; \alpha)$: If $x(t)$ is localized about (t_1, f_1) , then $(\mathbf{A}_{\alpha}x)(t)$ will be localized about $(t_2, f_2) = e_{\mathbf{A}}(t_1, f_1; \alpha)$. In order to find (t_2, f_2) , we consider the (generalized) eigenfunctions $u_b^{\mathbf{A}}(t)$ of \mathbf{A}_{α} . These are defined by $(\mathbf{A}_{\alpha}u_b^{\mathbf{A}})(t) \propto u_b^{\mathbf{A}}(t)$ and indexed by a parameter $b \in (\tilde{\mathcal{A}}, \bullet)$, where $(\tilde{\mathcal{A}}, \bullet)$ is again a commutative group isomorphic to $(\mathbb{R}, +)$. The (t, f) locus of $u_b^{\mathbf{A}}(t)$ is characterized by the IF $f_i\{u_b^{\mathbf{A}}\}(t)$ or the spectral delay (SD) $\tau_g\{u_b^{\mathbf{A}}\}(f)$, whichever exists.² Here, for example, we assume existence of $f_i\{u_b^{\mathbf{A}}\}(t)$. Let us choose b_1 such that the (t, f) curve defined by $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$ passes through (t_1, f_1) , that is,

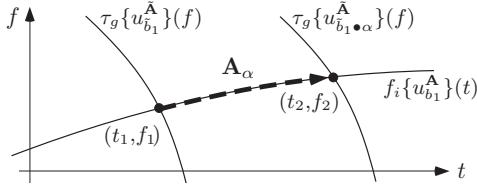
$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_1) = f_1. \quad (4.3.13)$$

This is shown in Fig. 4.3.1. Now since $(\mathbf{A}_{\alpha}u_b^{\mathbf{A}})(t) \propto u_b^{\mathbf{A}}(t)$, \mathbf{A}_{α} preserves the (t, f) locus of $u_{b_1}^{\mathbf{A}}(t)$. Therefore, *under the action of \mathbf{A}_{α} , all (t, f) points on the curve $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$ —including (t_1, f_1) —are mapped again onto (t, f) points on $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$.* Hence, $(t_2, f_2) = e_{\mathbf{A}}(t_1, f_1; \alpha)$ must lie on $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$ (see Fig. 4.3.1), so that

$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_2) = f_2. \quad (4.3.14)$$

¹For simplicity of notation, we assume that the parameter transformation $\psi : (\alpha, \beta) \rightarrow (\alpha', \beta')$ described in Section 4.3.2 has already been performed, and we write α, β instead of α', β' .

²The IF of a signal $x(t)$ is defined as $f_i\{x\}(t) = \frac{1}{2\pi} \frac{d}{dt} \arg\{x(t)\}$. The SD of $x(t)$ is defined as $\tau_g\{x\}(f) = -\frac{1}{2\pi} \frac{d}{df} \arg\{X(f)\}$ with $X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$. See the definitions in Chapter 1, Eqs. (1.3.1) and (1.3.22), and in [27].


FIGURE 4.3.1

Construction of the extended DF of \mathbf{A}_α .

In order to find the exact position of (t_2, f_2) on the (t, f) curve defined by $f_i\{u_{b_1}^A\}(t)$, we use the fact that for any partial DO \mathbf{A}_α there exists a *dual operator* $\tilde{\mathbf{A}}_{\tilde{\alpha}}$ with $\tilde{\alpha} \in (\tilde{\mathcal{A}}, \bullet)$ that is defined by the “almost commutation relation” $\tilde{\mathbf{A}}_{\tilde{\alpha}}\mathbf{A}_\alpha = e^{j2\pi\psi_{\mathcal{A}}(\alpha)\psi_{\tilde{\mathcal{A}}}(\tilde{\alpha})}\mathbf{A}_\alpha\tilde{\mathbf{A}}_{\tilde{\alpha}}$ [26,28]. For example, the dual operator of \mathbf{T}_τ is \mathbf{F}_v and vice versa. Let $u_{\tilde{b}}^{\tilde{A}}(t)$ with $\tilde{b} \in (\mathcal{A}, \bullet)$ denote the (generalized) eigenfunctions of $\tilde{\mathbf{A}}_{\tilde{\alpha}}$ and assume, for example, that the SD $\tau_g\left\{u_{\tilde{b}}^{\tilde{A}}\right\}(f)$ exists. Let us choose \tilde{b}_1 such that the (t, f) curve defined by $\tau_g\left\{u_{\tilde{b}_1}^{\tilde{A}}\right\}(f)$ passes through (t_1, f_1) (see Fig. 4.3.1), that is

$$\tau_g\left\{u_{\tilde{b}_1}^{\tilde{A}}\right\}(f_1) = t_1. \quad (4.3.15)$$

Now assuming a suitable parameterization of $u_{\tilde{b}}^{\tilde{A}}(t)$, it can be shown [26] that

$$(\mathbf{A}_\alpha u_{\tilde{b}}^{\tilde{A}})(t) \propto u_{\tilde{b} * \alpha}^{\tilde{A}}(t) \quad \text{for all } \alpha, \tilde{b} \in (\mathcal{A}, \bullet). \quad (4.3.16)$$

Thus, \mathbf{A}_α maps all (t, f) points on $\tau_g\left\{u_{\tilde{b}_1}^{\tilde{A}}\right\}(f)$ —including (t_1, f_1) —onto (t, f) points on $\tau_g\left\{u_{\tilde{b}_1 * \alpha}^{\tilde{A}}\right\}(f)$. So $(t_2, f_2) = e_{\mathbf{A}}(t_1, f_1; \alpha)$ must lie on $\tau_g\left\{u_{\tilde{b}_1 * \alpha}^{\tilde{A}}\right\}(f)$ (see Fig. 4.3.1), that is,

$$\tau_g\left\{u_{\tilde{b}_1 * \alpha}^{\tilde{A}}\right\}(f_2) = t_2. \quad (4.3.17)$$

The construction of $e_{\mathbf{A}}$ can now be summarized as follows (see Fig. 4.3.1):

1. For any given (t_1, f_1) , we calculate associated eigenfunction parameters $b_1 \in (\tilde{\mathcal{A}}, \bullet)$ and $\tilde{b}_1 \in (\mathcal{A}, \bullet)$ as the solutions of Eqs. (4.3.13) and (4.3.15), respectively:

$$f_i\{u_{b_1}^A\}(t_1) = f_1, \quad \tau_g\left\{u_{\tilde{b}_1}^{\tilde{A}}\right\}(f_1) = t_1. \quad (4.3.18)$$

2. The extended DF $e_{\mathbf{A}}$ is defined by the identity $(t_2, f_2) \equiv e_{\mathbf{A}}(t_1, f_1; \alpha)$, where (t_2, f_2) is obtained as the solution of the system of Eqs. (4.3.14) and (4.3.17):

$$f_i\{u_{b_1}^A\}(t_2) = f_2, \quad \tau_g\left\{u_{\tilde{b}_1 * \alpha}^{\tilde{A}}\right\}(f_2) = t_2. \quad (4.3.19)$$

A similar construction of e_A can be used if, for example, $\tau_g\{u_b^A\}(f)$ and $f_i\left\{u_{\tilde{b}}^{\tilde{A}}\right\}(t)$ exist instead of $f_i\{u_b^A\}(t)$ and $\tau_g\left\{u_{\tilde{b}}^{\tilde{A}}\right\}(f)$. An example is provided in [Section 4.3.6](#).

After construction of the extended DF of A_α as detailed above, the extended DF of B_β is constructed by means of an analogous procedure, and finally the extended DF of $D_{\alpha,\beta}$ is obtained by composing e_A and e_B according to Eq. (4.3.12).

4.3.4.2 The displacement function

The above discussion has shown how to construct the *extended* DF $e_D(t,f;\alpha,\beta)$. We go on to define the DF $d_D(\alpha,\beta)$ by fixing t,f in $e_D(t,f;\alpha,\beta)$:

$$d_D(\alpha,\beta) \triangleq e_D(t_0,f_0;\alpha,\beta) \text{ with } t_0,f_0 \text{ arbitrary but fixed.}$$

The DF is a mapping $(\alpha,\beta) \rightarrow (t,f)$, that is, from the displacement parameter (or group) domain to the (t,f) domain. If the inverse DF $d_D^{-1}(t,f)$ exists, then it can be shown [26] that the extended DF can be written as

$$e_D(t,f;\alpha,\beta) = d_D\left(d_D^{-1}(t,f) \circ (\alpha,\beta)\right). \quad (4.3.20)$$

4.3.4.3 Examples

The application of the construction explained above to the DOs $S_{\tau,v}$ and $R_{\sigma,\tau}$ yields the extended DFs $e_S(t,f;\tau,v) = (t + \tau, f + v)$ and $e_R(t,f;\sigma,\tau) = \left(\sigma t + \tau, \frac{f}{\sigma}\right)$. Note that this agrees with Eq. (4.3.11). Corresponding DFs are obtained by setting, for example, $t = f = 0$ in $e_S(t,f;\tau,v)$ and $t = 0, f = f_0 \neq 0$ in $e_R(t,f;\sigma,\tau)$, which yields $d_S(\tau,v) = (\tau,v)$ and $d_R(\sigma,\tau) = \left(\tau, \frac{f_0}{\sigma}\right)$. Another example is discussed in detail in [Section 4.3.6](#).

4.3.5 COVARIANT SIGNAL REPRESENTATIONS: TIME-FREQUENCY DOMAIN

[Section 4.3.3](#) derived covariant linear and bilinear/quadratic signal representations that are functions of the displacement parameter (α,β) . Using the inverse DF mapping $(\alpha,\beta) = d_D^{-1}(t,f)$, these covariant (α,β) representations can now be converted into covariant (t,f) representations (TFRs).

4.3.5.1 Covariance in the (t,f) domain

A linear TFR $\tilde{L}_x(t,f)$ is called *covariant to a DO* $D_{\alpha,\beta}$ if we have

$$\tilde{L}_{D_{\alpha',\beta'}x}(t,f) = e^{j\phi\left(d_D^{-1}(t,f) \circ (\alpha',\beta')^{-1}; \alpha',\beta'\right)} \tilde{L}_x\left(e_D(t,f;(\alpha',\beta')^{-1})\right), \quad (4.3.21)$$

for all $x(t)$ and for all $(\alpha',\beta') \in \mathcal{D}$ [26]. Similarly, a bilinear/quadratic TFR $\tilde{B}_{x,y}(t,f)$ is called covariant to a DO $D_{\alpha,\beta}$ if

$$\tilde{B}_{D_{\alpha',\beta'}x,D_{\alpha',\beta'}y}(t,f) = \tilde{B}_{x,y}\left(e_D(t,f;(\alpha',\beta')^{-1})\right), \quad (4.3.22)$$

for all $x(t), y(t)$ and for all $(\alpha', \beta') \in \mathcal{D}$ [26]. In view of Eq. (4.3.20), these covariance properties are simply the (α, β) -domain (group-domain) covariance properties (Eqs. (4.3.7) and (4.3.8)) with the transformation $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$.

For example, for $\mathbf{S}_{\tau, v}$, the covariance properties (Eqs. (4.3.21) and (4.3.22)) become

$$\begin{aligned}\tilde{L}_{\mathbf{S}_{\tau', v'} x}(t, f) &= e^{-j2\pi(f-v')\tau'} \tilde{L}_x(t - \tau', f - v'), \\ \tilde{B}_{\mathbf{S}_{\tau', v'} x, \mathbf{S}_{\tau', v'} y}(t, f) &= \tilde{B}_{x, y}(t - \tau', f - v').\end{aligned}$$

These relations are equivalent to Eqs. (4.3.2) and (4.3.4), respectively. For $\mathbf{R}_{\sigma, \tau}$, we obtain

$$\tilde{L}_{\mathbf{R}_{\sigma', \tau'} x}(t, f) = \tilde{L}_x\left(\frac{t - \tau'}{\sigma'}, \sigma'f\right), \quad \tilde{B}_{\mathbf{R}_{\sigma', \tau'} x, \mathbf{R}_{\sigma', \tau'} y}(t, f) = \tilde{B}_{x, y}\left(\frac{t - \tau'}{\sigma'}, \sigma'f\right).$$

4.3.5.2 Construction of covariant (t, f) representations

The (t, f) covariance properties (Eqs. (4.3.21) and (4.3.22)) were shown above to be equivalent to the (α, β) -domain covariance properties (Eqs. (4.3.7) and (4.3.8), respectively), apart from the transformation $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$. From this equivalence, it follows that all covariant linear TFRs $\tilde{L}_x(t, f)$ are obtained from corresponding covariant linear (α, β) representations $L_x(\alpha, \beta)$ as given by Eq. (4.3.9) simply by setting $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$. Consequently, all linear TFRs covariant to a DO $\mathbf{D}_{\alpha, \beta}$ are given by

$$\tilde{L}_x(t, f) = L_x(\alpha, \beta) \Big|_{(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)} = \left\langle x, \mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)} h \right\rangle = \int_{\mathcal{I}} x(t') \left(\mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)} h \right)^* (t') dt', \quad (4.3.23)$$

where $h(t)$ is an arbitrary function and \mathcal{I} is the time interval on which $\mathbf{D}_{\alpha, \beta}$ is defined. Similarly, all covariant bilinear/quadratic TFRs $\tilde{B}_{x, y}(t, f)$ are obtained from corresponding covariant bilinear/quadratic (α, β) representations $B_{x, y}(\alpha, \beta)$ as given by Eq. (4.3.10) by setting $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$. Thus, all covariant bilinear/quadratic TFRs are given by

$$\begin{aligned}\tilde{B}_{x, y}(t, f) &= B_{x, y}(\alpha, \beta) \Big|_{(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)} = \langle x, \mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)} \mathbf{H} \mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)}^{-1} y \rangle \\ &= \int_{\mathcal{I}} \int_{\mathcal{I}} x(t_1) y^*(t_2) \left[\mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)} \mathbf{H} \mathbf{D}_{d_{\mathbf{D}}^{-1}(t, f)}^{-1} \right]^* (t_1, t_2) dt_1 dt_2,\end{aligned} \quad (4.3.24)$$

where \mathbf{H} is an arbitrary linear operator. Equations (4.3.23) and (4.3.24) provide canonical expressions for all covariant linear and bilinear/quadratic TFRs [26].

For example, the classes of all linear and bilinear/quadratic TFRs covariant to $\mathbf{S}_{\tau, v}$ follow from Eqs. (4.3.23) and (4.3.24) as

$$\begin{aligned}\tilde{L}_x(t, f) &= \int_{-\infty}^{\infty} x(t') h^*(t' - t) e^{-j2\pi f t'} dt', \\ \tilde{B}_{x, y}(t, f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1) y^*(t_2) h^*(t_1 - t, t_2 - t) e^{-j2\pi f(t_1 - t_2)} dt_1 dt_2;\end{aligned}$$

these are the STFT in Eq. (4.3.1) and the quadratic class in Eq. (4.3.3), respectively. Similarly, the classes of all linear and bilinear/quadratic TFRs covariant to $\mathbf{R}_{\sigma,\tau}$ are obtained as

$$\begin{aligned}\tilde{L}_x(t,f) &= \sqrt{\left|\frac{f}{f_0}\right|} \int_{-\infty}^{\infty} x(t') h^* \left(\frac{f}{f_0}(t' - t) \right) dt', \\ \tilde{B}_{x,y}(t,f) &= \left| \frac{f}{f_0} \right| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1) y^*(t_2) h^* \left(\frac{f}{f_0}(t_1 - t), \frac{f}{f_0}(t_2 - t) \right) dt_1 dt_2,\end{aligned}$$

they are (t,f) versions of the wavelet transform and the affine class [11,22] (see Section 7.1). Thus, the STFT, the wavelet transform, the quadratic class in Eq. (4.3.3), and the affine class are all obtained by the systematic construction provided by covariance theory (see Section 4.1).

4.3.6 EXAMPLE: HYPERBOLIC WAVELET TRANSFORM AND HYPERBOLIC CLASS

So far, we have considered the elementary DOs $\mathbf{S}_{\tau,v}$ and $\mathbf{R}_{\sigma,\tau}$ as illustrative examples. Let us now apply the covariance theory to a situation that is less elementary. We consider the DO $\mathbf{V}_{\sigma,\gamma}$ defined by

$$(\mathbf{V}_{\sigma,\gamma}x)(t) = (\mathbf{G}_\gamma \mathbf{C}_\sigma x)(t) = \frac{1}{\sqrt{\sigma}} x\left(\frac{t}{\sigma}\right) e^{j2\pi\gamma \ln(t/t_0)}, \quad t > 0, \sigma > 0, \gamma \in \mathbb{R}.$$

Here, \mathbf{C}_σ is the (t,f) scaling operator defined by $(\mathbf{C}_\sigma x)(t) = \frac{1}{\sqrt{\sigma}} x\left(\frac{t}{\sigma}\right)$ with $\sigma > 0$; furthermore, \mathbf{G}_γ is the *hyperbolic frequency-shift operator* defined by $(\mathbf{G}_\gamma x)(t) = x(t) e^{j2\pi\gamma \ln(t/t_0)}$ with $t_0 > 0$ arbitrary but fixed. Note that \mathbf{C}_σ and \mathbf{G}_γ are dual operators since $\mathbf{G}_\gamma \mathbf{C}_\sigma = e^{j2\pi\gamma \ln \sigma} \mathbf{C}_\sigma \mathbf{G}_\gamma$. A comparison of the relation $\mathbf{V}_{\sigma_2,\gamma_2} \mathbf{V}_{\sigma_1,\gamma_1} = e^{-j2\pi\gamma_1 \ln \sigma_2} \mathbf{V}_{\sigma_1\sigma_2,\gamma_1+\gamma_2}$ with Eq. (4.3.5) shows that (σ, γ) belongs to the commutative group $(\mathbb{R}^+ \times \mathbb{R}, \circ)$ with group law $(\sigma_1, \gamma_1) \circ (\sigma_2, \gamma_2) = (\sigma_1\sigma_2, \gamma_1 + \gamma_2)$, identity element $(1, 0)$, and inverse elements $(\sigma, \gamma)^{-1} = (1/\sigma, -\gamma)$. This group is isomorphic to the group $(\mathbb{R}^2, +)$. Furthermore, the cocycle phase function is given by $\phi(\sigma_1, \gamma_1; \sigma_2, \gamma_2) = -2\pi\gamma_1 \ln \sigma_2$.

We now begin our construction of TFRs covariant to the DO $\mathbf{V}_{\sigma,\gamma}$. In the (σ, γ) domain, the covariance properties (Eqs. (4.3.7) and (4.3.8)) read as

$$\begin{aligned}L_{\mathbf{V}_{\sigma',\gamma'}x}(\sigma, \gamma) &= e^{-j2\pi(\gamma-\gamma') \ln \sigma'} L_x\left(\frac{\sigma}{\sigma'}, \gamma - \gamma'\right), \\ B_{\mathbf{V}_{\sigma',\gamma'}x, \mathbf{V}_{\sigma',\gamma'}y}(\sigma, \gamma) &= B_{x,y}\left(\frac{\sigma}{\sigma'}, \gamma - \gamma'\right),\end{aligned}$$

and the covariant linear and bilinear/quadratic (σ, γ) representations are obtained from Eqs. (4.3.9) and (4.3.10) as

$$\begin{aligned}L_x(\sigma, \gamma) &= \frac{1}{\sqrt{\sigma}} \int_0^{\infty} x(t) h^* \left(\frac{t}{\sigma} \right) e^{-j2\pi\gamma \ln(t/t_0)} dt, \quad \sigma > 0, \\ B_{x,y}(\sigma, \gamma) &= \frac{1}{\sigma} \int_0^{\infty} \int_0^{\infty} x(t_1) y^*(t_2) h^* \left(\frac{t_1}{\sigma}, \frac{t_2}{\sigma} \right) e^{-j2\pi\gamma \ln(t_1/t_2)} dt_1 dt_2, \quad \sigma > 0.\end{aligned}$$

Next, we construct the DF of $\mathbf{V}_{\sigma,\gamma} = \mathbf{G}_\gamma \mathbf{C}_\sigma$. We first consider the extended DF of \mathbf{C}_σ . Although clearly $e_{\mathbf{C}}(t,f;\sigma) = (\sigma t, f/\sigma)$, we derive $e_{\mathbf{C}}(t,f;\sigma)$ using the systematic construction procedure of Section 4.3.4. The eigenfunctions of \mathbf{C}_σ are $u_\gamma^\mathbf{C}(t) = \frac{1}{\sqrt{t}} e^{j2\pi\gamma \ln(t/t_0)}$, with IF $f_i\{u_\gamma^\mathbf{C}\}(t) = \gamma/t$. The eigenfunctions of the dual operator \mathbf{G}_γ are given by $u_s^\mathbf{G}(t) = \delta(t-s)$, with SD $\tau_g\{u_s^\mathbf{G}\}(f) \equiv s$. (It can be

verified that Eq. (4.3.16) is satisfied: $(\mathbf{C}_\sigma u_s^G)(t) = \sqrt{\sigma} \delta(t - \sigma s) \propto u_{\sigma s}^G(t)$. Thus, Eq. (4.3.18) becomes $\gamma_1/t_1 = f_1$ and $s_1 = t_1$, which yields the eigenfunction parameters associated with the (t, f) point (t_1, f_1) as $\gamma_1 = t_1 f_1$, $s_1 = t_1$. Similarly, Eq. (4.3.19) becomes $\gamma_1/t_2 = f_2$ and $s_1 \sigma = t_2$, which yields $t_2 = s_1 \sigma = \sigma t_1$ and $f_2 = \gamma_1/(s_1 \sigma) = f_1/\sigma$. Hence, the extended DF of \mathbf{C}_σ is finally obtained as

$$e_C(t_1, f_1; \sigma) = (t_2, f_2) = \left(\sigma t_1, \frac{f_1}{\sigma} \right).$$

The extended DF of \mathbf{G}_γ is obtained similarly. The eigenfunctions of \mathbf{G}_γ are $u_s^G(t) = \delta(t - s)$, with SD $\tau_g\{u_s^G\}(f) \equiv s$. The eigenfunctions of the dual operator \mathbf{C}_σ are $u_\gamma^C(t) = \frac{1}{\sqrt{t}} e^{j2\pi\gamma \ln(t/t_0)}$, with IF $f_i\{u_\gamma^C\}(t) = \gamma/t$. (We verify that Eq. (4.3.16) is satisfied: $(\mathbf{G}_\gamma u_\gamma^C)(t) = \frac{1}{\sqrt{t}} e^{j2\pi\gamma \ln(t/t_0)} e^{j2\pi\gamma' \ln(t/t_0)} \propto u_{\gamma+\gamma'}^C(t)$.) Thus, Eq. (4.3.18) (with the roles of IF and SD as well as time and frequency interchanged) becomes $s_1 = t_1$ and $\gamma_1/t_1 = f_1$, which yields the eigenfunction parameters $s_1 = t_1$, $\gamma_1 = t_1 f_1$. Similarly, Eq. (4.3.19) (with the same interchange of roles) becomes $s_1 = t_2$ and $(\gamma_1 + \gamma)/t_2 = f_2$, whence $t_2 = s_1 = t_1$ and $f_2 = (\gamma_1 + \gamma)/s_1 = f_1 + \gamma/t_1$. Hence, the extended DF of \mathbf{G}_γ is obtained as

$$e_G(t_1, f_1; \gamma) = (t_2, f_2) = \left(t_1, f_1 + \frac{\gamma}{t_1} \right).$$

The extended DF of $\mathbf{V}_{\sigma, \gamma} = \mathbf{G}_\gamma \mathbf{C}_\sigma$ can now be calculated by composing $e_C(t, f; \sigma)$ and $e_G(t, f; \gamma)$ according to Eq. (4.3.12), which yields

$$e_V(t, f; \sigma, \gamma) = e_G(e_C(t, f; \sigma); \gamma) = \left(\sigma t, \frac{f + \gamma/t}{\sigma} \right).$$

Finally, the DF (and inverse DF) of $\mathbf{V}_{\sigma, \gamma}$ follow upon setting $t = t_0 > 0$ and $f = 0$:

$$d_V(\sigma, \gamma) = e_V(t_0, 0; \sigma, \gamma) = \left(\sigma t_0, \frac{\gamma}{\sigma t_0} \right), \quad d_V^{-1}(t, f) = \left(\frac{t}{t_0}, tf \right).$$

With the DF at our disposal, we are ready to pass from the (σ, γ) domain into the (t, f) domain. The (t, f) covariance properties (Eqs. (4.3.21) and (4.3.22)) become

$$\begin{aligned} \tilde{L}_{\mathbf{V}_{\sigma', \gamma'} x}(t, f) &= e^{-j2\pi(tf - \gamma') \ln \sigma'} \tilde{L}_x \left(\frac{t}{\sigma'}, \sigma' \left(f - \frac{\gamma'}{t} \right) \right), \\ \tilde{B}_{\mathbf{V}_{\sigma', \gamma'} x, \mathbf{V}_{\sigma', \gamma'} y}(t, f) &= \tilde{B}_{x, y} \left(\frac{t}{\sigma'}, \sigma' \left(f - \frac{\gamma'}{t} \right) \right), \end{aligned}$$

and the covariant linear and bilinear/quadratic TFRs are obtained from Eqs. (4.3.23) and (4.3.24) as

$$\begin{aligned} \tilde{L}_x(t, f) &= \sqrt{\frac{t_0}{t}} \int_0^\infty x(t') h^* \left(t_0 \frac{t'}{t} \right) e^{-j2\pi tf \ln(t'/t_0)} dt', \quad t > 0, \\ \tilde{B}_{x, y}(t, f) &= \frac{t_0}{t} \int_0^\infty \int_0^\infty x(t_1) y^*(t_2) h^* \left(t_0 \frac{t_1}{t}, t_0 \frac{t_2}{t} \right) e^{-j2\pi tf \ln(t_1/t_2)} dt_1 dt_2, \quad t > 0. \end{aligned}$$

These TFRs are analogous to, respectively, the hyperbolic wavelet transform and the hyperbolic class introduced in [21], the difference being that in [21] the hyperbolic time-shift operator is used instead of the hyperbolic frequency-shift operator \mathbf{G}_γ .

4.3.7 SUMMARY AND CONCLUSIONS

TFRs that are covariant to practically important signal transformations—such as time and frequency shifts, time-frequency scaling (dilation/compression), or dispersive time and frequency shifts—are of great relevance in applications. We have presented a unified and coherent *covariance theory of TFRs* that allows the systematic construction of TFRs covariant to two-parameter transformations. Relations of covariance theory with the principle of unitary equivalence are discussed in Section 4.5 and in [26]; see also Sections 7.1 and 15.3 for discussions of affine TFRs and of power-class TFRs, respectively. A more detailed and mathematically rigorous discussion is provided in [26], where also the extension to groups not isomorphic to $(\mathbb{R}, +)$ is outlined.

4.4 UNCERTAINTY IN TIME-FREQUENCY ANALYSIS⁰

In 1946, Gabor [29] introduced the Heisenberg uncertainty relations to the signal processing community, with the simple but powerful inequality

$$\sigma_t \cdot \sigma_f \geq \frac{1}{4\pi}, \quad (4.4.1)$$

where σ_t and σ_f are the time and frequency standard deviations of the signal $s(t)$. That is (we will assume unit energy signals, centered at zero time and frequency):

$$\sigma_t = \sqrt{\int_{-\infty}^{\infty} t^2 |s(t)|^2 dt}, \quad (4.4.2)$$

$$\sigma_f = \sqrt{\int_{-\infty}^{\infty} f^2 |S(f)|^2 df}, \quad (4.4.3)$$

$S(f)$ being the FT of $s(t)$. The impact of this result in the field of spectrum estimation was immediate, since it seems to imply the existence of bounds on the achievable frequency resolution when analyzing finite length segments of data. Satisfaction of Eq. (4.4.1) implies that, for any given (nonzero) σ_t , there will be a minimum σ_f . However, some care must be exercised when following this line of reasoning, since standard deviations are not, in general, acceptable measures of broadness, and do not bear an easy relation with the concept of resolution. As an example, consider a long rectangular window. This window has a spectrum with a very thin main lobe and is, thus, capable of high-frequency resolution. However, the standard deviation of its power spectrum is ∞ , as can be seen using $S(f) = \sin(af)/(af)$ in Eq. (4.4.3). This example shows that Eq. (4.4.1), being a limit on standard deviations, does not necessarily limit the achievable frequency resolution. High-resolution waveforms may have large standard deviations. Other limitations of the use of standard deviations as measures of broadness can be found in [30,31]. We will introduce a different measure of broadness, which avoids these limitations.

With the development of joint time-frequency (t,f) analysis, the issue of uncertainty and/or frequency resolution has to be rethought. In (t,f) analysis, we often deal with a single bivariate function $\rho_s(t,f)$, desirably representing the distribution of energy in the (t,f) plane. For acceptance of a given

⁰Authors: **P.M. Oliveira**, Escola Naval, Almada, Portugal (pmonica@mail.telepac.pt) and **V. Barroso**, Instituto Superior Técnico, Lisboa, Portugal (victor.barroso@tecnico.ulisboa.pt). Reviewers: M.G. Amin and X.-G. Xia.

$\rho_s(t,f)$ as a valid TFD, one often requires satisfaction of the time and frequency marginal conditions (see p. 70):

$$\int_{-\infty}^{\infty} \rho_s(t,f) df = |s(t)|^2, \quad (4.4.4)$$

$$\int_{-\infty}^{\infty} \rho_s(t,f) dt = |S(f)|^2. \quad (4.4.5)$$

The Heisenberg-Gabor relations (Eq. 4.4.1) will imply limits to the joint behavior of the marginals, $|s(t)|^2$ and $|S(f)|^2$. But they do not imply limits on $\rho_s(t,f)$ at other (t,f) locations [32]. Do such limits exist? Are there uncertainty limits *within* the (t,f) plane, or can we achieve infinite (t,f) resolution, under the constraint that the marginals must satisfy Eq. (4.4.1)? To fully appreciate the scope of the question, let us consider an hypothetical $\rho_s(t,f)$, constituted by an infinitely narrow (t,f) ridge (see Fig. 4.4.1). Even though the marginals are broad, and, for any given σ_t can always be made to satisfy Eq. (4.4.1) by adjusting the frequency slope of the ridge, such a $\rho_s(t,f)$ possesses an infinite (t,f) concentration capability, which allows us to know the exact frequency that the signal occupied at any precise time (t_0 and f_0 , in the example figure). The question to be answered thus becomes: is such a $\rho_s(t,f)$ physically acceptable? Is infinite (t,f) concentration capability an achievable goal? And is infinite (t,f) resolution (the ability to separate two closely spaced components) also an achievable goal? Or are there limits (certainly other than Eq. (4.4.1)) to the concentration and/or resolution of any physically acceptable $\rho_s(t,f)$?

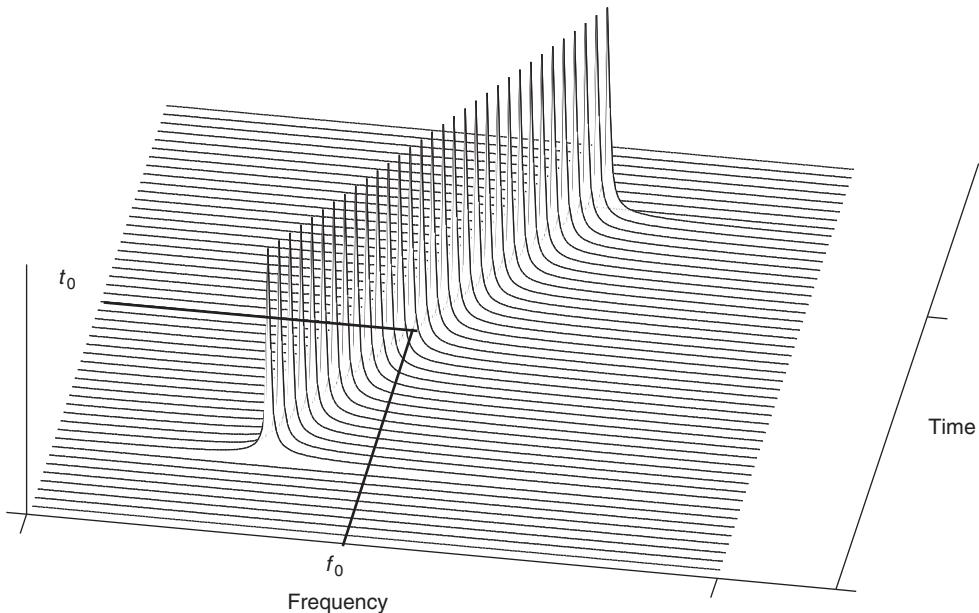


FIGURE 4.4.1

Example of a thin $\rho_s(t,f)$ with broad marginals.

4.4.1 THE TIME-FREQUENCY PLANE

Several attempts have been made to obtain uncertainty relations applicable to regions of the (t, f) plane other than the marginals. Each one of these attempts starts by defining a concentration/dispersion measure in the (t, f) plane (Δ_{tf}). Examples of such measures [11,33] are, for arbitrary T ,

$$\Delta_{tf} = \iint \left(\frac{t^2}{T^2} + T^2 f^2 \right) \rho_s(t, f) dt df, \quad (4.4.6)$$

$$\Delta_{tf} = \iint \left(1 - e^{-(t^2/T^2 + T^2 f^2)} \right) \rho_s(t, f) dt df, \quad (4.4.7)$$

$$\Delta_{tf} = \iint t^2 f^2 \rho_s(t, f) dt df, \quad (4.4.8)$$

$$\Delta_{tf} = \iint t^2 (f - f_i(t))^2 \rho_s(t, f) dt df. \quad (4.4.9)$$

All these measures turn out to imply a particular set of uncertainty relations when $\rho_s(t, f)$ is taken as being the WVD (Section 2.1.4, p. 69) or, more generally, one of the quadratic TFDs defined in Chapter 3 [33]. A different and promising approach can also be found in [11], where the Slepian-Landau-Pollack approach is extended to ellipsoidal regions (axis-aligned) in the (t, f) plane. However, careful analysis of these measures will show that none of them is still an appropriate measure of the concentration capabilities of a general $\rho_s(t, f)$. Let us take Eq. (4.4.6), as an example. When applied to the WVD, it implies [11] that $\Delta_{tf} \geq \frac{1}{2\pi}$. However, this distribution is known to generate infinitely thin ridges in the (t, f) plane for linear chirps. This shows that a given $\rho_s(t, f)$ can have the infinite (t, f) location capabilities of Fig. 4.4.1, even if it is lower bounded by measure Eq. (4.4.6). Furthermore, the known uncertainty relations implied by Eqs. (4.4.6)–(4.4.9) have been obtained for particular TFDs, and cannot be interpreted as general physical laws.

To avoid this restriction of the results to the scope of the particular (t, f) tool being used to derive them, we will take a different approach, based on the general concept of spectral information. This approach leads to general results that are not specific to any one TFD.

4.4.2 INFORMATION AND SPECTRAL ESTIMATION

In the following discussion, let us assume that we have a second degree *ensemble view*. That is, we will assume that, by observing the signal at time t_1 , we apprehend the expected value of the ensemble of signals ($E\{s(t_1)\}$), and not only the value of one particular realization; by also observing the signal at time t_2 , we now not only apprehend the value of $E\{s(t_2)\}$, but also $R_s(t_1, t_2) = E\{s(t_1)s^*(t_2)\}$. Since we are, under this assumption, directly observing expected values, and not mere realizations of the process, we may, in what follows, safely ignore the practical aspects of real estimators, such as variance. This idealization will enable us to focus only on the aspects which are, for our purposes, most relevant.

- *Stationary signals.* From this idealized point of view, let us now consider the estimation of the power spectrum of a stationary signal. To perform the estimate, one must extract the needed information out of the signal. But how much information must we collect? How much observation time do we need? At first, increasing the observation time will provide better and better estimates, in a process converging to the true power spectrum. But, after convergence (assuming it ever happens), will further increases in observation time provide more information about the power

spectrum? It clearly does not. Once this convergence process is essentially completed, no further observation time is needed, because no further information is required. The total amount of information that must be obtained for proper spectral estimation will be referred to as the *spectral complexity of the signal* (C_s). The point to be made here is that, if the observation time available to us is less than the time needed to reach C_s , our spectral estimate will be degraded, since we will be missing part of the information needed for a correct estimate. On the other hand, if the observation time available is greater than the time needed to reach C_s , the last part of the signal will be informationless. This definition of complexity is different in nature from the one proposed and used in [34–36], where *complexity* is interpreted as a measure of the number of components in the signal's (t, f) representation. Here, *spectral complexity* is a measure of the amount of information required for proper estimation of the signal's spectrum (and, hence, of the required effective observation time), and does not depend on the number of components in the (t, f) plane. Denoting by $D(t)$ the density of information contained in the signal (the function $D(t)$ depends on the type of estimator being used, an issue which we will not pursue here), and by $I(t)$ the amount of collected information, the collection procedure can be summarized as follows (assuming that the observation started at t_0):

$$I(t) = \int_{t_0}^t D(\xi) d\xi \leq \int_{t_0}^{t_0 + \tau_R} D(\xi) d\xi = C_s, \quad (4.4.10)$$

τ_R being the time needed to reach complexity. This immediately raises the question of how to quantify C_s . There are good reasons (some of which will become apparent) to use a measure of spectral narrowness as a measure of spectral complexity. While any acceptable measure of narrowness is basically equivalent, leading to a rescaled version of the quantity, we must be careful to avoid the pitfalls of global measures of width such as the standard deviation, as previously discussed. Let us then define

$$C_s = \sqrt{\max_f \left(-\frac{\partial^2}{\partial f^2} P_s(f) \right)}, \quad (4.4.11)$$

where $P_s(f) = E \{ |S(f)|^2 \}$ is the power spectrum density of the signal. This measure, based on the Hessian of the power spectrum, relates directly to the idea of frequency resolution, and is closely related with the definition of Fisher information.

From Eq. (4.4.11), the spectral complexity of a sinusoid is infinite. Although an impulsive spectrum may seem simple, the perfect localization capability needed for a proper estimate implies an infinite amount of information. Hence, the infinite C_s . Also, we must consider the fact that the spectrum of a pure sinusoid will always become narrower and narrower with increasing observation time, without ever stabilizing. Additional observation time will always improve our spectral estimate and, thus, provide new information. Complexity will never be reached for finite observation times. This is thus the high end of spectral complexity, where all observation time becomes useful and brings additional information. In the low end, we have white noise. An instantaneous ensemble observation fully characterizes its very low complexity (zero, in fact) power spectrum. Further observation of the noise ensemble will not contribute with any new information concerning its spectrum.

We must now address the next logical question: How do we determine τ_R ? The amount of information needed to estimate the power spectrum is clearly the same amount of information needed to estimate its inverse FT, the autocorrelation function. Hence, we only need to observe the signal for the amount of time needed to determine all (relevant) lags of its autocorrelation function. The time to reach complexity is thus the time support of the autocorrelation function. This is a very gratifying conclusion, since the spectral complexity equation (4.4.11) and the time support of the autocorrelation function $R_s(\tau)$ can be shown to be, in fact, directly related to each other through yet another “uncertainty relation”:

$$\frac{1}{C_s} \cdot \sigma_R \geq \frac{1}{2\pi}, \quad (4.4.12)$$

where $\sigma_R^2 = \int_{-\infty}^{\infty} \tau^2 |R_s(\tau)| d\tau$. In general, signals with narrowband components (and, thus, of high spectral complexity) require a high collection time τ_R (the limit is again a phase randomized sinusoid, known to have an autocorrelation function with infinite time support [37, p. 33]). Signals without narrowband components (hence, of low spectral complexity) have small collection times (in the limiting case of white noise, the autocorrelation function has zero time support). In any case, observing the ensemble for periods longer than τ_R is not useful, since no additional information about the power spectrum will be obtained.

- *Nonstationary signals.* Let us now consider nonstationary signals, and assume that we desire to estimate the power spectrum of a nonstationary signal at time t_1 . This instantaneous spectrum will have a given amount of spectral complexity ($C_s^{t_1}$), and to properly estimate it, we need to collect this very same amount of information about the spectrum (or the autocorrelation function) at time t_1 . But to represent time t_1 , all we have $s(t_1)$ itself, and no finite amount of spectral information can be extracted from an instantaneous value of the signal, since it would imply the acceptance of an infinite information density. On the other hand, information collected at times other than t_1 will only be useful if and only if it is correlated with the spectral information at time t_1 . In the previous case of stationary signals, the spectral information at any time was totally correlated with the spectral information at any other time. In the nonstationary case, however, we must weight the information collected at times t other than t_1 with the nonunitary correlation factor that determines how useful the collected information is for estimates at time t_1 (we will denote this weighting factor by $u^{t_1}(t - t_1)$). We now have to distinguish between useful past and future ($u^{t_1}(t - t_1) \neq 0$), and nonuseful past and future ($u^{t_1}(t - t_1) \approx 0$). The collection procedure (Eq. 4.4.10), using superscripts to denote the particular time for which the spectrum estimate is intended, becomes

$$\begin{aligned} I^{t_1}(t) &= \int_{t_0}^t D(\xi) u^{t_1}(\xi - t_1) d\xi, \text{ with} \\ I^{t_1}(t) &\leq C_s^{t_1} = \int_{-\infty}^{\infty} D(\xi) u^{t_1}(\xi - t_1) d\xi. \end{aligned} \quad (4.4.13)$$

The factor $u^{t_1}(t - t_1)$ is thus just formalizing the fact that observing a nonstationary signal *now* does not tell us much concerning the spectrum of the signal a fortnight ago, and may even be misleading, and decrease our state of knowledge (in the case of negative $u^{t_1}(t - t_1)$). The exception lies, of course, in the case of signals with deterministic and known frequency dynamics, since in

in these cases the information collected at any time can always be made useful, by taking the dynamics into proper account. Knowledge of the frequency dynamics thus makes the utility factor constant and unitary, bringing the case of nonstationary signals to the very same situation one encounters with stationary signals.

As an example, consider the estimation of the power spectrum of a constant amplitude linear chirp (let us assume a positive chirping rate), with uniformly distributed random phase

$$s(t) = e^{j(\alpha t^2 + \theta)}.$$

Its autocorrelation function is easily seen to be

$$R_s(t, t - \tau) = R_s(t, \tau) = e^{-j(\alpha\tau^2 - 2\alpha\tau t)}.$$

To determine $u(t)$, we can now determine how correlated are the autocorrelation functions at times t_1 and t_2 , with $t_2 > t_1$. Due to the infinite energy of these autocorrelation functions, in the computation of their correlation factor we will limit the integration region to an arbitrarily large region $(-l$ to l) centered at the zeroth lag. That is,

$$u^{t_1}(t_2 - t_1) = R_R(t_2, t_1, l) = \frac{1}{2l} \int_{-l}^l R_s(t_2, \xi) R_s^*(t_1, \xi) d\xi.$$

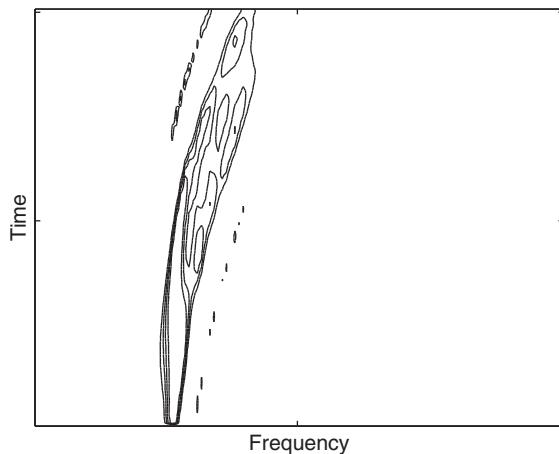
This means that, in our case,

$$u^{t_1}(t_2 - t_1) = \frac{\sin(2\alpha l(t_2 - t_1))}{2\alpha l(t_2 - t_1)}. \quad (4.4.14)$$

The inclusion of $u^{t_1}(t_2 - t_1)$ in Eq. (4.4.13) (in this case, a sinc function) will limit the amount of collectible information relative to time t_1 . Thus, it will set an upper bound on the achievable spectral complexity, hence the achievable frequency resolution.

This is thus the answer we have been trying to obtain. There are limits to the achievable frequency resolution within the (t, f) plane, due to the fact that the period of time during which one can collect information concerning the spectrum at a given time is decreasing as the spectral dynamics increase. For increasing dynamics (α , in our example) the useful neighborhood (and, hence, the amount of useful information) will continuously decrease, and so will the achievable spectral complexity and, consequently, the achievable spectral resolution. This namely means that the faster a chirp moves, the broader it becomes in the (t, f) plane. This predicted broadening of the power spectrum (and consequent decrease in resolution capabilities) with the increase of the chirping rate is, in fact, observed in most bilinear TFDs. To illustrate it, we computed the Margenau-Hill distribution of a cubic chirp. The results can be seen in Fig. 4.4.2, where the broadening of the main lobe of this TFD with the increase in chirp rate is clearly apparent.

Let us now try to determine, for the case of the chirp, the best observation time. From Eq. (4.4.14), we see that the best strategy is to limit the observation time to the main lobe of the sinc function, since we will then avoid negative values of $u^{t_1}(t - t_1)$. That is, observe the signal between $t_1 - \tau$ and $t_1 + \tau$,

**FIGURE 4.4.2**

Cubic chirp, Margenau-Hill distribution.

where $\tau = \pi/2\alpha l$. But this implies that τ is the maximum lag of the observed autocorrelation function and, thus, $l = \tau$. Thus we conclude that the best observation time for a linear chirp is

$$\tau = \sqrt{\frac{\pi}{2\alpha}} = \sqrt{\frac{1}{2\frac{\partial}{\partial t}[f_i(t)]}}, \quad (4.4.15)$$

$f_i(t)$ being the chirp IF (in this particular case, we may safely identify the concept of IF with the derivative of the phase function). With hindsight, it is now interesting to observe that: (a) τ , given by Eq. (4.4.15), is the optimum observation time for short-time Fourier analysis of a linear chirp; (b) τ is the effective time support of the optimum data-independent smoothing window to use with the WVD [38]; (c) τ is also basically the same quantity defined by Rihaczek as the signal’s “relaxation time” [39]. These separate results can now easily be understood as particular manifestations of Eq. (4.4.15).

A last comment must be made, concerning the use of models. Assuming a model for the frequency dynamics, such as the linear model implicit in the WVD (or higher order models in the polynomial WVD [40]), is an attempt to increase the size of what we called the “useful neighborhood,” by projecting all collected spectral information to the time of interest, according to the spectral dynamics of the assumed model. If, by inspiration or mere chance, the assumed model is, in fact, the correct one, we will overcome the limits imposed by the nonstationarity, and fall within the traditional, stationary uncertainty relations, as previously discussed. If, on the other hand, the model is incorrect, we must be prepared to pay for the wrong assumption. We may have apparently improved our frequency resolution, but we must pay for it with bias and artifacts. Another, more subtle, type of assumption is made whenever we arbitrarily decide that the “true” distribution is the one maximizing some chosen criterion. It may or may not be a sensible, supported assumption. It is an assumption, nonetheless. It may get us a better frequency resolution, but only if verified by the signal under analysis. This is exactly the type of trade-off one finds in parametric spectrum estimation.

4.4.3 SUMMARY AND CONCLUSIONS

This section addresses the issue of determining if there are lower bounds to the achievable (t,f) resolution within the (t,f) plane. The approach based on the informational aspects of the estimation achieves results independently of the specific joint power spectrum. The limits to the achievable (t,f) resolution are a direct consequence of the spectral dynamics of the signal and the implied decorrelation of the power spectrum from one moment to the next. Increasing spectral dynamics imply decreasing (t,f) resolution capabilities. The optimum observation time depends on the spectral dynamics of the signal and is tool-independent. [Section 4.6](#) further explores these issues.

4.5 GENERALIZED TFRs via UNITARY TRANSFORMS⁰

Despite the broad applicability of time-frequency representations such as the QTFDs defined in [Chapter 2](#), there exist situations where an analysis in terms of time and frequency coordinates may not be appropriate (see, e.g., [21, 41, 42]). These problems require joint distributions of other variables.

Joint distributions generalize single-variable distributions that measure the energy content of some physical quantity in a signal. Given a quantity a represented by the Hermitian (self-adjoint) operator \mathcal{A} , we obtain the density $|(\mathbb{F}_{\mathcal{A}} s)(a)|^2$ measuring the “ a content” of the signal s simply by squaring the projection of s onto the formal eigenfunctions $\mathbf{u}_a^{\mathcal{A}}$ of \mathcal{A} [32]

$$|(\mathbb{F}_{\mathcal{A}} s)(a)|^2 = \left| \int s(x) [\mathbf{u}_a^{\mathcal{A}}(x)]^* dx \right|^2. \quad (4.5.1)$$

(When the operator \mathcal{A} is Hermitian, the eigenequation is merely algebraic, and the eigenfunctions are actually tempered distributions. More rigorously, we could employ projection-valued measures for the eigenfunctions [43].) Classical examples of single variable densities include the time density $|(\mathbb{F}_{\mathcal{T}} s)(t)|^2 = |s(t)|^2$ and frequency density $|(\mathbb{F}_{\mathcal{F}} s)(f)|^2 = |S(f)|^2$ obtained by projecting onto the Dirac eigenfunctions of the time operator $(\mathcal{T}s)(x) = xs(x)$ and the sinusoidal eigenfunctions of the frequency operator $(\mathcal{F}s)(x) = \frac{1}{j2\pi} \dot{s}(x)$. (We will use both S and $\mathbb{F}s$ to denote the FT of the signal s .)

Joint distributions attempt to measure the simultaneous signal energy content of multiple physical quantities. Given two quantities a and b , a joint distribution $(\rho_{a,b}s)(a,b)$ measuring the joint a - b content in the signal s has, as marginals, the respective \mathcal{A} and \mathcal{B} energy densities

$$\int (\rho_{a,b}s)(a,b) db = |(\mathbb{F}_{\mathcal{A}} s)(a)|^2, \quad (4.5.2)$$

$$\int (\rho_{a,b}s)(a,b) da = |(\mathbb{F}_{\mathcal{B}} s)(b)|^2. \quad (4.5.3)$$

(Alternatively, we can define joint distributions in terms of their covariance properties under certain unitary transformations. For more details, see [7, 44] and [Sections 4.1](#) and [7.1](#).) The WVD, from the quadratic class of TFDs defined in [Chapters 2](#) and [3](#), supplies a classical example of a joint distribution that marginalizes to the time and frequency densities.

⁰Author: **R. Baraniuk**, Rice University, Houston, TX, USA (richb@rice.edu, <http://www.dsp.rice.edu>). Reviewers: P. Flandrin and F. Auger.

Many different constructions have been proposed for generating joint distributions. The various approaches fall into two broad categories: *general methods* and *coordinate change methods*. General methods can create distributions for every possible pairing of physical quantities by working from first principles, as discussed in Chapter 2 and in [45]. Coordinate change methods, on the contrary, sacrifice some flexibility for simplicity by bootstrapping existing distributions into new contexts using signal or axis transformations (see Sections 4.1 and 7.1) [21,41,46,47].

In the following, we review these two categories of joint distribution constructions and address their linkage.

4.5.1 THREE APPROACHES TO JOINT DISTRIBUTIONS

4.5.1.1 General quadratic method

Given two operators \mathcal{A} and \mathcal{B} representing two arbitrary physical quantities a and b , we can form the joint distribution of a and b as (see [32] for more details)

$$(\rho_{a,b}s)(a, b) = \iiint s^*(x) \left(e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})} s \right) (x) e^{-j2\pi(\alpha a + \beta b)} dx d\alpha d\beta. \quad (4.5.4)$$

The above construction is general, but it requires that we solve a sometimes complicated operator equation to express the exponentiated operator $e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}$. The time-frequency case is well understood; using the time and frequency operators \mathcal{T} and \mathcal{F} yields the quadratic class of TFDs discussed in Chapter 3 and in [32, p. 136].

4.5.1.2 Axis transformation method

Joint a - b distributions are easily obtained when we can functionally relate these variables to time and frequency by $a = \alpha^{-1}(t)$ and $b = \beta^{-1}(f)$. In this special case, we can derive an a - b distribution simply by warping the coordinates of a TFD $\rho_{t,f}$ [45]

$$(\rho_{a,b}s)(a, b) = |\dot{\alpha}(a)\dot{\beta}(b)| (\rho_{t,f}s) [\alpha(a), \beta(b)]. \quad (4.5.5)$$

It is easily verified that all $\rho_{a,b}s$ obtained in this way correctly marginalize to $|(\mathbb{F}_{\mathcal{A}}s)(a)|^2 = |\dot{\alpha}(a)| |(\mathbb{F}_{\mathcal{T}}s)(\alpha(a))|^2 = |\dot{\alpha}(a)| |s(\alpha(a))|^2$ and $|(\mathbb{F}_{\mathcal{B}}s)(b)|^2 = |\dot{\beta}(b)| |(\mathbb{F}_{\mathcal{F}}s)(\beta(b))|^2 = |\dot{\beta}(b)| |S(\beta(b))|^2$.

Example. Distributions $\rho_{t,r}s$ of time t and “inverse frequency” r (represented by the operator $\mathcal{R} = \frac{f_0}{f}$) can be constructed from the quadratic class (as defined in Chapter 3 on p. 113) through the change of variable $r = \frac{f_0}{f}$ [32, p. 238]. The resulting class of distributions, obtained as

$$(\rho_{t,r}s)(t, r) = \frac{f_0}{r^2} (\rho_{t,f}s) \left(t, \frac{f_0}{r} \right), \quad (4.5.6)$$

marginalizes to time $|s(t)|^2$ and inverse frequency $|(\mathbb{F}_{\mathcal{R}}s)(r)|^2 = \frac{f_0}{r^2} |S(f_0/r)|^2$.

4.5.1.3 Signal transformation method (unitary equivalence)

Joint a - b distributions are also easily generated when the quantities a and b are *unitarily equivalent* to time and frequency [41,46], with

$$\mathcal{A} = \mathbf{U}^{-1} \mathcal{T} \mathbf{U}, \quad \mathcal{B} = \mathbf{U}^{-1} \mathcal{F} \mathbf{U}, \quad (4.5.7)$$

and \mathbf{U} a unitary transformation. In this case, a joint a - b distribution can be obtained by preprocessing a TFD with \mathbf{U}

$$(\rho_{a,b} s)(a, b) = (\rho_{t,f} \mathbf{U}s)(a, b). \quad (4.5.8)$$

The signal transformation \mathbf{U} can be interpreted as “rotating” the coordinates of the TFD to the new variables. This can be useful to take into account specific directions of energy concentration.

All $\rho_{a,b}s$ obtained in this way correctly marginalize to $|(\mathbb{F}_{\mathcal{A}}s)(a)|^2 = |(\mathbb{F}_{\mathcal{T}}\mathbf{U}s)(a)|^2 = |(\mathbf{U}s)(a)|^2$ and $|(\mathbb{F}_{\mathcal{B}}s)(b)|^2 = |(\mathbb{F}_{\mathcal{F}}\mathbf{U}s)(b)|^2 = |(\mathbf{U}s)(b)|^2$; see [41].

Example. Define the logarithmic time operator $(\mathcal{L}s)(x) = \log(x)s(x)$, and define the Mellin operator $\mathcal{H} = \frac{1}{2}(\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T})$. (\mathcal{H} is also called the “scale” operator in [32,48].) These operators are unitarily equivalent to \mathcal{T} and \mathcal{F} as in Eq. (4.5.7) with $(\mathbf{U}s)(x) = e^{x/2}s(e^x)$. Therefore, we can construct distributions marginalizing to logarithmic time $|(\mathbb{F}_{\mathcal{L}}s)(l)|^2 = |e^{l/2}s(e^l)|^2$ and Mellin transform

$$|(\mathbb{F}_{\mathcal{H}}s)(m)|^2 = \left| \int s(x) e^{-j2\pi m \log x} |x|^{-1/2} dx \right|^2 \quad (4.5.9)$$

simply by preprocessing the signal by \mathbf{U} before computing a TFD [21,41,46,47].

4.5.2 LINKING SIGNAL AND AXIS TRANSFORMATIONS

Due to the individual limitations of the axis and signal transformation methods [49], only general operator methods can generate joint distributions for all possible operator pairs. However, when used in *tandem*, the axis and signal transformation methods yield a powerful method for generating a large number of joint distributions [49]. By executing first a unitary preprocessing transformation on the signal and then an axis warping transformation on the distribution, we can remap TFDs to a large class of different joint distributions.

Consider two variables a and b whose Hermitian operator representations \mathcal{A} and \mathcal{B} satisfy the following two conditions:

1. \mathcal{A} and \mathcal{B} can be related to the time and frequency operators as

$$\mathcal{A} = \mathbf{Z}^{-1}\mathcal{T}\mathbf{Z}, \quad \mathcal{B} = \mathbf{U}^{-1}\mathcal{F}\mathbf{U}, \quad (4.5.10)$$

with \mathbf{Z} and \mathbf{U} unitary transformations.

2. The composition $\mathbf{V} = \mathbf{Z}\mathbf{U}^{-1}$ is an axis warping operator of the form

$$(\mathbf{V}g)(x) = |\dot{v}(x)|^{1/2} g[v(x)], \quad (4.5.11)$$

with v a smooth, one to one function.

In this case, a joint a - b distribution $(\rho_{a,b}s)(a, b)$ can be constructed from any TFD $(\rho_{t,f}s)(t, f)$ through [49]

$$(\rho_{a,b}s)(a, b) = |\dot{v}(a)| (\rho_{t,f} \mathbf{U}s) [v(a), b]. \quad (4.5.12)$$

The interpretation of Eq. (4.5.12) is simple: Transformation \mathbf{U} rotates the (t, f) coordinates of the TFD to the new coordinates (u, b) . Transformation \mathbf{V} then warps the rotated time axis u to align it with the quantity a .

In other words, if \mathcal{A} and \mathcal{B} relate to \mathcal{T} and \mathcal{F} as in Eqs. (4.5.10) and (4.5.11), then we can obtain a large class of a - b distributions using a simple three-step procedure:¹

1. transform the signal $s \mapsto \mathbf{U}s$;
2. compute a conventional TFD of the transformed signal;
3. warp the remapped time axis of the resulting distribution.

The advantage of the double transformation method is that it breaks the severe restrictions placed on the quantities a and b by both the axis and signal transformation methods described in Section 4.5.1. By allowing the choice of both \mathbf{U} and \mathbf{Z} , we gain access to a much larger class of distributions. However, completely free choice is still not possible, because \mathbf{U} and \mathbf{Z} must have the structural property of canceling modulo the warping operator \mathbf{V} .

4.5.3 EXAMPLES OF LINKED SIGNAL/AXIS TRANSFORMATIONS

Linked signal/axis transformations are especially useful for transforming TFDs to distributions of time versus a “warped frequency” variable. In this case, we set \mathbf{Z} to the identity operator, choose \mathbf{U} to be a warping operator-based round the axis warping function v^{-1} , and set $\mathbf{V} = \mathbf{U}^{-1}$ to warp the rotated time axis back to the time variable. The resulting distributions marginalize to time

$$\int (\rho_{t,b}s)(t, b) db = |s(t)|^2 \quad (4.5.13)$$

and the integral transform

$$\int (\rho_{t,b}s)(t, b) dt = \left| \int s(x) e^{-j2\pi bv(x)} |\dot{v}(x)|^{1/2} dx \right|^2 \quad (4.5.14)$$

that projects the signal onto the eigenfunctions of the operator $\mathbf{U}^{-1}\mathcal{T}\mathbf{U}$ [41]

$$\mathbf{u}_b^{\mathbf{U}^{-1}\mathcal{T}\mathbf{U}}(x) = (\mathbf{U}^{-1}\mathbf{u}_b^{\mathcal{T}})(x) = e^{j2\pi bv(x)} |\dot{v}(x)|^{1/2}. \quad (4.5.15)$$

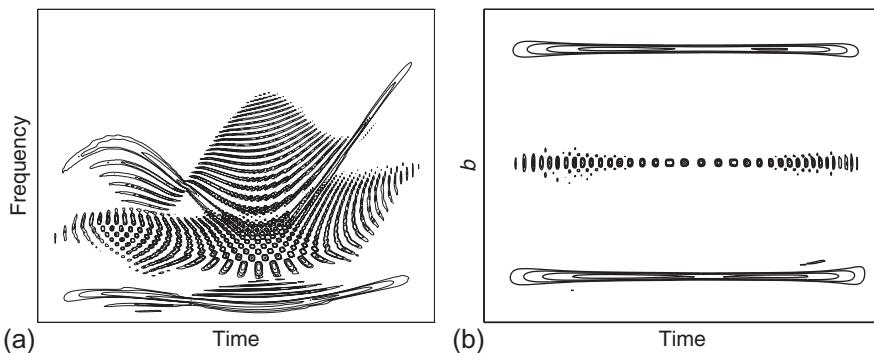
Each choice of v results in a different warped frequency transform matched to a distinct class of IF characteristics. The Fourier, Mellin, and chirp transforms result from the choices $v(x) = x$, $v(x) = \log x$, and $v(x) = |x|^c \operatorname{sgn}(x)$, respectively.

To continue the example of Section 4.5.1.3, applying the warp \mathbf{U}^{-1} to the logarithmic time axis of the logarithmic time versus Mellin distribution remaps that axis back to true time. The resulting distributions lie in Cohen’s class of time-Mellin distributions (time-scale in his terminology) [32,48]. This class contains the Marinovich-Altes (warped Wigner) distribution [47,50]. It is important to note that this class is out of the grasp of either signal or axis warping alone.

¹If we define the 2D transformation $\bar{\mathbf{V}}$ as the area-preserving change of variables

$$(\bar{\mathbf{V}}G)(x, y) = |\dot{v}(x)| G[v(a), b],$$

then we can write Eq. (4.5.12) in the standard form $\rho_{a,b} = \bar{\mathbf{V}}\rho_{t,f}\mathbf{U}$ found in [41]. However, whereas [41] emphasized using $\bar{\mathbf{V}}$ only to warp both axes of $\rho_{a,b}$ back to indicate time and frequency, in this section, we exploit a range of different $\bar{\mathbf{V}}$.

**FIGURE 4.5.1**

(a) WVD of synthetic test signal composed of two FM tones. (b) Warped WVD with time and “composite linear/sinusoidal IF” (variable “ b ” in Eq. (4.5.14)) marginals. The variable b measures the variation of the linear/sinusoidal IF path in time-frequency.

Figure 4.5.1 shows the plots of two TFDs of a signal consisting of two components concentrated along composite linear/sinusoidal IFs. Since the WVD does not match signals of this type as well as sinusoids, impulses, and linear chirps, it exhibits copious cross-components. Prewarping the signal to account for the form of the signal yields a postwarped TFD that marginalizes to time and “composite linear/sinusoidal IF” content (corresponding to variable “ b ” in Eq. (4.5.14)) and therefore better matches the signal.

Reversing the rôles of time and frequency in the warping procedure will yield frequency versus warped time distributions that match different classes of group delay (dispersion) characteristics.

4.5.4 SUMMARY AND CONCLUSIONS

Signal and axis transformations provide a simple framework for studying certain distributions of variables beyond time-frequency and time-scale. When applicable, warping results turn the theory of joint distributions of arbitrary variables into an easy exercise of coordinate transformation. For additional details, see [Sections 4.3](#) and [5.6](#) that discuss unitary transforms and [Section 7.1](#) that discussed the Mellin transform and the affine group.

4.6 MEASUREMENT CONCEPTS IN THE TIME-FREQUENCY PLANE⁰

A goal of time-frequency (t,f) analysis is to characterize and visualize the behavior of nonstationary signals. This is achieved by distilling both the amplitude and phase information of a signal to present

⁰Authors: **G. Jones**, Accipiter Radar Technologies Inc., Fenwick, ON, Canada (gjones@accipiterradar.com) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: D.L. Jones, W.J. Williams, and G. Putland.

an image of the variation of the frequency content of the signal with respect to time. The resulting (t,f) distribution (TFD) attempts to display this information with the highest possible resolution (energy concentration) and integrity (correct boundaries in time and frequency). One interesting aspect of the TFD [51,52] is the derivation of additional measurements from within the (t,f) plane that provide more information about the signal under analysis. This section provides an introduction to some measurement concepts in the (t,f) domain that allow for the definition of additional (t,f) features in applications such as pattern recognition, identification, modeling, and classification.

4.6.1 QUADRATIC TFDs AND SPECTROGRAM

The class of bilinear or quadratic TFDs (QTFDs; see [Chapter 3](#)) provides a convenient formulation in which to study and visualize signals jointly in time and frequency. QTFDs may be expressed as

$$\rho_{\gamma_z}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t-t_1, f-f_1) \gamma(t_1, f_1) dt_1 df_1, \quad (4.6.1)$$

where $W_z(t,f)$ is the WVD defined in [Chapter 2](#) as

$$W_z(t,f) = \int_{-\infty}^{\infty} z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau. \quad (4.6.2)$$

The quantity, $\gamma(t_1, f_1)$, is the kernel function and is used as a smoothing function, reducing the perceived undesirable properties of the WVD for visualization, such as negativity and signal component interference. When the smoothing function is itself a WVD, the resulting smoothed distribution, becomes the spectrogram and may be alternately expressed as

$$\left| F_z^h(t,f) \right|^2 = S_z^h(t,f) = \left| \int_{-\infty}^{\infty} z(\tau) h^*(t-\tau) e^{-j2\pi f\tau} d\tau \right|^2, \quad (4.6.3)$$

where $h(t)$ is the analyzing window (with $\gamma(t,f) = W_h(t,f)$ in Eq. (4.6.1)). By changing the smoothing or kernel function, many distributions may be defined, such as the RID [15] or the modified-B distribution (MBD) [53]. Other TFDs can be defined that employ a signal dependent/adaptive smoothing function (see [Section 5.3](#)).

4.6.2 SIGNAL MEASUREMENTS IN THE (t,f) DOMAIN

Basic signal analysis in one dimension often characterizes a signal by separating it into its amplitude and phase. For example, the IF is a common way to describe the phase variation of a signal, as is its mathematical dual in the frequency domain, the spectral delay (SD) (see definition in [Chapter 1](#), p. 52). Both of these quantities are first moments of the signal, in time and frequency, respectively. We can generalize these moment measures as

$$\int_{-\infty}^{\infty} f^m |Z(f)|^2 df = \int_{-\infty}^{\infty} \left[\frac{1}{z(t)} \left(\frac{1}{2\pi j} \right)^m \frac{d^m}{dt^m} z(t) \right] |z(t)|^2 dt = \int_{-\infty}^{\infty} F^m(t) |z(t)|^2 dt, \quad (4.6.4)$$

and

$$\int_{-\infty}^{\infty} t^n |z(t)|^2 dt = \int_{-\infty}^{\infty} \left[\frac{1}{Z(f)} \left(\frac{-1}{2\pi j} \right)^n \frac{d^n}{df^n} Z(f) \right] |Z(f)|^2 df = \int_{-\infty}^{\infty} T^n(f) |Z(f)|^2 df. \quad (4.6.5)$$

The above expressions can reveal how (t, f) measures can be derived. A special case of Eqs. (4.6.4) and (4.6.5) shows that the IF and SD are derived as $F^1(t)$ and $T^1(f)$, respectively.

4.6.2.1 Extension of the Fourier moment relations to two dimensions

The above measures in Eqs. (4.6.4) and (4.6.5) can be merged in a single equation. Standard Fourier theory thus yields an expression of the form

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m |v(t, f)|^2 dt df = \left(\frac{-1}{2\pi j} \right)^n \left(\frac{1}{2\pi j} \right)^m \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V^*(t_1, f_1) \frac{\partial^m}{\partial t_1^m} \frac{\partial^n}{\partial f_1^n} V(t_1, f_1) dt_1 df_1, \quad (4.6.6)$$

where

$$V(t_1, f_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v(t, f) e^{-j2\pi t f_1} e^{j2\pi t_1 f_1} dt df \quad (4.6.7)$$

for any 2D function $v(t, f)$. By using Fourier relations for QTFDs in two dimensions (see [54] for the complete derivation), sets of local moments in two dimensions can be expressed as

$$T^n F^m \sigma(t, f) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m \rho_{\sigma_z}(t_1, f_1) \rho_{\sigma_h}(t - t_1, f - f_1) dt_1 df_1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_{\sigma_z}(t_1, f_1) \rho_{\sigma_h}(t - t_1, f - f_1) dt_1 df_1} = \frac{1}{|F_z^h(t, f)|^2} \left(\frac{-1}{2\pi j} \right)^n \left(\frac{1}{2\pi j} \right)^m \quad (4.6.8)$$

$$\frac{\partial^m}{\partial \tau^m} \frac{\partial^n}{\partial \nu^n} \left[e^{-j\pi \nu t} e^{j\pi \tau f} F_z^h \left(t + \left(\frac{1}{2} - \sigma \right) \tau, f + \left(\frac{1}{2} + \sigma \right) \nu \right) F_z^h * \left(t - \left(\frac{1}{2} + \sigma \right) \tau, f - \left(\frac{1}{2} - \sigma \right) \nu \right) \right] \Big|_{(\nu, \tau)=(0,0)}, \quad (4.6.9)$$

where h is an even window; and

$$\rho_{\sigma_z}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t - t_1, f - f_1) e^{j2\pi t_1 f_1 / \sigma} dt_1 df_1, \quad 0 < \sigma \leq \frac{1}{2}, \quad (4.6.10)$$

and $F_z^h(t, f)$ is the STFT of the signal $z(t)$ with smoothing function $h(t)$ given by:

$$F_z^h(t, f) = \int_{-\infty}^{\infty} z \left(\frac{t}{2} + \tau \right) h^* \left(\frac{t}{2} - \tau \right) e^{-j2\pi f \tau} d\tau. \quad (4.6.11)$$

Note that h is even, so that $h^*(\frac{t}{2} - \tau) = h(\tau - \frac{t}{2})$. In the limit $h(t) \rightarrow \delta(t)$, the conventional moment quantities, in the single dimensions of time and frequency, are extracted. For example, with $n = 0$, the relation of Eq. (4.6.8) reduces to:

$$T^0 F^m \sigma(t) = \frac{\int_{-\infty}^{\infty} f^m \rho_{\sigma_z}(t, f) df}{\int_{-\infty}^{\infty} \rho_{\sigma_z}(t, f) df}. \quad (4.6.12)$$

Furthermore, for $\sigma = 0$ and $m = 1$, the classical measure of IF ($F^1(t)$ in Eq. (4.6.4)) is returned. (Similarly, the SD is obtained when $\sigma = 0$, $n = 1$, and $m = 0$.) This observation reinforces the notion that the above quantities are a natural 2D extension of the (1D) local moments.

4.6.2.2 Generation of (t, f) local measurements

When $\sigma = 0$ in Eq. (4.6.8), this relation may be reformulated as

$$T^n F^m_{(z,h)}(t, f) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m W_z(t_1, f_1) W_h(t - t_1, f - f_1) dt_1 df_1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t_1, f_1) W_h(t - t_1, f - f_1) dt_1 df_1} \quad (4.6.13)$$

with the denominator being a spectrogram. For these measurements to characterize a specific location in the (t, f) plane, a desirable property of the window function's WVD is that it possesses strong localization. The next section reveals how the WVD of a Gaussian function is an ideal choice for this task, and also presents a special property of Gaussian-based local measures that ultimately lead to a valid interpretation and application.

4.6.3 PROPERTIES AND INTERPRETATION OF LOCAL (t, f) MEASUREMENTS

This section seeks to associate the set of first and second order measures, as defined in Eq. (4.6.13), with our intuitive notions of signal extent and spread. In classical time-series analysis, the measures of extent for Fourier pairs are bounded by the Gabor-Heisenberg uncertainty principle. Its classical form is stated here as [55]:

$$\int_{-\infty}^{\infty} (t - \bar{t})^2 |z_N(t)|^2 dt \int_{-\infty}^{\infty} (f - \bar{f})^2 |Z_N(f)|^2 df \geq \frac{1}{16\pi^2}, \quad (4.6.14)$$

where $z_N(t)$ and $Z_N(f)$ are the normalized version of the signal in time and frequency, respectively, and \bar{t} and \bar{f} are the time and frequency signal averages. As has been shown in [56,57], by re-expressing the above as:

$$\left(\int_{-\infty}^{\infty} (t - \bar{t})^2 |z_N(t)|^2 dt \right)^{1/2} \left(\int_{-\infty}^{\infty} (f - \bar{f})^2 |Z_N(f)|^2 df \right)^{1/2} \geq \frac{1}{4\pi}, \quad (4.6.15)$$

and by applying the algebraic uncertainty relation ($a^2 + b^2 \geq 2ab$), one may obtain an additional uncertainty expression, represented as:

$$\int_{-\infty}^{\infty} (t - \bar{t})^2 |z_N(t)|^2 dt + \int_{-\infty}^{\infty} (f - \bar{f})^2 |Z_N(f)|^2 df \geq \frac{1}{2\pi}. \quad (4.6.16)$$

These uncertainty principles provide bounds on how a signal is spread in the time and the frequency domain, and both are satisfied with equality for Gaussian functions.

To further examine (t,f) signal bounds, we can expand time domain signals into series of Hermite functions to derive expressions for local moments. For example, the characteristics of Hermite functions [58] allow the expansion of the scalar time operator as

$$\begin{aligned} tP_{mn}(t,f) = & \frac{j}{8\pi} \left[(2m+2)^{1/2} P_{m+1,n}(t,f) - (2m)^{1/2} P_{m-1,n}(t,f) \right. \\ & \left. + (2n)^{1/2} P_{m,n-1}(t,f) - (2n+2)^{1/2} P_{m,n+1}(t,f) \right], \end{aligned} \quad (4.6.17)$$

where $P_{mn}(t,f)$ is the cross-WVD of an m th and n th order Hermite function. Other expressions for powers of t and f follow. By using relations such as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_{mn}(t,f) P_{qr}(t,f) dt df = \langle \psi_m, \psi_q \rangle \langle \psi_n, \psi_r \rangle \quad (4.6.18)$$

(where the left-hand side of the above equation represents the inner products of Hermite functions in $L^2(\mathcal{R})$), as well as the orthogonality of the Hermite functions:

$$\int_{-\infty}^{\infty} \psi_m(x) \psi_n(x) dx = \delta_{mn}, \quad (4.6.19)$$

local moments in two dimensions may be easily derived in a succinct form. When the WVD of a Gaussian signal, $h(t) = e^{-\pi t^2}$ (a zeroth-order Hermite function), is used as the localizing window to calculate the moment measurements as defined in Eq. (4.6.13), the following theorem is obtained [54]:

$$\left[T^2 F^0(t,f) - T^1 F^0(t,f)^2 \right] + \left[T^0 F^2(t,f) - T^0 F^1(t,f)^2 \right] = \frac{1}{4\pi}, \quad \forall z(t), h(t) = e^{-\pi t^2}. \quad (4.6.20)$$

This theorem, also shown to be true in [54] for any LFM signal, has some interesting implications. It is similar in form to the uncertainty principle of Eq. (4.6.16), and it justifies the wide use of a Gaussian window as a credible localizing function. A remarkable property forthcoming from the above result is that it holds even if one of the t or f local spread measures is negative.

One reason that measures in the (t,f) plane as well as conventional instantaneous quantities have been cited as nonintuitive is that they often yield values that do not appear related to the physical reality (e.g., a negative IF for a signal that has no negative frequency content). The relations given above show a logical balance exists within this measurement framework. The measures do not just possess nice theoretical properties, however, as shown in the next section which applies these local (t,f) measures to improve the resolution of the spectrogram.

4.6.4 APPLICATION OF LOCAL (t,f) MEASUREMENTS TO SPECTROGRAMS

Despite many advances in designing high-resolution QTFDs, spectrograms remain an important tool for displaying (t,f) signal content in many industries and applications. Their main disadvantage is that they often smear the actual (t,f) energy content, since they employ a window that necessarily possesses its own (t,f) signature. A window function that is circularly symmetric in the (t,f) plane is one solution, but this provides mediocre resolution in all directions at the expense of any smearing.

A more appropriate solution would be to adapt the analyzing window so that it captures the local (t,f) character of the signal (see [59–61] for some other examples). See also [51] for a summary of modern approaches and state of the art.

Let us look then at how we may obtain a local signal estimate for (smoothing window) matching. Taylor series expansion of the signal's TFD about the region of interest is one approach, but there is no guarantee that the local window generated from such an approximation will produce a valid spectrogram. Prompted by the above results, let us instead match the first couple of local moments, which are equivalently bandwidth and extent measures. The quantities of interest thus are the local time spread, LTB ($T^2 F^0(t,f) - T^1 F^0(t,f)^2$), the local frequency bandwidth, LFB ($T^0 F^2(t,f) - T^0 F^1(t,f)^2$), and the local cross (t,f) product LTF ($T^1 F^1(t,f)$). After applying a matching algorithm (described below), the resulting spectrogram becomes

$$S_z^{h_M}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t_1, f_1) W_h(t - t_1, f - f_1, t, f) dt_1 df_1 \quad (4.6.21)$$

$$= \left| \int_{-\infty}^{\infty} z(\tau) h_f(t, \tau) e^{-j2\pi f \tau} d\tau \right|^2, \quad (4.6.22)$$

where $W_h(t_1, f_1, t, f)$ is the WVD of the location-dependent window function h_f .

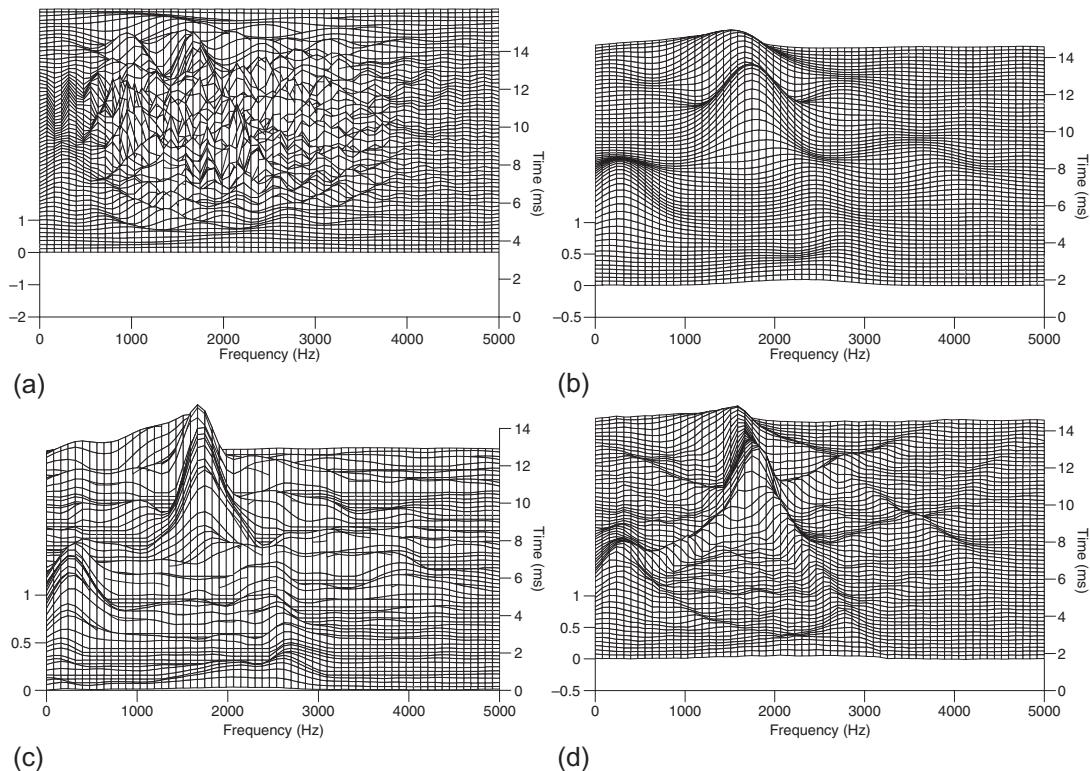
Note that the notation h_f is an emphasis to reflect the variation of h as a function of τ (which is the dual of f). The matching process used to realize the above is simple, and yields excellent signal visualization in the (t,f) plane. It uses the spread measures to determine the most appropriate values in time for α and in frequency for β (of a Gaussian window, $h(t) = e^{-\pi\alpha t^2} e^{j\pi\beta t^2}$) at each (t,f) point of interest. An iterative approach is used, leveraging the equality relation of Eq. (4.6.20) [54]. An example of such window matching follows next.

4.6.5 EXAMPLE: RESULT FOR AN ADAPTIVE SPECTROGRAM

The algorithm described above is applied here to a nonstationary real signal—a sample from a Humpback whale song. Its WVD, in Fig. 4.6.1(a), is hard to interpret in an intuitive way due to the many cross products. Although the spectrogram of Fig. 4.6.1(b) (using a (t,f) symmetric Gaussian window) improves the visualization, the information is highly smoothed. For comparison, a ZAM distribution is shown in Fig. 4.6.1(c). Finally, Fig. 4.6.1(d) displays the adaptive spectrogram, providing the analyst with a more intuitive and well-resolved QTFD.

4.6.6 SUMMARY AND CONCLUSIONS

An approach to deriving useful measurements from the (t,f) plane has been presented here. It was shown that basic Fourier relations, used to define measures such as IF and SD, can be extended to the (t,f) plane and the attendant distributions. Some tools for deriving and analyzing these local (t,f) measurements were introduced, and an interesting property of the second-order moment measures was presented. An example was provided to demonstrate potential applications of these local measures in the (t,f) plane. Such measures can be useful as additional (t,f) features in pattern recognition for automatic diagnosis applications such as those described in Section 16.3 and Chapter 16. Section 4.4 complements this material with a discussion of the uncertainty principle in (t,f) analysis.

**FIGURE 4.6.1**

Time-frequency distributions of a humpback whale sound. (a) WVD; (b) spectrogram; (c) ZAM distribution; and (d) adaptive spectrogram.

4.7 TIME-FREQUENCY TRANSFER FUNCTION CALCULUS OF LINEAR TIME-VARYING SYSTEMS⁰

4.7.1 LINEAR TIME-VARYING SYSTEMS

Due to their generality, *linear time-varying* (LTV) systems have important advantages over linear time-invariant (LTI) systems. Applications include mobile communications (see Sections 9.5, 13.2, and 13.3 as well as [62,63]), machine monitoring (Sections 15.2 and 15.6), and nonstationary statistical signal processing (see Sections 9.2, 9.4, 12.1, and 12.4 as well as [64]). An LTV system \mathbf{H} maps an input signal $x(t)$ to an output signal $y(t)$ according to

⁰Authors: **G. Matz**, Vienna University of Technology, Vienna, Austria (g.matz@ieee.org) and **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at). Reviewers: S. Barbarossa and A. Papandreou-Suppappola.

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t, t') x(t') dt', \quad (4.7.1)$$

where $h(t, t')$ is the kernel (impulse response) of \mathbf{H} . LTI systems and their dual, *linear frequency-invariant* (LFI) systems, are special cases of LTV systems. For an LTI system, $y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} g(t - t') x(t') dt'$ and thus $h(t, t') = g(t - t')$. For an LFI system, $y(t) = (\mathbf{H}x)(t) = w(t)x(t)$ and thus $h(t, t') = w(t)\delta(t - t')$.

For LTI and LFI systems, intuitive and efficient analysis and design methods are based on the spectral transfer function $G(f) = \int_{-\infty}^{\infty} g(\tau) e^{-j2\pi f\tau} d\tau$ and on the temporal transfer function $w(t)$, respectively. For LTV systems, this section shows that under a practically relevant *underspread* assumption, the *generalized Weyl symbol* (GWS) is an approximate time-frequency (briefly (t, f)) transfer function. We note that other (t, f) symbols are discussed in [Section 9.2](#).

4.7.2 THE GENERALIZED WEYL SYMBOL

The GWS of an LTV system \mathbf{H} is a family of linear (t, f) representations defined as [65]

$$L_{\mathbf{H}}^{(\alpha)}(t, f) \triangleq \int_{-\infty}^{\infty} h^{(\alpha)}(t, \tau) e^{-j2\pi f\tau} d\tau$$

with

$$h^{(\alpha)}(t, \tau) \triangleq h\left(t + \left(\frac{1}{2} - \alpha\right)\tau, t - \left(\frac{1}{2} + \alpha\right)\tau\right), \quad (4.7.2)$$

where α is a real-valued parameter. The GWS reduces to the ordinary *Weyl symbol* [66–68] for $\alpha = 0$, to Zadeh's *time-varying transfer function* [69] for $\alpha = 1/2$, and to Bello's *frequency-dependent modulation function* [70] (also known as *Kohn-Nirenberg symbol*) for $\alpha = -1/2$:

$$\begin{aligned} L_{\mathbf{H}}^{(0)}(t, f) &= \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau, \\ L_{\mathbf{H}}^{(1/2)}(t, f) &= \int_{-\infty}^{\infty} h(t, t - \tau) e^{-j2\pi f\tau} d\tau, \\ L_{\mathbf{H}}^{(-1/2)}(t, f) &= \int_{-\infty}^{\infty} h(t + \tau, t) e^{-j2\pi f\tau} d\tau. \end{aligned}$$

In what follows, α will be considered fixed. The GWS $L_{\mathbf{H}}^{(\alpha)}(t, f)$ is a *linear* (t, f) representation of the LTV system \mathbf{H} . It contains all information about \mathbf{H} since the kernel of \mathbf{H} can be recovered from the GWS:

$$h(t, t') = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}\left(\left(\frac{1}{2} + \alpha\right)t + \left(\frac{1}{2} - \alpha\right)t', f\right) e^{j2\pi f(t-t')} df. \quad (4.7.3)$$

Also, the input-output relation ([Eq. 4.7.1](#)) can be reformulated in terms of the GWS. This reformulation becomes especially simple for $\alpha = \pm 1/2$:

$$y(t) = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(1/2)}(t, f) X(f) e^{j2\pi ft} df, \quad Y(f) = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(-1/2)}(t, f) x(t) e^{-j2\pi ft} dt.$$

For a rank-one system with impulse response $h(t, t') = u(t)u^*(t')$, $L_{\mathbf{H}}^{(\alpha)}(t, f)$ reduces to the generalized Wigner distribution [13] of the signal $u(t)$. Other interesting properties of the GWS can be found in [65,67,71] and for $\alpha = 0$ in [66,68].

Next, we consider the GWS of some simple specific systems. The results obtained suggest that (under assumptions to be discussed later) the GWS can be interpreted as a “ (t, f) transfer function” that characterizes the “ (t, f) weighting” produced by the LTV system \mathbf{H} , that is, the way in which a component of the input signal $x(t)$ located about some (t, f) point is attenuated ($|L_{\mathbf{H}}^{(\alpha)}(t, f)| < 1$), amplified ($|L_{\mathbf{H}}^{(\alpha)}(t, f)| > 1$), or passed without amplification/attenuation ($|L_{\mathbf{H}}^{(\alpha)}(t, f)| = 1$).

- The GWS of the identity operator \mathbf{I} with kernel $h(t, t') = \delta(t - t')$ is given by $L_{\mathbf{I}}^{(\alpha)}(t, f) \equiv 1$ (i.e., no attenuation/amplification anywhere in the (t, f) plane).
- The GWS of the (t, f) shift operator $\mathbf{S}_{v, \tau}^{(\alpha)}$ defined by¹ $(\mathbf{S}_{v, \tau}^{(\alpha)}x)(t) = x(t - \tau) e^{j2\pi vt} e^{j2\pi v\tau(\alpha-1/2)}$ is a 2D complex sinusoid, $L_{\mathbf{S}_{v, \tau}^{(\alpha)}}^{(\alpha)}(t, f) = e^{j2\pi(vt-\tau f)}$, and thus $\left|L_{\mathbf{S}_{v, \tau}^{(\alpha)}}^{(\alpha)}(t, f)\right| = 1$ (i.e., no amplification/attenuation anywhere in the (t, f) plane).
- The GWS of an LTI system with kernel $h(t, t') = g(t - t')$ reduces to the ordinary transfer function $G(f)$ for all t , that is, $L_{\mathbf{H}}^{(\alpha)}(t, f) \equiv G(f)$.
- The GWS of an LFI system with kernel $h(t, t') = w(t)\delta(t - t')$ reduces to the temporal transfer function $w(t)$ for all f , that is, $L_{\mathbf{H}}^{(\alpha)}(t, f) \equiv w(t)$.

The last two examples show the GWS’s consistency with conventional transfer functions. In what follows, we investigate the conditions under which the GWS can be interpreted as a (t, f) transfer function. The key element in this investigation is an analysis of the (t, f) shifts produced by \mathbf{H} . For this analysis, we need another linear (t, f) representation of \mathbf{H} , to be discussed next.

4.7.3 THE GENERALIZED SPREADING FUNCTION

Besides a (t, f) -dependent weighting, LTV systems can also introduce (t, f) shifts of various input signal components. Here, the (t, f) shift operator $\mathbf{S}_{v, \tau}^{(\alpha)}$ mentioned above—with v denoting frequency (Doppler) shift and τ denoting time shift/delay—is an elementary example. A joint description of the (t, f) shifts introduced by a linear system \mathbf{H} is given by the *generalized spreading function* (GSF) [65]

$$S_{\mathbf{H}}^{(\alpha)}(v, \tau) \triangleq \int_{-\infty}^{\infty} h^{(\alpha)}(t, \tau) e^{-j2\pi vt} dt,$$

with $h^{(\alpha)}(t, \tau)$ as in Eq. (4.7.2). Like the GWS, the GSF $S_{\mathbf{H}}^{(\alpha)}(v, \tau)$ is a *linear* (t, f) representation of the LTV system \mathbf{H} , and it contains all information about \mathbf{H} since the kernel of \mathbf{H} can be recovered from the GWS. The input-output relation (Eq. 4.7.1) can be reformulated in terms of the GSF according to

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_{\mathbf{H}}^{(\alpha)}(v, \tau) (\mathbf{S}_{v, \tau}^{(\alpha)}x)(t) dv d\tau.$$

¹ The parameter α in $\mathbf{S}_{v, \tau}^{(\alpha)}$ corresponds to different ways of defining a joint (t, f) shift by combining time shifts and frequency shifts. In particular, $\alpha = 1/2$ corresponds to first shifting in time and then shifting in frequency, whereas $\alpha = -1/2$ corresponds to the reverse.

This represents the output signal $y(t) = (\mathbf{H}x)(t)$ as a weighted superposition of (t, f) shifted versions $(\mathbf{S}_{v,\tau}^{(\alpha)}x)(t) = x(t - \tau) e^{j2\pi\nu t} e^{j2\pi\nu\tau(\alpha-1/2)}$ of the input signal $x(t)$. The (ν, τ) -dependent weights in this superposition are given by the GSF, thus establishing the GSF's interpretation as a “ (t, f) shift distribution” of \mathbf{H} . The extension of $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ about the origin of the (ν, τ) plane indicates the amount of (t, f) shifts caused by \mathbf{H} . In particular, a large extension of $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ in the ν direction indicates large frequency/Doppler shifts (equivalently, fast time variation), and a large extension of $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ in the τ direction indicates large time shifts or delays.

GSFs with different α values differ merely by a phase factor, that is, $S_{\mathbf{H}}^{(\alpha_2)}(\nu, \tau) = S_{\mathbf{H}}^{(\alpha_1)}(\nu, \tau) e^{j2\pi(\alpha_1-\alpha_2)\nu\tau}$. Thus, the GSF magnitude is independent of α , $|S_{\mathbf{H}}^{(\alpha_1)}(\nu, \tau)| = |S_{\mathbf{H}}^{(\alpha_2)}(\nu, \tau)|$, so we may simply write $|S_{\mathbf{H}}(\nu, \tau)|$. The GSF is related to the GWS by a 2D FT,

$$S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}(t, f) e^{-j2\pi(\nu t - \tau f)} dt df. \quad (4.7.4)$$

Further properties of the GSF are described in [65,67,71].

Again, it is instructive to consider a few examples (see Fig. 4.7.1):

- The GSF of the identity operator \mathbf{I} is given by $S_{\mathbf{I}}^{(\alpha)}(\nu, \tau) = \delta(\nu)\delta(\tau)$, which is zero for $(\nu, \tau) \neq (0, 0)$ (i.e., neither frequency shifts nor time shifts).
- The GSF of a (t, f) shift operator $\mathbf{S}_{v_0, \tau_0}^{(\alpha)}$ is obtained as $S_{\mathbf{S}_{v_0, \tau_0}^{(\alpha)}}^{(\alpha)}(\nu, \tau) = \delta(\nu - v_0)\delta(\tau - \tau_0)$, which is zero for $(\nu, \tau) \neq (v_0, \tau_0)$ (i.e., no frequency and time shifts other than by v_0 and τ_0 , respectively).
- The GSF of an LTI system \mathbf{H} with kernel $h(t, t') = g(t - t')$ is given by $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) = \delta(\nu)g(\tau)$ (i.e., only time shifts whose distribution is characterized by the impulse response $g(\tau)$).
- The GSF of an LFI system \mathbf{H} with $h(t, t') = w(t)\delta(t - t')$ is given by $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) = W(\nu)\delta(\tau)$, where $W(\nu)$ is the FT of $w(t)$ (i.e., only frequency shifts whose distribution is characterized by $W(\nu)$).

4.7.4 UNDERSpread LTV SYSTEMS

Using the GSF, we now define the class of *underspread* LTV systems [67,71] for which, as we will see in Section 4.7.5, the GWS acts as a “ (t, f) transfer function.” Conceptually, an LTV system is

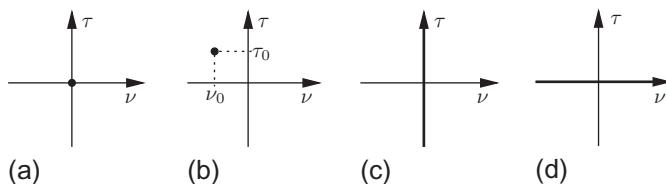


FIGURE 4.7.1

Schematic representation of the GSF magnitude of some (classes of) linear systems: (a) identity operator, (b) (t, f) shift operator, (c) LTI system, and (d) LFI system.

underspread if its GSF is well concentrated about the origin of the (ν, τ) plane, which indicates that the system introduces only small (t, f) shifts, that is, the system's time variations are slow and/or its memory is short. In contrast, systems introducing large (t, f) shifts are termed *overspread* [67,71].

There are two alternative mathematical characterizations of the GSF extension and, in turn, of underspread systems. The first characterization [71] requires that the support of the GSF $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ is confined to a compact region $\mathcal{G}_{\mathbf{H}}$ about the origin of the (ν, τ) plane, that is, $|S_{\mathbf{H}}(\nu, \tau)| = 0$ for $(\nu, \tau) \notin \mathcal{G}_{\mathbf{H}}$. Let $\nu_{\mathbf{H}} \triangleq \max_{(\nu, \tau) \in \mathcal{G}_{\mathbf{H}}} |\nu|$ and $\tau_{\mathbf{H}} \triangleq \max_{(\nu, \tau) \in \mathcal{G}_{\mathbf{H}}} |\tau|$ denote the maximum frequency shift and time shift, respectively, introduced by the system \mathbf{H} . We define the *Doppler-delay spread* of \mathbf{H} as $\sigma_{\mathbf{H}} \triangleq 4\nu_{\mathbf{H}}\tau_{\mathbf{H}}$, which is the area of the rectangle $[-\nu_{\mathbf{H}}, \nu_{\mathbf{H}}] \times [-\tau_{\mathbf{H}}, \tau_{\mathbf{H}}]$ enclosing $\mathcal{G}_{\mathbf{H}}$. Underspread LTV systems are then defined by the condition $\sigma_{\mathbf{H}} \ll 1$.

As the GSF of practical LTV systems rarely has compact support, an alternative, flexible characterization of GSF extension and underspread systems (without compact support) is based on the normalized *weighted GSF integrals* [67]

$$m_{\mathbf{H}}^{(\phi)} \triangleq \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\nu, \tau) |S_{\mathbf{H}}(\nu, \tau)| d\nu d\tau}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S_{\mathbf{H}}(\nu, \tau)| d\nu d\tau}. \quad (4.7.5)$$

Here, $\phi(\nu, \tau)$ is a non-negative weighting function that satisfies $\phi(\nu, \tau) \geq \phi(0, 0) = 0$ and penalizes GSF contributions away from $(0, 0)$. We define the *GSF moments* $m_{\mathbf{H}}^{(k,l)} \triangleq m_{\mathbf{H}}^{(\phi_{k,l})}$ as special cases of $m_{\mathbf{H}}^{(\phi)}$ using weighting functions $\phi_{k,l}(\nu, \tau) = |\nu|^l |\tau|^k$ with $k, l \in \{0, 1, \dots\}$. Thus, without assuming compact GSF support, a system \mathbf{H} can now be considered underspread if suitable GSF integrals/moments are “small.” This definition of underspread systems has the advantage of being more flexible than the previous definition based on the area of the (assumedly) compact support of the GSF. It can be shown that if the GSF of \mathbf{H} does have compact support with maximum frequency shift $\nu_{\mathbf{H}}$ and maximum time shift $\tau_{\mathbf{H}}$, then $m_{\mathbf{H}}^{(k,l)} \leq \nu_{\mathbf{H}}^l \tau_{\mathbf{H}}^k$ and, in particular, $m_{\mathbf{H}}^{(k,k)} \leq (\sigma_{\mathbf{H}}/4)^k$. Thus, LTV systems that are underspread in the compact-support sense are a special case of the underspread framework based on weighted GSF integrals and moments.

Examples of various types of underspread systems are shown in Fig. 4.7.2. Note that the concept of underspread systems is not equivalent to that of slowly time-varying (quasi-LTI) systems. A quasi-LTI system (i.e., small $m_{\mathbf{H}}^{(0,l)}$) may be overspread if its memory is very long (i.e., very large $m_{\mathbf{H}}^{(k,0)}$), and a

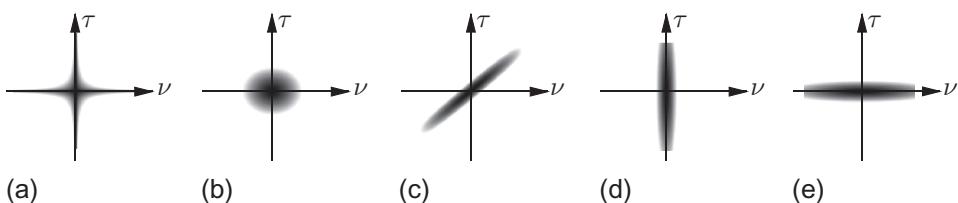


FIGURE 4.7.2

GSF magnitude of (a) an underspread system with small $m_{\mathbf{H}}^{(1,1)}$, (b) an underspread system with small $m_{\mathbf{H}}^{(1,0)} m_{\mathbf{H}}^{(0,1)}$, (c) a “chirpy” underspread system [67], (d) a quasi-LTI system (slowly time-varying; small $m_{\mathbf{H}}^{(0,1)}$), and (e) a quasi-LFI system (short memory; small $m_{\mathbf{H}}^{(1,0)}$).

system with faster time-variations (i.e., larger $m_{\mathbf{H}}^{(0,l)}$) may be underspread if its memory is short enough (i.e., very small $m_{\mathbf{H}}^{(k,0)}$). Finally, Eq. (4.7.4) implies that the GWS of an underspread LTV system is a *smooth* function.

4.7.5 TIME-FREQUENCY TRANSFER FUNCTION CALCULUS

For *underspread* LTV systems as defined above, the GWS acts as an approximate “ (t,f) transfer function” that generalizes the spectral (temporal) transfer function of LTI (LFI) systems. Indeed, if specific weighted GSF integrals $m_{\mathbf{H}}^{(\phi)}$ and/or moments $m_{\mathbf{H}}^{(k,l)}$ are small, one can show the validity of several transfer function approximations [67,71], some of which are discussed in the following. Applications of these approximations include time-varying spectral analysis, linear (t,f) filter design, and detection/estimation of nonstationary random processes (for references see [Section 4.7.6](#)).

4.7.5.1 Adjoint system

For an LTI system \mathbf{H} with transfer function $G(f)$, the transfer function of the *adjoint* \mathbf{H}^+ is $G^*(f)$. For general LTV systems, the GWS of the adjoint \mathbf{H}^+ (with kernel $h^+(t,t') = h^*(t',t)$) is not equal to the conjugate of the GWS of \mathbf{H} unless $\alpha = 0$. However, for an underspread LTV system \mathbf{H} this is approximately true:

$$L_{\mathbf{H}^+}^{(\alpha)}(t,f) \approx \left[L_{\mathbf{H}}^{(\alpha)}(t,f) \right]^*, \quad (4.7.6)$$

as the associated approximation error is upper bounded according to [67]

$$\left| L_{\mathbf{H}^+}^{(\alpha)}(t,f) - \left[L_{\mathbf{H}}^{(\alpha)}(t,f) \right]^* \right| \leq 4\pi|\alpha| \|S_{\mathbf{H}}\|_1 m_{\mathbf{H}}^{(1,1)} \quad (4.7.7)$$

with $\|S_{\mathbf{H}}\|_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S_{\mathbf{H}}(\nu, \tau)| d\nu d\tau$. Thus, for an underspread system with $m_{\mathbf{H}}^{(1,1)}$ small, the approximation (Eq. 4.7.6) is quite good. It follows from Eq. (4.7.6) that the GWS of a self-adjoint, underspread LTV system is approximately real-valued:

$$\mathbf{H}^+ = \mathbf{H} \implies \left[L_{\mathbf{H}}^{(\alpha)}(t,f) \right]^* \approx L_{\mathbf{H}}^{(\alpha)}(t,f), \quad \text{Im} \left\{ L_{\mathbf{H}}^{(\alpha)}(t,f) \right\} \approx 0.$$

In addition, if the underspread LTV system \mathbf{H} is positive (semi-)definite, then $\text{Re} \left\{ L_{\mathbf{H}}^{(\alpha)}(t,f) \right\} \approx L_{\mathbf{H}}^{(\alpha)}(t,f)$ is approximately nonnegative.

4.7.5.2 Composition of systems

The transfer function of a composition (series connection) of two LTI systems with transfer functions $G_1(f)$ and $G_2(f)$ is given by the product $G_1(f)G_2(f)$. A similar composition property no longer holds true for the GWS in the case of general LTV systems. However, the GWS of the composition $\mathbf{H}_2\mathbf{H}_1$ of two jointly underspread [67] LTV systems \mathbf{H}_1 and \mathbf{H}_2 is approximately equal to the product of their GWSs, that is,

$$L_{\mathbf{H}_2\mathbf{H}_1}^{(\alpha)}(t,f) \approx L_{\mathbf{H}_1}^{(\alpha)}(t,f)L_{\mathbf{H}_2}^{(\alpha)}(t,f). \quad (4.7.8)$$

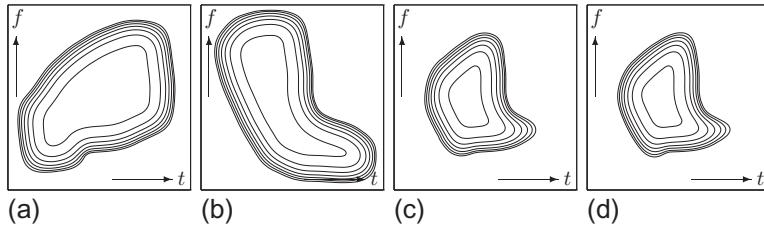


FIGURE 4.7.3

Transfer function approximation of the Weyl symbol (GWS with $\alpha = 0$) for a composition of underspread systems: (a) Weyl symbol of \mathbf{H}_1 , (b) Weyl symbol of \mathbf{H}_2 , (c) Weyl symbol of $\mathbf{H}_2\mathbf{H}_1$, and (d) product of the individual Weyl symbols $L_{\mathbf{H}_1}^{(0)}(t, f)$ and $L_{\mathbf{H}_2}^{(0)}(t, f)$. Note the similarity of (c) and (d). (Time duration = 128 samples, normalized frequency range = $[-1/4, 1/4]$.)

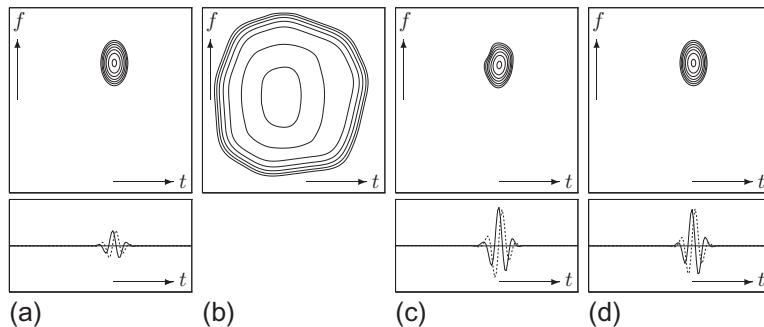
An upper bound on the associated approximation error (similar to Eq. (4.7.7)) can be provided [67]. Combining Eq. (4.7.6) with Eq. (4.7.8), we obtain $L_{\mathbf{HH}^+}^{(\alpha)}(t, f) \approx L_{\mathbf{H}^+\mathbf{H}}^{(\alpha)}(t, f) \approx |L_{\mathbf{H}}^{(\alpha)}(t, f)|^2$. Figure 4.7.3 shows an example illustrating the approximation (Eq. 4.7.8) for $\alpha = 0$. In this example, the maximum normalized error $\max_{t,f} |L_{\mathbf{H}_2\mathbf{H}_1}^{(0)}(t, f) - L_{\mathbf{H}_1}^{(0)}(t, f)L_{\mathbf{H}_2}^{(0)}(t, f)| / \max_{t,f} |L_{\mathbf{H}_2\mathbf{H}_1}^{(0)}(t, f)|$ is 0.045, which means that the approximation is quite good.

4.7.5.3 Approximate eigenfunctions and eigenvalues

The response of an LTI system with transfer function $G(f)$ to a complex sinusoid $e^{j2\pi f_0 t}$ is $G(f_0) e^{j2\pi f_0 t}$, and the response of an LFI system with temporal transfer function $w(t)$ to a Dirac impulse $\delta(t - t_0)$ is given by $w(t_0)\delta(t - t_0)$. Hence, complex sinusoids and Dirac impulses are the eigenfunctions of LTI and LFI systems, respectively, with the eigenvalues given by corresponding values of the transfer function, $G(f_0)$ and $w(t_0)$. In contrast, the eigenfunctions of general LTV systems are not localized or structured in any sense. However, for underspread LTV systems, well (t, f) localized functions are approximate eigenfunctions and the GWS $L_{\mathbf{H}}^{(\alpha)}(t, f)$ constitutes an approximate eigenvalue distribution over the (t, f) plane (for related results see Section 13.3 and [72]). Indeed, consider the family of signals $s_{t_0, f_0}(t) = s(t - t_0) e^{j2\pi f_0 t}$, where $s(t)$ is well (t, f) localized about $(0, 0)$. Evidently, $s_{t_0, f_0}(t)$ is well (t, f) localized about (t_0, f_0) . For an underspread LTV system \mathbf{H} , one can show [67]

$$(\mathbf{H}s_{t_0, f_0})(t) \approx L_{\mathbf{H}}^{(\alpha)}(t_0, f_0)s_{t_0, f_0}(t), \quad (4.7.9)$$

that is, the signals $s_{t_0, f_0}(t)$ are “approximate eigenfunctions” of \mathbf{H} , with the associated “approximate eigenvalues” given by the GWS values $L_{\mathbf{H}}^{(\alpha)}(t_0, f_0)$. Equation (4.7.9) shows that a signal well (t, f) localized about some (t, f) point passes through \mathbf{H} nearly undistorted; it is merely weighted by the GWS value at that point, thus corroborating the GWS’s interpretation as a (t, f) transfer function. Figure 4.7.4 shows the approximation (Eq. 4.7.9) for the case $\alpha = 0$; the normalized error $\|\mathbf{H}s_{t_0, f_0} - L_{\mathbf{H}}^{(0)}(t_0, f_0)s_{t_0, f_0}\|^2 / \|\mathbf{H}s_{t_0, f_0}\|^2$ (with $\|x\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 dt$) is 0.06.

**FIGURE 4.7.4**

Eigenfunction/eigenvalue approximation of the Weyl symbol (GWS with $\alpha = 0$) of an underspread LTV system:
 (a) Wigner distribution (top) and real and imaginary parts (bottom) of input signal $s_{t_0, f_0}(t)$, (b) Weyl symbol of \mathbf{H} ,
 (c) output signal $(\mathbf{H}s_{t_0, f_0})(t)$, and (d) input signal $s_{t_0, f_0}(t)$ multiplied by $L_{\mathbf{H}}(t_0, f_0)$. Note the similarity of (c)
 and (d).

4.7.5.4 Approximate uniqueness of the GWS

For an underspread LTV system \mathbf{H} , one can show [67] that the GWS is only weakly dependent on the parameter α , that is,

$$L_{\mathbf{H}}^{(\alpha_1)}(t, f) \approx L_{\mathbf{H}}^{(\alpha_2)}(t, f)$$

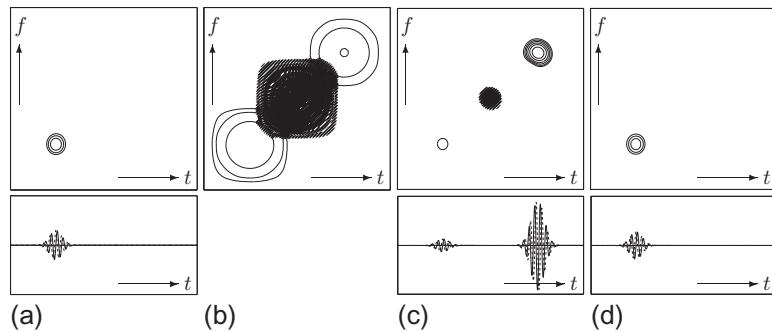
for moderate values of $|\alpha_1 - \alpha_2|$. This means that the (t, f) transfer function provided by the GWS is approximately unique. In particular, the Weyl symbol $L_{\mathbf{H}}^{(0)}(t, f)$, Zadeh's time-varying transfer function $L_{\mathbf{H}}^{(1/2)}(t, f)$, and Bello's frequency-dependent modulation function $L_{\mathbf{H}}^{(-1/2)}(t, f)$ are approximately equivalent for an underspread LTV system \mathbf{H} .

4.7.5.5 Discussion

As mentioned before, the above approximate relations extend analogous (exact) relations satisfied by the conventional transfer function of LTI/LFI systems. In this sense, the GWS is an approximate (t, f) transfer function of underspread LTV systems. Explicit upper bounds on the associated approximation errors have been formulated in terms of the GSF parameters $m_{\mathbf{H}}^{(\phi)}$, $m_{\mathbf{H}}^{(k,l)}$, or $\sigma_{\mathbf{H}}$ defined in Section 4.7.4 [67,71]. If specific such parameters are small (indicating that \mathbf{H} is underspread), the upper bounds on specific approximation errors are small and thus the respective approximation will be good. On the other hand, the above approximations and, thus, the GWS's interpretation as a (t, f) transfer function are *not* valid for overspread LTV systems. This is shown in Fig. 4.7.5.

4.7.6 SUMMARY AND CONCLUSIONS

For the practically important class of *underspread* LTV systems, there exists a simple and intuitively appealing time-frequency transfer function calculus. Indeed, the GWS can be used as an approximate *time-frequency transfer function* in a similar way as the conventional transfer function of time-invariant

**FIGURE 4.7.5**

Violation of the eigenfunction/eigenvalue approximation of the Weyl symbol in the case of an overspread LTV system: (a) Wigner distribution (top) and real and imaginary parts (bottom) of input signal $s_{t_0,f_0}(t)$, (b) Weyl symbol of \mathbf{H} , (c) output signal $(\mathbf{H}s_{t_0,f_0})(t)$, and (d) input signal $s_{t_0,f_0}(t)$ multiplied by $L_{\mathbf{H}}(t_0, f_0)$. It is instructive to compare this figure with Fig. 4.7.4; in particular, now the signals in (c) and (d) are very different. The overspread character of \mathbf{H} is indicated by the rapid oscillation of the Weyl symbol in (b) and by the fact that the signal $s_{t_0,f_0}(t)$ is partly (t, f) shifted by the system \mathbf{H} (see (a) and (b)).

systems. Applications of this time-frequency transfer function calculus to nonstationary signals and systems are considered in Sections 9.4, 9.5, 11.1, and 12.4.

4.8 WIGNER DISTRIBUTION AND FRACTIONAL FOURIER TRANSFORM⁰

4.8.1 TIME-FREQUENCY REPRESENTATIONS

The FT maps a 1D time signal $x(t)$ into a 1D spectrum denoted (in the notation to be explained in this section) by $X_{\pi/2}(f)$. A shortcoming of the FT is that it does not give an obvious indication of the time location of the spectral components, which is an important characteristic of nonstationary or time-varying signals (see Chapter 2 for more details). To describe such a signal we may use a time-frequency (t, f) representation (TFR), which maps a 1D time signal into a 2D function of time *and* frequency. Several such approaches are documented in Chapter 2 and a general formulation is given in Chapter 3. In this section, we consider the fractional FT, which belongs to the class of *linear* TFRs, and establish its connection to the WD, which is one of the most widely used *quadratic* TFRs in electrical engineering (see Section 2.1.3). In particular, we use the RWT, which relates projections of TFRs to the squared modulus of the fractional FT, and we show that all frequently used moments of the WD can be obtained from the RWT in an optimal way. Moreover, we introduce some fractional FT moments that can be useful for signal analysis.

⁰Authors: **T. Alieva**, Universidad Complutense de Madrid, Madrid, Spain (talieva@ucm.es) and **M. J. Bastiaans**, Eindhoven University of Technology, Eindhoven, The Netherlands (m.j.bastiaans@tue.nl). Reviewers: G. F. Boudreax-Bartels and L.J. Stanković.

4.8.2 WIGNER DISTRIBUTION AND AMBIGUITY FUNCTION

The WD is defined (see [Section 2.1.3](#)) as

$$\begin{aligned} W_x(t, f) &= \int_{-\infty}^{\infty} x\left(t + \frac{1}{2}\tau\right) x^*\left(t - \frac{1}{2}\tau\right) \exp(-j2\pi f\tau) d\tau \\ &= \int_{-\infty}^{\infty} X_{\pi/2}\left(f + \frac{1}{2}\nu\right) X_{\pi/2}^*\left(f - \frac{1}{2}\nu\right) \exp(j2\pi\nu t) d\nu, \end{aligned} \quad (4.8.1)$$

where $x(t)$ is a time signal and $X_{\pi/2}(f)$ its FT. The WD is always real-valued, but not necessarily positive; it preserves time and frequency shifts, and satisfies the marginal properties, which means that the frequency and time integrals of the WD, $\int W_x(t, f) df$ and $\int W_x(t, f) dt$, correspond to the signal's instantaneous power $|x(t)|^2$ and its spectral energy density $|X_{\pi/2}(f)|^2$, respectively. The WD can roughly be considered as the signal's energy distribution over the (t, f) plane, although the uncertainty principle prohibits the interpretation as a point (t, f) energy density.

If in Eq. (4.8.1) the integrations are carried out over the common variable (t or f) instead of over the difference variable (τ or ν), we get the AF $A_x(\tau, \nu)$, which is related to the WD by means of a Fourier transformation [[73](#)]:

$$A_x(\tau, \nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(t, f) \exp[-j2\pi(\nu t - f\tau)] dt df \quad (4.8.2)$$

(see [Sections 3.2.3](#) and [5.1](#) for a full treatment of the AF).

4.8.3 FRACTIONAL FOURIER TRANSFORM

The *fractional Fourier transform* of a signal $x(t)$ is defined as [[73](#)]

$$X_{\alpha}(u) = \mathbb{F}^{\alpha}[x(t)](u) = \int_{-\infty}^{\infty} K(\alpha, t, u)x(t) dt, \quad (4.8.3)$$

where the kernel $K(\alpha, t, u)$ is given by

$$K(\alpha, t, u) = \frac{\exp\left(j\frac{1}{2}\alpha\right)}{\sqrt{j \sin \alpha}} \exp\left(j\pi \frac{(t^2 + u^2) \cos \alpha - 2ut}{\sin \alpha}\right) = K(\alpha, u, t). \quad (4.8.4)$$

The fractional FT with parameter α can be considered as a generalization of the ordinary FT; thus the fractional FT for $\alpha = \frac{1}{2}\pi$ and $\alpha = -\frac{1}{2}\pi$ reduces to the ordinary and inverse FT, respectively. For $\alpha = 0$ the fractional FT corresponds to the identity operation, $X_0(u) = \mathbb{F}^0[x(t)](u) = x(u)$, and for $\alpha = \pm\pi$ to the axis reversal operation, $X_{\pm\pi}(u) = \mathbb{F}^{\pm\pi}[x(t)](u) = x(-u)$. With respect to the parameter α , the fractional FT is continuous, periodic [$\mathbb{F}^{\alpha+2\pi n} = \mathbb{F}^{\alpha}$, with n an integer] and additive [$\mathbb{F}^{\alpha}\mathbb{F}^{\beta} = \mathbb{F}^{\alpha+\beta}$], and has the symmetry relation $\mathbb{F}^{\alpha}[x^*(t)](u) = \{\mathbb{F}^{-\alpha}[x(t)](u)\}^*$. The inverse fractional FT can thus be written as

$$x(t) = \mathbb{F}^{-\alpha}[X_{\alpha}(u)](t) = \int_{-\infty}^{\infty} K(-\alpha, t, u)X_{\alpha}(u) du. \quad (4.8.5)$$

Since the Hermite-Gauss functions $\Psi_n(t) = (2^{n-1/2}n!)^{-1/2} \exp(-\pi t^2)H_n(\sqrt{2\pi}t)$, with $H_n(t)$ the Hermite polynomials, are eigenfunctions of the fractional FT with eigenvalues $\exp(-jn\alpha)$, and since

they compose a complete orthonormal set, it is possible to write the fractional FT kernel in the alternative form

$$K(\alpha, t, u) = \sum_{n=0}^{\infty} \exp(-j n \alpha) \Psi_n(t) \Psi_n(u). \quad (4.8.6)$$

The important property of the fractional FT, which establishes a connection between it, the AF, and the WD and other quadratic (t, f) distributions, is that a fractional FT produces a *rotation* of these functions in the (t, f) plane:

$$\begin{array}{ccc} x(t) & \longleftrightarrow & W_x(t, f) \text{ and } A_x(\tau, v) \\ \downarrow \text{fractional FT} & & \downarrow \text{rotation of WD and AF} \\ X_\alpha(t) = \mathbb{F}^\alpha [x] & \longleftrightarrow & W_{X_\alpha}(t, f) = W_x(t \cos \alpha - f \sin \alpha, t \sin \alpha + f \cos \alpha) \text{ and} \\ & & A_{X_\alpha}(\tau, v) = A_x(\tau \cos \alpha - v \sin \alpha, \tau \sin \alpha + v \cos \alpha). \end{array}$$

Hence, we conclude that $W_{X_\alpha}(u, v) = W_x(t, f)$, where the coordinates (u, v) in the rotated frame are related to (t, f) via the matrix relationship

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} t \\ f \end{bmatrix}. \quad (4.8.7)$$

A similar relation holds for the AF.

The main properties of the fractional FT are listed in [Table 4.8.1](#), and the fractional FT of some common functions are given in [Table 4.8.2](#).

Table 4.8.1: Fractional Fourier Transform Properties

Linearity:

$$\mathbb{F}^\alpha [ax(t) + by(t)](u) = a\mathbb{F}^\alpha [x(t)](u) + b\mathbb{F}^\alpha [y(t)](u)$$

Parseval's equality:

$$\int_{-\infty}^{\infty} x(t)y^*(t) dt = \int_{-\infty}^{\infty} X_\alpha(u)Y_\alpha^*(u) du$$

Shift theorem (real τ):

$$\mathbb{F}^\alpha [x(t - \tau)](u) = X_\alpha(u - \tau \cos \alpha) \exp[j\pi \sin \alpha (\tau^2 \cos \alpha - 2u\tau)]$$

Modulation theorem (real v):

$$\mathbb{F}^\alpha [x(t) \exp(j2\pi vt)](u) = X_\alpha(u - v \sin \alpha) \exp[-j\pi \cos \alpha (v^2 \sin \alpha - 2uv)]$$

Scaling theorem (real c and β , where $\tan \beta = c^2 \tan \alpha$):

$$\mathbb{F}^\alpha [x(ct)](u) = \sqrt{\frac{\cos \beta}{\cos \alpha}} \exp\left(j\frac{\alpha - \beta}{2}\right) \exp\left[j\pi u^2 \cot \alpha \left(1 - \frac{\cos^2 \beta}{\cos^2 \alpha}\right)\right] X_\beta\left(\frac{u \sin \beta}{c \sin \alpha}\right)$$

Table 4.8.2: Fractional Fourier Transforms of Some Common Functions

| $x(t)$ | $X_\alpha(u)$ |
|------------------------------------|---|
| $\delta(t - \tau)$ | $\frac{\exp\left(j\frac{1}{2}\alpha\right)}{\sqrt{j \sin \alpha}} \exp\left[j\pi \frac{(\tau^2 + u^2) \cos \alpha - 2u\tau}{\sin \alpha}\right]$ |
| $\exp(j2\pi t\nu)$ | $\frac{\exp\left(j\frac{1}{2}\alpha\right)}{\sqrt{\cos \alpha}} \exp\left[-j\pi(v^2 + u^2) \tan \alpha + j2\pi uv \sec \alpha\right]$ |
| $\exp(jc\pi t^2)$ | $\frac{\exp\left(j\frac{1}{2}\alpha\right)}{\sqrt{\cos \alpha + c \sin \alpha}} \exp\left(j\pi u^2 \frac{c - \tan \alpha}{1 + c \tan \alpha}\right)$ |
| $H_n(\sqrt{2\pi}t) \exp(-\pi t^2)$ | $H_n(\sqrt{2\pi}u) \exp(-\pi u^2) \exp(-jn\alpha),$ where H_n are the Hermite polynomials |
| $\exp(-c\pi t^2)$ ($c \geq 0$) | $\frac{\exp\left(j\frac{1}{2}\alpha\right)}{\sqrt{\cos \alpha + jc \sin \alpha}} \exp\left[\pi u^2 \frac{j(c^2 - 1) \cot \alpha - c \csc^2 \alpha}{c^2 + \cot^2 \alpha}\right]$ |

4.8.4 FRACTIONAL POWER SPECTRUM AND RADON-WIGNER TRANSFORM

If we introduce the *fractional power spectrum* $|X_\alpha(t)|^2$ as the squared modulus of the corresponding fractional FT, we find that these fractional power spectra are the projections of the WD upon a direction at an angle α in the (t, f) plane [73],

$$|X_\alpha(t)|^2 = \int_{-\infty}^{\infty} W_{X_\alpha}(t, f) df = \int_{-\infty}^{\infty} W_x(t \cos \alpha - f \sin \alpha, t \sin \alpha + f \cos \alpha) df, \quad (4.8.8)$$

and that they are related to the AF by a Fourier transformation:

$$|X_\alpha(t)|^2 = \int_{-\infty}^{\infty} A_x(f \sin \alpha, -f \cos \alpha) \exp(-j2\pi ft) df. \quad (4.8.9)$$

The set of fractional power spectra for the angles $\alpha \in [0, \pi]$ is called the *RWT*, because it defines the Radon transform [74] of the WD. The WD can be obtained from the RWT by applying an inverse Radon transformation. Note also that the AF can be reconstructed from the RWT by a simple inverse Fourier transformation (see Eq. (4.8.9)) and that quadratic (t, f) distributions other than the WD can be constructed subsequently.

The RWT can be considered as a quadratic TFR of $x(t)$, with advantageous properties. It is positive, invertible up to a constant phase factor, and ideally combines the concepts of the instantaneous power $|x(t)|^2$ and the spectral energy density $|X_{\pi/2}(f)|^2$. The association of the RWT with the power distributions allows its direct measurement in optics and quantum mechanics, which opens further perspectives for optical signal processing and quantum state characterization.

4.8.5 FRACTIONAL FOURIER TRANSFORM MOMENTS

The application of the different TFRs often depends on how informative their moments are and how easily these moments can be measured or calculated. The established connection between the WD and the RWT permits to find an optimal way for the calculation of the known WD moments and to introduce *fractional FT moments* that can be useful for signal analysis.

By analogy with time and frequency moments [74], given by

$$\begin{aligned} \int_{-\infty}^{\infty} t^n |x(t)|^2 dt &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n W_x(t,f) dt df, \\ \int_{-\infty}^{\infty} f^n |X_{\pi/2}(f)|^2 df &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^n W_x(t,f) dt df = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n W_{X_{\pi/2}}(t,f) dt df, \end{aligned}$$

the fractional FT moments can be introduced:

$$\frac{1}{(4\pi j)^m} \int_{-\infty}^{\infty} t^n \left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} \right)^m X_{\alpha}(t_1) X_{\alpha}^*(t_2) \Big|_{t_1=t_2=t} dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m W_{X_{\alpha}}(t,f) dt df.$$

The *zero-order* fractional FT moment E , given by

$$E = \int_{-\infty}^{\infty} |X_{\alpha}(t)|^2 dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{X_{\alpha}}(t,f) dt df = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (4.8.10)$$

is invariant under fractional Fourier transformation, which expresses the energy conservation law of a unitary transformation, also known as Parseval's relation.

The normalized *first-order* fractional FT moment m_{α} , given by

$$m_{\alpha} = \frac{1}{E} \int_{-\infty}^{\infty} t |X_{\alpha}(t)|^2 dt = \frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t W_{X_{\alpha}}(t,f) dt df \quad (4.8.11)$$

is related to the center of gravity of the fractional power spectrum. One can write the simple connection

$$m_{\alpha} = m_0 \cos \alpha + m_{\pi/2} \sin \alpha \quad (4.8.12)$$

between the first-order fractional FT moments. It is easy to see that the pair $(m_{\alpha}, m_{\alpha+\pi/2})$ is connected to $(m_0, m_{\pi/2})$ through the rotation transformation (Eq. 4.8.7), and that $m_{\alpha}^2 + m_{\alpha+\pi/2}^2 = m_0^2 + m_{\pi/2}^2$ is invariant under fractional Fourier transformation. The fractional domain corresponding to the zero-centered fractional power spectrum can be found as $\tan \alpha = -m_0/m_{\pi/2}$.

The normalized *second-order central* fractional FT moment p_{α} , given by

$$\begin{aligned} p_{\alpha} &= \frac{1}{E} \int_{-\infty}^{\infty} (t - m_{\alpha})^2 |X_{\alpha}(t)|^2 dt = \frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (t - m_{\alpha})^2 W_{X_{\alpha}}(t,f) dt df \\ &= \frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^2 W_{X_{\alpha}}(t,f) dt df - m_{\alpha}^2, \end{aligned} \quad (4.8.13)$$

is related to the effective width of the signal in the fractional FT domain. The normalized *mixed second-order central* fractional FT moment μ_{α} is given by

$$\begin{aligned}
\mu_\alpha &= \frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (t - m_\alpha)(f - m_{\alpha+\pi/2}) W_{X_\alpha}(t, f) dt df \\
&= \frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} tf W_{X_\alpha}(t, f) dt df - m_\alpha m_{\alpha+\pi/2} \\
&= \frac{1}{4\pi j E} \int_{-\infty}^{\infty} \left[\frac{\partial X_\alpha(t)}{\partial t} X_\alpha^*(t) - X_\alpha(t) \frac{\partial X_\alpha^*(t)}{\partial t} \right] t dt - m_\alpha m_{\alpha+\pi/2}.
\end{aligned} \tag{4.8.14}$$

Between the second-order fractional FT moments, the following relations hold:

$$\begin{aligned}
p_\alpha &= p_0 \cos^2 \alpha + p_{\pi/2} \sin^2 \alpha + \mu_0 \sin 2\alpha, \\
\mu_\alpha &= -\frac{1}{2}(p_0 - p_{\pi/2}) \sin 2\alpha + \mu_0 \cos 2\alpha.
\end{aligned} \tag{4.8.15}$$

In general, all second-order moments p_α and μ_α can be obtained from any three second-order moments p_α taken for three different angles α from the region $[0, \pi)$. The mixed moment μ_0 , for instance, can be expressed as $\mu_0 = -\frac{1}{2}(p_0 + p_{\pi/2}) + p_{\pi/4}$.

From Eq. (4.8.15), we conclude that the sum of the signal widths in the time and the frequency domain is invariant under fractional Fourier transformation:

$$p_\alpha + p_{\alpha+\pi/2} = p_0 + p_{\pi/2}. \tag{4.8.16}$$

Note also that the fractional domain corresponding to the extremum signal width p_α can be found by solving the equation $\tan 2\alpha = 2\mu_0/(p_0 - p_{\pi/2})$, that is, $\mu_\alpha = 0$. Due to the invariance relationship (Eq. 4.8.16), the solution of this equation corresponds to the domain with the smallest p_α and the largest $p_{\alpha+\pi/2}$, or *vice versa*.

For the product of the signal widths we find

$$p_\alpha p_{\alpha+\pi/2} = p_0 p_{\pi/2} + \frac{1}{4} \left[(p_0 - p_{\pi/2})^2 - 4\mu_0^2 \right] \sin^2 2\alpha - \frac{1}{2}\mu_0(p_0 - p_{\pi/2}) \sin 4\alpha. \tag{4.8.17}$$

This expression is not, in general, invariant under fractional Fourier transformation; invariance does, however, occur, for instance, in the case of eigenfunctions of the Fourier transformation, $g(t)$, say, for which $G_{\alpha+\pi/2}(t) = \exp(-jn\frac{1}{2}\pi)G_\alpha(t)$ (cf. the eigenvalues $\exp(-jn\alpha)$ in Eq. (4.8.6)), in which case $p_\alpha = p_0 = p_{\pi/2}$ and $\mu_\alpha = 0$. Note that, due to the uncertainty principle, we have $p_\alpha p_{\alpha+\pi/2} \geq \frac{1}{4}$. The fractional FT domain where the product $p_\alpha p_{\alpha+\pi/2}$ has an extremum value can be found by solving the equation $\tan 4\alpha = 4\mu_0(p_0 - p_{\pi/2})/[(p_0 - p_{\pi/2})^2 - 4\mu_0^2]$.

From Eq. (4.8.15) we obtain the following property for the mixed moment μ_α :

$$\mu_\alpha = \frac{1}{2}(p_{\alpha+\pi/4} - p_{\alpha-\pi/4}) = -\mu_{\alpha \pm \pi/2}. \tag{4.8.18}$$

The fractional FT moments may be helpful in the search for the most appropriate fractional domain to perform a filtering operation; in the special case of noise that is equally distributed throughout the time-frequency plane, for instance, the fractional domain with the smallest signal width is then evidently the most preferred one.

Instead of *global* moments, which we considered above, one can consider *local* fractional FT moments, which are related to such signal characteristics as the instantaneous power and IF (for $\alpha = 0$) or the spectral energy density and group delay (for $\alpha = \frac{1}{2}\pi$) in the different fractional FT domains.

The local frequency in the fractional FT domain with parameter α is defined as

$$U_{X_\alpha}(t) = \frac{\int_{-\infty}^{\infty} f W_{X_\alpha}(t, f) df}{\int_{-\infty}^{\infty} W_{X_\alpha}(t, f) df} = \frac{1}{2|X_\alpha(t)|^2} \int_{-\infty}^{\infty} \left. \frac{\partial |X_\beta(\tau)|^2}{\partial \beta} \right|_{\beta=\alpha} \operatorname{sgn}(\tau - t) d\tau. \quad (4.8.19)$$

The local frequency $U_{X_\alpha}(t)$ is related to the phase $\varphi_\alpha(t) = \arg X_\alpha(t)$ of the fractional FT $X_\alpha(t)$ through $U_{X_\alpha}(t) = (1/2\pi) d\varphi_\alpha(t)/dt$. This implies that the derivative of the fractional power spectra with respect to the angle α defines the local frequency in the fractional domain, and that it can be used for solving the phase retrieval problem by measuring intensity functions only.

We finally mention the relationship between the central local fractional second-order moment and the instantaneous power in the fractional FT domain:

$$V_{X_\alpha}(t) = \frac{\int_{-\infty}^{\infty} [f - U_{X_\alpha}(t)]^2 W_{X_\alpha}(t, f) df}{\int_{-\infty}^{\infty} W_{X_\alpha}(t, f) df} = -\frac{1}{4} \frac{d^2}{dt^2} (\ln |X_\alpha(t)|^2). \quad (4.8.20)$$

We conclude that all widely used moments of the WD are obtainable from the RWT.

4.8.6 APPLICATIONS

The fractional FT and the WD are applied in quantum mechanics, optics, and signal processing. While here only the 1D fractional FT has been considered, its generalization to two or more dimensions is straightforward. Increasing the signal dimension yields different transformations, which are inherently related to the fractional FT and also describe the WD rotation in phase space [74].

The wide application of the fractional FT in optics is based on the fact that—in the paraxial approximation of the scalar diffraction theory—it describes the optical field evolution during propagation through a quadratic-refractive-index medium (e.g., fibers) or a system of lenses [73,75]. The RWT, associated with the intensity distributions, is used, in particular, for the reconstruction of the WD and subsequently of the complex field amplitude (in the case of coherent light) [76] or the two-point correlation function (in the case of partially coherent light) [77].

In signal processing, the RWT was primarily developed for detection and classification of multicomponent LFM signals in noise [78,79]. Since the fractional FT of the chirp-type signal $\exp(-j\pi t^2/\tan \beta)$ reads

$$\exp\left(j\frac{1}{2}\alpha\right) \sqrt{\sin \beta / \sin(\beta - \alpha)} \exp\left[-j\pi u^2 / \tan(\beta - \alpha)\right],$$

it becomes proportional to a Dirac-function $\delta(u)$ for $\alpha \rightarrow \beta$, and it can be detected as a local maximum on the RWT map. Analogously, in order to remove chirp-type noise, a notch filter, which minimizes the signal information loss, can be placed at the proper point of the corresponding fractional FT domain [73].

Instead of performing, as usual, filtering operations in the frequency or the time domain, it can be done in a more appropriate fractional domain, for instance, the one that corresponds to the best signal/noise time-frequency separation [73,80]. Proposed applications of the fractional FT include motion analysis [81,82], neural network implementations [83], image encoding [84], and watermarking [85], to name a few.

The complexity of computation for the fractional FT is $O(N \log N)$, where N is the time-bandwidth product of the signal [86]. This is relatively low, allowing wide use in digital signal and image processing.

4.8.7 SUMMARY AND CONCLUSIONS

We have described the relationship between the fractional FT and the WD by using the RWT, which is a set of projections of the WD as well as a set of squared moduli of the fractional FT. We have introduced the concept of fractional FT moments and have proposed a way for the calculation of the well-known global and local moments of the WD, based on the knowledge of a few fractional power spectra. The application of the results in optics and signal processing has been discussed briefly. [Section 5.8](#) further explores the relationship between the fractional FT and quadratic TFDs.

4.9 A TIME-FREQUENCY PERSPECTIVE ON SYSTEMS: FROM SISO TO MIMO⁰

Signals in nature are commonly modeled as the output of systems defined by differential equations. Transforms are a powerful tool to design and analyze such systems and, ultimately, to understand the nature of the signals that they model. A highly effective transformation method dates back to the beginning of 1800, when Fourier solved the heat equation by decomposing the solution in a sum of sinusoids. Today, the Fourier transformation to the frequency domain is an essential tool in engineering. The analysis and design of electrical and electronic devices, vibrating structures, and mechanical systems, just to mention a few, benefit immensely from the transformation to the frequency domain. Nevertheless, in the last decades, it has become clear that signals in nature are inherently nonstationary, and their frequency content changes with time. Time-frequency analysis provides an effective representation of such time-varying frequency content. Similarly to frequency analysis, time-frequency analysis can be applied to the system that models the signal.

In this section, we illustrate how systems defined by differential equations can be transformed to the time-frequency domain. We consider both single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems. First, we show that the transformation generates a time-frequency system described by a differential equation. Then, we discuss the properties of the time-frequency system. Furthermore, we describe calculation techniques to obtain the time-frequency output when the input of the system in time is nonstationary. Finally, we show the time-frequency impulse response of an SISO system, and the transient spectrum of an MIMO system subject to a nonstationary noise at the input. Our results prove that the time-frequency representation of systems reveals the inner mechanisms involved in the generation of nonstationary signals, and fosters the development of novel system analysis and design methods.

⁰Author: **Lorenzo Galleani**, Politecnico di Torino, Torino, Italy (galleani@polito.it). Reviewers: A. Napolitano and Y. Zhang.

4.9.1 SISO SYSTEMS

First, we discuss the transformation of an SISO system from time to time-frequency. Then, we describe techniques for calculating the time-frequency output when the input is nonstationary. Finally we analyze the time-frequency impulse response of a system with two resonances.

4.9.1.1 Transformation to the time-frequency domain

We consider an SISO dynamical system defined by the differential equation

$$a_n \frac{d^n y(t)}{dt^n} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = x(t), \quad (4.9.1)$$

where a_0, \dots, a_n are deterministic coefficients, $x(t)$ is the forcing term or input, and $y(t)$ is the solution or output. The state of the system is the vector with components $y(t), \dots, y^{(n-1)}(t)$, where $y^{(k)}(t) = d^k y(t)/dt^k$, and the evolution rule is the differential equation. When $x(t)$ is deterministic, the output $y(t)$ is deterministic and, given an initial state, we can compute the state at any time by using the evolution rule. Consequently, as Eq. (4.9.1) is deterministic, we refer to the SISO system as deterministic. When $x(t)$ is random, the output $y(t)$ is random and we can make only probabilistic predictions about its future values. In this case, Eq. (4.9.1) is a stochastic differential equation, and we refer to the SISO system as random.

Wigner distribution. We factor Eq. (4.9.1) operationally as

$$a_n \left[\prod_{m=1}^n (D - \lambda_m) \right] y(t) = x(t), \quad (4.9.2)$$

where $D = \frac{d}{dt}$, and the complex quantities $\lambda_1, \dots, \lambda_n$ are the poles obtained by solving $P_n(\lambda) = 0$, where $P_n(\lambda) = a_n \lambda^n + \cdots + a_1 \lambda + a_0$. The corresponding equation for the WD is¹ [87,88]

$$\frac{|a_n|^2}{4^n} \left[\prod_{m=1}^n (\partial_t - p_m(f)) (\partial_t - p_m^*(f)) \right] W_y(t,f) = W_x(t,f), \quad (4.9.3)$$

where $\partial_t = \frac{\partial}{\partial t}$, the star indicates complex conjugation,

$$p_m(f) = 2\alpha_m + 2j(\beta_m - 2\pi f) \quad (4.9.4)$$

is a time-frequency pole [89], and α_m, β_m are the real and imaginary parts of λ_m , respectively. We point out that we do not use the analytic signal in the evaluation of $W_x(t,f)$ and $W_y(t,f)$. The results still apply to the WVD though, since there is a one-to-one relationship between the real signal and the analytic signal, therefore between the WD and the WVD. See Sections 2.1.3 and 2.1.4 for the definition and properties of the WD and WVDs, respectively. When $x(t)$ is a nonstationary random process, $y(t)$ is, in general, a nonstationary random process. The corresponding time-frequency system is given by [87]

$$\frac{|a_n|^2}{4^n} \left[\prod_{m=1}^n (\partial_t - p_m(f)) (\partial_t - p_m^*(f)) \right] \overline{W}_y(t,f) = \overline{W}_x(t,f), \quad (4.9.5)$$

¹In the referenced sections written by the author, the results are given with respect to the angular frequency $\omega = 2\pi f$, and with a slightly different definition of the general class of QTDFs.

where $\overline{W}_x(t, f)$ and $\overline{W}_y(t, f)$ are Wigner spectra [90], defined as the expected values of the corresponding WDs. Note that both $\overline{W}_x(t, f)$ and $\overline{W}_y(t, f)$ are deterministic quantities defined in Section 2.1.

It is useful to discuss some properties of the time-frequency system for the deterministic case, Eq. (4.9.3). These properties hold also for the random case, Eq. (4.9.5), because its structure is identical to the deterministic case.

Input/output. We note that the input of the time-frequency system is the WD of the input in time, and the output is the WD of the output in time.

Linearity. Although the WD is a nonlinear transformation, the time-frequency system is linear with respect to the WD of the input.

Differential structure. The differential equation defining the time-frequency system contains derivatives with respect to time only. Therefore, it can be solved as an ordinary differential equation, with frequency playing the role of a parameter. This property, together with the linearity of the equation, paves the way for a variety of tools, such as the Laplace transform.

Order of the system. The order of the time-frequency system is $2n$, whereas the order of the system in time is n . This property is due to the quadratic nature of the WD.

Time-frequency poles. Since the transformation to the time-frequency domain doubles the order of the system, the number of time-frequency poles is $2n$, whereas the number of poles of the system in time is n . A pair of complex conjugate poles $p_m(f), p_m^*(f)$ is in fact associated to every pole λ_m .

Coefficients. The time-frequency system can also be written as

$$b_{2n} \partial_t^{2n} \overline{W}_y(t, f) + \cdots + b_1 \partial_t \overline{W}_y(t, f) + b_0 \overline{W}_y(t, f) = \overline{W}_x(t, f), \quad (4.9.6)$$

where

$$b_{2n-k} = (-1)^k b_{2n} \Pi_k(\bar{p}_1, \dots, \bar{p}_{2n}), \quad k = 1, \dots, 2n-1, \quad (4.9.7)$$

$$b_{2n} = \frac{|a_n|^2}{4^n}, \quad b_0 = \frac{1}{|H(f)|^2}, \quad (4.9.8)$$

with $H(f) = 1/P_n(j2\pi f)$ being the system transfer function, $\Pi_k(\bar{p}_1, \dots, \bar{p}_{2n}) = \sum_{1 \leq m_1 < \dots < m_k \leq k} \bar{p}_{m_1} \cdots \bar{p}_{m_k}$ the k th symmetric sum, and $\bar{p}_{2k-1} = p_k$, $\bar{p}_{2k} = p_k^*$. By replacing the definition of the time-frequency poles, we can relate the coefficients of the time-frequency system to the system poles $\lambda_1, \dots, \lambda_n$.

Quadratic class. We rewrite the system in the polynomial notation

$$P_n(D)y(t) = x(t). \quad (4.9.9)$$

The time-frequency system for an arbitrary time-frequency representation from the quadratic class (as defined in Section 3.2.2) is [91]

$$P_n^*(\mathcal{A}_\rho)P_n(\mathcal{B}_\rho)\rho_y(t, f) = \rho_x(t, f), \quad (4.9.10)$$

where

$$\mathcal{A}_\rho = \frac{1}{2} \frac{\partial}{\partial t} + j2\pi f - \frac{\partial}{\partial \tau} \log g \left(\frac{1}{j2\pi} \frac{\partial}{\partial t}, -\frac{1}{j2\pi} \frac{\partial}{\partial f} \right), \quad (4.9.11)$$

$$\mathcal{B}_\rho = \frac{1}{2} \frac{\partial}{\partial t} - j2\pi f + \frac{\partial}{\partial \tau} \log g \left(\frac{1}{j2\pi} \frac{\partial}{\partial t}, -\frac{1}{j2\pi} \frac{\partial}{\partial f} \right). \quad (4.9.12)$$

Table 4.9.1: Operators \mathcal{A}_ρ and \mathcal{B}_ρ for Some Common QTFDs

| Distribution | $g(v, \tau)$ | \mathcal{A}_ρ | \mathcal{B}_ρ |
|-----------------|---|---|---|
| Wigner | 1 | $\frac{1}{2} \frac{\partial}{\partial t} + j2\pi f$ | $\frac{1}{2} \frac{\partial}{\partial t} - j2\pi f$ |
| Smoothed Wigner | $\frac{1}{2\pi} \sigma_t \sigma_f e^{-\frac{(2\pi v)^2}{2/\sigma_t^2} - \frac{\tau^2}{2/\sigma_f^2}}$ | $\frac{1}{2} \frac{\partial}{\partial t} + j2\pi f + j\frac{\sigma_f^2}{2\pi} \frac{\partial}{\partial f}$ | $\frac{1}{2} \frac{\partial}{\partial t} - j2\pi f - j\frac{\sigma_f^2}{2\pi} \frac{\partial}{\partial f}$ |
| Rihaczek | $e^{-j\pi v \tau}$ | $\frac{\partial}{\partial t} + j2\pi f$ | $-j2\pi f$ |
| Exponential | $e^{-v^2 \tau^2 / \sigma}$ | $\frac{1}{2} \frac{\partial}{\partial t} + j2\pi f - \frac{j}{4\pi^3 \sigma} \frac{\partial^2}{\partial t^2} \frac{\partial}{\partial f}$ | $\frac{1}{2} \frac{\partial}{\partial t} - j2\pi f + \frac{j}{4\pi^3 \sigma} \frac{\partial^2}{\partial t^2} \frac{\partial}{\partial f}$ |

For the random case, the equation becomes

$$P_n^*(\mathcal{A}_\rho) P_n(\mathcal{B}_\rho) \bar{\rho}_y(t, f) = \bar{\rho}_x(t, f), \quad (4.9.13)$$

where $\bar{\rho}_x(t, f)$ and $\bar{\rho}_y(t, f)$ are the expected values of $\rho_x(t, f)$ and $\rho_y(t, f)$, respectively. In [Table 4.9.1](#), we give the operators \mathcal{A}_ρ and \mathcal{B}_ρ for some common TFDs. We observe that the operators \mathcal{A}_ρ and \mathcal{B}_ρ contain, in general, derivatives with respect to both time and frequency. Therefore, for the quadratic class, the time-frequency system is defined, in general, by a partial differential equation. Nevertheless, the time-frequency system is still linear and, therefore, the partial differential equation can be solved with the analytic tools available for this class of equations.

4.9.1.2 Techniques for computing the time-frequency output when the input is nonstationary

By using the properties of the time-frequency system, we can compute the time-frequency output for a series of common nonstationary inputs. We first discuss the case of the WD, and then we show how to smooth it.

Laplace transform. The Laplace transform of the time-frequency system for the WD, Eq. (4.9.3), is

$$\frac{|a_n|^2}{4^n} \left[\prod_{m=1}^n (s - p_m)(s - p_m^*) \right] W_y(s, f) = W_x(s, f), \quad (4.9.14)$$

where

$$W_y(s, f) = \int_0^{+\infty} W_y(t, f) e^{-st} dt \quad (4.9.15)$$

and $W_x(s, f)$ is defined accordingly. Since the obtained equation is algebraic, we can readily invert it, obtaining

$$W_y(s, f) = \frac{4^n}{|a_n|^2} \frac{W_x(s, f)}{\prod_{m=1}^n (s - p_m)(s - p_m^*)}. \quad (4.9.16)$$

For several types of $W_x(s, f)$, we can expand the right-hand side into a sum of partial fractions, which can be straightforwardly inverted to give the Wigner output $W_y(t, f)$. The time-frequency poles greatly simplify this operation.

Generalized input/output method. The calculation of the time-frequency output can be considerably simplified by a method that takes advantage of the linearity of the time-frequency system [92]. We start by decomposing the Wigner input as

$$W_x(t, f) = \sum_{k=1}^N X_k(t, f), \quad (4.9.17)$$

where the quantities $X_k(t, f)$, $k = 1, \dots, N$ are called the *generalized* inputs, recognizing the fact that they are not, in general, proper WDs. The corresponding Wigner output is

$$W_y(t, f) = \mathcal{L}W_x(t, f), \quad (4.9.18)$$

where \mathcal{L} is a linear operator representing the convolution in time between the impulse response $W_0(t, f)$ of the time-frequency system, obtained by solving

$$\frac{|a_n|^2}{4^n} \left[\prod_{m=1}^n (\partial_t - p_m(f)) (\partial_t - p_m^*(f)) \right] W_0(t, f) = \delta(t), \quad (4.9.19)$$

and the Wigner input. Note that the convolution is with respect to time only, because the differential equation defining the time-frequency system does not contain derivatives with respect to f , as discussed under “Differential structure” in Section 4.9.1.1 above. By replacing Eq. (4.9.17), we have

$$W_y(t, f) = \sum_{k=1}^N Y_k(t, f), \quad (4.9.20)$$

where

$$Y_k(t, f) = \mathcal{L}X_k(t, f) \quad (4.9.21)$$

are the generalized outputs, that is, the outputs of the time-frequency system corresponding to the generalized inputs.

The key idea is to choose generalized inputs so that it is easy to compute the corresponding generalized outputs. The Wigner output is obtained by simply adding the obtained generalized outputs. The generalized inputs provided in [92] allow exact calculation of the time-frequency output for a delta function, a linear chirp, a causal sinusoid, and a short-duration sinusoid. These nonstationary inputs are commonly observed in physical systems. Without the generalized input/output method, direct calculation of these time-frequency outputs leads to intractable expressions.

Smoothing the time-frequency output. To reduce the interference terms of the WD, we can smooth the Wigner outputs obtained by the Laplace transform or the generalized input/output method. An effective smoothing is

$$\rho_x(t, f) = \int_{-\infty}^{+\infty} w(t - t') W_x(t', f) dt', \quad (4.9.22)$$

where $w(t)$ is the smoothing window and $\rho_x(t, f)$ the corresponding smoothed WD. We consider the smoothing window of duration T_w

$$w(t) = \left(a + b \cos \frac{2\pi t}{T_w} \right) \text{rect} \left(\frac{t}{T_w} \right). \quad (4.9.23)$$

By changing the parameters a and b , we obtain some of the most common smoothing windows: for $a = 0.5$, $b = 0.5$ the Hann window; for $a = 0.54$, $b = 0.46$ the Hamming window; and for $a = 1$, $b = 0$ the rectangular window.

This smoothing is particularly effective with the generalized input/output method. If \mathcal{L}_s is the linear operator corresponding to the smoothing in Eq. (4.9.22), we can compute the smoothed generalized outputs as

$$\bar{Y}_k(t,f) = \mathcal{L}_s Y_k(t,f). \quad (4.9.24)$$

Then, we combine the smoothed generalized outputs according to Eq. (4.9.20), and we obtain the smoothed Wigner outputs

$$\rho_y(t,f) = \sum_{k=1}^N \bar{Y}_k(t,f). \quad (4.9.25)$$

4.9.1.3 Example: time-frequency impulse response

We consider a system with two resonances with poles $\lambda_1 = -2.5 + 6\pi j$, $\lambda_2 = \lambda_1^*$, $\lambda_3 = -0.8 + 15\pi j$, and $\lambda_4 = \lambda_3^*$. Note that we use dimensionless quantities for simplicity. The time-frequency impulse response, that is, the smoothed Wigner output when the input in time is a delta function, can be computed exactly with the Laplace transform and the time-frequency poles [92]. Figure 4.9.1 shows the obtained smoothed Wigner output for a Hamming window with $T_w = 0.51$. At $t = 0$ the delta function at the input excites all frequencies. The system reacts with a stronger response at the resonant frequencies $f_1 \approx 3$ and $f_2 \approx 7.5$. The response at the higher resonant frequency f_2 lasts longer because of the smaller damping of this resonance.

4.9.2 MIMO SYSTEMS

We first discuss the transformation of an MIMO system to the time-frequency domain and then we illustrate our results with an example.

4.9.2.1 Transformation to the time-frequency domain

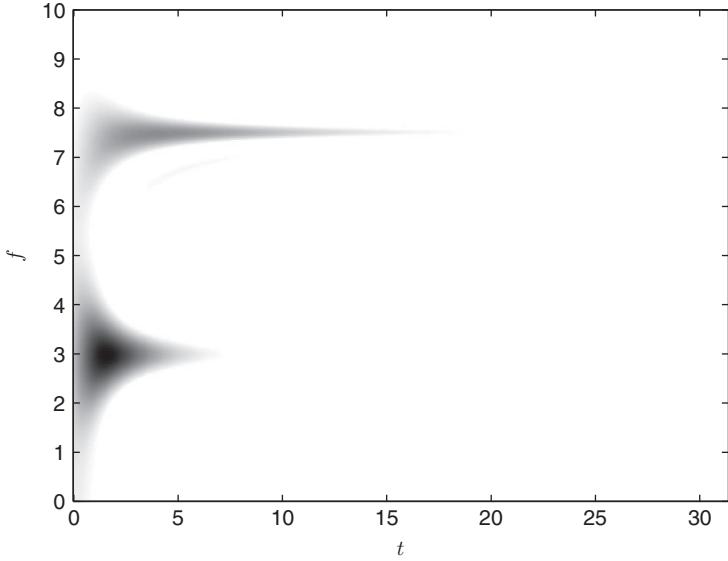
Consider the MIMO dynamical system described by the matrix differential equation

$$\dot{\mathbf{y}}(t) = \mathbf{M}\mathbf{y}(t) + \mathbf{x}(t), \quad (4.9.26)$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix}, \quad \mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_N(t) \end{bmatrix} \quad (4.9.27)$$

are, respectively, the forcing term, or input, and the solution, or output. Moreover, $\dot{\mathbf{y}}(t)$ is an N -by-1 vector whose k th component is $d^k y(t)/dt^k$, and \mathbf{M} is an N -by- N complex matrix. The vector $\mathbf{y}(t)$ is the state of the system, and the law of evolution is the matrix differential equation. When $\mathbf{x}(t)$ is deterministic, the output $\mathbf{y}(t)$ is deterministic and the MIMO system is deterministic, whereas when $\mathbf{x}(t)$ is random, the output $\mathbf{y}(t)$ is random and the MIMO system is random.

**FIGURE 4.9.1**

Time-frequency impulse response of a system with two resonances.

Deterministic case. The MIMO system in Eq. (4.9.26) can be transformed to the time-frequency domain, and the result of this transformation is the time-frequency system [93]

$$(\mathcal{A}_\rho \mathbf{I} - \mathbf{M}^*) \boldsymbol{\rho}_y(t,f) (\mathcal{B}_\rho \mathbf{I} - \mathbf{M}^T) = \boldsymbol{\rho}_x(t,f), \quad (4.9.28)$$

where \mathbf{I} is the identity matrix, \mathbf{M}^T is the transpose of \mathbf{M} ,

$$\boldsymbol{\rho}_y(t,f) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(v, \tau) y(u + \frac{\tau}{2}) y^*(u - \frac{\tau}{2}) e^{j2\pi(vt - vu - f\tau)} du dv d\tau \quad (4.9.29)$$

is the quadratic time-frequency representation for a vector signal $\mathbf{x}(t)$ [94], and the operators \mathcal{A}_ρ , \mathcal{B}_ρ are defined in Eqs. (4.9.11) and (4.9.12).

Similarly to the SISO case, the input of the time-frequency system is the time-frequency representation of the input in time, and the output is the time-frequency representation of the output in time. The time-frequency system is again linear, although the time-frequency representation of Eq. (4.9.29) is defined by a nonlinear transformation.

Surprisingly, the time-frequency representation corresponds to N^2 scalar differential equations, whereas the MIMO system in time corresponds to N equations only. As a matter of fact, $\boldsymbol{\rho}_y(t,f)$ is defined as

$$\boldsymbol{\rho}_y(t,f) = \begin{bmatrix} \rho_{y_1,y_1}(t,f) & \cdots & \rho_{y_1,y_N}(t,f) \\ \vdots & \ddots & \vdots \\ \rho_{y_N,y_1}(t,f) & \cdots & \rho_{y_N,y_N}(t,f) \end{bmatrix}. \quad (4.9.30)$$

Therefore, Eq. (4.9.28) defines a differential equation for each of the N^2 entries $\rho_{y_k,y_l}(t,f)$, $k,l = 1,\dots,N$, of this matrix. Clearly, our main interest is in the N time-frequency representations $\rho_{y_1,y_1}(t,f), \dots, \rho_{y_N,y_N}(t,f)$ of the components $y_1(t), \dots, y_N(t)$ of the state vector, but to obtain them we need to solve the N^2 time-frequency equations. Each of the N^2 time-frequency equations is, in general, a partial differential equation, because the operators \mathcal{A}_ρ and \mathcal{B}_ρ contain derivatives with respect to both time and frequency.

For the case of the WD, the time-frequency system becomes

$$\left[\left(\frac{1}{2} \partial_t + j2\pi f \right) \mathbf{I} - \mathbf{M}^* \right] \mathbf{W}_y(t,f) \left[\left(\frac{1}{2} \partial_t - j2\pi f \right) \mathbf{I} - \mathbf{M}^T \right] = \mathbf{W}_x(t,f). \quad (4.9.31)$$

Since this matrix differential equation has derivatives with respect to time only, the corresponding time-frequency system is made by N^2 ordinary differential equations, and the frequency f plays the role of a parameter in each of them.

Random case. The time-frequency system for a random MIMO system is

$$(\mathcal{A}_\rho \mathbf{I} - \mathbf{M}^*) \bar{\rho}_y(t,f) (\mathcal{B}_\rho \mathbf{I} - \mathbf{M}^T) = \bar{\rho}_x(t,f), \quad (4.9.32)$$

where the time-frequency representation $\bar{\rho}_y(t,f)$ is defined as

$$\bar{\rho}_y(t,f) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(v,\tau) E \{ y(u + \frac{\tau}{2}) y^*(u - \frac{\tau}{2}) \} e^{j2\pi(vt - vu - f\tau)} du dv d\tau \quad (4.9.33)$$

and $E\{\cdot\}$ is the expected value. The matrix $\bar{\rho}_x(t,f)$ is defined in an identical way. Both the time-frequency input $\bar{\rho}_x(t,f)$ and output $\bar{\rho}_y(t,f)$ are deterministic quantities. We note that the time-frequency system for the random case is identical to the time-frequency system of the deterministic case. The equation for the Wigner spectrum is

$$\left[\left(\frac{1}{2} \partial_t + j2\pi f \right) \mathbf{I} - \mathbf{M}^* \right] \bar{\mathbf{W}}_y(t,f) \left[\left(\frac{1}{2} \partial_t - j2\pi f \right) \mathbf{I} - \mathbf{M}^T \right] = \bar{\mathbf{W}}_x(t,f). \quad (4.9.34)$$

4.9.2.2 Example: the transient spectrum of a random MIMO system

We consider the MIMO system with order $N = 4$ shown on the left of Fig. 4.9.2. This system is governed by the equation

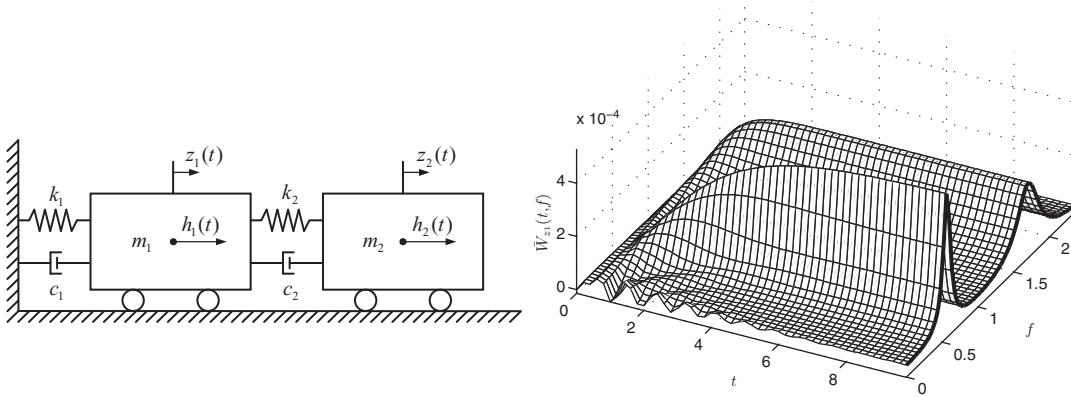
$$\mathbf{M}_2 \ddot{\mathbf{z}}(t) + \mathbf{C}_2 \dot{\mathbf{z}}(t) + \mathbf{K} \mathbf{z}(t) = \mathbf{h}(t), \quad (4.9.35)$$

where

$$\mathbf{z}(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}, \quad \mathbf{h}(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix}, \quad (4.9.36)$$

and

$$\mathbf{M}_2 = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} c_1 + c_2 & -c_2 \\ -c_2 & c_2 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix}. \quad (4.9.37)$$

**FIGURE 4.9.2**

MIMO system with two vibrational modes (left), and its “(1, 1)” transient spectrum (right), in which the thick line represents, up to a multiplicative constant, the power spectrum computed with classical techniques for WSS random processes.

The components $h_1(t), h_2(t)$ of $\mathbf{h}(t)$ are the input forces, whereas the displacements $z_1(t), z_2(t)$ represent the two outputs. The quantities c_1, c_2 and k_1, k_2 are the damping and stiffness coefficients of the masses m_1, m_2 , respectively. This MIMO system has two vibrational modes at the approximate natural frequencies obtained by solving

$$\det\{\mathbf{K} - 4\pi^2 f^2 \mathbf{M}_2\} = 0, \quad (4.9.38)$$

where \det is the matrix determinant. We compute the time-frequency output of the system when the input in time is a noise that is turned on at $t = 0$, namely

$$\mathbf{h}(t) = u(t) \begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \end{bmatrix}, \quad (4.9.39)$$

where $u(t)$ is the Heaviside step function, and each component $\epsilon_k(t)$ is a white Gaussian noise with zero mean and autocorrelation function

$$R_{\epsilon_k}(t_1, t_2) = E\{\epsilon_k(t_1)\epsilon_k(t_2)\} = \delta(t_1 - t_2), \quad (4.9.40)$$

for $k = 1, 2$, and we assume $\epsilon_1(t)$ and $\epsilon_2(t)$ to be uncorrelated.

Equation (4.9.35) can be put in the standard form of Eq. (4.9.26) by setting

$$\mathbf{y}(t) = \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}, \quad \mathbf{x}(t) = \begin{bmatrix} 0 \\ 0 \\ \mathbf{M}_2^{-1}\mathbf{h}(t) \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\mathbf{M}_2^{-1}\mathbf{K} & -\mathbf{M}_2^{-1}\mathbf{C}_2 \end{bmatrix}. \quad (4.9.41)$$

The Wigner spectrum of the input is given by [93]

$$\overline{W}_x(t,f) = \frac{1}{2\pi} u(t) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/m_1^2 & 0 \\ 0 & 0 & 0 & 1/m_2^2 \end{bmatrix}. \quad (4.9.42)$$

Because of the structure of the input Wigner spectrum and of the properties of the time-frequency system, the output Wigner spectrum $\overline{W}_y(t,f)$ can be obtained exactly. On the right of Fig. 4.9.2, we show its (1, 1) entry, corresponding to $\overline{W}_{z_1}(t,f)$, the Wigner spectrum of the output $z_1(t)$, for the dimensionless parameters $m_1 = m_2 = c_1 = c_2 = 1$, $k_1 = 100$, and $k_2 = 15$. From Eq. (4.9.38), the corresponding approximate natural frequencies of the two modes are $f_1 \approx 0.57$ and $f_2 \approx 1.72$, respectively. At $t = 0$ the input noise starts injecting energy in the system. Consequently, we see that the two components corresponding to the vibrational modes rise during a transient phase and eventually reach a steady state. The time-frequency spectrum of the random system during the transient phase is the transient spectrum [87]. For large times, the transient spectrum equals, up to a constant factor, the power spectrum obtained for the wide-sense stationary (WSS) case (thick line), namely, assuming that the input noise is turned on at $t = -\infty$.

4.9.3 SUMMARY AND CONCLUSIONS

Transforming a system in the time-frequency domain is advantageous because we better understand how the system generates a nonstationary output by processing a nonstationary input. In this section, we have discussed the transformation for SISO and MIMO systems, both deterministic and random. The transformation generates a time-frequency system whose input is the time-frequency representation of the input and whose output is the time-frequency representation of the output. We have described techniques for the computation of the time-frequency output when the input in time is nonstationary, and we have illustrated our method with two examples.

In addition to a deeper understanding of the spectral mechanisms involved in the generation of nonstationary processes, the time-frequency system and the solution techniques developed can lead to improved system design and analysis methods that operate directly in the time-frequency domain.

4.10 TEAGER-KAISER ENERGY OPERATORS IN TIME-FREQUENCY ANALYSIS⁰

4.10.1 CONCEPT OF ENERGY

Estimation of the signal energy is a commonly used measurement in engineering. Traditionally in signal processing, the energy of a signal is defined using the square of the signal magnitude, the envelope of squared signal magnitude or the integral of squared signal magnitude. Time-frequency (t,f) methods

⁰Authors: **A.-O. Boudraa**, Institut de Recherche de l'École Navale, Almada, France (boudra@ecole-navale.fr) and **F. Salzenstein**, University of Strasbourg, Strasbourg, France (f.salzenstein@unistra.fr). Reviewers: J. O'Toole and B. Jawad.

can be used to determine the signal energy content in both time and frequency bins and particularly for the study of nonstationary signals (see [Chapter 2](#) and [Chapters 1–3](#)). All these approaches estimate the signal energy without referring to what the signal represents. An alternative approach is to define the energy of the signal based on physical principles. This approach is well evidenced in the speech signal production principle. Indeed, in nonlinear speech modeling, the dominance of modulation as a process in the production of speech points to the importance of analyzing signals from the point of view of the energy required to generate them [95]. It was also observed that in the conventional view of energy, tones at 10 and 1000 Hz of the same amplitude have the same energy. However, the energy required to generate the signal of 1000 Hz is much greater than that for the 10 Hz signal. One observation was that speech resonances can vary rapidly both in frequency and amplitude, essentially due to the separated air flow in the vocal tract [96]. This alternative notion of energy can be understood by focusing attention on the generation process of the signal [97]. Based on the theory of oscillation of a simple spring-mass system ([Fig. 4.10.1](#)), an energy operator is defined to measure the energy of a speech signal, which is produced by a nonlinear process. Hearing can then be considered as detection of energy.

4.10.2 TEAGER-KAISER ENERGY OPERATOR

A simple and elegant form of the energy operator can be derived using basic physics of motion for a simple spring and mass oscillator [97]. Let us consider the law of motion of a mass m suspended by a spring with a constant $K > 0$:

$$\ddot{x}(t) + \frac{K}{m}x(t) = 0. \quad (4.10.1)$$

Equation (4.10.1) represents an undamped harmonic oscillator because it ignores the properties of the surrounding medium. We view it as an *approximate* model of a mechanical-acoustical system, in which the object oscillates, creating pressure waves in the medium. The solution of Eq. (4.10.1) is the oscillating signal $x(t) = A \cos(2\pi ft + \phi_0)$, where A is the amplitude, $f = \frac{1}{2\pi}\sqrt{K/m}$ is the frequency, and ϕ_0 is the initial arbitrary phase. The total energy E of this system equals the sum of the potential energy (in the spring) and the kinetic energy (of the mass); that is,

$$E = \frac{K}{2}x^2(t) + \frac{m}{2}\dot{x}^2(t). \quad (4.10.2)$$

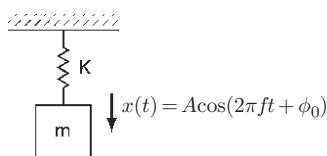


FIGURE 4.10.1

A spring-mass harmonic oscillator; $x(t)$ is the displacement from equilibrium.

By substituting for $\dot{x}(t)$ and $x(t)$ in Eq. (4.10.2), we get

$$E = \frac{m}{2} (2\pi A f)^2 \propto A^2 f^2, \quad (4.10.3)$$

which is the true energy in the harmonic oscillator system, that is, the energy required to generate the signal. Equation (4.10.3) shows that the energy is related to both A and f . Equation (4.10.2) indicates that the conventional energy approximation $x^2(t)$ only takes into account the *potential* energy of the signal source.

Consider a discrete-time signal representing the motion of the oscillator body:

$$x(n) = A \cos \left(\frac{2\pi f}{f_s} n + \phi_0 \right), \quad (4.10.4)$$

where f is the analog frequency and f_s the sampling rate. Using equally spaced samples $x(n - 1)$, $x(n)$, and $x(n + 1)$ and appropriate trigonometric identities, we find

$$x^2(n) - x(n - 1)x(n + 1) = A^2 \sin^2(2\pi f/f_s). \quad (4.10.5)$$

Equation (4.10.5) is exact and is unique provided that $2\pi f/f_s$ is restricted to values less than $\pi/2$. Equation (4.10.5) can be approximated by noting that if $2\pi f/f_s$ is small, $\sin(2\pi f/f_s) \approx 2\pi f/f_s$. The largest deviation from $\sin(2\pi f/f_s)$ is at $f/f_s = 1/8$. The ratio between $2\pi f/f_s$ and $\sin(2\pi f/f_s)$ is $\frac{\pi/4}{\sqrt{2}/2} \approx 1.11$ at this angle, which gives 11% approximation error. Thus, Eq. (4.10.5) can be reduced to

$$4\pi^2 A^2 (f/f_s)^2 \approx x^2(n) - x(n - 1)x(n + 1). \quad (4.10.6)$$

Equation (4.10.6) holds provided that $f_s > 8f$. The basic definition of the discrete TK energy operator is given by

$$\Psi_d[x(n)] = x^2(n) - x(n - 1)x(n + 1). \quad (4.10.7)$$

According to Eq. (4.10.7), the TK energy operator is nonlinear. It is *nearly* instantaneous in that only three samples are required in the energy computation at each time instant. The operator also has the following properties [97]:

- $\Psi_d[\cdot]$ is independent of the initial phase ϕ_0 ;
- $\Psi_d[x(n)] = \Psi_d[x(-n)]$;
- $\Psi_d[x(n)]$ is robust even when the signal passes through zero, as no division operation is required;
- $\Psi_d[x(n)]$ is capable of responding rapidly (in two sampling instants) to changes in both A and f .

One can understand the behavior of the TK energy operator by analyzing its output in the frequency domain [98], showing that this operator belongs to the family of the quadratic operators (class of Volterra filters) defined by

$$y(n) = \sum_{l=0}^{L-1} x(n+l)x(n-l)h(l), \quad (4.10.8)$$

where $h(l)$ is an L -point set of quadratic filter coefficients. One can show that the impulse response of the TK energy operator is given by

$$h(l) = \begin{cases} 1 & \text{if } l = 0 \\ -1 & \text{if } l = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The frequency response of the filter is given by

$$H(f) = e^{j\pi f} \sqrt{2(1 - \cos(2\pi f))}. \quad (4.10.9)$$

Relation (Eq. 4.10.9) indicates that the TK energy operator behaves like a high-pass filter, making it efficient, for example, for spike detection [99]. However, due to its quadratic nature, the operator is not as simple as a usual linear high-pass filter.

In the continuous time domain, the TK energy operator is defined as

$$\Psi_c[x(t)] = \dot{x}^2(t) - x(t)\ddot{x}(t). \quad (4.10.10)$$

This continuous version can be defined as the *Lie bracket* $\mathbb{L}[., .]$:

$$\Psi_c[x(t)] = \mathbb{L}[x(t), \dot{x}(t)], \quad (4.10.11)$$

which is proportional to the difference between the fractional rates of change of the real signal $x(t)$ and its first derivative $\dot{x}(t)$ [98]. When applied to an AM-FM signal as defined in Section 1.4.3 [27]

$$x(t) = a(t) \cos \left(2\pi \int_0^t f_i(\tau) d\tau \right), \quad (4.10.12)$$

with time-varying amplitude envelope $a(t)$ and IF $f_i(t)$, the Ψ_c operator yields

$$\Psi_c[x(t)] \approx a^2(t) f_i^2(t), \quad (4.10.13)$$

assuming that $a(t)$ and $f_i(t)$ do not vary too fast or too much with respect to the average value of $f_i(t)$ [96]. Equation (4.10.13) shows that Ψ_c (like Ψ_d) depends upon the instantaneous amplitude and IF which are attributed to the physics of the signal generation. Given the simplicity of the TK energy operator (and its extended versions) and the broad applicability of the AM-FM model in signal processing and communication systems, this operator leads to the *energy separation algorithm* (ESA) defined by [100]

$$f_i(t) \approx \frac{1}{2\pi} \sqrt{\frac{\Psi_c[\dot{x}(t)]}{\Psi_c[x(t)]}}, \quad |a(t)| \approx \frac{\Psi_c[x(t)]}{\sqrt{\Psi_c[\dot{x}(t)]}}. \quad (4.10.14)$$

The ESA has an excellent time resolution and low complexity but its main disadvantage is a moderate sensitivity to noise. To have smoother estimates of noisy signal time-derivatives (Eq. 4.10.14), the input signal can be approximated with a smoothing spline [101,102], Gabor filter [101], or Savitzky-Golay (SG) filter [103]. However, the ESA cannot handle wideband and multicomponent signals. In order to take advantage of its capabilities, the input signal is filtered through, for example, a filter bank constructed by Gabor band-pass filtering. Another way to deal with such limitation is to use the empirical mode decomposition (EMD) as a band-pass filter [102]. EMD is a data-driven technique (Sections 4.12 and 16.5). Unlike the Gabor filter, EMD does not require *a priori* knowledge of the input filtering parameters or the number of narrowband components to be extracted. The conjunction of ESA

and EMD is an interference-free (t, f) representation designated as the Teager-Huang transform (THT), which is suited for multicomponent AM-FM signal analysis [104].

The TK energy operator, limited to second order, has been extended into higher-order differential energy operators (DEOs) [98]:

$$\Psi_k[x(t)] = x^{(1)}(t)x^{(k-1)}(t) - x^{(0)}(t)x^{(k)}(t), \quad (4.10.15)$$

where k is the order of the operator and $x^{(k)}$ stands for the k th time derivative of $x(t)$. The DEOs can be generalized as [105]

$$\Psi_{p,q,m,l}[x(t)] = x^{(p)}(t)x^{(q)}(t) - x^{(m)}(t)x^{(l)}(t), \quad (4.10.16)$$

with integers $p + q = m + l$, $(p, q) \neq (m, l)$. For derivative-order combinations $(p = 1, q = k - 1$, $m = 0, l = k)$ and $(p = 1, q = 1, m = 0, l = 2)$, the generalized operator is reduced to the DEOs [98] and to the TK energy operator [97], respectively. All these operators are useful for estimating the instantaneous energy in a signal and for its demodulation [98,105]. Compared to TK operator, these extensions are more robust to noise and subsampling. Finally, all these operators can be extended to the multidimensional case as k th-order energy tensors Ψ_H (H stands for higher) [106]:

$$\Psi_{H,k,p,m}[s(\mathbf{u})] = \frac{d^m s}{d\mathbf{u}^m} \otimes \left(\frac{d^l s}{d\mathbf{u}^l} \right)^T - \frac{d^p s}{d\mathbf{u}^p} \otimes \left(\frac{d^q s}{d\mathbf{u}^q} \right)^T, \quad (4.10.17)$$

with $k = p + q = m + l$, $(m, l) \neq (p, q)$. Vector $\mathbf{u} = (x_1 x_2 \dots x_n)$ is n -dimensional, $s(\mathbf{u})$ is a multidimensional signal, and \otimes means the Kronecker product. The operator defined in Eq. (4.10.17) is a generalization of the operators introduced in [96]. The second-order operator ($k = 2$) gives the directional TK energy operator, which extends the 1D classical TK energy operator using directional derivatives along any vector \mathbf{v} as follows:

$$\mathbf{v}^T \Psi_2[s(\mathbf{u})]^T \mathbf{v} = \left(\frac{\partial s}{\partial \mathbf{v}}(\mathbf{u}) \right)^2 - s(\mathbf{u}) \frac{\partial^2 s}{\partial \mathbf{v}^2}(\mathbf{u}). \quad (4.10.18)$$

4.10.3 CROSS TEAGER-KAISER ENERGY OPERATORS

The cross TK (CTK) operator provides a measure of interaction between two real signals $x(t)$ and $y(t)$ [97]:

$$\mathbb{L}[x(t), y(t)] = \dot{x}(t)y(t) - x(t)\dot{y}(t). \quad (4.10.19)$$

If $x(t)$ and $y(t)$ represent displacements in some motions, $\mathbb{L}[x(t), y(t)]$ or $\mathbb{L}[y(t), x(t)]$ has dimension of energy per unit mass [98]. This energy-like quantity can measure the output of $\mathbb{L}[x(t) + y(t), \dot{x}(t) + \dot{y}(t)]$ for the detection of transient signals, where $x(t)$ is an AM-FM and $y(t)$ is a transient [107]. TK and CTK energy operators can be extended to complex-valued signals, yielding a new operator denoted by Ψ_B and defined by [108]:

$$\begin{aligned} \Psi_B[x(t), y(t)] &= \frac{1}{2} (\dot{x}^*(t)\dot{y}(t) + \dot{x}(t)\dot{y}^*(t)) - \frac{1}{4} (x(t)\ddot{y}(t) + x^*(t)\ddot{y}(t) + y(t)\ddot{x}(t) + y^*(t)\ddot{x}(t)) \\ &= \frac{1}{4} (\mathbb{L}[x(t), \dot{y}^*(t)] + \mathbb{L}[x^*(t), \dot{y}(t)] + \mathbb{L}[y(t), \dot{x}^*(t)] + \mathbb{L}[y^*(t), \dot{x}(t)]), \end{aligned} \quad (4.10.20)$$

where $x(t) = x_r(t) + jx_i(t)$ and $y(t) = y_r(t) + jy_i(t)$. Based on the Lie bracket, Ψ_B can measure the instantaneous interaction between two complex-valued signals [107]. One can show that Ψ_B is a real function, as expected for an energy operator [108]:

$$\Psi_B(x(t), y(t)) = \Psi_B(x_r(t), y_r(t)) + \Psi_B(x_i(t), y_i(t)). \quad (4.10.21)$$

The output of Ψ_B of two complex-valued signals is the sum of CTK energies of their real and imaginary parts. In the real case if $x(t) = y(t)$, Ψ_B reduces to the energy operator defined in Eq. (4.10.19) [109]. For $x(t) = y(t)$, the notation $\Psi_B(x(t), y(t)) \equiv \Psi_B(x(t))$ is used. To discretize Eq. (4.10.21), time derivatives can be approximated by time differences. It is easy to see that for two-sample forward or backward differences the discrete version of Ψ_B , denoted by Ψ_{Bd} , is given by

$$\Psi_{Bd}(x_k(n), y_k(n)) = \left(x_k(n)y_k(n) - \frac{1}{2} [x_k(n)y_k(n-1) + y_k(n+1)x_k(n-1)] \right) / T_s^2, \quad (4.10.22)$$

where $k \in \{i, r\}$, “ r ” and “ i ” indicate the real and imaginary parts, respectively, and $x_k(n)$ and $y_k(n)$ are the discrete-time counterparts of $x(t)$ and $y(t)$, respectively; the instant t is replaced by nT_s where $n \in \mathbb{N}$ and T_s is the sampling period.

The operator Ψ_B has found applications in various domains such as gene time-series expression data-clustering [110], transient detection [107], time delay estimation [111], and phase retrieval in white-light scanning interferometry [112].

More properties of Ψ_B are detailed in [104, 107, 108]. An interesting property of Ψ_B is its relationship to instantaneous cross-correlation (CC) [108]:

$$\Psi_B[x(t), y(t)] = -\frac{\partial^2 K_{xy}(t, \tau)}{\partial \tau^2} \Big|_{\tau=0} - \frac{\partial^2 K_{xy}^*(t, \tau)}{\partial \tau^2} \Big|_{\tau=0}, \quad (4.10.23)$$

where

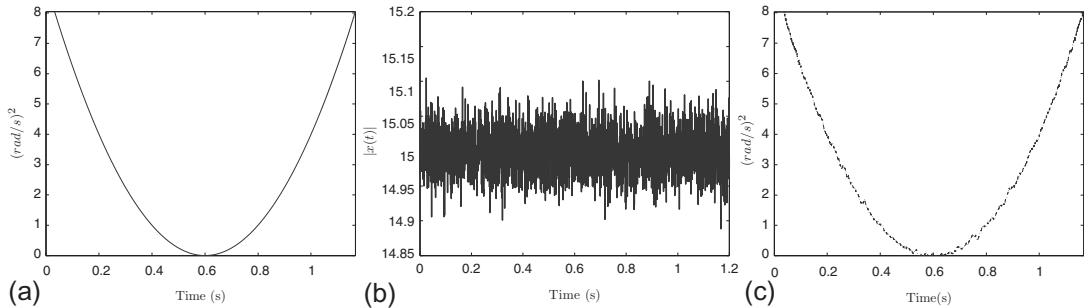
$$K_{xy}(t, \tau) = x\left(t + \frac{\tau}{2}\right) y^*\left(t - \frac{\tau}{2}\right). \quad (4.10.24)$$

Thus links to transforms using the concept of instantaneous CC, such as (t, f) distributions [27], can be established and show the interest of Ψ_B for analyzing nonstationary signals [104]. For example, it has been shown that [104]:

$$\Psi_B[x(t), y(t)] = 8\pi^2 \int f^2 W_{xy}(t, f) df \quad (4.10.25)$$

$$= \mathcal{F}^{-1} \left\{ -\frac{\partial^2}{\partial \tau^2} \left[A_{xy}^*(f, \tau) + A_{xy}(-f, \tau) \right]_{\tau=0} \right\}, \quad (4.10.26)$$

where \mathcal{F}^{-1} is the inverse FT, W_{xy} is the cross WVD, and A_{xy} is the cross AF. Useful relationships between Ψ_B and some (t, f) distributions are provided in [104]. One of the major advantages of this class of TK operators is that they account for the instantaneous behavior of the signals by using the derivatives relative to time [102].

**FIGURE 4.10.2**

Moment of LFM signal: (a) true $\langle f^2 \rangle_t$; (b) noisy signal; (c) $\langle f^2 \rangle_t$ estimated by Ψ_B operator.

Using the link to the cross WVD, for $x(t) = y(t)$, the resulting relation is a simple way to calculate the second-order moment in frequency, $\langle f^2 \rangle_t$, of the WVD:

$$\langle f^2 \rangle_t = \frac{\int f^2 W_{xx}(t,f) df}{\int W_{xx}(t,f) df} \quad (4.10.27)$$

$$= \frac{\Psi_B(x(t))}{8\pi^2 |x(t)|^2}. \quad (4.10.28)$$

This moment is a useful feature for signal classification [113]. For simulation purposes we generate a constant amplitude ($A = 15$) LFM signal of the form

$$x(t) = A e^{j2\pi\phi(t)} + n(t), \quad (4.10.29)$$

where $\phi(t) = \alpha t^2 + \beta t + \gamma$ with $\alpha = -2.5$, $\beta = 3$, and $\gamma = 20$, and $n(t)$ is a complex Gaussian white noise. The SNR is set to 30 dB, and the derivatives in the operator $\Psi_B(x(t))$ (Eq. 4.10.28) are calculated using a 5th-order SG filter and 127-point window.¹ Figure 4.10.2(c) shows that the moment $\langle f^2 \rangle_t$ is well extracted from the noisy signal (Fig. 4.10.2(b)) and agrees well with the true moment (Fig. 4.10.2(a)).

4.10.4 LINK OF Ψ_B TO DYNAMIC SIGNAL

The dynamic signal, introduced in [27], describes the rate of change of the log-magnitude and phase. It is similar to the complex cepstrum in the frequency domain and can be viewed as a generalization of the instantaneous phase. For the AM-FM signal $x(t) = a(t) e^{j\phi(t)}$, whose IF is $f_i(t) = \frac{1}{2\pi} \dot{\phi}(t)$, the dynamic signal is [27]

$$\beta(t) = \frac{d}{dt} \log x(t) = \frac{\dot{a}(t)}{a(t)} + j\dot{\phi}(t). \quad (4.10.30)$$

¹In this high SNR case, Eq. (4.10.22) can also be used. But for noisy signals and to avoid spikes (due to derivations) in the estimates, SG filter or B-splines smoothed functions are used [102].

It is easy to show that Ψ_B can be written in term of dynamic signal as follows [103]:

$$\Psi_B(x(t)) = -\frac{|x(t)|^2}{2} \left[(\beta^*(t) - \beta(t))^2 + (\dot{\beta}^*(t) + \dot{\beta}(t)) \right] \quad (4.10.31)$$

$$= |x(t)|^2 \left[8\pi^2 f_i^2(t) + B_i^2(t) - \frac{\ddot{a}(t)}{a(t)} \right], \quad (4.10.32)$$

where $B_i(t) = \left| \frac{\dot{a}(t)}{a(t)} \right|$ is the instantaneous bandwidth (IB). The above equations show that Ψ_B conveys precious local information about the instantaneous behavior over the time of $x(t)$ such as the IF and the IB. Thus for an FM signal with $a(t) = A$, the IF is retrieved without integral transforms, and with no *a priori* knowledge of the phase $\phi(t)$ of the signal, as follows:

$$f_i(t) = \frac{\sqrt{\Psi_B(x(t))}}{2\pi A \sqrt{2}}. \quad (4.10.33)$$

Let us consider a constant-amplitude ($A = -2.5$) polynomial-phase signal such as

$$x(t) = A \exp \left(j \sum_{k=0}^4 a_k t^k \right) + n(t), \quad (4.10.34)$$

where $a_0 = 1, a_1 = 10, a_2 = 10, a_3 = -3, a_4 = -2$, and $n(t)$ is a complex Gaussian white noise. The SNR is set to 25 dB and the operator $\Psi_B(x(t))$ in Eq. (4.10.33) is implemented using a 5th-order SG filter and 69-point window. The IF law calculated by the Ψ_B operator (Fig. 4.10.3(c)) is very close to the true law (Fig. 4.10.3(a)) with little oscillations attributed to noise (Fig. 4.10.3(b)). Note that the computation of the IF (Eq. 4.10.33) and the second moment (Eq. 4.10.28) by the Ψ_B operator has very low cost compared to that of the WVD.

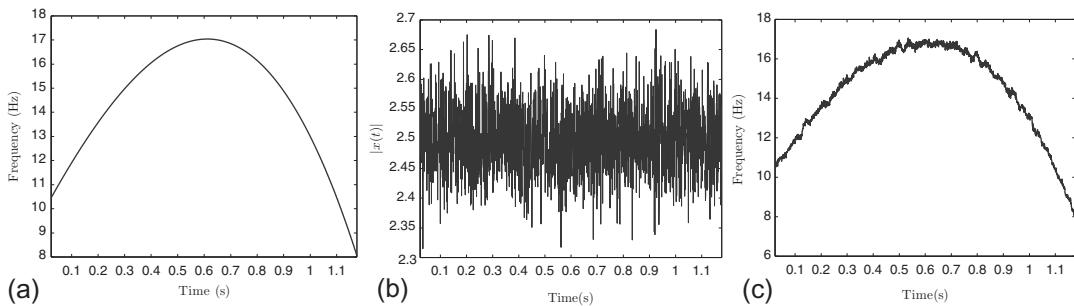


FIGURE 4.10.3

IF estimate for 4th-order polynomial-phase signal: (a) true IF; (b) noisy signal; and (c) IF estimated by Ψ_B operator. (For complex FM signals the method described is more efficient and more robust than the ESA (Eq. 4.10.14), which is based on the TK operator. The IF estimation method using the ESA involves the division of the first derivative of $x(t)$ by the output of the TK operator which can introduce instabilities; note, however, that the method described is also sensitive to noise as shown by the oscillations that appear in graph (c).)

4.10.5 SUMMARY AND CONCLUSIONS

This overview of the TK energy operator shows the importance of the concept of energy from the point of view of the generation process of the signal. Understanding the genesis of this operator from this perspective is key to explaining its behavior and its potential as a signal-analysis tool. One of the advantages of the class of TK operators is that they account for the instantaneous behavior of the signals using the derivatives w.r.t. time.² The output of these operators depends on the frequency and the amplitude, which are related to the physics of signal generation. An important application of the TK operator is the ESA used for AM-FM demodulation. The TK operator has been extended to deal with multidimensional signals and complex-valued signals. This complex version is related to the instantaneous correlation. Thus, links to transforms using the concept of the instantaneous correlation, such as (t,f) distributions (WVD, AF, etc.) are established, showing the application of this extension to nonstationary signal analysis.³ For example, the link to the WVD allows a simple calculation of the corresponding second moment in frequency, which is a useful feature for signal classification [113]. The relationship to dynamic signals shows that the complex version conveys precise local information, such as IF and IB, about the behavior over time of an AM-FM signal (see Eq. (4.10.33)). For an FM signal, this relationship provides a simple way to calculate the IF at high SNR without using integral transforms and with no *a priori* knowledge about the phase of the signal.

4.11 GABOR SPECTROGRAM, WVD, AND ENERGY ATOMS⁰

4.11.1 WVD PROPERTIES AND GAUSSIAN FUNCTIONS

As several approaches to time-frequency (t,f) analysis seem quite different, it is important to show the relationship between them, as discussed in [Chapters 2](#) and [3](#). This section extends previous discussions and presents more relationships with focus on the Gabor spectrogram, its properties and the links with other methods.

For an analytic signal $s(t) = a(t) \cos(\phi(t))$ with FT $S(f)$, a desirable TFD $\rho_s(t,f)$ should possess the following three properties ([Section 1.3](#)):

$$\int_{-\infty}^{+\infty} \rho_s(t,f) dt = |S(f)|^2; \quad \int_{-\infty}^{+\infty} \rho_s(t,f) df = |s(t)|^2 = |a(t)|^2, \quad (4.11.1)$$

$$\frac{\int_{-\infty}^{+\infty} f \rho_s(t,f) df}{\int_{-\infty}^{+\infty} \rho_s(t,f) df} = \frac{\int_{-\infty}^{+\infty} f \rho_s(t,f) df}{|a(t)|^2} = f_i(t), \quad (4.11.2)$$

²The effectiveness of the approaches based on the TK operators and their extensions have been shown on demodulation by tracking instantaneous amplitude and frequency for different dimensions (1D [102,103], 2D [112], and nD [106]); and the results were compared to well-established methods, such as HT and WVD. Also, in terms of signal detection the Ψ_B operator has been compared successfully with matched filter in the time domain [107,111]. This operator was shown to be efficient for the analysis of the LFM signals in the (t,f) plane and the results compared with WVD and pseudo-WVD [114].

³The authors have established links to quadratic TFDs that satisfy the spectral delay and IF conditions. One of the relationships links the quadratic TFDs via the generalized cross-AF and the kernel function. Another result shows that the calculated second frequency moment of the quadratic TFDs, using Ψ_B operator, is independent of the kernel used.

⁰Second edition updated and extended by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa), **Md A. Awal**, University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (m.awal@uq.edu.au), and **S. Dong**, University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (s.dong@uq.edu.au). First edition [115] by S. Qian. Reviewers: S. Ouelha and N.A. Khan.

where $f_i(t)$ is the IF discussed in Section (1.3). The properties (Eqs. (4.11.1) and (4.11.2)) relate to the average of $\rho_s(t,f)$.

The spectrogram based on the STFT cannot simultaneously meet properties (Eqs. (4.11.1) and (4.11.2)); the WVD, defined by Eq. (2.1.17), satisfies them, but its application is limited due to the cross-term interference. One effective approach to reduce the cross-term interference is to apply a 2D low-pass filter to the WVD resulting in some specific TFDs of the quadratic class defined in Chapter 3 without destroying the desirable properties possessed by the WVD (see Chapter 3, [116], and Section 4.2).

An alternative is to first write the signal in terms of the Gabor expansion and then take the WVD which becomes then a superposition of concentrated, symmetrical, and oscillating 2D Gaussians. The energy (computed by averaging a 2D Gaussian) contained in each individual 2D Gaussian is inversely proportional to the rate of its oscillation. The higher the oscillation, the smaller the energy (or average). Since the desirable properties are all computed by averaging $\rho_s(t,f)$, the highly oscillated Gaussians have limited influence on the properties (Eqs. (4.11.1) and (4.11.2)), but instead they have significant contribution to the cross-term interference. Consequently, we can remove the highly oscillated terms, without significantly altering desirable properties, to lessen the cross-term interference. The resulting representation is called the Gabor spectrogram to distinguish it from the spectrogram computed from the STFT [116]. See Section 2.3.1 for a presentation of the STFT.

4.11.2 GABOR DECOMPOSITION OF THE WVD

When $s(t)$ is a time- and frequency-shifted Gaussian function $h_{m,n}(t)$, that is,

$$s(t) = h_{m,n}(t) = (\alpha\pi)^{-0.25} e^{\left\{-\frac{(t-mT)^2}{2\alpha} + j2\pi n\Omega t\right\}} \quad (4.11.3)$$

the corresponding WVD of this monotonic signal has the form

$$W_h(t,f) = 2 e^{\left\{-\frac{(t-mT)^2}{\alpha} - \alpha(2\pi)^2(f-n\Omega)^2\right\}}, \quad (4.11.4)$$

which is non-negative and symmetric about the point $(mT, n\Omega)$. The contour plot of Eq. (4.11.4) forms conic ellipses. Because Eq. (4.11.4) is a WVD, it satisfies properties (Eqs. (4.11.1) and (4.11.2)). When $s(t)$ contains multiple monotonic signals, such as multiple Gaussians with different time- and frequency-shifts, unexpected cross-term interferences appear; the resulting WVD still possesses the properties (Eqs. (4.11.1) and (4.11.2)), but the unwanted cross-terms can destroy the features of interest. Since cross-terms always occur midway between two monotonic signals, one can reduce the cross-terms by decomposing the analyzed signal as monotonic components.

The Gabor expansion decomposes a signal $s(t)$ as the sum of monotonic time- and frequency-shifted Gaussian elementary functions $h_{m,n}(t)$ in Eq. (4.11.3), that is,

$$s(t) = (\alpha\pi)^{-0.25} \sum_m \sum_n c_{m,n} h_{m,n}(t), \quad (4.11.5)$$

where $c_{m,n}$ is called the Gabor coefficient (see Section 2.3.3).

Taking the WVD of both sides of Eq. (4.11.5) yields

$$W_s(t, f) = \sum_{m,m'} \sum_{n,n'} c_{m,n} c_{m',n'}^* W_{h,h'}(t, f). \quad (4.11.6)$$

The function $W_{h,h'}(t, f)$ represents the cross-WVD of elementary functions $h_{m,n}(t)$ and $h_{m',n'}(t)$. It is a 2D Gaussian function,

$$\begin{aligned} W_{h,h'}(t, f) = & 2 e^{\{-\alpha^{-1}(t - \frac{m+m'}{2}T)^2 - \alpha(2\pi)^2(f - \frac{n+n'}{2}\Omega)^2\}} \\ & \times e^{\{j2\pi[(n-n')\Omega t - (m-m')T(f - \frac{n+n'}{2}\Omega)]\}}. \end{aligned} \quad (4.11.7)$$

It oscillates with symmetrical amplitude concentrated about the point

$$\left(\frac{m+m'}{2}T, \frac{n+n'}{2}\Omega \right).$$

While the rate of oscillation in the frequency domain is determined by $(m - m')T$, the rate of oscillation in the time domain is dependent on $(n - n')\Omega$. Therefore, the further $h_{m,n}(t)$ and $h_{m',n'}(t)$ are apart, the higher the oscillation. If the WVD is thought of as a signal energy distribution, then Eq. (4.11.7) can be considered as an *energy atom*. Equation (4.11.6) indicates then that the WVD is a linear combination of energy atoms. Finally, $W_{h,h'}(t, f)$ and $W_{h',h}(t, f)$ form a pair of complex conjugates, that is, $W_{h,h'}(t, f) + W_{h',h}(t, f) = 2 \operatorname{Re}\{W_{h,h'}(t, f)\}$.

4.11.3 GABOR SPECTROGRAM

The desired properties are related to the average of energy atoms, and, the average is inversely proportional to the rate of oscillation; as a consequence, the slowly oscillating energy atoms contribute to the properties more than the fast-oscillating energy atoms. Moreover, the rate of oscillation is completely determined by the distance between two elementary functions. Hence, the energy atom whose corresponding elementary functions $h_{m,n}(t)$ are close to each other dominates the desirable properties. The term whose elementary functions are further apart causes cross-term interference, but has limited contribution to the useful features.

Based on these observations, let us now re-group Eq. (4.11.7) into a new representation, named as the Gabor spectrogram, that is,

$$\begin{aligned} \text{GS}_D(t, f) = & \sum_{|m-m'|+|n-n'| \leq D} c_{m,n} c_{m',n'}^* 2 e^{\{-\alpha^{-1}(t - \frac{m+m'}{2}T)^2 - \alpha(2\pi)^2(f - \frac{n+n'}{2}\Omega)^2\}} \\ & e^{j2\pi[(n-n')Tt - (m-m')T(\omega - \frac{n+n'}{2}\Omega)]}, \end{aligned} \quad (4.11.8)$$

where the parameter D denotes the order of the Gabor spectrogram [117]. When $D = 0$, Eq. (4.11.8) becomes non-negative and has no cross-term interference like the spectrogram; but, it does not have the desired properties (Eq. (4.11.1) or (4.11.2)) either. As the order D increases, however, the Gabor spectrogram becomes similar to the WVD and gets closer to meeting these properties. Consequently, adjusting the order of the Gabor spectrogram results in a good compromise between cross-term interference and the desired properties. For example, when $D = 4$, the error between the Gabor

spectrogram-based results of Eqs. (4.11.1) and (4.11.2) and the expected values are usually less than 1%, whereas the cross-terms are negligible in most applications.

4.11.4 TESTS ON SIMULATED AND REAL SIGNALS

We first tested the performance of the Gabor spectrogram using a simulated signal composed of two close LFM components in comparison with three other TFDs (i.e., spectrogram, WVD, and EMBD; see Chapters 2 and 3 for details of these TFDs). The results are plotted in Fig. 4.11.1 where the Gabor spectrogram shows better (t,f) resolution and cross-term-free signature while the WVD and EMBD

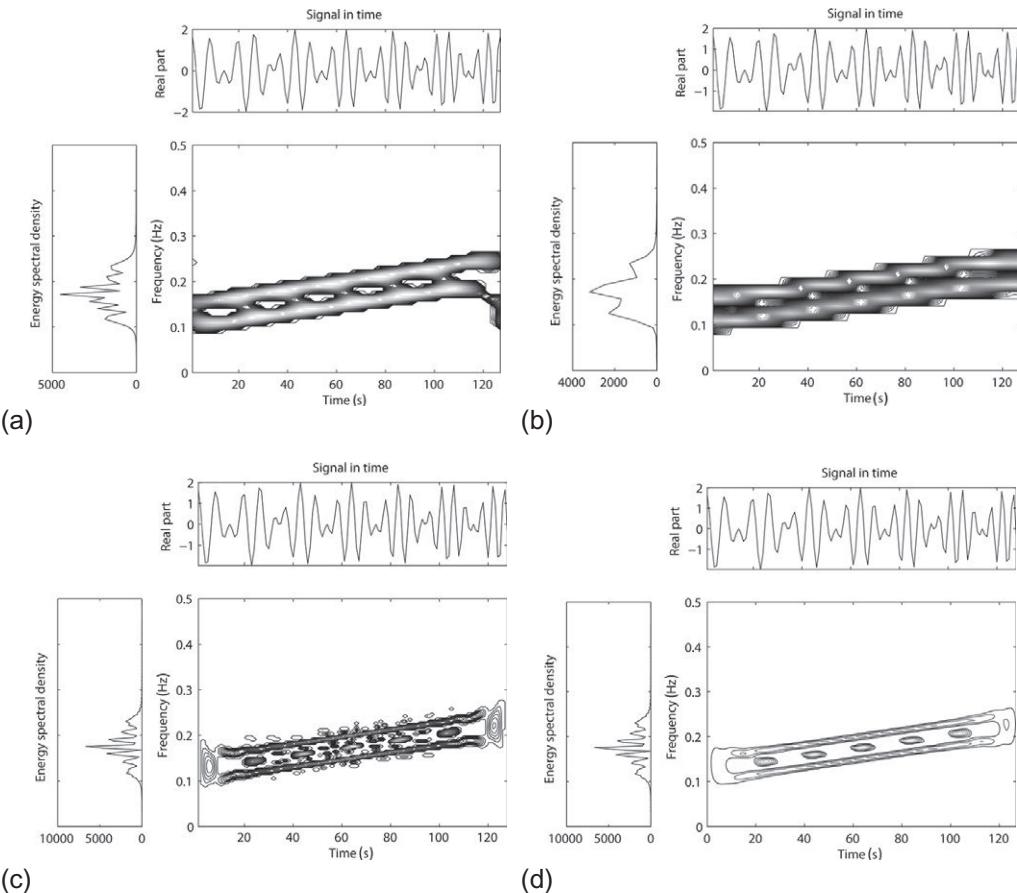


FIGURE 4.11.1

(a) Gabor spectrogram, (b) spectrogram (Hanning window, length 101), (c) WVD, (d) EMBD ($\alpha = 0.3$, $\beta = 0.85$) of the sum of two-closed LFM signals with frequency ranges 0.1-0.2 and 0.15-0.25 Hz, and duration of 128 samples (sampling rate 1 Hz).

render sharper characterization but suffer from cross-terms at different degrees. A bat sound is also used to test the Gabor spectrogram. Figure 4.11.2 shows the fourth-order Gabor spectrogram of the bat sound. The zeroth- and second-order Gabor spectrograms are plotted in Fig. 4.11.3, showing that the (t, f) resolution improves as the order increases. The Gabor spectrogram eventually converges to the WVD as the order increases. Comparatively, the STFT-based spectrogram in Fig. 4.11.4 has poor (t, f) resolution but no cross-term interference. The EMBD provides a good tradeoff between the two constraints of improved resolution and cross-terms reduction.

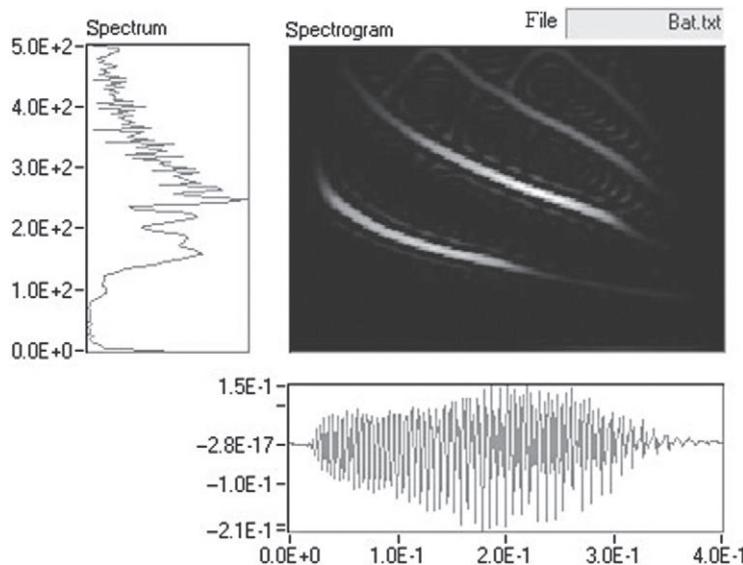


FIGURE 4.11.2

Fourth-order Gabor spectrogram of bat sound.

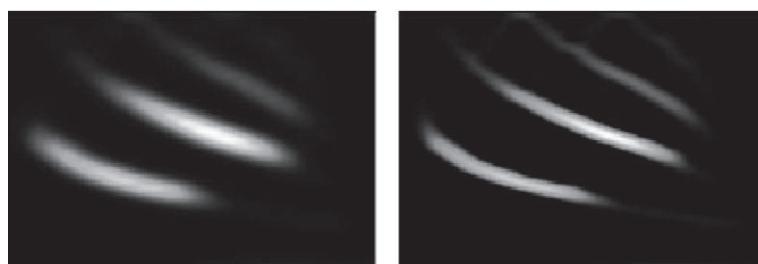


FIGURE 4.11.3

Zeroth- (left) and second (right)-order Gabor spectrogram of bat sound.

4.11.5 RELATIONSHIP BETWEEN GABOR SPECTROGRAM AND MATCHING PURSUIT

The Gabor expansion-based spectrogram and the matching pursuit (MP)-based spectrogram form a class of linear (t, f) methods as they decompose a signal by a linear combination of a family of elementary functions that are localized both in time and frequency (see Section 2.7.7). Like Eq. (4.11.3) for the Gabor expansion of a signal, the MP decomposes a signal as a weighted superposition of elementary functions, that is (see Sections 12.2, 7.6, and 6.3) [118]:

$$s(t) = \sum_k A_k h_k(t), \quad (4.11.9)$$

where $h_k(t)$ denotes the elementary function and A_k denotes weight or expansion coefficient similar to $c_{m,n}$ in Gabor expansion. If a set of elementary functions $\{h_k(t)\}_{k \in \mathbb{Z}}$ (called “dictionary”) forms an orthogonal basis, A_k can be easily found by regular inner product, that is $A_k = \langle s, h_k \rangle$. The MP satisfies energy conservation, that is, $\|s(t)\|^2 = \sum_{k=0}^{\infty} A_k^2$, because $\{h_k(t)\}_{k \in \mathbb{Z}}$ is set to have unit energy ($\|h_k(t)\|^2 = 1$). Like Eq. (4.11.6) where the WVD is applied to Gabor expansion, the adaptive spectrogram is obtained by applying the WVD to both sides of Eq. (4.11.9) as:

$$W_s(t, f) = \sum_k A_k^2 W_{h_k}(t, f) + \sum_{k \neq q} A_k A_q^* W_{h_k h_q}(t, f). \quad (4.11.10)$$

The first term represents the auto terms whereas the second one represents the cross-terms. Since $h_k(t)$ has unit energy and the WVD is energy conserving, it can be proved that the second term equals zero (see Section 4.11.8). As a result, a new time-varying and cross-term-free spectrum, known as adaptive spectrogram (A-Spec), can be defined as [116]:

$$\text{A-Spec}(t, f) = \sum_k |A_k|^2 W_{h_k}(t, f). \quad (4.11.11)$$

In each iteration of calculation, only one elementary function (atom) is selected to compose the signal using the MP algorithm and the WVD of this atom is then multiplied by $|A_k|^2$. The weighted WVDs are superimposed to obtain the A-Spec. As only one elementary function is used at a time, the TFD signature of A-Spec becomes free of cross-terms. In contrast, the Gabor spectrogram enhances

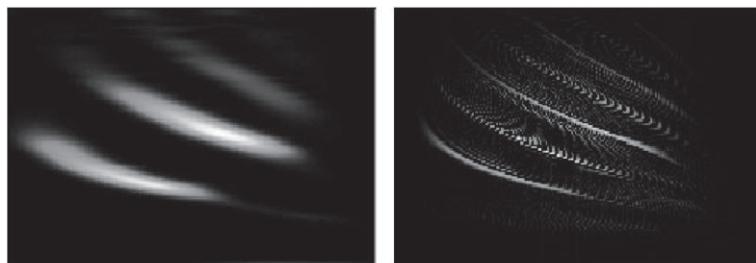
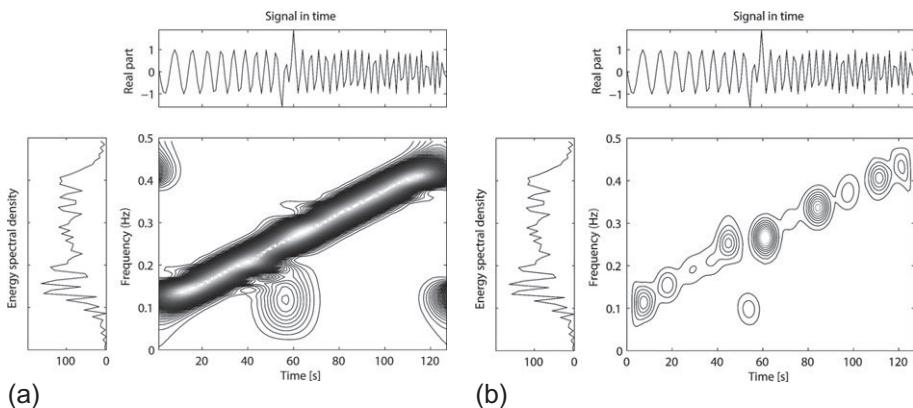


FIGURE 4.11.4

STFT-based spectrogram (left) and WVD of bat sound.

**FIGURE 4.11.5**

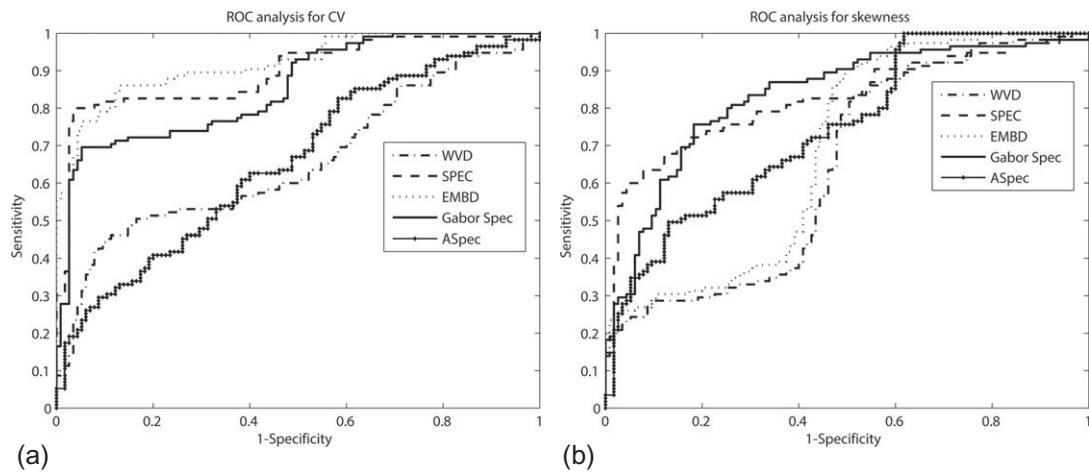
(a) Gabor spectrogram, (b) adaptive spectrogram of a signal with both short- and long-duration components (128 samples and sampling rate 1 Hz). The short-duration component appears between 55 and 60 samples whereas the long-duration component is in the entire duration.

the resolution at the expense of increased cross-terms by increasing D whereas the A-Spec produces no cross-terms but decreased resolution when $D = 0$ (see [Section 4.11.3](#)).

To sum up, both Gabor expansion and MP are linear (t,f) decompositions. The WVD of both decompositions produces an alternative estimate of the spectrogram, that is, Gabor spectrogram and adaptive spectrogram which provide high-resolution (t,f) representations. The MP can achieve higher resolution by using a large dictionary so that there is a high probability of finding an element from the dictionary whose shape closely matches that of the signal under analysis. On the other hand, the Gabor Spectrogram only uses Gaussian atoms as dictionary elements. The MP reduces to the Gabor spectrogram for order $D = 0$ if Gaussian atoms are used in the dictionary. The Gabor spectrogram has a limited performance for signals where a short- and a long-duration components are present together. On the contrary, the A-Spec can provide a better (t,f) representation for this type of signals (see [Fig. 4.11.5](#)). A disadvantage is that the computation load for the A-Spec is relatively high due to the excessive number of elementary functions in the dictionary; the higher the accuracy is, the more the number of elementary functions is needed.

4.11.6 EEG BURST-SUPPRESSION DETECTION USING GABOR SPECTROGRAM

Burst-suppression (B-S) is an abnormal pattern observed in newborn EEG associated with adverse neurodevelopmental outcome in the case of hypoxic-ischemic brain injury [[119](#)]. Accurate detection of this pattern is important for accurate prognosis and effective treatment. Given the nonstationary nature of EEG signals, the Gabor spectrogram offers a useful tool for characterizing the signals with B-S patterns precisely. The experiment used 115 epochs of burst and 115 epochs of suppression patterns which were extracted from artifact-free multichannel EEG signals. The EEG data were recorded from newborns admitted to the Royal Brisbane and Women's Hospital, Brisbane, Australia using a MEDELEC Profile System (Medelec, Oxford Instruments, UK). The sampling frequency was

**FIGURE 4.11.6**

The ROC plots of CV and skewness features based on Gabor spectrogram, A-Spec, Spectrogram, WVD, and EMBD for newborn EEG B-S detection. The AUC values for the features (a) CV 0.85, 0.67, 0.90, 0.66, 0.92; (b) skewness 0.83, 0.75, 0.82, 0.64, 0.69; were extracted from Gabor spectrogram, A-Spec, Spectrogram, WVD, and EMBD, respectively.

256 Hz. Multichannel signals were averaged into one channel to reduce the computational complexity. The signals are represented using Gabor spectrogram, A-Spec, Spectrogram, WVD, and EMBD, respectively. Features such as coefficient of variation (CV) and skewness were extracted from both representations [120]. The CV of TFDs accounts for the variability of burst and suppression patterns whereas the skewness of TFD accounts for the degree of deviation from the symmetry when EEG waveforms are asymmetrical with respect to the baseline. For each feature, the receiver operating characteristic (ROC) plot was produced as the sensitivity versus (1-specificity) and then the area under ROC curve (AUC) was calculated. The AUC value (ranging from 0.5 to 1)¹ is used as a measure of how well a feature (or classifier) can discriminate between two classes (i.e., burst and suppression patterns in this application). A classifier with an AUC value of 1 is regarded as perfect, and an AUC value of 0.5 corresponds to a random-guessing classifier. It can be seen from the ROC plots (Fig. 4.11.6) that Gabor spectrogram-based features exhibit good AUC values overall, suggesting a strong differentiating power between burst and suppression patterns.

4.11.7 SUMMARY AND CONCLUSIONS

The Gabor spectrogram can balance the resolution and cross-term interference by altering the order D . It is closer to the STFT-based spectrogram for lower orders, but converges to the WVD as the order approaches infinity. The advantage of the Gabor spectrogram is that by altering the order D ,

¹The AUC value is between 0 and 1 but when $AUC < 0.5$, we normally use $1-AUC$, that is, the significant value is between 0.5 and 1.

one can balance the resolution and cross-term interference. The adaptive spectrogram is implemented using the MP algorithm by searching suitable elementary functions (atoms) from an overcomplete dictionary. The advantage of adaptive spectrogram is that it contains no cross-terms. The disadvantage is increased computational cost. Both the Gabor spectrogram and adaptive spectrogram are based on linear decomposition of signals, where Gabor spectrogram is the WVD of the Gabor expansion while adaptive spectrogram is the WVD of the atomic expansion of a signal. The adaptive spectrogram reduces to the Gabor spectrogram of order $D = 0$ when a dictionary of Gaussian atoms is used. [Section 6.3](#) further explores these notions.

4.11.8 PROOF OF EQ. (4.11.11)

Since $h_k(t)$ has unit energy and the WVD is energy conserving, that is, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{h_k}(t,f) dt df = 1$, therefore, we can write [116]:

$$\|s(t)\|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_s(t,f) dt df = \sum_k |A_k|^2 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{k \neq q} A_k A_q^* W_{h_k h_q}(t,f) dt df.$$

The energy conservation of MP (i.e., $\|s(t)\|^2 = \sum_{k=0}^{\infty} |A_k|^2$) yields $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{k \neq q} A_k A_q^* W_{h_k h_q}(t,f) dt df = 0$; So, the cross-terms in Eq. (4.11.10) contain zero energy. This defines a time-dependent spectrum as $A\text{-Spec}(t,f) = \sum_k |A_k|^2 W_{h_k}(t,f)$.

4.12 EMPERICAL MODE DECOMPOSITION AND HILBERT SPECTRUM⁰

4.12.1 EMPIRICAL MODE DECOMPOSITION

The EMD, as detailed in [Section 16.5](#), is an empirically adaptive signal decomposition technique. It decomposes nonlinear and nonstationary signals into a number of natural “intrinsic” functions that lead to a well-behaved Hilbert transforms (HTs) from which the IFs can be calculated. The method is data-driven [121].

The principle of decomposition assumes that any signal consists of different intrinsic modes of oscillation. Each mode represents an oscillation that (1) has the same number of extrema and zero-crossings and (2) is symmetric with respect to the local mean. Each oscillatory mode is represented by an intrinsic mode function (IMF). The first condition is related to the classical narrowband requirement while the second one is a modification of the first requirement from global to a local one to deal with the characteristics of nonlinear and nonstationary signals [122]. Strictly speaking, the second requirement forces the IMFs to have a zero local mean.

These requirements on the IMFs are linked to the requirements of having a meaningful interpretation of the IF using the HT. This is in turn relates to Bedrosian’s theorem [121], which states that if $A(t)$ is a low-pass signal and $\Phi(t)$ is a high-pass signal (see [Section 1.2.3](#)), then

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **A. Akan**, Istanbul University, Avcilar, Istanbul, Turkey (akan@istanbul.edu.tr). Reviewers: A. Mert and N. Stevenson.

$$\mathcal{H}(A(t)\Phi(t)) = A(t)\mathcal{H}(\Phi(t)), \quad (4.12.1)$$

where \mathcal{H} is the Hilbert operator. Using a sifting process, the EMD decomposes a given signal $x(t)$ as follows:

$$x(t) = \sum_{m=1}^M \text{IMF}_m(t) + r_M(t), \quad (4.12.2)$$

where IMF_m are IMF functions, as defined above, and r_M is a residue. Each sifting iteration consists of several steps leading to the extraction of one IMF. Starting with the signal $x(t)$ to be decomposed (Fig. 4.12.1), at each step the mean of upper and lower envelopes, obtained by cubic spline interpolations of the local maxima (Fig. 4.12.2(a)) and local minima (Fig. 4.12.2(b)), respectively, is subtracted from the current signal until obtaining an IMF. Then a new sifting process is repeated and so on until obtaining Eq. (4.12.2). The sifting process will stop when the numbers of zero crossings and extrema stay the same and are equal or differ at most by one. The decomposition process can be

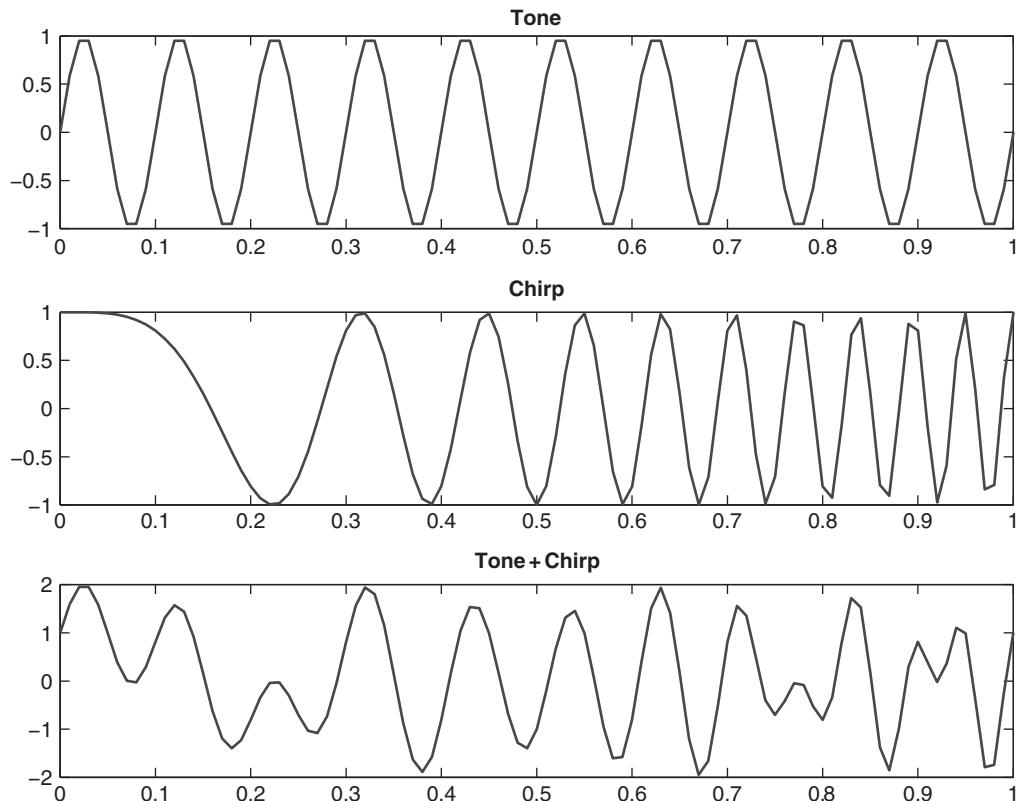
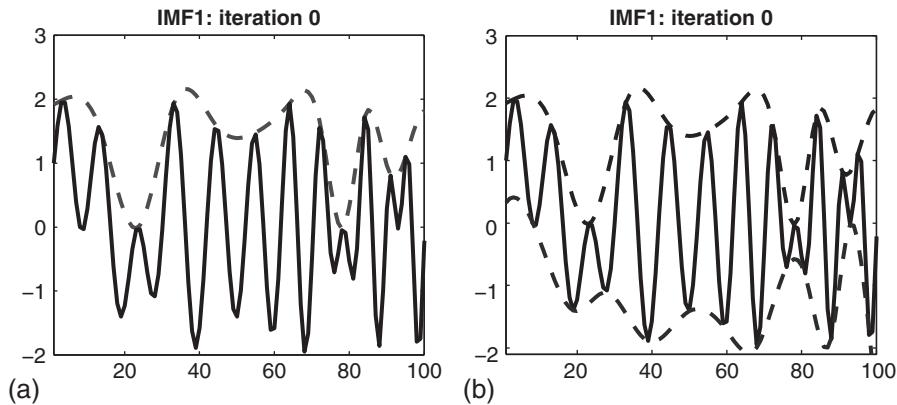


FIGURE 4.12.1

Signal to be decomposed using EMD.

**FIGURE 4.12.2**

(a) Linking local maxima using cubic spline to create the upper envelope; and (b) linking local minima using cubic spline to create the lower envelope.

stopped if: (1) the $\text{IMF}_M(t)$ or the residue $r_M(t)$ becomes smaller than a predefined threshold, (2) the residue $r_M(t)$ becomes a monotonic function (trend); even for data with zero mean, the final residue still can be different from zero (see Figs. 4.12.3 and 4.12.4 for an illustration of the process). The EMD algorithm lacks a firm theoretical foundation and it is not uniquely defined since it depends on user-controlled parameters such as the type of interpolation for envelope construction and the stopping criteria (for sifting and decomposition). As a consequence, the best way to analyze it is through numerical simulations [123].

4.12.2 HILBERT SPECTRAL ANALYSIS

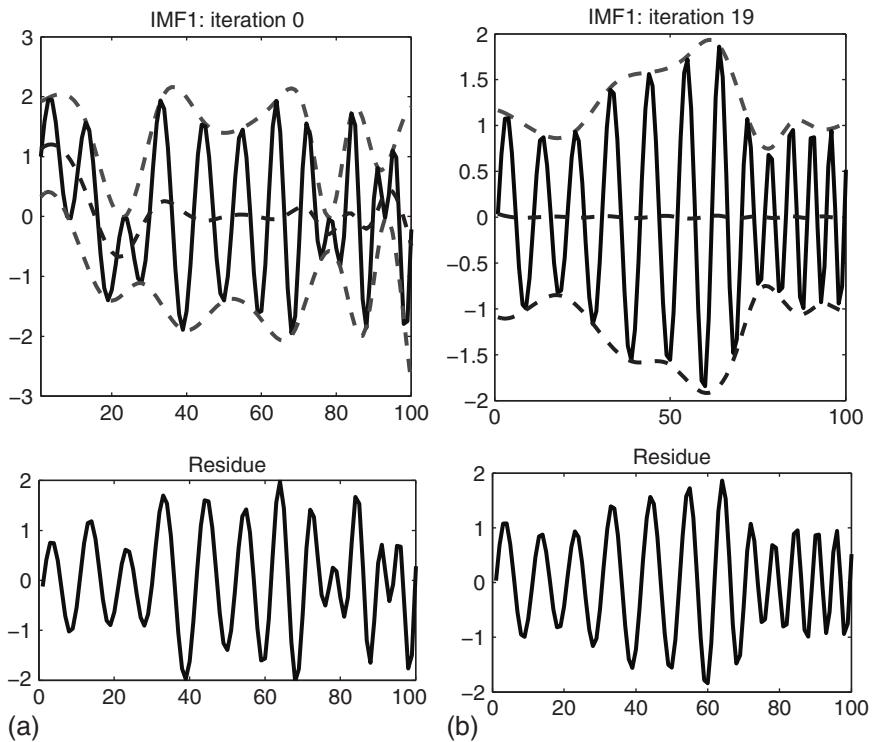
As defined in Chapter 1, the analytic signal associated with $x(t)$ is given by¹

$$z(t) = \sum_{m=1}^M a_m(t) \exp\left(j2\pi \int_{-\infty}^t f_m(\theta) d\theta\right), \quad (4.12.3)$$

where $a_m(t)$ and $f_m(t)$ are, respectively, the instantaneous amplitude and IF of the m th IMF. This expression can be interpreted as a generalization of the Fourier expansion with variable (instantaneous) amplitudes and variable (instantaneous) frequencies. The (t,f) representation thus obtained is based in a_m and f_m is called the Hilbert spectrum $H_z(t,f)$. It may be also expressed as²:

¹The residue has been removed as it is either monotonic or constant. The energy involved in the residual trend can be large compared with the IMFs. This may create an issue in the spectral representation as spectral information is omitted in the residue.

²Note that, from Eq. (4.12.3), the analytic signal $z(t)$ can be viewed in the (t,f) plane as a sum of Dirac functions with location $(t, f_m(t))$ and amplitude $a_m(t)$, which corresponds to Eq. (4.12.4). This (t,f) distribution of the amplitude, presented in a 3D plot, defines the Hilbert amplitude spectrum or simply the Hilbert spectrum. Therefore, the proof of Eq. (4.12.4) follows directly as it is the definition of the Hilbert spectrum itself derived from Eq. (4.12.3), which is obtained by applying Hilbert operator on the underlying multicomponent signal (sum of IMFs).

**FIGURE 4.12.3**

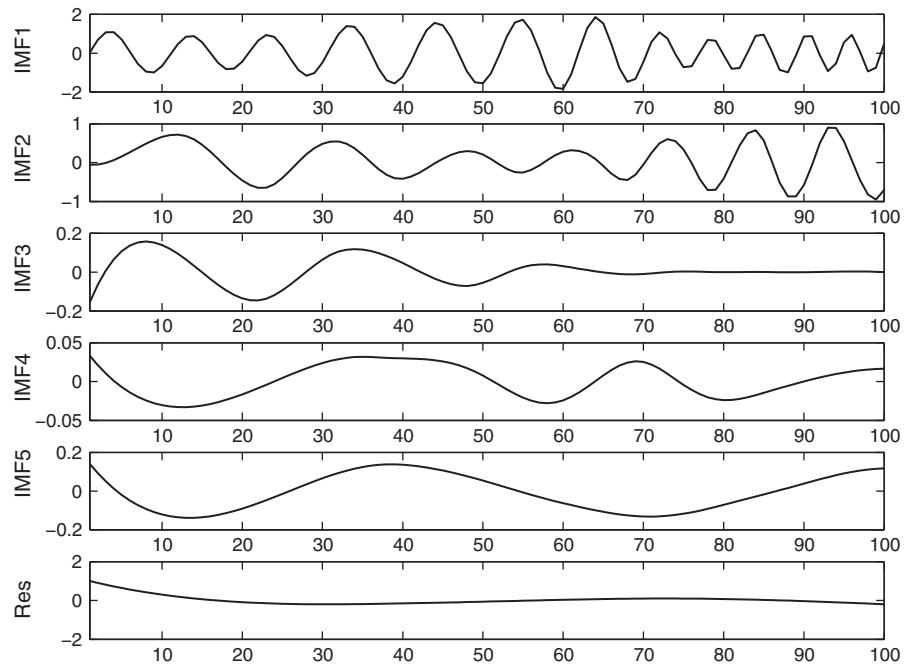
(a) Computing the local mean and the residue; and (b) the last iteration for the IMF1 and the residue.

$$H_z(t, f) = \sum_{m=1}^M a_m(t) \delta(t, f - f_m(t)). \quad (4.12.4)$$

To represent the energy density, the squared values of the amplitude can be substituted to produce the Hilbert energy spectrum. The representation can also be enhanced from $H_z(t, f)$ by using a 2D (t, f) smoothing [122]. The frequency in $H_z(t, f)$ has a different interpretation from stationary spectral (Fourier) analysis. Here, the existence of “energy” at the frequency f means only that, in the whole time span of the data, there is a higher likelihood for such a wave to appear at time t .

Thus, as the Hilbert spectrum is not a quadratic transform, it cannot be an energy distribution. A way to overcome this problem is to map the IMFs using the WVD and add these WVDs to obtain a (t, f) representation as follows:

$$\check{W}_z(t, f) = \sum_{m=1}^M W_{\text{IMF}_m}(t, f). \quad (4.12.5)$$

**FIGURE 4.12.4**

The full set of IMFs with the residue.

However, $\check{W}_z(t,f)$ does not equal $W_z(t,f)$ because the cross-terms in $W_z(t,f)$ do not appear in $\check{W}_z(t,f)$. Also, as the IFs $f_m(t)$ may be nonlinear, other reduced interference quadratic TFDs may be preferred to cancel out the inner cross-terms (see also [124]).

4.12.3 RELATIONSHIP BETWEEN THE EMD AND QTFDs

Let us consider the model of additive multicomponent signals given by [125]:

$$s(t) = \sum_{k=1}^N s_k(t) + n(t), \quad (4.12.6)$$

where $n(t)$ is a random process and $s_k(t)$ are time-varying monocomponent signals described by the amplitude envelope $a_k(t)$ and the IF $f_k(t)$; that is

$$s_k(t) = a_k(t) \exp \left(j2\pi \int_{-\infty}^t f(\theta) d\theta \right). \quad (4.12.7)$$

Let us compare Eqs. (4.12.6) and (4.12.2). The main similarity between the EMD and Eq. (4.12.6) is that both are signal decompositions whose aims are to have subcomponents (monocomponents for the

case of [125], and IMFs for [122]) that lead to physically meaningful IFs. The two expressions differ in the following ways:

1. How are these subcomponents obtained? In Boashash's model [125], these components are identified in the (t,f) domain from their energy concentrations. Relevant (t,f) masks can be applied around these TF components to isolate them and then a synthesis process can recover these time domain components. An automated process can estimate the multicomponent IFs prior to using (t,f) masks [126]. An alternative is to use blind source separation methods to automatically separate the monocomponents. In the EMD model [122], the components are obtained using the sifting process, described earlier, directly on the original signal. The aim of the sifting process is to obtain narrow pass-band components. The process is fully automated subject to prior predefined thresholds and parameters related to the decomposing algorithm.
2. In terms of the extra term: in [125], the extra term is the noise term, which is defined as a random noise process. In [122], the extra term is the residue, which can be a signal mean or a trend.

4.12.4 RELATIONSHIP BETWEEN EMD AND OTHER (t,f) /TIME-SCALE REPRESENTATIONS

The EMD has a major limitation, in that it lacks a sound theoretical foundation [121,123]. Performing a careful and detailed analysis of the EMD seems to be only possible through extensive numerical simulations [123]. This situation makes comparing the EMD, or even Hilbert spectrum (t,f) of IMFs representation, to other (t,f) representations difficult given the absence of clear criteria.

In addition, the EMD has lower resolution as compared to most (t,f) methods [127]. Furthermore, the EMD cannot take into account short-duration signals. It assumes that all signal components have same starting and ending times.

The EMD can be interpreted as a dyadic filter bank. This indicates a link between EMD and WT [123,128]. However, the filter bank properties of the EMD are data-driven: they depend on the statistical properties of the signal. In other words, the bandwidths of the filters and overlapping properties are data dependent [129–131].

4.12.5 SUMMARY AND CONCLUSIONS

The EMD is simple but its lack of mathematical rigor affects its application. Although the “basis” is obtained *a posteriori*, a number of parameters and thresholds need to be fixed beforehand. In addition, the method is not guaranteed to result in physically meaningful components. Therefore, adding to the deficiencies listed above in Section 4.12.4 and despite considering all these problems, EMD can still be useful in obtaining quantitative features of the signal under analysis as has been shown in a number of engineering applications such as seizure detection (Section 16.5); but more studies are needed to clarify the link with TFDs that have a firmer theoretical foundation.

Let us consider the cases where the EMD fails. For brevity, we will mention only two here. First, the EMD was shown to have some issues when decomposing multicomponent signals when a low-frequency component has nested transient high-frequency oscillations (known as the riding waves problem). Several possible solutions have been proposed but these result in a compromised EMD [132]. Other authors have shown that the decomposition is also variant with respect to sampling frequency [133]. All these findings suggest the need for extra care in using this method.

REFERENCES

- [1] M. Bastiaans, “Gabor’s expansion of a signal into Gaussian elementary signals”, *Proc. IEEE* 68 (4) (1980) 538-539, ISSN 0018-9219, <http://dx.doi.org/10.1109/PROC.1980.111686>.
- [2] I. Daubechies, Ten Lectures on Wavelets, second ed., Society for Industrial and Applied Mathematics, Philadelphia, 1992, <http://dx.doi.org/10.1137/1.9781611970104>, URL <http://pubs.siam.org/doi/abs/10.1137/1.9781611970104>.
- [3] J.G. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters”, *J. Opt. Soc. Am. A* 2 (7) (1985) 1160-1169, <http://dx.doi.org/10.1364/JOSAA.2.001160>, URL <http://josaa.osa.org/abstract.cfm?URI=josaa-2-7-1160>.
- [4] L. Liu, H. Hsu, “Inversion and normalization of time-frequency transform”, in: 2011 International Conference on Multimedia Technology (ICMT), 2011, pp. 2164-2168, <http://dx.doi.org/10.1109/ICMT.2011.6002490>.
- [5] P.S. Addison, The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, IOP Publishing, Bristol, UK, 2002.
- [6] M. Misiti, Y. Misiti, G. Oppenheim, J. Poggi, Wavelet Toolbox 4: User’s Guide. 1997-2007.
- [7] O. Rioul, P. Flandrin, “Time-scale energy distributions: a general class extending wavelet transforms”, *IEEE Trans. Signal Process.* 40 (7) (1992) 1746-1757.
- [8] O. Rioul, M. Vetterli, “Wavelets and signal processing”, *IEEE Signal Process. Mag.* 8 (1991) 14-38.
- [9] C. Torrence, G.P. Compo, “A practical guide to wavelet analysis”, *Bull. Am. Meteorol. Soc.* 79 (1) (1998) 61-78.
- [10] W.W. Boles, B. Boashash, “A human identification technique using images of the iris and wavelet transform”, *IEEE Trans. Signal Process.* 46 (4) (1998) 1185-1188, <http://dx.doi.org/10.1109/78.668573>.
- [11] P. Flandrin, Time-Frequency/Time-Scale Analysis, Academic Press, San Diego, 1999, translated by J. Stöckler from the French editions, *Temps-fréquence* (Paris: Hermès, 1993, 1998).
- [12] F. Hlawatsch, “Interference terms in the Wigner distribution”, in: Proc. Int. Conf. on Digital Signal Processing, 1984, pp. 363-367.
- [13] F. Hlawatsch, P. Flandrin, “The interference structure of the Wigner distribution and related time-frequency signal representations”, in: W. Mecklenbräuker, F. Hlawatsch (Eds.), The Wigner Distribution—Theory and Applications in Signal Processing, Elsevier, Amsterdam, 1997, pp. 59-133.
- [14] P. Flandrin, “Some features of time-frequency representations of multicomponent signals”, in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP’84), vol. 3, 1984, pp. 41B.4.1-41B.4.4.
- [15] H.-I. Choi, W.J. Williams, “Improved time-frequency representation of multicomponent signals using exponential kernels”, *IEEE Trans. Acoust. Speech Signal Process.* 37 (6) (1989) 862-871.
- [16] R.G. Baraniuk, D.L. Jones, “A signal-dependent time-frequency representation: optimal kernel design”, *IEEE Trans. Signal Process.* 41 (4) (1993) 1589-1602.
- [17] P. Flandrin, P. Borgnat, “Time-frequency energy distributions meet compressed sensing”, *IEEE Trans. Signal Process.* 58 (6) (2010) 2974-2982.
- [18] P. Flandrin, P. Gonçalvès, “Geometry of affine time-frequency distributions”, *Appl. Comput. Harmonic Anal.* 3 (1996) 10-39.
- [19] F. Hlawatsch, F. Auger (Eds.), Time-Frequency Analysis: Concepts and Methods, ISTE/Wiley, London, UK, 2008.
- [20] F. Hlawatsch, A. Papandreou-Suppappola, G.F. Boudreux-Bartels, “The power classes—quadratic time-frequency representations with scale covariance and dispersive time-shift covariance”, *IEEE Trans. Signal Process.* 47 (11) (1999) 3067-3083.
- [21] A. Papandreou, F. Hlawatsch, G.F. Boudreux-Bartels, “The hyperbolic class of quadratic time-frequency representations—part I: constant- Q warping, the hyperbolic paradigm, properties, and members”, *IEEE Trans. Signal Process.* 41 (12) (1993) 3425-3444.

- [22] J. Bertrand, P. Bertrand, "Affine time-frequency distributions", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, chap. 5, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 118-140.
- [23] F. Hlawatsch, H. Bölcseki, "Unified theory of displacement-covariant time-frequency analysis", in: Proc. IEEE-SP Int. Symp. on Time-Frequency & Time-Scale Analysis, 1994, pp. 524-527.
- [24] F. Hlawatsch, T. Twaroch, "Covariant (α, β) , time-frequency, and (a, b) representations", in: Proc. IEEE-SP Int. Symp. on Time-Frequency & Time-Scale Analysis, 1996, pp. 437-440.
- [25] A.M. Sayeed, D.L. Jones, "A canonical covariance-based method for generalized joint signal representations", *IEEE Signal Process. Lett.* 3 (4) (1996) 121-123.
- [26] F. Hlawatsch, G. Tauböck, T. Twaroch, "Covariant time-frequency analysis", in: L. Debnath (Ed.), Wavelets and Signal Processing, Birkhäuser, Boston, 2003.
- [27] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals", *Proc. IEEE* 80 (4) (1992) 520-538.
- [28] A.M. Sayeed, D.L. Jones, "Integral transforms covariant to unitary operators and their implications for joint signal representations", *IEEE Trans. Signal Process.* 44 (6) (1996) 1365-1377.
- [29] D. Gabor, "Theory of communication", *J. IEE* 93(III) (26) (1946) 429-457.
- [30] J. Hilgevoord, J. Uffink, "The mathematical expression of the uncertainty principle", in: A. van der Merwe, G. Tarozzi, F. Selleri (Eds.), Microphysical Reality and Quantum Formalism, no. 25-26 in Fundamental Theories of Physics, Kluwer, Dordrecht, 1988, pp. 91-114.
- [31] P.M. Oliveira, V. Barroso, "Uncertainty in the time-frequency plane", in: Proc. 10th IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000), 2000, pp. 607-611.
- [32] L. Cohen, Time-Frequency Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [33] N.G. de Bruijn, "A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence", *Nieuw Archief voor Wiskunde* (3) 21 (1973) 205-280.
- [34] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, O.J.J. Michel, "Measuring time-frequency information content using the Rényi entropies", *IEEE Trans. Inf. Theory* 47 (4) (2001) 1391-1409.
- [35] V. Sucic, N. Saulig, B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy", *EURASIP J. Adv. Signal Process.* 2011 (1) (2011) <http://dx.doi.org/10.1186/1687-6180-2011-125>, article ID 2011:125, 11 pp.
- [36] N. Saulig, Ž. Milanović, I. Orović, S. Stanković, V. Sucic, "Entropy based extraction and classification of frequency hopping signals from their TFDs", in: Proc. 55th International Symposium ELMAR-2013, 2013, pp. 169-172.
- [37] A.D. Whalen, Detection of Signals in Noise, Academic Press, New York, 1971.
- [38] J.C. Andrieux, R. Feix, G. Mourguès, P. Bertrand, B. Izrar, V.T. Nguyen, "Optimum smoothing of the Wigner-Ville distribution", *IEEE Trans. Acoust. Speech Signal Process.* 35 (6) (1987) 764-769.
- [39] A.W. Rihaczek, "Signal energy distribution in time and frequency", *IEEE Trans. Inf. Theory* 14 (3) (1968) 369-374.
- [40] B. Boashash, P.J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.* 42 (1) (1994) 216-220.
- [41] R.G. Baraniuk, D.L. Jones, "Unitary equivalence: a new twist on signal processing", *IEEE Trans. Signal Process.* 43 (10) (1995) 2269-2282.
- [42] M. Coates, W.J. Fitzgerald, "Time-frequency signal decomposition using energy mixture models", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000), vol. 2, 2000, pp. 633-636.
- [43] G.B. Folland, Harmonic Analysis in Phase Space, Princeton University Press, Princeton, NJ, 1989.
- [44] J. Bertrand, P. Bertrand, "A class of affine Wigner functions with extended covariance properties", *J. Math. Phys.* 33 (7) (1992) 2515-2527.
- [45] R.G. Baraniuk, "Beyond time-frequency analysis: energy densities in one and many dimensions", *IEEE Trans. Signal Process.* 46 (9) (1998) 2305-2314.

- [46] R.G. Baraniuk, L. Cohen, "On joint distributions of arbitrary variables", *IEEE Signal Process. Lett.* 2 (1) (1995) 10-12.
- [47] R.A. Altes, "Wideband, proportional-bandwidth Wigner-Ville analysis", *IEEE Trans. Acoust. Speech Signal Process.* 38 (6) (1990) 1005-1012.
- [48] L. Cohen, "The scale representation", *IEEE Trans. Signal Process.* 41 (12) (1993) 3275-3292.
- [49] R.G. Baraniuk, "Joint distributions of arbitrary variables made easy", *J. Multidim. Syst. Signal Process.* 9 (4) (1998) 341-348, special issue on time-frequency analysis.
- [50] G. Eichmann, N.M. Marinovich, "Scale-invariant Wigner distribution", in: Proc. SPIE: Analog Optical Processing and Computing, vol. 519, Soc. of Photo-Optical Instrumentation Engineers, 1984, pp. 18-25.
- [51] S.-C. Pei, S.-G. Huang, "STFT with adaptive window width based on the chirp rate", *IEEE Trans. Signal Process.* 60 (8) (2012) 4065-4080.
- [52] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [53] Z.M. Hussain, B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions", *IEEE Trans. Signal Process.* 50 (8) (2002) 1866-1876.
- [54] G. Jones, B. Boashash, "Generalized instantaneous parameters and window matching in the time-frequency plane", *IEEE Trans. Signal Process.* 45 (5) (1997) 1264-1275.
- [55] C.W. Helstrom, The Statistical Theory of Signal Detection, Oxford University Press, New York, 1968.
- [56] A.J.E.M. Janssen, "Optimality property of the Gaussian window spectrogram", *IEEE Trans. Signal Process.* 39 (1) (1991) 202-204.
- [57] N.G. De Bruijn, "Uncertainty principles in Fourier analysis", in: O. Shisha (Ed.), Inequalities, Academic Press, New York, NY, 1967, pp. 57-71.
- [58] S. Thangavelu, Lectures on Hermite and Laguerre expansions, Princeton University Press, Princeton, NJ, 1993.
- [59] D.L. Jones, R.G. Baraniuk, "A simple scheme for adapting time-frequency representations", *IEEE Trans. Signal Process.* 42 (12) (1994) 3530-3535.
- [60] R.N. Czerwinski, D.L. Jones, "Adaptive cone-kernel time-frequency analysis", *IEEE Trans. Signal Process.* 43 (7) (1995) 1715-1719.
- [61] J. Zhong, Y. Huang, "Time-frequency representation based on an adaptive short-time Fourier transform", *IEEE Trans. Signal Process.* 58 (10) (2010) 5118-5128.
- [62] G. Matz, F. Hlawatsch, "Fundamentals of time-varying communication channels", in: F. Hlawatsch, G. Matz (Eds.), Wireless Communications over Rapidly Time-Varying Channels, chap. 1, Academic Press, Oxford, UK, 2011, pp. 1-63.
- [63] G. Matz, H. Bölskei, F. Hlawatsch, "Time-frequency foundations of communications: concepts and tools", *IEEE Signal Process. Mag.* 30 (6) (2013) 87-96.
- [64] F. Hlawatsch, G. Matz, "Time-frequency methods for non-stationary statistical signal processing", in: F. Hlawatsch, F. Auger (Eds.), Time-Frequency Analysis: Concepts and Methods, chap. 10, ISTE/Wiley, London, UK, 2008, pp. 279-320.
- [65] W. Kozek, "On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems", in: Proc. IEEE-SP Int. Symp. on Time-Frequency & Time-Scale Analysis, 1992, pp. 167-170.
- [66] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework", *Signal Process.* 29 (1) (1992) 77-92.
- [67] G. Matz, F. Hlawatsch, "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory", *J. Math. Phys.* 39 (8) (1998) 4041-4070.
- [68] R.G. Shenoy, T.W. Parks, "The Weyl correspondence and time-frequency analysis", *IEEE Trans. Signal Process.* 42 (2) (1994) 318-331.

- [69] L.A. Zadeh, "Frequency analysis of variable networks", *Proc. IRE* 3A-8 (1950) 291-299.
- [70] P.A. Bello, "Characterization of randomly time-variant linear channels", *IEEE Trans. Commun. Syst.* 11 (4) (1963) 360-393.
- [71] W. Kozek, "On the transfer function calculus for underspread LTV channels", *IEEE Trans. Signal Process.* 45 (1) (1997) 219-223.
- [72] S. Barbarossa, A. Scaglione, "On the capacity of linear time-varying channels", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'99), 1999, pp. 2627-2630.
- [73] H.M. Ozaktas, Z. Zalevsky, M.A. Kutay, *The Fractional Fourier Transform: With Applications in Optics and Signal Processing*, Wiley, New York, 2001.
- [74] B.H.M. Testorf, J. Ojeda-Castañeda, *Phase-Space Optics: Fundamentals and Applications*, McGraw-Hill, New York, 2010.
- [75] J.A. Rodrigo, T. Alieva, M.L. Calvo, "Programmable two-dimensional optical fractional Fourier processor", *Opt. Expr.* 17 (7) (2009) 4976-4983.
- [76] M.G. Raymer, M. Beck, D.F. McAlister, "Complex wave-field reconstruction using phase-space tomography", *Phys. Rev. Lett.* 72 (8) (1994) 1137-1140.
- [77] A. Cámara, J.A. Rodrigo, T. Alieva, "Optical coherenscopy based on phase-space tomography", *Opt. Expr.* 21 (11) (2013) 13169-13183.
- [78] J.C. Wood, D.T. Barry, "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear FM signals", *IEEE Trans. Signal Process.* 42 (1994) 2094-2104.
- [79] L. Stanković, T. Alieva, M.J. Bastiaans, "Time-frequency signal analysis based on the windowed fractional Fourier transform", *Signal Process.* 83 (11) (2003) 2459-2468.
- [80] T. Alieva, M.J. Bastiaans, M.L. Calvo, "Fractional transforms in optical information processing", *EURASIP J. Appl. Signal Process.* 2005 (10) (2005) 1498-1519, <http://dx.doi.org/10.1155/ASP.2005.1498>.
- [81] H.-B. Sun, G.-S. Liu, H. Gu, W.-M. Su, "Application of the fractional Fourier transform to moving target detection in airborne SAR", *IEEE Trans. Aerosp. Electron. Syst.* 38 (4) (2002) 1416-1424.
- [82] D. Sazbona, Z. Zalevsky, E. Rivlin, D. Mendlovic, "Using Fourier/Mellin-based correlators and their fractional versions in navigational tasks", *Pattern Recogn.* 35 (12) (2002) 2993-2999.
- [83] B. Barshan, B. Ayrulu, "Fractional Fourier transform pre-processing for neural networks and its application to object recognition", *Neural Netw.* 15 (1) (2002) 131-140.
- [84] B.M. Hennelly, J.T. Sheridan, "Optical image encryption by random shifting in fractional Fourier domains", *Opt. Lett.* 28 (4) (2003) 269-271.
- [85] S. Stanković, I. Djurović, I. Pitas, "Watermarking in the space/spatial-frequency domain using two-dimensional Radon-Wigner distribution", *IEEE Trans. Image Process.* 10 (4) (2001) 650-658.
- [86] H.M. Ozaktas, O. Arıkan, M.A. Kutay, G. Bozdağı, "Digital computation of the fractional Fourier transform", *IEEE Trans. Signal Process.* 44 (9) (1996) 2141-2150.
- [87] L. Galleani, "The transient spectrum of a random system", *IEEE Trans. Signal Process.* 58 (10) (2010) 5106-5117.
- [88] L. Galleani, L. Cohen, "Direct time-frequency characterization of linear systems governed by differential equations", *IEEE Signal Process. Lett.* 11 (9) (2004) 721-724.
- [89] L. Galleani, "The time-frequency poles of a random system", in: Proc. IEEE Statistical Signal Processing Workshop (SSP 2011), 2011, pp. 793-796.
- [90] W. Martin, P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes", *IEEE Trans. Acoust. Speech Signal Process.* 33 (6) (1985) 1461-1470.
- [91] L. Cohen, L. Galleani, "Nonlinear transformation of differential equations into phase space", *EURASIP J. Appl. Signal Process.* 2004 (12) (2004) 1770-1777.
- [92] L. Galleani, "Response of dynamical systems to nonstationary inputs", *IEEE Trans. Signal Process.* 60 (11) (2012) 5775-5786.

- [93] L. Galleani, "Time-frequency representation of MIMO dynamical systems", *IEEE Trans. Signal Process.* 61 (17) (2013) 4309-4317.
- [94] A. Belouchrani, M.G. Amin, "Blind source separation based on time-frequency signal representation", *IEEE Trans. Signal Process.* 46 (11) (1998) 2888-2897.
- [95] H.M. Teager, S.M. Teager, "Evidence for nonlinear sound production mechanisms in the vocal tract", in: W.J. Hardcastle, A. Marchal (Eds.), *Speech Production and Speech Modelling*, NATO Advanced Study Inst. Series D: Behavioural and Social Sciences, vol. 55, Kluwer, 1990, pp. 241-261.
- [96] P. Maragos, J.F. Kaiser, T.F. Quatieri, "On amplitude and frequency demodulation using energy operators", *IEEE Trans. Signal Process.* 41 (4) (1993) 1532-1550.
- [97] J.F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'90), vol. 1, 1990, pp. 381-384.
- [98] P. Maragos, A. Potamianos, "Higher order differential energy operators", *IEEE Signal Process. Lett.* 2 (8) (1995) 152-154.
- [99] S. Mukhopadhyay, G.C. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection", *IEEE Trans. Biomed. Eng.* 45 (2) (1998) 180-187.
- [100] P. Maragos, J.F. Kaiser, T.F. Quatieri, "Energy separation in signal modulations with application to speech analysis", *IEEE Trans. Signal Process.* 41 (10) (1993) 3024-3051.
- [101] D. Dimitriadis, P. Maragos, "Continuous energy demodulation methods and application to speech analysis", *Speech Commun.* 48 (7) (2006) 819-837.
- [102] A. Bouchikhi, A.-O. Boudraa, "Multicomponent AM-FM signals analysis based on EMD-B-splines ESA", *Signal Process.* 92 (9) (2012) 2214-2228.
- [103] A.-O. Boudraa, "Instantaneous frequency estimation of FM signals by Ψ_B -energy operator", *Electron. Lett.* 47 (10) (2011) 623-624.
- [104] A.-O. Boudraa, "Relationships between Ψ_B -energy operator and some time-frequency representations", *IEEE Signal Process. Lett.* 17 (6) (2010) 527-530.
- [105] F. Salzenstein, A.-O. Boudraa, J.-C. Cexus, "Generalized higher-order nonlinear energy operators", *J. Opt. Soc. Am. A* 24 (12) (2007) 3717-3727.
- [106] F. Salzenstein, A.-O. Boudraa, T. Chonavel, "A new class of multi-dimensional Teager-Kaiser and higher order operators based on directional derivatives", *J. Multidim. Syst. Signal Process.* 24 (3) (2013) 543-572.
- [107] A.-O. Boudraa, S. Benramdane, J.-C. Cexus, T. Chonavel, "Some useful properties of cross- Ψ_B -energy operator", *Int. J. Electron. Commun.* 63 (9) (2009) 728-735.
- [108] J.-C. Cexus, A.-O. Boudraa, "Link between cross-Wigner distribution and cross-Teager energy operator", *Electron. Lett.* 40 (12) (2004) 778-780.
- [109] R. Hamila, J. Astola, F.A. Cheikh, M. Gabbouj, M. Renfors, "Teager energy and the ambiguity function", *IEEE Trans. Signal Process.* 47 (1) (1999) 260-262.
- [110] W.-F. Zhang, C.-C. Liu, H. Yan, "Clustering of temporal gene expression data by regularized spline regression and an energy based similarity measure", *Pattern Recogn.* 43 (12) (2010) 3969-3976.
- [111] A.-O. Boudraa, J.-C. Cexus, K. Abed-Meraim, "Cross Ψ_B -energy operator-based signal detection", *J. Acoust. Soc. Am.* 123 (6) (2008) 4283-4289.
- [112] F. Salzenstein, P. Montgomery, A.-O. Boudraa, "Local frequency and envelope estimation by Teager-Kaiser energy operators in white-light scanning interferometry", *Opt. Expr.* 22 (15) (2014) 18325-18334.
- [113] B. Tacer, P.J. Loughlin, "Non-stationary signal classification using the joint moments of time-frequency distributions", *Pattern Recogn.* 31 (11) (1998) 1635-1641, ISSN 0031-3203, [http://dx.doi.org/10.1016/S0031-3203\(98\)00031-4](http://dx.doi.org/10.1016/S0031-3203(98)00031-4), URL <http://www.sciencedirect.com/science/article/pii/S0031320398000314>.
- [114] A. Bouchikhi, A.-O. Boudraa, J.-C. Cexus, T. Chonavel, "Analysis of multicomponent LFM signals by Teager Huang-Hough transform", *IEEE Trans. Aerosp. Electron. Syst.* 50 (2) (2014) 1222-1233, ISSN 0018-9251, <http://dx.doi.org/10.1109/TAES.2014.120202>.

- [115] S. Qian, “Gabor spectrogram”, in: B. Boashash (Ed.), Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, first ed., chap. 4, Article 4.9, Elsevier, Oxford, 2003, pp. 153-158.
- [116] S. Qian, D. Chen, “Joint time-frequency analysis”, *IEEE Signal Process. Mag.* 16 (2) (1999) 52-65.
- [117] H. Shao, W. Jin, S. Qian, “Order tracking by discrete Gabor expansion”, *IEEE Trans. Instrument. Measur.* 52 (3) (2003) 754-761.
- [118] L. Rankine, M. Mesbah, B. Boashash, “A matching pursuit-based signal complexity measure for the analysis of newborn EEG”, *Med. Biol. Eng. Comput.* 45 (3) (2007) 251-260, URL <http://www.springerlink.com>.
- [119] M.A. Awal, M.M. Lai, G. Azemi, B. Boashash, P.B. Colditz, “EEG background features that predict outcome in term neonates with hypoxic ischaemic encephalopathy: a structured review”, *Clin. Neurophysiol.* (2015) ISSN 1388-2457, <http://dx.doi.org/10.1016/j.clinph.2015.05.018>, URL <http://www.sciencedirect.com/science/article/pii/S1388245715006136>.
- [120] M.A. Awal, P.B. Colditz, B. Boashash, G. Azemi, “Detection of neonatal EEG burst-suppression using a time-frequency approach”, in: 8th International Conference on Signal Processing and Communication Systems (ICSPCS), 2014, pp. 1-6, <http://dx.doi.org/10.1109/ICSPCS.2014.7021073>.
- [121] N.E. Huang, “Introduction to Hilbert-Huang transform and its related mathematical problems”, in: N.E. Huang, S. Shen (Eds.), The Hilbert-Huang Transform and Its Applications, chap. 1, World Scientific Publishing Company, Singapore, 2005, pp. 1-24.
- [122] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”, *Proc. R. Soc. A* 454 (1971) (1998) 903-995.
- [123] P. Flandrin, P. Gonçalvès, G. Rilling, “EMD equivalent filter banks, from interpretation to applications”, in: N.E. Huang, S. Shen (Eds.), The Hilbert-Huang Transform and Its Applications, chap. 3, World Scientific Publishing Company, Singapore, 2005, pp. 77-87.
- [124] N.J. Stevenson, M. Mesbah, B. Boashash, “Multiple-view time-frequency distribution based on the empirical mode decomposition”, *IET Signal Process.* 4 (4) (2010) 447-456.
- [125] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), Advances in Spectrum Analysis and Array Processing, vol. 1, chap. 9, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517.
- [126] V. Sucic, J. Lerga, B. Boashash, “Multicomponent noisy signal adaptive instantaneous frequency estimation using components time support information”, *IET Signal Process.* 8 (3) (2014) 277-284.
- [127] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, “Time-frequency reassignment and synchrosqueezing: an overview”, *IEEE Signal Process. Mag.* 30 (6) (2013) 32-41.
- [128] Z. Wu, N.E. Huang, “A study of the characteristics of white noise using the empirical mode decomposition method”, *Proc. R. Soc. A* 460 (2046) (2004) 1597-1611.
- [129] G. Rilling, P. Flandrin, “One or two frequencies? The empirical mode decomposition answers”, *IEEE Trans. Signal Process.* 56 (1) (2008) 85-95, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2007.906771>.
- [130] A. Mert, A. Akan, “Detrended fluctuation thresholding for empirical mode decomposition based denoising”, *Digital Signal Process.* 32 (2014) 48-56, ISSN 1051-2004, <http://dx.doi.org/10.1016/j.dsp.2014.06.006>, URL <http://www.sciencedirect.com/science/article/pii/S1051200414001936>.
- [131] N. Rehman, D. Mandic, “Filter bank property of multivariate empirical mode decomposition”, *IEEE Trans. Signal Process.* 59 (5) (2011) 2421-2426, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2011.2106779>.
- [132] Z. Yang, L. Yang, C. Qing, D. Huang, “A method to eliminate riding waves appearing in the empirical AM/FM demodulation”, *Digital Signal Process.* 18 (4) (2008) 488-504, ISSN 1051-2004, <http://dx.doi.org/10.1016/j.dsp.2007.07.003>, URL <http://www.sciencedirect.com/science/article/pii/S105120040700111X>.
- [133] N. Stevenson, M. Mesbah, B. Boashash, “A sampling limit for the empirical mode decomposition”, in: Proc. 8th Int. Conf. on Information Science, Signal Processing and Their Applications (ISSPA’05).

ADVANCED DESIGN AND SPECIFICATIONS OF TFDs

5

INTRODUCTION AND OVERVIEW

This chapter describes in detail specific examples of design of time-frequency distributions (TFDs), extending the material described in [Chapters 2](#) and [3](#) and covering it in more depth and at a more advanced level. This key time-frequency (t,f) topic is covered in 11 sections with appropriate internal and external cross-referencing.

Ambiguity functions (AFs) are traditionally used in radar and sonar. As dual of TFDs via a 2D Fourier transform (FT), they are a key basis for TFD kernel design methodologies ([Section 5.1](#)). One of the first TFDs constructed on this basis is the Gaussian kernel distribution, a TFD designed for its reduced interference properties ([Section 5.2](#)). To better adapt to the signal under analysis and achieve higher resolution and concentration performance, we can design adaptive TFDs using optimization theory ([Section 5.3](#)). TFDs can also be designed and adapted to a specific class of signals such as polynomial frequency modulated (FM) signals, leading to the formulation of polynomial Wigner-Ville distributions (WVDs) ([Section 5.4](#)). The design of such methods can be related mathematically to higher-order spectra (HOS) ([Section 5.5](#)). Another example of TFDs adapted to a particular class of signals is that of TFDs adapted to signals with dispersive group delay ([Section 5.6](#)). A step-by-step methodology for the design of a specific TFD is provided using separable kernels ([Section 5.7](#)). The class of generalized marginal TFDs is related to the fractional FT, allowing further design possibilities ([Section 5.8](#)). [Section 5.9](#) discusses fixed kernel high-resolution TFDs that improve performance by accounting for directions of energy concentration of signal components in the (t,f) domain. Adaptive methods that estimate the direction of the smoothing kernel on a point-by-point basis are presented in [Section 5.10](#). [Section 5.11](#) uses linear (t,f) methods based on a modified version of the S-transform (ST) to estimate the phase synchrony. Finally, [Section 5.12](#) reviews the requirements for designing positive TFDs.

5.1 AMBIGUITY FUNCTIONS⁰

5.1.1 THE RADAR/SONAR PROBLEM

Let us consider the typical radar/sonar problem in which the detection of a target (and the estimation of its relative range d and velocity v with respect to the emitter/receiver) is achieved from the analysis of the returning echo $r(t)$ associated to a given emitted waveform $x(t)$. Assuming a perfect reflection in the echo formation process and a constant radial velocity between the emitter/receiver and the target, $r(t)$ can be modeled as an attenuated replica of $x(t)$, up to a (range encoding) round trip delay, a (velocity encoding) modification due to the Doppler effect and some observation noise. On the basis of various criteria (maximum likelihood, Neyman-Pearson strategy, maximum contrast, etc.), it is known [1] that a basic ingredient for solving the detection problem is a measure of (linear) similarity, in the L^2 -sense of a correlation, between the signal to detect and the actual echo (the “matched filter” principle). Given the assumed model, it is therefore natural to compare the received echo with a battery of templates $(\mathbf{T}_{d',v'}x)(t)$, where $\mathbf{T}_{d',v'}$ stands for the range-velocity transformation attached to the candidate pair (d', v') , so that estimates of d and v can be inferred from:

$$(\hat{d}, \hat{v}) := \arg \max_{(d', v')} |\langle r, \mathbf{T}_{d',v'}x \rangle|. \quad (5.1.1)$$

Ideally, the deterministic part of the above inner product would be zero for all range-velocity pairs except for $(d', v') \equiv (d, v)$. This, however, proves not to be achievable (as explained below), leading to a joint determination of range and velocity which is intrinsically ambiguous: this is the reason why a quantity of the type $\langle x, \mathbf{T}_{d,v}x \rangle$ is loosely referred to as an ambiguity function (AF).

5.1.2 DEFINITIONS OF AMBIGUITY FUNCTIONS

To be more specific in defining AFs, we must take care with physical considerations about the Doppler effect, which causes a time-stretching of the returning echo.

Narrowband ambiguity functions. In the general case, the Doppler factor expresses as $\eta := (c + v)/(c - v)$, where c stands for the celerity of the propagating waves in the considered medium. In the radar case, the celerity of electromagnetic waves is $c \approx 3 \times 10^8$ m/s and, even if we assume a relative target velocity as large as $v = 3600$ km/h, we end up with a Doppler factor such that $\eta - 1 = 6.66 \times 10^{-6} \ll 1$, justifying the approximation $\eta \approx 1 + 2v/c$. It follows that, if the emitted signal is of the form $x(t) := \tilde{x}(t) \exp\{j2\pi f_0 t\}$, where $\tilde{x}(t)$ is a complex envelope that is narrowband with respect to the carrier f_0 , the deterministic part $r_d(t)$ of the returning echo $r(t)$ admits the approximation $r_d(t) \propto x(t - \tau) \exp\{j2\pi v t\}$, with τ a round trip delay such that $\tau := 2d/v$ and v a Doppler shift such that $v := 2f_0 v/c$. The corresponding inner product

$$\langle x, r_d \rangle \propto \int_{-\infty}^{+\infty} x(t) x^*(t - \tau) e^{-j2\pi v t} dt \quad (5.1.2)$$

⁰Author: **P. Flandrin**, École Normale Supérieure de Lyon, Lyon Cedex 07, France (flandrin@ens-lyon.fr). Reviewers: J.P. Ovarlez and A. Papandreou-Suppappola.

is therefore proportional to a quantity referred to as the *narrowband* AF of $x(t)$. Whereas this formulation is the one initially introduced by Woodward [2], it often proves useful to adopt the *symmetrized* definition used by Sussman [3], that is,

$$A_x(v, \tau) := \int_{-\infty}^{+\infty} x(t + \tau/2)x^*(t - \tau/2) e^{-j2\pi vt} dt, \quad (5.1.3)$$

which differs from Woodward's by the (unit-amplitude) factor $e^{j\pi v\tau}$.

Wideband ambiguity functions. The above definitions (5.1.2) and (5.1.3) are based on approximations that may prove not to be relevant in contexts different from radar. This is especially the case in airborne sonar, where the celerity of acoustic waves is $c = 340$ m/s, thus leading to $\eta \approx 1.2$ for relative radial velocities $v \approx 100$ km/h. A similar situation can also be observed (although to a smaller extent) in underwater sonar, where the sound celerity in water $c = 1500$ m/s and typical relative velocities $v \approx 2.6$ m/s lead to $\eta \approx 1.0034$. In such cases, the previous approximation of a Doppler shift may no longer be valid for wideband signals, and the more general form

$$\tilde{A}_x(\eta, \tau) := \sqrt{\eta} \int_{-\infty}^{+\infty} x(t)x^*(\eta(t - \tau)) dt \quad (5.1.4)$$

may be preferred as a definition of a *wideband* AF [4].

Such a wideband definition naturally reduces to the narrowband one when the analyzed signal is narrowband.

Ambiguity functions and time-frequency distributions. AFs can be viewed as two-variable generalizations of correlation functions. In this respect, they are dual of energy distributions. As seen in Sections 3.2.1 and 3.2.3, it follows from the definition (5.1.3) that

$$\iint_{-\infty}^{+\infty} A_x(v, \tau) e^{j2\pi(vt - \tau f)} dv d\tau = W_x(t, f), \quad (5.1.5)$$

where $W_x(t, f)$ is the WVD.¹ More generally, the whole time- and frequency-covariant quadratic class of TFDs can be obtained as the 2D FT (half forward, half inverse) of weighted (narrowband) AFs $g(v, \tau)A_x(v, \tau)$ (see Eq. (3.2.17) in p. 114).

Similarly, a properly symmetrized version of the wideband AF (Eq. 5.1.4) can be shown [6] to be in Mellin-Fourier duality with a wideband TFD, known as *Altes' Q-distribution* $Q_x(t, f)$. More precisely, if we let

$$\hat{A}_x(\eta, \tau) := \int_{-\infty}^{+\infty} x\left(\eta^{-1/2}(t + \tau/2)\right)x^*\left(\eta^{+1/2}(t - \tau/2)\right) dt = \tilde{A}_x\left(\eta, \eta^{-1/2}\tau\right), \quad (5.1.6)$$

we have [6]

$$\int_0^{+\infty} \int_{-\infty}^{+\infty} \hat{A}_x(\eta, \tau) e^{-j2\pi f\tau} \eta^{j2\pi t-1} d\eta d\tau = Q_x(t, f), \quad (5.1.7)$$

with the warping equivalence $Q_x(t, f) \equiv W_{\check{x}}(t, \log f)$, if $\check{X}(f) := X(e^f)$.

¹It is worth noting that the symmetrized AF (Eq. 5.1.3) has in fact been pioneered by Ville [5] as a form of “time-frequency characteristic function.”

Another interesting connection can be pointed out between AFs and linear time-frequency or time-scale representations. In fact, the right-hand side of Eq. (5.1.2) can be viewed as the *short-time Fourier transform* (STFT) of $x(t)$, with window $h(t) := x(t)$, whereas the right-hand side of Eq. (5.1.4) is nothing but the *wavelet transform* of $x(t)$, with wavelet $\psi(t) := x(t)$ and scale $a := 1/\eta$. In both cases, the AF is exactly identical to the *reproducing kernel* of the corresponding linear transform [7].

5.1.3 PROPERTIES OF NARROWBAND AMBIGUITY FUNCTIONS

Invariances and covariances. Whereas members of the quadratic class are *covariant* with respect to time and frequency shifts, the squared modulus of the AF (a quantity referred to as the *ambiguity surface*) is *invariant* to such transformations (i.e., $|A_y(v, \tau)|^2 = |A_x(v, \tau)|^2$ for any shifted version $y(t) := x(t - \theta) \exp[j2\pi\xi t]$ of a given signal $x(t)$). In a similar way, the AF inherits (by Fourier duality) a number of properties satisfied by the WVD, such as covariance with respect to dilations, rotations, or chirp modulations [7,8].

Cross-sections. As has been mentioned, the narrowband AF can be seen as a *correlation function* with respect to time and frequency shifts. As such, it is Hermitian symmetric, that is, $A_x(-v, -\tau) = A_x^*(v, \tau)$, and it satisfies the inequality

$$|A_x(v, \tau)| \leq |A_x(0, 0)| = \|x\|_2^2. \quad (5.1.8)$$

Although this interpretation cannot be pushed too far (in particular, the AF is not a non-negative definite quantity, since its 2D FT—namely, the WVD—can attain negative values), cross-sections of the AF are meaningful 1D correlation functions, since we have

$$A_x(0, \tau) = \int_{-\infty}^{+\infty} x(t + \tau/2)x^*(t - \tau/2) dt \quad (5.1.9)$$

and

$$A_x(v, 0) = \int_{-\infty}^{+\infty} X(f + v/2)X^*(f - v/2) df. \quad (5.1.10)$$

This idea of a time-frequency correlation function (which is shown in Fig. 5.1.1) is instrumental in the design of reduced-interference distributions (RIDs) within the quadratic class (cf. Section 4.2).

Volume invariance and self-transformation. Using Parseval's relation and Moyal's formula [7], we readily find that for any two signals x and y ,

$$\iint_{-\infty}^{+\infty} A_x(v, \tau)A_y^*(v, \tau) dv d\tau = \left| \int_{-\infty}^{+\infty} x(t)y^*(t) dt \right|^2. \quad (5.1.11)$$

For $y \equiv x$, this becomes

$$\iint_{-\infty}^{+\infty} |A_x(v, \tau)|^2 dv d\tau = \|x\|_2^4; \quad (5.1.12)$$

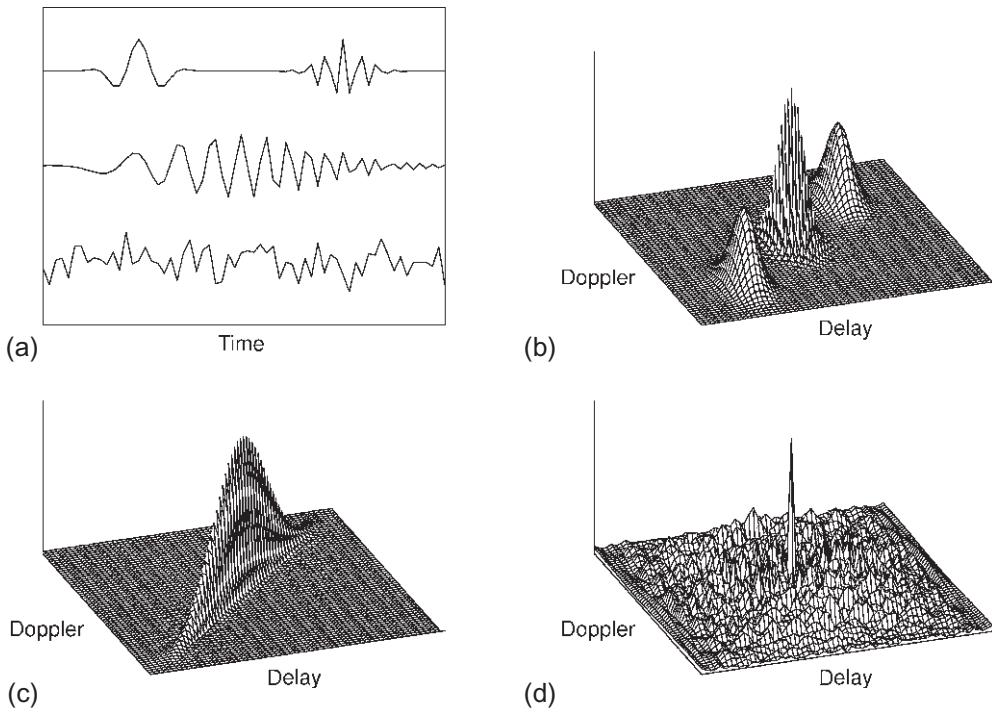


FIGURE 5.1.1

Ambiguity functions as time-frequency correlation functions. Subplots (b) to (d) display the ambiguity surfaces attached, respectively, to the three signals plotted in (a), namely (from top to bottom): two Gabor logons, a linear chirp, and a sample of white Gaussian noise. Each surface attains its maximum value at the origin, with values off the origin that reveal correlations in the signal structure, with respect to both time shifts and frequency shifts (delay and Doppler).

that is, the ambiguity surface has an *invariant volume* fixed by the signal's energy. More remarkably, Eq. (5.1.12) is just a special case of Siebert's *self-transformation* property [9]:

$$\iint_{-\infty}^{+\infty} |A_x(v, \tau)|^2 e^{-j2\pi(vt - \tau f)} dv d\tau = |A_x(f, t)|^2, \quad (5.1.13)$$

from which it can be inferred that an AF is a highly structured function.

Uncertainty principles. If we combine the “correlation” inequality Eq. (5.1.8) with “volume invariance” Eq. (5.1.12), it is clear that an ambiguity surface cannot be perfectly concentrated at the origin of the plane. This limitation, which is sometimes called the *radar uncertainty relation* [10], admits a more precise L^p -norm formulation ($p > 2$) as follows [11]:

$$I_x(p) := \iint_{-\infty}^{+\infty} |A_x(v, \tau)|^p dv d\tau \leq \frac{2}{p} \|x\|_2^{2p}, \quad (5.1.14)$$

with equality if and only if x is a linear chirp with a Gaussian envelope.

The common interpretation of this inequality is that an ambiguity surface cannot be zero everywhere except at the origin of the plane. In the case where all of the ambiguity surface is supposed to be concentrated around the origin, it has necessarily to extend over a domain (whose area is nonzero), which defines the joint accuracy of any delay-Doppler measurement [1,12]. However, ambiguity surfaces, which are more sharply peaked at the origin can be found, provided that nonzero values are accepted somewhere off the origin in the plane: for unit energy signals, ambiguity surfaces with null values except in $(0,0)$ can be obtained on convex domains whose *clear area* cannot, however, be greater than 4; see [13]. An example is given in Fig. 5.1.2.

Delay/Doppler estimation. The best achievable performance in joint estimation of delay and Doppler is bounded. The actual Cramér-Rao bounds on variances and covariances can be derived from the Fisher information matrix of the problem, whose terms are themselves expressible as partial derivatives of the ambiguity surface in the case of additive white Gaussian noise [1]. Since the ambiguity surface is basically the maximum likelihood estimator for delay and Doppler, and since this estimator can be shown to be asymptotically efficient, it follows that the ambiguity surface geometry is a direct indicator of the expected accuracy in the estimation. Roughly speaking, variances in delay and Doppler estimations are given by the effective widths of the central peak of the ambiguity surface.

Signal design. In active problems, in which the emitted signal can be freely chosen (up to a certain extent), an important issue is to design waveforms with a prescribed AF (or ambiguity surface), so that some desired performance can be guaranteed. From a purely theoretical point of view, a signal is entirely determined (up to a pure phase term) by its AF, since we can invert the definition (5.1.3) according to:

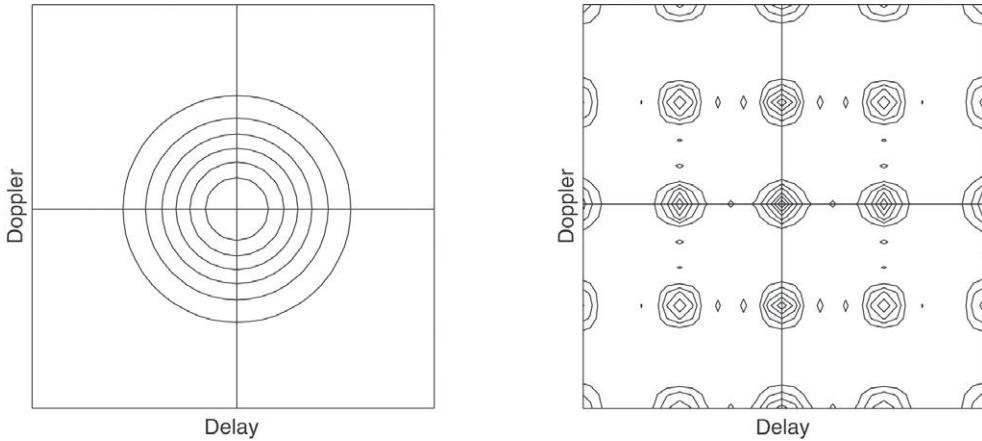


FIGURE 5.1.2

Sharpening the central peak of an ambiguity surface. *Left:* A single Gaussian pulse has an ambiguity surface whose central peak cannot have an effective area \mathcal{A} smaller than a limit fixed by the “radar uncertainty principle.” *Right:* A signal defined by the superposition of a number of replicae of such a pulse, periodically shifted in time and frequency, may guarantee a sharper central peak. But this comes at the price of auxiliary peaks off the origin, with a “clear area” of the order of \mathcal{A} .

$$x(\tau = t) = \frac{1}{x^*(0)} \int_{-\infty}^{+\infty} A_x(v, \tau) e^{j\pi v \tau} dv. \quad (5.1.15)$$

Unfortunately, because an AF (as noted above) is a highly structured function, an arbitrary 2D function has in general no reason to be admissible—that is, no reason to be the actual AF of some signal. Different approaches have been proposed to overcome this limitation. One can first think of looking for the signal $\hat{x}(t)$ whose AF best approaches a given time-frequency function $F(v, \tau)$ according to, for example, an L^2 -distance [3]:

$$\hat{x}(\tau = t) = \arg \min_x \iint_{-\infty}^{+\infty} |A_x(v, \tau) - F(v, \tau)|^2 dv d\tau. \quad (5.1.16)$$

One can also rely on the physical interpretation of the AF as a time-frequency correlation function and promote waveforms with adapted time-frequency characteristics. In this respect, a large bandwidth (or long duration) is required for an accurate estimation of delay (or Doppler, respectively). The simultaneous consideration of these two design principles advocates the use of a *chirp* signal with a large bandwidth-duration (BT) product (see [12] and [Section 1.2.4](#)).

5.1.4 REMARKS ON WIDEBAND AMBIGUITY FUNCTIONS

In many respects, properties of the wideband AF can be seen as natural generalizations of the narrowband case (although some properties, such as volume invariance, may no longer be satisfied), reducing them to the narrowband limit.

Parallel to what is mentioned above in the narrowband case, the best achievable performance in the joint estimation of delay and Doppler can be expressed, in the wideband case, in terms of geometrical properties of the wideband ambiguity surface [14].

A companion problem is that of *Doppler tolerance*, which consists in obtaining an unbiased estimate of delay in the presence of any unknown Doppler [15]. As Doppler stretches the emitted signal, the condition for no bias translates into the condition that the effective time-frequency structure of the emitted waveform is invariant under stretching. It turns out that the hyperbola is the only curve of the plane which is invariant under dilation/compression transformations: in terms of chirps, assumed to be conveniently described on the plane by a time-frequency skeleton, this justifies the use of *logarithmic* phases [7], that is, of hyperbolic chirps resembling those commonly observed in natural sonar systems, for example, of bats [15].

Finally, it must be pointed out that computing an AF proves to be more involved in the wideband case than in the narrowband case. Efficient solutions, based on the Mellin transform, have been proposed in [14].

5.1.5 SUMMARY AND CONCLUSIONS

AFs measure the degree of similarity between a signal and its delayed and Dopplerized versions; as such, they are central in the evaluation and design of radar/sonar systems. Since they are basically (2D) time-frequency correlation functions, AFs also offer a dual perspective on quadratic TFDs that proves especially useful for kernel design issues.

5.2 REDUCED INTERFERENCE TIME-FREQUENCY DISTRIBUTIONS⁰

5.2.1 NONSTATIONARITY, RESOLUTION, AND INTERFERENCE

Following from Chapter 3 and Sections 4.2 and 5.1, this section presents the concept and design methodology for Reduced Interference Distributions (RIDs) in order to handle nonstationary signals such as chirps satisfactorily. The spectrogram often presents serious difficulties when used to analyze rapidly varying signals. If the analysis window is made short enough to capture rapid changes in the signal it becomes impossible to resolve frequency components of the signal which are close in frequency during the analysis window duration (see Sections 2.3.1 and 2.3.2).

The WD has been employed as an alternative to overcome this shortcoming. It is also called WVD when it uses the analytic form of the signal (see Sections 2.1.3 and 2.1.4). Among its many important and interesting properties (see Chapter 2 and [16]), it provides a high-resolution representation in time and in frequency for a nonstationary signal such as a chirp. However, its energy distribution is not non-negative and it often possesses severe cross-terms, or interference terms, between components in different (t, f) regions, potentially leading to confusion and misinterpretation (see Section 4.2). RIDs effect a desirable compromise, which retains a number of desirable TFD properties, yet provides reduced cross-term interference, making them a good choice for the analysis of nonstationary signals.

Both the spectrogram and the WVD are members of the quadratic class of distributions [16]. For these TFDs a time shift in the signal is reflected as an equivalent time shift in the (t, f) distribution and a shift in the frequency of the signal is reflected as an equivalent frequency shift in the (t, f) distribution. The spectrogram, the WVD, and the RID all have this property. Different TFDs in the quadratic class can be obtained by selecting different kernel functions (see Section 2.7). The author and co-workers introduced one of the earliest “new” distributions [17], which they called the exponential distribution or ED to overcome several drawbacks of the spectrogram and WD, providing high resolution with suppressed interferences [17,18]. The method used in developing the ED is general and is the basis of this section.

5.2.2 THE REDUCED INTERFERENCE DISTRIBUTION

5.2.2.1 Ambiguity function relationships

The key to understanding (t, f) relationships and manipulations is a thorough understanding of the ambiguity domain. Let $Z(f)$ be the FT of the signal $z(t)$:

$$Z(f) = \mathcal{F}_{t \rightarrow f} \{z(t)\} = \int z(t) e^{-j2\pi ft} dt \quad (5.2.1)$$

and

$$z(t) = \mathcal{F}_{f \leftarrow t}^{-1} \{Z(f)\} = \int Z(f) e^{j2\pi ft} df. \quad (5.2.2)$$

⁰Author: **W.J. Williams**, The University of Michigan, Ann Arbor, MI, USA (wjw@umich.edu). Reviewers: B. Boashash and P. Flandrin.

Let $R_z(t, \tau)$ be the *instantaneous autocorrelation* of a complex signal $z(t)$; that is,

$$K_z(t, \tau) = z(t + \tau/2)z^*(t - \tau/2), \quad (5.2.3)$$

where z^* denotes the complex conjugate of z . The WD of $z(t)$ is the FT of $R_z(t, \tau)$ with respect to the lag variable τ :

$$W_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{z(t + \tau/2)z^*(t - \tau/2)\} = \mathcal{F}_{\tau \rightarrow f} \{K_z(t, \tau)\}. \quad (5.2.4)$$

Similarly, but with a different physical meaning, the symmetrical AF is defined as the inverse Fourier transform (IFT) of $R_z(t, \tau)$ with respect to t :

$$A_z(v, \tau) = \mathcal{F}_{t \rightarrow v} \{z(t + \tau/2)z^*(t - \tau/2)\} = \mathcal{F}_{t \rightarrow v} \{K_z(t, \tau)\}. \quad (5.2.5)$$

Thus, $W_z(t, f)$ and $A_z(v, \tau)$ are related by the 2D FT:

$$W_z(t, f) = \iint A_z(v, \tau) e^{-j2\pi(-tv+f\tau)} dv d\tau. \quad (5.2.6)$$

Equation (5.2.6) may be altered with a kernel to obtain $\rho_z(t, f; g)$ as:

$$\rho_z(t, f; g) = \iint g(v, \tau) A_z(v, \tau) e^{-j2\pi(-tv+f\tau)} dv d\tau, \quad (5.2.7)$$

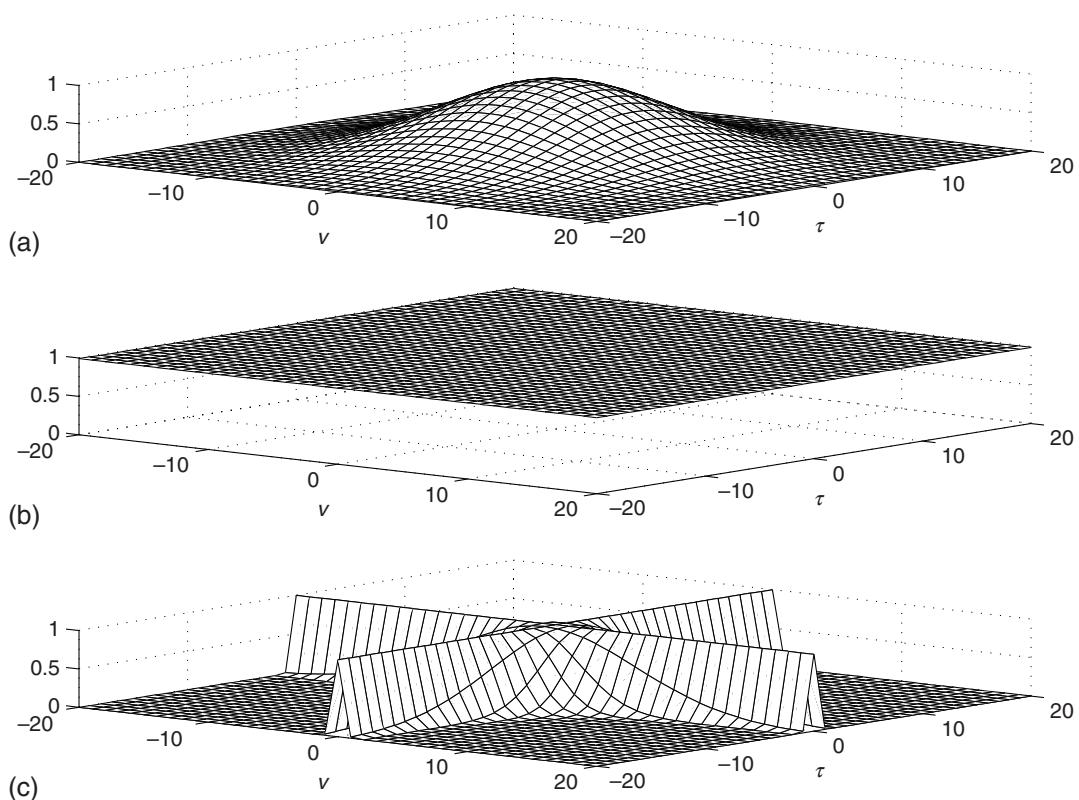
which is an expression of the quadratic class of distributions, in general. Thus, while the WD may be found from the symmetric AF by means of a double FT, any member of the quadratic class of distributions may be found by first multiplying the kernel $g(v, \tau)$ by the symmetric AF and then carrying out the double FT. The generalized AF, $g(v, \tau)A_z(v, \tau)$ [19], is a key concept in (t, f) which aids one in clearly seeing the effect of the kernel in determining $\rho_z(t, f; g)$.

The kernel for the WD is unity, so the generalized AF is identical to the AF, and its (t, f) representation (the double FT) preserves both the auto-terms and the cross-terms. The kernels of the spectrogram and the RID emphasize the auto-terms and deemphasize the cross-terms, but in very different ways.

5.2.2.2 The exponential distribution

Each TFD discussed is valuable under certain conditions. The ED, developed in [17], is an attempt to improve on the WD. It has a kernel $g(v, \tau) = \exp\left(\frac{-v^2\tau^2}{\sigma}\right)$, and it proves to be quite effective in suppressing the interferences while retaining high resolution. A comparison of the kernels of the spectrogram, the WD, and the ED are shown in Fig. 5.2.1.

Interference terms tend to lie away from the axes in the ambiguity plane, while auto-terms tend to lie on the axes. The spectrogram kernel attenuates everything away from the $(0, 0)$ point, the WD kernel passes everything, and the ED kernel passes everything on the axes and attenuates what is away from the axes. Thus, the ED generally has a reduced interference characteristic. The spectrogram reduces interference also, but at a cost to the auto-terms. Reduced interference may be achieved while maintaining a number of nice mathematical properties with just a few constraints. The ED is just one

**FIGURE 5.2.1**

Comparison of (a) spectrogram, (b) WD, and (c) ED kernels in the ambiguity plane.

example of the RID class. A methodology which guarantees an RID yet allows a considerable flexibility in design is discussed in the following section.

5.2.3 KERNEL SELECTION FOR REDUCED INTERFERENCE DISTRIBUTION (RID)

A more formal description of RIDs is appropriate at this point. RID requirements and properties are discussed in comparison with the WD. The unity value of the WD kernel guarantees the desirable properties of the WD. However, it is not necessary to require the kernel to be unity for all ν in order to maintain most of its desirable properties. It is sufficient to ensure that the kernel is unity along $\nu = 0$ and $\tau = 0$ and that the kernel is such that $g^*(\nu, \tau) = g(-\nu, -\tau)$, the latter property ensuring realness. The RID kernel is cross-shaped and acts as a low-pass filter in both ν and τ . The spectrogram has the virtue of suppressing cross-terms as does the RID, and has the further advantage of being non-negative which is not the case for the WD and the RID. The RID possesses almost all of the desirable properties of the WD except for its unitary property ($|g(\nu, \tau)| = 1$ for all ν, τ).

Table 5.2.1: TFD Properties and Associated Kernel Requirements

- P0. Non-negativity: $\rho_z(t, f; g) \geq 0 \quad \forall t, f$
- R0. $g(v, \tau)$ is a positively weighted sum of AFs
- P1. Realness: $\rho_z(t, f; g) \in R$
- R1. $g(v, \tau) = g^*(-v, -\tau)$
- P2. Time shift: $s(t) = z(t - t_0) \Rightarrow \rho_s(t, f; g) = \rho_z(t - t_0, z; g)$
- R2. $g(v, \tau)$ does not depend on t
- P3. Frequency shift: $s(t) = z(t) e^{j2\pi f_0 t} \Rightarrow \rho_s(t, f; g) = \rho_z(t, f - f_0; g)$
- R3. $g(v, \tau)$ does not depend on f
- P4. Time marginal: $\int \rho_z(t, f) df = z(t) z^*(t)$
- R4. $g(v, 0) = 1 \quad \forall v$
- P5. Frequency marginal: $\int \rho_z(t, f; g) dt = Z(f) Z^*(f)$
- R5. $g(0, \tau) = 1 \quad \forall \tau$
- P6. Instantaneous frequency: $\frac{\int f \rho_z(t, f; g) df}{\int \rho_z(t, f; g) df} = f_i(t)$
- R6. R4 and $\frac{\partial g(v, \tau)}{\partial \tau}|_{\tau=0} = 0 \quad \forall v$
- P7. Group delay: $\frac{\int t \rho_z(t, f; g) dt}{\int \rho_z(t, f; g) dt} = t_g(f)$
- R7. R5 and $\frac{\partial g(v, \tau)}{\partial v}|_{v=0} = 0 \quad \forall \tau$
- P8. Time support: $z(t) = 0$ for $|t| > t_c \Rightarrow \rho_z(t, f; g) = 0$ for $|t| > t_c$
- R8. $G(t, \tau) \stackrel{\Delta}{=} \int g(v, \tau) e^{-j2\pi v t} dv = 0$ for $|\tau| < 2|t|$
- P9. Frequency support: $Z(f) = 0$ for $|f| > f_c \Rightarrow \rho_z(t, f; g) = 0$ for $|f| > f_c$
- R9. $\int g(v, \tau) e^{j2\pi f \tau} d\tau = 0$ for $|v| < 2|f|$
- P10. Reduced interference
- R10. $g(v, \tau)$ is a 2D low-pass filter type
- P11. Scale: $s(t) = \frac{1}{\sqrt{|a|}} z(at) \Rightarrow \rho_s(t, f; g) = \rho_z(at, f/a; g)$
- R11. $g(v, \tau) = g(v\tau)$ is a product kernel

It is quite desirable for a TFD to possess the time and frequency support property. This property ensures that the TFD does not extend beyond the support of the signal in time or the support of its FT in frequency. The time support property (see Table 5.2.1) may be maintained for the RID by insuring that

$$G(t, \tau) = \int g(v, \tau) e^{-j2\pi v t} dv = 0 \quad \text{if } |\tau| < 2|t|. \quad (5.2.8)$$

This forms a “cone-shaped” region in (t, τ) . The WD obviously satisfies this support property since the FT of unity is an impulse function, clearly staying within the (t, τ) limits. The form of the kernel in (v, f) is also cone shaped, ensuring the frequency support property. Another approach [20] suggests a cone-shaped kernel for nonstationary signal analysis, but imposes restrictions such that time support only is ensured. The ED can be brought into the RID requirements by imposing an RID window as suggested above. The RID is not a totally new TFD since the Born-Jordan kernel, $g(v, \tau) = \text{sinc}(v\tau) = \frac{\sin(\pi v\tau)}{\pi v\tau}$, meets all of the RID requirements [19].

Table 5.2.2: Comparison of Properties for Several TFDs

| Distribution | $g(v, \tau)$ | P0 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P11 |
|------------------|--|----|----|----|----|----|----|----|----|----|----|-----|-----|
| Wigner | 1 | | x | x | x | x | x | x | x | x | x | | x |
| Rihaczek | $e^{j\pi v\tau}$ | | | x | x | x | x | x | x | x | x | | x |
| Exponential (ED) | $e^{-v^2\tau^2/2\sigma}$ | | x | x | x | x | x | x | x | | x | x | |
| Spectrogram | $A_w(v, \tau)$ | x | x | x | x | | | | | | x | | |
| Born-Jordan | $\frac{\sin(v\tau/2)}{v\tau/2}$ | | x | x | x | x | x | x | x | x | x | x | x |
| Windowed-ED | $e^{-u^2/2\sigma} * W(u) _{u=v\tau}$ | | | x | x | x | x | x | x | x | x | x | x |
| Cone (ZAM) | $g(\tau) \tau \frac{\sin(av\tau)}{av\tau}$ | | x | x | x | x | | | x | | x | | |

Notes: The BJD and the windowed ED are RIDs. $A_w(v, \tau)$ is the AF of the spectrogram window $w(t)$. $W(u)$ is the FT of $\text{rect}(t)$.

Table 5.2.1 summarizes desirable properties (P) and related kernel requirements (R) of TFDs. No known, practical TFD with a fixed kernel (not dependent on the signal or time) is able to meet all of these requirements.

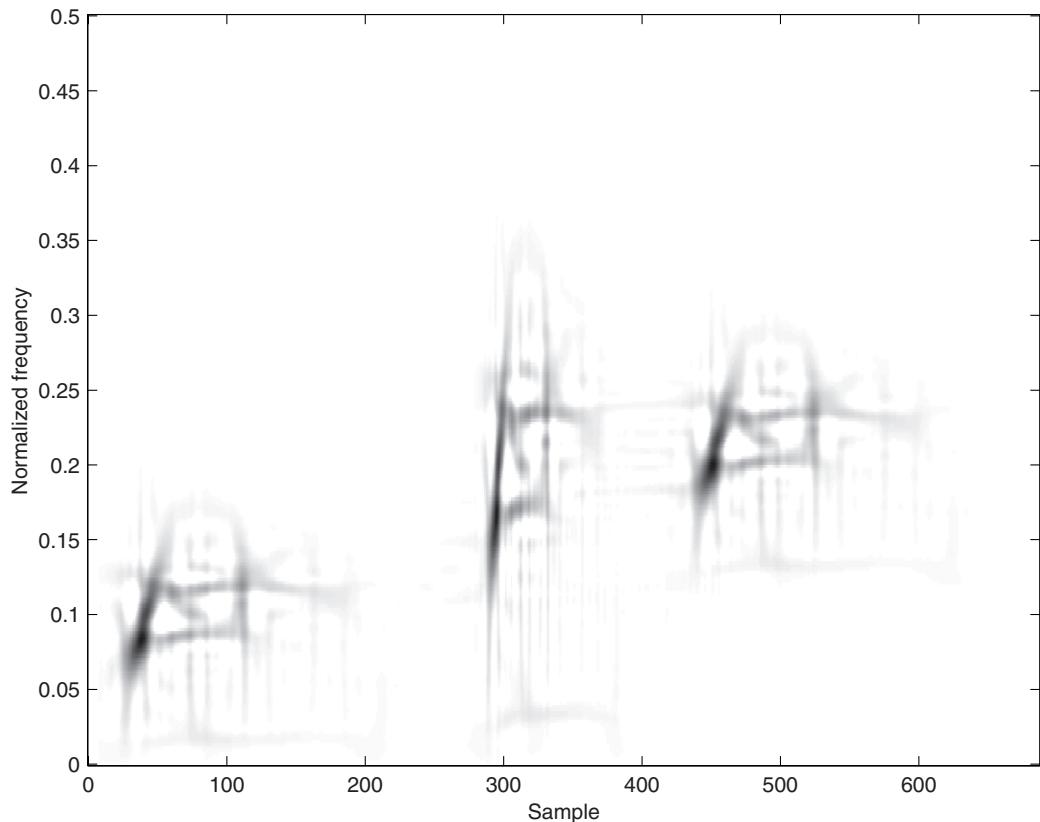
Table 5.2.2 illustrates how several common TFDs satisfy the desirable properties. Two TFDs whose kernels meet the RID requirements (Born-Jordan distribution [BJD] and windowed ED) are included. Note that the RID exhibits time-shift, frequency-shift, and scale covariance [21]. Some prefer to call this *invariance* because the TFD signature does not change in shape under these transformations. **Figure 5.2.2** shows the ED for a succession of signals consisting of a dolphin click, its 2:1 scaled and energy-corrected version, and its frequency-shifted version. This shows the time-shift, frequency-shift, and scale covariance of the ED. Any true RID would exhibit similar covariance properties.

5.2.3.1 Design procedures for effective RID kernels

Much more can be done in terms of kernel design by bringing windows and digital filters design methodology to bear in designing effective RID kernels [22]. The starting point for the kernel design is to consider a primitive function $h(t)$. This function is designed to have unit area ($\int h(t) dt = 1$) and is symmetrical; that is, $h(-t) = h(t)$. It is limited such that $h(t) = 0$ for $|t| > 1/2$, and is tapered so that it has little high-frequency content; that is, it is a low-pass impulse response. Then the kernel is

$$g(v, \tau) = H(v\tau), \quad (5.2.9)$$

where $H(v)$ is the FT of $h(t)$. It can be shown that $H(v\tau)$ satisfies the RID requirements. Desirable characteristics of $h(t)$ have been the subject of intensive study in terms of windows in time domain terms or filters in frequency domain terms. All of this knowledge can be used to select effective RID kernels. Note that in the (t, τ) domain the form of the RID kernel is $\frac{1}{|\tau|}h\left(\frac{t}{|\tau|}\right)$. In order to compute the RID, one convolves this form of the kernel with $K(t, \tau)$ along the t dimension and then FTs that result. Thus the kernel is scaled in t and normalized by $\frac{1}{|\tau|}$ for each τ . This is a wavelet-like characteristic. The RID thus exhibits time-shift, frequency-shift, and scale covariance—unlike wavelets, which commonly exhibit only scale and time-shift covariance, and many other TFDs which exhibit only time-shift and frequency-shift covariance.

**FIGURE 5.2.2**

ED with original, scaled, and frequency shifted dolphin click.

5.2.3.2 Optimum and adaptive RIDs

The RID constraints permit a considerable amount of flexibility in RID kernel design. Several TFDs have full-fledged RIDs; in other cases, some of the properties in Table 5.2.1 have been abandoned or relaxed in order to achieve a more specific effect for a narrow class of signals. Space does not permit a full discussion of these variations on the theme, but a more comprehensive discussion is available in Chapter 3 and [23]. RID kernels may also be adapted to minimize the entropy of the resulting TFD [24]. (Several entropy measures are discussed in Chapter 7.)

5.2.3.3 Discrete RID kernels

Requirements for discrete forms of the RID are similar to those of the discrete WD. The discrete RID may be formed by

$$\text{RID}_z(n, f) = \sum_{m=-\infty}^{\infty} K_z(n, m) * G(n, m) e^{-j2\pi fm}, \quad (5.2.10)$$

where

$$\gamma(m, f) = \sum_{n=-\infty}^{\infty} G(n, m) e^{-j2\pi fn} \quad (5.2.11)$$

is the discrete RID kernel. The discrete RID may thus be conveniently formed by obtaining the local autocorrelation $K_z(n, m)$, convolving it with $G(n, m)$ along n , and taking the DTFT of the result with respect to m . A fully discrete form of the kernel is more desirable. A convenient discrete RID kernel has been discovered based on the binomial distribution [18]. The form of the kernel is

$$\psi(n, m) = \frac{1}{2^{|m|}} \sum_{k=0}^{|m|} \binom{|m|}{k} \delta\left(n + \frac{|m|}{2} - k\right) \quad (5.2.12)$$

The correlation shift index m takes the values $-\infty, \dots, -1, 0, 1, \dots, \infty$, and the time shift index n takes the values $-\infty, \dots, -1, -0.5, 0, 0.5, 1, \dots, \infty$. It can be shown that the signal structure of the discrete local autocorrelation and the discrete form of the kernel can be easily formulated to include the half-integers [25].

5.2.4 COMPARISONS OF TFD RESULTS

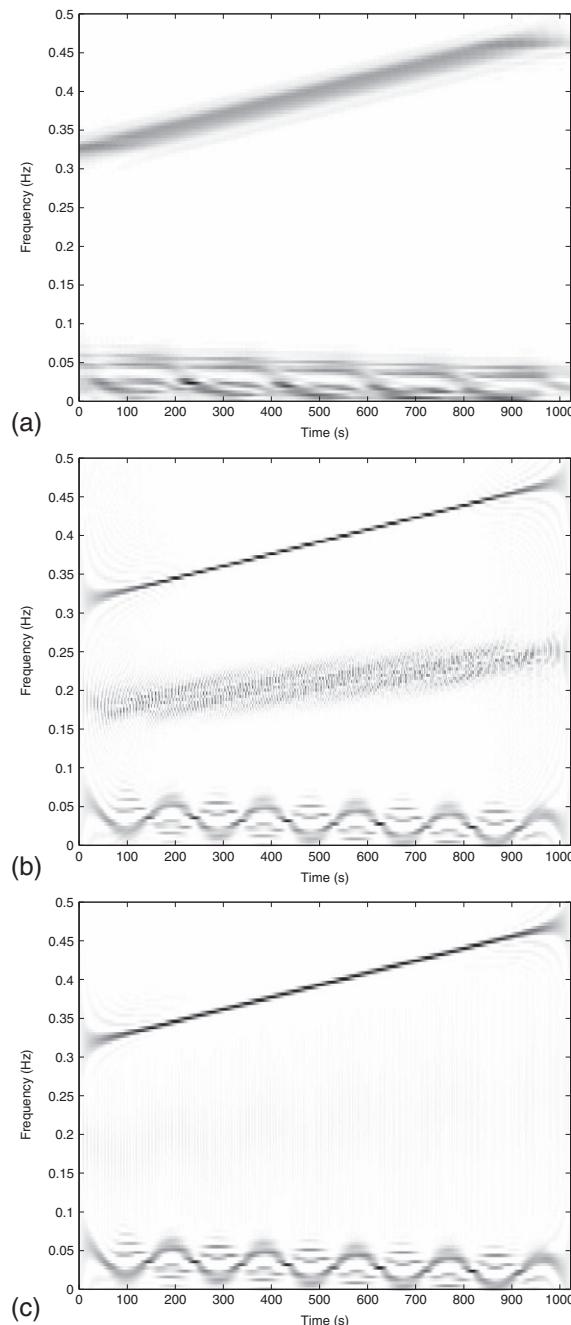
A few comparisons of TFD results are warranted. The spectrogram, the discrete WD (or WVD), and the binomial TFD just described are chosen for this purpose. A synthetic signal composed of a low-frequency, frequency-modulated sinusoid (warble) and a rising chirp serves to illustrate some of the differences between these TFDs. The alias-free form of the discrete TFD was used in each case [25]. The sampling frequency was assumed to be one sample per second.

One can readily see that the chirp is evident in the spectrogram of Fig. 5.2.3(a), but rather smeared in frequency as one might expect when using a 512-point window. The warble is a continuous smear with little evident fine structure. The WVD result is shown in Fig. 5.2.3(b).

A 512-point analysis window was used for the WVD. The resolution for the WVD is dramatically better than is the case for the spectrogram, however. One can see that the chirp is very well resolved and the frequency-modulated “warble” is also well resolved. However, there is an interlaced band of cross-term activity between the chirp and the warble. Finally, the binomial TFD result is shown in Fig. 5.2.3(c). Here, the warble and the chirp are well resolved with little evident cross-term activity between them, but there is still a bit of interlaced cross-term activity within the sinusoidal result.

5.2.5 SUMMARY AND CONCLUSIONS

The RID approach provides an advantage over the WD and the spectrogram. It may be efficiently computed, particularly in its discrete forms (see Sections 6.1 and 6.5). The RID retains almost all of the desirable properties of the WD, but with considerably reduced interference. The RID has proved to be a useful investigative tool in many applications, revealing previously unobservable (t, f) signal characteristics in many cases. Despite the constraints which guarantee the RID properties, the kernel retains a great deal of flexibility and may be tailored to specific applications with considerable benefit, as shown in other sections in this book and [27].

**FIGURE 5.2.3**

Comparison of binomial TFDs for a 512 pt. window: (a) spectrogram, (b) WVD, and (c) binomial TFD. From [26].

5.3 ADAPTIVE TIME-FREQUENCY ANALYSIS⁰

The lack of a single time-frequency representation (TFR) that is “best” for all applications has resulted in a proliferation of TFRs, each corresponding to a different, fixed mapping from signals to the time-frequency (t,f) plane (see Chapters 2 and 3 for a review of those methods). A major drawback of all fixed mappings is that, for each mapping, the resulting (t,f) representation is satisfactory only for a limited class of signals. Adaptive TFRs tune their representation to each signal to offer good performance for a large class of signals.

A natural approach to tuning is via *optimization*; this requires an objective function or performance measure expressing what we “want to see” and a class of representations to choose from. Below we overview several optimized TFRs in two broad categories: adaptive TFRs based on the linear short time Fourier transform (STFT), and those based on the quadratic WD. Proper tuning yields a high-precision representation, as we see in Fig. 5.3.1.

5.3.1 ADAPTIVE SHORT-TIME FOURIER TRANSFORMS

The simplicity, efficiency, robustness, and generally good performance of the STFT make it the initial primary tool for routine (t,f) analysis (see Section 2.3.1). However, in some cases, the spectrogram (the squared magnitude of the STFT) shows considerably inferior concentration relative to more sophisticated quadratic (t,f) representations. Studies have shown that inferior resolution in the spectrogram is primarily due to *mismatch* between the window and signal components; that is, the spectrogram performs best when the duration and orientation of the analysis window match those of the local signal components (see Section 2.3.1 and [28]). When these are not known *a priori* or when they vary between multiple components in a complex signal, an adaptive-window STFT can adjust the window parameters over time and frequency to obtain near-optimal performance for all components.

An adaptive-window STFT of the signal x takes the form

$$F_x^w(t,f) = \int x(\tau) w_{t,f}(\tau - t) e^{-j2\pi f\tau} d\tau, \quad (5.3.1)$$

where the only difference between the adaptive-window STFT and the conventional STFT is the potential variation of the window w with time and frequency. A family of unit-energy Gaussian windows of the form

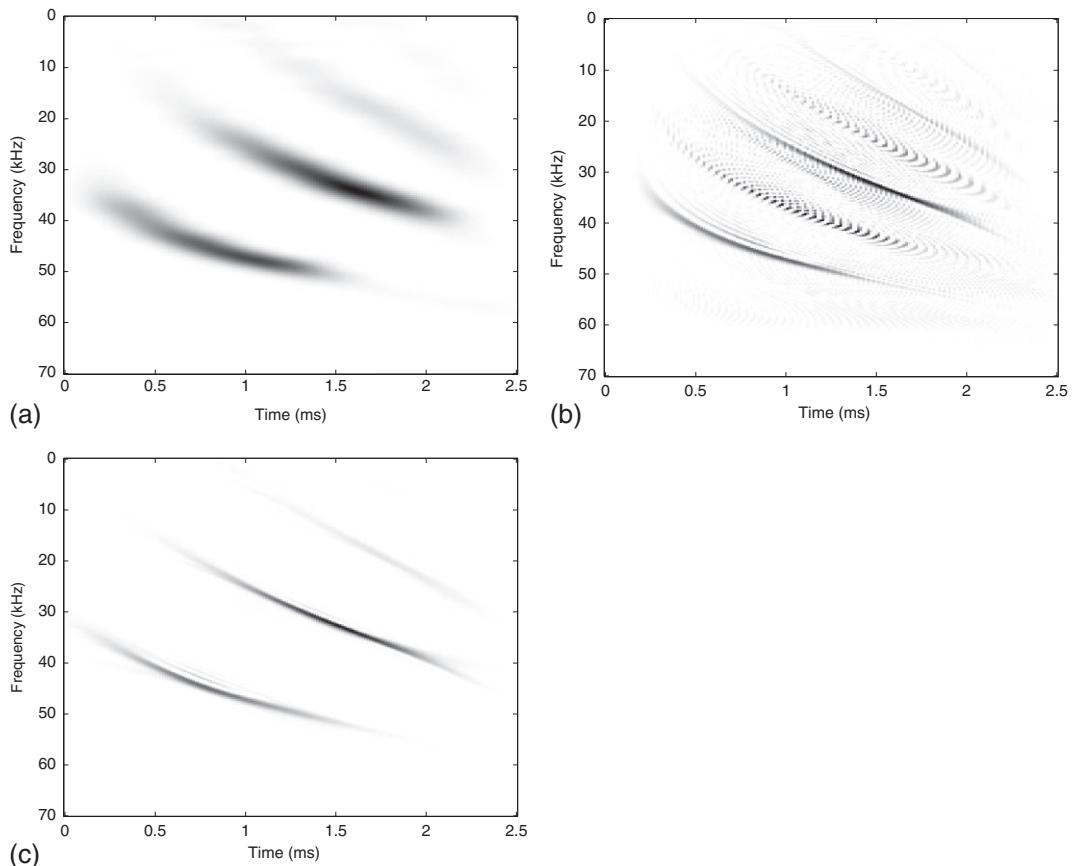
$$w_{t,f}(\tau) = (-2 \operatorname{Re}[c_{t,f}] / \pi)^{1/4} \exp \left[c_{t,f} (\tau - t)^2 \right] e^{-j2\pi f\tau} \quad (5.3.2)$$

is adopted in [29], where the real part of the window parameter $c_{t,f}$ controls the effective time-duration of the window, and the imaginary part determines the chirp rate (or (t,f) tilt).

The window parameter function $c_{t,f}$ is adapted at each time-frequency location (t,f) to maximize local concentration as measured by

$$\max_c \frac{\iint |F_x^w(u,v)\kappa(t-u,f-v)|^{2p} du dv}{\left(\iint |F_x^w(u,v)\kappa(t-u,f-v)|^p du dv \right)^2}, \quad (5.3.3)$$

⁰Authors: **R.G. Baraniuk**, Rice University, Houston, TX, USA (richb@rice.edu) and **D.L. Jones**, University of Illinois at Urbana-Champaign, Urbana, IL, USA (dl-jones@illinois.edu). Reviewers: P. Flandrin and F. Auger.

**FIGURE 5.3.1**

Time-frequency analysis of 2.5 ms of an echolocation pulse emitted by the large brown bat, *Eptesicus fuscus*. Pictured are the (a) spectrogram, (b) Wigner distribution, and (c) adaptive radially Gaussian optimal-kernel distribution (Section 5.3.2.3). The 1/0 optimal-kernel distribution looks very similar. (Thanks to Curtis Condon, Ken White, and Al Feng of the Beckman Institute of the University of Illinois for this data and for permission to use it in this publication.)

with $\kappa(t,f)$ a (t,f) window that tapers to zero away from $(t,f) = (0,0)$ and $p \geq 1$ ([29] takes $p = 2$). Maximizing this concentration measure is equivalent to maximizing several other measures of sharpness, focus, peakiness, or kurtosis or minimizing a measure of Rényi entropy [30].

A fairly efficient algorithm in [29] makes extensive use of fast Fourier transforms (FFTs) both to compute the local concentration measure and the necessary STFTs. In essence, the algorithm computes in parallel a number of STFTs with fixed windows spanning the range of acceptable window parameters, computes the local concentration metric for each window at each (t,f) location via FFT-based fast convolution, and then selects at each (t,f) location the value corresponding to the STFT with the optimal window. The computational cost is approximately two orders of magnitude greater than a single fixed-window STFT.

Support of both time and frequency adaptivity leads to a relatively costly algorithm. Often, adaptation of the window with frequency is unnecessary. For such cases, a very efficient adaptive window spectrogram can be developed [31,32]. This method adjusts one or more window parameters (most commonly, window duration) with time to best match the local signal structure. The same local concentration measure (Eq. 5.3.3) can be optimized. The efficient algorithm computes a few STFTs spanning the range of the adapting window parameter(s), computes the resulting concentration metric for each parameter, and interpolates this data to estimate the exact value of the parameter maximizing the concentration. A window with that optimal value is then applied at the current time to compute the optimal-window STFT using a single FFT.

This simplified time-adaptive-window STFT performs well on many signals and often provides nearly all of the potential benefits of adaptive (t,f) analysis at a cost only a few times that of a conventional fixed-window spectrogram. A similar procedure creates an adaptive-wavelet continuous wavelet transform [31].

An adaptive spectrogram developed for speech signals can adapt the window based on a user-provided segmentation of the signal into pitch periods.

5.3.2 ADAPTIVE QUADRATIC (t,f) REPRESENTATIONS

As shown in Chapters 2 and 3, QTFRs can be interpreted as filtered versions of the WD. Taking FTs, this corresponds to weighting the AF $A(\nu, \tau)$ of the signal by a 2D kernel function $g(\nu, \tau)$ (for the AF, see Section 5.1 and Chapter 3).

In an adaptive QTFR, we adapt the form of the kernel to match the shape of the signal's components in the ambiguity domain. In particular, we typically want g to be close to 1 in the vicinity of the signal's auto-components, and close to 0 in the vicinity of the cross-components. Again optimization is a key tool.

5.3.2.1 I/O Kernel method

Given a signal and its AF, we define the optimal 1/0 kernel as the real, non-negative function g_{opt} that solves the following optimization problem [33]:

$$\max_g \iint |A(\nu, \tau)g(\nu, \tau)|^2 d\nu d\tau \quad (5.3.4)$$

subject to

$$g(0, 0) = 1 \quad (5.3.5)$$

$$g(\nu, \tau) \text{ is radially nonincreasing} \quad (5.3.6)$$

$$\iint |g(\nu, \tau)|^2 d\nu d\tau \leq \alpha, \quad \alpha \geq 0. \quad (5.3.7)$$

The radially nonincreasing constraint (Eq. 5.3.6) can be expressed explicitly as

$$g(r_1, \psi) \geq g(r_2, \psi) \quad \forall r_1 < r_2, \quad \forall \psi, \quad (5.3.8)$$

where r_i and ψ correspond to the polar coordinates radius and angle, respectively.

The constraints (Eqs. 5.3.5–5.3.7) and performance measure (Eq. 5.3.4) are formulated so that the optimal kernel passes auto-components and suppresses cross-components. The constraints force the optimal kernel to be a low-pass filter of fixed volume α ; maximizing the performance measure encourages the passband of the kernel to lie over the auto-components (see Sections 3.2.3, 4.2, 5.1, and 5.2). Both the performance measure and the constraints are insensitive to the orientation angle and aspect ratio (scaling) of the signal components in the (v, τ) plane. Analysis in [33] shows that the optimal kernel solving Eqs. (5.3.4)–(5.3.7) takes on essentially only the values of 1 and 0.

By controlling the volume under the optimal kernel, the parameter α controls the tradeoff between cross-component suppression and smearing of the auto-components. Reasonable bounds are $1 \leq \alpha \leq 5$. At the lower bound, the optimal kernel shares the same volume as a spectrogram kernel, while at the upper bound, the optimal kernel smooths only slightly. In fact, as $\alpha \rightarrow \infty$, the optimal-kernel distribution converges to the WD of the signal.

Analyzing a signal with this optimal-kernel TFR entails a three-step procedure: (1) compute the AF of the signal; (2) solve the linear program (Eqs. 5.3.4–5.3.7) in variables $|g|^2$ (a fast algorithm is given in [34]); and (3) FT the AF-kernel product $A(v, \tau)g_{\text{opt}}(v, \tau)$ to obtain the optimized TFR $\rho_{\text{opt}}(t, f)$.

5.3.2.2 Radially Gaussian kernel method

Although the 1/0 kernel is optimal according to the criteria (Eqs. 5.3.4–5.3.7), its sharp cutoff may introduce ringing (Gibbs phenomena) into the optimized TFR, especially for small values of the volume parameter α . For an alternative, direct approach to smooth optimal kernels, explicit smoothness constraints can be appended to the kernel optimization formulation (Eqs. 5.3.4–5.3.7). In [35], the authors constrain the kernel to be Gaussian along radial profiles:

$$g(v, \tau) = \exp\left(-\frac{(vt_0)^2 + (\tau/t_0)^2}{2\sigma^2(\psi)}\right) \quad (5.3.9)$$

with t_0 a dimension parameter (typically $t_0 = 1$ s).

The term $\sigma(\psi)$ represents the dependence of the Gaussian spread on radial angle ψ . Any kernel of the form Eq. (5.3.9) is bounded and radially nonincreasing and, furthermore, smooth if the function $\sigma(\psi)$ is smooth. Since the shape of a radially Gaussian kernel is completely parameterized by this function, finding the optimal, radially Gaussian kernel for a signal is equivalent to finding the optimal function σ_{opt} for the signal. A hybrid gradient/Newton ascent algorithm solving the (nonlinear) system (Eqs. (5.3.4), (5.3.7), and (5.3.9)) is detailed in [35]. See Fig. 5.3.1 for an example on a bat chirp signal.

5.3.2.3 Adaptive optimal kernel method

While the 1/0 and radially Gaussian TFRs generally perform well, they are block-oriented techniques that design only one kernel for the entire signal. For analyzing signals with characteristics that change over time; for real-time, online operation; or for very long signals, adaptive signal-dependent TFRs are required.

Adaptation of the kernel to track the local signal characteristics over time requires that the kernel optimization procedure consider only the local signal characteristics. An ambiguity-domain design procedure such as the radially Gaussian kernel optimization technique described above does not immediately admit such time localization, since the AF includes information from all times and

frequencies in the signal. This difficulty has been surmounted, however, by the development of a time-localized, or short-time, AF [36]. Application of the radially Gaussian kernel optimization procedure to the short-time AF localized at time t_0 produces an optimal kernel $g_{\text{opt}}(\nu, \tau; t_0)$ and an optimal-kernel TFR frequency slice $\rho_{\text{opt}}(t_0, f)$ at time t_0 . Since the algorithm alters the kernel at each time to achieve optimal local performance, better tracking of signal changes results.

5.3.2.4 Adaptive cone-kernel TFR

The cone-kernel quadratic TFR reported in [20] has many desirable properties, including relatively good (t, f) resolution, good cross-term suppression, and global time-support and time-marginal preservation along with high computational efficiency. However, the performance and inner time-support behavior depend on a cone-length parameter; in addition, mismatch between this parameter and signal component duration can lead to poor resolution and cross-terms [37].

An adaptive cone-kernel distribution has been developed that adjusts the cone-length parameter over time to best match the local signal structure according to a short-time energy metric [37]. This technique in effect computes cone-kernel representations with several cone lengths in parallel and determines the optimal cone length at each time (reminiscent of [Section 5.3.1](#)). A very efficient recursive algorithm allows the parallel computations to be performed with very little overhead relative to a single cone-length TFR. The net result is an adaptive algorithm with the same order of complexity as the standard STFT. The technique has shown excellent performance on many signals.

5.3.2.5 Adaptive reduced-interference TFR

Adaptive TFRs based on other approaches have also been developed. The design of an adaptive kernel via a projection-based approach using generalized (chirp) time-frequency marginals is described in [38] (see [Section 5.2](#)). Cross-term-reduced marginals are generated from STFTs. A cross-term-reduced distribution is then reconstructed from these marginals via tomographic principles.

In [39], an “iterated projections distribution” is developed by iteratively extracting the negative part of the current TFR, estimating the cross-term from it via a signal-dependent mask function that adaptively identifies cross-term regions in the ambiguity domain, and subtracting the current cross-term estimate from the current TFR estimate. The procedure iterates the above steps until sufficient convergence is achieved. The method shows excellent (t, f) concentration and cross-term suppression in the examples shown in [39].

5.3.3 SUMMARY AND CONCLUSIONS

The drive for higher TFR performance (resolution, concentration, etc.) can be fully satisfied only by adaptive techniques. Optimal TFRs rest on a well-grounded mathematical theory, and different performance metrics or required properties result in different optimization formulations and different TFRs. Moreover, many nonlinear TFRs, such as the reassignment method ([Section 7.2](#)) and the image-processing-based adaptive TFD ([Section 5.10](#)), can be regarded as signal-dependent, optimized TFRs. The adaptive TFR concept has also been extended to random signals ([Section 12.1](#)). Finally, the same adaptivity principles can be applied to time-scale representations and continuous wavelet transforms.

5.4 POLYNOMIAL WIGNER-VILLE DISTRIBUTIONS⁰

5.4.1 POLYNOMIAL FM SIGNALS

For the complex signal

$$z(t) = a(t) e^{j\phi(t)}, \quad (5.4.1)$$

where the amplitude $a(t)$ and instantaneous phase $\phi(t)$ are real, let us define

$$f_i(t) = \frac{1}{2\pi} \phi'(t). \quad (5.4.2)$$

If $z(t)$ is analytic, $f_i(t)$ is its instantaneous frequency (IF). In this section, we use Eqs. (5.4.1) and (5.4.2) as definitions *without* assuming that $z(t)$ is analytic; if it *is* analytic, results concerning $f_i(t)$ will also be valid for the IF.

The factor $a(t)$ allows amplitude modulation, phase inversion, and time limiting. If $\phi(t)$ is a polynomial function of degree p , so that $f_i(t)$ is a polynomial of degree $p - 1$, then $z(t)$ is a *polynomial phase* or *polynomial FM* signal. If $p > 2$, then $f_i(t)$ is nonlinear, so that $z(t)$ is an example of a *nonlinear FM* signal.

Nonlinear FM signals occur both in nature and in man-made applications [40]. For example, the sonar systems of some bats use *hyperbolic* and *quadratic* FM signals for echo-location. Some radar systems use *quadratic* FM pulse compression signals. Earthquakes and underground nuclear tests may generate nonlinear FM seismic signals in some long-propagation modes. The altitude and speed of an aircraft may be estimated from the nonlinear IF of the engine noise reaching the ground. IF estimation for nonlinear FM signals is thus an important practical problem.

For a deterministic linear FM signal, the WVD gives an unbiased IF estimate. To obtain the same effect with higher-order polynomial FM signals, an extension of the WVD called the *polynomial Wigner-Ville distribution* (PWVD) was defined [41,42]. If the instantaneous phase $\phi(t)$ is a polynomial of degree p or less, the IF estimate based on a PWVD of order p is unbiased [43]. An alternative approach for estimating the IF of a nonlinear FM signal utilizes the windowed WVD, where the length of the window is estimated using the intersection of confidence interval (ICI) rule (see Section 10.2).

5.4.2 PRINCIPLES OF FORMULATION OF PWVDs

We seek a TFD

$$P_z(t,f) = \mathcal{F}_{\tau \rightarrow f} \{R_z(t,\tau)\}, \quad (5.4.3)$$

where $R_z(t,\tau)$, called the *signal kernel*,¹ depends on $z(t)$. If $a(t) = 1$, then, for the best possible representation of the IF law, we would like $P_z(t,f)$ to equal $\delta(f - f_i(t))$. Making this substitution in Eq. (5.4.3), taking inverse FTs and using Eq. (5.4.2), we find

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **G.R. Putland**, Prosper Australia, Melbourne, VIC, Australia (grpputland@gmail.com). Reviewers: B. Barkat and LJ. Stanković.

¹This notation is consistent with the convention that $R_z(t,\tau)$ (Section 2.7) is a generalization of $K_z(t,\tau)$ (Section 2.1.2). But here the generalization is in a new direction.

$$R_z(t, \tau) = e^{j2\pi f_i(t)\tau} = e^{j\phi'(t)\tau}. \quad (5.4.4)$$

So, while $z(t)$ may have a variable frequency with respect to t , we want $R_z(t, \tau)$ to have a constant frequency w.r.t. τ , namely $f_i(t)$. That is, we want the signal kernel to *dechirp* the signal, yielding a sinusoid of constant frequency, for which the FT is the optimal representation and gives a delta function at $f_i(t)$.

To estimate $\phi'(t)$ in Eq. (5.4.4), we use the central finite-difference (CFD) approximation

$$\widehat{\phi}'(t) = \frac{1}{\tau} \sum_{l=1}^q b_l [\phi(t + c_l \tau) - \phi(t - c_l \tau)], \quad (5.4.5)$$

where the dimensionless real coefficients b_l and c_l are chosen to make the approximation *exact* if $\phi(t)$ is a polynomial of degree p . Let that polynomial be

$$\phi(t) = \sum_{i=0}^p a_i t^i, \quad (5.4.6)$$

so that

$$\phi'(t) = \sum_{i=1}^p i a_i t^{i-1}. \quad (5.4.7)$$

A polynomial of degree p remains a polynomial of degree p if the argument is shifted and scaled. Shifting the time origin so that $t = 0$ in Eqs. (5.4.5) to (5.4.7), we see that if $\phi(t)$ contains only even-power terms, both $\phi'(t)$ and $\widehat{\phi}'(t)$ are zero, so that the even-power terms do not introduce any errors into the estimate. So we may assume that p is even and consider only the odd-power terms in $\phi(t)$. There are $p/2$ such terms, hence $p/2$ degrees of freedom in the coefficients a_i , suggesting that the CFD estimate can be made exact by using only $p/2$ finite differences in Eq. (5.4.5), with uniform sampling intervals. But we shall retain the generality of the sampling instants because we can do so without further algebraic complexity, and because the extra degrees of freedom are useful. With $q = p/2$, Eq. (5.4.5) becomes

$$\phi'(t) = \frac{1}{\tau} \sum_{l=1}^{p/2} b_l [\phi(t + c_l \tau) - \phi(t - c_l \tau)], \quad (5.4.8)$$

where the use of ϕ' instead of $\widehat{\phi}'$ asserts the exactness of the estimate. Substituting this into Eq. (5.4.4), and renaming the signal kernel as $R_z^{(p)}(t, \tau)$ to acknowledge the dependence on p , we obtain

$$R_z^{(p)}(t, \tau) = \prod_{l=1}^{p/2} [e^{j\phi(t+c_l\tau)} e^{-j\phi(t-c_l\tau)}]^{b_l}. \quad (5.4.9)$$

Then, substituting from Eq. (5.4.1) with $a(t) = 1$, we find

$$R_z^{(p)}(t, \tau) = \prod_{l=1}^{p/2} [z(t + c_l \tau) z^*(t - c_l \tau)]^{b_l}. \quad (5.4.10)$$

The TFD, denoted by $W_z^{(p)}(t, f)$ and given by Eq. (5.4.3) as

$$W_z^{(p)}(t, f) = \mathcal{F}_{\tau \rightarrow f} \left\{ R_z^{(p)}(t, \tau) \right\}, \quad (5.4.11)$$

is called a *polynomial Wigner distribution* (or *polynomial WD*) of order p .

Thus, we arrive at a general definition: A *polynomial WD of order p* of the signal $z(t)$ is a function $W_z^{(p)}(t, f)$ given by Eqs. (5.4.10) and (5.4.11), such that the coefficients b_l and c_l satisfy Eq. (5.4.8) when $\phi(t)$ is a polynomial of degree not exceeding p . In the special case in which $z(t)$ is analytic, the polynomial WD becomes the PWVD.

If we put $p = 2$, $c_1 = \frac{1}{2}$ and $b_1 = 1$, then $R_z^{(p)}(t, \tau)$ reduces to $z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$, which is the instantaneous autocorrelation function (IAF), denoted by $K_z(t, \tau)$. We might therefore describe $R_z^{(p)}(t, \tau)$ as a *polynomial IAF* or *higher-order IAF*.

Equation (5.4.10) shows that for any $z(t)$, the polynomial IAF is Hermitian in τ . It follows that the polynomial WD is *real*.

5.4.3 IF ESTIMATES WITH ZERO DETERMINISTIC BIAS

We now show that when the IF is a polynomial of degree not exceeding $p - 1$, the PWVD of order p gives an *unbiased* estimate of the IF law.

Theorem 5.4.1. If $z(t)$ and $f_i(t)$ are given by Eqs. (5.4.1) and (5.4.2), where $\phi(t)$ is a polynomial of degree not exceeding p , and if $W_z^{(p)}(t, f)$ satisfies the general definition of a p th-order polynomial WD of $z(t)$, then $W_z^{(p)}(t, f)$ is symmetrical in f about $f = f_i(t)$.

Proof: Substituting Eq. (5.4.1) into Eq. (5.4.10) and simplifying, we find

$$R_z^{(p)}(t, \tau) = R_a^{(p)}(t, \tau) \exp \left(j \sum_{l=1}^{p/2} b_l [\phi(t + c_l \tau) - \phi(t - c_l \tau)] \right), \quad (5.4.12)$$

where $R_a^{(p)}(t, \tau)$, as the notation suggests, is the polynomial IAF for $a(t)$, given by

$$R_a^{(p)}(t, \tau) = \prod_{l=1}^{p/2} [a(t + c_l \tau) a(t - c_l \tau)]^{b_l}. \quad (5.4.13)$$

Because $\phi(t)$ is a polynomial function of degree not exceeding p , and because $W_z^{(p)}(t, f)$ is a p th-order polynomial WD, Eq. (5.4.8) is applicable, so that Eq. (5.4.12) becomes

$$R_z^{(p)}(t, \tau) = R_a^{(p)}(t, \tau) e^{j\phi'(t)\tau} = R_a^{(p)}(t, \tau) e^{j2\pi f_i(t)\tau}. \quad (5.4.14)$$

Taking FTs ($\tau \rightarrow f$), we find

$$W_z^{(p)}(t, f) = W_a^{(p)}(t, f) * \delta(f - f_i(t)) = W_a^{(p)}(t, f - f_i(t)), \quad (5.4.15)$$

where $W_a^{(p)}(t, f)$ is the corresponding polynomial WD of $a(t)$:

$$W_a^{(p)}(t, f) = \mathcal{F}_{\tau \rightarrow f} \left\{ R_a^{(p)}(t, \tau) \right\}. \quad (5.4.16)$$

By Eq. (5.4.13), $R_a^{(p)}(t, \tau)$ is real and even in τ , so $W_a^{(p)}(t, f)$ in Eq. (5.4.16) is real and even in f . Hence by Eq. (5.4.15), $W_z^{(p)}(t, f)$ is real and symmetrical in f about $f_i(t)$. \square

Because a symmetrical distribution is symmetrical about its first moment, **Theorem 5.4.1** has the following corollary: *If $z(t)$ has polynomial phase of degree not exceeding p , and if $W_z^{(p)}(t, f)$ is a p th-order polynomial WD of $z(t)$, then the first moment of $W_z^{(p)}(t, f)$ w.r.t. f is equal to $f_i(t)$.*

Being unbiased for higher-degree polynomial FM signals, PWVDs can solve problems that quadratic TFDs cannot (e.g., unbiased estimation of the IF) [41]. PWVDs also give optimal frequency resolution in the sense that they are FTs of maximal-length polynomial IAFs derived from the full-length signal.

5.4.4 CALCULATION OF COEFFICIENTS

Applying Eq. (5.4.6) in Eq. (5.4.8) gives

$$\phi'(t) = \frac{1}{\tau} \sum_{l=1}^{p/2} \left[b_l \sum_{i=0}^p a_i \left[(t + c_l \tau)^i - (t - c_l \tau)^i \right] \right]. \quad (5.4.17)$$

Equations (5.4.7) and (5.4.17) give two expressions for $\phi'(t)$; equating them, and shifting and scaling the time variable so that $t = 0$ and $\tau = 1$, we find

$$a_1 = \sum_{l=1}^{p/2} \left[b_l \sum_{i=0}^p a_i \left[c_l^i - (-c_l)^i \right] \right] = 2 \sum_{l=1}^{p/2} \left[\sum_{i=1,3,\dots}^{p-1} a_i c_l^i \right] b_l. \quad (5.4.18)$$

On all sides of this equation, the coefficient of a_i is 0 for even values of i , justifying the decision to consider only odd values. Equating coefficients of a_i for odd i gives

$$\frac{1}{2} = \sum_{l=1}^{p/2} c_l b_l \text{ for } i = 1, \quad (5.4.19)$$

$$0 = \sum_{l=1}^{p/2} c_l^i b_l \text{ for } i = 3, 5, \dots, p-1. \quad (5.4.20)$$

The last two equations can be written in matrix form as

$$\begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_{p/2} \\ c_1^3 & c_2^3 & c_3^3 & \cdots & c_{p/2}^3 \\ c_1^5 & c_2^5 & c_3^5 & \cdots & c_{p/2}^5 \\ \vdots & \vdots & \vdots & & \vdots \\ c_1^{p-1} & c_2^{p-1} & c_3^{p-1} & \cdots & c_{p/2}^{p-1} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{p/2} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (5.4.21)$$

and solved algebraically or numerically.

Equation (5.4.21) is underdetermined: $p/2$ equations in p unknowns leave $p/2$ degrees of freedom. Various ways of exploiting the degrees of freedom give rise to various forms of the PWVD. Here, we mention two forms described in [41, p. 217].

If we want uniform sampling intervals, we choose the values of c_l and accept the resulting values of b_l , which in general are fractions. The resulting form of the PWVD is called the *fractional-powers form* or *Form I*; an example is given in [42, p. 550]. The need to compute fractional powers is a cause of inefficiency.

Alternatively, if we decide that the polynomial IAF must contain only positive integer powers, we choose the values of b_l and accept the resulting values of c_l , which in general give nonuniform sampling intervals. The resulting form of the PWVD is called the *integer-powers form* or *Form II*. In a discrete-time implementation, nonuniform sampling requires interpolation, but this is more efficient than computing noninteger powers. The number of interpolations required can be reduced by time-scaling [41, p. 218]. (There is also a “complex-time form,” or “Form III,” based on an analytic extension of the signal with a complex time argument [44].)

Notice that given one solution to Eq. (5.4.21), we can obtain another solution by changing the sign of any c_l and of the corresponding b_l . Thus, from one solution, we can always find another solution in which all the indices b_l are positive, ensuring that the polynomial IAF does not become unbounded as the signal approaches zero.

5.4.5 EXAMPLES

In the trivial case for which $p = 2$ (quadratic phase, linear FM), Eq. (5.4.21) reduces to $c_1 b_1 = 1/2$. If we fix the sampling points by choosing $c_1 = 1/2$, that leaves $b_1 = 1$. Alternatively, if we choose $b_1 = 1$ for unit powers that leaves $c_1 = 1/2$. Thus, we have a degenerate case in which Forms I and II overlap. Substituting for p , b_1 , and c_1 in Eq. (5.4.8), we obtain the simplest possible CFD approximation:

$$\phi'(t) = \frac{1}{\tau} \left[\phi \left(t + \frac{\tau}{2} \right) - \phi \left(t - \frac{\tau}{2} \right) \right]. \quad (5.4.22)$$

The same substitutions in Eq. (5.4.10) yield the polynomial IAF

$$R_z^{(2)}(t, \tau) = z \left(t + \frac{\tau}{2} \right) z^* \left(t - \frac{\tau}{2} \right). \quad (5.4.23)$$

This is just the ordinary (quadratic) IAF $K_z(t, \tau)$, which when substituted into Eq. (5.4.11) yields the ordinary (quadratic) WVD. So for $p = 2$, Theorem 5.4.1 confirms that if $z(t)$ is a quadratic-phase (i.e., linear FM) signal, the WVD is symmetrical about $f_i(t)$ and therefore gives an unbiased estimate of the IF.

If $p = 4$ (quartic phase, cubic FM), Eq. (5.4.21) reduces to the 2×2 system

$$c_1 b_1 + c_2 b_2 = 1/2; \quad c_1^3 b_1 + c_2^3 b_2 = 0. \quad (5.4.24)$$

Form I: If we take $c_1 = 1/4$ and $c_2 = -1/2$, Eq. (5.4.24) become a linear system with solutions $b_1 = 8/3$, $b_2 = 1/3$. Substituting these values into Eq. (5.4.10) gives

$$R_z^{(4)}(t, \tau) = \left[z \left(t + \frac{\tau}{4} \right) z^* \left(t - \frac{\tau}{4} \right) \right]^{8/3} \left[z \left(t - \frac{\tau}{2} \right) z^* \left(t + \frac{\tau}{2} \right) \right]^{1/3}, \quad (5.4.25)$$

where the superscript “(4)” indicates order 4.

Form II: If we take $b_1 = 2$ and $b_2 = 1$, Eq. (5.4.24) become a nonlinear system with solutions

$$\begin{aligned} c_1 &= \left[2(2 - 2^{1/3}) \right]^{-1} \approx 0.6756, \\ c_2 &= -2^{1/3}c_1 \approx -0.8512. \end{aligned} \quad (5.4.26)$$

Substituting for b_1 and b_2 in Eq. (5.4.10) gives

$$R_z^{(4)}(t, \tau) = [z(t + c_1\tau)z^*(t - c_1\tau)]^2 z(t + c_2\tau)z^*(t - c_2\tau), \quad (5.4.27)$$

where c_1 and c_2 are as in Eq. (5.4.26). This is one of an infinite number of Form II solutions. More details on the design process are given in [40,41,45] and in Section 5.5.

Figure 5.4.1 shows the effectiveness of the PWVD in suppressing artifacts caused by errors in the CFD estimate of the IF. In this case, the IF is

$$f_i(t) = f_c + f_d C_5 \left(\frac{t}{T} \right); \quad 0 \leq t \leq T, \quad (5.4.28)$$

where f_c is the center frequency, f_d is the minimax frequency deviation, and C_5 is the Chebyshev polynomial of degree 5. Trace (a) shows the ordinary WVD, which clearly cannot handle the nonlinear IF law. The Form II PWVD shown in trace (b) was computed by the “TFSA” software toolbox (described in Chapter 17). It is of 6th order, so that its signal kernel exactly dechirps the 5th-degree IF law. The superiority of the PWVD is evident, as is its symmetry about the IF law.

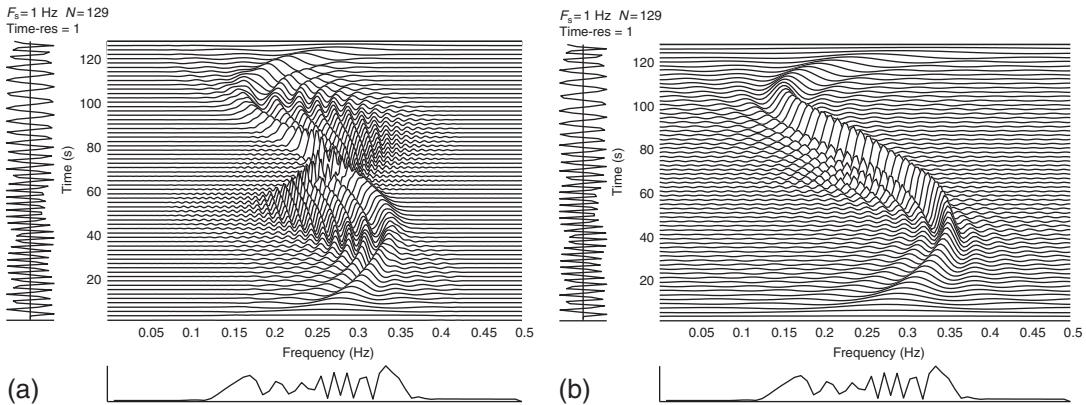


FIGURE 5.4.1

Time-frequency representations of a fifth-degree Chebyshev polynomial FM signal (Eq. 5.4.28) with center frequency $f_c = 0.25$ Hz, minimax frequency deviation $f_d = 0.1$ Hz, duration $T = 128$ s: (a) WVD; (b) sixth-order Form II PWVD. Both TFDs are unwindowed. Each plot shows time vertically (range 0-128 s; resolution 2 s) and frequency horizontally (range 0-0.5 Hz), and has the time trace at the left and the magnitude spectrum at the bottom.

5.4.6 SUMMARY AND CONCLUSIONS

A PWVD of degree p is derived from a p th-order CFD approximation to the derivative of the instantaneous phase. It reduces to the ordinary WVD if $p = 2$. It is real and symmetrical about the IF law of an FM signal whose instantaneous phase is a polynomial of degree not exceeding p (i.e., whose IF is a polynomial of degree not exceeding $p - 1$), even if that signal is also amplitude-modulated.

This topic is developed further in [Section 5.5](#). More properties of PWVDs are given in [40]. Some implementation issues are discussed in [43,45]. The effect of noise is considered in [Section 10.4](#), and Cramer-Rao lower bounds are discussed in [46].

5.5 DESIGN OF POLYNOMIAL TFDs, TIME-VARYING HIGHER ORDER SPECTRA, AND COMPLEX-LAG TFDs⁰

5.5.1 INTRODUCTION

Many applications such as radar, sonar, and communications involve nonstationary signals that can be modeled as higher-order polynomial FM signals (PFM signals) with constant or time-varying amplitude (see [Section 5.4.1](#)), where FM is short for “frequency modulation.” PFM signals are also called polynomial phase signals (PPS); Cramer-Rao lower bounds for such signals are discussed in [46]. This section is a short review of time-frequency (t,f) approaches suited to such signals.

5.5.2 FOUNDATIONS OF PTFDs

The WVD can process monocomponent PFM signals having polynomials of degree less than or equal to 2. However, many signals in radar, sonar, and communication involve nonlinear FM signals (e.g., hyperbolic FM), with the phase and frequency conveniently modeled by higher-order polynomials. A fundamental problem is to answer two questions. The first is: how can one construct a TFD that can achieve an optimal (delta function like) concentration for such signals? The Polynomial WVD (PWVD) defined in the preceding [Section 5.4](#) was introduced to answer this question for monocomponent signals [40,43], and several designs were proposed [40,47]. The kernel of this transform is based on an expression of a polynomial derivative using the principle of the finite difference as defined in the original contribution in [40,43] (see also [Section 5.4](#)). Using an exact decomposition of the derivatives of any order of a polynomial, the studies in [40,43] showed that the signal phase $\phi(t)$ can be expressed in terms of shifted versions of this polynomial, that is, $\phi(t - t_0), \dots, \phi(t - t_n)$. This decomposition can be used to design TFDs that generalize the energy concentration property of the classical WVD and ambiguity function (AF). However, for multicomponent signals, the PWVD creates even more cross-terms than the WVD. The second question is therefore: can we define methods analogous to the RIDs defined in [Chapter 3](#) that reduce these cross-terms. Several answers exist which involve time-varying higher-order spectra (TVHOS) [40,48,49].

This relationship with TVHOS was discussed in [40,43,47]. In particular, the Wigner-Ville trispectrum (WVT) was studied in detail in [40,47]. An interesting application is the analysis of signals in multiplicative noise using the PWVD, TVHOS methods [40], and methods based on higher-order

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: B. Ristic and N.A. Khan.

statistics [50]. In [51], the polynomial phase transform (PPT) was introduced for the estimation of constant amplitude PFM signals. The relationships between these TFDs, the PPT and the PWVD are discussed in [40] and Sections 10.4, 10.5, and 10.7. The estimation of the coefficients appearing in the model of a PPS affected by multiplicative and additive noise using the high-order AF is discussed in Sections 10.4, 10.5, and 10.7, including performance analysis.

5.5.3 MULTICOMPONENT SIGNALS AND PTFDs

The use of QTFDs for multicomponent signal analysis requires a reliable method of suppressing cross-terms (see Sections 3.1.2, 4.2, and 5.2). For QTFDs, the mechanisms of generation and suppression of cross-terms are well understood. In the WVD, each cross-term is generated midway between the interacting components and oscillates at a rate proportional to the separation between them (Section 4.2). Other QTFDs can be obtained from the WVD by 2D low-pass filtering, which attenuates the oscillating cross-terms.

For higher-order *polynomial* TFDs, an added difficulty is that *cross-terms are not necessarily high-pass*, so that it may not be possible to suppress all cross-terms by convolving them with a simple smoothing function in the (t,f) plane. However, if PTFDs are implemented using approaches such as those described in Section 6.2, the generation of cross-terms can be minimized.

5.5.4 IF ESTIMATION OF PFM SIGNALS USING PTFDs AND TVHOS

One main application of (t,f) methods is IF estimation. Several methods were developed by exploiting the properties of higher-order PTFDs. In particular, a cross-PWVD-based method was designed to estimate the IF of both quadratic and cubic FM signals in [52]. Other results are discussed in Sections 10.4 and 10.5.

5.5.5 COMPLEX ARGUMENT TFDs FOR POLYNOMIAL FM SIGNALS ANALYSIS

Another approach to mitigate the limitations of the WVD for the analysis of nonlinear FM signals is to first reconsider the derivation of the signal kernel given in Section 2.1. The signal kernel $K_z(t, \tau)$ uses a central finite difference approximation to estimate the phase derivative, as follows:

$$\phi'(t) \approx \widehat{\phi}'(t) = \frac{\phi(t + \frac{\tau}{2}) - \phi(t - \frac{\tau}{2})}{\tau}, \quad (5.5.1)$$

where $\widehat{\phi}'(t)$ is the estimate of the derivative of phase. Let us define the approximation error as $Q(t, \tau) = \phi'(t) - \widehat{\phi}'(t)$. For the WVD, this error can be obtained by taking the Taylor series expansion of $\phi'(t)$, that is [44]:

$$Q(t, \tau) = \phi^{(3)}(t) \frac{\tau^2}{2^2 3!} + \phi^{(5)}(t) \frac{\tau^5}{2^4 5!} + \dots \quad (5.5.2)$$

The above equation indicates that the approximation error depends only on the third-¹ and higher-order derivatives, so the WVD gives ideal estimates for linear frequency modulated (LFM) signals as the third-order phase derivative of these signals is zero. The WVD fails to concentrate the energy for nonlinear FM signals as its approximation error, that is, $Q(t, \tau)$ depends on the third-order phase derivative of the signal. This problem of the WVD can be solved to some extent by using complex lag distributions that are based on the Taylor series expansion of $\phi(t)$ using an imaginary step $j\tau$ [53]:

$$\phi(t + j\tau) = \phi(t) + j\tau\phi'(t) - \frac{\tau^2\phi''(t)}{2!} - j\frac{\tau^3\phi'''(t)}{3!} + \dots \quad (5.5.3)$$

The derivative of the phase can be estimated by using the following approximation:

$$\phi^{(1)}(t) \approx \frac{\phi(t + j\frac{\tau}{4}) - \phi(t - j\frac{\tau}{4}) + j\phi(t + \frac{\tau}{4}) - j\phi(t - \frac{\tau}{4})}{j\tau}. \quad (5.5.4)$$

This approximation leads to the following complex lag distribution kernel:

$$K_z^{\text{CL}}(t, \tau) = z^{-j} \left(t + j\frac{\tau}{4} \right) z^j \left(t - j\frac{\tau}{4} \right) z \left(t + \frac{\tau}{4} \right) z \left(t - \frac{\tau}{4} \right), \quad (5.5.5)$$

where the complex time lag, that is, $z(t + j\tau)$ is given by: $z(t + j\tau) = \int Z(f) e^{-2\pi f\tau} e^{2\pi ft} df$. The complex lag TFD can then be obtained by transforming the signal kernel, that is, $K_z^{\text{CL}}(t, \tau)$ to the frequency domain, that is:

$$\rho_z^{\text{CL}}(t, f) = \int K_z^{\text{CL}}(t, \tau) e^{-2\pi jf\tau} d\tau. \quad (5.5.6)$$

Note that the approximation error of Eq. (5.5.4) is given by:

$$Q(t, \tau) = \phi^{(5)}(t) \frac{\tau^5}{4^4 5!} + \phi^{(9)}(t) \frac{\tau^9}{4^8 9!} + \phi^{(13)}(t) \frac{\tau^{13}}{4^{12} 13!} + \dots \quad (5.5.7)$$

The approximation error for complex lag TFDs depends on the fifth-order derivatives of the phase and all its higher-order terms are twice smaller than the WVD. So, complex lag distributions achieve higher energy concentration for nonlinear FM as compared to the WVD; this makes them an effective tool for IF estimation [53].

5.5.6 MULTICOMPONENT SIGNAL ANALYSIS AND IMPLEMENTATION ISSUES

Like other high-order TFDs, complex argument TFDs generate more cross-terms than quadratic TFDs as $z(t)$ appears more than twice in the signal kernel. A reduced interference complex argument TFD can be obtained by taking the convolution of the WVD with a corrective term [54], as follows:

$$\rho_z^{\text{CL}}(t, f) = 2W_z(t, 2f) * c(t, f), \quad (5.5.8)$$

¹For convenience, for the third derivative and higher, we use the notation ⁽³⁾ shown in Eq. (5.5.2) to avoid using too many apostrophes.

where $c(t,f)$ is the corrective term defined as:

$$c(t,f) = \int z^{-j} \left(t + j \frac{\tau}{4} \right) z^j \left(t - j \frac{\tau}{4} \right) e^{-2\pi j f \tau} d\tau. \quad (5.5.9)$$

Cross-terms in the complex-lag TFD can then be reduced by estimating interference free realizations of the $W_z(t,f)$ and $c(t,f)$ followed by the computation of $\rho_z^{\text{CL}}(t,f)$ using Eq. (5.5.8). A reduced interference WVD can be estimated using any high-resolution TFD (see Section 3.3.3). Interference free corrective term can be obtained by first estimating the correction term for each component separately [54].

5.5.7 SUMMARY AND CONCLUSIONS

Decompositions of polynomial derivatives were used to design TFDs in Section 5.4 [40,43]. Applications to PFM signals affected by multiplicative and additive noise are discussed in Sections 10.4 and 10.5. Extending the fundamental introduction to PWVDs given in Section 5.4 [41], additional results and derivations can be found in [40,43]. In addition to PTFDs, complex argument TFDs can also be used to achieve high energy concentration for nonlinear FM signals.

5.6 TIME-FREQUENCY REPRESENTATIONS COVARIANT TO SPECTRAL DELAY SHIFTS⁰

5.6.1 SPECTRAL DELAY SHIFT COVARIANCE PROPERTY

As discussed in Chapter 3, no single quadratic time-frequency representation (QTFR) exists that can be used effectively in all possible applications. Each QTFR is best suited for analyzing signals with specific types of properties and time-frequency (t,f) structures. To select the appropriate analysis tool, QTFRs are often classified based on the various properties they satisfy including covariance properties. A QTFR satisfies a covariance property if it preserves, or is covariant to, certain (t,f) changes on the signal. For example, for a signal $x(t)$ with FT $X(f)$, the QTFRs class $\rho_X(t,f)$ (with signal-independent kernels; defined on p. 113) are covariant to constant time shifts (see Chapter 3). Specifically,

$$Y(f) = (\mathcal{Y}_\tau X)(f) = e^{-j2\pi \tau f} X(f) \Rightarrow \rho_Y(t,f) = \rho_X(t - \tau, f), \quad (5.6.1)$$

where \mathcal{Y}_τ is the constant-time-shift operator. (Table 5.6.1 summarizes various operators used in this section and their effect on signals.) Constant time shifts are important, for example, in shallow-water sonar signal processing. Due to boundary interactions, a bottom-bounce path may be received several milliseconds after the direct path. Thus, a QTFR analyzing the received signal must preserve the delay associated with the difference in path lengths. The class of QTFRs also preserve constant frequency shifts: $\rho_Y(t,f) = \rho_X(t,f - v)$ when $Y(f) = (\mathcal{Q}_v X)(f) = X(f - v)$. On the other hand, affine-class QTFRs $\rho_X^{(A)}(t,f)$ preserve constant time shifts as in Eq. (5.6.1) and scale changes (dilations) on the signal; that is, $\rho_Y^{(A)}(t,f) = \rho_X^{(A)}(at, f/a)$ when $Y(f) = (\mathcal{C}_a X)(f) = X(f/a)/\sqrt{|a|}$ (see [55] and Section 7.1).

⁰Author: A. Papandreou-Suppappola, Arizona State University, Tempe, AZ, USA (papandreou@asu.edu). Reviewers: P. Flandrin and D.L. Jones.

Table 5.6.1: Operators With Their Effects on a Signal $x(t)$

| Operator Name | Operator | Effect of Operator |
|--------------------------|---|--|
| Identity | \mathcal{I} | $X(f), f \in \Re$ |
| Scaling | \mathcal{C}_a | $\frac{1}{\sqrt{ a }} X\left(\frac{f}{a}\right)$ |
| Constant frequency shift | \mathcal{Q}_v | $X(f - v)$ |
| Constant time shift | \mathcal{Y}_τ | $e^{-j2\pi\tau f} X(f)$ |
| Hyperbolic time shift | $\mathcal{D}_c^{(\Lambda)}, \Lambda(b) = \ln b$ | $e^{-j2\pi c \ln(f/f_r)} X(f), f \in \Re^+$ |
| Power time shift | $\mathcal{D}_c^{(\Lambda)}, \Lambda(b) = \text{sgn}(b) b ^\kappa$ | $e^{-j2\pi c \text{sgn}(f) f/f_r ^\kappa} X(f), f \in \Re$ |
| Exponential time shift | $\mathcal{D}_c^{(\Lambda)}, \Lambda(b) = e^{kb}$ | $e^{-j2\pi c e^{kf/f_r}} X(f), f \in \Re$ |
| Spectral delay shift | $\mathcal{D}_c^{(\Lambda)}$ | $e^{-j2\pi c \Lambda(f/f_r)} X(f), f \in \wp$ |
| Hyperbolic warping | $\mathcal{U}_\Lambda, \Lambda(b) = \ln b$ | $e^{\frac{f}{2f_r}} X(f_r e^{f/f_r}), f \in \Re$ |
| Power warping | $\mathcal{U}_\Lambda, \Lambda(b) = \text{sgn}(b) b ^\kappa$ | $\frac{1}{\sqrt{ \kappa }} \left \frac{f}{f_r} \right ^{\frac{1-\kappa}{2\kappa}} X\left(f_r \text{sgn}(f) \left \frac{f}{f_r} \right ^{\frac{1}{\kappa}}\right), f \in \Re$ |
| Exponential warping | $\mathcal{U}_\Lambda, \Lambda(b) = e^{kb}$ | $\sqrt{\frac{f_r}{ \kappa f}} X\left(\frac{f_r}{\kappa} \ln \frac{f}{f_r}\right), f \in \Re^+$ |
| Dispersive warping | \mathcal{U}_Λ | $\left \Lambda' \left(\Lambda^{-1} \left(\frac{f}{f_r} \right) \right) \right ^{-1/2} X\left(f_r \Lambda^{-1} \left(\frac{f}{f_r} \right)\right), f \in \Re$ |

Notes: Some operators depend on a differentiable one-to-one function $\Lambda(b)$. Here, $\Lambda'(b) = \frac{d}{db} \Lambda(b)$, $\Lambda^{-1}(\Lambda(b)) = b$, $\text{sgn}(b)$ yields the sign (± 1) of b , and $f_r > 0$ is a reference frequency. Also, \wp and \Re are the domain and range, respectively, of $\Lambda(\cdot)$.

In some applications, it is important to preserve signal time shifts caused by the signal propagating through systems with dispersive (t, f) characteristics or, equivalently, with nonlinear spectral delay. (The spectral (or group) delay of a signal $x(t)$, with FT $X(f) = a(f) e^{-j2\pi\vartheta(f/f_r)}$, is defined in Chapter 1, p. 35 as the derivative of the phase modulation $\vartheta(f/f_r)$ of the signal; that is, $r(f) = \frac{d}{df} \vartheta(f/f_r)$. Here, $a(f) \geq 0$ is the signal's amplitude modulation (AM), and $f_r > 0$ is a reference frequency.) A dispersive system is one which delays, in time, different frequencies by different amounts. For example, in underwater acoustics, backscattering from immersed targets such as spherical shells may result in dispersive waves, thus leading to echoes with frequency-dependent spectral delay. If a QTFR preserves times shifts by an amount equal to the change in spectral delay in a signal or system, then the corresponding QTFR property is referred to as *spectral delay shift (SDS) covariance* (also called generalized time shift in [56]). If $X(f)$ is passed through an allpass dispersive system with output $Y(f) = e^{-j2\pi\Lambda(f/f_r)} X(f)$, the change in spectral delay, $\tau_d(f) = \frac{d}{df} \Lambda(f/f_r)$, is proportional to the derivative of the phase function $\Lambda(f/f_r)$. Since spectral delay is a measure of the time delay introduced in each sinusoidal component of the signal at frequency f , the ideal QTFR $\rho_Y(t, f)$ should preserve this frequency-dependent, spectral delay change $\tau(f)$ [56].

Property definition. The effect of the SDS operator¹ on a signal $x(t)$ is given by

$$Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \Lambda\left(\frac{f}{f_r}\right)} X(f), \quad f \in \wp, \quad (5.6.2)$$

where \wp is the domain of $\Lambda(\cdot)$ and $c \in \Re$ (see [56,57] and Section 15.3). A QTFR ρ is SDS covariant if the QTFR of the output, ρ_Y , corresponds to the QTFR of the input, ρ_X , shifted in time by an amount equal to the change in spectral delay, $c\tau_d(f) = c \frac{d}{df} \Lambda\left(\frac{f}{f_r}\right)$, that is introduced in Eq. (5.6.2). For $f \in \wp$, the property states

$$Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \Lambda\left(\frac{f}{f_r}\right)} X(f) \Rightarrow \rho_Y(t, f) = \rho_X(t - c\tau_d(f), f), \quad (5.6.3)$$

when $\Lambda(b)$ is a differentiable one-to-one function. Figure 5.6.1 demonstrates the QTFR transformation in Eq. (5.6.3) due to the signal transformation in Eq. (5.6.2) that results in a nonlinear change in spectral delay. The parameter c expresses the amount of dispersion or nonlinear time modulation on the signal as shown in Fig. 5.6.2. The SDS operator $\mathcal{D}_c^{(\Lambda)}$ in Eq. (5.6.2) is unitarily equivalent to the constant time-shift operator \mathcal{Y}_{c/f_r} in Eq. (5.6.1) (see [56,57] and Sections 4.5 and 15.3):

$$\mathcal{D}_c^{(\Lambda)} = \mathcal{U}_\Lambda^{-1} \mathcal{Y}_{c/f_r} \mathcal{U}_\Lambda, \quad (5.6.4)$$

where $(\mathcal{U}_\Lambda^{-1}(\mathcal{U}_\Lambda X))(f) = X(f)$, and the dispersive warping operator is given by [56]

$$Z(f) = (\mathcal{U}_\Lambda X)(f) = \left| \Lambda' \left(\Lambda^{-1}(f/f_r) \right) \right|^{-1/2} X \left(f_r \Lambda^{-1}(f/f_r) \right), \quad f \in \aleph. \quad (5.6.5)$$

Here, \aleph denotes the range of $\Lambda(\cdot)$. The unitary warping operator [56] in Eq. (5.6.5) preserves inner products; that is, $\int_{f \in \aleph} (\mathcal{U}_\Lambda X)(f) (\mathcal{U}_\Lambda X)^*(f) df = \int_{f \in \wp} X(f) X^*(f) df$. Note that since the SDS covariance in Eq. (5.6.3) follows directly from the constant-time-shift covariance in Eq. (5.6.1) via Eq. (5.6.4), it

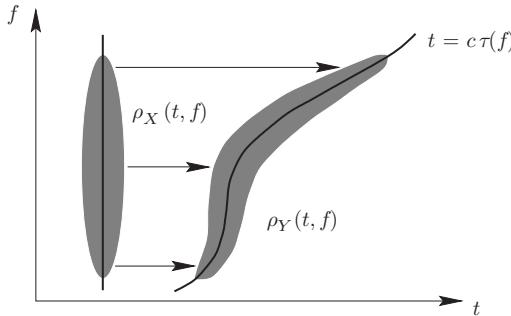
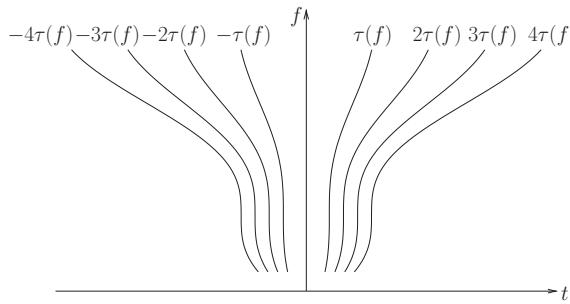


FIGURE 5.6.1

A QTFR preserves a dispersive change in a signal's spectral delay $c\tau(f)$ as the QTFR is shifted along the time axis by an amount equal to $c\tau(f)$. Here, $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f)$ as in Eq. (5.6.2).

¹One could also define the *dual* to this operator, that is $(\mathcal{D}_c^{(\xi)} x)(t) = e^{j2\pi c \xi(t/t_r)} x(t)$, $t_r > 0$, to represent time-varying IF shifts (see Section 9.2).

**FIGURE 5.6.2**

Various spectral delay curves $t = c\tau(f)$ corresponding to positive and negative changes in the parameter c in Eq. (5.6.2).

experiences the same problems as the constant-time-shift covariance when the time origin is not known. An example of such a case is the estimation of the range of a target in an active sonar application.

Specific examples of the SDS-covariance property.

The SDS-covariance property in Eq. (5.6.3) simplifies to a *particular* covariance property (satisfied by a different class of QTFRs) when the differentiable one-to-one function $\Lambda(b)$ and the time shift $\tau(f) = \frac{d}{df} \Lambda(\frac{f}{f_r})$ are fixed. Depending on the choice of $\Lambda(b)$, the time shift $\tau(f)$ may be constant, linear, or nonlinear (dispersive). For example,

- for $\Lambda(b) = b$ (with $\wp = \aleph = \aleph$), the SDS covariance in Eq. (5.6.3) simplifies to the constant time shift, $\tau(f) = \frac{1}{f_r}$, covariance in Eq. (5.6.1) of the quadratic class or the affine class [55].
- for $\Lambda(b) = \ln b$, $b > 0$ (with $\wp = \aleph_+$ and $\aleph = \aleph$), the SDS covariance in Eq. (5.6.3) simplifies to the (dispersive) hyperbolic time shift, $\tau(f) = \frac{1}{f}$, $f > 0$, covariance of the hyperbolic class [58].
- for $\Lambda(b) = \text{sgn}(b)|b|^\kappa$, $\kappa \neq 0$ (with $\wp = \aleph = \aleph$), the covariance in Eq. (5.6.3) simplifies to the (possibly dispersive) κ th power time shift, $\tau(f) = \frac{d}{df} \Lambda(\frac{f}{f_r}) = \frac{\kappa}{f_r} \left| \frac{f}{f_r} \right|^{\kappa-1}$, covariance of the κ th power class (see [57] and Section 15.3). The SDS is constant for $\kappa = 1$, and linearly dependent on frequency for $\kappa = 2$.
- for $\Lambda(b) = e^{\kappa b}$ (with $\wp = \aleph$ and $\aleph = \aleph_+$), the covariance in Eq. (5.6.3) simplifies to the (dispersive) κ th exponential time shift, $\tau(f) = \frac{\kappa}{f_r} e^{\kappa f/f_r}$, covariance of the κ th exponential class [56].

Matched signal analysis. For effective (t, f) analysis, one needs to match the SDS of a QTFR in Eq. (5.6.3) with the signal's spectral delay. Thus, QTFRs that satisfy Eq. (5.6.3) are ideally suited to analyze signals with spectral delay equal to that SDS. The (t, f) geometry underlying the SDS-covariance property is related to the *generalized impulse* function

$$I_c^{(\vartheta)}(f) = \sqrt{|r(f)|} e^{-j2\pi c\vartheta\left(\frac{f}{f_r}\right)}, \quad f \in \wp, \quad (5.6.6)$$

where \wp is the domain of the phase $\vartheta(\frac{f}{f_r})$. The spectral delay $r(f) = \frac{d}{df} \vartheta(\frac{f}{f_r})$ reflects the dispersion characteristics of the class which is covariant to the SDS $\tau(f)$ in Eq. (5.6.3) only when $r(f) = \tau(f)$ (or, equivalently, when $\vartheta(b) = \Lambda(b)$ in Eq. (5.6.3)).

5.6.2 CLASSES OF SDS-COVARIANT QTFRs

As shown in Eq. (5.6.4), the SDS operator in Eq. (5.6.2) can be acquired by unitarily warping the time-shift operator in Eq. (5.6.1) [56] (see Section 4.5). Thus, SDS-covariant QTFRs can be obtained by appropriately warping dispersively constant-time-shift covariant QTFRs, such as quadratic-class or affine-class QTFRs, using

$$\rho_X^{(D\text{class})}(t,f) = \rho_Z^{(\text{class})} \left(\frac{t}{f_r \tau_d(f)}, f_r \Lambda \left(\frac{f}{f_r} \right) \right), \quad f \in \wp, \quad (5.6.7)$$

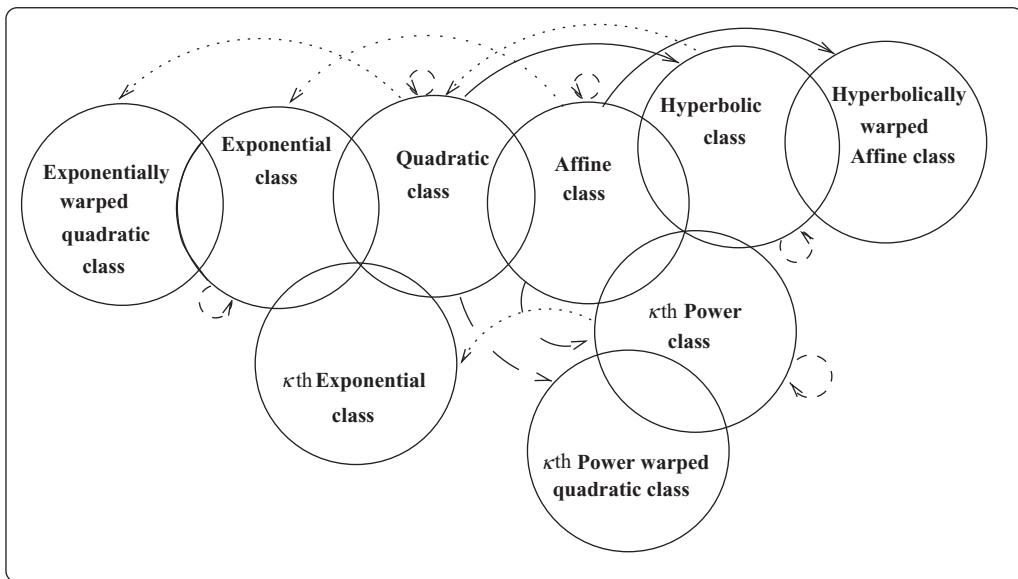
where \wp is the domain of $\Lambda \left(\frac{f}{f_r} \right)$ and $Z(f) = (\mathcal{U}_\Lambda X)(f)$ is the dispersively warped signal in Eq. (5.6.5). The superscript “class” indicates which QTFR class undergoes the warping in Eq. (5.6.7), and D class indicates the resulting QTFR class with possible dispersive characteristics. For example, with class = Q , $\rho^{(Q)}$ is a quadratic class of TFRs that can be warped to yield a new QTFR $\rho^{(DQ)}$ (see below). A more specific example follows. Consider the WD, defined in Section 2.1.4.2 as $W_Z(t,f) = \int_{-\infty}^{\infty} Z(f + \frac{v}{2}) Z^*(f - \frac{v}{2}) e^{j2\pi tv} dv$, which is a quadratic TFR. By warping the WD as in Eq. (5.6.7), one obtains the SDS-covariant version of the WD or frequency dispersively warped WD (DWD) [56]

$$W_X^{(\Lambda)}(t,f) = W_Z \left(\frac{t}{f_r \tau(f)}, f_r \Lambda \left(\frac{f}{f_r} \right) \right), \quad f \in \wp, \quad (5.6.8)$$

which is a specific QTFR example of $\rho^{(DQ)}$. For a fixed $\Lambda(b)$, this transformation can be achieved in three steps. First, the signal is warped as $Z(f) = (\mathcal{U}_\Lambda X)(f)$, and then the WD of the warped signal $Z(f)$ is computed. Lastly, the (t,f) axes are transformed for correct (t,f) localization using $t \rightarrow t/(f_r \tau(f))$ and $f \rightarrow f_r \Lambda(f/f_r)$. Note that this method ensures that the SDS-covariant QTFRs, including the DWD, always satisfy Eq. (5.6.3) for a given differentiable one-to-one function $\Lambda(b)$.

SDS-covariant QTFRs from the quadratic class. The SDS-covariant Cohen’s-class QTFRs, $\rho_X^{(DQ)}(t,f)$, are obtained by dispersively warping QTFRs, $\rho_X(t,f)$, using Eq. (5.6.7) with class = C and a warping function $\Lambda(b)$ chosen to give the desired SDS covariance in Eq. (5.6.3). Note that in p. 113 QTFRs are given by $\rho_X(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma_\rho(t - \hat{t}, f - \hat{f}) W_X(\hat{t}, \hat{f}) d\hat{t} d\hat{f}$, where $W_X(t,f)$ is the WD. The kernel $\gamma_\rho(t,f)$ uniquely characterizes the QTFR ρ as well as its corresponding warped version $\rho^{(DQ)}$ in Eq. (5.6.7) [56]. Due to the warping in Eq. (5.6.5), the constant time-shift and frequency-shift covariances that all QTFRs satisfy are transformed into the SDS and warped frequency-shift covariances, respectively, that all DQ QTFRs satisfy [56]. An important member of the DQ class that satisfies many desirable properties is the DWD, $W_X^{(\Lambda)}(t,f)$, in Eq. (5.6.8). For example, it satisfies a specific set of marginal properties and provides a highly concentrated representation for the generalized impulse $I_c^{(\vartheta)}(f)$ provided $\vartheta(b) = \Lambda(b)$ in Eq. (5.6.6). Specifically, for $X(f) = I_c^{(\vartheta)}(f)$, $W_X^{(\Lambda)}(t,f) = |\tau(f)|\delta(t - c\tau(f))$. For other QTFR members, see [56].

Different DQ classes can be obtained, suitable in different applications, simply by choosing a differentiable one-to-one function $\Lambda(b)$ to match a signal’s (t,f) characteristics. Once $\Lambda(b)$ is fixed, it can be replaced in Eq. (5.6.7) to obtain the corresponding formulation of the specified class; in Eq. (5.6.8) to obtain the DWD; and in Eq. (5.6.3) to obtain a fixed SDS (possibly dispersive) covariance property. Some class examples obtained by choosing the function $\Lambda(b)$ are summarized below, and are also pictorially demonstrated in Fig. 5.6.3.

**FIGURE 5.6.3**

Some SDS-covariant QTFR classes. An arrow points from a QTFR class being warped to a new QTFR class. Different types of warping are demonstrated depending on $\Lambda(b)$ in Eq. (5.6.7). They include hyperbolic warping with $\Lambda(b) = \ln b$ (—), power warping with $\Lambda(b) = \text{sgn}(b)|b|^\kappa$ (— — —), or exponential warping with $\Lambda(b) = e^{\kappa b}$ (---). The warping (---) maps a QTFR class back to itself (“self-mapping”). For example, the hyperbolic class in Eq. (5.6.9) is obtained by warping the quadratic class using the hyperbolic warping with $\Lambda(b) = \ln b$ (—) in Eq. (5.6.7).

Quadratic class. When $\Lambda(b) = b$ and $\tau(f) = 1/f_r$, the warping in Eq. (5.6.7) with $\varphi = \mathfrak{R}$ maps the quadratic class back to itself since $(\mathcal{U}_\Lambda X)(f) = (\mathcal{I}X)(f) = X(f)$ in Eq. (5.6.5) becomes the (nondispersive) identity operator. The SDS-covariance property in Eq. (5.6.3) simplifies to the constant time-shift covariance in Eq. (5.6.1).

Hyperbolic class. When $\Lambda(b) = \ln b$ and $\tau_d(f) = 1/f$, $f > 0$, the DQ class in Eq. (5.6.7) (with $\varphi = \mathfrak{R}_+$ and class = Q) is the hyperbolic class [58]. Thus, any hyperbolic-class QTFR, $\rho_X^{(H)}(t,f)$, $f > 0$, can be obtained by warping a corresponding QTFR, $\rho_X(t,f)$, using the transformation

$$\rho_X^{(H)}(t,f) = \rho_Z \left(\frac{tf}{f_r}, f_r \ln \frac{f}{f_r} \right), \quad (5.6.9)$$

where $Z(f) = (\mathcal{U}_\Lambda X)(f) = e^{\frac{f}{2f_r}} X \left(f_r e^{\frac{f}{f_r}} \right)$ in Eq. (5.6.5). The SDS covariance in Eq. (5.6.3) simplifies to the hyperbolic time-shift covariance $\rho_Y^{(H)}(t,f) = \rho_X^{(H)}(t - c/f, f)$, where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \ln \frac{f}{f_r}} X(f)$. The warped frequency-shift covariance simplifies to scale covariance [56]. These two covariance properties defining the hyperbolic class are important for the analysis of Doppler-invariant signals similar to the signals used by bats for echolocation, and for the analysis of self-similar random

processes. Members of the hyperbolic class include the Altes Q-distribution [6] (the DWD in Eq. (5.6.8) with $\Lambda(b) = \ln b$), the unitary Bertrand P_0 -distribution (see Section 7.1), and the hyperbologram [58]. Other class examples include the power-warped quadratic class [56] for which $\Lambda(b) = \text{sgn}(b)|b|^\kappa$ in Eq. (5.6.7), and the exponentially warped quadratic class [56] for which $\Lambda(b) = e^b$ in Eq. (5.6.7) (see Fig. 5.6.3).

SDS-covariant QTFRs from the affine class. The SDS-covariant affine-class QTFRs, $\rho_X^{(\text{DA})}(t,f)$, are obtained by warping the affine-class QTFRs $\rho_X^{(A)}(t,f)$ using Eq. (5.6.7) with class = A. The affine-class QTFRs are given by (see Section 7.1)

$$\rho_X^{(A)}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_\rho(f(t - \hat{t}), -\hat{f}/f) W_X(\hat{t}, \hat{f}) d\hat{t} d\hat{f},$$

where $W_X(t,f)$ is the WD and the kernel $\psi_\rho(a,b)$ uniquely characterizes $\rho^{(A)}$ and its corresponding warped version $\rho^{(\text{DA})}$. The warping transforms the two covariance properties defining the affine class, the constant-time-shift covariance and scale covariance, into two new covariance properties. These properties, which define the DA class, are the SDS covariance in Eq. (5.6.3), and the warped scale covariance [56]. The latter property simplifies to known covariance properties based on $\Lambda(b)$ as shown below. The DWD in Eq. (5.6.8) is a member of both the DC and the DA classes since the WD is a member of both quadratic class and the affine class.

Various QTFR classes useful in different applications are obtained by fixing $\Lambda(b)$ in Eq. (5.6.7). Some examples are listed below and are also demonstrated in Fig. 5.6.3.

Affine class. The affine class is an example of a “self-mapping” since, when $\Lambda(b) = b$ and $\tau(f) = 1/f_r$, the warping in Eq. (5.6.7) maps the affine class back to itself. Some important affine QTFRs include the WD, the Bertrand P_κ -distributions (see Section 7.1), and the scalogram [55].

κ th power class. When $\Lambda(b) = \text{sgn}(b)|b|^\kappa$ and $\tau(f) = \frac{\kappa}{f_r} \left| \frac{f}{f_r} \right|^{\kappa-1}$, $\kappa \neq 0$, the DA in Eq. (5.6.7) (with $\wp = \Re$ and class = A) is the κ th power class (see [56,57] and Section 15.3). Different power classes are obtained by varying κ , and the affine class is obtained when $\kappa = 1$. The κ th power QTFRs, $\rho_X^{(\kappa P)}$, are obtained by warping corresponding members of the affine class (cf. Eq. 5.6.7)

$$\rho_X^{(\kappa P)}(t,f) = \rho_Z^{(A)} \left((t/\kappa) |f/f_r|^{1-\kappa}, f_r \text{sgn}(f) |f/f_r|^\kappa \right),$$

where $Z(f) = (\mathcal{U}_\Lambda X)(f) = |\kappa|^{-1/2} |f/f_r|^{1-\kappa} X \left(f_r \text{sgn}(f) |f/f_r|^\frac{1}{\kappa} \right)$ in Eq. (5.6.5). The SDS covariance in Eq. (5.6.3) simplifies to the κ th power time-shift covariance property $\rho_Y^{(\kappa P)}(t,f) = \rho_X^{(\kappa P)} \left(t - c \frac{\kappa}{f_r} \left| \frac{f}{f_r} \right|^{\kappa-1}, f \right)$ where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \text{sgn}(f) \frac{f}{f_r}} X(f)$. The warped scale covariance simplifies to scale covariance [56]. These two properties defining the κ th power class are important for analyzing signals propagating through systems with power-law dispersive (t,f) characteristics. Members of the κ th power class include the power WD, the Bertrand P_κ -distributions (Section 7.1), and the powergram ([57] and Section 15.3).

κ th Exponential class. The κ th exponential class QTFRs, $\rho^{(\kappa E)}$, [56] are obtained when $\Lambda(b) = e^{\kappa b}$ and $\tau(f) = \frac{\kappa}{f_r} e^{\kappa f/f_r}$, $\kappa \neq 0$, in the DA formulation in Eq. (5.6.7) (with $\wp = \Re$ and class = A). Any such QTFRs can be written as warped affine QTFRs, $\rho_X^{(\kappa E)}(t,f) = \rho_Z^{(A)} \left(\frac{1}{\kappa} t e^{-\kappa f/f_r}, f_r e^{\kappa f/f_r} \right)$, where

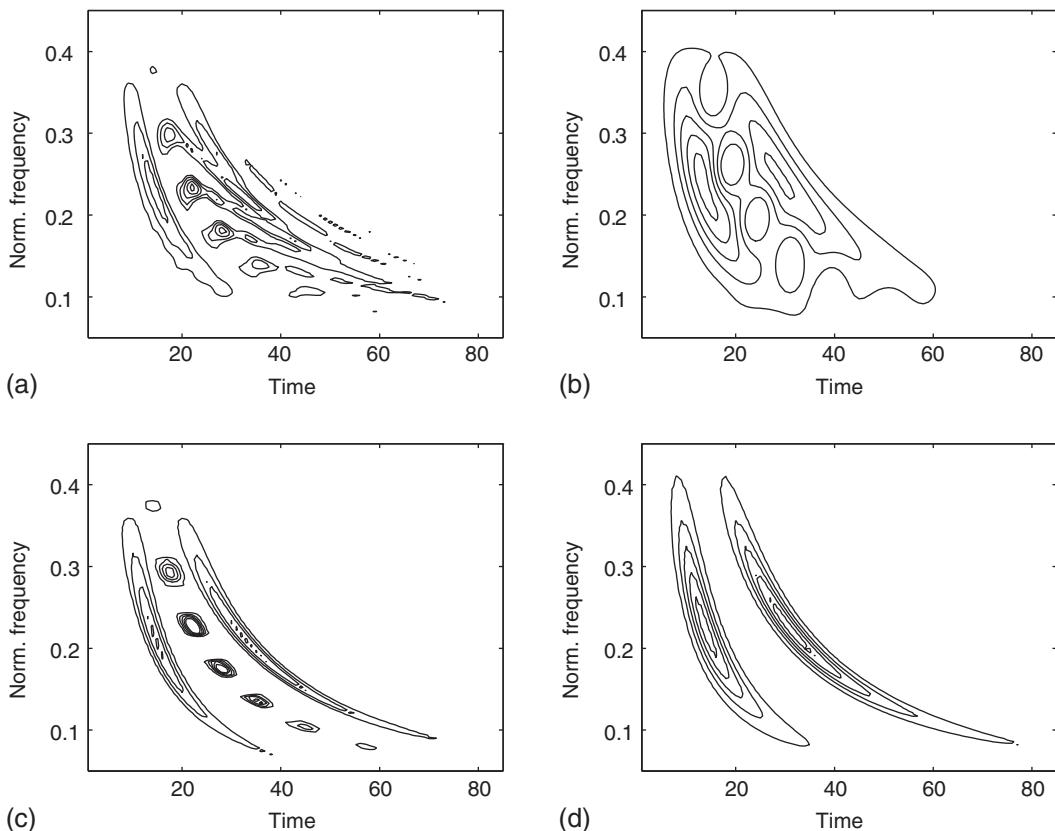
$Z(f) = (\mathcal{U}_\Lambda X)(f) = \sqrt{\frac{f_r}{|\kappa|f}} X(\frac{f_r}{\kappa} \ln \frac{f}{f_r})$ in Eq. (5.6.5). These QTFRs satisfy the κ th exponential SDS covariance given as $\rho_Y^{(\kappa E)}(t, f) = \rho_X^{(\kappa E)} \left(t - c \frac{\kappa}{f_r} e^{\kappa f/f_r}, f \right)$, where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c e^{\kappa f/f_r}} X(f)$. This is an important property for analyzing signals passing through exponentially dispersive systems. The warped scale covariance simplifies to constant-frequency-shift covariance [56]. Class members include the κ th exponential WD (the DWD in Eq. (5.6.8) with $\Lambda(b) = e^{\kappa b}$), and the κ th Bertrand P_0 -distribution [56]. When $\kappa = 1$, the κ th exponential class yields the exponential class [56] in Fig. 5.6.3. Another DA class example is the hyperbolically warped affine class [56] obtained when $\Lambda(b) = \ln b$ in Eq. (5.6.7) (see Fig. 5.6.3).

5.6.3 SIMULATION EXAMPLE

In real-life applications, some preprocessing or *a priori* knowledge is necessary to identify the (t, f) characteristics of the signal so as to match them to the SDS covariance of the QTFR. For example, signals with linear (t, f) characteristics, such as Dirac impulses $X(f) = e^{-j2\pi cf/f_r}/\sqrt{f_r}$ or dolphin clicks, are best analyzed using the quadratic or affine class QTFRs, which preserve constant time shifts $\tau(f) = 1/f_r$ that equal the constant spectral delay $r(f) = 1/f_r$ of the signal. Signals with hyperbolic (t, f) characteristics, such as hyperbolic impulses (the signal in Eq. (5.6.6) with $\vartheta(b) = \ln b$ and $r(f) = 1/f$) or Doppler-invariant signals such as bats echolocation signals, are best analyzed using hyperbolic QTFRs since they preserve hyperbolic time shifts $\tau(f) = 1/f$. Figure 5.6.4 analyzes the sum of two windowed hyperbolic impulses given (before windowing) as $X(f) = \frac{1}{\sqrt{f}} \left(e^{-j2\pi 3 \ln \frac{f}{f_r}} + e^{-j2\pi 7 \ln \frac{f}{f_r}} \right)$, $f > 0$. As expected, the hyperbolic Altes Q-distribution [6] (the DWD in Eq. (5.6.8) with $\Lambda(b) = \ln b$) in Fig. 5.6.4(c) results in high (t, f) concentration along the two hyperbolae $t = 3/f$ and $t = 7/f$. However, it also results in cross terms (CTs) along the mean hyperbola $t = (3 + 7)/(2f)$ [58]. The smoothed pseudo Altes Q-distribution (the hyperbolically warped version of the smoothed pseudo WD) in Fig. 5.6.4(d) removes the CTs with some loss of (t, f) resolution. QTFRs, such as the WD in Fig. 5.6.4(a) and the smoothed pseudo WD in Fig. 5.6.4(b), are not well-matched to hyperbolic impulses. The WD results in complicated CTs between the two signal components as well as inner interference terms. In comparison to the smoothed pseudo Altes Q-distribution in Fig. 5.6.4(d), the smoothed pseudo WD in Fig. 5.6.4(b) has a larger loss of (t, f) resolution, and it is not as successful at removing all the CTs.

5.6.4 SUMMARY AND CONCLUSIONS

This section emphasizes the importance of matching a QTFR to the analysis signal for successful processing, such as when the dispersive spectral delay or changes in spectral delay of a signal match the frequency-varying time-shift covariance property of the QTFRs. The QTFRs described in this section are important in analyzing signals propagating through systems with dispersive, nonlinear characteristics including hyperbolic, power, or exponential. Examples of systems where these QTFRs can be used are discussed in [59,60]. The QTFR classes considered here are obtained based on covariance properties, and can thus be considered as special cases of the covariant QTFRs in Section 4.3. See further discussions in Section 15.3.

**FIGURE 5.6.4**

(t, f) analysis of the sum of two windowed hyperbolic impulses. The first row shows QTFRs, and the second row shows QTFRs from the hyperbolic class in Eq. (5.6.9). (a) WD, (b) smoothed pseudo WD, (c) Altes Q-distribution (QD), and (d) smoothed pseudo QD.

5.7 DESIGN OF HIGH-RESOLUTION QUADRATIC TFDs WITH SEPARABLE KERNELS⁰

5.7.1 RIDs AND QUADRATIC TFDs

Reduced Interference Distributions (RIDs) smooth out the unwanted cross-terms in the time-frequency (t, f) domain (see [Chapter 3](#) and [Section 5.2](#)). The spectrogram is the best-known RID, but not the only one; [Section 5.2.3](#) shows that a (t, f) distribution (TFD) is an RID if its kernel has a

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **G.R. Putland**, Prosper Australia, Melbourne, VIC, Australia (grpputland@gmail.com). Reviewers: N.A. Khan, Z.M. Hussain, and V. Sucic.

2D low-pass characteristic in the Doppler-lag (v, τ) domain [18,27]. A suitably chosen *separable kernel*, that is, a kernel with the form $G_1(v)g_2(\tau)$, can meet this requirement while giving higher (t, f) resolution than the spectrogram; for example, the “smoothed WVDs” discussed in [61,62] have separable kernels. The superior performance of the separable kernels as compared to the spectrogram is due to the flexibility to independently adjust smoothing along the time or frequency axes, or both. Note that the spectrogram does not have this flexibility (see [62]). It has also been shown that if the kernel is *independent of lag* (i.e., a function of Doppler alone in the Doppler-lag domain, or of time alone in the time-lag domain), then the resulting TFD can exhibit improved frequency resolution and high attenuation of cross-terms for a certain class of signals [63,64].

This section explores the properties of TFDs with separable kernels, including lag-independent kernels, characterizes the signals for which such kernels can be recommended, and gives examples of RID designs using such kernels. It builds on the argument presented in Sections 3.2.1 and 3.2.2 (pp. 110–114), and extends Section 3.2.5 (p. 117).

5.7.2 SEPARABLE KERNEL FORMULATIONS

In the Doppler-lag domain, a *separable kernel* was defined in Section 3.2.5.1 as

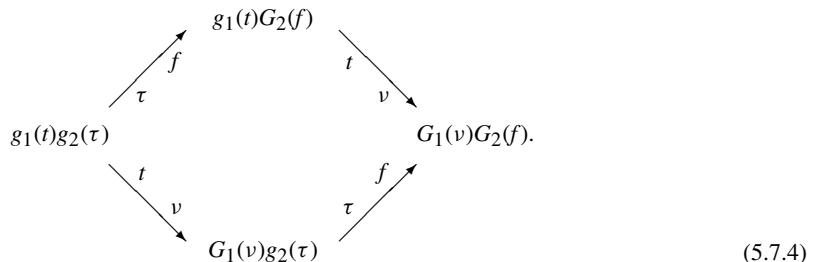
$$g(v, \tau) = G_1(v)g_2(\tau). \quad (5.7.1)$$

If we let

$$G_1(v) = \mathcal{F}_{t \rightarrow v} \{g_1(t)\} \quad (5.7.2)$$

$$G_2(f) = \mathcal{F}_{\tau \rightarrow f} \{g_2(\tau)\}, \quad (5.7.3)$$

then the relationships shown in the graphical Eq. (3.2.6) become



Substituting Eq. (5.7.1) into Eq. (3.2.11), we obtain the filtered AF

$$\mathcal{A}_z(v, \tau) = G_1(v)g_2(\tau)A_z(v, \tau) \quad (5.7.5)$$

(also called the “generalized ambiguity function” by some authors). Then, using the convolution properties and the notations of the graphical Eq. (3.2.10), we find

$$r_z(v, f) = G_1(v)G_2(f) * k_z(v, f) \quad (5.7.6)$$

$$R_z(t, \tau) = g_2(\tau) \underset{t}{\ast} K_z(t, \tau) \quad (5.7.7)$$

$$\rho_z(t, f) = g_1(t) \underset{f}{\ast} W_z(t, f) \underset{f}{\ast} G_2(f), \quad (5.7.8)$$

where $K_z(t, \tau) = z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$ is the (unsmoothed) IAF; $k_z(v, f)$ is its 2D-FT (the spectral autocorrelation function); $W_z(t, f)$ is the WVD; $r_z(v, f)$ is the smoothed spectral autocorrelation function; $R_z(t, \tau)$ is the smoothed IAF, and $\rho_z(t, f)$ is the quadratic TFD.

Equation (5.7.7) shows that the effect of the lag-dependent factor on the TFD is simply “lag windowing,” that is, multiplication by the same factor in the (t, τ) domain before transforming to the (t, f) domain. In Eq. (5.7.8), the two convolutions are associative: it does not matter whether we perform the t -convolution or the f -convolution first. Thus, we may consider the Doppler-dependent and lag-dependent factors in Eq. (5.7.5) as leading to separate convolutions in time and frequency, respectively.

A *Doppler-independent* (DI) kernel is a special case of a separable kernel obtained by putting $G_1(v) = 1$ in Eqs. (5.7.1) and (5.7.2), which then become $g(v, \tau) = g_2(\tau)$ and $g_1(t) = \delta(t)$.

Making these substitutions in Eqs. (5.7.4) to (5.7.8), we obtain

$$(5.7.9)$$

$$A_z(v, \tau) = g_2(\tau) A_z(v, \tau) \quad (5.7.10)$$

$$r_z(v, f) = G_2(f) \underset{f}{\ast} k_z(v, f) \quad (5.7.11)$$

$$R_z(t, \tau) = g_2(\tau) K_z(t, \tau) \quad (5.7.12)$$

$$\rho_z(t, f) = G_2(f) \underset{f}{\ast} W_z(t, f). \quad (5.7.13)$$

As seen in Eq. (5.7.9), a “Doppler-independent” kernel is indeed independent of Doppler in all four domains. In the Doppler-lag domain, it is a function of lag alone. Equation (5.7.13) shows that a DI kernel causes smoothing (or “smearing”) of the WVD in the frequency direction only. The graphical Eq. (3.2.10) defined the notation

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{R_z(t, \tau)\}. \quad (5.7.14)$$

Substituting Eq. (5.7.12) into Eq. (5.7.14) gives

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{g_2(\tau) K_z(t, \tau)\}, \quad (5.7.15)$$

which shows that a quadratic TFD with a DI kernel is a *windowed WVD*; the “windowing” is applied in the lag direction before Fourier transformation from lag to frequency.

A *lag-independent* (LI) kernel is another special case of a separable kernel, obtained by putting $g_2(\tau) = 1$ in Eqs. (5.7.1) and (5.7.3), which then become $g(v, \tau) = G_1(v)$ and $G_2(f) = \delta(f)$.

Making these substitutions in Eqs. (5.7.4) to (5.7.8), we obtain

$$\begin{array}{ccc}
 & g_1(t)\delta(f) & \\
 & \swarrow f \quad \searrow t & \\
 g_1(t) & & G_1(v)\delta(f) \\
 & \swarrow v \quad \searrow \tau & \\
 & G_1(v) &
 \end{array} \tag{5.7.16}$$

$$\mathcal{A}_z(v, \tau) = G_1(v)A_z(v, \tau) \tag{5.7.17}$$

$$r_z(v, f) = G_1(v)k_z(v, f) \tag{5.7.18}$$

$$R_z(t, \tau) = g_1(t) * \int_t K_z(t, \tau) \tag{5.7.19}$$

$$\rho_z(t, f) = g_1(t) * \int_t W_z(t, f). \tag{5.7.20}$$

The last equation shows that an LI kernel causes smoothing of the WVD in the time direction only; the result may be called a *filtered WVD*.

As seen in Eq. (5.7.16), a “lag-independent” kernel is indeed independent of lag in all four domains. In the time-lag domain, it is a function of time alone; for this reason, such kernels have also been called “time-only kernels” [63,64].

The WVD kernel is $g(v, \tau) = 1$, which is both Doppler-independent and lag-independent; it may be regarded as DI with $g_2(\tau) = 1$ or as LI with $G_1(v) = 1$.

5.7.3 PROPERTIES

Table 5.7.1 is extracted from **Table 3.3.1** (p. 122). The properties of time-shift invariance and frequency-shift invariance are omitted, being common to all quadratic TFDs. Positivity is omitted because, in practice, the design of a separable kernel involves a deliberate sacrifice of non-negativity in favor of higher resolution and the interpretation of the TFD as a (t, f) gradient of energy. Let us define a *proper* DI or LI kernel as one that is *nonconstant* (yielding a TFD that is *not* a WVD, with or without amplitude scaling), and let us define a *proper* separable kernel as one that is neither DI nor LI. Then **Table 5.7.1** shows that a TFD with a DI kernel can satisfy the realness, time marginal, time support, and IF properties; but no proper DI kernel satisfies the frequency marginal, frequency support, or spectral delay property due to the smoothing along the frequency axis. Similarly, a TFD with an LI kernel can satisfy the realness, frequency marginal, frequency support, and spectral delay properties; but no proper LI kernel satisfies the time marginal, time support, or IF property due to the smoothing along the time axis.

Table 5.7.1: TFD Properties and Associated Kernel Requirements for Separable, Doppler-Independent, and Lag-Independent Kernels

| Property | Kernel Constraints | | |
|-----------------|--|---------------------------|------------------------------|
| | Separable $g(v, \tau) = G_1(v)g_2(\tau)$ | DI $G_1(v) = 1$ | LI $g_2(\tau) = 1$ |
| Realness | $G_1(v)g_2(\tau)$ $= G_1^*(-v)g_2^*(-\tau)$ | $G_2(f)$ is real | $g_1(t)$ is real |
| Time marginal | $G_1(v)g_2(0) = 1 \forall v$ | $g_2(0) = 1$ | WVD only |
| Freq. marginal | $G_1(0)g_2(\tau) = 1 \forall \tau$ | WVD only | $G_1(0) = 1$ |
| Inst. freq. | $G_1(v)g_2(0) = \text{const.}$ $g_2'(0) = 0$ | $g_2'(0) = 0$ | WVD only* |
| Spectral delay | $G_1(0)g_2(\tau) = \text{const.}$ $G_1'(0) = 0$ | WVD only* | $G_1'(0) = 0$ |
| Time support | DI only | Always | WVD only* |
| Freq. support | LI only | WVD only* | Always |
| RID potential | Unrestricted | Inner artifacts | Cross-terms |

Notes: Explanation of properties: Time marginal: The integral of the TFD over frequency is the instantaneous power. Freq. marginal: The integral of the TFD over time is the energy spectrum. Inst. freq.: The IF is the first moment of the TFD w.r.t. frequency. Spectral delay: The spectral delay is the first moment of the TFD w.r.t. time. Time support: If the nonzero values of the signal are confined to a certain time interval, so are the nonzero values of the TFD. Freq. support: If the nonzero values of the spectrum are confined to a certain frequency range, so are the nonzero values of the TFD. “WVD only” (with asterisk) means a WVD multiplied by an arbitrary constant.*

The reduced-interference property (“RID potential”) requires further explanation. The WVD may contain interference terms of two kinds. *Inner artifacts* or “inner interference terms” (see [Section 4.2](#)) are caused by nonlinear frequency modulation laws, and cause the WVD to oscillate as we move normal to the expected feature(s) in the (t, f) plane. In the case of a multicomponent signal, *cross-terms* or “outer interference terms” (see [Section 4.2](#)) are caused by cross-product terms in the IAF $K_z(t, \tau)$, and cause the WVD to oscillate as we move parallel to the expected features in the (t, f) plane.

Cross-terms oscillate in the direction orthogonal to the line joining two signal components ([Section 4.2](#)). For small-duration spikes, the direction of oscillation of cross-terms thus becomes parallel to the frequency axis ([Section 5.9](#)). Such cross-terms can be suppressed by smoothing along the frequency axis by using a DI kernel with a sufficiently long $G_2(f)$ (see Eqs. [\(5.7.13\)](#) and [\(3.2.31\)](#)), which corresponds to a sufficiently short $g_2(\tau)$. For pure tones, the direction of oscillation of cross-terms becomes parallel to the time axis. Such cross-terms can be easily removed by smoothing along the time axis, that is, using an LI kernel with a sufficiently long $g_1(t)$, which corresponds to a sufficiently short $G_1(v)$ (see Eq. [\(5.7.20\)](#)). To deal with more general signals, a proper separable kernel would need to cause a convolution in both time and frequency so as to suppress or reduce both kinds of interference terms.

For an LI kernel, the suppression of cross-terms is facilitated if the components are of slowly varying frequencies, so that the cross-terms extend (and oscillate) approximately in the time direction.

Furthermore, the loss of frequency resolution caused by convolution with $g_1(t)$ is proportional to the rate of change of the IF; for constant frequency, the components run parallel to the time axis, so that there is no loss of resolution apart from that caused by the time-variation of frequency resolution in the WVD.

5.7.4 DESIGN EXAMPLES OF SEPARABLE-KERNEL TFDs

Early experience suggested that the kernel of an RID must exhibit a *two-dimensional* low-pass characteristic in the Doppler-lag domain (e.g., [18, pp. 79-81]). The *B-distribution* (BD) defined in Table 3.3.2 (p. 123) [65] has the separable time-lag kernel $G_B(t, \tau) = |\tau|^\beta \cosh^{-2\beta} t$, where β is a positive real parameter that controls the degree of smoothing. This kernel is low-pass in the Doppler dimension but *not* in the lag dimension. Nevertheless, the BD has shown impressive RID properties for certain signals [65].

The best results from the BD are consistently obtained for small positive values of β , for which the lag-dependent factor is nearly constant apart from a “slot” at $\tau = 0$. Moreover, any desired lag-dependence can be introduced later by windowing prior to Fourier transformation from τ to f , as is often done for computational economy or improved time resolution. Accordingly, the BD was modified in [63,64] by making the “lag-dependent” factor *exactly* constant. The resulting *modified B-distribution* (MBD) had an LI kernel. This inspired a more thorough inquiry into the properties of LI kernels, and the findings explained the behavior of the BD, whose kernel may fairly be described as “nearly LI” for small positive β .

The time-lag kernel of the MBD is given in Table 3.3.2 (p. 123) as $G_{MB}(t, \tau) = g_\beta(t) = \frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi}$, where β is a positive real parameter and the denominator is for normalization. The graph of $g_\beta(t)$ versus t is a bell-shaped curve whose spread is inversely related to β .

If the signal is a tone or sum of tones, we can obtain closed-form expressions showing that the MBD has optimal concentration about the IF law. In particular, the MBD of a single tone is a delta function of frequency, while the MBD of the sum of two tones comprises two delta functions plus a cross-term whose amplitude is controlled by β (see [64] and Section 10.3). For most signals, there is no closed-form expression for the MBD, so we must resort to numerical computations with discretized variables. For discrete time n and discrete lag m , the MBD kernel becomes

$$G_{MB}(n, m) = g_\beta(n) = \frac{\cosh^{-2\beta} n}{\sum_i \cosh^{-2\beta} i}. \quad (5.7.21)$$

The following numerical examples include one MBD for comparison. They also include TFDs with separable kernels using two types of window functions that are well known in digital signal processing and spectral analysis. The *M*-point *Hamming* function, where M is odd, is

$$\text{hamm}_M(i) = 0.54 + 0.46 \cos \frac{2\pi i}{M}; \quad -\frac{M-1}{2} \leq i \leq \frac{M-1}{2}, \quad (5.7.22)$$

where i is discrete time or lag. The *L*-point *Hanning* function, where L is odd, is

$$\text{hann}_L(i) = 0.5 + 0.5 \cos \frac{2\pi i}{L}; \quad -\frac{L-1}{2} \leq i \leq \frac{L-1}{2}. \quad (5.7.23)$$

In the numerical examples given below, the Hamming function is preferred in the lag direction and the Hanning in the time direction; and whenever a proper separable kernel is compared with DI and LI kernels, it has the same lag factor as the DI kernel and the same Doppler factor (or time factor) as the LI kernel.

5.7.5 RESULTS AND DISCUSSION

Some numerical computations of separable-kernel TFDs, and of other TFDs for purposes of comparison, are presented in the following graphs. Complete specifications of signals and kernels are given in the figure captions for ease of reference. Each graph includes the TFD (main panel, with time axis vertical), time plot (left panel), and magnitude spectrum (bottom panel).

[Figure 5.7.1](#) shows two TFDs of a signal comprising a tone (constant frequency) and a chirp (linearly increasing frequency). Part (a) shows the WVD, with the prominent cross-term. Part (b) shows the effect of an LI kernel, which smooths the WVD in the time direction, suppressing the oscillatory cross-term. For the tone, this smoothing causes only a slight loss of frequency resolution, and only because the frequency resolution of the WVD varies with time. For the chirp, the loss of frequency resolution is greater because the direction of smoothing is not parallel to the IF law; that is, the smoothing is “across” as well as “along” the component. The effect on resolution due to the angle between the FM law and the time axis is discussed further in [Section 5.9](#).

[Figure 5.7.2](#) compares six TFDs of a two-component signal with slowly varying frequencies; the lower-frequency component is a pure tone. In [Fig. 5.7.3](#), the tone is replaced by a faster-varying sinusoidal FM signal (nonlinear FM). In each figure, part (a) shows the WVD, while part (b) shows the effect of a DI kernel, part (c) an LI kernel, and part (d) a separable kernel combining the lag and time functions of parts (b) and (c). Note that (b) and (d) are related by the same time-smoothing as (a) and (c), while (c) and (d) are related by the same frequency-smoothing as (a) and (b). In each figure,

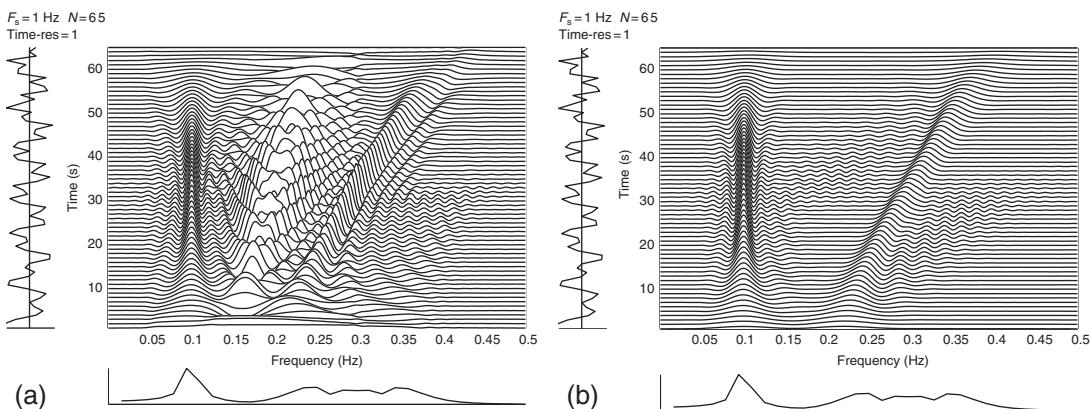
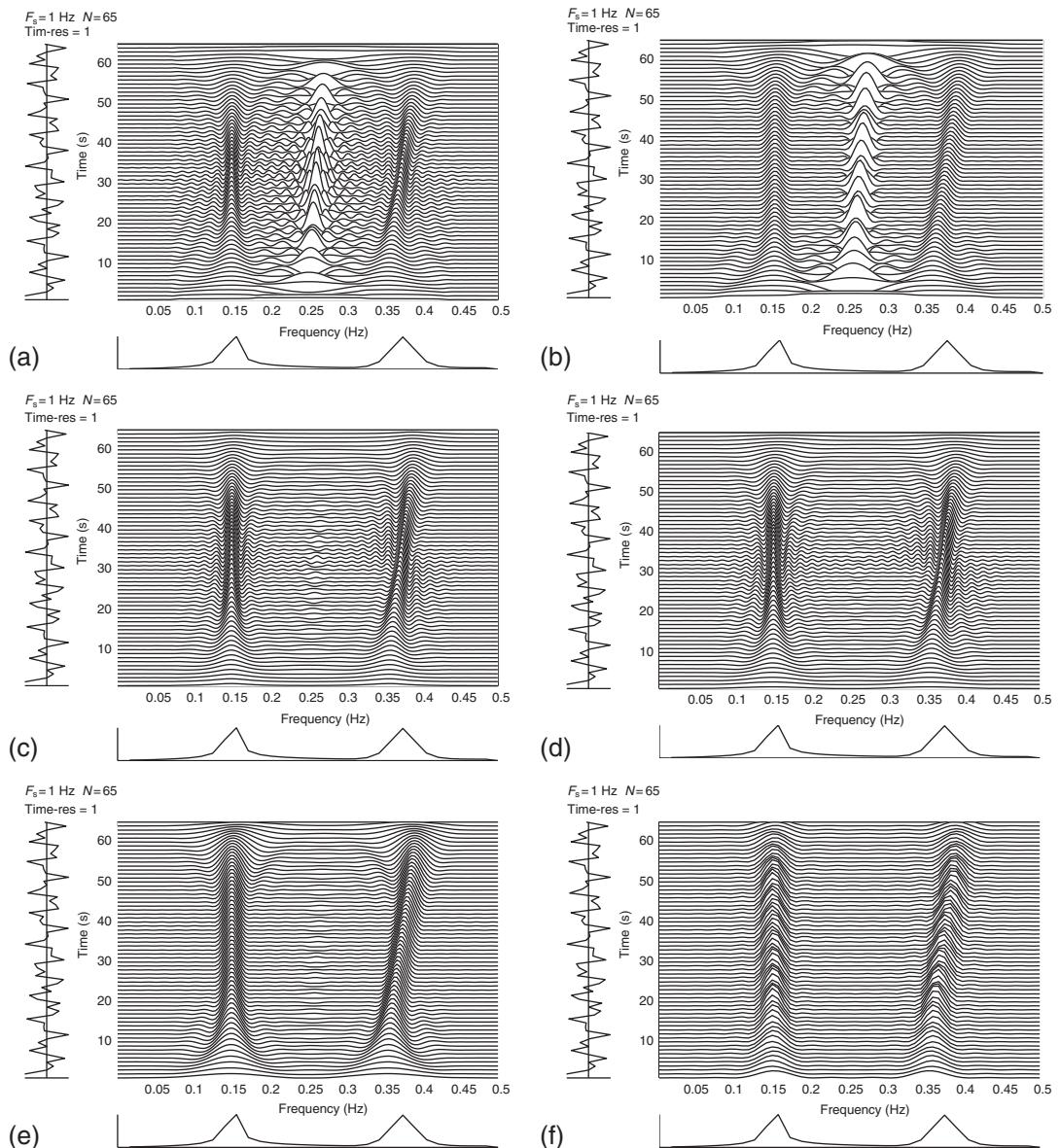
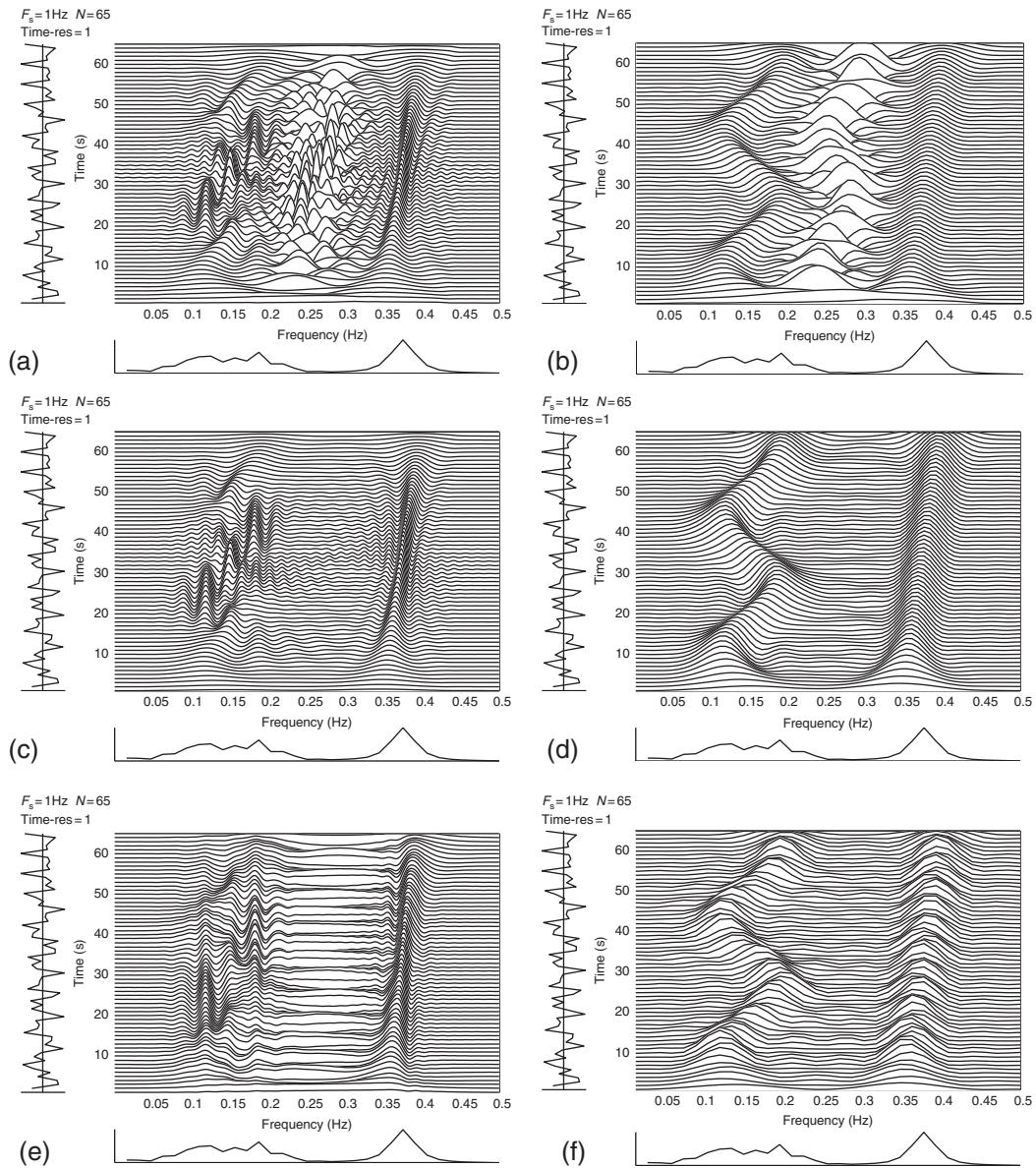


FIGURE 5.7.1

TFDs of the sum of a tone (frequency 0.1 Hz) and a linear FM signal (frequency range 0.2–0.4 Hz), unit amplitudes, duration 65 samples, sampling rate 1 Hz: (a) WVD; (b) lag-independent, $G(n, m) = \text{hann}_{15}(n)$.

**FIGURE 5.7.2**

TFDs of the sum of a tone (frequency 0.15 Hz) and a linear FM signal (frequency range 0.35-0.4 Hz), unit amplitudes, duration 65 samples, sampling rate 1 Hz in time, 0.5 Hz in lag: (a) WVD; (b) Doppler-independent, $G(n, m) = \delta(n) \text{ hamm}_{47}(m)$; (c) lag-independent, $G(n, m) = \text{hann}_{11}(n)$; (d) separable, $G(n, m) = \text{hann}_{11}(n) \text{ hamm}_{47}(m)$; (e) modified B, $\beta = 0.2$; (f) spectrogram, 35-point rectangular window (sampling rate 1 Hz).

**FIGURE 5.7.3**

TFDs of the sum of a sinusoidal FM signal (frequency 0.15 ± 0.05 Hz, two cycles of modulation) and a linear FM signal (frequency range 0.35-0.4 Hz), unit amplitudes, duration 65 samples, sampling rate 1 Hz in time, 0.5 Hz in lag: (a) WVD; (b) Doppler-independent, $G(n, m) = \delta(n) \text{ hamm}_{23}(m)$; (c) lag-independent, $G(n, m) = \text{hann}_{11}(n)$; (d) separable, $G(n, m) = \text{hann}_{11}(n) \text{ hamm}_{23}(m)$; (e) ED, $\sigma = 1$; (f) spectrogram, 17-point rectangular window (sampling rate 1 Hz).

part (f) shows a spectrogram for comparison with the separable-kernel TFD. (For further explanation of the sampling rates and their significance, see Sections 6.1 and 6.1.3.3.)

In Fig. 5.7.2, a cross-term is prominent in the WVD (a) and is not suppressed by the DI kernel (b). It is suppressed by the Hanning LI kernel (c) and the MBD kernel (e), which is also LI. Both LI kernels are bell-shaped functions of time, with parameters chosen to give similar degrees of smoothing in time; that accounts for the similarity between graphs (c) and (e). The proper separable kernel (d) has lower frequency resolution, but less ripple about the IF laws, than the LI kernels. The spectrogram (f) gives the best suppression of artifacts and the lowest resolution. Comparing Figs. 5.7.1(b) with Fig. 5.7.2(c), we see that the faster-varying IF, which causes a faster-varying cross-term frequency and lower minimum beat frequency, needs a longer time-smoothing function for adequate suppression of cross-terms.

In Fig. 5.7.3, both cross-terms and inner artifacts are visible. The cross-terms appear as “rough terrain” between the components, while the inner artifacts appear as spurious ridges in the sinusoidal FM component; both types are prominent in the WVD (a). The DI kernel (b) is effective against the inner artifacts. The LI kernel (c) is effective against the cross-terms. The separable kernel (d) is effective against both. The well-known ED (e), with the parameter σ visually optimized, spreads the inner artifacts in time and the cross-terms in frequency but does not satisfactorily suppress either. In this example, the separable-kernel TFD (d) appears to offer a better compromise between resolution and cross-term suppression than either the spectrogram (f) or the ED (e), neither of which has a separable kernel. Comparing Fig. 5.7.2(d) with Fig. 5.7.3(d), we may confirm that a more nonlinear IF requires a shorter lag window (for the suppression of inner artifacts), giving coarser frequency resolution.

5.7.6 SUMMARY AND CONCLUSIONS

A separable kernel gives separate control of the frequency-smoothing and time-smoothing of the WVD which is an improvement over the spectrogram which does not have flexibility to independently adjust smoothing along the time and/or the frequency axis [62]. The lag-dependent factor causes a convolution in the frequency direction in the (t,f) plane, while the Doppler-dependent factor causes a convolution in the time direction. A Doppler-independent (DI) kernel smooths the WVD in the frequency direction only, reducing the inner artifacts and preserving the time marginal. A lag-independent (LI) kernel smooths the WVD in the time direction only, reducing the cross-terms and preserving the frequency marginal.

For an LI kernel, slower variations in component frequencies allow easier suppression of cross-terms and higher frequency resolution in the TFD; such kernels should therefore be considered for representing multicomponent signals with slowly varying IFs. For a multicomponent signal with at least one highly nonlinear IF law, a lag-dependent factor is also needed to suppress the inner artifacts.

The separable-kernel approach allows a complete understanding and appraisal of the properties and behavior of the smoothed WVD [61], the BD [65], the MBD [63,64], and the EMBD. It also enables the construction of high-resolution quadratic TFDs using the classical smoothing functions commonly encountered in digital signal processing and spectral analysis. Further improvements in (t,f) resolution can be achieved by accounting for the main direction of energy concentration in the (t,f) plane, as discussed in [27,62] and Section 5.9.

5.8 FRACTIONAL FOURIER TRANSFORM AND GENERALIZED-MARGINAL TFDs⁰

5.8.1 FRACTIONAL FOURIER TRANSFORM

The fractional Fourier transform (FrFT) is a rotation of the time-frequency plane. For a real α , the FrFT \mathbb{F}^α with angle α is defined by

$$(\mathbb{F}^\alpha s)(u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j\frac{u^2}{2}\cot\alpha} \int_{-\infty}^{\infty} s(t) e^{j\frac{t^2}{2}\cot\alpha} e^{-jut\csc\alpha} dt, & \text{if } \alpha \text{ is not a multiple of } \pi, \\ s(u), & \text{if } \alpha \text{ is a multiple of } 2\pi, \\ s(-u), & \text{if } \alpha + \pi \text{ is a multiple of } 2\pi. \end{cases} \quad (5.8.1)$$

Also, one can see that $\mathbb{F}^{2n\pi}$ for an integer n is the identity transformation and $\mathbb{F}^{\frac{\pi}{2}}$ is the traditional FT \mathbf{F} . Moreover, the following rotation property holds:

$$\mathbb{F}^{\alpha+\beta} = \mathbb{F}^\alpha \mathbb{F}^\beta.$$

For more details, see, for example, [66,67]. A numerical example of the FrFT of a signal is shown in Fig. 5.8.1.

With the FrFT, it was proved in [67,68] that a rotation of a WVD is still a WVD as explained below.

In this section, the standard time-frequency notations (t,f) is replaced for convenience by (t,ω) where $\omega = 2\pi f$ to align with angle rotations notations. Let $W_s(t,\omega)$ denote the WVD of a signal $s(t)$, that is,

$$W_s(t,\omega) = \int s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau. \quad (5.8.2)$$

Let α be an angle and $(\tilde{t}, \tilde{\omega})$ be a rotation of (t, ω) with angle α :

$$\begin{cases} \tilde{t} &= t \cos \alpha + \omega \sin \alpha, \\ \tilde{\omega} &= -t \sin \alpha + \omega \cos \alpha, \end{cases} \quad (5.8.3)$$

and

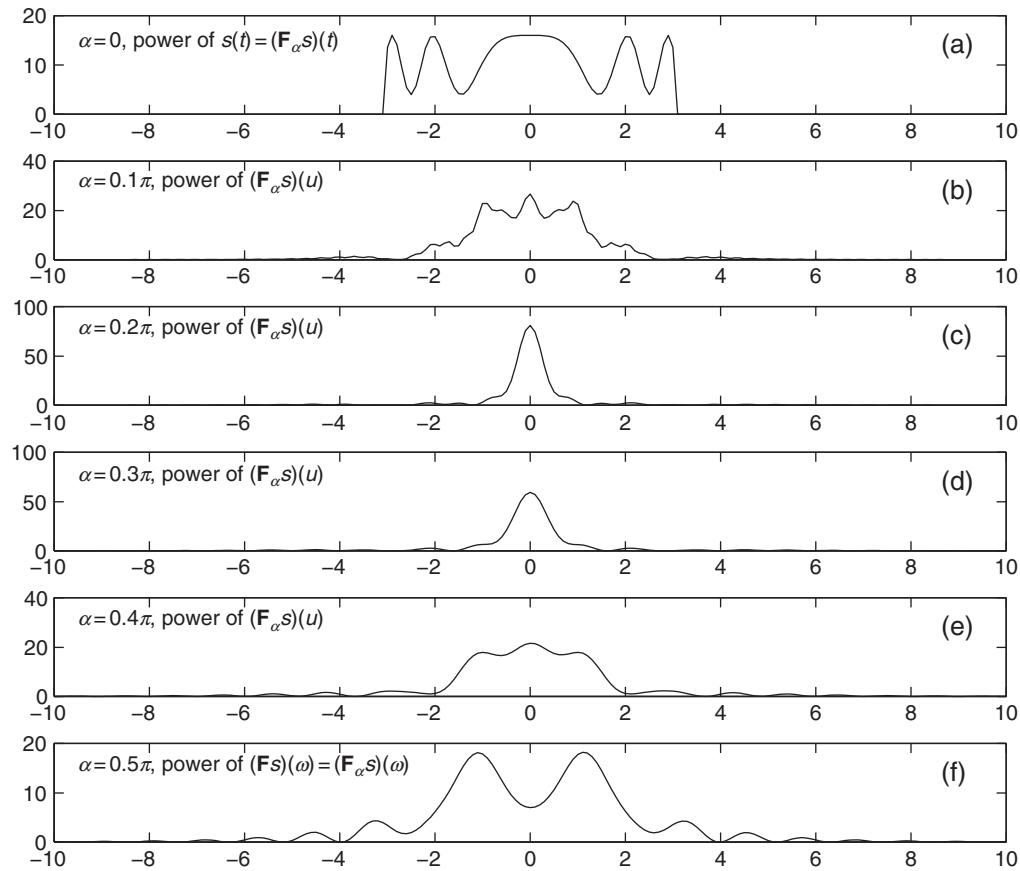
$$\tilde{W}_s(\tilde{t}, \tilde{\omega}) = W_s(\tilde{t} \cos \alpha - \tilde{\omega} \sin \alpha, \tilde{t} \sin \alpha + \tilde{\omega} \cos \alpha).$$

Then, applying the results in [67] yields:

$$\tilde{W}_s(\tilde{t}, \tilde{\omega}) = \int (\mathbb{F}^\alpha s)\left(\tilde{t} + \frac{\tau}{2}\right) (\mathbb{F}^\alpha s)^*\left(\tilde{t} - \frac{\tau}{2}\right) e^{-j\tilde{\omega}\tau} d\tau. \quad (5.8.4)$$

Equation (5.8.4) tells us that the rotation $\tilde{W}_s(\tilde{t}, \tilde{\omega})$ of the WVD $W_s(t, \omega)$ of a signal s is the WVD of the signal $\mathbf{F}_\alpha s$. It was also proved in [68] that a rotation of a Radon-Wigner distribution is also a Radon-Wigner distribution by using the FrFT technique. It is known that a WVD satisfies the conventional marginal properties, that is,

⁰Author: X.-G. Xia, University of Delaware, Newark, DE, USA (xxia@ee.udel.edu). Reviewers: P.M. Oliveira and S. Qian.

**FIGURE 5.8.1**

The FrFT of a signal $s(t)$ in (a) showing the transition from t to ω ; for (a) $\alpha = 0\pi$; (b) $\alpha = 0.1\pi$; (c) $\alpha = 0.2\pi$; (d) $\alpha = 0.3\pi$; (e) $\alpha = 0.4\pi$; and (f) $\alpha = 0.5\pi$. (Note the progression from time t to frequency ω through the transition α).

$$\int \tilde{W}_s(\tilde{t}, \tilde{\omega}) d\tilde{\omega} = |(\mathbb{F}^\alpha s)(\tilde{t})|^2 \quad (5.8.5)$$

and

$$\int \tilde{W}_s(\tilde{t}, \tilde{\omega}) d\tilde{t} = |\mathbb{F}(\mathbb{F}^\alpha s)(\tilde{\omega})|^2 = |(\mathbb{F}^{\alpha+\pi/2} s)(\tilde{\omega})|^2. \quad (5.8.6)$$

5.8.2 GENERALIZED-MARGINAL TIME-FREQUENCY DISTRIBUTION

For joint TFDs, one often imposes the marginal properties: the integrals of a TFD along the time t and the frequency ω are the powers of the signal in the frequency and the time domains, respectively. For

a TFD of the quadratic class (as defined on p. 113), satisfaction of the time and frequency marginals is equivalent to the kernel property $g(0, \tau) = g(\nu, 0) = 1$ for all real ν and τ , where $g(\nu, \tau)$ is the Doppler-lag kernel¹ (see Table 3.3.1). This fact will later be seen to be just a special case of a more general property.

It is normally accepted that, when signals are chirp signals, the TFDs should be concentrated on lines in the TF plane. The question is the following. If we have some prior information about a signal, can we take advantage of it in the design of a TFD? Or, can we impose some requirements on a TFD along these lines? If so, how? The generalized-marginal TFDs proposed in [69] give an answer to these questions.

5.8.2.1 Generalized marginals

Generalized marginals are the marginals beyond the usual time and frequency marginals. Let $\rho_s(t, \omega)$ be a TFD of a signal $s(t)$. Let L_α denote the straight line

$$\omega \cos \alpha - t \sin \alpha = 0, \quad (5.8.7)$$

that is, the line on the time-frequency plane through the origin with angle α . Let $L_\alpha(u)$ denote the line

$$\omega \cos \alpha - t \sin \alpha = u, \quad (5.8.8)$$

that is, the general member of the family of lines parallel to L_α , parameterized by the real number u , where u is the signed perpendicular distance from L_α . We call $\rho_s(t, \omega)$ a *generalized-marginal TFD*, if the line integrals of $\rho_s(t, \omega)$ along the lines $L_{\alpha_k}(u)$, for $k = 1, 2, \dots, N$, are the powers of the FrFT with angles $\alpha_k + \pi/2$, $k = 1, 2, \dots, N$, of the signal s , respectively. In other words,

$$\int_{L_{\alpha_k}(u)} \rho_s(t, \omega) dx = |(\mathbb{F}^{\alpha_k + \pi/2} s)(u)|^2, \quad k = 1, 2, \dots, N, \quad (5.8.9)$$

where the dummy variable x is on the line $L_{\alpha_k}(u)$. This is written more simply as

$$\int_{L_{\alpha_k}} \rho_s(t, \omega) dx = |\mathbb{F}^{\alpha_k + \pi/2} s|^2, \quad k = 1, 2, \dots, N.$$

It is clear that, when $\alpha_1 = 0, \alpha_2 = \pi/2$, and $N = 2$, the above-generalized marginals are the conventional marginals. Also, the angles α_k may be chosen to be close to the angles of chirp signals in the TF plane.

5.8.2.2 Generalized-marginal quadratic TFDs

In this section, we study TFDs with kernels $g(\nu, \tau)$ in the quadratic class [19], which are generalized-marginal TFDs (Eq. 5.8.9). We show that a TFD with kernel $g(\nu, \tau)$ in the quadratic class is a generalized-marginal one (Eq. 5.8.9) if and only if its kernel $g(\nu, \tau)$ is equal to 1 on the lines that are perpendicular to L_{α_k} , $k = 1, 2, \dots, N$, and pass through the origin. This implies that the WVD satisfies all the generalized marginals and it is the only one in the quadratic class with this property.

¹In this section, $\nu = 2\pi\nu$ and $\omega = 2\pi f$.

A TFD for a signal $s(t)$ in the quadratic class is defined by

$$\rho_s(t, \omega) = \frac{1}{4\pi^2} \iint \mathcal{A}_s(\nu, \tau) e^{+j\nu t - j\tau \omega} d\nu d\tau,$$

where $\mathcal{A}_s(\nu, \tau)$ is the generalized ambiguity function (AF) of the signal $s(t)$ with a kernel $g(\nu, \tau)$:

$$\mathcal{A}_s(\nu, \tau) = g(\nu, \tau) \int s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) e^{j\nu u} du.$$

The TFD $\rho_s(t, \omega)$ can be also written as

$$\rho_s(t, \omega) = \iiint e^{+j\nu t - j\tau \omega + j\nu u} g(\nu, \tau) s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) d\nu d\tau du. \quad (5.8.10)$$

Then, $\rho(t, \omega)$ is a generalized-marginal TFD if and only if the following holds. Let

$$\begin{aligned} \tilde{t} &= t \cos \alpha + \omega \sin \alpha, & t &= \tilde{t} \cos \alpha - \tilde{\omega} \sin \alpha, \\ \tilde{\omega} &= -t \sin \alpha + \omega \cos \alpha, & \omega &= \tilde{t} \sin \alpha + \tilde{\omega} \cos \alpha. \end{aligned}$$

Then, the condition (5.8.9) is equivalent to

$$\int \rho_s(\tilde{t} \cos \alpha_k - \tilde{\omega} \sin \alpha_k, \tilde{t} \sin \alpha_k + \tilde{\omega} \cos \alpha_k) d\tilde{t} = |(\mathbb{F}^{\alpha_k + \pi/2} s)(\tilde{w})|^2, \quad k = 1, 2, \dots, N. \quad (5.8.11)$$

In other words, a TFD $\rho_s(t, \omega)$ is a generalized-marginal distribution if and only if it satisfies Eq. (5.8.11). We now focus on the TFDs $\rho_s(t, \omega)$ in Eq. (5.8.10), for a given angle α in Eq. (5.8.11). Let us see what the left-hand side of Eq. (5.8.11) with angle α for $\rho_s(t, \omega)$ in Eq. (5.8.10) is.

$$\begin{aligned} &\int \rho_s(\tilde{t} \cos \alpha - \tilde{\omega} \sin \alpha, \tilde{t} \sin \alpha + \tilde{\omega} \cos \alpha) d\tilde{t} \\ &= \frac{1}{\cos \alpha} \iint e^{-j\tilde{\omega}\tau(\sin \alpha \tan \alpha + \cos \alpha) - j\tilde{t}\tau \tan \alpha} g(-\tau \tan \alpha, \tau) s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) d\tau du \\ &= \int e^{-j\tilde{\omega}\tau} g(-\tau \sin \alpha, \tau \cos \alpha) A_s(-\tau \sin \alpha, \tau \cos \alpha) d\tau, \end{aligned}$$

where A_s is the AF of s . It was proved (e.g., [70]) that

$$A_s(-\tau \sin \alpha, \tau \cos \alpha) = A_{\mathbb{F}^{\alpha + \pi/2} s}(\tau, 0).$$

Therefore,

$$\int \rho_s(\tilde{t} \cos \alpha - \tilde{\omega} \sin \alpha, \tilde{t} \sin \alpha + \tilde{\omega} \cos \alpha) d\tilde{t} = \iint e^{j(u - \tilde{\omega})\tau} g(-\tau \sin \alpha, \tau \cos \alpha) |\mathbb{F}^{\alpha + \pi/2} s(u)|^2 du d\tau.$$

Therefore, the generalized-marginal property holds if and only if

$$\int e^{j(u - \tilde{\omega})\tau} g(-\tau \sin \alpha, \tau \cos \alpha) d\tau = \delta(u - \tilde{\omega}),$$

that is, $g(-\tau \sin \alpha, \tau \cos \alpha) = 1$.

Although the above discussion is for one angle only, it is straightforward to generalize it to several angles α_k for $k = 1, 2, \dots, N$. Therefore, we have obtained the following result.

Table 5.8.1: Generalized-Marginal Kernels

| Name | Kernel $g(\nu, \tau)$ | Generalized-Marginal $\tilde{g}(\nu, \tau)$ With Angles α_k , $1 \leq k \leq N$ |
|---------------|-----------------------------------|--|
| Margenau-Hill | $\cos \frac{\nu\tau}{2}$ | $\cos \left[0.5 \prod_{k=1}^N (\nu \cos \alpha_k + \tau \sin \alpha_k) \right]$ |
| Rihaczek | $e^{0.5j\nu\tau}$ | $e^{0.5j \prod_{k=1}^N (\nu \cos \alpha_k + \tau \sin \alpha_k)}$ |
| sinc | $\frac{\sin a\nu\tau}{a\nu\tau}$ | $\frac{\sin [a \prod_{k=1}^N (\nu \cos \alpha_k + \tau \sin \alpha_k)]}{a \prod_{k=1}^N (\nu \cos \alpha_k + \tau \sin \alpha_k)}$ |
| Page | $e^{0.5\nu \tau }$ | $e^{0.5j \prod_{k_1=1}^{N_1} \nu \cos \alpha_{k_1} + \tau \sin \alpha_{k_1} \prod_{k_2=1}^{N_2} (\nu \cos \alpha_{k_2} + \tau \sin \alpha_{k_2})}$ |
| Exponential | $e^{-\frac{\nu^2\tau^2}{\sigma}}$ | $e^{-\frac{1}{\sigma} \prod_{k=1}^N (\nu \cos \alpha_k + \tau \sin \alpha_k)^2}$ |

Theorem 5.8.1. A TFD $\rho_s(t, \omega)$ in Eq. (5.8.10) in the quadratic class with a kernel $g(\nu, \tau)$ is a generalized-marginal TFD with angles α_k , $k = 1, 2, \dots, N$, as in Eq. (5.8.11) if and only if

$$g(-\tau \sin \alpha_k, \tau \cos \alpha_k) = 1, \quad \text{for all real } \tau, \text{ and } k = 1, 2, \dots, N, \quad (5.8.12)$$

in other words, g is 1 on the lines perpendicular to the lines L_{α_k} , $k = 1, 2, \dots, N$, and passing through the origin.

With this result, one can easily modify the well-known kernels so that the corresponding modified TFDs are generalized marginals. We now list them in Table 5.8.1.

Theorem 5.8.1 also tells us that the WVD satisfies all marginal properties for all angles because $g(\nu, \tau) = 1$ for all real ν and τ . There is, however, a tradeoff between the number of generalized marginals you want to impose and the freedom of choosing a kernel $g(\nu, \tau)$ in the quadratic class. As more generalized marginals are required, there is less freedom in choosing kernels.

5.8.3 SUMMARY AND CONCLUSIONS

FrFTs can be used to introduce generalized-marginal TFDs. Such distributions have been shown in the literature to be among the best TFDs in the sense of having the highest resolution for chirp-type signals.

FrFTs are further developed in Section 4.8. Some applications are described in [71–73].

5.9 DESIGN OF HIGH-RESOLUTION QUADRATIC TFDs WITH MULTIDIRECTIONAL KERNELS⁰

5.9.1 LIMITATIONS OF SEPARABLE KERNELS FOR TFD RESOLUTION

As shown in Section 5.7, the properties of separable kernels offer an elegant solution for designing high-resolution TFDs with suppressed cross-terms. For instance, the Modified B Distribution (MBD)

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa), **B.K. Jawad**, Qatar University, Doha, Qatar (brahim.jawad@gmail.com), and **N.A. Khan**, Qatar University, Doha, Qatar (nabeelalikhian@qu.edu.qa). Reviewers: T. Ben-Jabeur, S. Ouelha, and S. Dong.

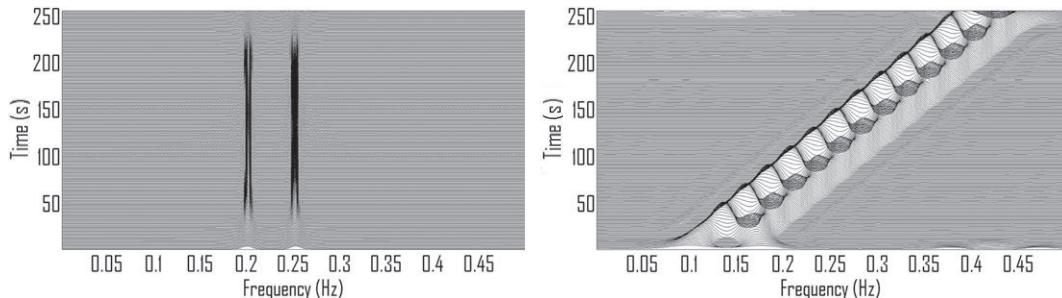


FIGURE 5.9.1

Separable-kernel-based TFDs of: (left) two tones, using the compact kernel distribution (CKD) with parameters $c = 1, D = 0.1, E = 1$; (right) two parallel LFM signals, using the CKD with $c = 1, D = E = 0.07$ (see definition of CKD in Chapter 3).

in Section 5.7, and the extended modified B-distribution (EMBD) in Section 2.7.5 and [27,62], which are both extensions of the BD (Section 5.7), were shown to be among the best quadratic TFD (QTFD) kernels for a wide range of signals, particularly newborn electroencephalography (EEG) seizure signals [27]. These QTFD kernels, however, significantly lose performance when signals have several different directions of energy concentration in the (t,f) domain. The reason is that separable kernels operate efficiently only if the signal energy concentration directions are quasi parallel to either the time axis or frequency axis in the (t,f) domain. An attempt to define a separable directional kernel is given in [74], but its domain of application is restricted to a small region centered around the ambiguity domain origin; this standard approximation ignores the presence of auto-terms away from the origin for signals modeled as multicomponent piecewise (PW) LFM signals. Figure 5.9.1 shows the limitations of separable kernels.

In order to design a high-resolution QTFD that is more generally applicable, let us analyze the location and behavior of auto-terms and cross-terms of a multicomponent PW-LFM signal in the ambiguity (ν, τ) domain.

5.9.2 LOCATION OF AUTO-TERMS AND CROSS-TERMS

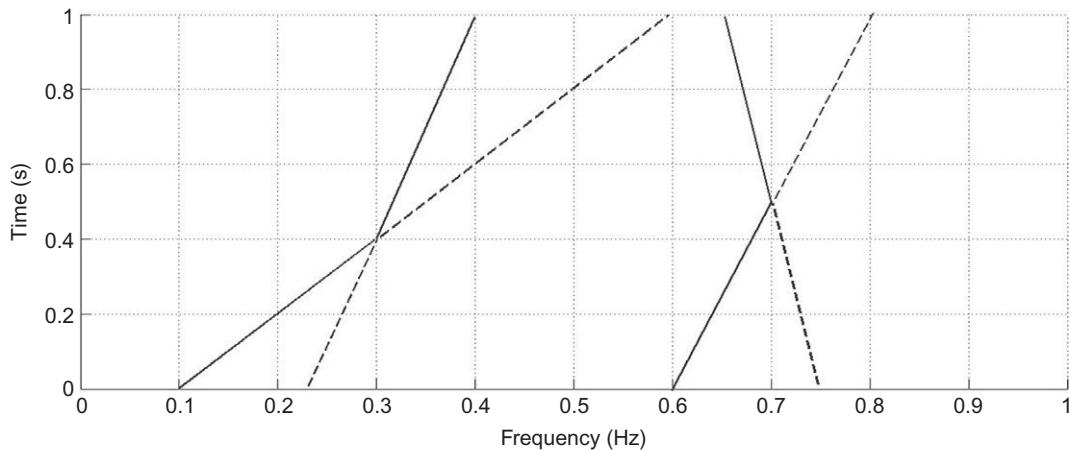
Let us consider a noise-free multicomponent PW-LFM signal $x(t)$, and its analytic associate $z(t)$ expressed as [27]:

$$z(t) = \sum_{i=1}^M z_i(t) = \sum_{i=1}^M \sum_{k=1}^N z_{ki}(t), \quad (5.9.1)$$

where z_{ki} is the k th branch of the i th signal component z_i , with each z_{ki} being expressed by

$$z_{ki}(t) = \begin{cases} e^{j2\pi(\alpha_{ki}t^2/2+f_{0ki}t)+j\phi_{0ki}} & \text{for } t_{0ki} < t < t_{0ki} + T_{ki} \\ 0 & \text{otherwise,} \end{cases} \quad (5.9.2)$$

where $\alpha_{ki}, f_{0ki}, \phi_{0ki}, t_{0ki}$, and T_{ki} are the frequency rate, the starting frequency, the initial phase, the initial time, and the duration of z_{ki} . Figure 5.9.2 shows an example.

**FIGURE 5.9.2**

IF laws of a PW-LFM signal with two components, each with two branches (solid lines). The signal has four distinct directions, as shown by the continuations of the branches (dashed lines).

Before giving the exact model of a PW-LFM signal, let us first introduce a simpler model, based on the ideal case where all the signal components have infinite durations and follow a simple LFM law. The more refined model then considers the general case of PW-LFM signal components for which all the signal components have finite duration and can have different start and stop times.

5.9.2.1 Simplified model: infinite-duration signal, LFM components

To simplify the analysis, all the signal branches are considered as slices of individual LFM components of infinite duration. If there are N components with M branches per component, then the total number of branches is $N_B = N \times M$, and each branch can be re-indexed by a unique integer i where $i = 1, \dots, N_B$. Following Eq. (3.2.4) on p. 111, the ambiguity function of $z(t)$ is then given by [27]

$$A_z(\nu, \tau) = \sum_{i,k=1}^{N_B} \mathcal{F}_{t \rightarrow \nu} \{ z_i(t + \tau/2) z_k^*(t - \tau/2) \}. \quad (5.9.3)$$

In the above, the expressions of the contributions of the auto-terms can be separated from those of the cross-terms by writing

$$A_z(\nu, \tau) = A_z^{(\text{auto})}(\nu, \tau) + A_z^{(\text{cross})}(\nu, \tau) \quad (5.9.4)$$

with

$$A_z^{(\text{auto})}(\nu, \tau) = \sum_{i=1}^{N_B} \mathcal{F}_{t \rightarrow \nu} \{ z_i(t + \tau/2) z_i^*(t - \tau/2) \}$$

$$A_z^{(\text{cross})}(\nu, \tau) = \sum_{\substack{i,k=1 \\ i \neq k}}^{N_B} \mathcal{F}_{t \rightarrow \nu} \left\{ z_i(t + \tau/2) z_k^*(t - \tau/2) \right\}.$$

After some calculations, the last two expressions become:

$$A_z^{(\text{auto})}(\nu, \tau) = \sum_{i=1}^{N_B} e^{j2\pi f_{0i}\tau} \delta(\nu - \alpha_i \tau) \quad (5.9.5)$$

$$\begin{aligned} A_z^{(\text{cross})}(\nu, \tau) &= \sum_{\substack{i,k=1 \\ \alpha_i = \alpha_k}}^{N_B} e^{j\pi(f_{0i} + f_{0k})\tau} [\delta(\nu - \alpha_i \tau - (f_{0i} - f_{0k})) + \delta(\nu - \alpha_i \tau + (f_{0i} - f_{0k}))] \\ &\quad + \sum_{\substack{i,k=1 \\ i < k \\ \alpha_i \neq \alpha_k}}^{N_B} \left(\frac{2}{\sqrt{\alpha_i - \alpha_k}} e^{j2\pi \frac{f_{0i}(\nu - \alpha_k \tau) - f_{0k}(\nu - \alpha_i \tau)}{\alpha_i - \alpha_k}} \right. \\ &\quad \left. \times \cos \left[\frac{\pi}{\alpha_i - \alpha_k} \left((f_{0i} - f_{0k})^2 + (\nu - \alpha_i \tau)(\nu - \alpha_k \tau) \right) - \frac{\pi}{4} \right] \right). \end{aligned} \quad (5.9.6)$$

Inspection of Eqs. (5.9.5) and (5.9.6) leads to the following key observations:

- The auto-terms' support is located on the lines $\nu = \alpha_i \tau$ (also reported in [75]).
- The auto-terms are piecewise-constant functions in the ambiguity domain.
- When $\alpha_i = \alpha_k$, cross-terms behave like auto-terms except that cross-terms support is translated in opposite directions along the ν axis by $|f_{0i} - f_{0k}|$.
- When $\alpha_i \neq \alpha_k$, cross-terms' magnitudes are globally inversely proportional to $\sqrt{|\alpha_i - \alpha_k|}$ and defined by a cosine function that is constant on the support of auto-terms.

Hence, for infinite-duration signals, optimal (t, f) resolution can be achieved by using a kernel that takes into account the (t, f) directions of energy distribution.

5.9.2.2 Exact model: finite duration LFM components starting and ending at different times

Typical PW-LFM signals are composed of finite-duration LFM components appearing as segments in the (t, f) domain. In this section, we use the more concise notation $\Pi_T(t)$ to represent the signal rect $(\frac{t}{T})$. Their AF is given by

$$\begin{aligned} A_z(\nu, \tau) &= \sum_{i,k=1}^{N_B} \mathcal{F}_{t \rightarrow \nu} \left\{ z_i(t + \frac{\tau}{2}) \Pi_{T_i} \left(t - t_{0i} - \frac{T_i}{2} + \frac{\tau}{2} \right) z_k^*(t - \frac{\tau}{2}) \Pi_{T_k} \left(t - t_{0k} - \frac{T_k}{2} - \frac{\tau}{2} \right) \right\} \\ &= \sum_{i,k=1}^{N_B} \mathcal{F}_{t \rightarrow \nu} \left\{ K_{z_i, z_k}(t, \tau) \right\} *_{\nu} \mathcal{F}_{t \rightarrow \nu} \left\{ K_{\Pi_{T_i}, \Pi_{T_k}}(t, \tau) \right\}, \end{aligned} \quad (5.9.7)$$

where $K_{z_i, z_k}(t, \tau)$ and $K_{\Pi_{T_i}, \Pi_{T_k}}(t, \tau)$ are cross-correlation functions defined by:

$$K_{z_i, z_k}(t, \tau) = z_i(t + \tau/2) z_k^*(t - \tau/2), \quad (5.9.8)$$

$$K_{\Pi_{T_i}, \Pi_{T_k}}(t, \tau) = \Pi_{T_i} \left(t - t_{0i} - \frac{T_i}{2} + \frac{\tau}{2} \right) \Pi_{T_k} \left(t - t_{0k} - \frac{T_k}{2} - \frac{\tau}{2} \right). \quad (5.9.9)$$

The FT $\mathcal{F}_{t \rightarrow v} \{ K_{z_i, z_k}(t, \tau) \}$ was already calculated and used for getting Eqs. (5.9.5) and (5.9.6), while the FT $\mathcal{F}_{t \rightarrow v} \{ K_{\Pi_{T_i}, \Pi_{T_k}}(t, \tau) \}$ is given by

$$\mathcal{F}_{t \rightarrow v} \{ K_{\Pi_{T_i}, \Pi_{T_k}}(t, \tau) \} = p_{i,k}(\tau) \text{sinc} (\pi v p_{i,k}(\tau)) e^{-j\pi v q_{i,k}(\tau)}, \quad (5.9.10)$$

where $p_{i,k}(\tau)$ and $q_{i,k}(\tau)$ are PW polynomials introduced to account for all the different cases of cross-correlation between the windows Π_{T_i} and Π_{T_k} . For instance, if two components are defined with $t_{0i} < t_{0k}$, $T_i < T_k$ and $t_{0i} < t_{0k} < t_{0i} + T_i < t_{0k} + T_k$, the polynomials $p_{i,k}$ and $q_{i,k}$ are given by

$$p_{i,k}(\tau) = \begin{cases} t_{0k} + T_k - t_{0i} + \tau & \text{if } -(t_{0k} - t_{0i} + T_k) < \tau < -(t_{0k} - t_{0i} + T_k - T_i), \\ T_i & \text{if } -(t_{0k} - t_{0i} + T_k - T_i) \leq \tau \leq -(t_{0k} - t_{0i}), \\ t_{0i} + T_i - t_{0k} - \tau & \text{if } -(t_{0k} - t_{0i}) < \tau < t_{0i} + T_i - t_{0k}, \\ 0 & \text{otherwise} \end{cases} \quad (5.9.11)$$

and

$$q_{i,k}(\tau) = \begin{cases} t_{0i} + t_{0k} + T_k & \text{if } -(t_{0k} - t_{0i} + T_k) < \tau < -(t_{0k} - t_{0i} + T_k - T_i), \\ 2t_{0i} + T_i - \tau & \text{if } -(t_{0k} - t_{0i} + T_k - T_i) \leq \tau \leq -(t_{0k} - t_{0i}), \\ t_{0i} + t_{0k} + T_i & \text{if } -(t_{0k} - t_{0i}) < \tau < t_{0i} + T_i - t_{0k}, \\ 0 & \text{otherwise.} \end{cases} \quad (5.9.12)$$

Using Eq. (5.9.5), Eq. (5.9.6) for $\alpha_i = \alpha_k$, Eq. (5.9.10), and the expression of the cross-AF A_{z_i, z_k} in the ideal case that leads to Eq. (5.9.6) for $\alpha_i \neq \alpha_k$, the AF of a PW-LFM signal can be decomposed as

$$A_z(v, \tau) = A_z^{(\text{auto})}(v, \tau) + A_z^{(\text{cross})}(v, \tau), \quad (5.9.13)$$

where

$$\begin{aligned} A_z^{(\text{auto})}(v, \tau) &= \sum_{i=1}^{N_B} \mathcal{F}_{t \rightarrow v} \{ K_{z_i, z_i}(t, \tau) K_{\Pi_{T_i}, \Pi_{T_i}}(t, \tau) \} \\ &= \sum_{i=1}^{N_B} e^{j2\pi f_{0i}\tau} \delta(v - \alpha_i \tau) * \left\{ T_i \text{tri}(\tau/T_i) \text{sinc} (\pi v T_i \text{tri}(\tau/T_i)) e^{-j\pi v (2t_{0i} + T_i)} \right\} \\ &= \sum_{i=1}^{N_B} T_i \text{tri}(\tau/T_i) \text{sinc} (\pi(v - \alpha_i \tau) T_i \text{tri}(\tau/T_i)) e^{j2\pi (f_{0i}\tau - (v - \alpha_i \tau)(t_{0i} + \frac{T_i}{2}))}, \end{aligned}$$

$\text{tri}(t)$ being the triangular function defined by $\text{tri}(t) = \max(1 - |t|, 0)$, and where (after more calculations)

$$\begin{aligned} A_z^{(\text{cross})}(v, \tau) &= \sum_{\substack{i,k=1; i < k \\ \alpha_i = \alpha_k}}^{N_B} \left(p_{i,k}(\tau) \text{sinc} (\pi[v - \alpha_i \tau - (f_{0i} - f_{0k})] p_{i,k}(\tau)) \cdot e^{-j\pi [v - \alpha_i \tau - (f_{0i} - f_{0k})] q_{i,k}(\tau)} \right. \\ &\quad \left. + p_{k,i}(\tau) \text{sinc} (\pi[v - \alpha_i \tau + (f_{0i} - f_{0k})] p_{k,i}(\tau)) \cdot e^{-j\pi [v - \alpha_i \tau + (f_{0i} - f_{0k})] q_{k,i}(\tau)} \right) e^{j\pi (f_{0i} + f_{0k}) \tau} \end{aligned}$$

$$\begin{aligned}
& + \frac{e^{-j\frac{\pi}{4}}}{2} \sum_{\substack{i,k=1 \\ \alpha_i \neq \alpha_k}}^{N_B} \frac{e^{-j\pi \frac{(f_{0i}-f_{0k})^2}{\alpha_i - \alpha_k}}}{\sqrt{\alpha_i - \alpha_k}} e^{\frac{j2\pi}{\alpha_i - \alpha_k} (-\alpha_i \alpha_k \tau^2 + \frac{\alpha_i \alpha_k}{2} \tau v + (f_{0i} - f_{0k})v - \frac{v^2}{2})} \\
& \cdot \left(\operatorname{erfi} \left[\sqrt{\frac{\pi}{\alpha_i - \alpha_k}} e^{j\frac{\pi}{4}} \frac{(p_{i,k}(\tau) - q_{i,k}(\tau))}{2} (\alpha_i - \alpha_k) - \frac{\alpha_i \alpha_k}{2} \tau - (f_{0i} - f_{0k}) + v \right] \right. \\
& \left. - \operatorname{erfi} \left[\sqrt{\frac{\pi}{\alpha_i - \alpha_k}} e^{j\frac{\pi}{4}} \left(-\frac{p_{i,k}(\tau) + q_{i,k}(\tau)}{2} (\alpha_i - \alpha_k) - \frac{\alpha_i \alpha_k}{2} \tau - (f_{0i} - f_{0k}) + v \right) \right] \right).
\end{aligned}$$

5.9.2.3 Exact model: behavior and support of auto- and cross-terms

1. Support and behavior of (magnitudes of) auto-terms of a PW-LFM signal:
 - (a) Their support is located around the lines $v = \alpha_i \tau$ with $|\tau| < T_i$.
 - (b) They are defined by sinc functions with maxima on the lines $v = \alpha_i \tau$ with $|\tau| < T_i$.
 - (c) As we move away from the origin of the ambiguity domain along the lag axis, the sinc functions main lobes are characterized by
 - (i) an increasing width as τ varies from 0 to T_i , and
 - (ii) a decreasing maximum from the value T_i to 0 at $\tau = \pm T_i$.
2. Support and behavior of (magnitudes of) cross-terms of a PW-LFM signal:
 - (a) When $\alpha_i = \alpha_k$, cross-terms behave like the auto-terms except that the cross-terms support is translated in opposite directions by $|f_{0i} - f_{0k}|$ along the Doppler axis, and by a distance dependent on the value of the polynomial $p_{i,k}(\tau)$ along the lag axis.
 - (b) When $\alpha_i \neq \alpha_k$, cross-terms' magnitudes are globally inversely proportional to $\sqrt{|\alpha_i - \alpha_k|}$; cross-terms in this case are defined by oscillatory functions whose characteristics need to be further analyzed.

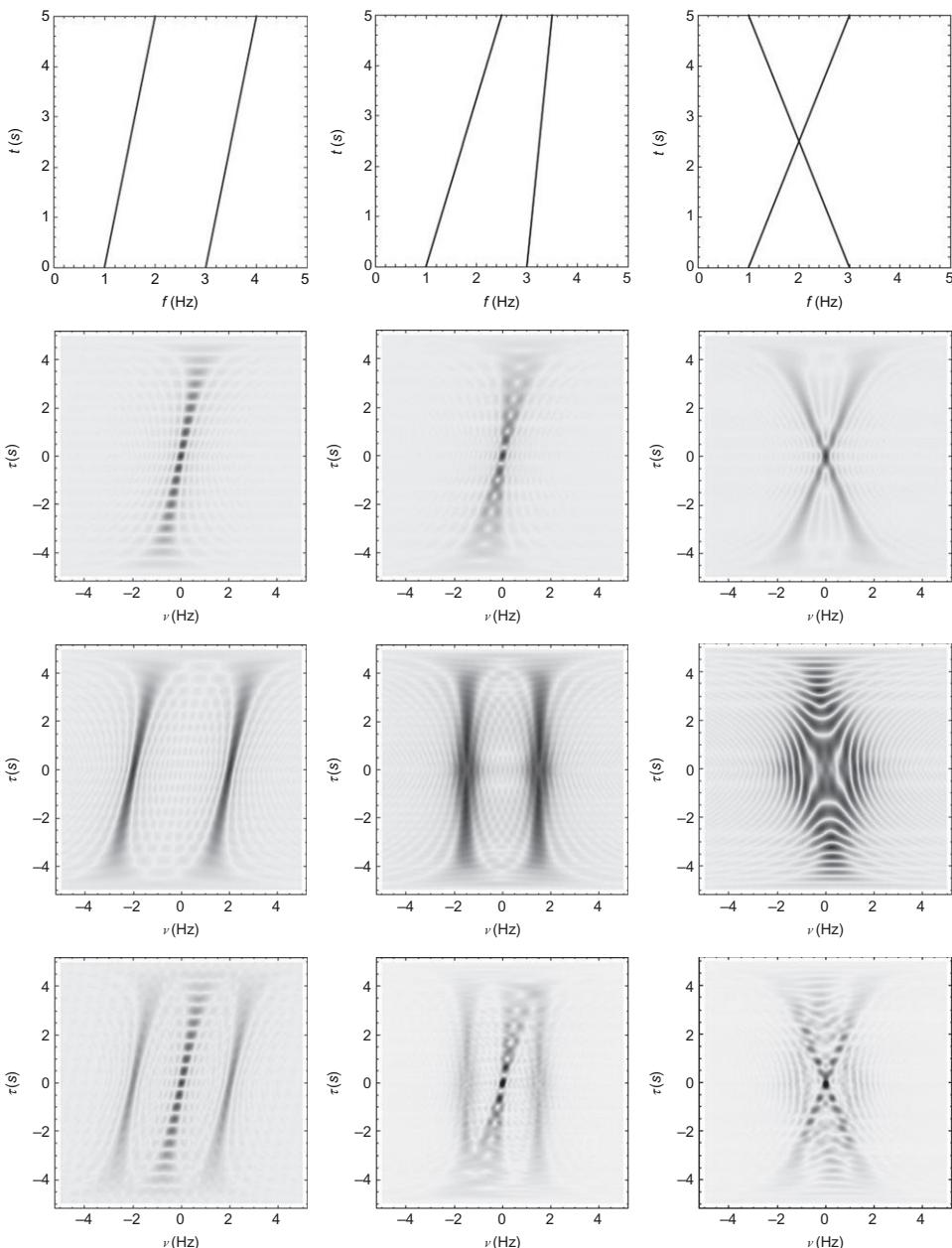
Figure 5.9.3 shows the support and behavior of both auto-terms and cross-terms for three configurations of a multicomponent signal having two LFM components. From the considered cases of the exact model, three main observations can be made:

1. When two signal components are parallel in the (t, f) domain, cross-terms are well detached from auto-terms in the ambiguity domain.
2. When two signal components are neither secant, nor parallel in the (t, f) domain, cross-terms and auto-terms overlap in the ambiguity domain at the level of the support ends of the auto-terms.
3. When two signal components intersect in the (t, f) domain, cross-term energy near the ambiguity-domain origin is more significant than in the two other cases, and will be visible in the (t, f) domain, but remains relatively small by comparison with the energy of auto-terms.

5.9.2.4 Application to kernel design in the ambiguity domain

The following are key properties of auto-terms and cross-terms in the ambiguity domain from the perspective of kernel design, using the above insights.

1. The energy of auto-terms is concentrated along lines passing through the origin in the ambiguity domain.
2. The support of auto-terms in the ambiguity domain along their major and minor axes depends on the signal components duration; for example, for a given value of frequency rate, shorter-duration

**FIGURE 5.9.3**

Density plots of $|A_z^{(\text{auto})}|$ and $|A_z^{(\text{cross})}|$ for a finite-duration signal with two directions, modeled by LFM components starting and ending at the same times. *Left column:* parallel directions; *center column:* nonparallel, nonsecant directions; *right column:* secant directions. *Top row:* IF law; *second row:* $|A_z^{(\text{auto})}|$; *third row:* $|A_z^{(\text{cross})}|$; *fourth row:* $|A_z^{(\text{auto})} + A_z^{(\text{cross})}|$.

signals have relatively smaller support along their major axes and larger half-height widths along their minor axes.

3. The auto-terms' magnitudes peak at the origin and decrease as we move away from the origin along the auto-term lines.
4. Cross-terms caused by the interaction between *parallel* LFM signal components do not intersect with the corresponding auto-terms in the ambiguity domain. Note that they may intersect with other auto-terms that are not parallel with them. Such a scenario is illustrated in [Section 5.9.4](#).
5. Cross-terms caused by the interaction between *nonparallel* LFM signal components intersect with the corresponding auto-terms in the ambiguity domain.

Most previous kernel design methods simply assume that auto-terms are located around the origin while cross-terms are away from the origin. This is a rough approximation of the distribution of auto-term energy and cross-term energy in the ambiguity domain and it is only valid for signals that are well separated in the (t,f) domain. In order to design high-resolution TFDs for signals that are “closely spaced” in the (t,f) domain, it is important to take into account the first three properties that precisely describe the location and orientation of auto-terms in the ambiguity domain. The last two properties indicate scenarios when cross-terms overlap with auto-terms in the ambiguity domain. For such cases, no kernel-based TFD can lead to optimal resolution as it becomes impossible to completely eliminate cross-terms while retaining all auto-term energy.

In general, an optimal fixed-kernel QTFD for a PW-LFM model should therefore verify that:

1. It has a multidirectional kernel; that is, it should have parameters that allow and control filtering along multiple directions in the ambiguity domain.
2. It has parameters to adjust the support of its kernel along both the major axis and the minor axis of auto-terms.

5.9.3 DEFINING A HIGH-RESOLUTION MULTIDIRECTIONAL TFD

In general, most real-life signals are multicomponent and may have several specific directions of energy concentration in the (t,f) domain (see, e.g., [Fig. I.1.1](#), which shows the TFD of a whale song). To account for such directionality, let us define a multidirectional TFD (MDD), which can be formulated as [\[62\]](#)

$$\rho_z(t,f) = W_z(t,f) \underset{(t,f)}{\ast\ast} \gamma_\theta(t,f), \quad (5.9.14)$$

where γ_θ is here a directional nonseparable kernel in the (t,f) domain, dependent on the set $\theta = (\theta_1, \theta_2, \dots, \theta_P)$ of direction angles of the signal, defined as the angles between the lag axis and the auto-term directions in the ambiguity domain. Using the “diamond diagrams” in Eqs. [\(3.2.5\)](#), [\(3.2.6\)](#), and [\(3.2.10\)](#), we can express Eq. [\(5.9.14\)](#) in the ambiguity domain as:

$$\mathcal{A}_z(v, \tau) = A_z(v, \tau) g_\theta(v, \tau), \quad (5.9.15)$$

where here $g_\theta(v, \tau)$ is similarly a nonseparable kernel in the Doppler-lag domain, also dependent on the set of signal direction angles θ .

Let us construct a multidirectional kernel (MDK) using first the compact support kernel (CSK) [76] as a building block, given its ability to fix its support while changing its shape. The MDK is then defined as [62]:

$$g_{\theta}(\nu, \tau) = \frac{e^{c_0+c}}{P} \sum_{i=1}^P g_{\theta_i}(\nu, \tau) h_{\theta_i}(\nu, \tau), \quad (5.9.16)$$

where the factor e^{c_0+c}/P in front of the summation is a normalization coefficient, c_0 and c are slope-adjustment parameters, and θ_i is related to the frequency rate α_i by $\alpha_i = \tan(\theta_i)$. The factor $g_{\theta_i}(\nu, \tau)$ is the i th branch of the MDK, which is rotated in the ambiguity domain by angle θ_i ,

$$g_{\theta_i}(\nu, \tau) = \begin{cases} e^{\left(\frac{c}{D_i} \frac{\cos(\theta_i)\nu - \sin(\theta_i)\tau}{\sqrt{1 + \cos^2(\theta_i)}}\right)^2 - 1}, & |\cos(\theta_i)\nu - \sin(\theta_i)\tau| < D_i, \\ 0, & \text{otherwise,} \end{cases}$$

where D_i is the half-support of $g_{\theta_i}(\nu, \tau)$ along the direction perpendicular to the i th branch of the MDK, and $h_{\theta_i}(\nu, \tau)$ is the Doppler-lag window for the i th branch, that is,

$$h_{\theta_i}(\nu, \tau) = \begin{cases} e^{\left(\frac{c_0}{E_i} \frac{\sin(\theta_i)\nu + \cos(\theta_i)\tau}{\sqrt{1 + \sin^2(\theta_i)}}\right)^2 - 1}, & |\sin(\theta_i)\nu + \cos(\theta_i)\tau| < E_i, \\ 0, & \text{otherwise,} \end{cases}$$

where E_i is related to either the time duration of LFM components or the bandwidth of spike components; that is, “ E_i ” controls the length of the i th branch of the MDD kernel along its major axis. This MDK is shown in [Fig. 5.9.4](#) for different values of c .

5.9.4 MDD PERFORMANCE: RESULTS AND DISCUSSIONS

5.9.4.1 Performance validation using a synthetic signal

Let us consider a simulated signal with two different directions of energy concentration in the (t, f) domain. The signal is composed of two tones and two parallel LFM signals. [Figure 5.9.5](#) shows that

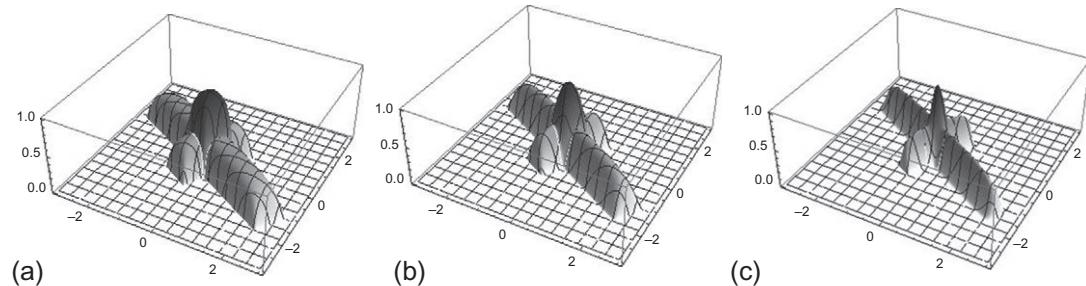
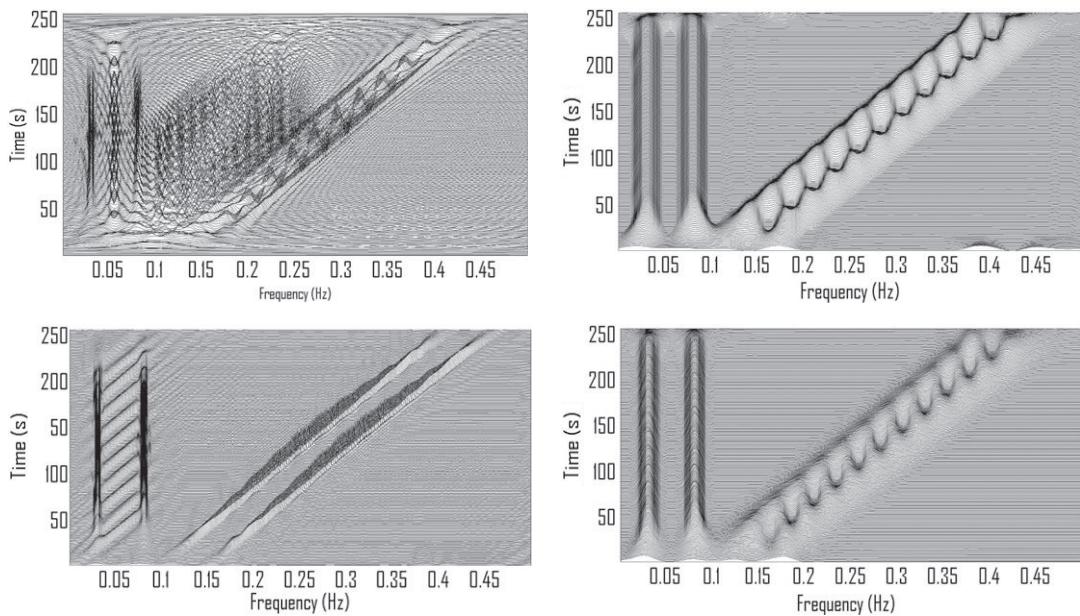


FIGURE 5.9.4

The multidirectional kernel for $P = 2$, $\theta_1 = \pi/6$, $E_1 = 2/\sqrt{3}$, $\theta_2 = 2\pi/3$, $E_2 = 3$, $D_1 = D_2 = 0.5$, $T_1 = 1$ s, $T_2 = 1.5$ s, $c_0 = 0.1$, and (left) $c = 0.25$, (center) $c = 1$, and (right) $c = 5$.

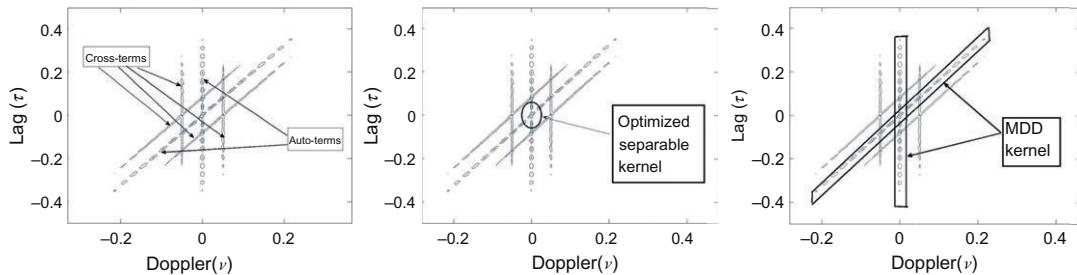
**FIGURE 5.9.5**

Comparison between MDD and reference TFDs of a four-component signal. *Top left:* WVD; *Top right:* CKD ($c = 1$, $D = 0.09$, $E = 0.05$); *Bottom left:* MDD ($\theta_1 = 0^\circ$, $\theta_2 = 58^\circ$, $c = 1$, $D = 0.015$); *Bottom right:* S-method (4 s Hamming window).

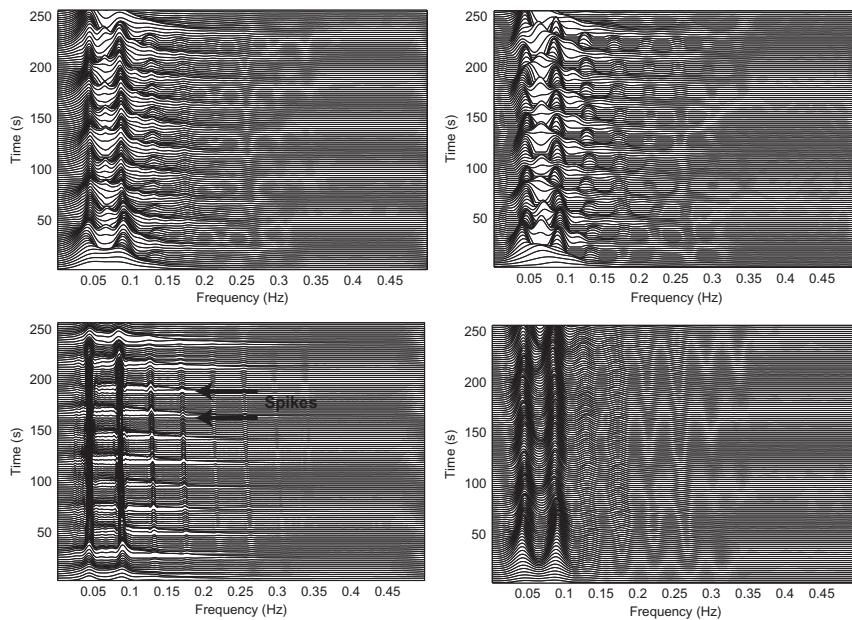
for this signal, the MDD outperforms the CKD, the S-method and of course the WVD (Chapters 2 and 3). Most undesirable cross-terms here mainly stem from the interaction of parallel tones and LFM components, as shown in the ambiguity domain in the left panel of Fig. 5.9.6. The optimized CKD kernel support depicted in the middle panel shows that the CKD cannot give a high concentration of auto-terms' energy without being affected by cross-terms. The MDD kernel performs better here because it can follow the directions of auto-terms. Its support can be directed along the lines of auto-terms, and can be adjusted to retain all their energy without being too much affected by the cross-terms due to interaction of closely spaced parallel signal components that are tones and LFM components (Fig. 5.9.6, right panel); however, the MDD does not completely remove the cross-terms; some of them are retained due to the overlapping of cross-terms with auto-terms in the ambiguity domain.

5.9.4.2 Application to EEG seizure signals

Let us consider an EEG seizure signal with both spike and LFM characteristics. This signal has two orthogonal directions of energy concentration in the (t,f) domain: one is almost parallel to the time axis due to the LFM characteristics while the other one is parallel to the frequency axis due to spikes (which are a train of impulses). Figure 5.9.7 shows that the EMBD and CKD are unable to concentrate the energy of the auto-terms of the tones and the spikes after the suppression of the cross-terms. The S-method gives high energy concentration for tones but fails to concentrate energy for the spikes. The MDD here again outperforms the three reference TFDs as it concentrates the energy of both tones and spikes more while rejecting the cross-terms caused by interaction among all the components.

**FIGURE 5.9.6**

Signal characteristics and kernel representation in the ambiguity domain. *Left*: modulus of the signal's AF; *Middle*: optimal CKD kernel; and *Right*: MDD kernel.

**FIGURE 5.9.7**

Comparison between MDD and reference TFDs of an EEG seizure signal. *Top left*: EMBD (separable kernel); $\alpha = 0.1, \beta = 0.1$; *Top right*: CKD ($c = 1, D = 0.1, E = 0.1$); *Bottom left*: MDD ($\theta_1 = 0^\circ$ (tones), $\theta_2 = 90^\circ$ (spikes), $c = 5, D = 0.03$); and *Bottom right*: S-Method (4 s Hamming window).

5.9.5 SUMMARY AND CONCLUSIONS

This section presents a detailed analysis of the geometry of cross-terms and auto-terms for PW-LFM signals in the ambiguity domain. The PW-LFM model is selected as it can be used to represent many real-life signals such as heart rate variability and EEG seizure signals. The ambiguity-domain analysis of the PW-LFM signals shows that the auto-terms energy is concentrated along truncated lines in the

ambiguity domain, while cross-terms can intersect with auto-terms for nonparallel signal components. Based on these observations, a high-resolution MDD is defined, which has the flexibility to adapt the shape of its kernel along all principal directions in the ambiguity domain. The MDD outperforms separable-kernel TFDs, such as the EMBD and CKD in terms of resolution and cross-term suppression for both synthetic and real-life EEG seizure signals.

5.10 ADAPTIVE DIRECTIONAL TIME-FREQUENCY DISTRIBUTIONS⁰

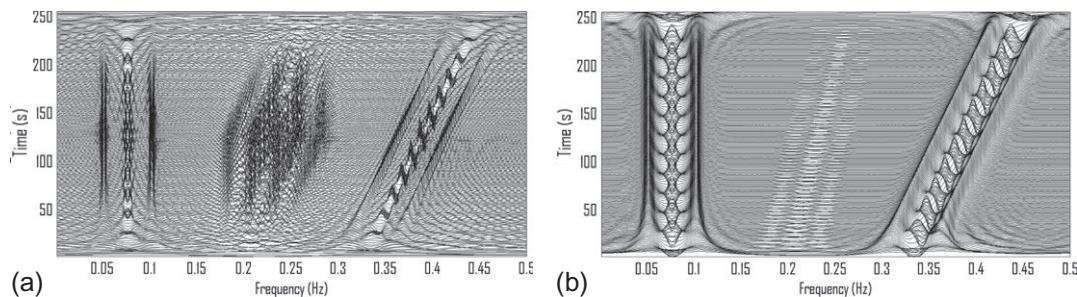
QTFDs are often assessed in terms of their ability to resolve closely spaced signal components while suppressing undesired cross-terms. The cross-terms are usually suppressed by 2D smoothing [75], but this reduces the resolution of auto-terms. Section 5.7 has shown that TFDs based on separable kernels offer higher resolution as compared to the spectrogram, due to the extra degree of freedom, namely independent control of smoothing along time and frequency axes; however, these methods fail to provide optimal energy concentration when the signal energy is concentrated in a certain direction in the (t,f) domain. In this scenario, (t,f) kernels rotated along the direction of Energy distribution can significantly suppress cross-terms while retaining the resolution of auto-terms [77]. This method of rotating a kernel along the major axis of the auto-terms cannot be applied to signals having multiple directions of ED in the (t,f) domain [77]. To deal with such cases, the shape of the smoothing kernel could be adapted locally either at each time instant (e.g., the adaptive optimal kernel approach [34]) or at each (t,f) point (e.g., the adaptive fractional spectrogram [78]).

5.10.1 THE NEED FOR ADAPTIVE DIRECTIONAL KERNELS

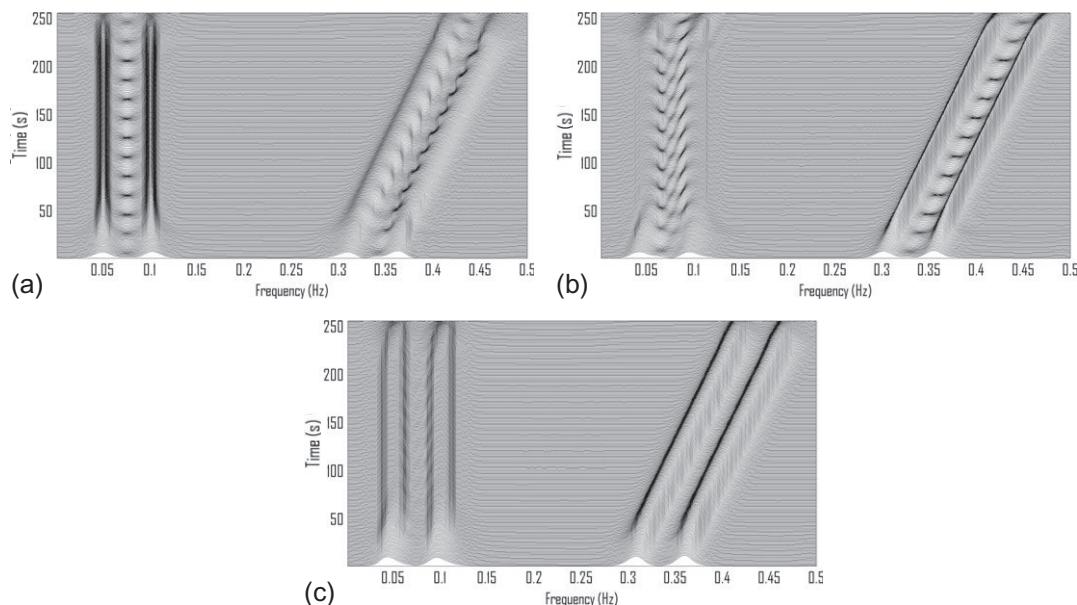
Let us consider two sample cases where signal representations in the (t,f) domain can be improved if adaptive kernels are used. Let us first consider a signal composed of two parallel LFM chirps and two tones. Figure 5.10.1 shows the WVD and the EMBD of the given signal. The (t,f) representation given by the WVD is more complicated to interpret due to the presence of cross-terms. The Extended Modified B Distribution (EMBD) has improved the signal representation by reducing cross-terms generated by the interaction between the LFM components and tones; however, the EMBD still contains cross-terms generated by the interaction between two close components; that is, cross-terms appearing between the IFs of the two close LFM signals and between the IFs of the two close tones.

To further suppress cross-terms in the EMBD, a single directional kernel cannot be used for directional smoothing as the LFM signals and two tones are not in the same direction (Fig. 5.10.2(a)). For example, directional smoothing in the direction of the major axes of the two tones suppresses the cross-terms due to the tone interaction, but not the cross-terms between the LFM signals (Fig. 5.10.2(b)). In this case, successive application of two directional filters, one parallel to the tones and another parallel to the LFM signals, can remove the cross-terms without degrading the resolution of auto-terms (Fig. 5.10.2(c)).

⁰Authors: **N.A. Khan**, Qatar University, Doha, Qatar (nabeelalikhan@qu.edu.qa), **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa), and **E. Sejdic**, University of Pittsburgh, Pittsburgh, PA, USA (esejdic@ieee.org). Reviewers: S. Djukanovic and S. Aviyente.

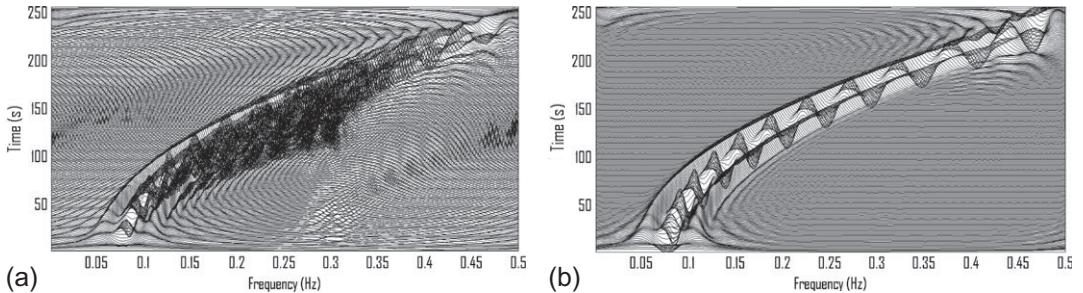
**FIGURE 5.10.1**

TFDs of a four-component signal composed of two tones and two LFM components: (a) WVD and (b) EMBD.

**FIGURE 5.10.2**

Directionally smoothed TFDs of a four-component signal composed of two tones and two LFM components: (a) EMBD filtered by directional Gaussian filter parallel to time axis; (b) EMBD filtered by directional Gaussian filter parallel to major axes of LFM signals; and (c) TFD obtained by first applying directional filter along time axis, then along major axes of LFM signals.

A more complicated case involves a signal composed of two quadratic frequency modulated components as shown in Fig. 5.10.3. The WVD is difficult to interpret because of both inner and outer interferences. The EMBD, by suppressing inner interference terms, shows the two quadratic FM signal components and oscillatory cross-terms between the two components. It is observed that the direction of oscillation of cross-terms changes with time; the cross-terms cannot therefore be suppressed by the

**FIGURE 5.10.3**

TFDs of a multicomponent signal composed of two quadratic chirps: (a) WVD and (b) EMBD.

application of one or two directional filters as done in earlier examples. In this scenario, we need a smoothing kernel that adapts itself locally at each time instant or at each (t,f) point, as discussed in the following sections.

5.10.2 ADAPTIVE DIRECTIONAL (t,f) DISTRIBUTIONS

Adapting the orientation and the shape of a smoothing kernel can be performed using the following two approaches:

1. *Time-varying local adaptation*: The shape of a smoothing kernel is adapted for each time instant, excluding the f variable.
2. *(t,f) -varying local adaptation*: The shape of a smoothing kernel is adapted for each (t,f) point [62].

Accounting for both cases, and others, a general formulation of an adaptive TFD may be given as:

$$\rho^{(\text{adapt})}(t,f) = \rho(t,f) *_{t,f} \gamma(t,f), \quad (5.10.1)$$

where $\gamma(t,f)$ is a smoothing kernel, $\rho(t,f)$ is a TFD, and $\rho^{(\text{adapt})}(t,f)$ is the adaptive TFD.

5.10.2.1 Time-varying local adaptation

The adaptive optimal kernel (AOK) TFD, defined in [Section 5.3](#), adapts the shape of its smoothing kernel at each time instant using the short-time AF. In this case, the windowed WVD, as computed from the short-time AF, is the time-varying (t,f) smoothing kernel related to the Doppler-lag kernel by the following expression (see [34] or [Section 5.3](#)):

$$\gamma_t(t,f) = \mathcal{F}_{t \leftarrow v}^{-1} \mathcal{F}_{\tau \rightarrow f} \exp \left(-\frac{v^2 + \tau^2}{2\sigma^2(\phi)} \right), \quad (5.10.2)$$

where $\sigma^2(\phi)$ controls the spread of the kernel in each direction based on the constraints described in [Section 5.3](#). \mathcal{F} and \mathcal{F}^{-1} are the forward and inverse Fourier operators, respectively. The optimization of the kernel shape is carried out by maximizing the correlation of the kernel with the short-time AF

of the given multicomponent signal at each time instant. The computational cost of this algorithm is $O(M^3)$ [76], where M is the number of samples in the frequency domain.

5.10.2.2 (t, f) -Varying local adaptation

5.10.2.2.1 Adaptive fractional spectrogram

Another approach is to define an adaptive fractional spectrogram by analyzing a signal using an adaptive window such that the window parameters are optimized at each (t, f) point [78]:

$$\text{AFS}(t, f) = \left| \int s(\tau) h_{\alpha, \sigma}(\tau - t) e^{-j2\pi f \tau} d\tau \right|^2, \quad (5.10.3)$$

where

$$h_{\alpha, \sigma}(t) = \frac{e^{j\alpha/2}}{\sqrt{j \sin \alpha}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\sigma^2}} e^{j\pi \frac{(u^2 + t^2) \cos \alpha - 2ut}{\sin \alpha}} du \quad (5.10.4)$$

is the signal-dependent analysis window, α is the chirp rate or the rotation order of the analysis window, and σ is the standard deviation of the Gaussian window. Both α and σ are adapted based on the observation that the ideal window for all those points that lie along the IF of a signal component is the one that maximizes the correlation of the signal with the window. This criterion is expressed as:

$$h_{\alpha, \sigma}(t) = \arg \max_{h_{\alpha, \sigma}(t)} \left| \int s(\tau) h_{\alpha, \sigma}(\tau - t) e^{-j2\pi f \tau} d\tau \right|^2. \quad (5.10.5)$$

This criterion is not optimal for those points that do not exactly lie on the IF of the signal component. Hence, the adaptive fractional spectrogram leaks the signal energy at those (t, f) points where no signal is present. As it is difficult to find an analytical solution to Eq. (5.10.5), this equation is numerically solved using the procedure outlined in [78].

In general, the implementation of the adaptive fractional spectrogram requires the computation of K spectrograms. The computational cost of computing a single spectrogram is $O(NM \log M)$, where M is the number of samples in the frequency domain and N is the length of the signal. The total computational cost of the adaptive fractional spectrogram becomes then $O(KNM \log M)$.

5.10.2.2.2 Directional Gaussian filter-based TFD

Based on the sample signals shown in Figs. 5.10.1–5.10.3, there are two major observations:

- *Cross-terms oscillate along their major axes.* The main implication of this observation is that smoothing the cross-terms along their major axes will reduce them, as smoothing acts as a low-pass filter. The cut-off frequency of such a low-pass filter is dependent on the separation of the two closest signal components, as the rate of oscillation of cross-terms is directly proportional to the spacing between signal components. Note that a narrowband filter with a longer duration would be better in resolving close frequency components, but such a filter spreads the signal energy for short-time components.
- *Auto-terms vary slowly along their major axes.* The major implication of this observation is that the smoothing (i.e., low-pass filtering) operation along the auto-terms' major axes does not significantly degrade the auto-terms' appearance as these terms are in the pass region of the

smoothing (filtering) operation. A general 2D smoothing has more severe effects on the appearance of auto-terms.

The above observations imply that smoothing along the major axes of auto- and cross-terms will reduce cross-terms without severely degrading the auto-terms. Hence, the problem of estimating the direction of the smoothing kernel is now reduced to the estimation of local directions of major axes (i.e., at each (t,f) point). Both auto-terms and cross-terms appear as ridges in the magnitude (t,f) domain, as taking the absolute value of a TFD removes cross-term oscillations. The local directions can be estimated by maximizing the convolution value for a directional kernel convolved with a magnitude of a quadratic TFD, as a directional kernel would have a maximum value when parallel to ridges (i.e., the major axes of cross-terms and auto-terms). Mathematically,

$$\theta(t,f) = \arg \max_{\theta} \left| |\rho(t,f)| *_{tf} \gamma_{\theta}(t,f) \right|. \quad (5.10.6)$$

Based on the above expression, a smoothing kernel should therefore have the following properties:

- *Good localization.* The kernel should have the maximum response when aligned with ridges. This implies that the kernel should have low-pass characteristics along its major axis. The kernel should also be symmetric along its major axis to match the shape of a ridge.
- *Minimal number of false responses.* The kernel output should be zero for nonridge points, as otherwise spectral leakages would be observed at (t,f) points where no signal is present.

The double-derivative directional Gaussian filter (DGF), commonly used in image processing, approximately fulfills the above properties. It is defined as [79]:

$$\gamma_{\theta}(t,f) = \frac{d^2}{df_{\theta}^2} \left\{ e^{-a^2 t_{\theta}^2 - b^2 f_{\theta}^2} \right\}, \quad (5.10.7)$$

where $t_{\theta} = t \cos(\theta) + f \sin(\theta)$ and $f_{\theta} = -t \sin(\theta) + f \cos(\theta)$. Note that the DGF has low-pass characteristics along its major axis (i.e., $e^{-a^2 t_{\theta}^2}$) while it performs second-order differentiation along its minor axis (i.e., $d^2 \{e^{-b^2 f_{\theta}^2}\} / df_{\theta}^2$). This filter has a maximum response when parallel to ridges, and zero response when orthogonal to ridges. The parameters a and b control smoothing along the major and minor ridge axes. Small parameter values imply more smoothing along the respective axes, while large values imply less smoothing. Optimal parameter values depend on the nature of a signal (further details regarding the trade-offs involved in the selection of optimal parameters for the DGF-TFD can be found in [80]). For example, we have experimentally observed that $a = 3$ and $b = 8$ provide good results for the analysis of real-life EEG seizure signals.

5.10.2.3 Algorithm for implementing a DGF-based TFD

To implement a DGF-based TFD, the following steps should be adopted:

1. Analyze a given signal using a quadratic TFD.
2. Convolve the magnitude of the quadratic TFD using K DGFs:

$$\rho_k(t,f) = |\rho(t,f)| *_{tf} \gamma_{\theta_k}(t,f) \quad (5.10.8)$$

with $\theta_k = 2\pi k/K$ for $0 \leq k \leq K - 1$, where K is the total number of quantization levels for the angle θ . In this study, 16 quantization levels were used.

3. Estimate $\theta(t,f)$ at each (t,f) point by choosing k that maximizes the magnitude of a directionally smoothed TFD:

$$\theta(t,f) = \frac{2\pi}{k} \arg \max_k |\rho_k(t,f)|. \quad (5.10.9)$$

4. Using $\theta(t,f)$, calculate the DGF-based adaptive TFD by convolving the adaptive DGF, $\gamma_\theta(t,f)$, with the quadratic TFD as expressed in Eq. (5.10.1).

To understand the computational cost of implementing the DGF-based TFD, let us consider the steps involved:

1. Computation of a quadratic TFD with the computational cost equal to $O(NM \log M)$.
2. Estimation of the direction of the smoothing kernel: This step involves K convolution operations, so the computational cost of this step is $O(KMN)$, as the computational cost of the single convolution operation is $O(MN)$ if there are MN points in the (t,f) domain.
3. Convolving the quadratic TFD with the adaptive DGF: The computational cost of this step is $O(MN)$, as it involves a single convolution operation.

Therefore, the total computation cost is equal to the sum of the computational costs of all the above-mentioned steps, that is, $O(NM \log M + (K + 1)MN)$.

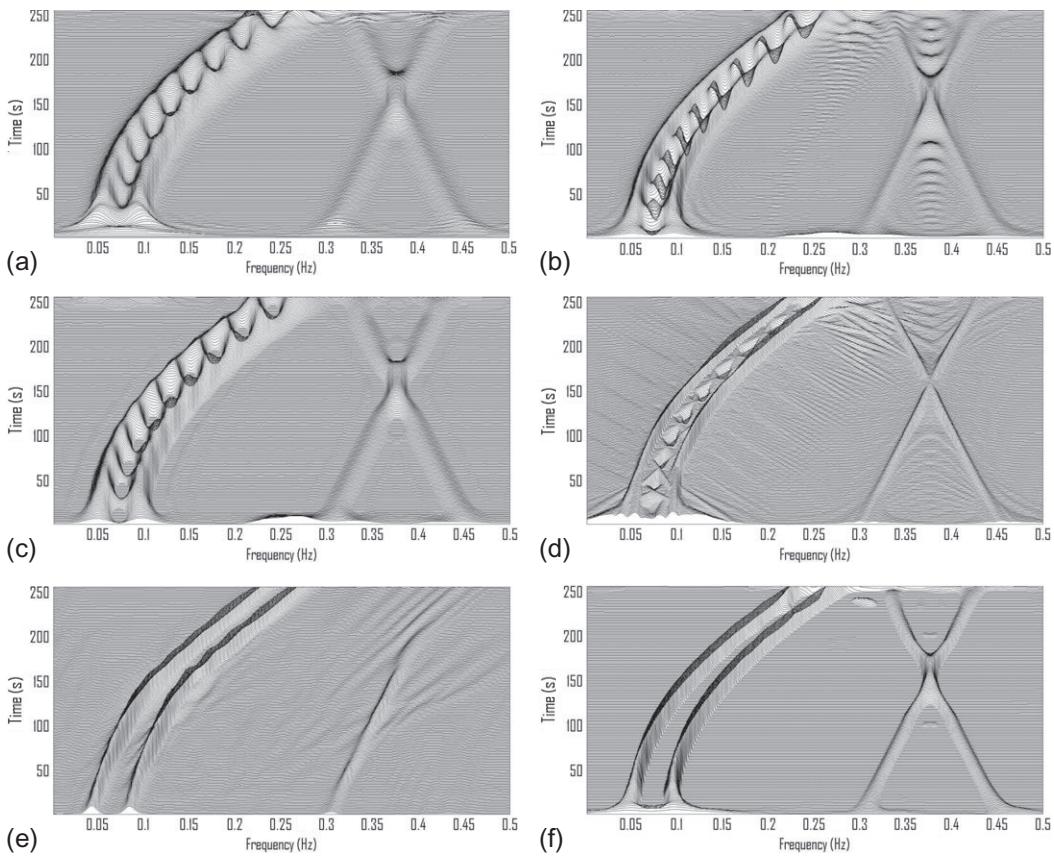
5.10.3 ILLUSTRATIVE EXAMPLES

To assess the performance of these adaptive algorithms, we consider a multicomponent signal composed of two quadratic chirps and two crossing LFM signals:

$$x(n) = x_1(n) + x_2(n) + x_3(n) + x_4(n), \quad (5.10.10)$$

where $x_1(n) = 2 \cos(0.1\pi n + 0.000002\pi n^3)$, $x_2(n) = 2 \cos(0.18\pi n + 0.000002\pi n^3)$, $x_3(n) = \cos(0.6\pi n + 0.0005\pi n^2)$, and $x_4(n) = \cos(0.9\pi n - 0.0005\pi n^2)$ for $0 \leq n \leq 255$. The signal is challenging for most TFDs as the instantaneous frequencies of its components have nonlinear variations. Furthermore, the signal energy is not uniformly distributed among all signal components; the quadratic chirps carry more energy than the crossing LFM chirps. To illustrate the performance of various TFDs, this signal is analyzed using several TFD approaches as shown in Fig. 5.10.4.

Figure 5.10.4 indicates that the fixed-kernel TFDs including the spectrogram, the EMBD, and the compact-support-kernel TFD have failed to resolve the two closely spaced FM quadratic chirps, as these methods cannot adapt the direction of the smoothing kernel. The adaptive fractional spectrogram has resolved closely spaced quadratic chirps, but it suffers from the energy leakage problem; we can observe the signal energy between the instantaneous frequencies of the two quadratic chirps. The AOK-TFD resolves the two quadratic chirps, but fails to concentrate energy for the crossing LFM components as the shape of the adaptive optimal kernel is adapted at each time instant, but not at each (t,f) point. Hence, the AOK-TFD fails to concentrate energy for weak signal components. On the other hand, the DGF-based TFD has resolved closely spaced signal components and it gives clear representation for crossing LFM components.

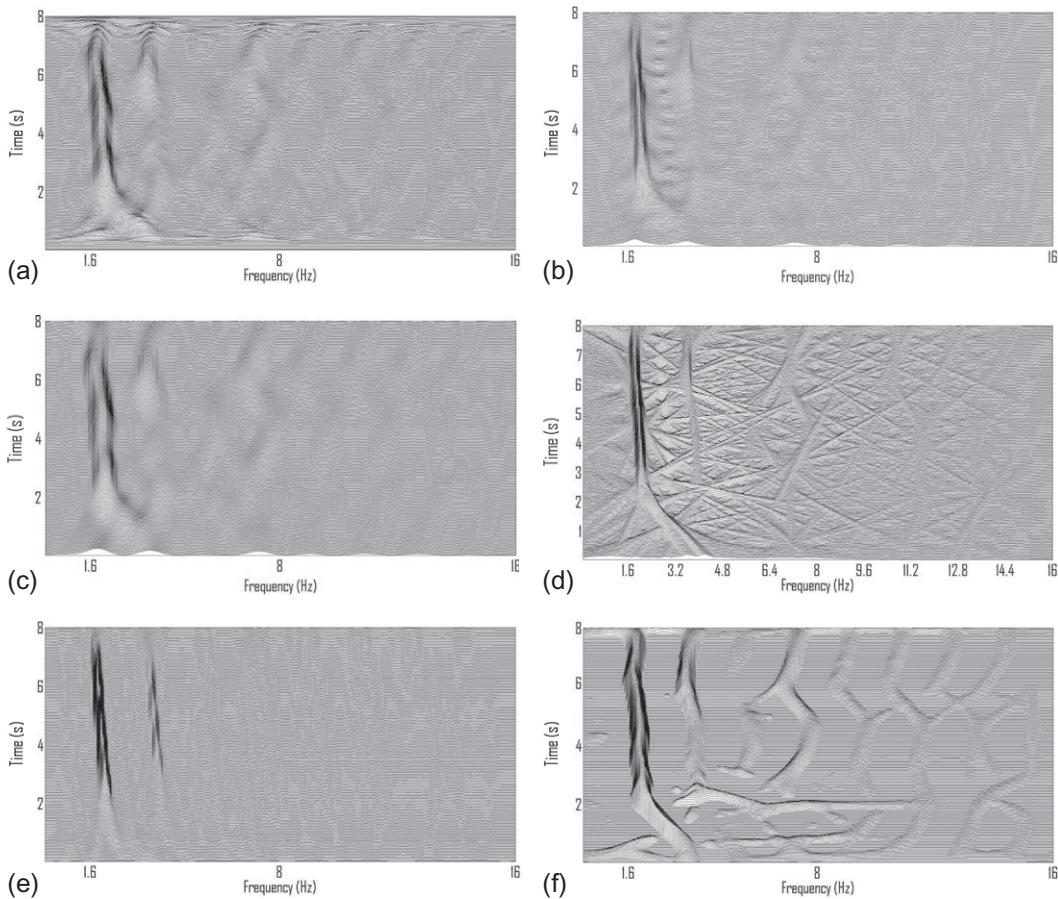
**FIGURE 5.10.4**

TFDs of a four-component signal composed of two quadratic chirps and two crossing LFM components:
 (a) spectrogram (65-point Hamming window); (b) EMBD ($\alpha = 0.07$, $\beta = 0.15$); (c) compact-support-kernel TFD ($C = 1$, $D = 0.06$, $E = 0.1$); (d) adaptive fractional spectrogram; (e) AOK-TFD; and (f) DGF-based TFD ($a = 3$, $b = 12$).

Let us now consider a real-life 8-s long seizure EEG signal. Figure 5.10.5 shows that the spectrogram, EMBD, compact-support-kernel TFD, and AOK-TFD provide high-energy concentration for the two strong signal components, but fail to concentrate signal energy for the weak (higher-frequency) components. The adaptive fractional spectrogram is difficult to interpret because of the leakage of the signal energy at those (t, f) points where there is no signal. The DGF-based TFD provides a clear representation of all the signal components including the weak ones.

5.10.4 SUMMARY AND CONCLUSIONS

A class of adaptive directional TFDs is described in this section. These TFDs adapt the direction of a smoothing kernel based on the distribution of the signal energy in the (t, f) domain.

**FIGURE 5.10.5**

TFDs of a sample EEG signal: (a) spectrogram (95-point Hamming window); (b) EMBD ($\alpha = 0.15, \beta = 0.05$); (c) compact-support-kernel TFD ($C = 1, D = 0.06, E = 0.1$); (d) adaptive fractional spectrogram; (e) AOK-TFD; and (f) DGF-based TFD ($a = 3, b = 8$).

The analysis of the location and orientation of cross-terms in quadratic TFDs defines several constraints for adapting the direction of a smoothing kernel [80], that is: (1) For parallel LFM signals, a global directional kernel aligned with the major axes of the auto-terms can reduce cross-terms without deteriorating the resolution of auto-terms. (2) For LFM signals with more than one direction of energy concentration, successive applications of multiple directional filters should be performed. (3) For nonlinearly frequency-modulated signals, local adaptation of the smoothing kernel at each (t,f) point is needed.

To achieve a high-resolution (t,f) distribution with significantly reduced cross-terms, one can adapt the direction of the DGF along the major axes of ridges in the (t,f) domain to exploit the fact that such smoothing removes cross-terms without degrading the resolution of auto-terms. Experimental results

have confirmed that the adaptive TFD based on the DGF outperforms other adaptive and fixed-kernel TFDs in terms of its ability to resolve closely spaced signal components while removing cross-terms for both synthetic signals and real-life EEG seizure signals.

5.11 THE MODIFIED S-TRANSFORM APPLIED TO TIME-FREQUENCY PHASE SYNCHRONY⁰

5.11.1 PHASE SYNCHRONY MEASURES

Phase synchrony quantifies the coherence and phase coupling between signals or components of signals. A suitable phase synchrony measure can separate phase and amplitude effects from each other while taking the nonstationary nature of signals into account [81]. Phase synchrony assesses the interdependence between the instantaneous phases of two signals; such instantaneous phases (IPs) can be strongly synchronized even when the corresponding signal amplitudes of two signals are statistically independent and uncorrelated. The synchrony is usually quantified for a particular frequency of interest. In the case of narrowband signals, the IP of a signal can be estimated using the Hilbert transform; however, when the signal is broadband and time-varying, the estimation of the IP is more difficult. Alternatively, the instantaneous phase could be inferred from the time-frequency (t,f) representation of the signal given that the derivative of the IP is the IF, which is easily estimated by the first moment of a TFD or by its peaks (see Chapter 1 and [82–84]). However, even where phase is accessible and meaningful, extracting phase information from the (t,f) plane is not a straightforward task. This section discusses how to estimate phase synchrony from the (t,f) plane using linear and quadratic distributions. Important properties given by the phase information of linear (t,f) transforms, such as the Wavelet transform (WT) and S-Transform (ST), and by quadratic (t,f) distributions (TFDs) are related. The concept of cross (t,f) analysis is introduced for linear and quadratic transforms to provide time- and frequency-localized phase difference, and is linked to the spectral delay or IF; hence phase information can be extracted using cross-TFDs, for example, to exploit the synchrony. The methodology is illustrated using simulated signals and borehole acoustic array waveforms to extract phase velocity of different modes for inferring acoustical properties of the medium.

5.11.1.1 Phase synchrony for narrowband signals

Phase *synchrony* is intended to measure the relation between the temporal structures of signals regardless of amplitude. Two signals $x_1(t)$ and $x_2(t)$ are said to be synchronous if their rhythms coincide. The amount of synchrony between them is usually quantified by estimating their instantaneous phases ϕ_{x_1} and ϕ_{x_2} around the frequency of interest. The instantaneous phases may be strongly synchronized even when the amplitudes of $x_1(t)$ and $x_2(t)$ are statistically independent. Synchrony should not be confused with *coherence*. The coherence function quantifies linear correlations in the frequency domain, where the magnitude-squared coherence function C_{ff} and the phase coherence function ϕ_f are distinguishable [81].

⁰Authors: **S. Assous** Geoscience Development, Weatherford Laboratories Ltd, Gotham Road, East Leake LE12 6JX, UK (said.assous@eu.weatherford.com), **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa), and **S. Ventosa**, Institut de Physique du Globe de Paris, Paris, France (ventosa@ipgp.fr). Reviewers: A. Gholami and S. Touati.

Consider a signal $x_1(t) = a_1(t) \cos(2\pi f_1 t + \psi_1)$, with instantaneous amplitude $a_1(t)$, constant frequency f_1 , and initial phase ψ_1 (cf. [Section 1.4](#)). If the argument of the cos function is called $\phi_1(t)$ (the *instantaneous phase*), then $f_1 = \frac{\phi'_1(t)}{2\pi}$. If $\phi_1(t)$ is allowed to be nonlinear, f_1 becomes a function of time, known as the IF: $f_1(t) = \frac{\phi'_1(t)}{2\pi}$. The analytical signal associated with $x_1(t)$, according to [Eq. \(1.4.4\)](#), is $z_1(t) = a_1(t) e^{j\phi_1(t)}$, where the imaginary part $y_1(t) = \text{Im}\{z_1(t)\}$ is the Hilbert transform of the signal $x_1(t)$; however, in the case of *narrowband* amplitude and frequency modulation, the form $x_1(t) = a_1(t) \cos(2\pi f_0 t + \psi_1(t))$ ([Eq. 1.4.5](#)) is used instead of $x_1(t) = a_1(t) \cos(\phi_1(t))$ ([Eq. 1.4.1](#)), in order to separate the contribution of the carrier frequency.

One approach for isolating the instantaneous phase of the signal $x_1(t)$ is by using the Hilbert transform and forming the analytic signal; the signal is first band-pass filtered around the frequency of interest, and then the instantaneous phase is estimated from the analytic signal as follows (see [[82,83](#)] and [Chapter 1](#)):

$$\phi_{x_1}(t) \equiv \arg \{x_1(t) + jy_1(t)\}. \quad (5.11.1)$$

Signals $x_1(t)$ and $x_2(t)$ are said to be *phase-locked* of order $m:n$ if [[81](#)]

$$\Delta\phi_{x_1,x_2}(t) = m\phi_{x_1}(t) - n\phi_{x_2}(t) = \text{Const.} \quad (5.11.2)$$

Usually, the following relaxed relation is considered for real-life signals where the phase-locking ratio between signals is not necessarily always rational [[81](#)]:

$$|m\phi_{x_1}(t) - n\phi_{x_2}(t)| < \text{Const.} \quad (5.11.3)$$

In the case of discrete signals with phase-locking of order 1:1 ($m = n = 1$), the phase synchrony measure is given using [[81](#)]

$$R = \left| \frac{1}{N} \sum_{k=1}^{N-1} e^{j(\phi_{x_1}(k\Delta t) - \phi_{x_2}(k\Delta t))} \right|, \quad (5.11.4)$$

where N is the length of the two signals $x_1(t)$ and $x_2(t)$, and Δt is the sampling period. R refers usually to the mean phase coherence (MPC) or phase-locking value (PLV); see [[81,82](#)] and [Eq. \(16.4.4\)](#).

The convenience of using the analytical signal for defining the instantaneous phase and quantifying the synchrony between narrowband signals does not extend to broadband frequency-modulated signals ([Section 1.4](#)). Moreover, bivariate measures are only able to detect pair-wise phase synchrony and are not adapted to the practical situation of multivariate signals measured by an array of sensors. Hence, a general form of phase synchrony based on the extraction of the instantaneous phase in the (t,f) plane is desirable.

5.11.2 PHASE SYNCHRONY IN TIME-FREQUENCY REPRESENTATIONS

When signals are said to be broadband (time-varying) and multivariate, the amount of synchrony between them can be quantified by estimating the instantaneous phases of individual signals in the (t,f) plane.

Consider two periodic signals $x_1(t)$ and $x_2(t)$, phase locked of order $m:n$ where m and n are integers. If the starting point of both signals is the same and the phase-locking ratio is rational, the

two signals will cross each other periodically at the same initial common value, where the period is related to the least common multiple between m and n . The rational phase-locking order is associated with an intuitive physical meaning for periodic signals. In contrast, the sum of two periodic signals is not periodic in the case of an irrational $m:n$ ratio. Moreover, phase synchrony for time-varying nonstationary signals is not a straightforward process.

For a signal $x(t)$ with a high-resolution complex TFD $\rho_x(t,f)$, the estimated time-varying phase in the (t,f) plane is defined as [82]

$$\Phi_x(t,f) = \arg \{ \rho_x(t,f) \}, \quad (5.11.5)$$

where $\arg\{z\}$ is taken as zero if $z = 0$. Once $\Phi_x(t,f)$ is defined for signal $x(t)$, the (t,f) phase difference (TFPD) between $x_1(t)$ and $x_2(t)$ can be expressed as [82]

$$\Phi_{xy}(t,f) = \arg \{ \rho_{x_1}(t,f) \rho_{x_2}^*(t,f) \}, \quad (5.11.6)$$

where $*$ denotes the complex conjugate.

5.11.2.1 Quadratic TFDs

QTFDs are defined as smoothed versions of the WVD and expressed as

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t-u, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau, \quad (5.11.7)$$

where $z(t)$ is the analytic associate of the real signal $x(t)$, and $G(t, \tau)$ is a time-lag kernel specifying the QTFD (see Section 3.2.2).

For polynomial IFs with degree 2 or higher (i.e., IPs of degree 3 or higher), the WVD introduces pseudo-information in the (t,f) domain in the form of inner cross-terms [85]. These become problematic when the IP is highly varying over time. To suppress cross-terms while keeping a high (t,f) resolution, different QTFDs have been proposed such as the smoothed pseudo WVD (SPWVD) (Section 7.2), the cone-shaped-kernel distribution (ZAMD) and the exponential (Choi-Williams) distribution (Section 2.7.5), and the MBD (Section 5.7). See Chapter 3 for a full discussion.

Since for most QTFDs, it is the signal's real-valued *energy*, not amplitude, that is distributed over time and frequency, there is no direct access to the phase information. Unless complex-valued QTFDs are used, QTFDs cannot be directly used for phase estimation and phase synchrony [82].

But the phase information can be extracted by using the cross-WVD [85, 86] to exploit the phase synchrony between signals [84]. If $z(t) = z_1(t) + z_2(t)$, the WVD of $z(t)$ is $W_z(t,f) = W_{z_1}(t,f) + W_{z_2}(t,f) + 2 \operatorname{Re}\{W_{z_1 z_2}(t,f)\}$, where $W_{z_1 z_2}(t,f)$ is the XWVD of $z_1(t)$ and $z_2(t)$, defined (in Eq. (3.1.21), p. 108) as

$$W_{z_1 z_2}(t,f) = \int_{-\infty}^{\infty} z_1\left(t + \frac{\tau}{2}\right) z_2^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau.$$

5.11.2.2 Continuous wavelet transform

The continuous wavelet transform (CWT) is defined as a set of inner products between the time signal x and a family of functions $\psi_{\tau,\lambda}$:

$$\text{WT}_x(\tau, \lambda) = \frac{1}{\sqrt{|\lambda|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-\tau}{\lambda} \right) dt, \quad (5.11.8)$$

where the mother wavelet ψ is a zero-mean function normalized to $\|\psi\| = 1$.

Strictly, the CWT provides a representation in the space of *scale* (dilation) λ and translation τ , but with the appropriate choice of the mother wavelet $\psi(t)$, it can be used to measure the power spectrum locally.

To explore the instantaneous phase information in the time-scale plane, an approximately analytic complex mother wavelet is desirable. A common choice is the complex Morlet wavelet,

$$\psi(t) = \pi^{-1/4} e^{-j2\pi f_0 t} e^{-t^2/2}, \quad (5.11.9)$$

where f_0 is the central frequency. Additional parameters may be introduced to allow further control of bandwidth.

Translation can be interpreted as time; however, scale cannot be interpreted in general as the inverse of frequency; this intuitive relationship is valid only when the spectrum of the mother wavelet is well localized and has a clear main maximum. In the particular case of the Morlet wavelet, the spectrum of an individual wavelet is a Gaussian function with a center frequency of $f = f_0/\lambda$.

The role of phase in complex wavelet transforms is less clear than in the FT. Being relative to each individual wavelet within the family of functions, instead of having an absolute reference at the origin, hinders its interpretation. But the information is not lost. The actual phase is relative to the center of each wavelet. In consequence, the reference point of the phase translates with τ . At least when using the Morlet wavelet, we can obtain absolute phase measurements using a simple correcting factor. The actual phase is relative because we apply the translation operator on the full mother wavelet function, in particular on the first exponential term in Eq. (5.11.9). So, we can obtain absolute phase observations by (1) applying the translation operation only on the Gaussian term (see under “S-transform” below) or (2) multiplying $\text{WT}_x(\tau, \lambda)$ by $e^{-j2\pi f_0 \tau / \lambda}$ [87]. The first approach has physical intuition while the second has computational efficiency.

Finally, the (t,f) phase difference (TFPD) between two signals $x_1(t)$ and $x_2(t)$ can be performed as defined in Eq. (5.11.6) after extracting the instantaneous phases of the two signals from the coefficients of their wavelet transforms at a given f .

5.11.2.3 S-transform

The ST of a signal $x(t)$ is defined [88] as a CWT wavelet multiplied by a phase factor of frequency f :

$$S(t, f) = e^{-j2\pi f t} \text{WT}_x(t, f), \quad (5.11.10)$$

where the frequency f is inversely related to scale in order to make the width of the used Gaussian window controlled by the scale factor, with better self-adaptability to different frequency components, and where

$$\text{WT}_x(t, f) = \int_{-\infty}^{\infty} x(\tau) g(t - \tau, f) d\tau \quad (5.11.11)$$

and

$$g(t - \tau, f) = \frac{|f|}{\sqrt{2\pi}} e^{-f^2(t-\tau)^2/2} e^{-j2\pi f(t-\tau)}. \quad (5.11.12)$$

Substituting Eqs. (5.11.12) and (5.11.11) into Eq. (5.11.10), we find:

$$S(t, f) = \int_{-\infty}^{\infty} x(\tau) \frac{|f|}{\sqrt{2\pi}} e^{-(t-\tau)^2 f^2/2} e^{-j2\pi f\tau} d\tau. \quad (5.11.13)$$

Note that the ST, unlike the CWT, is normalized by the integral of $g(t)$.

Two important terms in the ST formulation are (1) the phase function $e^{-j2\pi f\tau}$ and (2) the normalization factor $|f|/\sqrt{2\pi}$. The factor in Eq. (5.11.10) is actually a phase correction in the definition of the CWT. It is different from the wavelet analysis concept as it separates the mother wavelet into two parts: the slowly varying frequency-dependent Gaussian envelope, which localizes in time, and the oscillatory exponential term $e^{-j2\pi f\tau}$, which selects the frequency being localized. In other words, it is the time localization Gaussian window that translates, while the oscillatory term remains stationary. The fact that ST retains the absolute phase information (hidden in the CWT) makes ST more suitable for the study of phase synchrony. In the ST, the sinusoidal component of the basis function remains stationary, while the Gaussian envelope translates in time. The reference point for the phase thus remains stationary and the phase has the same meaning as in the FT. This feature makes the ST an intuitive tool for phase analysis in multiple nonstationary signals.

5.11.2.4 Modified S-transform

The original ST places two unnecessary restrictions on the window function. First, only a Gaussian window is considered. Second, that window has no parameters to allow its width in time or frequency to be adjusted. This is remedied by introducing a generalized ST parameterization by a width control coefficient γ [89], as follows:

$$S(t, f, \gamma) = \int_{-\infty}^{\infty} x(\tau) g(t - \tau, f, \gamma) e^{-j2\pi f\tau} d\tau, \quad (5.11.14)$$

where g is the window function of the ST. The parameter τ controls the position of the generalized window g on the time axis. The parameter γ controls the width of the Gaussian window:

$$g(t - \tau, f, \gamma) = \frac{|f|}{\gamma \sqrt{2\pi}} e^{-\frac{-f^2(t-\tau)^2}{2\gamma^2}}. \quad (5.11.15)$$

Based on the generalized form of the ST, a modified ST (MST) is introduced in [90] using extra parameters to adjust the width of the analyzing Gaussian window in time or frequency, rather than taking the frequency as inversely proportional to the standard deviation of the Gaussian window. The scaling function γ , defined above, is made to vary linearly with frequency:

$$\gamma(f) = mf + k, \quad (5.11.16)$$

where m is the slope and k is the intercept. The ST becomes [90]

$$S(t, f, m, k) = \int_{-\infty}^{\infty} x(\tau) \frac{|f|}{(mf+k)\sqrt{2\pi}} e^{-\frac{-f^2(t-\tau)^2}{2(mf+k)^2}} e^{-j2\pi f\tau} d\tau. \quad (5.11.17)$$

The parameter f/γ represents the number of cycles (periods) of a frequency that can be contained within one standard deviation $\sigma = 1/f$ of the Gaussian window given in Eq. (5.11.12). Hence, a progressive resolution improvement is achieved in this case but still governed by the Heisenberg principle and controlled by the factor γ . The values of m and k are tuned to get the desired resolution.

5.11.2.5 Resolution comparison between ST and MST

Let us consider a three-component signal $x_1(t)$ with a low-frequency (100 Hz), a medium-frequency (200 Hz), and a high-frequency burst (400 Hz). Figure 5.11.1(a) and (b) represents the ST and the MST (for $m = 0.25$ and $k = 2$) of the signal, showing that the (t,f) resolution is improved almost twofold by using the MST rather than the ST.

5.11.3 CROSS-MST APPLIED TO PHASE SYNCHRONY

Consider two signals measured by two sensors separated by a known distance. These two sensors see the same signal, but with a time shift in the arrivals and with the addition of noise. Since the MST localizes the spectral components in time, the cross-correlation of specific events on two spatially separated MSTs gives the phase difference, and hence the coherence can be analyzed. The amplitude of the cross-MST (XMST) indicates coincident or partly overlapped signals. The phase of the XMST at local maxima indicates the phase difference between signals. The XMST of $x_1(t)$ and $x_2(t)$ is defined as [91]:

$$\text{XMST}(t,f) = \text{MST}_{x_1}(t,f) \text{MST}_{x_2}^*(t,f). \quad (5.11.18)$$

The phase of the XMST is given by

$$\arg\{\text{XMST}\} = \phi_{x_1}(t,f) - \phi_{x_2}(t,f). \quad (5.11.19)$$

Three key properties of the MST that make the XMST useful are detailed below.

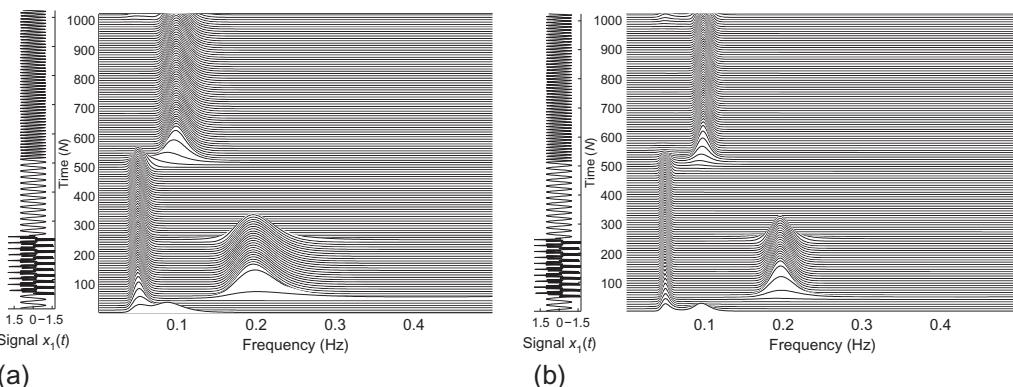


FIGURE 5.11.1

(a) The ST and (b) the MST of the three-component signal $x_1(t)$.

1. *MST time and frequency translation:* if $x(t) \iff \text{MST}(t, f)$, then, as the MST (or ST) preserves the same translation property as the FT [91],

$$x(t - r) \iff \text{MST}(t - r, f) e^{-j2\pi fr}. \quad (5.11.20)$$

2. *MST phase-shift:* for $x(t) = e^{j2\pi f_0 t}$, with spectrum $X(\alpha) = \delta(\alpha - f_0)$, the MST of $x(t)$ becomes $\text{MST}_{x(t)}(t, f) = \int \delta(\alpha + f - f_0) e^{-2\pi^2 \alpha^2 / f^2} e^{j2\pi \alpha t} d\alpha$ (see [91], or use Eq. (13) in [90]). This results in:

$$\text{MST}_{x(t)}(t, f) = e^{-(f-f_0)^2 / (\frac{f}{2\pi})^2} e^{j2\pi(f-f_0)t}. \quad (5.11.21)$$

3. *Phase in the XMST:* a constant phase shift $y(t) = x(t) e^{j\phi} = e^{j2\pi f_0 t + j\phi}$ results in a spectrum $Y(\alpha) = \delta(\alpha - f_0) e^{j\phi}$, which also introduces a phase shift in the modified ST:

$$\text{MST}_y(t, f) = e^{-(mf+k-f_0)^2 / (\frac{mf+k}{2\pi})^2} e^{j2\pi(mf+k-f_0)t} e^{j\phi}. \quad (5.11.22)$$

The above three properties indicate that we can perform a cross spectral analysis between $x(t)$ and $y(t)$ by multiplying the MST of $x(t)$ with the complex conjugate of the modified ST of $y(t)$, obtaining

$$\text{XMST}(t, f) = e^{-(mf+k-f_0)^2 / (\frac{mf+k}{2\pi})^2} e^{j2\pi(mf+k-f_0)t} e^{(mf+k-f_0)^2 / (\frac{mf+k}{2\pi})^2} e^{-j2\pi(mf+k-f_0)t} e^{-j\phi},$$

which simplifies to $\text{XMST}(t, f) = e^{-j\phi}$. Hence, the phase of XMST is

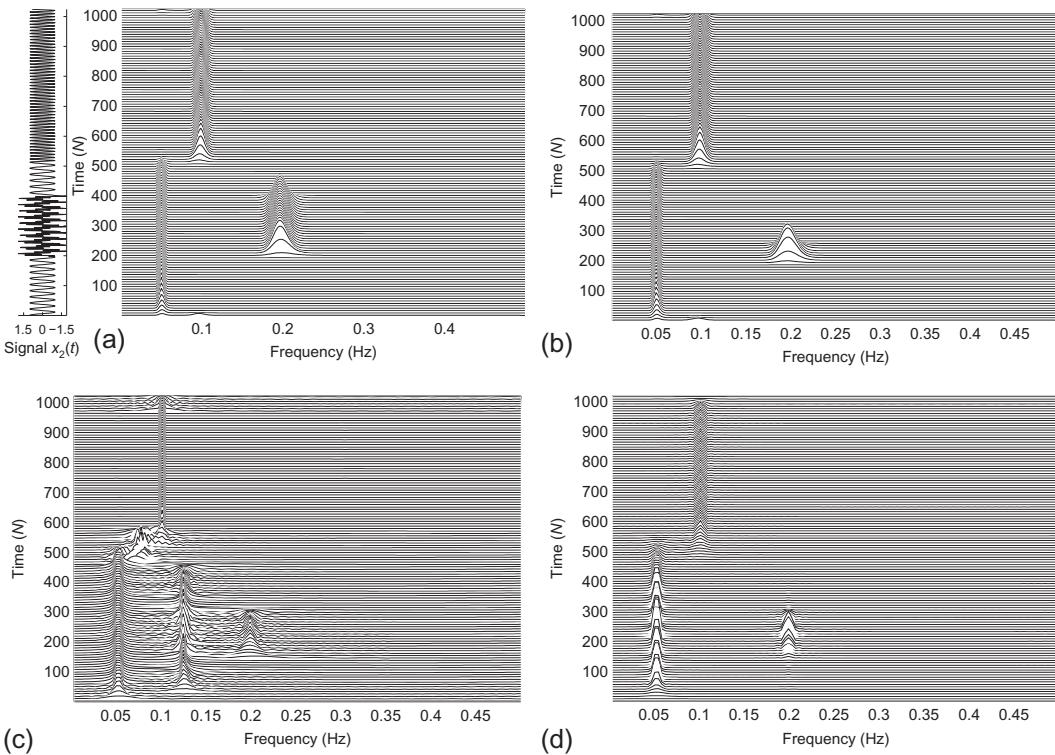
$$\text{Phase} \left\{ \text{MST}_x \text{MST}_y^* \right\} = -\phi. \quad (5.11.23)$$

Thus, the XMST can detect a time lag between signals as a function of t and f .

5.11.3.1 Illustration with simulated examples

Consider the same signal $x_1(t)$ in Fig. 5.11.1, with low-, medium-, and high-frequency components, and another signal $x_2(t)$ with similar components except that the higher component is translated and partly overlapped in time with the higher component of $x_1(t)$. The higher components are out of phase (a cosine in $x_1(t)$ and a sine in $x_2(t)$). The low and medium components of the two signals are in phase. Figure 5.11.2(a) shows the MST of the signal $x_2(t)$; Fig. 5.11.2(b), (c), and (d) represents, respectively, the XMST, the XWVD, and the cross complex-valued Rihaczek-Margenau-Hill distribution (XRMH) (in the *TFSAP* package) between signals $x_1(t)$ and $x_2(t)$. It is clear from both (b) and (d) that for the higher component, only the overlapped part in time is detected, as expected. This overlap is also detected in the XWVD but with an extra added cross-term artifact.

In a similar way to co-spectrum analysis, the real part of the XMST gives the in-phase components of the local spectra, and the imaginary part of the XMST gives the in-quadrature components. Figure 5.11.3(a) and (b) shows the co-XMST and the quadrature XMST amplitudes, respectively. In the co-XMST, only the features that are in phase are present—those being the low- and medium-frequency components. The higher frequency (which is out of phase) does not appear. In the quadrature XMST, only the out-of-phase component appears, which represents the higher-frequency component. Similarly Fig. 5.11.3(c) and (d) shows the co-XWVD and quadrature XWVD, respectively, which have less clarity due to the cross-term artifacts. Figure 5.11.3(e) and (f) represents the co-XRMH and quadrature XRMH, respectively, but with less accuracy than the co-XMST and quadrature XMST.

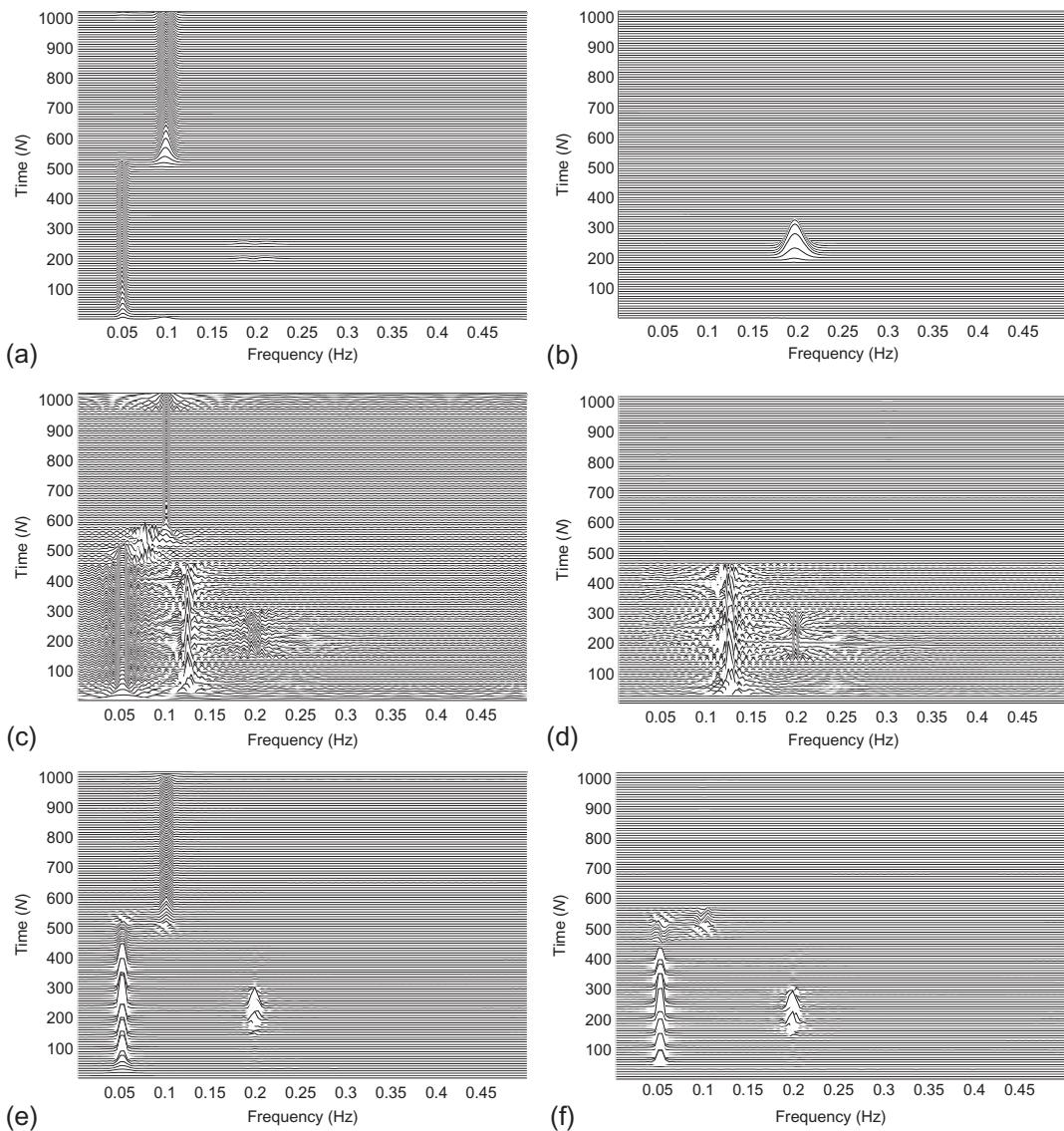
**FIGURE 5.11.2**

(a) MST of the three-component signal $x_2(t)$; (b) XMST of $x_1(t)$ and $x_2(t)$; (c) XWVD of $x_1(t)$ and $x_2(t)$; and (d) XRMH of $x_1(t)$ and $x_2(t)$.

These results show that we can study the coherence using the XMST, the XWVD, or cross-complex-valued QTFDs, whether for amplitude or phase coupling among two or more different spatially recorded components (signals). Careful consideration needs to be made for the XWVD and cross-complex-valued QTFDs to avoid misinterpretation due to the cross-term artifacts.

5.11.3.2 Phase synchrony estimation using cross-MST (borehole array)

Borehole acoustic logging for reservoir characterization consists of transmitting acoustic waveforms in rocks surrounding a borehole to study the effects of variations in rock properties on acoustic waves. Depending on the source of excitation, different monopole or dipole modes can be observed, such as compressional, shear, Stoneley, and flexural. The motion of a flexural wave along the borehole behaves in a similar way to the disturbance that travels up a tree-trunk of uniform diameter when it is shaken while fixed at the top. In this case, the wave shape stretches out in time from the near receiver (bottom) to the far receiver (top). The change in wave shape is caused by structural dispersion in the borehole wall. Flexural waves are dispersive, meaning that their slowness (inverse of velocity) varies

**FIGURE 5.11.3**

(a) Co-XMST amplitude, and (b) quadrature XMST amplitude, of XMST in Fig. 5.11.2(b); (c) co-XWVD amplitude, and (d) quadrature XWVD amplitude, of the XWVD in Fig. 5.11.2(c); (e) co-XRMH amplitude, and (f) quadrature XRMH amplitude, of the XRMH in Fig. 5.11.2(d).

with frequency. In many sets of flexural waveforms, it is possible to see the wave shape change across the receiver array as different frequency components propagate at different speeds. The velocities of the compressional and shear modes can be related to formation porosity and lithology. In the reflection mode, for example, acoustic measurements can yield images of the borehole shape and formation impedance; analysis of the casing flexural wave can be used to measure the integrity of the casing and cement. Measuring the flexural mode with a dipole source, where its low-frequency velocity is an asymptote to the shear velocity, is an indirect measurement of the shear mode. Techniques of extracting shear velocity from the dispersive flexural mode provide important geomechanical inputs regarding the near-borehole stress field. These are used in drilling programs to avoid borehole break-outs or drilling-induced fractures.

The signal array in Fig. 5.11.4(a) was recorded by an acoustic dipole tool via eight evenly spaced receivers (spacing = 0.2 m) when an impulse was fired from a source located 2.6 m away from the first receiver. A (t, f) analysis using the MST was performed for each signal, followed by an XMST between each successive MST pair, to measure the time-lag between different arrivals. Finally, a (t, f) representation was built, showing the dominant dispersive flexural mode (see Fig. 5.11.4(b)). The lower-frequency part of this flexural mode is used to calculate the shear velocity in the rock and assess the lithology. After converting the (t, f) plot to slowness versus frequency (inverse of velocity vs. frequency), knowing the receiver's spacing, a peak-picking algorithm is applied to the flexural curve to get a shear velocity estimate in the low-frequency limit. In theory, only at zero frequency is the flexural velocity equal to the shear velocity. This estimated shear can be used later for different applications such as: an input in the geomechanics model to calculate the wellbore stability; an input in the seismic model to correct for migration effects; and to calculate the acoustic impedance of the lithology knowing its density. These measurements provide valuable information on rock types, gas zones, porosity, rock elastic properties, stress fields around the borehole, and permeability.

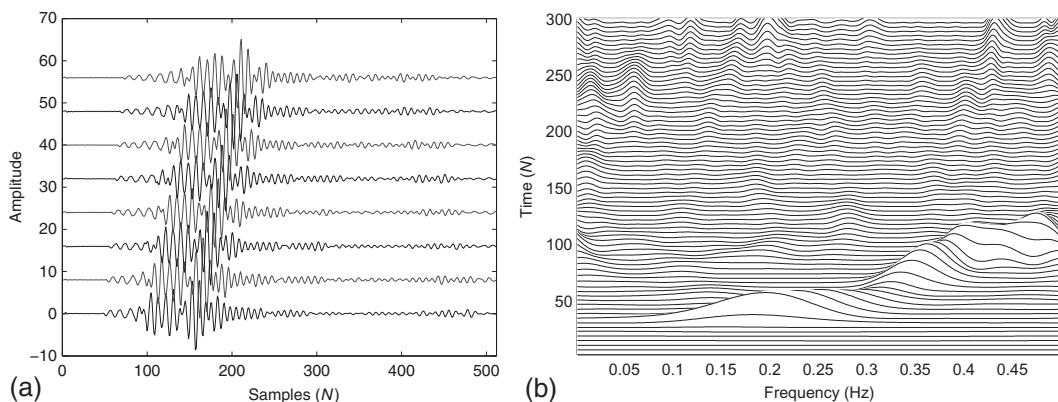


FIGURE 5.11.4

(a) Dipole acoustic receiver array waveforms; (b) XMST of the array, showing the dispersive flexural mode whose low-frequency part is an asymptote to the shear mode.

5.11.4 SUMMARY AND CONCLUSIONS

The relevance of cross (t,f) analysis has been established using the XMST, the XWVD, and cross-complex-valued QTFDs for quantifying the time-lag and phase synchronization within nonstationary signals. Measuring the (t,f) coherence using the XMST is a time-varying cross-spectral analysis obtained by extending the ST and using the absolutely referenced phase property. Defining the coherence using the XWVD and cross-complex-valued QTFDs requires a careful interpretation of the cross-term artifacts. The performance of phase synchrony and time-lag estimation using the XMST is illustrated on synthetic examples and acoustic data. In general, any complex TFD with good (t,f) resolution and accessible meaningful instantaneous phase can be used.

5.12 KERNEL CONDITIONS FOR THE POSITIVITY OF QUADRATIC TFDs⁰

5.12.1 INTRODUCTION

This section shows that a necessary and sufficient condition of a positive-valued time-frequency (t,f) representation is that the weighting function or kernel $g(v,\tau)$ is an ambiguity function (AF), or equivalently, that the (t,f) smoothing kernel $\gamma(t,f)$ is the WVD of some window. Positive TFDs can then be seen as the squared modulus of a cross-AF.

5.12.2 NECESSARY AND SUFFICIENT CONDITION

5.12.2.1 Problem formulation

Section 3.4 established that two positive QTFDs (spectrogram and sonogram) were defined by a kernel $g(v,\tau)$ that is an AF. It is desired to generalize the above result and define a sub-class of kernels $g(v,\tau) \in \{\mathcal{P}_g\}$ for which the resulting QTFD is positive. According to Eq. (3.2.18), the general expression for a QTFD of a finite-energy signal is

$$\left\{ \begin{array}{l} \rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(v,\tau) e^{j2\pi v(t-u)} z\left(u + \frac{\tau}{2}\right) z^* \left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} dv du d\tau; \\ g \in \{\mathcal{P}_g\}, z \in \mathcal{L}^2(\mathbb{C}); \end{array} \right. \quad (5.12.1)$$

where $\mathcal{L}^2(\mathbb{C})$ denotes the set of complex-valued square-integrable signals, and $\{\mathcal{P}_g\} = \{|g(v,\tau)| \leq 1, g = g^\# = g^*(-v,-\tau), |g(0,0)| = 1\}$. The QTFD verifies:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_z(t,f) dt df = E_z = \|z\|^2, \quad \forall z \in \mathcal{L}^2(\mathbb{C}), \quad (5.12.2)$$

as the WVD does (cf. Eq. 3.1.7), so that $\rho_z(t,f)$ can be interpreted as an energy density [92] analogous to a probability density function. The problem may therefore be posed as follows [92]: given Eq. (5.12.1), what is the general condition on $g(v,\tau)$ that will cause $\rho_z(t,f)$ to satisfy Eq. (5.12.2) and be non-negative for all $z \in \mathcal{L}^2(\mathbb{C})$?

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: B.K. Jawad, S. Touati, N.A. Khan, and S. Ouelha.

From the definition of the AF, Eq. (5.12.1) can be expressed as:

$$\rho_z(t, f) = \mathcal{F}_{t \leftarrow v}^{-1} \{ \mathcal{F}_{\tau \rightarrow f} \{ g(v, \tau) A_z(v, \tau) \} \}. \quad (5.12.3)$$

It follows that for a QTFD to be positive, $\rho_z(t, f)$ and $[gA_z](v, \tau) = g(v, \tau)A_z(v, \tau)$ need to satisfy the following two related properties:

$$\begin{aligned} \rho_z(t, f) &= |l(t, f)|^2, & \left(\Rightarrow l(t, f) = \mathcal{F}_{t \leftarrow v}^{-1} \{ \mathcal{F}_{\tau \rightarrow f} \{ L(v, \tau) \} \} \right), \\ \rho_z(t, f) &= \mathcal{F}_{t \leftarrow v}^{-1} \{ \mathcal{F}_{\tau \rightarrow f} \{ [gA_z](v, \tau) \} \}, & \left(\Rightarrow \mathcal{F}_{t \rightarrow v} \left\{ \mathcal{F}_{\tau \leftarrow f}^{-1} \left\{ |l(t, f)|^2 \right\} \right\} = \Gamma_L(v, \tau) \right), \end{aligned} \quad (5.12.4)$$

where $l(t, f)$ defines an amplitude joint (t, f) representation generalizing those previously proposed, for example, in [93,94]; an example is the STFT defined in Eq. (2.3.2), for which $\rho(t, f)$ becomes the spectrogram. The function $\Gamma_L(v, \tau)$ is a 2D autocorrelation function. This may be considered as a 2D extension of the well-known result that the IFT of the power spectral density (of a wide-sense stationary random process) is the autocorrelation function of the signal $z(t)$.

The problem posed is therefore to define the characteristics of $l(t, f)$ for Eq. (5.12.4) to hold; or in other words, let us seek the necessary and sufficient condition on $g \in \{\mathcal{P}_g\}$ so that the above is valid and $\rho(t, f)$ results in a “positive” (actually non-negative) TFD.

To simplify derivations, and following the original approach presented in [92], let us use operator theory in this section [95,96]. Recall that a linear operator, or simply an operator \mathbf{O} , is a linear function that maps $\mathcal{L}^2(\mathbb{C})$ into itself.¹ That is, to each signal $|z\rangle$ in $\mathcal{L}^2(\mathbb{C})$, \mathbf{O} assigns another element $\mathbf{O}|z\rangle$ in $\mathcal{L}^2(\mathbb{C})$ such that

$$\mathbf{O}(\alpha|z_1\rangle + \beta|z_2\rangle) = \alpha\mathbf{O}(|z_1\rangle) + \beta\mathbf{O}(|z_2\rangle) \quad (5.12.5)$$

whenever $|z_1\rangle$ and $|z_2\rangle$ are any two signal elements of $\mathcal{L}^2(\mathbb{C})$ and α and β are complex numbers [95,96]. The notation $|z\rangle$ follows the Dirac notation and is known as a ket, which is a vector decomposed in a functional basis $(|u\rangle)$ as

$$|z\rangle = \int du |u\rangle \langle u|z\rangle = \int du |u\rangle z(u), \quad (5.12.6)$$

where $\langle u|z\rangle$ is the inner/scalar product of ket $|u\rangle$ and ket $|z\rangle$.

As defined in Eqs. (5.12.1) or (5.12.3), $\rho(t, f)$ naturally satisfies the time-frequency shift covariance property; that is, if the signal $s(t)$ or $z(t)$ is translated in t or f or both, then $\rho(t, f)$ will be translated by the same quantities [92] (see proof in Section 5.12.5).

In addition, the IAF representation $R_z(t, \tau)$ can be considered as the mean value of the operator $\mathbf{T}_{-\tau/2}\mathbf{G}_t\mathbf{T}_{-\tau/2}$ (see proof Section 5.12.5 and [92]).

5.12.2.2 Necessary condition on $g(v, \tau)$ for a TFD to be positive

As per Eq. (3.2.6), let us define $G(t, \tau) = \mathcal{F}_{t \leftarrow v}^{-1} \{ g(v, \tau) \}$; then Eq. (5.12.1) leads to:

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-2j\pi\tau f} z^*(u - \tau/2) G(t - u, \tau) z(u + \tau/2) du d\tau. \quad (5.12.7)$$

¹Another notation for operators is $(\mathbf{O}_{ax})(t) = \dots$, which is used, for instance, in Section 4.3.

Let $e^{-2j\pi\tau f} = e^{-2j\pi(u+\tau/2-(u-\tau/2))f}$, then we can write:

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2j\pi(u-\tau/2)f} z^*(u - \tau/2) G(t - u, \tau) e^{-2j\pi(u+\tau/2)f} z(u + \tau/2) du d\tau. \quad (5.12.8)$$

By performing the change of variables $v = u - \tau/2$ and $w = u + \tau/2$, we get:

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2j\pi vf} z^*(v) G\left(t - \frac{(v+w)}{2}, w-v\right) e^{-2j\pi wf} z(w) dv dw. \quad (5.12.9)$$

Multiplying the first complex exponential by $e^{-2j\pi tf}$ and the second by $e^{2j\pi tf}$, this latter expression becomes:

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2j\pi(v-t)f} z^*(v) G\left(t - \left(\frac{v+w}{2}\right), w-v\right) e^{2j\pi(t-w)f} z(w) dv dw. \quad (5.12.10)$$

Now, assuming that in Eq. (5.12.10), $\rho_z(t,f)$ is positive; we can then write:

$$\rho_z(t,f) = |l_z(t,f)|^2 = l_z(t,f) l_z^*(t,f). \quad (5.12.11)$$

Given Eq. (5.12.10), a function \mathbf{O} exists, such that $l_z(t,f)$ can be expressed as:

$$l_z(t,f) = \langle \mathbf{O}|z\rangle_{(t,f)} = \int_{-\infty}^{\infty} \mathbf{O}_{(t,f)}^*(u) z(u) du, \quad (5.12.12)$$

$$\mathbf{O}_{(t,f)}^* = (\mathbf{T}_u \mathbf{O}'_{(t,f)})(t) = h^*(t-u) e^{2j\pi(t-u)f}. \quad (5.12.13)$$

From this latter expression and Eqs. (5.12.10)–(5.12.12), we identify that:

$$G(t - (v+w)/2, w-v) = h(t-v)h^*(t-w). \quad (5.12.14)$$

For $\frac{(v+w)}{2} = 0, w-v = \tau$, Eq. (5.12.14) becomes $G(t, \tau) = h(t + \frac{\tau}{2})h^*(t - \frac{\tau}{2})$. Thus the Doppler-lag kernel of a positive TFD needs to be an AF.

$$g(v, \tau) = \int h\left(t + \frac{\tau}{2}\right) h^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi v t} dt. \quad (5.12.15)$$

If $\{A\}$ denotes the class of AFs, one deduces [92]:

Theorem 5.12.1 (Necessary condition). $\rho_z(t,f) \geq 0$ implies that g is an AF of some window $h(t)$.

$$g \in \{A\} \subset \{\mathcal{P}_g\}. \quad (5.12.16)$$

5.12.2.3 Proof for Eq. (5.12.13)

The change of variable $w = t-u$ in Eq. (5.12.10) allows us to write: $l_z(t,f) = \int_{-\infty}^{\infty} \mathbf{O}'_{(t,f)}^*(u) (\mathbf{T}_u z)(t) du$; and from Section 5.12.5.1, we deduce that:

$$l_z(t,f) = \int_{-\infty}^{\infty} (\mathbf{T}_u \mathbf{O}'_{(t,f)})(t) z(u) du. \quad (5.12.17)$$

Taking for $\mathbf{O}'_{(t,f)}$ the general expression: $\mathbf{O}'_{(t,f)} = h(t) e^{-j2\pi tf}$, we deduce from Eqs. (5.12.12) and (5.12.17) that Eq. (5.12.13) holds.

5.12.2.4 Sufficient condition

To write $g(v, \tau) = A_h(v, \tau)$, $\mathcal{F}_{t \rightarrow u} \{h(t)\} = H(u) \in \mathcal{L}^2(\mathbb{C})$ is the starting assumption for searching the reciprocal condition. Let us use the relation (5.12.3):

$$\rho_z(t,f) = \mathcal{F}_{t \leftarrow v}^{-1} \{ \mathcal{F}_{\tau \rightarrow f} \{ A_h(v, \tau) A_z(v, \tau) \} \}. \quad (5.12.18)$$

Replacing the AFs by their explicit expressions gives:

$$\rho_z(t,f) = \iiint e^{-j2\pi vv} h(v + \frac{\tau}{2}) h^*(v - \frac{\tau}{2}) z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) e^{j2\pi v(t-u)} e^{-j2\pi f\tau} dv du dv d\tau.$$

Rearranging the order of integration as

$$\rho_z(t,f) = \iiint h(v + \frac{\tau}{2}) h^*(v - \frac{\tau}{2}) z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) e^{-j2\pi f\tau} \underbrace{\left(\int e^{-j2\pi v(v+u-t)} dv \right)}_{\delta(v+u-t)} du dv d\tau \quad (5.12.19)$$

yields

$$\rho_z(t,f) = \iint h(v + \frac{\tau}{2}) h^*(v - \frac{\tau}{2}) z(t - v + \frac{\tau}{2}) z^*(t - v - \frac{\tau}{2}) e^{-j2\pi f\tau} dv d\tau. \quad (5.12.20)$$

Proceeding to the change of variables $\theta_1 = v + \tau/2$, $\theta_2 = v - \tau/2$ gives a Jacobian matrix with determinant 1 and separable conjugate integrals (with $\tau = \theta_1 - \theta_2$):

$$\rho_z(t,f) = \int h(\theta_1) z^*(t - \theta_1) e^{-j2\pi f\theta_1} d\theta_1 \int h^*(\theta_2) z(t - \theta_2) e^{j2\pi f\theta_2} d\theta_2. \quad (5.12.21)$$

So ρ_z is positive, as with a simple change of variable, it can be written as

$$\rho_z(t,f) = \left| \int z(u) h^*(t-u) e^{j2\pi f(t-u)} du \right|^2. \quad (5.12.22)$$

By applying Parseval's theorem, Eq. (5.12.22) can also be rewritten as:

$$\rho_z(t,f) = \left| \int Z(v) H^\#(v-f) e^{j2\pi vt} dv \right|^2 = |A_{ZH^-}(-t,f)|^2 = |l(t,f)|^2, \quad (5.12.23)$$

where $Z(v)$ is the FT of $z(t)$ and $H(v)$ is the FT of $h(t)$.

Note that in the above, $H^-(u) = H(-u)$, and $A_{xy}(v, \tau)$ is the nonsymmetric cross-AF of the signals x and $y \in \mathcal{L}^2(\mathbb{C})$ defined as:

$$A_{xy}(v, \tau) = \int_{-\infty}^{\infty} x(t) y^*(t - \tau) e^{-j2\pi vt} dt. \quad (5.12.24)$$

Theorem 5.12.2 (Sufficient condition).

$$g \in \{A\} \rightarrow \rho_z(t,f) = |l(t,f)|^2 \geq 0. \quad (5.12.25)$$

It follows [92]:

Theorem 5.12.3. A necessary and sufficient condition for $\rho_z(t,f)$ to be positive is that g , the weighting function, that is, the Doppler-lag kernel be an AF:

$$g \in \{A\} \Leftrightarrow \rho_z(t,f) \geq 0. \quad (5.12.26)$$

Corollary 5.12.4. Given that a sum of positive TFDs will be positive, it logically follows that the kernel $g(v, \tau)$ can also be a sum of AFs [92].

Proof: The result follows directly from Eq. (5.12.18), using the FT linearity property. \square

A consequence of Theorem 5.12.2 (Eq. 5.12.25) is that if $g(v, \tau)$ defined in Eq. (5.12.15) is an AF, then it follows that $\gamma(t,f)$ is a WVD, as stated below.

Corollary 5.12.5. A necessary and sufficient condition for $\rho_z(t,f)$ to be positive is that the (t,f) smoothing function $\gamma(t,f)$ defined in Eq. (3.2.6) is a WVD.

The proof follows directly from Eq. (3.2.5) (see also Eq. (3.2.13)).

5.12.3 IF IMPLICATIONS, POSITIVITY, AND (t, f) RESOLUTION

We notice also that according to Eq. (5.12.26) and the definition of the AF, we cannot have the conditions

$$A_h(v, 0) = \text{constant}, \quad A_h(0, \tau) = \text{constant}, \quad (5.12.27)$$

which are needed for estimation of the IF and group delay parameters, f_i and τ_g , using the moments of ρ_z (see Table 3.3.1 on p. 122, and [97]). This shows, therefore, that the positivity of ρ_z is incompatible with directly estimating f_i and τ_g using the moments of ρ_z .

5.12.4 SUMMARY AND CONCLUSIONS

The most popular and easiest-to-use quadratic TFD is the spectrogram, which has the property of positivity and therefore allows an interpretation of the TFD as an energy density. Unfortunately, the spectrogram suffers from a poor resolution and window sensitivity, as shown in Fig. 2.3.1 (p. 78). It is therefore desirable to study the conditions on the kernel $g(v, \tau)$ of QTFDs, which allow $\rho(t,f)$ to be positive and have a high resolution. This section is part of the answer to this problem as it finds that the kernel filter $g(v, \tau)$ must be an AF or equivalently that the (t,f) smoothing kernel $\gamma(t,f)$ is a WVD. The next step is then to find out among all $g(v, \tau)$ that are ambiguity functions, which ones lead to the highest resolution and in which circumstances. Other related contributions to the issue of positivity of TFDs include [98].

5.12.5 PROOFS

5.12.5.1 Proof for (t, f) covariance

Let $(\mathbf{S}_{\tau,v}z)(t)$ be the time-frequency shift operator such that

$$(\mathbf{S}_{\tau,v}z)(t) = z(t - \tau) e^{j2\pi vt} = e^{j2\pi vt} (\mathbf{T}_\tau z)(t) = z_{\tau,v}, \quad (5.12.28)$$

where τ and v are the time delay and the frequency shift, and where \mathbf{T}_τ is the translation operator defined by

$$(\mathbf{T}_\tau z)(t) = z(t - \tau). \quad (5.12.29)$$

Using Eq. (5.12.1), and replacing t by $(t - \tau)$ and f by $(f - v)$, the FT properties lead to

$$\rho_{z_{\tau,v}}(t,f) = \rho_{\mathbf{S}_{\tau,v}z}(t,f) = \rho_z(t - \tau, f - v) \quad (5.12.30)$$

as claimed. \square

5.12.5.2 Proof of $R_z(t, \tau)$ as an operator mean-value

According to the definition of $R_z(t, \tau)$,

$$R_z(t, \tau) = \mathcal{F}_{\tau \leftarrow f}^{-1} \{\rho(t, f)\}. \quad (5.12.31)$$

Let us define the linear operator \mathbf{G}_t such that

$$(\mathbf{G}_t z)(u) = G(t - u, \tau)z(u) \quad (5.12.32)$$

with

$$G(t - u, \tau) = \mathcal{F}_{t-u \leftarrow v}^{-1} \{g(v, \tau)\}. \quad (5.12.33)$$

Using Eq. (5.12.1), the definition of R_z and the last equation, we have

$$R_z(t, \tau) = \int_{-\infty}^{\infty} z^*(u - \frac{\tau}{2}) G(t - u, \tau) z(u + \frac{\tau}{2}) du, \quad (5.12.34)$$

which can be rewritten more concisely using the Dirac notation and Eq. (5.12.29):

$$R_z(t, \tau) = \left\langle \mathbf{T}_{\frac{\tau}{2}} z | \mathbf{G}_t | \mathbf{T}_{-\frac{\tau}{2}} z \right\rangle. \quad (5.12.35)$$

Now using the properties

$$\begin{aligned} \mathbf{T}_\tau^+ &= \mathbf{T}_{-\tau}; \\ \langle z | \mathbf{O} | z \rangle &= \langle z | \mathbf{O} z \rangle = \langle \mathbf{O}^+ z | z \rangle; \\ \langle z | \mathbf{O}_1 \mathbf{O}_2 | z \rangle &= \langle z | \mathbf{O}_1 | \mathbf{O}_2 z \rangle = \langle \mathbf{O}_1^+ z | \mathbf{O}_2 z \rangle, \end{aligned}$$

where \mathbf{O}_1 and \mathbf{O}_2 are two arbitrary linear operators, and where \mathbf{T}^+ is the Hermitian conjugate of the operator \mathbf{T} , we can write

$$\begin{aligned}
R_z(t, \tau) &= \left\langle \mathbf{T}_{-\frac{\tau}{2}}^+ z | \mathbf{G}_t | \mathbf{T}_{-\frac{\tau}{2}}^- z \right\rangle \\
&= \left\langle \mathbf{T}_{-\frac{\tau}{2}}^+ z | \mathbf{G}_t \mathbf{T}_{-\frac{\tau}{2}}^- z \right\rangle \\
&= \left\langle z | \mathbf{T}_{-\frac{\tau}{2}}^- \mathbf{G}_t \mathbf{T}_{-\frac{\tau}{2}}^+ z \right\rangle \\
&= \left\langle z | \mathbf{T}_{-\frac{\tau}{2}}^- \mathbf{G}_t \mathbf{T}_{-\frac{\tau}{2}}^+ | z \right\rangle,
\end{aligned} \tag{5.12.36}$$

which completes the proof. \square

REFERENCES

- [1] H.L.L. Van Trees, Detection, Estimation, and Modulation Theory, vol. I: Detection, Estimation, and Linear Modulation Theory, Wiley, New York, 1968, reprinted 2001.
- [2] P.M. Woodward, Probability and Information Theory with Applications to Radar, Pergamon, London, 1953.
- [3] S.M. Sussman, “Least-squares synthesis of radar ambiguity functions”, *IRE Trans. Inf. Theory* 8 (1962) 246-254.
- [4] E.J. Kelly, R.P. Wishner, “Matched filter theory for high-velocity, accelerating targets”, *IEEE Trans. Mil. Electron.* 9 (1965) 56-59.
- [5] J. Ville, “Théorie et applications de la notion de signal analytique”, *Cables et Transmissions* 2A (1) (1948) 61-74 (in French. English translation: I. Selin, Theory and Applications of the Notion of Complex Signal, Rand Corporation Report T-92, Santa Monica, CA, August 1958).
- [6] R.A. Altes, “Wideband, proportional-bandwidth Wigner-Ville analysis”, *IEEE Trans. Acoust. Speech Signal Process.* 38 (6) (1990) 1005-1012.
- [7] P. Flandrin, Time-Frequency/Time-Scale Analysis, Academic Press, San Diego, 1999 (translated by J. Stöckler from the French editions, *Temps-fréquence*, Paris: Hermès, 1993, 1998).
- [8] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.
- [9] W.M. Siebert, “Studies of Woodward’s uncertainty function”, *Q. Progr. Rep.* (1958) 90-94.
- [10] D.E. Vakman, Sophisticated Signals and the Uncertainty Principle in Radar, Springer, New York, 1968 (translated by K.N. Trirogoff; edited by E. Jacobs).
- [11] E.H. Lieb, “Integral bounds for radar ambiguity functions and Wigner distributions”, *J. Math. Phys.* 31 (3) (1990) 594-599.
- [12] A.W. Rihaczek, Principles of High-Resolution Radar, McGraw-Hill, New York, 1969 (reprinted Los Altos, CA: Peninsula Publishing, 1985).
- [13] R. Price, E.M. Hofstetter, “Bounds on the volume and height distributions of the ambiguity function”, *IEEE Trans. Inf. Theory* 11 (1965) 207-214.
- [14] J. Bertrand, P. Bertrand, J.-P. Ovarlez, “The Mellin transform”, in: A.D. Poularikas (Ed.), The Transforms and Applications Handbook, first ed., CRC Press, Boca Raton, FL, 1996, pp. 829-885 (Chapter 11).
- [15] R.A. Altes, E.L. Titlebaum, “Bat signals as optimally Doppler tolerant waveforms”, *J. Acoust. Soc. Am.* 48 (4B) (1970) 1014-1020.
- [16] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), Advances in Spectrum Analysis and Array Processing, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [17] H.-I. Choi, W.J. Williams, “Improved time-frequency representation of multicomponent signals using exponential kernels”, *IEEE Trans. Acoust. Speech Signal Process.* 37 (6) (1989) 862-871.

- [18] W.J. Williams, J. Jeong, "Reduced interference time-frequency distributions", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 74-97 (Chapter 3).
- [19] L. Cohen, Time-Frequency Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [20] Y. Zhao, L.E. Atlas, R.J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals", *IEEE Trans. Acoust. Speech Signal Process.* 38 (7) (1990) 1084-1091.
- [21] F. Hlawatsch, R.L. Urbanke, "Bilinear time-frequency representations of signals: the shift-scale invariant class", *IEEE Trans. Signal Process.* 42 (1994) 357-366.
- [22] J. Jeong, W.J. Williams, "Kernel design for reduced interference distributions", *IEEE Trans. Signal Process.* 40 (2) (1992) 402-412.
- [23] W.J. Williams, "Reduced interference distributions: biological applications and interpretations", *Proc. IEEE* 84 (9) (1996) 1264-1280.
- [24] W.J. Williams, T.-H. Sang, "Adaptive RID kernels which minimize time-frequency uncertainty", in: Proc. IEEE-SP Int. Symp. on Time-Frequency & Time-Scale Analysis, 1994, pp. 96-99.
- [25] J. Jeong, W.J. Williams, "Alias-free generalised discrete-time time-frequency distributions", *IEEE Trans. Signal Process.* 40 (11) (1992) 2757-2765.
- [26] W.J. Williams, "Recent advances in time-frequency representations: some theoretical foundation", in: M. Akay (Ed.), Time Frequency and Wavelets in Biomedical Signal Processing, IEEE/Wiley, New York, 1998 (Chapter 1).
- [27] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [28] D.L. Jones, T.W. Parks, "A resolution comparison of several time-frequency representations", *IEEE Trans. Signal Process.* 40 (2) (1992) 413-420.
- [29] D.L. Jones, T.W. Parks, "A high-resolution data-adaptive time-frequency representation", *IEEE Trans. Acoust. Speech Signal Process.* 38 (12) (1990) 2127-2135.
- [30] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, O.J.J. Michel, "Measuring time-frequency information content using the Rényi entropies", *IEEE Trans. Inf. Theory* 47 (4) (2001) 1391-1409.
- [31] D.L. Jones, R.G. Baraniuk, "A simple scheme for adapting time-frequency representations", *IEEE Trans. Signal Process.* 42 (12) (1994) 3530-3535.
- [32] T.K. Hon, A. Georgakis, "Enhancing the resolution of the spectrogram based on a simple adaptation procedure", *IEEE Trans. Signal Process.* 60 (10) (2012) 5566-5571.
- [33] R.G. Baraniuk, D.L. Jones, "A signal-dependent time-frequency representation: optimal kernel design", *IEEE Trans. Signal Process.* 41 (4) (1993) 1589-1602.
- [34] R.G. Baraniuk, D.L. Jones, "A signal-dependent time-frequency representation: fast algorithm for optimal kernel design", *IEEE Trans. Signal Process.* 42 (1) (1994) 134-146.
- [35] R.G. Baraniuk, D.L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel", *Signal Process.* 32 (3) (1993) 263-284.
- [36] D.L. Jones, R.G. Baraniuk, "Adaptive optimal-kernel time-frequency representation", *IEEE Trans. Signal Process.* 43 (10) (1995) 2361-2371.
- [37] R.N. Czerwinski, D.L. Jones, "Adaptive cone-kernel time-frequency analysis", *IEEE Trans. Signal Process.* 43 (7) (1995) 1715-1719.
- [38] S. Krishnamachari, W.J. Williams, "Adaptive kernel design in the generalized marginals domain for time-frequency analysis", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94), vol. 3, 1994, pp. 341-344.

- [39] R.M. Nickel, T.-H. Sang, W.J. Williams, "A new signal adaptive approach to positive time-frequency distributions with suppressed interference terms", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98), vol. 3, 1998, pp. 1777-1780.
- [40] B. Boashash, B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise", *Signal Process.* 67 (1) (1998) 1-23.
- [41] B. Boashash, P.J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.* 42 (1) (1994) 216-220.
- [42] L. Stanković, S. Stanković, "An analysis of instantaneous frequency representation using time-frequency distributions-generalized Wigner distribution", *IEEE Trans. Signal Process.* 43 (2) (1995) 549-552.
- [43] B. Boashash, "Time-frequency signal analysis: past, present and future trends", in: C.T. Leondes (Ed.), Digital Control and Signal Processing Systems and Techniques, in: Control and Dynamic Systems: Advances in Theory and Applications, vol. 78, Academic Press, San Diego, 1996, pp. 1-69 (Chapter 1).
- [44] S. Stanković, L. Stanković, "Introducing time-frequency distribution with a 'complex-time' argument", *Electron. Lett.* 32 (14) (1996) 1265-1267.
- [45] B. Barkat, B. Boashash, "Design of higher order polynomial Wigner-Ville distributions", *IEEE Trans. Signal Process.* 47 (9) (1999) 2608-2611.
- [46] B. Ristic, B. Boashash, "Comments on 'The Cramer-Rao lower bounds for signals with constant amplitude and polynomial phase'", *IEEE Trans. Signal Process.* 46 (6) (1998) 1708-1709, ISSN 1053-587X, <http://dx.doi.org/10.1109/78.678501>.
- [47] B. Ristic, B. Boashash, "Relationship between the polynomial and higher order Wigner-Ville distribution", *IEEE Signal Process. Lett.* 2 (12) (1995) 227-229.
- [48] S. Shamsunder, G.B. Giannakis, B. Friedlander, "Estimating random amplitude polynomial phase signals: a cyclostationary approach", *IEEE Trans. Signal Process.* 43 (2) (1995) 492-505.
- [49] L. Stanković, "A multitime definition of the Wigner higher order distribution: L-Wigner distribution", *IEEE Signal Process. Lett.* 1 (7) (1994) 106-109.
- [50] A.H. Nuttal, "Efficient evaluation of polynomials and exponentials of polynomials for equispaced arguments", *IEEE Trans. Acoust. Speech Signal Process.* 35 (10) (1987) 1486-1487.
- [51] S. Peleg, B. Porat, "Estimation and classification of polynomial-phase signals", *IEEE Trans. Inf. Theory* 37 (2) (1991) 422-430.
- [52] B. Ristic, B. Boashash, "Instantaneous frequency estimation of quadratic and cubic polynomial FM signals using the cross polynomial Wigner-Ville distribution", *IEEE Trans. Signal Process.* 44 (1996) 1549-1553.
- [53] A. Omidvarnia, G. Azemi, J.M.O. Toole, B. Boashash, "Robust estimation of highly-varying nonlinear instantaneous frequency of monocomponent signals using a lower-order complex-time distribution", *Signal Process.* 93 (11) (2013) 3251-3260, ISSN 0165-1684, <http://dx.doi.org/10.1016/j.sigpro.2013.03.041>, URL <http://www.sciencedirect.com/science/article/pii/S0165168413001497>.
- [54] L. Stanković, "Time-frequency distributions with complex argument", *IEEE Trans. Signal Process.* 50 (3) (2002) 475-486.
- [55] F. Hlawatsch, G.F. Boudreux-Bartels, "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.* 9 (2) (1992) 21-67.
- [56] A. Papandreou-Suppappola, R.L. Murray, B.G. Iem, G.F. Boudreux-Bartels, "Group delay shift covariant quadratic time-frequency representations", *IEEE Trans. Signal Process.* 49 (11) (2001) 2549-2564.
- [57] F. Hlawatsch, A. Papandreou-Suppappola, G.F. Boudreux-Bartels, "The power classes—quadratic time-frequency representations with scale covariance and dispersive time-shift covariance", *IEEE Trans. Signal Process.* 47 (11) (1999) 3067-3083.
- [58] A. Papandreou, F. Hlawatsch, G.F. Boudreux-Bartels, "The hyperbolic class of quadratic time-frequency representations—part I: constant- Q warping, the hyperbolic paradigm, properties, and members", *IEEE Trans. Signal Process.* 41 (12) (1993) 3425-3444.

- [59] Y. Jiang, A. Papandreou-Suppappola, "Discrete time-frequency characterizations of dispersive linear time-varying systems", *IEEE Trans. Signal Process.* 55 (5) (2007) 2066-2076.
- [60] N.F. Josso, J.J. Zhang, A. Papandreou-Suppappola, C. Ioana, T.M. Duman, "Nonstationary system analysis methods for underwater acoustic communications", *EURASIP J. Adv. Signal Process.* 2011 (2011) doi:10.1155/2011/807472, article ID 807472, 14 pp.
- [61] E.F. Velez, H. Garudadri, "Speech analysis based on smoothed Wigner-Ville distribution", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 351-374 (Chapter 15).
- [62] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digital Signal Process.* 40 (2015) 1-30, doi:10.1016/j.dsp.2014.12.015.
- [63] Z.M. Hussain, B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions", *IEEE Trans. Signal Process.* 50 (8) (2002) 1866-1876.
- [64] Z.M. Hussain, B. Boashash, "The T-class of time-frequency distributions: time-only kernels with amplitude estimation", *J. Franklin Inst.* 343 (7) (2006) 661-675.
- [65] B. Barkat, B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis", *IEEE Trans. Signal Process.* 49 (10) (2001) 2232-2239.
- [66] A.C. McBride, F.H. Kerr, "On Namias's fractional Fourier transforms", *J. Appl. Math.* 39 (2) (1987) 159-175.
- [67] L.B. Almeida, "The fractional Fourier transform and time-frequency representations", *IEEE Trans. Signal Process.* 42 (11) (1994) 3084-3091.
- [68] A.W. Lohmann, B.H. Soffer, "Relationships between the Radon-Wigner and fractional Fourier transforms", *J. Opt. Soc. Am. A* 11 (6) (1994) 1798-1801.
- [69] X.-G. Xia, Y. Owechko, B.H. Soffer, R.M. Matic, "On generalized-marginal time-frequency distributions", *IEEE Trans. Signal Process.* 44 (11) (1996) 2882-2886.
- [70] H.L.L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: Radar-Sonar Signal Processing and Gaussian Signals in Noise, Wiley, New York, 1971, reprinted Krieger, Malabar, FL, 1992 (reprinted Wiley, New York, 2001).
- [71] E. Sejdić, I. Djurović, L. Stanković, "Fractional Fourier transform as a signal processing tool: an overview of recent developments", *Signal Process.* 91 (6) (2011) 1351-1369.
- [72] H. Hashemi, "Fuzzy clustering of seismic sequences: segmentation of time-frequency representations", *IEEE Signal Process. Mag.* 29 (3) (2012) 82-87.
- [73] L. Wu, S. Xu, G. Li, Z. Chen, "Ballistic missile precession frequency extraction by spectrogram's texture", in: Proc. SPIE, MIPPR 2013: Automatic Target Recognition and Navigation, vol. 8918, Soc. of Photo-Optical Instrumentation Engineers, 2013, article ID 891815, 8 pp.
- [74] V. Lakshmanan, "A separable filter for directional smoothing", *IEEE Geosci. Remote Sens. Lett.* 1 (3) (2004) 192-195.
- [75] B. Ristic, B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform", *IEEE Trans. Signal Process.* 41 (5) (1993) 1996-2008.
- [76] M. Abed, A. Belouchrani, M. Cheriet, B. Boashash, "Time-frequency distributions based on compact support kernels: properties and performance evaluation", *IEEE Trans. Signal Process.* 60 (6) (2012) 2814-2827.
- [77] M.J. Bastiaans, T. Alieva, L. Stanković, "On rotated time-frequency kernels", *IEEE Signal Process. Lett.* 9 (11) (2002) 378-381.
- [78] N.A. Khan, B. Boashash, "Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram", *IEEE Signal Process. Lett.* 20 (2) (2013) 157-160.
- [79] Y. Liu, L. Mejias, Z. Li, "Fast power line detection and localization using steerable filter for active UAV guidance", in: Proc. 22nd Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS2012), 2012, pp. 491-496.

- [80] N.A. Khan, B. Boashash, "Multi-component instantaneous frequency estimation using locally adaptive directional time frequency distributions", *Int. J. Adapt. Control Signal Process.* (2015) doi:10.1002/acs.2583.
- [81] S. Aviyente, E.M. Bernat, W.S. Evans, S.R. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain", *Hum. Brain Mapp.* 32 (1) (2011) 80-93.
- [82] S. Aviyente, A.Y. Mutlu, "A time-frequency-based approach to phase and phase synchrony estimation", *IEEE Trans. Signal Process.* 59 (7) (2011) 3086-3098.
- [83] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals", *Proc. IEEE* 80 (4) (1992) 520-538.
- [84] A. Omidvarnia, G. Azemi, P.B. Colditz, B. Boashash, "A time-frequency based approach for generalized phase synchrony assessment in nonstationary multivariate signals", *Digital Signal Process.* 23 (3) (2013) 780-790.
- [85] B. Boashash, P.J. O'Shea, "Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency", *IEEE Trans. Signal Process.* 41 (3) (1993) 1439-1445.
- [86] P. Boles, B. Boashash, "Practical use of the cross Wigner-Ville distribution to process vibroseis VSP data", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'90), vol. 4, 1990, pp. 1941-1944.
- [87] S. Ventosa, C. Simon, M. Schimmel, J.J. Dañobeitia, A. Mànuel, "The S-transform from a wavelet point of view", *IEEE Trans. Signal Process.* 56 (7) (2008) 2771-2780.
- [88] R.G. Stockwell, L. Mansinha, R.P. Lowe, "Localization of the complex spectrum: the S transform", *IEEE Trans. Signal Process.* 44 (4) (1996) 998-1001.
- [89] C.R. Pinnegar, L. Mansinha, "The S-transform with windows of arbitrary and varying shape", *Geophysics* 68 (1) (2003) 381-385.
- [90] S. Assous, B. Boashash, "Evaluation of the modified S-transform for time-frequency synchrony analysis and source localisation", *EURASIP J. Adv. Signal Process.* 2012 (2012) doi:10.1186/1687-6180-2012-49, article ID 49, 18 pp.
- [91] R.G. Stockwell, "Why use the S-transform?", in: L. Rodino, B.-W. Schulze, M.W. Wong (Eds.), *Pseudo-Differential Operators: Partial Differential Equations and Time-Frequency Analysis*, Fields Institute Communications, vol. 52, American Mathematical Soc., Providence, RI, 2007, pp. 279-309.
- [92] B. Bouachache, B. Escudié, P. Flandrin, J. Gréa, Sur une condition nécessaire et suffisante de positivité de la représentation conjointe en temps et fréquence des signaux d'énergie finie, vol. 288, Ser. A, Comptes rendus de l'Académie des Sciences, Paris, 1979, pp. 307-309 (in French).
- [93] C.W. Helstrom, "An expansion of a signal in Gaussian elementary signals", *IEEE Trans. Inf. Theory* 12 (1) (1966) 81-82.
- [94] L.K. Montgomery, I.S. Reed, "A generalization of the Gabor-Helstrom transform", *IEEE Trans. Inf. Theory* 13 (2) (1967) 344-345.
- [95] R.B. Griffiths, *Consistent Quantum Theory*, University Press, Cambridge, UK, 2002.
- [96] G. Grubb, *Distributions and Operators*, Graduate Texts in Mathematics, vol. 252, Springer, New York, 2009.
- [97] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications", *Proc. IEEE* 80 (4) (1992) 519-568.
- [98] M. Davy, A. Doucet, "Copulas: a new insight into positive time-frequency distributions", *IEEE Signal Process. Lett.* 10 (7) (2003) 215-218.

This page intentionally left blank

PART

TIME-
FREQUENCY
METHODS

III

This page intentionally left blank

ADVANCED IMPLEMENTATION AND REALIZATION OF TFDs

6

INTRODUCTION AND OVERVIEW

Algorithms and computational issues are the keys to efficiently utilizing the properties of time-frequency distributions (TFDs) for real-life applications. This chapter presents therefore the needed procedures, techniques, and methodologies for the effective implementation of such time-frequency (t,f) methods. The topic is covered in six sections with appropriate cross-referencing.

The discrete-time equivalent formulation of quadratic TFDs is defined for the purpose of digital computation ([Section 6.1](#)). An alternative method for realization of quadratic TFDs uses the short-time Fourier transform (STFT) as a basis ([Section 6.2](#)). The Gabor time-frequency representation may be expanded on a rectangular lattice, using the Fourier and Zak transforms for direct implementations ([Section 6.3](#)). The computation of other quadratic TFDs can also be done by using a spectrogram decomposition method ([Section 6.4](#)). Finally, the computational procedure for implementing quadratic time-frequency methods directly is outlined, along with the required algorithms and MATLABTM code fragments ([Section 6.5](#)). The last section focuses on the design of memory-efficient algorithms to implement discrete-time TFDs to deal with the issues of memory limitations when processing large amount of data in applications such as biomedicine, telecommunications, or geophysics ([Section 6.6](#)).

6.1 DISCRETE TIME-FREQUENCY DISTRIBUTIONS⁰

For the purposes of digital storage and processing, any real-life signal that is not discrete-time and time-limited must be made so by sampling and windowing. Moreover, if we wish to evaluate a continuous time-frequency distribution (TFD) numerically in a finite number of operations, we must be content with computing a sampled form of the TFD from a finite number of samples of the signal. For such reasons, we need to define discrete, time-limited equivalents of continuous TFDs. This section derives in a simple way discrete forms of the Wigner-Ville distribution (WVD), the windowed WVD, and the general quadratic TFD, and gives 13 examples of discrete-time kernels. Thus, it extends the material present in [Chapters 2 and 3](#).

If the signal $z(t)$ is ideally sampled [1, p. 299] at times $t = n/f_s$, where n is an integer and f_s (the sampling rate) is a positive constant, it becomes

$$z(t) \sum_{n=-\infty}^{\infty} \delta\left(t - \frac{n}{f_s}\right) = \sum_{n=-\infty}^{\infty} z\left(\frac{n}{f_s}\right) \delta\left(t - \frac{n}{f_s}\right), \quad (6.1.1)$$

where $\delta(\cdot)$ is the unit impulse. A wide caret ($\widehat{\cdot}$) will indicate that a TFD or kernel has been modified by sampling. In the lag (τ) domain, we shall consider *ideal* sampling at $\tau = 2m/f_s$, where m is an integer. This will draw attention to the discrete time values $t = n/f_s$, where n is an integer, so that f_s will be the effective sampling rate in t , although the latter sampling will not be specified as ideal.

6.1.1 THE DISCRETE WIGNER-VILLE DISTRIBUTION

The Wigner distribution (WD) of a continuous-time signal $z(t)$ is

$$W_z(t, f) = \mathcal{F}_{t \rightarrow f} \{z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})\} = \int_{-\infty}^{\infty} z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau, \quad (6.1.2)$$

where $z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$ is called the instantaneous autocorrelation function (IAF). The case in which $z(t)$ is analytic will be considered later; for the present, we merely allow $z(t)$ to be complex.

Theorem 6.1.1. *If $W_z(t, f)$ is modified by ideally sampling the IAF at $\tau = 2m/f_s$, where m is an integer, and if $z(t)$ is time-limited so that*

$$z(t) = 0 \quad \text{for } |t| \geq \frac{N}{2f_s}, \quad (6.1.3)$$

where N is a positive integer, and if the modified TFD is denoted by $\hat{W}_z(t, f)$, then

$$\hat{W}_z\left(\frac{n}{f_s}, \frac{kf_s}{2N}\right) = \sum_{|m| < N/2} z\left(\frac{n+m}{f_s}\right) z^*\left(\frac{n-m}{f_s}\right) e^{-j2\pi km/N} \quad (6.1.4)$$

and the time support for \hat{W}_z in terms of n is the same as for $z(n/f_s)$.

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **G.R. Putland**, Prosper Australia, Melbourne, VIC, Australia (grpputland@gmail.com). Reviewers: G. Matz, LJ. Stanković, and J.M. O'Toole.

Proof/explanation. After sampling, the integrand in Eq. (6.1.2) becomes

$$z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} \sum_{m=-\infty}^{\infty} \delta\left(\tau - \frac{2m}{f_s}\right) \quad (6.1.5)$$

and the WD becomes

$$\hat{W}_z(t, f) = \sum_{m=-\infty}^{\infty} z\left(t + \frac{m}{f_s}\right) z^*\left(t - \frac{m}{f_s}\right) e^{-j4\pi fm/f_s}. \quad (6.1.6)$$

Now consider $t = n/f_s$ where n is an integer. From Eq. (6.1.3), the time support of $z(n/f_s)$ is $|n| < N/2$. Hence, from Eq. (6.1.6), the time support of \hat{W}_z is

$$|n \pm m| < N/2, \quad (6.1.7)$$

where both signs must be satisfied. Because the sign can be chosen so as to increase the magnitude, we must have $|n| + |m| < N/2$, hence $|n| < N/2$ and $|m| < N/2$. So the time support for \hat{W}_z is the same as for $z(n/f_s)$, while the summation in Eq. (6.1.6) is confined to $|m| < N/2$, giving N terms (if we include a zero-padding term for even N). The sampling in τ makes $\hat{W}_z(t, f)$ periodic in f with period $f_s/2$, while the time-limiting gives a frequency resolution of N bins per period. So it is convenient to let

$$f = \frac{k f_s}{2N}, \quad (6.1.8)$$

where k is an integer. With these restrictions, Eq. (6.1.6) reduces to Eq. (6.1.4). \square

With a change of notation, Eq. (6.1.4) becomes

$$W_z[n, k] = \sum_{|m| < N/2} z[n+m] z^*[n-m] e^{-j2\pi km/N}. \quad (6.1.9)$$

This $W_z[n, k]$ is the *discrete WD*. If $z[n]$ has duration not exceeding N samples, the discrete WD is represented as an $N \times N$ real matrix. If the summand is extended periodically in m with period N (i.e., extended periodically in τ with period $2N/f_s$), we obtain

$$W_z[n, k] = \text{DFT}_{m \rightarrow k} \left\{ z[n+m] z^*[n-m] \right\}; \quad m \in \langle N \rangle \quad (\text{and } n \in \langle N \rangle), \quad (6.1.10)$$

where $\langle N \rangle$ means any set of N consecutive integers.

N is not unique; if condition (6.1.3) is true for one value of N , it is true for any higher value. Increasing N is equivalent to zero-padding before taking the DFT, hence increasing the number of frequency bins, that is, upsampling (interpolating) in frequency. This does not extract more information, but gives a smoother TFD graph and perhaps a more convenient fast Fourier transform (FFT) length.

It remains to find the minimum value of f_s that avoids aliasing in the frequency and Doppler domains. As usual, we define the IAF as $K_z(t, \tau) = z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2})$, the spectrum as $Z(f) = \mathcal{F}_{t \rightarrow f} \{z(t)\}$, the WD as $W_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \{K_z(t, \tau)\}$, and the ambiguity function as $A_z(v, \tau) = \mathcal{F}_{t \rightarrow v} \{K_z(t, \tau)\}$. Using these notations and the familiar properties of the FT, it can be shown that

$$W_z(t, f) = \left[2Z(2f) e^{j4\pi ft} \right]_f^* \left[2Z^*(2f) e^{-j4\pi ft} \right] \quad (6.1.11)$$

$$A_z(v, \tau) = \left[Z(v) e^{j\pi v\tau} \right]_v * \left[Z^*(-v) e^{-j\pi v\tau} \right]. \quad (6.1.12)$$

If $Z(f)$ is zero outside the band $|f| < B/2$, then the spectrum given by Eq. (6.1.11), that is, the WD, is zero outside the band $|f| < B/2$, and the spectrum given by Eq. (6.1.12) is zero outside the band $|v| < B$. Aliasing is avoided if the sampling rate in τ , namely $f_s/2$, is at least B and the sampling rate in t , namely f_s , is at least $2B$; that is, aliasing is avoided if $f_s \geq 2B$. This represents $2\times$ oversampling compared with the usual sampling rate needed to represent $z(t)$. The minimum sampling rate in τ makes the WD periodic in f with period B .

As $z(t)$ is considered in general *analytic* (Chapter 2), $W_z(t, f)$ in Eq. (6.1.2) becomes the WVD, and $W_z[n, k]$ becomes the *discrete WVD* or DWVD. Using the analytic signal not only avoids cross-terms between positive and negative frequencies but also avoids the need for oversampling [2]. If $z(t)$ is the analytic associate of the real signal $s(t)$, whose spectrum $S(f)$ is zero outside the band $|f| < B/2$, then $Z(f)$ is zero outside the band $0 \leq f < B/2$ so that the spectrum on the right-hand side of Eq. (6.1.11), that is, the WVD, is zero outside the band $0 \leq f < B/2$, and the spectrum on the right-hand side of Eq. (6.1.12) is zero outside the band $|v| < B/2$. Aliasing will be avoided if the sampling rate in τ is at least $B/2$ and the sampling rate in t is at least B ; that is, aliasing is avoided if $f_s \geq B$, as usual. The minimum sampling rate in τ makes the WVD periodic in f with period $B/2$.

Now consider $W_z[n, k]$ as the element on the n th row and k th column of a matrix W_z . If the sampling rate in t is B , there is no zero-padding in the frequency domain for an analytic signal, and all the columns of W_z are needed for the positive frequencies. If the sampling rate in t is $2B$, only half the columns of W_z are needed to represent the positive frequencies. The negative-frequency elements may be assumed to be zero if $z[n]$ is analytic, provided of course that the assumption of analyticity is not invalidated by short-segment effects, such as windowing of the signal in the n domain and/or windowing of the IAF in the m domain. The latter kind of windowing is discussed in the next section.

6.1.2 THE WINDOWED DWVD

Nonlinear instantaneous frequency (IF) laws and multiple signal components give rise to artifacts (interference terms, cross-terms) in the WD and WVD (see Section 4.2). The effect of nonlinear IF laws can be reduced, with a concomitant loss of frequency resolution, by windowing the IAF in the lag direction (τ or m) before taking the FT. For continuous time, if the window is $g(\tau)$, the resulting TFD is

$$W_z^g(t, f) = \int_{-\infty}^{\infty} g(\tau) z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \quad (6.1.13)$$

and is called the *windowed WD* or *pseudo-WD* or, if $z(t)$ is *analytic*, the *windowed WVD* or *pseudo-WVD*. The effects of sampling and time-limiting on $W_z^g(t, f)$ are described by the following theorem.

Theorem 6.1.2. *If $W_z^g(t, f)$ is modified by ideally sampling $g(\tau)$ at $\tau = 2m/f_s$, where m is an integer, and if $g(\tau)$ is time limited so that*

$$g(\tau) = 0 \quad \text{for } |\tau| \geq M/f_s, \quad (6.1.14)$$

where M is a positive integer, and if the modified TFD is denoted by $\hat{W}_z^g(t, f)$, then

$$\hat{W}_z^g\left(\frac{n}{f_s}, \frac{kf_s}{2M}\right) = \sum_{|m| < M/2} g\left(\frac{2m}{f_s}\right) z\left(\frac{n+m}{f_s}\right) z^*\left(\frac{n-m}{f_s}\right) e^{-j2\pi km/M} \quad (6.1.15)$$

and the time support for \hat{W}_z^g in terms of n is the same as for $z(n/f_s)$.

Proof: Apart from the limits on τ and m , which lead to the substitution $f = \frac{kf_s}{2M}$, the explanation is similar to that of [Theorem 6.1.1](#). \square

With a change of notation, Eq. (6.1.15) becomes

$$W_z^g[n, k] = \sum_{|m| < M/2} g[m] z[n+m] z^*[n-m] e^{-j2\pi km/M}, \quad (6.1.16)$$

where the sampling interval for g is understood to be *twice* the sampling interval for z ; this is significant for interpreting the duration of the “ M -point” window g . We shall call $W_z^g[n, k]$ the *windowed discrete WD* or the *discrete pseudo-WD* or, if $z(t)$ is *analytic*, the *windowed DWVD* or *pseudo-DWVD*. If $z[n]$ has a duration not exceeding N samples, $W_z^g[n, k]$ is represented as an $N \times M$ matrix. If the summand is extended periodically in m with period M (i.e., extended periodically in τ with period $2M/f_s$), padded with a zero term if M is even, we obtain

$$W_z^g[n, k] = \text{DFT}_{m \rightarrow k} \left\{ g[m] z[n+m] z^*[n-m] \right\}; \quad m \in \langle M \rangle. \quad (6.1.17)$$

Again the DFT length (M) may be increased further by zero-padding.

6.1.3 THE DISCRETE QUADRATIC TFD

In terms of continuous variables, the general quadratic TFD (QTFD) is

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} \left\{ G(t, \tau) *_t [z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2})] \right\} \quad (6.1.18)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(u, \tau) z(t - u + \frac{\tau}{2}) z^*(t - u - \frac{\tau}{2}) du e^{-j2\pi f \tau} d\tau, \quad (6.1.19)$$

where $G(t, \tau)$ is the time-lag kernel.

Theorem 6.1.3. *If $\rho_z(t, f)$ is modified by ideally sampling $G(u, \tau)$ at $\tau = 2m/f_s$ and $u = p/f_s$, where m and p are integers, and if $G(u, \tau)$ is time-lag limited so that*

$$G(u, \tau) = 0 \quad \text{for } \left(|\tau| \geq \frac{M}{f_s} \text{ or } |u| \geq \frac{P}{2f_s} \right), \quad (6.1.20)$$

where M and P are positive integers, and if the new TFD is called $\hat{\rho}_z(t, f)$, then

$$\hat{\rho}_z \left(\frac{n}{f_s}, \frac{kf_s}{2M} \right) = \sum_{|m| < \frac{M}{2}} \sum_{|p| < \frac{P}{2}} G \left(\frac{p}{f_s}, \frac{2m}{f_s} \right) z \left(\frac{n-p+m}{f_s} \right) z^* \left(\frac{n-p-m}{f_s} \right) e^{-j2\pi km/M}. \quad (6.1.21)$$

Proof/explanation. The sampled version of $G(u, \tau)$ is

$$\hat{G}(u, \tau) = G(u, \tau) \sum_{p=-\infty}^{\infty} \delta \left(u - \frac{p}{f_s} \right) \sum_{m=-\infty}^{\infty} \delta \left(\tau - \frac{2m}{f_s} \right) = G(u, \tau) \sum_{p=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \delta_2 \left(u - \frac{p}{f_s}, \tau - \frac{2m}{f_s} \right),$$

where $\delta_2(u, \tau)$ is the 2D unit impulse. If $G(u, \tau)$ is replaced by $\hat{G}(u, \tau)$, Eq. (6.1.19) becomes

$$\hat{\rho}_z(t, f) = \sum_{m=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} G\left(\frac{p}{f_s}, \frac{2m}{f_s}\right) z\left(t - \frac{p}{f_s} + \frac{m}{f_s}\right) z^*\left(t - \frac{p}{f_s} - \frac{m}{f_s}\right) e^{-j4\pi fm/f_s}. \quad (6.1.22)$$

By Eq. (6.1.20), the support of $G(u, \tau)$ restricts the summation to $|m| < M/2$ and $|p| < P/2$. The sampling in τ makes $\rho_z(t, f)$ periodic in f with period $f_s/2$, while the sum over m has M terms (if we include a zero-padding term for even M), giving a frequency resolution of M bins per period. So it is convenient to let

$$f = \frac{k f_s}{2M} \quad (6.1.23)$$

where k is an integer. With these restrictions, and with $t = n/f_s$, Eq. (6.1.22) reduces to Eq. (6.1.21). \square

With a change of notation, Eq. (6.1.21) becomes

$$\rho_z[n, k] = \sum_{|m| < \frac{M}{2}} \sum_{|p| < \frac{P}{2}} G[p, m] z[n - p + m] z^*[n - p - m] e^{-j2\pi km/M} \quad (6.1.24)$$

$$= \sum_{|m| < \frac{M}{2}} G[n, m] * (z[n + m] z^*[n - m]) e^{-j2\pi km/M}, \quad (6.1.25)$$

where the sampling interval for G in the lag direction is understood to be twice that in the time direction, and twice the sampling interval for z . This $\rho_z[n, k]$ is the generalized *discrete quadratic TFD* (discrete QTDF).

If the summand is extended in m with period M , Eq. (6.1.25) becomes

$$\rho_z[n, k] = \text{DFT}_{m \rightarrow k} \left\{ G[n, m] * (z[n + m] z^*[n - m]) \right\}; \quad m \in \langle M \rangle. \quad (6.1.26)$$

Thus, the computation involves construction of the discrete IAF $z[n + m] z^*[n - m]$, followed by convolution in n with the time-lag kernel, followed by discrete Fourier transformation. As the time support for the discrete IAF is the same as for $z[n]$, the time support for $\rho_z[n, k]$ and for the convolution in Eqs. (6.1.25) and (6.1.26) is the same as for $G[n, m] * z[n]$. If the convolution in Eqs. (6.1.25) and (6.1.26) is evaluated using an N -point FFT, it will be a modulo- N convolution. But if N is not less than the duration of $G[n, m] * z[n]$, the modulo- N convolution will be equivalent to a linear convolution (i.e., alias-free). The nonzero elements of $\rho_z[n, k]$ will then be represented in an $N \times M$ matrix.

If $G[n, m]$ is real and even in m , then the argument of the DFT is Hermitian in m so that $\rho_z[n, k]$ is real. Symmetries like this can be exploited in the implementation (see Refs. [3,4] and Section 6.6).

It remains to determine the effect of $G(t, \tau)$ on the required sampling rate. If $\mathcal{F}_{\tau \rightarrow f} \{G(t, \tau)\} = \gamma(t, f)$ and $\mathcal{F}_{t \rightarrow v} \{G(t, \tau)\} = g(v, \tau)$, we have the familiar results

$$\mathcal{F}_{\tau \rightarrow f} \left\{ G(t, \tau) * K_z(t, \tau) \right\} = \gamma(t, f) * * W_z(t, f) \quad (6.1.27)$$

$$\mathcal{F}_{t \rightarrow v} \left\{ G(t, \tau) * K_z(t, \tau) \right\} = g(v, \tau) A_z(v, \tau). \quad (6.1.28)$$

Comparing the above with Eqs. (6.1.11) and (6.1.12), we see that the kernel causes an additional spreading of the spectrum in f but not in ν . If $\gamma(t, f)$ has the frequency range $|f| < B_G/2$, then the total bandwidth of the WD or WVD is increased by B_G so that the required sampling rate in τ is increased by B_G and the corresponding sampling rate in t is increased by $2B_G$.

6.1.3.1 Special cases

If $G[n, m] = \delta[n]g[m]$, Eq. (6.1.26) reduces to Eq. (6.1.17); that is, the discrete QTFD reduces to the windowed discrete WD. If, in addition, $g[m] = 1$ (so that $G[n, m] = \delta[n]$), Eqs. (6.1.17) and (6.1.26) reduce to Eq. (6.1.10): the discrete WD.

Two of the above three theorems concern the sampling of a window or kernel function. [Theorem 6.1.1](#) looks different in that the entire IAF is sampled (which is possible only in theory; one cannot compute a continuous IAF in practice). But, because the WD may be considered as a windowed WD with $g(\tau) = 1$, [Theorem 6.1.1](#) can be restated in terms of sampling the lag window, like [Theorem 6.1.2](#).

6.1.3.2 Doppler-frequency form

The Doppler-frequency form of the general quadratic TFD is

$$\rho_z(t, f) = {}_{t \leftarrow \nu} \mathcal{F}^{-1} \left\{ \mathcal{G}(\nu, f) * \left[Z(f + \frac{\nu}{2}) Z^*(f - \frac{\nu}{2}) \right] \right\} \quad (6.1.29)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{G}(\nu, \eta) Z(f - \eta + \frac{\nu}{2}) Z^*(f - \eta - \frac{\nu}{2}) d\eta e^{j2\pi\nu t} d\nu, \quad (6.1.30)$$

where $\mathcal{G}(\nu, f)$ is the Doppler-frequency kernel. As the time-lag form leads to [Theorem 6.1.3](#), so the Doppler-frequency form leads to the following result.

Theorem 6.1.4. *If $\rho_z(t, f)$ is modified by ideally sampling $\mathcal{G}(\nu, \eta)$ at $\nu = \frac{lf_s}{N}$ and $\eta = \frac{qf_s}{2N}$, where l and q are integers and N is a positive integer, and if*

$$\mathcal{G}(\nu, \eta) = 0 \quad \text{for } \left(|\nu| \geq \frac{Lf_s}{2N} \text{ or } |\eta| \geq \frac{Qf_s}{4N} \right), \quad (6.1.31)$$

where L and Q are positive integers, and if the new TFD is called $\hat{\rho}_z(t, f)$, then

$$\hat{\rho}_z \left(\frac{n}{f_s}, \frac{kf_s}{2N} \right) = \sum_{|l| < \frac{L}{2}} \sum_{|q| < \frac{Q}{2}} \mathcal{G} \left(\frac{lf_s}{N}, \frac{qf_s}{2N} \right) Z \left(\frac{|k-q+l|f_s}{2N} \right) Z^* \left(\frac{|k-q-l|f_s}{2N} \right) e^{j2\pi ln/N}. \quad (6.1.32)$$

Proof: Parallel to the proof of [Theorem 6.1.3](#) and the subsequent discussion. □

With a change of notation, Eq. (6.1.32) becomes

$$\rho_z[n, k] = \sum_{|l| < \frac{L}{2}} \sum_{|q| < \frac{Q}{2}} \mathcal{G}[l, q] Z[k - q + l] Z^*[k - q - l] e^{j2\pi ln/N} \quad (6.1.33)$$

$$= \sum_{|l| < \frac{L}{2}} \mathcal{G}[l, k] * (Z[k + l] Z^*[k - l]) e^{j2\pi ln/N}, \quad (6.1.34)$$

where the sampling interval for \mathcal{G} in the Doppler direction is understood to be twice that in the frequency direction, and twice the sampling interval for Z . If the summand is extended in l with period L (including a zero-padding term if L is even), Eq. (6.1.34) becomes

$$\rho_z[n, k] = \text{IDFT}_{n \leftarrow l} \left\{ \frac{\mathcal{G}[l, k]}{k} * (Z[k+l]Z^*[k-l]) \right\}; \quad l \in \langle L \rangle. \quad (6.1.35)$$

Again the DFT length (L) may be increased further by zero-padding. The frequency support for $\rho_z[n, k]$ is the same as for $G[l, k] * Z[k]$. If the latter support has a width not exceeding K frequency samples, then the nonzero elements of $\rho_z[n, k]$ may be represented in an $L \times K$ matrix and the linear convolution in Eqs. (6.1.34) and (6.1.35) may be computed as a modulo- K convolution.

6.1.3.3 Special case: Direct form of the discrete spectrogram

The short-time Fourier transform (STFT) of the continuous-time signal $x(t)$ with real window $w(t)$ is defined (in Section 2.3.1) as

$$F_x^w(t, f) \stackrel{\Delta}{=} \mathcal{F}_{\tau \rightarrow f} \{x(\tau)w(\tau-t)\} = e^{-j2\pi ft} \mathcal{F}_{\tau \rightarrow f} \{x(\tau+t)w(\tau)\} \quad (6.1.36)$$

$$= e^{-j2\pi ft} \int_{-\infty}^{\infty} x(\tau+t)w(\tau) e^{-j2\pi f\tau} d\tau. \quad (6.1.37)$$

It is shown in Chapter 2 that the spectrogram $S_x^w(t, f)$, which is simply the squared magnitude of the STFT, can also be considered as a quadratic TFD with kernel $w(t + \frac{\tau}{2})w(t - \frac{\tau}{2})$. The discrete form of this kernel is $w[n+m]w[n-m]$. Hence, the discrete spectrogram can be conveniently evaluated using the general procedure described in Section 6.5.1 above. But it is simpler and more efficient to discretize the continuous spectrogram directly.

Theorem 6.1.5. *If the spectrogram S_x^w is modified by ideally sampling $w(\tau)$ at*

$$\tau = m/f_s, \quad (6.1.38)$$

where m is an integer and f_s is the sampling rate, and if

$$w(\tau) = 0 \quad \text{for } |\tau| \geq \frac{M}{2f_s}, \quad (6.1.39)$$

where M is a positive integer, and if the modified TFD is denoted by \hat{S}_x^w , then

$$\hat{S}_x^w \left(\frac{n}{f_s}, \frac{kf_s}{M} \right) = \left| \sum_{|m| < M/2} x \left(\frac{m+n}{f_s} \right) w \left(\frac{m}{f_s} \right) e^{-j2\pi km/M} \right|^2. \quad (6.1.40)$$

Proof/explanation. When $w(\tau)$ is sampled, the integrand in Eq. (6.1.37) becomes

$$x(\tau+t)w(\tau) e^{-j2\pi f\tau} \sum_{m=-\infty}^{\infty} \delta \left(\tau - \frac{m}{f_s} \right) \quad (6.1.41)$$

so that the STFT becomes

$$\hat{F}_x^w(t, f) = e^{-j2\pi ft} \sum_{m=-\infty}^{\infty} x \left(\frac{m}{f_s} + t \right) w \left(\frac{m}{f_s} \right) e^{-j2\pi fm/f_s}. \quad (6.1.42)$$

By Eqs. (6.1.38) and (6.1.39), the summation is restricted to $|m| < M/2$, giving M terms (if we include a zero-padding term for even M). The sampling in τ makes $\hat{F}_x^w(t, f)$ periodic in f with period f_s , while the time-limiting in τ gives a frequency resolution of M bins per period. So it is convenient to let

$$f = kf_s/M, \quad (6.1.43)$$

where k is an integer. With these restrictions, Eq. (6.1.42) becomes

$$\hat{F}_x^w\left(t, \frac{kf_s}{M}\right) = e^{-j2\pi kf_st/M} \sum_{|m| < M/2} x\left(\frac{m}{f_s} + t\right) w\left(\frac{m}{f_s}\right) e^{-j2\pi km/M}. \quad (6.1.44)$$

Putting $t = n/f_s$ to match the quantization of τ , then taking the squared magnitude of the discrete STFT, we obtain Eq. (6.1.40).

6.1.4 DESIRABLE PROPERTIES: KERNEL CONSTRAINTS

The desirable properties of continuous TFDs (defined in Section 3.3.1, p. 119ff) are easily redefined for discrete TFDs. Some important examples are given below.

1. *Realness (RE)* says simply that the TFD is real.
2. The *marginal conditions*, which may be considered optional for signal-processing purposes, are the *time marginal (TM)*

$$\sum_k \rho_z[n, k] = N |z[n]|^2, \quad (6.1.45)$$

where N is the FFT length, and the *frequency marginal (FM)*

$$\sum_n \rho_z[n, k] = |Z[k]|^2 \quad (6.1.46)$$

(which differs from Eq. (9) in Ref. [5] due to different scaling of the DFT).

3. The *IF property* says that the periodic first moment of the TFD w.r.t. frequency is the IF. Its dual, which seems to be less used, is the *group delay property (GD)*; it says the periodic first moment of the TFD w.r.t. time is the GD.
4. The *time support (TS)* property says that if $z[n] = 0$ everywhere except $n_1 \leq n \leq n_2$, then $\rho_z[n, k] = 0$ everywhere except $n_1 \leq n \leq n_2$.
5. Similarly, the *frequency support (FS)* property says that if $Z[k] = 0$ everywhere except $k_1 \leq k \leq k_2$, then $\rho_z[n, k] = 0$ everywhere except $k_1 \leq k \leq k_2$.

The class of TFDs satisfying realness, the time and frequency marginals, the IF property, and the time and frequency support properties is called class \mathcal{P} (see discussion in Section 3.1.1). Equations (6.1.45) and (6.1.46) can be approximated for most signal types [6].

The properties listed above are selected from Table 3.3.1 on p. 122. For each property, Table 3.3.1 gives necessary and sufficient conditions on the kernel of the general continuous quadratic TFD. To obtain the corresponding conditions for discrete TFDs, we first express the conditions entirely in the time-lag and Doppler-frequency domains, obtaining the “Continuous” column of Table 6.1.1. Then, we sample the kernels as specified in Theorems 6.1.3 and 6.1.4, obtaining the “Discrete” column of Table 6.1.1. The remaining columns are obtained by specialization.

The sampling of the time-lag kernel will be free of aliasing if the kernel is first band-limited to $\pm f_s/2$ in the t direction (so that $\delta(t)$ is discretized as $\delta[n]$) and to $\pm f_s/4$ in the τ direction. Similarly, the

Table 6.1.1: Kernel Requirements for Selected Properties of Quadratic TFDs of an Analytic Signal, in the Time-Lag and Doppler-Frequency Domains, for General Continuous and Discrete TFDs, and Discrete TFDs With Doppler-Independent (DI) and Lag-Independent (LI) Kernels

| Property | Kernel Constraints | | | |
|----------|--|---|--|--|
| | Continuous | Discrete | Discrete DI $G[n,m] = \delta[n]g_2[m]$ $\mathcal{G}[l,k] = G_2[k]$ | Discrete LI $G[n,m] = g_1[n]$ $\mathcal{G}[l,k] = G_1[l]\delta[k]$ |
| RE | $G(t, \tau) = G^*(t, -\tau)$ | $G[n,m] = G^*[n, -m]$ | $g_2[m] = g_2^*[-m]$ | $g_1[n]$ is real |
| TM | $G(t, 0) = \delta(t)$ | $G[n, 0] = \delta[n]$ | $g_2[0] = 1$ | WVD only |
| FM | $\mathcal{G}(0, f) = \delta(f)$ | $\mathcal{G}[0, k] = \delta[k]$ | WVD only | $G_1[0] = 1$ |
| IF | $G(t, 0) = \delta(t);$ $\int f \mathcal{G}(v, f) dv = 0.$ | $G[n, 0] = \delta[n];$ $\sum_k k \mathcal{G}[l, k] = 0.$ | $g_2[0] = 1;$ $\sum_k k G_2[k] = 0$ | WVD only |
| GD | $\mathcal{G}(0, f) = \delta(f);$ $\int t \mathcal{G}(t, \tau) dt = 0$ | $\mathcal{G}[0, k] = \delta[k];$ $\sum_n n G[n, m] = 0$ | WVD only | $G_1[0] = 1;$ $\sum_n n g_1[n] = 0$ |
| TS | $G(t, \tau) = 0$ if $ \tau < 2 t $. | $G[n, m] = 0$ if $ m < n $ | Always | WVD only |
| FS | $\mathcal{G}(v, f) = 0$ if $ v < 2 f $ | $\mathcal{G}[l, k] = 0$ if $ l < k $ | WVD only | Always |

sampling of the Doppler-frequency kernel will be free of aliasing if the kernel is first time-limited to $\pm N/f_s$ in the f direction (so that $\delta(f)$ is discretized as $\delta[k]$) and to $\pm N/(2f_s)$ in the v direction.

6.1.5 EXAMPLES

By sampling the kernels of common continuous quadratic TFDs (Table 3.3.2, p. 123), we obtain Table 6.1.2, which lists the kernels for the discrete versions of those TFDs. The convolutions with sinc functions are performed *before* restricting the variable to integer values; this requires oversampling and would be computationally inefficient. The convolutions in n and m arise from the band-limiting of the time-lag kernel prior to sampling. This band-limiting of the kernel does *not* affect the result of convolving the kernel with the IAF because the IAF is assumed to be similarly band-limited. Similarly, the convolutions in k and l arise from the time-limiting of the Doppler-frequency kernel prior to sampling. Where no convolution appears in the kernel, either a sinc function has been converted to a discrete delta function by the sampling or a window function is assumed to provide sufficient filtering.

6.1.5.1 Computational issues

The reader is encouraged to make an early use of the (t, f) toolbox TFSAP described in Section 6.5.7 and Chapter 17. TFSAP allows for computing modulated signals, quadratic and polynomial TFDs, ambiguity functions, wavelet transforms and scalograms, and various estimates of instantaneous

Table 6.1.2: Kernels of Selected TFDs in Time-Lag and Doppler-Frequency Domains

| Distribution | Kernel | | Property | | | | | | |
|---------------------|--|--|-----------------|----|----|----|----|----|----|
| | $G[n, m]$ | $\mathcal{G}[l, k]$ | RE | TM | FM | IF | SD | TS | FS |
| WVD | $\delta[n]$ | $\delta[k]$ | ! | ! | ! | ! | ! | ! | ! |
| Levin | $\frac{1}{2}\delta[n+m]$ $+ \frac{1}{2}\delta[n-m]$ | $\frac{1}{2}\delta[k+l]$ $+ \frac{1}{2}\delta[k-l]$ | ! | ! | ! | ! | ! | ! | ! |
| Born-Jordan | $\left[\frac{1}{4 \alpha m } \text{rect}\left(\frac{n}{4\alpha m}\right) \right]$ $* * [\text{sinc } n \text{ sinc } m]$ | $\left[\frac{1}{4 \alpha l } \text{rect}\left(\frac{k}{4\alpha l}\right) \right]$ $* * [\text{sinc } k \text{ sinc } l]$ | ! | ! | ! | ! | ! | * | * |
| Modified B | $\frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n}$ | | ! | | ! | | ! | | ! |
| Windowed WVD | $\delta[n]w[m]$ | $W[k]$ | * | * | | * | | ! | |
| w-Levin | $\frac{1}{2}w[m]\delta[n+m]$ $+ \frac{1}{2}w[m]\delta[n-m]$ | $\frac{1}{2}W[k+l]$ $+ \frac{1}{2}W[k-l]$ | ! | * | | * | | ! | |
| Zhao-Atlas-Marks | $[w[m]\text{rect}\left(\frac{an}{4m}\right)]$ $* * [\text{sinc } n \text{ sinc } m]$ | | * | * | | * | | * | |
| Rihaczek | $\delta[n-m]$ | $\delta[k+l]$ | | ! | ! | | | ! | ! |
| w-Rihaczek | $w^*[-m]\delta[n-m]$ | $W^*[k+l]$ | | * | | | | ! | |
| Page | $\delta[n- m]$ | $\frac{1}{2}\delta[k+l] + \frac{1}{2}\delta[k-l]$ $+ \frac{j l}{\pi(k^2-l^2)} * \text{sinc } l$ | ! | ! | ! | ! | | ! | |
| Exponential | $\frac{\sqrt{\pi\sigma}}{2 m } \exp\left(\frac{-\pi^2\sigma n^2}{4m^2}\right)$ $* * [\text{sinc } n \text{ sinc } m]$ | $\frac{\sqrt{\pi\sigma}}{2 l } \exp\left(\frac{-\pi^2\sigma k^2}{4l^2}\right)$ $* * [\text{sinc } k \text{ sinc } l]$ | ! | ! | ! | ! | ! | ! | |
| B | $\left(\frac{ 2m }{\cosh^2 n}\right)^\beta * \text{sinc } m$ | | ! | | | | | | |
| Spectrogram | $w[n+m]w[n-m]$ | $W[k+l]W[k-l]$ | ! | | | | | | |

Notes: For the spectrogram and windowed Levin (w-Levin) distributions, the window $w[n]$ is assumed to be real and even. $W[k]$ denotes the sampled spectrum of the window $w(\tau)$. In the Property column, an exclamation (!) means that the property is always satisfied, while an asterisk (*) means that the property is satisfied subject to nondegenerate constraints on the window and/or parameter.

frequency. As this is a production package rather than an experimental package, computationally intensive functions are precompiled and optimized for efficiency, and an interactive user interface is added as a supplement. The (t, f) toolbox has been distributed as a MATLAB/Octave toolbox so that TFSAP functions can be used with other computational and graphical functions of MATLAB or Octave.

Information on downloading the package is available from <http://time-frequency.net> (or qspace.qu.edu.qa/handle/10576/10904 or espace.library.uq.edu.au/view/UQ:318062 or elsevier.com). For more details, see Chapter 17.

6.1.6 SUMMARY AND CONCLUSIONS

“Ideal” sampling of window or kernel functions leads to discrete-time forms of the WVD, the windowed WVD, and other quadratic TFDs. Use of the analytic signal minimizes the required sampling rate. Further theoretical details may be found in Refs. [5–8]. Some practical computational issues relating to speed and memory are examined in Sections 6.5 and 6.6.

6.2 QUADRATIC AND HIGHER ORDER TIME-FREQUENCY ANALYSIS BASED ON THE STFT⁰

The oldest, simplest, and most commonly used tool for time-frequency (TF) analysis of a signal $x(t)$ is the spectrogram, defined as the squared magnitude of the short-time Fourier transform (STFT) (Chapter 2). The STFT is

$$F_x(t, f) = \int_{-\infty}^{\infty} x(t + \tau) w(\tau) e^{-j2\pi f\tau} d\tau, \quad (6.2.1)$$

where $w(\tau)$ is a real-valued even lag window (cf. Section 2.3.1). Implementations (hardware and software) of this transform are already widely present in practice. The STFT is linear and very simple for realization. However, it has some serious drawbacks. The most important one lies in its low concentration in the TF plane, when highly nonstationary signals are analyzed. In order to improve TF representation, various quadratic distributions have been introduced. The most important member of this class is the pseudo Wigner distribution (WD)

$$W_x(t, f) = \int_{-\infty}^{\infty} w\left(\frac{\tau}{2}\right) w\left(-\frac{\tau}{2}\right) x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau. \quad (6.2.2)$$

The WD itself has a drawback. Namely, in the case of multicomponent signals, $x(t) = \sum_{p=1}^P x_p(t)$, it produces emphatic cross-terms that can completely mask the auto-terms and make this distribution useless for analysis. This is why many other quadratic reduced interference distributions have been introduced: exponential, Zhao-Atlas-Marks, Born-Jordan, Butterworth, Zhang-Sato, etc. (Section 6.4). The cross-term reduction in these distributions is based on WD smoothing, which inherently leads to the auto-terms’ degradation (Chapter 3; Section 4.2). In contrast to these TF representations, which are focused on the preservation of marginal properties and the cross-term reduction, the S-method (SM), which is the topic of this section, is derived with the primary goal of preserving the auto-terms’ quality as in the WD, while avoiding (reducing) the cross-terms. The software or hardware realization of this method is very efficient, as it is completely based on the STFT. The SM can, in a straightforward manner, be extended to cross-term-free (reduced) realizations of higher order TF representations, time-scale representations, and multidimensional space/spatial-frequency representations.

⁰Author: Lj. Stanković, University of Montenegro, Podgorica, Montenegro (ljubisa@ac.me). Reviewers: J.F. Böhme, S. Carstens-Behrens, and B. Ristic.

6.2.1 STFT-BASED REALIZATION OF THE QUADRATIC REPRESENTATIONS

6.2.1.1 Basic S-method form

The relation between the STFT and the WD, namely [9]

$$W_x(t, f) = 2 \int_{-\infty}^{\infty} F_x(t, f + \theta) F_x^*(t, f - \theta) d\theta, \quad (6.2.3)$$

has led to the definition of a TF representation, referred to as the SM,

$$\text{SM}_x(t, f) = 2 \int_{-\infty}^{\infty} P(\theta) F_x(t, f + \theta) F_x^*(t, f - \theta) d\theta. \quad (6.2.4)$$

The special cases of the SM are two most important TF distributions: (1) for $P(\theta) = 1$ we obtain the WD, $\text{SM}_x(t, f) = W_x(t, f)$; and (2) for $P(\theta) = \delta(\theta)/2$ we obtain the spectrogram, $\text{SM}_x(t, f) = |F_x(t, f)|^2 = S_x(t, f)$. Between these two extreme cases, we can get a gradual transition from the spectrogram to the WD by changing the width of the window $P(\theta)$; if $P(\theta) = 0$ for $|\theta| > L_P$, the width is $2L_P$. The best choice of L_P would be the value when $P(\theta)$ is wide enough to enable complete integration over the auto terms, but narrower than the distance between the auto-terms, in order to avoid the cross terms (Fig. 6.2.1). Then, *the SM produces the sum of the WDs of individual signal components, avoiding cross-terms.*

Proposition. Consider the signal $x(t) = \sum_{p=1}^P x_p(t)$, where $x_p(t)$ are monocomponent signals. Assume that the STFT of each component lies inside the region $D_p(t, f)$, $p = 1, 2, \dots, P$. Denote the length of the p th region along f , for a given t , by $2B_p(t)$, and its central frequency by $f_{0p}(t)$. The SM of $x(t)$ produces the sum of the WDs $W_{x_p}(t, f)$ of each signal component $x_p(t)$,

$$\text{SM}_x(t, f) = \sum_{p=1}^P W_{x_p}(t, f), \quad (6.2.5)$$

if the regions $D_p(t, f)$, $p = 1, 2, \dots, P$, do not overlap (i.e., $D_p(t, f) \cap D_q(t, f) = \emptyset$ for $p \neq q$, meaning a cross-term-free spectrogram), and if the width of the rectangular window $P(\theta)$, for a point (t, f) , is defined by $L_P(t, f) = B_p(t) - |f - f_{0p}(t)|$ for $(t, f) \in D_p(t, f)$, and 0 elsewhere.

Proof: Consider a point (t, f) inside a region $D_p(t, f)$. The integration interval in Eq. (6.2.4), for the p th signal component, is symmetrical with respect to $\theta = 0$. It is defined by the smallest absolute value of θ for which $f + \theta$ or $f - \theta$ falls outside $D_p(t, f)$, that is, $|f + \theta - f_{0p}(t)| \geq B_p(t)$ or $|f - \theta - f_{0p}(t)| \geq B_p(t)$. For $f > f_{0p}(t)$ and positive θ , the integration limit is reached first in $|f + \theta - f_{0p}(t)| \geq B_p(t)$ for $\theta = B_p(t) - (f - f_{0p}(t))$. For $f < f_{0p}(t)$ and positive θ , the limit is reached first in $|f - \theta - f_{0p}(t)| \geq B_p(t)$ for $\theta = B_p(t) - (f_{0p}(t) - f)$. Thus, having in mind the interval symmetry, an integration limit which produces the same value of integral Eq. (6.2.4) as the value of Eq. (6.2.3), over the region $D_p(t, f)$, is given by $L_P(t, f)$ in the Proposition. Therefore, for $(t, f) \in D_p(t, f)$ we have $\text{SM}_x(t, f) = W_{x_p}(t, f)$, since $L_P(t, f) = 0$ for $(t, f) \notin D_p(t, f)$, $p = 1, 2, \dots, P$. \square

Note: any window $P(\theta)$ with constant width $L_P \geq \max_{(t,f)}\{L_P(t,f)\}$ produces $\text{SM}_x(t, f) = \sum_{p=1}^P W_{x_p}(t, f)$, if the regions $D_p(t, f)$, $p = 1, 2, \dots, P$, are at least $2L_P$ apart along the frequency axis, that is, $|f_{0p}(t) - f_{0q}(t)| > B_p(t) + B_q(t) + 2L_P$, for each p, q , and t . This is the SM with constant window width (Eq. 6.2.4). If two components overlap for some time instants t , then a cross-term will appear, but only between these two components and for those time instants.

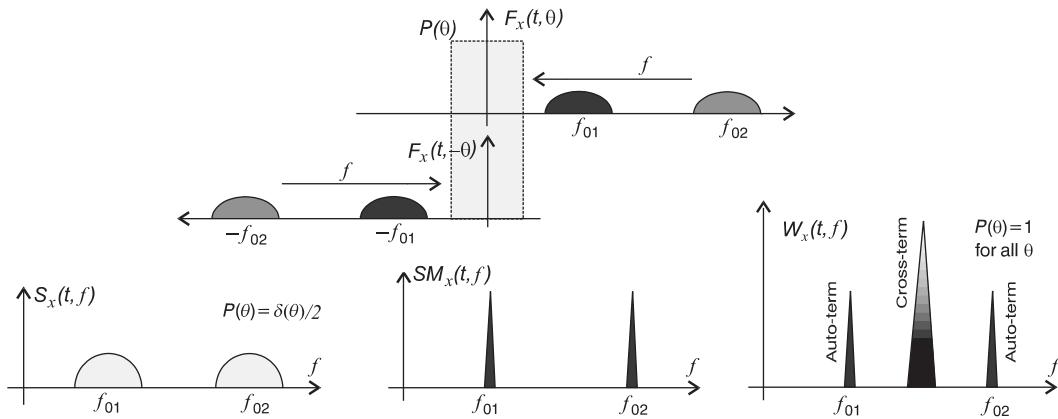
**FIGURE 6.2.1**

Illustration of the SM calculation including two special cases: the WD and the spectrogram.

The SM belongs to the general class of QTFDs, whose inner product form reads

$$\rho_{\tilde{x}}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{G}(t_1, t_2) [x(t + t_1) e^{-j2\pi f t_1}] [\overline{x(t + t_2) e^{-j2\pi f t_2}}]^* dt_1 dt_2. \quad (6.2.6)$$

If the inner product kernel $\tilde{G}(t_1, t_2)$ is factorized in the Hankel form $\tilde{G}(t_1, t_2) = 2w(t_1)p(t_1 + t_2)w(t_2)$, then by substituting its value into Eq. (6.2.6), with $P(-f) = \mathcal{F}_{t \rightarrow f}\{p(t)\}$, we get Eq. (6.2.4). Note that the Toeplitz factorization of the kernel $\tilde{G}(t_1, t_2) = 2w(t_1)p(t_1 - t_2)w(t_2)$ results in the smoothed spectrogram. The smoothed spectrogram composes two STFTs in the same direction, resulting in the distribution spread, in contrast to the SM, where two STFTs are composed in counter direction, resulting in the concentration improvement (Fig. 6.2.1; Section 9.1). The SM has been used to define an efficient signal-decomposition method in Ref. [10].

The Doppler-lag kernel of the SM is given by $g(v, \tau) = P(v/2) *_v A_{ww}(v, \tau)$, where $A_{ww}(v, \tau)$ is the ambiguity function of $w(\tau)$, and $*_v$ denotes a convolution in v . Generally, this kernel is not a separable function.

6.2.1.2 Other forms of the S-method in quadratic representations

Time direction form of the SM is

$$\text{SM}_x(t, f) = 2 \int_{-\infty}^{\infty} P(\tau) F_x(t + \tau, f) F_x^*(t - \tau, f) e^{-j4\pi f \tau} d\tau. \quad (6.2.7)$$

It results from the same analysis as Eq. (6.2.4), based on the frequency domain windowed WD, $W_x(t, f) = \int_{-\infty}^{\infty} W(\theta/2) W(-\theta/2) X(f + \theta/2) X^*(f - \theta/2) \exp(j2\pi t\theta) d\theta$.

Fractional domain form. The frequency and time direction forms of the SM can be generalized to any direction in the time-frequency plane. Consider the fractional FT of $x(t)$, denoted by $X^\alpha(u)$ (Sections 4.8 and 5.8). Its STFT is

$$F_x^\alpha(u, v) = \int_{-\infty}^{\infty} X^\alpha(u + \tau) h(\tau) \exp(-j2\pi v\tau) d\tau, \quad (6.2.8)$$

where $h(\tau)$ is the lag window. The SM in the fractional domain is defined by

$$\text{SM}_x^\alpha(u, v) = 2 \int_{-\infty}^{\infty} P(\theta) F_x^\alpha(u, v + \theta) F_x^{\alpha*}(u, v - \theta) d\theta. \quad (6.2.9)$$

It can be easily realized based on the signal's fractional FT and Eq. (6.2.8).

Using the STFT rotational property [11], that is, $F_x^\alpha(u, v) e^{-j\pi uv} = F_x^0(t, f) e^{-j\pi tf}$ with $u = t \cos \alpha + f \sin \alpha$ and $v = -t \sin \alpha + f \cos \alpha$, we can rewrite Eq. (6.2.9) as

$$\text{SM}_x^\alpha(t, f) = 2 \int_{-\infty}^{\infty} P(\theta) F_x^0(t - \theta \sin \alpha, f + \theta \cos \alpha) F_x^{0*}(t + \theta \sin \alpha, f - \theta \cos \alpha) e^{j4\pi f\theta \sin \alpha} d\theta. \quad (6.2.10)$$

For $\alpha = 0$, it gives Eq. (6.2.4), while Eq. (6.2.7) follows for $\alpha = -\pi/2$. For the derivation of Eq. (6.2.10), the lag window $h(\tau)$ is formally assumed as $(W^\alpha(-\tau))^*$. Optimal direction for the fractional SM calculation can be obtained based on the fractional-moment analysis in Section 4.8. It has been used in Ref. [11].

Affine SM form. The continuous wavelet transform (WT) is defined by $\text{WT}_x(t, f) = (1/\sqrt{|f_0/f|}) \times \int_{-\infty}^{\infty} x(\tau) h^*((\tau - t)f/f_0) d\tau$. As in Section 4.3, we used frequency instead of scale $a = f_0/f$. Consider $h(t) = w(t) \exp(j2\pi f_0 t)$, which provides a strong formal connection of the WT with the STFT. The pseudo affine WD is defined by

$$W_x^a(t, f) = \int_{-\infty}^{\infty} w\left(\frac{\tau}{2f_0}f\right) w\left(-\frac{\tau}{2f_0}f\right) x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi \tau f} d\tau. \quad (6.2.11)$$

The affine SM form reads:

$$\text{SM}_x^a(t, f) = 2 \int_{-\infty}^{\infty} P(\theta) \text{WT}_x(t, f; f_0 + \theta) \text{WT}_x^*(t, f; f_0 - \theta) d\theta, \quad (6.2.12)$$

where $\text{WT}_x(t, f; f_0 + \theta)$ is the WT calculated with $h(t) = w(t) \exp(j2\pi(f_0 + \theta)t)$. If $P(\theta) = \delta(\theta)/2$, then $\text{SM}_x^a(t, f)$ is equal to the scalogram of $x(t)$, while for $P(\theta) = 1$ it produces $W_x^a(t, f)$ defined by Eq. (6.2.11). This form of the SM has been extended to other time-scale representations in Ref. [12].

6.2.2 DISCRETE REALIZATION OF THE BASIC S-METHOD FORM

The discrete SM, for a rectangular window $P(\theta)$, follows from Eq. (6.2.4):

$$\text{SM}_x(n, k) = \frac{2}{T_w} \left[|F_x(n, k)|^2 + 2 \operatorname{Re} \left\{ \sum_{i=1}^{L_p} F_x(n, k+i) F_x^*(n, k-i) \right\} \right], \quad (6.2.13)$$

where $F_x(n, k) = \text{DFT}_{i \rightarrow k} \{x(n+i)w(i)\Delta t\}$, Δt is the sampling interval, $T_w = N\Delta t$ is the width of $w(\tau)$, and $2L_p + 1$ is the width of $P(\theta)$ in the discrete domain. For notational simplicity, we will assume the normalization $2/T_w = 1$. The recursive relation for the SM calculation is [13]

$$\text{SM}_x(n, k; L_P) = \text{SM}_x(n, k; L_P - 1) + 2 \operatorname{Re} \{F_x(n, k + L_P)F_x^*(n, k - L_P)\}, \quad (6.2.14)$$

where $\text{SM}_x(n, k; 0) = |F_x(n, k)|^2$, and $\text{SM}_x(n, k; L_P)$ denotes $\text{SM}_x(n, k)$ in Eq. (6.2.13) calculated with L_P . In this way, we start from the spectrogram, and gradually make the transition toward the WD. The calculation in Eqs. (6.2.13) and (6.2.14) need not be done for each point (n, k) separately. It can be performed for the whole matrix of the SM and the STFT. This can significantly save time in some matrix-based calculation tools. In the SM calculation: (1) There is no need for analytic signal calculation since the cross-terms between negative and positive frequency components are removed in the same way as are the other cross-terms [14]. (2) If we take that $F_x(n, k) = 0$ outside the basic period, that is, when $k < -N/2$ or $k > N/2 - 1$, then there is no aliasing when the STFT is alias-free (in this way, we can calculate the alias-free WD by taking $L_P = N/2$ in Eq. (6.2.13)).

For the SM realization, we have to implement the STFT first, based on either the FFT routines or the recursive approaches suitable for hardware realizations [9,14]. After we get the STFT we have to “correct” the obtained values, according to Eq. (6.2.13), by adding few terms $2 \operatorname{Re} \{F_x(n, k + i)F_x^*(n, k - i)\}$ to the spectrogram values.

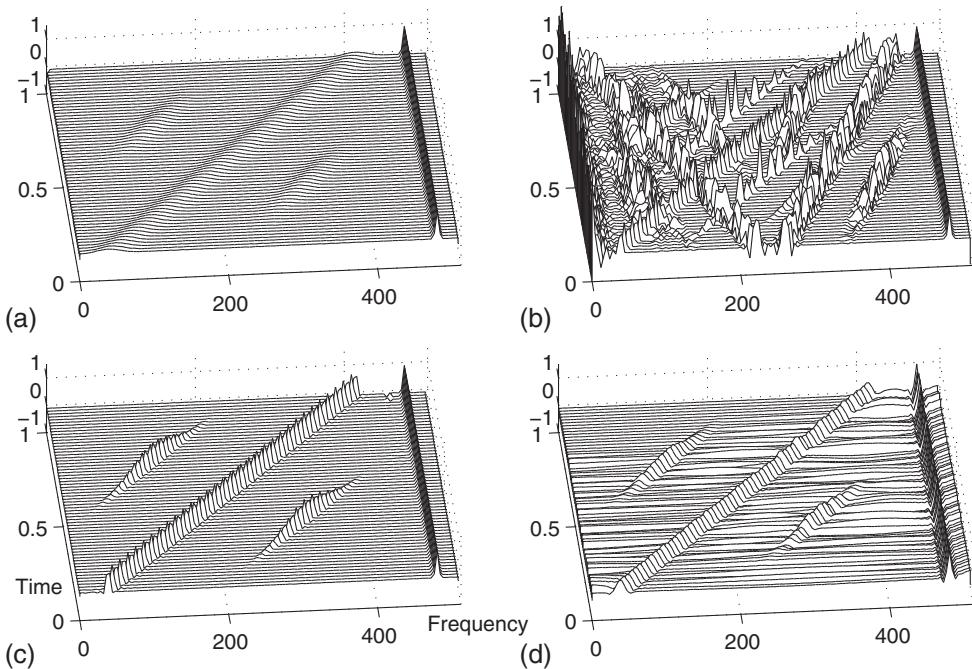
There are two possibilities to implement the summation in Eq. (6.2.13):

- (1) With a signal independent L_P . Theoretically, in order to get the WD for each individual component, the length L_P should be such that $2L_P$ is equal to the width of the widest auto-term. This will guarantee a cross-term-free distribution for all components which are at least $2L_P$ samples apart. For components and time instants where this condition is not satisfied, cross-terms will appear, but still in a reduced form (see also Section 7.3).
- (2) With a signal dependent $L_P = L_P(n, k)$ where the summation, for each point (n, k) , lasts until the absolute square value of $F_x(n, k + i)$ or $F_x(n, k - i)$ is smaller than an assumed reference level R . If a 0 value may be expected within a single auto-term, then the summation lasts until two subsequent 0 values of $F_x(n, k + i)$ or $F_x(n, k - i)$ are detected. The reference level is defined as a few percent of the spectrogram’s maximal value at a considered instant n :
 $R_n = \max_k \{S_x(n, k)\}/Q^2$, where Q is a constant. Index n is added to show that the reference level R is time dependent. Note that if $Q^2 \rightarrow \infty$, the WD will be obtained, while $Q^2 = 1$ results in the spectrogram. A choice of an appropriate value for design parameter Q^2 will be discussed in Example 6.2.2.

Example 6.2.1. Consider a real-valued multicomponent signal

$$\begin{aligned} x(t) = & \cos(1200(t + 0.1)^2) + e^{-36(t-1/3)^2} \cos(1200(t + 1/2)^2) \\ & + e^{-36(t-2/3)^2} \cos(1200(t - 1/3)^2) + \cos(960\pi t) \end{aligned}$$

within the interval $[0, 1]$, sampled at $\Delta t = 1/1024$. This sampling rate is very close to the Nyquist rate for this signal, that is $1/960$. The Hanning window of the width $T_w = 1/4$ is used. The spectrogram is shown in Fig. 6.2.2(a). Its “corrected” version (the SM), according to Eq. (6.2.13), with five terms, $L_P = 5$, is shown in Fig. 6.2.2(c). The auto-terms are concentrated almost as in the WD, Fig. 6.2.2(b). The exponential distribution (ED), whose kernel reads $g(v, \tau) = \exp(-(\nu\tau)^2)$, is shown in Fig. 6.2.2(d). Normalized values, $-\sqrt{\pi N/2} \leq |2\pi\nu| \leq \sqrt{\pi N/2}$, $-\sqrt{\pi N/2} \leq |\tau| \leq \sqrt{\pi N/2}$, and 128 samples within that interval, are used. If the analytic part of $x(t)$ were used, similar results would be obtained (see Fig. 9.1.1 on p. 530).

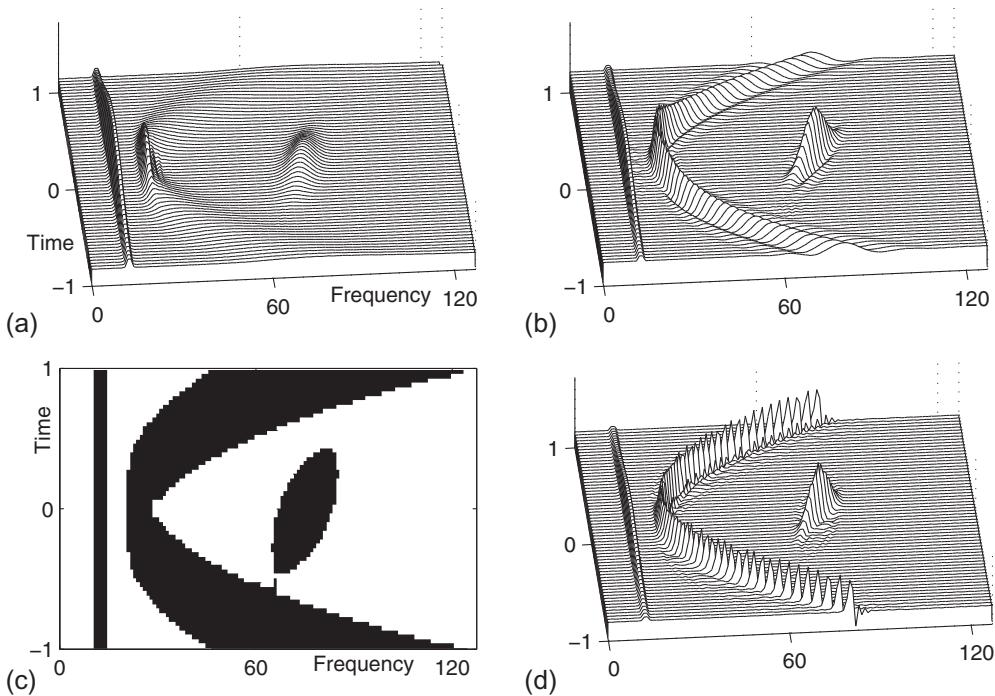
**FIGURE 6.2.2**

Time-frequency representation of a real-valued multicomponent signal: (a) spectrogram; (b) pseudo Wigner distribution; (c) S-method with five “correcting terms,” $L_P = 5$; and (d) exponential distribution, as the representative of reduced interference distributions.

Example 6.2.2. The adaptive SM realization will be illustrated on a three-component real signal, with a nonlinear FM component,

$$x(t) = e^{-t^2} \cos(25\pi t) + \cos(120t^3 + 45\pi t) + 1.5 e^{-25t^2} \cos(40\pi t^2 + 150\pi t)$$

with the sampling interval $\Delta t = 1/256$. The signal is considered within the time interval $[-1, 1]$. The Hanning window of the width $T_w = 1$ is used. The spectrogram is presented in Fig. 6.2.3(a), while the SM with the constant $L_P = 3$ is shown in Fig. 6.2.3(b). The concentration improvement with respect to the case $L_P = 0$, Fig. 6.2.3(a), is evident. Further increase of L_P would improve concentration, but it would also cause some cross-terms to appear. Some small changes are already noticeable between the components with quadratic and constant IF. An improved concentration, without cross-terms, can be achieved by using the variable window width L_P . The regions $D_i(n, k)$, determining the summation limit $L_P(n, k)$ for each point (n, k) , are obtained by imposing the reference level R_n corresponding to $Q^2 = 50$. They are defined as: $D_i(n, k) = 1$ when $S_x(n, k) \geq R_n = \max_k \{S_x(n, k)\}/Q^2$, and $D_i(n, k) = 0$ elsewhere (Fig. 6.2.3(c)). Clear regions mean that the value of the spectrogram is below 2% of its maximal value at that time instant n , meaning that the concentration improvement is not performed at these points. The signal-dependent SM is given in Fig. 6.2.3(d). The method sensitivity, with respect to the value of Q^2 , is low.

**FIGURE 6.2.3**

Time-frequency analysis of a multicomponent signal: (a) spectrogram; (b) the S-method with a constant window, with $L_P = 3$; (c) regions of support for the S-method with a variable window width calculation, corresponding to $Q^2 = 50$; and (d) the S-method with the variable window width calculated using regions in (c).

6.2.3 STFT-BASED REALIZATION OF HIGHER ORDER REPRESENTATIONS

In order to improve distribution concentration in the case of nonlinear FM signals, the higher order time-varying spectra have been defined (Wigner higher order spectra, multitime Wigner distributions). For practical realizations, the most interesting are the versions of these spectra reduced to the 2D TF plane [15]. Here, we will present the L-Wigner distribution (LWD) and the fourth-order polynomial Wigner-Ville distribution (PWVD).

6.2.3.1 The L-Wigner distribution

The LWD is defined by [16]

$$\text{LW}_L(t, f) = \int_{-\infty}^{\infty} w_L(\tau) x^*{}^L(t - \frac{\tau}{2L}) x^L(t + \frac{\tau}{2L}) e^{-j2\pi f\tau} d\tau. \quad (6.2.15)$$

For $L = 1$, it reduces to the WD.

The LWD is a windowed slice of the multitime WDs

$$W_x^K(t_1, \dots, t_k, f) = \int_{-\infty}^{\infty} x^* \left(\sum_{i=1}^K t_i + \frac{\tau}{K+1} \right) \prod_{i=1}^{L-1} x^* \left(-t_i + \frac{\tau}{K+1} \right) \prod_{i=L}^K x \left(t_i - \frac{\tau}{K+1} \right) e^{j2\pi f \tau} d\tau,$$

along the line $t_1 = t_2 = \dots = t_{L-1} = -t$, $t_L = t_{L+1} = \dots = t_K = t$, where the auto-terms in $W_x^K(t_1, \dots, t_k, f)$ are located, for $L = (K+1)/2$ [17].

Similarly, starting from the Wigner higher order spectra, dual to $W_x^K(t_1, t_2, \dots, t_k, f)$, we get a distribution dual to Eq. (6.2.15),

$$\text{LW}_L^f(t, f) = \int_{-\infty}^{\infty} W_L(\theta) X^*{}^L(f + \frac{\theta}{2L}) X^L(f - \frac{\theta}{2L}) e^{-j2\pi \theta t} d\theta, \quad (6.2.16)$$

studied in detail in Ref. [18]. Its realization is formally the same as for the time-domain LWD.

For a frequency modulated signal $x(t) = \exp(j\phi(t))$, the LWD produces [14]

$$\text{LW}_L(t, f) = W_L(f - \phi'(t)/2\pi) * \underset{f}{\text{FT}} \left\{ \exp \left(j \frac{\phi^{(3)}(t+\tau_1) + \phi^{(3)}(t-\tau_2)}{48L^2} \right) \right\},$$

where τ_1, τ_2 are the values of τ within the lag window $w_L(\tau)$, and $W_L(f) = \mathcal{F}_{\tau \rightarrow f}\{w_L(\tau)\}$. For $L \rightarrow \infty$, the LWD tends to a distribution completely concentrated along the IF, that is, $\text{LW}_L(t, f) \rightarrow W_L(f - \phi'(t)/2\pi)$.

The relationship between $\text{LW}_{2L}(t, f)$ and $\text{LW}_L(t, f)$ is of form Eq. (6.2.3):

$$\text{LW}_{2L}(t, f) = 2 \int_{-\infty}^{\infty} \text{LW}_L(t, f + \theta) \text{LW}_L(t, f - \theta) d\theta.$$

The realization of the cross-term-free and alias-free version of the LWD may be efficiently done in the discrete domain, by using the SM form Eq. (6.2.13), as

$$\text{LW}_{2L}(n, k) = \text{LW}_L^2(n, k) + 2 \sum_{i=1}^{L_P} \text{LW}_L(n, k+i) \text{LW}_L(n, k-i), \quad (6.2.17)$$

with $\text{LW}_1(n, k) = W_x(n, k)$, and $W_x(n, k)$ calculated according to Eq. (6.2.13). Form Eq. (6.2.17) is very convenient for software and hardware realizations since the same blocks, connected in cascade, can provide a simple and efficient system for higher order TF analysis, based on the STFT in the initial step, and the signal sampled at the Nyquist rate. Numerical examples and illustrations of the LWD can be found in Refs. [14,16–19].

6.2.3.2 Polynomial Wigner-Ville distribution

Modification of the presented method for the realization of the PWVD is straightforward. The fourth-order PWVD can be written in a frequency scaled form (Section 5.4)

$$\text{PW}_x(t, f) = \frac{1}{27} \int_{-\infty}^{\infty} x^2 \left(t + \frac{\tau}{4} \right) x^{*2} \left(t - \frac{\tau}{4} \right) x^* \left(t + A \frac{\tau}{2} \right) x(t - A \frac{\tau}{2}) e^{-j \frac{2\pi f}{27} \tau} d\tau, \quad (6.2.18)$$

where $A = 0.85/1.35$ and $f' \equiv f/2.7$. Note that $\text{PW}_x(t, f') = \frac{1}{2.7} \text{LW}_2(t, f') * W_x^A(t, f')$, where $W_x^A(t, f') = \text{FT}\{x^*(t + A\frac{\tau}{2})x(t - A\frac{\tau}{2})\}$ is the scaled and reversed version of the WD. The cross-term-free realization of the WD and LWD is already presented. In the discrete implementation of the above relation, the only remaining problem is to evaluate $W_x^A(t, f')$ on the discrete set of points on the frequency axis, $f' = -k\Delta f'$. Since $W_x^A(t, f')$ is, by definition, a scaled and reversed version of $W_x(t, f')$, its values at $f' = -k\Delta f'$ are the values of $W_x(t, f')$ at $f' = k\Delta f'/A$. However, these points do not correspond to any sample location along the frequency axis grid. Thus, the interpolation has to be done: one way of doing it is in an appropriate zero padding of the signal. A discrete form of convolution (Eq. 6.2.18), including the rectangular window $P(\theta)$ and the above considerations, is

$$\text{PW}_x(n, k) = \sum_{i=-L_p}^{L_p} \text{LW}_2(n, k + i) \hat{W}_x(n, k + i/A), \quad (6.2.19)$$

where $2L_p + 1$ is the width of $P(\theta)$ in the discrete domain, while $\hat{W}_x(n, k + i/A)$ is the WD approximation. We can simply use $\hat{W}_x(n, k + i/A) = \text{SM}_x(n, k + [i/A])$ where $[i/A]$ is the nearest integer to i/A , or use the linear interpolation of the SM values at two nearest integers. The terms in Eq. (6.2.19), when $k + i$ or $k + [i/A]$ is outside the basic period, are considered as being zero in order to avoid aliasing.

Example 6.2.3. Consider real-valued multicomponent signal

$$x(t) = \cos(20 \sin(\pi t) + 30\pi t) + \sin(20 \cos(\pi t) + 100\pi t)$$

within $-1 \leq t < 1$, with $\Delta t = 1/128$. In the realization, a Hanning window of the width $T_w = 2$ is used. Based on the STFT (using its positive frequencies), the cross-term-free WD is obtained from Eq. (6.2.13) with $L_p = 15$, and denoted by SM (Fig. 6.2.4(a)). Then the LWD, with $L = 2$, is calculated according to Eq. (6.2.17). It is combined with the linearly interpolated SM value into the PWVD Eq. (6.2.19), shown in Fig. 6.2.4(b). For the precise implementation of $[i/A]$, the lag window has been zero-padded by a factor of 2. An SM implementation of a higher order TFD-based IF estimator is presented in Ref. [20].

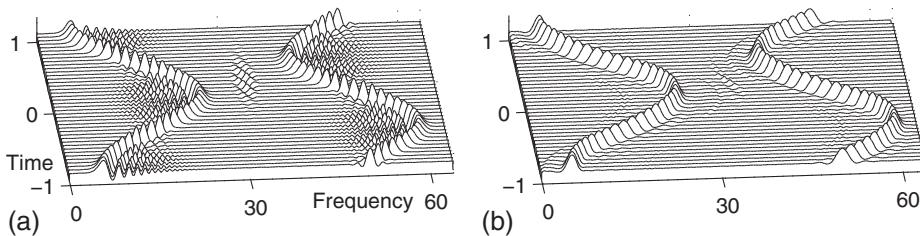


FIGURE 6.2.4

Time-frequency representation of a real-valued multicomponent signal: (a) the SM (cross-term-free and alias-free version of the WD); (b) polynomial Wigner-Ville distribution realization based on the STFT, using the SM and its order-recursive form.

6.2.4 SUMMARY AND CONCLUSIONS

The STFT-based realization of quadratic TF representations, having auto-terms close to or the same as the ones in the WD, but without (or with reduced) cross-terms, is presented. For this realization, the S-method is used. The method is generalized, in an order-recursive form, for the realization of higher order TF representations. Applications of the presented method on, for example, time-scale representations [12], and multidimensional space/spatial-frequency analysis [21], are straightforward. Hardware realization of the S-method is also simple and direct [14]. A general approach to the cross-term-free calculation of convolutions, based on the SM principle, is described in Sections 3.3.4.1 and 5.5 [16,20].

6.3 GABOR'S SIGNAL EXPANSION FOR A NONORTHOGONAL SAMPLING GEOMETRY⁰

6.3.1 HISTORICAL PERSPECTIVE

In 1946, Gabor [22] suggested the representation of a time signal in a combined time-frequency domain (see Section 2.3.3); in particular, he proposed to represent the signal as a superposition of shifted and modulated versions of a so-called elementary signal or synthesis window $g(t)$. Moreover, as a synthesis window $g(t)$, he chose a Gaussian signal because such a signal has good localization in both the time domain and the frequency domain. The other choice that Gabor made was that his signal expansion was formulated on a rectangular lattice in the time-frequency domain $(mT, k\Omega)$, where the sampling distances T and Ω satisfied the relation $\Omega T = 2\pi$. (In this section, the standard time-frequency notation (t, f) is replaced by (t, ω) where $\omega = 2\pi f$ so as to align with angle rotations notations.)

The coefficients in Gabor's signal expansion can be determined by using an analysis window $w(t)$. In the case of critical sampling, that is, $\Omega T = 2\pi$, the analysis window $w(t)$ follows uniquely from the given synthesis window $g(t)$. However, such a unique analysis window appears to have some mathematically very unattractive properties. For this reason, the expansion should be formulated on a denser lattice, $\Omega T < 2\pi$. This makes the analysis window no longer unique and thus allows for finding an analysis window that is optimal in some way. We can, for instance, look for the analysis window that resembles best the synthesis window; a better resemblance can then be reached for a higher degree of oversampling.

A better resemblance can also be reached if we adapt the structure of the lattice to the form of the window as represented in the time-frequency domain. For the Gaussian window, for instance, the time-frequency representation has circular contour lines, and circles are better packed on a hexagonal lattice than on a rectangular lattice. Gabor's signal expansion on such a hexagonal, nonorthogonal lattice then leads to a better resemblance between the window functions $g(t)$ and $w(t)$ than the expansion on a rectangular, orthogonal lattice does.

⁰Authors: M.J. Bastiaans, Eindhoven University of Technology, Eindhoven, The Netherlands (m.j.bastiaans@tue.nl) and A.J. van Leest, ASML, Veldhoven, The Netherlands. Reviewers: J.M. Morris and S. Qian.

6.3.2 GABOR'S SIGNAL EXPANSION ON A RECTANGULAR LATTICE

We start with the usual Gabor expansion [22–26] on a rectangular time-frequency lattice, in which case a signal $\varphi(t)$ can be expressed as a linear combination of properly shifted and modulated versions $g_{mk}(t) = g(t - mT) \exp(jk\Omega t)$ of a synthesis window $g(t)$:

$$\varphi(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{mk} g_{mk}(t). \quad (6.3.1)$$

The time step T and the frequency step Ω satisfy the relationship $\Omega T \leq 2\pi$; note that the factor $2\pi/\Omega T$ represents the degree of oversampling, and that in his original paper [22] Gabor considered the case of critical sampling, that is, $\Omega T = 2\pi$. The expansion coefficients a_{mk} follow from sampling the windowed Fourier transform with analysis window $w(t)$, $\int_{-\infty}^{\infty} \varphi(t)w^*(t - \tau) \exp(-j\omega t) dt$, on the rectangular lattice ($\tau = mT, \omega = k\Omega$):

$$a_{mk} = \int_{-\infty}^{\infty} \varphi(t)w_{mk}^*(t) dt. \quad (6.3.2)$$

This relationship is known as the Gabor transform.

The synthesis window $g(t)$ and the analysis window $w(t)$ are related to each other in such a way that their shifted and modulated versions constitute two sets of functions that are biorthogonal:

$$\sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g_{mk}(t_1)w_{mk}^*(t_2) = \delta(t_1 - t_2). \quad (6.3.3)$$

If the biorthogonality condition (6.3.3) is satisfied, the Gabor transform (Eq. 6.3.2) and Gabor's signal expansion (Eq. 6.3.1) form a transform pair in the following sense: if we start with an arbitrary signal $\varphi(t)$ and determine its expansion coefficients a_{mk} via the Gabor transform (Eq. 6.3.2), the signal can be reconstructed via the Gabor expansion (Eq. 6.3.1).

The biorthogonality relation (Eq. 6.3.3) leads immediately to the equivalent but simpler expression

$$\frac{2\pi}{\Omega} \sum_{m=-\infty}^{\infty} g(t - mT)w^*\left(t - \left[mT + n\frac{2\pi}{\Omega}\right]\right) = \delta_n, \quad (6.3.4)$$

where δ_n is the Kronecker delta. In the case of critical sampling, that is, $\Omega T = 2\pi$, the biorthogonality relation (Eq. 6.3.4) reduces to

$$T \sum_{m=-\infty}^{\infty} g(t - mT)w^*(t - [m+n]T) = \delta_n \quad (6.3.5)$$

and the analysis window $w(t)$ follows uniquely from a given synthesis window $g(t)$, or vice versa. An elegant way to find the analysis window if the synthesis window is given is presented in the next section.

6.3.3 FOURIER TRANSFORM AND ZAK TRANSFORM

It is well known (see, for instance, Refs. [23–26]) that in the case of critical sampling, $\Omega T = 2\pi$, Gabor's signal expansion (Eq. 6.3.1) and the Gabor transform (Eq. 6.3.2) can be transformed into

product form. We therefore need the Fourier transform $\bar{a}(t/T, \omega/\Omega)$ of the 2D array of Gabor coefficients a_{mk} , defined by

$$\bar{a}(x, y) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{mk} \exp[-j2\pi(my - kx)], \quad (6.3.6)$$

and the Zak transforms $\tilde{\varphi}(xT, 2\pi y/T; T)$, $\tilde{g}(xT, 2\pi y/T; T)$, and $\tilde{w}(xT, 2\pi y/T; T)$ of the signal $\varphi(t)$ and the window functions $g(t)$ and $w(t)$, respectively, where the Zak transform $\tilde{s}(t, \omega; \tau)$ of a function $s(t)$ is defined as (see, for instance, [23,26])

$$\tilde{s}(t, \omega; \tau) = \sum_{n=-\infty}^{\infty} s(t + n\tau) \exp(-jn\tau\omega). \quad (6.3.7)$$

Note that the Fourier transform $\bar{a}(x, y)$ is periodic in x and y with period 1, and that the Zak transform $\tilde{s}(t, \omega; \tau)$ is periodic in ω with period $2\pi/\tau$ and quasi-periodic in t with period τ : $\tilde{s}(t + m\tau, \omega + 2\pi k/\tau; \tau) = \tilde{s}(t, \omega; \tau) \exp(jm\omega\tau)$.

Upon substituting from the Fourier transform (Eq. 6.3.6) and the Zak transforms (cf. Eq. 6.3.7) into Eqs. (6.3.1) and (6.3.2), it is not too difficult to show that Gabor's signal expansion (Eq. 6.3.1) can be transformed into the product form

$$\tilde{\varphi}\left(xT, y\frac{2\pi}{T}; T\right) = \bar{a}(x, y) \tilde{g}\left(xT, y\frac{2\pi}{T}; T\right), \quad (6.3.8)$$

while the Gabor transform (Eq. 6.3.2) can be transformed into the product form

$$\bar{a}(x, y) = T \tilde{\varphi}\left(xT, y\frac{2\pi}{T}; T\right) \tilde{w}^*\left(xT, y\frac{2\pi}{T}; T\right). \quad (6.3.9)$$

In particular, the product form Eq. (6.3.9) is useful for determining Gabor's expansion coefficients. Since a Zak transform is merely a Fourier transform (cf. Eq. 6.3.7), the expansion coefficients can be determined by Fourier transformations and multiplications; and if things are formulated for discrete-time signals, we can use the *fast* Fourier transform to formulate a fast algorithm for the Gabor transform [24,25].

The relationship between the Zak transforms of the analysis window $w(t)$ and the synthesis window $g(t)$ then follows from substituting from Eq. (6.3.9) into Eq. (6.3.8) and reads

$$T \tilde{g}\left(xT, y\frac{2\pi}{T}; T\right) \tilde{w}^*\left(xT, y\frac{2\pi}{T}; T\right) = 1. \quad (6.3.10)$$

From the latter relationship, we conclude that (the Zak transform of) the analysis window $w(t)$ follows uniquely from (the Zak transform of) the given synthesis window $g(t)$. In general, however, the unique analysis window $w(t)$ has some very unattractive mathematical properties. We are therefore urged to consider Gabor's signal expansion on a denser lattice, in which case the analysis window is no longer unique. This enables us to choose an analysis window that is better suited to our purpose of determining Gabor's expansion coefficients.

6.3.4 RATIONAL OVERSAMPLING

In the case of oversampling by a rational factor, $2\pi/\Omega T = p/q \geq 1$, with p and q relatively prime, positive integers and $p > q \geq 1$, Gabor's expansion (Eq. 6.3.1) and the Gabor transform (Eq. 6.3.2) can be transformed into the sum-of-products forms [24,25], cf. Eqs. (6.3.8) and (6.3.9),

$$\varphi_s(x, y) = \frac{1}{p} \sum_{r=0}^{p-1} g_{sr}(x, y) a_r(x, y) \quad (s = 0, 1, \dots, q-1) \quad (6.3.11)$$

$$a_r(x, y) = \frac{pT}{q} \sum_{s=0}^{q-1} w_{sr}^*(x, y) \varphi_s(x, y) \quad (r = 0, 1, \dots, p-1), \quad (6.3.12)$$

respectively, where we have introduced the shorthand notations

$$\begin{aligned} a_r(x, y) &= \tilde{a}(x, y + r/p) \\ \varphi_s(x, y) &= \tilde{\varphi}((x+s)pT/q, 2\pi y/T; pT) \\ g_{sr}(x, y) &= \tilde{g}((x+s)pT/q, 2\pi(y+r/p)/T; T) \\ w_{sr}(x, y) &= \tilde{w}((x+s)pT/q, 2\pi(y+r/p)/T; T), \end{aligned}$$

with $0 \leq x < 1$ and $s = 0, 1, \dots, q-1$ (and hence $0 \leq (x+s)/q < 1$), and $0 \leq y < 1/p$ and $r = 0, 1, \dots, p-1$ (and hence $0 \leq y + r/p < 1$). The relationship between the Zak transforms of the analysis window $w(t)$ and the synthesis window $g(t)$ then follows from substituting from Eq. (6.3.12) into Eq. (6.3.11) and reads (cf. Eq. 6.3.10)

$$\frac{T}{q} \sum_{r=0}^{p-1} g_{s_1 r}(x, y) w_{s_2 r}^*(x, y) = \delta_{s_1 - s_2}, \quad (6.3.13)$$

with $s_1, s_2 = 0, 1, \dots, q-1$. The latter relationship represents a set of q^2 equations for pq unknowns, which set is underdetermined since $p > q$, and we conclude that the analysis window does not follow uniquely from the synthesis window.

After combining the p functions $a_r(x, y)$ into a p -dimensional column vector $\mathbf{a}(x, y)$, the q functions $\varphi_s(x, y)$ into a q -dimensional column vector $\boldsymbol{\phi}(x, y)$, and the $q \times p$ functions $g_{sr}(x, y)$ and $w_{sr}(x, y)$ into the $q \times p$ -dimensional matrices $\mathbf{G}(x, y)$ and $\mathbf{W}(x, y)$, respectively, the sum-of-products forms can be expressed as matrix-vector and matrix-matrix multiplications:

$$\boldsymbol{\phi}(x, y) = \frac{1}{p} \mathbf{G}(x, y) \mathbf{a}(x, y) \quad (6.3.14)$$

$$\mathbf{a}(x, y) = \frac{pT}{q} \mathbf{W}^*(x, y) \boldsymbol{\phi}(x, y) \quad (6.3.15)$$

$$\mathbf{I}_q = \frac{T}{q} \mathbf{G}(x, y) \mathbf{W}^*(x, y), \quad (6.3.16)$$

where \mathbf{I}_q denotes the $q \times q$ -dimensional identity matrix and where, as usual, the asterisk in connection with vectors and matrices denotes complex conjugation and transposition.

The latter relationship again represents q^2 equations for pq unknowns, and the $p \times q$ matrix $\mathbf{W}^*(x, y)$ cannot be found by a simple inversion of the $q \times p$ matrix $\mathbf{G}(x, y)$. An “optimum” solution that is often used, is based on the generalized inverse and reads $\mathbf{W}_{\text{opt}}^*(x, y) = (q/T) \mathbf{G}^*(x, y) [\mathbf{G}(x, y) \mathbf{G}^*(x, y)]^{-1}$. This solution for $\mathbf{W}(x, y)$ is optimum in the sense that (i) it yields the analysis window $w(t)$ with the

lowest L^2 norm, (ii) it yields the Gabor coefficients a_{mk} with the lowest L^2 norm, and (iii) it yields the analysis window that—in an L^2 sense, again—best resembles the synthesis window.

The “optimum” solution gets better if the degree of oversampling p/q becomes higher. However, there is another way of finding a better solution, based on the structure of the lattice. If the lattice structure is adapted to the form of the window function as it is represented in the time-frequency domain, the “optimum” solution will be better, even for a lower degree of oversampling. We will therefore consider the case of a nonorthogonal sampling geometry, but we will do that in such a way that we can relate this nonorthogonal sampling to orthogonal sampling. In that case, we will still be able to use product forms of Gabor's expansion and the Gabor transform, and benefit from all the techniques that have been developed for them.

6.3.5 NONORTHOGONAL SAMPLING

The rectangular (or orthogonal) lattice that we considered in the previous sections, where sampling occurred on the lattice points ($\tau = mT, \omega = k\Omega$), can be obtained by integer combinations of two orthogonal vectors $[T, 0]^t$ and $[0, \Omega]^t$ (see Fig. 6.3.1(a)), which vectors constitute the lattice generator matrix

$$\begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix}.$$

We now consider a time-frequency lattice that is no longer orthogonal. Such a lattice is obtained by integer combinations of two linearly independent, but no longer orthogonal vectors, which we express in the forms $[aT, c\Omega]^t$ and $[bT, d\Omega]^t$, with a, b, c and d integers, and which constitute the lattice generator matrix

$$\begin{bmatrix} aT & bT \\ c\Omega & d\Omega \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Without loss of generality, we may assume that the integers a and b have no common divisors, and that the same holds for the integers c and d ; possible common divisors can be absorbed in T and Ω . Note

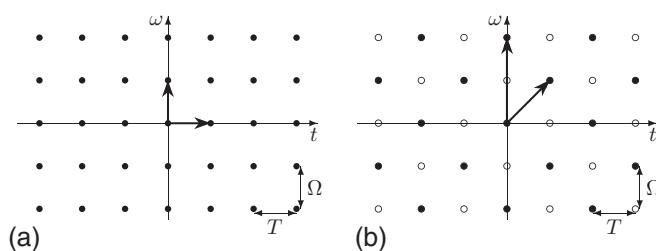


FIGURE 6.3.1

(a) A rectangular lattice with lattice vectors $[T, 0]^t$ and $[0, \Omega]^t$, and thus $R = 0$ and $D = 1$; and (b) a hexagonal lattice with lattice vectors $[T, \Omega]^t$ and $[0, 2\Omega]^t$, and thus $R = 1$ and $D = 2$.

that we only consider lattices that have samples on the time and frequency axes and that are therefore suitable for a discrete-time approach, as well.

The area of a cell (a parallelogram) in the time-frequency plane, spanned by the two vectors $[aT, c\Omega]^t$ and $[bT, d\Omega]^t$, is equal to the determinant of the lattice generator matrix, which determinant is equal to ΩTD , with $D = |ad - bc|$. To be usable as a proper Gabor sampling lattice, this area should satisfy the condition $D \leq 2\pi/\Omega T$.

There are a lot of lattice generator matrices that generate the same lattice. We will use the one that is based on the Hermite normal form, unique for any lattice,

$$\begin{bmatrix} T & 0 \\ R\Omega & D\Omega \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix} \begin{bmatrix} 1 & 0 \\ R & D \end{bmatrix},$$

where R and D are relatively prime integers and $0 \leq |R| < D$. Sampling then occurs on the lattice points ($\tau = mT, \omega = [mR + nD]\Omega$), and it is evident that these points of the nonorthogonal lattice form a subset of the points ($\tau = mT, \omega = k\Omega$) of the orthogonal lattice. To be more specific: the nonorthogonal lattice is formed by those points of the rectangular (orthogonal) lattice for which $k - mR$ is an integer multiple of D . Note that the original rectangular lattice arises for $R = 0$ and $D = 1$ (see Fig. 6.3.1(a)), and that a hexagonal lattice occurs for $R = 1$ and $D = 2$ (Fig. 6.3.1(b)).

6.3.6 GABOR'S SIGNAL EXPANSION ON A NONORTHOGONAL LATTICE

If we define the 2D array λ_{mk} as

$$\lambda_{mk} = \sum_{n=-\infty}^{\infty} \delta_{k-mR-nD}, \quad (6.3.17)$$

Gabor's signal expansion on a nonorthogonal lattice can be expressed as (cf. Eq. 6.3.1)

$$\varphi(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \lambda_{mk} a_{mk} g_{mk}(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a'_{mk} g_{mk}(t), \quad (6.3.18)$$

while—with a different analysis window $w(t)$, though—the expansion coefficients a_{mk} are still determined by the Gabor transform (Eq. 6.3.2). Of course, since we only need the limited array $a'_{mk} = \lambda_{mk} a_{mk}$ —which is, in fact, a properly sampled version of the full array a_{mk} —we need only calculate the coefficients a_{mk} for those values of m and k for which $k - mR$ is an integer multiple of D . We note that the Fourier transform $\bar{a}'(x, y)$ of the limited array a'_{mk} is related to the Fourier transform $\bar{a}(x, y)$ of the full array a_{mk} via the periodization relation

$$\bar{a}'(x, y) = \frac{1}{D} \sum_{n=0}^{D-1} \bar{a}\left(x - \frac{n}{D}, y - \frac{nR}{D}\right) \quad (6.3.19)$$

and thus

$$a'_r(x, y) = \frac{1}{D} \sum_{n=0}^{D-1} a_r\left(x - \frac{n}{D}, y - \frac{nR}{D}\right). \quad (6.3.20)$$

In the nonorthogonal case, the biorthogonality condition takes the form (cf. Eq. 6.3.3)

$$\sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \lambda_{mk} g_{mk}(t_1) w_{mk}^*(t_2) = \delta(t_1 - t_2) \quad (6.3.21)$$

and leads to the equivalent but simpler expression (cf. Eq. 6.3.4)

$$\frac{2\pi}{D\Omega} \sum_{m=-\infty}^{\infty} g(t - mT) w^* \left(t - \left[mT + n \frac{2\pi}{D\Omega} \right] \right) \exp \left(j2\pi m \frac{nR}{D} \right) = \delta_n. \quad (6.3.22)$$

Note that for $R = 0$ and $D = 1$, for which we have a rectangular lattice (see Fig. 6.3.1(a)), Eq. (6.3.22) reduces to Eq. (6.3.4), and that for $R = 1$ and $D = 2$, for which we have a hexagonal lattice (see Fig. 6.3.1(b)), Eq. (6.3.22) takes the form

$$\frac{\pi}{\Omega} \sum_{m=-\infty}^{\infty} g(t - mT) w^* \left(t - \left[mT + n \frac{\pi}{\Omega} \right] \right) (-1)^{mn} = \delta_n. \quad (6.3.23)$$

The biorthogonality condition expressed in terms of the Zak transforms of the window functions now takes the form (cf. Eq. 6.3.13)

$$\frac{T}{Dq} \sum_{r=0}^{p-1} g_{s_1,r}(x,y) w_{s_2,r}^* \left(x - \frac{n}{D}, y - \frac{nR}{D} \right) = \delta_n \delta_{s_1-s_2}, \quad (6.3.24)$$

with $s_1, s_2 = 0, 1, \dots, q-1$ and $n = 0, 1, \dots, D-1$, and allows an easy determination of the analysis window $w(t)$ for a given synthesis window $g(t)$. For $R = 0$ and $D = 1$, for instance, relation Eq. (6.3.24) reduces to Eq. (6.3.13), while for $R = 1, D = 2, q = 1$, and p an even integer—which corresponds to the integer ($p/2$ -times) oversampled hexagonal case—it reduces to

$$\frac{T}{2} \sum_{r=0}^{p-1} g_{0,r}(x,y) w_{0,r-np/2}^*(x,y) (-1)^{nr} = \delta_n \quad (n = 0, 1; p \text{ even}), \quad (6.3.25)$$

from which the Zak transform $\tilde{w}(t, \omega; T)$ and hence the window function $w(t)$ can easily be determined.

Since we have related Gabor's signal expansion on a nonorthogonal lattice to sampling on a denser but orthogonal lattice, followed by restriction to a sublattice that corresponds to the nonorthogonal lattice, we can still use all the techniques that are developed for rectangular lattices, in particular the technique of determining Gabor's expansion coefficients via the Zak transform (cf. Eq. 6.3.12).

6.3.7 SUMMARY AND CONCLUSIONS

Gabor's signal expansion and the Gabor transform on a rectangular lattice have been introduced, along with the Fourier transform of the array of expansion coefficients and the Zak transforms of the signal and the window functions. Based on these Fourier and Zak transforms, the sum-of-products forms for the Gabor expansion and the Gabor transform, which hold in the rationally oversampled case, have been derived.

We have then studied Gabor's signal expansion and the Gabor transform based on a nonorthogonal sampling geometry. We have done this by considering the nonorthogonal lattice as a sublattice of an

orthogonal lattice. This procedure allows us to use all the formulas that hold for the orthogonal sampling geometry. In particular, we can use the sum-of-products forms that hold in the case of a rationally oversampled rectangular lattice.

We finally note that if everything remains to be based on a rectangular sampling geometry, it will be easier to extend the theory of the Gabor scheme to higher-dimensional signals; see, for instance, Ref. [27], where the multidimensional case is treated for continuous-time as well as discrete-time signals.

6.4 SPECTROGRAM DECOMPOSITIONS OF TIME-FREQUENCY DISTRIBUTIONS⁰

Cohen's class of bilinear time-frequency distributions (i.e., the quadratic class; see Chapter 3, p. 113) has yielded a proliferation of methods that are becoming increasingly popular in applications; but their progress has been slowed, at least in part, by the computational burdens which are generally much greater than for spectrograms or wavelet-based methods. This section demonstrates methods for greatly reducing the computational burden, by decomposing the TFDs into linear weighted sums of spectrograms. The computational efficiency comes from economizing the number of STFTs that must be computed to form these spectrograms. Surprisingly, one can reduce the number of STFTs to a small number by using Haar windows for the STFT computation [28,29]. Finally, for certain kernels, special windows may be designed which can represent a TFD very well by using only three or four windows [30,31].

It should be noted that several attempts have been made to decompose *time-frequency distributions* themselves, in order to isolate specific TFD components or to synthesize TFDs using several components. While this is an interesting and valuable topic, it is beyond the scope of this section, which deals specifically with decomposition of the *kernel* of the TFD.

6.4.1 DECOMPOSITION-BASED APPROACHES

Cunningham and Williams have shown that a TFD of the quadratic class can be expressed as a weighted sum of spectrograms, where the spectrogram windows are an orthonormal set [32,33], and have given a rigorous base for further work. They were motivated by some ideas from White [34] and Amin [35], who has continued his work in this area. The windows result from an eigensystem decomposition of the kernel. This approach has been used to approximate various TFDs [33]. A twist on this concept is to use windows which are shifted and scaled versions of each other [28,29]. In these studies, windows from the Haar basis set were used. It can be shown that any TFD can be represented by $n + 1$ Haar windows, where $n = \log_2(N - 1)$, and N is the number of sample points. Thus, a 257-point kernel can be represented by 9 Haar functions, one at each scale. It is important to note that the cross-spectrograms, that is, spectrograms formed by STFTs using different windows, are used as well as the auto-spectrograms to form a basis for the TFD.

The reduced interference distribution (RID) (Section 5.2) or is a well-defined set of quadratic TFDs that can be constructed using a simple set of constraints, while retaining a number of desirable

⁰Authors: **W.J. Williams**, The University of Michigan, Ann Arbor, MI, USA (wjw@umich.edu) and **S. Aviyente**, Michigan State University, East Lansing, MI, USA (avyente@egr.msu.edu). Reviewers: B. Boashash and P. Flandrin.

mathematical properties. Faster computation of this class of TFDs was a particular motivation for the development of the spectrogram decomposition methods discussed in this section.

6.4.2 DECOMPOSITION OF TIME-FREQUENCY KERNELS

For a given time-frequency kernel, the first step in the spectrogram decomposition algorithm is to decompose the kernel in terms of spectrogram windows. The approach suggested by Cunningham and Williams used an eigensystem type approach to decompose the 2D kernel expressed in time-lag (n, m) as the outer product of a set of orthonormal basis functions. These basis functions become the windows of the spectrogram decomposition and the eigenvalues are the weightings of the spectrogram summations. Let $C(n, k)$ be a TFD with a real, bounded kernel of interest; then it can be expressed as

$$C(n, k) = \sum_{i=1}^M \lambda_i |\text{STFT}_i(n, k)|^2. \quad (6.4.1)$$

Denote the spectrogram windows as $w_i(n)$, and the signal as $x(n)$; then the STFTs can be represented as¹:

$$\text{STFT}_i(n, k) = \sum_{m=0}^{N-1} x(m - n) w_i(m) e^{-j2\pi mk/N}. \quad (6.4.2)$$

This is the Cunningham-Williams decomposition in a nutshell. The problem is that while the eigendecomposition gives an orthonormal set of windows, there is no orderly relationship between the windows as is the case for a principal component analysis or Karhunen-Loëve transform (KLT) for a set of signals. This is the “best basis” for the kernel in the same sense as is the KLT for representation of a set of signals. There is no redundancy or relationship between windows that allows more efficient computation. The principal eigenvector does the best job of representing the kernel in terms of an outer product with itself, the next eigenvector does the next best job when its contribution is compared to the remaining others, and so on.

Generally, the kernel can be adequately represented using a small number of the potential windows. However, a spectrogram must be computed for each window. In the present section, we suggest methods which use specially designed windows, potentially increasing the speed and efficiency of representation.

6.4.3 DEVELOPMENT OF THE METHOD

Time- and frequency-shift-covariant bilinear discrete TFRs are specified by a discrete kernel, and can be rewritten in the inner product form of

$$\begin{aligned} \text{TFR}_x(n, \omega; \psi) &= \sum_{n_1} \sum_{n_2} x(n + n_1) e^{-j\omega(n+n_1)} \psi\left(-\frac{n_1 + n_2}{2}, n_1 - n_2\right) [x(n + n_2) e^{-j\omega(n+n_2)}]^* \\ &= \langle \tilde{\psi} \bar{S}_{-n} \bar{M}_{-\omega} x, \bar{S}_{-n} \bar{M}_{-\omega} x \rangle, \end{aligned} \quad (6.4.3)$$

¹Other sections in this book denote the STFT of a signal $z(t)$ with window $w(t)$ as $F_z^w(t, f)$.

where \bar{S}_{-n} and $\bar{M}_{-\omega}$ are, respectively, the time and frequency shift operators on ℓ_2 , the space of finite-energy discrete-time signals, and $\tilde{\psi}$ is a bounded linear operator on ℓ_2 [32,33].²

The spectral representation of $\tilde{\psi}$ may be used to express the TFR as a weighted sum of spectrograms or “projectograms”. If the kernel is associated with a bounded, self-adjoint linear operator, then the kernel may be decomposed by an eigendecomposition such that one can represent the TFR as being composed of a finite series of spectrograms. The orthonormal windows forming the spectrograms are the eigenfunctions of the decomposition. The eigenvalues of the decomposition provide the weights for summing the set of spectrograms. The viewpoint may be taken that the projections of the signal on the eigenvectors of the kernel decomposition are then time and frequency shifted by the time and frequency shift operators, yielding, essentially, the STFT. The magnitude squared STFT is the “projectogram” or spectrogram associated with that particular window. A total of N windows are required to completely represent an $N \times N$ kernel.

The spectrogram decomposition can more generally be expressed as

$$\text{TFR}_x(n, \omega; \psi) = \sum_{n_1} x(n + n_1) e^{-j\omega(n+n_1)} \sum_{n_1} x^*(n + n_2) e^{j\omega(n+n_2)} \sum_{l=1}^N \sum_{k=1}^N \lambda_{k,l} w_k(n_1) w_l(n_2). \quad (6.4.4)$$

Here, w_k and w_l are the windows and $\lambda_{k,l}$ is the coefficient for each k, l . Then, $\sum_{n_1} w_k(n_1) x(n + n_1) e^{-j\omega(n+n_1)}$ is recognized as $\text{STFT}_k(n, \omega)$ and the kernel is realized by the outer product $\sum_{l=1}^N \sum_{k=1}^N \lambda_{k,l} w_k(n_1) w_l(n_2)$.

By realizing that the spectrogram (cross and auto) is

$$\text{SP}_{k,l}(n, \omega) = \text{STFT}_k(n, \omega) \text{STFT}_l^*(n, \omega) \quad (6.4.5)$$

one may write Eq. (6.4.4) as

$$\text{TFR}_x(n, \omega; \psi) = \sum_{l=1}^N \sum_{k=1}^N \lambda_{k,l} \text{SP}_{k,l}(n, \omega). \quad (6.4.6)$$

6.4.4 WIGNER EXAMPLE

The kernel matrix for an 8×8 discrete Wigner matrix is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (6.4.7)$$

²In this section, the standard discrete frequency notation ks replaced by ω where $\omega = 2\pi k$ to align with angle rotations notations.

Using the Haar basis vectors

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \quad (6.4.8)$$

one may obtain the projection of the kernel matrix on the outerproduct formed by all possible pairs of Haar basis vectors. The projection value is obtained by multiplying the kernel and the outerproduct matrices element by element and summing over all elements of the result. This provides the weighting $\lambda(k, l)$ for combining the spectrograms of Eq. (6.4.4). The array of weights is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}. \quad (6.4.9)$$

Generally, good TFR representation is possible using only a fraction of the windows required for full representation. The spectrogram itself, of course, requires only one term, since it has only one window in its decomposition. In general, N^2 cross-spectrograms are required, but for many kernels only a few weights are nonzero and other weights are small enough to set to zero. However, even better results may be obtained: by using a special set of orthogonal windows, one may represent the TFR using fewer cross-spectrograms.

6.4.5 OPTIMUM ORTHOGONAL WINDOWS

The kernel for the binomial TFD is shown here for illustration for $N = 9$:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & \frac{4}{16} & \frac{3}{8} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{6}{16} & \frac{3}{8} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{16} & \frac{1}{8} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (6.4.10)$$

The method used by Cunningham and Williams would require the decomposition

$$\mathbf{V}\lambda = \mathbf{BV}, \quad (6.4.11)$$

where λ is a diagonal matrix of eigenvalues. The corresponding rows and columns of \mathbf{V} form outer products which, weighted by the appropriate eigenvalues and summed, form the kernel

$$\mathbf{B} = \mathbf{V}\lambda\mathbf{V}' \quad (6.4.12)$$

One would wish the eigenvalues to decline quickly, indicating that a truncated outerproduct reconstruction could suffice. Generally, it has been found that about 17 outerproducts are sufficient to well represent a 256×256 kernel matrix [32].

More efficiency can be gained taking advantages of symmetry. The upper-right and lower-left portions of \mathbf{B} exhibit symmetry, so all of the kernel information is carried by the submatrix

$$\mathbf{C} = \begin{bmatrix} \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{4}{16} & 0 & 0 \\ 0 & \frac{4}{16} & \frac{1}{8} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{3}{8} \\ 0 & 0 & 0 & \frac{3}{8} & \frac{6}{16} \\ 0 & 0 & 0 & \frac{6}{16} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}. \quad (6.4.13)$$

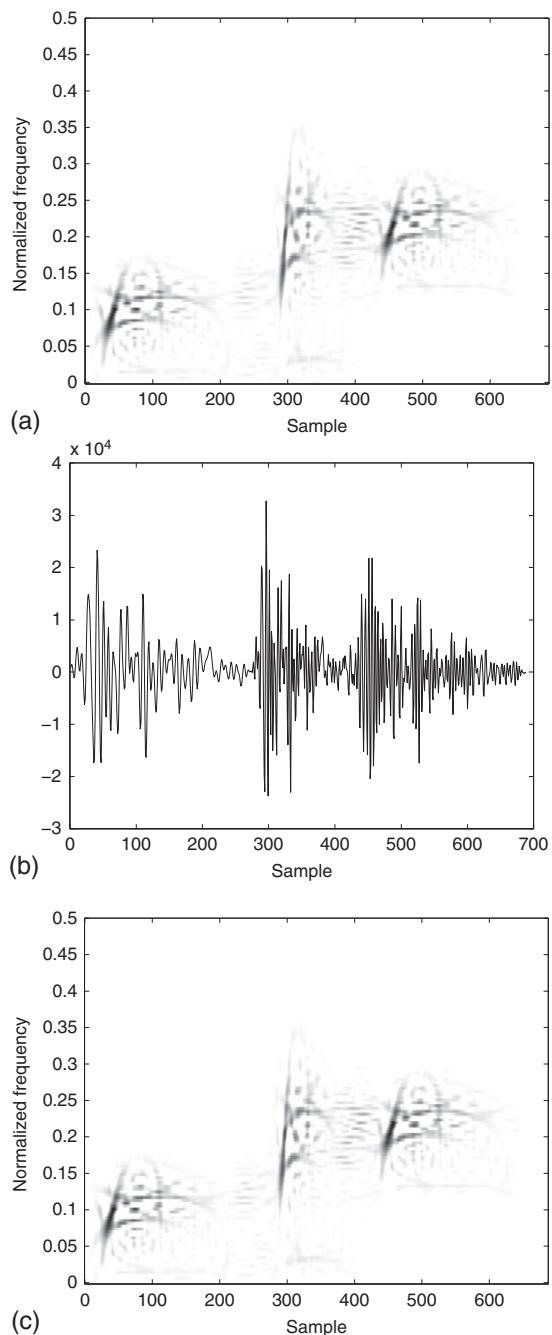
Much greater efficiency can be gained by rotating matrix \mathbf{C} to form

$$\mathbf{D} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ \frac{1}{16} & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (6.4.14)$$

where the 1 in the center of the kernel matrix has been replaced by 0 and a single companion matrix with the missing 1 (and 0 elements otherwise) is created. These matrices can be augmented with additional zeros in their rows and columns to a size equal to the original kernel. Then, a third matrix can be formed from \mathbf{C} by flipping \mathbf{C} along the antidiagonal. These three zero-augmented matrices can be summed to form the original kernel. The eigenvectors of \mathbf{D} can now be used as windows once the matrix is rotated back to its original position. However, the windows resulting from the eigenvectors now come in pairs such that the proper outerproducts are formed using a window and its time-reversed pair. Due to symmetry, only the upper right quadrant of \mathbf{B} as represented by \mathbf{C} , plus the impulse window which provides the 1 in the center, is required to compute the STFTs. Details about the construction of TFDS from the STFTs are given elsewhere [28–30].

6.4.6 KERNEL DECOMPOSITION RESULTS

Computation of TFDs by this method is efficient and fast. An example (Fig. 6.4.1) is provided for a 129×129 binomial kernel matrix. Only four distinct windows plus the impulse window are required to provide a very good representation of a complex signal. The kernel reconstructed from the truncated outer product series has only about 7% residual error. The signal consists of a dolphin click followed by a time-shifted and scaled version of the click (2:1 compressed in time and normalized for energy)

**FIGURE 6.4.1**

Comparison of results for dolphin clicks: (a) true binomial TFD, (b) time series, and (c) approximated binomial TFD using a four-spectrogram approximation.

followed by a time-shifted and frequency-shifted version of the click. This signal has been used elsewhere to illustrate the time-shift, scale, and frequency shift covariance of RIDs. Nine STFTs must be computed for this approximation. One additional trivial computation is required for the impulse window. Clearly, the true and approximated TFD results are very similar. If one looks closely, it appears that the cross-term activity is decreased in the approximate result. This is consistent with previous observations that noise and cross-term activity are abated in the approximations. One could compute each of the STFTs required in parallel. The formation of the cross-spectrograms from the STFTs and the summation of those results would require additional hardware or software, but these operations involve only multiplication of STFT points, multiplication of the resulting cross-spectrograms by the required eigenvalue coefficients and finally, summation of the results. In addition, a number of signal processing “tricks” may be employed to speed up the computations due to symmetry, the realness of the TFD result, etc.

Some TFDs that fall into the RID class may be approximated with a smaller number of windows. The Born-Jordan TFD requires only three distinct windows for a good representation. In fact, one may compute a legitimate RID using an impulse and one other distinct window [30]. However, this minimum window TFD may not have desirable representation properties in some cases. The windows derived by these methods have an interesting self-scaling property which accounts for the scale covariance being retained [36].

6.4.7 SUMMARY AND CONCLUSIONS

The liability of quadratic-class TFDs in terms of computational burden can be overcome to a considerable degree by employing various means of spectrogram decomposition. In addition to fast computation, noise and cross-terms may be decreased in the approximated form. It is clear that such TFDs can now be computed very rapidly using parallel computation or dedicated hardware, bringing their use into a wider range of practical applications.

6.5 COMPUTATION OF DISCRETE QUADRATIC TFDs⁰

6.5.1 GENERAL COMPUTATIONAL PROCEDURE

Section 6.1 deals with definitions and properties of the discrete WD and other discrete quadratic TFDs. It shows that the general discrete quadratic TFD of a complex signal $z[n]$ is

$$\rho_z[n, k] = \sum_{|m| < \frac{M}{2}} \sum_{|p| < \frac{P}{2}} G[p, m] z[n - p + m] z^*[n - p - m] e^{-j2\pi km/M} \quad (6.5.1)$$

$$= \text{DFT}_{m \rightarrow k} \left\{ G[n, m] * (z[n + m] z^*[n - m]) \right\}; \quad m \in \langle M \rangle, \quad (6.5.2)$$

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **G.R. Putland**, Land Values Research Group, Prosper Australia, Melbourne, VIC, Australia (grpputland@gmail.com). Reviewers: S.L. Marple, A. Reilly, and J.M. O’Toole.

where the support dimensions of the kernel $G[n, m]$ do not exceed M samples in lag (m) and P samples in time (n), and $\langle M \rangle$ is any set of M consecutive integers (cf. [37, p. 444]). So the general procedure for computing such a TFD is:

1. Formation of the instantaneous autocorrelation function (IAF)

$$K_z[n, m] = z[n + m]z^*[n - m];$$
2. Discrete convolution in n (time) with the smoothing function $G[n, m]$;
3. Discrete Fourier transformation mapping m (lag) to k (frequency).

For the discrete WD or WVD, which has $G[n, m] = \delta[n]$, step 2 reduces to an identity operation and may be omitted. The windowed discrete WD or WVD has $G[n, m] = \delta[n]g[m]$ so that step 2 reduces to multiplication by $g[m]$. Some other quadratic TFDs, however, have special forms leading to computational procedures which are *not* degenerate cases of the above, and which may be simpler or faster.

This section addresses some of the practical issues in computing such quadratic TFDs of a real signal, examines various cases of the above procedure, and offers the spectrogram as an example of a special form leading to a simpler, faster algorithm.

6.5.2 COMPUTATION OF THE ANALYTIC SIGNAL

The usual definitions of quadratic TFDs, especially the WVD and the windowed WVD, use an analytic signal in order to avoid interference terms between positive and negative frequencies. For computational purposes, an analytic signal also avoids the need for $2\times$ oversampling prior to computation of the IAF (see Section 6.1.1 and Refs. [2,3,37]). So, given a real signal $s(t)$, we must first compute the analytic signal $z(t)$ associated with $s(t)$. The simplest method is the direct approach of filtering out the negative frequencies in the frequency domain. If a real signal $s[n]$ is given for $n = 0, 1, 2, \dots, N - 1$ and periodically extended with period N , where N is even (or is made even by zero-padding), the algorithm is:

1. Compute $S[k] = \text{DFT}\{s[n]\}$ for $k = 0, 1, \dots, N - 1$;
2. Compute

$$Z[k] = \begin{cases} S[k] & \text{for } k = 0, \frac{N}{2} \\ 2S[k] & \text{for } k = 1, 2, \dots, \frac{N}{2} - 1 \\ 0 & \text{otherwise;} \end{cases} \quad (6.5.3)$$

3. Compute $z[n] = \text{IDFT}\{Z[k]\}$, where $\text{IDFT}\{\dots\}$ denotes the inverse DFT.

The treatment of the Nyquist term ($k = N/2$) and the meaning of “analytic” for a discrete-time periodic signal are explained in Ref. [38]; these issues become significant if the signal’s spectrum is nonzero at the Nyquist frequency. Further implementation details, including computation of the analytic signal, are given in Ref. [4]. A time-domain algorithm for computing the analytic signal using FIR filters is described in Ref. [39]. Note that there are also different formulations of the analytic signal that result in more or less aliasing in QTFDs [40].

6.5.3 REAL-TIME COMPUTATION OF TFDs

The formula for the discrete quadratic TFD (Eq. 6.5.2) involves the expression $z[n + m]$ where m is allowed to be positive, together with $z^*[n - m]$ where m is allowed to be negative. The same applies to the DWVD

$$W_z[n, k] = \text{DFT}_{m \rightarrow k} \left\{ z[n + m]z^*[n - m] \right\}; \quad m \in \langle N \rangle \quad (6.5.4)$$

and the windowed DWVD

$$W_z^g[n, k] = \text{DFT}_{m \rightarrow k} \left\{ g[m]z[n + m]z^*[n - m] \right\}; \quad m \in \langle M \rangle \quad (6.5.5)$$

(both of these expressions are derived in Section 6.1). All these cases involve *time-advanced* signals; for any value of n , the definition of the TFD involves signal samples as late as $z[n + \lambda]$, where λ is some positive integer. In real-time computation, we cannot compute the TFD for time n until we know the signal up to time $n + \lambda$; so λ is the *latency* (ignoring further delays due to central processing unit [CPU] time, etc.). In the case of the DWVD (Eq. 6.5.4), the range of m for which the IAF can be nonzero is maximized when n is at the center of the time-support of the signal; so the latency reaches a peak of half the signal duration. For the windowed DWVD, the peak latency is half the window duration. For the general discrete quadratic TFD, the peak latency is half the sum of the lag and time dimensions of the kernel. The latency of the analytic signal computation must be added to that of the TFD computation. In all cases, a smaller value of M not only reduces latency but also produces shorter FFTs, hence shorter computation times; but the cost is reduced frequency resolution.

Latency is one of two basic measures of speed for real-time computation of TFDs. The other measure is *throughput*, which depends on the efficiency of the numerical algorithms. Equation (6.5.2) can be written

$$\rho_z[n, k] = \text{DFT}_{m \rightarrow k} \left\{ R_z[n, m] \right\}, \quad (6.5.6)$$

where $R_z[n, m] = G[n, m] * (z[n + m]z^*[n - m])$. Similarly,

$$\rho_z[n + 1, k] = \text{DFT}_{m \rightarrow k} \left\{ R_z[n + 1, m] \right\}. \quad (6.5.7)$$

The above two equations represent successive time slices of the TFD. Multiplying the second equation by “ j ,” adding the result to the first equation and using the linearity of the DFT, we obtain

$$\rho_z[n, k] + j\rho_z[n + 1, k] = \text{DFT}_{m \rightarrow k} \left\{ R_z[n, m] + jR_z[n + 1, m] \right\}. \quad (6.5.8)$$

If, as usual, the TFD is known to be *real*, Eq. (6.5.8) means that the successive time slices of the TFD are, respectively, the real and imaginary parts of the right-hand side, which involves only one FFT [3]. So realness can reduce computation time by halving the number of FFTs from the m domain to the k domain. It also implies that $R_z[n, m]$ has Hermitian symmetry, which can halve the number of FFTs needed for the convolution in n . These opportunities are exploited in Section 6.6.

6.5.4 COMPUTATIONAL APPROXIMATIONS FOR DISCRETE-TIME KERNELS

Table 6.5.1: Computational Forms of Time-Lag Kernels of Selected Discrete Quadratic TFDs

| Distribution | $G[n, m]$ | Approx. |
|-------------------|---|---|
| WVD | $\delta[n]$ | |
| Levin | $\frac{1}{2}\delta[n+m] + \frac{1}{2}\delta[n-m]$ | |
| BJ | $\left[\frac{1}{ 4\alpha m } \text{rect}\left(\frac{n}{4\alpha m}\right) \right] * * [\text{sinc } n \text{ sinc } m]$ | $\approx \begin{cases} \frac{1}{ 4\alpha m +1} & \text{if } 2n \leq 4\alpha m + 1 \\ 0 & \text{otherwise} \end{cases}$ |
| $\text{BJ}_{1/2}$ | $\left[\frac{1}{ 2m } \text{rect}\left(\frac{n}{2m}\right) \right] * * [\text{sinc } n \text{ sinc } m]$ | $\approx \begin{cases} \frac{1}{ 2m +1} & \text{if } n \leq m \\ 0 & \text{otherwise} \end{cases}$ |
| Modified B | $\frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n}$ | |
| w-WVD | $\delta[n]w[m]$ | |
| w-Levin | $\frac{1}{2}w[m](\delta[n+m] + \delta[n-m])$ | |
| ZAM | $[w[m]\text{rect}\left(\frac{an}{4m}\right)] * * [\text{sinc } n \text{ sinc } m]$ | $\approx \begin{cases} w[m] & \text{if } an \leq 2m \\ 0 & \text{otherwise} \end{cases}$ |
| ZAM_2 | $[w[m]\text{rect}\left(\frac{n}{2m}\right)] * * [\text{sinc } n \text{ sinc } m]$ | $\approx \begin{cases} w[m] & \text{if } n \leq m \\ 0 & \text{otherwise} \end{cases}$ |
| Rihaczek | $\delta[n-m]$ | |
| w-Rihaczek | $w^*[-m]\delta[n-m]$ | |
| Page | $\delta[n - m]$ | |
| CW (or ED) | $\frac{\sqrt{\pi\sigma}}{ 2m } \exp\left(\frac{-\pi^2\sigma n^2}{4m^2}\right) * * [\text{sinc } n \text{ sinc } m]$ | $\approx \begin{cases} \delta[n] & \text{if } m = 0 \\ \sqrt{\frac{\pi\sigma}{4m^2+\pi\sigma}} \exp\left(\frac{-\pi^2\sigma n^2}{4m^2+\pi\sigma}\right) & \text{otherwise} \end{cases}$ |
| B | $\left[\frac{ 2m }{\cosh^2 n} \right]^\beta * \text{sinc } m$ | $\approx \left[\frac{\sqrt{4m^2+1}}{\cosh^2 n} \right]^\beta$ |
| Spectrogram | $w[n+m]w[n-m]$ | |

Notes: In the “Distribution” column, subscripts indicate parameter values while the prefix “w-” means “windowed” by the function $w[m]$. For the spectrogram and w-Levin distributions, the window w is assumed to be real and even. The “ $G[n, m]$ ” column shows the exact kernels required for the avoidance of aliasing in the Doppler-frequency domain. If $G[n, m]$ cannot be computed as written, the “Approx.” column shows the suggested computational approximation; otherwise “Approx.” is left blank.

Table 6.5.1 reproduces the “ $G[n, m]$ ” column of **Table 6.1.2** (p. 341) and adds two special cases often found in the literature: $\text{BJ}_{1/2}$ denotes the Born-Jordan distribution with $\alpha = 1/2$, while ZAM_2 denotes the Zhao-Atlas-Marks distribution with $a = 2$. Some entries in the “ $G[n, m]$ ” column of **Table 6.5.1** call for continuous convolution prior to sampling. At best, the evaluation of such a convolution in the time-lag domain requires oversampling. At worst, it requires the numerical evaluation of an improper integral arising from a singularity in $G(t, \tau)$. In either case, computational inefficiencies will arise unless the smoothing effect of the convolution can be approximated in some other way. The chosen approximations, shown in the right-hand column of **Table 6.5.1**, are explained below.

In the case of the B-distribution, the purpose of the convolution is to avoid aliasing. Without the convolution, and for typical values of the parameter β (e.g., $\beta = 0.01$), the time-lag kernel would be a

continuous function with a narrow slot at $m = 0$ caused by the factor $|2m|^\beta$. This factor is approximately unity for small nonzero values of m . The convolution fills in the slot so that the factor is approximately unity at $m = 0$ also. This effect can be approximated by replacing $|2m|$ with $[4m^2 + 1]^{1/2}$, as is done in [Table 6.5.1](#).

In the case of ZAM distribution, the convolution is also meant to ensure that even if $w(\tau)$ has a singularity, $G[n, 0] = \delta[n]$. This in turn verifies the TM property. Without the convolution, however, we would have

$$G_{\text{ZAM}}[n, m] = w[m] \text{rect}\left(\frac{an}{4m}\right) = \begin{cases} w[m] & \text{if } |an| \leq |2m| \\ 0 & \text{otherwise.} \end{cases} \quad (6.5.9)$$

This gives $G[n, 0] = w[0]\delta[n]$, which verifies the TM property provided that $w[0] = 1$. Accordingly, Eq. (6.5.9) is used in [Table 6.5.1](#), although other approximations are possible. For example, we could sacrifice the TM property in favor of some smoothing by using the approximation

$$G_{\text{ZAM}}[n, m] \approx \frac{1}{2}w[m][1 + \tanh(|4m| - |2an|)], \quad (6.5.10)$$

and we could salvage the TM property by using a separate definition for $m = 0$.

For the Born-Jordan (BJ) and exponential distributions (EDs), the convolutions are needed to remove singularities at $m = 0$ and ensure that $G[n, 0] = \delta[n]$. For the BJ distribution, we can remove the singularity and approximate the spreading in the $[n, m]$ plane by replacing $|4\alpha m|$ with $|4\alpha m| + 1$. The result is

$$G_{\text{BJ}}[n, m] \approx \frac{1}{|4\alpha m| + 1} \text{rect}\left(\frac{n}{|4\alpha m| + 1}\right), \quad (6.5.11)$$

which is equivalent to the rule given in [Table 6.5.1](#). For the ED (or CW distribution), a similar effect is obtained by replacing $|2m|$ with $[4m^2 + \pi\sigma]^{1/2}$. This step, by itself, gives the kernel

$$G_{\text{CW}}[n, m] \approx \sqrt{\frac{\pi\sigma}{4m^2 + \pi\sigma}} \exp\left(\frac{-\pi^2\sigma n^2}{4m^2 + \pi\sigma}\right). \quad (6.5.12)$$

For $n = m = 0$, this reduces to $G[0, 0] = 1$, which is consistent with the requirement that $G[n, 0] = \delta[n]$. However, for $m = 0$, Eq. (6.5.12) reduces to $G[n, 0] = e^{-\pi n^2}$, which is only an approximation to $\delta[n]$. Accordingly, a two-part definition of the kernel is used in the “Approx.” column of the table. With $n = 0$, the kernel as defined in the table reduces to

$$G_{\text{CW}}[0, m] = \sqrt{\frac{\pi\sigma}{4m^2 + \pi\sigma}}, \quad (6.5.13)$$

which takes the value 1 at $m = 0$ and $1/\sqrt{2}$ at $m = \pm\sqrt{\pi\sigma/4}$. For realistic values of σ (e.g., $\sigma \geq 1$), this gives a reasonable degree of smoothing in the m direction.

An alternative approach to the singularity problem, as detailed in [Section 6.6](#), is to evaluate the kernel in the ambiguity domain, that is, the Doppler-lag or $[l, m]$ domain. This is efficient because the ambiguity function is merely *multiplied* by, not convolved with, the Doppler-lag form of the kernel (cf. [Eq. 3.2.11](#) on p. 113). Indeed, if we evaluate the time-convolution in the time-lag domain by the FFT method (as in the present section), we implicitly perform that multiplication in the Doppler-lag domain, in which case we might as well define the kernel in the same domain if we can. Because the

Doppler-lag formulation of the kernel tends to be low-pass, the resulting algorithm tends to be not only time-efficient but also memory-efficient (see [Section 6.6](#)).

6.5.5 SPECIAL CASE: DIRECT FORM OF THE DISCRETE SPECTROGRAM

With a change of notation, Eq. (6.1.40) becomes

$$S_x^w[n, k] = \left| \sum_{|m| < M/2} x[m + n] w[m] e^{-j2\pi km/M} \right|^2. \quad (6.5.14)$$

This $S_x^w[n, k]$ is the *discrete spectrogram* of the discrete-time signal $x[n]$ with window $w[m]$. If the summand is extended periodically in m with period M (i.e., in τ with period M/f_s), we obtain

$$S_x^w[n, k] = \left| \sum_{m \in \langle M \rangle} x[m + n] w[m] e^{-j2\pi km/M} \right|^2, \quad (6.5.15)$$

where $\langle M \rangle$ denotes any set of M consecutive integers. This may be written

$$S_x^w[n, k] = \left| \text{DFT}_{m \rightarrow k} \{x[m + n] w[m]\} \right|^2; \quad m \in \langle M \rangle. \quad (6.5.16)$$

The time support of $S_x^w[n, k]$ is that of $x[n] * w[n]$, corresponding to $x(t) * w(t)$. If this has a duration not exceeding N samples, the nonzero elements of the discrete spectrogram may be represented by an $N \times M$ real matrix. Only half of the M columns are needed for the non-negative frequencies, which suffice if $x(t)$ is real.

Equation (6.5.14) involves $x[n + m]$ where $|m| < M/2$ and M is the window length in samples. So, in real-time computations, the latency of the discrete spectrogram computed by this formula is half the window length.

6.5.6 SAMPLE CODE FRAGMENTS

In view of the current popularity of MATLAB™ and GNU Octave, we illustrate this section with some code fragments from the experimental MATLAB/Octave function `t1kern.m`, which computed all of the TFDs plotted in [Section 5.7](#). As the code is experimental, efficiency is sometimes sacrificed for generality and clarity. The input parameters of the function specify the kernel in terms of a time-dependent factor $g_1[n]$, a lag-dependent factor $g_2[m]$, and an “auxiliary factor” $g_3[n, m]$. The overall time-lag kernel $G[n, m]$ is then computed as

$$G[n, m] = g_2[m] \left(\sum_n g_1[n] * g_3[n, m] \right) = (g_2[m] g_1[n]) * g_3[n, m]. \quad (6.5.17)$$

This scheme allows the computation of a wide variety of quadratic TFDs in under 320 lines of code.

6.5.6.1 Example 6.2.1: MBD (general algorithm)

For a separable kernel, the auxiliary factor would normally be omitted (i.e., taken as $\delta[n]$), while the time-dependent and lag-dependent factors would have input parameters specifying their types (e.g.,

Hamming or Hanning) and durations (in samples). As per [Section 5.7](#) and [Table 6.5.1](#), the MBD has a lag-independent kernel that is specified using the factors

$$g_1[n] = \delta[n]; \quad g_2[m] = 1; \quad g_3[n, m] = \frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n}. \quad (6.5.18)$$

Notice that the auxiliary factor is the complete kernel.

To compute the MBD in [Fig. 5.7.2\(e\)](#) on p. 281, the function `t1kern` is called with the following significant parameters:

```
s = signal vector
N = 128 = assumed period
tr = 1 = time resolution
tf = 'delta' = string specifying g1[n]
lf = '1' = string specifying g2[m]
af = 'mb' = string specifying form of g3[n, m]
ap = 0.2 = auxiliary parameter ( $\beta$ ).
```

All internal computations, including IAF generation, are designed to be valid for periodic signals. Therefore, to compute the IAF of a *nonperiodic* signal such as the one in [Fig. 5.7.2\(e\)](#), the assumed period N must be at least twice the signal length to avoid wrap-around effects (aliasing). Because the time support of the IAF is identical to that of the signal, the same value of N is also sufficient to avoid wrap-around in the subsequent convolution with $g_1[n]$.

The output is the real matrix `tfd(1:Mpad+1, 1:Nsel)`, whose dimensions `Mpad+1` and `Nsel` are assigned early by the statements

```
Mpad = 2^ceil(log(2*M)/log(2)); % lag-to-frequency FFT length
Ncut = min(N,length(s)); % duration of TF plot
Nsel = ceil(Ncut/tr); % no. traces in TF plot
```

where M is the support length of the kernel in the lag direction; in this case M has been set to `length(s)` because of the constant “lag-dependent” factor.

6.5.6.1.1 Preliminaries

The analytic signal is computed by the frequency-domain method. If N is even, the Nyquist term has Matlab/Octave index $N/2+1$ and the amplitude at that frequency is left unchanged [38]. If N is odd, there is no Nyquist term. The following code handles both cases:

```
Noff = fix(N/2);
z = fft(real(s),N); % s truncated or padded
z(2:N-Noff) = 2*z(2:N-Noff); % positive frequencies
z(Noff+2:N) = 0; % negative frequencies
z = ifft(z);
```

For this kernel, the time-dependent factor $g1$ and the lag-dependent factor $g2$ are computed by the statements

```

g1(1:N) = 0;
...
g1(1) = 1;
...
g2(1:Mpad) = 1;

```

where “...” denotes one or more line(s) of control code, or code that is skipped in this case. The auxiliary factor g_3 (the whole kernel in this case) is computed by

```

Moff = fix(M/2);
...
g3(1:N,1:Mpad) = 0;
...
temp(1:N) = 0;
for n = -Noff:Noff
    temp(1+rem(N+n,N)) = (cosh(n))^(−2*ap);
end
temp = temp/sum(temp); % normalize
for m = -Moff:Moff
    g3(:,1+rem(Mpad+m,Mpad)) = temp.';
end

```

where ap denotes the auxiliary parameter (β), and the remainder (`rem`) function causes high array indices to represent negative values of time and lag.

6.5.6.1.2 Step 1: Formation of the IAF

The IAF matrix $K(1:N, 1:Mpad)$ is formed by

```

for n = 1:N
    for m = -Moff:Moff
        % K(n,m) = z(n+m)z^*(n-m), with corrected indices:
        K(n,1+rem(Mpad+m,Mpad)) = z(1+rem(2*N+n+m-1,N))*conj(z(1+rem(2*N+n-m-1,N)));
    end
end

```

where the “corrected” indices allow handling of periodic signals.

6.5.6.1.3 Step 2: Convolution in time

The assembly of the time-lag kernel and the convolution in time with the IAF are performed together. The smoothed IAF is

$$R_z[n, m] = K_z[n, m] * G[n, m] = K_z[n, m] * \left(\sum_n g_2[n] g_1[n] * g_3[n, m] \right). \quad (6.5.19)$$

The above convolutions may be taken as circular if the assumed period is sufficiently long, in which case

$$R_z[n, m] = \text{IDFT} \left\{ \text{DFT}_{n \rightarrow l} \left\{ K_z[n, m] \right\} \text{DFT}_{n \rightarrow l} \left\{ g_3[n, m] \right\} \text{DFT}_{n \rightarrow l} \left\{ g_1[n] g_2[m] \right\} \right\}. \quad (6.5.20)$$

This is implemented by the following code, in which $K(:, mcorr)$ is initially the m th column of the IAF, but is overwritten by the m th column of the *smoothed* IAF:

```

for m = -Moff:Moff
    mcorr = 1+rem(Mpad+m,Mpad);
    ...
    K(:,mcorr) = ifft(fft(K(:,mcorr)).*fft(g3(:,mcorr)).*fft(g1.'*g2(mcorr)));
    ...
end

```

(The factor $g_2[m]$ could be taken outside the IDFT, but this would not improve the efficiency of the code because $g_2(mcorr)$ is a scalar.) The FFT method of convolution is useful in *experimental* code because of its generality, but is not necessarily the most efficient method, especially if one of the convolved sequences is short.

Now we apply the time-resolution (tr):

```

for nsel = 1:Nsel
    % nselth column of r is selected row of K:
    n = 1+tr*(nsel-1);
    r(:,nsel) = K(n,:).';
end

```

6.5.6.1.4 Step 3: DFT

The final DFT (lag to frequency) is computed by $r = \text{fft}(r)$, which, for the sake of maximum generality, does *not* exploit realness of the TFD.

6.5.6.1.5 Final adjustments

The following code scales the TFD and repeats its first row (the zero-frequency row) so that the TFD spans a full cycle in the frequency domain:

```
tfd = [real(r);real(r(1,:))].*(Ncut/Nsel/Mpad);
```

The scaling ensures that the sum of the matrix elements is close to the signal energy regardless of the time resolution.

6.5.6.2 Example 6.2.2: Spectrogram (special case)

The spectrogram in Fig. 5.7.3(f) on p. 282 was computed by the same function `tlkern`. For the spectrogram, the parameters s , N , and tr are the same as for the MBD, while tf is ignored. Other significant parameters are

```

lf = 'rect' = string specifying type of window;
M = 17 = window length (in samples); and
af = 'sg' = string calling for spectrogram.

```

The output is $\text{tfd}(1:\text{Mpad}/2+1, 1:\text{Nsel})$, where the dimensions are assigned as for the MBD, except that the window duration M is read as an input parameter and *not* overwritten.

The analytic signal is computed as for the MBD, although this is not strictly necessary for the spectrogram. The rectangular window is computed by

```
Moff = fix(M/2);
g2(1:Mpad) = 0;
...
for m = -Moff:Moff
    g2(1+rem(Mpad+m,Mpad)) = 1;
end
```

The matrix $K(1:N, 1:Mpad)$ normally represents the IAF, but for the spectrogram it is assigned differently:

```
for n = 1:N
    for m = -Moff:Moff
        % K(n,m) = z(n+m)g2(m), with corrected indices:
        K(n,1+rem(Mpad+m,Mpad)) = z(1+rem(2*N+n+m-1,N))*g2(1+rem(Mpad+m,Mpad));
    end
end
```

The code that applies the time resolution and performs the final DFT (lag to frequency) is the same as for the MBD. But the final adjustment is different:

```
tfd = (abs(r(1:Mpad/2+1,:))).^2.*((Ncut/Nsel/Mpad/sum(g2.^2)));
```

The magnitude-squared operation alters the relationship between the window and the scaling of the TFD. Also note that the above step uses only half the columns of the Fourier-transformed r matrix, namely those corresponding to the non-negative frequencies. Efficiency could be further improved by exploiting the analytic signal to halve the sampling rate [3,4].

6.5.7 THE TFSAP TOOLBOX

The *Time-Frequency Signal Analysis and Processing (TFSAP)* toolbox (formerly called *TFSA*) is a set of algorithms developed over more than two decades at the CRISSP research unit of the University of Queensland (1984-1991), then at the Signal Processing Research Centre at Queensland University of Technology. The version current at the time of publication is distributed as a MATLAB/Octave toolbox. Note that TFSAP functions can be used with other computational and graphical functions of MATLAB or Octave. Information on downloading the package is given in [Section 6.1](#) and [Chapter 17](#).

6.5.8 SUMMARY AND CONCLUSIONS

High-level programming languages with built-in FFT functions and matrix operations have made it possible to construct compact yet highly versatile functions for computing quadratic TFDs. Use of a common algorithm for all TFDs is convenient for the programmer. But, as illustrated by the direct form of the spectrogram, efficiency can sometimes be improved by using different algorithms in special cases. In addition, there is the option of using optical implementations [41].

6.6 MEMORY-EFFICIENT ALGORITHMS FOR QUADRATIC TFDs⁰

6.6.1 ALGORITHMS TO REDUCE COMPUTATION AND MEMORY

The era of ‘big data’ presents significant challenges for signal processing. Large datasets are now the norm. In intensive care units, for example, continuous neurological monitoring of patients with electroencephalography (EEG) generates multiple long-duration signals. Processing these immense datasets requires computationally efficient implementations of the signal processing methods.

Time-frequency distributions (TFDs) are 2D functions and hence require significant computational resources for large datasets. For example, the fast Fourier transform (FFT), for a discrete signal of length N , requires $N \log_2 N$ numerical operations and N sample points of computer memory; quadratic TFDs, in contrast, require $N^2 \log_2 N$ numerical operations and N^2 sample points of computer memory. An increase in numerical operations for an algorithm will increase computational time. But an increase in computer memory can be more problematic—if the memory is not available, the algorithm will fail to compute.

The TFD algorithms presented here focus primarily on reducing memory, and also on reducing computational load [42]. Many TFDs are over-sampled because often the Doppler-lag kernel is concentrated around the origin with a significant region of zeros away from the origin. By computing only the nonzero region, as Fig. 6.6.1 illustrates, we can reduce over-sampling, and hence reduce memory consumption and CPU time without diminishing time-frequency resolution. Also presented are algorithms that compute under-sampled (decimated) TFDs and require even less memory. Time-frequency resolution is not maintained, as the algorithms compute only a subset of the full TFD. These algorithms are useful when analyzing large datasets or when the hardware platform has limited memory.

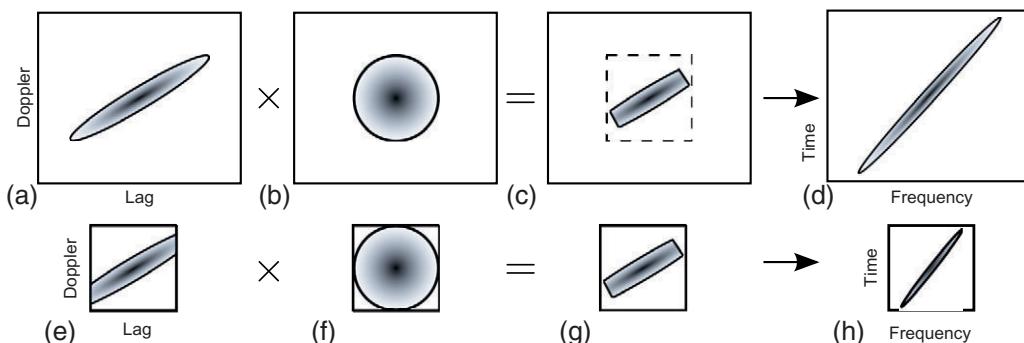


FIGURE 6.6.1

Reducing over-sampling in the TFD. The ambiguity function in (a) is multiplied by the kernel in (b) to produce the filtered ambiguity function in (c), which is then transformed to the time-frequency domain in (d). As the area outside the circle in (b) is zero, so is the area outside the dotted box in (c), resulting in the over-sampled TFD in (d). Computing only the area inside the dotted box, as shown in (e)-(g), reduces computation and memory without loss of time-frequency resolution.

⁰Authors: **J.M. O’Toole**, University College Cork, Corcaigh, Ireland (j.otoole@ieee.org) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: G.R. Putland, E. Sejdic, and S. Aviyente.

6.6.2 TFD ALGORITHMS FOR SPECIFIC KERNEL TYPES

As discussed in [Section 6.5.2](#), the first step in all TFD algorithms is to generate an analytic associate of the real-valued signal $s[n]$. To minimize aliasing in the Wigner-Ville distribution (WVD) [40], the length- N signal $s[n]$ must be zero-padded to length- $2N$ as follows:

1. zero-pad $s[n]$ from length N to length $2N$;
2. zero negative frequencies of $s[n]$ to produce $z[n]$ using the approach described in [Chapter 1](#) and [Section 6.5](#); and
3. let $z[n] = 0$ for $N \leq n \leq 2N - 1$.

The discrete TFD from [Sections 6.1](#) and [6.5](#) is

$$\rho_z[n, k] = W_z[n, k] * \gamma[n, k], \quad (6.6.1)$$

where \ast_k denotes circular convolution in discrete-time n and discrete-frequency k , and $\gamma[n, k]$ is the time-frequency kernel: $\gamma[n, k] = \text{IDFT}_{n \leftarrow l} \{\text{DFT}_{m \rightarrow k} \{g[l, m]\}\}$, where DFT means the discrete Fourier transform and IDFT the inverse DFT. The WVD $W_z[n, k]$ is defined in [Section 6.5.3](#) as the DFT of $K_z[n, m]$. This discrete TFD in Eq. (6.6.1) will not satisfy all mathematical properties of the continuous TFD [43], but is a popular discrete definition because of its simple formulation and easy application. Algorithms for alternative TFD definitions can be found elsewhere [42].

When computing TFDs, the number of DFT operations dominates computational cost. The goal, where possible, is to reduce the number and length of the FFTs (as FFT algorithms compute the DFT operations). This goal is achieved by developing specific algorithms for four kernel types: the nonseparable kernel $g[l, m]$, the separable kernel $G_1[l]g_2[m]$, the lag-independent kernel $G_1[l]$, and the Doppler-independent kernels $g_2[m]$. The algorithms are only applicable to TFDs with kernels that have analytical or numerical forms in the Doppler-lag domain. [Chapter 2](#) and [Sections 5.2](#) and [5.7](#) discuss the importance of these kernel types, and [Section 6.5](#) describes discrete versions of the kernels.

[Table 6.6.1](#) outlines the process of transformation from one domain to another for the four kernel types. Lag- and Doppler-independent kernel TFDs require only one transformation, and algorithms for these kernel types have a smaller computational load than separable and nonseparable kernel algorithms, which require three transformations each. Because we assume that the nonseparable kernel

Table 6.6.1: Computing the TFD for Each Kernel Type

| Kernel | Process to Compute the TFD $\rho_z[n, k]$ |
|---------|---|
| Nonsep. | $K_z[n, m] \xrightarrow{\text{DFT}_{n \rightarrow l}} A_z[l, m] \xrightarrow{\times g[l, m]} \mathcal{A}_z[l, m] \xrightarrow{\text{IDFT}_{n \leftarrow l}} R_z[n, m] \xrightarrow{\text{DFT}_{m \rightarrow k}} \rho_z[n, k]$ |
| Sep. | $K_z[n, m]g_2[m] \xrightarrow{\text{DFT}_{n \rightarrow l}} A_z[l, m] \xrightarrow{\times G_1[l]} \mathcal{A}_z[l, m] \xrightarrow{\text{IDFT}_{n \leftarrow l}} R_z[n, m] \xrightarrow{\text{DFT}_{m \rightarrow k}} \rho_z[n, k]$ |
| DI | $K_z[n, m]g_2[m] \xrightarrow{\text{DFT}_{m \rightarrow k}} \rho_z[n, k]$ |
| LI | $K_z[l, k]G_1[l] \xrightarrow{\text{IDFT}_{n \leftarrow l}} \rho_z[n, k]$ |

Notes: Labeled arrows indicate the steps. The process starts by generating $K_z[n, m]$ or $K_z[l, k]$ and then flows left-to-right, using the transformations described in [Section 3.2.1](#). Kernel types: nonseparable (Nonsep.) $g[l, m]$, separable (Sep.) $G_1[l]g_2[m]$, Doppler-independent (DI) $g_2[m]$, and lag-independent (LI) $G_1[l]$. Four domains: Doppler-frequency $[l, k]$, Doppler-lag $[l, m]$, time-lag $[n, m]$, and time-frequency $[n, k]$.

is nonzero throughout the entire Doppler-lag domain, memory savings are not possible for this kernel type. We present this algorithm here, however, to detail two basic rules that are incorporated into the subsequent algorithms.

6.6.2.1 Algorithm: TFD with nonseparable kernel

Input: Analytic signal $z[n]$ of length $2N$ (where N is the length of the real-valued signal $s[n]$); $N \times N$ kernel $g[l, m]$.

Output: $N \times N$ TFD $\rho[n, k]$.

1. Let $N_h = \lceil N/2 \rceil$; iterate over values $m = 0, 1, \dots, N_h - 1$:

- (a) Generate lag-slice: $K_m[n] = z[n + m]z^*[n - m]$ for $n = 0, 1, \dots, N - 1$.
- (b) Multiply by the kernel in the Doppler-lag domain:

$$R[n, m] = \text{IDFT}_{n \leftarrow l} \left\{ g[l, m] \text{DFT}_{n \rightarrow l} \left\{ K_m[n] \right\} \right\}. \quad (6.6.2)$$

2. Recover negative-lag values: $R[n, N - m] = R^*[n, m]$ for $1 \leq m \leq N_h - 1$, over all time samples $n = 0, 1, \dots, N - 1$.

3. Transform to the time-frequency domain; iterate over $n = 0, 2, 4, \dots, N - 2$:

$$\begin{aligned} \rho[n, k] &= \text{Re}\{B[n, k]\} \quad \text{and} \quad \rho[n + 1, k] = \text{Im}\{B[n, k]\} \\ \text{where } B[n, k] &= \text{DFT}_{m \rightarrow k} \left\{ R[n, m] + jR[n + 1, m] \right\}. \end{aligned}$$

Two simple rules halve the number of FFTs required for the domain transformations shown in [Table 6.6.1](#) [4]. First, because $R[n, m]$ is conjugate symmetrical in the lag direction m , Eq. (6.6.2) computes $R[n, m]$ for only one-half of the m values, thus reducing the number of FFTs by one-half [4]. And second, as described in [Section 6.5.3](#), step 3 combines two successive time slices into one FFT, thus reducing the number of FFTs again by one-half.

6.6.2.2 Algorithm: TFD with separable kernel

For kernel $G_1[l]g_2[m]$ of size $Q \times P$, this algorithm produces a TFD of size $N_{\text{time}} \times N_{\text{freq}}$, where the parameters $N_{\text{time}} \geq Q$ and $N_{\text{freq}} \geq P$ control the level of over-sampling in the time and frequency directions. Reducing oversampling reduces memory and computational load, as detailed in [Table 6.6.2](#). And because $N_{\text{time}} \geq Q$ and $N_{\text{freq}} \geq P$, time-frequency resolution remains unchanged (see [Fig. 6.6.1](#)).

Input: Analytic signal $z[n]$ of length $2N$; length- Q Doppler window $G_1[l]$ and length- P lag window $g_2[m]$ (Q and P are odd integers); parameters N_{time} and N_{freq} , with $N_{\text{time}} \geq Q$ and $N_{\text{freq}} \geq P$.

Output: $N_{\text{time}} \times N_{\text{freq}}$ TFD $\rho[n, k]$.

1. Let $P_h = \lceil P/2 \rceil$ and $Q_h = \lceil Q/2 \rceil$; iterate over values $m = 0, 1, \dots, P_h - 1$:

- (a) Generate lag-slice: $K_m[n] = g_2[m]z[n + m]z^*[n - m]$, for $0 \leq n \leq N - 1$.
- (b) Transform to the Doppler domain: $A_m[l] = \text{DFT}_{n \rightarrow l} \{K_m[n]\}$.

Table 6.6.2: Computational Costs of TFDs

| Kernel Type | Computation | Memory |
|---------------------|---|----------------------------------|
| Nonseparable | $\frac{3}{2}N^2 \log_2 N$ | N^2 |
| Lag-independent | $\frac{1}{2}NN_{\text{time}} \log_2 N_{\text{time}}$ | NN_{time} |
| Doppler-independent | $\frac{1}{2}NN_{\text{freq}} \log_2 N_{\text{freq}}$ | NN_{freq} |
| Separable | $P_h(N \log_2 N + N_{\text{time}} \log_2 N_{\text{time}}) + \frac{1}{2}N_{\text{time}}N_{\text{freq}} \log_2 N_{\text{freq}}$ | $N_{\text{time}}N_{\text{freq}}$ |

Notes: The number of FFTs determines the computational load (middle column), assuming that an FFT of a length- N complex-valued signal requires $N \log_2 N$ arithmetical operations. The total number of real-valued sample points needed to compute and store the TFD determines the memory load (right column). Parameters: N = signal length; P = lag-window length; $P_h = \lceil P/2 \rceil$; N_{time} controls over-sampling in the time direction, with $N_{\text{time}} \geq Q$ (where Q is the Doppler-window length); N_{freq} controls over-sampling in the frequency direction, with $N_{\text{freq}} \geq P$. As $N_{\text{freq}} \leq N$ and $N_{\text{time}} \leq N$, the computational and memory loads for three algorithms (separable, lag-, and Doppler-independent) are less than the loads for the nonseparable algorithm.

- (c) Multiply by the Doppler window and zero-pad if $N_{\text{time}} > Q$:

$$\begin{aligned} S_m[l] &= A_m[l]G_1[l], & 0 \leq l \leq Q_h - 1 \\ S_m[l] &= 0, & Q_h \leq l \leq N_{\text{time}} - Q_h \\ S_m[N_{\text{time}} - l] &= A_m[N - l]G_1[Q - l], & 1 \leq l \leq Q_h - 1. \end{aligned}$$

- (d) Transform back to the time-domain: $R[n, m] = \text{IDFT}_{n \leftarrow l}\{S_m[l]\}$.

2. Zero-pad and recover the negative-lag values:

$$\begin{aligned} R[n, m] &= 0, & P_h \leq m \leq N_{\text{freq}} - P_h \\ R[n, N_{\text{freq}} - m] &= R^*[n, m], & 1 \leq m \leq P_h - 1 \end{aligned}$$

over $n = 0, 1, \dots, N_{\text{time}} - 1$.

3. Transform to the time-frequency domain using step 3 in [Algorithm 6.6.2.1](#) (above), substituting N_{time} for N .

The Doppler-independent algorithm is a subset of the separable-kernel algorithm: all steps excluding the transformation to the Doppler-lag domain (steps 1(b)-(d) in the preceding algorithm). The lag-independent algorithm is similar to the Doppler-independent algorithm, except the instantaneous autocorrelation function (IAF) is formed in the Doppler-frequency domain as

$$\mathcal{K}[l, k] = Z[k + l]Z^*[k - l] + Z[k + l + N]Z^*[k - l - N],$$

where $Z[k] = \text{DFT}\{z[n]\}$. Because $Z[k]$ has nonzero energy at negative frequencies, the second term in the above expression is necessary to ensure that $W_z[n, k] = \text{DFT}_{m \rightarrow k}\{\mathcal{K}_z[n, m]\} = \text{IDFT}_{n \leftarrow l}\{\mathcal{K}[l, k]\}$ [40]. [Table 6.6.2](#) describes the computational costs for all four algorithms, showing that three algorithms have costs independent of the N^2 value.

6.6.3 FURTHER MEMORY REDUCTION: DECIMATED TFDs

A simple solution to reducing memory is to generate an under-sampled TFD by computing only a subset of TFD samples. This decimation must occur within the algorithm and not by simply discarding samples after generating the full TFD—otherwise the algorithm does not save memory. Computational load can also be reduced when decimation occurs within the algorithm.

The absence of time-frequency information implies that the decimated TFD only approximates the full TFD and therefore reduces the time-frequency resolution; this level of approximation depends on the signal type and extent of decimation. The decimated TFD differs, however, from a TFD generated from a decimated signal; using an under-sampled (below Nyquist frequency) signal for the TFD results in aliasing, whereas the decimated TFD will be alias-free. Although the decimated TFD is inexact, it does provide a simple memory-reducing solution that may suit applications with large datasets.

We present two approaches to decimation. The first approach decimates by computing only specific time slices (or frequency slices) of the TFD: generating, for example, the time-decimated TFD $\rho[n_i, k]$ for $\{n_i \mid 1 \leq i \leq U\}$, such that $0 \leq n_i \leq N - 1$ and $U \leq N$; thus $\rho[n_i, k]$ is size $U \times N$ and not $N \times N$. This is the most flexible approach, as $\{n_i\}$ can be any combination of the full set $\{0, 1, \dots, N - 1\}$, but is only applicable to lag- and Doppler-independent kernel TFDs.

The second approach also generates time slices (or frequency slices) of the TFD but with a constant interval between slices: generating, for example, the time-decimated TFD $\rho[an, k]$, where a is an integer; thus $\rho[an, k]$ is size $(N/a) \times N$. This approach applies the Fourier transform property that folding the time-domain signal $x[n]$ results in decimating the frequency domain:

$$X[ak] = \text{DFT}_{n \rightarrow k} \left\{ \sum_{p=0}^{a-1} x[pN/a + n] \right\}, \quad \text{for } 0 \leq n \leq (N/a) - 1, \quad (6.6.3)$$

where $x[n]$ is length N and $X[ak]$ is length N/a ; we refer to the summation term inside the DFT operation as the folded signal.

6.6.3.1 Algorithm: Decimated TFD with nonseparable kernel

The algorithm applies the property in Eq. (6.6.3) to fold the filtered ambiguity function $A_z[l, m]g[l, m]$ in both the Doppler and lag directions, producing the decimated TFD $\rho[an, bk]$, where a and b are integers. This TFD is size $(N/a) \times (N/b)$ and requires only $N^2/(ab)$ sample points of memory. For example if $N = 10,000$, then N^2 requires 763 MB of memory (assuming 1 point requires 8 bytes); but if $a = b = 8$, then $N^2/(ab)$ requires the more modest 12 MB.

Input: Analytic signal $z[n]$ of length $2N$; integer decimation factors a and b ; kernel type and kernel parameters.

Output: $(N/a) \times (N/b)$ TFD $\rho[n, k]$.

1. Let $N_{bh} = \lceil N/(2b) \rceil$; iterate over the values $m = 0, 1, \dots, N_{bh}$:

(a) Let $d = pN/b + m$ and fold the Doppler-lag function in the lag direction:

$$S_m[l] = \sum_{p=0}^{b-1} A_d[l]g[l, d] \quad 0 \leq l \leq N - 1, \text{ where}$$

$$A_d[l] = \begin{cases} \text{DFT}_{n \rightarrow l} \left\{ z[n+d]z^*[n-d] \right\} & \text{for } d \leq \lceil N/2 \rceil \\ \text{DFT}_{n \rightarrow l} \left\{ z^*[n+N-d]z[n-N+d] \right\} & \text{for } d > \lceil N/2 \rceil. \end{cases}$$

To avoid storing a large $N \times N$ kernel $g[l, m]$, generate lag slices of the kernel as required; the kernel $g[l, d]$, for each value of p , is size $1 \times N$.

- (b) Next, fold the filtered Doppler-lag function $S_m[l]$ in Doppler, that is:

$$S_m^{\text{fold}}[l] = \sum_{q=0}^{a-1} S_m[qN/a + l], \quad 0 \leq l \leq (N/a) - 1.$$

- (c) Then transform to the time-lag domain: $R[n, m] = \text{IDFT}_{n \leftarrow l} \{S_m^{\text{fold}}[l]\}$.
2. Expand $R[n, m]$ to include negative lag values: $R[n, N/b - m] = R^*[n, m]$ for $1 \leq m \leq N_{\text{bh}} - 1$ over $n = 0, 1, \dots, N/a - 1$.
 3. Transform $R[n, m]$ to the time-frequency domain as in step 3 of the preceding [Algorithm 6.6.2.1](#), substituting N/a for N .

The separable-kernel algorithm follows simply from this algorithm, decimating the $N_{\text{time}} \times N_{\text{freq}}$ TFD to compute the $(N_{\text{time}}/a) \times (N_{\text{freq}}/b)$ TFD, using a similar procedure to the preceding nonseparable algorithm.

6.6.3.2 Algorithm: Decimated TFD with Doppler-independent kernel

This algorithm, however, differs from the separable and nonseparable algorithms as the decimated TFD is of the form $\rho[n_i, bk]$, where $\{n_i\}$ is any combination of the set $\{0, 1, \dots, N - 1\}$. Thus, we can select any combination of time slices and manipulate the frequency sampling rate by the integer b .

Input: Analytic signal $z[n]$ of length $2N$; frequency-decimation factor b (integer); subset of time samples $\{n_i | 1 \leq i \leq U\}$; lag-window $g_2[m]$ of length P .

Output: $U \times (N_{\text{freq}}/b)$ TFD $\rho[n_i, bk]$.

1. Let $N_{\text{fh}} = \lceil N_{\text{freq}}/(2b) \rceil$ and $P_{\text{h}} = \lceil P/2 \rceil$; iterate over $i = 1, 2, \dots, U$:

- (a) Generate time slice: $R_i[m] = z[n_i + m]z^*[n_i - m]g_2[m]$ for $0 \leq m \leq P_{\text{h}} - 1$.
(b) Zero-pad and recover the negative-lag values:

$$\begin{aligned} R_i[m] &= 0, & P_{\text{h}} \leq m \leq N_{\text{freq}} - P_{\text{h}} \\ R_i[N_{\text{freq}} - m] &= R_i^*[m], & 1 \leq m \leq P_{\text{h}} - 1. \end{aligned}$$

- (c) Fold in the lag direction: $R_{\text{fold}}[n_i, m] = \sum_{p=0}^{b-1} R_i[pN_{\text{freq}}/b + m]$ for $0 \leq m \leq N_{\text{fh}}$. And again zero-pad and recover the negative-lag values:

$$\begin{aligned} R_{\text{fold}}[n_i, m] &= 0, & N_{\text{fh}} + 1 \leq m \leq N_{\text{freq}} - N_{\text{fh}} \\ R_{\text{fold}}[n_i, N_{\text{freq}}/b - m] &= R_{\text{fold}}^*[n_i, m], & 1 \leq m \leq N_{\text{fh}} - 1. \end{aligned}$$

Table 6.6.3: Computational Cost for Decimated TFDs, Using Metrics From Table 6.6.2

| Kernel Type | Computation | Memory | Grid |
|---------------------|--|---|-----------------|
| Nonseparable | $\frac{N^2}{2} \log_2 N + \frac{N^2}{2ab} \log_2 \frac{N}{b}$ | $\frac{N^2}{ab}$ | $\rho[an, bn]$ |
| Lag-independent | $\frac{VN_{\text{time}}}{2a} \log_2 \frac{N_{\text{time}}}{a}$ | $\frac{N_{\text{time}}V}{a}$ | $\rho[an, k_i]$ |
| Doppler-independent | $\frac{UN_{\text{freq}}}{2b} \log_2 \frac{N_{\text{freq}}}{b}$ | $\frac{UN_{\text{freq}}}{b}$ | $\rho[n_i, bk]$ |
| Separable | $\frac{NN_{\text{freq}}}{b} \log_2 N + \frac{N_{\text{time}}N_{\text{freq}}}{ab} \log_2 \frac{N_{\text{time}}N_{\text{freq}}}{ab}$ | $\frac{N_{\text{time}}N_{\text{freq}}}{ab}$ | $\rho[an, bn]$ |

Notes: Parameters: N , signal length; a , integer decimation factor in time direction; b , integer decimation factor in the frequency direction; $\{n_i \mid 1 \leq i \leq U\}$, subset of time samples, with $0 \leq n_i \leq N - 1$ and $U < N$; $\{k_i \mid 1 \leq i \leq V\}$, subset of frequency samples, with $0 \leq k_i \leq N - 1$ and $V < N$; N_{time} and N_{freq} , defined in Table 6.6.2. Parameters U, V, a, b control the degree of TFD decimation and hence influence the computation and memory loads for the algorithms.

2. Transform to the time-frequency domain; iterate over $i = 0, 2, 4, \dots, U - 2$:

$$\rho[n_i, k] = \operatorname{Re}\{B[n_i, k]\} \quad \text{and} \quad \rho[n_{i+1}, k] = \operatorname{Im}\{B[n_i, k]\}$$

where $B[n_i, k] = \operatorname{DFT}_{m \rightarrow k} \{R_{\text{fold}}[n_i, m] + jR_{\text{fold}}[n_{i+1}, m]\}$.

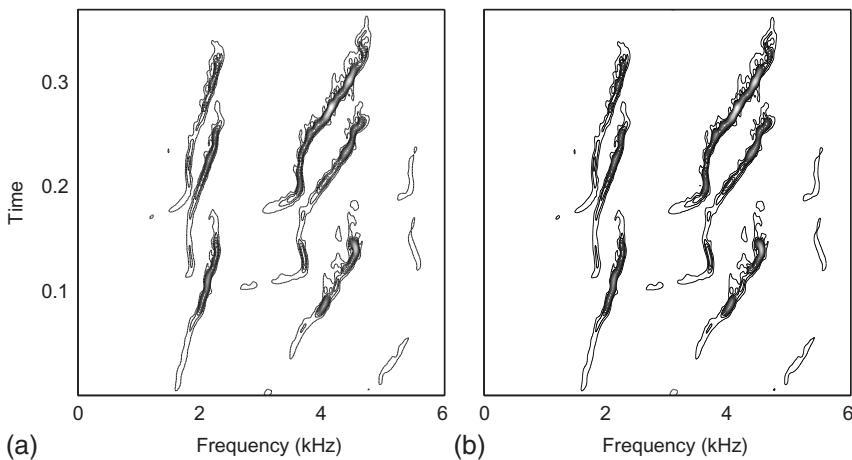
The algorithm folds in the lag direction, using the property from Eq. (6.6.3), to decimate in frequency for a specific set of time slices. The lag-independent algorithm is similar but folds in the Doppler direction to decimate in time for a specific set of frequency slices. Table 6.6.3 lists computational costs for the four algorithms.

6.6.4 EXAMPLE

The following example generates a TFD for a bird-song recording, shown on the book's cover of the first edition of this book and shown in Figure 6.6.2. The signal is sampled at 12,208 Hz and is 4500 samples long. Using Algorithm 6.6.2.2, we generate two separable kernel TFDs with $Q = 111$ (length of Doppler function) and $P = 221$ (length of lag function); both Doppler and lag functions use Hanning windows. For the first TFD, we set $N_{\text{time}} = 256$ and $N_{\text{freq}} = 512$ to compute the 256×512 TFD; this computes, on a 2013 desktop computer (2.8 GHz Intel Xeon) running MATLAB R2013a, in approximately 0.2 s with 1 MB of memory. For the second TFD, we set $N_{\text{time}} = N$ and $N_{\text{freq}} = N$, where $N = 4500$ is length of the signal to generate the 4500×4500 TFD; this computes in approximately 5 s with 155 MB of memory. Although having different sizes, both TFDs are oversampled (because $N_{\text{time}} > Q$ and $N_{\text{freq}} > P$) and therefore contain the same signal information, as Fig. 6.6.2 illustrates. This example illustrates how memory and computational load is reduced without reducing time-frequency resolution.

6.6.5 PRECISION OF DECIMATED TFDs: NUMERICAL EXAMPLES

The decimated TFD algorithms presented in Section 6.6.3 reduce (t, f) resolution by computing an approximation to the full TFD. The level of approximation, and thus (t, f) resolution, depends on the signal type. Oversampling, bandwidth-duration (BT) product, and sparsity will determine the extent of the loss of (t, f) resolution. Another important consideration is that in many applications, reducing (t, f) resolution is an acceptable trade-off for computation and memory efficiency. For example, reducing memory requirements can avoid the need for use of super computing facilities.

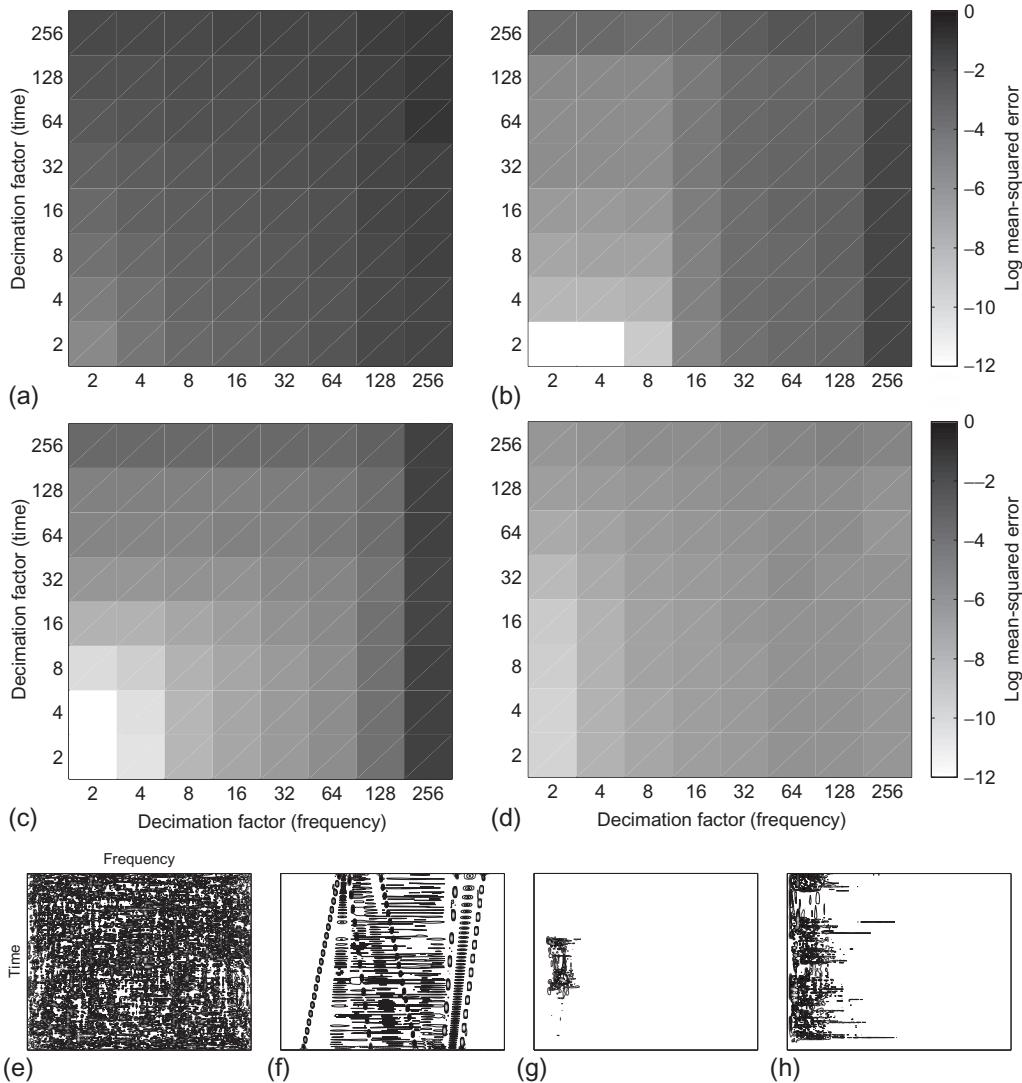
**FIGURE 6.6.2**

Comparing over-sampled TFDs of a bird song (used on the cover of the first edition of this book with data available as supplementary material): TFD of size (a) 256×512 and (b) 4500×4500 , for signal of length 4500 samples. The TFDs have separable kernels of the form: Hanning window of length 111 (samples) for the Doppler function $G_1[l]$, Hanning window of length 221 for the lag function $g_2[m]$. Although of different sizes, both TFDs have the same time-frequency resolution.

The following numerical analysis illustrates the heterogeneous effects of decimation. Using the algorithm in [Section 6.6.3.1](#), we generate decimated TFDs for four different signals over a range of decimation factors. The loss of (t,f) resolution is assessed using the mean-squared error (MSE) between the decimated TFD and the full TFD. To estimate the MSE, we interpolate the decimated TFD, using 2D cubic spline interpolation, so the decimated TFD is the same size ($N \times N$) as the full TFD. Increased decimation in the TFD will decrease (t,f) resolution and thus increase the error between the interpolation parts and the real (full) TFD.

The TFD is a nonseparable product kernel, generated using a Hamming window of length 4001 samples. The TFD is decimated both in time and frequency over the range $\{2, 4, 8, 16, 32, 64, 128, 256\}$ of decimation factors (a and b in [Algorithm 6.6.3.1](#)). The set of four signals are comprised of two synthetic signals and two real-world signals: a signal of zero-mean white Gaussian noise; a signal consisting of five linear frequency-modulation signal components; a heart rate variability (HRV) signal, recorded from a piglet during hypoxic-ischemic brain injury; and an EEG signal with multiple epileptiform discharges recorded from a patient with suspected brain injury. Each signal contains 4064 samples.

[Figure 6.6.3](#) shows the differences in loss of (t,f) resolution, as quantified by the MSE among the decimated and full TFDs, for the four different signals. As expected, the white Gaussian noise signal has the highest MSE throughout, as this signal has the largest BT product of all four signals. For the three other signals, different patterns emerge as all three signals have different time and frequency supports, but it is clear that a low (normalized) MSE of 2.5×10^{-3} is possible using decimation factors up to 32 in both the time and frequency directions.

**FIGURE 6.6.3**

Error representing the loss of (t, f) resolution for the decimated nonseparable-kernel TFD. The images (a-d) display log of the MSE between the true distribution and the interpolated decimated distribution. The four signal types are (a) white Gaussian noise, (b) a combination of linear frequency-modulated signals, (c) heart-rate variability recorded from a hypoxic piglet, and (d) an electroencephalogram recording with epileptiform discharges; TFDs for each signal are plotted in (e-h). Each signal is 4064 samples long, and thus decimated TFDs vary from size 2032×2032 to 16×16 .

6.6.6 SUMMARY AND CONCLUSIONS

The presented algorithms aim to reduce memory requirements and speed up computation. This has the potential to significantly benefit real-life signal processing applications, such as biomedicine [44], where there is often a need to process large sets of data. For example the separable-kernel algorithm ([Algorithm 6.6.2.2](#)), with $N_{\text{freq}} \ll N$ and $N_{\text{time}} \ll N$, will compute the TFD with a computational load that is independent of N^2 and do so requiring only $N_{\text{time}}N_{\text{freq}}$ sample points of memory. This reduced-size TFD retains all signal information and equal time-frequency resolution. Alternatively, the algorithm for the decimated TFD with a nonseparable kernel ([Algorithm 6.6.3.1](#)) will compute the TFD with $N^2/(ab)$ sample points of memory. The decimated TFD does lose time-frequency resolution but would be useful for long-duration signals or for hardware with limited memory resources.

The algorithms could be easily adapted to take advantage of computing trends toward parallel processing Ref. [42]. Multicore CPUs or GPUs (graphical processing units) could compute each vector of the algorithm's 2D FFTs in parallel, thus computing $N \times$ FFTs in approximately the time taken to compute one FFT operation.

REFERENCES

- [1] A.B. Carlson, Communication Systems, second ed., McGraw-Hill, Tokyo, 1975.
- [2] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis", *IEEE Trans. Acoust. Speech Signal Process.* 36 (9) (1988) 1518-1521.
- [3] B. Boashash, P.J. Black, "An efficient real-time implementation of the Wigner-Ville distribution", *IEEE Trans. Acoust. Speech Signal Process.* 35 (1987) 1611-1618.
- [4] B. Boashash, A.P. Reilly, "Algorithms for time-frequency signal analysis", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 163-181 (Chapter 7).
- [5] M.S. Richman, T.W. Parks, R.G. Shenoy, "Discrete-time, discrete-frequency, time-frequency analysis", *IEEE Trans. Signal Process.* 46 (6) (1998) 1517-1527.
- [6] J.M. O'Toole, M. Mesbah, B. Boashash, "Accurate and efficient implementation of the time-frequency matched filter", *IET Signal Process.* 4 (4) (2010) 428-437.
- [7] J.C. O'Neill, W.J. Williams, "Shift covariant time-frequency distributions of discrete signals", *IEEE Trans. Signal Process.* 47 (1) (1999) 133-146.
- [8] A.H. Costa, G.F. Boudreax-Bartels, "An overview of aliasing errors in discrete-time formulations of time-frequency representations", *IEEE Trans. Signal Process.* 47 (5) (1999) 1463-1474.
- [9] L. Stanković, "A method for time-frequency analysis", *IEEE Trans. Signal Process.* 42 (1) (1994) 225-229.
- [10] L. Stanković, T. Thayaparan, M. Daković, "Signal decomposition by using the S-method with application to the analysis of HF radar signals in sea-clutter", *IEEE Trans. Signal Process.* 54 (11) (2006) 4332-4342.
- [11] L. Stanković, T. Alieva, M.J. Bastiaans, "Fractional-Fourier-domain weighted Wigner distribution", in: Proc. Eleventh IEEE Workshop on Statistical Signal Processing, 2001, pp. 321-324.
- [12] P. Gonçalvès, R.G. Baraniuk, "Pseudo affine Wigner distributions: definition and kernel formulation", *IEEE Trans. Signal Process.* 46 (6) (1998) 1505-1517.
- [13] L. Stanković, S. Stanković, M. Daković, "From the STFT to the Wigner distribution", *IEEE Signal Process. Mag.* 31 (3) (2014) 163-174.
- [14] L. Stanković, "A method for improved distribution concentration in the time-frequency analysis of the multicomponent signals using the L-Wigner distribution", *IEEE Trans. Signal Process.* 43 (5) (1995) 1262-1268.

- [15] B. Ristic, B. Boashash, "Relationship between the polynomial and higher order Wigner-Ville distribution", *IEEE Signal Process. Lett.* 2 (12) (1995) 227-229.
- [16] L. Stanković, M. Daković, T. Thayaparan, Time-Frequency Signal Analysis with Applications, Artech House, Boston, 2013.
- [17] L. Stanković, "A multitime definition of the Wigner higher order distribution: L-Wigner distribution", *IEEE Signal Process. Lett.* 1 (7) (1994) 106-109.
- [18] L. Stanković, "An analysis of Wigner higher order spectra of multicomponent signals", *Ann. Télécommun.* 49 (3/4) (1994) 132-136.
- [19] L. Stanković, "An analysis of some time-frequency and time-scale distributions", *Ann. Télécommun.* 49 (9/10) (1994) 505-517.
- [20] L. Stanković, M. Daković, T. Thayaparan, "A real-time time-frequency based instantaneous frequency estimator", *Signal Process.* 93 (5) (2013) 1392-1397.
- [21] S. Stanković, L. Stanković, Z. Uskoković, "On the local frequency, group shift, and cross-terms in some multidimensional time-frequency distributions: a method for multidimensional time-frequency analysis", *IEEE Trans. Signal Process.* 43 (7) (1995) 1719-1725.
- [22] D. Gabor, "Theory of communication", *J. IEE* 93(III) (26) (1946) 429-457.
- [23] M.J. Bastiaans, "Gabor's signal expansion and its relation to sampling of the sliding-window spectrum", in: R.J. Marks II (Ed.), Advanced Topics in Shannon Sampling and Interpolation Theory, Springer, New York, 1993, pp. 1-35.
- [24] M.J. Bastiaans, "Gabor's expansion and the Zak transform for continuous-time and discrete-time signals: critical sampling and rational oversampling", Research Report 95-E-295, Eindhoven University of Technology, Eindhoven, The Netherlands, 1995.
- [25] M.J. Bastiaans, M.C.W. Geilen, "On the discrete Gabor transform and the discrete Zak transform", *Signal Process.* 49 (3) (1996) 151-166.
- [26] H.G. Feichtinger, T. Strohmer (Eds.), Gabor Analysis and Algorithms: Theory and Applications, Birkhäuser, Berlin/Boston, 1998.
- [27] A.J. van Leest, M.J. Bastiaans, "Implementations of non-separable Gabor schemes", in: Proc. 12th European Signal Processing Conf. (EUSIPCO-04), 2004, pp. 1565-1568.
- [28] W.J. Williams, J.C. O'Neill, "Decomposition of time-frequency distributions using scaled window spectrograms", in: Proc. SPIE: Advanced Signal Processing Algorithms, vol. 2563, Soc. of Photo-Optical Instrumentation Engineers, 1995, pp. 44-58.
- [29] W.J. Williams, T.-H. Sang, J.C. O'Neill, E.J. Zalubas, "Wavelet windowed time-frequency distribution decompositions", in: Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations VII, vol. 3162, Soc. of Photo-Optical Instrumentation Engineers, 1997, pp. 149-160.
- [30] W.J. Williams, S. Aviyente, "Optimum window time-frequency distribution decompositions", in: Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers, 1998, pp. 817-821.
- [31] S. Aviyente, W.J. Williams, "A centrosymmetric kernel decomposition for time-frequency distribution computation", *IEEE Trans. Signal Process.* 52 (6) (2004) 1574-1584.
- [32] G.S. Cunningham, W.J. Williams, "Kernel decomposition of time-frequency distributions", *IEEE Trans. Signal Process.* 42 (6) (1994) 1425-1442.
- [33] G.S. Cunningham, W.J. Williams, "Fast implementations of generalized discrete time-frequency distributions", *IEEE Trans. Signal Process.* 42 (6) (1994) 1496-1508.
- [34] L.B. White, "Transition kernels for bilinear time-frequency signal representations", *IEEE Trans. Signal Process.* 39 (2) (1991) 542-544.
- [35] M.G. Amin, "Spectral decomposition of time-frequency distribution kernels", *IEEE Trans. Signal Process.* 42 (5) (1994) 1156-1165.

- [36] W.J. Williams, S. Aviyente, “Minimal window time-frequency distributions”, in: Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX, vol. 3807, Soc. of Photo-Optical Instrumentation Engineers, 1999, pp. 446-457.
- [37] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), Advances in Spectrum Analysis and Array Processing, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [38] S.L. Marple Jr., “Computing the discrete-time ‘analytic’ signal via FFT”, *IEEE Trans. Signal Process.* 47 (9) (1999) 2600-2603.
- [39] A.P. Reilly, G.J. Frazer, B. Boashash, “Analytic signal generation—tips and traps”, *IEEE Trans. Signal Process.* 42 (11) (1994) 3241-3245.
- [40] J.M. O'Toole, M. Mesbah, B. Boashash, “A new discrete analytic signal for reducing aliasing in the discrete Wigner-Ville distribution”, *IEEE Trans. Signal Process.* 56 (11) (2008) 5427-5434.
- [41] O.P. Kenny, B. Boashash, “An optical signal processor for time-frequency signal analysis using the Wigner-Ville distribution”, *J. Electr. Electron. Eng. Aust.* 8 (3) (1988) 152-158.
- [42] J.M. O'Toole, B. Boashash, “Fast and memory-efficient algorithms for computing quadratic time-frequency distributions”, *Appl. Comput. Harmonic Anal.* 35 (2) (2013) 350-358.
- [43] J.M. O'Toole, M. Mesbah, B. Boashash, “Improved discrete definition of quadratic time-frequency distributions”, *IEEE Trans. Signal Process.* 58 (2) (2010) 906-911.
- [44] B. Boashash, G. Azemi, J.M. O'Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.

This page intentionally left blank

MEASURES, PERFORMANCE ASSESSMENT, AND ENHANCEMENT OF TFDs

7

INTRODUCTION AND OVERVIEW

This chapter describes a number of time-frequency (t, f) performance quality measures specifically developed as criteria for performance enhancement for a given application. The adopted performance measures are defined using objective criteria followed by time-frequency distribution (TFD) enhancement methods to improve the (t, f) concentration, resolution, and readability of TFDs. The topic is covered in nine sections.

Hyperbolic FM signals are well described by the affine Wigner-Ville distribution (WVD), a method related to time-scale analysis and the wavelet transform (Section 7.1). A general procedure for enhancing the time-frequency resolution and readability of TFDs is the reassignment principle described in Section 7.2. Techniques for measuring the concentration of TFDs and for automatic optimization of their parameters are presented based on entropy measures (Section 7.3). Another approach defines a resolution performance measure using local measurements in the (t, f) domain, such as relative amplitudes of auto-terms and cross-terms (Section 7.4). Then, a comparison of postprocessing methods in the (t, f) domain is presented (Section 7.5). The last four topics focus on time-frequency compressive sensing (Section 7.6); signal complexity estimation using (t, f) entropy measures (Section 7.7); time-frequency analysis using neural networks (Section 7.8); and attempts to unify time-frequency, time-scale, filter banks, wavelets, and the discrete-time Gabor transform using product functions and cascaded frames are presented briefly as they may assist in the selection of the best-performing method for a given application (Section 7.9).

7.1 TFD DESIGN BASED ON THE AFFINE GROUP⁰

7.1.1 SIGNAL SCALE TRANSFORMATIONS TO CONSTRUCT TFDs

In digital signal processing (DSP), the reference variable is time t . Transformations of that variable are naturally interpreted as *clock* changes, that is, changes in the origin or the unit of time, and are therefore expected to have the form

$$t \longrightarrow t' = at + b; \quad t = a^{-1}(t' - b), \quad (7.1.1)$$

where t and t' are the time variables and a and b are real numbers with $a > 0$. The set of all such transformations constitutes the *affine group*.

These transformations affect physical units (which involve powers of the time unit) by multiplying each of them by some power of the dilation a . Thus, a change of clock of the form (Eq. 7.1.1) will induce in a signal $s(t)$ a transformation of the form [2]

$$s(t) \longrightarrow s'(t) = a^r s(a^{-1}(t - b)), \quad (7.1.2)$$

where r , known as the *index of (dimensional) conformity*, is a real exponent depending on the physical nature of the signal.

In time-frequency (t, f) analysis, signals are described not by their instantaneous values in time, but as modulations of amplitude and frequency. In practice, this description is summarized by a real function $\rho(t, f)$ which gives an image of signal spreading in the (t, f) half-plane ($f > 0$). The basic (t, f) problem is to set up the correspondence between the physical signal $s(t)$ and its (t, f) representation $\rho(t, f)$. It is well known that the operation cannot be linear. The usual method consists of defining $\rho(t, f)$ as a Hermitian functional in a Hilbert space whose elements are directly related to the physical signals. In affine theory, this classical approach is followed, the specific point being the introduction of a Hilbert space where an irreducible unitary representation of the affine group exists [3].

A *real* signal $s(t)$ is fully characterized by the positive-frequency part of its Fourier transform (FT). In this section, we define the positive-frequency part as

$$Z(f) \equiv U(f) \int_{-\infty}^{\infty} e^{-j2\pi ft} s(t) dt, \quad (7.1.3)$$

where $U(f)$ is the Heaviside step function and the integral is the FT.

In agreement with Eq. (7.1.2), a general affine transformation on the signal $s(t)$ acts equivalently on $Z(f)$ according to

$$Z(f) \longrightarrow Z'(f) = a^{r+1} e^{-j2\pi bf} Z(af), \quad (7.1.4)$$

where r is the index of conformity of the physical signal. In contrast with Eq. (7.1.2), the transformation (Eq. 7.1.4) corresponds to an irreducible representation of the affine group on the Hilbert space of the functions $Z(f)$ such that $Z(f)f^{(r+1/2)}$ is square integrable, equipped with the scalar product

⁰Second edition updated by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Touati**, King Saud University, Riyadh, Kingdom of Saudi Arabia (stouati@ksu.edu.sa). First edition [1] by J. Bertrand and P. Bertrand. Reviewers: B.K. Jawad, S. Assous, and S. Ouelha.

$$\langle Z_1, Z_2 \rangle \equiv \int_0^\infty Z_1(f) Z_2^*(f) f^{2r+1} df. \quad (7.1.5)$$

We can easily show that the transformation (Eq. 7.1.4) conserves this scalar product, and therefore conserves the norm $\|Z\|$ defined by $\|Z\|^2 = \langle Z, Z \rangle$, that is

$$\|Z\|^2 \equiv \int_0^\infty |Z(f)|^2 f^{2r+1} df. \quad (7.1.6)$$

In the following, we will adopt this norm for which the representation (Eq. 7.1.4) of the affine group is unitary, that is, conserves the norm $\|Z\|$.

A change of reference variables will change the TFD of a signal in a way that depends on the change of variables (cf. Eq. 7.1.1) and on the meaning of the function $\rho(t, f)$ as a quantitative representation of signal spreading in the (t, f) half-plane. This leads to the general transformation

$$\rho(t, f) \longrightarrow \rho'(t, f) \equiv a^q \rho(a^{-1}(t - b), af), \quad (7.1.7)$$

where q is real and can be considered as an index of significance. The integral of $\rho(t, f)$ on the half-plane that is invariant under transformation (Eq. 7.1.7) has the form

$$\iint f^q \rho(t, f) dt df. \quad (7.1.8)$$

For the common choice $q = 0$, a probabilistic interpretation of the spreading is possible, as Eq. (7.1.8) reduces to $\iint \rho(t, f) dt df$; however, in special applications, the choice of other values for q can be pertinent.

The dimensional indices r and q have been introduced by the consideration of scaling transformations. They do not appear in the case of the WVD, which is based, in an essential way, on translations in both time and frequency through the Weyl-Heisenberg group. Indeed, it can be seen that the invariance of the scalar product (Eq. 7.1.5) is possible only if $r = -1/2$ (i.e., $2r + 1 = 0$) and that the invariance of the integral Eq. (7.1.8) under frequency translations is possible only if $q = 0$.

7.1.2 TOMOGRAPHIC METHODS AND THE AFFINE WVD

Tomographic methods are now routinely used in signal analysis. In (t, f) analysis, they arose from the recognition that the *generalized marginals* of the WVD (along straight lines of arbitrary slope) are positive and easily interpretable in terms of the signal. One immediate result was the construction of the WVD by relying only on the Weyl-Heisenberg group, which performs (t, f) translations [4,5]. Applications have arisen in quantum optics, where tomography [6] allows precise determination of the state of a system from experimental data [7]. A related method is the so-called fractional Fourier transform, which arises naturally in the expression for the Radon transform of the Wigner-Ville function (see Ref. [8] and Sections 4.8 and 5.8).

In fact, the relation between tomography and WVD is connected in an essential way to the underlying group. In traditional signal analysis, the straight lines in a given direction in the (t, f) plane arise as families of curves invariant under a subgroup of the translations group. In the affine case, as we show next, an analogous tomography arises from the subgroups acting as dilations around a fixed time.

Linear chirps, which correspond to straight lines in the (t, f) plane, play a central role in the usual tomography because they are invariant, except for phase, under the subgroup of translations in a given direction. In the affine case, the same role is played by signals that are invariant, except for phase, in dilations centered at the instant $t = \xi$. These are the hyperbolic chirps defined by

$$\psi_\beta^\xi(f) = f^{-j2\pi\beta-r-1} e^{-j2\pi\xi f}, \quad \beta \text{ real.} \quad (7.1.9)$$

It can be shown that such signals conserve the same form in the transformation of type Eq. (7.1.4) subject to the constraint

$$b = \xi(1 - a). \quad (7.1.10)$$

These transformations are just those of the subgroup of dilations of origin $t = \xi$.

A signal $Z(f)$, with index of conformity r , can be decomposed on the basis $\psi_\beta^\xi(f)$, with fixed ξ . The coefficients of the development are

$$\langle Z, \psi_\beta^\xi \rangle = \int_0^\infty Z(f) f^{j2\pi\beta+r} e^{j2\pi\xi f} df, \quad (7.1.11)$$

where the scalar product is defined from Eq. (7.1.5). The expression (7.1.11) defines the general Mellin transform whose inverse can be easily obtained.

The group delays of the hyperbolic chirps (Eq. 7.1.9) at frequency f are found to be

$$t = \xi + \beta/f, \quad \beta \text{ real,} \quad (7.1.12)$$

and describe curves in the half-plane $f > 0$ that are invariant by dilation (Eq. 7.1.10).

Let $\rho^q(t, f; Z)$ be the TFD of signal $Z(f)$ that is under construction. The tomographic condition relates the integral of $\rho^q(t, f)$ along hyperbolas (Eq. 7.1.12) to the coefficients of the hyperbolic-chirp decomposition (Eq. 7.1.11). It takes the form [9]

$$\int_{-\infty}^\infty \int_0^\infty \rho^q(t, f; Z) \delta((t - \xi)f - \beta) f^q dt df = \left| \langle Z, \psi_\beta^\xi \rangle \right|^2, \quad (7.1.13)$$

where each member is separately invariant by affine transformations. Considered for all real values of β and ξ , this relation has the form of a generalized Radon transform for $\rho^q(t, f)$. Its inversion gives the expression for the *affine WVD* [4,10,11]:

$$\rho^q(t, f; Z) = f^{2r-q+2} \int_{-\infty}^\infty e^{j2\pi u ft} Z\left(\frac{fu e^{u/2}}{2 \sinh(u/2)}\right) Z^*\left(\frac{fu e^{-u/2}}{2 \sinh(u/2)}\right) \left(\frac{u}{2 \sinh(u/2)}\right)^{2r+2} du. \quad (7.1.14)$$

The integrand can be seen to depend only on the functions $\lambda(u)$ and $\lambda(-u)$ where

$$\lambda(u) = \frac{u e^{u/2}}{2 \sinh(u/2)}. \quad (7.1.15)$$

In formula (7.1.14), the index of significance q can be chosen according to the type of TFD needed, but r is necessarily the index of conformity attached to $Z(f)$.

When an affine change is performed on $Z(f)$ according to Eq. (7.1.4), the function (7.1.14) is clearly transformed as in Eq. (7.1.7). Hence, the affine WVD satisfies

$$\rho^q(t, f; Z') = a^q \rho^q(a^{-1}(t - b), af; Z), \quad (7.1.16)$$

where $Z'(f)$ is defined in Eq. (7.1.4). This property of invariance of the correspondence between a signal $Z(f)$ and its TFD $\rho^q(t, f)$ is essential for an analysis founded on the affine group; however, it does not characterize univocally the affine WVD. In fact, an infinite family of functions satisfying condition (7.1.16) can be determined. They form the affine analog of the quadratic class. But the affine WVD (Eq. 7.1.14) is the only member with a tomographic construction based on the affine group alone. In this sense, it occupies the same position in the affine class as the usual Wigner-Ville function does in the (time- and frequency-covariant) quadratic class.

7.1.3 COMPARISON OF THE AFFINE AND USUAL WIGNER-VILLE FUNCTIONS

When the affine Wigner function (7.1.14) is used for analysis of *narrowband signals*, the factor f^{2r-q+1} is approximately constant, with f replaced by the central frequency f_0 of the band. Moreover, the integrand in Eq. (7.1.14) is different from zero only when the arguments of Z and Z^* are both in the band. This requires u to remain close to zero. In fact, it can be observed that the extent of the domain of u will decrease along with the bandwidth of the signal. This allows the function $\lambda(u)$ to be approximated by its first-order expansion about $u = 0$:

$$\lambda(u) \simeq 1 + u/2. \quad (7.1.17)$$

A change of variables from u to $v \equiv uf$ then leads to the approximate form

$$\rho^q(t, f) \approx f_0^{2r-q+1} \int_{-\infty}^{\infty} e^{j2\pi vt} Z(f + \frac{v}{2}) Z^*(f - \frac{v}{2}) dv. \quad (7.1.18)$$

Apart from a scale factor, this is the WVD expressed in terms of $Z(f)$ (Eq. 2.1.32, p. 72). Thus, for narrowband signals, affine WVD equals standard WVD.

In the wideband case, let us study the properties of the affine WVD as defined in Eq. (7.1.14) in comparison with those of the usual Wigner-Ville function. The total integral of $\rho^q(t, f)$ over the (t, f) space is now written in the invariant form

$$\int_{-\infty}^{\infty} \int_0^{\infty} \rho^q(t, f; Z) f^q dt df = \|Z\|^2, \quad (7.1.19)$$

where $\|Z\|$ is defined in Eq. (7.1.6). Using Eq. (7.1.18) and integrating over t yields

$$\int_{-\infty}^{\infty} \rho^q(t, f; Z) dt = f^{2r+1-q} |Z(f)|^2. \quad (7.1.20)$$

If $q = 2r + 1$, Eq. (7.1.20) reduces to the frequency marginal (Section 3.1.1).

Since $\rho^q(t, f)$ represents a spreading of the signal in the (t, f) plane, it can be used to compute the mean value of the epoch t for a fixed value of f . The resulting expression, whatever the value of q , is the spectral delay of the signal, as in the case of usual Wigner-Ville function

$$\left(\int_{-\infty}^{\infty} t \rho^q(t, f; Z) dt \right) \left(\int_{-\infty}^{\infty} \rho^q(t, f; Z) dt \right)^{-1} = -\frac{1}{2\pi} \frac{d\theta(f)}{df}, \quad (7.1.21)$$

where the phase $\theta(f)$ of the analytic signal $Z(f)$ is defined by

$$Z(f) = |Z(f)| e^{j\theta(f)}. \quad (7.1.22)$$

It may also be appropriate to compare the affine and usual WVDs in terms of frequency spreading. One must be particularly careful with the indices r and q when studying the effects of transducing filters on (t, f) representations. Suppose an input signal $Z(f)$, having an index of conformity equal to r , excites a linear space-invariant (LSI) system whose output has a different physical nature from its input. The output can be written

$$\tilde{Z}(f) = H(f)Z(f), \quad (7.1.23)$$

where $H(f)$ is the system transfer function. In general, $\tilde{Z}(f)$ will transform under dilations with an index of conformity \tilde{r} different from r . Hence, for consistency, the function $H(f)$ must be assigned an index α determined by relation (Eq. 7.1.23) as

$$\alpha = \tilde{r} - r - 1. \quad (7.1.24)$$

Note that the identity filter corresponds here to $\alpha = -1$.

These operations have a counterpart in the domain of (t, f) functions. Let $\rho_{(r)}^q(t, f; Z)$, $\rho_{(\alpha)}^\gamma(t, f; H)$, and $\rho_{(\tilde{r})}^{\tilde{q}}(t, f; \tilde{Z})$ be the TFDs corresponding to functions $Z(f)$, $H(f)$, and $\tilde{Z}(f)$, respectively. The dependence of the representations on the indices of conformity of the signals has been shown explicitly, to avoid ambiguity. The (t, f) analog of relation (Eq. 7.1.23) has the form of a convolution in time provided the indices are related according to

$$\tilde{q} = q + \gamma + 1. \quad (7.1.25)$$

In practical situations, where $\tilde{q} = q$ and $\gamma = -1$, the relation is

$$\rho_{(\tilde{r})}^q(t, f; \tilde{Z}) = \int_{-\infty}^{\infty} \rho_{(\alpha)}^{-1}(t - t', f; H) \rho_{(r)}^q(t', f; Z) dt' = \rho_Z(t, f) * \rho_H(t, f), \quad (7.1.26)$$

where the index α is given by Eq. (7.1.24).

Consider now the case of a transducer defined by

$$H(f) = f^{-h}, \quad h \text{ real}, \quad (7.1.27)$$

with conformity index $\alpha = h - 1$. The only action of such a device is to replace the signal $Z(f)$, whose index of conformity is r , by the signal $Z'(f) = f^{-h}Z(f)$, whose index is $\tilde{r} = r + h$. Starting from Eq. (7.1.14) with a narrowband signal and a first-order expansion about $u = 0$, relation (Eq. 7.1.26) becomes

$$\rho_{(r+h)}^q(t, f; f^{-h}Z) \equiv \rho_{(r)}^q(t, f; Z). \quad (7.1.28)$$

Thus, once q has been chosen, computing $\rho^q(t, f)$ for a signal $Z(f)$ or for any of its transforms (as defined by Eqs. (7.1.23) and (7.1.27)) yields the same result, provided care is taken to use the correct conformity indices. In the present context, the property leads to considering the device defined by Eq. (7.1.27) as a perfect transducer.

7.1.4 HYPERBOLIC CHIRPS AND AFFINE GROUP EXTENSION

The WVD properties of marginalization, localization, and extended invariance have direct analogs in the case of affine WVD. The marginal condition, leading to a density in f , is derived from Eq. (7.1.20).

In addition, marginal densities in β are obtained for each value of ξ when integrating $\rho^q(t, f)$ along hyperbolas $(t - \xi)f = \beta$. The tomographic condition (7.1.13) implies that these densities have the form

$$\rho_\xi(\beta) = |\langle Z, \psi_\beta^\xi \rangle|^2 \quad (7.1.29)$$

and satisfy

$$\int_{-\infty}^{\infty} \rho_\xi(\beta) d\beta = \|Z\|^2. \quad (7.1.30)$$

Localization in the (t, f) space arises for general hyperbolic chirps. For these signals, a direct computation gives

$$\rho^q(t, f; \psi_\beta^\xi) = f^{-q} \delta((t - \xi)f - \beta). \quad (7.1.31)$$

The special case $\beta = 0$ corresponds to a localized signal $\psi_0^\xi(f) = f^{-r-1} e^{-j2\pi\xi f}$ attached to the instant $t = \xi$. The latter form can be obtained directly by requiring that after a clock change, labeled by (a, b) and acting as in Eq. (7.1.4), the signal keeps a localized form at the transformed instant $t' = a\xi + b$.

Another case of localization, which can be considered as a limit of the hyperbolic chirp behavior, concerns the pure frequency signal with index of conformity r

$$Z_{f_0}(f) \equiv f^{-r} \delta(f - f_0), \quad (7.1.32)$$

which is represented by

$$\rho^q(t, f; Z_{f_0}) = f^{1-q} \delta(f - f_0). \quad (7.1.33)$$

The property of invariance, under affine transformations, of the correspondence between a signal $Z(f)$ and its representation $\rho^q(t, f; Z)$ can be extended. To this end, we introduce the transformations performing a translation on the β -parameter of the hyperbolic chirps according to

$$\psi_\beta^\xi(f) \longrightarrow \psi_{\beta+c}^\xi(f) = f^{-j2\pi(\beta+c)-r-1} e^{-j2\pi\xi f}, \quad (7.1.34)$$

where c is a real number. These transformations act on an arbitrary signal as

$$Z(f) \longrightarrow Z_c(f) = f^{-j2\pi c} Z(f). \quad (7.1.35)$$

Moreover, they combine with the affine transformations to form a three-parameter group G_0 , which is the largest group conserving the family of hyperbolas (Eq. 7.1.12) as a whole and the family of hyperbolic chirps. These general transformations, labeled by (a, b, c) , act on a signal according to

$$Z(f) \longrightarrow Z_{a,b,c}(f) = a^{r+1} e^{-j2\pi(bf+c \ln af)} Z(f), \quad (7.1.36)$$

and on its affine Wigner function according to

$$\rho^q(t, f; Z) \longrightarrow \rho^q(t, f; Z_{a,b,c}) = a^q \rho^q(a^{-1}(t - b - cf^{-1}), af; Z). \quad (7.1.37)$$

Thus the correspondence between the signal $Z(f)$ and the function $\rho^q(t, f; Z)$ is invariant not only by the affine group but also by its extension G_0 .

7.1.5 UNITARITY PROPERTY AND SOME OF ITS CONSEQUENCES

The affine Wigner function (7.1.14) satisfies the *unitarity* (or *Moyal*) property

$$\int_{-\infty}^{\infty} \int_0^{\infty} \rho^q(t, f; Z) \rho^q(t, f; Z') f^{2q} dt df = |\langle Z, Z' \rangle|^2, \quad (7.1.38)$$

where the scalar product $\langle Z, Z' \rangle$, issued from definition (7.1.5), depends on the index r of the signal.

A special case of relation (Eq. 7.1.38) is obtained when $Z'(f)$ is the hyperbolic chirp $\psi_{\beta}^{\xi}(f)$ (cf. Eq. 7.1.9) so that $\rho^q(t, f; Z')$ has the form (7.1.31). The result is just the tomographic condition, which was introduced in Eq. (7.1.13).

A more general form of the unitarity property (see Ref. [12] formulas (III.15-17)) finds a direct application in the reconstruction of the signal from its affine Wigner function. In fact, it allows writing directly the formula

$$\begin{aligned} Z(f_1) Z^*(f_2) &= (f_1 f_2)^{-2r-1} \int f^{2r+2+q} \rho^q(t, f; Z) e^{j2\pi u ft} \left(\frac{u}{2 \sinh \frac{u}{2}} \right)^{2r+2} \\ &\quad \cdot \delta \left(f_1 - \frac{fu e^{-u/2}}{2 \sinh \frac{u}{2}} \right) \delta \left(f_2 - \frac{fu e^{u/2}}{2 \sinh \frac{u}{2}} \right) du dt df, \end{aligned}$$

where δ denotes the Dirac distribution. Knowledge of $\rho(f_1, f_2) = Z(f_1) Z^*(f_2)$ allows reconstruction of the analytic signal up to a constant phase. Explicitly, we have

$$|Z(f)| = \sqrt{\rho(f, f)}; \quad e^{i(\theta_1 - \theta_2)} = \frac{\rho(f_1, f_2)}{\rho(f_1, f_1) \rho(f_2, f_2)}, \quad (7.1.39)$$

where the decomposition (7.1.22) of $Z(f)$ has been used. So the affine WVD (Eq. 7.1.14) represents the signal without losing any information except a constant phase.

The unitarity relation can also be used to define a regularized version $\tilde{\rho}^q(t, f)$ of $\rho^q(t, f)$. To this end, a basic function $\Phi(f)$ is chosen and its representation $\rho^q(t, f; \Phi)$ written down. The smoothed version $\tilde{\rho}^q(t, f; Z)$ of the representation $\rho^q(t, f; Z)$ is then defined by the convolution on the affine group

$$\tilde{\rho}^q(t, f; Z) = f^{-q} \int_{-\infty}^{\infty} \int_0^{\infty} \rho^q(t', f'; Z) \rho^q(f(t' - t), f'/f; \Phi) (f')^{2q} dt' df', \quad (7.1.40)$$

where the kernel is the TFD of the function Φ . A more practical form of the smoothed function $\tilde{\rho}^q$ is obtained when taking into account the transformation law Eq. (7.1.16) of $\rho^q(t, f; \Phi)$ and the unitarity property (Eq. 7.1.38)

$$\tilde{\rho}(t, f; Z) = |\langle Z, \Phi_{(t, f)} \rangle|^2, \quad (7.1.41)$$

where

$$\Phi_{(t, f)}(f') = f^{-r-1} e^{-j2\pi f' t} \Phi(f'/f). \quad (7.1.42)$$

The set of functions $\Phi_{(t, f)}(f')$ is recognized as a family of wavelets obtained from the mother wavelet $\Phi(f')$ by an affine transformation such that $a = f^{-1}$, $b = t$. Thus, the right-hand side of relation (Eq. 7.1.41) is the squared modulus of the wavelet coefficient of $Z(f)$. Conversely, the above developments allow the squared modulus of the wavelet coefficient to be interpreted as the result of a smoothing in the (t, f) half-plane. This provides a guide to assessing the properties of a mother wavelet.

For example, the function $\Phi(f)$ may be chosen so that its representation $\rho^q(t, f; \Phi)$ has the best possible concentration in the (t, f) plane. An optimal choice for $\Phi(f)$ is found to be the Klauder wavelet [10]. Using as parameters the mean frequency $f_0 = \langle f \rangle$, the mean square deviation $\sigma_f \equiv \langle f^2 \rangle - f_0^2$ (where $\langle \cdot \rangle$ denotes the expected value), and the relative bandwidth $B = \sigma_f/f_0$, that wavelet can be written

$$\Phi(f) \equiv K(B)f_0^{-r-1} (f/f_0)^{-r-1+1/(2B^2)} \exp\left(-\frac{1}{2B^2}\frac{f}{f_0}\right), \quad (7.1.43)$$

where $K(B)$ is a normalization constant.

The function (7.1.43) has a (t, f) representation approximately localized around the point $t = 0$, $f = f_0$. By varying σ_f , one can stretch the representation spread in the t or f direction without changing the central point.

7.1.6 SUMMARY AND CONCLUSIONS

The group of affine transformations of time t , or clock changes, leads to an analysis of real signals independent of the system of reference and the system of units employed. The approach defines an adapted (t, f) representation, the *affine Wigner function* or *affine WVD*, which has many properties similar to those of the standard WVD.

The effects of time dilations on signals and (t, f) distributions are, respectively, characterized by two real indices r and q . The index of conformity r depends on the physical origin of the signal and controls its behavior in a change of time unit. The index of significance q , attached to the TFD, can be chosen according to the application. The indices r and q are special features coming from the introduction of dilations and allowing expression of their effects. These indices are particularly important in the (t, f) analysis of problems involving transductions. They can be overlooked only in the limiting case of narrow relative bandwidth, which reduces to the WVD.

There are several ways to build a (t, f) analysis satisfying the above constraints of independence relative to clock changes. They result in an affine analog of the quadratic class; however, what is called the affine WVD in the present work stands out as the unique TFD obtained by a tomographic method fitted to the affine group. This distribution is unitary, gives a realistic description of hyperbolic chirps, and does not discard any information contained in the original signal except a constant phase factor.

7.2 TIME-FREQUENCY REASSIGNMENT⁰

As discussed in Chapters 2 and 1–3, time-frequency (t, f) and time-scale representations aim to extract relevant information from a signal by representing it over a two-dimensional (2D) plane. These tools have been extensively studied in the past 20 years, resulting today in many useful analysis methods. Among them, the spectrogram and the smoothed versions of the WVD are probably the most widely

⁰Authors: **F. Auger**, L'Université Nantes Angers Le Mans (LUNAM), Saint Nazaire Cedex, France (francois.auger@univ-nantes.fr), **P. Flandrin**, École Normale Supérieure de Lyon, Lyon Cedex 07, France (flandrin@ens-lyon.fr), and **E. Chassande-Mottin**, Université Paris Diderot, Paris Cedex 13, France (ecm@apc.univ-paris7.fr). Reviewers: S. Barbarossa, R. Baraniuk, F. Hlawatsch, and A. Papandreou-Suppappola.

used, but their applicability is limited by localization trade-offs, which may be troublesome in some applications. For the spectrogram, a shorter analysis window yields a better time resolution and henceforth a poorer frequency resolution, as a consequence of the Gabor-Heisenberg inequality [13]. For the smoothed versions of the WVD, a larger smoothing kernel yields reduced cross-terms, but also a poorer localization of the signal components (see [Section 4.2](#)).

These shortcomings must be overcome in order to obtain (t, f) “pictures” that can be both easily read by non-experts and easily included in a signal processing application. This is exactly what the *reassignment principle* has been devised for. Initially introduced in 1976 by Kodera, Gendrin and de Villeddy [14], this idea first remained little known and rarely used. But, advances made during the 1980s in (t, f) analysis have made its rebirth in the early 2000s possible [15], which considerably extended its applicability, both conceptually and computationally.

This section first reviews the basic principle of reassignment in [Section 7.2.1](#). Two other extensions are then presented in [Section 7.2.2.1](#), whereas connections with independently proposed and closely connected approaches are pointed out in [Section 7.2.2.2](#). MATLAB™ implementations of the algorithms discussed here are included in a freeware package available online at tftb.nongnu.org.

7.2.1 BASIC PRINCIPLE

For a sake of simplicity, we will first present the basics of reassignment in the case of the spectrogram, which was the only case considered by Kodera et al. [14]. Its application to other representations will be discussed afterward. The spectrogram, which is the squared modulus of the short-time Fourier transform

$$S_x^h(t, f) = |F_x^h(t, f)|^2, \quad (7.2.1)$$

$$F_x^h(t, f) = \int x(u) h^*(t - u) e^{-i2\pi fu} du, \quad (7.2.2)$$

can also be expressed as a 2D smoothing of the WVD [13]:

$$S_x^h(t, f) = \iint W_x(u, v) W_h(t - u, f - v) du dv. \quad (7.2.3)$$

In these expressions, t and f are, respectively, the time and frequency running variables, $x(t)$ is the analyzed signal, and $h(t)$ is the analyzing window. All integrals have integration bounds running from $-\infty$ to $+\infty$. The latter expression shows explicitly that the value of the spectrogram at a given point (t, f) is a weighted sum of all the WVD values at the neighboring points $(t - u, f - v)$. The number $S_x^h(t, f)$ is therefore the sum of a whole energy distribution located around its geometrical center (t, f) . Reasoning with a mechanical analogy, the situation is as if the total mass of an object was assigned to its geometrical center, an arbitrary point which except in the very specific case of an homogeneous distribution, has no reason to suit the actual distribution. A much more meaningful choice is to assign the total mass of an object—as well as the spectrogram value $S_x^h(t, f)$ —to the *center of gravity* of their respective distribution. This is exactly what the reassignment performs: at each time-frequency

point (t, f) where a spectrogram value is computed, we also compute the coordinates (\hat{t}, \hat{f}) of the local centroid of the WVD W_x , as seen through the (t, f) window W_h centered at (t, f) :

$$\hat{t}_x(t, f) = \frac{1}{S_x^h(t, f)} \iint u W_x(u, v) W_h(t - u, f - v) du dv \quad (7.2.4)$$

$$\hat{f}_x(t, f) = \frac{1}{S_x^h(t, f)} \iint v W_x(u, v) W_h(t - u, f - v) du dv. \quad (7.2.5)$$

Then, the spectrogram value $S_x^h(t, f)$ is moved from (t, f) to (\hat{t}, \hat{f}) . This leads us to define the reassigned spectrogram as

$$\check{S}_x^h(t, f) = \iint S_x^h(u, v) \delta(t - \hat{t}_x(u, v)) \delta(f - \hat{f}_x(u, v)) du dv. \quad (7.2.6)$$

Originally, the reassignment operators \hat{t} and \hat{f} have been equivalently related to the phase of the STFT, an information which is generally discarded when computing the spectrogram:

$$\hat{t}_x(t, f) = -\frac{1}{2\pi} \frac{\partial \varphi}{\partial f}(t, f) \quad (7.2.7)$$

$$\hat{f}_x(t, f) = f + \frac{1}{2\pi} \frac{\partial \varphi}{\partial t}(t, f), \quad (7.2.8)$$

with $\varphi(t, f) = \arg F_x^h(t, f)$. These expressions may be interpreted, respectively, as the local group delay and the local instantaneous frequency of the signal observed inside the (t, f) domain imposed by the analysis window h . But it has been shown in Ref. [15] that a much more efficient implementation is possible thanks to a third expression involving two additional STFTs with particular analysis windows:

$$\hat{t}_x(t, f) = t - \operatorname{Re} \left\{ \frac{F_x^{th}(t, f)}{F_x^h(t, f)} \right\}, \quad (7.2.9)$$

$$\hat{f}_x(t, f) = f + \operatorname{Im} \left\{ \frac{F_x^{dh/dt}(t, f)}{2\pi F_x^h(t, f)} \right\}. \quad (7.2.10)$$

As presented here, the reassignment principle can be used with a large number of distributions, beyond the spectrogram case. For example, if the WVD of the short-time window $h(t)$ in Eq. (7.2.3) is replaced by an arbitrary (low-pass) kernel $\Pi(u, v)$, one recognizes the general form of the quadratic (t, f) energy distributions that are covariant under time and frequency shifts, referred to as Cohen's class¹ [13]:

$$\rho_x^\Pi(t, f) = \iint W_x(u, v) \Pi(t - u, f - v) du dv. \quad (7.2.11)$$

The local centroids are then given by

$$\hat{t}_x(t, f) = \frac{1}{\rho_x^\Pi(t, f)} \iint u W_x(u, v) \Pi(t - u, f - v) du dv \quad (7.2.12)$$

¹That is, the quadratic class as defined on p. 113.

$$\hat{f}_x(t, f) = \frac{1}{\rho_x^{\Pi}(t, f)} \iint v W_x(u, v) \Pi(t - u, f - v) \, du \, dv \quad (7.2.13)$$

and the corresponding reassigned distribution becomes

$$\check{\rho}_x^{\Pi}(t, f) = \iint \rho_x^{\Pi}(u, v) \delta(t - \hat{t}_x(u, v)) \delta(f - \hat{f}_x(u, v)) \, du \, dv. \quad (7.2.14)$$

From a theoretical point of view, this reassigned representation is no longer bilinear, but it still remains an energy distribution covariant under time and frequency shifts. One of the most important properties of the reassignment principle is that the application of the reassignment process defined by Eqs. (7.2.12), (7.2.13), and (7.2.14) to any distribution of Cohen's class yields perfectly localized distributions for chirp signals, frequency tones, and impulses, since the WVD does so, and since the centroid of a linear distribution necessarily lies on the line. When applied to multicomponent signals, reassignment improves readability by overcoming—to a certain extent—the usual trade-off between cross-term level and localization: the underlying *smoothing* of the standard distribution guarantees some cross-term reduction, whereas reassignment acts as a *squeezing* that refocuses the signal terms that had been spread out by smoothing (see Fig. 7.2.1).

Among the examples of Cohen's class members studied in Ref. [15], the case of the smoothed pseudo WVD yields a very versatile signal analysis tool, with independently adjustable time and frequency smoothings:

$$\text{SPWV}_x^{g,h}(t, f) = \iint g(t - u) H(f - v) W_x(u, v) \, du \, dv. \quad (7.2.15)$$

Its reassigned version can be computed easily with two additional SPWDs:

$$\hat{t}_x(t, f) = t - \frac{\text{SPWV}_x^{tg,h}(t, f)}{\text{SPWV}_x^{g,h}(t, f)}, \quad (7.2.16)$$

$$\hat{f}_x(t, f) = f + i \frac{\text{SPWV}_x^{g, dh/dt}(t, f)}{2\pi \text{SPWV}_x^{g,h}(t, f)}. \quad (7.2.17)$$

A different kind of generalization can be obtained when switching to time-scale energy distributions of the affine class [13], that is, the quadratic distributions covariant under time shifts and dilations:

$$\Omega_x^{\Pi}(t, a) = \iint W_x(u, v) \Pi\left(\frac{t-u}{a}, av\right) \, du \, dv. \quad (7.2.18)$$

Within this framework, the reassignment operator in time is given directly by

$$\hat{t}_x(t, a) = \frac{1}{\Omega_x^{\Pi}(t, a)} \iint u W_x(u, v) \Pi\left(\frac{t-u}{a}, av\right) \, du \, dv, \quad (7.2.19)$$

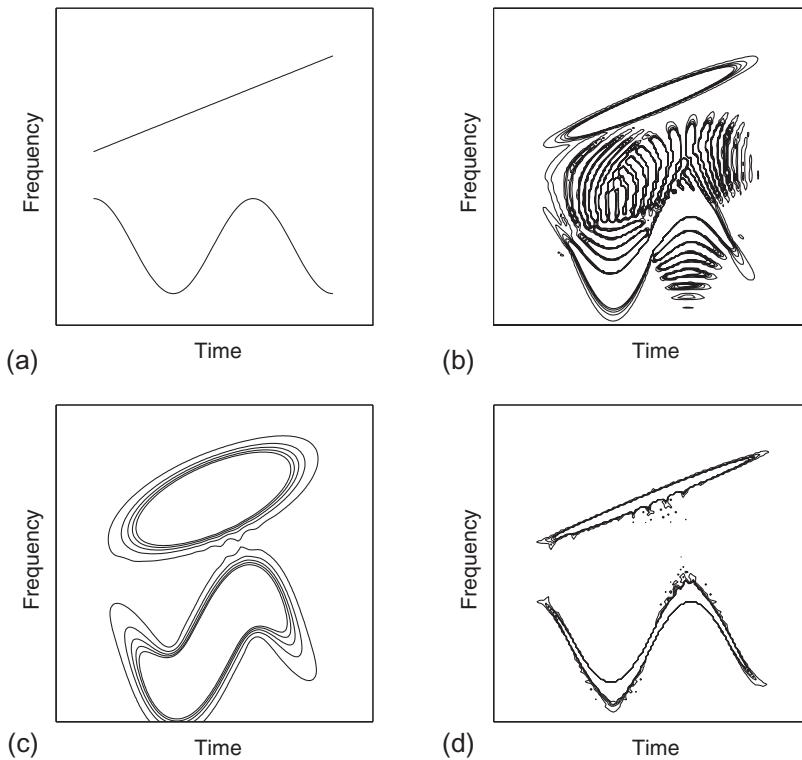


FIGURE 7.2.1

Cross-term level and localization trade-off in Cohen's class. We consider here the (t, f) representation of a signal composed of two different chirps whose instantaneous frequencies are detailed in (a). The kernel of the (t, f) distribution in Cohen's class defines the degree of smoothing which is applied [13]: a weak smoothing favors a sharp localization at the expense of the presence of cross-terms (an example is the WVD displayed in a contour plot in (b)). Conversely, a stronger one leads not only to a lower cross-term level but also to a poorer localization (such as the spectrogram in (c)). Thanks to its smoothing-squeezing scheme, the reassignment method overcomes this trade-off as shown in (d) with the reassigned spectrogram.

whereas the reassignment operator in scale requires an intermediate step in the frequency domain

$$\hat{a}_x(t, a) = \frac{f_0}{\hat{f}_x(t, a)}, \quad \text{with} \quad f_0 = \iint f \Pi(t, f) dt df \quad (7.2.20)$$

$$\text{and} \quad \hat{f}_x(t, a) = \frac{1}{\Omega_x^\Pi(t, a)} \iint v W_x(u, v) \Pi\left(\frac{t-u}{a}, av\right) du dv.$$

The most important case among this class is the scalogram (the squared modulus of the wavelet transform) [13], obtained by choosing for Π the WVD of the chosen wavelet. Simple and efficient expressions of the reassignment operators also exist in this case [15,16].

7.2.2 VARIATIONS AND RELATED APPROACHES

7.2.2.1 Two variations

The original purpose of the reassignment principle was the design of (t, f) distributions with increased readability. But some useful information on the signal structure can also be directly extracted from the reassignment operators, as shown by the following two extensions:

Signal/noise discrimination and supervised reassignment. When the analyzed signal includes broadband noise, the reassignment process yields peaked areas in noise-only regions, whereas rather smooth energy distributions are expected there. For such situations, an improved reassignment algorithm referred to as *supervised reassignment* [17] has been designed. This approach first attempts to discriminate between “signal + noise” and “noise only” regions in the (t, f) plane by means of a detector applied to the reassignment operators. Reassignment is then only performed for the points considered to belong to “signal + noise” regions.

Signal component extraction and differential reassignment. Many signal processing problems such as denoising and signal classification can be solved by a relevant tiling of the (t, f) plane so as to isolate each signal “component” (although this concept is not clearly defined). For such applications, another reassignment process called *differential reassignment* [18] has been considered. Whereas the original reassignment principle moves each value by one finite jump, differential reassignment considers each (t, f) point as the starting point of an elementary particle whose velocity field is deduced from the reassignment operators. The final points called *asymptotic reassignment points* are gathered and lead to a (t, f) map in which each cell indicates a signal component.

7.2.2.2 Related approaches

Although original in many respects, the concept of reassignment is obviously connected with several approaches that have been proposed independently. We lack space to discuss these interactions precisely, but we cite:

- The instantaneous frequency density [19], which yields at each time sample a histogram of the frequency reassignment operator of the spectrogram.
- The extraction of ridges and skeletons out of the phase structure of the wavelet transform [20,21]. These ridges are made of the fixed points of the reassignment operators, either horizontally ($\hat{a}_x(t, a) = a$) or vertically ($\hat{t}_x(t, a) = t$).
- The synchrosqueezed plane [22], which also moves the scalogram values, but by a scale displacement only. A comparison between reassignment and synchrosqueezing may be found in Ref. [23].

7.2.3 SUMMARY AND CONCLUSIONS

Reassignment can be viewed as the second step of a process whose goal is to build a readable (t, f) representation. It consists of:

1. a *smoothing*, whose main purpose is to rub out oscillatory interferences, but whose drawback is to smear localized components;
2. a *squeezing*, whose effect is to refocus the contributions which survived the smoothing.

As a result, this approach yields—without a drastic increase in computational complexity—enhanced contrast (when compared to smoothed distributions such as spectrograms) with a much reduced level of interferences (when compared to the WVD). This is especially true when the signal-noise ratio is not too low, and when the signal components are not “too close” to each other.

7.3 MEASURING TIME-FREQUENCY DISTRIBUTIONS CONCENTRATION⁰

Efficient measurement of concentration of TFDs can provide a quantitative criterion for evaluation of their performance. It can be used for adaptive and automatic parameter selection in time-frequency (t, f) analysis, without supervision of a user. Measures for distribution concentration of monocomponent signals date back to Refs. [24,25]. For more complex signals, some quantities from statistics and information theory were the inspiration for defining measures of TFD concentration [26,27]. They provided good quantitative measures of the auto-terms’ concentration. Various and efficient modifications are used to account for the appearance of oscillatory cross-terms.

The application of concentration measures will be demonstrated on automatic determination of the “best window length” for the spectrogram or the “best number of terms” in the method that provides transition from the spectrogram toward the pseudo Wigner distribution (pseudo WD) (Section 6.2).

7.3.1 CONCENTRATION MEASUREMENT

The basic idea for measuring TFD concentration can be explained with a simple example motivated by probability theory. Consider a set of N non-negative numbers $p_1, p_2, \dots, p_N \geq 0$ such that $p_1 + p_2 + \dots + p_N = 1$ (the *unit-sum* constraint). Form the test function $M(p_1, p_2, \dots, p_N) = p_1^2 + p_2^2 + \dots + p_N^2$. It is easy to conclude that under the unit-sum constraint, the test function is minimized if $p_1 = p_2 = \dots = p_N = 1/N$, that is, if the sequence p_1, p_2, \dots, p_N is maximally spread. Under the same constraint, the test function is maximized if only one p_i is nonzero, that is, if $p_i = \delta(i - i_0)$, where i_0 is an arbitrary integer in the range $1 \leq i_0 \leq N$. In this case the sequence p_1, p_2, \dots, p_N is maximally concentrated, at a single $p_{i_0} = 1$. Therefore, the function $M(p_1, p_2, \dots, p_N)$ can be used as a measure of concentration of the set of numbers p_1, p_2, \dots, p_N , under the unity sum constraint.¹ In general, the constraint can be included in the function itself by using the form $M(p_1, p_2, \dots, p_N) = (p_1^2 + p_2^2 + \dots + p_N^2) / (p_1 + p_2 + \dots + p_N)^2$. For non-negative p_1, p_2, \dots, p_N , this function has its minimum for $p_1 = p_2 = \dots = p_N$, and reaches its maximum when only one p_i is nonzero.

In (t, f) analysis, this idea has been used to measure the concentration. Several forms of the concentration measure, based on this fundamental idea, are introduced.

⁰Author: L.J. Stanković, University of Montenegro, Podgorica, Montenegro (ljubisa@ac.me). Reviewers: W.J. Williams and P. Flandrin.

¹In probability theory, the famous Shannon entropy $-\sum_i p_i \log(p_i)$ is commonly used for the same purpose. It produces the maximal value for the lowest concentration of probabilities p_i , $p_1 = p_2 = \dots = p_N = 1/N$, and the minimal value for the highest concentration $p_i = \delta(i - i_0)$.

1. *Measure based on the ratio of norms:* For the WD of energy-normalized signals, the relation $\sum_n \sum_k \rho_x^2(n, k) \equiv 1$ holds. Therefore, the substitution $p_i \rightarrow \rho_x^2(n, k)$, in the basic example, gives a function that can be used for measuring the concentration of the time-frequency representation (TFR) $\rho_x(n, k)$:

$$M_{\text{JP}} = (L_4/L_2)^4 = \frac{\sum_n \sum_k \rho_x^4(n, k)}{\left[\sum_n \sum_k \rho_x^2(n, k)\right]^2}. \quad (7.3.1)$$

This form is just the fourth power of the ratio of the L_4 and L_2 norms of $\rho_x(n, k)$.² It has been introduced by Jones and Parks in Ref. [26]. They have used the magnitude of the signal's short-time Fourier transform (STFT) as the TFR $\rho_x(n, k)$ in Eq. (7.3.1). High values of M_{JP} indicate that the representation $\rho_x(n, k)$ is highly concentrated, and vice versa. In general, any other ratio of norms L_p and L_q , where $p > q > 1$, can also be used for measuring the concentration of $\rho_x(n, k)$ [26].

When there are two or more components (or regions in the time-frequency plane for a single component) of approximately equal energies (importance), whose concentrations are very different, the norm-based measures will favor the distribution with a "peaky" component, due to raising of distribution values to a high power. It means that if one component (region) is "extremely highly" concentrated, and all the others are "very poorly" concentrated, then the measure will not look for a trade-off in which all components are "well" concentrated. To deal with this kind of problem, common in (t, f) analysis, a concentration measure could be applied to smaller local time-frequency regions [26]:

$$M_{\text{JPL}}(n, k) = \frac{\sum_m \sum_l Q^2(m - n, l - k) \rho_x^4(m, l)}{\left[\sum_m \sum_l Q(m - n, l - k) \rho_x^2(m, l)\right]^2}. \quad (7.3.2)$$

The localization weighting function $Q(n, k)$ determines the region where the concentration is measured. In Ref. [26], the Gaussian form of this function is used.

2. *Rényi entropy-based measures:* The second class of TFD measures, defined by analogy with the Rényi entropy, was introduced in (t, f) analysis by Williams et al. [27,28], with significant contributions by others [29,30] in establishing the properties of the measure. The Rényi entropy, applied on the TFD $\rho_x(n, k)$, has the form

$$R_\alpha = \frac{1}{1 - \alpha} \log_2 \left[\sum_n \sum_k \rho_x^\alpha(n, k) \right] \quad (7.3.3)$$

with $\alpha > 2$ being recommended for the TFD measures [30]. For $\alpha = 2$ and the WD of energy-normalized signals ($\sum_n \sum_k \rho_x^2(n, k) \equiv 1$), we have $R_2 = 0$ for all signals. Note that the logarithm is a monotone function. Thus, the behavior of R_α is determined by the behavior of the argument $\sum_n \sum_k \rho_x^\alpha(n, k)$, as explained at the beginning of this section. Unlike the measure Eq. (7.3.1), the entropy Eq. (7.3.3) has larger values for *less* concentrated distributions due to a negative coefficient $1/(1 - \alpha)$ for $\alpha > 2$. This is the case for all measures presented in the remainder of this section.

²In statistics, a similar form (known as kurtosis) is used as a measure of the flatness or peakedness of a distribution. Kurtosis is 0 for a Gaussian distribution. Values greater than 0 mean that the distribution has more of a peak than a Gaussian distribution, while values less than 0 mean flatter distributions.

It is interesting to note [30] that the *Shannon entropy*

$$H = - \sum_n \sum_k [\rho_x(n, k) \log_2 \rho_x(n, k)]$$

is a limiting case of the Rényi entropy as $\alpha \rightarrow 1$. The Shannon entropy cannot be used for general TFDs $\rho_x(n, k)$, which can assume negative values [30].

- 3.** *Normalized Rényi entropy measures.* To avoid the problem caused by the Rényi-entropy-based measure with $\alpha = 3$ ignoring the presence of oscillatory cross-terms (when the auto-terms are well separated), some kind of normalization should be done. It can be done in various ways, leading to a variety of possible measure definitions [27].

Normalization with the distribution volume is performed as:

$$\text{RV}_3 = -\frac{1}{2} \log_2 \sum_n \sum_k \left(\rho_x(n, k) \middle/ \sum_n \sum_k |\rho_x(n, k)| \right)^3. \quad (7.3.4)$$

If the distribution contains oscillatory values, then summing their absolute values means that large cross-terms will decrease the measure RV_α . This is the expected behavior of a measure, since it will seek for a balance between the cross-terms suppression and auto-terms enhancement.

The volume normalized form of measure has been used for adaptive kernel design in Ref. [27].

- 4.** Concentration measures based on the ratio of higher-order norms dominated for decades in signal processing. However, the concentration measures with lower-order norms, introduced in Ref. [31], are mainly used, as a result of their suitability for measuring signal sparsity, which is the crucial parameter in compressive sensing [32–36]. The key concentration measure used in realizations in this field is based on the norm-one, while the theory is developed based on the norm-zero concentration measure. The basic idea for the measure that will be presented next comes from an obvious *classical definition of the time-limited signal duration*. If a signal $x(n)$ is time-limited to the interval $n \in [n_1, n_2 - 1]$, that is, $x(n) \neq 0$ only for $n \in [n_1, n_2 - 1]$, then the duration of $x(n)$ is $d = n_2 - n_1$. It can be written as $d = \lim_{p \rightarrow \infty} \sum_n |x(n)|^{1/p}$. This is the so-called norm-zero form of the concentration (sparsity) measure. The same definition applied to a 2D function $\rho_x(n, k) \neq 0$ only for $(n, k) \in D_x$, gives

$$N_D = \lim_{p \rightarrow \infty} \sum_n \sum_k |\rho_x(n, k)|^{1/p}, \quad (7.3.5)$$

where N_D is the number of points within D_x . In reality, there is no sharp edge between $\rho_x(n, k) \neq 0$ and $\rho_x(n, k) = 0$, so the value of Eq. (7.3.5) could, for very large p , be sensitive to small values of $\rho_x(n, k)$. The robustness may be achieved by using lower order forms, for example, with $p = 2$. Therefore, the concentration can be measured with the function of the form

$$M_p^p = \left(\sum_n \sum_k |\rho_x(n, k)|^{1/p} \right)^p, \quad (7.3.6)$$

with $\sum_n \sum_k \rho_x(n, k) = 1$, and $p > 1$. When $p \rightarrow \infty$, the norm-zero-based measure is obtained. If, for example, the distribution is the squared STFT (spectrogram), then the case with $p = 2$ corresponds to the norm-one measure of the signal transform.

After we have presented several possible forms for measuring the concentration of TFDs, we can summarize *a procedure for constructing a TFD measure* based on definitions from 1D classical signal analysis, or definitions from probability, quantum mechanics, or information theory:

- (i) In the classical signal analysis definitions, consider the signal power $|x(t)|^2$ (or spectral energy density $|X(f)|^2$) as the probability density function in time (or frequency). This idea comes from quantum mechanics, where the absolute square of the wave function is the position's probability density function.
- (ii) Assume that the TFD $\rho_x(t, f)$ can be treated as a joint 2D probability density function.
- (iii) According to these assumptions, reintroduce the 1D definition into the joint 2D time-frequency domain.
- (iv) Additional modifications, interpretations, constraints, and normalizations are needed in order to get forms that can be used in (t, f) analysis. For example, several possible forms of the Rényi entropy measure in the time-frequency domain have been proposed and used in various problems.

Example. Consider the classic *Leipnik entropy measure* [25], and *Zakai's entropy parameter* $\delta_t = - \int_{-\infty}^{\infty} |x(t)|^2 \ln |x(t)|^2 dt$ of the signal $x(t)$ [25]. According to the procedure for constructing a time-frequency form, based on a classical signal processing relation, we get

$$\delta_t = - \int_{-\infty}^{\infty} |x(t)|^2 \ln |x(t)|^2 dt \rightarrow - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x(t, f) \ln \rho_x(t, f) dt df. \quad (7.3.7)$$

This is exactly the well-known *Shannon entropy*. It has already been discussed in Ref. [30] with respect to its (non)applicability in (t, f) problems. Similarly, a logarithm of the general Zakai's signal duration (uncertainty)

$$Z_\alpha = \log_2 T_{2\alpha} = \frac{1}{1-\alpha} \log_2 \frac{\int_{-\infty}^{\infty} |x(t)|^{2\alpha} dt}{\left(\int_{-\infty}^{\infty} |x(t)|^2 dt\right)^\alpha},$$

according to the proposed procedure, transforms into the Rényi entropy measure,

$$Z_\alpha \rightarrow \frac{1}{1-\alpha} \log_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x^\alpha(t, f) dt df = R_\alpha,$$

where $|x(t)|^2$ has been replaced by $\rho_x(t, f)$, and the unit signal energy is assumed.

Remark. In the probability theory, all results are derived for the probability values p_i , assuming that $\sum_i p_i = 1$ and $p_i \geq 0$. The same assumptions are made in classical signal analysis for the signal power. Since a general TFD commonly does not satisfy both $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x(t, f) dt df = 1$ and $\rho_x(t, f) \geq 0$, the obtained measures of TFD concentration may just *formally look like* the original entropies or classical signal-analysis forms, while they can have different behavior and properties.³

³Quantum-mechanical forms can also be used for the definition of highly concentrated signal representations. One of them is the "pseudo quantum" signal representation [37] in the form of $SD_x(t, \wp) = \int_{-\infty}^{\infty} x^{[L]}(t + \tau/(2L)) x^{*[L]}(t - \tau/(2L)) e^{-j\wp\tau} d\tau$, with $x^{[L]}(t) = A(t) \exp(jL\phi(t))$ for $x(t) = A(t) \exp(j\phi(t))$. For example, for $x(t) = A \exp(-at^2/2 + jbt^2/2 + jct)$ we get $SD_x(t, \wp) = A^2 \exp(-at^2) \sqrt{4\pi/(a/L^2)} \exp[-(\wp - bt - c)^2/(a/L^2)]$. For $a/L^2 \rightarrow 0$ it results in $SD(t, \wp) = 2\pi A^2 \exp(-at^2) \delta(\wp - bt - c)$, which is just an ideally concentrated distribution along the instantaneous frequency. For a large a , if L^2 is large enough so that $a/L^2 \rightarrow 0$, we get the distribution highly concentrated in a very small region around the point $(t, \wp) = (0, c)$.

7.3.2 NUMERICAL EXAMPLES

Consider the spectrogram

$$S_x^w(n, k) = |F_x^w(n, k)|^2 / E,$$

where $F_x^w(n, k) = \text{DFT}_{m \rightarrow k}\{w(m)x(n+m)\}$ is the short-time Fourier transform (STFT); E is the energy of the lag window $w(m)$. Among several spectrograms, calculated with different window lengths or forms, the best according to one of the proposed concentration measures, denoted by $\mathcal{M}[\rho_x(n, k)]$, will be that which minimizes (or maximizes, depending on the used measure form):

$$w^+ = \arg \min_w \{\mathcal{M}[S_x^w(n, k)]\}. \quad (7.3.8)$$

Let us illustrate this by an example. Consider the signal

$$x(t) = \cos(50 \cos(\pi t) + 10\pi t^2 + 70\pi t) + \cos(25\pi t^2 + 180\pi t) \quad (7.3.9)$$

sampled at $\Delta t = 1/256$, within $-1 \leq t < 1$ normalized to 1 s. The Hanning window $w(m)$ with different lengths is used in the spectrogram calculation. Here, we have used the measure Eq. (7.3.6) with $p = 2$ (norm-one of the STFT measure), although for this signal all presented measures would produce similar results [31]. Note that the presented measures would significantly differ if, for example, the second component were the pure sinusoid $\cos(180\pi t)$ instead of $\cos(25\pi t^2 + 180\pi t)$.

For wide lag windows, signal nonstationarity makes the spectrogram spread in the time-frequency plane, having relatively large measure $\mathcal{M}[S_x^w(n, k)] = M_2^2$ (Fig. 7.3.1(a) and (b)). For narrow lag windows, its Fourier transform is very wide, causing spread distributions and large M_2^2 (Fig. 7.3.1(d) and (e)). Obviously, between these two extreme situations there is a window that produces an acceptable trade-off between the signal nonstationarity and small-window-length effects. The measure M_2^2 is calculated for a set of spectrograms with window length N ranging from 32 to 256 (Fig. 7.3.1(f)). The minimal measure value, meaning the best concentrated spectrogram according to this measure, is achieved for $N = 88$. The spectrogram with $N = 88$ is shown in Fig. 7.3.1(c).

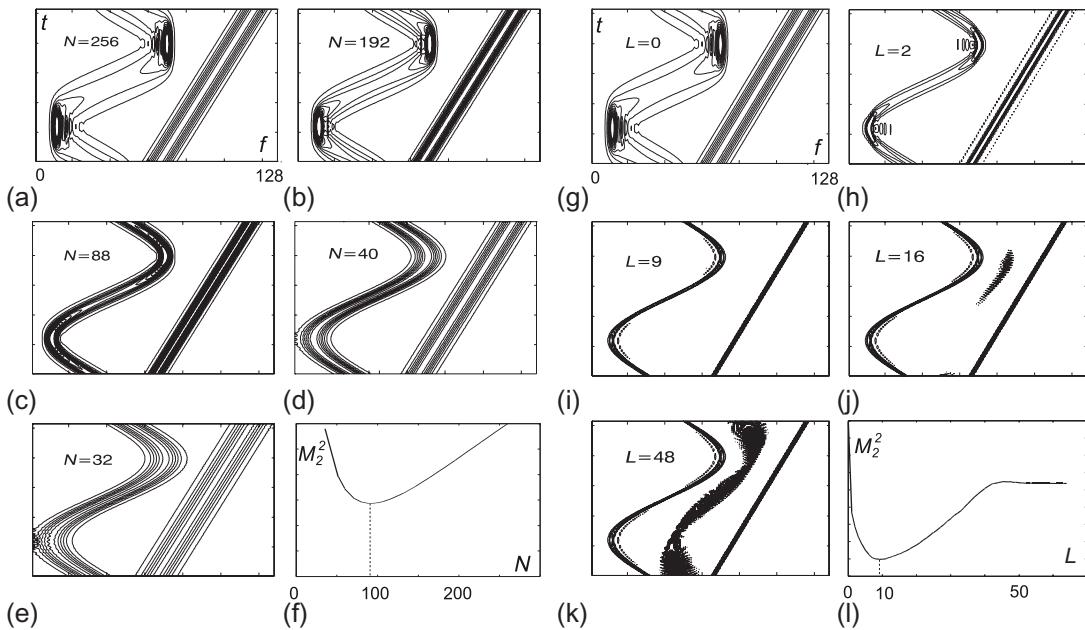
The same procedure will be used for determination of the optimal number of terms L , in a transition from the spectrogram to the pseudo WD, according to the recursive form of the S-method (SM) (Section 6.2):

$$\text{SM}_x(n, k; L) = \text{SM}_x(n, k; L - 1) + 2 \operatorname{Re} \{F_x(n, k + L) F_x^*(n, k - L)\}, \quad (7.3.10)$$

where $\text{SM}_x(n, k; 0) = |F_x^w(n, k)|^2$, and both $k + L$ and $k - L$ are within the basic frequency period. Note that $\text{SM}(n, k; N/2)$ is equal to the pseudo WD. The optimal distribution $\text{SM}_x^+(n, k; L)$, on our way from $L = 0$ (the spectrogram) toward $L = N/2$ (the pseudo WD), is the one calculated with L producing the minimal value of $\mathcal{M}[\text{SM}_x(n, k; L)]$,

$$L^+ = \arg \min_L \{\mathcal{M}[\text{SM}_x(n, k; L)]\}. \quad (7.3.11)$$

Here, instead of $|\text{SM}_x(n, k; L)|$, a non-negative part of $\text{SM}_x(n, k; L)$ will be used. Distributions $\text{SM}_x(n, k; L)$ should be properly scaled to satisfy the unbiased energy condition. The same signal is used for the illustration of the SM. Since this method is based on the WD, the best results will be achieved with a wide lag window in the STFT calculation, $N = 256$. The spectrogram ($L = 0$) is shown in

**FIGURE 7.3.1**

(a)-(e) Spectrogram for various window lengths, and (f) its measure $\mathcal{M}[S_x^W(n, k)] = M_2^2$. The lowest M_2^2 is achieved for $N = 88$, which is the best window length choice according to this measure. (g)-(k) The S-method for various values of parameter L , and (l) its measure $\mathcal{M}[\text{SM}_x(n, k; L)] = M_2^2$. The lowest M_2^2 is obtained for $L = 9$.

[Fig. 7.3.1\(g\)](#). By increasing L , the SM improves concentration of the spectrogram toward the pseudo WD quality, meaning lower measure $\{\mathcal{M}[\text{SM}_x(n, k; L)]\} = M_2^2$ ([Fig. 7.3.1\(h\)](#) and [\(i\)](#)). After L has reached the value equal to the distance between the auto-terms, cross-terms start to appear, increasing M_2^2 ([Fig. 7.3.1\(j\)](#) and [\(k\)](#)). Minimal M_2^2 means a trade-off between the auto-terms' concentration and the cross-terms' appearance ([Fig. 7.3.1\(k\)](#)). The SM with L corresponding to minimal M_2^2 is shown in [Fig. 7.3.1\(l\)](#).

The concentration measure is illustrated on (t, f) analysis of a pressure signal in a BMW engine with speed 2000 rev/min, in [Fig. 7.3.2](#) (cf. [Section 15.2](#)).

7.3.3 PARAMETER OPTIMIZATION

Parameter optimization may be done by a straightforward computation of a distribution measure $\mathcal{M}[\rho_x(n, k)]$, for various parameter values. The best choice according to this criterion (optimal distribution with respect to this measure) is the distribution which produces the minimal value of $\mathcal{M}[\rho_x(n, k)]$. In the cases when one has to consider a wide region of possible parameter values for the distribution calculation (like, for example, window lengths in the spectrogram), this approach can be numerically

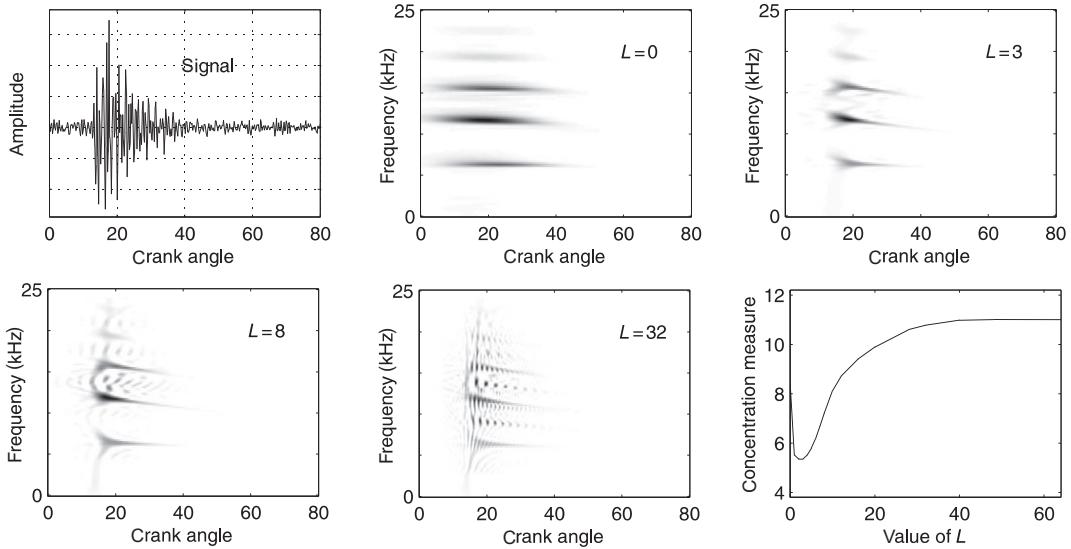


FIGURE 7.3.2

Concentration measure illustration on time-frequency analysis of a car-engine pressure signal, showing the signal (top left), its SM-based TFRs for various values of the parameter L , and the concentration measure vs L (bottom right) indicating that the best choice was $L = 3$. Time is represented by the corresponding crank-angle.

inefficient. Then, some more sophisticated optimization procedures, like the one using the steepest descent approach described in Ref. [27], can be used. Its simplified version will be presented here [31].

The gradient of a measure $\mathcal{M}[\rho_x(n, k)]$, with respect to a distribution's generalized optimization parameter denoted by ξ , is

$$\frac{\partial \mathcal{M}[\rho_x(n, k)]}{\partial \xi} = \frac{\partial \mathcal{M}[\rho_x(n, k)]}{\partial \rho_x(n, k)} \frac{\partial \rho_x(n, k)}{\partial \xi}.$$

Iterations, starting from a very low concentrated distribution toward the maximally concentrated one, that is, toward the measure minimum, can be done according to

$$\xi_{m+1} = \xi_m - \mu \frac{\partial \mathcal{M}[\rho_x(n, k)]}{\partial \xi}, \quad (7.3.12)$$

where μ is the step, which should be chosen in the same way as the step in the other adaptive algorithms. The step should not be too small (since the convergence would be too slow), and not too large (so as to miss the minimum, or cause divergence).

In discrete implementations, the gradient $\partial \mathcal{M}[\rho_x(n, k)]/\partial \xi$ can be approximated based on $\mathcal{M}[\rho_x(n, k; \xi_m)]$ calculated with ξ_m and its previous value ξ_{m-1} :

$$\xi_{m+1} = \xi_m - \mu \frac{\mathcal{M}[\rho_x(n, k; \xi_m)] - \mathcal{M}[\rho_x(n, k; \xi_{m-1})]}{\xi_m - \xi_{m-1}}. \quad (7.3.13)$$

Example. The optimization procedure will be illustrated on the signal $x(t)$ described in Eq. (7.3.9), its spectrogram, and the measure form 2. The optimal window length is obtained in few iterations by using Eq. (7.3.13), starting from the very narrow window. Initial values of $\xi_0 = N = 16$ and $\xi_1 = N = 20$ are assumed. The next value of $\xi_{m+1} \equiv N$ is calculated according to Eq. (7.3.13). During the iterations, we get $\xi_m = 16, 20, 76$, and 90. The algorithm is stopped at $\xi_m = 90$, when $|\xi_{m+1} - \xi_m| < 2$, since an even number of samples is used in the realization. Note that the obtained optimal value is within ± 2 of the value obtained by direct calculation. The parameter value $\mu = 1/3$ has been used in all examples.

7.3.4 SUMMARY AND CONCLUSIONS

Measurement of TFD concentration, with application to an automatic optimization of distribution parameters, is presented. It is based on the forms borrowed from classical signal analysis, probability, or information theory, with appropriate interpretations and adjustments. The presented concentration measures can be used not only to optimize the transformation parameters but also to recover missing or corrupted signal samples, by minimizing the appropriate concentration measure, within the compressive sensing framework [32–36].

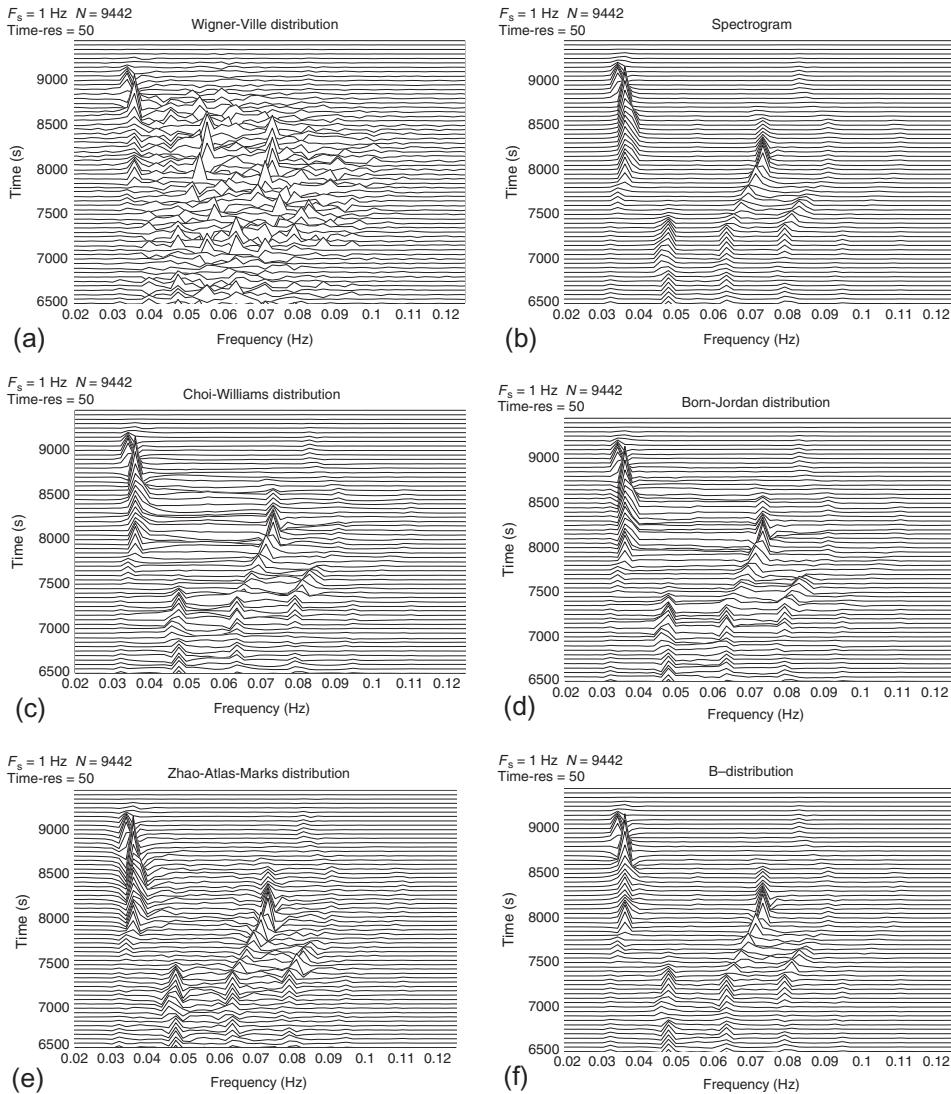
7.4 RESOLUTION PERFORMANCE ASSESSMENT FOR COMPARING AND SELECTING QUADRATIC TFD⁰

7.4.1 SELECTING AND COMPARING TFDs

Quadratic time-frequency distributions (TFDs) are effective tools for extracting information from a nonstationary signal, such as the number of components and their durations, bandwidths, relative amplitudes, and instantaneous frequency (IF) laws. Performance of these TFDs depends on the type of signal (as detailed in Ref. [38] and Chapter 3). For example, in the case of a monocomponent linear frequency modulated (FM) signal, the Wigner-Ville distribution (WVD) is known to be optimal in the sense that it achieves the best energy concentration around the IF law (Chapter 2). In applications involving multicomponent signals, choosing the right TFD to analyze the signals is an immediate critical task for the analyst [39,40]. A methodology for making this choice and assessment, using current knowledge, is the subject of this section.

Consider, for example, a multicomponent whale signal, represented in the (t, f) domain using the WVD, spectrogram, exponential distribution, Born-Jordan distribution, Zhao-Atlas-Marks distribution, and B-distribution (see Fig. 7.4.1; for definitions and properties of these TFDs, see Chapter 3). To determine which of the TFDs in Fig. 7.4.1 “best” represents the whale signal (i.e., which one gives the best component-energy concentration and best interference-term suppression, and allows the most accurate estimation of the components’ IF laws), one could *visually* compare the six plots and choose the most appealing one. The spectrogram and the B-distribution, being almost free from the cross-terms, seem to perform best.

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **V. Sucic**, University of Rijeka, Rijeka, Croatia (vsucic@riteh.hr). Reviewers: W.J. Williams and LJ. Stanković.

**FIGURE 7.4.1**

TFDs of a multicomponent whale signal. (a) Wigner-Ville, (b) spectrogram (Hanning, $L = 511$), (c) exponential ($\sigma = 10$), (d) Born-Jordan, (e) Zhao-Atlas-Marks ($a = 2$), and (f) B-distribution ($\beta = 0.01$).

The performance comparison based on the visual inspection of the plots becomes more difficult and unreliable; however, when the signal components are closely spaced in the (t, f) plane. To objectively compare the plots in Fig. 7.4.1 requires a quantitative performance measure for TFDs. There have been several attempts to define objective measures of “complexity” for TFDs (see Sections 7.3 and 7.7). One

of these measures, the Rényi entropy, has been modified to estimate the local number of components present in the signal (details are given in Ref. [41] and Section 7.7).

The performance measure described in this section is a local measure of the TFD resolution, and is thus suited to the selection problem shown by Fig. 7.4.1. The measure takes into account the characteristics of TFDs that influence their resolution, such as energy concentration, component separation, and interference minimization. Methodologies for choosing a TFD which best suits a given signal are developed below by optimizing the resolution performance of considered TFDs and modifying their parameters to better meet application-specific requirements.

7.4.2 PERFORMANCE CRITERIA FOR TFDs

In the case of *monocomponent* FM signals, the best TFD is the one that maximizes energy concentration about the IF. This is achieved by minimizing the component-sidelobe (inner-artifact) amplitude A_s relative to the mainlobe amplitude A_m , and the mainlobe bandwidth B relative to the central frequency f (see Fig. 7.4.2). The instantaneous concentration performance of a TFD may thus be quantified by the measure p expressed as

$$p(t) = \left| \frac{A_s(t)}{A_m(t)} \right| \frac{B(t)}{f(t)}. \quad (7.4.1)$$

A good performance is characterized by a small value of the measure p . For example, for the WVD of a linear FM signal with infinite duration, both bandwidth B and sidelobe amplitude A_s are zero so that we obtain $p = 0$.

For *multicomponent* FM signals, the performance of a TFD can be *quantitatively* assessed in terms of:

- energy concentration of the TFD about the IF of each component, as expressed by Eq. (7.4.1), and
- component resolution, as measured by the frequency separation of the components' mainlobes, including the effect of cross-terms (outer artifacts).

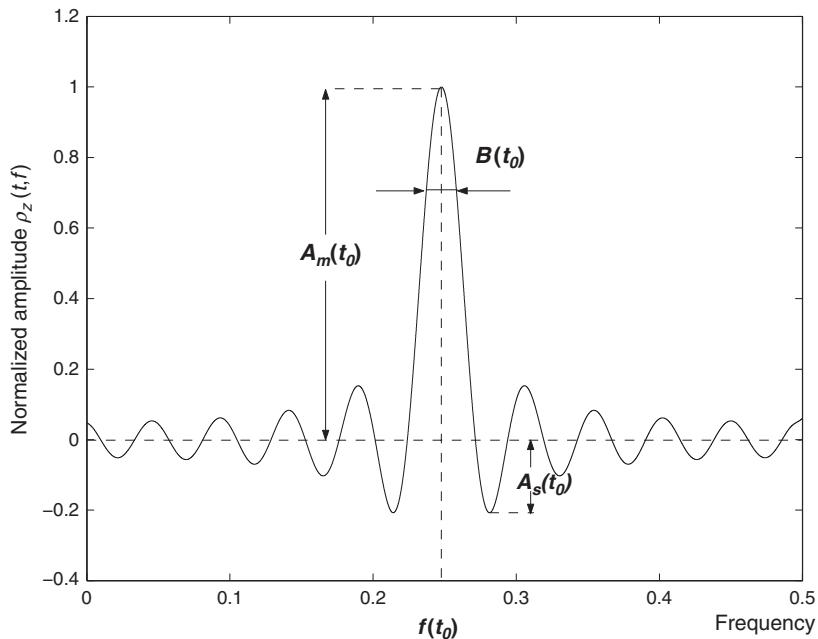
The frequency resolution in a spectrum estimate of a stationary signal composed of two tones, f_1 and f_2 (see Fig. 7.4.3), may be defined as the minimum difference $f_2 - f_1$ for which the following inequality holds:

$$f_1 + B_1/2 < f_2 - B_2/2, \quad f_1 < f_2, \quad (7.4.2)$$

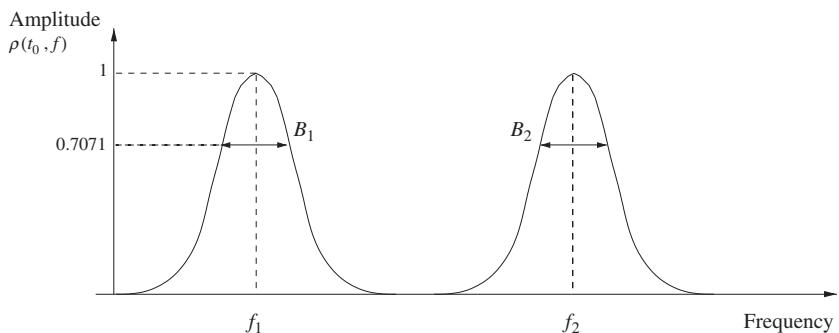
where B_1 and B_2 are the respective bandwidths of the two components.

For a TFD $\rho_z(t, f)$ of a two-component *nonstationary* signal, the above definition of resolution is valid for every time slice of a cross-term-free TFD, such as the spectrogram. However, for other TFDs, we need to take into account the effect of cross-terms on resolution.

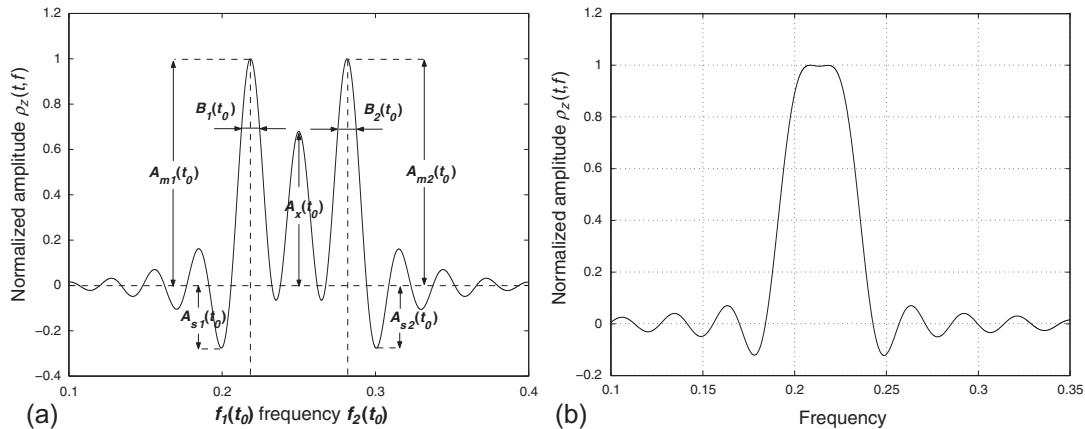
Figure 7.4.4(a) shows a time-slice of a typical quadratic TFD, with the components being clearly resolved, where $B_1(t_0)$, $f_1(t_0)$, $A_{s1}(t_0)$, and $A_{m1}(t_0)$ represent, respectively, the instantaneous bandwidth, IF, sidelobe amplitude, and mainlobe amplitude of the first component at time $t = t_0$. Similarly, $B_2(t_0)$, $f_2(t_0)$, $A_{s2}(t_0)$, and $A_{m2}(t_0)$ represent the instantaneous bandwidth, IF, sidelobe amplitude, and mainlobe amplitude of the second component at the same time $t = t_0$. The amplitude $A_x(t_0)$ is for the cross-term. An example of nonresolved components is shown in Fig. 7.4.4(b), where the two components and the cross-term have merged into a single lobe.

**FIGURE 7.4.2**

Time-slice of a TFD of a monocomponent signal at $t = t_0$. The dominant peak is the component, while the other peaks are the sidelobes. For clarity of presentation, we measure the mainlobe bandwidth at $1/\sqrt{2}$ of the component's normalized amplitude A_m .

**FIGURE 7.4.3**

Time-slice of a TFD of a two-component signal in the absence of cross-terms. The lobes are clearly distinguishable from each other; hence, the components are said to be resolved.

**FIGURE 7.4.4**

TFD time-slices of two-component signals with resolved and unresolved components. (a) *Resolved* components. (b) *Unresolved* components.

7.4.3 RESOLUTION PERFORMANCE MEASURE FOR TFDs

Equation (7.4.2) and Fig. 7.4.4(a) suggest that the resolution performance of a TFD for a neighboring pair of components may be defined by the minimum value of the IF separation, $D(t) = f_2(t) - f_1(t)$, for which we have a positive frequency separation measure $S(t)$ between the component mainlobes centered about their respective IFs, $f_1(t)$ and $f_2(t)$. For best performance, $S(t)$ should be as close as possible to $D(t)$, the true difference between the IFs. The component separation measure $S(t)$ is expressed as [42]:

$$S(t) = [f_2(t) - B_2(t)/2] - [f_1(t) + B_1(t)/2]. \quad (7.4.3)$$

The resolution also depends on the following set of variables, all of which should be as small as possible:

- (a) the normalized *instantaneous bandwidth* of the signal component, $B_k(t)/f_k(t)$, $k = 1, 2$, which is accounted for by $S(t)$ (Eq. 7.4.3);
- (b) the ratio of the *sidelobe amplitude* $|A_{s_k}(t)|$ to the *mainlobe amplitude* $|A_{m_k}(t)|$ of each component, where $k = 1, 2$; and
- (c) the ratio of the *cross-term amplitude* $|A_x(t)|$ to the *mainlobe amplitude* $|A_{m_k}(t)|$ of each component, where $k = 1, 2$.

It follows that the best TFD for multicomponent signal analysis is the one that concurrently *minimizes* the positive quantities (a), (b), (c), and *maximizes* $S(t)$.

Hence, by combining the above variables, expressions for a measure $P(t)$ of the resolution performance of a given TFD can be defined. Two definitions have been proposed in Ref. [42], one of them being a normalized performance measure

$$P(t) = 1 - \frac{1}{3} \left\{ \left| \frac{A_s(t)}{A_m(t)} \right| + \frac{1}{2} \left| \frac{A_x(t)}{A_m(t)} \right| + \left(1 - \frac{S(t)}{D(t)} \right) \right\}, \quad (7.4.4)$$

where, for a pair of signal components, $A_m(t)$ and $A_s(t)$ are the average amplitudes of the components' mainlobes and sidelobes, respectively; $A_x(t)$ is the cross-term amplitude; $S(t)$ is the separation measure defined by Eq. (7.4.3); and $D(t) = f_2(t) - f_1(t)$ is the difference between the actual IFs. In practice, the parameters in Eq. (7.4.4) can be measured using the algorithm presented in Ref. [43].

The measure $P(t)$ is close to 1 for the best TFD, and 0 for poor ones. Therefore, when designing a TFD we want to maximize $P(t)$ in order to reduce the cross-terms while preserving the component resolution.

In some applications involving real-life signals, we may need to better discriminate between different TFDs' resolution performances in a given set of K TFDs. In this case, a suitable alternative to $P(t)$ is defined in Ref. [42] as:

$$M_j(t) = 1 - \frac{1}{3} \left\{ \frac{\left| \frac{A_{s_j}(t)}{A_{m_j}(t)} \right|}{\max_{1 \leq k \leq K} \left(\left| \frac{A_{s_k}(t)}{A_{m_k}(t)} \right| \right)} + \frac{\left| \frac{A_{x_j}(t)}{A_{m_j}(t)} \right|}{\max_{1 \leq k \leq K} \left(\left| \frac{A_{x_k}(t)}{A_{m_k}(t)} \right| \right)} + \frac{\left(\frac{B_j(t)}{D_j(t)} \right)}{\max_{1 \leq k \leq K} \left(\frac{B_k(t)}{D_k(t)} \right)} \right\}, \quad (7.4.5)$$

where $M_j(t)$, $1 \leq j \leq K$, is the resolution performance measure of the j th TFD, and B is the average instantaneous bandwidth of the components mainlobes. The measure $M(t)$ is applied in Section 7.4.5 to TFDs of real-life signals.

7.4.4 SELECTING THE OPTIMAL TFD FOR A GIVEN MULTICOMPONENT SIGNAL

A methodology for selecting the optimal TFD for resolving closely spaced components in a multicomponent signal involves the following steps:

1. Define a set of comparison criteria describing the information sought from TFDs (as described in Section 7.4.2).
2. Objectively measure the resolution performance of TFDs based on those criteria (use the measure P defined by Eq. (7.4.4)).
3. Optimize each TFD to match the criteria as closely as possible [42,43]: select as the optimal kernel-parameter value the one which maximizes the overall performance measure P_{overall} , taken as the mean of the instantaneous measures P in a time interval of interest.
4. Quantitatively compare TFDs and select the best one: an optimized TFD with the largest value of P_{overall} is selected as the *best* TFD for representing the given signal in the joint (t, f) domain.

Example. We define the following two-component signal in noise:

$$\begin{aligned} s(t) &= s_1(t) + s_2(t) + n(t) \\ &= \cos \left(2\pi \left(0.1 t + \frac{\alpha}{2} t^2 \right) \right) + \cos \left(2\pi \left(0.2 t + \frac{\alpha}{2} t^2 \right) \right) + n(t), \end{aligned} \quad (7.4.6)$$

where $\alpha = 0.0016$ is the component bandwidth-duration ratio (duration $T = 128$), and $n(t)$ is additive white Gaussian noise (AWGN) with signal-to-noise ratio (SNR) = 10 dB. The sampling frequency is $f_s = 1$ Hz.

The signal $s(t)$ is analyzed in the (t, f) domain using the following TFDs: spectrogram, WVD, exponential distribution, Born-Jordan distribution, Zhao-Atlas-Marks distribution, and modified B-distribution (MBD). For definitions and properties of these TFDs refer to [Chapter 3](#).

To find the optimal TFD for resolving the two components $s_1(t)$ and $s_2(t)$, we first find the optimal values of the TFDs' kernel parameters. The WVD and Born-Jordan distribution have no “smoothing” parameters, hence do not need optimizing. The optimized TFDs are plotted in [Fig. 7.4.5](#). The TFD with the largest P_{overall} among the considered TFDs is then selected as best for representing $s(t)$. [Table 7.4.1](#) lists the results of the optimization process, which show that the optimal TFD for the signal is the MBD with parameter $\beta = 0.04$. From the selected TFD, signal parameters are measured (see [Table 7.4.2](#)). In addition, by optimizing component concentration and resolution, more accurate estimates of component IFs can be obtained from the peaks of dominant ridges of the best-performing TFD. IF estimates for the components of the signal $s(t)$ are shown in [Fig. 7.4.6](#).

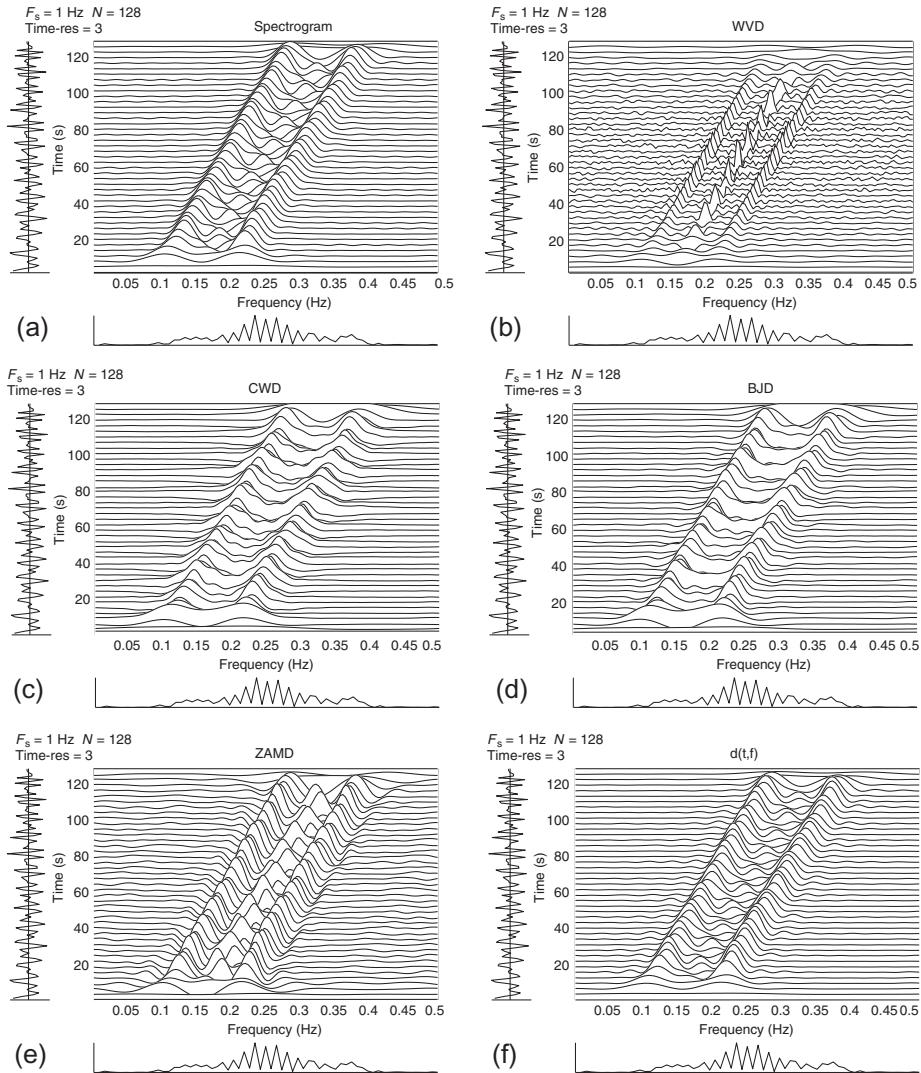
7.4.5 APPLYING THE PERFORMANCE MEASURE IN REAL-LIFE SITUATIONS

The aim is to select the best-performing TFD of a real-life signal in an objective, automatic way. This will make (t, f) techniques more applicable in practice (see Part V of the book for examples of potential engineering and medical applications). The TFD selection methodology consists of the following four steps:

1. Represent the signal in the (t, f) domain with the smoothed WVD. Following the approach described in [Section 5.7](#), we smooth the WVD in both time t and lag τ with the Hanning window of length equal to a quarter of the signal time duration. This (t, f) smoothing is intended to suppress the WVD inner and outer artifacts, while preserving the key (t, f) features of the components.
2. For different time instants of the smoothed WVD, select the two closest dominant peaks in the f direction. For the best resolution of the signal components, it is sufficient to optimize the two closest components' resolution at an observed time instant [44]. Note that if a signal is monocomponent or no components exist at a particular time, this time instant is not considered.
3. For the selected pairs of components, optimize different TFDs using the resolution performance measure M defined by Eq. [\(7.4.5\)](#). The measure M is used instead of P since it is a better discriminator of resolution performance in the case of TFDs of real-life signals [45]. The kernel parameter value which, of a set of different values considered, maximizes the overall performance measure (the mean of M over the observed times) is selected as its optimal value.
4. Calculate the measure M of the optimized TFDs for each of the selected pairs of signal components. The TFD that maximizes the average M (over time) is selected as the signal's best-performing TFD of the considered set of TFDs.

Example. To illustrate how to use this methodology in practice, let us find the best-performing TFD for the Noisy Miner (*Manorina melanocephala*) song signal.

We start by representing the signal in the (t, f) domain with the WVD smoothed in time and lag with Hanning windows of length $L = 3501$ ([Fig. 7.4.7\(a\)](#)). For each time t of the smoothed WVD, we

**FIGURE 7.4.5**

Optimized TFDs of the signal $s(t)$ defined by Eq. (7.4.6). (a) Spectrogram (Bartlett, $L = 31$); (b) Wigner-Ville; (c) exponential ($\sigma = 1$); (d) Born-Jordan; (e) Zhao-Atlas-Marks ($a = 2$); and (f) modified B ($\beta = 0.04$).

identify the pair of closest components. From Fig. 7.4.7(b), we observe that different components form such pairs at different times.

Next, as described in steps 3 and 4 above, the considered TFDs are first optimized, and their resolution performances are evaluated using the measure M . Table 7.4.3 shows the TFDs' optimal

Table 7.4.1: Optimization Results for TFDs of the Signal $s(t)$ Defined by Eq. (7.4.6)

| TFD | Optimal Value of the Kernel Parameter | P_{overall} |
|------------------|---------------------------------------|----------------------|
| Spectrogram | Bartlett window, length 31 | 0.86 |
| Wigner-Ville | N/A in this case | 0.62 |
| Exponential | $\sigma = 1$ | 0.82 |
| Born-Jordan | N/A in this case | 0.81 |
| Zhao-Atlas-Marks | $a = 2$ | 0.67 |
| Modified B | $\beta = 0.04$ | 0.88 |

Table 7.4.2: Parameters of $s_1(t)$ and $s_2(t)$ (see Eq. 7.4.6) Measured From the Optimal TFD (MBD, $\beta = 0.04$) of the Signal $s(t)$

| Parameter | Component $s_1(t)$ | Component $s_2(t)$ |
|------------------------------|--------------------|--------------------|
| Instantaneous bandwidth B | 0.0194 | 0.0195 |
| Mainlobe amplitude $ A_m $ | 1.0002 | 0.9574 |
| Sidelobe amplitude $ A_s $ | 0.0900 | 0.0858 |
| Cross-term amplitude $ A_x $ | | 0.1503 |

Note: The values shown are averages over $t \in [32, 96]$.

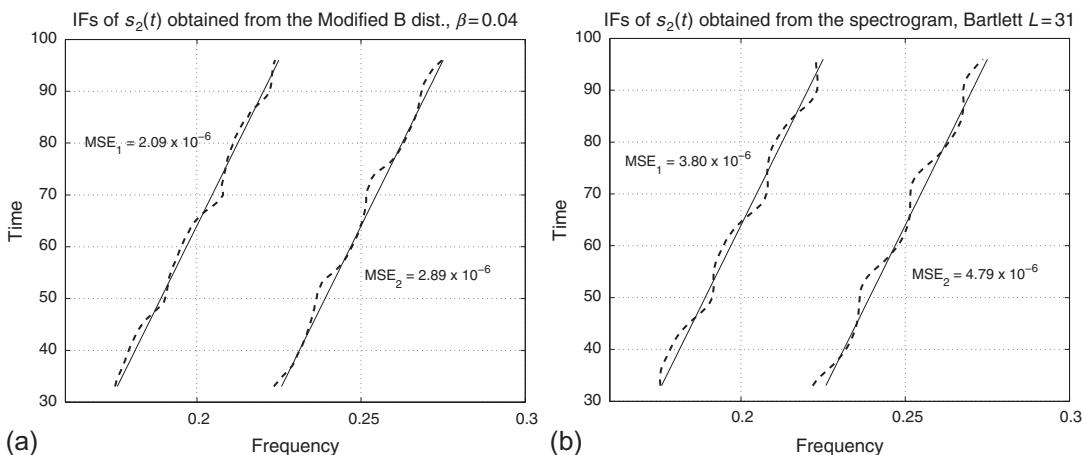
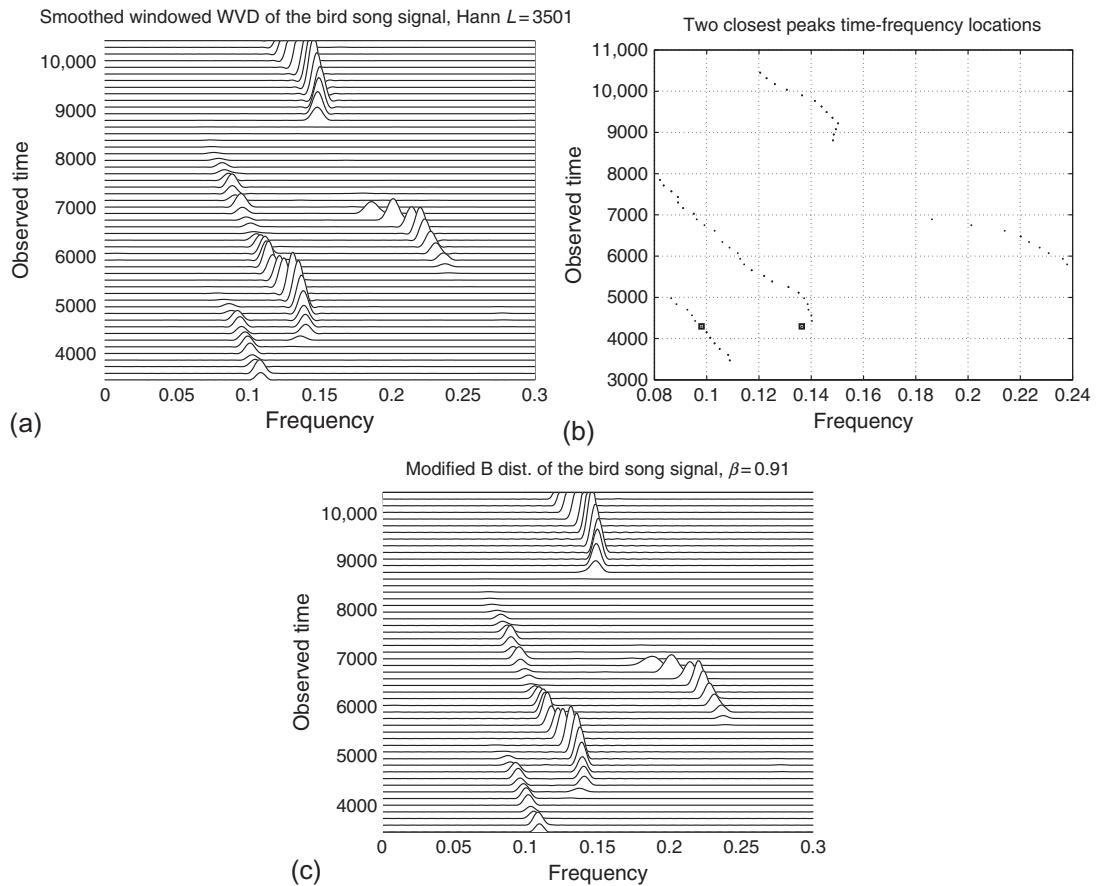


FIGURE 7.4.6

Comparison of the measured (dashed) and true (solid) IF laws of the components $s_1(t)$ (left in both graphs) and $s_2(t)$ (right) of the signal $s(t)$ defined by Eq. (7.4.6). The mean-square-errors (MSEs) of the IF estimates obtained from the peaks of the signal-optimized MBD (best-performing TFD) are given in (a), and from the signal-optimized spectrogram (second best TFD) in (b). (a) Modified B ($\beta = 0.04$) and (b) spectrogram (Bartlett, $L = 31$).

**FIGURE 7.4.7**

Optimization of TFDs of a bird-song signal. The smoothed WVD is shown in (a), and the pairs of its closest components in (b), with the overall closest pair (at time $t = 4295$) marked by the squares. The signal's best-performing TFD (MBD for $\beta = 0.91$) is shown in (c). (a) Smoothed WVD (Hanning, $L = 3501$); (b) pairs of closest components; and (c) modified B ($\beta = 0.91$).

kernel parameters and their resolution performance results. The MBD for $\beta = 0.91$, plotted in Fig. 7.4.7(c), is found to have the largest value of M_{overall} (the mean of M over the observed time instants). Therefore, we select the MBD as the best of the considered TFDs to represent the Noisy Miner song signal in the (t, f) plane.

7.4.6 SUMMARY AND CONCLUSIONS

The measure defined in this section assesses the performance of quadratic TFDs in resolving closely spaced components in the (t, f) domain. It takes into account key attributes of TFDs, such as component

Table 7.4.3: Optimization and Comparison Results for the TFDs of the Noisy Miner Song Signal

| TFD | Optimal Value of the Kernel Parameter | M_{overall} |
|------------------|---------------------------------------|----------------------|
| Spectrogram | Bartlett window, length 3501 | 0.90 |
| Wigner-Ville | N/A in this case | 0.50 |
| Exponential | $\sigma = 0.004$ | 0.74 |
| Born-Jordan | N/A in this case | 0.65 |
| Zhao-Atlas-Marks | $a = 2$ | 0.63 |
| Modified B | $\beta = 0.91$ | 0.93 |

Notes: The values of M_{overall} indicate that the spectrogram performs better than most other traditional TFDs. Only the MBD outperforms the spectrogram.

mainlobes, sidelobes, cross-terms, and instantaneous bandwidths. The measure allows us to quantify the performance of TFDs instead of relying solely on visual inspection of their (t, f) plots. It helps in selecting the optimal TFD in a given practical application, and provides a criterion for improving methodologies to design advanced high-resolution TFDs.

7.5 POSTPROCESSING IN THE (t, f) DOMAIN: METHODS AND PERFORMANCE COMPARISON⁰

When signals are processed using time frequency (t, f) methods, one is confronted with several possible approaches with most requiring some kind of postprocessing to improve the clarity of the (t, f) representation. Among the TFDs, the WVD provides an optimal energy concentration for monocomponent linear frequency modulation (LFM) signals, but it has suboptimal energy concentration for nonlinear FM and multicomponent signals due to the presence of cross-terms and inner interferences. Cross-terms that appear due to the multicomponent characteristics of a signal can be reduced by first extracting signal components using time-frequency filtering algorithms such as blind source separation (Section 8.5), fractional Fourier filtering [46], or matching pursuit algorithm (Section 12.2) and then analyzing each component separately using the WVD or the masked WVD [46]. Time-frequency methods based on signal extraction algorithms produce highly concentrated TFDs for signal components that are well separated in the (t, f) domain; however, to resolve close signal components, such methods require prior information regarding the (t, f) structure of the underlying signal; for example, the performance of the matching pursuit algorithm depends on how closely the (t, f) dictionary is matched with the signal to be analyzed (see Section 12.2 for details).

Another approach to reduce the cross-terms in the WVD is to employ 2D smoothing kernels, but such smoothing has the effect of deteriorating the energy concentration of auto-terms. Several methods

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **N. Khan**, Qatar University, Doha, Qatar (nabeelalikhan@qu.edu.qa). Reviewers: E. Sejdic, S. Ouelha, and S. Dong.

exist to ensure that the energy concentration of blurred auto-terms can be restored using postprocessing techniques. This section presents a brief review of such postprocessing methods including the reassignment method, synchrosqueezing techniques, S-method, and image processing techniques [23,47–49]. These methods are compared in terms of their ability to resolve close components in a high signal-to-noise ratio (SNR) situation, for example, physiological signals. Specifically, this study aims to answer the following questions:

- What improvements can be made to TFDs by using (t, f) postprocessing techniques?
- Can a postprocessing technique overcome the limitation of an underlying TFD to resolve close signal components?
- Which postprocessing method gives the best performance in terms of energy concentration and resolution?

Answers to these questions are provided in the next sections.

7.5.1 OVERVIEW OF POSTPROCESSING METHODS

7.5.1.1 Reassignment method

The spectrogram, like all kernel-based methods, can be interpreted as a smoothed WVD. Smoothing the WVD shifts auto-terms from their original locations in the (t, f) domain. The reassignment method aims to restore the energy concentration of the spectrogram by relocating the auto-terms to their original position (see [Section 7.2](#)). The reassignment method is based on the following principle: first estimate the original location of the auto-terms from the blurred representation and then relocate the auto-terms to their original positions. The original location of an auto-term point can be estimated from the local centroid of the spectrogram using the following expression ([Section 7.2.1](#)):

$$\hat{t}(t, f) = \frac{1}{S_s^h(t, f)} \iint u W_z(t - u, f - v) W_h(u, v) du dv, \quad (7.5.1)$$

$$\hat{f}(t, f) = \frac{1}{S_s^h(t, f)} \iint v W_z(t - u, f - v) W_h(u, v) du dv, \quad (7.5.2)$$

where $W_z(t, f)$ is the WVD of a signal $z(t)$, $W_h(t, f)$ is the WVD of the window function $h(t)$, and $S_s^h(t, f)$ is the spectrogram of the signal. The reassigned TFD is obtained by shifting each (t, f) point of the spectrogram to the corresponding centroid, that is, $(\hat{t}(t, f), \hat{f}(t, f))$

$$\check{S}_s^h(t, f) = \iint S_s^h(u, v) \delta(t - \hat{t}(u, v)) \delta(f - \hat{f}(u, v)) du dv. \quad (7.5.3)$$

7.5.1.2 Synchrosqueezing transformation

The synchrosqueezing transformation like the reassignment method improves the energy concentration of a TFD by relocating auto-terms along the frequency axis in the (t, f) domain. The key difference between the two methods is that synchrosqueezed TFDs apply shifting along the frequency axis only while the reassignment method applies shifting along both time and frequency axes [23]. The synchrosqueezing transformation follows the principle that frequency-only reallocation ensures perfect reconstruction of a signal for linear time-frequency methods such as the short-time Fourier transform (STFT). The key steps of a synchrosqueezing transformation are given below:

- For each (t, f) point, estimate the local centroid along the frequency axis using Eq. (7.5.2).
- Reallocate the auto-term points using the following expression.

$$\check{F}_s^h(t, f) = \frac{1}{2\pi h(0)} \int F_s^h(t, v)\delta(f - \hat{f}(t, v)) dv, \quad (7.5.4)$$

where $F_s^h(t, f)$ is the STFT of the window function $h(t)$.

7.5.1.3 S-method

The S-method aims to combine the advantages of both the spectrogram and the WVD such that the resultant distribution does not suffer from a cross-term problem and has a high-energy concentration for auto-terms (close to that of the WVD). The S-method is based on the principle that the WVD can be represented as the convolution of the STFT with its conjugate along the frequency axis (see Section 6.2 and Ref. [47])

$$W_z(t, f) = \int_{-\infty}^{\infty} F_s(t, f + v)F_s^*(t, f - v) dv. \quad (7.5.5)$$

A cross-term free TFD with high-energy concentration can be obtained by limiting the integration in the above expression to the support of auto-terms [47]. This can be achieved by introducing a narrow window $P(v)$ in Eq. (7.5.5).

$$SM(t, f) = \int_{-\infty}^{\infty} P(v)F_s(t, f + v)F_s^*(t, f - v) dv. \quad (7.5.6)$$

$P(v)$ controls the cross-term suppression and auto-term resolution properties of the S-method. For $P(v) = 1$, the above distribution equals the WVD, while for $P(v) = \delta(v)$, it equals the spectrogram.

7.5.1.4 Time-frequency image enhancement

This approach enhances a given TFD by retaining only the points corresponding to the instantaneous frequency (IF) peaks in the (t, f) domain and assigning zeros to all other locations. The image-processing approach follows the following principle: (1) estimate the IF of a given signal from the TFD; (2) retain only those points that lie along the IF curves.

The IF of a signal can be estimated using an image-processing algorithm that links each IF point in the (t, f) domain. The key steps of the image-processing IF estimation algorithm are (see Section 10.6 and Ref. [48]):

- Define a binary (t, f) image having 1 along (t, f) peaks and 0 at all other points using the first- and second-order derivatives in the (t, f) domain.

$$B(t, f) = 1 \quad \text{if } \frac{d}{df}\rho(t, f) = 0 \quad \text{and} \quad \frac{d^2}{df^2}\rho(t, f) < 0, \\ B(t, f) = 0 \quad \text{otherwise.} \quad (7.5.7)$$

- Link the detected peaks using a connectivity criterion. According to the connectivity criterion, a peak belongs to a signal component if there is at least one other peak for the same signal component in its 10-neighborhood. Given a pixel location (x, y) , pixels in the following set belong to its 10-neighborhood: $(x - 1, y), (x - 1, y - 1), (x - 1, y + 1), (x - 1, y + 2), (x - 1, y - 2), (x + 1, y), (x + 1, y - 1), (x + 1, y + 1), (x + 1, y + 2)$, and $(x + 1, y - 2)$.

7.5.2 RESULTS

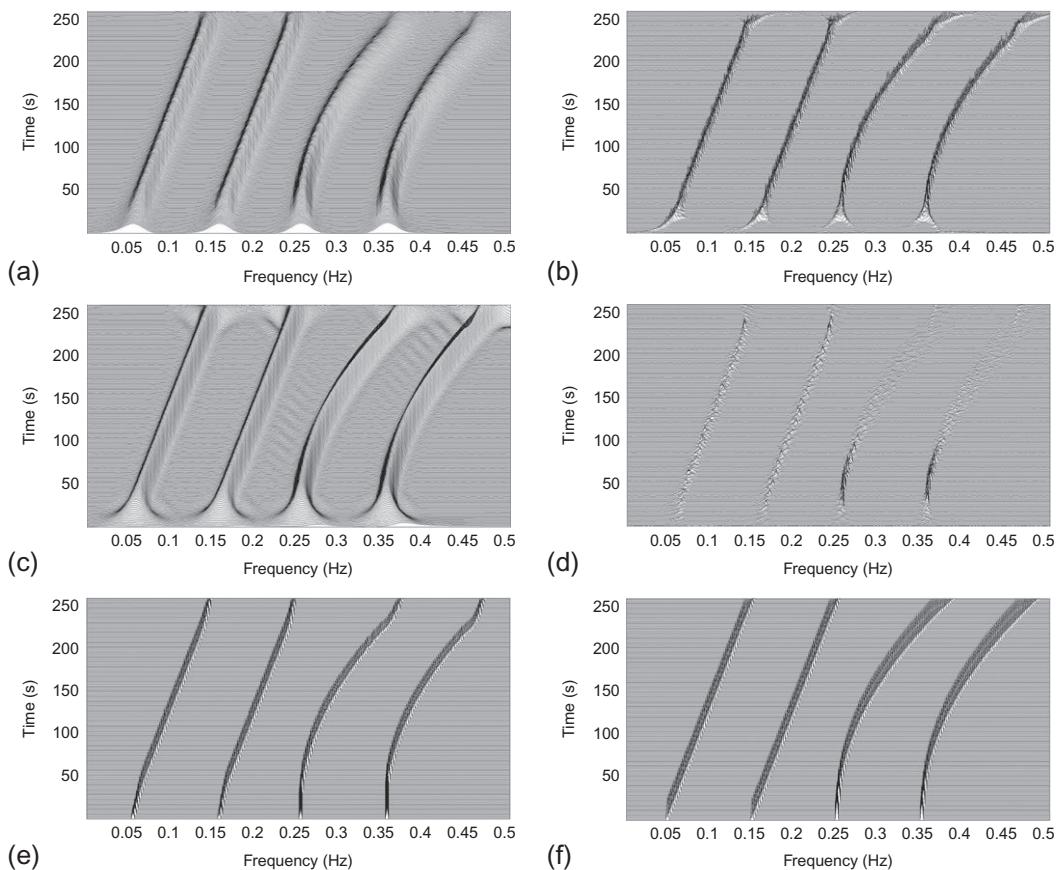
Let us compare the performance of the above (t, f) postprocessing methods in terms of their ability to accurately concentrate the energy of signal components along the corresponding IF curves. In order to have a fair comparison, we select the same underlying TFD for all the postprocessing methods. The spectrogram is selected as the underlying TFD (as the S-method, one of the postprocessing methods, can only improve the energy concentration of the spectrogram).

- Example 1: Let us consider a simple case of four well-separated signal components. [Figure 7.5.1](#) shows the results of postprocessing operations in comparison with an ideal TFD that has peaks only at the IF points and 0 elsewhere. All postprocessing methods improve the energy concentration of the spectrogram. The best energy concentration is obtained using image-processing and reassignment methods, as both methods are able to concentrate signal energy along the IF curves (see [Fig. 7.5.1\(b\)](#) and [\(e\)](#)).
- Example 2: Let us analyze the same four component signals in additive white noise with SNR of 5 dB using the same (t, f) methods. As [Fig. 7.5.2](#) shows, the peaks of the spectrogram are shifted from the true IF curves due to the presence of noise, therefore deteriorating the performance of all postprocessing methods. The performance of the S-method is similar to that of the spectrogram. The image-processing and reassignment methods are still able to improve the energy concentration, but the energy is not accurately localized along the IF curves for all TFDs, as expected.
- Example 3: Let us consider a more complicated case of four signal components such that the two LFM components are close in the (t, f) domain, while the other two components intersect each other in the (t, f) domain. [Figure 7.5.3](#) shows that the spectrogram, reassignment method, synchrosqueezed TFD, and S-method fail to resolve the close signal components. The image-processing method is the only one that resolves close components, although it does not accurately concentrate the signal energy along the true IF curves at all points (see [Fig. 7.5.3\(e\)](#)).

7.5.3 DISCUSSION AND INTERPRETATION OF RESULTS

The key findings of the experimental results can be summarized as follows:

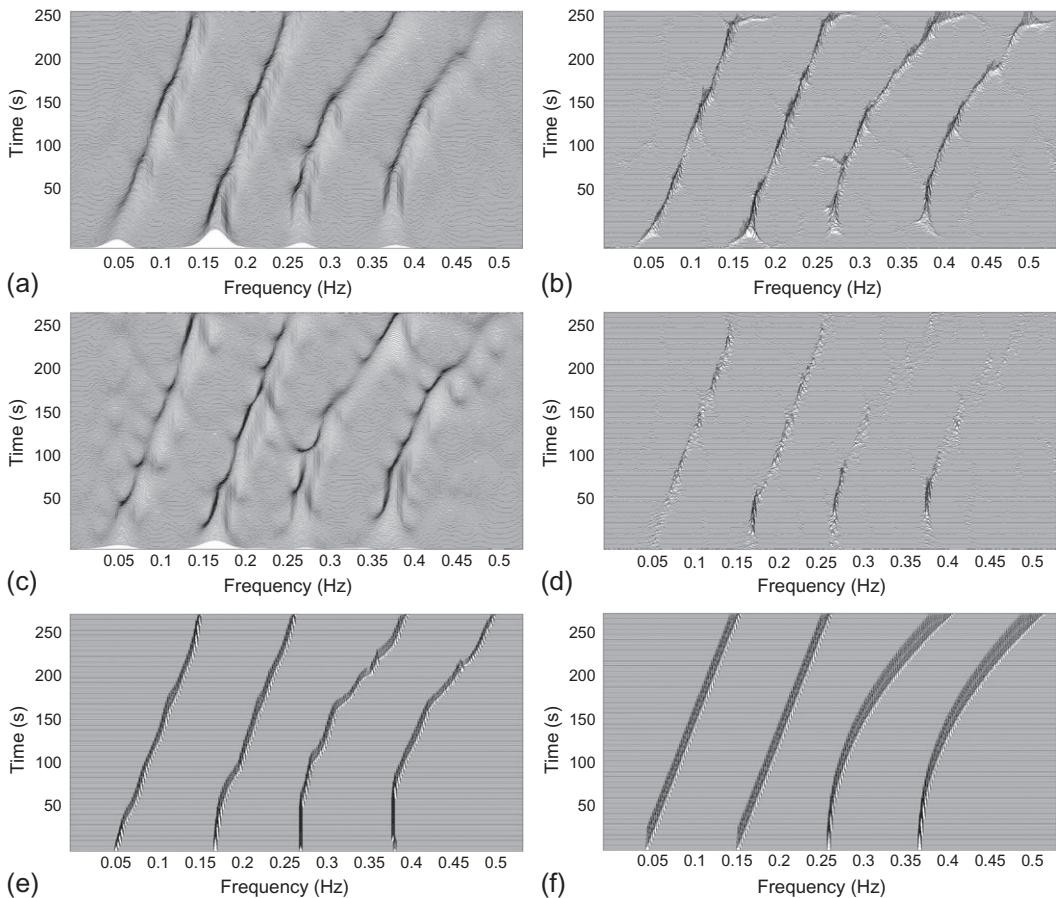
- Postprocessing methods generally improve the energy concentration of a given TFD provided the underlying method has resolved the signal components.
- Among all the postprocessing methods, only the image-processing-based IF estimation method can compensate for the low resolution of the underlying TFD. [Figure 7.5.3](#) indicates that the image-processing method is able to resolve close signal components even when the underlying TFD failed. This performance of the image-processing method is due to its ability to exploit the continuity of peaks in the (t, f) domain. The poor performance of the reassignment method is due to the inaccurate estimation of the true locations of the auto-term points. The performance of the reassignment method can be improved by using the snake-based ridge extraction algorithm for estimating the original location of auto-terms before applying the shifting operation [50].
- The performance of all postprocessing methods deteriorates in the presence of noise and it of course depends on the SNR conditions.

**FIGURE 7.5.1**

Comparison of (t, f) postprocessing techniques for the analysis of a four component signal: (a) spectrogram (Hamming window length = 85); (b) reassigned spectrogram; (c) S-method; (d) synchrosqueezed spectrogram; (e) IF estimated using image-processing method; and (f) original IFs.

These findings indicate that it is important to consider using high-resolution (t, f) methods as the performance of postprocessing methods is directly related to the ability of the underlying TFD to resolve close signal components. The image-processing method can compensate to some extent for the low resolution of the underlying TFD, but such improvement may not result in an accurate localization of the peaks of the signal components for all points.

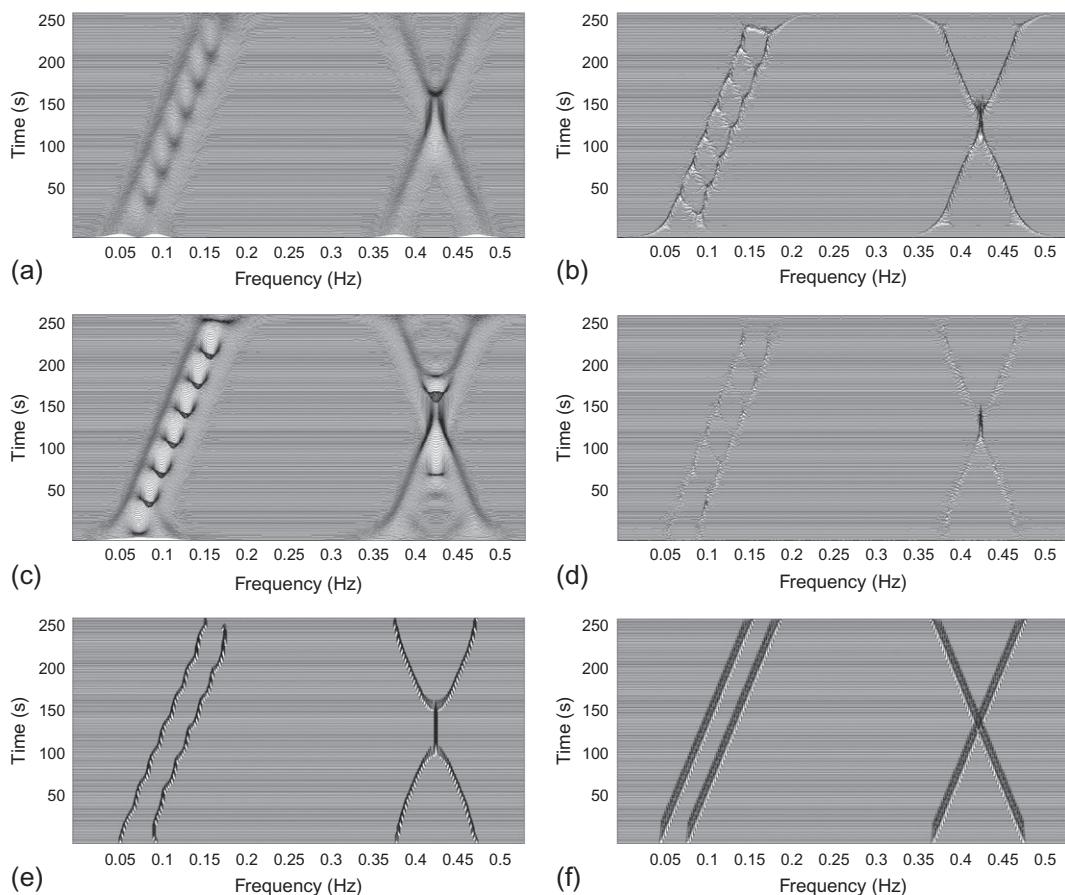
Such postprocessing methods are mainly used to improve the clarity and energy concentration of a TFD obtained in high SNR situations. For low SNR signals, a signal de-noising preprocessing stage such as time-frequency peak filtering (see [Section 11.4](#)) is needed before transforming the signal in the (t, f) domain.

**FIGURE 7.5.2**

Comparison of (t, f) postprocessing techniques for the analysis of a four-component signal corrupted by AWGN noise with SNR of 5 dB: (a) spectrogram (Hamming window length = 85); (b) reassigned spectrogram; (c) S-method; (d) synchrosqueezed spectrogram; (e) IF estimated using image-processing method; and (f) original IFs.

7.5.4 SUMMARY AND CONCLUSIONS

This section provides an insight into the advantages and limitations of some time-frequency postprocessing operations. Methods such as the S-method, reassignment method and synchrosqueezing technique can improve the energy concentration of a quadratic TFD for well separated signal components, but most of these methods fail to improve the resolution of a TFD when signal components are close to each other in the (t, f) domain. The image-processing-based IF estimation procedure can be used to resolve closely spaced signal component even in scenarios when an underlying TFD fails to resolve close signal components.

**FIGURE 7.5.3**

Comparison of (t, f) postprocessing techniques for close and crossing signal components: (a) spectrogram (Hamming window length = 85); (b) reassigned spectrogram; (c) S-method; (d) synchrosqueezed spectrogram; (e) IF estimated using image-processing method; and (f) original IFs.

7.6 TIME-FREQUENCY COMPRESSIVE SENSING⁰

Traditional sampling approaches rely on the Shannon-Nyquist sampling theorem, which mandates that the sampling frequency must be at least twice the highest frequency present in the signal in order to reconstruct it accurately. This approach often results in a large number of samples, and compression schemes are used shortly after sampling to reduce the number of data points that need to be transmitted and/or stored. This is clearly a redundant approach, as most samples are disregarded

⁰Author: E. Sejdić, University of Pittsburgh, Pittsburgh, PA, USA (esejdic@ieee.org). Reviewers: I. Djurović and A. Akan.

soon after acquiring them. An approach called compressive sensing (CS) has been proposed to avoid the aforementioned issues; it postulates that a signal can be reconstructed using a smaller number of randomly chosen signal samples [51–53].

7.6.1 SAMPLING WITH COMPRESSION

The idea behind CS is to diminish the number of steps during a data acquisition process by combining sampling and compression into a single step [52,53]. Specifically, CS approaches acquire signals at sub-Nyquist rates and tend to recover such signals accurately from sparse samples [52]. CS approaches are well suited for K -sparse signals, that is, signals that can be represented by significant K coefficients over an N -dimensional basis, as the CS approaches will encode a K -sparse, discrete-time signal of dimension N by computing a measurement vector y that consists of $M \ll N$ linear projections of the signal s [32,33]

$$y = \Phi s, \quad (7.6.1)$$

where Φ represents an $M \times N$ sensing matrix, and M represents the number of measurements [53]. A noise-free K -sparse signal can be recovered via l_0 minimization and noise-free $M = 2K$ measurements [53]. To recover s from sparse samples, various contributions suggested the use of a norm-minimization framework:

$$\min \|s\|_0 \text{ subject to } \|y - \Phi s\|_2 < \eta, \quad (7.6.2)$$

where η is the expected noise of measurements, $\|s\|_0$ counts the number of nonzero entries of s , and $\|\cdot\|_2$ is the Euclidian norm. Equations (7.6.1) and (7.6.2) will not always yield an accurate representation of sparse signals, and in some applications, such as medical diagnostic applications, it is desired to achieve a “precise” recovery of such signals. In other words, it is desirable to obtain a very small error, which can be accomplished by compressively sampling signals in a different domain. To denote the change of a domain, we can rewrite Eq. (7.6.1) as [54]:

$$y = \Phi s = \Phi \Psi x, \quad (7.6.3)$$

where $s = \Psi x$ represents a sparse representation of a signal in a domain given by Ψ and x representing expansion coefficients.

An important aspect of CS is to understand how the sparsity in the transform domain regulates the number of measurements needed to reconstruct a signal. This is assessed using the so-called coherence measure between the matrices Φ and Ψ :

$$\mu(\Phi, \Psi) = \sqrt{N} \max |\langle \phi_k, \psi_j \rangle|, \quad (7.6.4)$$

where N is the signal length, ϕ_k is the k th row of Φ , and ψ_j is the j th row of Ψ . A simple rule is that a lower coherence value indicates a smaller number of random measurements needed to reconstruct a signal.

7.6.2 COMPRESSIVE SENSING AND TIME-FREQUENCY ANALYSIS

Nonstationary signals typically do not have a sparse representation in the time or frequency domain, but may have a sparse representation in the time-frequency (t, f) domain (Section 14.6). A typical example

of a sparse signal in the (t, f) domain is a linear frequency modulated signal. If a (t, f) representation of such a signal is obtained using the WVD, the signal is concentrated along a straight line in the (t, f) domain and is close to zero at most other points. Hence, it would be advantageous to compressively sample such a signal in the (t, f) domain.

Compressed sensing of nonstationary signals can be accomplished using two different approaches. The first, which is limited to quadratic (t, f) representations, relies on the idea that a signal is sampled in the ambiguity domain [55]. The second approach is to utilize (t, f) dictionaries in order to find a sparse signal representation.

7.6.2.1 Compressive sensing based on the ambiguity domain

One method of obtaining the commonly used quadratic (t, f) distributions of a compressively sampled signal $x(t)$ is to use the so-called ambiguity domain [56]. Let us consider the WVD $W_x(t, f)$, and its 2D Fourier transform, that is, the ambiguity function

$$A_x(v, \tau) = \mathcal{F}_{t \rightarrow v} \left\{ \mathcal{F}_{\tau \leftarrow f}^{-1} \{ W_x(t, f) \} \right\}. \quad (7.6.5)$$

The advantage of using the ambiguity domain is that the cross-terms are usually dislocated from the origin in this domain, in which case they can be suppressed, or significantly attenuated, by the use of low-pass filtering. This is achieved by applying the kernel $g(v, \tau)$

$$\mathcal{A}_x(v, \tau) = g(v, \tau) A_x(v, \tau). \quad (7.6.6)$$

Improved (t, f) signal power localization can be achieved by using the compressed sensing approach and exploiting sparsity in the (t, f) domain. Namely, we can collect a set of samples from the ambiguity domain and solve the ℓ_1 -norm minimization problem to obtain the sparsest (t, f) distribution. By using the compressed sensing approach, the desired (t, f) distribution $\hat{\rho}_x$ can be obtained as

$$\hat{\rho}_x(t, f) = \arg \min_{\rho_x(t, f)} \|\rho_x(t, f)\|_1 \quad (7.6.7)$$

$$F_{2D}^{-1} \{ \rho_x(t, f) \} - \mathcal{A}_x^M = 0 \Big|_{(v, \tau) \in \Omega}, \quad (7.6.8)$$

where $\rho_x(t, f)$ is the (t, f) distribution to be optimized, \mathcal{A}_x^M is the set of samples from the ambiguity domain in the region defined by the mask $(v, \tau) \in \Omega$, and $\|\cdot\|_1$ denotes the ℓ_1 -norm. In the presence of noise, we may use the approximation

$$\hat{\rho}_x(t, f) = \arg \min_{\rho_x(t, f)} \|\rho_x(t, f)\|_1 \quad (7.6.9)$$

$$\left\| F_{2D}^{-1} \{ \rho_x(t, f) \} - \mathcal{A}_x^M \right\|_2 \leq \epsilon \Big|_{(v, \tau) \in \Omega}. \quad (7.6.10)$$

Here, it is important to select a suitable set of ambiguity domain samples, which is done by an appropriate ambiguity function masking. The mask can be formed as a small area around the origin of the ambiguity plane, as done for designing high-resolution (t, f) distributions [39]. In other words, the mask is designed to pass any auto-terms while reducing the cross-terms in the ambiguity domain.

7.6.2.2 Compressive sensing based on (t, f) dictionary and matching pursuit

The first step in this approach is to construct a (t, f) dictionary. Using the dictionary, the next challenge is to obtain a sparse signal representation via Eqs. (7.6.2) and (7.6.3); however, this approach can be extremely computationally expensive, making it unsuitable for many applications. Therefore, matching pursuit [57] or its different variants (e.g., Ref. [58]) can be utilized in order to avoid computational burdens associated with traditional CS approaches.

A matching pursuit approach begins with an initial approximation of the signal, $\hat{x}^{(0)}(m) = 0$, and the residual, $R^{(0)}(m) = x(m)$, where m represents the M time indices that are uniformly or nonuniformly distributed. Then, at each subsequent stage, k , the matching pursuit identifies a dictionary atom with the strongest contribution to the residual and adds it to the current approximation

$$\hat{x}^{(k)}(m) = \hat{x}^{(k-1)}(m) + \alpha_k \phi_k(m) \quad (7.6.11)$$

$$R^{(k)}(m) = x(m) - \hat{x}^{(k)}(m), \quad (7.6.12)$$

where $\alpha_k = \langle R^{(k-1)}(m), \phi_k(m) \rangle / \|\phi_k(m)\|^2$. The process continues until the norm of the residual $R^{(k)}(m)$ does not exceed the permitted margin of error $\varepsilon > 0$: $\|R^{(k)}(m)\| \leq \varepsilon$ [57], or as long as the number of bases $n_{\mathcal{B}}$ needed for signal approximation satisfies $n_{\mathcal{B}} \leq \mathcal{K}$, where \mathcal{K} is a predefined number of bases.

Regardless of a stopping criterion, a signal approximation is obtained using L bases as

$$x(n) = \sum_{l=1}^L \langle x(m), \phi_l(m) \rangle \phi_l(n) + R^{(L)}(n), \quad (7.6.13)$$

where ϕ_l are L bases from the dictionary with the strongest contributions.

The advantage of the proposed decomposition is that other signal operations can be easily carried out. Specifically, we can obtain a (t, f) representation of a signal using its L -bases approximation

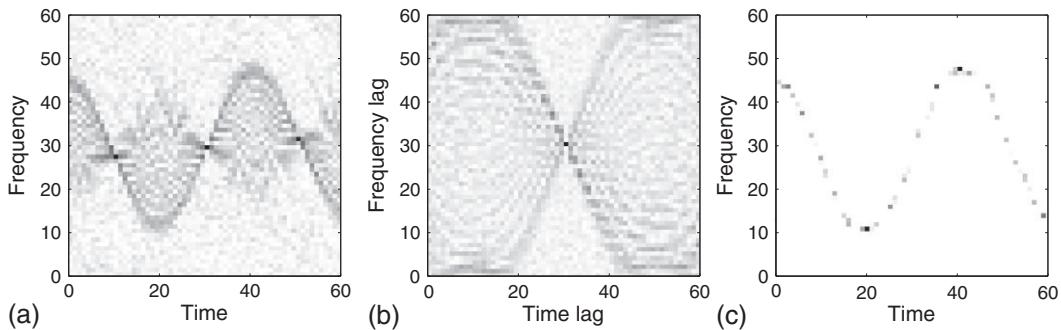
$$\mathcal{T}\mathcal{F}\{x(n)\} = \sum_{l=1}^L \langle x(m), \phi_l(m) \rangle \mathcal{T}\mathcal{F}\{\phi_l(n)\}, \quad (7.6.14)$$

where $\mathcal{T}\mathcal{F}\{\cdot\}$ is a (t, f) operator (e.g., short time Fourier transform) [56].

Note that these steps require knowledge of the sampling times in order to acquire proper values of members of the (t, f) dictionary. In real-life conditions, we do not always have such knowledge and there may exist a need to estimate the sampling time instances. To accomplish this task, the use of annihilating filters has been proposed in several contributions [51, 59, 60]. The annihilating filter approach relies on determining the roots of an autoregressive filter in order to estimate the sampling instances.

7.6.3 ILLUSTRATING EXAMPLES

For example, let us consider a sinusoidally modulated signal contaminated with white Gaussian noise. In order to provide faster computations, we use (t, f) representations of size 60×60 (3600 points). The mask is of size 7×7 (1.4% of the total number of points) in the ambiguity domain. The original (t, f) distribution and ambiguity function are shown in Fig. 7.6.1(a) and (b), respectively. The resulting sparse representation is illustrated in Fig. 7.6.1(c). Note that the compressed sensing approach provides

**FIGURE 7.6.1**

Time-frequency representations of a sinusoidally modulated signal using: (a) the Wigner distribution; (b) the ambiguity function; and (c) the resulting sparse (t, f) representation.

improved results, reducing the noise influence significantly. The number of nonzero points in the sparse (t, f) representation is approximately between 45 and 50 (estimated from different experiments), which is a quite small percentage of the total number of points in the (t, f) domain.

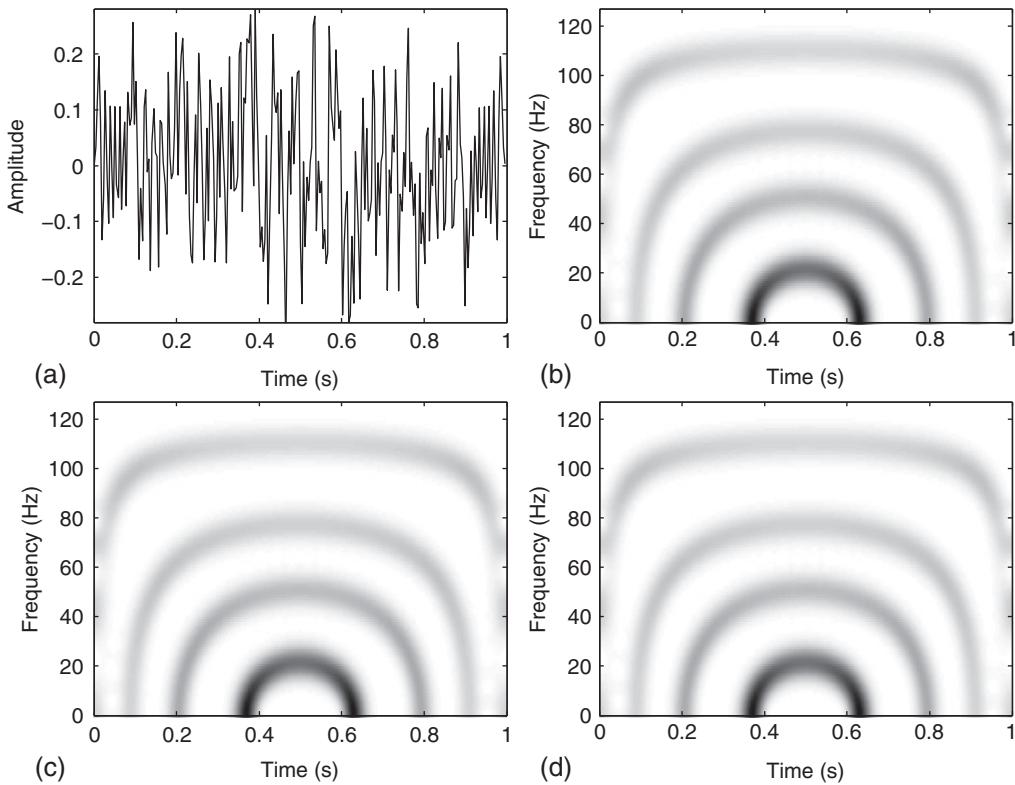
The second case, shown in Fig. 7.6.2(a), involves a signal consisting of four basis functions from a 25-band dictionary based on modulated discrete prolate spheroidal sequences (MDPSSs) [61,62] with normalized half-bandwidth $W = 0.495$ and $N = 256$. For both uniform and nonuniform sampling, only 42 samples (less than 17% of the total number) were needed to accurately recover the signal and its spectrograms based on regular and irregular sample times as shown in Fig. 7.6.2(c) and (d). A greater percentage of samples was required for this case in comparison to the first case, as the second case is recovered almost exactly.

Third, we examine the accuracy of the instantaneous frequency (IF) estimator based on the complete (t, f) representations and those representations obtained from compressed samples. As a sample signal, we considered a sinusoidally modulated signal defined as

$$x(t) = \sin(120\pi t + 2\pi \cos(8\pi t)), \quad (7.6.15)$$

where $0 < t < 1$ s, and the assumed sampling period is $T_s = 1/256$ s. Figure 7.6.3(a)-(c) shows sample results when 75% of samples were used for CS. Next, the signal was contaminated with an AWGN, and its variance was proportional to the considered SNR. The accuracy of the estimator was then examined for a number of SNR values as shown in Fig. 7.6.3(d), and the presented values were obtained using 500 realizations. As expected, the complete (t, f) representation obtained the lowest values of mean square errors, but (t, f) representations based on CS closely follow the trend of the complete (t, f) representation, especially the (t, f) representation based on uniformly CS. These results clearly demonstrate that even if we use compressive samples, we can achieve a reliable estimate of IF. Further studies are needed to understand the effects of the number of acquired samples on mean square errors.

Lastly, the accuracy of CS schemes based on (t, f) dictionaries is examined using swallowing accelerometry signals, which can be used to infer if a patient has swallowing difficulties. Here, vibrations in three anatomical directions associated with a sample swallow obtained during a videofluoroscopy

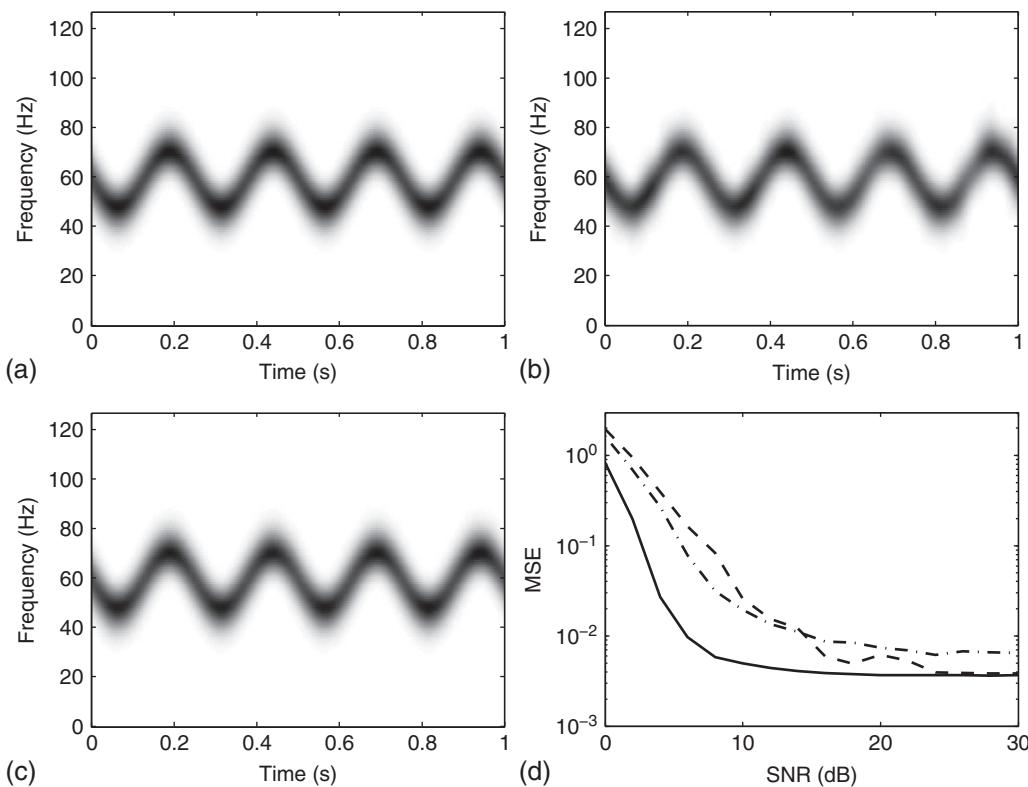
**FIGURE 7.6.2**

(a) Time-domain plot of original signal; spectrograms of (b) original signal, (c) signal based on equal-distance samples, and (d) signal based on irregular samples.

exam are considered. An MDPSS-based dictionary was used with $W = 0.25$ and $N = 128$. In this sample, case presented in Fig. 7.6.4, we considered the error when 33% of samples were available. These sample results show that the considered biomedical signals can be recovered almost exactly with a small number of samples. This is because bases in the (t, f) dictionary are able to capture these nonstationarities accurately.

7.6.4 SUMMARY AND CONCLUSIONS

Two techniques of CS using (t, f) analysis have been briefly summarized. One is based on the ambiguity function and the other on (t, f) dictionaries. As expected, the CS approach provided accurate results in the recovery of nonstationary signals. While the present section considered a (t, f) dictionary based on MDPSS, other bases such as wavelets or modulated Gaussian functions are expected to produce accurate results as well.

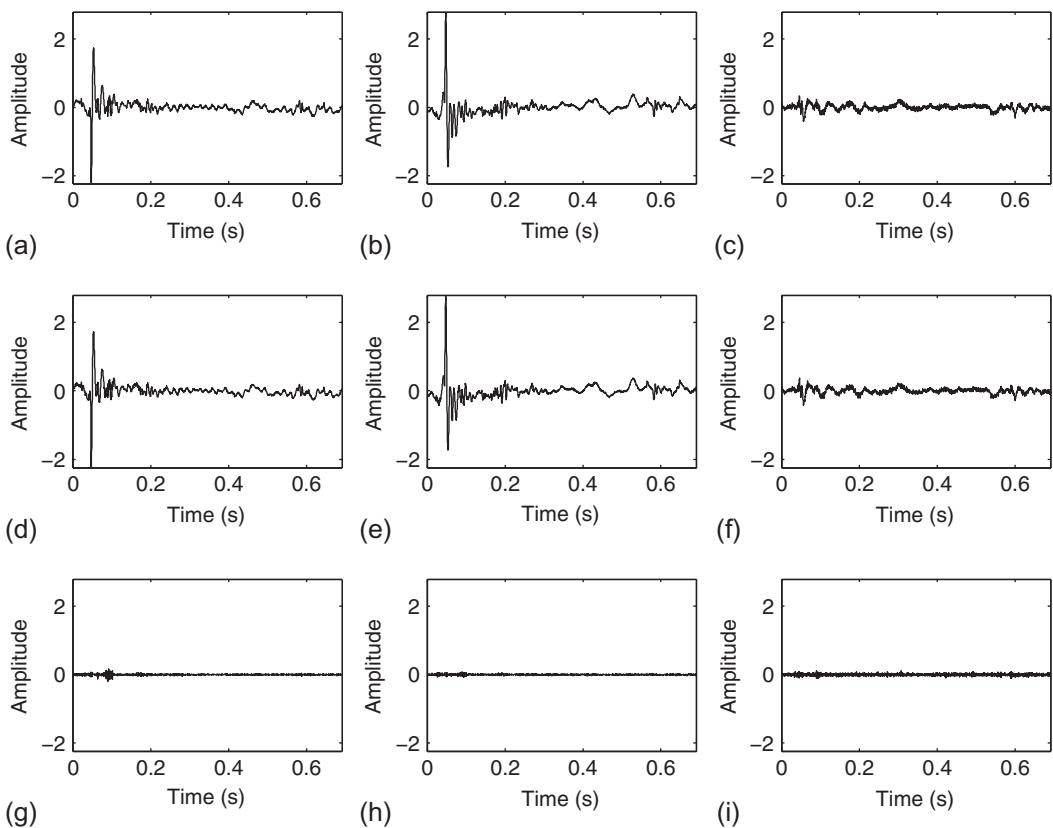
**FIGURE 7.6.3**

Spectrograms based on: (a) the original signal; (b) equal-distance samples; and (c) irregular samples. (d) Performance of the IF estimator based on the original signal (solid line), equal-distance compressed samples (dashed line), and nonuniform compressed samples (dash-dotted line).

7.7 SIGNAL COMPLEXITY ESTIMATION USING TIME-FREQUENCY SHORT-TERM ENTROPY⁰

Applications such as classification require (t, f) features that can be used for pattern recognition as an aid to identification and detection [39,63]. A simple but useful feature is the measure of complexity, which is reviewed below and applied to the estimation of the number of components in a signal. The particular TFD selected for illustration is the spectrogram with a Gaussian window as it allows analytical calculations to validate the results.

⁰Authors: **V. Sucic**, University of Rijeka, Rijeka, Croatia (vsucic@riteh.hr), **N. Saulig**, University of Rijeka, Rijeka, Croatia (nsaulig@riteh.hr), and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: LJ. Stanković and S. Aviyente.

**FIGURE 7.6.4**

Sample swallowing accelerometry signals from a patient with swallowing difficulties. Original signal in: (a) anterior-posterior direction, (b) superior-inferior direction, and (c) medio-lateral direction. Recovered signal in: (d) anterior-posterior direction, (e) superior-inferior direction, and (f) medio-lateral direction. Error between original and recovered signals in: (g) anterior-posterior direction, (h) superior-inferior direction, and (i) medio-lateral direction.

7.7.1 ENTROPY MEASURES AND TIME-FREQUENCY DISTRIBUTIONS

The concept of entropy was originally derived from thermodynamics as a measure of disorder of a thermodynamic system. Its introduction to information theory has allowed quantification of the information content of a probability density function (PDF) [64]. Entropy-based complexity estimation of nonstationary signals is founded on the analogy that TFDs can be interpreted as (pseudo-)PDFs of signal energy in the joint time-frequency plane [65].

A TFD $\rho_z(t, f)$ is said to satisfy energy preservation and marginal conditions if

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_z(t, f) dt df = E_z, \quad \int_{-\infty}^{\infty} \rho_z(t, f) df = |z(t)|^2, \quad \int_{-\infty}^{\infty} \rho_z(t, f) dt = |Z(f)|^2. \quad (7.7.1)$$

The instantaneous power $|z(t)|^2$ and the energy spectrum $|Z(f)|^2$ can be understood as 1D densities of signal energy in time or frequency, respectively, while the TFD $\rho_z(t, f)$ may be interpreted as a 2D energy density [65].

TFDs belonging to the quadratic class can locally assume negative values, thus limiting the application of Shannon entropy to those TFDs [66]. This is overcome by using an alternative complexity measure known as Rényi entropy, defined as [64]

$$H_{\alpha,z} = \frac{1}{1-\alpha} \log_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\rho_z(t, f)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_z(u, v) du dv} \right)^{\alpha} dt df, \quad (7.7.2)$$

with the parameter α being the order of Rényi entropy.

Effects of different entropy orders have been studied in Ref. [29], showing that for odd α , in particular for $\alpha = 3$, Rényi entropy can be used as a measure of signal complexity for a broad class of signals and TFDs, including the TFDs that may locally take negative values, known to be a characteristic of some popular high-resolution TFDs [39]. Also note that for α approaching 1, Rényi entropy reduces to Shannon entropy [29,66].

7.7.2 THE ENTROPY COUNTING PROPERTY

It is intuitive to expect that signals composed of a number of nonoverlapping elementary components achieve larger entropy values when compared to a single elementary component. In the case of “ideal” TFDs—in which each signal component contributes individually to the overall TFD, but with no interference terms between the components—Rényi entropy can be shown to exhibit a component “counting” property.

Consider an ideal TFD, $I_z(t, f)$, of a two-component signal $z(t) = z_1(t) + z_2(t)$, where $z_2(t)$ is a copy of $z_1(t)$ shifted in time by t_0 and in frequency by f_0 , and the components’ respective TFDs are $I_{z_1}(t, f)$ and $I_{z_2}(t, f)$. It follows that

$$I_z(t, f) = I_{z_1}(t, f) + I_{z_2}(t, f) = I_{z_1}(t, f) + I_{z_1}(t - t_0, f - f_0). \quad (7.7.3)$$

Under the assumption that $I_{z_1}(t, f)$ and $I_{z_1}(t - t_0, f - f_0)$ are compactly supported nonoverlapping components in the (t, f) plane, Rényi entropy carries exactly one additional bit of information when compared to Rényi entropy of the TFD of one of the components [29], that is,

$$H_{\alpha,z_1+z_2} = H_{\alpha,z_1} + 1 = H_{\alpha,z_2} + 1, \quad (7.7.4)$$

so the number of signal components can be estimated as

$$N = 2^{H_{\alpha,z_1+z_2} - H_{\alpha,z_1}} = 2^{H_{\alpha,z_1+z_2} - H_{\alpha,z_2}} = 2. \quad (7.7.5)$$

In general, the number of estimated components in an M -component signal with nonoverlapping components, represented by its ideal TFD, is

$$N = 2^{H_{\alpha,z_1+z_2+\dots+z_M} - H_{\alpha,z_m}} = M, \quad (7.7.6)$$

where $z_m(t)$ is an arbitrarily chosen component from the mixture.

In order for Eq. (7.7.6) to hold, all signal components must exhibit the same time and frequency supports—that is, all components must be time- and/or frequency-shifted replicas of the reference

component $z_m(t)$ —and the reference component must be known in advance [41]. Those assumptions, however, do not hold in general, in particular when it comes to real-life signals. So, a modification is needed for the method to be applicable in practice.

7.7.3 DEFINITION AND PROPERTIES OF THE SHORT-TERM ENTROPY

When observing only short-time intervals of the signal TFD, the components can be locally assumed to have the same time durations and bandwidths [41], hence removing the influence of different components on the counting property of Rényi entropy. So, even if signals may globally have different time and frequency supports, their Rényi entropies are locally comparable [41].

The counting property of the short-term Rényi entropy will next be analyzed on the spectrogram. In spite of its limited time-frequency resolution, the spectrogram's simple definition still makes it one of the most popular TFDs, both for initial comparative theoretical analysis and as a reference in practical applications. More details on the spectrogram window selection are given in [Chapter 2](#).

7.7.3.1 Calibration on stationary signals

The case of a pure tone. The behavior of the locally applied Rényi entropy is first illustrated on the signal

$$z(t) = e^{j2\pi f_0 t}, \quad (7.7.7)$$

the spectrogram of which is [67]

$$S_z^w(t, f) = K e^{-4\pi^2 \sigma_w^2 (f - f_0)^2}, \quad (7.7.8)$$

where $K = 2\sqrt{\pi}\sigma_w$ is related to the analysis window $w(t)$ with standard deviation σ_w .

The selected time interval is obtained by annulling the spectrogram outside the interval Δt , centered on the time instant p

$$Sp_z^w(t, f) = \begin{cases} S_z^w(t, f), & p - \Delta t/2 < t < p + \Delta t/2 \\ 0, & \text{otherwise.} \end{cases} \quad (7.7.9)$$

Using Eq. (7.7.2), the spectrogram's short-term Rényi entropy is found to be

$$H_{\alpha, z}(p) = \log_2 \Delta t + \left(\log_2 \frac{1}{2\sqrt{\pi}\sigma_w} \right) \alpha^{-\frac{1}{2(1-\alpha)}}. \quad (7.7.10)$$

Equation (7.7.10) indicates that the local entropy of the single-tone spectrogram time slice depends on the duration of the interval Δt , the standard deviation σ_w of the spectrogram window $w(t)$, and the order α of the Rényi entropy. The local Rényi entropy decreases as the parameter σ_w increases. Larger entropy-order values emphasize the peakiness of the spectrogram, hence decreasing the entropy. Note also that the second term in Eq. (7.7.10) decreases as the entropy order α increases, since $\lim_{\alpha \rightarrow \infty} \alpha^{-\frac{1}{2(1-\alpha)}} = 1$.

The case of a two-component signal. In the case of a signal whose components are shifted versions of the reference signal $z(t)$, that is, $z_1(t) = z(t - t_1) e^{j2\pi f_1' t}$ and $z_2(t) = z(t - t_2) e^{j2\pi f_2' t}$, where $f_0 + f_1' = f_1$ and $f_0 + f_2' = f_2$, the spectrogram of the two-component signal is

$$\begin{aligned} S_{z_1+z_2}^w(t, f) = & K e^{-4\pi^2 \sigma_w^2 (f-f_1)^2} + K e^{-4\pi^2 \sigma_w^2 (f-f_2)^2} \\ & - 2K e^{-2\pi^2 \sigma_w^2 [(f-f_1)^2 + (f-f_2)^2]} \cos [2\pi(f_1 - f_2)t + \Phi + \pi], \end{aligned} \quad (7.7.11)$$

where Φ is the initial phase difference of the components.

By assuming that the frequencies f_1 and f_2 are distant enough so that the spectral characteristics of $w(t)$ do not cause any significant overlapping of the components in the (t, f) plane, the last term in Eq. (7.7.11) can be ignored. So, the local entropy of the two-component signal spectrogram is approximated as

$$H_{\alpha, z_1+z_2}(p) \approx \log_2 \Delta t + \left(\log_2 \frac{1}{2\sqrt{\pi}\sigma_w} \right) \alpha^{-\frac{1}{2(1-\alpha)}} + 1. \quad (7.7.12)$$

Equations (7.7.10) and (7.7.12) imply that the local Rényi entropy of the spectrogram of the two-component signal carries approximately an additional bit of information when compared to the entropy of the spectrogram of the reference component. Therefore, in order to obtain the local number of components in a multicomponent signal, the counting property of Rényi entropy can be applied locally, that is

$$N(p) = 2^{H_{\alpha, z_1+z_2+\dots+z_M}(p) - H_{\alpha, z_m}(p)}. \quad (7.7.13)$$

Under the assumption that all signal components have equal amplitudes, $N(p) \approx M(p)$.

The case of completely overlapping components in the time-frequency plane. When the two stationary components, $z_1(t)$ and $z_2(t)$, completely overlap in the (t, f) plane, that is for $f_1 = f_2$, the spectrogram of their sum takes the form

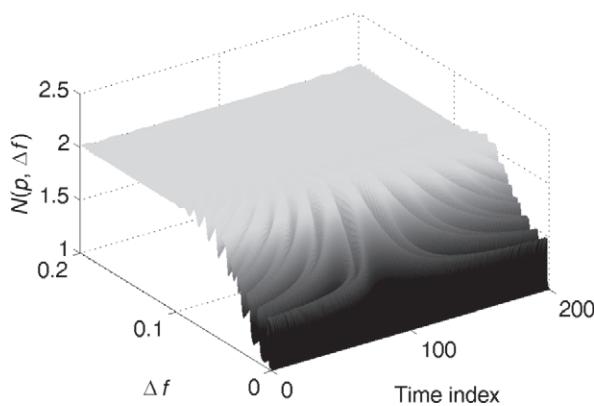
$$S_{z_1+z_2}^w(t, f) = 2KK_1 e^{-4\pi^2 \sigma_w^2 (f-f_1)^2}, \quad K_1 = 1 - \cos(\Phi + \pi). \quad (7.7.14)$$

This result is equivalent to the single-component case, since K_1 is a constant. Thus, for the two overlapped components, even for $\Phi = \pi(2n + 1)$, $n \in \mathbb{Z}$, we have

$$\lim_{\Phi \rightarrow \pi} H_{\alpha, z_1+z_2}(p) = \log_2 \Delta t + \left(\log_2 \frac{1}{2\sqrt{\pi}\sigma_w} \right) \alpha^{-\frac{1}{2(1-\alpha)}} = H_{\alpha, z}(p). \quad (7.7.15)$$

Equation (7.7.15) shows that independent of their initial phases, two completely overlapped components produce the entropy equivalent to that of a single component.

The case of partially overlapping components in the time-frequency plane. When the two components partially overlap in the (t, f) plane, based on Eq. (7.7.11), the local Rényi entropy will exhibit a periodic behavior. Figure 7.7.1 shows the estimated number of components, $N(p)$, for the case when the two components $z_1(t)$ and $z_2(t)$ have a varying frequency separation $\Delta f = |f_1 - f_2|$. From its initial value of $N(p, \Delta f) = 2$, when the components are distant enough so that the effect of interferences is minimized, the estimated number of components starts to present an oscillatory transition region, with the oscillation period being proportional to Δf , until it reaches the final value of $N(p, \Delta f) = 1$ as Δf approaches 0.

**FIGURE 7.7.1**

Number of components $N(p)$ w.r.t the frequency separation Δf between two tones.

7.7.3.2 Application to nonstationary signals

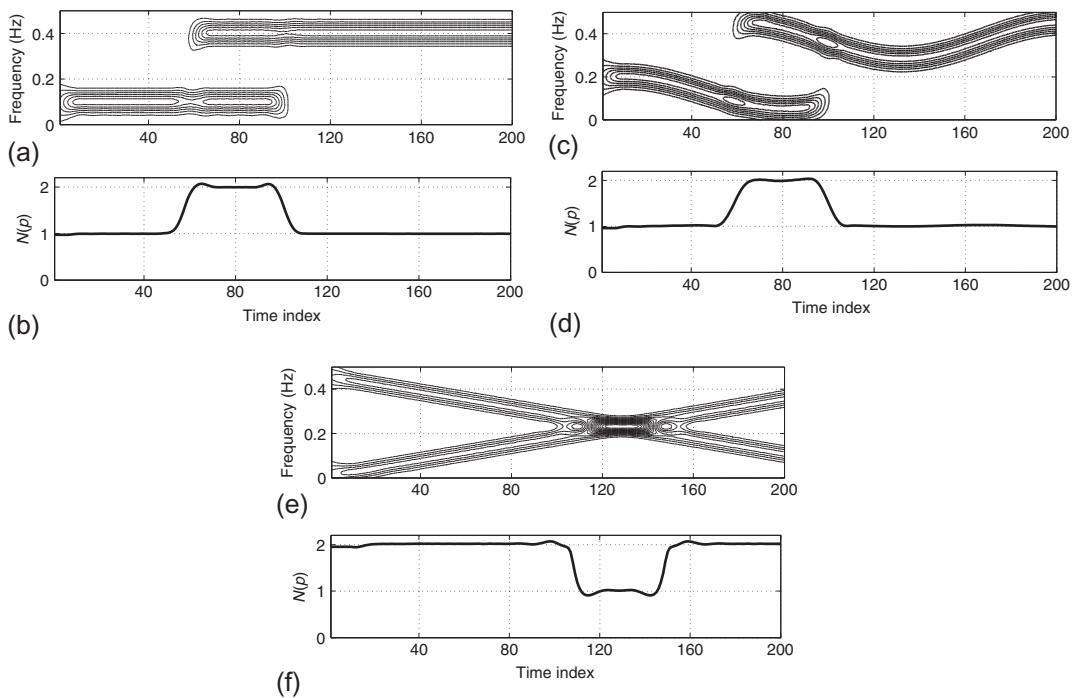
The nonstationary signal model. The local estimation can generally be applied to nonstationary signals, provided the spectrogram window is selected in such a way that the spectral characteristics of the signal components are locally stationary. In this case, the spectrogram of a nonstationary signal can be locally substituted by the spectrogram of the stationary signal model, given by Eq. (7.7.11). If the component separation is sufficient to ensure that interferences can be neglected, the counting property will still hold, since the time dependencies of the instantaneous frequencies (IFs) $f_1(t)$ and $f_2(t)$ vanish under the frequency integration of Eq. (7.7.11). The reference signal as given by Eq. (7.7.7) can be also kept.

The results for the estimated number of local components, obtained from the spectrograms of stationary and nonstationary signals with components having identical time supports, are shown in Fig. 7.7.2(a)-(d). These results confirm that the nonstationarity brings no inaccuracies in the estimation process.

The case of locally overlapping nonstationary components. If the intersection of two nonstationary components occurs at a time instant t_0 , and if the time interval Δt is short enough such that the components' IFs, $f_1(t)$ and $f_2(t)$, can be considered constant inside Δt (i.e., $f_1(t) \approx f_1$ and $f_2(t) \approx f_2$), then the function $N(p)$ can be ideally represented by its cross-section (see Fig. 7.7.1) in the plane $\Delta f(p) = |f_1(p) - f_2(p)|$; however, such a representation is a rough approximation, and at the instant of component intersection, $t = t_0$, the estimated number of components is approximately equal to one.

In general, when the intersection of the components' IFs occurs inside the interval Δt , the estimated number of components of a multicomponent signal will be $N(p) \approx M(p) - 1$, as shown in Fig. 7.7.2(e)-(f). Therefore, even when $M(p)$ components coexist at the same time interval and contribute to the local spectrum, the local entropy detects only the number of different regions in the (t, f) plane occupied by the signal.

The case of components with nonidentical amplitudes. The counting property of the local entropy estimation holds only when the signal components have equal amplitudes. If this is not the case, however, the following iterative procedure can be employed. First, the dominant spectral component ($\max_m A_m(p)$) is filtered out such that contributions from weaker spectral components can also be taken

**FIGURE 7.7.2**

Spectrogram of a signal with two stationary components and estimated number of components (a-b); spectrogram of a signal with two nonstationary components and estimated number of components (c-d); and spectrogram of a signal with two intersecting components and estimated number of components (e-f).

into account. With each new iteration “ i ,” the estimated number of components $N^{[i]}(p)$ decreases. For each time instant p , a component is removed from the signal TFD, and the integer counter of the number of components, $N_{\text{count}}(p)$, is increased by one. When $N^{[i]}(p) < 1$, the procedure is terminated [68].

The performance of the method is illustrated on a synthetic signal (see Fig. 7.7.3(a)-(b)), proving its capability to accurately detect different-amplitude components in moderate noise conditions. It has also been applied to a newborn EEG signal, shown in Fig. 7.7.3(c)-(d). Indeed, brain functional abnormalities, such as seizures, can be efficiently detected using (t, f) features of EEG signals [39]. The local number of components can be one such feature, obtainable with the entropy-based method described in this section. Finally, the method is tested on the bird-song signal, shown in Fig. 6.6.2, and the results obtained are presented in Fig. 7.7.4.

7.7.4 SUMMARY AND CONCLUSIONS

This section has presented a method for quantifying the information content of a nonstationary signal using the short-term Rényi entropy of the signal TFD. The standard global Rényi entropy provides reliable results only when the signal components have the same time durations and frequency

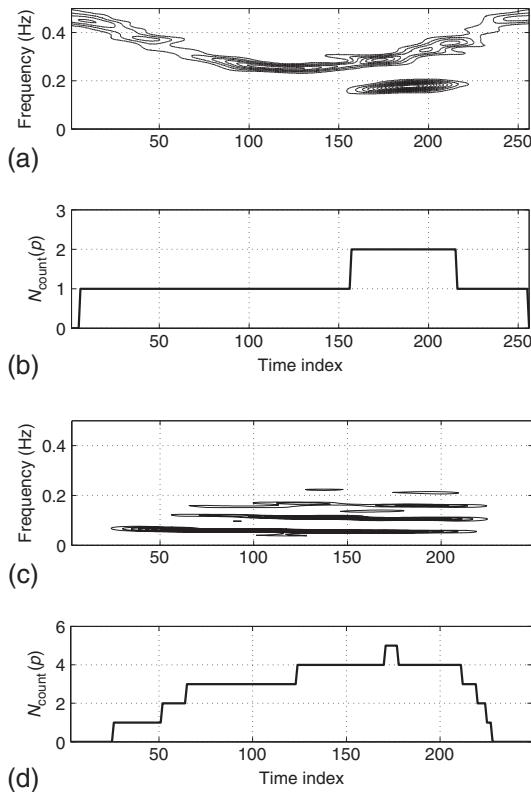
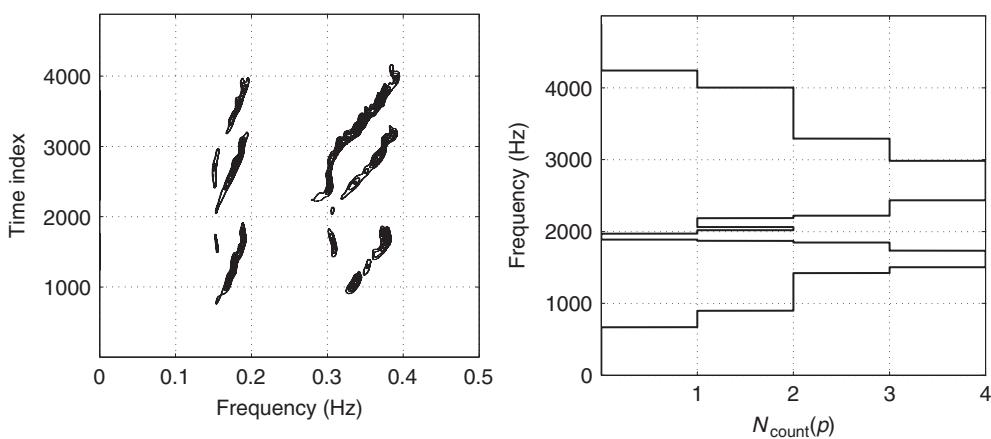


FIGURE 7.7.3

TFDs and local numbers of components of a synthetic nonstationary signal in 5 dB additive white noise (a-b), and an EEG signal (c-d).

bandwidths, and equal amplitudes, and when the reference signal is known *a priori*. These requirements are difficult to meet in practice. Therefore, an extension has been described using the entropy estimation over a short-time interval for components with different time and/or frequency supports, which is comparable to the local entropy of an arbitrary reference signal.

Furthermore, when extended to nonstationary signals, the counting property of the short-term Rényi entropy offers additional insights on the signal structure, quantifying the local number of components present in the signal. To efficiently deal with signals having components of varying amplitudes, an iterative algorithm has been described. The local component number estimation is a valuable information for different source separation/extraction approaches, as well as applications requiring dynamic memory allocation. Other (t, f) features that can be used for automatic classification are reported in Refs. [39,49,69].

**FIGURE 7.7.4**

TFD (left) and local number of components (right) of a bird-song signal.

7.8 TIME-FREQUENCY ANALYSIS USING NEURAL NETWORKS⁰

7.8.1 REQUIREMENTS AND MOTIVATION

Nonstationary real-life signals include human speech, animal sounds, propeller sounds, door slams, and inelastic acoustics. It is desirable to obtain temporal and spectral information on such signals with high resolution and without potential interference (see Chapter 3). These characteristics facilitate visual interpretation and discrimination between known patterns for classification [70].

The removal of distortions in a TFD can be considered as a case of image de-blurring. The distortions in a TFD may be caused by cross-terms (CTs), as in the WVD, and/or by the blur in auto-terms, as in the spectrogram. An artificial neural network (ANN)-based setup is suitable for de-blurring an image because: (a) the information on the source of blurring is insufficient; (b) it is not possible to get an accurate mathematical description of the blurring phenomenon, due to the complex combination of events causing it; and (c) enough data are available that, if given to an ANN, can disclose the fundamental relationship depicted in the blurred image.

The ANN-based framework to obtain highly concentrated TFDs is a data-driven method that does not assume any prior knowledge about the components present in the signal. This method is aimed at performing deconvolution once CTs are eliminated while trying to minimize the distortion of auto-terms. The focus is not on designing any optimal kernel for reducing CTs. The Bayesian regularized Levenberg-Marquardt backpropagation (LMB) training algorithm is used for training the ANN. The LMB algorithm is a variation of the Newton method that was designed for minimizing functions that are sums of squares of other nonlinear functions. The Bayesian framework of David Mackay [71] has

⁰Authors: **I. Shafi**, Abasyn University, Islamabad, Pakistan (imran.shafi@abasyn.edu.pk), **J. Ahmad**, Abasyn University, Islamabad, Pakistan (jamil.ahmad@abasyn.edu.pk), and **S.I. Shah**, Abasyn University, Islamabad, Pakistan (syedismailshah@gmail.com). Reviewers: S. Stankovic and A. Akan.

been used because it causes the network to have smaller weights and biases, thus forcing the network response to be smoother and less likely to overfit.

7.8.2 COMPUTING DE-BLURRED TIME-FREQUENCY DISTRIBUTIONS

The ANN-based method fundamentally involves selection and training of a set of suitable ANNs. The improved TFDs are obtained in the testing phase using these selected ANNs [72].

7.8.2.1 Preprocessing of the spectrogram and WVD images

The spectrogram and preprocessed WVD of various known signals constitute the input and target TFDs for the ANN. For training, two common but different types of signals are considered. These include signals with parallel chirps and sinusoidally modulated FM components in the (t, f) plane, given as

$$x_1(n) = \exp[j\xi_1(n)n] + \exp[j\xi_2(n)n], \quad (7.8.1)$$

$$x_2(n) = \exp([j\pi/2 - j\pi\xi(n)]n), \quad (7.8.2)$$

where we arbitrarily choose $\xi_1(n) = \pi n/(4N)$, $\xi_2(n) = \pi/3 + \pi n/(4N)$ in the first signal, and $\xi(n) = 0.1 \sin(2\pi n/N)$ in the second signal. The choice of these values is based on getting better-shaped (t, f) representations of the signals during simulation. Here, N refers to the total number of sampling points.

The WVD of a multicomponent signal contains CTs that can be removed by multiplying the WVD with the spectrogram of the same signal. The spectrogram and CT-free WVD of these known signals are converted to suitably sized vectors. This conversion is required to reduce complexity and to employ the ANN setup in an effective manner. The vectors are chosen to be 1×3 in size for spectrogram images and are extracted from each row. These vectors are grouped into clusters such that similar vectors belong to the same cluster. This is achieved by correlating the vectors from the blurred spectrogram with some common types of edges, for example, ascending, descending, and wedge-shaped, based on the “elbow criterion” (see below). The fuzzy partitioning algorithm validated by various objective measures is also used to confirm the results obtained by the elbow criterion (see Ref. [73] and Chapter 3).

7.8.2.2 Processing by the localized neural networks

Localized neural processing is considered important for numerous reasons. First, it is well known that different parts of the human brain are designated to perform different tasks. The nature of the task imposes certain structure for the region, resulting in a structure-function correspondence. Also, different regions in the brain compete to perform a task, and the task is assigned to the winning region. Mimicking the behavior of the brain, ANNs may also be employed based on these arguments. An image contains structural information with low- and high-frequency content. A blurred version has lost most of its high-frequency information. The objective of any de-blurring system is to restore this information by gaining sufficient knowledge about the blur function; however, information is generally lost at various scales in different regions, which must be taken into account; for example, the edges and the flat regions are blurred simultaneously but at different rates. This favors the idea of subdividing the data into appropriate groups. A second reason is the problem of overtraining the ANN, causing loss of generalization ability. If only a single ANN is used, it may end up memorizing the training data and may adjust its weights to any noise. Yet another reason is specific to the case of image processing; if an ANN is trained by an entire image containing different distribution characteristics for data corresponding to

different structures in the image, it may attempt to represent different structures by finding a common ground between the different data distributions, thus limiting its recognition ability. This forces one network to learn distant input patterns, causing training to slow down in attempting to represent input data that are significantly different [74].

We use the feed-forward back-propagation ANN architecture with 3 input neurons, 1 output neuron, and 40 hidden neurons. We use the LMB algorithm for training. The *hyperbolic tangent sigmoid* and *positive linear* transfer functions are used to represent the hidden layer of sigmoid neurons followed by an output layer of positive linear neurons. This architecture and topology are chosen after experimenting with various training algorithms, using different parameters such as activation functions between layers, numbers of hidden layers, and numbers of neurons [75].

During training, vectors x_i taken from the spectrogram are mapped to scalar values y_i taken from the preprocessed WVD. As an example, if $x_i = [x_{i1} \ x_{i2} \ x_{i3}]_{(1,3)}$ is a vector from the spectrogram, then the target value is obtained by averaging out the same sized vector from the WVD image like $y_i = [(y_{i1} + y_{i2} + y_{i3})/3]_{(1,1)}$, where y_i are corresponding pixel values from the same location in the WVD. It is desired that the ANN does well on data it has not seen before and is not overtrained. The weights are initialized to random values during training and may not converge in an optimal fashion. Additionally, the phenomenon of early stopping may result in poorly trained ANNs that are used to avoid overfitting the data. These facts favor the idea of training multiple ANNs for each type of region and choosing the best ANN for each subgroup of data. The performance parameters include the mean-square error (MSE) reached in the last epoch, maximum number of epochs, performance goal, maximum validation failures, and the performance gradient. These can be accessed to find out the most optimally trained ANN out of multiple ANNs employed for each cluster. These selected ANNs for all clusters are called the *localized neural networks* (LNNs).

In the testing phase, the spectrogram of an unknown signal is first converted to vectors that are clustered. These vectors are tested by the LNNs, and the results are postprocessed to produce the TFD image. The postprocessing involves zero-padding the resulting scalar values to form vectors, which are restored to the original positions using their known index values from the preprocessing stage, to form the 2D image matrix.

7.8.3 RESULTS AND DISCUSSION

7.8.3.1 Cluster analysis

The *elbow criterion* is a common rule of thumb for optimizing the number of clusters based on the information content. The percentage of variance is graphed against the number of clusters. The first cluster adds much information; but as more clusters are added, the marginal gain drops significantly and gives rise to an “elbow.” The number of subspaces into which vectors are distributed is selected based on the elbow criterion in relation to the edges present in the data. We consider the edge because it is one of the important underlying image features and characteristics. Moreover, it is well established that blurring mostly causes loss of *edge* information. Intuitively, an edge could be ascending (e.g., 1, 2, 3), descending (3, 2, 1), wedge (1, 3, 2), flat (1, 1, 1), triangular (1, 3, 1), etc. Considering the above-mentioned edges based on the elbow criterion, it is found that going from three to four clusters does not add sufficient information, as the end result has no significant change in the entropy values. The number of clusters to be chosen by this rule of thumb shall be 3, as suggested in Ref. [76].

Table 7.8.1: Values of Validity Measures for Different Numbers of Clusters

| | | PC | CE | SC | S _(1.0e-004 *) | XB | DI | ADI |
|--------------------------|----|-------|-------|-------|---------------------------|-------|--------|--------|
| Number of clusters | 2 | 0.969 | 0.050 | 0.243 | 0.064 | 9.256 | 0.0019 | 0.0159 |
| | 3 | 0.954 | 0.080 | 0.551 | 0.192 | 12.73 | 0.0003 | 0.0947 |
| | 4 | 0.950 | 0.095 | 0.657 | 0.199 | 5.794 | 0.0005 | 0.0061 |
| | 5 | 0.947 | 0.105 | 0.553 | 0.166 | 5.210 | 0.0005 | 0 |
| | 6 | 0.946 | 0.110 | 0.696 | 0.248 | 20.31 | 0.0003 | 0 |
| | 7 | 0.943 | 0.119 | 0.496 | 0.176 | 5.174 | 0.0005 | 0 |
| | 8 | 0.945 | 0.117 | 0.473 | 0.180 | 3.970 | 0.0007 | 0 |
| | 9 | 0.944 | 0.121 | 0.444 | 0.170 | 3.438 | 0.0007 | 0 |
| | 10 | 0.945 | 0.121 | 0.396 | 0.152 | 3.142 | 0.0007 | 0 |
| | 11 | 0.944 | 0.125 | 0.397 | 0.143 | 2.917 | 0.0003 | 0 |
| | 12 | 0.947 | 0.119 | 0.356 | 0.135 | 3.805 | 0.0005 | 0 |
| | 13 | 0.946 | 0.123 | 0.384 | 0.150 | 2.857 | 0.0003 | 0 |
| | 14 | 0.944 | 0.130 | 0.349 | 0.141 | 2.778 | 0.0018 | 0 |

In addition, fuzzy cluster validity measures are used to confirm the result. The values of the validity measures depending on the number of clusters are shown in Table 7.8.1. During this optimization process, all parameters are fixed to their default values, and the number of clusters c is varied in the range $2 \leq c \leq 14$. We refer to Ref. [73] for a detailed analysis on definition and interpretation of various validation index values. It is important to mention that no single validation index is perfect and reliable only by itself. The optimal value can be detected only with the comparison of all the results. We choose a number of clusters so that adding another cluster does not add sufficient information. This means that either the marginal gain drops or differences become insignificant between the values of a validation index. The partition coefficient (PC) and classification entropy (CE) suffer from drawbacks of their monotonic decrease with the number of clusters and the lack of direct connection to the data. The optimal number of clusters is at the maximum value for both the criteria. On the score of their values in Table 7.8.1, the number of clusters can be only rated to 3. The values of the separation index (S) and the partition index (SC) hardly decrease at the $c = 3$ point (lower values indicate a better partition). The Xie-Beni (XB) index reaches the local minimum at $c = 10$, which indicates a better partition and the optimal number of clusters; however, the optimal number of clusters is chosen to be 3 based on the fact that SC and S are more useful, as confirmed by the values of Dunn's index (DI) and the alternative Dunn's index (ADI) shown in Table 7.8.1.

7.8.3.2 Concentration and resolution performance

The results of the ANN-based approach have been compared to the results obtained by some traditional and advanced high-resolution (t, f) techniques. The list includes the WVD, the traditional reassignment method (Section 7.2), and the optimal-radially Gaussian kernel method (OKM) (Section 5.3). For an empirical judgment on TFDs' performance, various objective measures are used, including the ratio-of-norms-based measure [77], normalized Rényi entropy measure (Section 5.2), Stankovic measure

(Section 7.3), and Boashash-Sucic performance measures (Section 7.4). The visual quality of the resultant TFD images is also judged keeping in mind the limitations of these objective measures. The TFD with the largest value of the ratio-of-norms-based measure displays the best concentration. On the contrary, the *lesser* the value of the normalized Rényi entropy or the Stankovic measure, the better the concentration of the TFD. Performance is evaluated by considering both synthetic and real-life signals.

Example 7.8.1. The first example is the real-life bat-chirp signal TI1 (test image 1). Its spectrogram, shown in Fig. 7.8.1(a), is blurred and difficult to interpret. The results are obtained using other high-resolution (t, f) methods including the WVD, the traditional reassignment method (RAM) (Section 7.2), and the OKM [78]. The (t, f) plots for these methods and the neural-network-based TFD (NTFD) are shown in Fig. 7.8.1(b)-(e) for visual interpretation.

The reassigned spectrogram for the bat-echolocation chirp signal shows energy concentration, but often can diminish accuracy due to its way of approaching the problem (Section 7.2). Also, its performance deteriorates for low SNR values, and it contains discontinuities as evident from Fig. 7.8.1(c). On the other hand, the OKM proposes a signal-dependent kernel that changes shape for each signal, to offer improved (t, f) representation (Section 5.3); but, as shown in Fig. 7.8.1(d), it does not recover all components and thus loses some useful information about the signal.

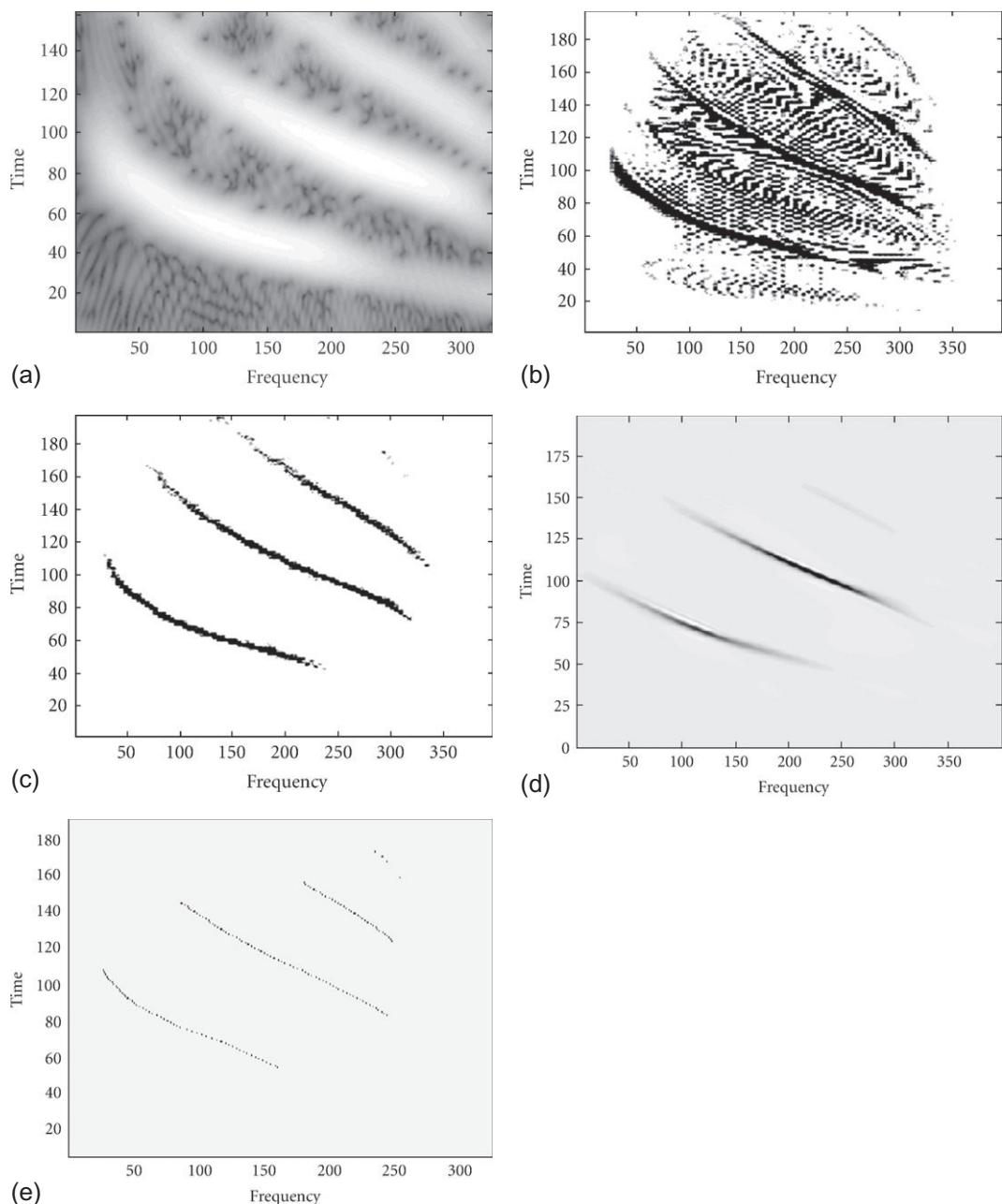
To each data point in the spectrogram, the ANN-based method assigns a (t, f) coordinate from the preprocessed WVD. The reassigned TFD obtained by this method for the test case is shown in Fig. 7.8.1(e), which has better resolution and is concentrated along the individual components. The evaluation by various objective criteria is presented in Table 7.8.2 for both examples, showing that the ANN-based method yields better results than the other approaches. Additionally, Fig. 7.8.1(e) is more informative, as the four components are clearly visible. For further analysis, slices of the spectrogram and the TFD obtained by the LNNs are taken at the time instants $n = 150$ and $n = 310$ and are plotted in Fig. 7.8.2. These instants are chosen because three chirps can be identified visually at these instants (see Fig. 7.8.2(b)). The peaky appearance of three different frequencies at these instants is obvious in Fig. 7.8.2. There are no CTs, and the narrower main lobe with no side lobes indicates better frequency resolution.

Example 7.8.2. The second example is a synthetic signal that contains two sinusoidal FM components and two chirps intersecting with each other. An additive Gaussian noise of variance 0.01 is also added to the signal to consider the performance of the approach under noise. This signal is chosen to confirm the ANN-based scheme's performance at the intersection of the IFs and closely spaced components under noise, with the view that estimation of the IF is rather difficult in these situations. Its discrete mathematical form is given as

$$\text{sig}_1(n) = \sin\left[\frac{3\pi}{2} + 0.1\pi \sin\left(\frac{2\pi n}{N}\right)n\right] + \exp\left(\frac{j\pi n}{4N}n\right) + \exp\left[j(4\pi - \frac{\pi n}{4N})n\right]. \quad (7.8.3)$$

The noisy spectrogram of the signal is shown in Fig. 7.8.3(a) and referred to as TI 2. The two components intersect at 2.3 s between 100 and 180 Hz. Also, the sinusoidal-FM and chirp signals are closely spaced around 825-900 Hz near 0.7 s.

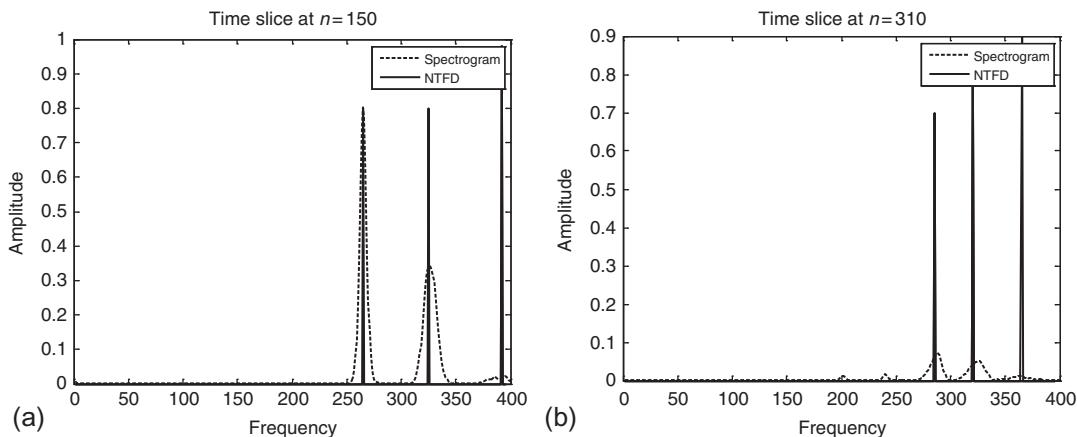
This synthetic test example is processed through the ANN-based method and the result is shown in Fig. 7.8.3(b). It achieves better resolution and concentration, as is verified by the objective assessment in Table 7.8.2. We further evaluate TFDs' instantaneous concentration and resolution performance by the measures defined by Boashash and Sucic in Section 7.4; however, these are computationally expensive

**FIGURE 7.8.1**

TFDs of the multicomponent bat-echolocation chirp signal: (a) spectrogram, (b) WVD, (c) reassigned spectrogram, (d) TFD obtained by the OKM, and (e) TFD obtained by the ANN method.

Table 7.8.2: Objective Assessment

| Description | Test TFD | NTFD | RAM | OKM |
|--|--------------|------------------|------------------|----------------|
| Ratio of norm-based measure ($\times 10^{-4}$) | TI 1 TI 2 | 76 21 | 68 15 | 8.32 4.33 |
| Volume-normalized Rényi entropy measure | TI 1 TI 2 | 6.21 7.20 | 7.30 9.99 | 11.77 11.34 |
| Stankovic measure ($\times 10^5$) | TI 1 TI 2 | 0.00019 0.143 | 0.00129 2.396 | 0.63 9.515 |

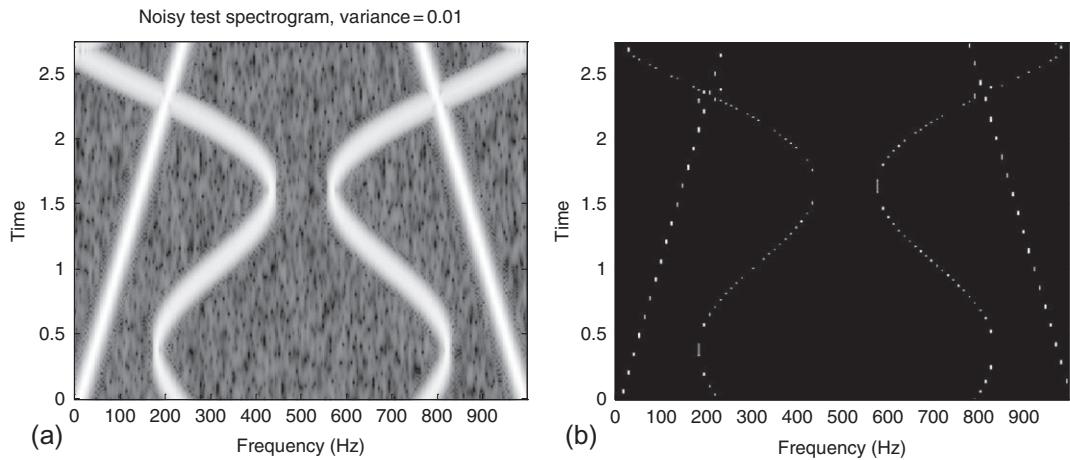
**FIGURE 7.8.2**

Time slices of the spectrogram (smooth dotted line) and the TFD obtained by the ANN-based approach (spiky solid line), for the bat-echolocation signal, at (a) $n = 150$ and (b) $n = 310$.

and require extensive calculations that are beyond the scope of this section. The results are recorded and described in [73] for the interested reader.

7.8.4 SUMMARY AND CONCLUSIONS

The ANN-based framework is found effective for (t, f) analysis of time-varying signals using both synthetic and real-life examples. This is demonstrated against some high-resolution (t, f) methods. The performance of the ANN-based scheme is satisfactory for signals corrupted with additive Gaussian noise with small variance, and is better at detecting the number of components in a signal. Although energy preservation and marginal properties are not satisfied, ANN analysis can be used for classification problems and similar applications, due to an easy visual interpretation.

**FIGURE 7.8.3**

TFDs of a synthetic signal consisting of two sinusoidal FM components and two chirp components: (a) spectrogram (TI 2) (Hamming, $L=90$) with additive Gaussian noise; and (b) TFD obtained by the ANN-based approach.

7.9 DEFINING JOINT-DOMAIN REPRESENTATIONS VIA DISCRETE-DOMAIN FRAMES⁰

Representation of a signal in joint domains such as the (t, f) domain is an active area in signal processing. Examples of such representation in the joint (t, f) domain are the short-time Fourier transform (STFT) (see [Chapter 2, Eq. 2.3.2](#)), the discrete-time Gabor transform (DTGT) ([Chapter 2, Eq. 2.3.14](#)), and the WVD ([Chapter 2, Eq. 2.1.17](#)) in the joint time-frequency domain, and filter banks ([Chapter 2, Eq. 2.4.1](#)) and wavelets ([Section 4.1](#)) in the joint time-scale domain. This section shows how a variety of joint-domain transforms can be unified in this section in the form of product-function frames and cascaded frames.

7.9.1 FRAMES AND RECONSTRUCTION COLLECTIONS

A *frame* is a collection of sequences, which we refer to as frame elements, that generates a finite representation of a signal via the inner product of the signal with the frame elements. An overview of frame theory can be found in Ref. [80]. Mathematically, a collection of sequences $\{v_m\}$ in a Hilbert space \mathbb{H} , where $m \in \mathbb{Z}$ (the set of integers), forms a frame for \mathbb{H} , if there exist two numbers A and B , called the lower and upper frame bound, respectively, such that $0 < A \leq B < \infty$ and

⁰Second edition condensed and updated by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). First edition [79] by J.M. Morris and S.M. Joshi. Reviewers: X.-G. Xia, R. Baraniuk, and M.J. Bastiaans.

$$A\|s_f\|^2 \leq \sum_m |\langle s_f, v_m \rangle|^2 \leq B\|s_f\|^2, \quad \forall s_f \in \mathbb{H}. \quad (7.9.1)$$

Every $s_f \in \mathbb{H}$ can be reconstructed using a corresponding dual frame $\{u_m\}$ as $s = \sum_m \langle s_f, v_m \rangle u_m = \sum_m \langle s_f, u_m \rangle v_m$. The corresponding dual frame $\{u_m\}$ in \mathbb{H} is defined by $u_m = F_o^{-1}v_m$, where F_o is the frame operator defined via $F_o s = \sum_m \langle s_f, v_m \rangle v_m$.

A collection $\{u_m\}$ in \mathbb{H} is defined to be a *reconstruction collection* [81] (RC) for $\{v_m\}$ if every $x \in \mathbb{H}$ can be written as $x = \sum_m \langle x, v_m \rangle u_m$. It is assumed that all the collections or sequences described in this section belong to the same Hilbert space of interest. An RC $\{u_m\}$ is related to a frame $\{v_m\}$ via the following theorem [81].

Theorem 7.9.1. A collection $\{u_m\}$ is a reconstruction collection for a frame $\{v_m\}$ for \mathbb{H} if and only if

$$\sum_m v_m^*(l) u_m(k) = \delta(k - l). \quad (7.9.2)$$

The spaces of interest are the spaces of periodic and nonperiodic square-summable sequences. The space $l^2(\mathbb{Z}/L)$ is the space of all sequences $s_f(k)$ periodic in k with integer period L . The inner product for $l^2(\mathbb{Z}/L)$ is defined as $\langle x, y \rangle = \sum_{k=0}^{L-1} x(k) y^*(k)$. The designation L represents the set $\{0, 1, 2, \dots, L-1\}$. The space $l^2(\mathbb{Z})$ is the space of all square-summable sequences, with the inner product defined as $\langle x, y \rangle = \sum_{k=-\infty}^{\infty} x(k) y^*(k)$.

7.9.2 PRODUCT-FUNCTION FRAMES

Product-function frames (PFFs) generalize the windowed transforms such as the discrete-time Gabor transforms.

7.9.2.1 PFF for periodic spaces

A PFF with two defining factor functions g_m and ϕ_n in $l^2(\mathbb{Z}/L)$ is defined via $\{g_m \phi_n\}$. The elements of this frame are $(g_m \phi_n)(k) = g_m(k) \phi_n(k)$, where $m \in \underline{M}$ and $n \in \underline{N}$. *Windowed transforms* are equivalent to oversampled (or critically sampled) *modulated filterbanks* [82]. Instead of using a single window for the Gabor transform, we can use a number of windows, resulting in a *multi-window Gabor transform* [83].

7.9.2.2 PFF for nonperiodic spaces

For $l^2(\mathbb{Z})$, the elements of a PFF $\{g_m \phi_n\}$ are $(g_m \phi_n)(k) = g_m(k) \phi_n(k)$, as in the periodic case. The indices m and n are integers and $m \in \mathbb{Z}$ and $n \in \underline{N}$. Examples of such collections are windowed transforms and a different type of transform using a *scale-modulation* structure. This different transform was used in Ref. [81] to analyze an exponentially decaying modulated signal, as found in magnetic resonance techniques. It was shown that the scale-modulation transform is likely to yield better signal estimation in the presence of noise.

The shift-modulation structure of the Gabor transforms has been in use for many years. The shift-scale structure of the wavelet transforms is another development. A third choice is the scale-modulation structure, which is provided by the scale-modulation transform.

7.9.3 CASCADED FRAMES

Although the product-function structure unifies the windowed transforms, it is not a good structure to represent the other popular joint time-frequency analysis scheme: wavelets. The discrete wavelet transform (DWT) and all of its variations can be unified under a second structure, cascaded frames. In simpler terms, a cascaded frame is a “transform of transforms,” although the discussion is presented in terms of the more general frames concept.

The following theorem proves that the cascaded structure corresponds to a frame [81].

Theorem 7.9.2. Suppose $\{u_{0,m}\}$, $m \in \underline{L}_1$, is a frame for $\ell^2(\mathbb{Z}/L_0)$, and $\{u_{1,n}\}$, $n \in \underline{L}_2$, is a frame for $\ell^2(\mathbb{Z}/L_1)$, where $L_0 \leq L_1 \leq L_2$. Then there exists a frame $\{u_p\}$, $p \in \underline{L}_2$, for $\ell^2(\mathbb{Z}/L_0)$ obtained by cascading $\{u_{0,m}\}$ and $\{u_{1,n}\}$.

When the number of elements in a frame for a finite-dimensional space is finite, frame theory is characterized in terms of matrix theory. Computation of an RC becomes an inversion problem of square matrices in the case of critically sampled frames, and that of rectangular matrices in the case of oversampled frames.

We now discuss some examples of cascaded frames in $\ell^2(\mathbb{Z})$. Consider a perfect reconstruction (PR) filterbank (FB) shown in Fig. 7.9.1. One of the analysis filters, say $G0$, is typically a low-pass filter and the other one, $G1$, is a high-pass filter. The output of each filter thus gives the information about the signal in a particular time period in the corresponding frequency band. In *discrete wavelet transforms* (DWTs), the output of $G0$ is further processed using the same filterbank, thus increasing the frequency resolution at the cost of reducing the time resolution for the transform coefficients. This cascading process is continued repeatedly to obtain the complete DWT. PR is guaranteed as long as the corresponding synthesis FB is present on the synthesis side.

In *wavelet packet transforms* (WPTs), the output of any filter may be subjected to another layer of cascading filters. This gives much more flexibility in choosing the time and frequency resolution for the transform coefficients. In *hybrid WPTs*, the filters $G0$ and $G1$ may be different for each cascading layer. All these results can be readily extended to $\ell^2(\mathbb{Z}/L)$ [84].

We need not restrict ourselves to cascaded filterbanks only. We can replace some of the filterbanks by any frame. Even a frame for a periodic space $\ell^2(\mathbb{Z}/L)$, when used as a filterbank, provides a frame for $\ell^2(\mathbb{Z})$, since it is effectively a windowed transform, where the window is a rectangular window of length L moved by L samples for each shift [81]. Thus, we can combine the WPT with the DFT to obtain different TF plane partitions. Figure 7.9.2 shows such a scheme.

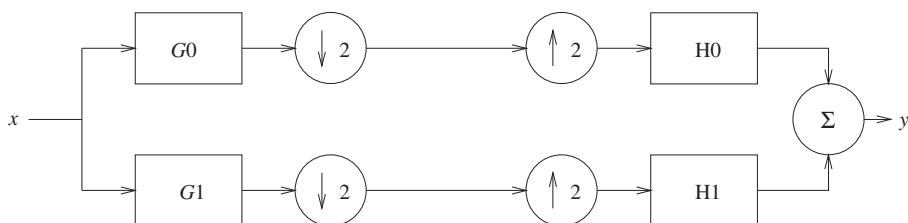
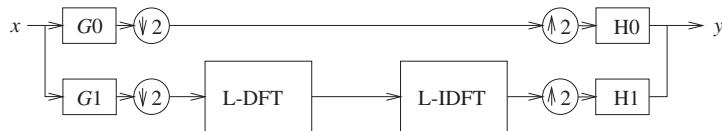
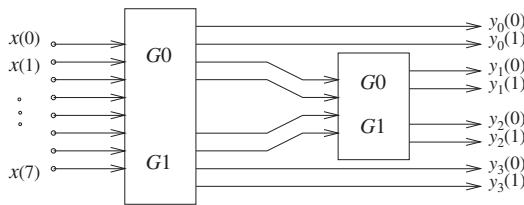


FIGURE 7.9.1

Perfect reconstruction filterbank.

**FIGURE 7.9.2**

Cascaded WPT and DFT.

**FIGURE 7.9.3**

An extension of a WPT: $x(0)$ through $x(7)$ are input signal samples; y_0 through y_3 are the output signals.

Instead of cascaded filterbanks, we can use the DFT structure repeatedly to obtain various TF plane partitions, resulting in *block transform packets*.

Different arrangements of the PRFBs, such as shown in Fig. 7.9.3, can be used to obtain more extensions of the WPT. The box labeled “G0 G1” is an analysis FB like the one shown in Fig. 7.9.1.

7.9.4 SUMMARY AND CONCLUSIONS

In this section, we have seen how two frame structures, namely product-function frames and cascaded frames, generalize the popular transform techniques for joint-domain representation: discrete-time Gabor transforms for periodic and nonperiodic spaces, and the discrete-time wavelet transforms. It was demonstrated that the popular discrete joint-domain representation techniques, namely the discrete-time Gabor and wavelet transforms, are special cases of these structures. Several other transforms, namely the product-transform, the scale-modulation transform, and various extensions of the discrete wavelet transforms, can be derived as special cases of these structures. These two structures result in a variety of discrete transforms, giving a user more freedom in choosing a transform better suited for a particular type of signal. For further details about the topic of product-function frames see Ref. [85]. Related material in this book includes Sections 4.1, 4.5, and 7.1.

REFERENCES

- [1] J. Bertrand, P. Bertrand, “Time-frequency analysis based on the affine group”, in: B. Boashash (Ed.), Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, first ed., Article 7.1, Elsevier, Oxford, 2003, pp. 280-289 (Chapter 7).

- [2] L. Feuvret, N. Georges, J.-J. Mazeran, P. Bey, "Conformity index: a review", *Int. J. Radiat. Oncol. Biol. Phys.* 64 (2) (2006) 333-342, ISSN 0360-3016, <http://dx.doi.org/10.1016/j.ijrobp.2005.09.028>.
- [3] M. Reed, B. Simon, Methods of Modern Mathematical Physics: vol. 2: Fourier Analysis, Self-Adjointness, Academic Press, San Diego, 1975.
- [4] J. Bertrand, P. Bertrand, Représentations temps-fréquence des signaux, vol. 299, Ser. 1, Comptes rendus de l'Académie des Sciences, Paris, 1984, pp. 635-638.
- [5] J. Bertrand, P. Bertrand, "A tomographic approach to Wigner's function", *Found. Phys.* 17 (1987) 397-405.
- [6] K. Vogel, H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase", *Phys. Rev. A* 40 (1989) 2847-2849.
- [7] D.T. Smithey, M. Beck, M.G. Raymer, A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum", *Phys. Rev. Lett.* 70 (1993) 1244-1247.
- [8] E. Sejdić, I. Djurović, L. Stanković, "Fractional Fourier transform as a signal processing tool: an overview of recent developments", *Signal Process.* 91 (6) (2011) 1351-1369.
- [9] M.A. Man'ko, V.I. Man'ko, R.V. Mendes, "Tomograms and other transforms: a unified view", *J. Phys. A Math. Gen.* 34 (40) (2001) 8321-8332.
- [10] J. Bertrand, P. Bertrand, "Affine time-frequency distributions", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 118-140 (Chapter 5).
- [11] P. Gonçalvès, R.G. Baraniuk, "Pseudo affine Wigner distributions: definition and kernel formulation", *IEEE Trans. Signal Process.* 46 (6) (1998) 1505-1517.
- [12] J. Bertrand, P. Bertrand, "Symbolic calculus on the time-frequency half-plane", *J. Math. Phys.* 39 (8) (1998) 4071-4090.
- [13] P. Flandrin, Time-Frequency/Time-Scale Analysis, Academic Press, San Diego, 1999, translated by J. Stöckler from the French editions, *Temps-fréquence* (Paris: Hermès, 1993, 1998).
- [14] K. Kodera, C. de Villedeary, R. Gendrin, "A new method for the numerical analysis of nonstationary signals", *Phys. Earth Planet. Inter.* 12 (1976) 142-150.
- [15] F. Auger, P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Process.* 43 (5) (1995) 1068-1089.
- [16] P. Flandrin, E. Chassande-Mottin, P. Abry, "Reassigned scalograms and their fast algorithms", in: Proc. SPIE: Wavelet Applications in Signal and Image Processing III, vol. 2569, Soc. of Photo-Optical Instrumentation Engineers, 1995, pp. 152-158.
- [17] E. Chassande-Mottin, F. Auger, P. Flandrin, "Supervised time-frequency reassignment", in: Proc. IEEE-SP Internat. Symp. on Time-Frequency and Time-Scale Analysis, 1996, pp. 517-520.
- [18] E. Chassande-Mottin, I. Daubechies, F. Auger, P. Flandrin, "Differential reassignment", *IEEE Signal Process. Lett.* 4 (10) (1997) 293-294.
- [19] D. Friedman, "Instantaneous frequency distribution vs. time: an interpretation of the phase structure of speech", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'85), 1985, pp. 1121-1124.
- [20] R. Carmona, W.L. Hwang, B. Torrésani, Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S, Academic Press, San Diego, 1998.
- [21] P. Guillemain, R. Kronland-Martinet, "Horizontal and vertical ridges associated to continuous wavelet transforms", in: Proc. IEEE-SP Internat. Symp. on Time-Frequency and Time-Scale Analysis, 1992, pp. 63-66.
- [22] S. Maes, "The synchrosqueezed representation yields a new reading of the wavelet transform", in: Proc. SPIE: Wavelet Applications II, vol. 2491, Soc. of Photo-Optical Instrumentation Engineers, 1995, pp. 532-559.
- [23] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: an overview", *IEEE Signal Process. Mag.* 30 (6) (2013) 32-41.
- [24] D. Gabor, "Theory of communication", *J. IEE* 93(III) (26) (1946) 429-457.

- [25] D.E. Vakman, Sophisticated Signals and the Uncertainty Principle in Radar, Springer, New York, 1968, translated by K.N. Trirogoff; edited by E. Jacobs.
- [26] D.L. Jones, T.W. Parks, “A high-resolution data-adaptive time-frequency representation”, *IEEE Trans. Acoust. Speech Signal Process.* 38 (12) (1990) 2127-2135.
- [27] T.-H. Sang, W.J. Williams, “Rényi information and signal-dependent optimal kernel design”, in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP’95), vol. 2, 1995, pp. 997-1000.
- [28] W.J. Williams, M.L. Brown, A.O. Hero III, “Uncertainty, information and time-frequency distributions”, in: Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations II, vol. 1566, Soc. of Photo-Optical Instrumentation Engineers, 1991, pp. 144-156.
- [29] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, O.J.J. Michel, “Measuring time-frequency information content using the Rényi entropies”, *IEEE Trans. Inf. Theory* 47 (4) (2001) 1391-1409.
- [30] P. Flandrin, R.G. Baraniuk, O. Michel, “Time-frequency complexity and information”, vol. 3, pp. 329-332.
- [31] L. Stanković, “A measure of some time-frequency distributions concentration”, *Signal Process.* 81 (3) (2001) 621-631.
- [32] R.G. Baraniuk, “Compressive sensing”, *IEEE Signal Process. Mag.* 24 (4) (2007) 118-121.
- [33] E.J. Candès, M.B. Wakin, “An introduction to compressive sampling”, *IEEE Signal Process. Mag.* 25 (2) (2008) 21-30.
- [34] L. Stanković, I. Orović, S. Stanković, M.G. Amin, “Compressive sensing based separation of non-stationary and stationary signals overlapping in time-frequency”, *IEEE Trans. Signal Process.* 61 (18) (2013) 4562-4572.
- [35] L. Stanković, M. Daković, T. Thayaparan, Time-Frequency Signal Analysis with Applications, Artech House, Boston, 2013.
- [36] L. Stanković, S. Stanković, M.G. Amin, “Missing samples analysis in signals for applications to L-estimation and compressive sensing”, *Signal Process.* 94 (2014) 401-408.
- [37] L. Stanković, “Highly concentrated time-frequency distributions: pseudo-quantum signal representation”, *IEEE Trans. Signal Process.* 45 (3) (1997) 543-551.
- [38] L. Stanković, “An analysis of some time-frequency and time-scale distributions”, *Ann. Télécommun.* 49 (9/10) (1994) 505-517.
- [39] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [40] M. Abed, A. Belouchrani, M. Cheriet, B. Boashash, “Time-frequency distributions based on compact support kernels: properties and performance evaluation”, *IEEE Trans. Signal Process.* 60 (6) (2012) 2814-2827.
- [41] V. Sucic, N. Saulig, B. Boashash, “Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy”, *EURASIP J. Adv. Signal Process.* 2011 (1) (2011) <http://dx.doi.org/10.1186/1687-6180-2011-125>, article ID 2011:125, 11 pp.
- [42] B. Boashash, V. Sucic, “Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions”, *IEEE Trans. Signal Process.* 51 (5) (2003) 1253-1263.
- [43] V. Sucic, B. Boashash, “Parameter selection for optimising time-frequency distributions and measurements of time-frequency characteristics of non-stationary signals”, in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP’01), vol. 6, 2001, pp. 3557-3560.
- [44] V. Sucic, B. Boashash, “Selecting the optimal time-frequency distribution for real-life multicomponent signals under given constraints”, in: Proc. Eleventh European Signal Processing Conf. (EUSIPCO-02), vol. 1, 2002, pp. 141-144.
- [45] V. Sucic, B. Boashash, “An approach for selecting a real-life signal best-performing time-frequency distribution”, in: Proc. Seventh Int. Symp. on Signal Processing and Its Applications (ISSPA’03), vol. 1, 2003, pp. 125-128.

- [46] N.A. Khan, I.A. Taj, M.N. Jaffri, S. Ijaz, "Cross-term elimination in Wigner distribution based on 2D signal processing techniques", *Signal Process.* 91 (3) (2011) 590-599.
- [47] L. Stanković, "A method for time-frequency analysis", *IEEE Trans. Signal Process.* 42 (1) (1994) 225-229.
- [48] L.J. Rankine, M. Mesbah, B. Boashash, "IF estimation for multicomponent signals using image processing techniques in the time-frequency domain", *Signal Process.* 87 (6) (2007) 1234-1250.
- [49] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Dig. Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [50] E. Sejdic, U. Ozertem, I. Djurovic, D. Erdogan, "A new approach for the reassignment of time-frequency representations", in: IEEE International Conference on Acoustics, Speech and Signal Processing, ISSN 1520-6149, 2009, pp. 2997-3000, <http://dx.doi.org/10.1109/ICASSP.2009.4960254>.
- [51] M. Vetterli, P. Marziliano, T. Blu, "Sampling signals with finite rate of innovation", *IEEE Trans. Signal Process.* 50 (6) (2002) 1417-1428.
- [52] D.L. Donoho, "Compressed sensing", *IEEE Trans. Inf. Theory* 52 (4) (2006) 1289-1306.
- [53] W. Dai, O. Milenković, "Subspace pursuit for compressive sensing signal reconstruction", *IEEE Trans. Inf. Theory* 55 (5) (2009) 2230-2249.
- [54] S. Stanković, I. Orović, E. Sejdić, *Multimedia Signals and Systems*, Springer, NY, 2012.
- [55] P. Flandrin, P. Borgnat, "Time-frequency energy distributions meet compressed sensing", *IEEE Trans. Signal Process.* 58 (6) (2010) 2974-2982.
- [56] E. Sejdić, I. Djurović, J. Jiang, "Time-frequency feature representation using energy concentration: an overview of recent advances", *Dig. Signal Process.* 19 (1) (2009) 153-183.
- [57] S.G. Mallat, Z. Zhang, "Matching pursuits with time-frequency dictionaries", *IEEE Trans. Signal Process.* 41 (12) (1993) 3397-3415.
- [58] D. Needell, J.A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples", *Appl. Comput. Harmonic Anal.* 26 (3) (2009) 301-321.
- [59] S. Senay, L.F. Chaparro, L. Durak, "Reconstruction of nonuniformly sampled time-limited signals using prolate spheroidal wave functions", *Signal Process.* 89 (12) (2009) 2585-2595.
- [60] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, L. Coulot, "Sparse sampling of signal innovations", *IEEE Signal Process. Mag.* 25 (2) (2008) 31-40.
- [61] E. Sejdić, M. Luccini, S. Primak, K. Baddour, T. Willink, "Channel estimation using DPSS based frames", in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'08), 2008, pp. 2849-2852.
- [62] E. Sejdić, A. Can, L.F. Chaparro, C.M. Steele, T. Chau, "Compressive sampling of swallowing accelerometry signals using time-frequency dictionaries based on modulated discrete prolate spheroidal sequences", *EURASIP J. Adv. Signal Process.* 2012 (1) (2012) <http://dx.doi.org/10.1186/1687-6180-2012-101>, article ID 2012:101, 14 pp.
- [63] D. Boutana, M. Benidir, B. Barkat, "Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis", *IET Signal Process.* 5 (6) (2011) 527-537.
- [64] A. Rényi, "On measures of entropy and information", in: Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1961, pp. 547-561.
- [65] S. Aviyente, W.J. Williams, "Minimum entropy time-frequency distributions", *IEEE Signal Process. Lett.* 12 (1) (2005) 37-40.
- [66] C.E. Shannon, "A mathematical theory of communication—parts I and II", *Bell Syst. Tech. J.* 27 (3) (1948) 379-423.
- [67] S. Kadambe, G.F. Boudreux-Bartels, "A comparison of the existence of 'cross terms' in the Wigner distribution and the squared magnitude of the wavelet transform and the short-time Fourier transform", *IEEE Trans. Signal Process.* 40 (10) (1992) 2498-2517.

- [68] N. Saulig, N. Pustelnik, P. Borgnat, P. Flandrin, V. Sucic, "Instantaneous counting of components in nonstationary signals", in: Proc. 21st European Signal Processing Conf. (EUSIPCO-13), 2013, Paper No. 1569742487, 5 pp.
- [69] B. Boashash, G. Azemi, N.A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: applications to newborn EEG abnormality detection", *Pattern Recogn.* 48 (3) (2015) 616-627.
- [70] I. Shafi, J. Ahmad, S.I. Shah, F.M. Kashif, "Techniques to obtain good resolution and concentrated time-frequency distributions: a review", *EURASIP J. Adv. Signal Process.* 2009 (2009) <http://dx.doi.org/10.1155/2009/673539>, article ID 673539, 43 pp.
- [71] D.J.C. MacKay, "A practical Bayesian framework for backpropagation networks", *Neural Comput.* 4 (3) (1992) 448-472.
- [72] I. Shafi, J. Ahmad, S.I. Shah, F.M. Kashif, "Evolutionary time-frequency distributions using Bayesian regularised neural network model", *IET Signal Process.* 1 (2) (2007) 97-106.
- [73] I. Shafi, J. Ahmad, S.I. Shah, A.A. Ikram, A.A. Khan, S. Bashir, "Validity-guided fuzzy clustering evaluation for neural network-based time-frequency reassignment", *EURASIP J. Adv. Signal Process.* 2010 (2010) <http://dx.doi.org/10.1155/2010/636858>, article ID 636858, 14 pp.
- [74] A.E. Ruano (Ed.), Intelligent Control Systems Using Computational Intelligence Techniques, IEE Control Engineering Series, vol. 70, IEE/IET, London, 2005.
- [75] I. Shafi, J. Ahmad, S.I. Shah, F.M. Kashif, "Impact of varying neurons and hidden layers in neural network architecture for a time frequency application", in: Proc. 10th IEEE International Multitopic Conference (INMIC'06), 2006, pp. 188-193.
- [76] I. Shafi, J. Ahmad, S.I. Shah, F.M. Kashif, "Computing de-blurred time-frequency distributions using artificial neural networks", *Circ. Syst. Signal Process.* 27 (3) (2008) 277-294.
- [77] D.L. Jones, T.W. Parks, "A resolution comparison of several time-frequency representations", *IEEE Trans. Signal Process.* 40 (2) (1992) 413-420.
- [78] I. Shafi, J. Ahmad, S.I. Shah, A.A. Ikram, A.A. Khan, S. Bashir, F.M. Kashif, "High-resolution time-frequency methods' performance analysis", *EURASIP J. Adv. Signal Process.* 2010 (2010) <http://dx.doi.org/10.1155/2010/806043>, article ID 806043, 7 pp.
- [79] J.M. Morris, S.M. Joshi, "Joint-domain representations via discrete-domain frames", in: B. Boashash (Ed.), Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, first ed., Article 7.5, Elsevier, Oxford, 2003, pp. 315-322 (Chapter 7).
- [80] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", *IEEE Trans. Inf. Theory* 36 (5) (1990) 961-1005.
- [81] S.M. Joshi, J.M. Morris, "Some results on product-function frames", *Signal Process.* 80 (4) (2000) 737-740.
- [82] H. Bölcskei, F. Hlawatsch, "Oversampled modulated filter banks", in: H.G. Feichtinger, T. Strohmer (Eds.), Gabor Analysis and Algorithms: Theory and Applications, chap. 9, Birkhäuser, Berlin/Boston, 1998, pp. 295-322.
- [83] M. Zibulski, Y.Y. Zeevi, "Discrete multiwindow Gabor-type transforms", *IEEE Trans. Signal Process.* 45 (6) (1997) 1428-1442.
- [84] G. Strang, T.Q. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996.
- [85] S.M. Joshi, J.M. Morris, "Product-function frames in $\ell^2(\mathbb{Z})$ ", *IEEE Trans. Inf. Theory* 49 (5) (2003) 1336-1342.

MULTISENSOR, MULTICHANNEL, AND TIME-SPACE PROCESSING

8

INTRODUCTION AND OVERVIEW

This chapter presents time-frequency (t,f) methods suitable for multichannel signal processing using multisensor and time-space processing methods. The topic is covered in seven sections with relevant cross-referencing.

A brief tutorial review of the topic of multichannel/multisensor (t,f) signal processing describes the extension of (t,f) methods to incorporate the spatial diversity information provided by multisensor recordings; this is illustrated on an application to brain electroencephalogram (EEG) abnormality source localization ([Section 8.1](#)). The multichannel data can then be processed with time-frequency distributions (TFDs) for channel estimation and equalization. In blind source separation (BSS) and direction of arrival (DOA) estimation problems, the (t,f) approach to array signal processing leads to improved spatial resolution and source separation performances. Methods include (t,f) multiple signal classification (MUSIC), and TFD-based BSS ([Section 8.2](#)). In sensor array processing, for source localization, TFDs provide a good framework for hypothesis testing, and they allow the optimal detector to be implemented naturally and efficiently ([Section 8.3](#)). In underwater acoustics and telecommunications, separation of signal mixtures is traditionally based on methods such as independent component analysis (ICA) or BSS. These can be formulated using TFDs for dealing with the case when the signals are nonstationary ([Section 8.4](#)). In the underdetermined case, the (t,f) formulations, methodologies, and algorithms for BSS are implemented using vector clustering and component extraction ([Section 8.5](#)). Then, [Section 8.6](#) describes a method where audio source localization and separation can be improved using multisensor (t,f) analysis. Finally, in [Section 8.7](#), a selection of basic algorithm and MATLAB code is provided so as to allow the reader to easily reproduce the results in this chapter.

8.1 MULTISENSOR TIME-FREQUENCY ANALYSIS AND PROCESSING: METHODS FOR MULTICHANNEL NONSTATIONARY DATA⁰

In many engineering applications, a collection of related measurements is recorded from multiple sensors. Examples include multiple antennas used in telecommunications and radar, multiple electrodes used in medical science, and multiple acoustic sensors used for source localization in sonar, geophysics, and speech processing. In these applications, the multiple sensors combine to improve the quality of information about the phenomena producing the signals. Such sensors are typically placed at different physical locations (spatially apart) to exploit spatial diversity and to determine spatial characteristics of the underlying process. For example, in a radar/sonar application, multiple sensors forming an array is used to determine the direction of a returning echo and locate the position of a target. Multiple sensors are also used to obtain multichannel electroencephalogram (EEG) recordings from the scalp. Previous studies found that the performance could be further enhanced by combining both the nonstationary characteristics of measured signals using quadratic time-frequency (t,f) distributions (QTFDs) (see Sections 2.1 and 3.1–3.4) and the spatial information provided by sensor arrays [1].

Multisensor or multichannel QTFDs are also often called spatial time-frequency distributions (STFDs). These techniques can solve array processing problems such as DOA estimation with improved resolution, using spatial information for the (t,f) processing of multichannel nonstationary signals.

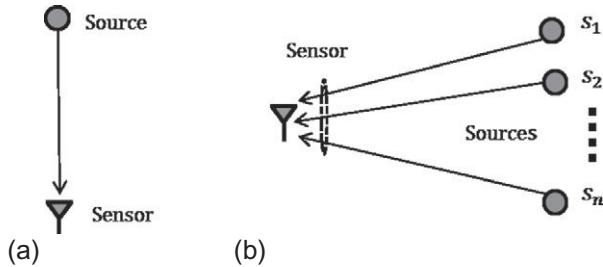
8.1.1 EXTENSION OF SINGLE-SENSOR QTFDs TO MULTISENSOR QTFDs

The previous chapters considered nonstationary signals recorded by a single sensor. In some cases, only one source produced the signal received by the sensor, noting that a single source can generate a multicomponent signal (e.g., see whale signal in Fig I.1.1 in Chapter I). In other cases, several sources generate different components that merge into one signal recorded by the sensor (see Fig. 8.1.1). These two cases are called “single input single output (SISO)” and “multiple input single output (MISO).” The (t,f) problem of analyzing multicomponent signals may therefore be considered as a problem of source separation in the case of just one sensor [2]. Multisensor (array processing) deals effectively with this case as well as the more complex case of multisource and multisensor. These cases are called “single input multiple output (SIMO)” and “multiple input multiple output (MIMO).” This section establishes the foundation for the extension of single-sensor QTFDs to multisensor QTFDs.

Let us consider a vector $\mathbf{z}(t) = [z_1(t), z_2(t), \dots, z_m(t)]^T$ constituted of the analytic associates of $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$. The class of quadratic multichannel TFDs (STFDs) is then defined as

$$\begin{aligned} \rho_{zz}(t,f) &= \underset{\tau \rightarrow f}{\mathcal{F}} \left\{ G(t,\tau) * \mathbf{K}_{zz}(t,\tau) \right\} \\ &= \begin{bmatrix} \rho_{z_1,z_1}(t,f) & \rho_{z_1,z_2}(t,f) & \cdots & \rho_{z_1,z_m}(t,f) \\ \rho_{z_2,z_1}(t,f) & \rho_{z_2,z_2}(t,f) & \cdots & \rho_{z_2,z_m}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{z_m,z_1}(t,f) & \rho_{z_m,z_2}(t,f) & \cdots & \rho_{z_m,z_m}(t,f) \end{bmatrix}, \end{aligned} \quad (8.1.1)$$

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Ali**, Qatar University, Doha, Qatar (sadiq@qu.edu.qa). Reviewers: A. Aissa-El-Bey and Y.D. Zhang.

**FIGURE 8.1.1**

(a) Single source single sensor (SISO) and (b) multiple sources single sensor (MISO).

where $G(t, \tau)$ is the time-lag kernel defined in Eq. (2.7.7) and the time convolution operator $*$ is applied to each entry of the spatial instantaneous correlation matrix $\mathbf{K}_{zz}(t, \tau)$ with elements $K_{z_i z_j}(t, \tau) = z_i(t + \frac{\tau}{2}) z_j^*(t - \frac{\tau}{2})$ $i, j = 1, 2, \dots, m$ being the instantaneous auto/cross-correlation functions defined in Eq. (2.1.18) such that

$$\mathbf{K}_{zz}(t, \tau) = \mathbf{z}(t + \tau/2) \mathbf{z}^H(t - \tau/2),$$

$$= \begin{bmatrix} K_{z_1 z_1}(t, \tau) & K_{z_1 z_2}(t, \tau) & \cdots & K_{z_1 z_m}(t, \tau) \\ K_{z_2 z_1}(t, \tau) & K_{z_2 z_2}(t, \tau) & \cdots & K_{z_2 z_m}(t, \tau) \\ \vdots & \vdots & \ddots & \vdots \\ K_{z_m z_1}(t, \tau) & K_{z_m z_2}(t, \tau) & \cdots & K_{z_m z_m}(t, \tau) \end{bmatrix}, \quad (8.1.2)$$

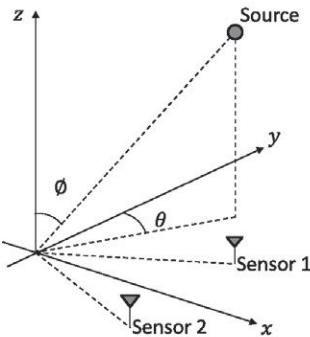
In the STFD matrix $\rho_{zz}(t, f)$ (Eq. 8.1.1), diagonal terms are called auto-TFDs and the quadratic class of auto-TFD of $x_i(t)$ is defined as (Section 3.1.2)

$$\rho_{z_i z_i}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t - u, \tau) z_i(u + \tau/2) z_i^*(u - \tau/2) e^{-j2\pi\tau f} du d\tau. \quad (8.1.3)$$

Similarly, the off-diagonal terms are called cross-TFDs. The cross-TFD of two signals $x_i(t)$ and $x_j(t)$ is defined as (see Eq. 3.1.21)

$$\rho_{z_i z_j}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t - u, \tau) z_i(u + \tau/2) z_j^*(u - \tau/2) e^{-j2\pi\tau f} du d\tau. \quad (8.1.4)$$

In addition to the auto-TFDs encountered in the monocomponent single channel, multicomponent or multichannel (t, f) analysis has both auto-TFDs and cross-TFDs. Thus, there are two types of cross-terms in multichannel TFDs. The first type of cross-terms is associated with auto-TFDs and they reside along the auto-terms on the main diagonal of the source TFD matrix. These cross-terms are formed by the interactions of components of the same source signal. The second type of cross-terms is associated with cross-TFDs of different source signals as they result from interactions between signal components of two different sources [3].

**FIGURE 8.1.2**

Two sensors and a point source.

8.1.2 FUNDAMENTALS OF ARRAY PROCESSING

A simple example of sensor array with two sensors and one source is shown in Fig. 8.1.2 where the angles θ and ϕ denote the azimuth and elevation angles of the source, respectively. It is assumed that the sources are points in space such as a transmitter in a wireless communications system, an electromagnetic energy from a target in a radar system or acoustic energy from a sonar system.

The signal energy corresponds to a wave propagating radially outward from the source location. If the physical size of the sensor array is very small compared to the distance between the source and sensor array (i.e., the source is in the far-field), the received wavefront may be considered as a plane across the array. Assuming the far-field scenario, the sources and array are coplanar, implying that the azimuth angle θ is the only relevant spatial parameter of a source.

8.1.2.1 Instantaneous mixing model

Figure 8.1.3 shows n source signals $\{s_1, s_2, \dots, s_n\}$ arriving at an m element multisensor array (uniform linear array, ULA). The linear data model is given as

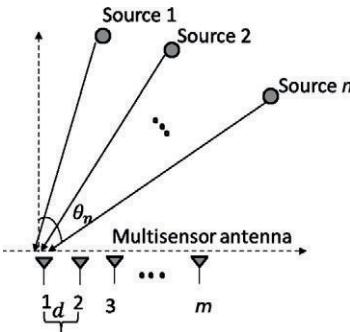
$$\mathbf{x}(t) = \mathbf{As}(t) + \mathbf{n}(t), \quad (8.1.5)$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ is the signal vector received at the M multiple sensors, vector $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$ contains the source signals, and $\mathbf{n}(t)$ is an additive noise vector whose entries are modeled generally as stationary, temporally, and spatially white random processes, and independent of the source signals. The $m \times n$ matrix \mathbf{A} represents the propagation matrix or mixing matrix, having n column vectors called steering vectors. The matrix \mathbf{A} contains information on the DOAs of the different signals

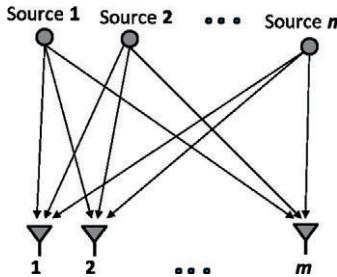
$$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_n)], \quad (8.1.6)$$

where $\mathbf{a}(\theta_i)$ for $i = 1, 2, \dots, n$ is the steering vector of the array for direction θ_i .

Equation (8.1.5) indicates that each sensor receives a contribution from each source resulting in the observations $\{x_i(t)\}_{i=1}^m$. This mixture of n signals is shown in Fig. 8.1.4; the elements $\{x_i(t)\}_{i=1}^m$ of vector $\mathbf{x}(t)$ are linear superpositions of the source signals. In the simple case where the sensors receive signals

**FIGURE 8.1.3**

Multisensor array and multiple sources.

**FIGURE 8.1.4**

Instantaneous mixing model.

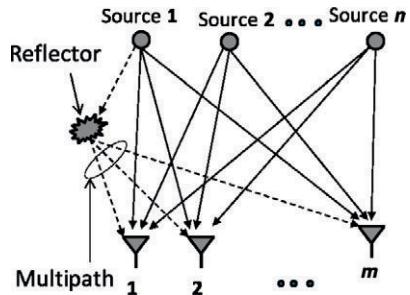
from sources only through a single path (i.e., line-of-sight) and the observations $\{x_i(t)\}_{i=1}^m$ depend on the values of \$n\$ source signals at the same time, such mixing process is referred to as instantaneous mixing. The covariance matrix for Eq. (8.1.5) has the structure [4]

$$\mathbf{R}_{xx}(\tau) = \mathbf{A}\mathbf{R}_{ss}(\tau)\mathbf{A}^H + \sigma^2\mathbf{I}_m, \quad (8.1.7)$$

where the superscript \$H\$ denotes conjugate transpose, \$\tau\$ is the time lag, \$\mathbf{A}\$ is the steering matrix, and \$\mathbf{I}\$ is the \$m \times m\$ identity matrix.

When the signals received at the sensors are stationary (case of Eq. (8.1.7)), the correlation function taken at different times remains constant and \$\mathbf{R}_{xx}(\tau)\$ is only a function of the time lag. For stationary signals, the *Wiener-Khintchine theorem* states that the Fourier transform (FT) of the auto-correlation \$\mathbf{R}_{xx}\$ is the power spectral density (PSD) \$S_{xx}(f)\$, which is related to \$S_{ss}(f)\$, the source PSD matrices as

$$S_{xx}(f) = \mathcal{F}_{\tau \rightarrow f} \{\mathbf{R}_{xx}(\tau)\} = \mathbf{A}S_{ss}(f)\mathbf{A}^H + \sigma^2\mathbf{I}_m, \quad (8.1.8)$$

**FIGURE 8.1.5**

Convulsive mixing model.

8.1.2.2 Convulsive mixing model

In some situations, the contributions to the mixture may have different delays as shown in Fig. 8.1.5. For example, wireless communication channel models often include multipath propagation and in audio-processing reverberation causes delays. Hence, in such cases the sensor array is convulsive, as signals picked up by the sensors consist of direct-path signals as well as their delayed (reflected) and attenuated versions in the presence of noise. The convulsive mixing model is then given as

$$\mathbf{x}(t) = \sum_{k=0}^K \mathbf{A}(k) \mathbf{s}(t-k) + \mathbf{n}(t) = \mathbf{A} * \mathbf{s}(t) + \mathbf{n}(t), \quad (8.1.9)$$

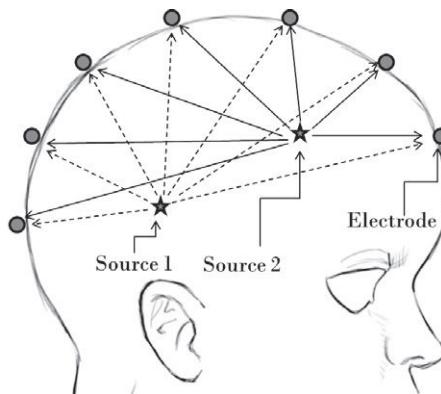
where $*$ denotes the convolution. The received signals are mixtures of filtered (i.e., K -tapped filter) versions of the original source signals; the coefficients $(\mathbf{A}(k))_{i,j}$ denote the response of sensor i at time k to an impulse in source j at time 0.

8.1.2.3 Example: Source localization of brain EEG abnormality

The EEG measures the electrical activity produced by the nerve cells in the brain. The recordings are made using multiple electrodes placed on the scalp. Based on these multichannel EEG measurements (shown in Fig. 8.1.6), source localization can be performed to determine the location and strength of the contributing sources as well as their propagation path. The relationship between neural current sources and multichannel EEG measurements is governed by quasistatic approximations of Maxwell's equations, which reduces to a Poisson equation as [5]

$$\nabla \cdot (\mu \nabla \varphi) = \nabla \cdot \mathbf{J}^s, \quad \text{in } \Omega, \quad (8.1.10)$$

where ∇ is a partial differential vector operator, μ is the electrical conductivity, φ are the electrical potentials, and \mathbf{J}^s are the electric current sources. Equation (8.1.10) represents the distribution of the electric potential φ over the head volume Ω due to the primary current \mathbf{J}^s in the cortex of the human brain; it indicates that for a given configuration of electric sources, the mapping from electric sources within the head to the scalp recordings on the outside of the scalp can be represented by a lead-field matrix [6]. The columns of the lead-field matrix represent the signal vectors that are observed at the scalp.

**FIGURE 8.1.6**

View of the electrode positions in an arbitrary display to get EEG measurements.

Let us define $x_i(t)$ as the electric field measured at sensor i at time instant t . The vector $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ is the set of measurements collected by m sensors. Let us assume n current dipole sources generate the electric field and $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ represent the locations of these sources. The magnitude of the $\{i\text{th}\}_{i=1}^n$ dipole source movement is $s_i(t)$ and the source magnitude vector is defined as $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. For m sensors and n dipole sources, the relationship between $\mathbf{x}(t)$ and $\mathbf{s}(t)$ can be expressed by Eq. (8.1.5) using matrix \mathbf{A} , which is a lead-field matrix of dimension $m \times n$ [6] that includes both the effect of location \mathbf{A} and orientation Φ of the dipoles as $\mathbf{A} = \mathbf{A}\Phi$. Each column of the lead-field matrix is called lead field and it defines the current flow for a given sensor through each dipole position [6].

Hence, the array model can be used to solve the problem of EEG source localization.

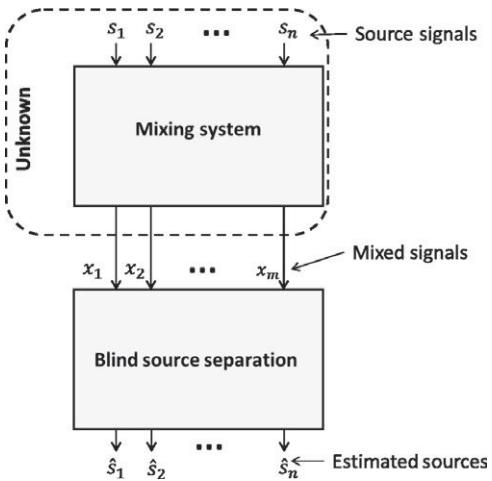
8.1.3 FOUNDATIONS OF SOURCE SEPARATION AND LOCALIZATION

8.1.3.1 *Blind source separation*

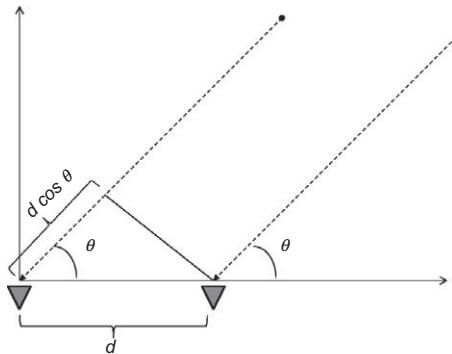
Considering the signal model of Eq. (8.1.5), the simple *source separation* problem is to extract the source signal vector \mathbf{s} by assuming that both mixing matrix \mathbf{A} and observation vector \mathbf{x} are known. Whereas, the blind source separation (BSS) problem (depicted in Fig. 8.1.7) is the joint estimation of both mixing matrix \mathbf{A} and source signal vector \mathbf{s} for a given observation vector \mathbf{x} [3]. BSS techniques can be used in EEG applications to recover unknown dipole current sources, localize abnormal EEG sources, and remove artifacts from EEG [5]. Moreover, BSS techniques can be extended to nonstationary signals using multichannel (or spatial) TFDs, resulting in improved performance when compared to conventional time domain techniques (see details in Sections 8.2, 8.4, and 8.5).

8.1.3.2 *Direction of arrival estimation*

(1) *Basic principle of DOA estimation.* Let us consider a simple scenario of an array with only two elements as shown in Fig. 8.1.8. Let the distance between the two elements be d , the speed of

**FIGURE 8.1.7**

Blind source separation.

**FIGURE 8.1.8**

Basic principle of DOA estimation.

light c , the time delay between the elements Δt , and the incident angle of the far-field signal is θ . The delay in the signal received at the two antennas' elements due to the path difference is

$$\Delta t = \frac{d \cos(\theta)}{c}. \quad (8.1.11)$$

In other words, it takes Δt seconds more for one signal to reach an antenna in the two-element array relative to the first one. In the frequency domain, the delays appear as phase shift in the received signals by the elements as

$$e^{-j\omega\Delta t} = e^{-j2\pi f_0 \left(\frac{d}{c} \right) \cos(\theta)} = e^{-j2\pi f_0 \left(\frac{d \cos(\theta)}{\lambda f_0} \right)} = e^{-j2\pi \left(\frac{d}{\lambda} \right) \cos(\theta)}, \quad (8.1.12)$$

where f_0 is the center frequency and λ is the wavelength of the signal [4]. Given the geometry of the array, the delays or phase differences can be used to estimate the incident angle. Therefore, if the time delay of the signal is known, the direction of the signal can be estimated from Eq. (8.1.11), which is the basic principle of spatial spectrum estimation techniques. Extending the discussion, the steering vector for m antennas array given in Eq. (8.1.6) can be defined as

$$\mathbf{a}(\theta_i) = \left[1, e^{-j2\pi\left(\frac{d}{\lambda}\right)\cos(\theta_i)}, e^{-j2\pi\left(\frac{2d}{\lambda}\right)\cos(\theta_i)}, \dots, e^{-j2\pi\left(\frac{(m-1)d}{\lambda}\right)\cos(\theta_i)} \right]^T, \text{ where } \{\theta_i\}_{i=1}^n \text{ is the angle of arrival of the } i\text{th source signal.}$$

- (2) *Subspace-based localization.* The covariance matrix of the observation vector

$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ is given by Eq. (8.1.7). Let us assume that $\mathbf{R}_{ss} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ is a full-rank matrix by assuming noncoherence of the n incoming signals, and that the set of n vectors in \mathbf{A} are linearly independent. The eigenvalue decomposition of the covariance matrix can exploit the fact that the DOAs determine the eigen structure of the matrix. Let

$\varrho_1 \geq \varrho_2 \geq \dots \geq \varrho_m$ denote the eigenvalues of the matrix \mathbf{R}_{xx} , and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ denote the eigenvalues of the matrix $\mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H$, respectively. Equation (8.1.7) yields

$\varrho_i = \lambda_i + \sigma^2$, $i = 1, 2, \dots, m$. Since, the matrix \mathbf{A} is full-column rank n , the $(m - n)$ smallest eigenvalues of \mathbf{R}_{xx} are equal to σ^2

$$\varrho_i = \begin{cases} \lambda_i + \sigma^2 & i = 1, 2, \dots, n, \\ \sigma^2 & i = n+1, n+2, \dots, m. \end{cases} \quad (8.1.13)$$

The eigenvalue decomposition of the covariance matrix \mathbf{R}_{xx} can then be written as

$$\mathbf{R}_{xx} = \sum_{i=1}^n (\lambda_i + \sigma^2) \mathbf{v}_i \mathbf{v}_i^H + \sum_{i=n+1}^m \sigma^2 \mathbf{v}_i \mathbf{v}_i^H, \quad (8.1.14)$$

where $\mathbf{v}_i^H \mathbf{v}_j = \delta_{ij}$ are the orthogonal eigenvectors of the matrix \mathbf{R}_{xx} (i.e., $\mathbf{R}_{xx} \mathbf{v}_i = \varrho_i \mathbf{v}_i$ for $i = 1, 2, \dots, m$). From Eq. (8.1.14), we have $\mathbf{R}_{xx} \mathbf{v}_i = \sigma^2 \mathbf{v}_i$, $i = n+1, n+2, \dots, m$ or, equivalently,

$$(\mathbf{R}_{xx} - \sigma^2 \mathbf{I}) \mathbf{v}_i = \mathbf{0}, \quad i = n+1, n+2, \dots, m. \quad (8.1.15)$$

From Eq. (8.1.7), we have $\mathbf{R}_{xx} - \sigma^2 \mathbf{I} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H$; hence, Eq. (8.1.15) becomes $\mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H \mathbf{v}_i = \mathbf{0}$, $i = n+1, n+2, \dots, m$, from which it follows that

$$\mathbf{A}^H \mathbf{v}_i = \mathbf{0}, \quad i = n+1, n+2, \dots, m. \quad (8.1.16)$$

Equation (8.1.16) indicates that the subspace spanned by the eigenvectors $\{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\}$ is the orthogonal complement of the subspace spanned by the steering vectors in matrix \mathbf{A} .

Therefore, given the eigenvectors of \mathbf{R}_{xx} , one can determine the signal DOAs by searching for those steering vectors that are orthogonal to the noise subspace.

8.1.3.3 Time domain MUSIC algorithm

Spectral-based DOA estimation methods are based on maximizing the power of the beam-forming output for a given input signal. One standard DOA technique known as MUSIC is based on the eigenvalue decomposition of the covariance matrix \mathbf{R}_{xx} . In practice, \mathbf{R}_{xx} is unknown and estimated as

$$\hat{\mathbf{R}}_{xx} = \frac{1}{T} \int_0^T \mathbf{x}(\tau) \mathbf{x}^H(\tau) d\tau, \quad (8.1.17)$$

where T is the duration of the data. The eigenvector estimates are then expressed as $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \dots, \hat{\mathbf{v}}_M\}$. Using Eq. (8.1.16), the MUSIC algorithm estimates the signal directions as the peaks of the spatial spectrum estimate given by [4]

$$P_{\text{MUSIC}}(\theta) = \frac{1}{\sum_{i=n+1}^M |\mathbf{a}(\theta) \hat{\mathbf{v}}_i|^2}, \quad (8.1.18)$$

where $\mathbf{a}(\theta)$ is a column vector of the steering matrix given in Eq. (8.1.6). Note that the MUSIC spectrum is based on a single realization of the stochastic process represented by the snapshots $\mathbf{x}(t)$ for $t = 1, 2, \dots, M$. MUSIC estimates are consistent and they converge to the true source bearings as the number of snapshots grows to infinity. Considering the context of the brain application, where the aim is to estimate the locations of EEG abnormalities, the MUSIC algorithm exploits the fact that the lead-field $\mathbf{A} = \mathbf{A}\Phi$ is orthogonal to the noise subspace of the received covariance matrix [6]. To find the locations of abnormalities, the covariance matrix (to be decomposed to find $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \dots, \hat{\mathbf{v}}_M\}$) is estimated from the received multichannel EEG signals, and instead of using steering vectors, in Eq. (8.1.18) the eigenvectors are projected onto the columns of lead-field matrix \mathbf{A} .

8.1.4 EXTENSION OF ARRAY SIGNAL MODEL TO THE NONSTATIONARY CASE

Conventional stationary array processing methods operate on the covariance matrix of the received observation at the array of sensors; however, in the case of nonstationary signal, the spectral content of the measured signals is time varying. The assumption of stationarity reduces the achievable performances with respect to those that would be achieved in a stationary environment. The following sections extend array processing methods for nonstationary signals.

8.1.4.1 Defining multichannel (spatial) TFDs

Multichannel (spatial) TFDs are TFDs for a vector signal representing the number of channels using a space variable. The three dimensions—namely space, time and frequency—are used to construct a matrix called the STFDs matrix as given in Eq. (8.1.1). This approach is also called space-time-frequency processing. It uses (t, f) information across sensors that are located at different spatial locations to characterize the nonstationary signals. By relaxing the assumption of stationarity, the time-dependent covariance matrix becomes

$$\mathbf{R}_{zz}(t, \tau) = \mathbf{A}\mathbf{R}_{ss}(t, \tau)\mathbf{A}^H + \sigma^2\mathbf{I}. \quad (8.1.19)$$

In such case, we use the extended *Wiener-Khintchine* theorem defined in Eq. (2.2.3); the time-varying power spectrum of a nonstationary signal $\mathbf{z}(t)$ is estimated as the FT of the smoothed time-dependent covariance matrix $\mathbf{R}_{zz}(t, \tau)$,

$$\begin{aligned}\rho_{zz}(t,f) &= \mathcal{F}_{\tau \rightarrow f} \left\{ G(t,\tau) * \mathcal{R}_{zz}(t,\tau) \right\} = \mathcal{F}_{\tau \rightarrow f} \left\{ G(t,\tau) * \left(\mathbf{A} \mathbf{R}_{ss}(t,\tau) \mathbf{A}^H + \sigma^2 \mathbf{I} \right) \right\} \\ &= \mathbf{A} \mathcal{F}_{\tau \rightarrow f} \left\{ G(t,\tau) * \mathcal{R}_{ss}(t,\tau) \right\} \mathbf{A}^H + \sigma^2 \mathbf{I} = \mathbf{A} \rho_{ss}(t,f) \mathbf{A}^H + \sigma^2 \mathbf{I},\end{aligned}\quad (8.1.20)$$

where $\rho_{ss}(t,f) = \mathcal{F}_{\tau \rightarrow f} \left\{ G(t,\tau) * \mathcal{R}_{ss}(t,\tau) \right\}$ is the signal TFD. By replacing the covariance matrices with $\mathbf{R}_{zz}(t,\tau)$ and defining $\rho_{zz}(t,f)$, then the multichannel TFD (i.e., STFD) matrices, Eqs. (8.1.19) and (8.1.20) have similar structures to Eq. (8.1.7).

8.1.4.2 Advantages of the STFD over the covariance matrix approach

In essence, this STFD-based approach increases the effective signal-to-noise ratio (SNR) and provides robustness with respect to noise by spreading the noise power while localizing the source signal power in the (t,f) plane [3]. Indeed, the TFD of the white noise is distributed all over the (t,f) domain, whereas the TFDs of the source waveforms are confined to much smaller regions, as shown in Fig. 8.1.9. Let us consider three sources 1, 2, and 3 incident on a multisensor array. Source 1 occupies the (t,f) region R_1 , source 2 occupies the (t,f) region R_2 , and source 3 occupies the (t,f) region R_3 . The (t,f) signatures of the three sources overlap (i.e., regions R_{12} , R_{13} , R_{23} and R_{123}), but each source still has its own particular (t,f) region that has not overlapped with other sources; however, the noise is spread over R_1 , R_2 , and R_3 , as well as the complement region R_c . When we select (t,f) points for either the averaging or joint diagonalization (JD) approaches that belong to the noise-only region R_c (such as (t_1,f_1)), then no useful information about the sources is available. On the other hand, if we confine the selection of (t,f) points to R_1 or R_2 or R_3 , such as (t_2,f_2) , then only the noise part in these regions is included. The result of leaving out the (t,f) points that are not part of the (t,f) signatures of the signal arrivals enhances the SNR. To be more specific, the STFDs' property of concentrating the input signal energy in its instantaneous bandwidth (IB) and around its instantaneous frequency (IF) while spreading the noise

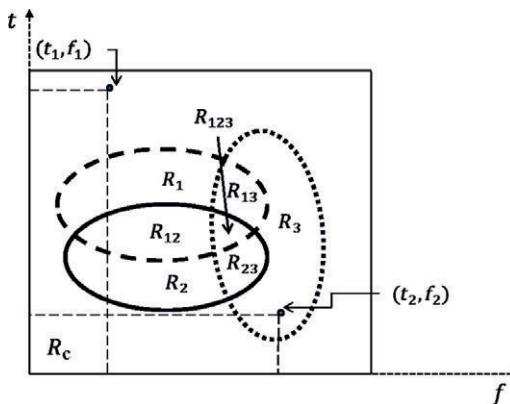


FIGURE 8.1.9

Depiction of three source signals having different (t,f) signatures.

over the entire (t,f) domain increases the effective SNR and proves valuable in many applications. A key point is therefore the selection of (t,f) points in the region of interest.

8.1.4.3 Time-frequency MUSIC algorithm

As discussed in [Section 8.1.4.1](#), the linear model given in [Eq. \(8.1.20\)](#) has the same structure as the covariance matrix-based linear model given in [Eq. \(8.1.7\)](#). In this section, the MUSIC algorithm is extended for direction-finding based on STFD matrices, where we consider the subspace decomposition of an averaged STFD matrix. The steps needed to find the averaged STFD matrix ρ_{zz} are given in [Table 8.1.1](#). After estimating the averaged STFD matrix ρ_{zz} , time-frequency MUSIC (TF-MUSIC) uses the sub-space technique given in [Section 8.1.3](#) to get

$$P_{\text{TF-MUSIC}}(\theta) = \frac{1}{\sum_{i=p+1}^M |\mathbf{a}(\theta) \hat{\mathbf{v}}_{\text{TF},i}|^2}, \quad (8.1.24)$$

where $\hat{\mathbf{v}}_{\text{TF},i}$ is the i th eigenvector of the averaged spatial TFD matrix ρ_{zz} . The difference between the averaged STFD matrix ρ_{zz} and the sample covariance $\hat{\mathbf{R}}_{xx}$ is that the former is achieved by averaging selected high-signal energy points by rejecting noise contributions and the latter is achieved by averaging all available points including noise contributions. Hence, the averaged STFD matrix ρ_{zz} -based directional estimation technique $P_{\text{TF-MUSIC}}(\theta)$ promises an improved performance. This gain is obtained at an additional computational cost of $\mathcal{O}(N_T N_f \log(N_f) m(m+1)/2)$, where N_T is the number of samples and N_f is the number of fast Fourier transform (FFT) points. In addition,

Table 8.1.1: Algorithm to Calculate the Averaged STFD Matrix

1. Calculate the STFD $\rho_{zz}(t,f)$ as given in [Eq. \(8.1.20\)](#).
2. Select auto-TFDs $\{\rho_{z_i,i}(t,f)\}_{i=1}^M$ and calculate the spatial averaged TFD as

$$\rho_{\text{avg}}(t,f) = \frac{1}{M} \sum_{i=1}^M \rho_{z_i,i}(t,f). \quad (8.1.21)$$

3. High-energy (t,f) points are selected as [3]

$$\text{Select the point } (t_i, f_i) \text{ if } \|\rho_{\text{avg}}(t_i, f_i)\| > \epsilon_t \quad (8.1.22)$$

and ignore the (t,f) points with negligible energy (e.g., due to noise) to improve the SNR. The threshold ϵ_t is a user-defined parameter; in this study, the selected value for ϵ_t such that $\epsilon_t \geq 0.05 \times \max(\rho_{\text{avg}})$.

4. Calculate the averaged STFD matrix ρ_{zz} by taking the average of the selected (t,f) points in the previous step as

$$\rho_{zz} = \frac{1}{n_{\text{points}}} \sum_{i=1}^{n_{\text{points}}} \rho_{zz}(t_i, f_i), \quad (8.1.23)$$

where n_{points} indicates the total number of selected (t,f) points.

the (t,f) approach allows us to estimate the covariance matrix of each source separately (i.e., using multicomponent IF approach), and therefore allowing finding estimates of DOAs, even in the case of more sources than sensors (see [Section 8.5](#)).

8.1.4.4 Numerical results

Let us compare the performance of TF-MUSIC and the conventional time-domain MUSIC, for the case of a ULA having $M = 4$ sensors. Two source signals impinge on this array from directions θ_1 and θ_2 , respectively. The SNRs for the two source signals are kept similar at -5 dB. In [Fig. 8.1.10](#), the estimated spatial spectra of the TF-MUSIC and conventional time-domain MUSIC are shown for the linear frequency modulated (LFM) signal. The results indicate that TF-MUSIC outperforms its time-domain counterpart in resolving two closely spaced sources. The MUSIC-based localization approach is general and can be applied to a variety of real-world source localization problems. For example, it can be used to localize source abnormalities in the human brain (e.g., radiated signals coming from directions θ_1 and θ_2) using EEG signals. In the case of seizures, previous studies have established that they have LFM characteristics, as illustrated by the above simulations. This application is presented next.

8.1.4.5 Application to source localization of brain EEG abnormalities

In this section, a methodology for the localization of EEG abnormality using TF-MUSIC is briefly presented. The steps involved in this methodology are shown in [Fig. 8.1.11](#). After initial preprocessing, the multichannel EEG recordings are used to calculate the STFDs. The diagonal elements of the STFDs are the auto-TFDs of the EEG channel recordings, and off-diagonal elements of STFDs are the cross-TFDs between different EEG channel recordings. Following the steps given in [Table 8.1.1](#), the matrix ρ_{zz} is calculated by taking the average of the TFDs at selected higher energy (t,f) points.

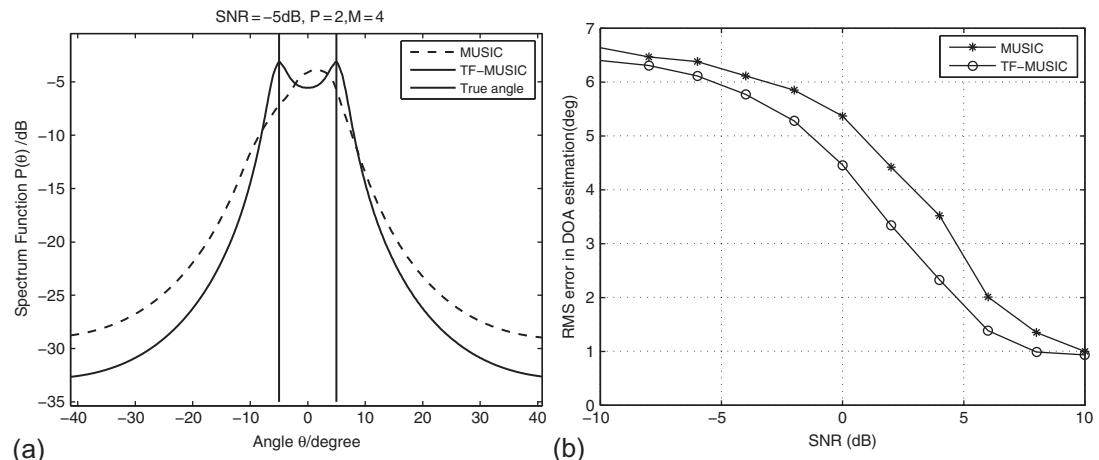


FIGURE 8.1.10

DOA estimation of two sources emitting linear frequency modulated (LFM) signals using MUSIC and TF-MUSIC algorithms. (a) Average spectrum (b) Effects of SNR on DOA estimation.

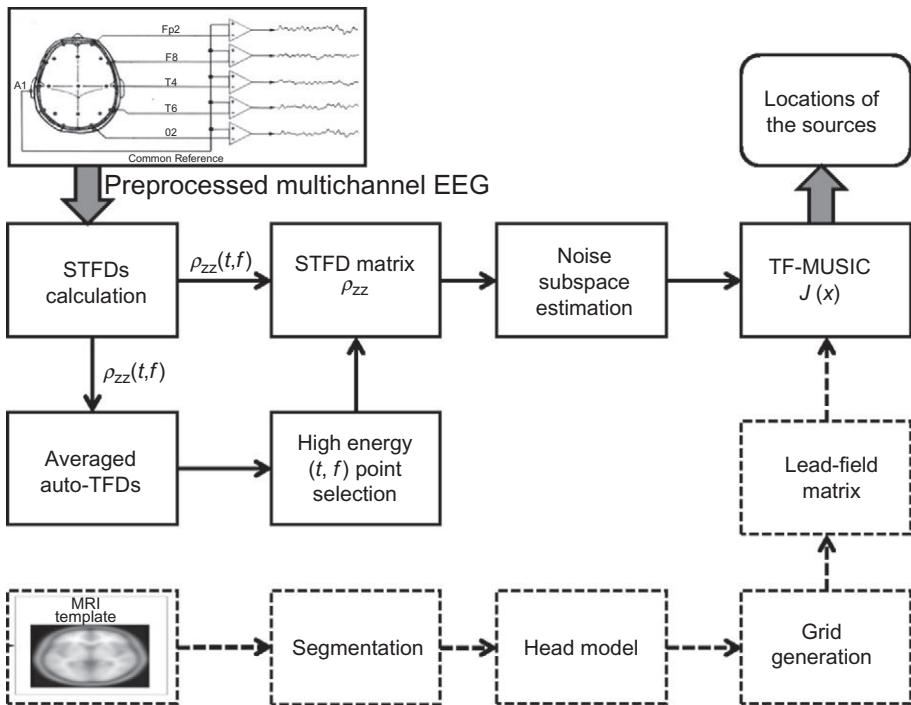
**FIGURE 8.1.11**

Illustration of the source localization of brain EEG abnormalities using TF-MUSIC.

Then, the matrix ρ_{zz} is decomposed into signal and noise spaces. To find the positions of the sources in the head, the TF-MUSIC algorithm exploits the fact that the source lead-field is orthogonal to the noise subspace of the matrix ρ_{zz} at the true source locations. Let $\mathbf{V}_E = [\mathbf{v}_{E,n+1}, \mathbf{v}_{E,n+2}, \dots, \mathbf{v}_{E,m}]$ indicate the eigenvectors corresponding to the noise subspace of the data covariance matrix ρ_{zz} ; then searching the lead-field vectors that are orthogonal to the noise subspace using the TF-MUSIC can be evaluated by using the following function [7]

$$J(\mathbf{x}) = \frac{1}{\lambda_{\min} \left(\mathbf{a}^H(\mathbf{x}) \mathbf{V}_E \mathbf{V}_E^H \mathbf{a}(\mathbf{x}), \mathbf{a}^H(\mathbf{x}) \mathbf{a}(\mathbf{x}) \right)}, \quad (8.1.25)$$

where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of the bracketed term and $\mathbf{a}(\mathbf{x})$ is a column of the lead-field matrix that represents the lead field at position \mathbf{x} in the head volume. Using Eq. (8.1.25), TF-MUSIC measures the orthogonality between the noise subspace and the lead-field matrix; therefore, acquiring the lead-field matrix is a critical step in calculating Eq. (8.1.25).

The lead-field matrix is obtained through constructing a forward model [6]. The forward model allows us to calculate an estimate of the electric field measured by the electrodes for a unit dipole current distribution [6]. For constructing the forward model, the brain surface is found from the MRI objects through a process termed as segmentation. Then, the head model is prepared from the segmented

brain surface, which is divided into a regular three-dimensional (3D) grid with a specified resolution, and for each grid point, the lead field is computed. Equation (8.1.25) is calculated on every grid point (where sources can exist), and each point in the 3D grid where $J(\mathbf{x})$ reaches a peak is chosen as the location of the source.

8.1.5 SUMMARY AND CONCLUSIONS

By extending the principles of single-sensor QTFDs to multisensor QTFDs, conventional stationary array processing can be combined with time-frequency methods. Numerical results show that TF-MUSIC outperforms the standard MUSIC in DOA applications. The method is illustrated on an application to source localization of brain EEG abnormalities. Additional details, applications, and extensions are described in the next sections of this chapter.

8.2 SPATIAL TIME-FREQUENCY DISTRIBUTIONS⁰

8.2.1 SPATIAL TIME-FREQUENCY DISTRIBUTIONS

The evaluation of quadratic TFDs of nonstationary signals impinging on a multisensor receiver yields spatial time-frequency distributions (STFDs). This permits the application of eigenstructure subspace techniques for solving a large class of channel estimation and equalization, blind source separation (BSS), and high-resolution direction of arrival (DOA) estimation problems [3,8–10]. STFD-based techniques are appropriate to handle sources of nonstationary waveforms that are highly localized in the time-frequency (t,f) domain. In the area of BSS, the use of the STFDs allows the separation of sources with identical spectral shape, but with different (t,f) localization properties, that is, different (t,f) signatures. For both source separation and DOA estimation problems, spreading the noise power while localizing the source energy in the (t,f) domain amounts to increasing the robustness of eigenstructure signal and noise subspace estimation methods with respect to channel and receiver noise. This in turn leads to an improvement of spatial resolution and source separation performance. The quadratic class of STFD of a signal vector $\mathbf{x}(t)$ extends from Eq. (3.2.18) as

$$\rho_{\mathbf{xx}}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(v,\tau) \mathbf{x}\left(u + \frac{\tau}{2}\right) \mathbf{x}^H\left(u - \frac{\tau}{2}\right) e^{j2\pi(vu-vt-f\tau)} \, d\tau \, du \, dv, \quad (8.2.1)$$

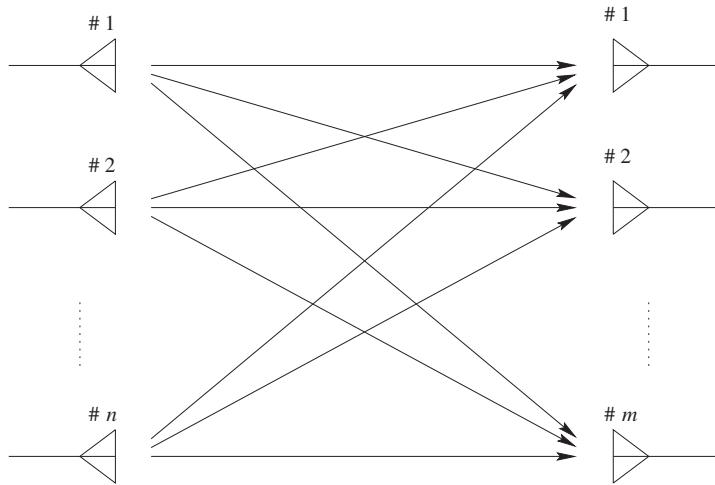
where $g(v,\tau)$ is the kernel function (defined in Eq. (3.2.6) of Chapter 3).

In narrowband array processing, when n signals arrive at an m -element array (see Fig. 8.2.1), the linear data model given in Eq. (8.1.5)

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{Ad}(t) + \mathbf{n}(t) \quad (8.2.2)$$

is commonly assumed, where $\mathbf{x}(t)$ is the $m \times 1$ data vector received at the array, $\mathbf{d}(t)$ is the $n \times 1$ source data vector, the $m \times n$ spatial matrix $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n]$ represents the mixing matrix, \mathbf{a}_i is the steering vector of i th signal, and $\mathbf{n}(t)$ is an additive noise vector whose elements are modeled as

⁰Authors: M.G. Amin, Villanova University, Villanova, PA, USA (moeness.amin@villanova.edu) and Y.D. Zhang, Villanova University, Villanova, PA, USA (yimin.zhang@villanova.edu). Reviewers: A. Gershman and K. Abed-Meraim.

**FIGURE 8.2.1**

m -Element array with n signal arrivals.

stationary, spatially and temporally white, zero-mean complex random processes, independent of the source signals.

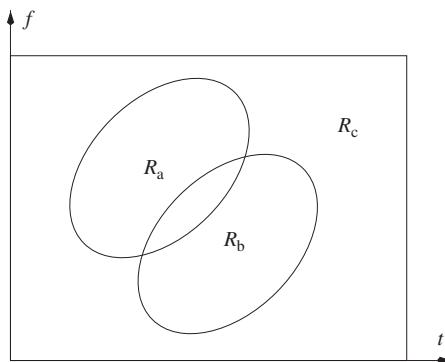
Under the uncorrelated signal and noise assumption and the zero-mean noise property, the expectation of the cross-term TFD matrices between the signal and noise vectors is zero, that is, $E[\rho_{yn}(t,f)] = E[\rho_{ny}(t,f)] = \mathbf{0}$, and it follows

$$E[\rho_{xx}(t,f)] = \rho_{yy}(t,f) + E[\rho_{nn}(t,f)] = \mathbf{A}\rho_{dd}(t,f)\mathbf{A}^H + \sigma^2\mathbf{I}, \quad (8.2.3)$$

where σ^2 is the noise power, and \mathbf{I} is the identity matrix. Equation (8.2.3) is the same as Eq. (8.1.20) and similar to that which has been commonly used in array processing based on second-order statistics, relating the signal correlation matrix to the data spatial correlation matrix [8]. This implies that key problems in various applications of array processing, specifically those dealing with nonstationary signal environments, can be approached using quadratic transformations. If $\rho_{dd}(t,f)$ is a full-rank matrix, the two subspaces spanned by the principal eigenvectors of $\rho_{xx}(t,f)$ and the columns of \mathbf{A} become identical. In this case, direction finding techniques based on eigenstructures can be applied. If $\rho_{dd}(t,f)$ is diagonal, that is, the signal cross-TFDs at the (t,f) point (t,f) are zeros, then both the mixing matrix and the signal waveforms can be recovered using BSS methods.

8.2.2 FUNDAMENTAL PROPERTIES

There are five key advantages of array signal processing using STFD. In order to properly explain these advantages, we use the diagram in Fig. 8.2.2. We consider two sources A and B incident on a multisensor array. Source A occupies the (t,f) region R_a , whereas source B occupies the (t,f) region

**FIGURE 8.2.2**

Signals with different time-frequency signature.

R_b . The (t,f) signatures of the two sources overlap, but each source still has a (t,f) region that is not intruded over by the other source.

- (1) Equation (8.2.3) can be easily derived for any arbitrary joint variables. Time and frequency are indeed the two most commonly used and physically understood parameters. However, by replacing the STFDs by spatial arbitrary joint-variable distributions, one can relate the sensor joint-variable distributions to the sources joint-variable distributions through the same mixing matrix \mathbf{A} . As shown in the “Examples” section below, there are situations where it is preferable to consider other domains such as the ambiguity lag-Doppler domain, where the locations of the signals and their cross-terms are guided by properties and mechanisms different than those associated with the (t,f) domain (see [Section 3.2](#)).
- (2) Equation (8.2.3) is valid for all (t,f) points. It is well known that direction finding techniques require $\rho_{dd}(t,f)$ to be full rank, preferably diagonal. On the other hand, BSS techniques demand the diagonal structure of the same matrix without degenerate eigenvalues. These properties along with high SNR requirements may be difficult to achieve using a single (t,f) point. Two different methods can be used for integrating several (t,f) points into Eq. (8.2.3). One method is based on a simple averaging performed over the signatures of the sources of interest, whereas the second method is based on incorporating several desired (t,f) points into joint diagonalization (JD) or joint block-diagonalization (JBD) schemes (details are given in [Sections 8.4.1](#) and [8.7](#)).
- (3) The TFD of the white noise is distributed all over the (t,f) domain, whereas the TFDs of the source waveforms are likely to be confined to much smaller regions. Referring to [Fig. 8.2.2](#), the noise is spread over both R_a and R_b as well as the complement region R_c . If the (t,f) points used in either the averaging or JD approaches belong to the noise-only region R_c , then no information of the incident waveforms is used, and therefore no reasonable source localization and signal separation outcomes can be obtained. On the other hand, if all points (t,f) in [Fig. 8.2.2](#) are used, and the employed TFD satisfies the marginal constraints, then it can be easily shown that only the signal average power is considered. As a result, the problem simplifies to the second-order covariance-based matrix approach, traditionally used in high-resolution DOA estimation.

This is an important feature, as it casts the conventional techniques as special cases of the (t,f) array signal processing framework. Finally, if we confine the (t,f) points to R_a and R_b , then only the noise part in these regions is included. The result of leaving out the points (t,f) that are not part of the (t,f) signatures of the signal arrivals is enhancing the input SNR, which is utilized by the source localization and signal separation techniques.

- (4) If we select only (t,f) points that belong to the (t,f) signature of *one* source, then this source will be the only one considered by Eq. (8.2.3). This selection, in essence, is equivalent to implicitly performing spatial filtering and removing other sources from consideration. It is important to note, however, that such removal does not come at the expense of reduction of the number of degrees-of-freedom (DOFs), as it is the case in beamspace processing, but the problem remains a sensor space processing with the original number of DOFs kept intact. This property represents a key contribution of TFDs to the direction finding and DOA estimation areas. An antenna array can be used to localize a number of sources equal or even greater than its number of sensors; this is an undetermined case and it is discussed in Sections 8.5 and 8.6. The fundamental condition is that there must be (t,f) regions over which the respective (t,f) signatures of the sources do not overlap. Referring to Fig. 8.2.2 and considering the case of two sensors, if all (t,f) points incorporated in direction finding belong to region R_a and not R_b , then the signal subspace defined by Eq. (8.2.3) is 1D. This concept will be further elaborated in Section 8.5. Thus, by excluding source *B*, a 1D noise subspace is established. This allows us to proceed with high-resolution techniques for localization of source *A*. In a general scenario, one can localize one source at a time or a set of selected sources, depending on the array size, overlapping and distinct (t,f) regions, and the dimension of the noise subspace necessary to achieve the required resolution performance. The same concepts and advantages of (t,f) point selection discussed above for direction finding can be applied to BSS problems.
- (5) The *a priori* knowledge of some temporal characteristics or the nature of time-varying frequency contents of the sources of interest may permit us to directly select the (t,f) regions used in Eq. (8.2.3). For instance, it is known that, in the ambiguity domain, all fixed frequency sinusoidal signals map to the time-lag axis. By only incorporating the points on this axis, we have, in fact, opted to separate and localize all narrowband signals in broadband communications platforms.

8.2.3 EXAMPLES

In this section, we present simulation examples to demonstrate the fundamental offerings discussed in the previous section. TF-MUSIC, ambiguity-domain MUSIC (AD-MUSIC), and the BSS (detailed discussion about the BSS is given in Section 8.4) based on STFDs are three different techniques chosen for the demonstration (Section 8.4). The algorithms involved in the implementation of the techniques are given in Tables 8.2.1–8.2.3 [8,9,11].

Example 8.2.1 (Amin et al. [11]). Consider the scenario of a four-element equi-spaced linear array spaced by half a wavelength, where one chirp signal and two sinusoidal signals are received. The data record has 128 samples. All three signals have the same SNR of 20 dB. The DOAs of the chirp signal and the two sinusoidal signals are 15° , 10° , and 0° , respectively. While the ambiguity function of the chirp signal sweeps the ambiguity domain with contribution at the origin, the exact auto-term

Table 8.2.1: Time-Frequency MUSIC

- Step I Form K matrices $\rho_{xx}(t_i, f_i)$ for the selected (t_i, f_i) points, $i = 1, \dots, K$ (see point selection in [Section 8.7.4](#)).
- Step II The eigenvectors $\mathbf{e}_1, \dots, \mathbf{e}_{m-n}$, corresponding to the $m - n$ smallest eigenvalues of $E[\rho_{xx}(t, f)]$, are obtained by joint block diagonalization, or the eigen-decomposition of averaged matrix $\frac{1}{K} \sum_{i=1}^K \rho_{xx}(t_i, f_i)$.
- Step III Estimate the number of signals from the eigenvalues, and estimate the DOAs from the peaks of the TF-MUSIC spectra $f(\theta) = |\hat{\mathbf{E}}_n^H \mathbf{a}(\theta)|^{-2}$, where $\hat{\mathbf{E}}_n = [\mathbf{e}_1, \dots, \mathbf{e}_{m-n}]$, and $\mathbf{a}(\theta)$ is the steering vector corresponding to DOA θ .

Table 8.2.2: Ambiguity-Domain MUSIC

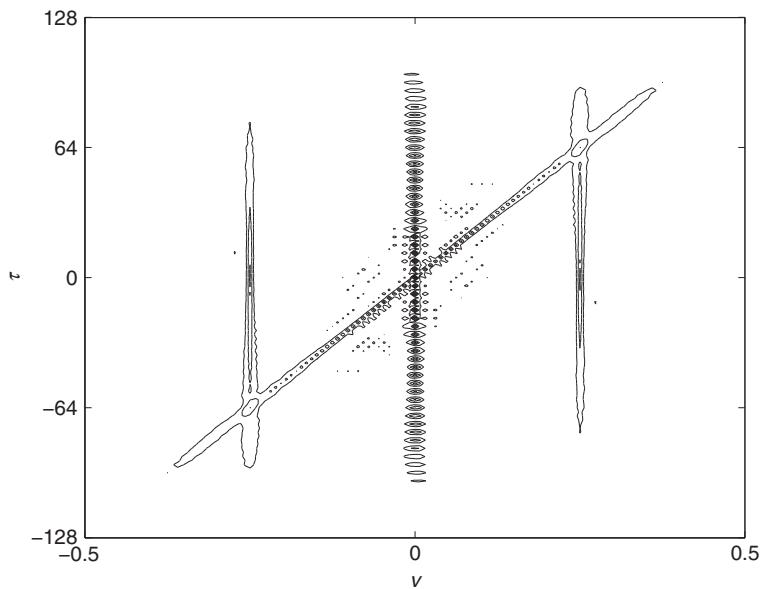
Ambiguity-domain MUSIC follows the same procedure as time-frequency MUSIC by using $\mathbf{A}_{xx}(\nu_i, \tau_i)$ instead of $\rho_{xx}(t_i, f_i)$, $i = 1, \dots, K$.

Table 8.2.3: Blind Source Separation Based on STFDs

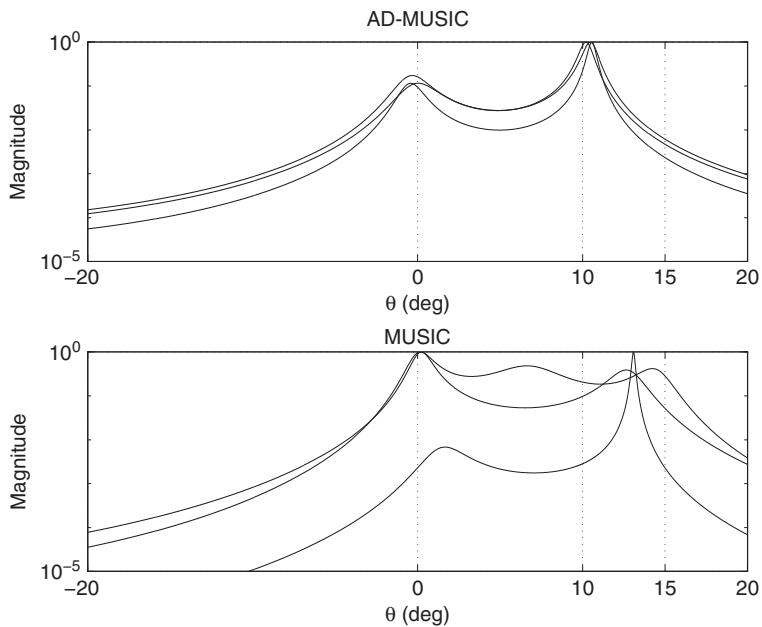
- Step I Estimate the auto-correlation matrix $\hat{\mathbf{R}}_{XX}$ from T data samples. Denote by $\lambda_1, \dots, \lambda_n$ the n largest eigenvalues, and by $\mathbf{h}_1, \dots, \mathbf{h}_n$ the corresponding eigenvectors, of $\hat{\mathbf{R}}_{XX}$.
- Step II An estimate $\hat{\sigma}^2$ of the noise variance is the average of the $m - n$ smallest eigenvalues of $\hat{\mathbf{R}}_{XX}$. The whitening matrix is formed as $\hat{\mathbf{W}} = [(\lambda_1 - \hat{\sigma}^2)^{-1/2} \mathbf{h}_1, \dots, (\lambda_n - \hat{\sigma}^2)^{-1/2} \mathbf{h}_n]^H$.
- Step III Form K matrices by computing the STFD of whitened vector $\mathbf{z}(t) = \hat{\mathbf{W}}\mathbf{x}(t)$ for a fixed set of (t_i, f_i) points, $i = 1, \dots, K$, corresponding to signal auto-terms.
- Step IV A unitary matrix $\hat{\mathbf{U}}$ is then obtained as joint diagonalizer of the set $\rho_{zz}(t_i, f_i)$, $i = 1, \dots, K$.
- Step V The source signals are estimated as $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^H \hat{\mathbf{W}}\mathbf{x}(t)$, and the mixing matrix \mathbf{A} is estimated as $\hat{\mathbf{A}} = \hat{\mathbf{W}}^H \hat{\mathbf{U}}$.

ambiguity function of the narrowband arrivals $s_1(t)$ and $s_2(t)$ is zero for nonzero frequency lags and may have nonzero values only along the vertical axis $\nu = 0$.

In this simulation example, we selected 24 points on the time-lag axis, excluding the origin, thereby emphasizing the narrowband components. [Figure 8.2.3](#) shows the ambiguity function where the two vertical lines away from the origin represent the cross-terms between the sinusoidal components. [Figure 8.2.4](#) shows the two estimated spatial spectra, one corresponding to the conventional method and the other corresponding to the AD-MUSIC, for three independent trials. There are two dominant eigenvalues for the case of the AD-MUSIC, since we have not deliberately considered the chirp signal

**FIGURE 8.2.3**

The ambiguity functions of the chirp signal and two sinusoidal signals.

**FIGURE 8.2.4**

The estimated spatial spectra of AD-MUSIC and conventional MUSIC.

through our careful selection of the ambiguity-domain points. It is clear that the AD-MUSIC resolves the two sinusoidal signals, while the conventional MUSIC could not separate the three signals.

Example 8.2.2 (Zhang et al. [12]). Consider a ULA of eight sensors separated by half a wavelength. Two chirp signals are emitted from two sources positioned at $(\theta_1, \theta_2) = (-10^\circ, 10^\circ)$, respectively. The data record has 1024 samples. The start and end frequencies of the chirp signal of the source at θ_1 are $f_{s1} = 0$ and $f_{e1} = 0.5$, while the corresponding two frequencies for the signal of the other source at θ_2 are $f_{s2} = 0.5$ and $f_{e2} = 0$, respectively.

Figure 8.2.5 shows the standard deviations of the DOA estimation $\hat{\theta}_1$ versus SNR. The curves in this figure show the theoretical and experimental results of the conventional MUSIC and TF-MUSIC. Pseudo Wigner-Ville distributions (WVDs) with window lengths $L = 33$ and 129 are considered. The Cramer-Rao bound (CRB) is also shown. Both signals are selected when performing TF-MUSIC. Simulation results are averaged over 100 independent trials of Monte Carlo experiments. The advantages of TF-MUSIC in low-SNR cases are evident from this figure. Figure 8.2.6 shows estimated spatial spectra at SNR = -20 dB based on TF-MUSIC ($L = 129$) and the conventional MUSIC. The TF-MUSIC spectral peaks are clearly resolved.

Example 8.2.3 (Belouchrani and Amin [8]). In Fig. 8.2.7, we show an example of the application of STFDs to the BSS problem (details about the process of BSS are given in Sections 8.4–8.6). A three-element equi-spaced linear array is considered where the interelement spacing is half a wavelength. Two chirp signals arrive at -10° and 10° , respectively. The number of data samples used to compute

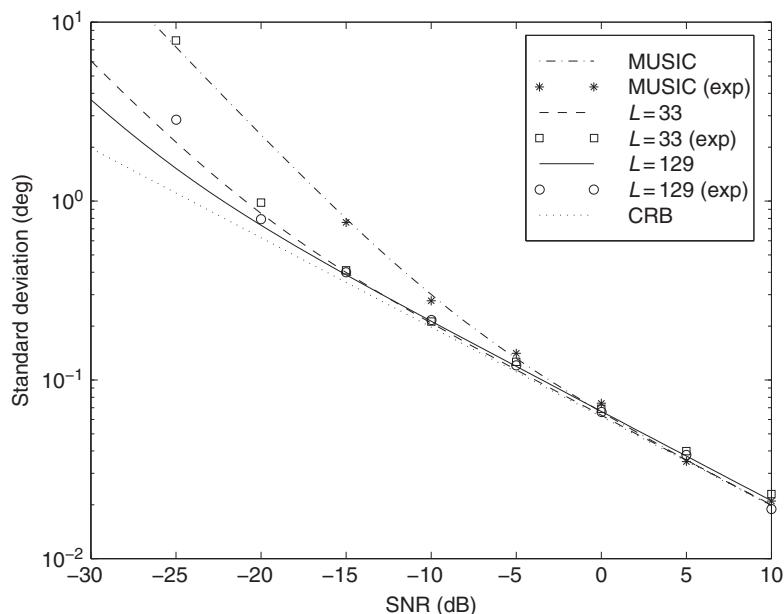
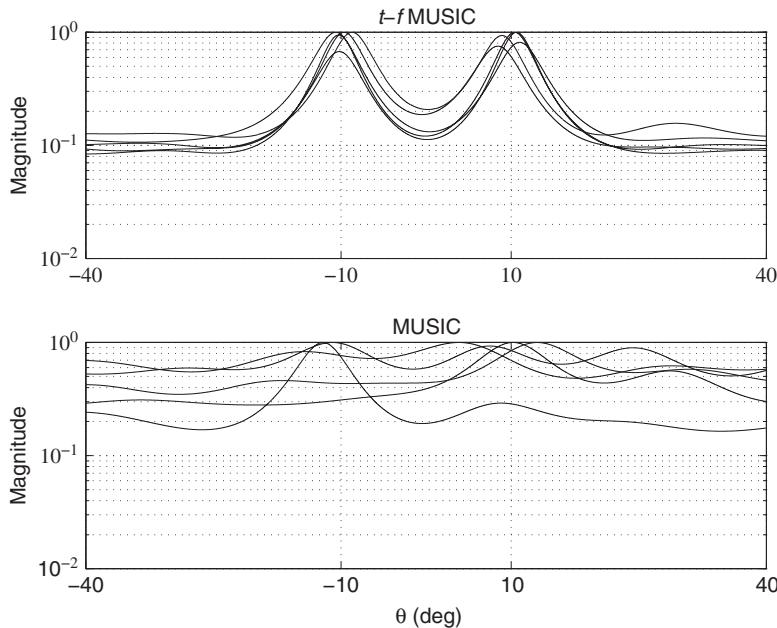


FIGURE 8.2.5

The standard deviations of DOA estimation $\hat{\theta}_1$ vs SNR.

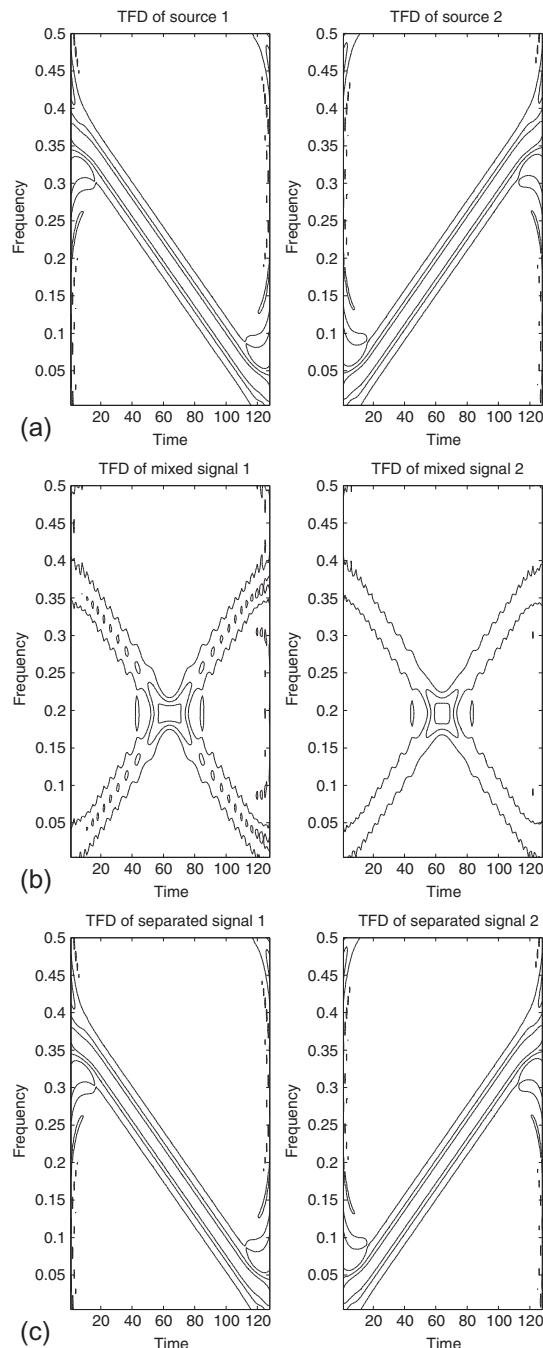
**FIGURE 8.2.6**

The estimated spatial spectra of TF-MUSIC and conventional MUSIC.

the STFD is 128. The number of (t,f) points employed in the JD is $p = 128$, with an equal number of points on each signature. Figure 8.2.7(b) shows the exponential distributions of two linear mixtures of the original chirp signals depicted in Fig. 8.2.7(a), corresponding to the data at the first and the second sensors. Using the STFDs, we are able to recover the original signals from their observed mixture, as shown in Fig. 8.2.7(c).

8.2.4 CROSS-TERM ISSUES IN STFD

As discussed in Section 3.1.2 in the case of a single sensor, there are two sources of cross-terms. The first type is the cross-terms that are the results of the interactions between the components of the same source signal. The other type of cross-terms is those generated from the interactions between two signal components belonging to two different sources. This second category of cross-terms is associated with cross-TFDs of the source signals and, at any given (t,f) point, it constitutes the off-diagonal entries of the source TFD matrices $\rho_{dd}(t,f)$ defined in Eq. (8.2.3). Although, the off-diagonal elements do not necessarily destroy the full-rank matrix property necessary for the direction-finding application [13], they violate the basic assumption in the problem of source separation regarding the diagonal structure of the source TFD matrix. We must therefore select the (t,f) points that belong to auto-term regions where cross-term contributions are at minimum, for example, by using *a priori* information of the source signals.

**FIGURE 8.2.7**

Blind source separation based on STFDs. (a) TFDs of the source signals; (b) TFDs of the mixed signals; and (c) TFDs of the separated signals.

The method of spatial averaging of the STFD introduced in Ref. [14] does not reduce the cross-terms as in the case with reduced interference distribution kernels (see [Section 3.1.2.3](#)), but rather move them from their locations on the off-diagonal matrix entries to be part of the matrix diagonal elements. The other parts of the matrix diagonal elements represent the contribution of the auto-terms at the same point. Therefore, not only are we able to set the off-diagonal elements of the source TFD matrix to zeros, but also we can improve performance by selecting the (t, f) points of peak values, irrespective of whether these points belong to auto-term or cross-term regions.

8.2.5 SUMMARY AND CONCLUSIONS

The STFD is an important tool for temporal and spatial separations of sources emitting nonstationary signals. It is a discriminatory tool that allows a consideration of only a subset of source signals impinging on a multisensor receiver. This property enhances signal parameter estimation and permits direction finding and signal separation to be applied to a number of sources that equals or even exceeds the number of sensors.

All material presented in this section is based on the model [\(8.2.2\)](#). When this strict model is relaxed, a direction-finding technique employing an STFD-based wideband root-MUSIC can be defined [\[15\]](#). Another approach is the utilization and integration of cross-terms into STFDs. Source separation can then be performed based on both auto-terms and cross-terms through JD and joint anti-diagonalization (JAD) schemes of STFD matrices [\[3\]](#). [Section 8.7](#) presents algorithms and sample code for these methods.

8.3 QUADRATIC DETECTION IN ARRAYS USING TFD⁰

The detection of signals in noise is a classical hypothesis-testing problem. We assume the very popular ULA geometry and a single signal source arriving from the far field at some angle θ with respect to the array. The goal is to detect the presence of this signal source from observations contaminated by additive noise at the sensors.

In many situations, the channel may induce unknown modifications to the time and frequency parameters of the transmitted signal. Such situations include the well-known delay-Doppler situation in radar/sonar detection problems. TFDs, which jointly describe the time and frequency characteristics of the signal, are powerful tools for designing the optimal detector in such situations. It has been shown that the optimal quadratic detector for the detection of nonstationary Gaussian signals (i.e., signals whose statistics vary rapidly but are Gaussian at each fixed time instance) with unknown time and frequency offsets in the presence of noise can be implemented naturally within the quadratic class of TFDs (see [Section 12.5](#) and Ref. [\[17\]](#)). Here, we show how this idea can be extended to detection using an array of sensors, even those which exhibit imperfect spatial coherence. Imperfect coherence (i.e., a loss in the spatial correlation for a fixed time instance) is often attributed to signal scattering in multipath channels. Specifically, we consider the problem of detecting arbitrary nonstationary second-order

⁰Second edition updated and extended by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Ali**, Qatar University, Qatar (sadiq@qu.edu.qa). First edition [\[16\]](#) by A.M. Rao and D.L. Jones. Reviewers: A. Aissa-El-Bey and Y. Zhang.

signals (signals described by their second-order statistics such as Gaussian signals [18]) with unknown time and frequency offsets arriving in a linear array with an unknown angle of arrival. We explicitly show how the optimal detector for such a problem can be implemented naturally and efficiently in the time-frequency (t, f) domain.

8.3.1 THE DETECTION PROBLEM

Consider the following composite hypothesis-testing problem in continuous time:

$$\begin{aligned} H_0: x(t) &= n(t) \\ H_1: x(t) &= s(t - \tau) e^{j2\pi\nu t} + n(t), \end{aligned} \quad (8.3.1)$$

where $t \in T$, the time interval of observation, $x(t)$ is the observed signal, $n(t)$ is a zero-mean complex white Gaussian noise with variance σ^2 , and $s(t)$ is a zero-mean complex arbitrary second-order signal with correlation function $R_s(t_1, t_2)$. The parameters (τ, ν) represent time and frequency offset parameters that are assumed to be unknown. These parameters arise in the classic radar delay/Doppler scenario where the delay is due to the range of the target and the Doppler is due to the velocity of the target. In statistical hypothesis testing, for each observation $x(t)$, a real-valued test statistic $L(x)$ is compared to a threshold to decide in favor of H_1 or H_0 ; that is, to decide whether the signal is present or not. TFDs provide a natural detection framework for such hypothesis-testing problems for two main reasons: first, detecting a second-order signal (such as a Gaussian signal) in the presence of Gaussian noise involves a quadratic function of the observations [17], and bilinear TFDs are quadratic in the observations; second, TFDs possess additional degrees of freedom provided by the time and frequency parameters.

8.3.2 QUADRATIC DETECTION IN A LINEAR ARRAY

The signal reaches the array of M sensors with spacing d at angle θ , where θ is assumed to be unknown, and $x_i(t)$, $i = 1, \dots, M$ denotes the signal at the i th sensor as shown in Fig. 8.3.1. For a ULA configuration, the signal at the i th sensor is a delayed version of the signal at the first sensor, and the value of the delay depends on the unknown angle of arrival, θ . That is, when the signal is present $x_i(t) = s(t - (i - 1)D) + n_i(t)$, where $D = \frac{d}{c} \sin(\theta)$ and c is the velocity of propagation

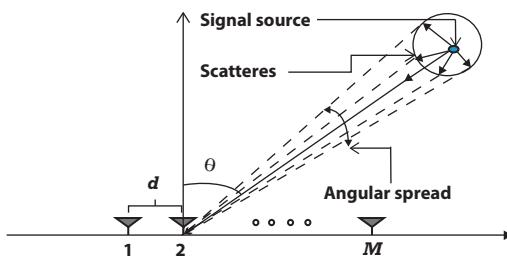


FIGURE 8.3.1

Geometrical illustration of the multipath angle spread.

in the medium. However, when considering the use of a very large array in order to achieve a high array gain, the signal received at widely separated sensors may have reduced coherence due to the complexity in the propagation of the signal from the source to spatially separated receivers (see Refs. [19–21] and references therein). Loss of spatial coherence can often be attributed to complex signal scattering arising in multipath propagation [22]. In particular, the larger the angular spread of the multipath the more rapid the loss in spatial coherence (the smaller the coherence length). Considering only the signal second-order statistics, the model for partial coherence used is given in terms of the correlation function. An exponential power-law model has been suggested [19] whereby the cross-correlation function between the i th and j th sensors is scaled by the coefficient $c_{ij} = e^{-\frac{|i-j|}{L}}$, where L is a dimensionless characteristic correlation length. We arrange the decorrelation coefficients in matrix form as $\mathbf{C} = \{c_{ij}\}$, for convenience to deal with the aligned sensor outputs; that is, let $y_i(t) = x_i(t + (i - 1)D)$. For convenience, let us also arrange the aligned sensor outputs in vector form as $\mathbf{Y}^\theta = [y_1^\theta(t) \ y_2^\theta(t) \ \dots \ y_M^\theta(t)]^T$, where θ denotes the dependence of aligning the sensor signals on the unknown angle of arrival. Deflection-optimal detectors can be interpreted as “maximum SNR” detectors because deflection is a measure of SNR. The deflection is defined as

$$q = \frac{E[L(x)|H_1] - E[L(x)|H_0]}{\text{Var}[L(x)|H_0]}.$$

Let us consider the optimal test statistic based on this criterion and using a generalized likelihood ratio test (GLRT) technique to deal with the unknown time, frequency, and angle parameters. This is given by [17,18]

$$L_{\text{opt}} = \max_{\tau, v, \theta} \left\langle \tilde{\mathbf{P}}_n^{-1} \left(\mathbf{C} \otimes \mathbf{R}_s^{(\tau, v)} \right) \tilde{\mathbf{P}}_n^{-1} \mathbf{Y}^\theta, \mathbf{Y}^\theta \right\rangle, \quad (8.3.2)$$

where \otimes represents the Kronecker product, $\langle ., . \rangle$ denotes the inner product defined as: $\langle x, y \rangle = \int x(t) y^*(t) dt$, and $\mathbf{R}_s^{(\tau, v)}$ denotes the linear operator defined by the corresponding correlation function $R_s^{(\tau, v)}$ as [18]:

$$\left(\mathbf{R}_s^{(\tau, v)} x \right)(t) = \int R_s^{(\tau, v)}(t, \tau) x(\tau) d\tau,$$

where $R_s^{(\tau, v)}(t_1, t_2) = R_s(t_1 - \tau, t_2 - \tau) e^{j2\pi v t_1} e^{-j2\pi v t_2}$ denotes the correlation function corresponding the signal source with the unknown time and frequency offsets. Similarly, in Eq. (8.3.2), $\tilde{\mathbf{P}}_n = \mathbf{I} \otimes \mathbf{R}_n$ is a matrix of linear operators \mathbf{R}_n corresponding to the correlation function of the noise process. Upon expanding the inner product in Eq. (8.3.2) we obtain [17,18]

$$\begin{aligned} L_{\text{opt}} &= \max_{\tau, v, \theta} \sum_{i=1}^M \sum_{j=1}^M c_{ij} \left\langle \mathbf{R}_n^{-1} \mathbf{R}_s^{(\tau, v)} \mathbf{R}_n^{-1} y_i^\theta(t), y_j^\theta(t) \right\rangle \\ &= \max_{\tau, v, \theta} \sum_{i=1}^M \sum_{j=1}^M c_{ij} \left\langle \mathbf{Q}^{(\tau, v)} y_i^\theta(t), y_j^\theta(t) \right\rangle, \end{aligned} \quad (8.3.3)$$

where $\mathbf{Q}^{(\tau, v)} = \mathbf{R}_n^{-1} \mathbf{R}_s^{(\tau, v)} \mathbf{R}_n^{-1}$ is a positive definite linear operator. In the following sections, Eq. (8.3.3) is represented in the (t, f) domain using the Weyl correspondence [17].

8.3.3 TFD-BASED ARRAY DETECTION

8.3.3.1 Principle

Let us consider the Weyl correspondence which relates inner products, positive definite linear operators, and the WVD. Using the fact that the Weyl correspondence involves a covariance to time, frequency, and scale offsets using the methods in Ref. [17], it can be verified that the test statistic in Eq. (8.3.3) can be expressed in terms of TFDs, allowing for a natural and efficient implementation of the optimal detector. In Eq. (8.3.3), the inner-product term can be expressed as [17,18]:

$$\left\langle \mathbf{Q}^{(\tau,v)} y_i^\theta(t), y_j^\theta(t) \right\rangle = \iint W_{y_i y_j}^\theta(u, v) \text{WS}_{\mathbf{Q}^{(\tau,v)}}(u, v) du dv, \quad (8.3.4)$$

where $\text{WS}_{\mathbf{Q}^{(\tau,v)}}(u, v)$ is the Weyl symbol of the operator $\mathbf{Q}^{(\tau,v)}$ defined as [17,18]:

$$\text{WS}_{\mathbf{Q}^{(\tau,v)}}(u, v) = \int \mathbf{Q}^{(\tau,v)}\left(u + \frac{\tau}{2}, u - \frac{\tau}{2}\right) e^{-j2\pi v \tau} d\tau. \quad (8.3.5)$$

Using Weyl correspondence given in Eq. (8.3.4), the test statistic in Eq. (8.3.3) can be expressed as

$$L_{\text{opt}} = \max_{\tau, v, \theta} \sum_{i=1}^M \sum_{j=1}^M c_{ij} \iint W_{y_i y_j}^\theta(u, v) \text{WS}_{\mathbf{Q}^{(\tau,v)}}(u, v) du dv. \quad (8.3.6)$$

Any bi-linear TFD from the QTFD class can be expressed as (see Chapter 3)

$$\rho_x(t, f; g) = \iint W_x(u, v) g(u - t, v - f) du dv, \quad (8.3.7)$$

where $W_x(u, v)$ is the WVD of the signal $x(t)$, defined as

$$W_x(t, f) = \int x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f \tau} d\tau. \quad (8.3.8)$$

Comparing Eqs. (8.3.6) and (8.3.7), we can write

$$L_{\text{opt}} = \max_{\tau, v, \theta} \sum_{i=1}^M \sum_{j=1}^M c_{ij} \iint W_{y_i y_j}^\theta(u, v) \text{WS}_{\mathbf{Q}}(u - \tau, v - v) du dv. \quad (8.3.9)$$

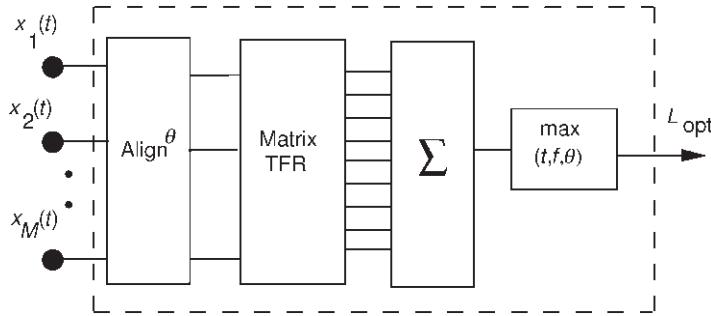
Equation (8.3.9) can be simplified as

$$L_{\text{opt}} = \max_{t, f, \theta} \sum_{i=1}^M \sum_{j=1}^M c_{ij} \rho_{y_i y_j}^\theta(t, f; g = \text{WS}_{\mathbf{Q}}). \quad (8.3.10)$$

The superscript θ denotes the dependance of the signal TFD for each hypothesized angle of arrival. Equation (8.3.10) shows that we must form the sum of all weighted cross-TFDs; we refer to this quantity as a matrix TFD.

8.3.3.2 Partially coherent case

In the above discussion, although the signals in the vector \mathbf{Y}^θ are aligned, they are not identical due to the coherence loss between sensors as a result of the partially coherent environment. Hence, Eq. (8.3.10) is the optimal detector for the partially coherent environment. Figure 8.3.2 shows the detector structure

**FIGURE 8.3.2**

TFD-based optimal quadratic array processor: partially coherent case.

for the partially coherent case. Because the detector involves forming TFDs of signals that are aligned to examine different spatial directions, we may view this detection structure in terms of time-frequency-space or time-scale-space. Since we have not assumed any spatial statistical characteristics of the signal, the kernel is the same here regardless of the angle of arrival being analyzed.

8.3.3.3 Coherent case

If the array environment is perfectly coherent, then $c_{ij} = 1, \forall i, j$. Using basic properties of QTFDs, it can be shown that the optimal test statistic involves first summing the sensor observations and then applying the TFD (with kernel g) for each hypothesized angle of arrival and choosing the maximum value. Using $c_{ij} = 1$ in Eq. (8.3.6), interchanging the order of summation and integration, and exploiting the properties of Weyl symbol, we get [18]:

$$L_{\text{opt}} = \max_{t, f, \theta} \rho_{\sum_{i=1}^M y_i}^\theta (t, f; g = \text{WS}_Q). \quad (8.3.11)$$

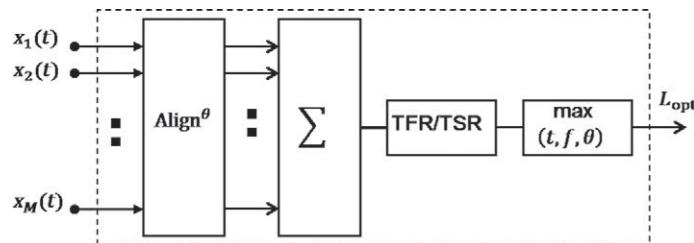
Hence, the matrix processing of Fig. 8.3.2 is replaced by a simple summation. The structure of the coherent detection scheme is shown in Fig. 8.3.3, where it can be seen that to design the optimal detector, first the sensor outputs are aligned and added together, then the TFD is applied to the resulting sum for each value of the hypothesized angle of arrival; and, finally the maximum value in the (t, f) domain is used to find L_{opt} .

8.3.3.4 Noncoherent case

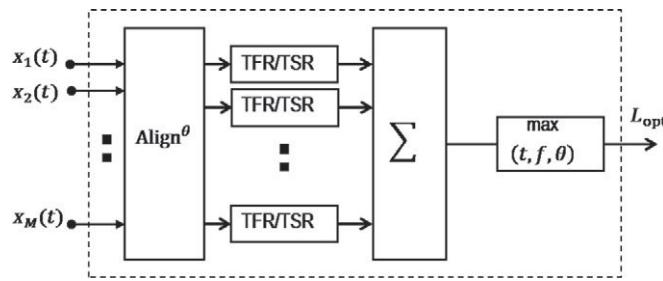
If the array environment is noncoherent, then $\mathbf{C} = \mathbf{I}$, as $c_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. Then for the unknown time and frequency shift, Eq. (8.3.10) becomes

$$L_{\text{opt}} = \max_{t, f, \theta} \sum_{i=1}^M \rho_{y_i}^\theta (t, f; g = \text{WS}_Q). \quad (8.3.12)$$

It means that the optimal test statistic includes first taking the TFD of each sensor observation (with kernel g), and then summing the resulting TFDs for each hypothesized angle of arrival and choosing the maximum value as explained in Fig. 8.3.4.

**FIGURE 8.3.3**

TFD-based optimal quadratic array processor: coherent case.

**FIGURE 8.3.4**

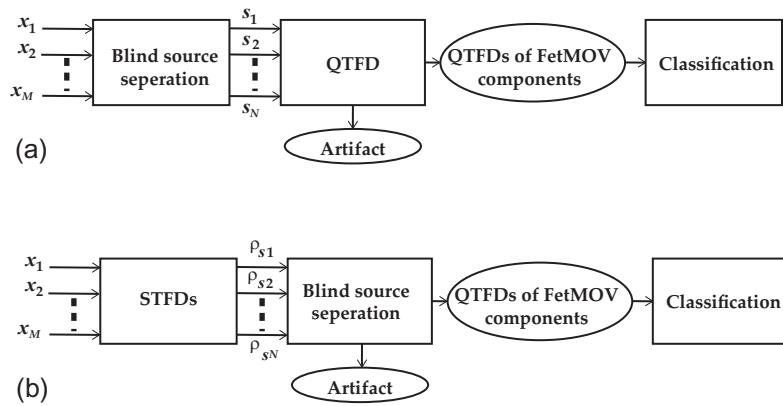
TFD-based optimal quadratic array processor: noncoherent case.

8.3.4 APPLICATION: MULTISENSOR FETAL MOVEMENT DETECTION SYSTEM

In the multisensor case, more precision can be obtained by optimally taking into account the extra diversity provided by all combined sensor information available for processing. As an application, this section describes a multisensor fetal movement (FetMov) detection system based on the (t, f) approach [23]. In this FetMov application, the received signals are corrupted by noise. In addition, the effect of channel transmission (fading problem) is also not negligible. In the multisensor FetMov system, this fading effect presents a source of diversity of information that can be exploited to separate different components and to improve the visualization of the TFD of the signal [23]. The multisensor FetMov detection system described in Ref. [23] records fetal activity using tri-axial accelerometers. In order to develop this system, a good synchronization between the sensors is required. To address this problem, a localization of the local energy maximum allows estimating the delay of the reception from the first sensor to the second sensor [23]. This approach can be formulated using the synchronized (aligned) signals received through different sensors. The resulting received signal can then be expressed as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (8.3.13)$$

where $M \times N$ matrix \mathbf{A} represents the diversity matrix (or mixing matrix), \mathbf{n} is the noise vector, and the elements of vector $\mathbf{s} = [s_1, s_2, \dots, s_N]^T$ are the signal components.

**FIGURE 8.3.5**

Flowchart of the FetMov detection and classification based on source separation: (a) quadratic TFD processing and (b) STFDs of FetMov signals.

The main functions of the system are to estimate different components of the FetMov and to estimate their characteristics prior to decision-making. Due to the lack of sufficient information concerning the mixing matrix and the nature of the FetMov signal which is nonstationary, a blind source separation (BSS) algorithm combined with a (t,f) analysis would be required to achieve the desired outcome with precision. As the information about the mixing matrix are insufficient, the noncoherent case is applicable where channels are processed independently using the (t,f) matched filtering (more details regarding the (t,f) matched filtering can be found in [Section 12.5](#)). Classical BSS techniques can be used to separate different components of the signals and also to separate the signal from the artifacts; then, for each separated source, a QTFD analysis is used to determine its characteristics. [Figure 8.3.5\(a\)](#) shows the flowchart of such a FetMov detection and classification scheme based on a combined source separation and QTFD processing. A possible way to improve the above results is to use source separation on STFDs as shown in [Fig. 8.3.5\(b\)](#) [23]. In this method, first, the STFDs of the received signals are calculated and then BSS is performed. The STFD-based approach provides robustness to the BSS against the noise due to spreading the noise power and localizing the source signal power in the (t,f) plane (see [Section 8.1.4.2](#)).

8.3.5 SUMMARY AND CONCLUSIONS

This section demonstrates that time-frequency-based detectors are naturally suited to quadratic detection in an array environment. By using a GLRT approach, the deflection-optimal test statistic is cast using TFDs. The TFD-based structure allows the optimal detector to be implemented naturally and efficiently by exploiting the many degrees of freedom available. In the general case of a partially coherent environment, the test statistic includes a weighted sum of all cross-TFDs of the aligned sensor outputs for each value of hypothesized angle of arrival. Completely coherent and noncoherent cases are shown to be special cases of the partially coherent model. In the coherent case, the optimal test statistic simplifies to include a single auto-TFD of the sum of the aligned sensor outputs for each hypothesized

angle of arrival. In the noncoherent case, the optimal test statistic simplifies to include the sum of auto-TFDs of the aligned sensor outputs for each hypothesized angle of arrival (see also [Section 12.1.3](#) for more details).

8.4 BLIND SOURCE SEPARATION USING TIME-FREQUENCY DISTRIBUTIONS⁰

In BSS, the underlying model is that of n “statistically” independent signals whose m (possibly noisy) mixtures are observed (see illustration in [Fig. 8.2.1](#)). Neither the structure of the mixtures nor the source signals are known to the receivers. In this environment, the problem is to identify and then decouple the mixtures blindly, as discussed in [Section 8.1.2](#) and as illustrated in [Fig. 8.1.7](#). Blind source separation (BSS) applications include: source localization and tracking by radar and sonar devices; speaker separation (cocktail party problem); multiuser detection in communication systems; medical signal processing (e.g., separation of EEG or ECG signals); industrial fault detection; extraction of meaningful features from data, etc.

Surprisingly, this difficult problem has elegant solutions that depend on the nature of the mixtures and the nature of the source statistical information.

Most approaches to BSS assume that each source signal is a sequence of independently and identically distributed (i.i.d.) variables [3]. In this context, BSS is possible only if at most one of the sources has a Gaussian distribution. In contrast, if the source sequences are *not* i.i.d., we can blindly separate the sources even for Gaussian processes. For the case where each source sequence is a temporally correlated stationary process [3], BSS is possible if the source signals have different spectra. The case where the second “i” of “i.i.d.” is failing, that is, the nonstationary case is considered in Refs. [3,24].

The focus is the exploitation of signal nonstationarity for BSS by using time frequency (t,f) analysis to separate and recover the incoming signals. The underlying problem can be posed as a signal synthesis from the (t,f) plane with the incorporation of the spatial diversity provided by the multisensor. Moreover, the effects of spreading the noise power while localizing the source energy in the (t,f) domain amount to increasing the SNR and hence improve performance. This section reviews the BSS techniques that exploit the joint eigenstructure of a combined set of (t,f) distributions, including (1) the BSS problem of instantaneous mixtures and (2) the general case of BSS of convolutive mixtures.

8.4.1 SEPARATION OF INSTANTANEOUS MIXTURES

8.4.1.1 Data model

Consider m sensors receiving an instantaneous linear mixture of signals (i.e., without delay) emitted from n sources (with $n \leq m$). The $m \times 1$ vector $\mathbf{x}(t)$ denotes the output of the sensors at time instant t which may be corrupted by additive noise $\mathbf{n}(t)$. Hence, the linear data model may be given by:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (8.4.1)$$

⁰Authors: **K. Abed-Meraim**, Université d’Orléans, Orléans, France (karim.abed-meraim@univ-orleans.fr), **A. Belouchrani**, Ecole Nationale Polytechnique, Algiers, Algeria (adel.belouchrani@enp.edu.dz), and **A.R. Leyman**, National University of Singapore, Singapore (larahim@cwc.nus.edu.sg). Reviewers: B. Barkat and A. Cickochi.

where the $m \times n$ matrix \mathbf{A} is called the “mixing matrix.” The n source signals are collected in an $n \times 1$ vector denoted $\mathbf{s}(t)$ which is referred to as the source signal vector. The sources are assumed to have different structures and localization properties in the (t, f) domain. The mixing matrix \mathbf{A} is full-column rank but is otherwise unknown. In contrast to traditional parametric methods, no specific structure of the mixture matrix is assumed.

8.4.1.2 Spatial time-frequency distributions

The discrete-time form of a quadratic-class (t, f) distribution (TFD), for a signal $x(t)$, may be expressed (see [Section 6.1](#)) as

$$\rho_{xx}(t, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} G(m, l) x(t+m+l) x^*(t+m-l) e^{-j4\pi fl}, \quad (8.4.2)$$

where t and f represent the time index and the frequency index, respectively. The kernel $G(m, l)$ characterizes the TFD and is a function of both the time and lag variables. The cross-TFD of two signals $x_1(t)$ and $x_2(t)$ is defined in [Chapter 3](#) by

$$\rho_{x_1 x_2}(t, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} G(m, l) x_1(t+m+l) x_2^*(t+m-l) e^{-j4\pi fl}. \quad (8.4.3)$$

Expressions (8.4.2) and (8.4.3) are now used to define the following data *spatial (t, f) distribution (STFD) matrix*,

$$\rho_{\mathbf{xx}}(t, f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} G(m, l) \mathbf{x}(t+m+l) \mathbf{x}^H(t+m-l) e^{-j4\pi fl}, \quad (8.4.4)$$

where $[\rho_{\mathbf{xx}}(t, f)]_{ij} = \rho_{x_i x_j}(t, f)$, for $i, j = 1, \dots, m$, and \mathbf{x}^H denotes the conjugate transpose of \mathbf{x} .

Under the linear data model of Eq. (8.4.1) and assuming a noise-free environment, the STFD matrix takes the following structure, for $\sigma = 0$:

$$\rho_{\mathbf{xx}}(t, f) = \mathbf{A} \rho_{\mathbf{ss}}(t, f) \mathbf{A}^H + \sigma^2 \mathbf{I}, \quad (8.4.5)$$

where $\rho_{\mathbf{ss}}(t, f)$ is the source TFD matrix whose entries are the auto- and cross-TFDs of the sources, respectively, defined as $\rho_{s_i s_i}(t, f)$ and $\rho_{s_i s_j}(t, f)$ for $i \neq j$.

8.4.1.3 Blind source separation based on STFD

Let us review the principle of BSS based on STFDs [3] (see also [Table 8.2.3](#)).

8.4.1.3.1 Whitening

Let \mathbf{W} denote an $n \times m$ matrix such that $(\mathbf{W}\mathbf{A})(\mathbf{W}\mathbf{A})^H = \mathbf{I}$, that is, $\mathbf{U} = \mathbf{W}\mathbf{A}$ is an $n \times n$ unitary matrix (this matrix is referred to as a whitening matrix, since it whitens the signal part of the observations). Pre- and postmultiplying the STFD matrices $\rho_{\mathbf{xx}}(t, f)$ by \mathbf{W} leads to the *whitened STFD-matrices*, defined as

$$\underline{\rho}_{\mathbf{xx}}(t, f) = \mathbf{W} \rho_{\mathbf{xx}}(t, f) \mathbf{W}^H. \quad (8.4.6)$$

From the definition of \mathbf{W} and Eq. (8.4.5), we can express $\underline{\rho}_{\mathbf{xx}}(t,f)$ as

$$\underline{\rho}_{\mathbf{xx}}(t,f) = \mathbf{U} \rho_{\mathbf{ss}}(t,f) \mathbf{U}^H. \quad (8.4.7)$$

Note that the whitening matrix \mathbf{W} can be computed in different ways (see also Section 8.7.3). It can be obtained, for example, as an inverse square root of the observation autocorrelation matrix or else computed from the STFD matrices [3].

8.4.1.3.2 Source separation using JD

By selecting auto-term (t,f) points (see Section 8.7.4), the whitened STFD matrices will have the following structure,

$$\underline{\rho}_{\mathbf{xx}}^{\mathbf{a}}(t,f) = \mathbf{U} \rho_{\mathbf{ss}}^{\mathbf{a}}(t,f) \mathbf{U}^H, \quad (8.4.8)$$

where $\rho_{\mathbf{ss}}^{\mathbf{a}}(t,f)$ is diagonal.¹ This expression shows that any whitened STFD matrix is diagonal in the basis of the columns of the matrix \mathbf{U} (the eigenvalues of $\underline{\rho}_{\mathbf{xx}}^{\mathbf{a}}(t,f)$ being the diagonal entries of $\rho_{\mathbf{ss}}(t,f)$).

If, for a given (t,f) point, the diagonal elements of $\rho_{\mathbf{ss}}^{\mathbf{a}}(t,f)$ are all distinct, the missing unitary matrix \mathbf{U} may be uniquely (up to permutation and scaling ambiguity²) retrieved by computing the eigendecomposition of $\rho_{\mathbf{xx}}^{\mathbf{a}}(t,f)$. Indeterminacy occurs in the case of degenerate eigenvalues, that is, when $\rho_{s_i s_j}(t,f) = \rho_{s_j s_i}(t,f)$, $i \neq j$. It does not seem possible to *a priori* choose the (t,f) point such that the diagonal entries of $\rho_{\mathbf{ss}}^{\mathbf{a}}(t,f)$ are all distinct. Moreover, when some eigenvalues of $\rho_{\mathbf{xx}}^{\mathbf{a}}(t,f)$ come to degeneracy, the robustness of determining \mathbf{U} from eigendecomposition of a single whitened STFD matrix is seriously impaired. The situation is more favorable when considering the joint diagonalization (JD) of a combined set $\{\rho_{\mathbf{xx}}^{\mathbf{a}}(t_i, f_i) | i = 1, \dots, p\}$ of p (source auto-term) STFD matrices. This amounts to incorporating several (t,f) points in the source separation problem which reduces the likelihood of having degenerate eigenvalues. Note that two source signals with identical (t,f) signatures cannot be separated even with the inclusion of all information in the (t,f) plane.

The JD [3] of a set $\{\mathbf{M}_k | k = 1, \dots, p\}$ of p matrices is defined as the maximization of the JD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^p \sum_{i=1}^n |\mathbf{v}_i^H \mathbf{M}_k \mathbf{v}_i|^2. \quad (8.4.9)$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ [3]. An efficient joint approximate diagonalization algorithm exists in Ref. [3] and it is a generalization of the Jacobi technique [25] for the exact diagonalization of a single normal matrix.

8.4.1.3.3 Source separation using JAD

By selecting cross-term (t,f) points, the whitened STFD matrices will have the following structure:

$$\underline{\rho}_{\mathbf{xx}}^{\mathbf{c}}(t,f) = \mathbf{U} \rho_{\mathbf{ss}}^{\mathbf{c}}(t,f) \mathbf{U}^H, \quad (8.4.10)$$

¹As the off-diagonal elements of $\rho_{\mathbf{ss}}^{\mathbf{a}}(t,f)$ are cross-terms, the source TFD matrix is quasi-diagonal for (t,f) points that correspond to a true power concentration, that is, a source auto-term.

²The BSS problem has inherent ambiguity concerning the order and amplitudes of the sources.

where $\underline{\rho}_{ss}^c(t,f)$ is anti-diagonal.³ The missing unitary matrix \mathbf{U} is retrieved by the joint anti-diagonalization (JAD) of a combined set $\{\underline{\rho}_{xx}^c(t_i, f_i) \mid i = 1, \dots, q\}$ of q source cross-term STFD matrices [3].

The JAD is explained by first noting that the problem of anti-diagonalization of a single $n \times n$ matrix \mathbf{N} is equivalent⁴ to maximizing

$$C(\mathbf{N}, \mathbf{V}) \stackrel{\text{def}}{=} - \sum_{i=1}^n |\mathbf{v}_i^H \mathbf{N} \mathbf{v}_i|^2 \quad (8.4.11)$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$. Hence, JAD of a set $\{\mathbf{N}_k \mid k = 1, \dots, q\}$ of q $n \times n$ matrices is defined as the maximization of the JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^q C(\mathbf{N}_k, \mathbf{V}) = - \sum_{k=1}^q \sum_{i=1}^n |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \quad (8.4.12)$$

under the same unitary constraint.

More generally, we can combine JD and JAD of two sets $\{\mathbf{M}_k \mid k = 1, \dots, p\}$ and $\{\mathbf{N}_k \mid k = 1, \dots, q\}$ of $n \times n$ matrices by maximizing the JD/JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{i=1}^n \left(\sum_{k=1}^p |\mathbf{v}_i^H \mathbf{M}_k \mathbf{v}_i|^2 - \sum_{k=1}^q |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \right) \quad (8.4.13)$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$. The combined JD/JAD criterion can be applied to a combined set of p (source auto-term) STFD matrices and q (source cross-term) STFD matrices to estimate the unitary matrix \mathbf{U} (see implementation in Section 8.7.5).

Notes: (1) The success of the JD or JAD of STFD matrices in determining the unitary matrix \mathbf{U} depends strongly on the correct selection of the auto-term and cross-term points [3]. Therefore, it is crucial to have a selection procedure that is able to distinguish between auto-term and cross-term points based only on the STFD matrices of the observations. A selection approach that exploits the anti-diagonal structure of the source cross-term STFD matrices is described in Ref. [3]. More precisely, for a source cross-term STFD matrix, we have

$$\text{Trace}(\underline{\rho}_{xx}^c(t,f)) = \text{Trace}(\mathbf{U} \underline{\rho}_{ss}^c(t,f) \mathbf{U}^H) = \text{Trace}(\underline{\rho}_{ss}^a(t,f)) \approx 0.$$

Based on this observation, the following testing procedure applies:

$$\begin{aligned} & \text{if } \frac{\text{Trace}(\underline{\rho}_{xx}^c(t,f))}{\text{norm}(\underline{\rho}_{xx}^c(t,f))} < \epsilon \longrightarrow \text{decide that } (t,f) \text{ is a cross-term} \\ & \text{if } \frac{\text{Trace}(\underline{\rho}_{xx}^c(t,f))}{\text{norm}(\underline{\rho}_{xx}^c(t,f))} > \epsilon \longrightarrow \text{decide that } (t,f) \text{ is an auto-term,} \end{aligned}$$

where ϵ is a “small” positive real scalar.

³As the diagonal elements of $\underline{\rho}_{ss}^c(t,f)$ are auto-terms, the source TFD matrix is quasi anti-diagonal (i.e., diagonal entries are close to zero) for each (t,f) point corresponding to a cross-term.

⁴This is due to the fact that the Frobenius norm of a matrix is constant under unitary transform, that is, $\text{norm}(\mathbf{N}) = \text{norm}(\mathbf{V}^H \mathbf{N} \mathbf{V})$.

- (2) In practice, the source cross-term STFD matrices are not purely anti-diagonal. This is because some auto-terms, through their side lobes or main lobes, will intrude over the cross-term regions. The cross-terms are, however, the dominant components. This is similar to the case of JD of STFD matrices selecting auto-term points [3], where the source auto-term STFD matrices are not purely diagonal because of cross-term intrusion. This impairment is mitigated by the joint approximation property of the JD/JAD algorithm and by its robustness.
- (3) Other classes of TFDs and techniques can also be used in BSS; for example, a cumulant-based fourth-order Wigner distribution or Wigner trispectrum is used for source separation [26]. Blind separation of *more* sources than sensors (underdetermined BSS) is solved using a (t,f) domain orthogonality concept [27] (see [Section 8.5](#)). Implementation details and the corresponding MATLABTM code of the above algorithm are given in [Section 8.7](#). Sample code for computation of TFDs is given in [Section 6.5](#).

8.4.2 SEPARATION OF CONVOLUTIVE MIXTURES

8.4.2.1 Data model

Consider now a convolutive (i.e., with delayed elements as in multipath) multiple input multiple output (MIMO) linear time invariant model given by:

$$x_i(t) = \sum_{j=1}^n \sum_{l=0}^L a_{ij}(l) s_j(t-l) \quad \text{for } i = 1, \dots, m, \quad (8.4.14)$$

where $s_j(t)$, $j = 1, \dots, n$, are the n source signals (model inputs), $x_i(t)$, $i = 1, \dots, m$, are the $m > n$ sensor signals (model outputs), a_{ij} is the transfer function between the j th source and the i th sensor with an overall extent of $(L + 1)$ taps. As before, the sources are assumed to have different structures and localization properties in the (t,f) domain and the channel matrix \mathbf{A} defined in Eq. (8.4.16) is full-column rank.

In matrix form, we have

$$\mathbf{x}(t) = \mathbf{As}(t), \quad (8.4.15)$$

where

$$\begin{aligned} \mathbf{s}(t) &= [s_1(t), \dots, s_1(t - (L + L') + 1), \dots, s_n(t), \dots, s_n(t - (L + L') + 1)]^T \\ \mathbf{x}(t) &= [x_1(t), \dots, x_1(t - L' + 1), \dots, x_m(t), \dots, x_m(t - L' + 1)]^T \end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mn} \end{bmatrix} \quad (8.4.16)$$

with

$$\mathbf{A}_{ij} = \begin{bmatrix} a_{ij}(0) & \cdots & a_{ij}(L) & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_{ij}(0) & \cdots & a_{ij}(L) \end{bmatrix}. \quad (8.4.17)$$

Note that \mathbf{A} is a $[mL' \times n(L + L')]$ matrix and \mathbf{A}_{ij} are $[L' \times (L + L')]$ matrices. The parameter L' is chosen such that $mL' \geq n(L + L')$.

The formalism is the same as in the instantaneous mixture case. The data STFD matrices still have the same expression as in Eq. (8.4.5). But the source auto-term (respectively, cross-term) matrices $\rho_{ss}(t, f)$ are no longer diagonal (respectively, anti-diagonal), but block-diagonal⁵ (respectively, block anti-diagonal) where each diagonal block is of size $(L + L') \times (L + L')$. It is this block diagonal or block anti-diagonal structure that we exploit, in the next section, to achieve BSS.

8.4.2.2 BSS using STFD matrices for convolution mixtures

Let us now generalize the above BSS method to the case of convolutive mixtures.

8.4.2.2.1 Whitening

The first step of the procedure is to whiten the data vector $\mathbf{x}(t)$ by applying to $\mathbf{x}(t)$ a *whitening matrix* \mathbf{W} , that is, a $[n(L' + L) \times mL']$ matrix verifying:

$$\mathbf{W} \lim_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T \mathbf{x}(t) \mathbf{x}(t)^H \right) \mathbf{W}^H = \mathbf{W} \mathbf{R}_x \mathbf{W}^H = (\mathbf{W} \mathbf{R}_s^{1/2}) (\mathbf{W} \mathbf{R}_s^{1/2})^H = \mathbf{I}, \quad (8.4.18)$$

where \mathbf{R}_x and \mathbf{R}_s denote the autocorrelation matrices⁶ of $\mathbf{x}(t)$ and $\mathbf{s}(t)$, respectively. Equation (8.4.18) shows that if \mathbf{W} is a whitening matrix, then

$$\mathbf{U} = \mathbf{W} \mathbf{R}_s^{1/2} \quad (8.4.19)$$

is an $[n(L' + L) \times n(L' + L)]$ unitary matrix where $\mathbf{R}_s^{1/2}$ (Hermitian square root matrix of \mathbf{R}_s) is block diagonal. The whitening matrix \mathbf{W} can be determined from the eigendecomposition of the data autocorrelation \mathbf{R}_x as in [3].

8.4.2.2.2 Separation using JBD

Consider the whitened STFD matrices $\underline{\rho}_{xx}(t, f)$ defined in Eq. (8.4.6). By Eqs. (8.4.15) and (8.4.19), we obtain the key relation:

$$\underline{\rho}_{xx}(t, f) = \mathbf{U} \mathbf{R}_s^{-1/2} \rho_{ss}(t, f) \mathbf{R}_s^{-1/2} \mathbf{U}^H = \mathbf{U} \rho(t, f) \mathbf{U}^H \quad (8.4.20)$$

where we have set $\rho(t, f) = \mathbf{R}_s^{-1/2} \rho_{ss}(t, f) \mathbf{R}_s^{-1/2}$.

Since the matrix \mathbf{U} is unitary and $\rho(t, f)$ is block diagonal, the latter just means that any whitened STFD matrix is block diagonal in the basis of the column vectors of matrix \mathbf{U} . The unitary matrix can be retrieved by computing the block diagonalization of some matrix $\underline{\rho}_{xx}(t, f)$. But to reduce the likelihood of indeterminacy and increase the robustness of determining \mathbf{U} , we consider the joint block diagonalization (JBD) of a set $\{\underline{\rho}_{xx}(t_i, f_i); i = 1, \dots, p\}$ of p whitened STFD matrices.⁷

⁵The block diagonal structure comes from the fact that the cross-terms between $s_i(t)$ and $s_j(t - d)$ are not zero and depend on the local correlation structure of the signal.

⁶In practice, \mathbf{R}_x and \mathbf{R}_s are replaced by time-averaged estimates, for example, $\mathbf{R}_x = (\sum_{t=1}^T \mathbf{x}(t) \mathbf{x}(t)^H) / T$.

⁷A similar procedure can be used with joint block anti-diagonalization of source cross-term STFD matrices.

This JBD is achieved by the maximization under unitary transform of the following criterion,

$$C(\mathbf{U}) \stackrel{\text{def}}{=} \sum_{k=1}^p \sum_{l=1}^n \sum_{i,j=(L'+L)(l-1)+1}^{(L'+L)l} \left| \mathbf{u}_i^* \underline{\rho}_{\mathbf{xx}}(t_k, f_k) \mathbf{u}_j \right|^2, \quad (8.4.21)$$

over the set of unitary matrices $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{n(L'+L)}]$. Note that an efficient Jacobi-like algorithm for JBD algorithm exists in Refs. [25,28].

Once the unitary matrix \mathbf{U} is determined (up to a block diagonal unitary matrix \mathbf{D} coming from the inherent indeterminacy of the JBD problem [29]), the recovered signals are obtained up to a filter⁸ by

$$\hat{\mathbf{s}}(t) = \mathbf{U}^H \mathbf{W} \mathbf{x}(t). \quad (8.4.22)$$

According to Eqs. (8.4.15) and (8.4.19), the recovered signals verify,

$$\hat{\mathbf{s}}(t) = \mathbf{D} \mathbf{R}_s^{-1/2} \mathbf{s}(t) \quad (8.4.23)$$

where the matrix $\mathbf{R}_s^{-1/2}$ is block diagonal and \mathbf{D} is a block diagonal unitary matrix.

- Notes:**
- (1) In practice, only n signals among the $n(L' + L)$ recovered ones are selected. This is done by choosing the signals which lead to the smallest cross-terms coefficients. Note that this information is a by-product of the JBD procedure and hence this selection needs no additional computations.
 - (2) Here, we have considered source separation up to a filter, instead of the full MIMO deconvolution procedure. Note that if needed a single input multi output (SIMO) deconvolution/equalization [30] can be applied to the separated sources.

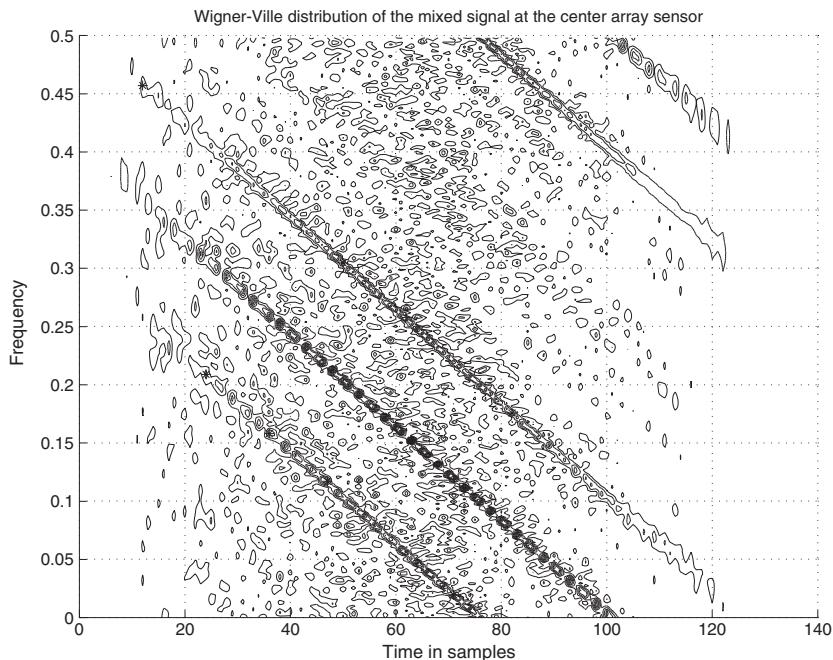
8.4.3 ILLUSTRATIVE EXAMPLES

The following two examples illustrate the effectiveness of the (t,f) approach in achieving blind separation of nonstationary sources.

8.4.3.1 Experiment 1: Separation of instantaneous mixtures

Let us consider two chirp signals ($n = 2$), $s_1(t) = \exp(-j0.004\pi t^2)$ and $s_2(t) = \exp(-j0.004\pi t^2 - j\pi 0.4t)$, impinging on an array of $m = 5$ sensors at 30° and 60° . White Gaussian noise is added, leading to an SNR of 0 dB. The Wigner-Ville distribution (WVD) of the mixture at the middle sensor is shown in Fig. 8.4.1, from which we selected eight arbitrary (t,f) points, among which one was a cross-term. By applying the JD/JAD algorithm, we obtain the results shown in Fig. 8.4.2 with a signal rejection level estimate of -26 dB.

⁸The separated sources correspond to filtered versions of the original ones, that is, $\hat{s}_i(t) = s_i(t) * h_i(t)$ where $h_i(t)$ is an unknown filter and $*$ stands for the convolution.

**FIGURE 8.4.1**

WVD of two mixed signals at 0dB SNR.

8.4.3.2 Experiment 2: Separation of convulsive mixtures of speech signals

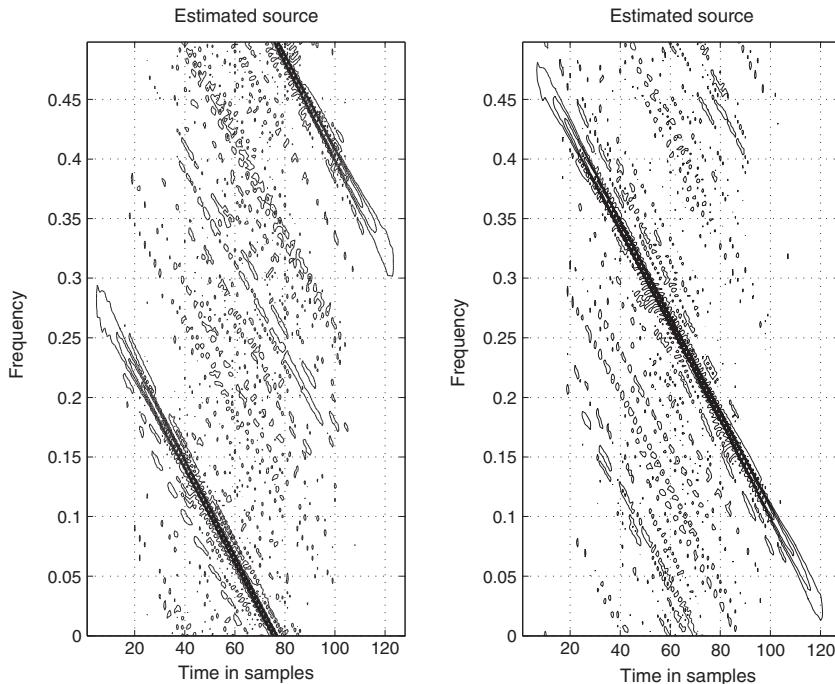
Let us consider $n = 2$ (speech signals sampled at 8 kHz), $m = 3$, $L = 1$, and $L' = 2$. The transfer function matrix of the simulated multichannel is given by:

$$\mathbf{A}(z) = \begin{bmatrix} 1 & 0.85 + 0.1z^{-1} \\ 0.7 + 0.4z^{-1} & 0.25 + z^{-1} \\ 1 + 0.5z^{-1} & 0.7 + 0.85z^{-1} \end{bmatrix}.$$

Figure 8.4.3 shows the original speech signals, their speech convulsive mixture, and the recovered speech signals using the (t,f) separation algorithm.

8.4.4 SUMMARY AND CONCLUSIONS

This section presents solutions to the problem of blind separation of linear mixtures of nonstationary source signals based on (t,f) distributions. Both instantaneous and convulsive mixture cases have been considered. In both cases, solutions based on the use of the joint matrix structure (i.e., diagonal, anti-diagonal, block diagonal, or block anti-diagonal) of a combined set of spatial (t,f) distribution matrices selected in both the auto-term and cross-term regions have been presented.

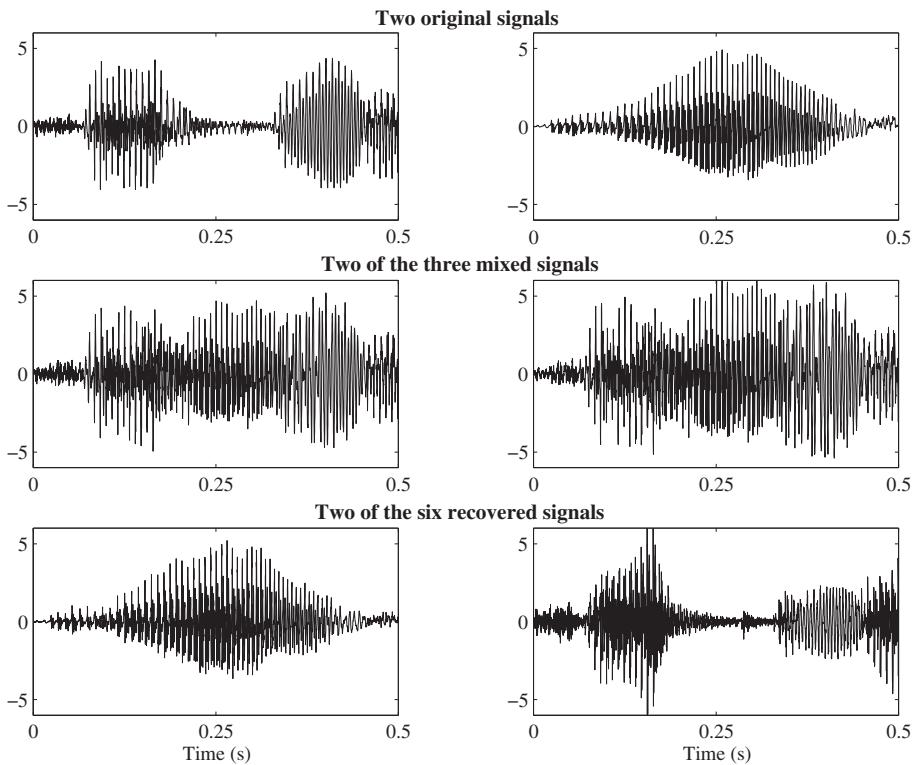
**FIGURE 8.4.2**

WVDs of the two chirps using JD/JAD with seven auto-terms and one cross-term.

8.5 UNDERDETERMINED BLIND SOURCE SEPARATION FOR FM-LIKE SIGNALS⁰

In many applications, including sonar and telecommunications, the signal under consideration may consist of a mixture of several signals. The user may not be interested in the whole mixture signal (also called a multicomponent signal), but rather in one or more particular components of it. For instance, in telecommunications, the received signal may be a mixture of several source signals (multiple access interference) but the user may wish to recover only one or a few individual source signals. The various component signals may be nonstationary (i.e., their spectral contents may vary with time); and since time-frequency (t,f) analysis has proved to be an effective tool in the analysis of such signals, we can use this technique to extract the desired source signal, and separate/recover any signal component (see [Section 8.4](#) for a detailed justification).

⁰Authors: **N. Linh-Trung**, Vietnam National University, Hanoi, Vietnam (nltrung@vnu.edu.vn), **K. Abed-Meraim**, Université d'Orléans, Orléans, France (karim.abed-meraim@univ-orleans.fr), **A. Belouchrani**, Ecole Nationale Polytechnique, Algiers, Algeria (belouchrani@hotmail.com), and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: V. Chandran and A. Beghdadi.

**FIGURE 8.4.3**

Speech signal separation.

The blind source separation (BSS) algorithms given in Section 8.4 have limitations in situations where there are more sources than sensors and they fail to separate sources. Such problems are known as the under-determined BSS.

This section therefore focuses on the problem of underdetermined blind source separation (UBSS) where the number of sources is larger than the number of sensors, and the sources are nonstationary frequency-modulated (FM) signals as encountered in radar and some wireless communication systems. For instance, linear FM chirp signaling is considered for frequency-hopped CDMA (Code Division Multiple Access) systems [31–33], where UBSS coincides with the blind multiuser detection problem.

Two classes of methods allow UBSS using source spatial signatures in conjunction with their (t,f) signatures. The two methods are based on a vector clustering technique and a mono-component extraction technique, respectively.

8.5.1 DATA MODEL AND ASSUMPTIONS

As in Section 8.4.1.1, let us assume that an n -dimensional vector $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T \in \mathbb{C}^{(n \times 1)}$ corresponds to n nonstationary complex source signals $s_i(t)$, $i = 1, \dots, n$. The source signals are

transmitted through a medium so that an array of m sensors picks up a set of mixed signals represented by an m -dimensional vector $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T \in \mathbb{C}^{(m \times 1)}$. For an instantaneous linear mixture medium, the observed signals can, then, be modeled as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}(t), \quad (8.5.1)$$

where $\mathbf{A} \in \mathbb{C}^{(m \times n)}$ is the mixing matrix and $\boldsymbol{\eta}(t) = [\eta_1(t), \eta_2(t), \dots, \eta_m(t)]^T \in \mathbb{C}^{(m \times 1)}$ is the observation noise vector. For the UBSS problem, that is, for $n > m$, the mixing matrix \mathbf{A} is not (left) invertible. However, the column vectors of matrix $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ are assumed to be *pairwise linearly independent*, that is, for any $i, j \in 1, 2, \dots, n$ and $i \neq j$, \mathbf{a}_i and \mathbf{a}_j are linearly independent.

Here, the sources are assumed to be multicomponent FM signals whose (t, f) representation shows multiple ridges. The k th source may be defined as,

$$s_k(t) = \sum_{l=1}^{M_k} s_{k,l}(t), \quad (8.5.2)$$

where each component $s_{k,l}(t)$, of the form

$$s_{k,l}(t) = a_{k,l}(t) e^{j\phi_{k,l}(t)}, \quad (8.5.3)$$

is assumed to have only one ridge in the (t, f) plane. An example of a multicomponent signal, consisting of three components, is shown in Fig. 8.5.1.

8.5.2 SEPARATION USING VECTOR CLUSTERING

In this approach, the sources are assumed to have different structures and localization properties in the (t, f) domain so that the sources are orthogonal in the (t, f) domain (Fig. 8.5.2) in the sense that their (t, f) supports¹ are disjoint.

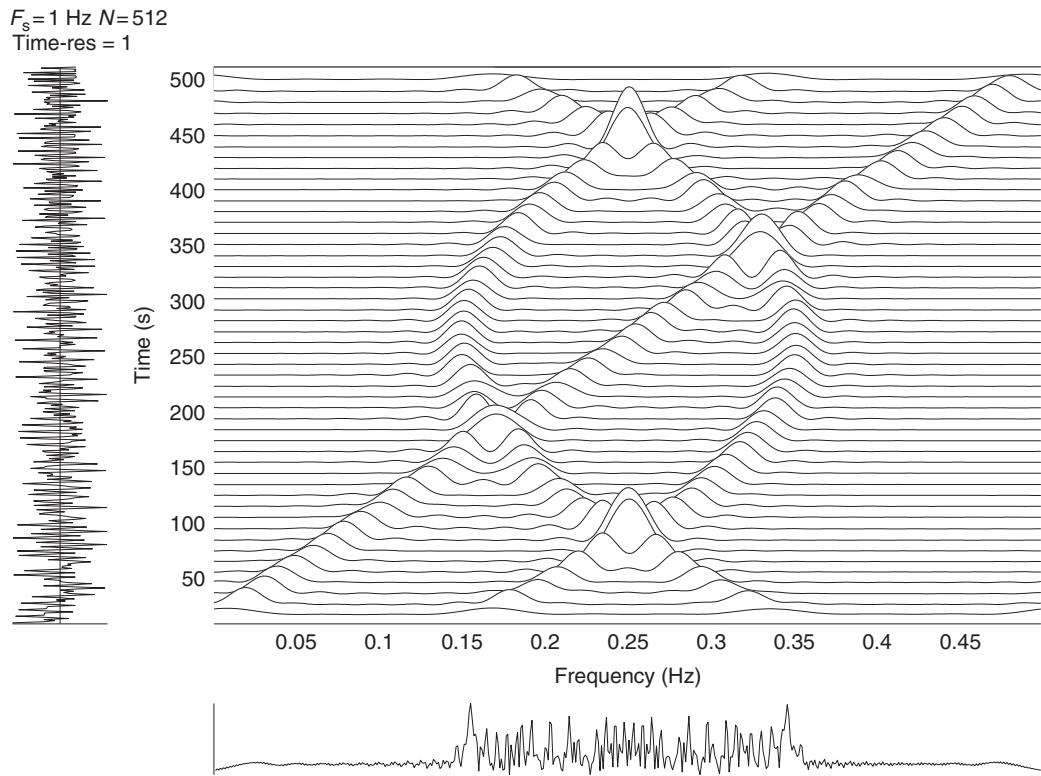
The above assumption can be applied to any TFD. It is clear that the (t, f) orthogonality is too restrictive and will almost never be satisfied exactly in practice. Fortunately, only approximate orthogonality, referred to as *quasi-orthogonality*, is needed in practice to achieve source separation [27]; for example, two FM signals with different FM laws satisfy the quasi-orthogonality assumption.

Under the above assumption, one can notice that two auto-term spatial TFDs (STFD) matrices (see Section 8.4 for a thorough definition) of the observation $\rho_{\mathbf{xx}}(t_1, f_1)$ and $\rho_{\mathbf{xx}}(t_2, f_2)$ corresponding to the same source $s_i(t)$ verify:

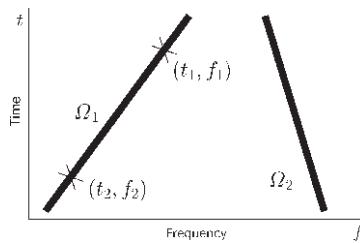
$$\begin{aligned} \rho_{\mathbf{xx}}(t_1, f_1) &= \rho_{s_i s_i}(t_1, f_1) \mathbf{a}_i \mathbf{a}_i^H, \\ \rho_{\mathbf{xx}}(t_2, f_2) &= \rho_{s_i s_i}(t_2, f_2) \mathbf{a}_i \mathbf{a}_i^H. \end{aligned} \quad (8.5.4)$$

Equation (8.5.4) indicates that $\rho_{\mathbf{xx}}(t_1, f_1)$ and $\rho_{\mathbf{xx}}(t_2, f_2)$ have the same principal eigenvector \mathbf{a}_i . The algorithm groups together the auto-term points associated with the same principal eigenvector representing a particular source signal. The TFD of this source ($\rho_{s_i s_i}(t, f)$) is obtained as the principal eigenvalues of the STFD matrices at the auto-term points. The algorithm shown in Fig. 8.5.3 has four steps:

¹The (t, f) support of a signal $s(t)$ is $\{(t, f) | \rho_{ss}(t, f) \neq 0\}$ where $\rho_{ss}(t, f)$ is the TFD of $s(t)$.

**FIGURE 8.5.1**

A time-frequency distribution of a multicomponent signal.

**FIGURE 8.5.2**

(t, f) Orthogonal sources; the (t, f) supports of two sources are disjoint.

- (1) *STFD computation and noise thresholding.* Given a finite set of observation vectors, the STFD matrices of the observation $\rho_{\mathbf{xx}}(t, f)$ can be estimated using time-lag domain discrete implementation as shown in [Sections 6.1](#) and [8.4](#) (see also [Section 8.7](#)). These STFD matrices are

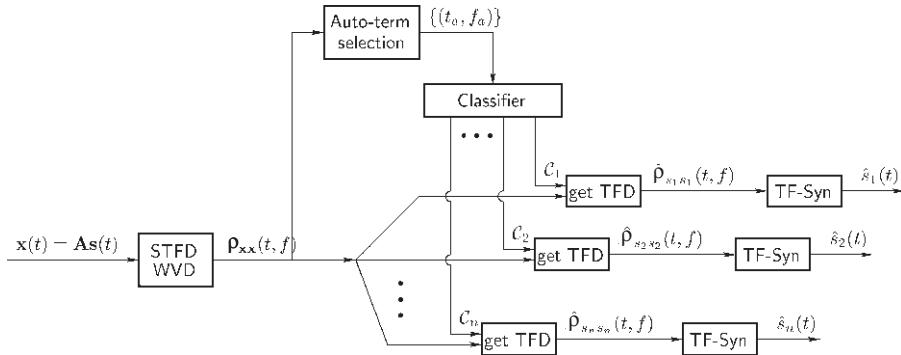
**FIGURE 8.5.3**

Diagram of the UBSS algorithm.

next processed to extract the source signals. In order to reduce the computational complexity by processing only “significant” STFD matrices, a noise thresholding is applied to the signal TFD. More precisely, a threshold ϵ_1 (typically, $\epsilon_1 = 0.05$ of the point with maximum energy) is used to keep only the points $\{(t_s, f_s)\}$ with sufficient energy:

$$\text{Keep } (t_s, f_s) \quad \text{if } \|\rho_{\mathbf{xx}}(t_s, f_s)\| > \epsilon_1. \quad (8.5.5)$$

- (2) *Auto-term selection.* The second algorithm step consists of separating the auto-term points from cross-term points using an appropriate testing criterion. Given the sources (t, f) orthogonality, each auto-term STFD matrix is of rank 1, or at least has one “large” eigenvalue compared to its other eigenvalues. Therefore, one can use rank selection criteria, such as minimum description length (MDL) or Akaike information criterion (AIC), to select auto-term points as those corresponding to STFD matrices of rank 1. For simplicity, we use the following criterion:

$$\text{if } \left| \frac{\lambda_{\max} \{\rho_{\mathbf{xx}}(t, f)\}}{\text{norm} \{\rho_{\mathbf{xx}}(t, f)\}} - 1 \right| > \epsilon_2 \longrightarrow \text{decide that } (t, f) \text{ is a cross-term point}$$

where ϵ_2 is a small positive scalar (typically, $\epsilon_2 = 10^{-4}$), and $\lambda_{\max} \{\rho_{\mathbf{xx}}(t, f)\}$ represents the largest eigenvalue of $\rho_{\mathbf{xx}}(t, f)$ (see p. 486).

- (3) *Clustering and source TFD estimation.* Once the auto-term points have been selected, a clustering step based on the spatial signatures of the sources is performed. This clustering is based on the observation that two STFD matrices corresponding to the same source signal have the same principal eigenvectors. Moreover, the corresponding principal eigenvalues are given by the desired source TFD. This leads to the core step of our source separation method that uses vector clustering. This is implemented by the following four operations:

- (a) For each auto-term point, (t_a, f_a) , compute the main eigenvector, $\mathbf{a}(t_a, f_a)$, and its corresponding eigenvalue, $\lambda(t_a, f_a)$, of $\rho_{\mathbf{xx}}(t_a, f_a)$.
- (b) As the vectors $\{\mathbf{a}(t_a, f_a)\}$ are estimated up to a random phase $e^{j\phi}$, $\phi \in [0, 2\pi)$, we force, without loss of generality, their first entries to be real and positive.

- (c) These vectors are then clustered into different classes $\{\mathcal{C}_i\}$. That is, $\mathbf{a}(t_i, f_i)$ and $\mathbf{a}(t_j, f_j)$ belong to the same class if:

$$d(\mathbf{a}(t_i, f_i), \mathbf{a}(t_j, f_j)) < \epsilon_3, \quad (8.5.6)$$

where ϵ_3 is a properly chosen positive scalar and d is a distance measure (different strategies for choosing the threshold ϵ_3 and the distance d or even the clustering method can be found in Ref. [34]). As an example, in the simulated experiment, we used the angle between the two vectors as a distance measure:

$$d(\mathbf{a}_i, \mathbf{a}_j) = \arccos(\tilde{\mathbf{a}}_i^T \tilde{\mathbf{a}}_j), \quad (8.5.7)$$

where $\tilde{\mathbf{a}} = [\text{Re}(\mathbf{a})^T, \text{Im}(\mathbf{a})^T]^T / \|\mathbf{a}\|$.

- (d) Set the number of sources equal to the number of “significant” classes and, for each source s_i (i.e., each class \mathcal{C}_i), estimate its TFD as

$$\hat{\rho}_{s_i s_i}(t, f) = \begin{cases} \lambda(t_a, f_a), & \text{if } (t, f) = (t_a, f_a) \in \mathcal{C}_i \\ 0, & \text{otherwise.} \end{cases} \quad (8.5.8)$$

- (4) *Source signal synthesis.* Use an adequate source synthesis procedure to estimate the source signals, $s_i(t)$ ($i = 1, \dots, n$), from their respective TFD estimates $\hat{\rho}_{s_i s_i}$ (see (t, f) synthesis algorithms in Chapter 11 and Ref. [35]). An example in Fig. 8.5.4 illustrates the performance of the four-step UBSS algorithm.

8.5.3 SEPARATION USING MONOCOMPONENT EXTRACTION

To improve UBSS, we introduce here a four-step procedure consisting of:

1. computation and spatial averaging of the observed signal TFDs;
2. component extraction to separate all signal (mono) components;
3. component clustering to group together components belonging to the same multicomponent source signal; and
4. (t, f) signal synthesis to recover the original source waveforms.

To have a “clean” TFD (i.e., one that can reveal the features of the signal clearly, without any “ghost” components [created by cross-terms]), we use a high resolution quadratic TFD called the B-distribution (see Sections 2.7, 3.3, and 5.7). In addition, we use a spatial averaging that mitigates further the sources cross-terms by a factor depending on their spatial signatures angle. More precisely, we compute the averaged TFD as (in Section 8.5.3.1 (2), line detection is applied to the averaged TFD)

$$\rho(t, f) = \text{Trace}(\boldsymbol{\rho}_{\mathbf{X}\mathbf{X}}(t, f)) = \sum_{l=1}^m \rho_{x_l x_l}(t, f). \quad (8.5.9)$$

For component clustering, we use the observation that the STFD matrices at two auto-term points corresponding to the same source signal have the same principal eigenvector. Therefore, the proposed component clustering procedure consists of grouping together components associated with the same spatial direction representing a particular source signal. This spatial direction is estimated as the

averaged value over all component points of the principal eigenvectors of the corresponding STFD matrices. More precisely, for each extracted component C , one estimates the corresponding spatial direction as

$$\mathbf{a}_C = \frac{1}{N_{\mathcal{I}_C}} \sum_{i \in \mathcal{I}_C} \mathbf{a}(t_i, f_i), \quad (8.5.10)$$

where \mathcal{I}_C denotes the set of points of component C , $N_{\mathcal{I}_C}$ denotes the number of points in \mathcal{I}_C , and $\mathbf{a}(t_i, f_i)$ is the estimated principal eigenvector of the i th component point STFD matrix $\rho_{xx}(t_i, f_i)$. These vectors are then clustered into different classes using the clustering algorithm in [Section 8.5.2](#).

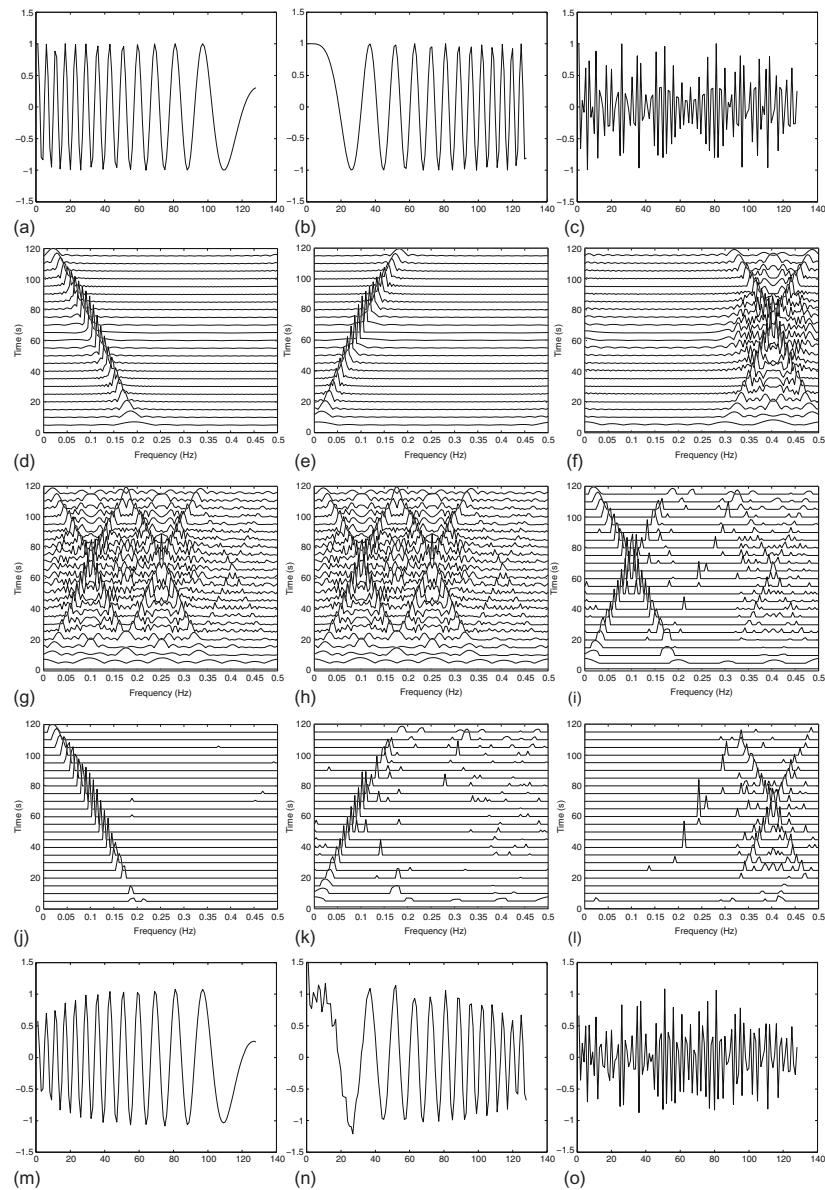
For the component extraction, two different techniques, presented next, can be used. The first is based on a “road extraction” procedure used for road detection in satellite imaging [36] while the second uses “peak detection and tracking” [2].

8.5.3.1 A “road network tracking” approach

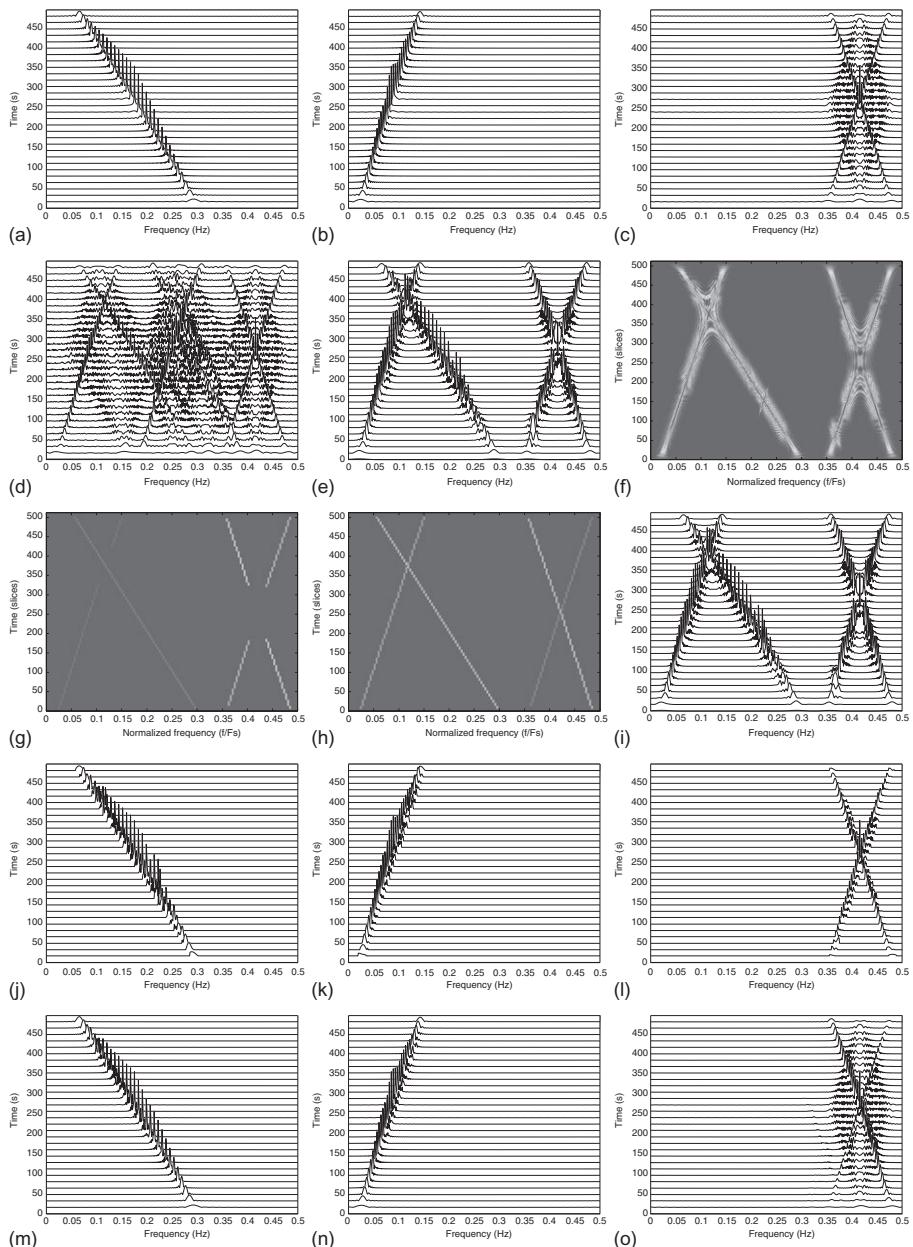
The component-extraction method has three main steps [27]: (i) preprocessing; (ii) line detection giving local binary detection of the potential linear structures (segments); and (iii) global optimization giving a set of labeled components [36].

- (1) *Preprocessing.* First the (t, f) image is transformed to a real positive-valued image by forcing to zero all negative values² of the TFD and by using a gray scale in the range from 1 to 256. Also, line detectors are usually limited to a line width of 5 pixels. If the researched components do not respect this limit (which is usually the case for a (t, f) image), an image subsampling by block-averaging is applied to reduce the pixel size. Despite the blurring effect, this filter presents the advantage of reducing the noise in the (t, f) image. Moreover, as the (t, f) image is unisotropic (i.e., it contains horizontal lines as observed in [Fig. 8.5.1](#)), this image downsampling (see [Fig. 8.5.4\(f\)](#)) removes this particular feature of the (t, f) image.
- (2) *Line detection (Local optimization).* A line detector used in radar imaging [36] is applied at each pixel of the image. For a given direction, its response is based on the ratio of the means computed on both sides of the suspected line and the mean of the line itself. Height directions are studied and the best response is kept. The resulting image is then binarized using a simple thresholding. If statistics on the image are available (noise distribution, additive or multiplicative noise, etc.), a statistical study of the line detector performance can be made to choose the more adapted threshold (for instance, the threshold corresponding to a fixed false alarm rate in homogeneous areas).
- (3) *Road detection (global optimization).* This step is a global step introducing constraints on the shape of the linear features to extract connected components and to suppress false alarms [36]. It works on segments extracted on the thresholded line response image by thinning and linearization. The previously detected segments are connected depending on proximity and alignment constraints (specially on the line curvature) to form coherent components. Small isolated segments are suppressed. The algorithm depends on three thresholds: the maximum gap between two segments to connect them, the allowed angular difference between the two segments, and the minimum size of a component. The result is a labeled image of components. [Figure 8.5.5](#) shows an example illustrating the separation technique performance.

²Negative values correspond mainly to undesired cross-terms or noise.

**FIGURE 8.5.4**

Testing the “vector clustering” UBSS algorithm for $m = 2$ mixtures of two monocomponent and one multicomponent LFM signals denoted by $s_1(t)$, $s_2(t)$, and $s_3(t)$ (a-c). The recovered source signals (m-o) indicated the success of the UBSS algorithm. Source $s_3(t)$ was not falsely separated into two monocomponent sources. (a) Source $s_1(t)$; (b) source $s_2(t)$; (c) source $s_3(t)$; (d) WVD of $s_1(t)$; (e) WVD of $s_2(t)$; (f) WVD of $s_3(t)$; (g) WVD of observed mix; (h) noise-reduced mix; (i) X-term-removed mix; (j) separated (t,f) signature of $s_1(t)$; (k) separated (t,f) signature of $s_2(t)$; (l) separated (t,f) signature of $s_3(t)$; (m) recovered $\hat{s}_1(t)$; (n) recovered $\hat{s}_2(t)$; and (o) recovered $\hat{s}_3(t)$.

**FIGURE 8.5.5**

(a-c) WVD of $s_1(t)$, $s_2(t)$, $s_3(t)$; (d, e) spatial-averaged TFD of the mixture outputs using WVD and MWVD; (f) STFD mixture converted to image; (g-h) extraction of source components using “road network tracking”; (i) auto-term points of known components; and (j-l) TFD estimates of the sources; (m-o) TFD of estimated sources after (t,f) synthesis: $m = 2$ sensors and SNR = 10 dB.

8.5.3.2 A “peak detection and tracking” approach

Figure 8.5.6 and Tables 8.5.1 and 8.5.2 illustrate a component separation algorithm [2] that assumes all components of the signal exist at almost all time instants. It has three key stages.

- (1) *Noise thresholding.* Noise thresholding is applied to remove the undesired “low” energy peaks in the (t,f) domain.³ We set to zero the TFD values smaller than a properly chosen threshold ϵ . In MATLABTM, this is written as

$$\text{rho_th}(t, f) = \text{rho}(t, f) \cdot (\text{rho}(t, f) > \text{epsilon}).$$
- (2) *Estimation of the number of components.* In general, for a noiseless and cross-terms-free TFD, the number of components at a given time instant t can be estimated as the number of peaks of the TFD slice $\rho(t,f)$. A simple technique to estimate the number of components is to efficiently evaluate the number of components (see Table 8.5.1) as the maximum argument of the histogram of the number of peaks computed for each time instant t in the range $[1, 2, \dots, t_{\max}]$ (where $t_{\max} \times f_{\max}$ is the dimension of the TFD matrix).
- (3) *Component separation procedure.* Assuming that (i) all components exist at all time instants in the (t,f) plane and (ii) any component intersection is a crossing point, we note that if, at a time instant t , two or more components are crossing, then the number of peaks (at this particular slice $\rho(t,f)$) is smaller than the total number of components d (see Table 8.5.2).

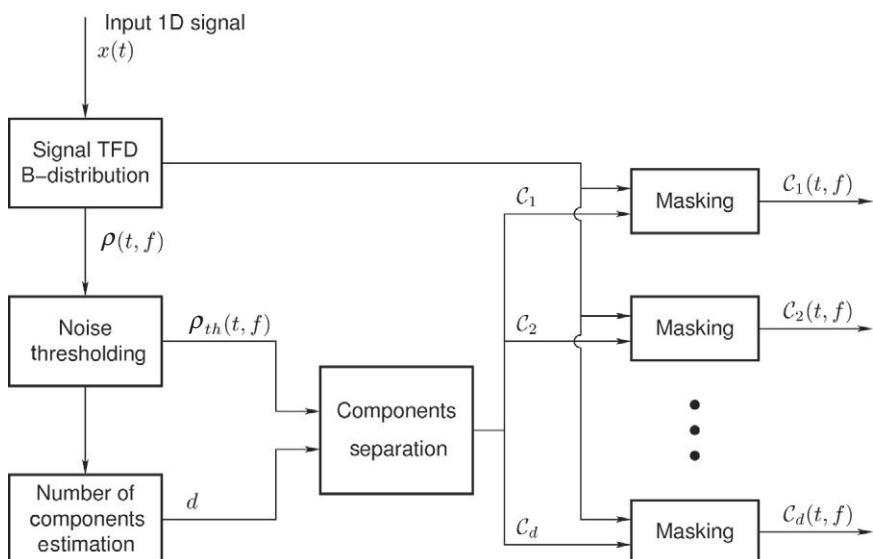


FIGURE 8.5.6

Flowchart of the “component extraction”-based algorithm.

³Noise energy is spread over the whole (t,f) domain while the components energies are well localized around their respective IFs leading to high energy peaks for the latter (assuming no cross-terms).

Table 8.5.1: Estimation of the Number of Components

1. For $t = 1, \dots, t_{\max}$
 $\text{number_components}(t) = \text{number of peaks of } \rho(t, :)$
end
2. Compute the histogram H of number_components
3. Evaluate the number of components as $d = \arg \max H$

Table 8.5.2: Component Separation Procedure for the “Peak Detection and Tracking” Algorithm

1. Assign an index to each of the d components in an orderly manner.
2. For each time instant t (starting from $t = 1$), find the component frequencies as the peak positions of the TFD slice $\rho(t, f)$.
3. Assign a peak to a particular component based on the smallest distance to the peaks of the previous slice $\rho(t - 1, f)$ (IFs continuous functions of time).
4. If at a time instant t a crossing point exists (i.e., number of peaks smaller than the number of components), identify the crossing components using the smallest distance criterion by comparing the distances of the actual peaks to those of the previous slice.
5. Permute the indices of the corresponding crossing components.

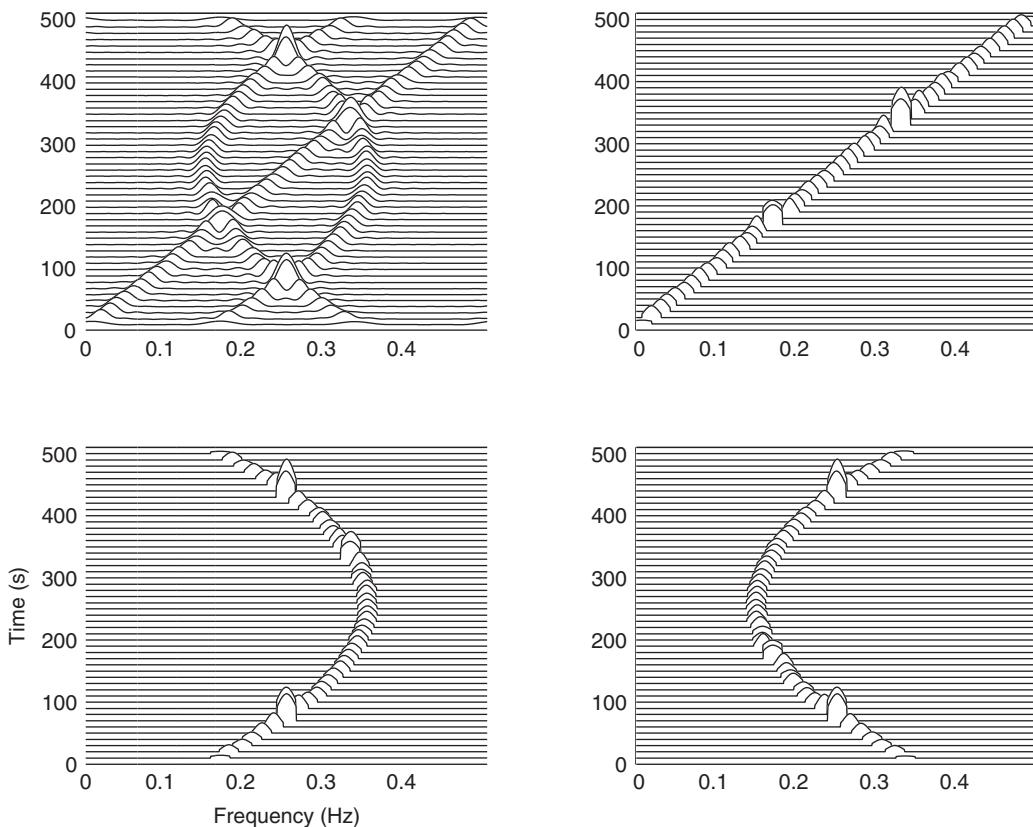
To validate the algorithm, we reconsider the same multicomponent signal analyzed in Fig. 8.5.1. This signal consists of a mixture of two quadratic FM components and a linear FM component. The mixture signal is added to a zero-mean white Gaussian noise, with SNR equal to 0 dB. The B-distribution of the noisy signal as well as the components resulting from the separation algorithm are shown in Fig. 8.5.7.

8.5.4 SUMMARY AND CONCLUSIONS

This section presents (t, f) methods for the underdetermined blind separation of FM-like signals. The first class of methods is based on a vector clustering approach while the second is based on a component extraction approach. Two component extraction techniques are introduced using a “road network tracking” algorithm and a “peak detection and tracking” algorithm, respectively.

Simulation results indicate that the UBSS method based on vector clustering is more general in the sense that it can be applied to separate a larger class of (not necessarily FM) signals as long as they satisfy the (t, f) orthogonality property. It is related to the “sparse decomposition”-based UBSS methods [37] which essentially use the sparse representation of the source signals in a given transformed domain (in our case, it is the (t, f) domain). The “vector clustering”-based method is, however, less performant than the second class of methods using component extraction when considering only FM-type signals. Several component extraction techniques exist in the literature [2, 27, 38] among which two have been presented in this section. The best promising technique (i.e., the one providing the best separation quality) is the “image-processing”-based one using a “road detection” algorithm. In terms of computational cost, it is, however, much more expensive than the “peak detection and tracking” technique.

Note: based on the proposed method of the time-frequency-based UBSS in this section, there have been a number of developments in this direction to facilitate the separation of other types of signals which are nondisjoint in the time-frequency domain (e.g., [27, 39, 40]).

**FIGURE 8.5.7**

The B-distribution of the original signal (top left) as well as the extracted components using the “peak detection and tracking” algorithm.

8.6 AUDIO SOURCE LOCALIZATION AND SEPARATION USING TIME-FREQUENCY REPRESENTATIONS⁰

Source localization using array signal processing is an important method for audio- and speech-processing systems. For example, knowledge about the location of a potential listener can help a system provide a better audio service by using beamforming techniques from an array of microphones. Similarly, beamforming can also be used to capture an audio signal from a desired source location and attenuate surrounding environmental noises or undesired signals. The location information of the source

⁰Authors: **A. Aïssa-El-Bey**, Institut Mines-Télécom, Télécom Bretagne, Brest Cedex 3, France (abdeldjalil.aissa-el-bey@telecom-bretagne.eu) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: Y. Zhang.

can also help in separating the desired audio sources from each other or from undesired audio source signals, using blind source separation (BSS) of the observations recorded by the microphone array.

Sections 8.1 and 8.4 describe how BSS techniques can recover original source signals from a mixture of different source signals. BSS is an essential preprocessing step for numerous practical applications. For instance, in hearing-aid systems, BSS estimates the original sound signals from mixtures, enhances desired signals, and suppresses undesired noise and interferences. In speech recognition, source signals are first separated, and relevant features are then extracted to perform the speech recognition stage.

The above-mentioned applications are convolutive in nature, as the environments are in general reverberant. Given that the separation of multiple audio sources is mostly needed in such reverberant cases, the main focus of this section is to formulate and model the BSS problem in the context of convolutive mixtures using time-frequency (t,f) methods. In addition, we also discuss how location information can be used to solve the BSS problem from convolutive audio signal mixtures.

8.6.1 DATA MODEL AND ASSUMPTIONS

Let us consider a mixture of n source signals $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T$ resulting in m observations received at an array of m sensors. The received signals are denoted $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T$. The convolutive model introduces the following relationship between the mixed signals and the original source signals

$$\mathbf{x}(t) = \sum_{l=0}^L \mathbf{A}(l) \mathbf{s}(t-l) + \mathbf{n}(t), \quad (8.6.1)$$

where $\mathbf{A}(l) = [\mathbf{a}_1(l), \dots, \mathbf{a}_n(l)]$ are $m \times n$ matrices for $l = 0, \dots, L$, which represent the impulse response coefficients of the propagation channel. In the case of anechoic (i.e., without echo) mixtures where no reflection is assumed from the objects (or reflection contribution is neglected), the model given by Eq. (8.6.1) simplifies to

$$x_i(t) = \sum_{k=1}^n \lambda_{ik} s_k(t - \tau_{ik}) + n_i(t), \quad (8.6.2)$$

where λ_{ik} and τ_{ik} are attenuation coefficients and time delays associated with the path from the i th source to the k th sensor.

Then, by taking the short time Fourier transform (STFT) of $\mathbf{x}(t)$, the mixing model given by Eqs. (8.6.1) and (8.6.2) can be written in a matrix form as

$$F_{\mathbf{x}}^W(t,f) = \mathbf{A}(f) F_{\mathbf{s}}^W(t,f) + F_{\mathbf{n}}^W(t,f), \quad (8.6.3)$$

where $\mathbf{A}(f) = [\mathbf{a}_1(f), \dots, \mathbf{a}_n(f)]$ and

$$a_{ik}(f) = \lambda_{ik} \exp(-j2\pi f \tau_{ik}). \quad (8.6.4)$$

The following sections consider the anechoic model case.

8.6.2 AUDIO SOURCE LOCALIZATION

8.6.2.1 Localization based on parametric method

Localization of the audio source using microphone arrays is described in [Section 8.1](#). Among various audio source localization techniques, subspace methods are widely used to exploit the information of DOAs of the source signals. In [Section 8.1](#), details are provided about the MUSIC algorithm, which is the most popular subspace method for DOA estimation. This MUSIC algorithm is robust against noise, but it assumes there are more sensors (e.g., antennas or microphones) than sources. In the underdetermined case, that is, there are more sources than sensors; hence, the traditional MUSIC algorithm cannot be used.

8.6.2.2 Localization with ad hoc sensor positioning

The DOA estimation methods presented in [Sections 8.1](#) and [8.2](#) assume controlled sensor positioning (where the relative positions of the sensors are known). In the case of *ad hoc* sensor positioning, the methods presented in Refs. [41,42] consider the underdetermined cases where $m < n$, and assume the signals are sufficiently sparse in the (t,f) domain, and that the sources are disjoint in the domain; in other words, there is at most, one source present at any point in the (t,f) domain. Then, when signal $s_i(t)$ is present at (t,f) , the convolutive model given by Eq. [\(8.6.3\)](#) can be approximated by [39,43]

$$F_{\mathbf{x}}^w(t,f) \approx \mathbf{a}_i(f) F_{s_i}^w(t,f). \quad (8.6.5)$$

Therefore, the phase difference between sensor observations at this (t,f) point holds the geometric information of the dominant source at each (t,f) point.

It follows that in order to estimate the DOAs of sources, under an anechoic model, the argument of the frequency response $\mathbf{a}_i(f)$ is expressed solely by the time-delay $\boldsymbol{\tau}_i = [\tau_{1i}, \dots, \tau_{mi}]^T$ with respect to the origin

$$\arg \mathbf{a}_i(f) = -j2\pi f \boldsymbol{\tau}_i. \quad (8.6.6)$$

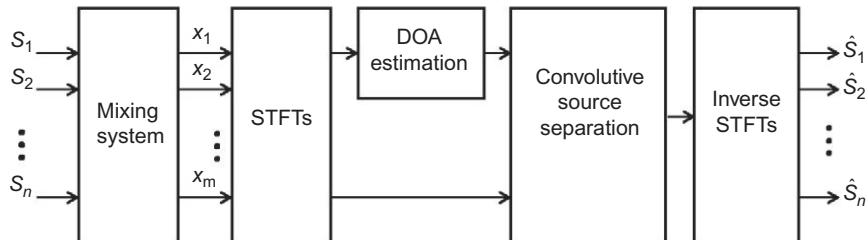
Let $\mathbf{q}_i = [\cos \theta_i \cos \phi_i, \sin \theta_i \cos \phi_i, \sin \phi_i]^T$ be the 3D vectors of a unit norm representing the direction of source s_i . The location of sensor k is given by a 3D vector \mathbf{r}_k . Then, the time-delay τ_{ki} can be expressed with respect to the origin by $\tau_{ki} = \mathbf{r}_k^T \mathbf{q}_i c^{-1}$, where c is the signal propagation velocity.

By clustering the geometric information obtained from each (t,f) point, we can estimate the DOAs by using the cluster centroids and Eq. [\(8.6.6\)](#) (for more details, see Ref. [42]).

8.6.3 UNDERDETERMINED AUDIO SOURCE SEPARATION OF CONVOLUTIVE MIXTURES BASED ON LOCALIZATION INFORMATION

8.6.3.1 Underdetermined audio source separation of convolutive mixtures

High-quality separation of speech sources is an important prerequisite for applications such as speech recognition in environments with several active speakers. Often, the underlying mixing process is unknown, thus requiring BSS. Its potential audio signal applications include speech enhancement for speech recognition, teleconferences, and hearing aids. In such applications, signals are mixed in a convolutive manner with reverberation. This makes the source separation problem difficult. In this section, we are assuming a simplifying assumption by considering an anechoic mixture model where no reflection is assumed from the objects (or reflection contribution is neglected). Another difficulty

**FIGURE 8.6.1**

Flowchart of the underdetermined BSS algorithm in the convulsive mixture case based on localization information.

stems from the fact that there may be more source signals of interest than microphones $m < n$. In this case, the problem is called underdetermined; it is one of the challenging problems of BSS.

Several underdetermined BSS methods for nonstationary sources have been developed, given that these sources are sparse in the (t,f) domain [44]. The main assumption used in these methods is that the sources are disjoint in the (t,f) domain; in other words, there is, at most, one source present at any point in the (t,f) domain. This assumption is rather restrictive, though the methods have also shown that they work well under a quasisparseness condition (i.e., sources are almost disjoint in (t,f) domain). In this method, the localization information can be used for the estimation of the binary masks or prior to achieving the separation (see Fig. 8.6.1).

Approach 1. In the binary masking approach, the sources are extracted by estimating the DOA $\theta(t,f)$ at each (t,f) point and designing a mask such that

$$M_i(t,f) = \begin{cases} 1, & |\theta(t,f) - \theta_i| \leq \delta \\ 0, & \text{otherwise,} \end{cases} \quad (8.6.7)$$

where δ is the extraction range parameter and θ_i is the DOA of the i th source. Then, the estimated mask is applied to the output of one microphone to select the desired signal

$$\hat{F}_{s_i}^w(t,f) = M_i(t,f) F_{x_k}^w(t,f). \quad (8.6.8)$$

Although we can extract each signal using this binary mask, such extracted signals are discontinuously zero-padded by the binary masks. Therefore, we hear considerable musical noise in the extracted output.

Approach 2. Another approach is to estimate the channel by up to a constant scalar by using the DOA information. Assuming that the attenuation coefficients are $\lambda_{1i} = \lambda_{2i} = \dots = \lambda_{mi}$, where $i = 1, 2, \dots, n$, the i th mixing vector becomes

$$\mathbf{a}_i(f) \approx \exp(-j2\pi f c^{-1} \mathbf{d}_i \cos \theta_i), \quad (8.6.9)$$

where $\mathbf{d}_i = [d_{1i}, \dots, d_{mi}]^T$ and d_{ki} is the distance between the i th source and the k th sensor, and c is the propagation velocity of the signals. Then, combine the binary mask with the estimated channel to extract the source

$$\hat{F}_{s_i}^w(t,f) = \begin{cases} \mathbf{a}_i^H(f) F_{\mathbf{x}}^w(t,f), & |\theta(t,f) - \theta_i| \leq \delta, \\ 0, & \text{otherwise.} \end{cases} \quad (8.6.10)$$

Clearly, the (t,f) disjoint assumption is a restrictive condition. When this condition does not hold, the separation performance degrades greatly at those overlapped (t,f) points. In order to overcome this drawback, a subspace projection-based separation method, with sparsity assumption relaxed, is used [39,43]. In this algorithm, the (t,f) representations of different sources are allowed to overlap to the extent that the number of active sources at any (t,f) point is strictly less than the number of sensors. As this is a more relaxing sparsity condition than the (t,f) disjoint assumption where the sources are allowed to overlap in the (t,f) domain, it will be called **(\mathbf{t}, \mathbf{f}) nondisjoint**.

Under the (t,f) nondisjoint assumption, consider a particular (t,f) point (t_0, f_0) at which there are \mathcal{K} sources $s_{\alpha_1}(t), \dots, s_{\alpha_K}(t)$ present, with $\mathcal{K} < m$, where $\alpha_1, \dots, \alpha_K \in \{1, 2, \dots, n\}$ denote the indices of the sources present at the point (t_0, f_0) . The goal is to identify the sources that are present at the point (t_0, f_0) , that is, $\alpha_1, \dots, \alpha_K$, and to estimate the STFT of each of these contributing sources.

Let us define the following:

$$\tilde{\mathbf{s}}(t) = [s_{\alpha_1}(t), \dots, s_{\alpha_K}(t)]^T, \quad (8.6.11a)$$

$$\tilde{\mathbf{A}}_\alpha(f) = [\mathbf{a}_{\alpha_1}(f), \dots, \mathbf{a}_{\alpha_K}(f)]. \quad (8.6.11b)$$

Then, the convolutive model given by Eq. (8.6.3) reduces to the following under the (t,f) nondisjoint assumption

$$F_{\mathbf{x}}^w(t,f) = \tilde{\mathbf{A}}_\alpha(f) F_{\tilde{\mathbf{s}}}^w(t_0, f_0). \quad (8.6.12)$$

Let the indices $\beta_1, \dots, \beta_K \in \{1, 2, \dots, n\}$, the matrix $\tilde{\mathbf{A}}_\beta(f) = [\mathbf{a}_{\beta_1}(f), \dots, \mathbf{a}_{\beta_K}(f)]$ and $\mathbf{Q}_\beta(f)$ be the orthogonal projection matrix onto the noise subspace of $\tilde{\mathbf{A}}_\beta(f)$ expressed by

$$\mathbf{Q}_\beta(f) = \mathbf{I} - \tilde{\mathbf{A}}_\beta(f) \left(\tilde{\mathbf{A}}_\beta^H(f) \tilde{\mathbf{A}}_\beta(f) \right)^{-1} \tilde{\mathbf{A}}_\beta^H(f). \quad (8.6.13)$$

We have the following observation:

$$\begin{cases} \mathbf{Q}_\beta(f) \mathbf{a}_i(f) = 0, & i \in \{\beta_1, \dots, \beta_K\}, \\ \mathbf{Q}_\beta(f) \mathbf{a}_i(f) \neq 0, & \text{otherwise.} \end{cases} \quad (8.6.14)$$

Consequently, as $F_{\mathbf{x}}^w(t,f) \in \text{Range}\{\tilde{\mathbf{A}}_\alpha(f)\}$, we have

$$\begin{cases} \mathbf{Q}_\beta(f) F_{\mathbf{x}}^w(t_0, f_0) = 0, & \text{if } \{\beta_1, \dots, \beta_K\} = \{\alpha_1, \dots, \alpha_K\}, \\ \mathbf{Q}_\beta(f) F_{\mathbf{x}}^w(t_0, f_0) \neq 0, & \text{otherwise.} \end{cases} \quad (8.6.15)$$

If $\mathbf{A}(f)$ has already been estimated by the method presented above based on localization information (see Eq. 8.6.9), then this observation gives us the criterion to detect the indices $\alpha_1, \dots, \alpha_K$; hence, the contributing sources at the considered (t_0, f_0) point. In practice, to take noise into account, the column vectors of $\tilde{\mathbf{A}}_\alpha(f)$ can be detected by minimizing

$$\{\alpha_1, \dots, \alpha_K\} = \arg \min_{\beta_1, \dots, \beta_K} \|\mathbf{Q}_\beta(f) F_{\mathbf{x}}^w(t_0, f_0)\|. \quad (8.6.16)$$

Table 8.6.1: Subspace Projection-Based Separation Algorithm

Input parameters: \mathcal{K}, θ_i with $i = 1, 2, \dots, n$ and STFT of $\mathbf{x}(t)$, $F_{\mathbf{x}}^w(t, f)$.

1. Estimate the channel frequency response based on localization information

$$\mathbf{a}_i(f) \approx \exp(-j2\pi f c^{-1} \mathbf{d}_i \cos \theta_i).$$

2. For all selected (t, f) points, detect the indices of the \mathcal{K} active sources $\{\alpha_1, \dots, \alpha_{\mathcal{K}}\}$ by using subspace projection

$$\{\alpha_1, \dots, \alpha_{\mathcal{K}}\} = \arg \min_{\beta_1, \dots, \beta_{\mathcal{K}}} \|\mathbf{Q}_{\beta}(f) F_{\mathbf{x}}^w(t_0, f_0)\|.$$

3. Form the local (t, f) mixing matrix $\tilde{\mathbf{A}}_{\alpha}(f)$ based on the estimated indices $\tilde{\mathbf{A}}_{\alpha}(f) = [\mathbf{a}_{\alpha_1}(f), \dots, \mathbf{a}_{\alpha_{\mathcal{K}}}(f)]$.

4. The (t, f) contribution of the \mathcal{K} active sources are estimated as $\hat{F}_{\tilde{\mathbf{s}}}^w(t, f) \approx \tilde{\mathbf{A}}_{\alpha}^{\#}(f) F_{\mathbf{x}}^w(t_0, f_0)$.

Output parameters: Estimated source STFT, $\hat{F}_{\tilde{\mathbf{s}}}^w(t, f)$.

Next, the values of the \mathcal{K} sources at the (t, f) point (t_0, f_0) are estimated by

$$\hat{F}_{\tilde{\mathbf{s}}}^w(t_0, f_0) \approx \tilde{\mathbf{A}}_{\alpha}^{\#}(f_0) F_{\mathbf{x}}^w(t_0, f_0), \quad (8.6.17)$$

where $(\#)$ represents the Moore-Penrose's pseudo inversion operator. **Table 8.6.1** shows the key stages of the subspace projection-based separation algorithm.

8.6.3.2 Solving the permutation problem based on DOA information

BSS of convulsive mixtures in the overdetermined case (more sensors than source signals of interest, $m \geq n$) can be carried out using a (t, f) approach, whereby the convulsive mixture is approximated as an instantaneous mixture at each frequency in the (t, f) domain. A filtering operation for generating separated signals is performed by applying $\mathbf{W}(f)$ at each frequency bin f , and then an inverse STFT is applied to the bin-wise separated signals to reconstruct the original signals, where $\mathbf{W}(f)$ is an $n \times m$ separating matrix, such that $\mathbf{W}(f) = [\mathbf{w}_1(f), \dots, \mathbf{w}_n(f)]^T$ (see [Section 2.3.1.1](#) for an introduction of the inverse STFT). In order to reconstruct proper separated signals in the time domain, (t, f) separated signals originating from the same source should be grouped together for each frequency bin f . This is known as the permutation problem.

Indeed, the BSS solution in each frequency bin has permutation and scaling ambiguity; even if we permute the rows of $\mathbf{W}(f)$ or multiply a row by a constant, it is still a BSS solution. The output of the separation filter is processed with the $n \times n$ permutation matrix $\Pi(f)$ and the $n \times n$ diagonal scaling matrix $\mathbf{C}(f)$, such that

$$\mathbf{C}(f) \Pi(f) \mathbf{W}(f). \quad (8.6.18)$$

The scaling ambiguity is removed by using the minimal distortion principle of audio signals in the frequency domain. The permutation ambiguity is solved by using the DOA estimation and analyzing the beamforming pattern formed by the separation matrix and DOA information [\[45\]](#). In order to achieve this, the beamforming pattern $P_{\Pi}(f, \theta_i)$ is estimated for each frequency bin f and source i as

$$P_{\Pi}(f, \theta_i) = \left\| \Pi(f) \mathbf{W}(f) e^{-j2\pi f c^{-1} \mathbf{d}_i \cos \theta_i} \right\|^2. \quad (8.6.19)$$

Then, one can estimate the permutation matrix $\Pi(f)$ at each frequency bin f in order to align the sources by minimizing the following criterion with respect to $\Pi(f)$

$$\arg \min_{\Pi} \sum_{i=1}^n P_{\Pi}(f, \theta_i). \quad (8.6.20)$$

8.6.4 ILLUSTRATIVE EXAMPLE

Let us consider an array of $m = 3$ sensors that receives signals from $n = 4$ independent audio sources (three speech signals corresponding to two men and one woman plus a guitar signal) from directions $\theta_1 = 15^\circ$, $\theta_2 = 25^\circ$, $\theta_3 = 50^\circ$, and $\theta_4 = 70^\circ$, respectively. The signal length corresponds approximately to 6-s recordings of speech signals sampled at 8 kHz.

By applying the subspace projection-based algorithm, we obtain the results shown in Fig. 8.6.2 at SNR 30 dB. Figure 8.6.2(a) represents the m mixture signals and Fig. 8.6.2(b) represents the original source signals. Plots shown in Fig. 8.6.2(c) represent the source estimates by the subspace projection-based algorithm. The results are obtained with a normalized mean square error level of -16 dB.

8.6.5 SUMMARY AND CONCLUSIONS

This section presents solutions to the problem of audio source localization and underdetermined BSS in the convolutive mixtures using localization information in the (t, f) domain. The main assumption of the method described above is that the sources are disjoint in the (t, f) domain; there is, at most, one source present at any point in the (t, f) domain. This section also presents a BSS algorithm based on subspace projection, which relaxes the (t, f) disjoint assumption by supposing that the (t, f) representation of different sources is allowed to overlap to the extent that the number of active sources at any (t, f) point is strictly less than the number of sensors.

8.7 IMPLEMENTATION AND CODE OF STFDs-BASED SOURCE SEPARATION ALGORITHMS⁰

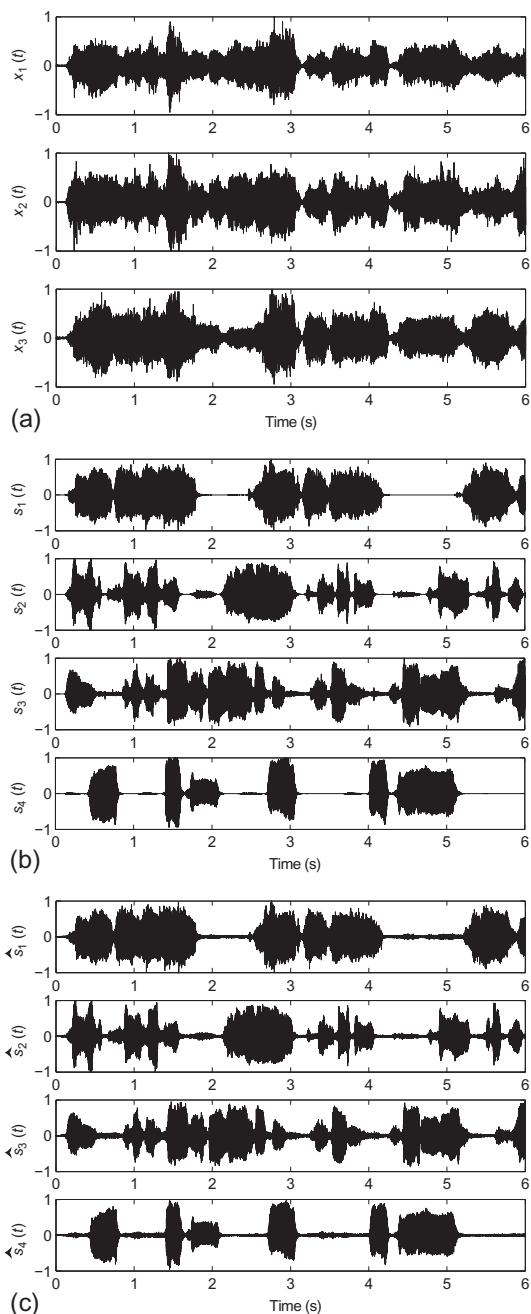
This section describes the implementation of STFDs-based source separation algorithms. Theoretical aspects of these algorithms are presented in Sections 8.2 and 8.4. The MATLABTM codes are explained and provided in the next sections for each stage of the algorithm.

8.7.1 IMPLEMENTATION OF THE STFD

This section considers the implementation of the methods described in Sections 8.2.1 and 8.4.1.2. Given an analytic signal vector $\mathbf{z}(t)$, the spatial instantaneous autocorrelation function (SIAF), which is the generalization of the instantaneous autocorrelation function (IAF) defined in Chapter 2, is given by

$$\mathbf{K}_{\mathbf{z}, \mathbf{z}}(t, \tau) = \mathbf{z}\left(t + \frac{\tau}{2}\right) \mathbf{z}^*\left(t - \frac{\tau}{2}\right). \quad (8.7.1)$$

⁰Author: A. Belouchrani, Ecole Nationale Polytechnique, Algiers, Algeria (adel.belouchrani@enp.edu.dz). Reviewers: K. Abed-Meraim and H. Bousbia-Salah.

**FIGURE 8.6.2**

Blind audio source separation of four sources from three mixtures: (a) mixture signals, (b) original source signals, and (c) estimated source signals.

Following the approach in Chapters 2 and 3, we define also the smoothed SIAF as

$$\mathbf{R}_{\mathbf{z},\mathbf{z}}(t, \tau) = G(t, \tau) * \int_t \mathbf{K}_{\mathbf{z},\mathbf{z}}(t, \tau), \quad (8.7.2)$$

where $G(t, \tau)$ is the time-lag kernel. The time convolution operator $*_t$ is applied to each entry of the matrix $\mathbf{K}_{\mathbf{z},\mathbf{z}}(t, \tau)$. The class of quadratic STFDs is then defined as

$$\rho_{\mathbf{z},\mathbf{z}}(t, f) = \mathcal{F}_{\tau \rightarrow f} \{\mathbf{R}_{\mathbf{z},\mathbf{z}}(t, \tau)\}, \quad (8.7.3)$$

where the FT \mathcal{F} is applied to each entry of the matrix $\mathbf{R}_{\mathbf{z},\mathbf{z}}(t, \tau)$.

Following Chapters 2, 3, and 6 and [46] the discrete-time definition equivalent to Eqs. (8.7.2) and (8.7.3), which leads to the simple implementation of an STFD, is

$$\rho_{\mathbf{z},\mathbf{z}}(n, k) = \text{DFT}_{m \rightarrow k} \left\{ G(n, m) * \int_n \mathbf{K}_{\mathbf{z},\mathbf{z}}(n, m) \right\}, \quad (8.7.4)$$

which can be expanded as

$$\rho_{\mathbf{z},\mathbf{z}}(n, k) = \sum_{m=-M}^M \sum_{p=-M}^M G(n-p, m) \mathbf{z}(p+m) \mathbf{z}^*(p-m) e^{-j4\pi \frac{mk}{N}}, \quad (8.7.5)$$

where the discrete Fourier transform (DFT) and the discrete time convolution operator $*_n$ are applied to each entry of the matrix $G(n, m) *_n \mathbf{K}_{\mathbf{z},\mathbf{z}}(n, m)$ and matrix $\mathbf{K}_{\mathbf{z},\mathbf{z}}(n, m)$, respectively. $N = 2M + 1$ is the signal length. Note that the STFD of a signal vector is a matrix whose diagonal entries are the classical auto-TFDs of the vector components and the off-diagonal entries are the cross-TFDs.

A more general definition of the STFD can be given as

$$\rho_{\mathbf{z},\mathbf{z}}(n, k) = \sum_{m=-M}^M \sum_{p=-M}^M \mathbf{G}(n-p, m) \odot \mathbf{z}(p+m) \mathbf{z}^*(p-m) e^{-j4\pi \frac{mk}{N}}, \quad (8.7.6)$$

where \odot denotes the Hadamard product, and $[\mathbf{G}(n, m)]_{ij} = G_{ij}(n, m)$ is the time-lag kernel associated with the pair of the sensor signals $z_i(n)$ and $z_j(n)$.

The following MATLAB code calculates half of the STFD of a vector signal $\mathbf{z}(n)$ (the other half is obtained by Hermitian symmetry)

```
% MATLAB code 1
% Initialization
D = zeros(fft_length, sample_size, sensor_number, sensor_number);
% STFD computation
for k = 1 : K; % K: sensor number
    for l = k : K;
        D(:,:,k,l) = Cross-TFD(z(k,:), z(l,:), window_length, fft_length);
    end
end
```

where the function

```
Cross-TFD(z(k,:), z(l,:), window_length, fft_length)
```

calculates the cross-TFD of the signals $z_k(n)$ and $z_l(n)$. (For the implementation of such functions, see Chapter 6 of this book and Ref. [46].)

8.7.1.1 Structure under linear model

Consider the following linear model of the multisensor signal $\mathbf{z}(n)$, as given in Eqs. (8.2.2), (8.4.2), and (8.4.15)

$$\mathbf{z}(n) = \mathbf{A}\mathbf{s}(n), \quad (8.7.7)$$

where \mathbf{A} is a $K \times L$ matrix ($K \geq L$) and $\mathbf{s}(n)$ is an $L \times 1$ vector which is referred to as the source signal vector, following the approaches in Sections 8.2.1 and 8.4.1.

Under this linear model, the STFDs take the following structure:

$$\rho_{\mathbf{zz}}(n, k) = \mathbf{A}\rho_{\mathbf{ss}}(n, k)\mathbf{A}^H, \quad (8.7.8)$$

where $\rho_{\mathbf{ss}}(n, k)$ is the source STFD of vector $\mathbf{s}(n)$ whose entries are the auto- and cross-TFDs of the source signals.

The auto STFD denoted by $\rho_{\mathbf{zz}}^a(n, k)$ is the STFD $\rho_{\mathbf{zz}}(n, k)$ evaluated at auto-term (t, f) points only. Correspondingly, the cross STFD $\rho_{\mathbf{zz}}^c(n, k)$ is the STFD $\rho_{\mathbf{zz}}(n, k)$ evaluated at cross-term (t, f) points.

Note that the diagonal (off-diagonal) elements of $\rho_{\mathbf{ss}}(n, k)$ are auto-terms (cross-terms). Thus, the auto (cross) STFD $\rho_{\mathbf{ss}}^a(n, k)$ ($\rho_{\mathbf{ss}}^c(n, k)$) is diagonal (off-diagonal) for each (t, f) point that corresponds to a source auto-term (cross-term), provided the window effect is neglected.

8.7.2 IMPLEMENTATION OF STFDs-BASED SOURCE SEPARATION

The multisensor signal $\mathbf{z}(n)$ is assumed to be nonstationary and to satisfy the linear model (8.7.7), following the approaches in Sections 8.2.3 and 8.4.1.3.

The problem of source separation consists of identifying the matrix \mathbf{A} and/or recovering the source signals $\mathbf{s}(n)$ up to a fixed permutation and some complex factors (see Section 8.4 for more details).

Denote by \mathbf{W} the $L \times K$ whitening matrix, such that

$$(\mathbf{W}\mathbf{A})(\mathbf{W}\mathbf{A})^H = \mathbf{U}\mathbf{U}^H = \mathbf{I}. \quad (8.7.9)$$

Pre- and postmultiplying the STFD $\rho_{\mathbf{zz}}(n, k)$ by \mathbf{W} leads to the whitened STFD, defined as

$$\underline{\rho}_{\mathbf{zz}}(n, k) = \mathbf{W}\rho_{\mathbf{zz}}(n, k)\mathbf{W}^H = \mathbf{U}\rho_{\mathbf{ss}}(n, k)\mathbf{U}^H, \quad (8.7.10)$$

where the second equality stems from the definition of \mathbf{W} and Eq. (8.7.8).

By selecting auto-term (t, f) points, the whitened auto STFD will have the following structure:

$$\underline{\rho}_{\mathbf{zz}}^a(n, k) = \mathbf{U}\rho_{\mathbf{ss}}^a(n, k)\mathbf{U}^H, \quad (8.7.11)$$

where $\rho_{\mathbf{ss}}^a(n, k)$ is diagonal. The missing unitary matrix \mathbf{U} is retrieved (up to permutation and phase shifts) by joint diagonalization (JD) of a combined set $\{\underline{\rho}_{\mathbf{zz}}^a(n_i, k_i) | i = 1, \dots, P\}$ of P auto STFDs. The incorporation of several auto-term (t, f) points in the JD reduces the likelihood of having degenerate eigenvalues and increases robustness to a possible additive noise.

The selection of cross-term (t, f) points leads the following whitened cross STFD:

$$\underline{\rho}_{\mathbf{zz}}^c(n, k) = \mathbf{U}\rho_{\mathbf{ss}}^c(n, k)\mathbf{U}^H, \quad (8.7.12)$$

where $\rho_{ss}^c(n, k)$ is off-diagonal. The unitary matrix \mathbf{U} is found up to permutation and phase shifts by joint off diagonalization (JOD) of a combined set $\{\underline{\rho}_{zz}^c(n_i, k_i) | i = 1, \dots, Q\}$ of Q auto STFDs.

The unitary matrix \mathbf{U} can also be found up to permutation and phase shifts by a combined JD/JOD of the two sets $\{\rho_{zz}^a(n_i, k_i) | i = 1, \dots, P\}$ and $\{\rho_{zz}^c(n_i, k_i) | i = 1, \dots, Q\}$. Once the unitary matrix \mathbf{U} is obtained from either the JD or the JOD or the combined JD/JOD, an estimate of the mixing matrix \mathbf{A} can be computed by the product $\mathbf{W}^\# \mathbf{U}$, where $\#$ denotes the pseudo-inverse operator. An estimate of the source signals $\mathbf{s}(n)$ is then obtained by the product $\mathbf{A}^\# \mathbf{z}(n)$.

8.7.3 IMPLEMENTATION OF THE WHITENING

The implementation of the whitening goes as follows:

- Estimate the sample covariance \mathbf{R} from T data samples.
- Denote by $\lambda_1, \dots, \lambda_L$ the L largest eigenvalues, and by $\mathbf{h}_1, \dots, \mathbf{h}_L$ the corresponding eigenvectors of \mathbf{R} .
- Under the white noise assumption, an estimate of the noise variance σ^2 is the average of the $K - L$ smallest eigenvalues of \mathbf{R} .
- The whitening matrix is computed by

$$\mathbf{W} = \left[(\lambda_1 - \sigma^2)^{-1/2} \mathbf{h}_1, \dots, (\lambda_L - \sigma^2)^{-1/2} \mathbf{h}_L \right]^T. \quad (8.7.13)$$

- The whitened signals are then computed by $\underline{\mathbf{z}}(n) = \mathbf{W}\mathbf{z}(n)$.

Note that in the last step, the dimension of the problem is reduced from K to L . The above procedure is valid in the case of more sensors than sources, that is, $K > L$. When $K = L$, no noise is assumed and the whitening matrix is computed as the matrix square root of the sample covariance matrix \mathbf{R} .

The following MATLAB code gives the implementation of the whitening:

```
% MATLAB code 2
% Computation of the covariance matrix
R = z*z'/sample_size;
if L < K, % Assumes white noise
    % K: sensor number, L: source number
% Compute the eigen decomposition of R
[U,d] = eig(R);
[power,k]= sort(diag(real(d)));
sigma = mean(power(1:K-L));
w1 = ones(L,1)./(sqrt(power(K-L+1:K))-sigma);
% Computation of the whitening matrix
W = diag(w1)*U(1:K,k(K-L+1:K))';
else % Assumes no noise
    W = inv(sqrtm(R));
end;
% Whitening the signals
z = W*z;
```

8.7.4 SELECTION OF AUTO-TERMS AND CROSS-TERMS

As discussed on p. 486, the selection procedure of the auto-terms and cross-terms exploits the off-diagonal structure of the source cross STFD matrices. More precisely, for a source cross STFD, we have

$$\text{Trace} \left(\underline{\rho}_{\mathbf{zz}}^c(n, k) \right) = \text{Trace} \left(\mathbf{U} \underline{\rho}_{\mathbf{ss}}^c(n, k) \mathbf{U}^H \right) = \text{Trace} \left(\underline{\rho}_{\mathbf{ss}}^c(n, k) \right) \approx 0. \quad (8.7.14)$$

Based on this observation, the following testing procedure applies:

if $\text{Trace} \left(\underline{\rho}_{\mathbf{zz}}(n, k) \right) / \text{norm} \left(\underline{\rho}_{\mathbf{zz}}(n, k) \right) < \varepsilon$ then decide that (n, k) is a cross-term
 if $\text{Trace} \left(\underline{\rho}_{\mathbf{zz}}(n, k) \right) / \text{norm} \left(\underline{\rho}_{\mathbf{zz}}(n, k) \right) > \varepsilon$ then decide that (n, k) is an auto-term,

where ε is a “small” positive real scalar.

The following MATLAB code implements the selection procedure:

```
% MATLAB code 3
% Initialization
M = []; % will contain the auto STFDs
N = []; % will contain the cross STFDs
mp = 2*L; % number of selected frequencies at a given time sample
           % where L is the source number
esp = 0.1; % threshold
% To select only (t-f) points with significant energy
[Dsort,Ii] = sort(D); % D is the STFD computed at MATLAB code 1
for tp=1:skip:sample_size % skip is some chosen step
    for k = 1:mp
        fp = Ii(fft_length-k,tp);
    % Selection of the STFDs associated to the (t-f) point tp,fp
        Ds = squeeze(D(fp,tp,:,:));
    % Form the other half of Ds by Hermitian symmetry
        Ds = Ds+ (Ds-diag(diag(Ds)))';
        if (trace(Ds)/norm(Ds))> esp, % Selection criterion
            M = [M Ds]; % Selection of auto STFDs
        else
            N = [N Ds]; % Selection of cross STFDs
        end;
    end;
end;
% Note that the matrices M and N are the concatenation matrices of
% the auto-STFD matrices and cross-STFD matrices, respectively.
```

8.7.5 IMPLEMENTATION OF JD AND JOD

As discussed in Sections 8.4.1.3 and 8.4.2.2, the JD [47] of a set $\{\mathbf{M}_k | k = 1, \dots, P\}$ of P matrices of dimensions $L \times L$ is defined as the maximization of the JD criterion

$$C_{\text{JD}}(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^P \sum_{i=1}^L |\mathbf{v}_i^* \mathbf{M}_k \mathbf{v}_i|^2 \quad (8.7.15)$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_L]$.

The JOD [3] of a set $\{\mathbf{N}_k | k = 1, \dots, Q\}$ of Q matrices of dimensions $L \times L$ is defined as the maximization of the JOD criterion

$$C_{\text{JOD}}(\mathbf{V}) \stackrel{\text{def}}{=} - \sum_{k=1}^P \sum_{i=1}^L |\mathbf{v}_i^* \mathbf{N}_k \mathbf{v}_i|^2 \quad (8.7.16)$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_L]$.

The combined JD and JOD [3] of two sets $\{\mathbf{M}_k | k = 1, \dots, P\}$ and $\{\mathbf{N}_k | k = 1, \dots, Q\}$ of $L \times L$ matrices is defined as the maximization, over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_L]$, of the JD/JOD criterion

$$C_{\text{JD/JOD}}(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{i=1}^L \left(\sum_{k=1}^P |\mathbf{v}_i^* \mathbf{M}_k \mathbf{v}_i|^2 - \sum_{k=1}^Q |\mathbf{v}_i^* \mathbf{N}_k \mathbf{v}_i|^2 \right). \quad (8.7.17)$$

Maximization of Eq. (8.7.17) (which is Eq. (8.4.13) in Section 8.4.1.3) is performed by successive Givens rotations as follows: a unitary matrix $\mathbf{V} = [v_{ij}]$ is sought such that the criterion in Eq. (8.7.17) is maximized w.r.t. the matrices $\mathbf{M}_k = [m_{ij}^k]$, $k = 1, \dots, P$, and $\mathbf{N}_k = [n_{ij}^k]$, $k = 1, \dots, Q$. At each Givens rotation step, a pivot pair r, s is chosen. With δ_{ij} denoting the Kronecker delta operator, v_{ij} is set to δ_{ij} , except for the elements $v_{rr} = v_{ss} = \cos(\theta)$, $v_{rs} = e^{j\phi} \sin(\theta)$ and $v_{sr} = -e^{-j\phi} \sin(\theta)$. One can show [47] that optimization of Eq. (8.7.17) is equivalent to the maximization of

$$\mathcal{Q} = \mathbf{v}^T \text{Re}(\mathbf{G}) \mathbf{v}, \quad (8.7.18)$$

where

$$\mathbf{G} = \sum_{k=1}^P \mathbf{g}_{1,k} \mathbf{g}_{1,k}^T - \sum_{k=1}^Q \mathbf{g}_{2,k} \mathbf{g}_{2,k}^T \quad (8.7.19)$$

$$\mathbf{v}^T = [\cos 2\theta, -\sin 2\theta \cos \phi, -\sin 2\theta \sin \phi] \quad (8.7.20)$$

$$\mathbf{g}_{1,k}^T = [m_{rr}^k - m_{ss}^k, m_{rs}^k + m_{sr}^k, j(m_{sr}^k - m_{rs}^k)] \quad (8.7.21)$$

$$\mathbf{g}_{2,k}^T = [n_{rr}^k - n_{ss}^k, n_{rs}^k + n_{sr}^k, j(n_{sr}^k - n_{rs}^k)]. \quad (8.7.22)$$

The next step is to recognize that the particular parameterization of \mathbf{v} is equivalent to the condition $\mathbf{v}^T \mathbf{v} = 1$. Maximization of a quadratic form under the unit norm constraint of its argument is obtained by taking \mathbf{v} to be the eigenvector of $\text{Re}(\mathbf{G})$ associated with the largest eigenvalue.

The reader may check that if we set $P = 1$ and $Q = 0$,¹ the above boils down to the standard Jacobi diagonalization procedure [48]. Also note that the main computational cost in this kind of technique is the update under Givens rotations of the various matrices involved in the diagonalization. This makes the computational cost of the JD/JOD procedure similar to $P + Q$ times the diagonalization of a single matrix.

The following MATLAB code gives the implementation of the JD/JOD:

```
% MATLAB code 4
% Initialization
L      = min(size(M)); % M: auto STFDs given at MATLAB code 3
nm1    = max(size(M));
nm2    = max(size(N)); % N: cross STFDs given at MATLAB code 3
V      = eye(L);        % L is the source number
threshold = 1/sqrt(sample_size)/100;
more   = 1;
while more, more=0;
for p=1:L-1,
  for q=p+1:L,
% Givens rotations
  g1=[M(p,p:L:nl)-M(q,q:L:nl);M(p,q:L:nl)+M(q,p:L:nl);
       i*(M(q,p:L:nl)-M(p,q:L:nl))];
  g2=[M(p,p:L:nl)-M(q,q:L:nl);M(p,q:L:nl)+M(q,p:L:nl);
       i*(M(q,p:L:nl)-M(p,q:L:nl))];
  [vcp,d] = eig(real(g1*g1'-g2*g2'));
  [la,Ki]=sort(diag(d));
  angles=vcp(:,Ki(1));angles=sign(angles(1))*angles;
  c=sqrt(0.5+angles(1)/2);
  sr=0.5*(angles(2)-j*angles(3))/c; sc=conj(sr);
  yes = abs(sr) > threshold;
  more=more | yes ;
  if yes, % Update of the M, N and V matrices
    colp1=M(:,p:L:nl);colq1=M(:,q:L:nl);
    M(:,p:L:nl)=c*colp1+sr*colq1;M(:,q:L:nl)=c*colq1-sc*colp1;
    rowp1=M(p,:);rowq1=M(q,:);
    M(p,:)=c*rowp1+sc*rowq1;M(q,:)=c*rowq1-sr*rowp1;
    colp2=N(:,p:L:nl);colq2=N(:,q:L:nl);
    N(:,p:L:nl)=c*colp2+sr*colq2;N(:,q:L:nl)=c*colq2-sc*colp2;
    rowp2=N(p,:);rowq2=N(q,:);
    N(p,:)=c*rowp2+sc*rowq2;N(q,:)=c*rowq2-sr*rowp2;
    temp=V(:,p);
    V(:,p)=c*V(:,p)+sr*V(:,q);V(:,q)=c*V(:,q)-sc*temp;
  end; % if
  end; % q loop
end; % p loop
end; % while
```

¹That is, only matrix \mathbf{M}_1 is incorporated in the criterion of Eq. (8.7.17).

8.7.6 SUMMARY AND CONCLUSIONS

The detailed implementation of nonstationary source separation algorithms based on STFDs has been presented together with the associated MATLAB codes. The implementation details included whitening, selection of auto-term and cross-term regions, and the combined JD and JOD.

REFERENCES

- [1] A. Belouchrani, K. Abed-Meraim, B. Boashash, “Time frequency and array processing of non-stationary signals”, *EURASIP J. Adv. Signal Process.* 2012 (230) (2012) 6, <http://dx.doi.org/10.1186/1687-6180-2012-230>.
- [2] J. Lerga, V. Sucic, B. Boashash, “An efficient algorithm for instantaneous frequency estimation of nonstationary multicomponent signals in low SNR”, *EURASIP J. Adv. Signal Process.* 2011 (2011) <http://dx.doi.org/10.1155/2011/725189>, article ID 725189, 16 pp.
- [3] A. Belouchrani, M.G. Amin, N. Thirion-Moreau, Y.D. Zhang, “Source separation and localization using time-frequency distributions: an overview”, *IEEE Signal Process. Mag.* 30 (6) (2013) 97-107.
- [4] H.L.L. Van Trees, Detection, Estimation, and Modulation Theory, vol. IV: Optimum Array Processing, Wiley Interscience, New York, 2002.
- [5] M.A. Latif, S. Sanei, J. Chambers, L. Shoker, “Localization of abnormal EEG sources using blind source separation partially constrained by the locations of known sources”, *IEEE Signal Process. Lett.* 13 (3) (2006) 117-120, ISSN 1070-9908.
- [6] R. Grech, T. Cassar, J. Muscat, K.P. Camilleri, S.G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, B. Vanrumste, “Review on solving the inverse problem in EEG source analysis”, *J. Neuroeng. Rehabil.* 5 (1) (2008) 25.
- [7] K. Sekihara, S. Nagarajan, D. Poeppel, Y. Miyashita, “Time-frequency MEG-MUSIC algorithm”, *IEEE Trans. Med. Imaging* 18 (1) (1999) 92-97, ISSN 0278-0062, <http://dx.doi.org/10.1109/42.750262>.
- [8] A. Belouchrani, M.G. Amin, “Blind source separation based on time-frequency signal representation”, *IEEE Trans. Signal Process.* 46 (11) (1998) 2888-2897.
- [9] A. Belouchrani, M.G. Amin, “Time-frequency MUSIC: a new array signal processing method based on time-frequency signal representation”, *IEEE Signal Process. Lett.* 6 (5) (1999) 109-110.
- [10] Y. Zhang, W. Mu, M.G. Amin, “Time-frequency maximum likelihood methods for direction finding”, *J. Franklin Instit.* 337 (4) (2000) 483-497.
- [11] M.G. Amin, A. Belouchrani, Y. Zhang, “The spatial ambiguity function and its applications”, *IEEE Signal Process. Lett.* 7 (6) (2000) 138-140.
- [12] Y. Zhang, W. Mu, M.G. Amin, “Subspace analysis of spatial time-frequency distribution matrices”, *IEEE Trans. Signal Process.* 49 (4) (2001) 747-759.
- [13] M.G. Amin, Y. Zhang, “Direction finding based on spatial time-frequency distribution matrices”, *Dig. Signal Process.* 10 (4) (2000) 325-339.
- [14] Y. Zhang, M.G. Amin, “Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays”, *IEEE Trans. Signal Process.* 48 (10) (2000) 2892-2902.
- [15] A.B. Gershman, M.G. Amin, “Wideband direction-of-arrival estimation of multiple chirp signals using spatial time-frequency distributions”, *IEEE Signal Process. Lett.* 7 (6) (2000) 152-155.
- [16] A.M. Rao, D.L. Jones, “Quadratic detection in arrays using TFDs”, in: B. Boashash (Ed.), Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, first ed., Article 8.3, Elsevier, Oxford, 2003, pp. 344-348, Chapter 8.
- [17] A.M. Sayeed, D.J. Jones, “Optimal detection using bilinear time-frequency and time-scale representations”, *IEEE Trans. Signal Process.* 43 (12) (1995) 2872-2883.

- [18] A.M. Rao, D.L. Jones, "Nonstationary array signal detection using time-frequency and time-scale representations", in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98), vol. 4, 1998, pp. 1989-1992.
- [19] D.R. Morgan, T.M. Smith, "Coherence effects on the detection performance of quadratic array processors with application to large-array matched-field beamforming", *J. Acoust. Soc. Am.* 87 (2) (1990) 737-747.
- [20] A.M. Rao, D.L. Jones, "A denoising approach to multisensor signal estimation", *IEEE Trans. Signal Process.* 48 (5) (2000) 1225-1234.
- [21] A.M. Rao, D.L. Jones, "Efficient detection with arrays in the presence of angular spreading", *IEEE Trans. Signal Process.* 51 (2) (2003) 301-312, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2002.806992>.
- [22] R.B. Ertel, P. Cardieri, K.W. Sowerby, T.S. Rappaport, J.H. Reed, "Overview of spatial channel models for antenna array communication systems", *IEEE Pers. Commun.* 5 (1) (1998) 10-22.
- [23] B. Boashash, M. Khelif, T. Ben-Jabeur, C. East, P. Colditz, "Passive detection of accelerometer-recorded fetal movements using a time-frequency signal processing approach", *Dig. Signal Process.* 25 (2014) 134-155, ISSN 1051-2004, <http://dx.doi.org/10.1016/j.dsp.2013.10.002>, URL <http://www.sciencedirect.com/science/article/pii/S1051200413002212>.
- [24] L. Parra, C. Spence, "Convulsive blind separation of non-stationary sources", *IEEE Trans. Speech Audio Process.* 8 (3) (2000) 320-327.
- [25] G. Chabriel, M. Kleinstuber, E. Moreau, H. Shen, P. Tichavsky, A. Yeredor, "Joint matrices decompositions and blind source separation: a survey of methods, identification, and applications", *IEEE Signal Process. Mag.* 31 (3) (2014) 34-43.
- [26] A.R. Leyman, Z.M. Kamran, K. Abed-Meraim, "Higher-order time frequency-based blind source separation technique", *IEEE Signal Process. Lett.* 7 (7) (2000) 193-196.
- [27] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, B. Boashash, "Separating more sources than sensors using time-frequency distributions", *EURASIP J. Appl. Signal Process.* 2005 (17) (2005) 2828-2847.
- [28] A. Belouchrani, M.G. Amin, K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices", *IEEE Signal Process. Lett.* 4 (9) (1997) 266-268.
- [29] H. Bousbia-Saleh, A. Belouchrani, K. Abed-Meraim, "Jacobi-like algorithm for blind signal separation of convulsive mixtures", *Electron. Lett.* 37 (16) (2001) 1049-1050.
- [30] K. Abed-Meraim, W. Qiu, Y. Hua, "Blind system identification", *Proc. IEEE* 85 (8) (1997) 1310-1322.
- [31] S. El-Khamy, S.E. Shaaban, E.A. Thabet, "Frequency-hopped multi-user chirp modulation (FH/M-CM) for multipath fading channels", in: Proc. Sixteenth National Radio Science Conference (NRSC'99), 1999, pp. C6/1-8.
- [32] C. Gupta, A. Papandreou-Suppappola, "Wireless CDMA communications using time-varying signals", in: Proc. Sixth Internat. Symp. on Signal Processing and Its Applications (ISSPA'01), vol. 1, 2001, pp. 242-245.
- [33] A. Persson, T. Ottosson, E. Strom, "Time-frequency localized CDMA for downlink multi-carrier systems", in: Proc. Seventh Internat. Symp. on Spread Spectrum Techniques and Applications (ISSSTA-02), vol. 1, 2002, pp. 118-122.
- [34] A. Gersho, R.M. Gray, Vector Quantization and Signal Compression, Kluwer, Norwell, MA, 1991.
- [35] G.F. Boudreux-Bartels, T.W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 34 (3) (1986) 442-451.
- [36] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, E. Pechersky, "Detection of linear features in SAR images: application to road network extraction", *IEEE Trans. Geosci. Remote Sens.* 36 (2) (1998) 434-453.
- [37] P. Bofill, M. Zibulevsky, "Underdetermined blind source separation using sparse representations", *Signal Process.* 81 (11) (2001) 2353-2362.
- [38] F.S. Cohen, S. Kadamb, G.F. Boudreux-Bartels, "Tracking of unknown nonstationary chirp signals using unsupervised clustering in the Wigner distribution space", *IEEE Trans. Signal Process.* 41 (11) (1993) 3085-3101.

- [39] A. Aissa-El-Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, Y. Grenier, “Underdetermined blind separation of nondisjoint sources in the time-frequency domain”, *IEEE Trans. Signal Process.* 55 (3) (2007) 897-907.
- [40] S. Xie, L. Yang, J.-M. Yang, G. Zhou, Y. Xiang, “Time-frequency approach to underdetermined blind source separation”, *IEEE Trans. Neural Netw. Learn. Syst.* 23 (2) (2012) 306-316.
- [41] C. Blandin, A. Ozerov, E. Vincent, “Multi-source TDOA estimation in reverberant audio using angular spectra and clustering”, *Signal Process.* 92 (8) (2012) 1950-1960, ISSN 0165-1684, <http://dx.doi.org/10.1016/j.sigpro.2011.09.032>, URL <http://www.sciencedirect.com/science/article/pii/S0165168411003525>.
- [42] S. Araki, H. Sawada, R. Mukai, S. Makino, “DOA estimation for multiple sparse sources with arbitrarily arranged multiple sensors”, *J. Signal Process. Syst.* 63 (3) (2011) 265-275, ISSN 1939-8018, <http://dx.doi.org/10.1007/s11265-009-0413-9>.
- [43] A. Aissa-El-Bey, K. Abed-Meraim, Y. Grenier, “Blind separation of underdetermined convolutive mixtures using their time-frequency representation”, *IEEE Trans. Audio Speech Lang. Process.* 15 (5) (2007) 1540-1550, ISSN 1558-7916, <http://dx.doi.org/10.1109/TASL.2007.898455>.
- [44] V. Reju, S.N. Koh, I. Soon, “Underdetermined convolutive blind source separation via time-frequency masking”, *IEEE Trans. Audio Speech Lang. Process.* 18 (1) (2010) 101-116, ISSN 1558-7916, <http://dx.doi.org/10.1109/TASL.2009.2024380>.
- [45] M. Ikram, D. Morgan, “Permutation inconsistency in blind speech separation: investigation and solutions”, *IEEE Trans. Speech Audio Process.* 13 (1) (2005) 1-13, ISSN 1063-6676, <http://dx.doi.org/10.1109/TSA.2004.834441>.
- [46] B. Boashash, A.P. Reilly, “Algorithms for time-frequency signal analysis”, in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 163-181, Chapter 7.
- [47] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, “Blind source separation using second order statistics”, *IEEE Trans. Signal Process.* 45 (2) (1997) 434-444.
- [48] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1989.

PART

TF STATISTICAL
TECHNIQUES

IV

This page intentionally left blank

NOISE ANALYSIS AND RANDOM PROCESSES IN THE (t, f) DOMAIN

9

INTRODUCTION AND OVERVIEW

Time-Frequency Distributions (TFDs) have been studied initially for deterministic signals. Noise is an ever permanent feature in signals and so there is a need to extend the formulation of TFDs to the case of random signals. This chapter describes time-frequency (t, f) methodologies developed for dealing with randomness in nonstationary signals and systems. The topic is covered in six sections with appropriate cross-referencing.

Noise analysis in the (t, f) domain is presented with derivations of mean and variance of TFDs of random signals (Section 9.1). Both cases of additive and multiplicative noise, including white and colored noise, are considered. Time-varying random processes and systems can be represented with dispersive time-frequency characteristics using the Weyl symbol (Section 9.2). This approach allows the adaptation of analysis tools to systems or processes that have specific (t, f) characteristics. TFDs such as the spectrogram and WVD are designed that are robust to impulse noise (Section 9.3). Time-varying power spectra can be defined based on generalizations of the Wigner-Ville spectrum and evolutionary spectrum. These are particularly suitable to represent underspread nonstationary processes (i.e., processes with small (t, f) correlations) (Section 9.4). Time-varying random channels are also described using a (t, f) approach (Section 9.5). The last section presents methods for spectral analysis of cyclostationary processes, i.e., processes whose statistical properties vary cyclically with time (Section 9.6).

9.1 ANALYSIS OF NOISE IN TIME-FREQUENCY DISTRIBUTIONS⁰

This section presents noise analysis for commonly used time-frequency distributions such as those presented in Chapters 2 and 3. The Wigner distribution, as a basic time-frequency representation, is studied first. The bias and variance in the case of complex white noise are derived. The analysis of noise is extended to other quadratic distributions, and to different types of additive and multiplicative noise, including stationary white noise, nonstationary white noise, and colored stationary noise. Exact expressions for the mean value and the variance of quadratic distributions for each point in the time-frequency plane are given.

9.1.1 WIGNER DISTRIBUTION

The window Wigner distribution (WD) of the discrete-time noisy signal $x(n) = s(n) + \epsilon(n)$ is defined by

$$W_x(n, f) = \sum_m w(m) w(-m) x(n+m) x^*(n-m) e^{-j4\pi fm}, \quad (9.1.1)$$

where $w(m)$ is a real-valued lag window such that $w(0) = 1$. In this section, the constant factor 2 is omitted from TFD definitions, including the above, and the notation \sum_m^∞ means $\sum_{m=-\infty}^\infty$.

Consider first the case when $s(n)$ is deterministic and the noise $\epsilon(n)$ is a white, Gaussian, complex, stationary, zero-mean process, with independent real and imaginary parts having equal variances. Its autocorrelation function is $R_{\epsilon\epsilon}(m) = \sigma_\epsilon^2 \delta(m)$. The WD mean for the noisy signal $x(n)$ is

$$\begin{aligned} E\{W_x(n, f)\} &= \sum_m w(m) w(-m) s(n+m) s^*(n-m) e^{-j4\pi fm} + \sum_m w(m) w(-m) R_{\epsilon\epsilon}(2m) e^{-j4\pi fm} \\ &= 2 \int_{-1/4}^{1/4} W_s(n, f - \alpha) F_w(2\alpha) d\alpha + \sigma_\epsilon^2, \end{aligned} \quad (9.1.2)$$

where $F_w(f) = \mathcal{F}_{m \rightarrow f}[w(m) w(-m)]$ is the Fourier transform (FT) of the product $w(m) w(-m)$, and $W_s(n, f)$ is the original WD of $s(n)$, without a lag window.

The lag window $w(m)$ causes the **WD bias**. The second term on the right-hand side in Eq. (9.1.2) is constant, so one can assume that it does not distort the WD. Expanding $W_s(n, f - \alpha)$ into a Taylor series, around f , we get

$$2 \int_{-1/4}^{1/4} W_s(n, f - \alpha) F_w(2\alpha) d\alpha \cong W_s(n, f) + \frac{1}{8} \frac{\partial^2 W_s(n, f)}{\partial f^2} m_2 + \dots \quad (9.1.3)$$

Thus, the bias can be approximated by

$$\text{bias}(n, f) \cong \frac{1}{8} \frac{\partial^2 W_s(n, f)}{\partial f^2} m_2 = \frac{1}{8} b(n, f) m_2,$$

where $m_2 = \int_{-1/2}^{1/2} f^2 F_w(f) df$. For the regions where the WD variations in the frequency direction are small, the bias is small, and vice versa.

⁰Author: **LJ. Stanković**, University of Montenegro, Podgorica, Montenegro (ljubisa@ac.me). Reviewers: M.G. Amin and M.J. Bastiaans.

The WD estimator **variance**, at a given point (n, f) , is defined by:

$$\sigma_{\text{WD}}^2(n, f) = E\{W_x(n, f) W_x^*(n, f)\} - E\{W_x(n, f)\} E\{W_x^*(n, f)\}. \quad (9.1.4)$$

For signals $x(n) = s(n) + \epsilon(n)$, it results in

$$\begin{aligned} \sigma_{\text{WD}}^2(n, f) &= \sum_{m_1} \sum_{m_2} \left\{ w(m_1) w(-m_1) w(m_2) w(-m_2) e^{-j4\pi f(m_1 - m_2)} \right. \\ &\quad \times [s(n + m_1) s^*(n + m_2) R_{\epsilon\epsilon}(n - m_2, n - m_1) \\ &\quad + s^*(n - m_1) s(n - m_2) R_{\epsilon\epsilon}(n + m_1, n + m_2) \\ &\quad + s(n + m_1) s(n - m_2) R_{\epsilon\epsilon^*}^*(n - m_1, n + m_2) \\ &\quad + s^*(n - m_1) s^*(n + m_2) R_{\epsilon\epsilon^*}(n + m_1, n - m_2) \\ &\quad + R_{\epsilon\epsilon}(n + m_1, n + m_2) R_{\epsilon\epsilon}(n - m_2, n - m_1) \\ &\quad \left. + R_{\epsilon\epsilon^*}(n + m_1, n - m_2) R_{\epsilon\epsilon^*}^*(n - m_1, n + m_2) \right] \Big\}. \end{aligned} \quad (9.1.5)$$

The fourth-order moment of noise is reduced to the correlation functions by using the relation $E\{z_1 z_2 z_3 z_4\} = E\{z_1 z_2\}E\{z_3 z_4\} + E\{z_1 z_3\}E\{z_2 z_4\} + E\{z_1 z_4\}E\{z_2 z_3\}$, which holds for Gaussian zero-mean random variables $z_i, i = 1, 2, 3, 4$. For the considered complex noise, $R_{\epsilon\epsilon}(n, m) = \sigma_\epsilon^2 \delta(n - m)$ and $R_{\epsilon\epsilon^*}(n, m) = 0$. The variance of the WD estimator reduces to

$$\sigma_{\text{WD}}^2(n, f) = \sigma_\epsilon^2 \sum_m w^2(m) w^2(-m) \left[2|s(n + m)|^2 + \sigma_\epsilon^2 \right].$$

It is frequency-independent. For constant-modulus signals $s(n) = a \exp[j\phi(n)]$, the variance is $\sigma_{\text{WD}}^2(n, f) = \sigma_\epsilon^2 E_w(2a^2 + \sigma_\epsilon^2)$, where $E_w = \sum_m [w(m) w(-m)]^2$ is the energy of $w(m) w(-m)$ window. Thus the variance is constant. A finite-energy lag window is sufficient to make the variance of $W_x(n, f)$ finite.

The optimal lag window width can be obtained by minimizing the error $e^2 = \text{bias}^2(n, f) + \sigma_{\text{WD}}^2(n, f)$. For example, for constant-modulus signals and the Hanning window $w(m) w(-m)$ of width N , when $E_w = 3N/8$ and $m_2 = 1/(2N^2)$, we get

$$e^2 \cong \frac{1}{256 N^4} b^2(n, f) + \frac{3N}{8} \sigma_\epsilon^2 (2a^2 + \sigma_\epsilon^2).$$

From $\partial e^2 / \partial N = 0$, the approximation of optimal window width follows:

$$N_{\text{opt}}(n, f) \cong \sqrt[5]{\frac{b^2(n, f)}{24 \sigma_\epsilon^2 (2a^2 + \sigma_\epsilon^2)}}.$$

An approach to the calculation of the estimate of $N_{\text{opt}}(n, f)$, without using the value of $b^2(n, f)$, is presented in Ref. [1] and Section 10.2. Other statistical properties of the Wigner distribution are studied in Refs. [2–6].

9.1.2 NOISE IN QUADRATIC TIME-FREQUENCY DISTRIBUTIONS

A discrete-time form of the quadratic class of distributions of noise $\epsilon(n)$ is defined by:

$$\rho_\epsilon(n,f) = \sum_l \sum_m G(m,l) \epsilon(n+m+l) \epsilon^*(n+m-l) e^{-j4\pi fl}, \quad (9.1.6)$$

where $G(m,l)$ is the kernel in the time-lag domain.

Its **mean value**, for a general nonstationary noise, is

$$E\{\rho_\epsilon(n,f;G)\} = \sum_l \sum_m G(m,l) R_{\epsilon\epsilon}(n+m+l, n+m-l) e^{-j4\pi fl},$$

where $R_{\epsilon\epsilon}(m,n)$ is the noise autocorrelation function. For special cases of noise, the values of $E\{\rho_\epsilon(n,f;G)\}$ follow.

(1) Stationary white noise, $R_{\epsilon\epsilon}(m,n) = \sigma_\epsilon^2 \delta(m-n)$,

$$E\{\rho_\epsilon(n,f;G)\} = \sigma_\epsilon^2 g(0,0).$$

(2) Nonstationary white noise, $R_{\epsilon\epsilon}(m,n) = I(n) \delta(m-n)$, $I(n) \geq 0$,

$$E\{\rho_\epsilon(n,f;G)\} = \sum_m G(m,0) I(n+m).$$

(3) Stationary colored noise, $R_{\epsilon\epsilon}(m,n) = R_{\epsilon\epsilon}(m-n)$,

$$E\{\rho_\epsilon(n,f;G)\} = \int_{-1/2}^{1/2} \mathcal{G}(0, 2(f-\alpha)) S_{\epsilon\epsilon}(\alpha) d\alpha,$$

where $S_{\epsilon\epsilon}(f) = \mathcal{F}_{m \rightarrow f}[R_{\epsilon\epsilon}(m)]$ is the noise power spectrum density, and the kernel forms in time-lag, Doppler-lag, and Doppler-frequency domains are denoted by:

$$\sum_m G(m,l) e^{-j2\pi vm} = g(v,l) = \int_{-1/2}^{1/2} \mathcal{G}(v,f) e^{j2\pi fl} df. \quad (9.1.7)$$

The **variance** of $\rho_\epsilon(n,f;G)$, is defined by

$$\sigma_{\epsilon\epsilon}^2(n,f) = E\{\rho_\epsilon(n,f;G) \rho_\epsilon^*(n,f;G)\} - E\{\rho_\epsilon(n,f;G)\} E\{\rho_\epsilon^*(n,f;G)\}.$$

For Gaussian noise, as in Eqs. (9.1.4) and (9.1.5), we get:

$$\begin{aligned} \sigma_{\epsilon\epsilon}^2(n,f) &= \sum_{l_1} \sum_{l_2} \sum_{m_1} \sum_{m_2} \left\{ G(m_1, l_1) G^*(m_2, l_2) \right. \\ &\quad \times [R_{\epsilon\epsilon}(n+m_1+l_1, n+m_2+l_2) R_{\epsilon\epsilon}^*(n+m_1-l_1, n+m_2-l_2) \\ &\quad + R_{\epsilon\epsilon}^*(n+m_1+l_1, n+m_2-l_2) R_{\epsilon\epsilon}(n+m_1-l_1, n+m_2+l_2)] \\ &\quad \left. \times e^{-j4\pi f(l_1-l_2)} \right\}. \end{aligned} \quad (9.1.8)$$

Forms of $\sigma_{\epsilon\epsilon}^2(n,f)$ for the specific noises will be presented next.

9.1.2.1 Complex stationary and nonstationary white noise

For nonstationary complex white noise, with independent real and imaginary parts of equal variance, $R_{\epsilon\epsilon}(m, n) = I(n)\delta(m - n)$, $R_{\epsilon\epsilon^*}(n, m) = 0$, we get

$$\sigma_{\epsilon\epsilon}^2(n, f) = \sum_l \sum_m |G(m, l)|^2 I(n + m + l) I^*(n + m - l) = \rho_I(n, 0; |G|^2). \quad (9.1.9)$$

For stationary white noise, $I(n) = \sigma_\epsilon^2$, the variance is proportional to the kernel energy,

$$\sigma_{\epsilon\epsilon}^2(n, f) = \sigma_\epsilon^4 \sum_l \sum_m |G(m, l)|^2. \quad (9.1.10)$$

9.1.2.2 Colored stationary noise

For complex colored stationary noise, the variance (9.1.8) can be written as

$$\begin{aligned} \sigma_{\epsilon\epsilon}^2(n, f) &= \sum_{l_1} \sum_{m_1} G(m_1, l_1) \left\{ \sum_{m_2} \sum_{l_2} G^*(m_2, l_2) \right. \\ &\quad \times R_{\epsilon\epsilon}(m_1 - m_2 + l_1 - l_2) R_{\epsilon\epsilon}^*(m_1 - m_2 - (l_1 - l_2)) e^{-j4\pi f(l_1 - l_2)} \left. \right\} \end{aligned}$$

or

$$\sigma_{\epsilon\epsilon}^2(n, f) = \sum_l \sum_m G(m, l) \left\{ G(m, l) *_l *_m \left[R_{\epsilon\epsilon}^*(m + l) R_{\epsilon\epsilon}(m - l) e^{j4\pi f l} \right] \right\}^*,$$

where $*_l*_m$ denotes a two-dimensional convolution in l, m . Consider the product of $G(m, l)$ and $Y^*(m, l) = \{G(m, l) *_l *_m [R_{\epsilon\epsilon}^*(m + l) R_{\epsilon\epsilon}(m - l) e^{j4\pi f l}]\}^*$ in the last expression. Two-dimensional FTs of these terms are $\mathcal{G}(\nu, \xi)$ and $y(\nu, \xi) = \mathcal{G}(\nu, \xi) S_{\epsilon\epsilon}(f - (\xi - \nu)/2) S_{\epsilon\epsilon}^*(f - (\xi + \nu)/2)/2$. According to Parseval's theorem, we get

$$\begin{aligned} \sigma_{\epsilon\epsilon}^2(n, f) &= \frac{1}{2} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |\mathcal{G}(\nu, \xi)|^2 S_{\epsilon\epsilon}^*(f - \frac{\xi}{2} + \frac{\nu}{2}) S_{\epsilon\epsilon}(f - \frac{\xi}{2} - \frac{\nu}{2}) d\nu d\xi \\ &= \rho_{S_{\epsilon\epsilon}}(0, f; |\mathcal{G}|^2) \end{aligned} \quad (9.1.11)$$

for $|f - (\xi - \nu)/2| < 1/2$ and $|f - (\xi + \nu)/2| < 1/2$. The transforms in Eq. (9.1.11) are periodic in ν and ξ with period 1. It means that we should take into account all ν and ξ when $|f - [(\xi + k_1) - (\nu + k_2)]/2| < 1/2$ and $|f - [(\xi + k_1) + (\nu + k_2)]/2| < 1/2$, where k_1 and k_2 are integers.

Note that the FT of a colored stationary noise is a white nonstationary noise, with autocorrelation in the frequency domain

$$R_{\Xi\Xi}(f_1, f_2) = \sum_m \sum_n E\{\epsilon(m)\epsilon^*(n)\} e^{-j2\pi f_1 m + j2\pi f_2 n} = S_{\epsilon\epsilon}(f_2) \delta_p(f_1 - f_2),$$

where $\delta_p(f)$ is a periodic delta function with period 1. Thus, Eq. (9.1.11) is just a form dual to Eq. (9.1.9).

9.1.2.3 Analytic noise

In the numerical implementation of quadratic distributions, an analytic part of the signal is commonly used, rather than the signal itself. The analytic part of noise can be written as $\epsilon_a(n) = \epsilon(n) + j\epsilon_h(n)$, where $\epsilon_h(n) = \mathcal{H}[\epsilon(n)]$ is the Hilbert transform of $\epsilon(n)$. The spectral power density of $\epsilon_a(n)$, within the basic period $|f| < 1/2$, for a white noise $\epsilon(n)$, is $S_{\epsilon_a\epsilon_a}(f) = 2\sigma_\epsilon^2 U(f)$, where $U(f)$ is the unit step function. The variance follows from Eq. (9.1.11) in the form

$$\sigma_{\epsilon\epsilon}^2(n,f) = 2\sigma_\epsilon^4 \int_{-1/2}^{1/2} \int_{-d(f,\xi)}^{d(f,\xi)} |\mathcal{G}(\nu, \xi)|^2 d\xi d\nu \quad \text{for } |2f| \leq \frac{1}{2}, \quad (9.1.12)$$

where the integration limits are defined by $d(f, \xi) = |\arcsin(\sin(\pi(2f - \xi)))| / \pi$ (for details see Ref. [7]).

The kernel $\mathcal{G}(\nu, \xi)$ is mainly concentrated at and around the (ν, ξ) origin and $\xi = 0$ axis. Having this in mind, as well as the fact that $|\mathcal{G}(\nu, \xi)|^2$ is always positive, we may easily conclude that the minimal value of $\sigma_{\epsilon\epsilon}^2(n,f)$ is for $f = 0$. The maximal value will be obtained for $|f| = 1/4$. It is very close to [7]:

$$\max\{\sigma_{\epsilon\epsilon}^2(n,f)\} \cong 2\sigma_\epsilon^4 \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |\mathcal{G}(\nu, \xi)|^2 d\xi d\nu = 2\sigma_\epsilon^4 \sum_l \sum_m |G(m, l)|^2.$$

9.1.2.4 Real noise

Now consider real stationary white Gaussian noise $\epsilon(n)$ with variance σ_ϵ^2 . In this case, the variance (9.1.8) contains all terms. It can be written as:

$$\sigma_{\epsilon\epsilon}^2(n,f) = \sigma_\epsilon^4 \sum_l \sum_m \left[|G(m, l)|^2 + G(m, l) G^*(m, -l) e^{-j8\pi f l} \right]. \quad (9.1.13)$$

For distributions whose kernel is symmetric with respect to l , $G(m, l) = G(m, -l)$ holds. The FT is therefore applied to the positive and even function $|G(m, l)|^2$. The transform's maximal value is reached at $f = 0$ and $|f| = 1/4$. Accordingly:

$$\max \left\{ \sigma_{\epsilon\epsilon}^2(n,f) \right\} = 2\sigma_\epsilon^4 \sum_l \sum_m |G(m, l)|^2. \quad (9.1.14)$$

The crucial parameter in all previous cases is the kernel energy $\sum_l \sum_m |G(m, l)|^2$. Its minimization is thoroughly studied in Ref. [8]. It has been concluded that, out of all the quadratic distributions satisfying the marginal and time-support conditions, the Born-Jordan distribution is optimal with respect to this parameter.

9.1.3 NOISY SIGNALS

Analysis of deterministic signals $s(n)$ corrupted by noise, $x(n) = s(n) + \epsilon(n)$, is highly signal dependent. It can be easily shown [7,8] that the distribution variance $\sigma_\rho^2(n,f)$ consists of two components:

$$\sigma_\rho^2(n,f) = \sigma_{\epsilon\epsilon}^2(n,f) + \sigma_{s\epsilon}^2(n,f). \quad (9.1.15)$$

The first variance component, and the distribution mean value, have already been studied in detail.¹ For the analysis of the second, signal dependent, component $\sigma_{se}^2(n,f)$, we will use the inner product form of the quadratic class of distributions:

$$\rho_x(n,f; \tilde{G}) = \sum_l \sum_m \tilde{G}(m,l) \left[x(n+m) e^{-j2\pi fm} \right] \left[x(n+l) e^{-j2\pi fl} \right]^*, \quad (9.1.16)$$

where $\tilde{G}(m,l) = G((m+l)/2, (m-l)/2)$. Calculation of $\tilde{G}(m,l)$ is described in the next section. For a real and symmetric $G(m,l)$, and complex noise, we get

$$\begin{aligned} \sigma_{se}^2(n,f) &= 2 \sum_{l_1} \sum_{m_1} \sum_{l_2} \sum_{m_2} \left\{ \tilde{G}(m_1, l_1) \tilde{G}^*(m_2, l_2) s(n+m_1) s^*(n+m_2) \right. \\ &\quad \times R_{ee}(n+l_2, n+l_1) e^{-j2\pi f(m_1-l_1-m_2+l_2)} \Big\}, \end{aligned}$$

which can be written as

$$\sigma_{se}^2(n,f) = 2 \sum_{m_1} \sum_{m_2} \tilde{\Phi}(m_1, m_2) [s(n+m_1) e^{-j2\pi fm_1}] [s(n+m_2) e^{-j2\pi fm_2}]^*, \quad (9.1.17)$$

where the new kernel $\tilde{\Phi}(m_1, m_2)$ reads

$$\tilde{\Phi}(m_1, m_2) = \sum_{l_1} \sum_{l_2} \tilde{G}(m_1, l_1) \tilde{G}^*(m_2, l_2) e^{-j2\pi f(l_2-l_1)} R_{ee}(n+l_2, n+l_1). \quad (9.1.18)$$

The signal-dependent part of the variance $\sigma_\rho^2(n,f)$ is a quadratic distribution of the signal, with the new kernel $\tilde{\Phi}(m_1, m_2)$, i.e., $\sigma_{se}^2(n,f) = 2\rho_s(n,f; \tilde{\Phi})$.

Special Case 1: White stationary complex noise, when $R_{ee}(n+l_1, n+l_2) = \sigma_\epsilon^2 \delta(l_1 - l_2)$, produces

$$\tilde{\Phi}(m_1, m_2) = \sigma_\epsilon^2 \sum_l \tilde{G}(m_1, l) \tilde{G}^*(m_2, l). \quad (9.1.19)$$

For time-frequency kernels, we assumed realness and symmetry throughout the section, i.e., $\tilde{G}^*(m_2, l) = \tilde{G}(l, m_2)$. Thus, for finite limits, Eq. (9.1.19) is a matrix multiplication form, $\tilde{\Phi} = \sigma_\epsilon^2 \tilde{\mathbf{G}} \cdot \tilde{\mathbf{G}}^* = \sigma_\epsilon^2 \tilde{\mathbf{G}}^2$. Boldface letters, without arguments, will be used to denote a matrix. For example $\tilde{\mathbf{G}}$ is a matrix with elements $\tilde{G}(m, l)$. Thus,

$$\sigma_{se}^2(n,f) = 2\rho_s(n,f; \sigma_\epsilon^2 \tilde{\mathbf{G}}^2). \quad (9.1.20)$$

Note: Any two distributions with kernels $\tilde{G}_1(m, l) = \tilde{G}_2(m, -l)$ have the same variance, since

$$\sum_l \tilde{G}_1(m_1, l) \tilde{G}_1^*(m_2, l) = \sum_l \tilde{G}_1(m_1, -l) \tilde{G}_1^*(-m_2, -l) = \sum_l \tilde{G}_2(m_1, l) \tilde{G}_2^*(m_2, l).$$

Corollary: A distribution with a real symmetric product kernel $g(\nu\tau)$ and the distribution with its dual kernel $g_d(\nu\tau) = \mathcal{F}_{\alpha \rightarrow \nu, \beta \rightarrow \tau}[g(\alpha\beta)]$ have the same variance.

¹An analysis of the bias of $\rho_s(n,f; G)$ due to the kernel may be found in Ref. [1].

Proof: Consider all coordinates in the analog domain. The time-lag domain forms of $g(v\tau)$ and $g_d(v\tau)$, i.e., $G(t, \tau) = \mathcal{F}_{v \rightarrow t}^{-1}[g(v\tau)]$ and $G_d(t, \tau) = \mathcal{F}_{v \rightarrow t}^{-1}[g_d(v\tau)]$, are related by $G(t, \tau) = G_d(\tau, t)$. In the rotated domain, this relation produces $\tilde{G}(t_1, t_2) = \tilde{G}_d(t_1, -t_2)$, which ends the proof according to the previous note.

Example: The WD has the kernel $g(v\tau) = 1$, $\tilde{G}(m, l) = \delta(m + l)$. According to the Corollary, the WD has the same variance as its dual kernel counterpart, with $g(v\tau) = \delta(v, \tau)$, $\tilde{G}(m, l) = \delta(m - l)$. This dual kernel corresponds to the signal energy $\sum_m |x(n + m)|^2$ [see Eq. (9.1.16)]. Thus, the WD and the signal energy have the same variance. The same holds for the smoothed spectrogram, and the enhanced spectrogram (or S-method) described in Sections 3.3.3 and 6.2, whose kernels are $\tilde{G}(m, l) = w(m)p(m + l)w(l)$ and $\tilde{G}(m, l) = w(m)p(m - l)w(l)$, respectively [1]. Their variance is the same.

Eigenvalue decomposition: Assume that both the summation limits and values of $\tilde{G}(m, l)$ are finite. It is true when the kernel $G(m, l)$ is calculated from the well-defined kernel in a finite Doppler-lag domain, $G(m, l) = \mathcal{F}_{v \rightarrow m}^{-1}[g(v, l)]$, using a finite number of samples. The signal-dependent part of the variance $\sigma_{se}^2(n, f)$ can be calculated, like other Cohen-class distributions, by using eigenvalue decomposition of the matrix $\tilde{\mathbf{G}}$ [9,10]. The distribution of the non-noisy signal (9.1.16) is

$$\rho_s(n, f) = \sum_{i=-N/2}^{N/2-1} \lambda_i S_s(n, f; q_i) = \rho_s(n, f; \lambda, q), \quad (9.1.21)$$

where λ_i and $q_i(m)$ are eigenvalues and eigenvectors of the matrix $\tilde{\mathbf{G}}$, respectively, and $S_s(n, f; q_i) = \left| \sum_{m=-N/2}^{N/2-1} s(n + m) q_i(m) e^{-j2\pi fm} \right|^2$ is the spectrogram of signal $s(n)$ calculated by using $q_i(m)$ as a lag window. Since $\tilde{\Phi} = \sigma_\epsilon^2 \tilde{\mathbf{G}}^2$, its eigenvalues and eigenvectors are $\sigma_\epsilon^2 |\lambda_i|^2$ and $q_i(m)$, respectively. Thus, according to Eq. (9.1.20)

$$\sigma_{se}^2(n, f) = 2\sigma_\epsilon^2 \sum_{i=-N/2}^{N/2-1} |\lambda_i|^2 S_s(n, f; q_i) = 2\sigma_\epsilon^2 \rho_s(n, f; |\lambda|^2, q). \quad (9.1.22)$$

Relation between the original kernel and variance $\sigma_{se}^2(n, f)$ kernel: According to Eq. (9.1.21), we can conclude that the original kernel in the Doppler-lag domain can be decomposed into $g(v, l) = \sum_{i=-N/2}^{N/2-1} \lambda_i a_i(v, l)$, where $a_i(v, l)$ are ambiguity functions of the eigenvectors $q_i(m)$. The kernel of $\rho_s(n, f; |\lambda|^2, q)$, in Eq. (9.1.22), is $g_\sigma(n, f) = \sum_{i=-N/2}^{N/2-1} |\lambda_i|^2 a_i(n, f)$. A detailed analysis of distributions, with respect to their eigenvalue properties, is presented in Ref. [10] and Section 6.4. In the sense of that analysis, the signal-dependent variance is just “an energetic map of the time-frequency distribution” of the original signal.

The mean value of variance (9.1.17) is

$$\overline{\sigma_{se}^2(n, f)} = \int_{-1/2}^{1/2} \sigma_{se}^2(n, f) df = 2\sigma_\epsilon^2 \sum_m \tilde{\Phi}(m, m) |s(n + m)|^2. \quad (9.1.23)$$

For frequency modulated signals $s(n) = a \exp[j\phi(n)]$, it is a constant proportional to the kernel energy [8].

Special Case 2: For **nonstationary white complex noise**, Eq. (9.1.18) results in:

$$\tilde{\Phi}(m_1, m_2) = \sum_{l=-N/2}^{N/2-1} I(n+l) \tilde{G}(m_1, l) \tilde{G}^*(m_2, l), \quad (9.1.24)$$

or $\tilde{\Phi} = \tilde{G} \mathbf{I}_n \tilde{G}$, where \mathbf{I}_n is a diagonal matrix, with the elements $I(n+l)$. For the quasi-stationary case $I(n+l_1) \delta(l_1 - l_2) \cong I(n) \delta(l_1 - l_2)$, we have $\mu_i = I(n) |\lambda_i|^2$, with all other parameters as in Eq. (9.1.22).

Special Case 3: In the case of **colored stationary complex noise**, relations dual to those in Special Case 2, hold (like Eqs. (9.1.9) and (9.1.11)).

Special Case 4: Let $x(n) = s(n)(1 + \mu(n))$, where $\mu(n)$ is **multiplicative noise**. We can write $x(n) = s(n) + s(n)\mu(n) = s(n) + \epsilon(n)$, where $\epsilon(n) = s(n)\mu(n)$ is an additive noise. Thus, the case of this kind of multiplicative noise can be analyzed in the same way as the additive noise. For example, if the noise $\mu(n)$ is a nonstationary white complex one with $R_{\mu\mu}(m, n) = I_\mu(n) \delta(n - m)$, then $R_{\epsilon\epsilon}(m, n) = I_\epsilon(n) \delta(n - m)$, where $I_\epsilon(n) = |s(n)|^2 I_\mu(n)$.

9.1.4 NUMERICAL EXAMPLE

Consider the signal

$$\begin{aligned} x(t) = & \exp(j1100(t+0.1)^2) + \exp(-25(t-0.25)^2) \exp(j1000(t+0.75)^2) \\ & + \exp(-25(t-0.67)^2) \exp(j1000(t-0.4)^2) + \exp(j2850t) + \epsilon(t), \end{aligned}$$

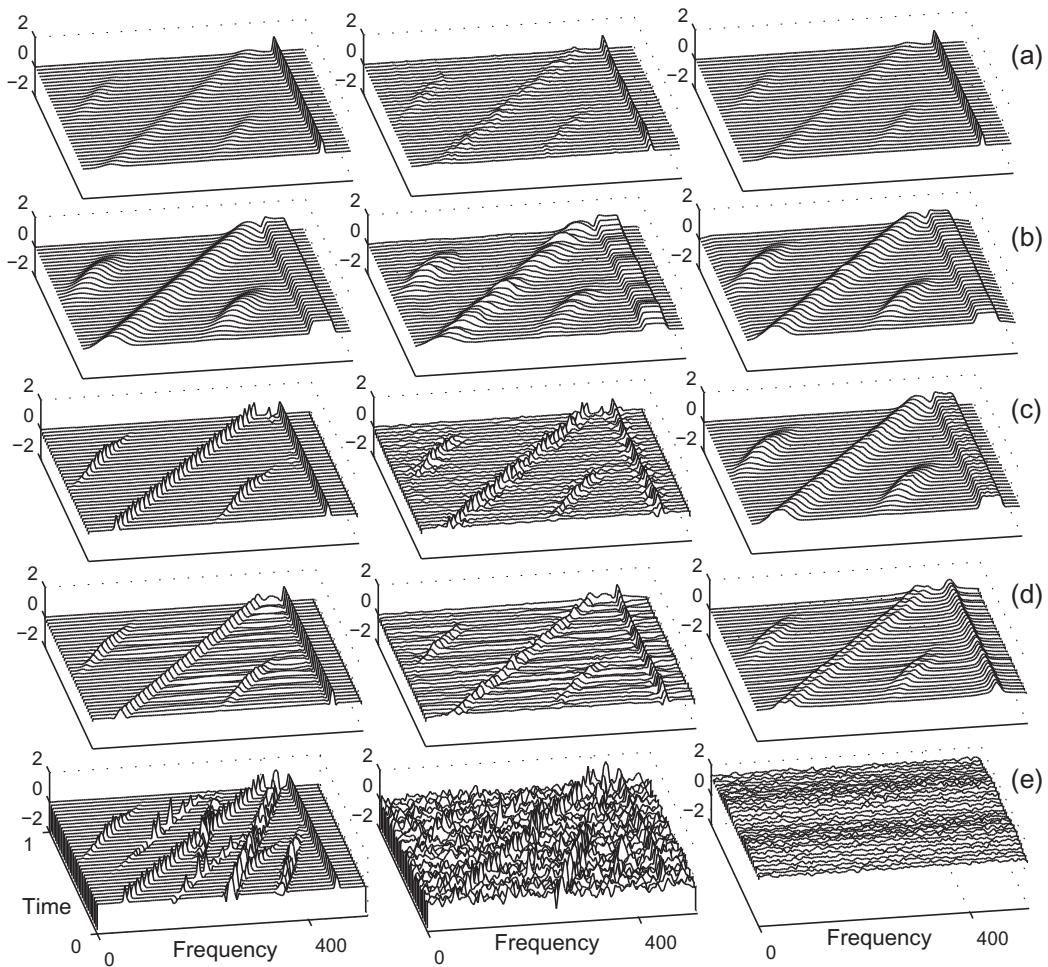
within the interval $[0, 1]$, sampled at $\Delta t = 1/1024$. A Hanning lag window of the width $T_w = 1/4$ is used. Stationary white complex noise with variance $\sigma_\epsilon^2 = 2$ is assumed. The spectrogram, smoothed spectrogram, S-method, Exponential distribution (ED), and the WD of the noise-free signal are presented in the first column of Fig. 9.1.1, respectively. For the ED, the kernel $g(v, \tau) = \exp(-(v\tau)^2)$ is used, with normalized coordinates $-\sqrt{\pi N/2} \leq |2\pi v| < \sqrt{\pi N/2}$, $-\sqrt{\pi N/2} \leq |\tau| < \sqrt{\pi N/2}$, and 128 samples within the intervals. Elements of the matrix \tilde{G} were calculated as [10]

$$\tilde{G}(m, l) = \sum_{p=-N/2}^{N/2} g(p\Delta v, (m-l)\delta\tau) \exp(-j2\pi(m+l)p/(2N)) \delta v. \quad (9.1.25)$$

The normalized eigenvalues of the matrix $\tilde{\Phi}$ were $\lambda_i = \{1, -0.87, 0.69, -0.58, 0.41, -0.30, \dots\}$ and $\mu_i = |\lambda_i|^2 = \{1, 0.76, 0.47, 0.33, 0.17, 0.09, \dots\}$. In the spectrogram and smoothed spectrogram, the whole signal-dependent part of the variance is “located” just on the signal components, while in the WD, it is “spread” over the entire time-frequency plane. Variance behavior in the other two distributions is between these two extreme cases. As has been shown, the variances in the smoothed spectrogram and the S-method are the same [Fig. 9.1.1(b), (c)].

9.1.5 SUMMARY AND CONCLUSIONS

The variances for white nonstationary complex noise with $R_{\epsilon\epsilon}(m, n) = I(n) \delta(m - n)$, $I(n) \geq 0$, may be summarized as follows:

**FIGURE 9.1.1**

Column 1: Time-frequency distributions (TFDs) of a non-noisy signal. *Column 2:* TFDs of one realization of the signal corrupted by white stationary complex noise. *Column 3:* Variances of the TFDs obtained numerically by averaging over 1000 realizations. *Rows:* (a) Spectrogram, (b) Smoothed spectrogram, (c) S-method, (d) Exponential distribution, (e) Pseudo Wigner distribution.

Pseudo Wigner distribution $W_s(n, f; w)$ with $\tilde{G}(m, l) = w(m) \delta(m + l) w(l)$:

$$\sigma_{\text{WD}}^2(n, f) = \sigma_{\epsilon\epsilon}^2(n, f) + \sigma_{s\epsilon}^2(n, f) = W_I(n, 0; w^2) + 2W_{I,|s|^2}(n, 0; w^2), \quad (9.1.26)$$

where $W_{I,|s|^2}$ denotes the cross Wigner distribution for $I(n)$ and $|s(n)|^2$.

Spectrogram $S_s(n, f; w)$ with $\tilde{G}(m, l) = w(m) w(l)$:

$$\sigma_{\text{SPEC}}^2(n, f) = S_I(n, 0; w^2) + 2F_I(n, 0; w^2) S_s(n, f; w), \quad (9.1.27)$$

where $F_I(n, f; w^2)$ is the STFT of $I(n)$ with the window $w^2(m)$.

General quadratic distribution with kernel $\tilde{G}(m, l) = G((m + l)/2, (m - l)/2)$, in Eq. (9.1.6) or Eq. (9.1.16), and $\tilde{\mathbf{G}}$ being a matrix with elements $\tilde{G}(m, l)$:

$$\sigma_\rho^2(n, f) = \rho_I(n, 0; |G|^2) + 2\rho_S(n, f; \tilde{\mathbf{G}}I_n\tilde{\mathbf{G}}). \quad (9.1.28)$$

The first two formulae are special cases of Eq. (9.1.28). Expressions for stationary white noise follow with $I(n) = \sigma_\epsilon^2$. Dual expressions hold for colored stationary noise. Further details can be found in Ref. [11].

9.2 STATISTICAL PROCESSING OF DISPERSIVE SYSTEMS AND SIGNALS⁰

9.2.1 PROCESSING TOOLS FOR TIME-VARYING SYSTEMS AND SIGNALS

In many applications, one can observe linear time-varying (LTV) systems (e.g., wireless communication or sonar channels) and nonstationary random processes (RPs) (e.g., noisy signals with random parameters). Due to their time fluctuations, both of these can exhibit time-frequency (t, f) characteristics that could be embedded in the model design of a system or in the autocorrelation function of a RP. Thus, it would be ideal to design transforms or (t, f) representations (TFRs) that would extract useful information from these characteristics such as important system or RP changes [12–14]. Specifically, signals propagating over LTV systems may be shifted by a constant amount in frequency or shifted in a nonlinear time-dependent manner due to a *dispersive* change in the instantaneous frequency (IF) such that the IF change or shift varies nonlinearly with time. For example, Information of a constant frequency change is very useful in improving the Doppler diversity of a communications channel (Section 13.2). Also, information of an IF change on the input signal to a dispersive system can be used to improve system performance.

An appropriate analysis tool for LTV systems and RPs is a class of TFRs called (t, f) *symbols* [12–14] which are 2-D functions that depend on an integral operator \mathcal{L} defined as $(\mathcal{L}x)(t) = \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}(t, \tau) x(\tau) d\tau$. The kernel $\Upsilon_{\mathcal{L}}(t, \tau)$ of the operator can be considered as the time-varying impulse response of an LTV system or the autocorrelation function of a nonstationary RP.

9.2.1.1 Narrowband Weyl symbol

The narrowband *Weyl symbol* (WS) of a linear operator \mathcal{L} on $L_2(\mathbb{R})$ is used to analyze LTV systems and nonstationary RPs characterized by constant (t, f) shifts and scale changes [13,14]. It is defined as

$$\text{WS}_{\mathcal{L}}(t, f) = \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi\tau f} d\tau. \quad (9.2.1)$$

⁰Authors: **A. Papandreou-Suppappola**, Arizona State University, Tempe, AZ, USA (papandreou@asu.edu), **B.-G. Iem**, Gangneung-Wonju National University, Gangwon-do, Republic of Korea (ibg@gwnu.ac.kr), and **G.F. Boudreax-Bartels**, University of Rhode Island, Kingston, RI, USA (boud@ele.uri.edu). Reviewers: A.M. Sayeed and G. Matz.

Table 9.2.1: Summary of Some Commonly Used Operators

| Operators | Operator Definitions |
|--|--|
| $(\mathcal{S}_\tau x)(t)$ | $x(t - \tau)$ |
| $(\tilde{\mathcal{S}}_\zeta^{(\xi)} x)(t)$ | $(\mathcal{W}_\xi^{-1} \mathcal{S}_{t_r \zeta} \mathcal{W}_\xi x)(t)$ |
| $(\mathcal{M}_v x)(t)$ | $x(t) e^{j2\pi v t}$ |
| $(\tilde{\mathcal{M}}_\beta^{(\xi)} x)(t)$ | $x(t) e^{j2\pi \beta \xi \left(\frac{t}{t_r} \right)} = (\mathcal{W}_\xi^{-1} \mathcal{M}_{\beta/t_r} \mathcal{W}_\xi x)(t)$ |
| $(\mathcal{H}_\beta x)(t)$ | $x(t) e^{j2\pi \beta \ln \left(\frac{t}{t_r} \right)}$ |
| $(\mathcal{P}_\beta^\kappa x)(t)$ | $x(t) e^{j2\pi \beta \operatorname{sgn}(t) \frac{t}{t_r} ^{\kappa-1}}$ |
| $(\mathcal{E}_\beta x)(t)$ | $x(t) e^{j2\pi \beta e^{t/t_r}}$ |
| $(\mathcal{C}_a x)(t)$ | $\frac{1}{\sqrt{ a }} x\left(\frac{t}{a}\right)$ |
| $(\tilde{\mathcal{C}}_a^{(\xi)} x)(t)$ | $(\mathcal{W}_\xi^{-1} \mathcal{C}_a \mathcal{W}_\xi x)(t)$ |
| $(\mathcal{W}_\xi x)(t)$ | $x\left(t_r \xi^{-1} \left(\frac{t}{t_r} \right)\right) \left t_r v \left(t_r \xi^{-1} \left(\frac{t}{t_r} \right) \right) \right ^{-1/2}$ |

Depending on the sign of t , $\operatorname{sgn}(t) = \pm 1$.

It can be interpreted as a time-varying transfer function of an LTV system (see Ref. [13] and Sections 4.7 and 12.4) or as a time-varying spectrum of a nonstationary RP (Section 9.4). When \mathcal{L} is the autocorrelation operator \mathcal{R}_x of a RP $x(t)$ whose kernel is the autocorrelation function $\Upsilon_{\mathcal{R}_x}(t, \tau) = E\{x(t)x^*(\tau)\}$, the WS of \mathcal{R}_x is the expected value of the Wigner distribution (WD), i.e., $\text{WS}_{\mathcal{R}_x}(t, f) = E\{\text{WD}_x(t, f)\}$ (called the Wigner-Ville spectrum in Section 2.2.1). Here, $E\{\cdot\}$ denotes statistical expectation and the WD is defined as $\text{WD}_x(t, f) = \int_{-\infty}^{\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi \tau f} dt$. When \mathcal{L} is an innovations system, the squared magnitude of Eq. (9.2.1) gives the Weyl spectrum, a concept useful for defining (t, f) concentration measures [14], and in (t, f) detection applications [15,16].

The WS satisfies desirable properties such as preserving constant (t, f) shifts and scale changes (Table 9.2.1) on LTV systems and nonstationary RPs. For example, the WS of the autocorrelation operator \mathcal{R}_x of a nonstationary RP $x(t)$ satisfies [13]

$$y(t) = (\mathcal{S}_\tau x)(t) = x(t - \tau) \Rightarrow \text{WS}_{\mathcal{R}_y}(t, f) = \text{WS}_{\mathcal{R}_x}(t - \tau, f), \quad (9.2.2)$$

$$y(t) = (\mathcal{M}_v x)(t) = x(t) e^{j2\pi v t} \Rightarrow \text{WS}_{\mathcal{R}_y}(t, f) = \text{WS}_{\mathcal{R}_x}(t, f - v), \quad (9.2.3)$$

$$y(t) = (\mathcal{C}_a x)(t) = \frac{1}{\sqrt{|a|}} x\left(\frac{t}{a}\right) \Rightarrow \text{WS}_{\mathcal{R}_y}(t, f) = \text{WS}_{\mathcal{R}_x}\left(\frac{t}{a}, af\right), \quad (9.2.4)$$

where \mathcal{S}_τ , \mathcal{M}_v , and \mathcal{C}_a are the constant time-shift, constant frequency-shift, and scale operators, respectively. The WS also satisfies the unitarity property given by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WS}_{\mathcal{L}}(t, f) \text{WS}_{\mathcal{N}}^*(t, f) dt df = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}(t, \tau) \Upsilon_{\mathcal{N}}^*(t, \tau) dt d\tau, \quad (9.2.5)$$

where $\Upsilon_{\mathcal{L}}(t, \tau)$ and $\Upsilon_{\mathcal{N}}(t, \tau)$ are the kernels of the operators \mathcal{L} and \mathcal{N} , respectively. This is an important property for preserving energy or norms. Since it is unitary, the WS is *associated* with the unitary WD and preserves the quadratic form

$$\int_{-\infty}^{\infty} (\mathcal{L}x)(t) x^*(t) dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WS}_{\mathcal{L}}(t, f) \text{WD}_x(t, f) dt df. \quad (9.2.6)$$

Here, the correlation of the WS of a system \mathcal{L} with the WD of the system's input $x(t)$ results in the correlation of the system's output $(\mathcal{L}x)(t)$ with $x(t)$ [13,14].

9.2.1.2 Narrowband spreading function

The spreading function (SF) of a linear operator \mathcal{L} is defined as the 2-D Fourier transform (FT) of the WS,

$$\begin{aligned} \text{SF}_{\mathcal{L}}(\tau, \nu) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WS}_{\mathcal{L}}(t, f) e^{-j2\pi(t\nu-f\tau)} dt df \\ &= \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi t\nu} dt. \end{aligned} \quad (9.2.7)$$

Using the SF, the output of an LTV system \mathcal{L} can be interpreted as a superposition of (t, f) shifted versions of the input signal each weighted by the SF [14]

$$(\mathcal{L}x)(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{SF}_{\mathcal{L}}(\tau, \nu) e^{-j\pi\tau\nu} x(t - \tau) e^{j2\pi\nu t} d\tau d\nu. \quad (9.2.8)$$

The system output interpretation in Eq. (9.2.8) is comparable to that of a linear time-invariant (LTI) system, and it can be used to formulate (t, f) receiver structures. For example, it can provide critical information on the output signal of a time and frequency selective wireless communications channel. The SF is also used in the classification of operators as underspread or overspread [13].

9.2.1.3 Wideband Weyl symbol

The (t, f) version of the affine WS, introduced in Ref. [14], is called P_0 -Weyl symbol (P_0 WS) in Ref. [12], and for $\lambda(\alpha) = \frac{\alpha/2}{\sinh(\alpha/2)}$, it is defined as

$$P_0\text{WS}_{\mathcal{B}}(t, f) = f \int_{-\infty}^{\infty} \Gamma_{\mathcal{B}}\left(f\lambda(\alpha)e^{\frac{\alpha}{2}}, f\lambda(\alpha)e^{-\frac{\alpha}{2}}\right) \lambda(\alpha) e^{j2\pi f t \alpha} d\alpha, \quad f > 0,$$

where $(\mathcal{B}X)(f) = \int_0^{\infty} \Gamma_{\mathcal{B}}(f, \nu) X(\nu) d\nu$, and $\Gamma_{\mathcal{B}}(f, \nu)$ is the kernel of the frequency domain operator \mathcal{B} on $L_2(\mathbb{R}^+)$. The P_0 WS is a unitary symbol related to the unitary Bertrand P_0 -distribution [17]. The wideband SF (WSF) is the 2-D modified FT of $P_0\text{WS}_{\mathcal{B}}(t, f)$ [14], and they are both important for analyzing RPs and LTV systems having constant or hyperbolic time shifts and scale changes [12,14].

9.2.2 DISPERSIVE TIME-FREQUENCY SYMBOLS

Systems and RPs characterized by dispersive time or frequency shifts (see Section 5.6) include the ocean that can cause echoes with power dispersive characteristics from acoustic waves reflected from immersed spherical shells [17,18]. The WS is not well suited to analyze such systems and processes as it is better matched to constant (t, f) shifts [12,13]. Instead, as will be shown next, modified versions of the narrowband WS and SF were developed for analyzing RPs and LTV systems characterized by dispersive (such as power) IF shifts [12].

9.2.2.1 Dispersive Weyl symbol

Let the linear operator \mathcal{Z} , defined in the time domain on $L_2([p, q])$, be $(\mathcal{Z}x)(t) = \int_p^q \Upsilon_{\mathcal{Z}}(t, \tau) x(\tau) d\tau$, with kernel $\Upsilon_{\mathcal{Z}}(t, \tau)$. The *dispersive Weyl symbol*, $DWS_{\mathcal{Z}}^{(\xi)}(t, f)$, of \mathcal{Z} is defined in Ref. [12] as

$$DWS_{\mathcal{Z}}^{(\xi)}(t, f) = \int_c^d \Upsilon_{\mathcal{Z}}(l(t, \zeta), l(t, -\zeta)) e^{-j2\pi \frac{f\zeta}{v(t)}} \frac{d\zeta}{|v(l(t, \zeta)) v(l(t, -\zeta))|^{\frac{1}{2}}} \quad (9.2.9)$$

$$= WS_{\mathcal{W}_{\xi} \mathcal{Z} \mathcal{W}_{\xi}^{-1}} \left(t_r \xi \left(\frac{t}{t_r} \right), \frac{f}{t_r v(t)} \right), \quad (9.2.10)$$

where $l(t, \zeta) = t_r \xi^{-1} \left(\xi \left(\frac{t}{t_r} \right) + \frac{\zeta}{2} \right)$, $\xi^{-1}(\xi(b)) = b$, $v(t) = \frac{d}{dt} \xi \left(\frac{t}{t_r} \right)$, and the integration range $[c, d]$ in Eq. (9.2.9) and $[p, q]$ above depend on the range and domain, respectively, of $\xi(\cdot)$. Note that $t_r > 0$ is a reference time point that is needed for unit precision (unless otherwise stated, it could be taken as $t_r = 1$). Thus, the DWS is obtained by warping the operator \mathcal{Z} and the WS in Eq. (9.2.1) using a unitary transformation (Refs. [12,17] and Sections 5.6 and 4.5) based on the one-to-one warping function $\xi(b)$. With $(\mathcal{W}_{\xi}^{-1} \mathcal{W}_{\xi} x)(t) = x(t)$, the unitary warping operator \mathcal{W}_{ξ} in Eq. (9.2.10) becomes

$$(\mathcal{W}_{\xi} x)(t) = x \left(t_r \xi^{-1} \left(\frac{t}{t_r} \right) \right) \left| t_r v \left(t_r \xi^{-1} \left(\frac{t}{t_r} \right) \right) \right|^{-1/2}.$$

The DWS preserves dispersive IF shifts on an RP $x(t)$; i.e.,

$$y(t) = x(t) e^{j2\pi \beta \xi \left(\frac{t}{t_r} \right)} \Rightarrow DWS_{\mathcal{R}_y}^{(\xi)}(t, f) = DWS_{\mathcal{R}_x}^{(\xi)}(t, f - \beta v(t)), \quad (9.2.11)$$

where¹ $y(t) = (\tilde{\mathcal{M}}_{\beta}^{(\xi)} x)(t) = x(t) e^{j2\pi \beta \xi \left(\frac{t}{t_r} \right)}$, $\tilde{\mathcal{M}}_{\beta}^{(\xi)}$ is the IF shift operator, and $\beta v(t)$ is the change in IF. This follows as the WS preserves constant frequency shifts $(\mathcal{M}_v x)(t)$ in Eq. (9.2.3), and the warping in Eq. (9.2.10) yields $\mathcal{W}_{\xi}^{-1} \mathcal{M}_{\beta/t_r} \mathcal{W}_{\xi} = \tilde{\mathcal{M}}_{\beta}^{(\xi)}$. Because of this important property, the DWS is potentially useful in analyzing RPs or LTV systems with characteristics that are constant (when $\xi(b) = b$) or dispersive (when $\xi(b)$ is nonlinear). The DWS also preserves warped time shifts $(\tilde{\mathcal{S}}_{\zeta}^{(\xi)} x)(t) = (\mathcal{W}_{\xi}^{-1} \mathcal{S}_{t_r \zeta} \mathcal{W}_{\xi} x)(t)$ [cf. Eq. (9.2.2)] and warped scale changes $(\tilde{\mathcal{C}}_a^{(\xi)} x)(t) = (\mathcal{W}_{\xi}^{-1} \mathcal{C}_a \mathcal{W}_{\xi} x)(t)$ [cf. Eq. (9.2.4)] as defined in Ref. [12].

The DWS is also useful when analyzing random nonlinear frequency modulated (FM) signals $x(t) = \mathbf{a} \varrho(t) e^{j2\pi \beta \phi \left(\frac{t}{t_r} \right)}$ with phase function $\phi(b)$ and random amplitude \mathbf{a} (Ref. [17] and Section 5.6). When the phase function of the nonlinear FM signal is matched to the warping function $\xi(b)$ used in the DWS, i.e., when $\phi(b) = \xi(b)$ and $\varrho(t) = |\frac{d}{dt} \phi \left(\frac{t}{t_r} \right)|^{1/2} = |v(t)|^{1/2}$, the DWS of the process autocorrelation produces an ideally localized representation, $DWS_{\mathcal{R}_x}^{(\xi)}(t, f) = E \{|\mathbf{a}|^2\} \delta(f - \beta v(t))$, provided that the range of $\xi(b)$ is \mathbb{R} . The DWS is localized along the IF $f = \beta v(t)$.

¹The tilde above operators indicate the warped versions of the operators in Eqs. (9.2.2)–(9.2.4).

9.2.2.2 Dispersive spreading function

The dispersive spreading function (DSF) is obtained as the SF in Eq. (9.2.7) of the warped operator [12] $\mathcal{W}_\xi \mathcal{Z} \mathcal{W}_\xi^{-1}$,

$$\text{DSF}_{\mathcal{Z}}^{(\xi)}(\zeta, \beta) = \text{SF}_{\mathcal{W}_\xi \mathcal{Z} \mathcal{W}_\xi^{-1}}(t_r \zeta, \beta/t_r). \quad (9.2.12)$$

For an LTV system \mathcal{Z} , the DSF provides an interpretation of the operator output as a weighted sum of warped (t, f) shifted versions of the input signal $x(t)$,

$$(\mathcal{Z}x)(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{DSF}_{\mathcal{Z}}^{(\xi)}(\zeta, \beta) e^{-j\pi\zeta\beta} (\tilde{\mathcal{M}}_\beta^{(\xi)} \tilde{\mathcal{S}}_\zeta^{(\xi)} x)(t) d\zeta d\beta. \quad (9.2.13)$$

Depending on the warping function $\xi(b)$, the formulation in Eq. (9.2.13) simplifies to a specific interpretation; for example, if $\xi(b) = \xi_{\ln}(b) = \ln b$, then $(\tilde{\mathcal{M}}_\beta^{(\xi_{\ln})} \tilde{\mathcal{S}}_\zeta^{(\xi_{\ln})} x)(t) = e^{-\frac{\zeta}{2}} x(te^{-\zeta}) e^{j2\pi\beta \ln\left(\frac{t}{t_r}\right)}$, and Eq. (9.2.13) describes the operator \mathcal{Z} as a weighted sum of hyperbolic IF shifts and scale changes (by e^ζ) on the input $x(t)$ [12]. The DSF was used for shallow water environment characterization in Ref. [18].

9.2.2.3 DWS unitarity and quadratic form

The DWS obeys Eq. (9.2.5) since [12]

$$\int_{-\infty}^{\infty} \int_p^q \text{DWS}_{\mathcal{Z}}^{(\xi)}(t, f) \text{DWS}_{\mathcal{X}}^{(\xi)*}(t, f) dt df = \int_p^q \int_p^q \Upsilon_{\mathcal{Z}}(t, \tau) \Upsilon_{\mathcal{X}}^*(t, \tau) dt d\tau. \quad (9.2.14)$$

The DWS relates to $\text{DWD}_x^{(\xi)}(t, f) = \text{WD}_{\mathcal{W}_\xi x}\left(t_r \xi\left(\frac{t}{t_r}\right), \frac{f}{t_r v(t)}\right)$, i.e., the IF shift covariant version of the WD (see Section 5.6). Using the DWS, Eq. (9.2.6) becomes

$$\int_p^q (\mathcal{Z}x)(t) x^*(t) dt = \int_{-\infty}^{\infty} \int_p^q \text{DWS}_{\mathcal{Z}}^{(\xi)}(t, f) \text{DWD}_x^{(\xi)}(t, f) dt df \quad (9.2.15)$$

with potential detection applications for nonlinear (t, f) processes [12].

9.2.3 SPECIAL CASES OF DISPERSIVE TIME-FREQUENCY SYMBOLS

Depending on the warping function $\xi(b)$, the DWS may simplify to (t, f) symbols that are matched to linear or non-linear (dispersive) (t, f) structures. The trivial case $\xi(b) = b$ simplifies the DWS to the WS. Other examples are described below.

9.2.3.1 Hyperbolic WS and SF

The hyperbolic Weyl symbol (HWS) is obtained as the warped version of the WS when $\xi(b) = \xi_{\ln}(b) = \ln b$ in the DWS formulation in Eqs. (9.2.10) and (9.2.2.1). That is, for a linear time domain operator \mathcal{J} on $L_2(\mathbb{R}^+)$, $\text{HWS}_{\mathcal{J}}(t, f) = \text{DWS}_{\mathcal{J}}^{(\xi_{\ln})}(t, f) = \text{WS}_{\mathcal{W}_{\xi_{\ln}} \mathcal{J} \mathcal{W}_{\xi_{\ln}}^{-1}}\left(t_r \ln\left(\frac{t}{t_r}\right), \frac{tf}{t_r}\right)$, $t > 0$. This yields

$$\text{HWS}_{\mathcal{J}}(t, f) = t \int_{-\infty}^{\infty} \Upsilon_{\mathcal{J}}\left(te^{\zeta/2}, te^{-\zeta/2}\right) e^{-j2\pi tf\zeta} d\zeta, \quad (9.2.16)$$

where $\Upsilon_{\mathcal{J}}(t, \tau)$ is the kernel of \mathcal{J} [12]. The HWS is significant for processing systems and nonstationary signals with hyperbolic (t, f) characteristics. Specifically, it preserves *hyperbolic* IF changes on an RP $x(t)$:

$$y(t) = x(t) e^{j2\pi\beta \ln\left(\frac{t}{t_r}\right)} \Rightarrow \text{HWS}_{\mathcal{R}_y}(t, f) = \text{HWS}_{\mathcal{R}_x}\left(t, f - \frac{\beta}{t}\right),$$

where $y(t) = (\tilde{\mathcal{M}}_{\beta}^{(\xi \ln)} x)(t) = (\mathcal{H}_{\beta} x)(t)$ is the hyperbolic IF shift operator (obtained when $\xi(b) = \ln b$ in Eq. (9.2.11)). The HWS also satisfies the scale covariance property in Eq. (9.2.4) since $\tilde{\mathcal{S}}_{\zeta}^{(\xi \ln)} = \mathcal{C}_{e^{\zeta}}$ [12]. The HWS satisfies the unitarity property in Eq. (9.2.14) with $\xi(b) = \ln b$ and $[p, q] = [0, \infty)$. It is *associated* with the dual form of the Altes Q-distribution (Ref. [17] and Section 5.6). As an example of the HWS, if the output of a system \mathcal{J} is the scale convolution of an input signal $x(t)$ and some function $g(t)$, then the HWS in Eq. (9.2.16) is naturally the Mellin transform of $g(t)$, given its scale property; i.e., $\text{HWS}_{\mathcal{J}}(t, f) = \int_0^{\infty} g(\tau) \frac{1}{\sqrt{\tau}} e^{-j2\pi t f \ln\left(\frac{\tau}{t_r}\right)} d\tau$. For comparison, the WS in Eq. (9.2.1), $\text{WS}_{\mathcal{J}}(t, f) = \int_0^{\infty} \frac{\sqrt{t_r}}{t-\tau/2} g\left(t_r \frac{t+\tau/2}{t-\tau/2}\right) e^{-j2\pi t f} d\tau$, of the *same* operator \mathcal{J} is difficult to interpret.

The hyperbolic spreading function (HSF) is obtained from Eq. (9.2.12) when $\xi(b) = \xi_{\ln}(b) = \ln b$ as $\text{HSF}_{\mathcal{J}}(\zeta, \beta) = \text{DSF}_{\mathcal{J}}^{(\xi_{\ln})}(\zeta, \beta) = \text{SF}_{\mathcal{W}_{\xi_{\ln}} \mathcal{J} \mathcal{W}_{\xi_{\ln}}^{-1}}(t_r \zeta, \beta/t_r)$ yielding [12]

$$\text{HSF}_{\mathcal{J}}(\zeta, \beta) = \int_0^{\infty} \Upsilon_{\mathcal{J}}\left(te^{\zeta/2}, te^{-\zeta/2}\right) e^{-j2\pi\beta \ln\left(\frac{t}{t_r}\right)} dt.$$

It is related to the HWS using a modified FT and a Mellin transform as $\text{HSF}_{\mathcal{J}}(\zeta, \beta) = \int_{-\infty}^{\infty} \int_0^{\infty} \text{HWS}_{\mathcal{J}}(t, f) e^{j2\pi\zeta tf} e^{-j2\pi\beta \ln\left(\frac{t}{t_r}\right)} dt df$ [12]. The HSF provides an alternative interpretation of the operator output as a weighted sum of hyperbolic IF shifted and scale changed versions of the input signal where the weight is the HSF, i.e., $(\mathcal{J}x)(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{HSF}_{\mathcal{J}}(\zeta, \beta) e^{-j\pi\zeta\beta} (\mathcal{H}_{\beta} \mathcal{C}_{e^{\zeta}} x)(t) d\zeta d\beta$, $t > 0$. The hyperbolic version of the quadratic form in Eq. (9.2.15) with $\xi(b) = \ln b$ is useful in detection when signals have hyperbolic (t, f) characteristics [12]. The HSF was used in designing hyperbolic signaling for underwater acoustic communications in Ref. [19].

9.2.3.2 Power WS and SF

When $\xi(b) = \xi_{\kappa}(b) = \text{sgn}(b) |b|^{\kappa}$ and $v_{\kappa}(t) = \frac{d}{dt} \xi_{\kappa}\left(\frac{t}{t_r}\right)$, the DWS in Eq. (9.2.9) and the DSF in Eq. (9.2.12) simplify, respectively, to the κ th power WS (PWS) and the κ th power SF (PSF) for $\mathcal{L} \in L_2(\mathbb{R})$

$$\begin{aligned} \text{PWS}_{\mathcal{L}}^{(\kappa)}(t, f) &= \text{DWS}_{\mathcal{L}}^{(\xi_{\kappa})}(t, f) = \text{WS}_{\mathcal{W}_{\xi_{\kappa}} \mathcal{L} \mathcal{W}_{\xi_{\kappa}}^{-1}}\left(t_r \xi_{\kappa}\left(\frac{t}{t_r}\right), \frac{f}{t_r v_{\kappa}(t)}\right), \\ \text{PSF}_{\mathcal{L}}^{(\kappa)}(\zeta, \beta) &= \text{DSF}_{\mathcal{L}}^{(\xi_{\kappa})}(\zeta, \beta) = \text{SF}_{\mathcal{W}_{\xi_{\kappa}} \mathcal{L} \mathcal{W}_{\xi_{\kappa}}^{-1}}(t_r \zeta, \beta/t_r), \end{aligned}$$

where $(\mathcal{W}_{\xi_{\kappa}} x)(t) = x\left(t_r \xi_{\kappa}^{-1}\left(\frac{t}{t_r}\right)\right) |t_r v_{\kappa}\left(t_r \xi_{\kappa}^{-1}\left(\frac{t}{t_r}\right)\right)|^{-1/2}$.

The PWS preserves power IF shifts on a RP $x(t)$; i.e.,

$$y(t) = x(t) e^{j2\pi\beta \xi_{\kappa}\left(\frac{t}{t_r}\right)} \Rightarrow \text{PWS}_{\mathcal{R}_y}^{(\kappa)}(t, f) = \text{PWS}_{\mathcal{R}_x}^{(\kappa)}(t, f - \beta v_{\kappa}(t)),$$

where $y(t) = (\tilde{\mathcal{M}}_{\beta}^{(\xi_k)} x)(t) = (\mathcal{P}_{\beta}^{\kappa} x)(t)$ is the power IF shift operator (the operator in Eq. (9.2.11) when $\xi(b) = \xi_k(b)$). The PWS also preserves scale changes since warped scale covariance simplifies to the scale covariance in Eq. (9.2.4), i.e., $\tilde{\mathcal{C}}_a^{(\xi_k)} = \mathcal{C}_{\xi_{1/\kappa}(a)}$. The corresponding operator output can be interpreted as a weighted sum of power IF shifted and power warped time-shifted versions of the input signal. Also, the PWS is unitary as it satisfies Eq. (9.2.14) and it is *associated* with the power WD (Ref. [17] and Sections 5.6 and 15.3) in the quadratic form in Eq. (9.2.15) when $\xi(b) = \xi_k(b)$ [12].

9.2.3.3 Exponential WS and SF

For a linear operator \mathcal{L} defined on $L_2(\mathbb{R})$, the exponential WS, $\text{EWS}_{\mathcal{L}}(t,f) = \text{WS}_{\mathcal{W}_{\xi_e}\mathcal{L}\mathcal{W}_{\xi_e}^{-1}}(t_r e^{t/t_r}, fe^{-t/t_r})$, and the exponential SF, $\text{ESF}_{\mathcal{L}}(\zeta, \beta) = \text{SF}_{\mathcal{W}_{\xi_e}\mathcal{L}\mathcal{W}_{\xi_e}^{-1}}(t_r \zeta, \beta/t_r)$ are given as exponentially warped versions of the narrowband WS and SF, respectively. Here, $(\mathcal{W}_{\xi_e}x)(t) = x\left(t_r \ln\left(\frac{t}{t_r}\right)\right) \sqrt{t_r/t}$, $t > 0$. The EWS and the ESF are obtained from the DWS in Eq. (9.2.9) and the DSF in Eq. (9.2.12), respectively, when $\xi(b) = e^b$. The EWS preserves exponential IF shifts.

$$y(t) = x(t) e^{j2\pi\beta e^{t/t_r}} \Rightarrow \text{EWS}_{\mathcal{R}_y}(t,f) = \text{EWS}_{\mathcal{R}_x}(t, f - \beta e^{t/t_r}/t_r),$$

where $y(t) = (\mathcal{D}_{\beta}^{(\xi_e)} x)(t) = (\mathcal{E}_{\beta} x)(t)$. Also, the EWS preserves constant time shifts in Eq. (9.2.2) since $\tilde{\mathcal{C}}_a^{(\xi_e)} = \mathcal{S}_{t_r \ln a}$ is unitary, and satisfies the quadratic form in Eq. (9.2.15) with *association* with the exponential WD (see Section 5.6 and Ref. [17]).

9.2.4 ANALYSIS APPLICATION EXAMPLES

As shown next, a dispersive WS produces an ideally localized spectrum of a process when they both have similar dispersive (t,f) characteristics. Figure 9.2.1 demonstrates the advantage of the HWS over the WS when used to analyze signals with hyperbolic (t,f) characteristics. The deterministic signal

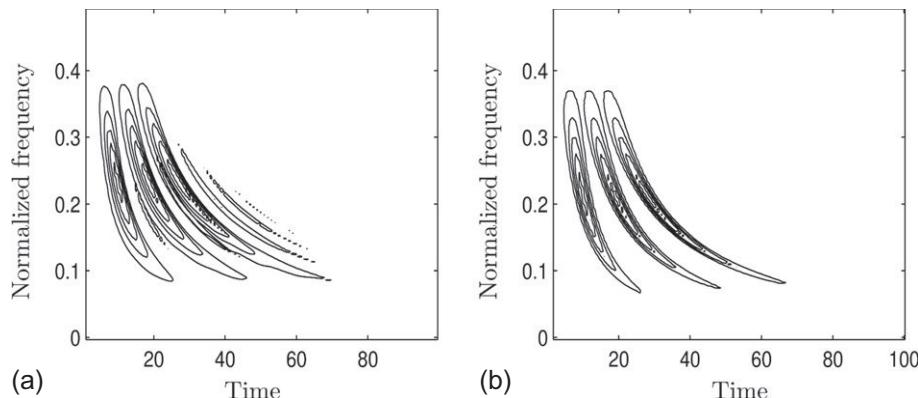
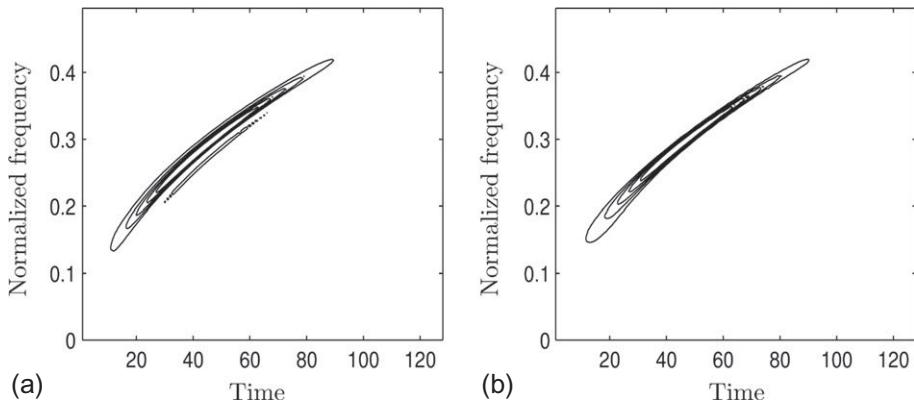


FIGURE 9.2.1

(a) Narrowband Weyl symbol and (b) hyperbolic Weyl symbol of a RP consisting of signal components with random amplitude and hyperbolic instantaneous frequency.

**FIGURE 9.2.2**

(a) Narrowband Weyl symbol and (b) power Weyl symbol of an RP consisting of a signal component with random amplitude and power instantaneous frequency.

components are weighted by random, zero mean amplitudes, i.e., $x(t) = \left[\sum_{m=1}^3 \mathbf{a}_m s_m(t) \right] w(t)$ where $w(t)$ is a shifted Hanning window, and \mathbf{a}_m are uncorrelated random variables with zero means and constant variances $\sigma_{\mathbf{a}_m}^2$. Also, $s_m(t) = \frac{1}{\sqrt{t}} e^{j4\pi m \ln\left(\frac{t}{t_r}\right)}$, $t > 0$ is a deterministic hyperbolic FM signal (see [Section 5.6](#)). For the ideal case without windowing, the theoretical HWS of the process $\tilde{x}(t) = \sum_{m=1}^3 \mathbf{a}_m s_m(t)$ is

$$\text{HWS}_{\mathcal{R}_{\tilde{x}}}(t, f) = \sum_{m=1}^3 E \left\{ |\mathbf{a}_m|^2 \right\} \text{HWS}_{\mathcal{R}_{s_m}}(t, f) = \frac{1}{t} \sum_{m=1}^3 \sigma_{\mathbf{a}_m}^2 \delta \left(f - \frac{2m}{t} \right), \quad t > 0.$$

It has three Diracs centered along the hyperbolic IF $f = 2m/t$ of $s_m(t)$, $m = 1, 2, 3$. Both WS and HWS produce no outer cross-terms between neighboring hyperbolae since the random amplitudes \mathbf{a}_m are uncorrelated. However, the WS in [Fig. 9.2.1\(a\)](#) has inner cross-terms inside each hyperbola since it does not match the hyperbolic (t, f) characteristics of the process. The HWS in [Fig. 9.2.1\(b\)](#) is an ideally concentrated TFR without any cross-terms as it is matched to the process.

[Figure 9.2.2](#) shows the advantage of the $\kappa = 3$ power WS over the narrowband WS when the analysis RP is a power FM signal $x(t) = \mathbf{a} v_\kappa(t) e^{j2\pi \beta \xi_\kappa \left(\frac{t}{t_r} \right)}$ (also with $\kappa = 3$) [see Ref. [\[17\]](#) and [Sections 5.6 and 15.3](#)]. The power WS in [Fig. 9.2.2\(b\)](#) results in an ideal TFR for this process since, unlike the WS in [Fig. 9.2.2\(a\)](#), it matches the signal's power (t, f) characteristics without any inner interference.

9.2.5 SUMMARY AND CONCLUSIONS

The (t, f) symbols are useful time-varying counterparts of transfer functions for systems and of power spectral densities for RPs. The Weyl symbol has been successfully used for narrowband systems with constant (t, f) structures, whereas the P_0 -Weyl symbol has been used for wideband systems with either constant or hyperbolic (t, f) characteristics. For systems with nonlinear dispersive structures,

new symbols are presented for matched processing suitable for analyzing tetherless underwater communication systems in dispersive mediums.

9.3 ROBUST TIME-FREQUENCY DISTRIBUTIONS⁰

Study of the additive Gaussian noise influence on time-frequency (TF) distributions is an important issue (see [Section 9.1](#)). However, in many practical applications, especially in communications, signals are disturbed by a kind of impulse noise. These noises are commonly modeled by heavy-tailed (long-tailed) probability density functions (pdfs) [20]. It is well known that the conventional TF distributions are quite sensitive to this kind of noise, which is able to destroy sensitive signal information. The minimax Huber M -estimates [21] can be applied in order to design the periodogram and TF distributions, robust with respect to the impulse noise. For nonstationary signals, the robust TF distributions are developed as an extension of the robust M -estimation approach.

9.3.1 ROBUST SPECTROGRAM

The standard short-time Fourier transform (STFT) at a given point (t, f) , introduced in [Section 2.3.1](#), can be alternatively defined as a solution of the following optimization problem:

$$F_z(t, f) = \arg \left\{ \min_m I(t, f, m) \right\}, \quad (9.3.1)$$

$$I(t, f, m) = \sum_{n=-N/2}^{N/2-1} w(n\Delta t) \mathbf{F}(e(t, f, n)). \quad (9.3.2)$$

Here, the loss function is given as $\mathbf{F}(e) = |e|^2$, and $w(n\Delta t)$ is a window function, and Δt is a sampling interval. The error function has the form

$$e(t, f, n) = z(t + n\Delta t) e^{-j2\pi fn\Delta t} - m, \quad (9.3.3)$$

where m is a complex-valued optimization parameter in Eq. (9.3.1). The error function can be considered as a residuum expressing the “similarity” between the signal and a given harmonic $\exp(j2\pi fn\Delta t)$.

The solution of Eq. (9.3.1) easily follows from

$$\frac{\partial I(t, f, m)}{\partial m^*} = 0 \quad (9.3.4)$$

in the form of the well-known standard STFT (cf. [Section 6.1.3.3](#)):

$$F_z(t, f) = \frac{1}{a_w} \sum_{n=-N/2}^{N/2-1} w(n\Delta t) z(t + n\Delta t) e^{-j2\pi fn\Delta t}, \quad (9.3.5)$$

⁰Authors: **V. Katkovnik**, Tampere University of Technology, Hervanta, Tampere, Finland (vladimir.katkovnik@tut.fi), **I. Djurović**, University of Montenegro, Podgorica, Montenegro (igordj@ac.me), and **LJ. Stanković**, University of Montenegro, Podgorica, Montenegro (ljubisa@ac.me). Reviewer: S. Stanković.

where

$$a_w = \sum_{n=-N/2}^{N/2-1} w(n\Delta t). \quad (9.3.6)$$

The corresponding spectrogram is defined by

$$S_z(t, f) = |F_z(t, f)|^2. \quad (9.3.7)$$

The maximum likelihood (ML) approach can be used for selection of the appropriate loss function $\mathbf{F}(e)$ if the pdf $p(e)$ of the noise is known. The ML approach suggests the loss function $\mathbf{F}(e) \sim -\log p(e)$. For example, the loss function $\mathbf{F}(e) = |e|^2$ gives the standard STFT, as the ML estimate of spectra for signals corrupted with the Gaussian noise, $p(e) \sim \exp(-|e|^2)$. The standard STFT produces poor results for signals corrupted by impulse noise. Additionally, in many cases, the ML estimates are quite sensitive to deviations from the parametric model and the hypothetical distribution. Even a slight deviation from the hypothesis can result in a strong degradation of the ML estimate. The minimax robust approach has been developed in statistics as an alternative to the conventional ML in order to decrease the ML estimate's sensitivity, and to improve the efficiency in an environment with the heavy-tailed pdfs. The loss function

$$\mathbf{F}(e) = |e| = \sqrt{\text{Re}^2\{e\} + \text{Im}^2\{e\}}, \quad (9.3.8)$$

where Re and Im denote the real and imaginary parts, is recommended by the robust estimation theory for a wide class of heavy-tailed pdfs. It is worth noting that the loss function

$$\mathbf{F}(e) = |\text{Re}\{e\}| + |\text{Im}\{e\}| \quad (9.3.9)$$

is the ML selection for the Laplacian distribution of independent real and imaginary parts of the complex valued noise.

Nonquadratic loss functions in Eq. (9.3.1) can improve filtering properties for impulse noises. Namely, in Refs. [22,23], it is proved that there is a natural link between the problem of spectra's resistance to impulse noise and the minimax Huber estimation theory. It has been shown that the loss function derived in this theory could be applied to the design of a class of robust spectra, inheriting properties of strong resistance to impulse noises.

In particular, the robust M -STFT has been derived by using the absolute error loss function $\mathbf{F}(e) = |e|$ in Eqs. (9.3.1)–(9.3.4) [23]. It is a solution of the nonlinear equation

$$F_z(t, f) = \frac{1}{a_w(t, f)} \sum_{n=-N/2}^{N/2-1} d(t, f, n) z(t + n\Delta t) e^{-j2\pi fn\Delta t}, \quad (9.3.10)$$

where

$$d(t, f, n) = \frac{w(n\Delta t)}{|z(t + n\Delta t) e^{-j2\pi fn\Delta t} - F_z(t, f)|}, \quad (9.3.11)$$

and

$$a_w(t, f) = \sum_{n=-N/2}^{N/2-1} d(t, f, n). \quad (9.3.12)$$

If real and imaginary parts of the additive noise are independent, the statistically optimal robust estimation theory requires replacement of Eqs. (9.3.1) and (9.3.2) with [22]:

$$F_z(t, f) = \arg \left\{ \min_m I_1(t, f, m) \right\}, \quad (9.3.13)$$

$$I_1(t, f, m) = \sum_{n=-N/2}^{N/2-1} w(n\Delta t) [\mathbf{F}(\text{Re}\{e_1\}) + \mathbf{F}(\text{Im}\{e_1\})], \quad (9.3.14)$$

where e_1 is an error function of the form

$$e_1(t, f, n) = z(t + n\Delta t) - m e^{j2\pi f n \Delta t}. \quad (9.3.15)$$

For $\mathbf{F}(e) = |e|$, the robust STFT Eq. (9.3.13) can be presented as a solution of Eq. (9.3.10), where $d(t, f, n)$ is given by

$$d(t, f, n) = w(n\Delta t) \frac{|\text{Re}\{e_1(t, f, n)\}| + |\text{Im}\{e_1(t, f, n)\}|}{|\text{Re}\{e_1(t, f, n)\}|^2 + |\text{Im}\{e_1(t, f, n)\}|^2}. \quad (9.3.16)$$

The robust spectrogram defined in the form

$$S_z(t, f) = I_1(t, f, 0) - I_1(t, f, F_z(t, f)) \quad (9.3.17)$$

is called the residual spectrogram, in order to distinguish it from the amplitude spectrogram (9.3.7). For the quadratic loss function $\mathbf{F}(e)$, the residual spectrogram (9.3.17) coincides with the standard amplitude spectrogram (9.3.7). In Ref. [22], it has been shown that, in a heavy-tailed noise environment, the residual robust spectrogram performs better than its amplitude counterpart.

The accuracy analysis of the robust spectrograms, as well as a discussion on further details on the minimax approach, can be found in Refs. [22,23]. A comprehensive review of the various robust DFT/STFT transformations can be found in Ref. [24].

9.3.2 REALIZATION OF THE ROBUST STFT

9.3.2.1 Iterative procedure

The expression (9.3.10) includes $F_z(t, f)$ on the right-hand side. Therefore, to get the robust STFT, we have to solve a nonlinear equation of the form $x = f(x)$. Here, we will use the fixed point iterative algorithm $x_i = f(x_{i-1})$, with the stopping rule $|x_i - x_{i-1}|/|x_i| < \eta$, where η defines the solution precision. This procedure, applied to Eq. (9.3.10), can be summarized as follows.

Step (0): Calculate the standard STFT (9.3.5): $F_z^{(0)}(t, f) = F_z(t, f)$, and $i = 0$.

Step (i): Set $i := i+1$. Calculate $d^{(i)}(t, f, n)$ for $F_z^{(i-1)}(t, f)$ determined from Eq. (9.3.11) or Eq. (9.3.16). Calculate $F_z^{(i)}(t, f)$ as

$$F_z^{(i)}(t, f) = \frac{1}{\sum_{n=-N/2}^{N/2-1} d^{(i)}(t, f, n)} \sum_{n=-N/2}^{N/2-1} d^{(i)}(t, f, n) z(t + n\Delta t) e^{-j2\pi fn\Delta t}. \quad (9.3.18)$$

Step (ii): If the magnitude of the fractional difference between two iterations is η or less, i.e., if

$$\frac{|F_z^{(i)}(t, f) - F_z^{(i-1)}(t, f)|}{|F_z^{(i)}(t, f)|} \leq \eta, \quad (9.3.19)$$

then the robust STFT is obtained as $F_z(t, f) = F_z^{(i)}(t, f)$.

9.3.2.2 Vector filter approach

Note that the standard STFT (9.3.5) can be treated as an estimate of the mean, calculated over a set of complex-valued observations:

$$\mathbf{E}^{(t,f)} = \left\{ z(t + n\Delta t) e^{-j2\pi fn\Delta t} : n \in [-N/2, N/2] \right\}. \quad (9.3.20)$$

If we restrict values of m in Eq. (9.3.1) to the set $\mathbf{E}^{(t,f)}$, the vector filter concept [25–27] can be applied to get a simple approximation of the robust estimate of the STFT. Here, the coordinates of the vector-valued variable are the real and imaginary parts of $z(t + n\Delta t) e^{-j2\pi fn\Delta t}$. The vector estimate of the STFT is defined as $F_z(t, f) = m$, where $m \in \mathbf{E}^{(t,f)}$, and for all $k \in [-N/2, N/2]$ the following inequality holds:

$$\sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \mathbf{F} \left(|m - z(t + n\Delta t) e^{-j2\pi fn\Delta t}| \right) \leq \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \mathbf{F} \left(|z(t + k\Delta t) e^{-j2\pi fk\Delta t} - z(t + n\Delta t) e^{-j2\pi fn\Delta t}| \right). \quad (9.3.21)$$

For $\mathbf{F}(e) = |e|$, this estimate is called the vector median.

The marginal median can be used for independent estimation of real and imaginary parts of $F_z(t, f)$. It results in

$$\begin{aligned} \text{Re}\{F_z(t, f)\} &= \text{median} \left\{ \text{Re}\{z(t + n\Delta t) e^{-j2\pi fn\Delta t}\} : n \in [-N/2, N/2] \right\}, \\ \text{Im}\{F_z(t, f)\} &= \text{median} \left\{ \text{Im}\{z(t + n\Delta t) e^{-j2\pi fn\Delta t}\} : n \in [-N/2, N/2] \right\}. \end{aligned} \quad (9.3.22)$$

The separate estimation of the real and imaginary parts of $F_z(t, f)$ assumes independence of the real and imaginary parts of $z(t + n\Delta t) e^{-j2\pi fn\Delta t}$, which in general does not hold here. However, in numerous experiments the accuracies of the median estimates (9.3.21) and (9.3.22) are of the same order. A simplicity of calculation is the advantage of these median estimates over the iterative procedures.

9.3.3 ROBUST WIGNER-VILLE DISTRIBUTION

The standard (windowed) Wigner-Ville distribution (WVD) of a discrete-time signal is defined as

$$W_z(t, f) = \frac{1}{a_w} \sum_{n=-N/2}^{N/2} w(n\Delta t) z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t}, \quad (9.3.23)$$

with the normalization factor

$$a_w = \sum_{n=-N/2}^{N/2} w(n\Delta t) \quad (9.3.24)$$

(cf. Section 6.1). It can be interpreted as a solution of the problem

$$W_z(t, f) = \arg \left\{ \min_m J(t, f, m) \right\}, \quad (9.3.25)$$

$$J(t, f, m) = \sum_{n=-N/2}^{N/2} w(n\Delta t) \mathbf{F} \left(\left| z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t} - m \right| \right), \quad (9.3.26)$$

where $\mathbf{F}(e) = |e|^2$. For the loss function $\mathbf{F}(e) = |e|$, the solution of Eqs. (9.3.25) and (9.3.26) is a WVD robust to impulse noise. It can be obtained as a solution of the nonlinear equation [28]

$$W_z(t, f) = \frac{1}{a_{we}(t, f)} \sum_{n=-N/2}^{N/2} d(t, f, n) z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t}, \quad (9.3.27)$$

with

$$d(t, f, n) = \frac{w(n\Delta t)}{\left| z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t} - W_z(t, f) \right|}, \quad (9.3.28)$$

$$a_{we}(t, f) = \sum_{n=-N/2}^{N/2} d(t, f, n).$$

An iterative procedure similar to the one described for the robust STFT can be used to find $W_z(t, f)$ from Eqs. (9.3.27) and (9.3.28).

9.3.3.1 Properties of the robust WVD

(1) The robust WVD is real-valued for a real and symmetric window function:

$$\begin{aligned} W_z^*(t, f) &= \frac{1}{a_{we}^*(t, f)} \sum_{n=-N/2}^{N/2} \frac{w^*(n\Delta t) z^*(t + n\Delta t) z(t - n\Delta t) e^{j4\pi fn\Delta t}}{\left| z^*(t + n\Delta t) z(t - n\Delta t) e^{j4\pi fn\Delta t} - W_z^*(t, f) \right|} \\ &= \frac{1}{a_{we}(t, f)} \sum_{n=-N/2}^{N/2} \frac{w^*(-n\Delta t) z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t}}{\left| z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi fn\Delta t} - W_z(t, f) \right|} = W_z(t, f). \end{aligned} \quad (9.3.29)$$

(2) TF invariance: If $y(t) = z(t - t_0) e^{j2\pi f_0 t}$, then $W_y(t, f) = W_z(t - t_0, f - f_0)$.

- (3) For linear FM signals $z(t) = \exp(jat^2/2 + jbt)$, when $w(n\Delta t)$ is a very wide window, the WVD is an almost ideally concentrated TF distribution.

9.3.3.2 Median WVD

For a rectangular window, the standard WVD can be treated as an estimate of the mean, calculated over a set of complex-valued observations

$$\mathbf{G} = \{z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi f n \Delta t} : n \in [-N/2, N/2]\}, \quad (9.3.30)$$

i.e.,

$$W_z(t, f) = \frac{1}{N+1} \sum_{n=-N/2}^{N/2} z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi f n \Delta t}. \quad (9.3.31)$$

From Eq. (9.3.29) follows that the robust WVD is real-valued; thus the minimization of $J(t, f, m)$ can be done with respect to the real part of $z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi f n \Delta t}$ only. A form of the robust WVD, the median WVD, can be introduced as

$$W_z(t, f) = \text{median} \left\{ \text{Re}\{z(t + n\Delta t) z^*(t - n\Delta t) e^{-j4\pi f n \Delta t}\} : n \in [-N/2, N/2] \right\}. \quad (9.3.32)$$

Generally, it can be shown that any robust TF distribution, obtained by using the Hermitian smoothed instantaneous autocorrelation function (IAF) $R_z(t, n\Delta t) = R_z^*(t, -n\Delta t)$ in the minimization, is real-valued. In the WVD case, this condition is satisfied, since $R_z(t, n\Delta t) = z(t + n\Delta t) z^*(t - n\Delta t)$. For a general time- and frequency-covariant quadratic TFD with a Hermitian smoothed IAF, the proposed robust version reads

$$\rho_z(t, f) = \text{median} \left\{ \text{Re}\{R_z(t, n\Delta t) e^{-j4\pi f n \Delta t}\} : n \in [-N/2, N/2] \right\}, \quad (9.3.33)$$

where $R_z(t, n\Delta t)$ includes the kernel in time-lag domain [cf. Eq. (3.2.7), p. 112].

Note that for a Gaussian noise input the resulting noise in the WVD has both Gaussian and impulse component, due to the WVD's quadratic nature. Thus, as it is shown in Ref. [29], robust WVD forms can improve performance of the standard WVD, even in an environment with high Gaussian input noise. Design of robust spectral analysis tools with high breakdown points (more robust to impulsive noise influence) is presented in Ref. [30]. The proposed approach for defining Robust TFDs can be extended for complex lag TFDs as well to obtain highly concentrated TFDs for nonlinear FM signals in high noise [31].

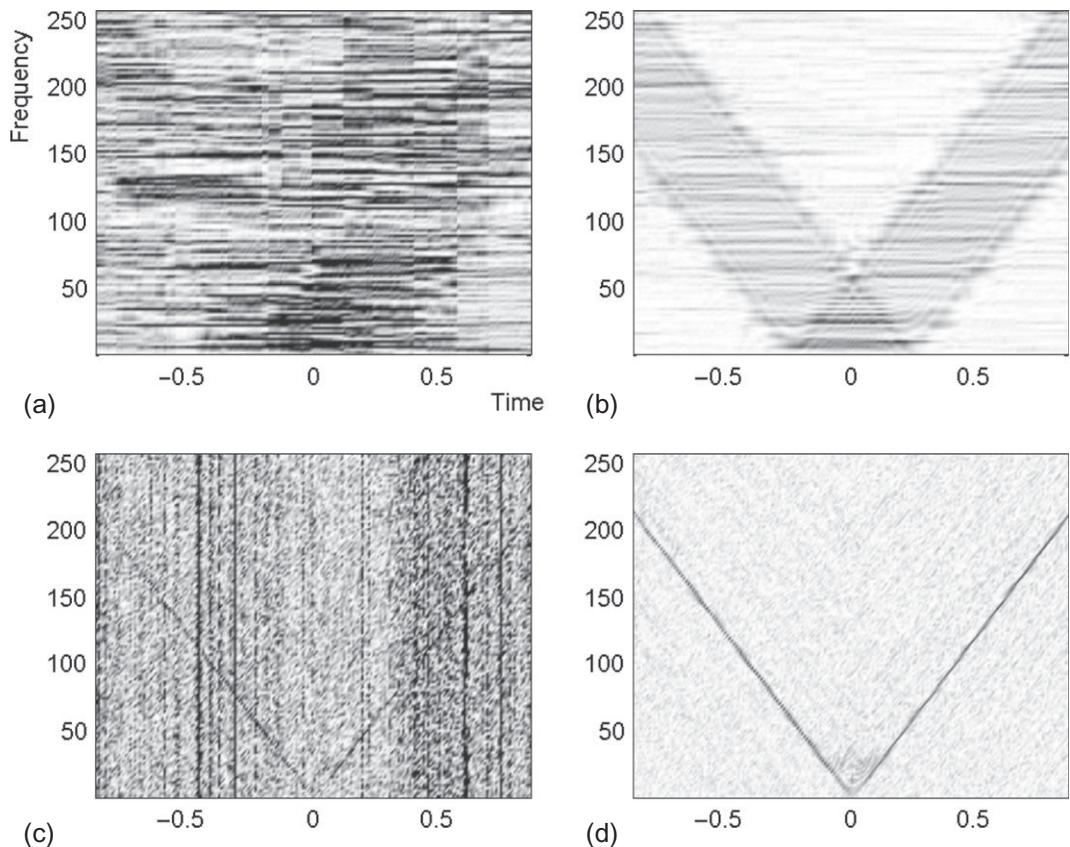
9.3.4 EXAMPLE

Consider the nonstationary FM signal

$$z(t) = \exp(j 204.8 \pi t |t|), \quad (9.3.34)$$

corrupted with a high amount of the heavy-tailed noise:

$$\varepsilon(t) = 0.5 \left[\varepsilon_1^3(t) + j \varepsilon_2^3(t) \right], \quad (9.3.35)$$

**FIGURE 9.3.1**

TF representations of signal corrupted by impulse noise: (a) Standard spectrogram (b) Robust spectrogram (c) Standard WVD (d) Robust WVD.

where $\varepsilon_i(t)$, $i = 1, 2$ are mutually independent Gaussian white noises $\mathcal{N}(0, 1)$. We consider the interval $t \in [-7/8, 7/8]$ with a sampling rate $\Delta t = 1/512$ for spectrograms, and $\Delta t = 1/1024$ for WVDs. The rectangular window width is $N = 256$ in all cases. The standard spectrogram and the WVD [Fig. 9.3.1(a),(c)] are calculated according to Eqs. (9.3.5) and (9.3.23). The robust spectrogram [Fig. 9.3.1(b)] is calculated using the iterative procedure (9.3.18) and (9.3.19). In this case, similar results would be produced by residual spectrogram (9.3.13) and (9.3.17), vector median (9.3.21), and marginal median (9.3.22). The robust WVD [Fig. 9.3.1(d)] is calculated by using expression (9.3.32) for the considered TF point. It can be concluded from Fig. 9.3.1 that the robust spectrogram and the robust WVD filter the heavy-tailed noise significantly better than the standard spectrogram and the standard WVD. Note that the standard and the robust WVD exhibit higher TF resolution in comparison with the corresponding spectrograms.

9.3.5 SUMMARY AND CONCLUSIONS

The TF distributions are defined within the Huber robust statistics framework. The loss function $F(e) = |e|$ gives distributions robust to the impulse noise influence. They can be realized by using the iterative procedures, the vector median, or the marginal median approach. All calculation procedures produce accuracy of the same order of magnitude. Some other robust forms of signal transforms are presented in Refs. [32–34]. Sections 6.1 and 6.5 further discuss the definition and computation of discrete-time TFDs.

9.4 TIME-VARYING POWER SPECTRA OF NONSTATIONARY RANDOM PROCESSES⁰

9.4.1 NONSTATIONARY RANDOM PROCESSES

We consider zero-mean nonstationary random processes that are real or circular complex. The second-order statistics of a nonstationary random process $x(t)$ are characterized by the correlation function $r_x(t, t') = E\{x(t)x^*(t')\}$ (with $E\{\cdot\}$ denoting expectation). For a (wide-sense) *stationary* random process, the correlation function is of the form $r_x(t, t') = \tilde{r}_x(t - t')$ and the Fourier transform (FT) of $\tilde{r}_x(\tau)$,

$$\mathcal{S}_x(f) = \int_{-\infty}^{\infty} \tilde{r}_x(\tau) e^{-j2\pi f\tau} d\tau \geq 0, \quad (9.4.1)$$

is known as the *power spectral density* (PSD) (see Eq. (I.1.7) of Chapter I). The PSD describes the distribution of the process' mean power over frequency f . The time-frequency (briefly (t, f)) dual of stationary processes is given by *white* processes with correlation functions of the form $r_x(t, t') = q_x(t)\delta(t - t')$. Here, the *mean instantaneous intensity* $q_x(t) \geq 0$ is the (t, f) dual of the PSD.

Possible extensions of the PSD to the nonstationary case are discussed in Chapter 2, Section 2.2; they are related to the *generalized Wigner-Ville spectrum* [35–40] and the *generalized evolutionary spectrum* [40,41]. This section briefly discusses these “time-varying power spectra” and shows that they yield satisfactory descriptions for the important class of *underspread* nonstationary processes.

9.4.2 THE GENERALIZED WIGNER-VILLE SPECTRUM

The *generalized Wigner-Ville spectrum* (GWVS) of a nonstationary process $x(t)$ is defined as [35–40]

$$\overline{W}_x^{(\alpha)}(t, f) \triangleq \int_{-\infty}^{\infty} r_x^{(\alpha)}(t, \tau) e^{-j2\pi f\tau} d\tau \quad (9.4.2)$$

⁰Authors: **G. Matz**, Vienna University of Technology, Vienna, Austria (g.matz@ieee.org) and **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at). Reviewers: M.G. Amin and A. Papandreou-Suppappola.

with

$$r_x^{(\alpha)}(t, \tau) \triangleq r_x\left(t + \left(\frac{1}{2} - \alpha\right)\tau, t - \left(\frac{1}{2} + \alpha\right)\tau\right), \quad (9.4.3)$$

where α is a real-valued parameter. The GWVS equals the generalized Weyl symbol (see [Section 4.7](#)) of the correlation operator \mathbf{R}_x (the linear operator whose kernel is the correlation function $r_x(t, t') = E\{x(t)x^*(t')\}$) and, under mild assumptions, it equals the expectation of the generalized Wigner distribution [[37](#)] of $x(t)$. For $\alpha = 0$, the GWVS becomes the ordinary *Wigner-Ville spectrum*, and for $\alpha = 1/2$ it reduces to the *Rihaczek spectrum* [[35–40,42](#)]:

$$\begin{aligned}\overline{W}_x^{(0)}(t, f) &= \int_{-\infty}^{\infty} r_x\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau, \\ \overline{W}_x^{(1/2)}(t, f) &= \int_{-\infty}^{\infty} r_x(t, t - \tau) e^{-j2\pi f\tau} d\tau.\end{aligned}$$

The GWVS $\overline{W}_x^{(\alpha)}(t, f)$ is a complete characterization of the second-order statistics of $x(t)$ since the correlation function $r_x(t, t')$ can be recovered from it. Integration of the GWVS gives the marginal properties

$$\int_{-\infty}^{\infty} \overline{W}_x^{(\alpha)}(t, f) dt = E\{|X(f)|^2\}, \quad \int_{-\infty}^{\infty} \overline{W}_x^{(\alpha)}(t, f) df = E\{|x(t)|^2\},$$

provided that the expectations on the right-hand sides exist. In this sense, the GWVS can be considered as a (t, f) distribution of the mean energy of $x(t)$. However, in general the GWVS is not real-valued; for $\alpha = 0$, $\overline{W}_x^{(0)}(t, f)$ is real-valued though possibly not everywhere nonnegative. Other properties of the GWVS are considered in Refs. [[35–40,42](#)]. We next discuss the GWVS of three fundamental types of processes.

- The GWVS of a stationary process with correlation function $r_x(t, t') = \tilde{r}_x(t - t')$ reduces to the PSD $S_x(f)$ for all t , i.e., $\overline{W}_x^{(\alpha)}(t, f) \equiv S_x(f)$.
- The GWVS of a (generally nonstationary) white process with correlation function $r_x(t, t') = q_x(t)\delta(t - t')$ reduces to the mean instantaneous intensity $q_x(t)$ for all f , i.e., $\overline{W}_x^{(\alpha)}(t, f) \equiv q_x(t)$.
- The GWVS of a stationary white process with correlation function $r_x(t, t') = \eta \delta(t - t')$ is given by $\overline{W}_x^{(\alpha)}(t, f) \equiv \eta$ (i.e., constant mean energy distribution over the entire (t, f) plane).

These results show that the GWVS is consistent with the PSD of stationary processes and the mean instantaneous intensity of white processes.

Next, we consider an alternative definition of “time-varying power spectrum.”

9.4.3 THE GENERALIZED EVOLUTIONARY SPECTRUM

The PSD of a stationary random process $x(t)$ can alternatively be defined using an innovations system representation. Here, $x(t)$ is viewed as the output of a linear, time-invariant system \mathbf{H} with impulse response $h(\tau)$ (innovations system) that is driven by stationary white noise $n(t)$ with PSD $S_n(f) \equiv 1$,

i.e., $x(t) = (\mathbf{H}n)(t) = \int_{-\infty}^{\infty} h(\tau) n(t - \tau) d\tau$. The PSD of $x(t)$ can then be written as $\mathcal{S}_x(f) = |H(f)|^2$, where $H(f) = \int_{-\infty}^{\infty} h(\tau) e^{-j2\pi f\tau} d\tau$ is the transfer function (frequency response) of \mathbf{H} .

A similar innovations system representation is also possible in the nonstationary case. The innovations system \mathbf{H} of a nonstationary random process $x(t)$ is a linear, *time-varying* system defined by $\mathbf{H}\mathbf{H}^+ = \mathbf{R}_x$ (here, the superscript $+$ denotes the adjoint). Note that \mathbf{H} is not uniquely defined; indeed, all innovation systems can be written as $\mathbf{H} = \mathbf{H}_p \mathbf{U}$ where \mathbf{H}_p is the *positive (semi-) definite* innovations system (which is unique) and \mathbf{U} is a linear operator satisfying $\mathbf{U}\mathbf{U}^+ = \mathbf{I}$ [41].

In analogy to the PSD expression $\mathcal{S}_x(f) = |H(f)|^2$, the *generalized evolutionary spectrum* (GES) of a nonstationary process $x(t)$ can be defined as [40,41]

$$G_x^{(\alpha)}(t,f) \triangleq |L_{\mathbf{H}}^{(\alpha)}(t,f)|^2. \quad (9.4.4)$$

Here, $L_{\mathbf{H}}^{(\alpha)}(t,f)$ is the generalized Weyl symbol (see Section 4.7) of an innovations system \mathbf{H} of $x(t)$, i.e.,

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h\left(t + \left(\frac{1}{2} - \alpha\right)\tau, t - \left(\frac{1}{2} + \alpha\right)\tau\right) e^{-j2\pi f\tau} d\tau, \quad (9.4.5)$$

where $h(t,t')$ is the kernel of \mathbf{H} . Note that the nonuniqueness of \mathbf{H} implies a corresponding nonuniqueness of the GES. For $\alpha = 1/2$, $\alpha = -1/2$, and $\alpha = 0$, the GES reduces to the ordinary *evolutionary spectrum*¹ [43,44], the *transitory evolutionary spectrum* [41,45], and the *Weyl spectrum* [41], respectively.

In contrast to the GWVS, the GES is a nonnegative real-valued function. However, it is not a complete second-order description of $x(t)$ since in general the correlation function $r_x(t,t')$ cannot be recovered from it. For $\alpha = \pm 1/2$ and a normal innovations system (i.e., \mathbf{H} satisfies $\mathbf{H}\mathbf{H}^+ = \mathbf{H}^+\mathbf{H}$; note, in particular, that \mathbf{H}_p is always normal), the GES satisfies the marginal properties, i.e.,

$$\int_{-\infty}^{\infty} G_x^{(\pm 1/2)}(t,f) dt = E\{|X(f)|^2\}, \quad \int_{-\infty}^{\infty} G_x^{(\pm 1/2)}(t,f) df = E\{|x(t)|^2\}.$$

Next, we consider the GES of our three fundamental types of processes, using the positive (semi-) definite innovations system \mathbf{H}_p in the GES definition (9.4.4).

- For a stationary process with PSD $\mathcal{S}_x(f)$, \mathbf{H}_p is time-invariant with frequency response $H_p(f) = \sqrt{\mathcal{S}_x(f)}$. Here, the GES reduces to the PSD $\mathcal{S}_x(f)$ for all t , i.e., $G_x^{(\alpha)}(t,f) \equiv \mathcal{S}_x(f)$.
- For a nonstationary white process with mean instantaneous intensity $q_x(t)$, \mathbf{H}_p is “frequency-invariant” with kernel $h_p(t,t') = \sqrt{q_x(t)} \delta(t-t')$. The GES here reduces to $q_x(t)$ for all f , i.e., $G_x^{(\alpha)}(t,f) \equiv q_x(t)$.
- For a stationary white process with correlation function $r_x(t,t') = \eta \delta(t-t')$, $\mathbf{H}_p = \sqrt{\eta} \mathbf{I}$ (\mathbf{I} is the identity operator), and the GES is given by $G_x^{(\alpha)}(t,f) \equiv \eta$.

Section 9.4.5 considers conditions allowing the interpretation of the GWVS and GES as a “time-varying power spectrum.” The formulation of these conditions is based on a further (t,f) representation of nonstationary processes, discussed next.

¹Note that Priestley’s original definition of the evolutionary spectrum was based on a conceptually different approach using “oscillatory processes” [43,44].

9.4.4 THE GENERALIZED EXPECTED AMBIGUITY FUNCTION

The *generalized expected ambiguity function* (GEAF) is defined as [39–41]

$$\bar{A}_x^{(\alpha)}(\nu, \tau) \triangleq \int_{-\infty}^{\infty} r_x^{(\alpha)}(t, \tau) e^{-j2\pi\nu t} dt,$$

with $r_x^{(\alpha)}(t, \tau)$ as in Eq. (9.4.3). The interpretation of the GEAF is quite different from that of a “time-varying power spectrum”: For a given frequency lag ν and a given time lag τ , the GEAF $\bar{A}_x^{(\alpha)}(\nu, \tau)$ quantifies the statistical correlations of all process components separated in frequency by ν and in time by τ [39]. Hence, the extension of $\bar{A}_x^{(\alpha)}(\nu, \tau)$ about $(0, 0)$ indicates the amount of “(t, f) correlations” of $x(t)$. In particular, if $\bar{A}_x^{(\alpha)}(\nu, \tau)$ extends far in the ν direction, $x(t)$ has a large spectral correlation width (i.e., $x(t)$ is highly nonstationary), and if $\bar{A}_x^{(\alpha)}(\nu, \tau)$ extends far in the τ direction, $x(t)$ has a large temporal correlation width.

The GEAF equals the generalized spreading function (see Section 4.7) of the correlation operator \mathbf{R}_x . Like the GWVS, the GEAF is a complete second-order statistic. GEAFs with different α values differ merely by a phase factor, i.e.,

$$\bar{A}_x^{(\alpha_2)}(\nu, \tau) = \bar{A}_x^{(\alpha_1)}(\nu, \tau) e^{j2\pi(\alpha_1 - \alpha_2)\nu\tau}.$$

The GEAF magnitude is thus independent of α , $|\bar{A}_x^{(\alpha_1)}(\nu, \tau)| = |\bar{A}_x^{(\alpha_2)}(\nu, \tau)|$, so we can simply write $|\bar{A}_x(\tau, \nu)|$. GWVS and GEAF are related by a 2-D FT,

$$\bar{W}_x^{(\alpha)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{A}_x^{(\alpha)}(\nu, \tau) e^{-j2\pi(\tau f - \nu t)} d\nu d\tau; \quad (9.4.6)$$

this extends the Wiener-Khintchine theorem (9.4.1) to the nonstationary case.

Let us again consider our three process types (see Fig. 9.4.1, cf. Fig. 4.7.1):

- The GEAF of a stationary process $x(t)$ with correlation function $r_x(t, t') = \tilde{r}_x(t - t')$ is given by $\bar{A}_x^{(\alpha)}(\nu, \tau) = \delta(\nu) \tilde{r}_x(\tau)$ (i.e., only temporal correlations, which are characterized by $\tilde{r}_x(\tau)$).
- The GEAF of a (generally nonstationary) white process $x(t)$ with correlation function $r_x(t, t') = q_x(t) \delta(t - t')$ is given by $\bar{A}_x^{(\alpha)}(\nu, \tau) = Q_x(\nu) \delta(\tau)$, where $Q_x(\nu)$ is the FT of the mean instantaneous intensity $q_x(t)$ (i.e., only spectral correlations, which are characterized by $Q_x(\nu)$).

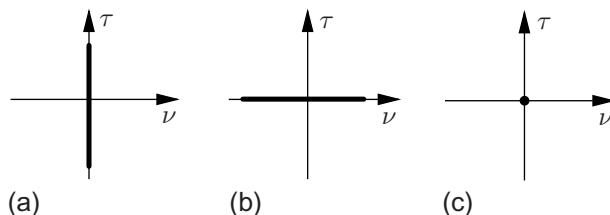


FIGURE 9.4.1

Schematic representation of the GEAF magnitude of some (classes of) random processes: (a) stationary process, (b) white process, (c) stationary white process.

- The GEAF of a stationary white process $x(t)$ with correlation function $r_x(t, t') = \eta \delta(t - t')$ is given by $\bar{A}_x^{(\alpha)}(\nu, \tau) = \eta \delta(\nu)\delta(\tau)$ (i.e., neither temporal nor spectral correlations).

9.4.5 UNDERSpread PROCESSES

A nonstationary random process is called *underspread* if its GEAF is well concentrated about the origin of the (ν, τ) plane, thus implying a small “ (t, f) correlation width.” In contrast, a process with large (t, f) correlation width is termed *overspread*. We will see in [Section 9.4.6](#) that the GWVS and GES of an underspread process are approximately equivalent and can be interpreted as “time-varying power spectra.”

There are two alternative mathematical definitions of underspread processes [39,40]. The first one [39] assumes that the GEAF $\bar{A}_x^{(\alpha)}(\nu, \tau)$ is supported in a compact region \mathcal{G}_x about the origin of the (ν, τ) plane, i.e., $|\bar{A}_x(\nu, \tau)| = 0$ for $(\nu, \tau) \notin \mathcal{G}_x$. Let $\nu_x \triangleq \max_{(\nu, \tau) \in \mathcal{G}_x} |\nu|$ and $\tau_x \triangleq \max_{(\nu, \tau) \in \mathcal{G}_x} |\tau|$ denote the maximum frequency lag and time lag, respectively, for which the process $x(t)$ features (t, f) correlations. The (t, f) correlation spread of $x(t)$ is defined as $\sigma_x \triangleq 4\nu_x\tau_x$, which is the area of the rectangle $[-\nu_x, \nu_x] \times [-\tau_x, \tau_x]$ enclosing \mathcal{G}_x . The process $x(t)$ is considered underspread if $\sigma_x \ll 1$ [39]. An alternative description of the GEAF’s extension that avoids the assumption of compact GEAF support uses the normalized weighted GEAF integrals, whose definition is analogous to Eq. (4.7.5) [40]

$$m_x^{(\phi)} \triangleq \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\nu, \tau) |\bar{A}_x(\nu, \tau)| d\nu d\tau}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\bar{A}_x(\nu, \tau)| d\nu d\tau}. \quad (9.4.7)$$

The discussion and interpretation of the above equation are similar to what follows Eq. (4.7.5), page 189. In particular, the GEAF moments are defined as $m_x^{(k,l)} \triangleq m_x^{(\phi_{k,l})}$ with weighting function $\phi_{k,l}(\nu, \tau) = |\nu|^k |\tau|^l$, where $k, l \in \mathbb{N}_0$. A random process $x(t)$ can now be considered underspread if suitable weighted GEAF intervals or moments are “small.”

Examples of underspread processes are illustrated in [Fig. 9.4.2](#) which is analogous to [Fig. 4.7.2](#) in [Section 4.7.4](#). The concept of underspread processes is not equivalent to that of quasi-stationary processes: e.g., a quasi-stationary process may be overspread if its temporal correlation width is very large (cf. the similar discussion following Eq. (4.7.5) on page 189).

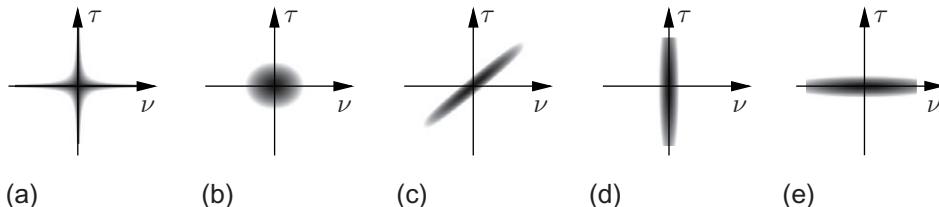


FIGURE 9.4.2

Schematic representation of the GEAF magnitude of various types of nonstationary processes: (a) underspread process with small $m_x^{(1,1)}$; (b) underspread process with small $m_x^{(1,0)} m_x^{(0,1)}$; (c) “chirpy” underspread process [40]; (d) quasi-stationary process (small $m_x^{(0,1)}$); (e) quasi-white process (small $m_x^{(1,0)}$).

9.4.6 TIME-VARYING SPECTRAL ANALYSIS OF UNDERSpread PROCESSES

For *underspread* nonstationary processes, the GWVS and GES can be interpreted as “time-varying power spectra” that generalize the PSD of stationary processes and the mean instantaneous intensity of white processes. Indeed, small weighted GEAF integrals $m_x^{(\phi)}$ (or small moments $m_x^{(k,l)}$ or a small (t,f) correlation spread σ_x) ensure the validity of the approximations described in what follows [39–41].

9.4.6.1 Approximate equivalence

In general, the GWVS and GES of a given process $x(t)$ may yield very different results which, moreover, may strongly depend on the parameter α used. However, for an underspread process $x(t)$, we have $\bar{W}_x^{(\alpha_1)}(t,f) \approx \bar{W}_x^{(\alpha_2)}(t,f)$ and $G_x^{(\alpha_1)}(t,f) \approx G_x^{(\alpha_2)}(t,f)$, as well as $\bar{W}_x^{(\alpha_1)}(t,f) \approx G_x^{(\alpha_2)}(t,f)$. Indeed, the approximation error $\bar{W}_x^{(\alpha_1)}(t,f) - \bar{W}_x^{(\alpha_2)}(t,f)$ is upper bounded as [40]

$$|\bar{W}_x^{(\alpha_1)}(t,f) - \bar{W}_x^{(\alpha_2)}(t,f)| \leq 2\pi |\alpha_1 - \alpha_2| \|\bar{A}_x\|_1 m_x^{(1,1)}, \quad (9.4.8)$$

with $\|\bar{A}_x\|_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\bar{A}_x(v,\tau)| dv d\tau$. Thus, for an underspread system where $m_x^{(1,1)}$ is small, $\bar{W}_x^{(\alpha_1)}(t,f)$ and $\bar{W}_x^{(\alpha_2)}(t,f)$ will be approximately equal as long as $|\alpha_1 - \alpha_2|$ is not too large. Similar bounds can be developed for the approximation errors $G_x^{(\alpha_1)}(t,f) - G_x^{(\alpha_2)}(t,f)$ and $\bar{W}_x^{(\alpha_1)}(t,f) - G_x^{(\alpha_2)}(t,f)$ [40].

We can conclude from these results that for an underspread process, the choice of a specific spectrum is not critical. An example is shown in Fig. 9.4.3. For this example, the maximum normalized differences between the spectra shown are all around 0.03 (e.g., $\max_{t,f} |\bar{W}_x^{(0)}(t,f) - G_x^{(0)}(t,f)| / \max_{t,f} |\bar{W}_x^{(0)}(t,f)| = 0.029$). A counterexample involving an *overspread* process is shown in Fig. 9.4.4, where the results obtained with the various spectra are dramatically different; indeed, the maximum normalized differences range from 1 to 8.5 (e.g., $\max_{t,f} |\bar{W}_x^{(1/2)}(t,f) - G_x^{(1/2)}(t,f)| / \max_{t,f} |\bar{W}_x^{(1/2)}(t,f)| = 2.13$). All spectra are seen to contain oscillating components (so-called statistical cross-terms) which are indicative of (t,f) correlations [40]. Such statistical

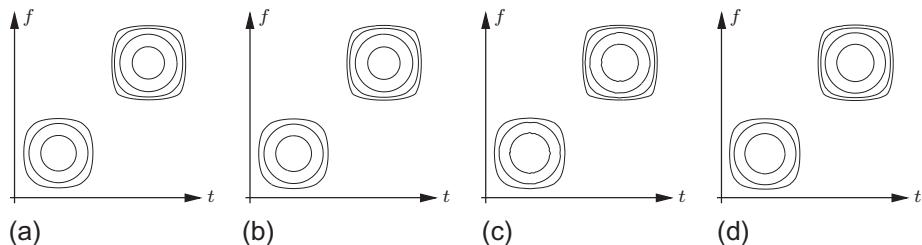
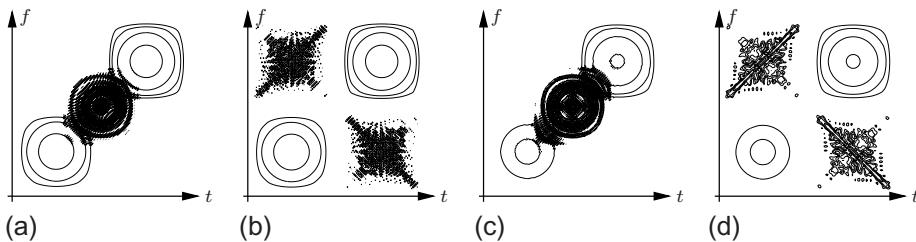


FIGURE 9.4.3

GWVS and GES of an underspread process $x(t)$: (a) Wigner-Ville spectrum $\bar{W}_x^{(0)}(t,f)$, (b) real part of Rihaczek spectrum $\bar{W}_x^{(1/2)}(t,f)$, (c) Weyl spectrum $G_x^{(0)}(t,f)$, (d) evolutionary spectrum $G_x^{(1/2)}(t,f)$. In this and subsequent simulations, signal duration is 256 samples and normalized frequency ranges from $-1/4$ to $1/4$.

**FIGURE 9.4.4**

GWVS and GES of an overspread process $x(t)$: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t, f)$, (b) real part of Rihaczek spectrum $\overline{W}_x^{(1/2)}(t, f)$, (c) Weyl spectrum $G_x^{(0)}(t, f)$, (d) evolutionary spectrum $G_x^{(1/2)}(t, f)$.

cross-terms are reduced in extensions of the GWVS and GES that contain a (t, f) smoothing [35–38,40]. For underspread processes, on the other hand, a (t, f) smoothing does not cause a big difference.

9.4.6.2 Approximate real-valuedness and positivity of the GWVS

The PSD of stationary processes and the mean instantaneous intensity of white processes are real-valued and nonnegative. This is also true for the GES of arbitrary processes. In contrast, the GWVS is real-valued only for $\alpha = 0$ and generally not nonnegative. In the case of underspread processes, however, it can be shown [40] that the imaginary part of the GWVS is approximately zero and the real part of the GWVS is approximately nonnegative, i.e., $\text{Im}\left\{\overline{W}_x^{(\alpha)}(t, f)\right\} \approx 0$ and $\text{Re}\left\{\overline{W}_x^{(\alpha)}(t, f)\right\} \gtrsim 0$. Upper bounds on the approximation errors (similar to Eq. (9.4.8)) can be found in Ref. [40]. For example, for the underspread process of Fig. 9.4.3, the normalized maximum of the imaginary part of the Rihaczek spectrum (the real part shown in Fig. 9.4.3(b)) is $\max_{t,f} |\text{Im}\{\overline{W}_x^{(1/2)}(t, f)\}| / \max_{t,f} |\overline{W}_x^{(1/2)}(t, f)| = 0.024$, and the normalized maximum of the negative real part is $\max_{t,f} \{-\text{Re}\{\overline{W}_x^{(1/2)}(t, f)\}\} / \max_{t,f} |\overline{W}_x^{(1/2)}(t, f)| = 0.006$.

9.4.6.3 Approximate input-output relations

If a stationary process $x(t)$ with PSD $S_x(f)$ is passed through a time-invariant linear system with impulse response $k(\tau)$ and transfer function $K(f)$, the output $y(t) = (x * k)(t)$ is also stationary and its PSD equals $S_y(f) = |K(f)|^2 S_x(f)$. Similarly, the response $y(t) = w(t)x(t)$ of a linear frequency-invariant system (see Section 4.7) to a white process $x(t)$ with mean instantaneous intensity $q_x(t)$ is again white with $q_y(t) = |w(t)|^2 q_x(t)$. A similar input-output relation does not exist for a general nonstationary process $x(t)$ that is passed through a general time-varying linear system \mathbf{K} . However, for an underspread process that is passed through an underspread system (i.e., a time-varying linear system introducing only small (t, f) shifts, see Section 4.7), one can show the following approximate input-output relations of the GWVS and GES:

$$\begin{aligned}\overline{W}_y^{(\alpha)}(t, f) &\approx |L_{\mathbf{K}}^{(\alpha)}(t, f)|^2 \overline{W}_x^{(\alpha)}(t, f), \\ G_y^{(\alpha)}(t, f) &\approx |L_{\mathbf{K}}^{(\alpha)}(t, f)|^2 G_x^{(\alpha)}(t, f),\end{aligned}$$

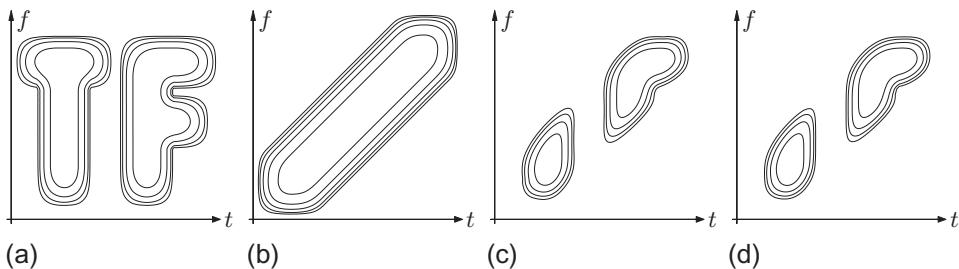


FIGURE 9.4.5

Approximate input-output relation for the GWVS: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t, f)$ of input process $x(t)$, (b) Weyl symbol $L_{\mathbf{K}}^{(0)}(t, f)$ of LTV system \mathbf{K} , (c) Wigner-Ville spectrum $\overline{W}_y^{(0)}(t, f)$ of filtered process $y(t) = (\mathbf{K}x)(t)$, (d) approximation $|L_{\mathbf{K}}^{(0)}(t, f)|^2 \overline{W}_x^{(0)}(t, f)$.

with $y(t) = (\mathbf{K}x)(t)$. Note that the generalized Weyl symbol $L_{\mathbf{K}}^{(\alpha)}(t, f)$ of \mathbf{K} (see Eq. (9.4.5)) takes the place of the transfer function $K(f)$ or $w(t)$. An example for the Wigner-Ville spectrum (GWVS with $\alpha = 0$) is shown in Fig. 9.4.5. In this example, the normalized maximum approximation error is $\max_{t,f} \left\{ \left| \overline{W}_y^{(0)}(t, f) - |L_{\mathbf{K}}^{(0)}(t, f)|^2 \overline{W}_x^{(0)}(t, f) \right| \right\} / \max_{t,f} |\overline{W}_y^{(0)}(t, f)| = 0.017$.

9.4.6.4 Discussion

The above approximations (see also Refs. [40–42]) corroborate the interpretation of the GWVS and GES of underspread processes as *time-varying power spectra*. A mathematical underpinning of these approximations is provided by explicit upper bounds on the associated approximation errors [40]; these bounds involve the GEAF parameters $m_x^{(\phi)}, m_x^{(k,l)}$ or σ_x defined in Section 9.4.5. In the underspread case, these GEAF parameters are small and thus the approximations are guaranteed to be good. On the other hand, we caution that the approximations are *not* valid for overspread processes (cf. Fig. 9.4.4).

9.4.7 SUMMARY AND CONCLUSIONS

For the practically important class of underspread processes (i.e., processes with small (t, f) correlations), the generalized Wigner-Ville spectrum and generalized evolutionary spectrum can be interpreted in a meaningful way as time-varying power spectra, as they (approximately) satisfy desirable properties that any reasonable definition of a time-varying power spectrum would be expected to satisfy. Applications of the generalized Wigner-Ville spectrum are considered in Sections 12.1 and 12.4. A more detailed discussion of time-varying power spectra (also addressing their estimation) is provided in Ref. [46].

9.5 TIME-FREQUENCY CHARACTERIZATION OF RANDOM TIME-VARYING CHANNELS⁰

9.5.1 TIME-VARYING CHANNELS

In many practical communication systems, the channel is modeled as linear but *time-varying* and *random*. Examples are the mobile radio, ionospheric, tropospheric, and underwater acoustic channels [47–52]. This section discusses time-frequency (briefly (t,f)) descriptions of both the channel (see Section 4.7 for more details) and its second-order statistics.

The input-output relation of a linear, time-varying, random channel \mathbf{H} is

$$r(t) = (\mathbf{H}s)(t) = \int_{-\infty}^{\infty} h(t, t') s(t') dt' = \int_{-\infty}^{\infty} \tilde{h}(t, \tau) s(t - \tau) d\tau, \quad (9.5.1)$$

where $s(t)$ is the transmit signal, $r(t)$ is the received signal, $h(t, t')$ is the (random) kernel of \mathbf{H} , and $\tilde{h}(t, \tau) = h(t, t - \tau)$ is the (random) impulse response of \mathbf{H} . Two major physical phenomena underlying practical channels are *multipath propagation* and *Doppler spreading*. Multipath propagation (i.e., several different propagation paths from the transmitter to the receiver via various scattering objects) causes the received signal to consist of several delayed versions of the transmit signal. Doppler spreading is due to the movement of transmitter and/or receiver and/or scatterers; for a narrowband transmit signal $s(t)$, it causes the multipath signals to be frequency-shifted. The received signal $r(t)$ thus consists of several (t, f) shifted (i.e., delayed and modulated) versions of the transmit signal $s(t)$ [48,49,51,52],

$$r(t) = \sum_{k=1}^N a_k s(t - \tau_k) e^{j2\pi v_k t}.$$

Here, N is the number of scatterers and τ_k , v_k , and a_k are respectively the (random) delay, Doppler frequency, and reflectivity of the k th scatterer. The above relation can be extended to a continuum of scatterers (corresponding to a continuum of delays τ and Doppler frequency shifts v) as

$$r(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_{\mathbf{H}}^{(\alpha)}(v, \tau) s_{v, \tau}^{(\alpha)}(t) d\tau dv, \quad (9.5.2)$$

with $s_{v, \tau}^{(\alpha)}(t) = s(t - \tau) e^{j2\pi vt} e^{j2\pi v\tau(\alpha-1/2)}$, where α is a real-valued parameter that is arbitrary but assumed fixed (see the footnote on page 187 in Section 4.7) and $S_{\mathbf{H}}^{(\alpha)}(v, \tau)$ denotes the *generalized (delay-Doppler) spreading function* (GSF) of the channel [48] (see also Section 4.7, especially the definition of the GSF in Eq. (4.7.2)). It can be shown that the input-output relation (9.5.2) is mathematically equivalent to Eq. (9.5.1).

In what follows, we also use the *generalized Weyl symbol* (GWS) defined in Eq. (9.4.5), which is related to the GSF according to

$$L_{\mathbf{H}}^{(\alpha)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_{\mathbf{H}}^{(\alpha)}(v, \tau) e^{j2\pi(vt - \tau f)} d\tau dv. \quad (9.5.3)$$

⁰Authors: **G. Matz**, Vienna University of Technology, Vienna, Austria (g.matz@ieee.org) and **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at). Reviewers: P. Flandrin and A.M. Sayeed.

The GWS can be interpreted (with certain precautions, see [Section 4.7](#)) as a “ (t,f) transfer function” of \mathbf{H} .

In the mobile communications literature, the parameter α is usually chosen as 1/2. In this case, $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ and $L_{\mathbf{H}}^{(\alpha)}(t, f)$ become

$$S_{\mathbf{H}}^{(1/2)}(\nu, \tau) = \int_{-\infty}^{\infty} \tilde{h}(t, \tau) e^{-j2\pi\nu t} dt, \quad L_{\mathbf{H}}^{(1/2)}(t, f) = \int_{-\infty}^{\infty} \tilde{h}(t, \tau) e^{-j2\pi f \tau} d\tau.$$

9.5.2 WSSUS CHANNELS

Since the channel \mathbf{H} is random, its GSF $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$, GWS $L_{\mathbf{H}}^{(\alpha)}(t, f)$, and impulse response $\tilde{h}(t, \tau)$ are 2-D random functions (random processes). Hereafter, these random processes are assumed zero-mean. The second-order statistics of \mathbf{H} are characterized by the 4-D correlation functions $e \left\{ S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) S_{\mathbf{H}}^{(\alpha)*}(\nu', \tau') \right\}$, $e \left\{ L_{\mathbf{H}}^{(\alpha)}(t, f) L_{\mathbf{H}}^{(\alpha)*}(t', f') \right\}$, and $e \left\{ \tilde{h}(t, \tau) \tilde{h}^*(t', \tau') \right\}$, which are all mathematically equivalent.

9.5.2.1 Definition and description of WSSUS channels

An important simplification results from the assumption of *wide-sense stationary uncorrelated scattering* (WSSUS) [[48,49,52](#)]. For WSSUS channels, by definition, the reflectivities of scatterers corresponding to paths with different delay or Doppler are uncorrelated. In terms of the GSF $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$, this means

$$E \left\{ S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) S_{\mathbf{H}}^{(\alpha)*}(\nu', \tau') \right\} = C_{\mathbf{H}}(\nu, \tau) \delta(\nu - \nu') \delta(\tau - \tau'), \quad (9.5.4)$$

i.e., $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ is a *wide-sense white* random process with mean intensity function $C_{\mathbf{H}}(\nu, \tau) \geq 0$, which is known as the *scattering function* [[48,49,51–53](#)]. Together with the Fourier transform (FT) relation [\(9.5.3\)](#), the WSSUS relation [\(9.5.4\)](#) implies

$$e \left\{ L_{\mathbf{H}}^{(\alpha)}(t, f) L_{\mathbf{H}}^{(\alpha)*}(t', f') \right\} = R_{\mathbf{H}}(t - t', f - f'), \quad (9.5.5)$$

with the *(t,f) correlation function* given by [[48,49,51,52](#)]

$$R_{\mathbf{H}}(\Delta t, \Delta f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau) e^{j2\pi(\nu\Delta t - \tau\Delta f)} d\tau d\nu. \quad (9.5.6)$$

This *(t,f) correlation* satisfies $|R_{\mathbf{H}}(\Delta t, \Delta f)| \leq R_{\mathbf{H}}(0, 0)$ and $R_{\mathbf{H}}^*(-\Delta t, -\Delta f) = R_{\mathbf{H}}(\Delta t, \Delta f)$. Equation [\(9.5.5\)](#) shows that the GWS $L_{\mathbf{H}}^{(\alpha)}(t, f)$ of a WSSUS channel is a 2-D *wide-sense stationary* process. According to Eq. [\(9.5.6\)](#), the scattering function $C_{\mathbf{H}}(\nu, \tau)$ is the FT of the correlation function $R_{\mathbf{H}}(\Delta t, \Delta f)$ of $L_{\mathbf{H}}^{(\alpha)}(t, f)$. Thus, $C_{\mathbf{H}}(\nu, \tau)$ can be interpreted as the *power spectral density* (PSD) of $L_{\mathbf{H}}^{(\alpha)}(t, f)$. The *path loss* [[51](#)] is defined as

$$\rho_{\mathbf{H}}^2 \triangleq R_{\mathbf{H}}(0, 0) = E \left\{ |L_{\mathbf{H}}^{(\alpha)}(t, f)|^2 \right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau) d\tau d\nu.$$

Finally, using the impulse response $\tilde{h}(t, \tau)$, the WSSUS property is expressed as

$$\mathbb{E} \left\{ \tilde{h}(t, \tau) \tilde{h}^*(t', \tau') \right\} = r_{\mathbf{H}}(t - t', \tau) \delta(\tau - \tau'),$$

where $r_{\mathbf{H}}(\Delta t, \tau)$ is related by FTs to $C_{\mathbf{H}}(\nu, \tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f)$. Thus, $\tilde{h}(t, \tau)$ is wide-sense stationary with respect to time t and uncorrelated for different delays τ . Note that this stationarity with respect to t refers to the second-order statistics of the channel and does not imply that the channel's realizations are time-invariant systems (cf. the first of the examples given below).

The 2-D functions $C_{\mathbf{H}}(\nu, \tau)$, $R_{\mathbf{H}}(\Delta t, \Delta f)$, and $r_{\mathbf{H}}(\Delta t, \tau)$ are *mathematically equivalent* descriptions of the second-order statistics of a WSSUS channel. They are related by FTs and do not depend on α .

The composition (series connection) $\mathbf{H}_2 \mathbf{H}_1$ of two statistically independent WSSUS channels \mathbf{H}_1 and \mathbf{H}_2 can be shown to be again a WSSUS channel. Its scattering function and (t, f) correlation function are given by

$$\begin{aligned} C_{\mathbf{H}_2 \mathbf{H}_1}(\nu, \tau) &= (C_{\mathbf{H}_2} * * C_{\mathbf{H}_1})(\nu, \tau) \\ R_{\mathbf{H}_2 \mathbf{H}_1}(\Delta t, \Delta f) &= R_{\mathbf{H}_2}(\Delta t, \Delta f) R_{\mathbf{H}_1}(\Delta t, \Delta f), \end{aligned}$$

where $**$ denotes 2-D convolution. Note that $\mathbf{H}_2 \mathbf{H}_1$ and $\mathbf{H}_1 \mathbf{H}_2$ have the same second-order statistics.

From the 2-D functions $C_{\mathbf{H}}(\nu, \tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f)$, several 1-D channel descriptions can be derived. In particular, the *delay power profile* and *Doppler power profile* are respectively defined as [51]

$$P_{\mathbf{H}}(\tau) \triangleq \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau) d\nu, \quad Q_{\mathbf{H}}(\nu) \triangleq \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau) d\tau.$$

Their FTs,

$$\begin{aligned} p_{\mathbf{H}}(\Delta f) &\triangleq \int_{-\infty}^{\infty} P_{\mathbf{H}}(\tau) e^{-j2\pi\tau\Delta f} d\tau = R_{\mathbf{H}}(0, \Delta f), \\ q_{\mathbf{H}}(\Delta t) &\triangleq \int_{-\infty}^{\infty} Q_{\mathbf{H}}(\nu) e^{j2\pi\nu\Delta t} d\nu = R_{\mathbf{H}}(\Delta t, 0), \end{aligned}$$

are known as *time correlation function* and *frequency correlation function*, respectively. Often, for the sake of simplicity, a separable model is assumed for the scattering function and the (t, f) correlation function, i.e., $C_{\mathbf{H}}(\nu, \tau) = \frac{1}{\rho_{\mathbf{H}}^2} Q_{\mathbf{H}}(\nu) P_{\mathbf{H}}(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = \frac{1}{\rho_{\mathbf{H}}^2} q_{\mathbf{H}}(\Delta t) p_{\mathbf{H}}(\Delta f)$.

9.5.2.2 Examples of WSSUS channels

Some special cases of WSSUS channels are the following:

- *Time-invariant WSSUS channel.* The impulse response of a time-invariant channel has the form $\tilde{h}(t, \tau) = g(\tau)$. The WSSUS property here implies $\mathbb{E}\{g(\tau) g^*(\tau')\} = P_{\mathbf{H}}(\tau) \delta(\tau - \tau')$. It follows that $C_{\mathbf{H}}(\nu, \tau) = \delta(\nu) P_{\mathbf{H}}(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = p_{\mathbf{H}}(\Delta f)$.
- *Frequency-invariant WSSUS channel.* Next, we consider a “frequency-invariant” channel with impulse response $\tilde{h}(t, \tau) = w(t) \delta(\tau)$, i.e., the input signal is simply multiplied by $w(t)$. With $W(\nu)$ denoting the FT of $w(t)$, the WSSUS property here implies $\mathbb{E}\{W(\nu) W^*(\nu')\} = Q_{\mathbf{H}}(\nu) \delta(\nu - \nu')$. It follows that $C_{\mathbf{H}}(\nu, \tau) = Q_{\mathbf{H}}(\nu) \delta(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = q_{\mathbf{H}}(\Delta t)$.

- *Random (t,f) shift.* The GSF of a channel effecting a random frequency shift by ν_0 and a random time shift by τ_0 is given by $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau) = \delta(\nu - \nu_0)\delta(\tau - \tau_0)$. This channel is WSSUS with scattering function $C_{\mathbf{H}}(\nu, \tau) = f_{\nu_0, \tau_0}(\nu, \tau)$, where $f_{\nu_0, \tau_0}(\nu, \tau)$ is the joint probability density function of (ν_0, τ_0) [37]. Furthermore, $R_{\mathbf{H}}(\Delta t, \Delta f) = \Psi_{\nu_0, \tau_0}(\Delta t, -\Delta f)$, with $\Psi_{\nu_0, \tau_0}(\Delta t, \Delta f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\nu_0, \tau_0}(\nu, \tau) e^{j2\pi(\nu\Delta t + \tau\Delta f)} d\nu d\tau$ being the characteristic function of (ν_0, τ_0) .
- *Typical mobile radio channel.* A channel model popular in the mobile radio literature [51] uses a separable scattering function $C_{\mathbf{H}}(\nu, \tau) = \frac{1}{\rho_{\mathbf{H}}^2} Q_{\mathbf{H}}(\nu) P_{\mathbf{H}}(\tau)$ with an exponential delay power profile

$$P_{\mathbf{H}}(\tau) = \begin{cases} \frac{\rho_{\mathbf{H}}^2}{\tau_0} e^{-\tau/\tau_0}, & \tau \geq 0, \\ 0, & \tau < 0, \end{cases} \quad (9.5.7)$$

and a *Jakes* Doppler power profile

$$Q_{\mathbf{H}}(\nu) = \begin{cases} \frac{\rho_{\mathbf{H}}^2}{\pi\sqrt{\nu_{\max}^2 - \nu^2}}, & |\nu| \leq \nu_{\max}, \\ 0, & |\nu| > \nu_{\max}. \end{cases} \quad (9.5.8)$$

The (t,f) correlation function is $R_{\mathbf{H}}(\Delta t, \Delta f) = \frac{1}{\rho_{\mathbf{H}}^2} q_{\mathbf{H}}(\Delta t) p_{\mathbf{H}}(\Delta f)$, with $p_{\mathbf{H}}(\Delta f) = \frac{\rho_{\mathbf{H}}^2}{1+j2\pi\tau_0\Delta f}$ and $q_{\mathbf{H}}(\Delta t) = \rho_{\mathbf{H}}^2 J_0(2\pi\nu_{\max}\Delta t)$, where $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind.

9.5.2.3 Statistical input-output relations for WSSUS channels

The scattering function and (t,f) correlation function are useful for formulating input-output relations that show how the second-order statistics of the channel output signal $r(t)$ depend on the second-order statistics of the input signal $s(t)$. Let $s(t)$ be a nonstationary random process that is statistically independent of the random channel \mathbf{H} . The second-order statistics of a nonstationary random process $x(t)$ with correlation operator \mathbf{R}_x can be described by the *generalized Wigner-Ville spectrum* $\bar{W}_x^{(\alpha)}(t, f) \triangleq L_{\mathbf{R}_x}^{(\alpha)}(t, f)$ or, alternatively, by the *generalized expected ambiguity function* $\bar{A}_x^{(\alpha)}(\nu, \tau) \triangleq S_{\mathbf{R}_x}^{(\alpha)}(\nu, \tau)$, which is the 2-D FT of $\bar{W}_x^{(\alpha)}(t, f)$ (see Section 9.4). Then,

$$\bar{W}_r^{(\alpha)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau) \bar{W}_s^{(\alpha)}(t - \tau, f - \nu) d\tau d\nu, \quad (9.5.9)$$

$$\bar{A}_r^{(\alpha)}(\nu, \tau) = R_{\mathbf{H}}(\tau, \nu) \bar{A}_s^{(\alpha)}(\nu, \tau). \quad (9.5.10)$$

For a wide-sense *stationary* transmitted signal $s(t)$ with PSD $P_s(f)$ and autocorrelation $r_s(\tau)$, the received signal $r(t)$ is stationary, and Eqs. (9.5.9) and (9.5.10) become

$$P_r(f) = \int_{-\infty}^{\infty} Q_{\mathbf{H}}(\nu) P_s(f - \nu) d\nu, \quad r_r(\tau) = q_{\mathbf{H}}(\tau) r_s(\tau).$$

Furthermore, $e\{|r(t)|^2\} = \rho_{\mathbf{H}}^2 e\{|s(t)|^2\}$. Dual results are obtained for a nonstationary white transmit signal $s(t)$. Finally, if $s(t)$ is *cyclostationary* with period T , cyclic correlation function $r_s^{(k)}(\tau)$, and cyclic

spectral density $P_s^{(k)}(f)$ [54], the received signal $r(t)$ is cyclostationary with the same period T and we have

$$P_r^{(k)}(f) = \int_{-\infty}^{\infty} Q_{\mathbf{H}}^{(k)}(\nu) P_s^{(k)}(f - \nu) d\nu, \quad r_r^{(k)}(\tau) = q_{\mathbf{H}}^{(k)}(\tau) r_s^{(k)}(\tau),$$

where $q_{\mathbf{H}}^{(k)}(\Delta t) = R_{\mathbf{H}}(\Delta t, k/T)$ and $Q_{\mathbf{H}}^{(k)}(\nu) = \int_{-\infty}^{\infty} q_{\mathbf{H}}^{(k)}(\Delta t) e^{-j2\pi\nu\Delta t} d\Delta t$.

9.5.3 UNDERSpread WSSUS CHANNELS

A fundamental classification of WSSUS channels is into *underspread* and *overspread* channels [49,52]. Underspread WSSUS channels have some interesting properties, which are analogous to those of underspread deterministic time-varying systems in Section 4.7.

9.5.3.1 Definition of underspread channels

A WSSUS channel is *underspread* [52] if its scattering function is highly concentrated about the origin. The underspread property is practically relevant as most mobile radio channels are underspread. A characterization of scattering function concentration without the restrictive assumption of compact support uses normalized *weighted integrals* of the form (4.7.5), i.e.,

$$\bar{m}_{\mathbf{H}}^{(\phi)} \triangleq \frac{1}{\rho_{\mathbf{H}}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\nu, \tau) C_{\mathbf{H}}(\nu, \tau) d\nu d\tau,$$

where $\phi(\nu, \tau) \geq 0$ is a weighting function as discussed in Section 4.7.4. Of particular importance are the *delay spread* $\tau_{\mathbf{H}}$ and *Doppler spread* $v_{\mathbf{H}}$, which are obtained with $\phi(\nu, \tau) = \tau^2$ and $\phi(\nu, \tau) = \nu^2$, respectively:

$$\tau_{\mathbf{H}}^2 \triangleq \frac{1}{\rho_{\mathbf{H}}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau^2 C_{\mathbf{H}}(\nu, \tau) d\nu d\tau, \quad (9.5.11)$$

$$v_{\mathbf{H}}^2 \triangleq \frac{1}{\rho_{\mathbf{H}}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu^2 C_{\mathbf{H}}(\nu, \tau) d\nu d\tau. \quad (9.5.12)$$

Within this framework, a WSSUS channel is called underspread if specific weighted integrals and moments of the scattering function are small.

9.5.3.2 Approximate eigenfunctions and eigenvalues of underspread channels

It is known [49,52] that signals with good time and/or frequency concentration can pass an underspread WSSUS channel almost undistorted, i.e., merely multiplied by a random complex factor. We analyze this effect using the approach in Ref. [55]. Similar results in a deterministic context are reported in Sections 4.7 and 13.3.

A normalized transmitted signal $s(t)$ that remains undistorted, i.e., $(\mathbf{H}s)(t) = \lambda s(t)$, is an eigenfunction of the system \mathbf{H} ; the associated eigenvalue is given by $\lambda = \langle \mathbf{H}s, s \rangle$. Since \mathbf{H} is random, the relation $(\mathbf{H}s)(t) = \langle \mathbf{H}s, s \rangle s(t)$ is formulated in the mean-square sense, i.e., $E \left\{ \| \mathbf{H}s - \langle \mathbf{H}s, s \rangle s \|^2 \right\} = 0$. The eigenfunctions of a WSSUS channel \mathbf{H} are random and generally do not possess a specific structure. However, in the *underspread* case, (t, f) translates of a function $g(t)$ with good (t, f) concentration are

approximate eigenfunctions. Specifically, consider the (t, f) translates $g_{t_0, f_0}(t) = g(t - t_0) e^{j2\pi f_0 t}$, where $g(t)$ is a normalized function that is well concentrated about $(t, f) = (0, 0)$. One can show

$$e \left\{ \| \mathbf{H}g_{t_0, f_0} - \langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle g_{t_0, f_0} \|^2 \right\} \leq \rho_{\mathbf{H}}^2 \bar{m}_{\mathbf{H}}^{(\phi)}, \quad (9.5.13)$$

with $\phi(v, \tau) = 1 - |A_g^{(\alpha)}(v, \tau)|^2$, where

$$A_g^{(\alpha)}(v, \tau) = \int_{-\infty}^{\infty} g \left(t + \left(\frac{1}{2} - \alpha \right) \tau \right) g^* \left(t - \left(\frac{1}{2} + \alpha \right) \tau \right) e^{-j2\pi v t} dt$$

denotes the generalized ambiguity function of $g(t)$. Therefore, if the channel is underspread, i.e., the channel's scattering function is concentrated about the origin (where $|A_g^{(\alpha)}(v, \tau)|^2 \approx |A_g^{(\alpha)}(0, 0)|^2 = 1$ and thus $\phi(v, \tau) \approx 0$), the weighted integral $\bar{m}_{\mathbf{H}}^{(\phi)}$ is small and one has the approximation (valid in the mean-square sense)

$$\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle \approx \langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle g_{t_0, f_0}(t). \quad (9.5.14)$$

This shows that $g_{t_0, f_0}(t)$ is an approximate eigenfunction of \mathbf{H} . Furthermore,

$$e \left\{ |\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle - L_{\mathbf{H}}^{(\alpha)}(t_0, f_0)|^2 \right\} \leq \rho_{\mathbf{H}}^2 \bar{m}_{\mathbf{H}}^{(\phi')}, \quad (9.5.15)$$

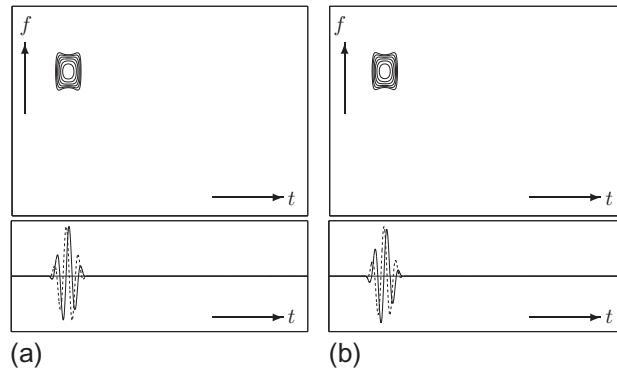
with $\phi'(v, \tau) = |1 - A_g^{(\alpha)}(v, \tau)|^2$. Thus, under the same conditions, we have

$$\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle \approx L_{\mathbf{H}}^{(\alpha)}(t_0, f_0) \quad (9.5.16)$$

(again valid in the mean-square sense), which shows that the approximate eigenvalue $\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle$ is approximately equal to the GWS at the (t, f) point (t_0, f_0) .

In contrast to the exact eigenfunctions of \mathbf{H} , the approximate eigenfunctions $g_{t_0, f_0}(t)$ are (t, f) translates of a single prototype function $g(t)$ and thus highly structured; they do not depend on the specific channel realization and their parameters t_0, f_0 have an immediate physical interpretation.

To illustrate the above eigenfunction/eigenvalue approximations, we simulate the transmission of a signal $g_{t_0, f_0}(t)$, with $g(t)$ a Hanning window of duration $T_g = 128 \mu\text{s}$, over a WSSUS channel. The channel's scattering function is $C_{\mathbf{H}}(v, \tau) = \frac{1}{\rho_{\mathbf{H}}^2} Q_{\mathbf{H}}(v) P_{\mathbf{H}}(\tau)$ with exponential $P_{\mathbf{H}}(\tau)$ (Eq. (9.5.7) with $\tau_0 = 1 \mu\text{s}$) and Jakes-type $Q_{\mathbf{H}}(v)$ (Eq. (9.5.8) with $v_{\max} = 305 \text{ Hz}$). Figure 9.5.1 illustrates the approximations (9.5.14) and (9.5.16) for a single channel realization. It is seen that the received signal $\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle$ and the approximation $L_{\mathbf{H}}^{(1/2)}(t_0, f_0) g_{t_0, f_0}(t)$ are practically identical. We used 500 realizations of \mathbf{H} to estimate the normalized mean-square error $E \left\{ \| \mathbf{H}g_{t_0, f_0} - \langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle g_{t_0, f_0} \|^2 \right\} / \rho_{\mathbf{H}}^2$ (see Eq. (9.5.13)). The result, $9 \cdot 10^{-4}$, confirms the validity of the eigenfunction/eigenvalue approximation. The associated upper bound $\bar{m}_{\mathbf{H}}^{(\phi)}$ (see Eq. (9.5.13)) is calculated as $5 \cdot 10^{-3}$. Finally, the normalized mean-square error $E \left\{ |\langle \mathbf{H}g_{t_0, f_0}, g_{t_0, f_0} \rangle - L_{\mathbf{H}}^{(1/2)}(t_0, f_0)|^2 \right\} / \rho_{\mathbf{H}}^2$ (see Eq. (9.5.15)) is estimated as $5 \cdot 10^{-6}$ and the associated upper bound $\bar{m}_{\mathbf{H}}^{(\phi')}$ is calculated as $2 \cdot 10^{-3}$.

**FIGURE 9.5.1**

Eigenfunction/eigenvalue approximation for an underspread WSSUS channel: (a) Wigner distribution (top) and real and imaginary parts (bottom) of received signal $r(t) = (\mathbf{H}g_{t_0, f_0})(t)$, (b) Wigner distribution (top) and real and imaginary parts (bottom) of approximation $L_{\mathbf{H}}^{(1/2)}(t_0, f_0)g_{t_0, f_0}(t)$.

9.5.3.3 Sampling approximation for underspread channels

Next, we consider 2-D sampling of the channel's transfer function (GWS) $L_{\mathbf{H}}^{(\alpha)}(t, f)$. This is important for simplified channel representations that are used, e.g., in the context of *orthogonal frequency division multiplexing (OFDM)* modulation [56,57].

Consider the representation of a WSSUS channel \mathbf{H} by the samples $L_{\mathbf{H}}^{(\alpha)}(k\delta t, l\delta f)$ of its GWS taken on the uniform sampling grid $(k\delta t, l\delta f)$. The reconstructed (interpolated) GWS is given by

$$\hat{L}_{\mathbf{H}}^{(\alpha)}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}(k\delta t, l\delta f) \operatorname{sinc} \frac{t-k\delta t}{\delta t} \operatorname{sinc} \frac{f-l\delta f}{\delta f},$$

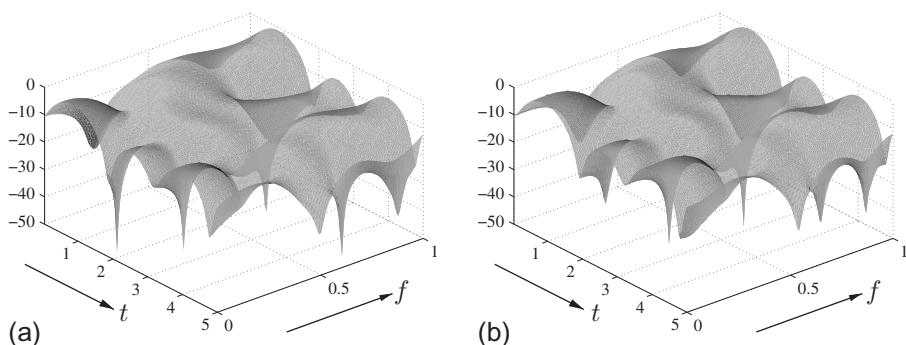
with $\operatorname{sinc} x = \frac{\sin \pi x}{\pi x}$. For WSSUS channels with scattering function $C_{\mathbf{H}}(\nu, \tau)$ compactly supported within a rectangular area $[-\nu_{\max}, \nu_{\max}] \times [-\tau_{\max}, \tau_{\max}]$ and for sampling grid constants satisfying $\delta t \leq 1/(2\nu_{\max})$ and $\delta f \leq 1/(2\tau_{\max})$, the above reconstruction can be shown to be exact in the sense that

$$\mathbb{E} \left\{ |\hat{L}_{\mathbf{H}}^{(\alpha)}(t, f) - L_{\mathbf{H}}^{(\alpha)}(t, f)|^2 \right\} = 0.$$

Note that a smaller channel spread $\sigma_{\mathbf{H}} = 4\nu_{\max}\tau_{\max}$ allows for a coarser sampling grid and thus for a more parsimonious channel representation.

If the above conditions are not satisfied, the reconstructed GWS $\hat{L}_{\mathbf{H}}^{(\alpha)}(t, f)$ will contain errors due to aliasing. However, these errors are bounded as

$$\mathbb{E} \left\{ |\hat{L}_{\mathbf{H}}^{(\alpha)}(t, f) - L_{\mathbf{H}}^{(\alpha)}(t, f)|^2 \right\} \leq 2\rho_{\mathbf{H}}^2 \left(\tau_{\mathbf{H}}^2 \delta f^2 + \nu_{\mathbf{H}}^2 \delta t^2 \right), \quad (9.5.17)$$

**FIGURE 9.5.2**

Sampling approximation for an underspread WSSUS channel: (a) $|L_{\mathbf{H}}^{(1/2)}(t, f)|^2$ and (b) $|\hat{L}_{\mathbf{H}}^{(1/2)}(t, f)|^2$ (in dB; time in ms, frequency in MHz).

where $\tau_{\mathbf{H}}$ and $\nu_{\mathbf{H}}$ are the delay spread and Doppler spread as defined in Eqs. (9.5.11) and (9.5.12). Thus, for WSSUS channels with small $\tau_{\mathbf{H}}$ and/or small $\nu_{\mathbf{H}}$, i.e., for *underspread* channels, a sampling of the transfer function results in negligible errors provided that the sampling periods δt and δf are chosen appropriately. Specifically, the upper error bound in Eq. (9.5.17) is minimized when $\delta t/\delta f = \tau_{\mathbf{H}}/\nu_{\mathbf{H}}$.

For the WSSUS channel with exponential/Jakes scattering function that was considered further above, and for sampling periods $\delta t = 138 \mu\text{s}$, $\delta f = 136.72 \text{ kHz}$, the normalized mean-square error $e \left\{ |\hat{L}_{\mathbf{H}}^{(1/2)}(t, f) - L_{\mathbf{H}}^{(1/2)}(t, f)|^2 \right\} / \rho_{\mathbf{H}}^2$ was estimated from 500 channel realizations as $6.4 \cdot 10^{-3}$, and the upper bound $2(\tau_{\mathbf{H}}^2 \delta f^2 + \nu_{\mathbf{H}}^2 \delta t^2)$ was calculated as $3.2 \cdot 10^{-2}$. Figure 9.5.2 shows the squared magnitude of the true channel transfer function $L_{\mathbf{H}}^{(1/2)}(t, f)$ and of the reconstruction $\hat{L}_{\mathbf{H}}^{(1/2)}(t, f)$ for a specific channel realization.

9.5.4 SUMMARY AND CONCLUSIONS

This section considered time-frequency characterizations of (the second-order statistics of) random linear time-varying channels satisfying the assumption of wide-sense stationary uncorrelated scattering (WSSUS). The practically important class of *underspread* WSSUS channels allows some interesting approximations. In particular, underspread WSSUS channels possess approximate eigenfunctions with time-frequency shift structure (which suggests the use of OFDM), and they can be discretized by means of a time-frequency sampling. Related considerations and results can be found in Sections 13.2 and 13.3.

9.6 CHARACTERIZATION OF CYCLOSTATIONARY SIGNALS AND THEIR GENERALIZATIONS⁰

In this section, cyclostationary signals are characterized and their spectral analysis is provided. Links between the considered statistical functions and quadratic time-frequency representations are highlighted. Generalizations of the concept of cyclostationarity are addressed. The considered classes of signals are suitable models in several fields of application including communications, radar/sonar, telemetry, climatology, astronomy, acoustics, mechanics, biology, bioengineering, econometrics, and finance (see Refs. [58–61] and references therein).

9.6.1 CYCLOSTATIONARY SIGNALS

9.6.1.1 Characterization

A finite-power complex-valued stochastic process $x(t)$ is said to be a second-order process if $E\{|x(t)|^2\}$ is finite for all t . Thus, its second-order moments

$$\mathcal{R}_x(t, \tau) \triangleq E \left\{ x(t + \tau) x^{(*)}(t) \right\} \quad (9.6.1)$$

exist for all t and τ , and $E\{|x(t)|\}$ is finite for all t . In Eq. (9.6.1), $(*)$ denotes an optional complex conjugation, and subscript $x = [xx^{(*)}]$. Two functions are defined in Eq. (9.6.1): the autocorrelation function if conjugation is present; and the conjugate autocorrelation function, also called the relation function or complementary correlation, if conjugation is absent. The autocorrelation function in Eq. (9.6.1) is the asymmetric counterpart of the symmetric autocorrelation in Eq. (2.2.1) in Section 2.2.1, p. 74.

If a process arises from the interaction of a periodic or almost-periodic phenomenon with a random one, then the (almost-)periodicity property is hidden in the process itself, but is present in its statistical functions [58,59]. A process $x(t)$ is said to be *second-order almost cyclostationary (ACS) in the wide sense* if its first- and second-order moments are almost-periodic functions of t in the sense of Besicovitch [61, Section 1.2.1], i.e., they are limits of uniformly convergent sequences of trigonometric polynomials. That is, these moments can be expressed by the (generalized) Fourier series expansions

$$E\{x(t)\} = \sum_{f_\eta \in F} x_{f_\eta} e^{j2\pi f_\eta t} \quad (9.6.2)$$

$$\mathcal{R}_x(t, \tau) = \sum_{f_\alpha \in A} R_x^{f_\alpha}(\tau) e^{j2\pi f_\alpha t} \quad (9.6.3)$$

(see Refs. [58, Part II], [59, Chap. 12], and [60,62]), where F is the countable set of the possibly incommensurate *first-order cycle frequencies* f_η , and A is the countable set (depending on $(*)$) of possibly incommensurate *second-order (conjugate) cycle frequencies* f_α . The Fourier coefficients in Eq. (9.6.2) are

$$x_{f_\eta} \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} E\{x(t)\} e^{-j2\pi f_\eta t} dt \quad (9.6.4)$$

⁰Author: **A. Napolitano**, University of Napoli “Parthenope”, Napoli, Italy (antnapol@gmail.com, antonio.napolitano@uniparthenope.it). Reviewers: I. Djurović and L. Galleani.

and those in Eq. (9.6.3) are

$$R_{\mathbf{x}}^{f_\alpha}(\tau) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathcal{R}_{\mathbf{x}}(t, \tau) e^{-j2\pi f_\alpha t} dt, \quad (9.6.5)$$

which are called the (*conjugate*) *cyclic autocorrelation functions*. If all elements of sets F and A are multiples of the same fundamental frequency $1/T_0$, then the mean value and (*conjugate*) autocorrelation function are periodic in t with period T_0 , and the process is said to be cyclostationary with period T_0 . If F and A contain only the zero-cycle frequency, then the mean value and (*conjugate*) autocorrelation function do not depend on t and the process is *wide-sense stationary* (WSS). More general classes of ACS processes can be defined if the almost-periodicity of the mean and (*conjugate*) autocorrelation is defined in a generalized sense [61, Section 1.2]. In such a case, the mean and (*conjugate*) autocorrelation function can be expressed as the sum of a (generalized) Fourier series (as in Eqs. (9.6.2) and (9.6.3)) and a residual term not containing any finite-strength additive sinewave component [61, Sections 1.2, 2.2.2].

A suitable characterization of stochastic processes in the frequency domain can be made by resorting to the concept of harmonizability. A second-order process $x(t)$ is said to be *harmonizable* if the (*conjugate*) autocorrelation function can be expressed by the Fourier-Stieltjes integral [62, Section 5.2]

$$\mathcal{R}_{\mathbf{x}}(t, \tau) = \int_{\mathbb{R}^2} e^{j2\pi[f_1(t+\tau)+(-)f_2t]} d\gamma_{\mathbf{x}}(f_1, f_2), \quad (9.6.6)$$

where $(-)$ is an optional minus sign that is linked to $(*)$, and $\gamma_{\mathbf{x}}(f_1, f_2)$ is a (*conjugate*) spectral correlation function of bounded variation. For a harmonizable process $x(t)$, the following Cramér representation holds [62, Sections 4.1.2, 5.2]:

$$x(t) = \int_{\mathbb{R}} e^{j2\pi ft} d\chi(f), \quad (9.6.7)$$

where $\chi(f)$ is the integrated spectrum of $x(t)$. If $\gamma_{\mathbf{x}}(f_1, f_2)$ contains no singular components and $\chi(f)$ contains no singular components with probability 1, in the sense of the generalized processes, the Fourier transform (FT) $X(f)$ of $x(t)$ can be defined so that $d\chi(f) = X(f) df$ and $d\gamma_{\mathbf{x}}(f_1, f_2) = e\{d\chi(f_1) d\chi^{(*)}(f_2)\} = S(f_1, f_2) df_1 df_2$, where

$$S(f_1, f_2) \triangleq E \left\{ X(f_1) X^{(*)}(f_2) \right\} \quad (9.6.8)$$

is the (*conjugate*) *Loève bifrequency spectrum* of $x(t)$ [62, p. 197] [61, Sections 1.1.2, 4.2.1].

The almost-periodicity of the (*conjugate*) autocorrelation function reflects in the frequency domain into correlation between spectral components of the process that are separated by quantities equal to the (*conjugate*) second-order cycle frequencies. In fact, for an ACS process, the Loève bifrequency spectrum can be expressed as

$$S_{\mathbf{x}}(f_1, f_2) = \sum_{f_\alpha \in A} S_{\mathbf{x}}^{f_\alpha}(f_1) \delta(f_2 + (-)(f_1 - f_\alpha)), \quad (9.6.9)$$

where $\delta(\cdot)$ denotes the Dirac delta function and

$$S_{\mathbf{x}}^{f_\alpha}(f) = \lim_{\Delta f \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \Delta f E \left\{ X_{1/\Delta f}(t, f) X_{1/\Delta f}^{(*)}(t, (-)(f_\alpha - f)) \right\} dt \quad (9.6.10)$$

is the (*conjugate*) *cyclic spectrum* or (*conjugate*) *spectral correlation density function*, with

$$X_Z(t,f) \triangleq \int_{t-Z/2}^{t+Z/2} x(s) e^{-j2\pi f s} ds \quad (9.6.11)$$

being the short-time Fourier transform (STFT) of $x(t)$. For $f_\alpha = 0$, $S_x^{f_\alpha}(f)$ is the power spectrum [see Eq. (2.2.2), p. 74].

The following Gardner relation (also called cyclic Wiener-Khinchin relation) [59] links the (*conjugate*) cyclic autocorrelation function and the cyclic spectrum

$$S_x^{f_\alpha}(f) = \int_{\mathbb{R}} R_x^{f_\alpha}(\tau) e^{-j2\pi f \tau} d\tau. \quad (9.6.12)$$

For $f_\alpha = 0$, Eq. (9.6.12) reduces to the Wiener-Khinchin relation [Eq. (2.2.3)].

9.6.1.2 Links with quadratic time-frequency representations

The FT with respect to (w.r.t.) τ of the (*conjugate*) autocorrelation function (9.6.1) is the (*conjugate*) *time-varying spectrum* $S_x(t,f)$. It is the asymmetric counterpart of the symmetric time-varying spectrum defined in Section 2.2.1. For an ACS process, using Eqs. (9.6.3) and (9.6.12), one has

$$S_x(t,f) = \sum_{f_\alpha \in A} S_x^{f_\alpha}(f) e^{j2\pi f_\alpha t}; \quad (9.6.13)$$

that is, the (*conjugate*) time-varying spectrum is an almost-periodic function whose (generalized) Fourier series has frequencies equal to the (*conjugate*) cycle frequencies and coefficients equal to the (*conjugate*) cyclic spectra. A similar representation is obtained for the expected Wigner distribution defined in accordance with the definition provided for stochastic processes in Section 9.4.2. In fact, it results in

$$E\{W_x(t,f)\} \triangleq \int_{\mathbb{R}} E\left\{x(t + \tau/2)x^{(*)}(t - \tau/2)\right\} e^{-j2\pi f \tau} d\tau = S_x(t,f + f_\alpha/2). \quad (9.6.14)$$

That is, the expected Wigner distribution of ACS signals is an almost-periodic function whose (generalized) Fourier series has frequencies equal to the (*conjugate*) cycle frequencies and coefficients equal to $S_x^{f_\alpha}(f + f_\alpha/2)$.

The generalized expected ambiguity function (GEAF) of ACS processes is given by (Section 9.4.4)

$$E\{A_x(v, \tau)\} = \int_{\mathbb{R}} \mathcal{R}_x(t, \tau) e^{-j2\pi v t} dt = \sum_{f_\alpha \in A} R_x^{f_\alpha}(\tau) \delta(v - f_\alpha). \quad (9.6.15)$$

From Eq. (9.6.15), it follows that ACS processes cannot be underspread (Section 9.4.5), since the GEAF is not concentrated around $(v, \tau) = (0, 0)$. This is in agreement with the fact that ACS processes exhibit the spectral correlation property. Note that in the special case of WSS processes, the set A contains only the element $f_\alpha = 0$ and the expression (9.6.15) reduces to the expression given in Section 9.4.4 for the WSS case.

Let $v(\tau)$ be a unit-width lag window with FT $V(f)$ such that $\int_{\mathbb{R}} V(f) df = 1$. In the notation of [Section 3.2](#), by using the kernels

$$\begin{aligned} G(t, \tau) &= \frac{1}{T} \operatorname{rect}\left(\frac{t}{T}\right) \sum_{f_\alpha \in A} e^{j2\pi f_\alpha t} v(\Delta f \tau) \\ \gamma(t, f) &= \underset{\tau \rightarrow f}{\mathcal{F}} \{G(t, \tau)\} = \frac{1}{T} \operatorname{rect}\left(\frac{t}{T}\right) \sum_{f_\alpha \in A} e^{j2\pi f_\alpha t} \frac{1}{\Delta f} V\left(\frac{f}{\Delta f}\right), \end{aligned} \quad (9.6.16)$$

the estimators of the symmetric time-varying autocorrelation function and of the expected Wigner spectrum in [Section 3.2.2](#) are reduced to

$$R_x(t, \tau) = x(t + \tau/2) x^*(t - \tau/2) *_t G(t, \tau) = \sum_{f_\alpha \in A} R_x^{(T)}(f_\alpha, \tau; t) e^{j2\pi f_\alpha t} \quad (9.6.17)$$

$$\rho_x(t, f) = W_x(t, f) \underset{(t, f)}{*} \gamma(t, f) = \sum_{f_\alpha \in A} \left[I_x^{(T)}(f_\alpha, f; t) *_f \frac{1}{\Delta f} V\left(\frac{f}{\Delta f}\right) \right] e^{j2\pi f_\alpha t}, \quad (9.6.18)$$

where $*$ is the convolution operation, and the expression

$$R_x^{(T)}(f_\alpha, \tau; t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} x(u + \tau/2) x^*(u - \tau/2) e^{-j2\pi f_\alpha u} du \quad (9.6.19)$$

is the cyclic correlogram; $I_x^{(T)}(f_\alpha, f; t)$ is its FT w.r.t. τ , which is called the cyclic periodogram (see related relationships in [Section 3.2.4](#)). Starting from the results of [Section 9.6.1.3](#), these estimators can be shown to be mean-square consistent as $T \rightarrow \infty$ and $\Delta f \rightarrow 0$ with $T\Delta f \rightarrow \infty$.

9.6.1.3 Statistical function estimation and applications

Under mild conditions on the finite or practically finite memory of the process expressed in terms of α -mixing conditions or summability of cumulants, the (conjugate) cyclic correlogram and the time- or frequency-smoothed (conjugate) cyclic periodogram of $x(t)$ are consistent estimators of the (conjugate) cyclic autocorrelation function and cyclic spectrum, respectively [63–65]. Moreover, these estimators are asymptotically complex normal.

Almost all man-made modulated signals adopted in communications, radar, sonar, and telemetry exhibit cyclostationarity properties that have been exploited for a variety of applications. In fact, (conjugate) cycle frequencies are related to parameters such as carrier frequency, baud rate, sampling frequency, scanning frequency, and coding rate [58]. Moreover, cyclic features of modulated signals depend on signal characteristics such as pulse shape, code, and statistical characteristics of the modulating signal. Consequently, cyclostationarity properties turn out to be signal-selective and can be suitably exploited to counteract the effects of noise and interference. Applications have been proposed in weak-signal detection problems, interference rejection, source location, synchronization, and signal classification [58, 59, 66–70]. Exploitation of cyclostationarity in mechanics and other fields is discussed in Refs. [60, 72].

The relative motion between transmitter and receiver modifies at the receiver the kind of nonstationarity of the transmitted signal [61, Chap. 7]. For this purpose, two different classes of nonstationary processes are considered in [Sections 9.6.2](#) and [9.6.3](#) to accurately model the received signal in several cases of interest when the transmitted signal is ACS.

9.6.2 GENERALIZED ALMOST-CYCLOSTATIONARY SIGNALS

9.6.2.1 Characterization

A continuous-time second-order process $x(t)$ with possibly almost-periodic mean value is said to be *generalized almost-cyclostationary* (GACS) in the wide sense if its (conjugate) autocorrelation function can be expressed by the two equivalent (generalized) Fourier series expansions

$$\mathcal{R}_x(t, \tau) = \sum_{f_\alpha \in A_\tau} R_x(f_\alpha, \tau) e^{j2\pi f_\alpha t} = \sum_{k \in \mathbb{I}} R_x^{(k)}(\tau) e^{j2\pi \alpha_x^{(k)}(\tau)t} \quad (9.6.20)$$

(see [73], [61, Section 2.2.2], and [71]). In Eq. (9.6.20), the (conjugate) cyclic autocorrelation function $R_x(f_\alpha, \tau)$ is defined as $R_x^{f_\alpha}(\tau)$ in Eq. (9.6.5), but with (conjugate) cycle frequency f_α ranging in the countable set A_τ depending on τ . \mathbb{I} is a countable set and both the coefficients $R_x^{(k)}(\tau)$, known as *generalized (conjugate) cyclic autocorrelation functions*, and the frequencies $\alpha_x^{(k)}(\tau)$, known as *lag-dependent (conjugate) cycle frequencies*, depend on the lag-parameter τ . We have [61, Section 2.2.2]

$$R_x(f_\alpha, \tau) = \sum_{k \in \mathbb{I}} R_x^{(k)}(\tau) \delta(f_\alpha - \alpha_x^{(k)}(\tau)), \quad (9.6.21)$$

where δ denotes the Kronecker delta, here defined as $\delta_\gamma = 1$ for $\gamma = 0$, and $\delta_\gamma = 0$ for $\gamma \neq 0$. Thus the equations $f_\alpha = \alpha_x^{(k)}(\tau)$, $k \in \mathbb{I}$, describe the support curves in the (f_α, τ) -plane of $R_x(f_\alpha, \tau)$.

ACS signals are obtained as a special case of the GACS process when the functions $\alpha_x^{(k)}(\tau)$ do not depend on τ , and the support curves are lines parallel to the τ axis (compare Eqs. (9.6.3) and (9.6.20)).

The GEAF of GACS processes are given by [74]

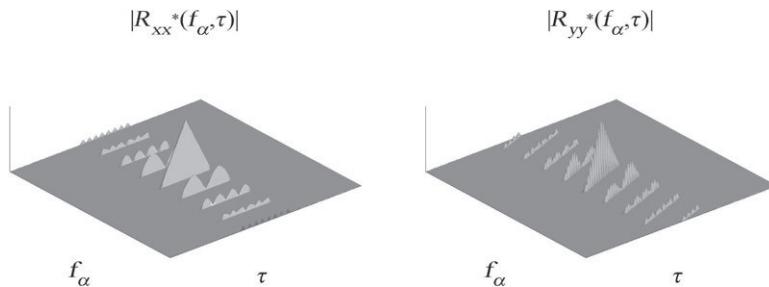
$$E\{A_x(v, \tau)\} = \sum_{k \in \mathbb{I}} R_x^{(k)}(\tau) \delta(v - \alpha_x^{(k)}(\tau)), \quad (9.6.22)$$

which reduces to Eq. (9.6.15) when the functions $\alpha_x^{(k)}(\tau)$ do not depend on τ . As in the ACS case, GACS processes are not underspread.

The characterization and analysis of GACS processes in the frequency domain is a challenging problem. Every GACS process that does not reduce to an ACS process is not harmonizable in the sense of Loève. A second-order characterization in the frequency domain of GACS processes is provided in Ref. [61, Sections 2.2.3, 3.2.1], in a generalized sense introducing a different distribution (generalized function).

9.6.2.2 Discrete-time, statistical function estimation, and applications

The discrete-time counterpart of continuous-time GACS processes does not exist [61, Section 2.5]. In fact, the discrete-time counterpart of the first (generalized) Fourier series expansion in Eq. (9.6.20) can always be written with a cycle frequency set that does not depend on the lag parameter. Similarly, the discrete-time counterpart of the second (generalized) Fourier series expansion in Eq. (9.6.20) can always be written with cycle frequencies that do not depend on the lag parameter. In addition, uniformly sampling GACS processes leads to discrete-time ACS processes.

**FIGURE 9.6.1**

Magnitude of the cyclic autocorrelation function vs. f_α and τ for (left) the input ACS signal $x(t)$, and (right) the output GACS signal $y(t)$, of the Doppler channel between transmitter and receiver with constant relative radial acceleration.

Under mild regularity assumptions of the memory of the process, expressed in terms of summability of second- and fourth-order cumulants, the (conjugate) cyclic correlogram is a mean-square consistent estimator of the (conjugate) cyclic autocorrelation function $R_x(f_\alpha, \tau)$ (see Refs. [61, Section 2.4.2], and [71]). Under further summability conditions on higher-order cumulants, a properly normalized version of the estimation error is asymptotically complex normal (see [71], [61, Section 2.4.3]). Mean-square consistent and asymptotically complex normal discrete-time estimators of samples of the continuous-time (conjugate) cyclic autocorrelation function of GACS processes are presented in Ref. [61, Section 2.6].

The GACS model finds application in mobile communications. Let us consider the case of relative motion between transmitter and receiver with constant relative radial acceleration under the so-called *narrow band condition*—that is, when the product of signal bandwidth and data-record length is much smaller than the ratio of medium propagation speed and radial speed [61, Chap. 7]. If the transmitted signal is ACS, then the received signal is GACS with lag-dependent cycle frequencies that are linear functions of τ [61, Section 7.8.1.1] (Fig. 9.6.1). A further application of the GACS signals is to model communications signals with slowly varying parameters such as carrier frequency or baud rate [61, Section 7.9.3].

The ACS and GACS signals can also be analyzed by the fraction-of-time or functional approach where signals are modeled as single functions of time rather than sample paths of stochastic processes (see Refs. [58, 68], and [61, Chap. 6]). A theory alternative to the classical theory can be constructed by replacing the ensemble expectation operator with the almost-periodic component extraction operator.

9.6.3 SPECTRALLY CORRELATED SIGNALS

9.6.3.1 Characterization

Let $x(t)$ be a second-order harmonizable process not containing any additive finite-strength sinewave component. The process $x(t)$ is said to be *spectrally correlated* (SC) if its Loève bifrequency spectrum has spectral masses in the bifrequency plane (f_1, f_2) concentrated on a countable set \mathbb{I} of support curves defined by the equations $f_2 = \Psi_x^{(k)}(f_1)$, with *spectral correlation density* $S_x^{(k)}(f_1)$ on each curve. That is (see Ref. [75], [61, Section 4.2.1]),

$$S_{\mathbf{x}}(f_1, f_2) = \sum_{k \in \mathbb{I}} S_{\mathbf{x}}^{(k)}(f_1) \delta(f_2 - \Psi_{\mathbf{x}}^{(k)}(f_1)). \quad (9.6.23)$$

For SC processes, the time-varying spectrum is an almost-periodic function of t whose (generalized) Fourier series expansion has both frequencies and coefficients depending on the spectral frequency f

$$S_{\mathbf{x}}(t, f) = \sum_{k \in \mathbb{I}} S_{\mathbf{x}}^{(k)}(f) e^{j2\pi[f + (-)\Psi_{\mathbf{x}}^{(k)}(f)]t}. \quad (9.6.24)$$

ACS processes are obtained as a special case of SC processes when the support curves of the Loève bifrequency spectrum in the bifrequency plane are lines with unit slope (compare Eqs. (9.6.9) and (9.6.23)). In such a case, the expression (9.6.24) of the time-varying spectrum reduces to Eq. (9.6.13). Note that, unlike the case of the ACS processes, the expected Wigner spectrum for SC processes cannot be easily linked to the time-varying spectrum. ACS processes are obtained as the intersection of the class of the GACS processes and that of the SC processes. That is, they are the subclass of GACS processes, i.e., of those with almost-periodic (conjugate) autocorrelation that also exhibit the spectral correlation property [61, Section 4.2.2].

9.6.3.2 Discrete-time, spectral correlation estimation, and applications

SC processes can be defined in discrete time, and uniform sampling of a continuous-time SC process leads to a discrete-time SC process [61, Sections 4.8–4.9].

For SC processes that do not reduce to ACS processes, the spectral correlation density can be consistently estimated only if the location of the support curves is known [61, Section 4.7]. In contrast, if this location is unknown, the spectral correlation density can be estimated only with a degree of uncertainty (see [75], [61, Section 4.5]). The larger the departure of the nonstationarity from almost-cyclostationarity, i.e., the larger the difference of the slope of the support curve from ± 1 , the larger the amount of bias that must be tolerated to keep the variance small.

In many communication and radar/sonar problems, the narrowband condition is assumed to hold, which allows the effect of the relative motion between transmitter and receiver with constant relative radial speed to be described just as a frequency shift of the carrier, which is a very common model for the Doppler effect [61, Section 7.5]. In contrast, if the narrowband condition is not satisfied, then also a nonunit time-scale (or stretch) factor in the argument of the received complex envelope should be accounted for. In such a case, as shown in Fig. 9.6.2, a transmitted ACS signal passing through a multipath Doppler channel transforms into an SC signal with a Loève bifrequency spectrum having spectral masses concentrated on lines without necessarily achieving unit slope (see [61, Sections 4.2.4.1, 7.7.2] and [76]). Moreover, if a moving source transmits an ACS signal, the signals received on two sensors are still singularly ACS but jointly SC. The adoption of the joint SC model for the joint characterization of the received signals gives rise to significant performance improvement in source parameter estimation [77].

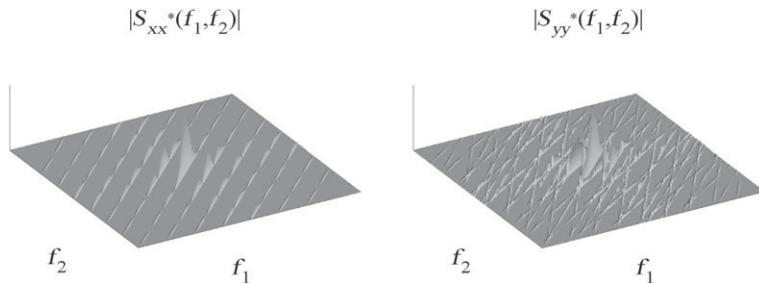


FIGURE 9.6.2

Magnitude of the density of the Loève bifrequency spectrum as a function of f_1 and f_2 for (left) the input ACS signal $x(t)$, and (right) the output SC signal $y(t)$, for the multipath Doppler channel between transmitter and receiver with constant relative radial speed for each path.

9.6.4 SUMMARY AND CONCLUSIONS

Almost-cyclostationary processes in the wide sense have first- and second-order moments that are almost-periodic functions of time. The (generalized) Fourier series expansion of the second-order moment has frequencies (not depending on the lag parameter) and coefficients (depending on the lag parameter) that are called (conjugate) cycle frequencies and cyclic autocorrelation functions, respectively. ACS processes have spectral components that are correlated when the frequency separation is equal to one of the cycle frequencies. Thus, the Loève bifrequency spectrum in the bifrequency plane is concentrated on support lines with unit slope.

The relative motion between transmitter and receiver modifies the nonstationarity of the transmitted signal and the joint nonstationary characterization of transmitted and received signals. Generalizations of cyclostationarity are suitable models to be exploited. In the case of constant relative radial speed between transmitter and receiver, a transmitted ACS signal is still ACS at the receiver, but with different cyclic features. In addition, if the narrowband condition is not satisfied, then transmitted and received signals are jointly spectrally correlated. In the presence of multipath propagation, the received signal is spectrally correlated. Its Loève bifrequency spectrum has spectral masses concentrated on a countable set of support lines in the bifrequency plane with nonunit slopes. In the case of constant relative radial acceleration, if the narrowband condition is satisfied, a transmitted ACS signal is received as generalized almost-cyclostationary. In such a case, second-order moments are almost-periodic functions of time whose (generalized) Fourier series expansion has both coefficients and frequencies depending on the lag parameter.

REFERENCES

- [1] L. Stanković, V. Katkovnik, “The Wigner distribution of noisy signals with adaptive time-frequency varying window”, *IEEE Trans. Signal Process.* 47 (4) (1999) 1099-1108.
- [2] P. Duvaut, D. Declerq, “Statistical properties of the pseudo-Wigner-Ville representation of normal random processes”, *Signal Process.* 75 (1) (1999) 93-98.

- [3] O. Kalyoncu, "Noise reduction in time-frequency domain", Master's thesis, Middle East Technical University, 2007.
- [4] T. Thayaparan, M. Daković, L. Stanković, "Mutual interference and low probability of interception capabilities of noise radar", *IET Radar Sonar Navig.* 2 (4) (2008) 294-305.
- [5] S. Djukanović, M. Daković, T. Thayaparan, L. Stanković, "Method for nonstationary jammer suppression in noise radar systems", *IET Signal Process.* 4 (3) (2010) 305-313.
- [6] I. Djurović, L. Stanković, "XWD-algorithm for the instantaneous frequency estimation revisited: statistical analysis", *Signal Process.* 94 (2014) 642-649.
- [7] L. Stanković, V. Ivanović, "Further results on the minimum variance time-frequency distributions kernels", *IEEE Trans. Signal Process.* 45 (6) (1997) 1650-1655.
- [8] M.G. Amin, "Minimum-variance time-frequency distribution kernels for signals in additive noise", *IEEE Trans. Signal Process.* 44 (9) (1996) 2352-2356.
- [9] M.G. Amin, "Spectral decomposition of time-frequency distribution kernels", *IEEE Trans. Signal Process.* 42 (5) (1994) 1156-1165.
- [10] G.S. Cunningham, W.J. Williams, "Kernel decomposition of time-frequency distributions", *IEEE Trans. Signal Process.* 42 (6) (1994) 1425-1442.
- [11] L. Stanković, "Analysis of noise in time-frequency distributions", *IEEE Signal Process. Lett.* 9 (9) (2002) 286-289.
- [12] B.G. Iem, A. Papandreou-Suppappola, G.F. Boudreux-Bartels, "Classes of smoothed Weyl symbols", *IEEE Signal Process. Lett.* 7 (7) (2000) 186-188.
- [13] W. Kozek, "On the transfer function calculus for underspread LTV channels", *IEEE Trans. Signal Process.* 45 (1) (1997) 219-223.
- [14] R.G. Shenoy, T.W. Parks, "The Weyl correspondence and time-frequency analysis", *IEEE Trans. Signal Process.* 42 (2) (1994) 318-331.
- [15] P. Flandrin, "A time-frequency formulation of optimum detection", *IEEE Trans. Acoust. Speech Signal Process.* 36 (9) (1988) 1377-1384.
- [16] G. Matz, F. Hlawatsch, "Time-frequency formulation and design of optimal detectors", in: Proceedings of the IEEE-SP International Symposium on Time-Frequency & Time-Scale Analysis, 1996, pp. 213-216.
- [17] A. Papandreou-Suppappola, R.L. Murray, B.G. Iem, G.F. Boudreux-Bartels, "Group delay shift covariant quadratic time-frequency representations", *IEEE Trans. Signal Process.* 49 (11) (2001) 2549-2564.
- [18] J.J. Zhang, A. Papandreou-Suppappola, B. Gottin, C. Ioana, "Time-frequency characterization and receiver waveform design for shallow water environments", *IEEE Trans. Signal Process.* 57 (8) (2009) 2973-2985.
- [19] M. Zhou, J.J. Zhang, A. Papandreou-Suppappola, "Hyperbolic frequency modulation for multiple users in underwater acoustic communications", in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014), 2014, pp. 3498-3502.
- [20] S.A. Kassam, *Signal Detection in Non-Gaussian Noise*, Springer, New York, 1988.
- [21] P.J. Huber, *Robust Statistics*, Wiley, New York, 1981.
- [22] V. Katkovnik, "Robust M -estimates of the frequency and amplitude of a complex-valued harmonic", *Signal Process.* 77 (1) (1999) 71-84.
- [23] V. Katkovnik, "Robust M-periodogram", *IEEE Trans. Signal Process.* 46 (11) (1998) 3104-3109.
- [24] A.A. Roenko, V.V. Lukin, I. Djurović, "An overview of the adaptive robust DFT", *EURASIP J. Adv. Signal Process.* 2010 (2010) article ID 595071, 17 pp.
- [25] J. Astola, P. Haavisto, Y. Neuvo, "Vector median filters", *Proc. IEEE* 78 (4) (1990) 678-689.
- [26] I. Djurović, V. Katkovnik, L. Stanković, "Median filter based realizations of the robust time-frequency distributions", *Signal Process.* 81 (8) (2001) 1771-1776.
- [27] I. Pitas, A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer, Dordrecht, 1990.

- [28] I. Djurović, L. Stanković, “Robust Wigner distribution with application to the instantaneous frequency estimation”, *IEEE Trans. Signal Process.* 49 (12) (2001) 2985-2993.
- [29] I. Djurović, L. Stanković, J.F. Böhme, “Estimates of the Wigner distribution in Gaussian noise environment”, *Int. J. Electron. Commun.* 56 (5) (2002) 337-340.
- [30] I. Djurović, V.V. Lukin, “Robust DFT with high breakdown point for complex-valued impulse noise environment”, *IEEE Signal Process. Lett.* 13 (1) (2006) 25-28.
- [31] N. Žarić, I. Orović, S. Stanković, “Robust time-frequency distributions with complex-lag argument”, *EURASIP J. Adv. Signal Process.* 2010 (2010) 9:1-9:10, ISSN 1110-8657, <http://dx.doi.org/10.1155/2010/879874>.
- [32] I. Djurović, L. Stanković, J.F. Böhme, “Robust L -estimation based forms of signal transforms and time-frequency representations”, *IEEE Trans. Signal Process.* 51 (7) (2003) 1753-1761.
- [33] I. Djurović, L. Stanković, M. Simeunović, “Robust time-frequency representation based on the signal normalization and concentration measures”, *Signal Process.* 104 (2014) 424-431, <http://dx.doi.org/10.1016/j.sigpro.2014.05.005>.
- [34] L. Stanković, S. Stanković, I. Orović, M.G. Amin, “Robust time-frequency analysis based on the L -estimation and compressive sensing”, *IEEE Signal Process. Lett.* 20 (5) (2013) 499-502.
- [35] P. Flandrin, “Time-dependent spectra for nonstationary stochastic processes”, in: G. Longo, B. Picinbono (Eds.), *Time and Frequency Representation of Signals and Systems*, Springer, Vienna, 1989, pp. 69-124.
- [36] M.G. Amin, “Time-frequency spectrum analysis and estimation for non-stationary random processes”, in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 208-232 (Chapter 9).
- [37] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, 1999, translated by J. Stöckler from the French editions, *Temps-fréquence* (Paris: Hermès, 1993, 1998).
- [38] P. Flandrin, W. Martin, “The Wigner-Ville spectrum of nonstationary random signals”, in: W. Mecklenbräuker, F. Hlawatsch (Eds.), *The Wigner Distribution—Theory and Applications in Signal Processing*, Elsevier, Amsterdam, 1997, pp. 211-267.
- [39] W. Kozek, F. Hlawatsch, H. Kirchauer, U. Trautwein, “Correlative time-frequency analysis and classification of nonstationary random processes”, in: Proceedings of the IEEE-SP International Symposium on Time-Frequency & Time-Scale Analysis, 1994, pp. 417-420.
- [40] G. Matz, F. Hlawatsch, “Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations”, *IEEE Trans. Inform. Theory* 52 (3) (2006) 1067-1086.
- [41] G. Matz, F. Hlawatsch, W. Kozek, “Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes”, *IEEE Trans. Signal Process.* 45 (6) (1997) 1520-1534.
- [42] W. Martin, P. Flandrin, “Wigner-Ville spectral analysis of nonstationary processes”, *IEEE Trans. Acoust. Speech Signal Process.* 33 (6) (1985) 1461-1470.
- [43] M.B. Priestley, *Spectral Analysis and Time Series—Part II*, Academic Press, London, 1981.
- [44] M.B. Priestly, “Evolutionary spectra and non-stationary processes”, *J. R. Stat. Soc. Ser. B* 27 (2) (1965) 204-237.
- [45] C.S. Detka, A. El-Jaroudi, “The transitory evolutionary spectrum”, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'94), vol. 4, 1994, pp. 289-292.
- [46] F. Hlawatsch, G. Matz, “Time-frequency methods for non-stationary statistical signal processing”, in: F. Hlawatsch, F. Auger (Eds.), *Time-Frequency Analysis: Concepts and Methods*, ISTE/Wiley, London, UK, 2008, pp. 279-320 (Chapter 10).
- [47] G. Matz, F. Hlawatsch, “Fundamentals of time-varying communication channels”, in: F. Hlawatsch, G. Matz (Eds.), *Wireless Communications over Rapidly Time-Varying Channels*, Academic Press, Oxford, UK, 2011, pp. 1-63 (Chapter 1).
- [48] P.A. Bello, “Characterization of randomly time-variant linear channels”, *IEEE Trans. Commun. Syst.* 11 (4) (1963) 360-393.

- [49] R.S. Kennedy, *Fading Dispersive Communication Channels*, Wiley, New York, 1969.
- [50] G. Matz, H. Bölcskei, F. Hlawatsch, “Time-frequency foundations of communications: concepts and tools”, *IEEE Signal Process. Mag.* 30 (6) (2013) 87-96.
- [51] J.D. Parsons, *The Mobile Radio Propagation Channel*, Pentech Press, London, 1992.
- [52] J.G. Proakis, *Digital Communications*, third ed., McGraw-Hill, New York, 1995.
- [53] H. Artés, G. Matz, F. Hlawatsch, “Unbiased scattering function estimators for underspread channels and extension to data-driven operation”, *IEEE Trans. Signal Process.* 52 (5) (2004) 1387-1402.
- [54] W.A. Gardner (Ed.), *Cyclostationarity in Communications and Signal Processing*, IEEE Press, New York, 1994.
- [55] W. Kozek, A.F. Molisch, “On the eigenstructure of underspread WSSUS channels”, in: Proceedings of the IEEE-SP Workshop on Signal Processing Advances in Wireless Communications (SPAWC’97), 1997, pp. 325-328.
- [56] J.A.C. Bingham, “Multicarrier modulation for data transmission: an idea whose time has come”, *IEEE Commun. Mag.* 28 (1990) 5-14.
- [57] W. Kozek, A.F. Molisch, “Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels”, *IEEE J. Sel. Areas Comm.* 16 (1998) 1579-1589.
- [58] W.A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [59] W.A. Gardner, *Introduction to Random Processes: With Applications to Signals and Systems*, first ed., Macmillan, New York, 1986, second ed.: McGraw-Hill, 1990.
- [60] W.A. Gardner, A. Napolitano, L. Paura, “Cyclostationarity: half a century of research”, *Signal Process.* 86 (4) (2006) 639-697.
- [61] A. Napolitano, *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*, Wiley/IEEE, New York, 2012.
- [62] H.L. Hurd, A. Miamee, *Periodically Correlated Random Sequences: Spectral Theory and Practice*, Wiley, Hoboken, NJ, 2007.
- [63] D. Dehay, H.L. Hurd, “Representation and estimation for periodically and almost periodically correlated random processes”, in: W.A. Gardner (Ed.), *Cyclostationarity in Communications and Signal Processing*, IEEE Press, New York, 1994, pp. 295-328 (Chapter 6).
- [64] A.V. Dandawaté, G.B. Giannakis, “Asymptotic theory of mixed time averages and k th-order cyclic-moment and cumulant statistics”, *IEEE Trans. Inform. Theory* 41 (1) (1995) 216-232.
- [65] A.V. Dandawaté, G.B. Giannakis, “Nonparametric polyspectral estimators for k th-order (almost) cyclostationary processes”, *IEEE Trans. Inform. Theory* 40 (1) (1994) 67-84.
- [66] P. Ciblat, P. Loubaton, E. Serpedin, G.B. Giannakis, “Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators”, *IEEE Trans. Inform. Theory* 48 (7) (2002) 1922-1934.
- [67] O.A. Dobre, A. Abdi, Y. Bar-Ness, W. Su, “Survey of automatic modulation classification techniques: classical approaches and new trends”, *IET Commun.* 1 (2) (2007) 137-156.
- [68] W.A. Gardner, C.M. Spooner, “The cumulant theory of cyclostationary time-series, Part I: foundation”, *IEEE Trans. Signal Process.* 42 (12) (1994) 3387-3408.
- [69] J. Lundén, V. Koivunen, A. Huttunen, H.V. Poor, “Collaborative cyclostationary spectrum sensing for cognitive radio systems”, *IEEE Trans. Signal Process.* 57 (11) (2009) 4182-4195.
- [70] C.M. Spooner, W.A. Gardner, “The cumulant theory of cyclostationary time-series, Part II: development and applications”, *IEEE Trans. Signal Process.* 42 (12) (1994) 3409-3429.
- [71] A. Napolitano, “Estimation of second-order cross-moments of generalized almost-cyclostationary processes”, *IEEE Trans. Inform. Theory* 53 (6) (2007) 2204-2228.
- [72] J. Antoni, “Cyclostationarity by examples”, *J. Mech. Syst. Signal Process.* 23 (4) (2009) 987-1036.
- [73] L. Izzo, A. Napolitano, “The higher-order theory of generalized almost-cyclostationary time-series”, *IEEE Trans. Signal Process.* 46 (11) (1998) 2975-2989.

- [74] L. Izzo, A. Napolitano, "Generalized almost-cyclostationary signals", in: P.W. Hawkes (Ed.), *Advances in Imaging and Electron Physics*, vol. 135, Academic Press, San Diego, 2005, pp. 103-223 (Chapter 3).
- [75] A. Napolitano, "Uncertainty in measurements on spectrally correlated stochastic processes", *IEEE Trans. Inform. Theory* 49 (9) (2003) 2172-2191.
- [76] K.-S. Lii, M. Rosenblatt, "Spectral analysis for harmonizable processes", *Ann. Stat.* 30 (1) (2002) 258-297.
- [77] A. Napolitano, "Generalizations of cyclostationarity: a new paradigm for signal processing for mobile communications, radar, and sonar", *IEEE Signal Process. Mag.* 30 (6) (2013) 53-63.

This page intentionally left blank

INSTANTANEOUS FREQUENCY ESTIMATION AND LOCALIZATION

10

INTRODUCTION AND OVERVIEW

In many applications, a critical feature of a nonstationary signal is provided by its instantaneous frequency (IF), which accounts for the signal spectral variations as a function of time. This chapter presents methods and algorithms for the localization and estimation of the signal IF using time-frequency (t,f) based methods. The topic is covered in eight sections with appropriate internal cross-referencing to this and other chapters.

In addition to filter banks and zero-crossings, one of the first conventional approaches for IF estimation used the spectrogram. To account for its major limitations related to accuracy, resolution, window dependence, and sensitivity, improvements were made by introducing iterative methodologies on the estimate provided by the first moment of the spectrogram (Section 10.1). Another approach uses an adaptive algorithm for IF estimation using the peak of suitable Time-Frequency Distributions (TFDs) with adaptive window length (Section 10.2). This method was extended to the case of multicomponent signals using high-resolution TFDs such as the modified B-distribution (Section 10.3). When the signals considered have polynomial FM characteristics, both the peak of the polynomial WVD and higher-order ambiguity functions can be used as IF estimators (Section 10.4). In the special case when the signals are subject to random amplitude modulation (or multiplicative noise), IF estimation procedures are described using the peak of the Wigner Ville Distribution (WVD) for linear Frequency Modulated (FM) signals, and the peak of the Polynomial Wigner Ville Distribution (PWVD) for nonlinear FM signals (Section 10.5). Then, a comparison of multicomponent IF estimation algorithms is provided (Section 10.6); and methods for instantaneous frequency and polynomial phase parameters estimation using linear time-frequency representations are presented (Section 10.7). Methods based on particle filtering and sequential Bayesian estimation are conceptually developed for multicomponent FM signals IF estimation (Section 10.8). Finally, Section 10.9 describes briefly the Viterbi algorithm for completeness.

10.1 ITERATIVE INSTANTANEOUS FREQUENCY ESTIMATION FOR RANDOM SIGNALS⁰

10.1.1 IF ESTIMATION: INTRODUCTION AND BACKGROUND

The instantaneous frequency (IF) is a basic parameter which may be used to describe the nonstationarity in a process (see Section 1.3). It is used in many areas such as seismic, radar, sonar, communications, and biomedical applications [1–6]; it is based on the analytic signal (see Section 1.3).

The concept of IF is often introduced in the context of frequency modulation in communication theory. The IF of a signal at time t_0 could be viewed as the frequency of the sinusoid which locally fits the signal at that time [1]. Using the definition of Gabor and later Ville [7], the IF, $f(t)$, of a signal $x(t)$ can be defined in terms of the derivative of the phase of its analytic signal $z(t)$; i.e.,

$$f(t) = \frac{1}{2\pi} \frac{d}{dt} \arg(z(t)). \quad (10.1.1)$$

Another interpretation of the IF uses a time-frequency distribution (TFD) and defines the IF as the weighted average of the frequencies existing in the signal at time t [1,8]; i.e.,

$$\hat{f}(t) = \frac{\int_{-\infty}^{\infty} f \hat{\rho}(t, f) df}{\int_{-\infty}^{\infty} \hat{\rho}(t, f) df}, \quad (10.1.2)$$

where $\hat{\rho}(t, f)$ is the TFD estimate.

The IF has many applications in time-frequency signal analysis. For example a matched spectrogram, which uses chirp windows, is shown to perform better than a regular spectrogram for chirp-like signals when the chirp rate of the window is matched to the IF of the signal [9]. The IF can also be utilized to obtain distributions mainly concentrated along the IF for monocomponent signals. For example, Ref. [10] shows how to obtain distributions infinitely concentrated along the IF for pure tones and for amplitude- and frequency-modulated signals, and considers kernel functions for bilinear distributions with these properties. Another method [11] describes how to utilize the IF to iteratively generate a time-frequency representation with satisfactory energy concentration.

There are many algorithms for estimating the IF of a signal. Current algorithms can be grouped as phase differencing methods, signal modeling methods (e.g., short-time AR modeling, time-varying AR modeling), phase modeling methods, and time-frequency methods. A good review of these methods can be found in Ref. [4].

One important time-frequency method is the iterative cross Wigner-Ville IF estimation technique [4], in which, at each iteration, we construct a constant-amplitude FM signal whose IF matches the previous IF estimate, then take the cross Wigner-Ville distribution between the constructed signal and the signal under analysis, and then use the cross-WVD to produce a new IF estimate. This method is shown to perform very well at low signal-to-noise ratios (cf. Section 3.1.2.3, especially at item 2).

In the next two sections, we present an iterative algorithm to estimate the IF and matched spectrogram of a nonstationary sinusoidal signal, and then analyze the convergence of the algorithm. Our algorithm is similar to the cross Wigner-Ville algorithm in that we use a TFD iteratively to obtain

⁰Authors: **A. El-Jaroudi**, University of Pittsburgh, Pittsburgh, PA, USA (amro@ee.pitt.edu) and **M.K. Emresoy**, University of Pittsburgh, Pittsburgh, PA, USA (emresoy@ee.pitt.edu). Reviewers: S. Ouelha and G.R. Putland.

an estimate of the IF. The iterative IF estimation method can be used to improve the performance of many time-frequency methods [12,13].

10.1.2 ITERATIVE ALGORITHM FOR IF ESTIMATION

The basic idea behind the iterative IF estimation is to use a TFD (in this case, the spectrogram) to obtain an IF estimate using Eq. (10.1.2) above, then use this IF estimate to recalculate the spectrogram. This process is repeated until convergence is reached. The principle at work here is that improving the IF estimate makes the matched spectrogram estimate better and vice versa. Consequently each iteration will lead to improvements in both quantities.

We will assume that the signal under analysis is a complex signal of the form

$$x(t) = a(t)e^{j\phi(t)}, \quad (10.1.3)$$

where $a(t)$ is the time-varying amplitude and $\phi(t)$ is the phase of the signal.

In the iterative algorithm given below, the instantaneous frequency and the spectrogram as a TFD are estimated by using the whole frequency axis. Then the signal is “demodulated” by subtracting the phase estimate (obtained by integrating the IF estimate) from the phase of the signal. This demodulation shifts the spectrum of the nonstationary signal around zero frequency. The demodulated signal is then reanalyzed. Consequently we do not restrict the definition of the IF to use only positive frequencies, since after the first iteration the signal under analysis has nonredundant information over positive and negative frequencies. But when the algorithm is applied to real signals, we use the analytic signal to start the iterations, ensuring that the first IF estimate is not trivially zero. Note that although we use the analytic signal at first, the demodulated signal used in the subsequent iteration is not analytic. The outputs of the algorithm are the estimated IF and the matched spectrogram of the signal.

Algorithm:

1. Take the spectrogram $S_z(t,f)$ of the analytic component of the signal.
2. Estimate $\hat{f}_i(t)$ using Eq. (10.1.2) and find the phase estimate

$$\hat{\phi}_i(t) = 2\pi \int_{-\infty}^t \hat{f}_i(\tau) d\tau. \quad (10.1.4)$$

3. Demodulate the signal along the estimated IF:

$$\tilde{x}_i(t) = a(t)e^{-j(\phi(t)-\hat{\phi}_i(t))} = a(t)e^{-j\tilde{\phi}_i(t)}. \quad (10.1.5)$$

4. Take the spectrogram $S_{xi}(t,f)$ of $\tilde{x}(t)$ and compensate for the demodulation to obtain a new matched spectrogram estimate

$$S_{zi+1}(t,f) = S_{xi}(t,f - \hat{f}_i(t)). \quad (10.1.6)$$

5. Go to step 2.

The algorithm is stopped when the difference between consecutive estimates of the IF is within a specified tolerance.

While the algorithm seems very simple, we show below that its convergence can be guaranteed given certain conditions on the analysis signal.

10.1.3 CONVERGENCE OF THE ESTIMATION ALGORITHM

In order to examine the convergence of the iterative algorithm, we derive an expression for the IF estimate after each iteration. The IF estimate of the complex signal in Eq. (10.1.3) obtained by using the first moment of a spectrogram with a real analysis window, $h(t)$, is given by [1,14]

$$\hat{f}(t) = \frac{1}{2\pi} \sum_{k=0}^{\infty} \frac{M_k(t)}{k!} \phi^{(k+1)}(t), \quad (10.1.7)$$

where

$$M_n(t) = \frac{\int a^2(t+\tau)h^2(\tau)\tau^n d\tau}{\int a^2(t+\tau)h^2(\tau) d\tau} \quad (10.1.8)$$

and $\phi^{(n+1)}(t)$ is the $(n+1)$ th derivative of the phase of the signal, with

$$\frac{M_0(t)}{0!} = 1. \quad (10.1.9)$$

Our hypothesis is that the general expression for the IF estimate at the end of the n th iteration is

$$\hat{f}_n(t) = f(t) + (-1)^{(n+1)} \frac{1}{2\pi} \sum_{i_n=1}^{\infty} \dots \sum_{i_1=1}^{\infty} \frac{M_{i_n}(t)}{i_n!} \left\{ \dots \left\{ \frac{M_{i_1}(t)}{i_1!} \phi^{(i_1+1)}(t) \right\}^{(i_2)} \dots \right\}^{(i_n)}, \quad (10.1.10)$$

where $f(t) = \frac{1}{2\pi} \phi'(t)$. We prove the hypothesis by induction.

To show that the Eq. (10.1.10) is true for the first iteration, we remove the summations i_2 to i_n and substitute $n = 1$ in Eq. (10.1.10), obtaining

$$\hat{f}_1(t) = f(t) + \frac{1}{2\pi} \sum_{i_1=1}^{\infty} \frac{M_{i_1}(t)}{i_1!} \phi^{(i_1+1)}(t). \quad (10.1.11)$$

This is the same as Eq. (10.1.7); i.e., Eq. (10.1.10) is true for $n = 1$.

For the inductive step, we assume that Eq. (10.1.10) is true for iteration $(n - 1)$ so that

$$\hat{f}_{n-1}(t) = f(t) + (-1)^n \frac{1}{2\pi} \sum_{i_{n-1}=1}^{\infty} \dots \sum_{i_1=1}^{\infty} \frac{M_{i_{n-1}}(t)}{i_{n-1}!} \left\{ \dots \left\{ \frac{M_{i_1}(t)}{i_1!} \phi^{(i_1+1)}(t) \right\}^{(i_2)} \dots \right\}^{(i_{n-1})}, \quad (10.1.12)$$

and show that it is also true for the n th iteration. Following the algorithm given in the previous section, the phase error $\tilde{\phi}(t)$ at the n th iteration is given by

$$\tilde{\phi}_n(t) = \phi(t) - 2\pi \int_{-\infty}^t \hat{f}_{n-1}(\tau) d\tau. \quad (10.1.13)$$

By combining the previous two equations, we obtain

$$\tilde{\phi}_n(t) = (-1)^{(n+1)} \sum_{i_{n-1}=1}^{\infty} \dots \sum_{i_1=1}^{\infty} \int_{-\infty}^t \frac{M_{i_{n-1}}(\tau)}{i_{n-1}!} \left\{ \dots \left\{ \frac{M_{i_1}(\tau)}{i_1!} \phi^{(i_1+1)}(\tau) \right\}^{(i_2)} \dots \right\}^{(i_{n-1})} d\tau. \quad (10.1.14)$$

The IF estimate from iteration n is the sum of the IF estimate from iteration $n - 1$ and the IF estimate of the signal $\tilde{x}(t)$ whose phase is given by Eq. (10.1.14); that is,

$$\hat{f}_n(t) = \hat{f}_{n-1}(t) + \frac{1}{2\pi} \sum_{i_n=0}^{\infty} \frac{M_{i_n}(t)}{i_n!} \tilde{\phi}_n^{(i_n+1)}(t). \quad (10.1.15)$$

Substituting for $\hat{f}_{n-1}(t)$ and $\tilde{\phi}_n(t)$ in the above equation, we find that the IF estimate at the end of the n th iteration is

$$\begin{aligned} \hat{f}_n(t) &= f(t) + (-1)^n \frac{1}{2\pi} \sum_{i_{n-1}=1}^{\infty} \dots \sum_{i_1=1}^{\infty} \frac{M_{i_{n-1}}(t)}{i_{n-1}!} \left\{ \dots \left\{ \frac{M_{i_1}(t)}{i_1!} \phi^{(i_1+1)}(t) \right\}^{(i_2)} \dots \right\}^{(i_{n-1})} \\ &\quad + (-1)^{(n+1)} \sum_{i_n=0}^{\infty} \dots \sum_{i_1=1}^{\infty} \frac{M_{i_n}(t)}{i_n!} \int_{-\infty}^t \left\{ \dots \left\{ \frac{M_{i_1}(\tau)}{i_1!} \phi^{(i_1+1)}(\tau) \right\}^{(i_2)} \dots \right\}^{(i_n+1)} d\tau, \end{aligned}$$

which simplifies to Eq. (10.1.10) as claimed. This completes the proof.

In Eq. (10.1.10), the first term represents the actual IF whereas the second term represents the error in the IF estimate at the end of the n th iteration. In general, the error term is very difficult to analyze, being affected by both the amplitude and phase variations in the signal. Assuming, however, that the spectrogram window is Gaussian, and that the phase function of the signal has bounded derivatives of all orders, and that the signal is dominantly FM modulated (i.e., has constant or slowly varying amplitude), we may show that the iteration in Eq. (10.1.10) converges to the true IF of the signal. Let the spectrogram window be

$$h(t) = \left(\frac{\alpha}{\pi} \right)^{1/4} e^{-\alpha t^2/2}. \quad (10.1.16)$$

Let the amplitude within the window be approximated by a constant a so that

$$x(t) = ae^{j\phi(t)}. \quad (10.1.17)$$

Making these two substitutions in Eq. (10.1.8), we find

$$M_n = \frac{(1 + (-1)^n)\gamma(n + 1/2)}{2\sqrt{\pi}\alpha^{n/2}} \quad (10.1.18)$$

$$= \begin{cases} 0 & \text{if } n \text{ is odd} \\ \prod_{k=0}^{n/2} \frac{(n+(2k-1))\sqrt{\pi}}{\alpha^{n/2}} & \text{if } n \text{ is even.} \end{cases} \quad (10.1.19)$$

Then, from Eq. (10.1.10), the general expression for the IF estimate is

$$\hat{f}_n(t) = f(t) + (-1)^n \frac{1}{2\pi} \sum_{i_1, i_2, \dots, i_n=2}^{\infty} \frac{M_{i_1} M_{i_2}, \dots, M_{i_n}}{i_1! i_2!, \dots, i_n!} \phi^{(i_1+i_2+\dots+i_n+1)}(t). \quad (10.1.20)$$

If we define

$$c_m = \sum_{i_1+i_2+\dots+i_n=m} \frac{M_{i_1} M_{i_2}, \dots, M_{i_n}}{i_1! i_2!, \dots, i_n!}, \quad (10.1.21)$$

then Eq. (10.1.20) can be written as

$$\hat{f}_n(t) = f(t) + \frac{(-1)^n}{2\pi} \sum_{m=2}^{\infty} c_m \phi^{(m+1)}(t). \quad (10.1.22)$$

Note that the $\{c_m\}$ are independent of the analysis signal and are mainly a function of the window parameter α . It is easy to show that after each iteration, two coefficients of the coefficient series $\{c_m\}$ become zero and overall the coefficients decay rapidly to zero. It is clear that, as $n \rightarrow \infty$, the coefficients go to zero. If we assume that the phase of the signal is continuous and has bounded derivatives, then the sum in Eq. (10.1.22) also goes to zero which implies $\hat{f}_n(t) \rightarrow f(t)$.

10.1.3.1 Convergence properties

- Equation (10.1.19) implies that the convergence rate of the algorithm is proportional to $\alpha^{n/2}$. That is, for small α (long window) the convergence is slow whereas for large α (short window) the convergence is very fast.
- The convergence rate is also affected by the rate of change of the IF. For signals whose IFs have nonzero higher-order derivatives, the convergence is slower. But for signals with polynomial phases, the algorithm theoretically converges with an iteration number equal to half of the degree of the polynomial. This can be seen by examining Eq. (10.1.20). At iteration n the first $2n$ coefficients in Eq. (10.1.22) become zero.
- If the phase of the signal has discontinuities at some points t_i , the derivatives of $\phi(t)$ go to $\pm\infty$. Then we cannot expect the algorithm to converge exactly to the $f(t)$. If $\phi(t)$ has bounded derivatives for all orders and for all t , then the algorithm converges to $f(t)$.

10.1.4 SUMMARY AND CONCLUSIONS

We have presented an iterative algorithm for estimating the instantaneous frequency (IF) and matched spectrogram of a nonstationary sinusoidal signal. For a monocomponent signal, the matched spectrogram thus obtained is concentrated along the IF. The estimated IF is shown to converge to the true IF under reasonable conditions.

10.2 ADAPTIVE INSTANTANEOUS FREQUENCY ESTIMATION USING TFDs⁰

Instantaneous frequency (IF) estimators based on maxima of time-frequency representations have variance and bias which are highly dependent on the lag window width. The optimal window width may be determined by minimizing the estimation mean squared error (MSE), provided that some signal and noise parameters are explicitly known. However, these parameters are not available in advance. This is especially true for the IF derivatives which determine the estimation bias. In this section, an adaptive algorithm for the lag window width determination, based on the confidence intervals intersection, will be presented [15–19]. This algorithm does not require knowledge of the estimation bias value.

⁰Author: **LJ. Stanković**, University of Montenegro, Podgorica, Montenegro (ljubisa@ac.me). Reviewers: B. Barkat, I. Djurović, and V. Ivanović.

10.2.1 OPTIMAL WINDOW WIDTH

Consider a noisy signal:

$$x(n\Delta t) = s(n\Delta t) + \epsilon(n\Delta t), \quad s(t) = a \exp(j\phi(t)), \quad (10.2.1)$$

with $s(n\Delta t)$ being a signal and $\epsilon(n\Delta t)$ being a white complex-valued Gaussian noise with mutually independent real and imaginary parts of equal variances $\sigma_\epsilon^2/2$. The sampling interval is denoted by Δt . Consider the problem of estimating the IF, $f_i(t) = \phi'(t)/2\pi$, from the discrete-time observations $x(n\Delta t)$, based on maxima of a time-frequency distribution $\rho_x(t, f)$:

$$\hat{f}(t) = \arg \max_f \rho_x(t, f). \quad (10.2.2)$$

Let $\Delta\hat{f}(t) = f_i(t) - \hat{f}(t)$ be the estimation error. The MSE, $E\{\Delta\hat{f}(t)\}^2$, is used for the accuracy characterization at a given time instant t . Asymptotically, the MSE for commonly used time-frequency representations (e.g., the spectrogram, the pseudo Wigner distribution (WD), and its higher order versions) can be expressed in the following form (see Refs. [1,16–19] and Sections 10.3 and 10.4):

$$E\left\{(\Delta\hat{f}(t))^2\right\} = \frac{V}{h^m} + B(t)h^n, \quad (10.2.3)$$

where h is the width of the lag window $w_h(t)$ so that $w_h(t) = 0$ for $|t| > h/2$, and is related to the number of samples N by $h = N\Delta t$. The variance and the bias of estimate, for a given h , are

$$\sigma^2(h) = V/h^m, \quad \text{bias}(t, h) = \sqrt{B(t)h^n}. \quad (10.2.4)$$

The expression for $B(t)$ is a function of the IF derivatives.

For example, for the WD with a rectangular lag window we have [17]

$$E\left\{\left(\Delta\hat{f}(t)\right)^2\right\} = \frac{6\sigma_\epsilon^2\Delta t}{(2\pi a)^2} \frac{1}{h^3} + \left(\frac{\phi^{(3)}(t)}{80\pi}\right)^2 h^4, \quad (10.2.5)$$

corresponding to $m = 3$ and $n = 4$ in Eq. (10.2.3). Values of m and n for some other distributions are indicated in Table 10.2.1, according to the results from Refs. [16–20].

The MSE in Eq. (10.2.3) has a minimum with respect to h . This minimum occurs for the optimal value of h given by

$$h_{\text{opt}}(t) = \left[mV / (nB(t)) \right]^{1/(m+n)}. \quad (10.2.6)$$

Note that this relation is not useful in practice because its right hand-side contains $B(t)$ which depends on derivatives of the unknown IF.

10.2.2 ADAPTIVE ALGORITHM

Here, we present an adaptive method which can produce an estimate of $h_{\text{opt}}(t)$ without having to know the value of $B(t)$. For the optimal window width, by differentiating Eq. (10.2.3), we obtain

$$\frac{\partial E\left\{(\Delta\hat{f}(t))^2\right\}}{\partial h} = -m \frac{V}{h^{m+1}} + nB(t)h^{n-1} = 0 \Big|_{h=h_{\text{opt}}}. \quad (10.2.7)$$

**Table 10.2.1: Parameters in the Adaptive Algorithm for Various m, n ,
 κ : $m = 3, n = 4$ for the IF Estimators Based on the Spectrogram, Wigner and L-Wigner Distributions; $m = 3, n = 8$ for the IF Estimators Based on the Fourth-Order Polynomial Wigner-Ville Distribution and Local Polynomial Distribution; $m = 1, n = 4$ for the Wigner Distribution as a Spectrum Estimator**

| | | | | | | | | |
|----------------|------|------|------|------|------|-------|-------|------|
| m | 1 | 1 | 3 | 3 | 3 | 3 | 3 | 3 |
| n | 4 | 4 | 4 | 4 | 4 | 8 | 8 | 8 |
| κ | 2 | 3 | 2 | 3 | 5 | 2 | 3 | 5 |
| $\Delta\kappa$ | 0.86 | 1.29 | 0.39 | 0.58 | 0.97 | 0.09 | 0.14 | 0.23 |
| p | 0.99 | 1.22 | 0.34 | 0.51 | 0.72 | -0.13 | -0.03 | 0.11 |
| p_1 | 1.18 | 1.41 | 0.59 | 0.76 | 0.97 | 0.19 | 0.30 | 0.43 |

Multiplying Eq. (10.2.7) by h , we get the relationship between the bias and standard deviation, Eq. (10.2.4), for $h = h_{\text{opt}}$:

$$\text{bias}(t, h_{\text{opt}}) = \sqrt{\frac{m}{n}} \sigma(h_{\text{opt}}). \quad (10.2.8)$$

It will be assumed, without loss of generality, that the bias is positive. The IF estimate $\hat{f}_h(t)$ (obtained from Eq. (10.2.2) by using the lag window of width h) is a random variable distributed around the true IF $f_i(t)$ with the bias “ $\text{bias}(t, h)$ ” and the standard deviation $\sigma(h)$. Thus, we may write the relation

$$\left| f_i(t) - (\hat{f}_h(t) - \text{bias}(t, h)) \right| \leq \kappa \sigma(h), \quad (10.2.9)$$

where the inequality holds with probability $P(\kappa)$ depending on parameter κ .¹ We will assume that κ is such that $P(\kappa) \rightarrow 1$.

Let us introduce a set of discrete dyadic window-width values, $h \in H$,

$$H = \{h_s \mid h_s = 2h_{s-1}, s = 1, 2, \dots, J\}. \quad (10.2.10)$$

Define the confidence intervals $D_s = [L_s, U_s]$ of the IF estimates, with the following upper and lower bounds

$$L_s = \hat{f}_{h_s}(t) - (\kappa + \Delta\kappa) \sigma(h_s), \quad U_s = \hat{f}_{h_s}(t) + (\kappa + \Delta\kappa) \sigma(h_s), \quad (10.2.11)$$

where $\hat{f}_{h_s}(t)$ is an estimate of the IF, for the window width $h = h_s$, and $\sigma(h_s)$ is its standard deviation. Assume that a window width denoted by $h_{s+} \in H$ is of h_{opt} order, $h_{s+} \sim h_{\text{opt}}$. Since h_{opt} does not correspond to any h_s from the set H , for the analysis that follows we can write $h_{s+} = 2^p h_{\text{opt}}$, where p

¹If we assume, for example, that the random variable $\hat{f}_h(t)$ is Gaussian, with the mean value $M = f_i(t) + \text{bias}(t, h)$ and the standard deviation $\sigma(h)$, then the probability that $\hat{f}_h(t)$ takes a value in the interval $[M - \kappa\sigma(h), M + \kappa\sigma(h)]$ is $P(\kappa) = 0.95$ for $\kappa = 2$, and $P(\kappa) = 0.997$ for $\kappa = 3$.

is a constant close to 0. According to Eq. (10.2.10), all other windows can be written as a function of h_{s^+} as

$$h_s = h_{s^+} 2^{(s-s^+)} = h_{\text{opt}} 2^{s-s^++p}, (s-s^+) = \dots, -2, -1, 0, 1, 2, \dots \quad (10.2.12)$$

With this notation, having in mind Eq. (10.2.8), the standard deviation and the bias from Eq. (10.2.4) can be expressed by

$$\begin{aligned} \sigma(h_s) &= \sqrt{V/h_s^m} = \sigma(h_{\text{opt}}) 2^{-(s-s^++p)m/2}, \\ \text{bias}(t, h_s) &= \sqrt{B(t)h_s^n} = \sqrt{m/n}\sigma(h_{\text{opt}}) 2^{(s-s^++p)n/2}. \end{aligned} \quad (10.2.13)$$

For small window widths h_s , when $s \ll s^+$, the bias of $\hat{f}_{h_s}(t)$ is negligible; thus $f_i(t) \in D_s$ (with probability $P(\kappa + \Delta\kappa) \rightarrow 1$). Then, obviously, $D_{s-1} \cap D_s \neq \emptyset$, since at least the true IF, $f_i(t)$, belongs to both confidence intervals. For $s \gg s^+$ the variance is small, but the bias is large. It is clear that for $\text{bias}(t, h_s) \neq 0$ there exists such a large s that $D_s \cap D_{s+1} = \emptyset$ for a finite $\kappa + \Delta\kappa$.

The **idea behind the algorithm** is that $\Delta\kappa$ in D_s can be found in such a way that the largest s , for which the sequence of the pairs of the confidence intervals D_{s-1} and D_s has at least a point in common, is $s = s^+$. Such a value of $\Delta\kappa$ exists because the bias and the variance are monotonically increasing and decreasing functions of h , respectively, (10.2.13). As soon as this value of $\Delta\kappa$ is found, an intersection of the confidence intervals D_{s-1} and D_s ,

$$\left| \hat{f}_{h_{s-1}}(t) - \hat{f}_{h_s}(t) \right| \leq (\kappa + \Delta\kappa) [\sigma(h_{s-1}) + \sigma(h_s)], \quad (10.2.14)$$

works as an indicator of the event $s = s^+$, i.e., the event $h_s = h_{s^+} \sim h_{\text{opt}}$. The value of h_{s^+} is the last h_s when Eq. (10.2.14) is still satisfied.

10.2.2.1 Parameters in the adaptive algorithm

There are three possible approaches to choosing the algorithm's parameters κ , $\Delta\kappa$, and p . Their performances do not differ significantly.

(1) When our knowledge about the variance and bias behavior, given by Eq. (10.2.3), is not quite reliable, an approximate approach for κ , $\Delta\kappa$, and p determination can be used. Then, we can assume a value of $\kappa \cong 2.5$ such that $P(\kappa) \cong 0.99$ for Gaussian distribution of estimation error. The value of $\Delta\kappa$ should take into account the bias for the expected optimal window width (10.2.8). It is common to assume that, for the optimal value of h , the bias and variance are of the same order, resulting in $\Delta\kappa \cong 1$. Then we can expect that the obtained value h_{s^+} is close to h_{opt} , thus $p \cong 0$, and all parameters for the key algorithm (10.2.14) are defined. This simple heuristic form has been successfully used in Refs. [16, 17], and it is highly recommended for most of practical applications. Estimation of the standard deviation $\sigma(h_s)$ will be discussed within the numerical example.

(2) When the knowledge about the variance and bias behavior is reliable, i.e., when Eq. (10.2.3) accurately describes estimation error, then we can calculate all algorithm parameters. According to the basic idea of the algorithm, only three confidence intervals, D_{s^+-1} , D_{s^+} , and D_{s^++1} , should be considered. The confidence intervals D_{s^+-1} and D_{s^+} should have, while D_{s^+} and D_{s^++1} should not have, at least one point in common. Assuming that relation Eq. (10.2.9) holds, and that the bias is positive, this condition means that the minimum possible value of the upper D_{s^+-1} bound, Eq. (10.2.11), denoted by $\min\{U_{s^+-1}\}$, is always greater than or equal to the maximum possible value of

the lower D_{s^+} bound, denoted by $\max\{L_{s^+}\}$, i.e., $\min\{U_{s^+-1}\} \geq \max\{L_{s^+}\}$. The condition that D_{s^+} and D_{s^++1} do not intersect is given by $\max\{U_{s^+}\} < \min\{L_{s^++1}\}$. According to Eqs. (10.2.9) and (10.2.11) the above analysis results in

$$\begin{aligned} \text{bias}(h_{s^+-1}) + \Delta\kappa\sigma(h_{s^+-1}) &\geq \text{bias}(h_{s^+}) - \Delta\kappa\sigma(h_{s^+}), \\ \text{bias}(h_{s^+}) + (2\kappa + \Delta\kappa)\sigma(h_{s^+}) &< \text{bias}(h_{s^++1}) - (2\kappa + \Delta\kappa)\sigma(h_{s^++1}). \end{aligned} \quad (10.2.15)$$

Since the inequalities are written for the worst case, we can calculate the algorithm parameters by using the corresponding equalities. With Eq. (10.2.13) we get

$$\begin{aligned} \Delta\kappa &= 2\kappa/[2^{(m+n)/2} - 1], \\ 2^p &= \left[\Delta\kappa \sqrt{n/m} \left(2^{m/2} + 1 \right) / \left(1 - 2^{-n/2} \right) \right]^{2/(m+n)}. \end{aligned} \quad (10.2.16)$$

Values of the parameters $\Delta\kappa$ and p for various distributions, i.e., for various values of m and n , are given in Table 10.2.1.

For further and very **fine tuning of the algorithm parameters**, one may want that the adaptive window is unbiased in logarithmic instead of in linear scale (due to definition (10.2.10)). The estimation bias and variance are exponential functions with respect to m and n as in Eq. (10.2.13). Thus the confidence interval limits vary as $2^{(s-s^+)(m+n)/2}$. The mean value for this exponential function, for two successive confidence intervals, for example $(s - s^+) = 0$ and $(s - s^+) = 1$, is $(1 + 2^{(m+n)/2})/2$. It is shifted with respect to the geometrical mean $\sqrt{2^{(m+n)/2}}$ of these two intervals, by approximately $\Delta p \cong [\log_2((1 + 2^{(m+n)/2})/2)] \frac{2}{m+n} - \frac{1}{2}$, resulting in the total logarithmic shift $p_1 = p + \Delta p$, presented in Table 10.2.1. Therefore the adaptive window width (as an estimate of the optimal window width) should be $\hat{h}_{\text{opt}} = h_{s^+}/2^{p_1}$.

Note that the set H of window widths h is assumed *a priori*. Therefore, as long as we can calculate p_1 , we can use it in the following ways: (a) To calculate the distribution with the new window width $h_a = h_{s^+}/2^{p_1}$ as the best estimate of h_{opt} ; (b) To remain within the assumed set of $h_s \in H$, and to decide only whether to correct the obtained h_{s^+} or not. For example, if $|p_1| \leq 1/2$, the correction is smaller than the window discretization step. Thus, we can use $h_a = h_{s^+}$. For $1/2 < p_1 \leq 3/2$, it is better to use $h_a = h_{s^+}/2 = h_{s^+-1}$, as the adaptive window width value. Fortunately, the loss of accuracy for the adaptive widths h_a , as far as they are of h_{opt} order, is not significant since the MSE varies slowly around its stationary point. Thus, in numerical implementations we can use only the lag windows from the given set H .

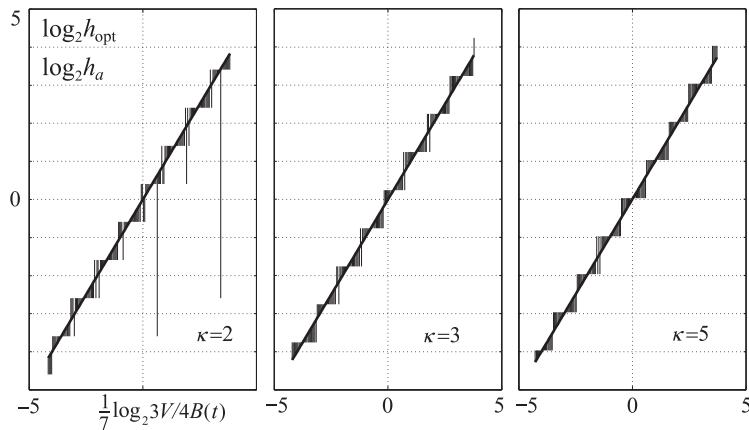
10.2.2.1.1 Illustration

We have simulated the IF estimates as a random variable

$$\hat{f}_h(t) = \mathbf{a} \sqrt{V/h^m} + \sqrt{B(t)h^n} + f_i(t), \quad (10.2.17)$$

having the MSE given by Eq. (10.2.3), where $\mathbf{a} = \mathcal{N}(0, 1)$ is a Gaussian (zero-mean, unity-variance) random variable, $m = 3$, $n = 4$, and $V = 1$. For the true IF value $f_i(t)$, at a given t , any constant can be assumed. The bias parameter $B(t)$ in $\hat{f}_h(t)$ logarithmically varies within $\frac{1}{7} \log_2(mV/nB(t)) \in [-4, 3]$, with step 0.05.

For each value of parameter $B(t)$, we have calculated the optimal window width according to Eq. (10.2.6), and plotted $\log_2(h_{\text{opt}})$ as a thick line in Fig. 10.2.1.

**FIGURE 10.2.1**

Optimal window width (straight thick line) and adaptive window widths (ends of vertical lines, starting from the optimal-window-width line) for $m = 3$, $n = 4$, $V = 1$. The variance to bias ratio $V/B(t)$ is logarithmically varied. The adaptive width $h_a = h_{s^+}/2^{p_1}$ is obtained from h_{s^+} , according to Eq. (10.2.14), after correction for the corresponding values of p_1 given in Table 10.2.1.

The value of $\hat{f}_h(t)$ was simulated for each $B(t)$ and $h_s \in H$. The assumed set of possible window widths was $H = \{1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8, 16, 32\}$, and $\kappa = 2$. The key algorithm relation (10.2.14) was tested each time, with the known standard deviation $\sigma(h_s) = \sqrt{V/h_s^m}$. The largest value of h_s when the key Eq. (10.2.14) was still satisfied was denoted by h_{s^+} . Value $\Delta\kappa = 0.39$, corresponding to $m = 3$, $n = 4$, $\kappa = 2$, was used (Table 10.2.1). The adaptive values $h_a = h_{s^+}/2^{p_1}$, $p_1 = 0.59$ (Table 10.2.1), produced in this way, are connected with the optimal window line by thin vertical lines in Fig. 10.2.1. The same simulation is repeated with $\kappa = 3$ and $\kappa = 5$. We can conclude that the presented algorithm almost always chooses the width h_s from H which is the nearest to the optimal one. However, for $\kappa = 2$ (a relatively small value), there are a few complete misses of the optimal window width, since Eq. (10.2.9) is satisfied only with probability $P(2) = 0.95$. For $\kappa = 2$, two successive confidence intervals do not intersect when the bias is small, producing a false result, with probability of $2(0.05)^2 \sim 10^{-2}$ order.

(3) The third approach for estimating the parameter $\kappa + \Delta\kappa$ is based on the statistical nature of confidence intervals, and *a posteriori* checking of the fitting quality [16]. This approach is beyond the scope of this section.

10.2.3 NUMERICAL EXAMPLE

In the example we assumed a signal of the form Eq. (10.2.1), with the given IF,

$$f_i(n\Delta t) = \frac{128}{\pi} \arctan(250(n\Delta t - 0.5)) + 128,$$

and the phase $\phi(n\Delta t) = 2\pi\Delta t \sum_{m=0}^n f_i(m\Delta t)$. The signal amplitude was $a = 1$, and $20 \log(a/\sigma_\epsilon) = 10$ [dB] ($a/\sigma_\epsilon = 3.16$). The time interval considered was $0 \leq n\Delta t \leq 1$, with $\Delta t =$

1/1024. The IF is estimated by using the discrete WD with a rectangular lag-window, $W_x^h(t,f) = \text{DFT}_{n \rightarrow f} [w_h(n\Delta t)x(t + n\Delta t)x^*(t - n\Delta t)]$, calculated with the standard FFT routines.

The algorithm is implemented as follows:

- (1) A set H of window widths h_s , corresponding to the following number of signal samples $N = \{4, 8, 16, 32, 64, 128, 256, 512\}$, is assumed. In order to have the same number of frequency samples, as well as to reduce the quantization error, all windows are zero-padded up to the maximal window width.
- (2) For a given time instant $t = n\Delta t$, the WDs are calculated starting from the smallest toward the wider window widths h_s .
- (3) The IF is estimated using Eq. (10.2.2) and $W_x^{h_s}(t,f)$.
- (4) The confidence intervals intersection, Eq. (10.2.14), is checked for the estimated IF, $\hat{f}_{h_s}(t)$, and $\sigma(h_s) = \sqrt{3\sigma_\epsilon^2\Delta t/(2\pi^2a^2h_s^3)}$ with, for example, $\kappa + \Delta\kappa = 6$, when $p_1 \cong 1$, and $P(\kappa) \rightarrow 1$ (see Table 10.2.1, and the comment that follows).
- (5) The adaptive window width $h_a = h_{s+}/2$ is obtained from the last $h_s = h_{s+}$ when Eq. (10.2.14) is still satisfied. Back to (2).

Comment: Estimation of the signal and noise parameters a and σ_ϵ^2 can be done by using $|\hat{a}|^2 + \hat{\sigma}_\epsilon^2 = \frac{1}{N} \sum_{n=1}^N |x(n\Delta t)|^2$. The variance is estimated by $\hat{\sigma}_\epsilon^2 = \hat{\sigma}_{\epsilon r}^2 + \hat{\sigma}_{\epsilon i}^2$, where

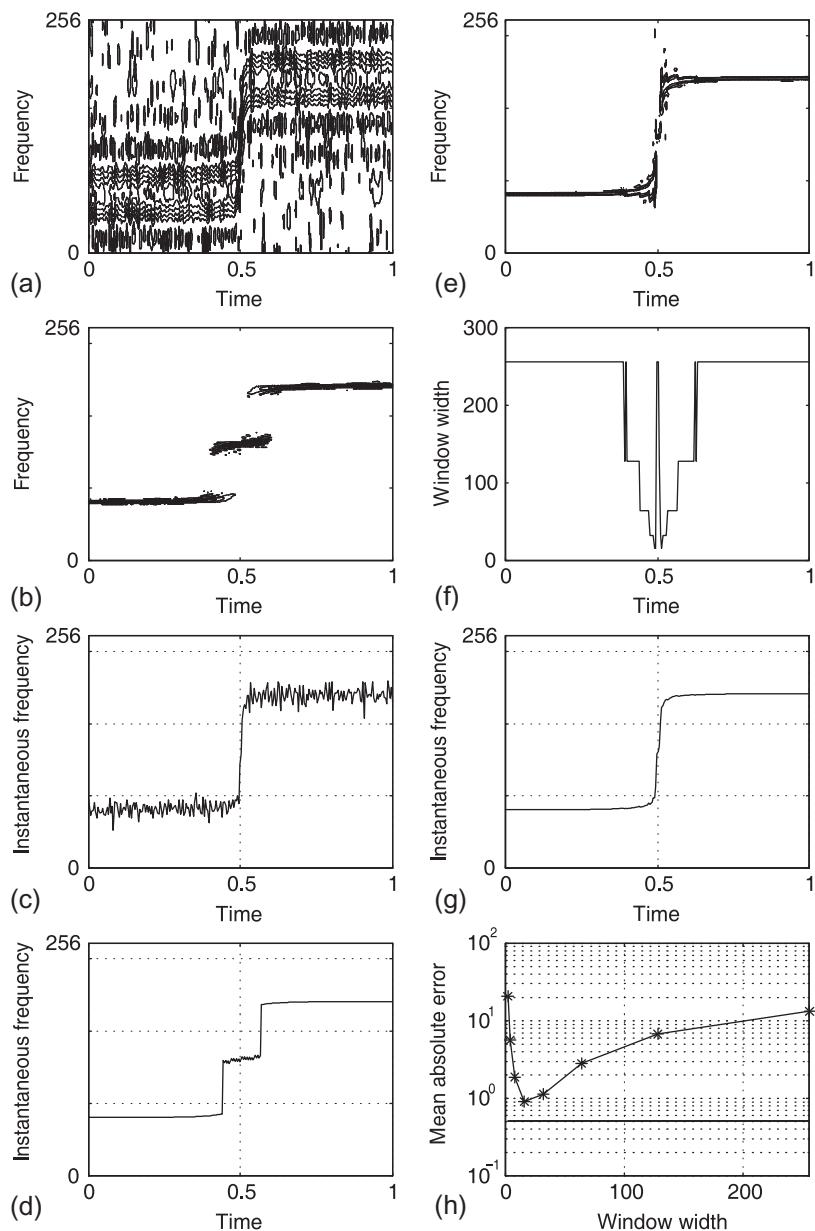
$$\hat{\sigma}_{\epsilon r,i} = \{ \text{median}(|x_{r,i}(n\Delta t) - x_{r,i}((n-1)\Delta t)| : n = 2, \dots, N) \} / (0.6745\sqrt{2}), \quad (10.2.18)$$

with $x_r(n\Delta t)$ and $x_i(n\Delta t)$ being the real and imaginary parts of $x(n\Delta t)$. It is assumed that N is large and Δt is small [16–19]. For this estimation, we oversampled the signal by a factor of four.

The WDs with constant window widths $N_s = 16$ and $N_s = 256$ are presented in Fig. 10.2.2(a), (b), respectively. The IF estimates using the WDs with constant window widths $N_s = 8$ and $N_s = 256$ are given in Fig. 10.2.2(c),(d). Figure 10.2.2(e) shows the WD with an adaptive window width. Values of the adaptive window width, determined by the algorithm, are presented in Fig. 10.2.2(f). We can see that when the IF variations are small, the algorithm uses the widest window in order to reduce the variance. Near the point $n\Delta t = 0.5$, where the IF variations are fast, the windows with smaller widths are used. The IF estimate with the adaptive window width is presented in Fig. 10.2.2(g). The mean absolute error, normalized to the discretization step, is shown in Fig. 10.2.2(h) for each considered window width. The line represents the value of the mean absolute error for the adaptive window width.

10.2.4 SUMMARY AND CONCLUSIONS

An algorithm that can produce an accurate estimate of the optimal window width, without using the bias value, is presented. The IF estimates obtained by using this algorithm and the WD have lower error than those obtained by using the best constant-window width, which also is not known in advance. Additional examples, including distributions with adaptive order, the WD as a spectrum estimator, algorithm application to sensor-array signal tracking, as well as other realization details can be found in Refs. [16–25]. The algorithm has various other applications [22–25]. The performance of the ICI rule can be further improved by taking into account the overlap between the current and previous confidence intervals [25].

**FIGURE 10.2.2**

Time-frequency analysis of a noisy signal: (a) Wigner distribution (WD) with $N = 16$; (b) WD with $N = 256$; (c) Estimated instantaneous frequency using the WD with $N = 8$; (d) Estimated instantaneous frequency using the WD with $N = 256$; (e) WD with adaptive window width; (f) Adaptive window width as a function of time; (g) Estimated instantaneous frequency using the WD with the adaptive window width; (h) Absolute mean error as a function of the window width; the line represents the mean absolute error value for the adaptive window width.

10.3 IF ESTIMATION FOR MULTICOMPONENT SIGNALS⁰

10.3.1 TIME-FREQUENCY PEAK IF ESTIMATION

There is a wide range of applications where we encounter signals comprising M components with different IF laws $f_m(t)$ and different envelopes $a_m(t)$, in additive noise. It is often desired from such an observed signal to determine the number of components M , the IF law of each component, and the corresponding envelope $a_m(t)$. This can be achieved by representing the observed signal $z(t)$ in a time-frequency (t,f) domain and use time-frequency filtering methods to recover the individual components [26]. Another approach involves extending algorithms for IF estimation of monocomponent FM signals to the case of multicomponent signals and design an algorithm that simultaneously tracks the various IF components of the observed signal [27,28]. Both approaches require the use of time-frequency distributions (TFDs) with very specific properties such as high time-frequency localization of the IF components and high reduction of cross-terms interferences.

The basic concept of instantaneous frequency is described in Part 1 of Ref. [29] and in Chapter 1 of this book. Methods of IF estimation are reported in this chapter and in Part 2 of Ref. [29]. Essential results are reproduced below for greater clarity.

10.3.1.1 Spectrogram peak IF estimation

Various approaches for IF estimation of monocomponent signals exist [29]. Most of these algorithms are suited to a particular class of signals, and both fixed and adaptive algorithms have been proposed. The aim here is to approach the problem from a general viewpoint in order to define a general IF estimation methodology that would be suitable for the largest class of signals found in practical applications. To illustrate this approach, we thus consider from the outset multicomponent signals in additive noise, which can be expressed as follows:

$$z(t) = \sum_{m=1}^M z_m(t) + \epsilon(t) = \sum_{m=1}^M a_m(t) e^{j\phi_m(t)} + \epsilon(t), \quad (10.3.1)$$

where $a_m(t)$ is the m th component amplitude, $\phi_m(t)$ is the m th component phase, and $\epsilon(t)$ is a complex-valued white Gaussian noise process with independent and identically distributed (i.i.d.) real and imaginary parts and with total variance σ_ϵ^2 . The individual IF laws for each component are given by [26]:

$$f_m(t) = \frac{1}{2\pi} \frac{d\phi_m(t)}{dt}; \quad m = 1, \dots, M. \quad (10.3.2)$$

A conventional approach to represent and analyze such signals for IF estimation is to take a TFD of $z(t)$ and search for the peaks in the (t,f) domain (see Section 10.1). Curves formed by a continuum of these peaks describe the IF laws of the individual components of the observed signal $z(t)$, as illustrated in Fig. 10.3.1 using the spectrogram of a bat signal. Analytically, this can be expressed as follows:

⁰Authors: **Z.M. Hussain**, University of Kufa (zahir.hussain@uokufa.edu.iq) and, Edith Cowan University, Australia (z.hussain@ecu.edu.au) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: LJ. Stanković and V. Katkovnik.

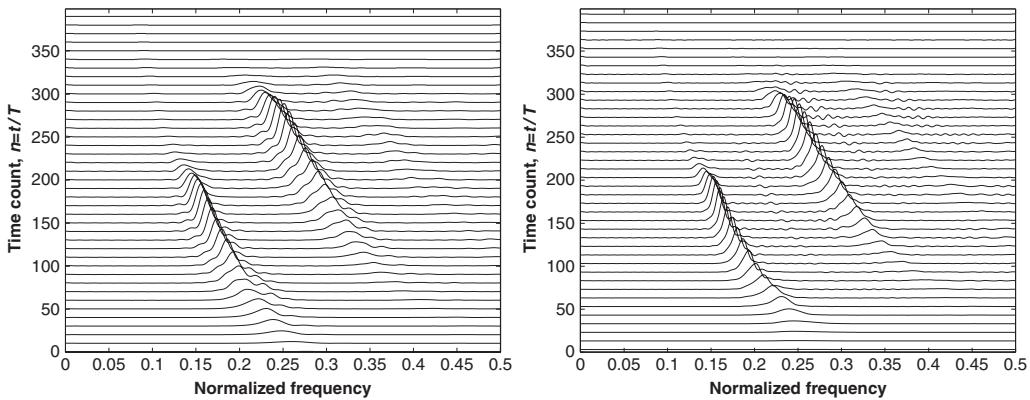


FIGURE 10.3.1

Left: The spectrogram of a bat signal using a small analysis window. Right: The modified B-distribution of the same signal with parameter $\beta = 0.05$. Total signal length is $N = 400$ and sampling interval $\Delta t = 1$. The spectrogram cannot show the weakest component.

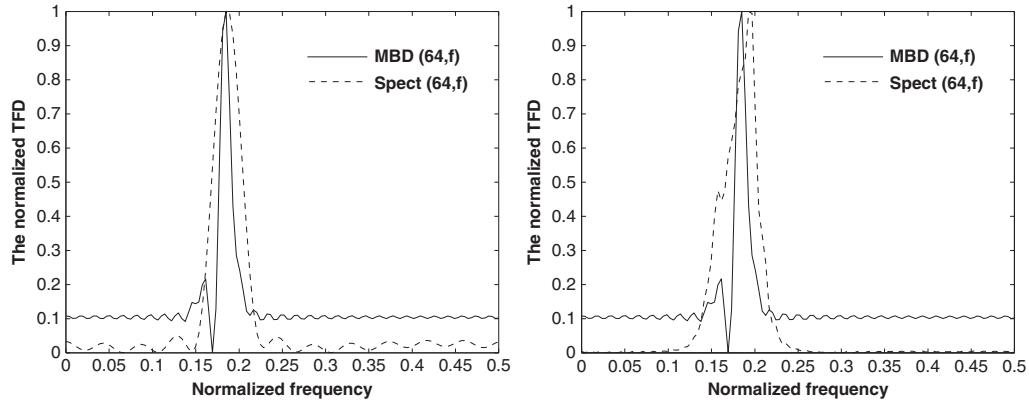
$$\hat{f}_m(t) = \arg \max_f \rho_m(t, f); \quad 0 \leq f \leq f_s/2, \quad (10.3.3)$$

where $\rho_m(t, f)$ is the m th peak of the TFD.

The spectrogram-based approach has several advantages: it is easy to understand, easy to use, and there are no cross-terms producing unwanted interferences. A major disadvantage though is that the (t, f) resolution of the spectrogram for closely spaced components is often poor, especially if one of the components is much weaker. This is illustrated in Fig. 10.3.2 for a two-component linear FM signal with one component weaker than the other. Figure 10.3.2 shows the comparison of the performance of the spectrogram with the modified B-distribution (MBD) that was designed specifically for multicomponent IF estimation, as discussed later.

10.3.1.2 Peaks of WVD, PWVD, and RIDs

To improve upon the resolution of the spectrogram, various TFDs were proposed for IF estimation, one of the most important being the Wigner-Ville distribution (WVD). IF estimation using the peak of the WVD is optimal for linear FM signals with high to moderate signal-to-noise ratios (SNRs) [29], but its performance degrades significantly at low SNRs, and in this case the cross WVD (XWVD) peak can be used as an IF estimator [30]. For polynomial FM signals, it was shown that the polynomial WVD (PWVD) gives the best performance, especially at high SNRs (see Ref. [31] and Section 10.4). However, both WVD and PWVD suffer from cross-terms when used to analyze multicomponent signals, especially at low SNR. Cross-terms generate artifacts that obscure the (t, f) representation of the signal. This led to the development of reduced interference distributions (RIDs) to remedy the problem [32]. Straightforward IF estimation using the peak of RIDs give an IF estimate that is biased from the true IF law, and this bias is different for different RIDs. Although reduced, cross-terms still exist and can obscure weak components, hence the need to define special purpose RIDs with efficient

**FIGURE 10.3.2**

Performance comparison between the spectrogram $\text{Spect}(t, f)$ and the modified B-distribution (MBD) for $\beta = 0.06$ using a two-component noise-free linear FM signal at the sampling instant $n = t/\Delta t = 64$. Total signal length is $N = 128$ and the sampling interval is $\Delta t = 1$. The right component is five times larger in amplitude than the left component. Left: Spectrogram with small analysis window length ($\Delta = 23$). Right: Spectrogram with large analysis window length ($\Delta = 83$). In both cases, the spectrogram fails to resolve the two components. In addition, time resolution is bad for a large window length.

cross-terms reduction, high time-frequency resolution, and minimum bias from the true IF laws, such as the MBD [27,33].

10.3.2 PROPERTIES OF IF ESTIMATES BASED ON QUADRATIC TFDs

10.3.2.1 IF estimates and window length

Let us consider an analytic signal $z(t)$ of the form $z(t) = ae^{j\phi(t)} + \epsilon(t)$ where the amplitude a is constant, and $\epsilon(t)$ is a complex-valued white Gaussian noise with independent identically distributed (i.i.d.) real and imaginary parts with total variance σ_ϵ^2 . The IF of $z(t)$ is given by Eq. (10.3.2), and it is assumed to be an arbitrary, smooth, and differentiable function of time with bounded derivatives of all orders. The general equation for quadratic TFDs of the signal $z(t)$ is given by [26]

$$\rho_z(t, f) = \mathcal{F}_{\tau \rightarrow f} [G(t, \tau) * K_z(t, \tau)], \quad (10.3.4)$$

where $G(t, \tau)$ is the time-lag kernel, $K_z(t, \tau) = z(t + \tau/2)z^*(t - \tau/2)$ is called the signal kernel or the instantaneous autocorrelation function (IAF), and $*$ denotes time convolution. For smoothing

(t)
and localization on the IAF, we apply a window function $w_h(\tau) = \frac{\Delta t}{h} w(\frac{\tau}{2h})$ on the instantaneous autocorrelation $K_z(t, \tau)$, where $w(t)$ is a real-valued symmetric window with unit length, i.e., $w(t) = 0$ for $|t| > \frac{1}{2}$; hence the window length is h .

The TFD is therefore now dependent on the window length h as follows:

$$\rho_{z,h}(t,f) = \underset{\tau \rightarrow f}{\mathcal{F}}[w_h(\tau) G(t, \tau) * K_z(t, \tau)]. \quad (10.3.5)$$

If $\rho_{z,h}(t,f)$ is discretized over time, lag, and frequency, then we have

$$\rho_{z,h}(n, k) = \sum_{l=-N_s}^{N_s-1} \sum_{m=-N_s}^{N_s-1} w_h(m\Delta t) K_z(l\Delta t, 2m\Delta t) G(n\Delta t - l\Delta t, 2m\Delta t) e^{-j2\pi \frac{kn}{2N_s}}, \quad (10.3.6)$$

where $2N_s$ is the number of samples and Δt is the sampling interval.

The IF estimate is a solution of the following optimization

$$\hat{f}_h(t) = \arg \max_f \rho_{z,h}(t,f); \quad 0 \leq f \leq f_s/2, \quad (10.3.7)$$

where $f_s = 1/\Delta t$ is the sampling frequency.

10.3.2.2 Bias and variance of the IF estimate

By extending the results in Ref. [17], the estimation bias and variance are found to be [27]

$$E[\Delta \hat{f}_h(t)] = \frac{L_h(t)}{2F_h}, \quad \text{var}(\Delta \hat{f}_h(t)) = \frac{\sigma_\epsilon^2}{2|a|^2} \left[1 + \frac{\sigma_\epsilon^2}{2|a|^2} \right] \frac{E_h}{F_h^2}, \quad (10.3.8)$$

where

$$\begin{aligned} \Delta \hat{f}_h(t) &= \frac{1}{2\pi} \phi'(t) - \hat{f}(t); \quad F_h = \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t) (2\pi m\Delta t)^2 G(u, 2m\Delta t) du \\ L_h(t) &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t) \Delta \phi(u, m\Delta t) (2\pi m\Delta t) G(t-u, 2m\Delta t) du \\ E_h &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t)^2 (2\pi m\Delta t)^2 G(u, 2m\Delta t) du, \end{aligned} \quad (10.3.9)$$

where $\Delta \phi(t, \tau) = \phi(t + \tau/2) - \phi(t - \tau/2) - \tau \phi'(t)$.

Equations (10.3.8) and (10.3.9) indicate that the bias and the variance of the estimate depend on the lag window length h for any kernel $G(t, \tau)$. To see how the bias and the variance vary with h , an asymptotic analysis as $\Delta t \rightarrow 0$ is necessary for the selected TFD.

10.3.2.3 TFD properties needed for multicomponent IF estimation

The results above indicate that if a general method for IF estimation of multicomponent FM signals in additive Gaussian noise is to be based on quadratic TFDs, those TFDs should satisfy the following conditions [27]:

- (1) $\rho(t, f)$ should have a high (t, f) resolution while suppressing cross-terms efficiently so as to give a robust IF estimate for mono- and multicomponent FM signals;

- (2) $\rho(t,f)$ should enable amplitude estimation for the individual components of the signal, as the amplitude is necessary for evaluating the variance of the IF estimate for each component [17,27,28], and to allow for the reconstruction of the individual components of the signal;
- (3) the choice of the lag window length should lead to a bias-variance tradeoff (see Eqs. (10.3.14) and (10.3.15)).

Although some TFDs, like Exponential distribution (ED) and the spectrogram, can satisfy some of these conditions, they do not meet the second requirement, i.e., allowing direct amplitude estimation. The design of TFDs which satisfies all of these required properties is considered next.

10.3.3 DESIGN OF QUADRATIC TFDs FOR MULTICOMPONENT IF ESTIMATION

10.3.3.1 Desirable time-lag kernel

A TFD referred to as the B-distribution (BD) was proposed and shown to be superior to other fixed-kernel TFDs in terms of cross-terms reduction and resolution enhancement [34]. As it does not allow direct component-amplitude estimation [28], as per the second condition on TFDs required for multicomponent IF estimation listed above, the BD kernel was modified to become the MBD kernel as [27]

$$G(t, \tau) = G_\beta(t) = k_\beta / \cosh^{2\beta}(t), \quad (10.3.10)$$

where β is a real positive number with range 0 to 1, and $k_\beta = \Gamma(2\beta)/(2^{2\beta-1} \Gamma^2(\beta))$, Γ stands for the gamma function (this modified B-distribution (MBD) was also referred to as the hyperbolic T-distribution (HTD) in Ref. [33]). See Table 3.3.2 for additional details.

10.3.3.2 Relevant properties of the modified B-distribution (MBD)

Most desirable properties of TFDs relevant to IF estimation (as explained in Refs. [26,29]) are satisfied by the MBD kernel. In particular, realness, time-shift and frequency shift invariance, frequency marginal and group delay, and the frequency support properties are satisfied. The time support property is not strictly satisfied, but it is approximately true [27]. The three required conditions listed in Section 10.3.2.3 are discussed in detail below.

- (1) **Reduced interference and resolution:** This property is satisfied by MBD. For example, consider the sum of two complex sinusoidal signals $z(t) = z_1(t) + z_2(t) = a_1 e^{j(2\pi f_1 t + \theta_1)} + a_2 e^{j(2\pi f_2 t + \theta_2)}$ where a_1, a_2, θ_1 and θ_2 are constants. The TFD of the signal $z(t)$ is obtained as [27]

$$\text{MBD}(t, f) = a_1^2 \delta(f - f_1) + a_2^2 \delta(f - f_2) + 2a_1 a_2 \gamma_\beta(t) \delta(f - (f_1 + f_2)/2), \quad (10.3.11)$$

where $\gamma_\beta(t) = |\Gamma(\beta + j\pi(f_1 - f_2))|^2 \cos(2\pi(f_1 - f_2)t + \theta_1 - \theta_2) / \Gamma^2(\beta)$. The cross-terms are oscillatory in time and depend on the frequency separation between signal components. If f_1 and f_2 are well separated then the term $|\Gamma(\beta + j\pi(f_1 - f_2))|^2$ can be substantially reduced, while $\Gamma^2(\beta)$ can be made high if β is small. When f_1 and f_2 are not well separated, the MBD still performs better than most quadratic TFDs (see Section 7.4).

- (2) **Direct amplitude and IF estimation:** The MBD allows direct IF estimation by peak localization, i.e., at any time instant t , it has an absolute maximum at $f = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$ for linear FM signals. For

nonlinear FM signals, this estimate is biased, but this bias can be accounted for in the adaptive IF estimation, as presented next.

For an FM signal of the form $z(t) = a e^{j\phi(t)}$, the MBD is approximated by [27]

$$\text{MBD}(t, f) \approx |a|^2 \int_{-\infty}^{\infty} G_{\beta}(t-u) \delta\left(\frac{1}{2\pi}\phi'(u) - f\right) du = |a|^2 G_{\beta}(t - \psi(f)) \psi'(f), \quad (10.3.12)$$

where ψ is the inverse of $\frac{1}{2\pi}\phi'$, i.e., $\frac{1}{2\pi}\phi'(\psi(f)) = f$. Assuming that $\psi'(f)$ is not a highly peaked function of f and knowing that $G_{\beta}(t - \psi(f))$ is peaked at $t = \psi(f)$, the absolute maximum of $\text{MBD}(t, f)$ for any time t would be at $\psi(f) = t$, or $f = \frac{1}{2\pi}\phi'(t)$, which is the IF of the FM signal $z(t)$. For nonlinear FM signals, the energy peak of the MBD is actually biased from the IF because of the extra term $\sum_{k=3(k \text{ odd})}^{\infty} \frac{\tau^{k-1}}{k! 2^{k-1}} \phi^{(k)}(u)$. The major contribution in this term is due to $\phi^{(3)}(u)$. Therefore at the instants of rapid change in the IF law the bias is not negligible and Eq. (10.3.12) would not be an accurate approximation to the MBD unless suitable windowing in the lag direction is used.

For linear FM signals, we have $\phi^{(k)}(t) = 0$ for $k \geq 3$. Assuming that $z(t) = a e^{j2\pi(f_o t + \frac{\beta_o}{2} t^2)}$, where f_o and β_o are constants, we have

$$\text{MBD}(t, f) = \frac{1}{\beta_o} |a|^2 G_{\beta}\left(t - \frac{1}{\beta_o}(f - f_o)\right), \quad (10.3.13)$$

which has an absolute maximum at $f = f_o + \beta_o t$, the IF of the linear FM signal $z(t)$. As $\beta_o \rightarrow 0$, i.e., $z(t)$ approaches a sinusoid, we have $\text{MBD}(t, f) \rightarrow |a|^2 \delta(f - f_o)$, in accordance with Eq. (10.3.11).

As for amplitude estimation, Eqs. (10.3.12) and (10.3.13) indicate that the MBD allows amplitude estimation [27].

- (3) Asymptotic Formulas Using MBD:** The asymptotic formulas for the variance and the bias as $\Delta t \rightarrow 0$ using a rectangular lag window are given by [27]:

$$\text{var}(\Delta \hat{f}_h(t)) = \frac{3\sigma_{\epsilon}^2}{2\pi^2 |a|^2} \left[1 + \frac{\sigma_{\epsilon}^2}{2|a|^2} \right] \frac{\Delta t}{h^3} \quad (10.3.14)$$

and

$$E(\Delta \hat{f}_h(t)) = \frac{h^2}{80} \int_{-\infty}^{\infty} \frac{\lambda(u) du}{\cosh^2 \beta(t-u)}; \quad E(\Delta \hat{f}_h(t)) \leq \frac{M_2}{40} h^2, \quad (10.3.15)$$

where $\lambda(t) = f^{(2)}(t + \tau_1) + f^{(2)}(t - \tau_1)$, $f^{(2)}(t)$ is the second derivative of the IF, and $\sup_t |f^{(2)}(t)| \leq M_2$. For small h , the optimal window length that minimizes the mean squared error is obtained by extending the result in Ref. [31] as:

$$h_{\text{opt}}(t) = \left[\frac{1800 \sigma_{\epsilon}^2 \Delta t \left(1 + \frac{\sigma_{\epsilon}^2}{2|a|^2} \right)}{\pi^2 |a|^2 \left(f^{(2)}(t) * \frac{1}{\cosh^2 \beta(t)} \right)^2} \right]^{\frac{1}{7}}. \quad (10.3.16)$$

Thus, the optimal window length depends on the second derivative of the IF, $f^{(2)}(t)$, which is time- and signal-dependent. Equations (10.3.14) and (10.3.15) indicate that the variance and bias of the IF estimate using the MBD have the same rates of change with respect to the window length h as those using the WVD [17].

10.3.3.3 Examples of quadratic TFDs suitable for multicomponent IF estimation

TFDs with time-only (or lag-independent) kernels constitute a subclass of the quadratic class of TFDs. These TFDs share the important properties of cross-terms suppression, high-resolution, and supporting amplitude estimation, making them well suited for multicomponent IF estimation. The MBD was defined earlier in this section. Another example is the exponential T-distribution, which is defined in terms of its time-lag kernel as [33]

$$G(t, \tau) = G_\eta(t) = \sqrt{\eta/\pi} \exp(-\eta t^2),$$

where η is a real parameter and $\sqrt{\eta/\pi}$ is a normalization factor. The resulting TFD used for multicomponent IF estimation is then given by Eq. (10.3.5).

10.3.4 AN ADAPTIVE ALGORITHM FOR MULTICOMPONENT IF ESTIMATION

Equation (10.3.16) shows that the optimal lag window length using the MBD is a function of time and depends on the second derivative of the IF law $f^{(2)}(t)$; it decreases when the IF law $f(t)$ has a high variation. Hence a time-varying window length is needed to optimize the estimation. The Stanković-Katkovnik adaptive algorithm developed in Ref. [17] for monocomponent FM signals can be used since the IF estimation variance is a continuously decreasing function of h while its bias is continuously increasing, as shown in Eqs. (10.3.14) and (10.3.15); see also Section 10.2. These conditions are necessary for bias-variance tradeoff such that the algorithm converges at the optimum window length that resolves this tradeoff. It is shown in Ref. [17] that, if h is small enough then the IF estimate $\hat{f}_h(t)$ is inside the confidence interval D defined as follows

$$D = \left[\hat{f}_h(t) - 2\kappa \sqrt{\text{var}(\Delta \hat{f}_h(t))}, \hat{f}_h(t) + 2\kappa \sqrt{\text{var}(\Delta \hat{f}_h(t))} \right] \quad (10.3.17)$$

with Gaussian probability $P(\kappa)$, κ being a parameter (usually 2); while for large h , $\hat{f}_h(t)$ is outside D . Hence, if we consider an increasing sequence of window lengths $\{h_r | r = 1 : N\}$ (N being the number of samples) and calculate the MBD (and hence $\hat{f}_{h_r}(t)$) for each h_r , then all $\{D_r\}$ have at least one point in common (which is $\hat{f}_{h_r}(t)$) if h_r is sufficiently small. The first h_r for which D_{r-1} and D_r have no point in common is considered optimal as it decides the bias-variance tradeoff.

The estimates for the amplitude of the signal a and the variance of noise σ_ϵ^2 [used in Eq. (10.3.14) and implicitly in Eq. (10.3.17)] were given in Ref. [17] as:

$$\hat{a}^2 + \hat{\sigma}_\epsilon^2 = \frac{1}{N} \sum_{n=1}^N |z(n\Delta t)|^2; \quad \hat{\sigma}_\epsilon^2 = \frac{1}{2N} \sum_{n=2}^N |z(n\Delta t) - z((n-1)\Delta t)|^2, \quad (10.3.18)$$

where N is the number of samples. For further details, see Refs. [17,35].

For a multicomponent analytic signal of the form stated in Eqs. (10.3.1) and (10.3.2), with $\{a_m\}$ constant, we can use the extension of the monocomponent IF estimation algorithm in Ref. [17] for

multicomponent signals as described in Refs. [27,28]. This algorithm tracks component maxima in the (t,f) plane and requires a threshold $\mathcal{T}_\rho(t)$ so as to ignore the local maxima caused by the cross-terms and windowing. In fact, $\mathcal{T}_\rho(t)$ is application and distribution dependent.

The algorithm requires the knowledge of the confidence intervals $D_{r,m}$ for each component, where r refers to the window length (h_r) and m refers to the signal component. The calculation of $D_{r,m}$ depends on the estimation of the individual amplitudes a_m of the components. Using the MBD, the actual amplitudes $|\hat{a}_m|$ can be estimated as shown in Ref. [28]. Using $|\hat{a}_m|^2$ and $\hat{\sigma}^2$ to calculate $\text{var}(\Delta\hat{f}_h(t))$ [given by Eq. (10.3.14) for $\text{MBD}(t,f)$], we can define the confidence intervals $\{D_{r,m}\}$ for all components as in Refs. [27,28]. The IF $f_m(t)$ is contained in at least one of the confidence intervals $\{D_{r,m}\}$ if h_r is sufficiently small, and the optimal window length is the first h_r (from the increasing sequence $\{h_r|r = 1 : N\}$) for which $D_{r-1,m}$ and $D_{r,m}$ have no point in common.

Example. Let us consider a three-component FM signal $z(n\Delta t)$ with amplitudes $a_1 = 0.5$, $a_2 = 1$, and $a_3 = 1.5$ and nonlinear IF laws: $f_1 = 47 + 2.5 \sinh^{-1}(20(n\Delta t - 0.2))$, $f_2 = 30 + 2.5 \text{sgn}(40(n\Delta t - 0.6))$, and $f_3 = 10 + 2 \sin(10(n\Delta t - 0.7))$, with SNR = 15 dB, $\beta = 0.1$, $\kappa = 2$, $0 \leq n\Delta t \leq 1$, and $\Delta t = 1/128$. Figure 10.3.3 shows the result of the tracking adaptive algorithm for IF estimation of $z(n\Delta t)$ using the peaks of the MBD and the spectrogram.

Figure 10.3.4 shows the conventional peak IF estimation for the same signal using MBD and the spectrogram. Both TFDs fail to give a robust IF estimation at the instants of rapid frequency change. In addition, the spectrogram has poor (t,f) resolution, both in adaptive and constant-window IF estimation.

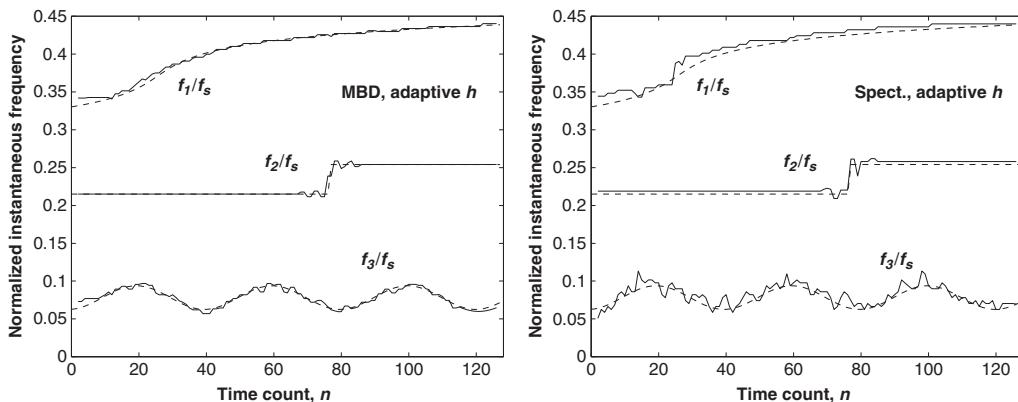
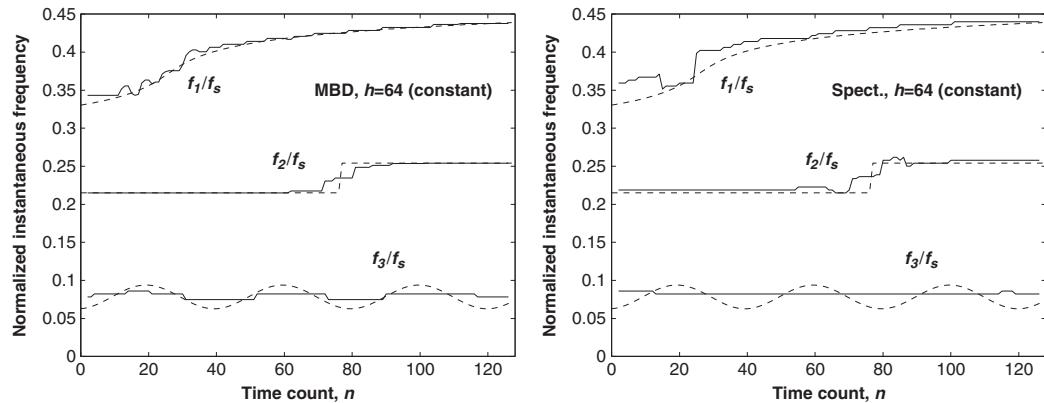


FIGURE 10.3.3

Left: Adaptive IF estimation using the peak of the MBD for a three-component FM signal with total length $N = 128$, SNR = 15 dB, and $\Delta t = 1/128$. Dashed lines represent the true IF laws. Right: Adaptive IF estimation using the peak of the spectrogram for the same signal, assuming that component amplitudes are known.

**FIGURE 10.3.4**

Left: IF estimation for the three-component FM signal as in Fig. 10.3.3 using the conventional (constant window) peak estimation. Left: MBD peak IF estimation. Right: Spectrogram peak IF estimation. (In addition to failure of IF estimation at instants of rapid frequency change, the spectrogram has poor tracking performance. Note also that both methods cannot track the continuously varying frequency of the third component.)

10.3.5 SUMMARY AND CONCLUSIONS

Concurrent IF estimation of the separate components of a multicomponent FM signal using TFD peaks location requires conditions for the selection of a suitable quadratic TFD. Required properties are: (1) (t, f) resolution is high with efficient suppression of cross-terms, (2) the TFD enables direct amplitude estimation for the individual components, and (3) the variance of the IF estimate using the TFD is a continuously decreasing function of the lag window length while the bias is continuously increasing. Quadratic TFDs that satisfy these conditions were presented and discussed. A constant-window tracking algorithm may not give a robust IF estimate if the IF changes rapidly with time due to the effect of the higher-order derivatives of the IF law. Hence an adaptive algorithm is used for robust multicomponent IF estimation. Further developments are reported in Ref. [36] and in other Sections in this chapter. Areas of applications include EEG and HRV analysis, as discussed in Section 16.6, [37,38] as well as in Section 10.9.

10.4 PARAMETER ESTIMATION FOR POLYNOMIAL FM SIGNALS IN ADDITIVE GAUSSIAN NOISE⁰

10.4.1 PROBLEM DEFINITION

This section presents two methods of parameter estimation, especially IF estimation, for monocomponent and multicomponent polynomial FM signals. The general principles followed by the two methods are:

⁰Second Edition updated by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **P.J. O'Shea**, Qatar University, Doha, Qatar; University of Queensland, Brisbane, QLD, Australia. First Edition [39] by P.J. O'Shea and B. Barkat.

1. In the time-frequency domain and its dual, the ambiguity (Doppler-lag) domain, a sinusoidal signal shows a ridge along a line characterizing the fundamental frequency; hence
2. We can estimate the coefficients of the polynomial recursively by transforming a polynomial of order P to a new polynomial of order $P - 1$, and so on.

The details are presented step-by-step.

Let us consider a noiseless complex polynomial frequency modulated (FM) signal, $z(t)$, expressed as

$$z(t) = A \exp(j\phi(t)) \operatorname{rect}(t - T/2) = A \exp \left\{ j \sum_{k=0}^P a_k t^k \right\} \operatorname{rect}(t - T/2), \quad (10.4.1)$$

where A is the signal amplitude, $\phi(t)$ is the signal phase, P is the polynomial phase order, the $\{a_k\}$ are arbitrary parameters and T is the signal duration. The FM law of the signal is the instantaneous frequency (IF) trajectory, given by

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \frac{1}{2\pi} \sum_{k=1}^P k a_k t^{k-1}, \quad t = [0, T]. \quad (10.4.2)$$

In additive noise, the signal model can be rewritten as:

$$y(t) = z(t) + w(t) = A \cdot \exp \left\{ j \sum_{k=0}^P a_k t^k \right\} + w(t), \quad t = [0, T], \quad (10.4.3)$$

where $w(t)$ is a complex, Gaussian noise.

If the order of the polynomial phase in Eq. (10.4.1) is equal to 1, then $z(t)$ is a complex sinusoid (with linear phase) and its IF reduces to a constant independent of t . It can be estimated as the frequency at which the FT of the data has its peak. The Fourier transform is useful in this case because it concentrates the signal energy in frequency about the IF, while dispersing the noise energy over a wide band. In the f domain, the signal energy is therefore peaked, while the noise contribution tends to be broad and relatively low. The FT peak provides an *optimal* estimator for the frequency of a complex sinusoidal data sequence, when the noise is white and Gaussian [40].

If the order of the polynomial phase in Eq. (10.4.1) is higher than 1, so that the phase of $z(t)$ is non-linear, then the IF is not constant but varies with t . The FT is then less effective for IF estimation; the signal energy is spread over a continuum of frequencies and the signal-to-noise ratio (SNR) in the spectrum is reduced compared with the complex sinusoidal case. Hence it would be desirable to define an operator which transforms the nonlinear phase signal into a linear-phase one with frequency equal to the IF so that the FT could be applied to estimate the IF.

10.4.2 THE POLYNOMIAL WIGNER-VILLE DISTRIBUTIONS

10.4.2.1 IF estimation

Transforming a polynomial FM signal into a sinusoid for a given time instant is discussed in Ref. [41] and Sections 5.4 and 5.5. This transform can be expressed as [41]:

$$K_z^P(t, \tau) = \prod_{i=1}^I z(t + c_i \tau)^{k_i} \cdot z^*(t - c_i \tau)^{k_i}, \quad (10.4.4)$$

where the c_i are the coefficients of the transformation K_z^P , and I is the total number of distinct c_i values. The order of the transformation is $q = \sum_{i=1}^I 2k_i$. The procedure to obtain the c_i , k_i , and q for a fixed polynomial phase order, P , is described in Ref. [42]. For example, for $P = 2$, we find $q = 2$ and $c_1 = 0.5$. For $P = 3$ or $P = 4$ we find $q = 6$, $c_1 = 0.62$, $c_2 = 0.75$, and $c_3 = -0.87$ [42].

The FT of the “kernel” K_z^P is the polynomial Wigner-Ville distribution (PWVD)

$$W_z^P(t, f) = \int_{-\infty}^{\infty} K_z^P(t, \tau) \cdot e^{-j2\pi f \tau} d\tau, \quad (10.4.5)$$

whose peak yields the IF of the signal [41]. The PWVD $W_z^P(t, f)$ is a generalization of the Wigner-Ville distribution (WVD), which is effective for IF estimation of quadratic-phase signals. In other words, the conventional WVD is a specific case of the PWVDs with $P = 2$, $q = 2$, and $c_i = 1/2$ [43,44].

Example. Figure 10.4.1 displays the PWVD of a quadratic FM signal ($P = 3$). The plot on the left is the PWVD of a noiseless signal, while the plot on the right is the PWVD of the same signal immersed in additive Gaussian noise. In the noisy case, the IF law of the original signal is still apparent, despite the degradation in the TFD caused by the noise. This suggests that the PWVD is useful for providing unbiased estimates of the IF of a polynomial phase signal with a high signal-to-noise ratio (SNR) [31], but it may be poor for low-SNR signals.

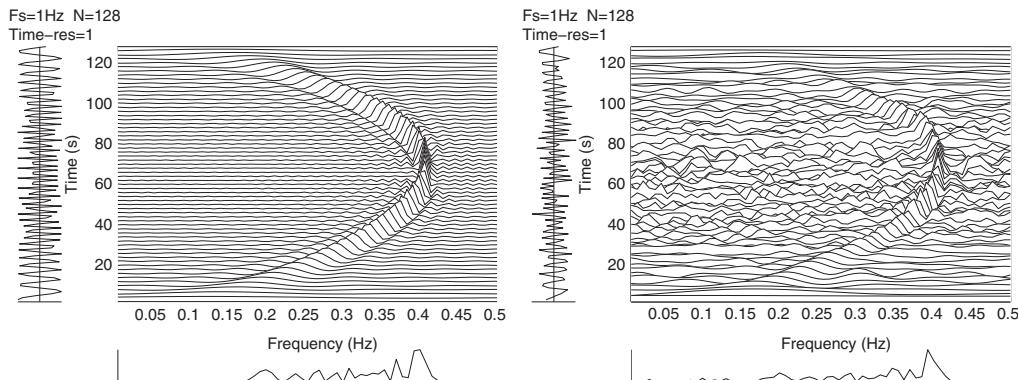


FIGURE 10.4.1

PWVDs of noiseless (left) and noisy (right) quadratic FM signals.

10.4.2.2 Iterative IF estimation

To achieve accurate IF estimation at low SNR values, iterative procedures are generally used [45]. Step 1 is to get an initial estimate of the IF law, $\hat{f}_i(t)$, with a method which has a low SNR threshold, such as the peak of the spectrogram. Step 2 is to reconstruct from $\hat{f}_i(t)$ an estimate, $\hat{y}(t)$, of the noisy observation $y(t)$. Step 3 is to use $\hat{y}(t)$ with $y(t)$ to form a Cross PWVD (XPWVD):

$$W_{y,\hat{y}}^P(t,f) = \int_{-\infty}^{\infty} K_{y,\hat{y}}^P(t,\tau) \cdot e^{-j2\pi f\tau} d\tau, \quad (10.4.6)$$

where

$$K_{y,\hat{y}}^P(t,\tau) = y(t + c_1\tau)^{k_1} \hat{y}^*(t - c_1\tau)^{k_1} \prod_{i=2}^I \hat{y}(t + c_i\tau)^{k_i} \cdot \hat{y}^*(t - c_i\tau)^{k_i}. \quad (10.4.7)$$

Step 4 is then to use XPWVD peak extraction to form a new estimate of the IF law, which in turn is used to form an updated estimate of $y(t)$.

The above process is repeated iteratively until there is an acceptably small difference between successive signal estimates. The SNR threshold of this method is typically similar to that of the preliminary IF estimation method, but the accuracy improves (for more details, see Ref. [45]).

10.4.2.3 Discrete-time formulation

Let us define $z(n)$, $y(n)$, and $K_y^P(n,m)$ as the discrete-time signals obtained by sampling $z(t)$, $y(t)$, and $K_y^P(t,\tau)$, respectively, with a normalized sampling rate of 1. The number of samples available from both $z(n)$ and $y(n)$ is denoted $N(\approx T)$. The discrete-time version of the PWVD is then:

$$W_y^P(n,f) = \underset{m \rightarrow f}{\text{DTFT}} K_y^P(n,m), \quad (10.4.8)$$

where the DTFT operation specifies discrete-time Fourier transformation from m to f . To form the kernel term, $K_y^P(n,m)$, in the above expression, one must obtain samples of $y(t)$ at non-integer values of time, e.g., by interpolation of the discrete-time observation, $y(n)$. To limit the number of samples obtained by interpolation, let us introduce some “lag scaling” and “frequency scaling” operations into the definition of the discrete-time PWVD, resulting in:

$$W_y^P(n,f) = \underset{m \rightarrow \frac{f}{c_{i_{\max}}}}{\text{DTFT}} K_{y_s}^P(n,m), \quad (10.4.9)$$

where

$$K_{y_s}^P(t,\tau) = \prod_{i=1}^I y \left(t + \frac{c_i}{c_{i_{\max}}} \tau \right)^{k_i} \cdot y^* \left(t - \frac{c_i}{c_{i_{\max}}} \tau \right)^{k_i}, \quad (10.4.10)$$

and $K_{y_s}^P(n,m)$ is obtained from $K_{y_s}^P(t,\tau)$ by sampling at a rate of 1. $c_{i_{\max}}$ is the c_i value with the largest magnitude. The discrete-time PWVD defined above can be used for obtaining low-variance IF estimates from discrete-time measurements. A derivation of the variance of such estimates is given next.

10.4.2.4 Asymptotic mean-square error of PWVD-based IF estimates

The PWVD-based IF estimate, evaluated at $n = 0$, is defined by:

$$\hat{f}_0 = \arg \max_f \{W_y^P(n, f)|_{n=0}\} = \arg \max_f \{W_y^P(f)\}, \quad (10.4.11)$$

where

$$W_y^P(f) = \underset{m \rightarrow \frac{f}{c_{i_{\max}}}}{\text{DTFT}} K_{y_s}^P(n = 0, m). \quad (10.4.12)$$

The PWVD of a noisy polynomial phase signal, evaluated at $n = 0$, is:

$$W_y^P(f) = \sum_{m=-(N-1)/2}^{(N-1)/2} \prod_{i=1}^I [y(mc_i/c_{i_{\max}}) \cdot y^*(-mc_i/c_{i_{\max}})]^{k_i} e^{-j2\pi fm/c_{i_{\max}}}. \quad (10.4.13)$$

In the absence of noise $W_y^P(f)$ has a global maxima at $f = f_0$, where f_0 is the value of the IF at $n = 0$. The resulting perturbation of $W_y^P(f)$ which occurs when noise is added to the observation is given by

$$\delta W_y^P(f) = \sum_{m=-(N-1)/2}^{(N-1)/2} z w(m) e^{-j2\pi fm/c_{i_{\max}}}, \quad (10.4.14)$$

where

$$z w(m) = \prod_{i=1}^I [y(mc_i/c_{i_{\max}}) y^*(-mc_i/c_{i_{\max}})]^{k_i} - [z(mc_i/c_{i_{\max}}) z^*(-mc_i/c_{i_{\max}})]^{k_i}. \quad (10.4.15)$$

With this perturbation, the peak of $W_y^P(f)$ shifts from $f = f_0$ to $f = f_0 + \delta f$. To derive the a.m.s.e. of δf (i.e., the a.m.s.e. of the IF estimate), let us calculate the asymptotic mean-square fluctuations of the maximum of a real valued random function using [46]

$$E \left\{ \delta f^2 \right\} = E \left\{ \left[\frac{\partial \delta W_y^P(f_0)}{\partial f} \right]^2 \right\} \left[\frac{\partial^2 W_y^P(f_0)}{\partial f^2} \right]^{-2}, \quad (10.4.16)$$

where $E\{\cdot\}$ denotes the expected value.

Under the simplifying assumption that all samples of the kernel are uncorrelated, relevant terms on the right-hand side of Eq. (10.4.16) are found to be:

$$\frac{\partial^2 W_y^P(f_0)}{\partial f^2} \approx \frac{-4\pi^2 A^q N^3}{12 c_{i_{\max}}^2}, \quad (10.4.17)$$

$$\frac{\partial \delta W_y^P(f_0)}{\partial f} \approx -j2\pi \sum_{m=-N/2}^{N/2} m \cdot z w(m) e^{-j2\pi fm/c_{i_{\max}}}, \quad (10.4.18)$$

$$E \left\{ \left[\frac{\partial \delta W_y^P(f_0)}{\partial f} \right]^2 \right\} \approx \frac{4\pi^2 A^{2q} N^3 V_{\text{kern}}}{12}, \quad (10.4.19)$$

where V_{kern} , the “noise to signal ratio” in the kernel, is [31,47]:

$$V_{\text{kern}} = \left[\left(\sum_{i=0}^{k_1} \binom{k_1}{i}^2 i! \frac{\sigma^{2i}}{A^{2i}} \right)^2 \times \dots \times \left(\sum_{i=0}^{k_l} \binom{k_l}{i}^2 i! \frac{\sigma^{2i}}{A^{2i}} \right)^2 \right] - 1. \quad (10.4.20)$$

Equations (10.4.17) and (10.4.19) can be substituted into Eq. (10.4.16) to obtain:

$$E \left\{ \delta f^2 \right\} \approx \frac{12 \cdot c_{l_{\max}}^2 \cdot V_{\text{kern}}}{4\pi^2 \cdot N^3}. \quad (10.4.21)$$

10.4.2.5 Polynomial FM parameter estimation

To estimate the a_k parameters themselves, rather than simply the IF, in white and Gaussian noise, the estimation procedure involves (a) forming P slices of the PWVD, (b) determining P IF estimates from these slices, (c) fitting a polynomial IF law to these P IF estimates, (d) calculating the a_k parameters from this polynomial IF law, and (e) performing a spectral zoom type technique to refine the parameter estimates. The procedure is specified in more details below:

1. For the measurement, $y(n)$, form the set of PWVD slices,

$$\mathbf{W}_y^P(f) = [W_y^P(n_1, f) \ W_y^P(n_2, f) \ \dots \ W_y^P(n_P, f)]^T, \quad (10.4.22)$$

where n_1, n_2, \dots, n_P are different, well-spaced discrete-time instants. A good selection of $\{n_1, n_2, \dots, n_P\}$ is to have them equally spaced between $0.25N$ and $0.75N$, i.e., $n_i = 0.25N + 0.5N(i-1)/(P-1)$, $i = 1, \dots, P$.

2. Estimate a vector of IF estimates according to:

$$\hat{\mathbf{f}}_i = \arg \max_f \{ \mathbf{W}_y^P(f) \}, \quad (10.4.23)$$

where **arg max** denotes the vector argument of a vector (row by row) maximization.

3. Determine the vector of initial coarse estimates for the polynomial phase parameters estimates, denoted $\hat{\mathbf{a}}_i = [\hat{a}_{i1} \ \hat{a}_{i2} \ \hat{a}_{i3} \ \dots \ \hat{a}_{iP}]^T$, by evaluating the matrix equation:

$$\hat{\mathbf{a}}_i = 2\pi \mathbf{X}^{-1} \hat{\mathbf{f}}_i, \quad (10.4.24)$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & 2n_1 & 3n_1^2 & \dots & Pn_1^{P-1} \\ 1 & 2n_2 & 3n_2^2 & \dots & Pn_2^{P-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 2n_P & 3n_P^2 & \dots & Pn_P^{P-1} \end{bmatrix}. \quad (10.4.25)$$

4. Dechirp the measurement by the estimated polynomial phase law:

$$y_d(n) = y(n) \exp(-j(\hat{a}_{i1}n + \hat{a}_{i2}n^2 + \dots + \hat{a}_{ip}n^p)). \quad (10.4.26)$$

- The dechirped measurement results in a spectral content highly localized around $f = 0$.
5. Filter $y_d(n)$ with an ideal low-pass filter whose bandwidth is B . Subsample the result by a factor, $1/B$, to obtain $y_o(n)$.
6. Unwrap the phase of $y_o(n)$. The vector of unwrapped phase values

$$\mathbf{V} = [V(0) \ V(1) \ \dots \ V(N/B - 1)]^T$$

is also modeled as a (noisy) polynomial phase signal with phase parameters $\mathbf{a}_d = [a_{d0} \ a_{d1} \ \dots \ a_{dP}]^T$. One can then estimate \mathbf{a}_d with the linear regression

$$\hat{\mathbf{a}}_d = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{V}, \quad (10.4.27)$$

where

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1^P \\ 1 & 2 & \dots & 2^P \\ \dots & \dots & \dots & \dots \\ 1 & N/B - 1 & \dots & (N/B - 1)^P \end{bmatrix}. \quad (10.4.28)$$

7. Use the results from Step 6 to refine the parameter estimates from Step 3:

$$\hat{\mathbf{a}}_f = \hat{\mathbf{a}}_i + \hat{\mathbf{a}}_d \cdot \mathbf{B}, \quad (10.4.29)$$

where $\mathbf{B} = [1 \ B \ B^2 \ \dots \ B^P]^T$, and $\hat{\mathbf{a}}_d \cdot \mathbf{B}$ denotes element-by-element multiplication of $\hat{\mathbf{a}}_d$ and \mathbf{B} .

At least $P + 1$ samples of $y_o(n)$ are required for the regression in Step 6 to be properly defined. In practice, it is recommended that a large number of samples be obtained. If $y(n)$ is white, then $y_o(n)$ will also be white, and being a filtered version of $y(n)$, will have a higher SNR (by a factor of $1/B$). The unwrapped phase of $y_o(n)$ will also be very close to white, provided that the SNR of $y_o(n)$ is above about 9 dB [48]. The overall algorithm yields estimates which are asymptotically optimal above a threshold defined in Ref. [49].

10.4.3 HIGHER ORDER AMBIGUITY FUNCTIONS

The first step in this alternative method is to convert an arbitrary P th order polynomial phase signal into a linear phase one whose frequency is proportional to the a_P parameter, using with “Higher order ambiguity functions (HAFs).” The idea is that after this first step, it becomes possible to estimate the a_P parameter with the FT. The second step is to reduce the original signal to a $(P - 1)$ th order polynomial phase signal by “de-chirping” the highest order phase term out of the signal. A similar process can then be used to estimate subsequently all lower order parameters, as described by the following algorithm [43,50]:

1. Let $p = P$ and $y^{(p)}(n) = y(n)$, $1 \leq n \leq N$.
2. Set $\tau_p = \frac{N}{p}$; and calculate the a_p estimate using the formula:

$$\begin{aligned}\hat{a}_p &= \frac{1}{p! \tau_p^{p-1}} \arg \max_{\omega} \left\{ \left| \sum_{n=(M-1)\tau_p+1}^N \Pi_{q=0}^{p-1} \left[y^{\$q}(n - q\tau_p)^{\binom{p-1}{q}} \right] e^{-j\omega n} \right| \right\} \\ &= \frac{1}{p! \tau_p^{p-1}} \arg \max_{\omega} \{ |\text{HAF}_y(\omega, \tau_p)| \},\end{aligned}\quad (10.4.30)$$

where $y^{\$q}(n) = \begin{cases} y(n) & \text{if } q \text{ is even} \\ y^*(n) & \text{if } q \text{ is odd} \end{cases}$, the * signifies complex conjugation, and $\text{HAF}_y(\omega, \tau_p)$ denotes the p th order “higher order ambiguity function.”

3. Set $y^{(p-1)}(n) = y^p(n) \exp(-j\hat{a}_p n^p)$; then, set $p = p - 1$.
4. If $p \geq 1$, go back to step 2. Else proceed.
5. $\hat{a}_0 = \arg\{y^0(n)\}$, $\hat{A} = |y^0(n)|$.

The parameter estimates obtained above are good but not optimal [51]. The estimates can be refined (to the point of optimality) with either a Newton algorithm or with the procedure outlined in Steps 4-7 of [Section 10.4.2.5](#).

Once the a_0, a_1, \dots, a_P parameters have been estimated, Eq. (10.4.1) can be used to obtain a reconstruction of the entire signal phase (and using Eq. (10.4.2), the IF). Like the PWVD approach this method performs very well at high SNR, but suffers from threshold effects at low SNR. As with the PWVDs, an iterative method can be used to improve performance at low SNR, as detailed in Ref. [52].

10.4.4 COMPARATIVE PERFORMANCE EVALUATION OF PWVDs AND HAFs

10.4.4.1 Single component analysis

For a single polynomial phase signal, the computational costs of the PWVD and HAF methods are similar. The asymptotic variances of the a_k parameter estimates using the PWVD are also similar. If the estimate refinement scheme outlined in Steps 4-7 of [Section 10.4.2.5](#) is used for both methods, the two techniques are both optimal above threshold. The SNR thresholds for the HAF and PWVD techniques are similar for lower order phase laws (i.e., third order and lower), while for higher order laws PWVDs tend to have lower thresholds. Threshold values are given in Refs. [49,50].

10.4.4.2 Multiple component analysis

Both PWVDs and HAFs, being nonlinear, produce “cross-terms” in the case of two or more polynomial phase signals, i.e., “multicomponent” signals. The PWVDs produce cross-terms which are often spectrally concentrated, while the HAFs give rise to cross-terms which are spectrally dispersed. The spectrally dispersed cross-terms prove to be less problematic because they appear more like white noise.

In analyzing multiple components with the HAFs, one can first estimate the parameters of the highest amplitude component and then remove the estimate of this component from the signal [53]. Then proceed with the next highest amplitude component. The process can be repeated until all components are accounted for. Conditions for the effective analysis of multicomponent signals include (1) higher required SNR thresholds than for single-component analysis, (2) good separation of the various components in the HAFs, and (3) phase polynomial orders of signal components which are

not too high. Requirement (2) is necessary because the level of dispersion of the cross-terms is related to the extent of the separation of components in the HAFs. Requirement (3) is necessary because the energy of the cross-terms relative to the auto-terms increases with increasing HAF order. In addition, component extraction methods discussed in [Section 10.6](#) can be used to obtain a coarse IF estimate, which can then be refined using the above methods.

10.4.5 SUMMARY AND CONCLUSIONS

This section discusses the problem of analyzing polynomial FM signals in additive noise using two different approaches: the polynomial Wigner-Ville distributions and the higher order ambiguity functions. Both methods perform well at high SNR, and can be modified to work well even at low SNRs.

10.5 IF ESTIMATION OF FM SIGNALS IN MULTIPLICATIVE NOISE⁰

10.5.1 RANDOM AMPLITUDE MODULATION

Most IF estimation techniques, such as those presented in the previous sections of this chapter, assume that the signal of interest has a constant amplitude. While this is a valid assumption in a wide range of scenarios, there are several important applications in which this assumption does not hold. Indeed, in many situations the signal may be subjected to a *random* amplitude modulation which behaves as multiplicative noise [54]. Examples include fading in wireless communications [55], fluctuating targets in radar [56], and structural vibration of a spacecraft during launch and atmospheric turbulence [57]. This section focuses on nonparametric methods; in particular, it shows that the Wigner-Ville distribution (defined in [Section 2.1.4](#)) is able to display the IF of a signal affected by multiplicative noise, and that this representation is optimal in the sense of maximum energy concentration for a linear FM signal. For higher-order polynomial FM signals, the use of the polynomial Wigner-Ville distribution (PWVD), presented in [Section 5.4](#), is shown to give optimal representations. Statistical performance of each case will be presented here.

10.5.2 LINEAR FM SIGNAL

This section presents the case of a linear FM signal, with a multiplicative noise that is a *real-valued* process.

10.5.2.1 Optimality of the Wigner-Ville spectrum

First let us show that the Wigner-Ville spectrum (WVS) is optimal, in the sense of IF localization, for the (t,f) analysis of linear FM signals affected by multiplicative noise. Consider the signal $y(t)$ given by

$$y(t) = a(t) \cdot z(t), \quad (10.5.1)$$

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **B. Barkat**, Petroleum Institute, Abu Dhabi, United Arab Emirates (bbarkat@pi.ac.ae). Reviewers: Dr. F. Sattar and S. Gulam Razul.

where $a(t)$ is a non-zero-mean real-valued stationary noise and $z(t)$ is a deterministic FM signal given by $z(t) = \exp\{j\phi(t)\}$. For a linear FM signal, $\phi(t)$ is a second-order polynomial. Using the expectation operator notation, the autocorrelation of the signal above can be expressed as [54]

$$\begin{aligned}\mathcal{K}_y(t, \tau) &= E\left[y\left(t - \frac{\tau}{2}\right) y^*\left(t + \frac{\tau}{2}\right)\right] \\ &= E\left[a\left(t - \frac{\tau}{2}\right) a\left(t + \frac{\tau}{2}\right)\right] \cdot \left\{z\left(t - \frac{\tau}{2}\right) z^*\left(t + \frac{\tau}{2}\right)\right\} \\ &= \mathcal{R}_a(\tau) K_z(t, \tau).\end{aligned}\quad (10.5.2)$$

The WVS of $y(t)$, which is defined as the FT of $\mathcal{K}_y(t, \tau)$ [26], can be expressed as

$$\mathcal{W}_y(t, f) = \mathcal{F}_{\tau \rightarrow f}\{\mathcal{K}_y(t, \tau)\} \quad (10.5.3)$$

$$= \mathcal{F}_{\tau \rightarrow f}\left\{z\left(t - \frac{\tau}{2}\right) z^*\left(t + \frac{\tau}{2}\right)\right\}_f * \mathcal{F}_{\tau \rightarrow f}\left\{E\left[a\left(t - \frac{\tau}{2}\right) a\left(t + \frac{\tau}{2}\right)\right]\right\} \quad (10.5.4)$$

$$= W_z(t, f) * S_a(f), \quad (10.5.5)$$

where $*$ is the convolution operation in the frequency space.

If we express the non-zero-mean random process $a(t)$ as $a(t) = \mu_a + a_0(t)$, where μ_a is a constant mean of $a(t)$ and $a_0(t)$ is a zero-mean noise with autocorrelation $R_{a_0}(\tau)$, we can rewrite the WVS of $y(t)$ as

$$\mathcal{W}_y(t, f) = \mu_a^2 W_z(t, f) + S_{a_0}(f) * W_z(t, f). \quad (10.5.6)$$

For the case of a linear FM signal, the Wigner-Ville distribution (WVD) is given in Section 2.1 by [26]

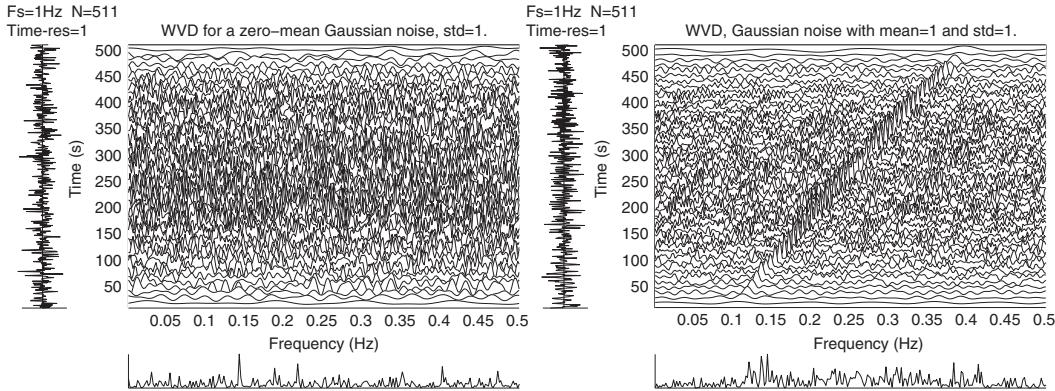
$$W_z(t, f) = \delta(f - f_i(t)), \quad (10.5.7)$$

where $f_i(t)$ is the signal IF and δ is the Dirac delta function. In this case, we obtain

$$\mathcal{W}_y(t, f) = \mu_a^2 \delta(f - f_i(t)) + S_{a_0}(f - f_i(t)). \quad (10.5.8)$$

Note that Eq. (10.5.8) exhibits the presence of a spectral line at the frequency $f_i(t)$ for all time instants. This means that, theoretically, the WVS always localizes the IF of a linear FM signal. This makes it a powerful tool in the analysis of linear FM signals affected by multiplicative noise. Also note that when $\mu_a = 0$, the WVS will not exhibit a peak at the signal IF, indicating a breakdown of the WVS in analyzing the noisy signal.

As an illustration, consider a unit-modulus linear FM signal sampled at 1 Hz, whose frequency range lies between 0.1 and 0.4 Hz. The signal length (in samples) is chosen as $N = 511$. This signal is multiplied by a real-valued i.i.d. Gaussian noise with a standard deviation $\sigma_a = 1$ and a mean equal to 0 and 1, respectively. Figure 10.5.1 displays the WVS (one realization) of the noisy signal for both values of the mean. As expected, the WVS for the zero-mean case cannot reveal the signal IF; however, it can do so for the other case. In this last case, the peak of the WVS can be used to estimate the IF of the signal. In what follows, we will evaluate the statistical performance of such an estimator.

**FIGURE 10.5.1**

The Wigner-Ville spectrum of a linear FM signal affected by a real-valued Gaussian multiplicative noise. The noise variance is 1 (both plots) and its mean is 0 (left plot) and 1 (right plot).

10.5.2.2 Statistical performance evaluation

Here, for a more complete study, we consider the presence of additive noise as well as multiplicative noise. The objective is to derive the asymptotic variance of the IF estimator, based on the peak of the WVS, for this case.

Let the discrete-time version of the noisy signal be

$$y(n) = a(n) e^{j\phi(n)} + w(n), \quad n = 0, \dots, N - 1. \quad (10.5.9)$$

The process $a(n)$ is considered to be a real-valued stationary Gaussian noise with mean and variance given by μ_a and σ_a^2 , respectively. The complex zero-mean additive process $w(n)$ is assumed to be stationary, white, circular, and Gaussian with variance equal to σ_w^2 . In addition, both noises are assumed to be independent.

The WVS used to estimate the signal IF is defined, in the discrete-time domain, using the expression given in Section 6.1, as [26]

$$\mathcal{W}_z(n, f) = E \left[2 \sum_{m=-M}^M y(n+m) \cdot y^*(n-m) e^{-j4\pi fm} \right]. \quad (10.5.10)$$

Straightforward derivations show that, for increasing window length $(2M + 1)$, the WVS converges in probability to $\mu_a^2 \delta(f - f_i(t))$ [58]. We can also show that the IF estimator asymptotic variance is approximately equal to [58]

$$\text{Var}(\hat{f}_i(n)) = \frac{3}{(2\pi)^2 \mathcal{S}_w (2M + 1)^3} \left(2 + \frac{2}{\mathcal{S}_a} + \frac{1}{\mathcal{S}_w} \right), \quad (10.5.11)$$

where $\mathcal{S}_a = \mu_a^2 / \sigma_a^2$ and $\mathcal{S}_w = \mu_w^2 / \sigma_w^2$. Note that:

- (i) When $\mu_a = 0$, the variance goes to infinity indicating that the WVS-based estimator breaks down. This result confirms the analysis presented earlier.
- (ii) When $\mu_a = A$ where A is a constant and $\sigma_a = 0$, i.e., the signal under consideration is just a constant amplitude linear FM signal embedded in noise, the variance expression can be rewritten as

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{3\sigma_w^2[2A^2 + \sigma_w^2]}{(2\pi)^2 A^4 (2M + 1)^3}. \quad (10.5.12)$$

If, in addition, we assume high signal-to-noise ratio (SNR) (i.e., $A^2 \gg \sigma_w^2$), the asymptotic variance expression reduces to

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{6\sigma_w^2}{(2\pi)^2 A^2 (2M + 1)^3}, \quad (10.5.13)$$

which is similar to the result obtained in Ref. [31].

The above theoretical results were confirmed by Monte-Carlo simulations. Specifically, we estimate the IF estimator variance using 5000 realizations of the signal given by Eq. (10.5.9). In Fig. 10.5.2, we display the theoretical (dashed curve) and the estimated (+) variances plotted against S_a for $S_w = 0$ and 5 dB, respectively.

If the signal under consideration is not a linear FM but a higher-order polynomial FM signal, the WVS becomes inappropriate because it introduces some artifacts which might hide the real features of the signal and its peak-based IF estimator is biased for such signals [31]. In this situation, PWVDs are used for the analysis, as discussed in the next section.

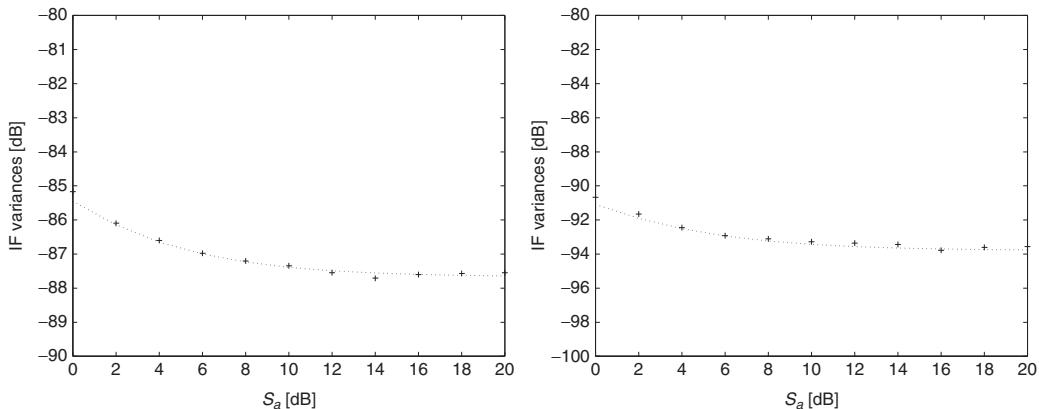


FIGURE 10.5.2

Theoretical (dashed curve) and estimated (+) variances of a linear FM signal corrupted by multiplicative and additive noise. Left plot is for $S_w = 0$ dB and right plot is for $S_w = 5$ dB.

10.5.3 POLYNOMIAL FM SIGNALS

The previous section is extended twofold: (i) the signal considered is assumed to be a polynomial FM signal of arbitrary order and (ii) the multiplicative noise is no longer limited to a real-valued process but is assumed to be a non-zero complex circular Gaussian process. The noisy signal $y(t)$ can then be written as

$$y(t) = a(t) z(t) + w(t), \quad (10.5.14)$$

where the stationary processes $a(t)$ and $w(t)$ are both assumed circular, complex, Gaussian, and independent with means and variances given by (μ_a, σ_a^2) and $(0, \sigma_w^2)$, respectively. The noiseless polynomial FM signal $z(t)$ is given by

$$z(t) = e^{j\phi(t)} = \exp \left\{ j \sum_{i=0}^P a_i t^i \right\}, \quad (10.5.15)$$

where the a_i are real coefficients and P is the order of the polynomial phase. Note that the derivation below does not require knowledge of the coefficients a_i ; it only assumes a polynomial FM signal.

The IF of the signal $z(t)$ is given in [Section 1.3](#) by

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \sum_{i=1}^P i a_i t^{i-1}$$

and the primary objective here is to estimate $f_i(t)$ from the noisy signal $y(t)$. For that purpose, we use the polynomial Wigner-Ville distribution (PWVD) defined in [Section 5.4](#) as

$$W_z^{(q)}(t, f) = \int_{-\infty}^{\infty} \prod_{i=1}^{q/2} z(t + c_i \tau) z^*(t + c_{-i} \tau) e^{-j2\pi f \tau} d\tau \quad (10.5.16)$$

$$= \int_{-\infty}^{\infty} K_z^{(q)}(t, \tau) \cdot e^{-j2\pi f \tau} d\tau, \quad (10.5.17)$$

where q is an even integer which indicates the order of nonlinearity of the PWVD. The coefficients c_i and $c_{-i}(i = 1, 2, \dots, q/2)$ are calculated so that the PWVD is real and equal to

$$W_z^{(q)}(n, f) = \delta(f - f_i(t)),$$

for signals given by Eq. (10.5.15). Note that the realness of the PWVD implies that $c_i = -c_{-i}$. Also note that the WVD is a member of the PWVDs class with parameters $q = 2$ and $c_1 = -c_{-1} = 0.5$. Full details of the design procedure may be found in Ref. [42] and [Sections 5.4](#) and [5.5](#).

The choice of the PWVD stems from the fact that it yields a continuum of delta functions around the IF for a given polynomial FM signal [54]. This property implies that the peak of the PWVD can be used as an IF estimator for polynomial FM in a noisy environment. In Ref. [59], the statistical performance of this estimator was evaluated for noisy signals described by Eq. (10.5.14). It shows that this estimator is unbiased and its asymptotic variance is approximately equal to [59]

$$\text{Var}(\hat{f}_i(n)) = \frac{6(\sigma_a^2 + \sigma_w^2) \sum_{i=1}^{n_1} k_i^2}{(2\pi)^2 |\mu_a|^2 (2M+1)^3}. \quad (10.5.18)$$

In the above expression, n_1 represents the number of different coefficients c_i in the PWVD kernel, while k_i (for $i = 1, \dots, n_1$) represents the multiplicity of each of these coefficients c_i , and $(2M + 1)$ is the window length considered in the PWVD discrete-time implementation. Note that:

- (i) When $\mu_a = 0$, the variance goes to infinity indicating that the PWVD-based estimator breaks down.
- (ii) When $\mu_a = A$ where A is a constant and $\sigma_a = 0$, i.e., the signal under consideration is just a constant amplitude polynomial FM signal embedded in complex Gaussian noise; the variance expression can then be rewritten as

$$\mathbf{Var}(\hat{f}_i) = \frac{6\sigma_w^2 \sum_{i=1}^{n_1} k_i^2}{(2\pi)^2 A^2 (2M + 1)^3}. \quad (10.5.19)$$

The above expression is exactly similar to the result obtained in Ref. [31], which treats *constant* amplitude polynomial FM signals only.

10.5.3.1 Monte-Carlo simulations

To confirm the validity of the above theoretical results, let us consider the IF estimation of a quadratic FM signal at the middle of the signal interval. The peak of the sixth-order PWVD, whose signal kernel is [42]

$$K_y^{(6)}(t, \tau) = y(t + 0.62\tau)y^*(t - 0.62\tau)y(t + 0.75\tau)y^*(t - 0.75\tau)y(t - 0.87\tau)y^*(t + 0.87\tau), \quad (10.5.20)$$

is used here as the IF estimator. The noisy signal $y(t)$ is generated as suggested by Eq. (10.5.14). For this example, the sampling period is $T = 1$, the signal length is $N = 129$, the window length is $2M + 1 = N = 129$, and the noise variances are equal (i.e., $\sigma_a^2 = \sigma_w^2$). In the simulations, the overall signal-to-noise ratio is defined as

$$\text{SNR}_{w_1} = 10 \log_{10} \left(|\mu_a|^2 / (\sigma_a^2 + \sigma_w^2) \right),$$

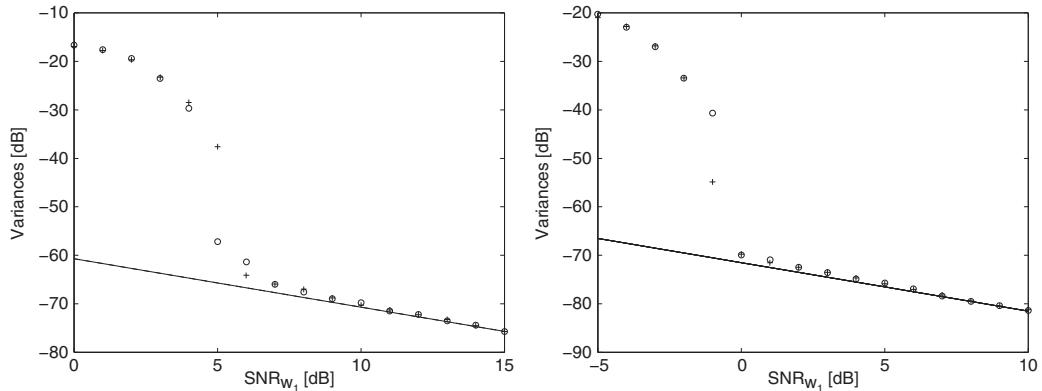
is varied in 1 dB steps from 0 to 15 dB. Monte-Carlo simulations for 1000 realizations are run for each value of SNR_{w_1} . The left plot of Fig. 10.5.3 shows the results of two different experiments, one for $|\mu_a| = 0.01$ and the other for $|\mu_a| = 1$. Observe that above a certain threshold, the estimated variances represented by “+” (for $|\mu_a| = 0.01$) and “o” (for $|\mu_a| = 1$) are in total agreement with the theoretical ones given by Eq. (10.5.18) and represented by the continuous lines (superimposed).

Simulations run under the same noise conditions for other polynomial FM signals using the appropriate PWVD order also confirm the theoretical results presented above, including when $y(t)$ is a linear FM signal and the PWVD considered is the second-order PWVD (see the right-hand plot of Fig. 10.5.3).

10.5.4 TIME-VARYING HIGHER-ORDER SPECTRA

A class of Time-Varying Higher-Order Spectra (TV-HOS) can be defined by taking the expected value of the PWVD [54], namely

$$\mathcal{W}_z^{(q)}(t, f) = E \left[W_z^{(q)}(t, f) \right], \quad (10.5.21)$$

**FIGURE 10.5.3**

Results of the experiments for a **quadratic** (left plot) and a **linear** (right plot) FM signal corrupted by complex Gaussian multiplicative and additive noise processes. The continuous lines (superimposed) represent the theoretical variances while “+” and “o” correspond to the estimated variances for $|\mu_a| = 0.01$ and $|\mu_a| = 1$, respectively.

where $W_z^{(q)}(t, f)$ is the q th-order PWVD defined by Eq. (10.5.16). Interchanging the expectation operator with the integration in the PWVD, one obtains [41,54]

$$\mathcal{W}_z^{(q)}(t, f) = \int_{-\infty}^{\infty} E \left[\prod_{i=1}^{q/2} z(t + c_i \tau) z^*(t + c_{-i} \tau) \right] e^{-j2\pi f \tau} d\tau \quad (10.5.22)$$

$$= \int_{-\infty}^{\infty} \mathcal{K}_z^{(q)}(t, \tau) \cdot e^{-j2\pi f \tau} d\tau, \quad (10.5.23)$$

where $\mathcal{K}_z^{(q)}(t, \tau)$ represents a slice of a time-varying q th-order moment function [54]. If the quantity $\mathcal{K}_z^{(q)}(t, \tau)$ is absolutely integrable, then $\mathcal{W}_z^{(q)}(t, f)$ can be interpreted as a form of time-varying higher-order moment spectrum. Such TV-HOS combine the advantages of classical time-frequency analysis with the benefits of higher-order spectra [54]. To avoid the problem of non-superposition of the higher-order moments, higher-order cumulants can be used instead [54].

As nonstationary random signals are non-ergodic, the ensemble averaging above cannot be replaced by time averaging; so *local* ergodicity needs to be assumed as done for the estimation of the WVS in section 2.2.2.

Readers interested in TV-HOS are referred to Ref. [54] for more details and examples of the efficacy of TV-HOS in the analysis of random FM signals affected by multiplicative noise. (See also Section 14.5.4 and Ref. [54].) A method for detecting IF laws can be found in Ref. [60]. This method uses an algorithm based on TV-HOS, which is then applied to frequency-modulated (FM) signals corrupted by multiplicative and additive noise.

10.5.5 SUMMARY AND CONCLUSIONS

The Wigner-Ville spectrum (WVS) and polynomial Wigner-Ville distributions (PWVDs) are applied to the analysis of polynomial FM signals corrupted by multiplicative and additive noise. In the noisy linear FM case, the WVS is shown to give optimal IF localization, and the peak of the WVS can therefore be proposed as an IF estimator. A statistical performance test shows that this estimator is very accurate even at low SNR. For the case of noisy higher-order polynomial FM signals, the peak of the PWVD is shown to be a consistent and accurate IF estimator. Although the methods presented here were validated using a single component case, they can be extended to multicomponent signals as well.

10.6 COMPONENT EXTRACTION FROM TFDs FOR MULTICOMPONENT SIGNAL IF ESTIMATION⁰

The instantaneous frequency (IF) of a signal is a component-specific feature; therefore, its estimation is preceded or accompanied by component separation and extraction. This section considers nonstationary multicomponent signals, as defined in [Chapter 1](#), and the problem of estimating the IFs corresponding to the individual components [4]. Two solutions are presented.

The first approach uses a blind source separation (BSS) algorithm whose aim is to recover unobserved signal components from their observed mixture. The term “blind” signifies that no *a priori* knowledge of signal components or the mixing process is required. Much research has been carried out in the field of BSS due to its broad range of applications (e.g., sonar array processing, data communication, seismic exploration, biological signal processing), leading to numerous approaches being proposed over the last few decades. One such algorithm for component localization and extraction from a noisy multicomponent signal TFD was introduced in Ref. [61], and its extension that additionally includes the components’ time support information was proposed in Ref. [62]. The method is summarized in [Section 10.6.1](#) and illustrated using an example of a multicomponent signal.

The two-dimensional nature of a TFD implies that image processing techniques, the second approach, are also suitable methods for extracting relevant time-frequency information, including the IFs of signal components. Pattern recognition techniques, such as the Hough transform, have been proposed as methods for obtaining an IF estimate from a TFD [63–65]. Edge detection using the matched filter technique [66] was incorporated in Ref. [65] to improve the IF estimation. A method of edge detection by linking edges through a sequential search [67] was adopted as the feature extraction method to estimate the IFs of multicomponent signals [68]. This algorithm is described in [Section 10.6.2](#), including an illustrative example.

Finally, the two methods are tested on a real-life signal, and the IF estimation results are presented.

⁰Authors: **V. Sucic**, University of Rijeka, Rijeka, Croatia (vsucic@riteh.hr), **J. Lerga**, University of Rijeka, Rijeka, Croatia (jlerga@riteh.hr), **L.J. Rankine**, Business Engineers Australia, QLD (formerly University of Queensland, Brisbane, QLD, Australia) (luke@beaustralia.com.au), and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: S. Stanković and Z.M. Hussain.

10.6.1 A BLIND SOURCE SEPARATION-BASED METHOD

10.6.1.1 The algorithm

We start with a BSS algorithm for component localization and separation based on a double-directional component tracking and extraction approach. The algorithm is complemented by the components' time-support information, which is obtained from the short-term Rényi entropy (STRE) of the signal TFD. The flowchart of the algorithm is presented in Fig. 10.6.1 and it involves the following steps:

Step 1: A multicomponent signal TFD, $\rho(t,f)$, is calculated.

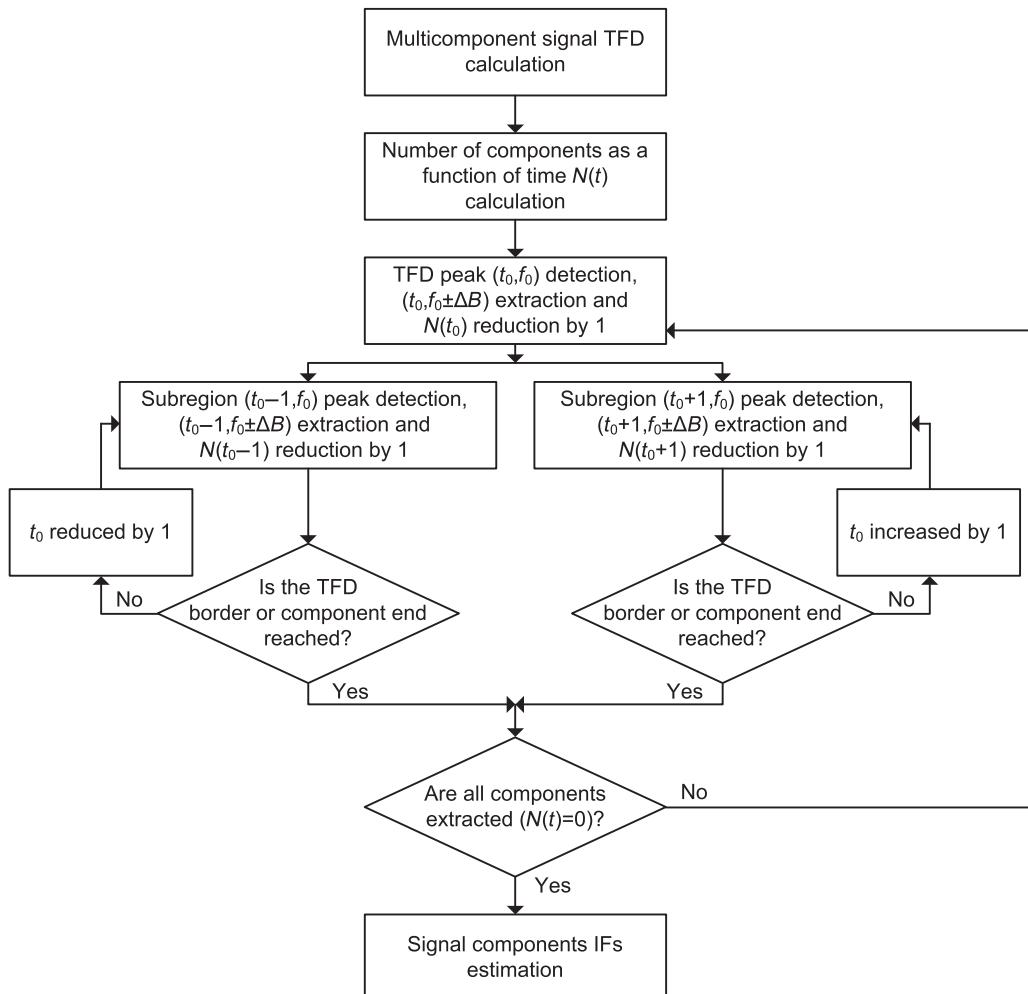


FIGURE 10.6.1

The flowchart of the BSS-STRE-based component-extraction algorithm.

Step 2: The local number of signal components, $N(t)$, is found as [69]

$$N(t) = 2^{H_\alpha(\rho(t,f)) - H_\alpha(\rho_{\text{ref}}(t,f))}. \quad (10.6.1)$$

The Rényi entropy of $\rho(t,f)$ is defined as

$$H_\alpha(\rho(t,f)) = \frac{1}{1-\alpha} \log_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^\alpha(t,f) dt df, \quad (10.6.2)$$

where α is the order of Rényi entropy, and $\rho_{\text{ref}}(t,f)$ is the TFD of the reference signal (see [Section 7.7](#) and Ref. [69] for discussions of the STRE and its properties).

Step 3: The largest TFD peak is detected, at (t_0, f_0) , with its neighboring frequency band, ΔB . The parameter ΔB should be selected as small as possible, but still large enough, in order to extract a whole component, and not just a part of it. The components' frequency separation in the signal TFD must also be taken into account when selecting the bandwidth ΔB [61]. For example, in high-resolution TFDs, having well-concentrated components in the (t,f) plane, smaller ΔB values can be used.

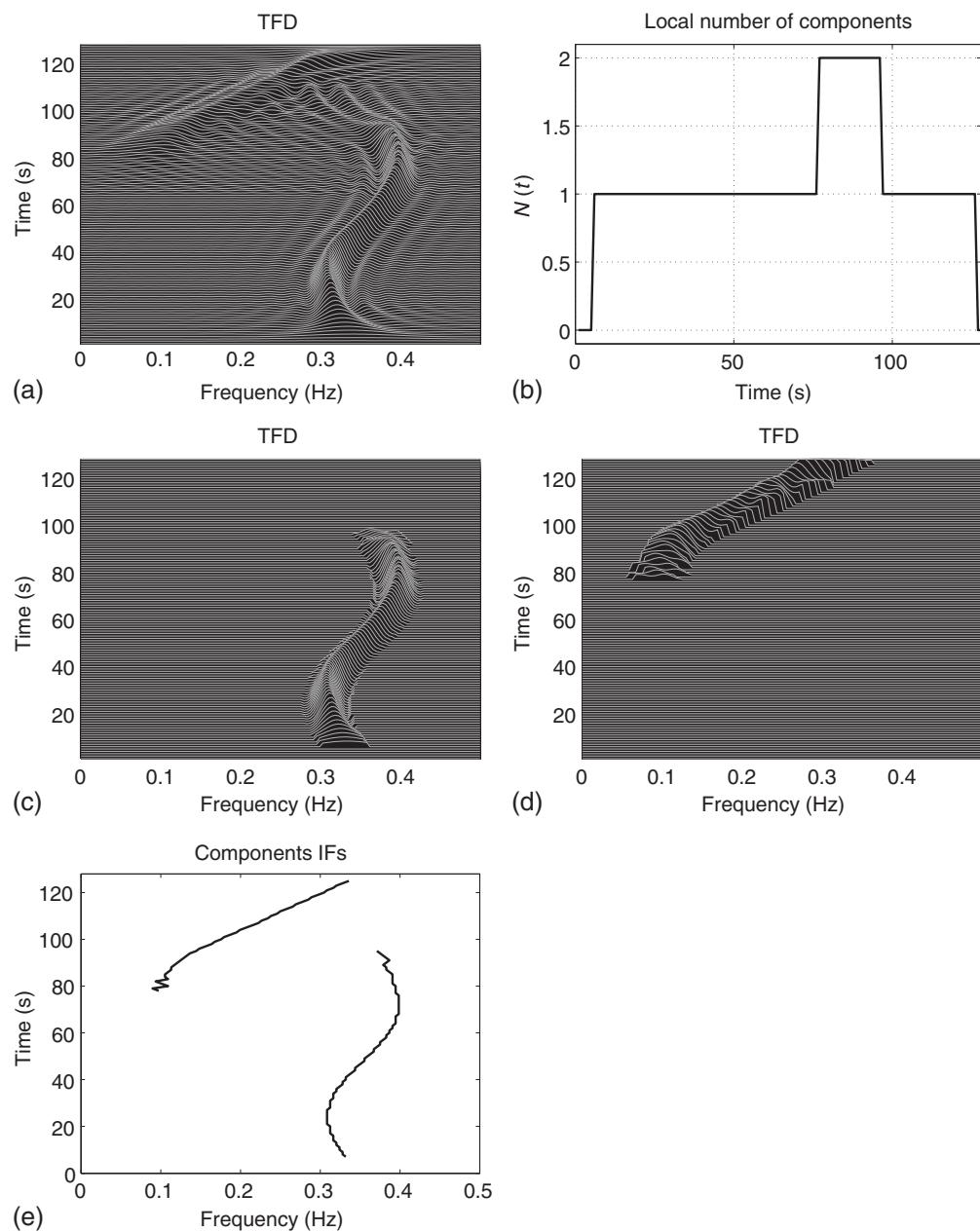
Once the component peak and its frequency band are extracted, the time-frequency interval $(t_0, f_0 \pm \Delta B)$ is set to zero, and $N(t_0)$ is reduced by one. Two new subregions are defined as $(t_0 - 1, f_0 \pm \Delta B)$ and $(t_0 + 1, f_0 \pm \Delta B)$, followed by the largest peak detection in each of them. The procedure is repeated until the TFD borders or the component end is reached for each subregion, resulting in a single component being extracted.

Step 4: Step three is repeated for the remaining components in the signal TFD, i.e., until $N(t)$ becomes zero. Once all components are extracted, the individual IFs can be estimated using TFD-based monocomponent IF estimation methods (see, e.g., [Section 10.2](#) and Ref. [25]).

10.6.1.2 A two-component signal example

Consider a signal $z[n] = z_1[n] + z_2[n]$, where $z_k[n] = A_k \exp(j\phi_k[n])$, $k = 1, 2$ ($A_k = 1$, signal length $N = 128$), containing a sinusoidal FM component and a linear FM component, whose time supports partly overlap. The signal TFD is shown in [Fig. 10.6.2\(a\)](#). Here we use a reduced interference distribution referred to as the modified B-distribution (MBD), as defined in [Chapter 3](#). The components' respective IF laws are defined as $f_1[n] = (1/2\pi)(0.35 + 0.05 \cos(2\pi[n - N_{1/2}]/N_1 - \pi/2))$, and $f_2[n] = (1/2\pi)(0.05 + 0.3[n - 1]/(N_2 - 1))$, where the components' durations are $N_1 = 96$ and $N_2 = 48$, respectively.

Accurate estimation of component time support plays a key role in the IF estimation error reduction [62]. In addition, the information on time support and the local number of components is used in the component-extraction procedure in order to avoid data-dependent thresholding. The local number of components of signal $z[n]$ obtained using the STRE, the two extracted components' MBDs, and their IFs are respectively given in [Fig. 10.6.2\(b\)-\(e\)](#).

**FIGURE 10.6.2**

(a) MBD ($\beta = 0.1$) of the signal $z[n]$; (b) the local number of components; (c) the first extracted component; (d) the second extracted component; (e) the estimated components' IFs.

10.6.1.3 Strengths and weaknesses of the method

The main advantage of the BSS-STRE-based method for component extraction is its double-directional component tracking capability. This ensures complete, rather than partial, component extraction. Therefore, the BSS-STRE-based method does not need classification procedures for component-segment linking. In moderate-noise environments, for which the STRE is known to perform well [69], the presented algorithm does not require any noise reduction thresholding [62].

The method is applicable to signals having components with overlapping time supports; however, when dealing with intersecting components, special attention is needed to ensure that the intersection is assigned to each of the extracted intersecting components.

10.6.2 AN IMAGE PROCESSING-BASED METHOD

We next present an alternative method for multicomponent IF estimation, which is based on the local peak detection in the signal TFD, followed by an image processing technique called component linking.

10.6.2.1 The algorithm

The algorithm is divided into three sequential steps as follows:

Step 1: A signal TFD, $\rho(t,f)$, is first calculated.

Step 2: Local peaks of the TFD are detected. The detection is based on the first and second partial derivatives of the TFD with respect to frequency, resulting in a binary image [68]

$$B(t,f) = \begin{cases} 1 & \text{if } \left\{ \frac{\partial \rho(t,f)}{\partial f} = 0 \right\} \otimes \left\{ \frac{\partial^2 \rho(t,f)}{\partial f^2} < 0 \right\} \\ 0 & \text{otherwise,} \end{cases} \quad (10.6.3)$$

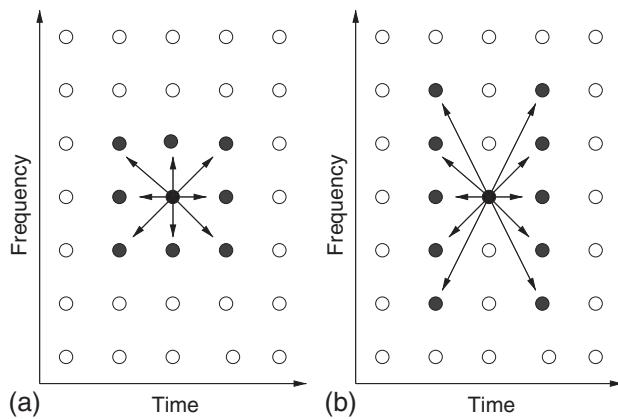
where \otimes is the Boolean AND operator. The ones in the binary image represent the local peaks in the TFD.

Step 3: Components are linked. The general method for component linking in image processing is to use the standard eight-connected neighboring pixels set [66], shown in Fig. 10.6.3(a); however, this is not a suitable set when it comes to IF estimation applications, since a signal component, by definition, can only have one IF at a given time instant. Hence, a more appropriate ten-connected pixels set, shown in Fig. 10.6.3(b), has been proposed [68]

$$\left\{ \begin{array}{l} (n-1, k-1), (n-1, k), (n-1, k+1), \\ (n+1, k-1), (n+1, k), (n+1, k+1), \\ (n-1, k-2), (n-1, k+2), (n+1, k-2), (n+1, k+2) \end{array} \right\}, \quad (10.6.4)$$

where n and k represent discrete time and discrete frequency, respectively (see Sections 6.1 and 6.5 for details on discrete forms of TFDs).

To finalize the component-linking algorithm, a threshold needs to be set for the number of pixels to define a true signal component. This number is directly related to the time support of the signal, and it is used to remove short, falsely linked components.

**FIGURE 10.6.3**

Connected neighboring sets: (a) 8 pixels (b) 10 pixels.

10.6.2.2 A three-component signal example

Consider a signal of length $N = 1024$ samples, and the sampling frequency of 1 Hz

$$s[n] = \sum_{k=1}^3 A_k[n] \cos(\phi_k[n]),$$

where

$$A_1[n] = 1 + 0.075 \cos(2\pi 0.00195 n); \quad \phi_1[n] = 2\pi(0.05 + 4.9 \times 10^{-5} n) n,$$

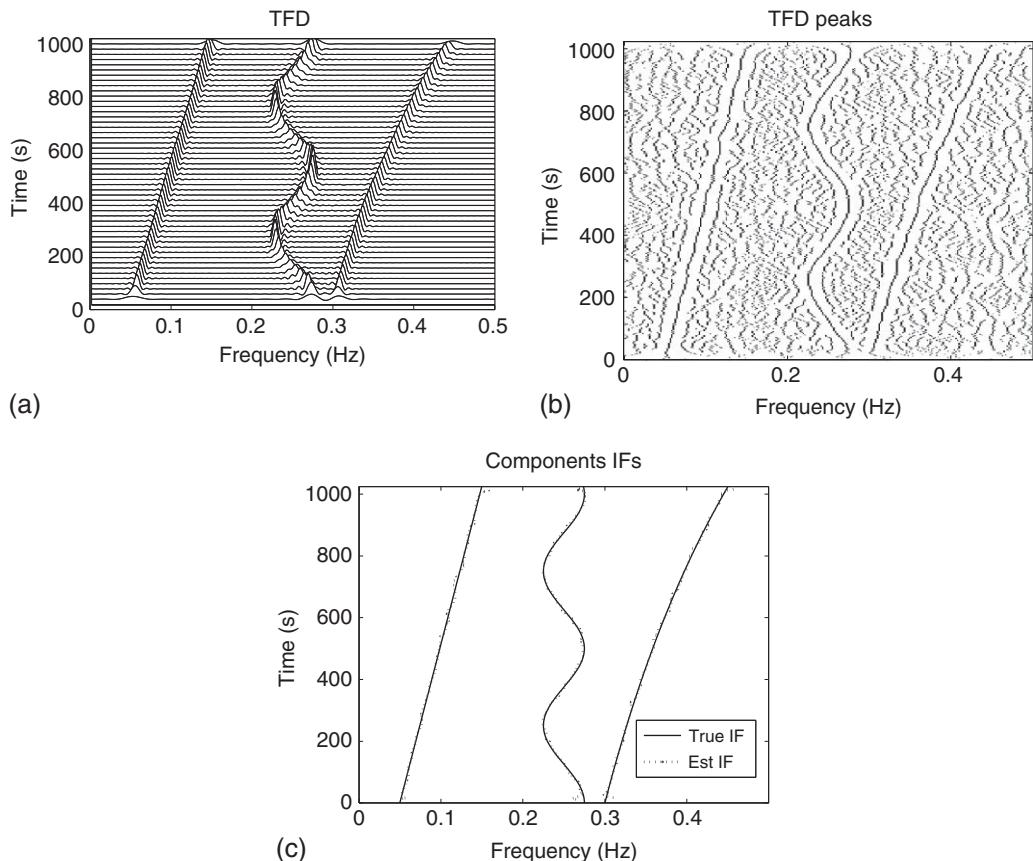
$$A_2[n] = 1 + 0.075 \cos(2\pi 0.00195 n + 1); \quad \phi_2[n] = 2\pi 0.25 n + 12.5 \sin(2\pi 0.002 n),$$

$$A_3[n] = 1 + 0.075 \cos(2\pi 0.00195 n + 2.5); \quad \phi_3[n] = \frac{2\pi 0.3}{-3.3 \times 10^{-4}} \ln(1 - 3.3 \times 10^{-4} n).$$

The MBD of the signal $s[n]$ is shown in Fig. 10.6.4(a), from which the three components are clearly identified. The binary representation of the local maxima is shown in Fig. 10.6.4(b). The components' IFs are estimated using the component linking algorithm and the ten-connected pixels set defined by Eq. (10.6.4). The estimated IFs are compared to the components' true IFs, and the results are plotted in Fig. 10.6.4(c).

10.6.2.3 Strengths and weaknesses of the method

The component-linking algorithm for estimating the IFs of multicomponent signals does not require *a priori* information regarding either the components' IF laws or their amplitudes. It can be also used with signals whose components have different IFs. Note that the components' IFs do not need to be represented by parametric functions.

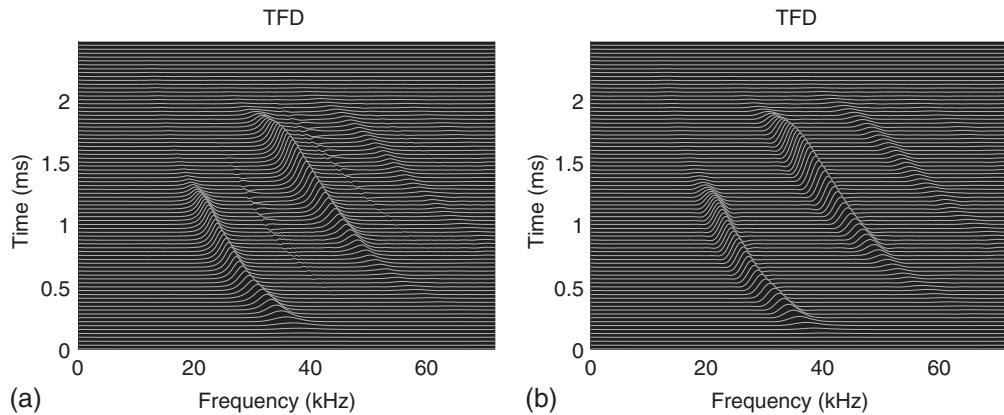
**FIGURE 10.6.4**

(a) MBD ($\beta = 0.2$) of the signal $s[n]$; (b) the binary image from detected local peaks; (c) the estimated (dashed) and true (solid) component IFs.

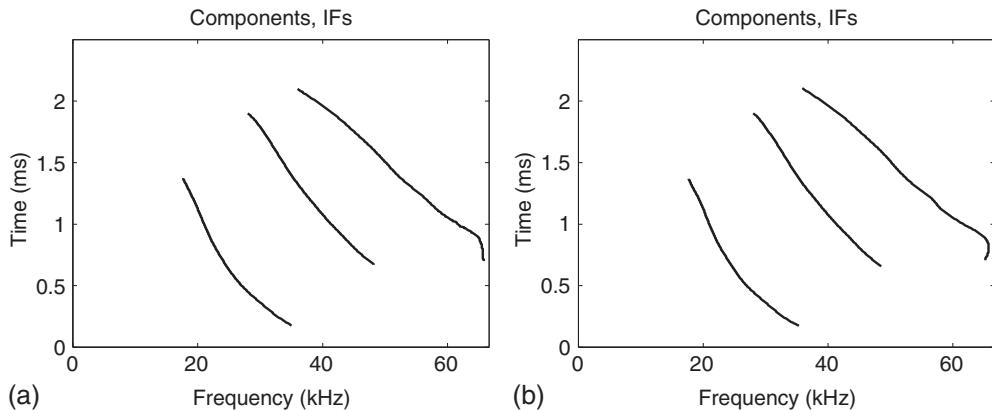
However, the algorithm is limited to the class of multicomponent signals referred to as time-frequency separable, i.e., signals whose components are clearly separated in the (t,f) domain, and which therefore possess a unique decomposition. In addition, it requires a threshold to be set for the minimum time duration of a component, which is directly related to the component time support information.

10.6.3 A REAL-LIFE SIGNAL EXAMPLE

We next illustrate the applicability of the two multicomponent IF estimation methods in real-life situations using a 2.5 ms echolocation pulse emitted by a bat (see Chapter 17). Two reduced interference TFDs, namely, the compact-kernel distribution (CKD) (Section 3.3.4) and the extended

**FIGURE 10.6.5**

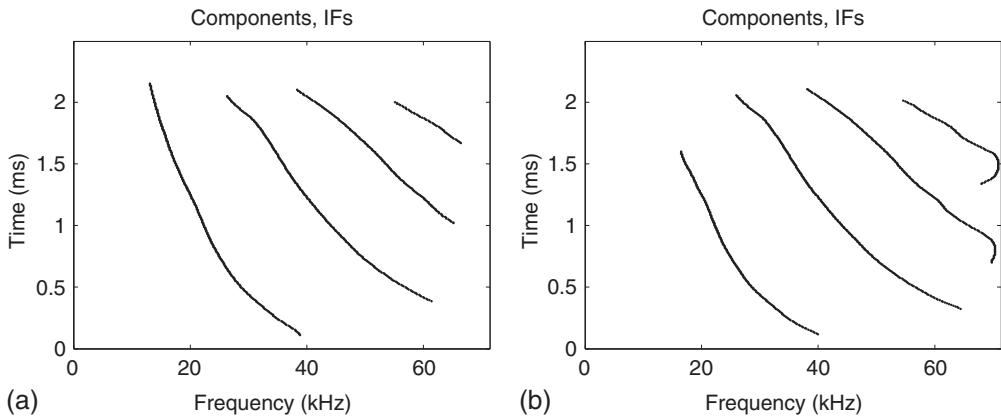
The bat signal: (a) EMBD ($\alpha = \beta = 0.1$) (b) CKD ($C = 4$, $D = 0.1$).

**FIGURE 10.6.6**

The bat signal components' IFs estimated by applying the BSS-STRE method to the signal: (a) EMBD (b) CKD.

separable-kernel MBD (EMBD) (Section 2.7.5), are used to represent the bat signal in the (t, f) plane (see Fig. 10.6.5). The signal components' IF estimates obtained using the BSS-STRE component extraction method are shown in Fig. 10.6.6, and the IF estimates obtained by applying the image-processing-based component-linking approach are given in Fig. 10.6.7.

From the presented results, we observe that both the EMBD and the CKD have produced interference-clean (t, f) representations of the bat signal, with the CKD being even more powerful in suppressing the cross-terms between the components (see Fig. 10.6.5(b)). The image-processing

**FIGURE 10.6.7**

The bat signal components' IFs estimated by applying the image-processing method to the signal: (a) EMBD (b) CKD.

method outperforms the BSS-STRE approach in detecting the bat signal components and estimating their IFs. Unlike the BSS-STRE method, the image-processing method identifies all four components in the signal, including the lowest energy component. For more details about the EMBD and CKD, see Refs. [36,70]. The failure of the BSS-STRE approach is due to its inability to count the total number of signal components.

10.6.4 SUMMARY AND CONCLUSIONS

The section has presented two methods for multicomponent IF estimation. The first method is a BSS-based one for localization and extraction of a signal's components from its TFD. It is a double-directional component-tracking technique, enhanced by information on the local number of components, as obtained from the signal STRE. The method is free from any additional component-classification procedures, as well as different thresholds that are commonly required by similar BSS approaches.

The second method for multicomponent IF estimation applies an image-processing-based technique to a TFD. It is applicable to time-frequency separable signals, and it does not require prior information regarding the total number of signal components. It also validates the use of image processing applied to TFDs for feature-extraction applications.

When SNR is low and the minimum duration of signal components is known in advance, the image-processing approach is a better choice for extracting signal components and estimating their respective IFs; however, when SNR is high and we do not have prior information regarding the minimum duration of signal components, the BSS-STRE method should be selected.

10.7 INSTANTANEOUS FREQUENCY AND POLYNOMIAL PHASE PARAMETER ESTIMATION USING LINEAR TIME-FREQUENCY REPRESENTATIONS⁰

10.7.1 POLYNOMIAL PHASE SIGNALS

The instantaneous frequency (IF) is an important feature of nonstationary signals. It can contain useful information regarding the signal's behavior and/or the phenomenon that generated it (i.e., its source). Communications based on frequency modulation (FM) were among the first applications of signals with time-varying frequency. In radar and sonar, the position and/or velocity of a target can be extracted from the IF of the returned signal. Seismological activities due to earthquakes and underground nuclear tests can be monitored using IF estimation techniques.

According to the Weierstrass theorem, signals with continuous phase can be successfully modeled by polynomial-phase signals (PPSs). Such signals are found in various applications, such as telecommunications, radar, sonar, seismic applications, and biomedicine (Section 5.4 and Ref. [54]). The P th order PPS is defined as

$$s(n) = A e^{j\phi(n)} = A e^{j \sum_{p=0}^P a_p (n\Delta)^p}, \quad |n| \leq N/2, \quad (10.7.1)$$

where $\phi(n)$ represents the phase with parameters $\{a_p | p \in [0, P]\}$, A the amplitude, Δ the sampling interval, and $N + 1$ the signal length (where N is assumed even). We assume that no aliasing occurs during sampling [71]. The IF of $s(n)$ is given by

$$f_i(n) = \frac{1}{2\pi} \sum_{p=1}^P p a_p (n\Delta)^{p-1} \quad (10.7.2)$$

and it comprises all the phase parameters except a_0 . In practice, the signal $s(n)$ is corrupted by a noise $\epsilon(n)$, so the signal to be analyzed can be written as

$$x(n) = s(n) + \epsilon(n). \quad (10.7.3)$$

Here we will assume that $\epsilon(n)$ is a complex zero-mean white Gaussian noise with variance σ_ϵ^2 .

The estimation of PPS parameters directly from the signal is practically infeasible due to the phase unwrapping ambiguity (emphasized with highly nonstationary signals) and the influence of noise (see Ref. [72] and references there in). In the last two decades, the most common direction in PPS parameter estimation is based on phase differentiation (PD) [51,73–75]. In the PD-based approach, the PPS order is decreased until a signal with low-phase order is obtained that can be processed using the standard techniques, for example the Fourier transform (FT). The PD reduces the PPS phase order by applying several successive auto-correlations, where one auto-correlation reduces the phase order by one. The main advantages of PD-based techniques are that they are unbiased and have low calculation complexity [51].

⁰Authors: **I. Djurović**, University of Montenegro, Podgorica, Montenegro (igordj@ac.me), **M. Simeunović**, University of Montenegro, Podgorica, Montenegro (markos@ac.me), and **S. Djukanović**, University of Montenegro, Podgorica, Montenegro (slobdj@ac.me). Reviewers: E. Sejdić and A. Napolitano.

PD-based estimators are closely related to estimators based on time-frequency distributions (TFDs), namely the polynomial Wigner-Ville distribution (Section 5.4 and Ref. [54]) and L-Wigner distribution [76]. These TFDs achieve high signal concentration in the time-frequency (t, f) plane around the IF and reduce estimation bias.

The auto-correlation is a nonlinear function, which introduces several fundamental drawbacks, such as increased signal-to-noise-ratio (SNR) threshold, spurious cross-terms when multicomponent PPSs are processed, and error propagation from higher- to lower-order PPS parameters [75]. Therefore, the PD-based techniques cannot be used with higher-order PPSs at low SNRs. Linear (t, f) representations (TFRs), such as the short-time Fourier transform (STFT) and local polynomial Fourier transform (LPFT), are biased for PPSs of order higher than two. Although this has been considered a significant drawback for a long time, linear TFRs are more robust to the noise influence than nonlinear techniques [77], making them good candidates for IF and PPS estimation in low-SNR scenarios.

In order to estimate the IF and PPS parameters in low-SNR environments, a linear TFR can be used to provide a coarse IF estimate, which can be further refined using techniques such as those in Ref. [78]. Such an approach is referred to as a quasi-maximum likelihood (QML) method, and it provides more robust and more accurate results than those based on the PD. In addition, the QML method can deal with signals having nonpolynomial phase.

In Section 10.7.2, a review of the most popular PD-based estimators is given. The QML method is introduced in Section 10.7.3, and its performance is compared to that of PD-based methods in Section 10.7.4.

10.7.2 PD-BASED ESTIMATORS

The maximum likelihood (ML) estimation of parameters of $s(n)$ is defined as

$$\begin{aligned} (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_P) &= \arg \max_{(b_1, \dots, b_P)} |\text{ML}(b_1, b_2, \dots, b_P)|, \\ \text{ML}(b_1, b_2, \dots, b_P) &= \sum_{n=-N/2}^{N/2} x(n) e^{-j \sum_{i=1}^P b_i (n\Delta)^i}. \end{aligned} \quad (10.7.4)$$

After the estimates \hat{a}_i for $i = 1, \dots, P$ are obtained, the amplitude and initial phase can be obtained from the amplitude and phase of the dechirped signal $x_d(n) = x(n) \exp(-j \sum_{i=1}^P \hat{a}_i (n\Delta)^i)$, respectively. The ML procedure requires a P -dimensional search, which is computationally daunting for $P > 3$. Therefore, alternative approaches for the estimation of higher order PPSs must be used.

Popular estimation approaches, based on the PD, include the high-order ambiguity function (HAF), product HAF (PHAF), cubic-phase function (CPF), and hybrid CPF-HAF (HCPF-HAF). The definitions of these transforms and their maxima positions, related to the PPS parameters, are given in Table 10.7.1.

In Table 10.7.1, $\text{PD}[\cdot]$ represents the PD operator defined as

$$\begin{aligned} \text{PD}_q[n, \tau_q] &= \text{PD}_{q-1}[n + \tau_q(q); \tau_{q-1}] \times \text{PD}_{q-1}^*[n - \tau_q(q); \tau_{q-1}], \\ \text{PD}_0[n] &= x(n), \end{aligned} \quad (10.7.5)$$

Table 10.7.1: Common PD Estimators: HAF, PHAF, CPF, and HCPF-HAF

| Definition | Maximum Position |
|--|--|
| $\text{HAF}(f) = \sum_n \text{PD}_{P-1}[n; \boldsymbol{\tau}_{p-1}] e^{-j2\pi f n \Delta}$ | $f = \frac{2^{P-2}}{\pi} P! a_p \Delta^{P-1} \prod_{i=1}^{P-1} \boldsymbol{\tau}_{p-1}(i)$ |
| $\text{PHAF}(f) = \prod_{l=1}^L \sum_n \text{PD}_{P-1}\left[n; \boldsymbol{\tau}_{p-1}^{(l)}\right] e^{-j\beta_l 2\pi f n \Delta}$ | $f = \frac{2^{P-2}}{\pi} P! a_p \Delta^{P-1} \prod_{i=1}^{P-1} \boldsymbol{\tau}_{p-1}^{(l)}(i)$ |
| $\beta_l = \prod_{i=1}^{P-1} \boldsymbol{\tau}_{p-1}^{(l)}(i) / \boldsymbol{\tau}_{p-1}^{(1)}(i)$ | |
| $\text{CPF}_n(\Omega) = \sum_m x(n+m)x(n-m) e^{-j\Omega(m\Delta)^2}$ | $\Omega = 6a_3 n \Delta + 2a_2$ |
| $\text{HCPF-HAF}_n(\Omega) =$ | $\Omega = 6C_3 a_p n \Delta + 2C_2 a_{p-1}$ |
| $\sum_m \left(\text{PD}_{P-3}[n+m; \boldsymbol{\tau}_{p-3}] \times \text{PD}_{P-3}[n-m; \boldsymbol{\tau}_{p-3}] e^{-j\Omega(m\Delta)^2} \right)$ | $C_3 = \frac{2^{P-4} P! \Delta^{P-3}}{3} \prod_{i=1}^{P-3} \boldsymbol{\tau}_{p-3}(i)$ |
| | $C_2 = \frac{3}{P} C_3$ |

where $\boldsymbol{\tau}_q = [\tau_1, \tau_2, \dots, \tau_q]$ is the vector of lag parameters used in the PD calculation, $\tau_q(i)$ its i th element, and * is the conjugation operator. In the PHAF definition, the superscript (l) denotes the l th vector of lag parameters, from a set of L vectors.

The parameter estimation is performed by maximizing the transform function over f or Ω . The maxima position is related to one (HAF, PHAF) or two (CPF, HCPF-HAF) highest-order PPS parameters, as indicated in the right column of Table 10.7.1. Once these parameters are estimated, the PPS is dechirped so the contribution of the estimated parameters is eliminated, and lower-order parameters such as initial phase are estimated by repeating the procedure on the dechirped signal.

The HAF uses the PD operator to transform the underlying PPS to a complex sinusoid [51]. From the sinusoid frequency, the parameter a_p can be estimated straightforwardly. The HAF is an acceptable solution for monocomponent lower-order PPSs ($P < 4$). For higher order PPSs, the HAF is characterized by higher SNR thresholds. In a multicomponent environment, the HAF produces undesirable cross-terms due to the nonlinear nature of the PD. The influence of noise and cross-terms is mitigated by the PHAF [73], which is obtained by multiplying several HAFs calculated with different lag sets. This product enhances desirable spectral components, and at the same time significantly attenuates the cross-terms; however, since the PHAF is also based on the PD, it is suitable for lower-order PPSs only. The number of cross-terms and noise terms in the PHAF grows exponentially with the PD order [75]. Each PD increases the SNR threshold by approximately 6 dB [51]. The transform that exploits this fact is the CPF [74], which reduces the PD order (for details, see Section 10.4). The CPF is proposed for the estimation of third-order PPSs (cubic-phase signals), requiring only one multiplication of the signal with its shifted version, as opposed to two multiplications required in the HAF and PHAF. In this way, the CPF gains in robustness against the noise compared to the latter ones. The CPF is generalized to higher-order PPSs in the HCPF-HAF [75]. Also, the product form of the HCPF-HAF is defined in Ref. [75].

To summarize, all mentioned transforms incorporate multiplications of the signal with its shifted version (and additionally conjugated in the PD operator), which results in an increase of number of

undesirable noise terms and the appearance of cross-terms when multiple PPSs are considered. The use of these techniques is restricted to low-order P and high SNR. An additional problem is the error propagation from higher- to lower-order PPS parameters due to the dechirping procedure.

The parameter estimation of higher-order PPSs can be improved by using a hybrid approach that combines nonparametric and parametric estimators. As nonparametric estimators, linear (t, f) methods, characterized by high robustness to noise, can be used to provide initial (coarse) IF estimates, from which the initial PPS parameter estimations can be obtained using the polynomial fitting. The final step is to refine the initial PPS estimates. The QML method, described in the following section, uses this strategy.

10.7.3 QML METHOD

In the QML method [78], the STFT is used to provide the coarse IF estimates, whereas the refinement is performed through signal dechirping, low-pass filtering and phase unwrapping [79]. The method can be summarized by the algorithm given below. The vector of parameters to be estimated is denoted as $\mathbf{a} = [a_1, \dots, a_P]^T$.

Step 1. Calculate the STFTs using the window function $w_h(m)$ with lengths h taken from a set \mathbf{H} , i.e.,

$$F_x^{w_h}(n, k) = \text{DFT}[x(n + m) w_h(m)],$$

where n and m represent discrete time variables, k the discrete frequency variable, $\text{DFT}[\cdot]$ the discrete FT operator, and the window function $w_h(m)$ satisfies $w_h(m) \neq 0$ for $|m| \leq h/2$ and $w_h(m) = 0$ elsewhere.

Step 2. Estimate the IF of $s(n)$ by locating the maxima position of $F_x^{w_h}(n, k)$ calculated in Step 1

$$\hat{f}_h(n) = \arg \max_k |F_x^{w_h}(n, k)|. \quad (10.7.6)$$

Step 3. Obtain the vector of PPS parameter estimates $\hat{\mathbf{a}}_h = [\hat{a}_1^h, \dots, \hat{a}_P^h]^T$, $h \in \mathbf{H}$, by polynomial fitting of the estimated IFs as follows:

$$\hat{\mathbf{a}}_h = (\mathbf{X}_h^T \mathbf{X}_h)^{-1} \mathbf{X}_h^T \mathbf{y}_h, \quad (10.7.7)$$

where $\mathbf{y}_h = 2\pi[\hat{f}_h(-N_h), \hat{f}_h(-N_h + 1), \dots, \hat{f}_h(N_h)]^T$, and $N_h = (N - h)/2$, and \mathbf{X}_h is a $(2N_h + 1) \times P$ matrix with $\mathbf{X}_h(i, j) = j[(-N_h + i - 1)\Delta]^{j-1}$, where $i = 1, \dots, 2N_h + 1$ and $j = 1, \dots, P$.

Step 4. Refine estimates $\hat{\mathbf{a}}_h$ using the approach proposed in Ref. [79] to obtain $\hat{\mathbf{a}}_h^r = [\hat{a}_1^{r,h}, \dots, \hat{a}_P^{r,h}]^T$, $h \in \mathbf{H}$.

Step 5. The final estimate is a vector $\hat{\mathbf{a}}_h^f = [\hat{a}_1, \dots, \hat{a}_P]^T$ that satisfies

$$\hat{\mathbf{a}}_h^f = \arg \max_{\hat{\mathbf{a}}_h^r} |\text{ML}(\hat{\mathbf{a}}_h^r)|, \quad (10.7.8)$$

$$\text{ML}(\hat{\mathbf{a}}_h^r) = \left| \sum_n x(n) e^{-j \sum_{p=1}^P \hat{a}_p^{r,h} (n\Delta)^p} \right|. \quad (10.7.9)$$

The STFT-based IF estimation accuracy depends on the window length h . In the QML algorithm, a set of predefined window lengths \mathbf{H} is considered. The optimal length is obtained by maximizing the ML function $\text{ML}(\hat{\mathbf{a}}_h^r)$.

Step 4 is necessary since the STFT is a biased IF estimator. The refinement approach [79] basically estimates the error due to the coarse PPS estimation, and it consists of:

- dechirping the original signal in order to concentrate its frequency spectrum around the direct-current component, i.e. $f = 0$
- low-pass filtering the dechirped signal to increase the SNR
- phase unwrapping of the filtered dechirped signal and
- estimating error by using linear regression

The PPS estimation based on the phase unwrapping is very accurate at high SNR, which is obtained by the low-pass filtering of the dechirped signal [79]. The mean squared error (MSE) of the refined estimates reaches the Cramér-Rao lower bound (CRLB) even for high PPS orders. In addition, the method [79] performs well even with less accurate initial parameter estimates.

The QML method provides both robustness to noise and high accuracy. The former comes from the use of the STFT which, being linear, does not incorporate the signal auto-correlation that increases the SNR threshold. The latter comes from the use of the least-squares polynomial fitting and the refinement method [79].

The implementation steps of the QML method imply that the QML-based estimation calculation complexity is mainly influenced by the STFT calculation; i.e., it requires $\mathcal{O}(|\mathbf{H}|N^2 \log N)$ operations, where $|\mathbf{H}|$ is the cardinality of the set \mathbf{H} . In contrast, the complexity of the HAF/PHAF-based procedures is $\mathcal{O}(PN \log N)$ operations [73]. Thus the cost of the excellent QML results is increased complexity.

The advantages and acceptable computational complexity of the QML make it a good candidate for PPS estimation in low SNR scenarios. In addition, it can be adopted to deal with signals whose phase is not polynomial, in which case an additional search over polynomial orders is required to obtain a polynomial that optimally fits the IF law [78].

10.7.4 SIMULATIONS

The QML method will be evaluated on PPS and non-PPS signals.

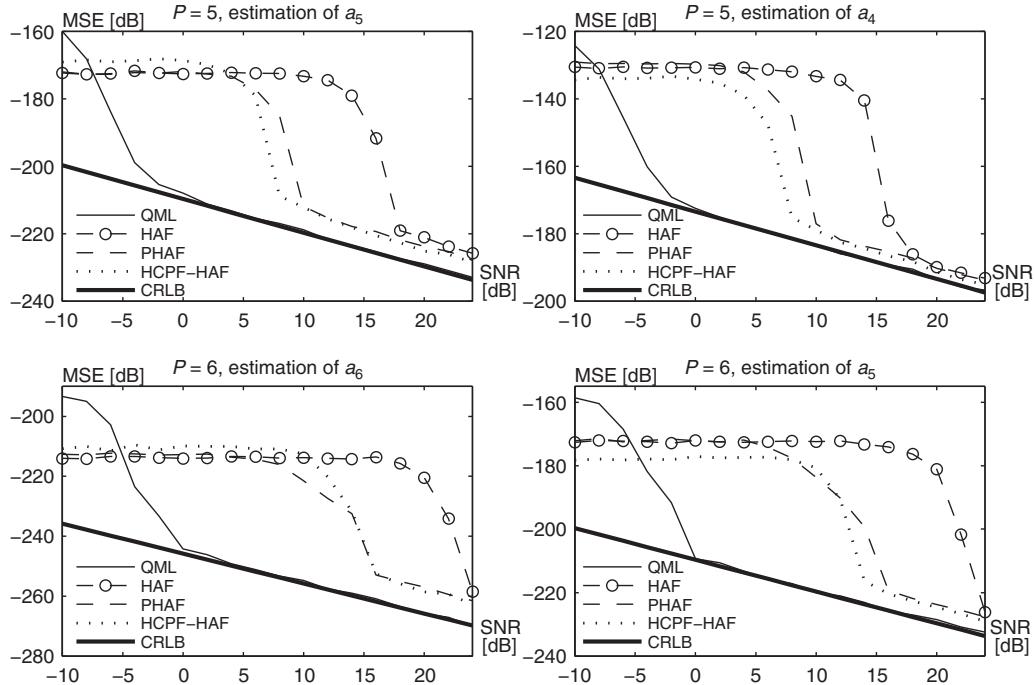
10.7.4.1 Polynomial phase signals

We consider fifth- and sixth-order PPSs, with parameters given in

$$\mathbf{a} = \left[0, 1, -1.9 \times 10^{-4}, 7.3 \times 10^{-6}, 4.1 \times 10^{-8}, -3.1 \times 10^{-10}, 7.3 \times 10^{-12} \right], \quad (10.7.10)$$

according to $a_p = \mathbf{a}(p + 1)$, $p \in [0, P]$. Note that for the fifth-order PPS, the last element in \mathbf{a} is omitted. In addition, $A = 1$, $N = 256$ and $\Delta = 1$. In the QML method, the adopted window set is $\mathbf{H} = [4, 6, 8, \dots, N/2]$.

The QML method is compared with the PD-based estimators, namely the HAF, PHAF and HCPF-HAF, whose setup is given in Ref. [75]. The MSE in the estimation of the highest two-phase parameters, obtained using 300 Monte-Carlo simulations, is plotted in Fig. 10.7.1, along with the corresponding

**FIGURE 10.7.1**

MSE of a_p and a_{p-1} estimates; *Top row*: estimation of a_5 and a_4 for $P = 5$; *Bottom row*: estimation of a_6 and a_5 for $P = 6$.

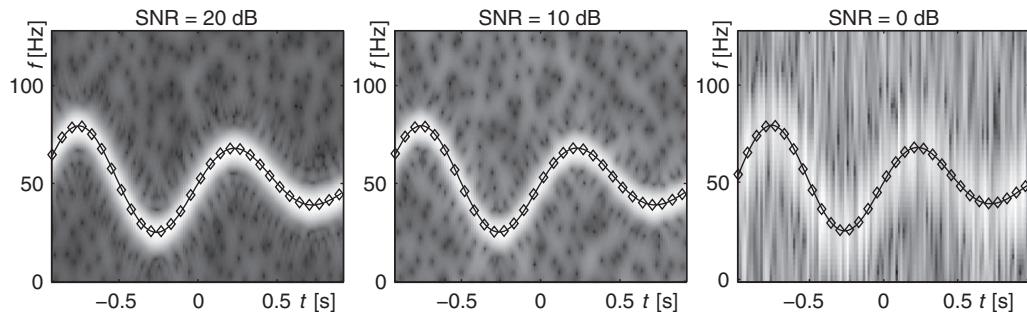
CRLBs. The top row corresponds to the fifth-order PPS, whereas the bottom one corresponds to the sixth-order PPS. The SNR is varied between -10 dB and 24 dB in steps of 2 dB. Clearly, the QML method outperforms all other methods in terms of both accuracy and robustness to noise. The SNR threshold is about 12 and 15 dB lower than the next best method (HCPF-HAF) for $P = 5$ and $P = 6$, respectively. Above the SNR threshold, the QML method practically attains the CRLB, which is not the case for the other considered estimators.

10.7.4.2 Nonpolynomial phase signals

The QML method can be successfully used for the estimation of FM signals with nonpolynomial IF laws. To that end, we consider a constant amplitude signal whose continuous IF is given by

$$f_i(t) = 50 + 30 \sin(2\pi t) e^{-\frac{(t+1)^2}{3}}, \quad (10.7.11)$$

where $t \in [-1, 1]$. The signal is sampled with $\Delta = 1/256$ giving $N + 1 = 513$ samples. In the QML method, the same window set \mathbf{H} as in the previous example is used. In addition, since the order of polynomial that optimally fits the considered IF is not known in advance, an additional search over predefined polynomial orders in the QML algorithm is required. In this example, we consider polynomial orders between 3 and 13 . Figure 10.7.2 depicts one realization of the spectrogram of the

**FIGURE 10.7.2**

Spectrogram of the signal with the IF defined in Eq. (10.7.11). The brightest shades correspond to the highest spectrogram values. The solid black line represents the true IF, whereas black diamonds represent the estimated IF.

considered signal for three SNR values (20 dB, 10 dB, and 0 dB) with solid lines representing the true IF and diamonds the estimated IF. The IF estimation is obtained from PPS estimates output by the QML method. High robustness to noise is evident from the rightmost figure (SNR = 0 dB), where the QML managed to accurately estimate the IF from the spectrogram of a signal heavily corrupted by noise.

10.7.5 SUMMARY AND CONCLUSIONS

Polynomial-phase signal estimators have been reviewed. A selected estimation technique, referred to as QML, based on linear time-frequency representations, has been described. With a proper refinement, this technique can provide a robust and statistically efficient estimation. Moreover, the generalization of the QML to other forms of FM signals produces very accurate IF estimation, even at low SNR. The QML method in its current form cannot be used for the estimation of multidimensional and multicomponent signals. The first step in dealing with multicomponent signals would be to separate components in the (t, f) plane. Due to its poor resolution, the STFT can fail in separating components with close IF trajectories. Also, the refinement strategy would need to be modified. Another issue is the estimation of signals with abrupt changes in phase, when polynomial fitting cannot provide accurate estimation. These issues are also addressed in Ref. [43].

10.8 SEQUENTIAL BAYESIAN ESTIMATION OF INSTANTANEOUS FREQUENCIES⁰

10.8.1 INTRODUCTION

In parallel with the developments of time-frequency techniques, statistical estimation in the sequential Bayesian framework has considerably matured in the last two decades. The consequence is a plethora

⁰Author: **B. Ristic**, RMIT University, Swanston Street, Melbourne, VIC 3000 Australia (branko.ristic@dsto.defence.gov.au).
Reviewers: S. Arulampalam and V. Climente-Alarcon.

of powerful algorithms for estimation of highly nonlinear and/or non-Gaussian stochastic systems. The most prominent is the particle filter: a Monte Carlo estimation technique that sequentially (over time) estimates the posterior probability density function (PDF) of the state vector, thereby providing the full statistical description of the system as it evolves [80–83].

The standard particle filters, however, are limited to a single dynamic system, which is permanently active. In addition, the measurements are assumed to be collected with perfect detection (i.e., no false or miss detections). The recent theoretical advances in sequential Bayesian estimation, carried out in the framework of random-set theory, have overcome these limitations and thus have dramatically widened the scope of applications of particle filters: from single to multiple systems; from permanently active systems to on/off systems; from perfect to possibly imperfect detection [84].

This section describes how the sequential Bayesian estimation framework, and the particle filters in particular, can be used to estimate the instantaneous frequencies (IFs) of a multicomponent frequency-modulated (FM) signal. The problem is not only to “track” the time-varying frequency and amplitude of each component, but also to detect the appearance of new components and possible disappearance of some of the existing components. A number of important papers have been published on this topic (see, e.g., Ref. [85–88] and the references therein). Throughout this section, the input signal for estimation (the observed multicomponent FM signal) is a complex-valued time series $\{y_k\}_{1 \leq k \leq K}$, where $t_k = kT$ is the sampling instance with index $k = 1, 2, \dots, K$.

10.8.2 THE SEQUENTIAL BAYESIAN ESTIMATION FRAMEWORK

We adopt the state-space approach, where the state vector $\mathbf{x}_k \in \mathbb{R}^{n_x}$ fully describes the dynamic system of interest at a discrete-time t_k , $k = 0, 1, 2, \dots, K$. In order to define the problem, we need to specify two mathematical models, the dynamic model and the measurement model, as well as the PDF of the state at the initial time t_0 , that is $p(\mathbf{x}_0)$.

10.8.2.1 Standard formulation

For the case of a single, always active dynamic system (e.g., a single FM component), the dynamics can be specified as follows:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{v}_k, \quad (10.8.1)$$

where \mathbf{f}_k is a known (possibly nonlinear) function of the state, and \mathbf{v}_k is a process noise [81,82]. The dynamic equation (10.8.1) defines the transitional density $\pi(\mathbf{x}_k | \mathbf{x}_{k-1})$.

The system is observed at time instances t_k , and the measurement model is specified as

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \boldsymbol{\epsilon}_k, \quad (10.8.2)$$

where $\mathbf{z}_k \in \mathbb{R}^{n_z}$. The function \mathbf{h}_k is a known (possibly nonlinear) function, while $\boldsymbol{\epsilon}_k$ is the measurement noise. The noise sequences \mathbf{v}_k and $\boldsymbol{\epsilon}_k$ are typically assumed white, mutually independent, and characterized by known PDFs. The measurement equation (10.8.2) defines the likelihood function $g(\mathbf{z}_k | \mathbf{x}_k)$.

Once the two models and the initial PDF $p(\mathbf{x}_0)$ are defined, one can formulate the optimal sequential Bayesian estimator. Suppose the posterior PDF at discrete-time $k - 1$ is known and denoted

$p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$, where $\mathbf{z}_{1:k-1} \equiv \mathbf{z}_1, \dots, \mathbf{z}_{k-1}$. Then the posterior PDF at k can be expressed via the prediction and update equations

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int \pi(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (\text{prediction}) \quad (10.8.3)$$

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{g(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{\int g(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) d\mathbf{x}_k} \quad (\text{update}). \quad (10.8.4)$$

In a general nonlinear/non-Gaussian case, there is no analytic solution to the optimal sequential Bayesian estimator, Eqs. (10.8.3)–(10.8.4). Particle filters are a family of Monte Carlo-based approximate solutions that are popular mainly due to the ever increasing computational power of computers and the ease of their implementation. Detailed discussion regarding the theory and implementation of particle filters is outside the scope of this section and the reader is referred to [80–83] for further study.

Let us illustrate the described framework by an example, originally formulated in [87], where the goal is to estimate the instantaneous frequency of a signal with a nonlinear frequency modulation. The state vector can be adopted as a vector $\mathbf{x}_k = [a_k, \phi_k, f_k]^T$, where a_k, ϕ_k and f_k are the signal amplitude, initial phase and the frequency at time t_k , respectively. The system dynamics can be modeled by a random walk model, i.e., $\mathbf{x}_{k+1} = \mathbf{x}_k + \boldsymbol{\epsilon}_k$, where $\boldsymbol{\epsilon}_k$ is zero-mean white Gaussian noise. Hence, the system dynamic is linear. The observation model can be expressed by Eq. (10.8.2), where $n_z = 2$ with measurement vector components $\mathbf{z}_k[1] = \text{Re}\{y_k\}$, $\mathbf{z}_k[2] = \text{Im}\{y_k\}$, and the measurement function

$$\mathbf{h}_k(\mathbf{x}_k) = \begin{bmatrix} \mathbf{x}_k[1] \cdot \cos(\mathbf{x}_k[2] + 2\pi\mathbf{x}_k[3]t_k) \\ \mathbf{x}_k[1] \cdot \sin(\mathbf{x}_k[2] + 2\pi\mathbf{x}_k[3]t_k) \end{bmatrix}.$$

Here, $\mathbf{x}_k[m]$ denotes the m th component of the vector \mathbf{x}_k , that is $\mathbf{x}_k[1] = a_k$, etc. The initial PDF $p(\mathbf{x}_0)$ can be adopted assuming the independence between a_0, ϕ_0 and f_0 , i.e., with $p(\mathbf{x}_0) = p(a_0)p(\phi_0)p(f_0)$. Then for $p(\phi_0)$, we can choose the uniform distribution $\mathcal{U}(-\pi, \pi)$ and similarly for $p(f_0) = \mathcal{U}(0, f_s/2)$, where f_s is the sampling frequency. The initial PDF for amplitude must be a positive distribution, e.g., the Gamma distribution.

For a slowly varying state \mathbf{x}_k , one can form a measurement vector \mathbf{z}_k from a sliding window $\{y_{k-\ell}, \dots, y_k, \dots, y_{k+\ell}\}$, with $\ell > 0$, rather than just from y_k . This can potentially improve the estimation accuracy at low SNRs.

10.8.2.2 Appearing/disappearing single FM component

In order to model the appearance and disappearance of an FM component during the observation period, it is convenient to introduce a binary random variable $\varepsilon_k \in \{0, 1\}$, referred to as the *existence*. By convention, $\varepsilon_k = 1$ means that the FM component exists at time t_k , and vice versa. The dynamics of ε_k is modeled by the first-order two-state Markov chain with a transitional probability matrix (TPM) Π . Its elements are $\Pi[i,j] = \Pr\{\varepsilon_k = j | \varepsilon_{k-1} = i - 1\}$ for $i, j \in \{1, 2\}$. The TPM is adopted as follows:

$$\Pi = \begin{bmatrix} (1-p_b) & p_b \\ (1-p_s) & p_s \end{bmatrix}, \quad (10.8.5)$$

where p_b and p_s are the probability of “birth” and the probability of “survival,” respectively, during the sampling interval $[t_{k-1}, t_k]$. They are defined as $p_b = \Pr\{\varepsilon_k = 1 | \varepsilon_{k-1} = 0\}$ and $p_s = \Pr\{\varepsilon_k = 1 | \varepsilon_{k-1} = 1\}$. Probabilities p_b , and p_s , together with the initial target existence probability

$q_0 = \Pr\{\varepsilon_0 = 1\}$, are assumed known. If an FM component appears during the sampling interval $(t_{k-1}, t_k]$, its “birth” distribution $b_{k|k-1}(\mathbf{x})$ is essentially the same as the initial distribution $p(\mathbf{x}_0)$ we discussed earlier.

During the period when the FM component is present in the signal, its dynamics are specified by Eq. (10.8.1), and its measurement equation by Eq. (10.8.2). When the FM component is not present in the signal, the measurement equation is simply $\mathbf{z}_k = \boldsymbol{\epsilon}_k$.

The optimal sequential Bayesian estimator for the described problem is known as the Bernoulli filter for the intensity measurement model [89]. This filter propagates two posteriors over time: the posterior probability of FM component existence $q_{k|k} = \Pr\{\varepsilon_k = 1 | \mathbf{z}_{1:k}\}$ and the (usual) posterior PDF of the state $s_{k|k}(\mathbf{x}) = p(\mathbf{x}_k | \mathbf{z}_{1:k})$. The prediction equations are given by [89]

$$q_{k|k-1} = p_b (1 - q_{k-1|k-1}) + p_s q_{k-1|k-1} \quad (10.8.6)$$

$$s_{k|k-1}(\mathbf{x}) = \frac{p_b (1 - q_{k-1|k-1}) b_{k|k-1}(\mathbf{x}) + p_s q_{k-1|k-1} \int \pi_{k|k-1}(\mathbf{x}|\mathbf{x}') s_{k-1|k-1}(\mathbf{x}') d\mathbf{x}'}{p_b (1 - q_{k-1|k-1}) + p_s q_{k-1|k-1}}. \quad (10.8.7)$$

The update equation for the probability of existence is

$$q_{k|k} = \frac{q_{k|k-1} \int \ell_k(\mathbf{z}_k | \mathbf{x}) s_{k|k-1}(\mathbf{x}) d\mathbf{x}}{1 - q_{k|k-1} + q_{k|k-1} \int \ell(\mathbf{z}_k | \mathbf{x}) s_{k|k-1}(\mathbf{x}) d\mathbf{x}}, \quad (10.8.8)$$

where $\ell_k(\mathbf{z}|\mathbf{x})$ is the likelihood ratio

$$\ell(\mathbf{z}|\mathbf{x}) = \frac{g(\mathbf{z}|\mathbf{x})}{g(\mathbf{z}|\emptyset)} \quad (10.8.9)$$

and $g(\mathbf{z}|\emptyset) \equiv p(\boldsymbol{\epsilon})$ is the PDF of the measurement noise. The update equation for the state is

$$s_{k|k}(\mathbf{x}) = \frac{\ell_k(\mathbf{z}_k | \mathbf{x}) s_{k|k-1}(\mathbf{x})}{\int \ell_k(\mathbf{z}_k | \mathbf{x}) s_{k|k-1}(\mathbf{x}) d\mathbf{x}}. \quad (10.8.10)$$

The Bernoulli filter can also be implemented as a particle filter (see [84,89] for details).

10.8.2.3 Appearing/disappearing multiple FM components

Sequential Bayesian estimation of instantaneous frequencies of a multicomponent FM signal can be formulated using the standard framework of Section 10.8.2.1, by introducing a discrete random variable $n_k \in \{0, 1, 2, \dots, n_{\max}\}$ that corresponds to the number of components at time t_k (n_{\max} is a user-defined parameter). The dynamics of n_k can be modeled by the first-order two-state Markov chain with a known TPM Π . For example, in Ref. [86], the transitional probabilities (the elements of the TPM) are adopted as

$$\Pr\{n_k = i | n_{k-1} = j\} = \begin{cases} p_s & i = j, \\ p_t & i = j \pm 1, \\ 0 & \text{otherwise} \end{cases} \quad (10.8.11)$$

with $p_s \gg p_t$ and $p_s + 2p_t = 1$. In this way, the number of components during the sampling interval $(t_{k-1}, t_k]$ can increase or decrease by one. The size of the state vector is therefore dependent on n_k , that is,

$$\mathbf{x}_k = [\mathbf{x}_{k,1}^T \ \mathbf{x}_{k,2}^T \ \cdots \ \mathbf{x}_{k,n_k}^T]^T, \quad (10.8.12)$$

where each component state vector can be chosen as before; e.g., $\mathbf{x}_{k,j} = [a_{kj}, \phi_{kj}, f_{kj}]^T$, for $j = 1, \dots, n_k$. The dynamics of each component can be specified by Eq. (10.8.1), while the measurement equation is given by Eq. (10.8.2). If the measurement is adopted as $\mathbf{z}_k[1] = \text{Re}\{y_k\}$, $\mathbf{z}_k[2] = \text{Im}\{y_k\}$, then

$$\mathbf{h}_k(\mathbf{x}_k) = \begin{bmatrix} \sum_{j=1}^{n_k} \mathbf{x}_{k,j}[1] \cdot \cos(\mathbf{x}_{k,j}[2] + 2\pi \mathbf{x}_{k,j}[3] t_k) \\ \sum_{j=1}^{n_k} \mathbf{x}_{k,j}[1] \cdot \sin(\mathbf{x}_{k,j}[2] + 2\pi \mathbf{x}_{k,j}[3] t_k) \end{bmatrix}.$$

Dubois and Davy [86] define the measurement as a vector within a sliding window of length $L > 1$ (thus assuming the signal is stationary within the window).

This approach is powerful, but has two drawbacks. First, the IF components are not tracked (as labeled sequences). Instead, at each time instant, a vector \mathbf{x}_k of Eq. (10.8.12) is estimated, but there is no link (association) between the components in \mathbf{x}_{k-1} with those in \mathbf{x}_k , for $k = 0, 1, 2, \dots$. Second, the computational complexity grows exponentially with the number of IF components.

The alternative is to cast the problem in the random-set Bayesian estimation framework and solve it using some of the principled approximations, such as the probability hypothesis density (PHD) filter or the multi-Bernoulli filter, assuming the intensity-measurement model (also known as *track-before-detect* model) [90].

10.8.2.4 Spectrogram-based IF estimation

Using the spectrogram or one of its variants to create the measurements for IF estimation is a common practice in time-frequency analysis [91]. For FM signals at low SNR, an approach to generate such measurements would be to find at every time t_k all local maxima (peaks) of the spectrogram. While most of these peaks (detections) would be due to narrowband FM components, some of them could also be due to noise, and therefore false. It is also possible that on some occasions, the low-SNR components are not detected. A collection of peaks from the spectrogram can be modeled as a random finite set. The optimal Bayesian estimator using this kind of measurements for a time-varying number of on/off dynamic stochastic systems has been formulated in the framework of random-set theory [92]. Being computationally intensive, the optimal Bayesian solution has been approximated in a principled manner, resulting in filters such as the PHD filter and the multi-Bernoulli filter for the “detector-output” measurement model [84]. Many standard and more recent multitarget tracking techniques are also directly applicable (see Ref. [93] for more details).

10.8.3 SUMMARY AND CONCLUSIONS

By defining an appropriate dynamic and measurement model, the problem of IF estimation of a noisy multicomponent FM signal can be cast in the sequential Bayesian estimation framework. This framework provides many powerful algorithms for estimation and tracking of highly nonlinear and/or non-Gaussian stochastic dynamic systems. While some of these algorithms have been already explored in the context of IF estimation, there is a large potential for future innovations in the time-frequency domain.

10.9 INSTANTANEOUS FREQUENCY ESTIMATION USING THE VITERBI ALGORITHM

For mono-component signals, the IF at each time instant can be estimated by detecting the locations of the peaks of time-frequency distributions (TFDs) $\rho(t,f)$ along the frequency axis. At low SNR, noise may cause the maxima of a TFD to shift outside the support region of signal auto-terms. In order to reduce such errors, the IF estimation algorithm can be extended to take into account the following criteria [94]:

1. The IF curve should have minimum variations along the selected interval. That is it should minimize $\sum_{n=n_1}^{n_2} |\hat{f}_i[n] - \hat{f}_i[n+1]|$, where $\hat{f}_i[n]$ is the estimated IF of the signal, n_1 and n_2 are starting and ending times of the given interval.
2. Given a TFD $\rho(t,f)$ and its discrete-time version $\rho(n,k)$, the IF curve should maximize the number of high energy (t,f) points. That is the curve should maximize $\sum_{n=n_1}^{n_2} g(\rho(n, \hat{f}_i[n]))$, where $g(x)$ is a non-decreasing function of $x = \rho(n, \hat{f}_i[n])$.

The direct search of the optimal curve among all the possible curves is computationally not feasible. The aforementioned criterion can be implemented recursively using the generalized Viterbi algorithm [95].

REFERENCES

- [1] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals”, *Proc. IEEE* 80 (4) (1992) 520-538.
- [2] B. Ferguson, “A ground based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller driven aircraft”, *J. Acoust. Soc. Am.* 92 (3) (1992) 1403-1407.
- [3] M.T. Taner, F. Koehler, R.E. Sheriff, “Complex seismic trace analysis”, *Geophysics* 44 (6) (1979) 1041-1063.
- [4] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 540-568.
- [5] G. Azemi, B. Senadji, B. Boashash, “Mobile unit velocity estimation based on the instantaneous frequency of the received signal”, *IEEE Trans. Veh. Technol.* 53 (3) (2004) 716-724, ISSN 0018-9545, <http://dx.doi.org/10.1109/TVT.2004.827157>.
- [6] M. Mesbah, J.M. O’Toole, P.B. Colditz, B. Boashash, “Instantaneous frequency based newborn EEG seizures characterization”, *EURASIP J. Adv. Signal Process.* 2012 (143) (2012) 22, <http://dx.doi.org/10.1186/1687-6180-2012-143>.
- [7] J. Ville, “Théorie et applications de la notion de signal analytique”, *Cables et Transmissions* 2A (1) (1948) 61-74 (in French). English translation: I. Selin, Theory and applications of the notion of complex signal, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).
- [8] L.B. White, B. Boashash, “A design methodology for the optimal estimation of the spectral characteristics of non-stationary random signals”, in: Proceedings of the Fourth ASSP Workshop on Spectrum Estimation and Modeling, 1988, pp. 65-70.
- [9] D.L. Jones, T.W. Parks, “A high-resolution data-adaptive time-frequency representation”, *IEEE Trans. Acoust. Speech Signal Process.* 38 (12) (1990) 2127-2135.
- [10] L. Cohen, “Distributions concentrated along the instantaneous frequency”, in: Proceedings of the SPIE: Advanced Signal-Processing Algorithms, Architectures, and Implementations, vol. 1348, Soc. of Photo-optical Instrumentation Engineers, 1990, pp. 149-157.

- [11] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, Z. He, “Matching demodulation transform and synchrosqueezing in time-frequency analysis”, *IEEE Trans. Signal Process.* 62 (1) (2014) 69-84.
- [12] M.K. Emresoy, A. El-Jaroudi, “Iterative instantaneous frequency estimation and adaptive matched spectrogram”, *Signal Process.* 64 (2) (1998) 157-165.
- [13] D.L. Jones, R.G. Baraniuk, “A simple scheme for adapting time-frequency representations”, *IEEE Trans. Signal Process.* 42 (12) (1994) 3530-3535.
- [14] L. Cohen, “Introduction: a primer on time-frequency analysis”, in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 3-42 (Chapter 1).
- [15] A. Goldenshluger, A. Nemirovski, “Adaptive denoising of signals satisfying differential inequalities”, *IEEE Trans. Inform. Theory* 43 (1997) 873-889.
- [16] V. Katkovnik, L. Stanković, “Periodogram with varying and data-driven window length”, *Signal Process.* 67 (3) (1998) 345-358.
- [17] V. Katkovnik, L. Stanković, “Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length”, *IEEE Trans. Signal Process.* 46 (9) (1998) 2315-2325.
- [18] L. Stanković, V. Katkovnik, “Instantaneous frequency estimation using the higher order L-Wigner distributions with the data driven order and window length”, *IEEE Trans. Inform. Theory* 46 (1) (2000) 302-311.
- [19] L. Stanković, V. Katkovnik, “The Wigner distribution of noisy signals with adaptive time-frequency varying window”, *IEEE Trans. Signal Process.* 47 (4) (1999) 1099-1108.
- [20] L. Stanković, “Performance analysis of the adaptive algorithm for bias-to-variance trade-off”, *IEEE Trans. Signal Process.* 52 (5) (2004) 1228-1234.
- [21] A.B. Gershman, L. Stanković, V. Katkovnik, “Sensor array signal tracking using a data-driven window approach”, *Signal Process.* 80 (12) (2000) 2507-2515.
- [22] E. Sejdíć, L. Stanković, M. Daković, J. Jiang, “Instantaneous frequency estimation using the S-transform”, *IEEE Signal Process. Lett.* 15 (2008) 309-312.
- [23] L. Stanković, M. Daković, T. Thayaparan, *Time-Frequency Signal Analysis with Applications*, Artech House, Boston, 2013.
- [24] V. Sucic, J. Lerga, M. Vrankić, “Adaptive filter support selection for signal denoising based on the improved ICI rule”, *Digital Signal Process.* 23 (1) (2013) 65-74.
- [25] J. Lerga, V. Sucic, “Nonlinear IF estimation based on the pseudo WVD adapted using the improved sliding pairwise ICI rule”, *IEEE Signal Process. Lett.* 16 (11) (2009) 953-956.
- [26] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [27] Z.M. Hussain, B. Boashash, “Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions”, *IEEE Trans. Signal Process.* 50 (8) (2002) 1866-1876.
- [28] Z.M. Hussain, B. Boashash, “The T-class of time-frequency distributions: time-only kernels with amplitude estimation”, *J. Franklin Inst.* 343 (7) (2006) 661-675.
- [29] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 519-568.
- [30] B. Boashash, P.J. O’Shea, “Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency”, *IEEE Trans. Signal Process.* 41 (3) (1993) 1439-1445.
- [31] B. Barkat, B. Boashash, “Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: statistical performance in the presence of additive Gaussian noise”, *IEEE Trans. Signal Process.* 47 (9) (1999) 2480-2490.
- [32] J. Jeong, W.J. Williams, “Kernel design for reduced interference distributions”, *IEEE Trans. Signal Process.* 40 (2) (1992) 402-412.
- [33] Z.M. Hussain, B. Boashash, “Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals”, in: Proceedings of the Sixth International Symposium on Signal Processing and its Applications (ISSPA’01), vol. 1, 2001, pp. 339-342.

- [34] B. Barkat, B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis", *IEEE Trans. Signal Process.* 49 (10) (2001) 2232-2239.
- [35] S.C. Sekhar, T.V. Sreenivas, "Signal-to-noise ratio estimation using higher-order moments", *Signal Process.* 86 (4) (2006) 716-732, ISSN 0165-1684, <http://dx.doi.org/10.1016/j.sigpro.2005.06.003>, URL <http://www.sciencedirect.com/science/article/pii/S0165168405001866>.
- [36] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [37] M.B. Malarvili, M. Mesbah, B. Boashash, "Time-frequency analysis of heart rate variability for neonatal seizure detection", *EURASIP J. Appl. Signal Process.* 2007 (1) (2007) 103-103, ISSN 1110-8657, <http://dx.doi.org/10.1155/2007/50396>.
- [38] M.B. Malarvili, M. Mesbah, B. Boashash, "Time-frequency analysis of heart rate variability for neonatal seizure detection", *Australas. Phys. Eng. Sci. Med.* 29 (1) (2006) 76-72.
- [39] P.J. O'Shea, B. Barkat, "Analysis of polynomial FM signals in additive noise", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, first ed., Elsevier, Oxford, 2003, pp. 447-456 (Chapter 10, Article 10.4).
- [40] D.C. Rife, R.R. Boorstyn, "Single-tone parameter estimation from discrete-time observations", *IEEE Trans. Inform. Theory* 20 (5) (1974) 591-598.
- [41] B. Boashash, P.J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.* 42 (1) (1994) 216-220.
- [42] B. Barkat, B. Boashash, "Design of higher order polynomial Wigner-Ville distributions", *IEEE Trans. Signal Process.* 47 (9) (1999) 2608-2611.
- [43] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, M. Daković, "Instantaneous frequency in time-frequency analysis: enhanced concepts and performance of estimation algorithms", *Digital Signal Process.* 35 (0) (2014) 1-13, ISSN 1051-2004, <http://dx.doi.org/10.1016/j.dsp.2014.09.008>, URL <http://www.sciencedirect.com/science/article/pii/S1051200414002838>.
- [44] A. Omidvarnia, G. Azemi, J.M. O'Toole, B. Boashash, "Robust estimation of highly-varying nonlinear instantaneous frequency of monocomponent signals using a lower-order complex-time distribution", *Signal Process.* 93 (11) (2013) 3251-3260, ISSN 0165-1684, <http://dx.doi.org/10.1016/j.sigpro.2013.03.041>, URL <http://www.sciencedirect.com/science/article/pii/S0165168413001497>.
- [45] B. Ristic, B. Boashash, "Instantaneous frequency estimation of quadratic and cubic polynomial FM signals using the cross polynomial Wigner-Ville distribution", *IEEE Trans. Signal Process.* 44 (1996) 1549-1553.
- [46] S. Peleg, B. Porat, "Linear FM signal parameter estimation from discrete-time observations", *IEEE Trans. Aerosp. Electron. Syst.* 27 (4) (1991) 607-616.
- [47] G. Reina, B. Porat, "Comparative performance analysis of two algorithms for instantaneous frequency estimation", in: Proceedings of the Eighth IEEE Workshop on Statistical Signal and Array Processing (SSAP-96), 1996, pp. 448-451.
- [48] S.A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression", *IEEE Trans. Inform. Theory* 31 (6) (1985) 832-835.
- [49] D.C. Reid, A.M. Zoubir, B. Boashash, "Aircraft flight parameter estimation based on passive acoustic techniques using the polynomial Wigner-Ville distribution", *J. Acoust. Soc. Am.* 102 (1) (1997) 207-23.
- [50] S. Peleg, B. Porat, "Estimation and classification of polynomial-phase signals", *IEEE Trans. Inform. Theory* 37 (2) (1991) 422-430.
- [51] B. Porat, B. Friedlander, "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial-phase signal", *IEEE Trans. Inform. Theory* 42 (3) (1996) 995-1001.
- [52] M.L. Farquharson, P.J. O'Shea, G.F. Ledwich, "A computationally efficient technique for estimating the parameters of polynomial-phase signals from noisy observations", *IEEE Trans. Signal Process.* 53 (8) (2005) 3337-3342.

- [53] S. Peleg, B. Friedlander, "Multicomponent signal analysis using the polynomial-phase transform", *IEEE Trans. Aerosp. Electron. Syst.* 32 (1) (1996) 378-386.
- [54] B. Boashash, B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise", *Signal Process.* 67 (1) (1998) 1-23.
- [55] R.S. Kennedy, *Fading Dispersive Communication Channels*, Wiley, New York, 1969.
- [56] H.L.L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: Radar-Sonar Signal Processing and Gaussian Signals in Noise, Wiley, New York, 1971, Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.
- [57] A.G. Piersol, "Power spectra measurements for space vibration data", *J. Spacecr. Rockets* 4 (1967) 1613.
- [58] M.R. Morelande, B. Barkat, A.M. Zoubir, "Statistical performance comparison of a parametric and a non-parametric method for IF estimation of random amplitude linear FM signals in additive noise", in: Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000), 2000, pp. 262-266.
- [59] B. Barkat, "Instantaneous frequency estimation of nonlinear frequency-modulated signals in the presence of multiplicative and additive noise", *IEEE Trans. Signal Process.* 49 (10) (2001) 2214-2222.
- [60] M. Colas, G. Gelle, "A multitime-frequency approach for detection and classification of neighboring instantaneous frequency laws in a noisy environment", *IEEE Signal Process. Lett.* 11 (2) (2004) 71-74.
- [61] J. Lerga, V. Sucic, B. Boashash, "An efficient algorithm for instantaneous frequency estimation of nonstationary multicomponent signals in low SNR", *EURASIP J. Adv. Signal Process.* 2011 (2011) <http://dx.doi.org/10.1155/2011/725189>, article ID 725189, 16 pp.
- [62] V. Sucic, J. Lerga, B. Boashash, "Multicomponent noisy signal adaptive instantaneous frequency estimation using components time support information", *IET Signal Process.* 8 (3) (2014) 277-284.
- [63] S. Barbarossa, "Analysis of multicomponent LFM signals by a combined Wigner-Hough transform", *IEEE Trans. Signal Process.* 43 (6) (1995) 1511-1515.
- [64] R.M. Rangayyan, S. Krishnan, "Feature identification in the time-frequency plane by using the Hough-Radon transform", *Pattern Recogn.* 34 (6) (2001) 1147-1158.
- [65] N.N. Bennett, N. Saito, "Using edge information in time-frequency representations for chirp parameter estimation", *Appl. Comput. Harmon. Anal.* 18 (2) (2005) 186-197.
- [66] R. Nevatia, K.R. Babu, "Linear feature extraction and description", *Comput. Graph. Image Process.* 13 (3) (1980) 257-269.
- [67] A.A. Farag, E.J. Delp, "Edge linking by sequential search", *Pattern Recogn.* 28 (5) (1995) 611-633.
- [68] L.J. Rankine, M. Mesbah, B. Boashash, "IF estimation for multicomponent signals using image processing techniques in the time-frequency domain", *Signal Process.* 87 (6) (2007) 1234-1250.
- [69] V. Sucic, N. Saulig, B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy", *EURASIP J. Adv. Signal Process.* 2011 (1) (2011) <http://dx.doi.org/10.1186/1687-6180-2011-125>, article ID 2011:125, 11 pp.
- [70] B. Boashash, G. Azemi, N.A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: applications to newborn EEG abnormality detection", *Pattern Recogn.* 48 (3) (2015) 616-627, <http://dx.doi.org/10.1016/j.patcog.2014.08.016>.
- [71] J. Ångeby, "Aliasing of polynomial-phase signal parameters", *IEEE Trans. Signal Process.* 48 (5) (2000) 1488-1491.
- [72] R.G. McKilliam, B.G. Quinn, I.V.L. Clarkson, B. Moran, B.N. Vellambi, "Polynomial phase estimation by least squares phase unwrapping", *IEEE Trans. Signal Process.* 62 (8) (2014) 1962-1975.
- [73] S. Barbarossa, A. Scaglione, G.B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling", *IEEE Trans. Signal Process.* 46 (3) (1998) 691-708.
- [74] P. O'Shea, "A fast algorithm for estimating the parameters of a quadratic FM signal", *IEEE Trans. Signal Process.* 52 (2) (2004) 385-393.

- [75] I. Djurović, M. Simeunović, S. Djukanović, P. Wang, "A hybrid CPF-HAF estimation of polynomial-phase signals: detailed statistical analysis", *IEEE Trans. Signal Process.* 60 (10) (2012) 5010-5023.
- [76] L. Stanković, "A multitime definition of the Wigner higher order distribution: L-Wigner distribution", *IEEE Signal Process. Lett.* 1 (7) (1994) 106-109.
- [77] L. Stanković, "Analysis of noise in time-frequency distributions", *IEEE Signal Process. Lett.* 9 (9) (2002) 286-289.
- [78] I. Djurović, L. Stanković, "Quasi-maximum-likelihood estimator of polynomial phase signals", *IET Signal Process.* (2013) <http://dx.doi.org/10.1049/iet-spr.2013.0104>, 13 pp.
- [79] P. O'Shea, "On refining polynomial phase signal parameter estimates", *IEEE Trans. Aerosp. Electron. Syst.* 46 (3) (2010) 978-987.
- [80] A. Doucet, N. de Freitas, N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001, <http://dx.doi.org/10.1007/978-1-4757-3437-9>.
- [81] M. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking", *IEEE Trans. Signal Process.* 50 (2) (2002) 174-188, ISSN 1053-587X, <http://dx.doi.org/10.1109/78.978374>.
- [82] B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, Norwood, MA, 2004.
- [83] O. Cappé, S. Godsill, E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo", *Proc. IEEE* 95 (5) (2007) 899-924, ISSN 0018-9219, <http://dx.doi.org/10.1109/JPROC.2007.893250>.
- [84] B. Ristic, *Particle Filters for Random Set Models*, vol. 20013, Springer, New York, 2013.
- [85] C. Andrieu, M. Davy, A. Doucet, "Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions", *IEEE Trans. Signal Process.* 51 (7) (2003) 1762-1770, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2003.810284>.
- [86] C. Dubois, M. Davy, "Joint detection and tracking of time-varying harmonic components: a flexible Bayesian approach", *IEEE Trans. Audio Speech Lang. Process.* 15 (4) (2007) 1283-1295, ISSN 1558-7916, <http://dx.doi.org/10.1109/TASL.2007.894522>.
- [87] E. Tsakonas, N. Sidiropoulos, A. Swami, "Optimal particle filters for tracking a time-varying harmonic or chirp signal", *IEEE Trans. Signal Process.* 56 (10) (2008) 4598-4610, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2008.927462>.
- [88] V. Clemente-Alarcon, J. Antonino-Daviu, A. Haavisto, A. Arkkio, "Particle filter-based estimation of instantaneous frequency for the diagnosis of electrical asymmetries in induction machines", *IEEE Trans. Instrum. Meas.* 63 (10) (2014) 2454-2463, ISSN 0018-9456, <http://dx.doi.org/10.1109/TIM.2014.2310113>.
- [89] B. Ristic, B.-T. Vo, B.-N. Vo, A. Farina, "A tutorial on Bernoulli filters: theory, implementation and applications", *IEEE Trans. Signal Process.* 61 (13) (2013) 3406-3430, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2013.2257765>.
- [90] B.-N. Vo, B.-T. Vo, N.-T. Pham, D. Suter, "Joint detection and estimation of multiple objects from image observations", *IEEE Trans. Signal Process.* 58 (10) (2010) 5129-5141, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2010.2050482>.
- [91] N.A. Khan, B. Boashash, "Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram", *IEEE Signal Process. Lett.* 20 (2) (2013) 157-160.
- [92] R.P. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, Inc., Norwood, MA, 2007.
- [93] S. Challa, M.R. Morelande, D. Mušicki, R.J. Evans, *Fundamentals of Object Tracking*, Cambridge University Press, Cambridge, 2011.
- [94] I. Djurović, L. Stanković, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment", *Signal Process.* 84 (2004) 631-643.
- [95] G. D. Forney Jr., "The Viterbi algorithm", *Proc. IEEE* 61 (1973) 268-278.

This page intentionally left blank

TIME-FREQUENCY SYNTHESIS AND FILTERING

11

INTRODUCTION AND OVERVIEW

This chapter presents methods and techniques to design time-varying linear systems such as filters with precise time-frequency (t,f) specifications; this capability can then allow one to model and predict accurately the effects of linear systems on nonstationary signals in applications such as signal cleansing and enhancement. This topic is covered in six sections with appropriate cross-referencing to other chapters.

The design of time-varying filters is useful in applications where it is desired to separate, suppress or reduce undesirable nonstationary signal components. This can be done with a number of methods such as the STFT and Gabor transform presented in [Section 11.1](#). In particular, the use of the Gabor expansion for time-varying filtering is illustrated on an application that involves monitoring machine vibrations ([Section 11.2](#)). Another illustration of the procedure for designing a time-varying filter is provided in the context of an application involving hands-free telephone speech signals ([Section 11.3](#)). Another important application of time-varying filtering, namely signal enhancement, is described using an iterative algorithm based on time-frequency peak filtering ([Section 11.4](#)). Then, a method for subspace noise filtering using a time-frequency distribution is described ([Section 11.5](#)); and finally a comparison of denoising algorithms for speech enhancement completes the chapter ([Section 11.6](#)).

11.1 LINEAR TIME-FREQUENCY FILTERS⁰

11.1.1 TIME-FREQUENCY DESIGN OF LINEAR, TIME-VARYING FILTERS

Linear time-varying (LTV) filters are useful in many applications, especially for weighting, suppressing, or separating nonstationary signal components. The input-output relation of an LTV filter \mathbf{H} with kernel (impulse response) $h(t, t')$ reads

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t, t') x(t') dt'. \quad (11.1.1)$$

The nonstationary nature of $x(t)$, $y(t)$, and \mathbf{H} justifies the use of time-frequency (briefly (t, f)) representations. There are two fundamentally different approaches to a (t, f) design of LTV filters, namely, the “explicit” and “implicit” design philosophies [1]. Both are based on a prescribed (t, f) weight function $M(t, f)$ that provides a (t, f) specification of the desired filtering characteristic.

- **Explicit design:** The impulse response $h(t, t')$ of the LTV filter \mathbf{H} is designed such that a (t, f) representation of \mathbf{H} equals or best approximates the (t, f) weight function $M(t, f)$. Let us choose as (t, f) representation of \mathbf{H} the *generalized Weyl symbol* (GWS) (see Section 4.7). The filtering itself may be performed in the time domain according to Eq. (11.1.1).
- **Implicit design:** The LTV filter \mathbf{H} is designed *implicitly* during the filtering, using a three-step analysis-weighting-synthesis procedure. First (analysis step), a *linear* (t, f) representation—such as the short time Fourier transform (STFT)—of the input signal $x(t)$ is calculated. Second (weighting step), this (t, f) representation is multiplied by the (t, f) weight function $M(t, f)$. Third (synthesis step), the output signal $y(t)$ is calculated in a linear manner from the (t, f) function obtained in Step 2. Since all processing steps are linear, the overall procedure amounts to an LTV filter.

This section describes explicit (t, f) filter designs based on the generalized Weyl symbol [1–3] and implicit (t, f) filter designs based on the STFT [1,4–7] and the Gabor transform [1,8–11]. We will show that the resulting filters tend to perform similarly if the (t, f) weight function $M(t, f)$ is sufficiently smooth.

11.1.2 EXPLICIT DESIGN—THE GENERALIZED WEYL FILTER

Section 4.7 defines the *generalized Weyl symbol* (GWS) of an LTV system \mathbf{H} as

$$L_{\mathbf{H}}^{(\alpha)}(t, f) \triangleq \int_{-\infty}^{\infty} h\left(t + \left(\frac{1}{2} - \alpha\right)\tau, t - \left(\frac{1}{2} + \alpha\right)\tau\right) e^{-j2\pi f\tau} d\tau, \quad (11.1.2)$$

where α is a real-valued parameter (in particular, $\alpha = 0$ and $\alpha = 1/2$ give the *Weyl symbol* and *Zadeh’s time-varying transfer function*, respectively). For *underspread* LTV systems (which produce only moderate (t, f) displacements), the GWS can be interpreted as a (t, f) transfer function describing the (t, f) weighting produced by the system (see Section 4.7). Hence, a conceptually simple (t, f) design of an LTV filter \mathbf{H} from a prescribed (t, f) weight function $M(t, f)$ is based on setting the filter’s GWS equal to $M(t, f)$ [1,3],

⁰Authors: **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at) and **G. Matz**, Vienna University of Technology, Vienna, Austria (g.matz@ieee.org). Reviewers: M.G. Amin and D.L. Jones.

$$L_{\mathbf{H}_{\text{GWF}}}^{(\alpha)}(t,f) \equiv M(t,f). \quad (11.1.3)$$

The impulse response of the LTV filter \mathbf{H}_{GWF} thus defined (called *generalized Weyl filter*) is obtained via the inverse of Eq. (11.1.2), i.e.,

$$h_{\text{GWF}}(t,t') = \int_{-\infty}^{\infty} M\left(\left(\frac{1}{2} + \alpha\right)t + \left(\frac{1}{2} - \alpha\right)t', f\right) e^{j2\pi f(t-t')} df. \quad (11.1.4)$$

The choices $\alpha = 0$ or $\alpha = 1/2$ lead to the *Weyl filter* or *Zadeh filter*. For $\alpha = 0$, a real-valued weight function $M(t,f)$ results in a self-adjoint Weyl filter.

11.1.2.1 Dependence on α

The dependence of \mathbf{H}_{GWF} on α effectively disappears in the case of a *smooth* (t,f) weight function $M(t,f)$ (yielding an *underspread* LTV system \mathbf{H}_{GWF} as discussed in Section 4.7; also note that smoothness of $M(t,f)$ is incompatible with a sharp (t,f) cutoff). Let $\mathbf{H}_{\text{GWF}}^{(1)}$ and $\mathbf{H}_{\text{GWF}}^{(2)}$ be two generalized Weyl filters designed according to Eq. (11.1.4) with GWS parameter α_1 and α_2 , respectively. Then, one can show that the difference $(\mathbf{H}_{\text{GWF}}^{(1)}x)(t) - (\mathbf{H}_{\text{GWF}}^{(2)}x)(t)$ of the output signals of these filters satisfies

$$\frac{\|\mathbf{H}_{\text{GWF}}^{(1)}x - \mathbf{H}_{\text{GWF}}^{(2)}x\|_2}{\|x\|_2} \leq \epsilon_1 \triangleq |\alpha_1 - \alpha_2| \left\| \frac{\partial^2 M}{\partial t \partial f} \right\|_2,$$

where $\|\cdot\|_2$ denotes the L_2 norm. Thus, the generalized Weyl filter design is almost independent of α if ϵ_1 is small, i.e., if $M(t,f)$ is a smooth function.

11.1.2.2 (t,f) projection filter

Formally, Eq. (11.1.3) can be viewed as the solution to the (unconstrained) minimization problem $\mathbf{H}_{\text{GWF}} = \arg \min_{\mathbf{H}} \|M - L_{\mathbf{H}}^{(\alpha)}\|_2$. Solving this minimization problem under the side constraint that \mathbf{H} be an orthogonal projection operator yields the (t,f) *projection filter* \mathbf{H}_P introduced in [2]. An orthogonal projection operator is characterized by being self-adjoint ($\mathbf{H}^\dagger = \mathbf{H}$) and idempotent ($\mathbf{H}\mathbf{H} = \mathbf{H}$). More specifically, let $u_k(t)$ and λ_k denote the eigenfunctions and eigenvalues, respectively, of $(\mathbf{H}_{\text{GWF}} + \mathbf{H}_{\text{GWF}}^\dagger)/2$, where $\mathbf{H}_{\text{GWF}}^\dagger$ denotes the adjoint of \mathbf{H}_{GWF} . Then, the impulse response of \mathbf{H}_P can be shown [2] to equal $h_P(t,t') = \sum_{k \in \mathcal{I}} u_k(t)u_k^*(t')$, where \mathcal{I} is the set of indices k for which $\lambda_k > 1/2$. The (t,f) projection filter is only able to pass or suppress signal components, with no other weights possible. It is advantageous in some situations since it is capable of realizing very sharp (t,f) cutoff characteristics. However, compared to \mathbf{H}_{GWF} , the calculation of \mathbf{H}_P requires the additional solution of an eigenproblem. (An efficient online implementation of the (t,f) projection filter is proposed in [12].)

11.1.3 IMPLICIT DESIGN I—THE STFT FILTER

An *STFT filter* consists of the following three steps [1,4,5,7]:

- *Analysis:* Calculation of the STFT [7] of the input signal $x(t)$, $F_x^\gamma(t,f) = \int_{-\infty}^{\infty} x(t')\gamma_{t,f}^*(t') dt'$, where $\gamma_{t,f}(t') = \gamma(t' - t)e^{j2\pi ft'}$ with $\gamma(t)$ being an analysis window (see Section 2.3.1 for details).
- *Weighting:* Multiplication of the STFT by the (t,f) weight function $M(t,f)$, i.e., calculation of $M(t,f)F_x^\gamma(t,f)$.

- Synthesis:* The output signal $y(t)$ is obtained via an inverse STFT [7],
 $y(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [M(t', f') F_x^{\gamma}(t', f')] g_{t,f}(t') dt' df'$. Here, $g_{t,f}(t') = g(t' - t) e^{j2\pi f t'}$, where $g(t)$ is a synthesis window that is assumed to satisfy $\int_{-\infty}^{\infty} g(t) \gamma^*(t) dt = 1$ (for perfect reconstruction when $M(t, f) \equiv 1$).

These steps implement an LTV filter—hereafter denoted $\mathbf{H}_{\gamma,g}$ —that depends on the (t,f) weight function $M(t,f)$ and the windows $\gamma(t)$ and $g(t)$.

11.1.3.1 Multiwindow STFT filter

An extension of the STFT filter $\mathbf{H}_{\gamma,g}$ is the *multiwindow STFT filter* [1,6] $\mathbf{H}_N \triangleq \sum_{i=1}^N \eta_i \mathbf{H}_{\gamma^{(i)}, g^{(i)}}$, with $\sum_{i=1}^N \eta_i = 1$, $\eta_i \in \mathbb{R}$. This is a linear combination of N STFT filters $\mathbf{H}_{\gamma^{(i)}, g^{(i)}}$ with the *same* (t,f) weight function $M(t,f)$ but different analysis windows $\gamma^{(i)}(t)$ and different synthesis windows $g^{(i)}(t)$. Note that the STFT filter $\mathbf{H}_{\gamma,g}$ is a special case with $N = 1$. Using a larger number N of STFT filters yields increased flexibility of design at the expense of increased computational complexity. The impulse response of the multiwindow STFT filter \mathbf{H}_N can be expressed as

$$h_N(t, \tilde{t}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(t', f') p(t - t', \tilde{t} - t') e^{j2\pi f'(t - \tilde{t})} dt' df',$$

with $p(t, t') = \sum_{i=1}^N \eta_i g^{(i)}(t) \gamma^{(i)*}(t')$. Furthermore, the GWS of \mathbf{H}_N is obtained as

$$L_{\mathbf{H}_N}^{(\alpha)}(t, f) = M(t, f) * * L_{\mathbf{P}}^{(\alpha)}(t, f), \quad (11.1.5)$$

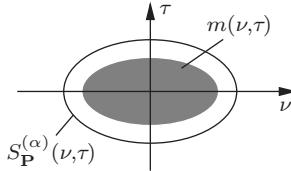
where $* *$ denotes 2D convolution and \mathbf{P} is the LTV system with impulse response $p(t, t')$.

11.1.3.2 Comparison with generalized Weyl filter

Comparing Eq. (11.1.5) with Eq. (11.1.3) shows that the multiwindow STFT filter \mathbf{H}_N using (t,f) weight function $M(t,f)$ is equivalent to a generalized Weyl filter using the modified (t,f) weight function $\tilde{M}(t,f) = M(t,f) * * L_{\mathbf{P}}^{(\alpha)}(t,f)$. For nonnegative coefficients η_i , $\tilde{M}(t,f)$ is a smoothed version of $M(t,f)$. However, for $\eta_i = (-1)^i$ and $N \rightarrow \infty$, it is possible to have $L_{\mathbf{H}_N}^{(\alpha)}(t, f) \rightarrow M(t, f)$ and thus $\mathbf{H}_N \rightarrow \mathbf{H}_{\text{GWF}}$, i.e., the multiwindow STFT filter approaches the generalized Weyl filter using the (t,f) weight function $M(t,f)$. It can be shown that the difference $(\mathbf{H}_N x)(t) - (\mathbf{H}_{\text{GWF}} x)(t)$ of the output signals of \mathbf{H}_N and \mathbf{H}_{GWF} (both based on the same (t,f) weight function $M(t,f)$) satisfies

$$\frac{\|\mathbf{H}_N x - \mathbf{H}_{\text{GWF}} x\|_2}{\|x\|_2} \leq \epsilon_2 \triangleq \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |m(v, \tau)|^2 |1 - S_{\mathbf{P}}^{(\alpha)}(v, \tau)|^2 dv d\tau \right]^{1/2}. \quad (11.1.6)$$

Here, $m(v, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(t, f) e^{-j2\pi(vt - \tau f)} dt df$, and $S_{\mathbf{P}}^{(\alpha)}(v, \tau)$ is the generalized spreading function of \mathbf{P} (see Section 4.7). The constant ϵ_2 is related to the operator \mathbf{P} that characterizes the effect of the windows $\gamma^{(i)}(t), g^{(i)}(t)$; in particular, ϵ_2 is small if the effective support of $S_{\mathbf{P}}^{(\alpha)}(v, \tau)$ covers the effective support of $m(v, \tau)$, so that $|1 - S_{\mathbf{P}}^{(\alpha)}(v, \tau)|^2 \approx 0$ on the support of $m(v, \tau)$ (see Fig. 11.1.1). This is favored by a *smooth* (t,f) weight function $M(t,f)$. Here, $m(v, \tau)$ is well concentrated about $(0,0)$; thus its effective support can easily be covered by $S_{\mathbf{P}}^{(\alpha)}(v, \tau)$, even using a small N . Hence, for a smooth $M(t,f)$, the generalized Weyl filter can be approximated by the multiwindow STFT filter (see Section 11.1.6).

**FIGURE 11.1.1**

Case where the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu, \tau)$ covers the effective support of $m(\nu, \tau)$.

11.1.4 IMPLICIT DESIGN II—THE GABOR FILTER

The Gabor transform is the STFT evaluated on a (t, f) lattice (nT, kF) with $n, k \in \mathbb{Z}$ [9]. A *Gabor filter* (see [1,8,10,11] and [Section 11.2](#)) performs the following steps:

- *Analysis:* Calculation of the Gabor coefficients of the input signal $x(t)$ [9],
 $c_{n,k} = \int_{-\infty}^{\infty} x(t) \gamma_{n,k}^*(t) dt$, where $\gamma_{n,k}(t) = \gamma(t - nT) e^{j2\pi k F t}$ with $\gamma(t)$ being a suitable analysis window.
- *Weighting:* Multiplication of the Gabor coefficients by the weights $M_{n,k} = M(nT, kF)$, i.e., calculation of $M_{n,k} c_{n,k}$.
- *Synthesis:* The output signal $y(t)$ is obtained via Gabor synthesis [9],
 $y(t) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k} c_{n,k} g_{n,k}(t)$, where $g_{n,k}(t) = g(t - nT) e^{j2\pi k F t}$ with $g(t)$ being a suitable synthesis window.

Let us denote this LTV filter by $\tilde{\mathbf{H}}_{\gamma,g}$. The windows $\gamma(t)$ and $g(t)$ are usually assumed to satisfy the perfect-reconstruction (biorthogonality) condition $\int_{-\infty}^{\infty} g(t) \gamma^*(t - \frac{n}{F}) e^{-j2\pi kt/F} dt = \delta_n \delta_k$, which presupposes critical sampling ($TF = 1$) or oversampling ($TF < 1$).

11.1.4.1 Multiwindow Gabor filter

The *multiwindow Gabor filter* is an extension of the Gabor filter $\tilde{\mathbf{H}}_{\gamma,g}$ defined as [1,11] $\tilde{\mathbf{H}}_N \triangleq \sum_{i=1}^N \eta_i \tilde{\mathbf{H}}_{\gamma^{(i)}, g^{(i)}}$, with $\sum_{i=1}^N \eta_i = 1$, $\eta_i \in \mathbb{R}$, i.e., a linear combination of Gabor filters $\tilde{\mathbf{H}}_{\gamma^{(i)}, g^{(i)}}$ with the same (t, f) weights $M_{n,k}$ but different analysis windows $\gamma^{(i)}(t)$ and different synthesis windows $g^{(i)}(t)$. The Gabor filter $\tilde{\mathbf{H}}_{\gamma,g}$ is reobtained with $N = 1$. Using a larger number N of Gabor filters allows a reduction of the (t, f) sampling density TF (cf. [13]). The impulse response of the multiwindow Gabor filter $\tilde{\mathbf{H}}_N$ is given by

$$\tilde{h}_N(t, t') = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k} p(t - nT, t' - nT) e^{j2\pi k F (t - t')},$$

with $p(t, t') = \sum_{i=1}^N \eta_i g^{(i)}(t) \gamma^{(i)*}(t')$, and the GWS of $\tilde{\mathbf{H}}_N$ is

$$L_{\tilde{\mathbf{H}}_N}^{(\alpha)}(t, f) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k} L_{\mathbf{P}}^{(\alpha)}(t - nT, f - kF),$$

where \mathbf{P} is the LTV system with impulse response $p(t, t')$.

11.1.4.2 Comparison with generalized Weyl filter

Next, we analyze how close the multiwindow Gabor filter $\tilde{\mathbf{H}}_N$ (using (t,f) weights $M_{n,k} = M(nT, kF)$) is to the generalized Weyl filter \mathbf{H}_{GWF} (using (t,f) weight function $M(t,f)$). One can show that $\frac{\|\tilde{\mathbf{H}}_{N,x} - \mathbf{H}_{\text{GWF},x}\|_2}{\|x\|_2} \leq \epsilon_2 + \epsilon_3$, where ϵ_2 was given in Eq. (11.1.6) and ϵ_3 is defined as

$$\epsilon_3 \triangleq \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| S_{\mathbf{P}}^{(\alpha)}(\nu, \tau) \right|^2 \left| \sum_{\substack{n \neq 0 \\ k \neq 0}} m\left(\nu - \frac{n}{F}, \tau - \frac{k}{T}\right) \right|^2 d\nu d\tau \right]^{1/2}.$$

As for the STFT filter, the term ϵ_2 is related to the operator \mathbf{P} that describes the effect of the windows $\gamma^{(i)}(t), g^{(i)}(t)$. If $M(t,f)$ is smooth so that $m(\nu, \tau)$ is well concentrated about the origin, a suitable choice of \mathbf{P} allows to cover the effective support of $m(\nu, \tau)$ by the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu, \tau)$, which results in a small value of ϵ_2 (see the discussion in Section 11.1.3). The additional term ϵ_3 is mainly due to potential aliasing errors which are caused by the sampling $M_{n,k} = M(nT, kF)$ that distinguishes the Gabor filter from the STFT filter. For ϵ_3 to be small, it is necessary that the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu, \tau)$ does not overlap with the periodic repetitions $m\left(\nu - \frac{n}{F}, \tau - \frac{k}{T}\right)$ of $m(\nu, \tau)$ (see Fig. 11.1.2). This is achieved if (i) $m(\nu, \tau)$ is well concentrated about $(0,0)$ and thus $m\left(\nu - \frac{n}{F}, \tau - \frac{k}{T}\right)$ is well localized about $(\frac{n}{F}, \frac{k}{T})$ and (ii) T and F are small enough to ensure that the periodic repetitions $m\left(\nu - \frac{n}{F}, \tau - \frac{k}{T}\right)$ are sufficiently separated. For $m(\nu, \tau)$ well concentrated, the latter condition can be met even for $TF > 1$. Hence, for a smooth $M(t,f)$, the generalized Weyl filter can be accurately approximated by the (multiwindow) Gabor filter.

11.1.5 THE DISCRETE-TIME CASE

While the discussion of (t,f) filters has so far been placed in a continuous-time framework, practical implementation of these filters requires a discrete-time formulation. The input-output relation of a *discrete-time LTV system* \mathbf{H} reads $y[n] = (\mathbf{H}x)[n] = \sum_{n'=-\infty}^{\infty} h[n, n']x[n']$, with $h[n, n']$ being the impulse response of \mathbf{H} . The GWS with arbitrary α is not easily reformulated in a discrete-time setting. However, for $\alpha = 0$ (Weyl symbol) and $\alpha = 1/2$ (Zadeh's time-varying transfer function), which are the main cases of practical interest, discrete-time formulations are given by

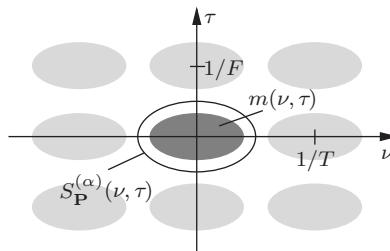


FIGURE 11.1.2

Windowing and aliasing effects involved in Gabor filter design. The dark gray ellipse shows the effective support of $m(\nu, \tau)$, and the light gray ellipses show the effective support of $m\left(\nu - \frac{n}{F}, \tau - \frac{k}{T}\right)$ for $(n, k) \neq (0, 0)$.

$$\begin{aligned} L_{\mathbf{H}}^{(0)}(n, \theta) &= 2 \sum_{m=-\infty}^{\infty} h[n+m, n-m] e^{-j4\pi\theta m}, \\ L_{\mathbf{H}}^{(1/2)}(n, \theta) &= \sum_{m=-\infty}^{\infty} h[n, n-m] e^{-j2\pi\theta m}. \end{aligned}$$

Here, θ denotes normalized frequency. For $L_{\mathbf{H}}^{(0)}(n, \theta)$ to be in one-to-one correspondence to $h[n, n']$, \mathbf{H} has to be a *halfband system*, i.e., an LTV system that accepts input signal components only within a specified halfband (e.g., $\theta \in [-1/4, 1/4]$) and maps them to a halfband output signal [1].

The (t, f) system representations $L_{\mathbf{H}}^{(0)}(n, \theta)$ and $L_{\mathbf{H}}^{(1/2)}(n, \theta)$ can be used to design discrete-time LTV filters from a (t, f) weight function $M(n, \theta)$ via an explicit filter design (cf. Section 11.1.2) [1]. The *discrete-time Zadeh filter* ($\alpha = 1/2$) is defined by setting $L_{\mathbf{H}}^{(1/2)}(n, \theta) = M(n, \theta)$; its impulse response is obtained as

$$h[n, n'] = \int_{-1/2}^{1/2} M(n, \theta) e^{j2\pi\theta(n-n')} d\theta.$$

In a similar manner, the *discrete-time Weyl filter* ($\alpha = 0$) is obtained as

$$h[n+m, n-m] = \int_{-1/4}^{1/4} M(n, \theta) e^{j2\pi\theta m} d\theta. \quad (11.1.7)$$

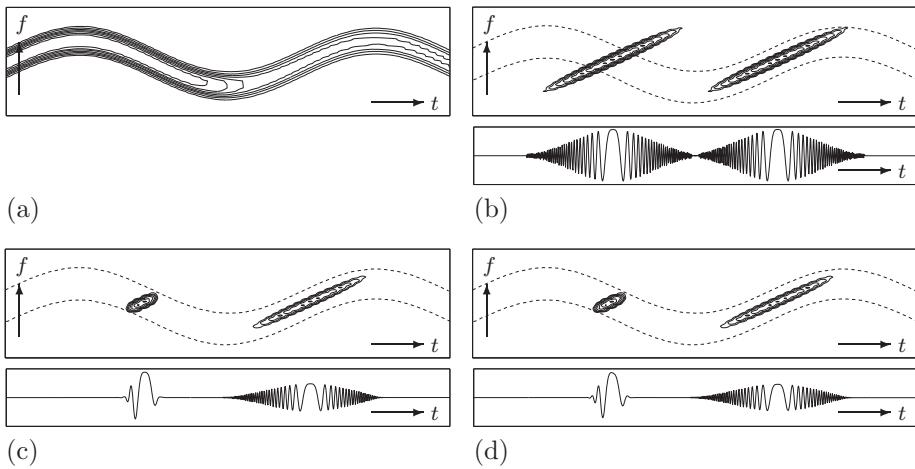
Since $L_{\mathbf{H}}^{(0)}(n, \theta)$ is meaningful only for halfband systems, $M(n, \theta)$ here is specified on the halfband $[-1/4, 1/4]$. According to Eq. (11.1.7), the impulse response $h[n_1, n_2]$ of the discrete-time Weyl filter is obtained only for $n_1 + n_2$ even (since $n_1 + n_2 = n + m + n - m = 2n$). For a halfband system \mathbf{H} , $h[n_1, n_2]$ is completely specified by these samples. The missing samples (for $n_1 + n_2$ odd) could be obtained by interpolation, but this is not necessary since the entire filtering can be performed using only the even-indexed samples [1]. In some cases (e.g., for “chirpy” (t, f) weight functions), the Weyl filter design results in better filtering performance than the Zadeh filter design [1].

Discrete-time versions of the implicit filter design methods discussed in Sections 11.1.3 and 11.1.4 can be obtained in a straightforward manner; see [1,4,7] for discrete-time STFT filters and [1,8] for discrete-time Gabor filters.

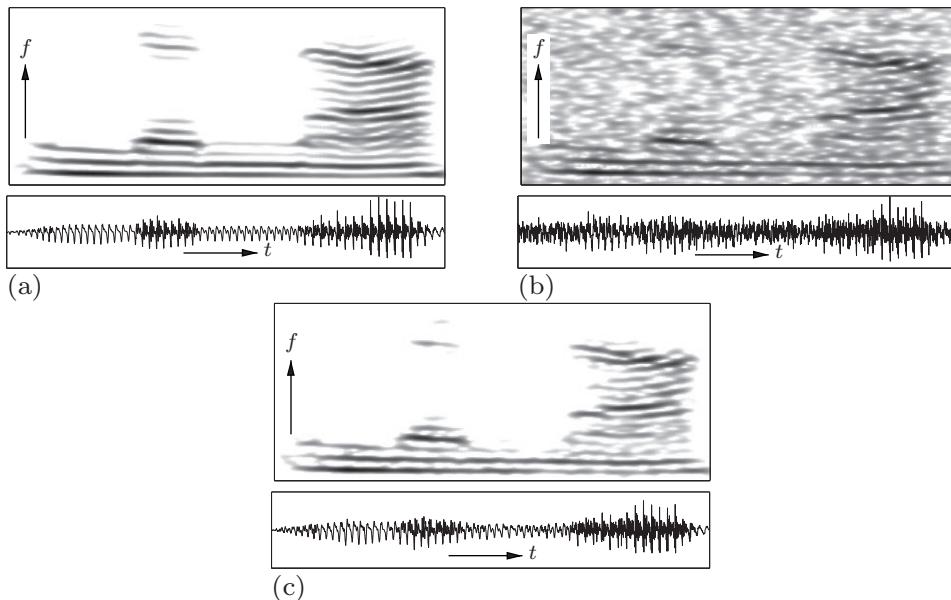
11.1.6 SIMULATION RESULTS

The first simulation example (see Fig. 11.1.3) compares the performance of the Zadeh filter \mathbf{H}_{GWF} (generalized Weyl filter with $\alpha = 1/2$) and the STFT filter $\mathbf{H}_{\gamma,g}$. The (t, f) weight function $M(t, f)$ (see Fig. 11.1.3(a)) models a bandpass filter with sinusoidally time-varying center frequency and time-varying gain. The gain is 1 in the first (earlier) half and $1/2$ in the second (later) half, with a roll-off in between. The two filters were applied to an input signal $x(t)$ consisting of two chirps (see Fig. 11.1.3(b)). The resulting output signals (shown in Fig. 11.1.3(c) and (d)) are seen to conform to the specified (t, f) weighting. Furthermore, they are effectively identical (we note that $\|\mathbf{H}_{\gamma,g}x - \mathbf{H}_{\text{GWF}}x\|_2 / \|\mathbf{H}_{\text{GWF}}x\|_2 = 0.047$), due to the smoothness of $M(t, f)$ and verifying the approximation in Eq. (11.1.6).

Figure 11.1.4 shows the application of a multiwindow Gabor filter $\tilde{\mathbf{H}}_N$ to speech enhancement (denoising). The speech signal $s(t)$ and its noisy version $x(t) = s(t) + n(t)$ (where $n(t)$ is white noise with

**FIGURE 11.1.3**

Comparison of explicit and implicit (t, f) filter designs: (a) (t, f) weight function $M(t, f)$, (b)–(d) spectrogram (top) and real part (bottom) of (b) input signal $x(t)$, (c) output signal $y(t)$ obtained with Zadeh filter, and (d) output signal $y(t)$ obtained with STFT filter (dashed lines in spectrograms indicate (t, f) pass region; time duration = 2048 samples; normalized frequency interval = $[-1/2, 1/2]$).

**FIGURE 11.1.4**

Speech enhancement using a multiwindow Gabor filter $\tilde{\mathbf{H}}_N$. The figure shows the smoothed pseudo-Wigner distribution (top) and the time-domain signal (bottom) of (a) the clean speech, (b) the noisy speech (input of $\tilde{\mathbf{H}}_N$), and (c) the enhanced speech (output of $\tilde{\mathbf{H}}_N$). (Time duration = 4096 samples; (normalized) frequency interval = $[0, 1/2]$.)

an SNR of 0 dB) are shown in Fig. 11.1.4(a) and (b). The multiwindow Gabor filter has $N = 5$ branches and lattice parameters $T = 5.8$ ms, $F = 172.25$ Hz. The analysis/synthesis windows $\gamma^{(i)}(t) = g^{(i)}(t)$ and the branch weights η_i were chosen as discussed in [1, Section 4.6.5]. The weights $M_{n,k}$ were computed from the multiwindow Gabor coefficients $c_{n,k}^{(i)} = \int_{-\infty}^{\infty} x(t) \gamma_{n,k}^{(i)*}(t) dt$ in a signal-adaptive, online manner that does not require knowledge about the clean speech signal or its statistics [1, Section 4.7.3]. The filter output is shown in Fig. 11.1.4(c); the SNR improvement is 4.92 dB.

11.1.7 SUMMARY AND CONCLUSIONS

Methods for “explicit” and “implicit” (t,f) designs of linear, time-varying filters are useful for filtering nonstationary signals if the filter characteristic can be specified in the (t,f) domain via a (t,f) weight function. All filters discussed (except the (t,f) projection filter) tend to perform similarly if the (t,f) weight function is sufficiently smooth.

The concept of implicit (t,f) filters can be generalized to a wide class of (t,f) representations based on frame theory; see Section 7.1 as well as [11,14] and references therein. Other (t,f) approaches to time-varying filtering are described in Sections 11.2–11.4. The application of (t,f) filtering to nonstationary signal detection/estimation is considered in Section 12.4 and in [2,6,15].

11.2 TIME-VARYING FILTERING USING THE STFT AND GABOR EXPANSION⁰

In many applications, like speech enhancement for mobile communications, underwater signals classification, or machine faults detection, it is essential to reduce the noise level. One suitable approach for nonstationary signals is based on time-frequency (t,f) filtering. Chapters 2 and 3 defined the short time Fourier transform (STFT), Spectrogram, Gabor expansion, and their mutual relationships. This section applies the STFT and Gabor expansion to such time-varying filtering for nonstationary signals whose frequency content changes rapidly. The methods are illustrated on two real signals: one speech signal and one engine sound signal. The first part of this section presents in detail the theory of the STFT and inverse STFT (ISTFT); the selected approach uses a vector formulation which allows an easy but rigorous formulation and implementation of the ISTFT. The second part of this section presents another solution to (t,f) filtering using the Gabor expansion.

11.2.1 THE SHORT-TIME FOURIER TRANSFORM APPROACH TO (t,f) FILTERING

The STFT-based (t,f) filtering includes three procedures in general: (1) computing the STFT, (2) filtering the STFT, and (3) signal reconstruction using filtered STFT. The ISTFT can reconstruct a time-domain signal from a filtered STFT. In most cases the filtered STFT is not valid in the sense that no

⁰Second Edition updated by **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Ouelha**, Qatar University, Doha, Qatar (samir.ouelha@qu.edu.qa). First Edition [16] by Shie Qian. Reviewers: A. Aïssa-El-Bey and S. Dong.

time-domain signal has the exact filtered STFT or desired (t, f) specifications. This requires algorithms that estimate the time-domain signal whose STFT is close in some sense to the filtered STFT. This section describes a vector formulation of the STFT, a derivation of the ISTFT and an illustration on a speech signal.

11.2.1.1 Windowing process for STFT

Let \underline{x} be a discrete-time signal $n = 0, \dots, N - 1$, such that: $\underline{x} = [x[0], \dots, x[N - 1]]^T$.

The nonstationary information in \underline{x} is observable using a window w of N_w samples, with $0 \leq N_w \leq N - 1$. All nonstationary information observed among the N samples is considered as stationary when it is observed among the N_w samples of the window used. Thus, the signal is observed with a weighted window \underline{w} , of length N_w , and with power such that

$$P_w = \frac{1}{N_w} \sum_{n=0}^{N_w-1} w^2[n]; \quad w[n] \neq 0, \quad \forall n = 0, \dots, N_w - 1. \quad (11.2.1)$$

In this case, let w^0 be a window with zero elements. Let $I_w^0 = \{n \text{ s.t. } w^0[n] = 0, \forall n = 0, \dots, N_w - 1\}$ be the set of indexes of zero elements of the window, and the number of elements in I_w^0 equals $N_w^0 < N_w$. The set $\overline{I_w^0}$ contains the indexes of nonzero elements of the window. In order to construct a window w without zero elements and retaining the power P_{w^0} , a small amount $\varepsilon_w \ll 1$ is added to the zero elements of w^0 and the penalty coefficient $\alpha_w > 0$ is applied to the elements of $\overline{I_w^0}$, such that $w[n] = \varepsilon_w$, if $n \in I_w^0$; $w[n] = \alpha_w w^0[n]$, if $n \in \overline{I_w^0}$. Its power is given by

$$\begin{aligned} P_w &= \frac{1}{N_w} \sum_{n=0}^{N_w-1} w^2[n] = \frac{1}{N_w} \sum_{n \in I_w^0} \varepsilon_w^2 + \frac{1}{N_w} \sum_{n \in \overline{I_w^0}} (\alpha_w w^0[n])^2 \\ &= \frac{N_w^0}{N_w} \varepsilon_w^2 + \frac{\alpha_w^2}{N_w} \sum_{n=0}^{N_w-1} (w^0[n])^2 = \frac{N_w^0}{N_w} \varepsilon_w^2 + \alpha_w^2 P_{w^0}. \end{aligned}$$

The added coefficient ε_w causes a change of power. So α_w is computed to conserve the power of w :

$$P_{w^0} = P_w = \frac{N_w^0}{N_w} \varepsilon_w^2 + \alpha_w^2 P_{w^0} \quad \leftrightarrow \quad \alpha_w = \sqrt{1 - \frac{N_w^0}{N_w} \frac{\varepsilon_w^2}{P_{w^0}}}.$$

11.2.1.2 Derivation of the STFT with overlap window

The overlap length between two consecutive windows is N_{overlap} , with $0 \leq N_{\text{overlap}} \leq N_w - 1$. The number of necessary windows to cover the N samples of the signal is the number L defined by

$$L = \text{ceil} \left[\frac{N - N_{\text{overlap}}}{N_w - N_{\text{overlap}}} \right], \quad (11.2.2)$$

where $\text{ceil}[\cdot]$ represents the upper rounded operator. The signal \underline{x} is padded with 0 in order to complete the L^{th} frame, so there are now N_0 samples in $\underline{x}_{\text{add}}$ with $N_0 = L(N_w - N_{\text{overlap}}) + N_w$, and the new signal $\underline{x}_{\text{add}}$ is defined by

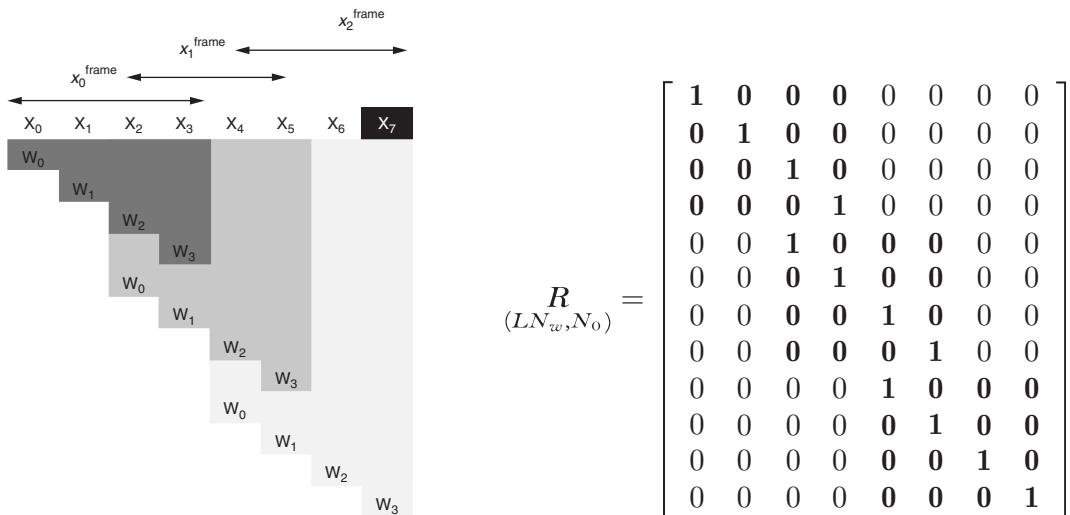
**FIGURE 11.2.1**

Illustration of the process of the windowing and overlap process, with $N = 7$, $N_w = 4$, $N_{\text{overlap}} = 2$, $N_0 = 8$; \mathbf{R} has dimension $(12,8)$.

$$\underline{x}_{\text{add}} = [x[0], \dots, x[N-1], \dots, x[N_0-1]]. \quad (11.2.3)$$

Let us define the l^{th} frame of the windowing process of $\underline{x}_{\text{add}}$:

$$x_l^{\text{frame}}[n] = \underline{x}_{\text{add}}[l(N_w - N_{\text{overlap}}) + n], \quad \forall l = 0, \dots, L-1, \forall n = 0, \dots, N_w - 1.$$

[Figure 11.2.1](#) illustrates the process described to window the signal. Then, the samples in each window are transformed by DFT using the fast Fourier transform (FFT) on N_{FFT} samples, with $N_{\text{FFT}} \geq N_w$ and $\exists q_{\text{FFT}}$ such that $N_{\text{FFT}} = 2^{q_{\text{FFT}}}$. Thus the STFT for each window is defined by

$$S_{lk} = \sum_{n=0}^{N_w-1} x_l^{\text{frame}}[n] w[n] e^{-\frac{-2j\pi nk}{N_{\text{FFT}}}}, \quad \forall l = 0, \dots, L-1, \forall k = 0, \dots, N_{\text{FFT}} - 1. \quad (11.2.4)$$

11.2.2 VECTOR FORMULATION OF THE STFT

In this part, we introduce four matrices that are used to construct the vector form of the STFT (see Eq. (11.2.11)).

11.2.2.1 Matrix of the windowed operator \mathbf{R}

Let us define a binary matrix \mathbf{R} with the dimension (LN_w, N_0) , which concatenates all the L frames when we apply it on $\underline{x}_{\text{add}}$. The general term and the indexes of non-null elements of \mathbf{R} are defined, $\forall l = 0, \dots, L-1, \forall n = 0, \dots, N_w - 1$, by

$$I(n, l) = lN_w + n; \quad J(n, l) = l(N_w - N_{\text{overlap}}) + n,$$

$$\mathbf{R}_{ij} = \delta[i - I(n, l); j - J(n, l)], \quad \forall i = 0, \dots, LN_w - 1, \quad \forall j = 0, \dots, N_0 - 1. \quad (11.2.5)$$

As $I(n, l) \geq J(n, l)$, the matrix \mathbf{R} is a lower triangular matrix; moreover $\mathbf{R}_{ij} \in \{0, 1\}$, so the sum of each row of \mathbf{R} is equal to 1:

$$\sum_{j=0}^{N_0-1} \mathbf{R}_{ij} = 1, \quad i = 0, \dots, LN_w - 1. \quad (11.2.6)$$

Figure 11.2.1 shows an example with $N = 7$, $N_w = 4$, $N_{\text{overlap}} = 2$, $N_0 = 8$, so $L = 3$;

In fact, \mathbf{R} consists of L identity matrices \mathbf{Id}_{N_w} (in bold in the matrix \mathbf{R}). Each identity matrix is shifted by $N_w - N_{\text{overlap}}$ columns to the right.

11.2.2.2 Weighted matrix operator \mathbf{W}

Let \underline{w} be the vector defined by $\underline{w} = [w[0], \dots, w[N_w - 1]]^T$. Then, \mathbf{W} is the diagonal matrix of dimension (N_w, N_w) , defined by

$$\mathbf{W}_{nm} = w[n]\delta[n - m], \quad \forall n = 0, \dots, N_w - 1, \quad \forall m = 0, \dots, N_w - 1 \quad (11.2.7)$$

so

$$\mathbf{W}_{(N_w, N_w)} = \begin{bmatrix} w[0] & & & 0 \\ & \ddots & & \\ 0 & & & w[N_w - 1] \end{bmatrix}.$$

11.2.2.3 Zero-padding Matrix \mathbf{P}

Let \mathbf{P} be the 0-padding operator matrix, defined by

$$\mathbf{P}_{(N_{\text{FFT}}, N_w)} = \begin{bmatrix} \mathbf{Id}_{N_w} \\ \mathbf{0} \end{bmatrix}_{(N_{\text{FFT}} - N_w, N_w)},$$

where $\mathbf{0}_{(N_{\text{FFT}} - N_w, N_w)}$ is a matrix of dimension $(N_{\text{FFT}} - N_w, N_w)$ with elements of only zeros. This matrix satisfies the following relation:

$$\mathbf{P}^T \mathbf{P} = \mathbf{Id}_{N_w}. \quad (11.2.8)$$

11.2.2.4 Discrete Fourier transform matrix \mathbf{F}

Let $\mathbf{F}_{(N_{\text{FFT}}, N_{\text{FFT}})}$ be the DFT matrix such that

$$\mathbf{F}_{km} = e^{-\frac{2j\pi mk}{N_{\text{FFT}}}}, \quad \forall m = 0, \dots, N_{\text{FFT}} - 1, \forall k = 0, \dots, N_{\text{FFT}} - 1. \quad (11.2.9)$$

This matrix satisfies the following relation:

$$\mathbf{F}^H \mathbf{F} = N_{\text{FFT}} \mathbf{Id}_{N_{\text{FFT}}}. \quad (11.2.10)$$

11.2.3 ANALYSIS, SYNTHESIS, AND (t, f) FILTERING USING THE STFT

11.2.3.1 Analysis using the STFT

From the formalism defined above, the expression of the STFT is given by

$$\underline{\rho}_{(LN_{\text{FFT}})} = (\mathbf{Id}_L \otimes \mathbf{FPW})(\mathbf{R}\underline{x}_{\text{add}}), \quad (11.2.11)$$

where $\underline{\rho}$ is a column vector of LN_{FFT} elements which represent the L DFT of the windowed signal $\underline{x}_l^{\text{frame}}$, computed with N_{FFT} samples and where $\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}]_{i,j}$ is the tensor product. The computation of Eq. (11.2.11) has two steps:

- the change of variable: $\hat{\underline{x}}_{(LN_w)} = \mathbf{R}\underline{x}_{\text{add}} = \begin{bmatrix} \underline{x}_0^{\text{frame}} \\ \vdots \\ \underline{x}_{L-1}^{\text{frame}} \end{bmatrix}$
- the computation of the matrix defined by

$$\mathbf{A}_{(LN_{\text{FFT}}, LN_w)} = \mathbf{Id}_L \otimes \mathbf{FPW}. \quad (11.2.12)$$

11.2.3.2 Synthesis from the STFT

To reconstruct the original time domain sequence from an STFT, Eq. (11.2.11) needs to be inverted. Clearly, the ISTFT estimation is a linear overdetermined problem for which the least squares (LS) approach is applicable. So, let $\mathbf{A}^\#$ and $\mathbf{R}^\#$ be the Moore-Penrose pseudo-inverse matrix of \mathbf{A} and \mathbf{R} defined respectively by

$$\mathbf{A}_{(LN_w, LN_{\text{FFT}})}^\# = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H, \quad (11.2.13)$$

and

$$\mathbf{R}_{(N_0, LN_w)}^\# = (\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H. \quad (11.2.14)$$

These matrices exist if: $\det \mathbf{A}^H \mathbf{A} > 0$; and $\det \mathbf{R}^H \mathbf{R} > 0$. To perform the synthesis the previous conditions must be respected. In this case, the synthesis of the time-domain signal from the STFT is given by $\hat{\underline{x}}_{(N_0)} = \mathbf{R}^\#(\mathbf{A}^\# \underline{\rho})$.

11.2.3.2.1 Expression of $\mathbf{A}^\#$

From Eqs. (11.2.12) and (11.2.13), and using the properties (11.2.7), (11.2.8), and (11.2.10) of the matrices \mathbf{F} , \mathbf{P} , and \mathbf{W} , we obtain the diagonal matrix:

$$\begin{aligned} \mathbf{A}_{(LN_w, LN_{\text{FFT}})}^H \mathbf{A}_{(LN_w, LN_{\text{FFT}})} &= (\mathbf{Id}_L \otimes \mathbf{FPW})^H (\mathbf{Id}_L \otimes \mathbf{FPW}) = (\mathbf{Id}_L \otimes \mathbf{WP}^T \mathbf{F}^H) (\mathbf{Id}_L \otimes \mathbf{FPW}) \\ &= \mathbf{Id}_L \otimes (\mathbf{WP}^T \mathbf{F}^H \mathbf{FPW}) = N_{\text{FFT}} \mathbf{Id}_L \otimes \mathbf{W}^2; \end{aligned}$$

and its determinant is $\det \mathbf{A}^H \mathbf{A} = (N_{\text{FFT}})^{LN_w} \left(\prod_{n=0}^{N_w-1} w[n]^2 \right)^L$. Given Eq. (11.2.1), the invertibility condition of $\mathbf{A}^H \mathbf{A}$ is always satisfied, i.e., $\det \mathbf{A}^H \mathbf{A} \neq 0 \Leftrightarrow w[n] \neq 0 \quad \forall n = 0, \dots, N_w - 1$; so that

$$(\mathbf{A}^H \mathbf{A})^{-1} = \frac{1}{N_{\text{FFT}}} \mathbf{Id}_L \otimes \mathbf{W}^{-2}; \text{ and finally,}$$

$$\begin{aligned} \mathbf{A}^\# &= (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H = \left(\frac{1}{N_{\text{FFT}}} \mathbf{Id}_L \otimes \mathbf{W}^{-2} \right) (\mathbf{Id}_L \otimes \mathbf{W} \mathbf{P}^T \mathbf{F}^H) \\ &= \frac{1}{N_{\text{FFT}}} \mathbf{Id}_L \otimes (\mathbf{W}^{-2} \mathbf{W} \mathbf{P}^T \mathbf{F}^H) = \frac{1}{N_{\text{FFT}}} \mathbf{Id}_L \otimes \mathbf{W}^{-1} \mathbf{P}^T \mathbf{F}^H. \end{aligned}$$

11.2.3.2.2 Expression of $\mathbf{R}^\#$

Using the definitions of \mathbf{R} and $\mathbf{R}^\#$ in Eqs. (11.2.5) and (11.2.14), and the property defined in Eq. (11.2.6), the general term of the matrix $\mathbf{R}^T \mathbf{R}$ is defined by

$$\begin{aligned} (\mathbf{R}^T \mathbf{R})_{ij} &= \sum_{k=0}^{LN_w-1} \mathbf{R}_{ik}^T \mathbf{R}_{kj} = \sum_{k=0}^{LN_w-1} \mathbf{R}_{ki} \mathbf{R}_{kj} \\ &= \left(\sum_{k=0}^{LN_w-1} \mathbf{R}_{ki}^2 \right) \delta[i-j] = \left(\sum_{k=0}^{LN_w-1} \mathbf{R}_{ki} \right) \delta[i-j]. \end{aligned}$$

This matrix is diagonal; each term is the sum of the correspondent column of \mathbf{R} , which is by construction strictly positive. So,

$$\det \mathbf{R}^T \mathbf{R} = \prod_{i=0}^{N_0-1} (\mathbf{R}^T \mathbf{R})_{ii} = \prod_{j=0}^{N_0-1} \left(\sum_{i=0}^{LN_w-1} \mathbf{R}_{ij} \right)$$

and $\left(\sum_{i=0}^{LN_w-1} \mathbf{R}_{ij} \right) > 0$ so $\det \mathbf{R}^T \mathbf{R} > 0$; thus the invertibility condition holds, and

$$(\mathbf{R}^T \mathbf{R})_{ij}^{-1} = \frac{\delta[i-j]}{\sum_{i=0}^{LN_w-1} \mathbf{R}_{ij}}. \quad (11.2.15)$$

Using Eqs. (11.2.5) and (11.2.15), we obtain

$$\begin{aligned} \mathbf{R}_{ij}^\# &= \sum_{k=0}^{LN_w-1} (\mathbf{R}^T \mathbf{R})_{ik}^{-1} \mathbf{R}_{kj}^T = \sum_{k=0}^{LN_w-1} \left(\frac{\delta[i-k]}{\sum_{m=0}^{LN_w-1} \mathbf{R}_{mk}} \right) \delta[j - I(n, l); k - J(n, l)] \\ &= \frac{1}{\sum_{k=0}^{LN_w-1} \mathbf{R}_{mi}} \delta[j - I(n, l); i - J(n, l)]. \end{aligned}$$

For example, if $N = 7$, $N_w = 4$, $N_{\text{overlap}} = 2$, $N_0 = 8$, so $L = 3$ and according to relation (11.2.5) we have $I(n, l) \in \{0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11\}$ and $J(n, l) \in \{0; 1; 2; 3; 2; 3; 4; 5; 4; 5; 6; 7\}$. And on the other hand we have $\sum_{m=0}^{11} \mathbf{R}_{mi} = \begin{cases} 1 & \text{if } i \in \{0; 1; 6; 7\} \\ 2 & \text{if } i \in \{2; 3; 4; 5; \} \end{cases}$.

Thus, $\mathbf{R}^\#$ has the dimension (8,12) and its expression is given by

$$\mathbf{R}^\# = \left[\begin{array}{cccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

Now, let us define

$$\check{x} = (\mathbf{A}^\# \underline{\rho}) = \frac{1}{N_{\text{FFT}}} (\mathbf{Id}_L \otimes \mathbf{W}^{-1} \mathbf{P}^T \mathbf{F}^H) \underline{\rho}. \quad (11.2.16)$$

Finally, the detailed expression of the synthesized signal, $\forall i = 0, \dots, N_0 - 1$, is given by

$$\hat{x}[i] = \frac{1}{\sum_{m=0}^{LN_w-1} \mathbf{R}_{mi}} \sum_{j=0}^{LN_w-1} \delta[j - I(n, l); i - J(n, l)] \check{x}[j].$$

Thus, the first N samples of \hat{x} are the samples of the original signal.

11.2.3.3 Example of (t, f) filtering using the STFT

In many applications, it is desired to filter the TFD and then estimate the time domain signal from the filtered TFD. We illustrate the STFT-based (t, f) filtering process using a clean speech signal. This signal is roughly 3 s, sampled at 8 kHz and added with a known noise. Firstly, the STFT of the signal (F_z^w) was implemented. Then, we applied the gain function G_{Wien} from the Wiener attenuation rule (see Section 11.6 and Fig. 11.2.2(a)) such that

$$F_{zd}^w[k, l] = F_z^w[k, l] G_{\text{Wien}}[k, l].$$

Then, the time-domain signal \underline{z}_r is reconstructed from the filtered STFT (F_{zd}^w), by applying the process described earlier. Finally, the STFT F_{zr}^w is computed (see Fig. 11.2.2(b)) and the error between F_{zd}^w and F_{zr}^w evaluated. The error is quantified by the normalized mean squared error (NMSE):

$$\text{NMSE} = 10 \log_{10} \frac{1}{N_{\text{tot}}} \sum_{k=0}^{N_{\text{FFT}}-1} \sum_{l=0}^{L-1} (|F_{zd}^w[k, l]| - |F_{zr}^w[k, l]|)^2,$$

where N_{tot} is the number of elements of the matrices F_{zd}^w and F_{zr}^w .

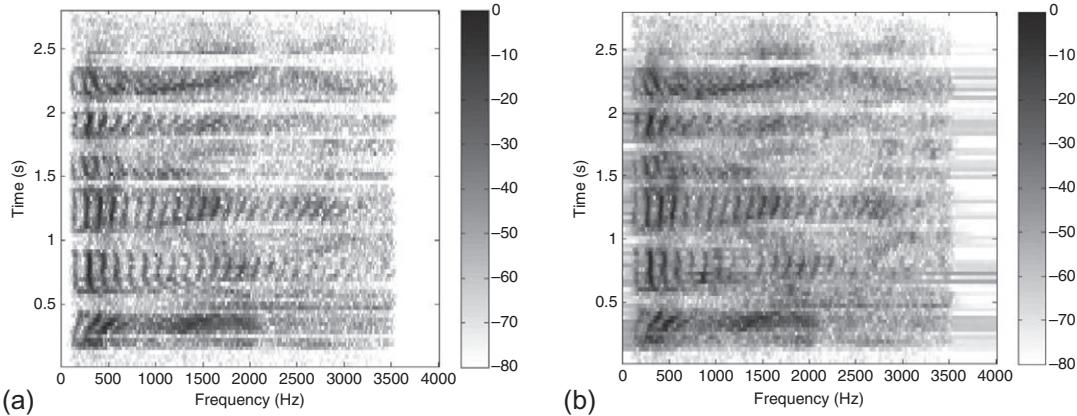
**FIGURE 11.2.2**

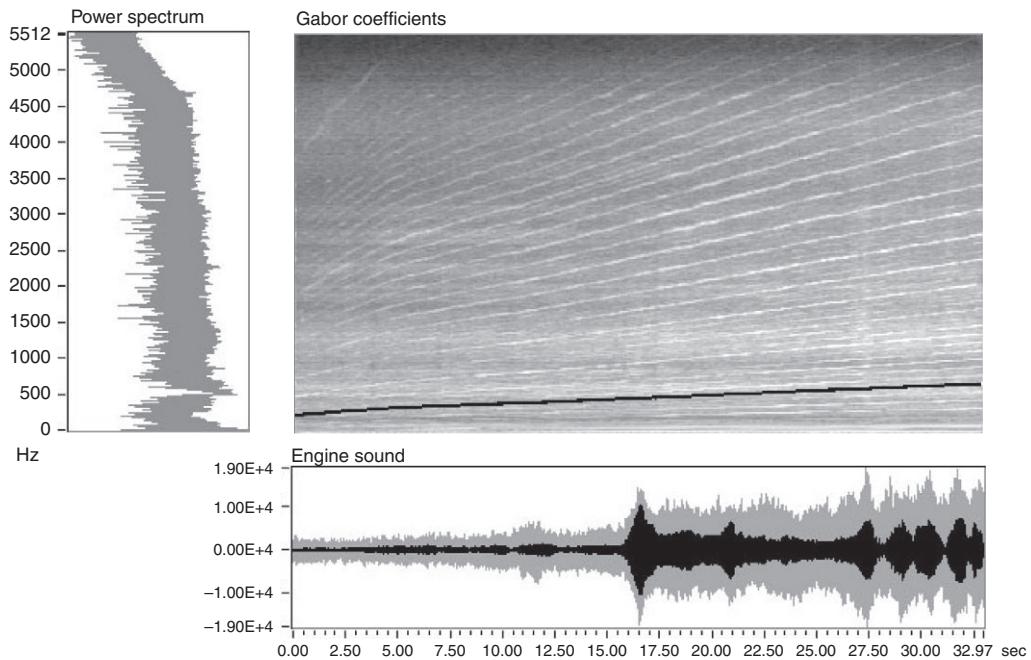
Illustration of (t, f) filtering on a speech signal using the inverse STFT. (a) True STFT, (b) reconstructed STFT; parameters: hanning window with $N_w = 512$, $N_{\text{overlap}} = 256$, $N_{\text{FFT}} = 512$.

The error between F_{z_d} and $F_{z_r}^w$ equals -18 dB. This result illustrates that the modified STFT F_{z_d} is not exactly valid in the sense that no time-domain signal may have the exact modified STFT; but it is valid in the sense of least mean square error.

11.2.4 THE GABOR EXPANSION APPROACH TO (t, f) FILTERING

11.2.4.1 Filtering a six-cylinder engine sound

Figure 11.2.3 illustrates an audio signal $s(t)$ recorded during the run-up of a six-cylinder engine. The sound is created by the engine rotation and the other parts that vibrate due to the engine rotation. The sound signal plotted in the bottom is a combination of all vibrations. These vibration frequencies are multiples of the fundamental frequency corresponding to the engine rotation speed. When the engine speed is constant, the classical FT can isolate the vibrations. When the engine runs up/down, the fundamental frequency and its multiples increase/decrease with time. The corresponding frequency bandwidths become wide and overlap each other, as shown in the left plot. Consequently, the FT is no longer able to distinguish different vibrations. In the automobile industry, such time-varying harmonics are called *orders* to distinguish them from time-invariant harmonics. To evaluate a vibration, engineers use adaptive filters to extract the signal $s(t)$ corresponding to a particular order. Using $s(t)$, engineers can then obtain further information, such as phase and amplitude. Given the signal nonstationarity, a (t, f) approach is appropriate. The order tracking can be performed by, e.g., the Gabor expansion (see Section 2.3.3 on the connection between the Gabor transform and the STFT). The middle plot of Fig. 11.2.3 shows the magnitudes of the Gabor coefficients for the sound signal. While neither $s(t)$ nor its spectrum gives information about the structure of the engine sound, orders have distinct signatures in the joint (t, f) plot. These allow us to select the desired Gabor coefficients and perform a Gabor expansion to obtain the corresponding modified signal, which is a time-varying-filtered version of the original signal $s(t)$.

**FIGURE 11.2.3**

Gabor-expansion-based time-varying filter for order tracking. The bottom plot shows the original engine run-up sound (light gray) vs. the extracted sixth-order time waveforms (dark gray). The thin dark line in the main plot marks the Gabor coefficients corresponding to the sixth order.

11.2.4.2 Discrete Gabor expansion

The definition of the Gabor expansion is given in [Section 2.3.3](#). For a discrete-time signal $s[n]$ with a selected window $h[n]$, the Gabor expansion is defined as [17,18]

$$s[n] = \sum_m \sum_{k=0}^{K-1} c_{m,k} h[n - mT] e^{j2\pi kn/K}, \quad (11.2.17)$$

where the Gabor coefficients $c_{m,k}$ are computed by the regular STFT, as per Eq. (2.3.19); i.e.,

$$c_{m,k} = \sum_n s[n] g^*[n - mT] e^{-j2\pi kn/K}, \quad (11.2.18)$$

where $g[n]$ is a window function and the parameters T and K denote the time sampling interval (or decimation) and the number of frequency bins (or bands). The ratio K/T determines the sampling rate. When $K/T = 1$, we have critical sampling; in this case there is no redundancy for the resulting Gabor coefficients. When $K/T > 1$, we have oversampling. For stable reconstruction, we need $K/T \geq 1$.

Note that the window functions $h[n]$ and $g[n]$ are exchangeable. In other words, either of them can be used as the analysis or the synthesis function. They satisfy the so-called dual relation. The core issue

of the Gabor expansion is how to compute the dual function for a given function (either $h[n]$ or $g[n]$). Many methods have been proposed, each with its pros and cons. The approach presented here ensures the resulting dual window function always has the same length as the given function.

Without loss of generality, let us assume that the given window function is $h[n]$ with N_h points, and unit energy. Then, the corresponding dual function $g[n]$ can be solved by T independent linear equations, i.e.,

$$\mathbf{A}_l \underline{g}_l = \underline{u}_l, \quad l = 0, 1, \dots, T - 1, \quad (11.2.19)$$

where the elements of the matrix \mathbf{A}_l , and vectors \underline{g}_l and \underline{u}_l are defined as

$$\begin{aligned} a_l[q, p] &= \tilde{a}[l + pT + qK], \\ g_l[p] &= g[l + pT], \\ u_l[q] &= (K^{-1}, 0, 0, \dots)^T, \end{aligned}$$

where $0 \leq p < N_h/T$. The periodic auxiliary function $\tilde{a}[n]$ is defined as

$$\tilde{a}[n + i(2N_h - K)] = \begin{cases} h[n] & 0 \leq n < N_h, \\ 0 & N_h \leq n < 2N_h - K, \end{cases} \quad 0 \leq q < \frac{2N_h}{K} - 1, \quad (11.2.20)$$

where i is the period index. If the window length is equal to the signal length, the periodic auxiliary function $\tilde{a}[n]$ is simply

$$\tilde{a}[n + iN_h] = h[n], \quad 0 \leq q < \frac{N_h}{K}.$$

Note that the solution of Eq. (11.2.19) is not unique for oversampling. To ensure that the analysis and synthesis window functions are both concentrated in the joint (t, f) domain, we require that the dual function is optimally similar to the given window function in the least mean square sense (LMS), where the LMS error is

$$\text{LMSE} = \min_{\mathbf{A}\underline{g}=\underline{u}} \left\| \frac{\underline{g}}{|\underline{g}|} - \frac{\underline{h}}{|\underline{h}|} \right\|^2.$$

When LMSE is small, i.e., $\underline{g} \approx \underline{h}$, Eq. (11.2.18) becomes

$$c_{m,k} = \sum_n s[n] h^*[n - mT] e^{-j2\pi nk/K}. \quad (11.2.21)$$

Then Eqs. (11.2.21) and (11.2.17) form an *orthogonal-like Gabor transform pair*. In this case, the Gabor coefficients $c_{m,k}$ are exactly the signal projection on the window function $h[n]$.

Since at critical sampling rate the pair of dual functions cannot be simultaneously concentrated in both the time and frequency domains, we usually employ the oversampling scheme. Consequently, the resulting Gabor transformation is redundant. In this case, the Gabor coefficients form a subspace of two-dimensional functions. In other words, for an arbitrary two-dimensional function, there may be no corresponding time waveform. For example, assume that we have a modified two-dimensional function

$$\hat{c}_{m,k} = \phi_{m,k} c_{m,k},$$

where $\phi_{m,k}$ denotes a binary mask function, equal to either zero or one. Apply the Gabor expansion (11.2.17) to obtain

$$\hat{s}[n] = \sum_m \sum_{k=0}^{K-1} \hat{c}_{m,k} h[n - mT] e^{j2\pi nk/K}. \quad (11.2.22)$$

Then we find that

$$\sum_n \hat{s}[n] g^*[n - mT] e^{-j2\pi nk/K} \neq \hat{c}_{m,k}.$$

The Gabor coefficients of the reconstructed time waveform $\tilde{s}[n]$ will not be inside the masked area determined by the mask function $\phi_{m,k}$.

11.2.4.3 Time-varying filtering

To overcome the above-mentioned problem, we can introduce an iterative approach. That is, for a set of two-dimensional Gabor coefficients, first determine a binary mask matrix. Apply the mask to the two-dimensional Gabor coefficients to preserve desirable coefficients and remove unwanted ones. Then, compute the Gabor expansion. Once the time waveform is obtained, compute the new Gabor coefficients. Repeat this process until the time waveforms converge. For the sake of clarity, let us rewrite the pair of Gabor expansion (11.2.17) and (11.2.18) in matrix form, i.e.,

$$\begin{aligned} \mathbf{C} &= \mathbf{G}\underline{s}, \\ \underline{s} &= \mathbf{H}^T \mathbf{G}\underline{s}, \end{aligned}$$

where \mathbf{H} and \mathbf{G} denote the analysis and synthesis matrices, respectively. Note that for oversampling

$$\mathbf{H}^T \mathbf{G} = \mathbf{I}, \quad \mathbf{G}\mathbf{H}^T \neq \mathbf{I}.$$

Then, the iterative process can be described as

$$\begin{aligned} \hat{\mathbf{C}}^{(1)} &= \Phi \mathbf{C} \\ \underline{s}^{(1)} &= \mathbf{H}^T \hat{\mathbf{C}}^{(1)} \\ \mathbf{C}^{(2)} &= \mathbf{G}\underline{s}^{(1)} = \mathbf{G}\mathbf{H}^T \Phi \mathbf{C} \\ \hat{\mathbf{C}}^{(2)} &= \Phi \mathbf{C}^{(2)} \\ \underline{s}^{(2)} &= \mathbf{H}^T \hat{\mathbf{C}}^{(2)} \\ \mathbf{C}^{(3)} &= \mathbf{G}\underline{s}^{(2)} = \mathbf{G}\mathbf{H}^T \Phi \mathbf{G}\mathbf{H}^T \Phi \mathbf{C} = (\mathbf{G}\mathbf{H}^T \Phi)^2 \mathbf{C} \\ &\vdots \\ \mathbf{C}^{(n)} &= (\mathbf{G}\mathbf{H}^T \Phi)^{n-1} \mathbf{C}. \end{aligned}$$

Where in the above the superscript (n) indicates the order. It can be shown [19] that *if and only if*

$$\sum_{i=0}^{N_h/K-1} g^*[iK + n] h[iK + n + mT] = \sum_{i=0}^{N_h/K-1} h^*[iK + n] g[iK + n + mT] \quad (11.2.23)$$

for $0 \leq n < K$ and $0 \leq m < N$, where N is equal to the number of time sampling points and N_h denotes the window length, *then*

- (1) $\mathbf{C}^{(n)}$ and $\underline{s}^{(n)}$ converge and
- (2) $\mathbf{C}^{(n)} = \Phi\mathbf{C}$, $n \rightarrow \infty$; that is, the support of $\mathbf{C}^{(n)}$ in the (t,f) domain is inside the masked area.

For more details on this iterative approach see [19]. Two trivial cases for Eq. (11.2.23) are

- (1) the critical sampling, $K = T$. Note that in this case the analysis and synthesis windows cannot both be localized in the joint (t,f) domain.
- (2) $g[n] = h[n]$. In this case the Gabor coefficients $\mathbf{C}^{(2)}$, after the first iteration, are closest in the LMSE sense to the masked Gabor coefficients $\Phi\mathbf{C}$ (that is, desirable Gabor coefficients).

The second case usually implies heavy oversampling (or huge redundancy) and is therefore impractical due to computation speed and memory consumption. Usually, we pursue the orthogonal-like representation introduced early. It has been found that for those commonly used window functions, such as Gaussian and Hanning windows, the difference between the analysis and synthesis windows would be negligible when the oversampling rate is four.

11.2.4.4 Application to car engine sound (t,f) filtering

Figure 11.2.3 illustrates the application of the Gabor-expansion-based time-varying filter to order tracking. The dark waveform in the bottom plot depicts the sixth order extracted by the Gabor-expansion-based time-varying filter. In this example, $h[n]$ is the Hanning window. The oversampling rate was selected high enough (four) so that the dual function $g[n]$ has a form that is almost identical to $h[n]$. It has been found that after a few iterations, the difference (LMSE) between $\underline{s}^{(n)}$ and $\underline{s}^{(n-1)}$ reduces to 10^{-4} . The resulting time waveform was found to be almost identical to that computed by other methods.

11.2.5 SUMMARY AND CONCLUSIONS

This section presents two algorithms for (t,f) filtering of nonstationary signals, i.e., the short-time Fourier transform and the Gabor expansion. A speech application of the STFT defined in Chapter 2 is described. In addition, another real-world application of the Gabor expansion defined in Section 2.3.3 is given to illustrate time-varying filtering. More details can be found in [19]. Other methods for (t,f) filtering are presented in Sections 11.3, 11.4, 11.5, 11.6, and 12.3. There is a wide range of other applications of the concepts for (t,f) filtering including biomedicine [20] and telecommunications [21,22].

11.3 TIME-FREQUENCY FILTERING OF SPEECH SIGNALS IN HANDS-FREE TELEPHONE SYSTEMS⁰

Time-varying filtering of noisy speech signals is a very attractive challenge, with the main question: What does the most appropriate time-varying filter scheme look like? Speech signals are of highly

⁰Author: S. Stanković, University of Montenegro, Podgorica, Montenegro (srdjan@ac.me). Reviewers: J. Tilp, LJ. Stanković, and S. Ouelha.

nonstationary and multicomponent nature. If we deal with filtering of noisy speech signals, as they occur in hands-free telephone systems, then the desired scheme would provide a signal-to-noise ratio (SNR) greater than approximately 12 dB. At the same time it should be suited for real-time implementation, with time delay less than 39 ms for mobile telephony, and 2 ms for circuit-switched telephony. The second requirement can cause additional difficulties and restrictions on finding an appropriate time-varying filter procedure.

11.3.1 LIMITATIONS OF THE QUASI-STATIONARY APPROACH

The most commonly used approach in the filtering of speech signals is the so-called quasi-stationary approach, where it is assumed that the signal is stationary in the time interval T , with T between 20 and 40 ms being often used [23,24]. In this interval of time, classical speech enhancement schemes such as those given in Table 11.3.1 are used [25,26] (these operators are also described in Section 11.6).

The noise is reduced by applying frequency-dependent suppression factors according to the various filtering rules given in Table 11.3.1. From the aspect of time-varying filtering we can say that this technique is quasi time-varying filtering. Thus, we have a sliding window of duration T along the signal where the filtering is performed after every T or after every $T/2$ (the second case is used in an overlap-add scheme in order to avoid block effects).

From the point of view of time-frequency (t,f) analysis, having in mind the high nonstationarity of speech signals, we can conclude that the quasi-stationary approach of filtering is approximate in nature, and that it will more or less satisfy subjective perception requirements. If we want to achieve more accurate and more objective filtering of speech signals, time-varying filtering needs to be applied. Since a unique definition of (t,f) spectra does not exist (as is shown by the variety of (t,f) distributions (TFDs) derived in Chapter 2), several approaches to time-varying filtering have been proposed. We will use the one based on the Wigner distribution. It uses the Wigner spectrum, where the statistically independent cross-terms in Wigner distribution are averaged out. However, in order to calculate the Wigner spectrum it is necessary to have many different realizations of the same random process at a given instant. Clearly, in the case of real-time applications, the processing has to be based on a single noisy speech realization. It is the reason for using an approximation, in the sense that the Wigner spectrum is replaced by a cross-term-free or cross-term-reduced TFD. According to the additional criterion of realization simplicity, special attention will be devoted to the filtering based on the spectrogram and TFDs whose realization is directly related to the spectrogram. The use of other reduced interference TFDs for filtering, in place of the spectrogram, is straightforward.

Table 11.3.1: Filter Transfer Function for Different Algorithms, Where $S_{xx}^2(f)$ and $S_{\epsilon\epsilon}^2(f)$ Are the Power Spectra of the Noisy Signal and Noise, Respectively, and λ Is an Overestimation Factor

| Algorithm | Wiener | Spectral Subtraction | Maximum Likelihood | Magnitude Subtraction |
|---------------------|---|--|---|--|
| Filter trans. func. | $1 - \frac{S_{\epsilon\epsilon}^2(f)}{S_{xx}^2(f)}$ | $\sqrt{1 - \frac{\lambda S_{\epsilon\epsilon}^2(f)}{S_{xx}^2(f)}}$ | $\frac{1}{2} \left[1 + \sqrt{1 - \frac{S_{\epsilon\epsilon}^2(f)}{S_{xx}^2(f)}} \right]$ | $1 - \sqrt{\frac{S_{\epsilon\epsilon}^2(f)}{S_{xx}^2(f)}}$ |

11.3.2 TIME-VARIANT FILTERING OF SPEECH SIGNALS

By analogy with the filtering of stationary signals, nonstationary time-varying filtering of a noisy signal can be defined by [27–29]

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2}) x(t + \tau) d\tau. \quad (11.3.1)$$

The noisy signal $x(t) = s(t) + \epsilon(t)$ includes the desired signal $s(t)$ and the noise $\epsilon(t)$. The impulse response of the time-varying filter is $h(t + \frac{\tau}{2}, t - \frac{\tau}{2})$. The optimal transfer function

$$L_H(t, f) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2}) x(t + \tau) e^{-j2\pi f \tau} d\tau$$

is defined by the relation [3,30–32]:

$$\bar{W}_{sx}(t, f) = L_H(t, f) \bar{W}_{xx}(t, f), \quad (11.3.2)$$

where

$$\bar{W}_{xx}(t, f) = E\{W_{xx}(t, f)\} = \int_{-\infty}^{\infty} E\left\{x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2})\right\} e^{-j2\pi f \tau} d\tau \quad (11.3.3)$$

is the mean value of the Wigner distribution $W_{xx}(t, f)$ of the signal $x(t)$ (i.e., the Wigner spectrum [33] of $x(t)$) (see Chapter 2, Section 2.2). We can conclude that Eq. (11.3.2) is of the same form as the Wiener filter for the stationary case.

If the signal and noise are not correlated, we have

$$L_H(t, f) = \frac{\bar{W}_{ss}(t, f)}{\bar{W}_{ss}(t, f) + \bar{W}_{\epsilon\epsilon}(t, f)}. \quad (11.3.4)$$

Consider now Eq. (11.3.4). Clearly, the mean value $E\{W_{ss}(t, f)\} = \bar{W}_{ss}(t, f)$ will eliminate uncorrelated cross-terms in the Wigner distribution, since

$$E\left\{s_i(t + \frac{\tau}{2}) s_j^*(t - \frac{\tau}{2})\right\} = 0 \text{ for } i \neq j,$$

as long as components $s_i(t)$ and $s_j(t)$ are not correlated [33]. However, if we have to perform filtering on the base of a single realization, the Wigner distribution should be used instead of the Wigner spectrum in Eq. (11.3.4). For filtering of multicomponent signals, Eq. (11.3.4) is useless because emphatic cross-terms will appear (see Chapter 3, Section 3.1.2). The problem of cross-terms will be partially overcome if we modify the definition (11.3.4) so that we apply some of the cross-term-reduced distributions $\rho(t, f)$ instead of the Wigner distribution. In this case we have

$$L_H(t, f) = \frac{\rho_{ss}(t, f)}{\rho_{ss}(t, f) + \rho_{\epsilon\epsilon}(t, f)}. \quad (11.3.5)$$

It is clear that definition (11.3.5) is an approximation of Eq. (11.3.4) with $\rho(t, f)$ approximating the Wigner spectrum.

11.3.3 TIME-VARYING FILTERING BASED ON THE SPECTROGRAM

In order to obtain a more efficient filter for numerical implementations, the previous definitions can be slightly modified by using their pseudo form:

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2}) w(\tau) x(t + \tau) d\tau. \quad (11.3.6)$$

Here, a lag window $w(\tau)$ is introduced. It can be shown [28] that for frequency modulated signals, $w(\tau)$ does not influence the output signal $(Hx)(t)$ if $w(0) = 1$. By using Parseval's theorem, Eq. (11.3.6) can be written in the form

$$(Hx)(t) = \int_{-\infty}^{\infty} L_H(t, f) F_x(t, f) df, \quad (11.3.7)$$

where

$$F_x(t, f) = \int_{-\infty}^{\infty} x(t + \tau) w(\tau) e^{-j2\pi f\tau} d\tau$$

is the short-time Fourier transform of the signal $x(t)$.

The choice of $\rho(t, f)$ in Eq. (11.3.5) will play a crucial role in the time-varying filter scheme. Clearly, for efficient time-varying filtering, it is desired that the chosen $\rho(t, f)$ satisfies three main conditions: (1) satisfactory noise reduction; (2) appropriateness for real-time realization; and (3) auto-terms close to those in the Wigner spectrum. The simplest and most commonly used $\rho(t, f)$, for which the real-time application is very well studied, is the spectrogram (see Section 2.3.1), which is the squared modulus of the short-time Fourier transform:

$$S_x(t, f) = |F_x(t, f)|^2 = \left| \int_{-\infty}^{\infty} x(t + \tau) w(\tau) e^{-j2\pi f\tau} d\tau \right|^2.$$

The main problem of using the short-time Fourier transform (and the spectrogram) is in determination of the window length $w(t)$. A narrow window produces a better time resolution, while a wider window gives a better frequency resolution. The window should be chosen by a compromise of these two opposite requirements.

Having in mind that a speech signal is approximately stationary within the interval T between 20 and 40 ms, for a sampling rate of $f_s = 8$ kHz, we conclude that we can use a lag window width of $N = 256$ samples, corresponding to $T = 32$ ms. In order to achieve a more accurate calculation of integral (11.3.7), zero padding up to 1024 samples is used.

Consider now the spectrogram-based filtering of noisy speech signals, recorded in a car cruising along the highway. Estimations of the spectrogram of noise are performed in only one time instant during a speech pause. This assumption is made in order to have the worst filtering situation as in a real case. Since the noisy signal contains significant noise components in the low frequency range (below 98 Hz) where, in our application, no speech components exist, the signal is prefiltered by using a high-pass filter with cut-off frequency 98 Hz. In this realization we will apply the time-varying Wiener filter definition (11.3.5), with the spectrogram instead of $\rho(t, f)$, and the time-varying version of the spectral subtraction definition:

$$L_{HW}(t, f) = 1 - \frac{S_{\epsilon}(t, f)}{S_x(t, f)}, \quad (11.3.8)$$

$$L_{\text{HSS}}(t,f) = \sqrt{1 - \lambda \frac{S_\epsilon(t,f)}{S_x(t,f)}}. \quad (11.3.9)$$

In Eq. (11.3.9), λ is an overestimation factor applied in order to give some correction of the errors caused by the assumption that the noise is stationary in the interval between two pause estimations. The value $\lambda = 4$ is used in this study. Modifications of Eqs. (11.3.8) and (11.3.9) are used after introducing a spectral floor β [24]:

$$L_{\text{HW}}(t,f) = \max \{L_{\text{HW}}(t,f), \beta\}, \quad (11.3.10)$$

$$L_{\text{HSS}}(t,f) = \max \{L_{\text{HSS}}(t,f), \beta\}, \quad (11.3.11)$$

where “=” indicates assignment.

11.3.4 RESULTS

In our examples the spectral floors are set to $\beta = 0.12$ and $\beta = 0.08$ in Eqs. (11.3.10) and (11.3.11), respectively.

Note that as λ increases, better noise reduction is obtained, but the distortion of the signal becomes significant. As β increases, more noise remains in the signal, but speech distortion is audible. Thus, these two factors are chosen by compromise. Observe that if

$$L_H(t,f) = \begin{cases} 1 & \text{for } \rho(t,f) > 0 \text{ (or } \rho(t,f) > \text{threshold}), \\ 0 & \text{otherwise,} \end{cases} \quad (11.3.12)$$

the (t,f) features of audio signals are obtained, and thus can be efficiently used in many applications such as watermarking [34–36] and signal classification [37].

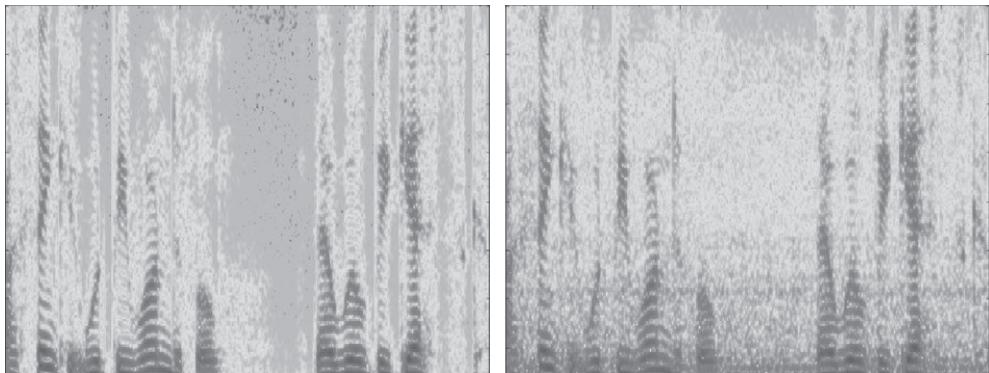
Figure 11.3.1 shows the (t,f) representations (TFRs) of a clean signal (left) and a noisy signal (right). Figure 11.3.2 shows TFRs of the denoised signal, filtered by using time-varying Wiener filtering (upper left) and time-varying spectral-subtraction filtering (upper right), both based on the spectrogram. The noise suppression is clearly better when the time-varying spectral-subtraction filter definition is used, because the overestimation factor λ provides better estimation of the spectrogram of the noise.

11.3.5 USING RIDs FOR TIME-FREQUENCY FILTERING

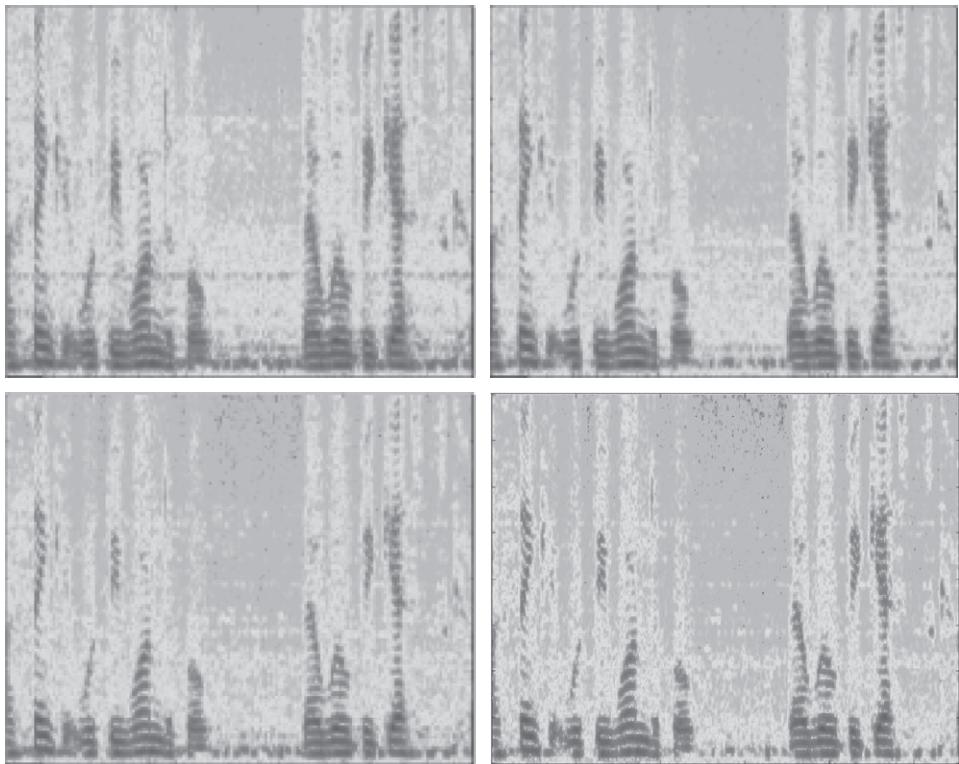
Now, there is the question of whether it is possible to use some other TFDs, in order to further improve the filtering results. The answer is yes. Namely, we can use reduced-interference distributions (RIDs) which belong to the general quadratic class of TFDs described in Chapters 2 and 3 and expressed as [38]

$$\rho_{xx}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\nu, \tau) x(u + \frac{\tau}{2}) x^*(u - \frac{\tau}{2}) e^{j2\pi\nu t} e^{-j2\pi f \tau} e^{-j2\pi \nu u} du dv d\tau, \quad (11.3.13)$$

where the kernel $g(\nu, \tau)$ specifies the distribution. The most commonly used distributions are described in Chapters 2 and 3 [39].

**FIGURE 11.3.1**

Spectrograms of (left) the clean speech signal; (right) the noisy signal filtered by a high-pass filter.

**FIGURE 11.3.2**

Denoised signal obtained by filtering based on (upper left) the time-varying Wiener filter using the spectrogram; (upper right) time-varying spectral subtraction using the spectrogram; (lower left) the time-varying Wiener filter using the SM; (lower right) time-varying spectral subtraction using the SM.

When we use reduced-interference distributions, it is important to know that, in the case of a noisy signal, the distance between two auto-terms during voiced segments of speech is approximately equal to the value of the fundamental frequency in the case of a signal without noise. In the noisy case, we also have harmonically shaped components of noise, which can occur between the auto-terms of speech, causing additional cross-terms and errors in filtering [40].

A very simple and flexible implementation can be obtained by using the S-method (SM), described in [Chapter 3](#) and [Section 6.2](#) [41], whose realization is straightforwardly based on the STFT. The consequence of this fact is that time-varying filtering based on the SM is a simple extension of the spectrogram-based filtering. Additionally, the SM of a multicomponent signal

$$x(t) = \sum_{i=1}^N x_i(t) \quad (11.3.14)$$

can assume the form $\text{SM}_{xx}(t,f) \cong \sum_{i=1}^N W_{x_i x_i}(t,f)$, being a desired approximation of the Wigner distribution auto-terms.

The SM is defined in [Section 3.3.4.1](#) as

$$\text{SM}_{xx}(t,f) = 2 \int_{-\infty}^{\infty} P(v) F_x(t,f+v) F_x^*(t,f-v) dv, \quad (11.3.15)$$

where $P(v)$ is a rectangular window in the frequency domain. Discretization of the SM in Eq. (11.3.15), taking a rectangular window for P , produces

$$\begin{aligned} \text{SM}_{xx}(n,k) &= \sum_{l=-L}^L F_x(n, k+l) F_x^*(n, k-l) \\ &= |F_x(n, k)|^2 + 2 \operatorname{Re} \left\{ \sum_{l=1}^L F_x(n, k+l) F_x^*(n, k-l) \right\}. \end{aligned} \quad (11.3.16)$$

From the previous equation we see that the SM realization is based on the spectrogram. Thus, filtering based on the SM is a straightforward extension of the previously considered filter schemes:

$$L_{HW}(t,f) = \max \left\{ 1 - \frac{\text{SM}_\epsilon(t,f)}{\text{SM}_x(t,f)}, \beta \right\}, \quad (11.3.17)$$

$$L_{HSS}(t,f) = \max \left\{ \sqrt{1 - \lambda \frac{\text{SM}_\epsilon(t,f)}{\text{SM}_x(t,f)}}, \beta \right\}. \quad (11.3.18)$$

In our experiments we have used the SM with $L = 3$, and spectral floors $\beta = 0.12$ and $\beta = 0.08$, respectively [42].

The lower half of [Fig. 11.3.2](#) shows the denoised signals obtained by using time-varying Wiener filtering (lower left) and time-varying spectral-subtraction filtering (lower right), both based on the SM. By comparing the results with the ones produced by using the spectrogram-based filtering, the improvements are obvious. It is important to note that the SM has a form very suitable for simple hardware realization. This property is attractive for on-line applications.

11.3.6 SUMMARY AND CONCLUSIONS

Time-varying filtering of speech signals disturbed by car noise is presented. On the base of the time-varying Wiener filter form, the time-varying spectral subtraction form of filtering is introduced. The filtering is performed on the base of the spectrogram and the S-method. The proposed filter schemes are efficient and suitable for hardware realization.

11.4 SIGNAL DENOISING BY TIME-FREQUENCY PEAK FILTERING⁰

11.4.1 DENOISING FOR SNR ENHANCEMENT

Time-frequency peak filtering (TFPF) is an innovative alternative to the filtering methods described earlier in this chapter.

The TFPF method was developed with the author's former PhD students Morgan Arnold and Mark Roessgen [43–45]. It is based on encoding the noisy signal as the instantaneous frequency (IF) of a unit-amplitude frequency-modulated (FM) analytic signal. The IF of the analytic signal is then estimated using standard time-frequency peak-detection methods [46] to obtain an estimate of the underlying deterministic signal. For some signals, TFPF using a windowed Wigner-Ville distribution (WVD) results in a significant enhancement of signals for SNR as low as −9 dB.

The rationale for introducing such a method is as follows. First, the signals considered are assumed to be sums of arbitrary numbers of band-limited nonstationary components in additive noise. For high-SNR situations many signal processing algorithms such as IF estimation operate at desirable performance levels, but most perform poorly when SNR falls below a given threshold [46]. In this case, signal denoising algorithms can be used as a preprocessing stage to improve the SNR and reduce the distorting effects of noise. For this purpose, both adaptive and fixed methods have been developed to deal with the case of nonstationary signals in noise. Adaptive techniques are naturally superior in performance to fixed methods, but they perform poorly in certain conditions, such as the case of nonstationary signals whose spectral content changes rapidly with time. This suggests the need for a more general filtering method when the SNR is low and the underlying signal statistics vary rapidly with time. Hence the introduction of time-frequency peak filtering.

11.4.2 TIME-FREQUENCY PEAK FILTERING (TFPF)

11.4.2.1 Basic principle of TFPF

Let us consider a signal $x(t)$ corrupted by additive noise $n(t)$ as expressed below:

$$s(t) = x(t) + n(t) = \sum_{k=1}^p x_k(t) + n(t), \quad (11.4.1)$$

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: M.G. Amin and D.L. Jones.

where $n(t)$ is an additive white Gaussian noise (WGN) and $x_k(t)$ are band-limited nonstationary deterministic components that may have overlapping frequency spectra. The stated problem is to recover the signal $x(t)$ given the observation of $s(t)$.

The IF of an analytic signal $z(t) = a(t) e^{j2\pi\phi(t)}$ is defined in [Chapter 1](#) and [46]:

$$f_z(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}, \quad (11.4.2)$$

where $\phi(t)$ is the instantaneous phase and $a(t)$ is the instantaneous amplitude of the analytic signal $z(t)$ which can be expressed as [47]

$$z(t) = a(t) e^{j2\pi \int_{-\infty}^t f_z(\lambda) d\lambda}. \quad (11.4.3)$$

The selected IF estimation [46] is based on taking the peak of the signal's TFD. The WVD is a natural first choice as a TFD for peak filtering given that the other quadratic TFDs are simply smoothed versions of the WVD [47]. The WVD of the analytic signal $z(t)$ is defined (in [Section 2.1.4](#)) as

$$W_z(t, f) = \int_{-\infty}^{\infty} z(t + \tau/2) z^*(t - \tau/2) e^{-j2\pi f\tau} d\tau. \quad (11.4.4)$$

When the signal's IF is linear, with infinite duration, delta functions will appear at the positions of the IF providing a perfect signal IF estimate. For other monocomponent nonlinear FM signals, the WVD still shows significant energy concentration around the signal's IF, although the delta function behavior is lost. The IF estimate is found by maximizing the WVD over frequency [46]; that is,

$$\hat{f}_z(t) = \operatorname{argmax}_f [W_z(t, f)]. \quad (11.4.5)$$

The IF estimate based on the peak of the WVD is unbiased, with variance approaching the Cramer-Rao lower bound, for signals with linear IF laws in *additive* white zero-mean noise with moderate to high SNR [46]. However, as the order of the polynomial IF increases, the delta functions will be replaced by less peaky functions, with peaks that lie away from the true IF, resulting in biased IF estimates. To remedy this, a windowed WVD is used, such that the signal IF is as close to linear as possible across the entire window length.

11.4.2.2 TFPF: Method and properties

Time-frequency peak filtering consists of a two-step procedure whereby the signal to be filtered is first encoded as the IF of a unit-amplitude FM analytic signal. Then, the IF is estimated by taking the peak of a time-frequency distribution (TFD) to recover the filtered signal. This may be summarized as follows:

Step 1. Encode the noisy signal $s(t)$ via FM modulation as $z_s(t) = e^{j2\pi\mu \int_0^t s(\lambda) d\lambda}$, where μ is a scaling parameter analogous to the FM modulation index.

Step 2. Estimate the peak of the WVD of the analytic signal $z_s(t)$:

$$\hat{x}(t) = \hat{f}_z(t) = \operatorname{argmax}_f [W_{z_s}(t, f)]/\mu.$$

The properties of the encoding and IF estimation steps of TFPF are derived for the case of WVD only. The use of other TFDs results in slightly different properties.

Property 1. The encoding step converts the additive noise $n(t)$ to multiplicative noise $z_n(t)$ that modulates the signal component $z_x(t)$; that is,

$$z_s(t) = e^{j2\pi\mu \int_0^t s(\lambda) d\lambda} = z_x(t) z_n(t), \quad (11.4.6)$$

where the encoded noise and deterministic signal components are given by

$$z_x(t) = e^{j2\pi\mu \int_0^t x(\lambda) d\lambda} \quad \text{and} \quad z_n(t) = e^{j2\pi\mu \int_0^t n(\lambda) d\lambda}. \quad (11.4.7)$$

Proof: Equation (11.4.6) is obtained by a direct substitution of Eq. (11.4.1) into Eq. (11.4.3). \square

Property 2. The Wigner-Ville Spectrum (WVS) of the signal $z_s(t)$ is given by

$$\text{WVS}_{z_s}(t, f) = E[W_{z_s}(t, f)] = \text{WVS}_{z_n}(t, f) * \int_f W_{z_x}(t, f), \quad (11.4.8)$$

where $E[\cdot]$ is the expectation operator, $\text{WVS}_{z_n}(t, f) = E[W_{z_n}(t, f)]$, and $*$ represents the convolution operation in the frequency domain.

Proof: This property follows from *Property 1* and the direct application of the expectation operator to the WVD of the encoded signal $z_s(t)$.

This latter is given by

$$E[W_{z_s}(t, f)] = \int_{-\infty}^{\infty} R_{z_n}(t, \tau) K_{z_x}(t, \tau) e^{-j2\pi f\tau} d\tau, \quad (11.4.9)$$

where the time-dependent autocorrelation function of $z_n(t)$ is

$$R_{z_n}(t, \tau) = E[z_n(t + \tau/2) z_n^*(t - \tau/2)] = E[e^{j2\pi\mu \int_{t-\tau/2}^{t+\tau/2} n(\lambda) d\lambda}], \quad (11.4.10)$$

and the time-dependent bilinear product function of $z_x(t)$ is

$$K_{z_x}(t, \tau) = z_x(t + \tau/2) z_x^*(t - \tau/2) = e^{j2\pi\mu \int_{t-\tau/2}^{t+\tau/2} x(\lambda) d\lambda}. \quad (11.4.11)$$

Equation (11.4.8) is then obtained by using the fact that the Fourier transform of a product in time is equivalent to the convolution in frequency. \square

Equation (11.4.8) shows that additive noise smears the encoded signal WVD, $W_{z_x}(t, f)$, through convolution. Therefore the bias of TFPF is dependent on the encoded signal $z_x(t)$ as well as the shape of the encoded noise spectrum $\text{WVS}_{z_n}(t, f)$. By restricting the shape of the encoded noise spectrum, a class of noise can be defined which does not introduce bias to the IF estimation. An example from this class is the WGN as will be seen next.

Property 3. The time-dependent autocorrelation function of the encoded noise, $R_{z_n}(t, \tau; \mu)$, is equal to the characteristic function of $q(t, \tau) = 2\pi \int_{t-\tau/2}^{t+\tau/2} n(\lambda) d\lambda$. That is,

$$R_{z_n}(t, \tau, \mu) = E[e^{j2\mu q(t, \tau)}] = \Phi_q(t, \tau, \mu), \quad (11.4.12)$$

where $\Phi_q(t, \tau, \mu)$ is the characteristic function of $q(t, \tau)$ defined by [48]

$$\Phi_q(t, \tau, \mu) = E[e^{j2\mu q(t, \tau)}] = \exp \left(\sum_{i=1}^{\infty} \frac{k_{qi}(t, \tau)(j\mu)^i}{i!} \right) \quad (11.4.13)$$

and $k_{qi}(t, \tau)$ is the i th cumulant of $q(t, \tau)$.

Proof: Equation (11.4.12) is obtained by forming the autocorrelation function of $z_n(t)$ in Eq. (11.4.7) and using the above definition of $q(t, \tau)$. \square

11.4.3 ACCURACY OF TFPF ESTIMATES

Equation (11.4.8) indicates that in general, a bias in IF estimation is introduced by the TFD of $z_x(t)$ (deterministic bias) and/or the noise (stochastic bias). In the case where the encoded signal $s(t)$ is composed of a deterministic signal $x(t)$ that is linear in time and embedded in stationary WGN $n(t)$, TFPF gives an unbiased estimate of the signal $x(t)$.

Proof: Consider the signal $s(t)$, given in Eq. (11.4.1), to be filtered using TFPF. If $n(t)$ is stationary WGN, the i th cumulant of $n(t)$ is such that $k_{ni} = 0$ for $i \geq 3$, and $q(t, \tau)$ is Gaussian with $k_{qi} = 0$ for $i \geq 3$. Furthermore, if the noise is a zero-mean independent process, i.e., if $R_n(\tau) = k_{n2}\delta(\tau)$, then [49, p. 369]

$$k_{q1}(t, \tau) = 0 \quad \text{and} \quad k_{q2}(t, \tau) = 4\pi^2|\tau|k_{n2}. \quad (11.4.14)$$

The characteristic function given in Eq. (11.4.13) becomes

$$\Phi_q(t, \tau, \mu) = e^{-2\pi^2\mu^2|\tau|k_{n2}}. \quad (11.4.15)$$

Taking the Fourier transform of this expression gives

$$E[W_n(f, t)] = \frac{4\pi^2k_{n2}\mu^2}{(2\pi^2k_{n2}\mu^2)^2 + (2\pi f)^2}. \quad (11.4.16)$$

This shows that the encoded noise spectrum is low-pass with a maximum at the frequency 0 Hz. Hence, WGN will not introduce any bias to the estimate of the IF. By substituting Eq. (11.4.16) in Eq. (11.4.8) we obtain

$$W_{z_s}(t, f) = W_{z_x}(t, f) * \frac{4\pi^2k_{n2}\mu^2}{(2\pi^2k_{n2}\mu^2)^2 + (2\pi f)^2}. \quad (11.4.17)$$

Equation (11.4.17) indicates that the bias in the IF estimate given by the peak of $W_{z_s}(f, t)$ could only be caused by $W_{z_x}(t, f)$. In the special case where the signal $x(t)$ is linear in time—that is, $x(t) = \alpha t + C$, where α and C are constants—Eq. (11.4.17) becomes

$$W_{z_s}(t, f) = \delta(f - x(t)) * \frac{4\pi^2k_{n2}\mu^2}{f(2\pi^2k_{n2}\mu^2)^2 + (2\pi f)^2} \quad (11.4.18)$$

$$= \frac{4\pi^2k_{n2}\mu^2}{(2\pi^2k_{n2}\mu^2)^2 + (2\pi f - 2\pi x(t))^2}. \quad (11.4.19)$$

The delta function ensures that the peak of this function occurs at $x(t)$. Therefore if the signal $x(t)$ is linear in time and embedded in stationary WGN, TFPF gives an unbiased estimate. This case is important because it allows the use of the piece-wise linear approximation. \square

Equation (11.4.17) shows that in the general case where the signal $x(t)$ is a nonlinear function of time, the WVD-based TFPF is biased, requiring an appropriate windowing of the data. The window is chosen such that the signal within this window behaves almost linearly [50] (see Section 11.4.4). Thus the WVD is replaced by the windowed WVD, also referred to as the pseudo-WVD. In the special

case where the signal $x(t)$ is a finite-order polynomial in time, the deterministic bias can be completely eliminated if the WVD is replaced by the polynomial WVD (PWVD) of an appropriate order, since the PWVD exhibits delta functions along the IF law for polynomial FM signals [51].

11.4.4 DISCRETE-TIME ALGORITHM FOR TFPF

The implementation of TFPF using the windowed WVD requires both signal scaling before encoding to prevent aliasing and the selection of the window length for reduced bias. These two aspects are discussed next.

11.4.4.1 Signal scaling

FM modulation of un-scaled discrete time signals can lead to aliasing which produces discontinuities in the estimated IF at the frequency boundaries of the time-frequency plane. This is avoided by amplitude scaling of the noisy signal before frequency encoding. Without loss of generality and unless otherwise specified, we assume that the signal $s(t)$ is sampled at a normalized sampling frequency of 1 Hz.

The scaled signal, $s_c(m)$, is obtained by using the following transformation:

$$s_c(m) = S[s(m)] = (a - b) \frac{s(m) - \min[s(m)]}{\max[s(m)] - \min[s(m)]} + b, \quad (11.4.20)$$

where $S[.]$ is the scaling operator and the parameters a and b , which satisfy the constraint $.5 \geq a = \max[s_c(m)] > b = \min[s_c(m)] \geq 0$, are chosen to provide suitable frequency limits on the encoded signal. The operators $\max[.]$ and $\min[.]$ are the maximizing and minimizing functions, respectively. The estimate of the desired signal, $\hat{x}(m)$, is recovered by an inverse scaling operation; that is

$$\hat{x}(m) = S^{-1}[\hat{x}_c(m)] = \frac{(\hat{x}_c(m) - b)(\max[s(m)] - \min[s(m)])}{a - b} + \min[s(m)], \quad (11.4.21)$$

where $\hat{x}_c(m)$ is the scaled signal obtained using TFPF on $s_c(m)$.

11.4.4.2 Reduced-bias window length selection

The bias-variance tradeoff is a key in the practical implementation of TFPF with the windowed WVD. Bias reduction requires a small window length to minimize the nonoptimal nature of the WVD for signals whose instantaneous phase is of degree higher than quadratic. On the other hand, variance reduction is achieved by increasing the window length. To reduce the variance of the estimate while maintaining bias performance it becomes necessary to increase the sampling rate. Thus there is a tradeoff between estimator bias and sampling frequency which results in a relationship between window length and bias, for a given sampling frequency. Results relating to TFPF window lengths are derived in [50]. The basic results for window length τ_w , sampling frequency f_s , and peak instantaneous frequency f_p are given below. For the case of signal estimation,

$$\tau_w \leq \frac{0.634f_s}{\pi f_p}. \quad (11.4.22)$$

This equation gives the maximum window length as a function of maximum signal IF and sampling frequency. In a typical application, a specified window length is required for a given SNR. The signal is sampled at a high enough rate to satisfy Eq. (11.4.22).

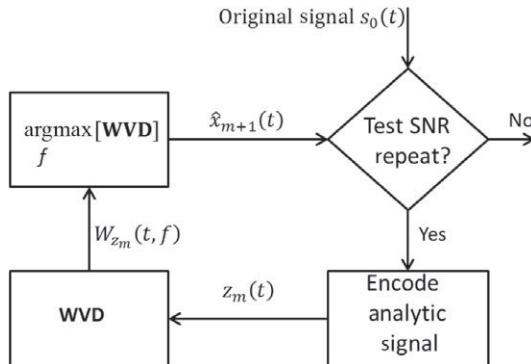
**FIGURE 11.4.1**

Illustration of the iterative TFPF scheme.

11.4.4.3 Iterative TFPF algorithm

The first time it is used, TFPF may not remove as much additive noise as desired. If this situation occurs, reapplication of the procedure to the filtered signal is recommended. This leads to a basic three-step iterative process:

1. Scale and encode the noisy signal.
2. Apply TFPF to yield a signal estimate $\hat{x}(t)$.
3. If $\hat{x}(t)$ contains substantial noise, go to step 1; else terminate the process.

The iterative process is also illustrated in Fig. 11.4.1. Simulations illustrate the convergence of the repeated scheme to a stable estimate $\hat{x}(t)$ [43].

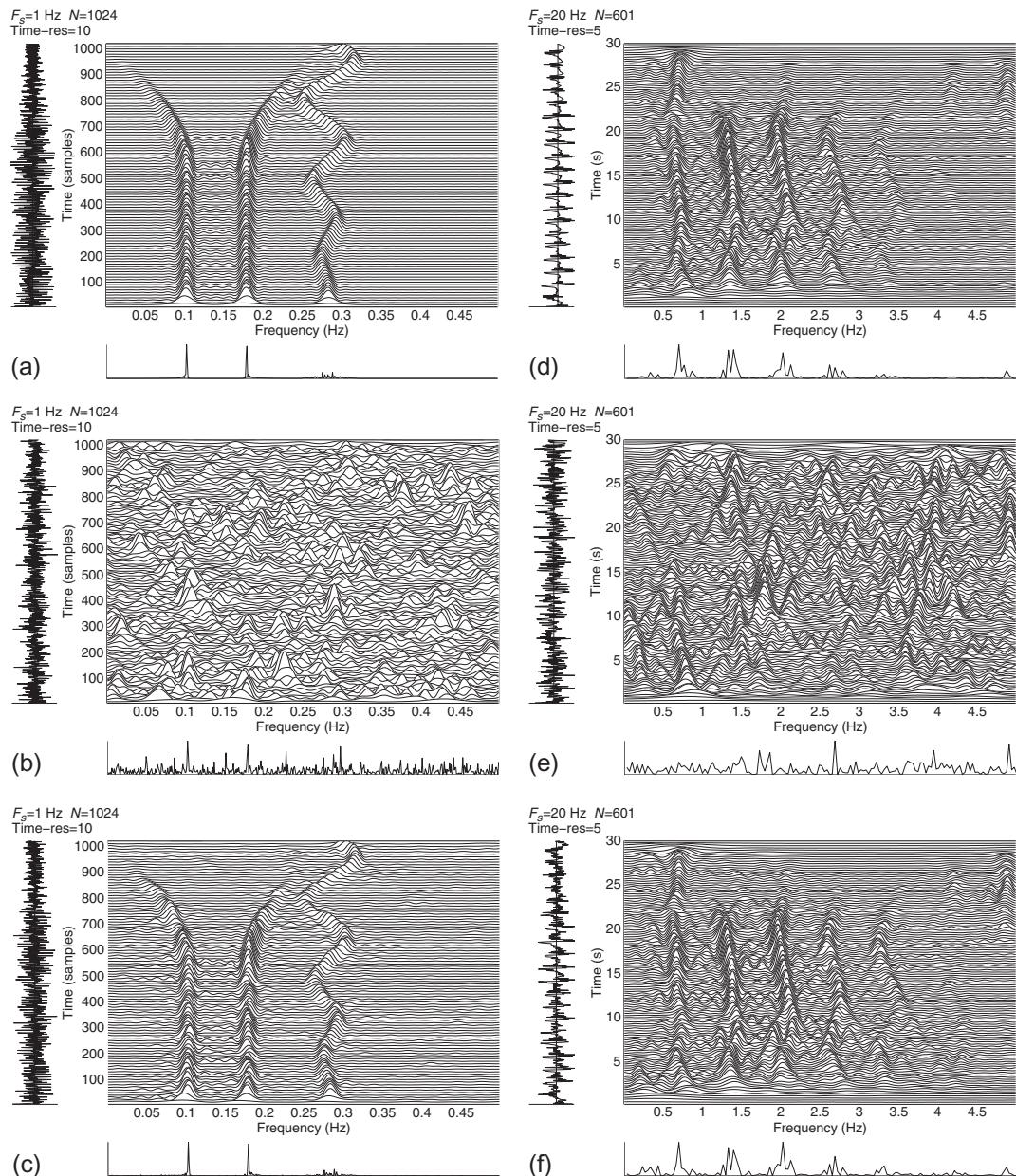
11.4.5 TFPF OF MULTICOMPONENT SIGNALS IN NOISE

Test case 1 (A multicomponent signal in WGN): Let us consider the multicomponent signal

$$x(m) = \begin{cases} 0.85 \sin(0.055m + 3.75 \times 10^{-4} \sin(0.000625m)m) + \sin(0.035m), \\ + \sin(0.020m), & 0 \leq m < L/2, \\ 0.85 \sin(0.055m + 3.75 \times 10^{-4} \sin(0.000625m)m), \\ + 2(1 - m/L) \sin(0.035m + 2.75 \times 10^{-11}(m - L/2)^3), \\ + 2(1 - m/L) \sin(0.020m - 2.75 \times 10^{-11}(m - L/2)^3), & L/2 \leq m < L, \end{cases}$$

where the data length L is taken as 32,768 data points. For a (t, f) illustration of this signal, the B-distribution (BD) with controlling parameter $\beta = 0.01$ is shown in Fig. 11.4.2(a).¹ WGN was added to the above signal giving an SNR of -9 dB; the BD of the noisy signal is shown in Fig. 11.4.2(b). The windowed WVD peak filter was then implemented to recover $x(m)$ from the noisy signal. A window length of 15 data points was chosen to satisfy the window length constraints given in Eq. (11.4.22).

¹See p. 92 for a definition of the BD and its parameter.

**FIGURE 11.4.2**

B-distributions of a synthetic signal (left column) and an EEG signal (right column), showing the original signal (top row), and the noisy signal before enhancement (middle row) and after enhancement (bottom row). (a) Original synthetic signal; (b) noisy synthetic signal; (c) enhanced synthetic signal; (d) original EEG signal; (e) noisy EEG signal; (f) enhanced EEG signal.

Figure 11.4.2(c) shows the clean recovery of the signal after three TFPF iterations. Note that the WVD is used as the vehicle for signal recovery while the B-distribution is used only for presentation of the results.

Test case 2 (Newborn EEG data in WGN): Fig. 11.4.2(d) shows a time-frequency representation of a real newborn EEG signal using the B-distribution with $\beta = 0.01$. WGN is then added to the signal at $SNR = -9$ dB. The noisy signal in Fig. 11.4.2(e) shows that the time-frequency patterns of the EEG signal are not clearly visible. Using a window length of 20 data points, four iterations of the TFPF were used to recover a cleaner signal. The filtered signal in Fig. 11.4.2(f) demonstrates the efficiency of TFPF.

11.4.6 SUMMARY AND CONCLUSIONS

TFPF is a signal enhancement tool, applicable to a large class of signals. The use of the windowed WVD based TFPF is used for reduced bias. The method is well suited to signals which can be represented as a sum of band-limited nonstationary processes in additive WGN. Testing on simulated and real data indicates that the method significantly enhances signals of this class by filtering out most of the additive noise. Further details of the time-frequency peak filtering method are provided in [50], and examples of applications can be found in [52–54].

11.5 SUBSPACE NOISE FILTERING USING TIME-FREQUENCY DISTRIBUTIONS⁰

11.5.1 INTRODUCTION

In developing a noise-reduction technique, keeping the essential structure of the original signal is often the main concern. For example, typical low-pass/high-pass filters assume that the noise is in the high-frequency/low-frequency regions and that the frequency bands of the noise and the clean signal are distinct. This assumption may not be acceptable in some conditions and may restrict the range of application of the filters. When noise or a noise-perturbed signal is nonstationary, signal enhancement becomes more challenging. In such cases, time-frequency distributions (TFDs) can be a suitable tool for noise attenuation while keeping the structure of the signal unchanged (see, e.g., [55–57] and Section 11.4).

In this section, additive noise in a nonstationary signal is reduced using a subspace process with several stages. First the noisy signal is mapped in the time-frequency (t, f) domain. Then the TFD is enhanced using the singular value decomposition (SVD) technique. Finally, the enhanced time series is produced from the enhanced TFD, resulting in an improved signal-to-noise ratio (SNR).

⁰Authors: **H. Hassanpour**, Shahrood University of Technology, Shahrood, Iran (h.hassanpour@shahroodut.ac.ir) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: A. Gholami, H. Ghassemian, and S. Ouelha.

11.5.2 THE NOISE-REDUCTION TECHNIQUE

Representing the signal in the (t, f) domain results in spreading any corruptive additive noise over the (t, f) plane. The SVD is applied to the TFD matrix in order to decompose the matrix into the signal subspace and the noise subspace. Figure 11.5.1(a) shows a TFD, namely a modified B-distribution (MBD) [55], of a signal containing two components (hyperbolic FM and linear FM (LFM)). The power spectrum density and the time domain representation of the signal are shown at the bottom and left sides of the figure, respectively. The TFD of the signal corrupted by white Gaussian noise (WGN) ($\text{SNR} = 0 \text{ dB}$) is shown in Fig. 11.5.1(b).

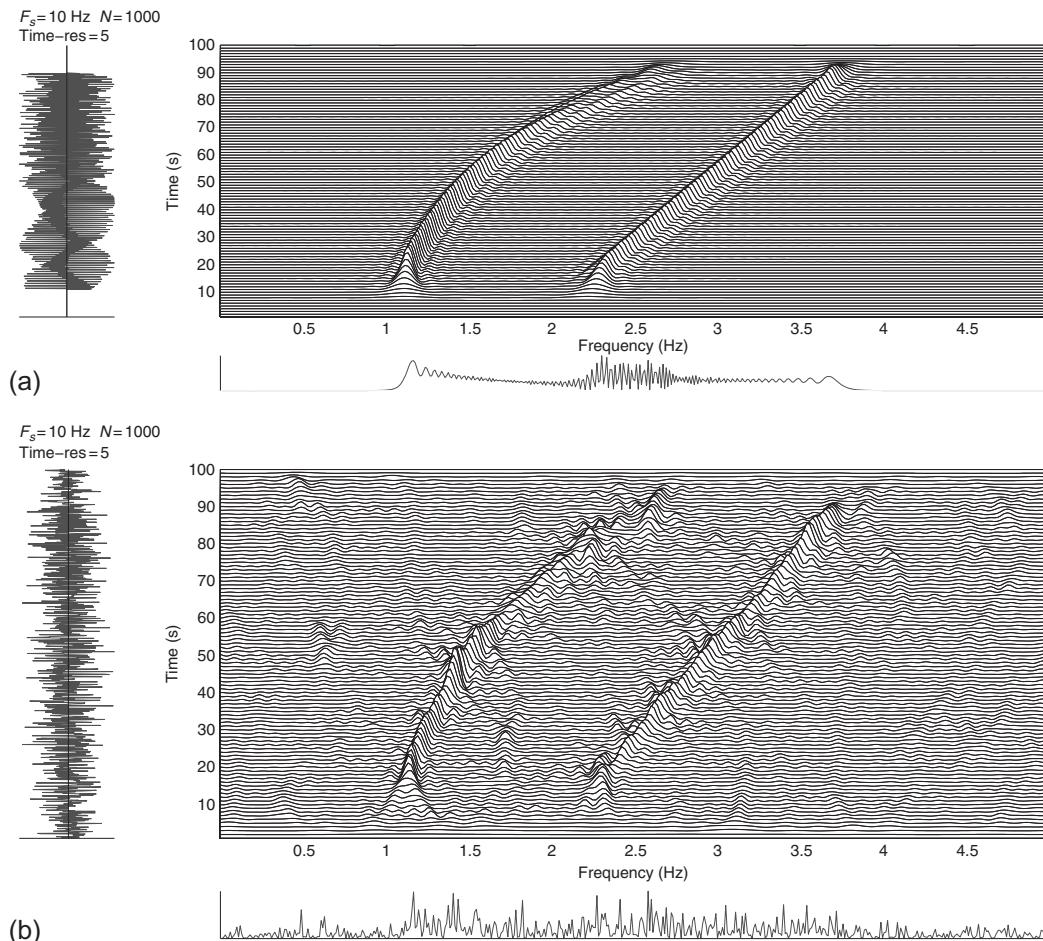


FIGURE 11.5.1

TFDs of a two-component signal with hyperbolic- and linear-FM components: (a) noiseless signal; (b) signal corrupted by white Gaussian noise ($\text{SNR} = 0 \text{ dB}$).

In the analysis of the noise influence, it is known that even in the case of white noise, the noise distribution over the (t,f) plane is not uniform, except in the WVD case. If the noise distribution over the (t,f) plane can be considered uniform, then the (t,f) plane is enhanced only by space division. To take into account the nonuniform effect of noise, we filter (repair) the singular vectors. The TFD matrix is decomposed into signal and noise subspaces based on the characteristics of its singular values. The noise is also reflected in the singular *vectors* (SVs) [56]. By restoring the SVs and reproducing the matrix therefrom, we can further enhance the TFD and the information embedded within it, since the SVs are the span bases of the matrix [58]. The resulting enhanced TFD can then be used for better analysis of the signal and for producing the signal with improved SNR.

11.5.2.1 Singular value decomposition

The SVD of a $P \times Q$ matrix ρ is a decomposition into the form

$$\rho = \mathbf{U}\Sigma\mathbf{V}^T, \quad (11.5.1)$$

where $\mathbf{U}(P \times P)$ and $\mathbf{V}(Q \times Q)$ are orthonormal matrices, Σ is a $P \times Q$ diagonal matrix of singular values with components $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ii} > 0$, and the superscript T indicates the transpose. Furthermore, there exist nonunique matrices \mathbf{U} and \mathbf{V} such that $\delta_{11} \geq \delta_{22} \geq \dots \geq 0$. The columns of the orthonormal matrices \mathbf{U} and \mathbf{V} are called the left and right SVs respectively. An important property of \mathbf{U} and \mathbf{V} is that they are mutually orthogonal [58].

11.5.2.2 Noise subspace subtraction

Following Eq. (9.107) in [47], under the assumption of additive noise ϵ corrupting a signal \mathbf{s} resulting in a noisy signal \mathbf{x} , the TFD¹ of the noisy signal can be written

$$\rho_x = \rho_s + \rho_\epsilon, \quad (11.5.2)$$

where ρ_s and ρ_ϵ are respectively associated with the TFD of the clean signal and the additive WGN. Mathematically, the subspace separation of the noisy signal can be expressed as [59]

$$\rho_x = \mathbf{U}\Sigma\mathbf{V}^T = (\mathbf{U}_s\mathbf{U}_\epsilon) \begin{bmatrix} \Sigma_s & \mathbf{0} \\ \mathbf{0} & \Sigma_\epsilon \end{bmatrix} \begin{pmatrix} \mathbf{V}_s^T \\ \mathbf{V}_\epsilon^T \end{pmatrix}. \quad (11.5.3)$$

Then

$$\rho_x = \mathbf{U}_s\Sigma_s\mathbf{V}_s^T + \mathbf{U}_\epsilon\Sigma_\epsilon\mathbf{V}_\epsilon^T, \quad (11.5.4)$$

where Σ_s and Σ_ϵ respectively represent the singular values that belong to the clean signal subspace and the noise subspace. Equations (11.5.2) and (11.5.4) yield

$$\rho_s = \mathbf{U}_s\Sigma_s\mathbf{V}_s^T; \rho_\epsilon = \mathbf{U}_\epsilon\Sigma_\epsilon\mathbf{V}_\epsilon^T. \quad (11.5.5)$$

As can be deduced from the above equations, we must determine a threshold point in the Σ matrix such that singular values lower than that point can be categorized as the singular values of the noise subspace.

¹Note that this relation is exact for the STFT; for quadratic TFDs such as the MBD, it is approximate as the right-hand side part would normally need to include the cross-TFD to have exact equality. In this section, the terminology adopted is that the second element ρ_ϵ includes both the noise TFD auto-terms and the cross-terms.

To determine this point, the singular values of matrix Σ for the given noisy signal (Fig. 11.5.1(a)) are plotted with respect to their indices in Fig. 11.5.2. In the left-hand side graph, a break point can be seen where the slope of the curve changes drastically. Thus the threshold can be determined by calculating the derivative of the curve (right-hand side graph) and finding the place of the maximum change between consecutive derivative values. Since the noise subspace is mainly related to those singular values that are lower than this threshold, one can set these singular values to zero for space division.

11.5.2.3 Enhancing the singular vectors

Figure 11.5.3 shows the first right SV of the TFD for the original and noisy signals represented in Fig. 11.5.1. It confirms that the existence of noise in the signal is reflected in the SVs related to the TFD of the signal, because singular vectors are the span bases of the matrix. As noted above, noise in a signal affects both the singular values and the singular vectors of the data (t, f) matrix. Hence, if we merely filter the singular values, some noisy data will still be available in the signal subspace. Consequently, the SVs also need to be restored for further noise reduction. In this restoration, the SVs are treated as time-series. To reduce the effect of noise, we use the Savitzky-Golay smoothing filter [60]. In this approach, a polynomial of degree d is fitted to n_w consecutive data points from the time-series, where n_w is the frame or window size. Filtered singular vectors can be obtained as follows:

$$\mathbf{U}_e^i = \Gamma(\mathbf{U}^i), \quad i = 1, \dots, P; \quad \mathbf{V}_e^i = \Gamma(\mathbf{V}^i), \quad i = 1, \dots, Q, \quad (11.5.6)$$

where $\Gamma(\cdot)$ is the Savitzky-Golay smoothing filter. The SV represented in Fig. 11.5.3 has been filtered using the Savitzky-Golay filter and (alternatively) a finite-impulse response (FIR) filter, and the results

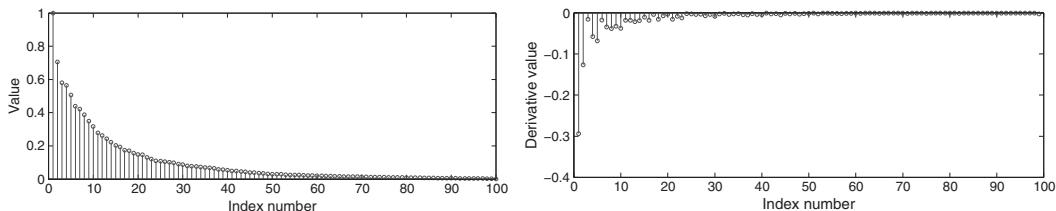


FIGURE 11.5.2

Normalized singular values (left) and their derivatives (right) for the TFD in Fig. 11.5.1(b).

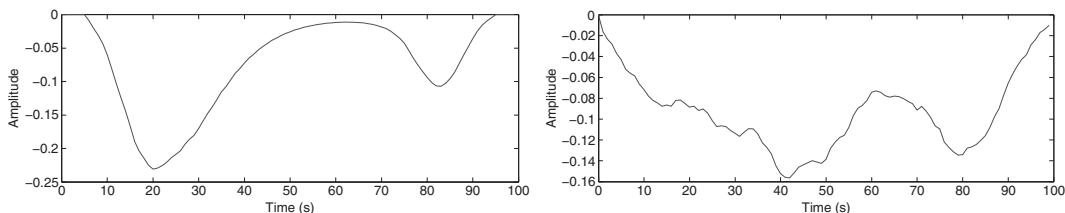
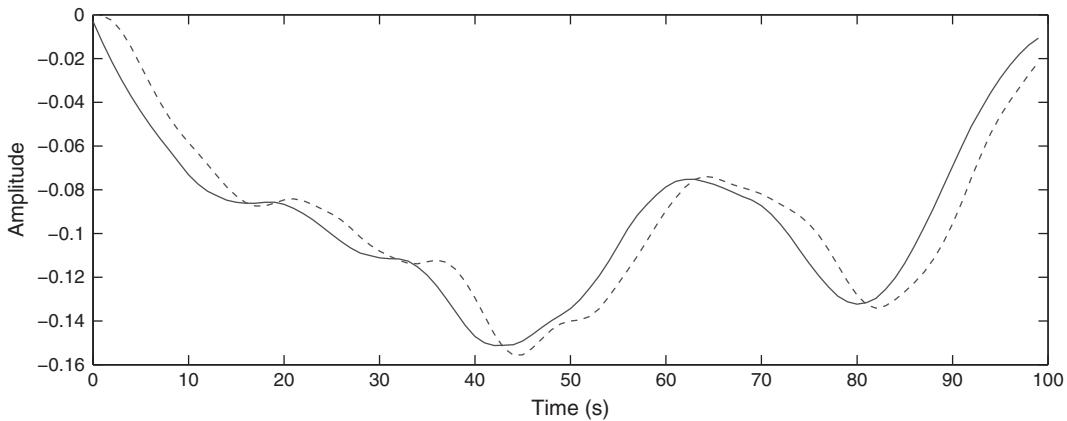


FIGURE 11.5.3

The first right SV related to the original signal (left) and the noisy signal (right).

**FIGURE 11.5.4**

The results of filtering the first right SV of the matrix associated with Fig. 11.5.1(b) using the Savitzky-Golay filter (solid line) and the FIR filter (dashed line).

are shown in Fig. 11.5.4. This is achieved by fitting successive subsets of adjacent data points with a low-degree polynomial into Savitzky-Golay filtering. Comparison of the graphs shows that the Savitzky-Golay filter has no shifting effect on the filtered SV. The Euclidian distance between the SVs related to the original and noisy signals is 0.6131; and the distance between the original and filtered SVs is 0.5267 for the Savitzky-Golay filter and 0.6547 for the FIR filter. Note that the higher Euclidian distance for the FIR filter is due to the shifting effect and not adding further noise to the SV. This shifting effect may damage the structure of the original pattern in the (t, f) plane.

11.5.2.4 Reconstruction of the TFD

The filtered SVs in \mathbf{U} and \mathbf{V} are used to reconstruct the enhanced TFD using the following process. Let ρ_s be the $P \times Q$ strong interference subspace containing the signal of interest as in Eq. (11.5.5). This matrix can be decomposed using SVD as

$$\rho_s = \mathbf{U}_s \Sigma_s \mathbf{V}_s^T, \quad (11.5.7)$$

where

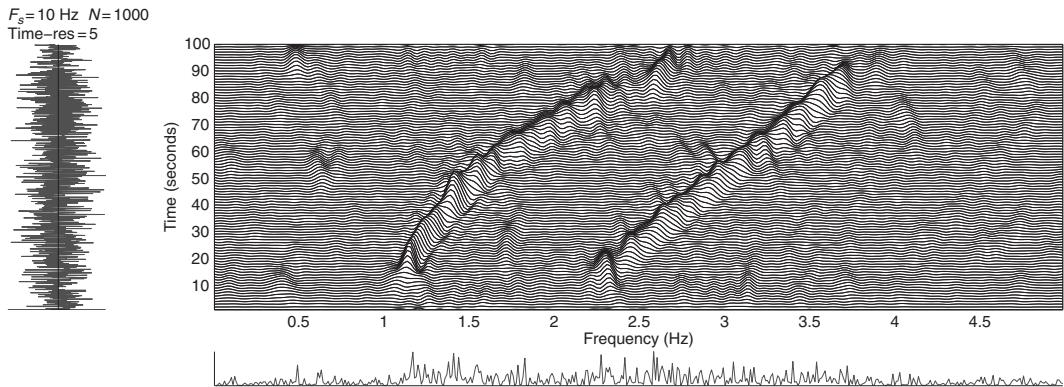
$$\mathbf{U}_s = [U_{s_1}, U_{s_2}, \dots, U_{s_p}] \in R^{P \times P} \quad \text{and} \quad \mathbf{V}_s = [V_{s_1}, V_{s_2}, \dots, V_{s_Q}] \in R^{Q \times Q}.$$

The filtered version of SVs is defined by

$$\mathbf{U}'_s = [U'_{s_1}, U'_{s_2}, \dots, U'_{s_p}] \in R^{P \times P} \quad \text{and} \quad \mathbf{V}'_s = [V'_{s_1}, V'_{s_2}, \dots, V'_{s_Q}] \in R^{Q \times Q},$$

where $U'_{s_p} = \Gamma(U_{s_p})$, $p = 1, 2, \dots, P$ and $V'_{s_q} = \Gamma(V_{s_q})$, $q = 1, 2, \dots, Q$. Matrices \mathbf{U}'_s and \mathbf{V}'_s are normalized as is necessary for SVs; the normalized matrices are called \mathbf{U}_r and \mathbf{V}_r , respectively. The noise-reduced TFD is then obtained as

$$\rho_{X_r} = \mathbf{U}_r \Sigma_r \mathbf{V}_r^T. \quad (11.5.8)$$

**FIGURE 11.5.5**

The enhanced TFD plot related to Fig. 11.5.1(b).

The above-mentioned (t,f) enhancement process applied to the matrix associated with Fig. 11.5.1(b) yields the result in Fig. 11.5.5. In the enhanced (t,f) plot, the background noise has been reduced and the pattern is more clearly recognizable.

11.5.2.5 Time-frequency signal synthesis

Time-frequency signal synthesis is the estimation of a signal using specifications of its (t,f) energy distribution. In this estimation, the signal is generated such that a given TFD of the signal is the best approximation for the desired (t,f) function.

Algorithms for synthesizing a signal from its TFD are described in [47,61]. The TFD used here is the WVD [47]; the algorithm is briefly described below.

Let $\rho_x(t,f)$ represent the WVD of a signal $x(t)$. The aim is to find the $\hat{x}(t)$ whose WVD, denoted by $\rho_{\hat{x}}(t,f)$, is closest to the desired (t,f) function $\rho_x(t,f)$. In other words, we wish to minimize the error given by

$$e(x) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\rho_x(t,f) - \rho_{\hat{x}}(t,f)| df dt. \quad (11.5.9)$$

The above minimization allows for the even and odd samples of the signal, denoted by \hat{x}_e and \hat{x}_o respectively, to be obtained [61]. The samples are equal to the normalized principal eigenvectors of the matrices \mathbf{M}_e and \mathbf{M}_o whose elements are

$$\mathbf{M}_e(i+1,j+1) = g(i+j, i-j) + g^*(i+j, j-i) \text{ for } i,j = 0, \dots, L_e - 1, \quad (11.5.10)$$

$$\mathbf{M}_o(i,j) = g(i+j+1, i-j) + g^*(i+j+1, j-i) \text{ for } i,j = 0, \dots, L_o, \quad (11.5.11)$$

where $g(p,q)$ is the discrete inverse Fourier transform of $\rho_x(t,f)$, and L_e and L_o are the lengths of \hat{x}_e and \hat{x}_o respectively. The phase of the estimated signal can be set using the original signal by initially computing

$$\phi_e = \arg \left\{ \sum_{p=0}^{L_e-1} x(2p) \hat{x}_e^*(p) \right\}, \quad (11.5.12)$$

$$\phi_o = \arg \left\{ \sum_{p=0}^{L_o} x(2p-1) \hat{x}_o^*(p) \right\}, \quad (11.5.13)$$

then applying the assignments $\hat{x}_e(p) = \hat{x}_e(p) e^{j\phi_e}$ and $\hat{x}_o(p) = \hat{x}_o(p) e^{j\phi_o}$.

11.5.3 PERFORMANCE EVALUATION

To show the performance of the (t,f) -based noise-reduction approach, several experiments have been carried out on both synthetic and real signals, and the results have been compared with those obtained using the Wiener filter, as follows.

11.5.3.1 Reduction of noise in time series

The (t,f) -based technique is applied to a number of stationary (sinusoidal) and nonstationary (LFM) signals in WGN with different SNRs. The frequency of the stationary signal in each experiment is randomly chosen between 0.1 and 0.4 Hz. The frequency of the LFM signal varies from 0.1 to 0.4 Hz. The number of samples in each realization of the stationary or nonstationary signal is set to 512 with unit sampling frequency. To assess the performance of the proposed technique, we use the normalized mean-squared error (MSE) to compare the enhanced signal with the original noise-free signal. Table 11.5.1 shows the comparison results of 100 realizations of stationary and nonstationary signals with differing SNRs. These results indicate that the (t,f) -based technique has a better performance in noise reduction. The results also show that the performance of the Wiener filter drops considerably as SNR is reduced.

The Wiener filter is able to reduce noise in a signal if the SNR is high (higher than 4). When the SNR of a signal is low, the Wiener filter may just transform the noise from one form to another [62] (see Fig. 11.5.6(b)). This issue limits the applicability of the Wiener filter. In the (t,f) -based approach,

Table 11.5.1: Comparison Results of the Two Noise-Reduction Techniques on Sinusoidal and LFM Signals in White Gaussian Noise

| SNR | Sinusoid | | | | LFM | | | |
|-----|----------|--------|---------|--------|--------|--------|---------|--------|
| | Wiener | | (t,f) | | Wiener | | (t,f) | |
| | Mean | Var. | Mean | Var. | Mean | Var. | Mean | Var. |
| 0 | 1.431 | 0.0002 | 0.292 | 0.0090 | 1.682 | 0.0088 | 0.635 | 0.045 |
| 2 | 0.693 | 0.0023 | 0.231 | 0.0063 | 1.409 | 0.0022 | 0.561 | 0.0066 |
| 4 | 0.421 | 0.0019 | 0.202 | 0.0075 | 1.210 | 0.0018 | 0.523 | 0.0034 |
| 6 | 0.214 | 0.006 | 0.165 | 0.0035 | 0.751 | 0.0054 | 0.401 | 0.0098 |

These results show the normalized mean square error and variance of the enhanced signals compared with the original noise-free signals.

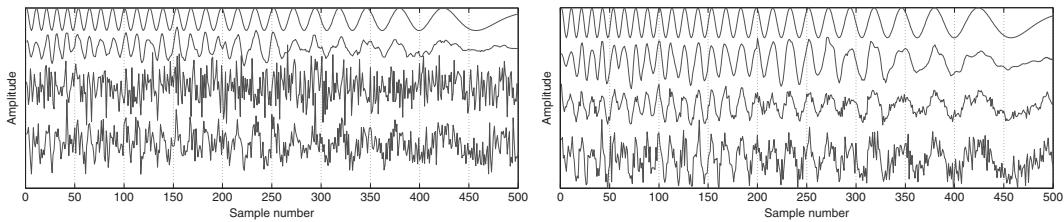


FIGURE 11.5.6

Comparing performance of noise-reduction techniques on LFM signals with $\text{SNR} = 1$ (left) and $\text{SNR} = 5$ (right); from top to bottom: the clean signal; output of the (t, f) -based technique; output of the Wiener filter; the signal in white Gaussian noise.

since all SVs from the signal subspace are employed, the structure of the original patterns in the (t, f) plane may not be damaged. By filtering all SVs and employing them in producing the enhanced TFD, the basic structure of the original pattern is kept.

Figure 11.5.6 shows the results of the experiment for two LFM signals with different SNRs. As can be seen, the (t, f) method considerably reduces the noise in the enhanced signals, whereas the Wiener filter may be less suitable than (t, f) methods for enhancing signals with low SNR, especially nonstationary signals.

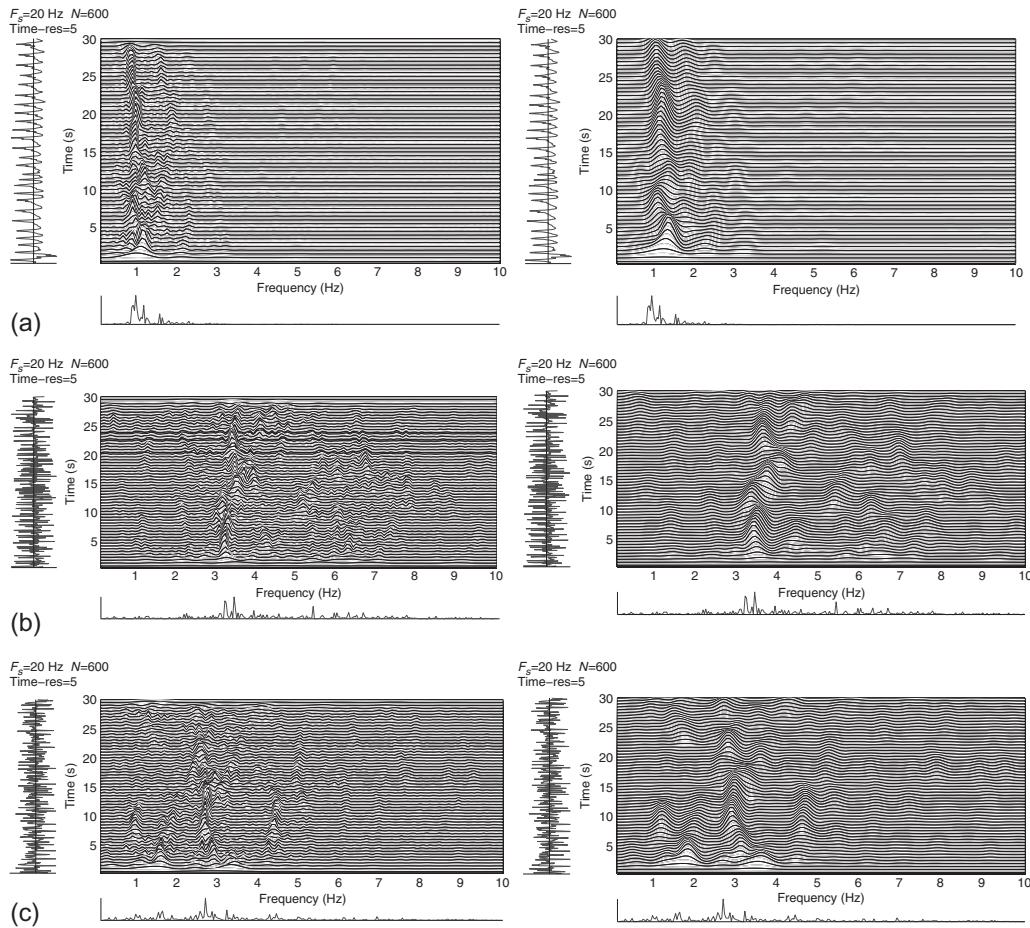
The (t, f) approach can be employed for both (t, f) signal enhancement and time-series signal enhancement. There is no restriction in choosing a TFD for the (t, f) approach; however, in time-series noise reduction, only an invertible TFD, such as the WVD, can be used, because the time-series must be computed from the TFD of the signal. (On the invertibility of the WVD, see Eq. (3.1.13) on p. 105.)

11.5.3.2 Reduction of noise in TFDs

As an application to real-life signals, it is desired to enhance the TFDs of three neonate EEG signals shown in Fig. 11.5.7. The first example (a) shows seizure activity (see [20,63] for more details on (t, f) EEG analysis for neonatal seizure detection). For each signal, the result of applying the noise-reduction process is shown in the right-hand side (t, f) plot. As can be seen, the TFDs are enhanced by showing more details, and the (t, f) patterns are made more easily recognizable. These examples clearly show the capability of (t, f) techniques in reducing noise in multicomponent nonstationary signals.

11.5.4 SUMMARY AND CONCLUSIONS

This section introduces a (t, f) -based approach for noise reduction that complements the one in Section 11.4. In this approach, the noise subspace is first excluded from the signal subspace in the (t, f) domain using the singular value decomposition technique. The singular vectors of the TFD matrix are then filtered using the Savitzky-Golay smoothing filter. Results show that applying the technique on the (t, f) plane may cause no changes to the basic structure of the patterns embedded in the signal. Any TFD may be chosen for (t, f) approach; but if the time series is to be reconstructed, the TFD must be invertible.

**FIGURE 11.5.7**

Using the (t, f) approach to noise reduction: original (left) and noise-reduced (right) (t, f) plots of three newborn EEG signals, labeled (a)–(c); for each signal, the time-domain plot is at the far left, and the spectrum is shown under each (t, f) plot.

11.6 EVALUATION OF TIME-FREQUENCY DENOISING ALGORITHMS FOR SPEECH ENHANCEMENT⁰

11.6.1 SIGNAL ENHANCEMENT BY TIME-FREQUENCY DENOISING

In many applications, such as geophysics, mobile communications or speech recognition [64], efficient signal enhancement techniques are needed. In mobile communications, the signal of interest is often

⁰Authors: **S. Ouelha**, Qatar University, Doha, Qatar (samir.ouelha@qu.edu.qa) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: A. Aïssa-El-Bey, V. Chandran, and J. O'Toole.

speech and it often arises from conversations that take place in noisy and nonstationary environments such as inside a car, in the street, or inside airports. Hence, methods are needed for improving the quality and intelligibility of such speech conversations by removing or reducing the effect of noise in such situations. Two time-frequency (t,f) denoising algorithms for speech enhancement are described and compared with standard methods. Although the methods' performances are illustrated on speech signals, they are in fact applicable to any other signals, including underwater acoustic signals (Section 14.7).

11.6.2 ESTIMATION OF NONSTATIONARY NOISE POWER SPECTRUM

The speech denoising algorithms require an estimate of the noise power level; the methods selected here use two such estimators as described below.

11.6.2.1 *The median absolute deviation estimator*

The median absolute deviation (MAD) [65] is a measure of the variance of a data sample; this estimator is popular because of its robustness with respect to outliers. Let y be the noisy observation of the signal, then the MAD is defined by

$$\hat{\sigma} = C \times \text{median}(|y - \text{median}(y)|), \quad (11.6.1)$$

where the median m is defined such that $P(y \leq m) \geq \frac{1}{2}$ and $P(y \geq m) \geq \frac{1}{2}$; and C is a constant scale factor, which depends on the probability density function (PDF) of the observation. For a Gaussian distribution, C equals to $1/\Phi^{-1}(0.75) = 0.6745^{-1}$ where Φ is the cumulative distribution function (CDF) of the Gaussian distribution [65]. The estimation of the noise variance is more robust than the classical moment based on the standard deviation estimator. To account for the lack of prior knowledge, we assume that the signal probabilities of occurrence are less than half.

11.6.2.2 *The extended-DATE*

The d-dimensional amplitude trimmed estimator (DATE) [66] was originally introduced for additive WGN power spectrum estimation. The purpose of DATE is to estimate the noise standard deviation. It performs the trimming by assuming that the signal norms are above some known lower bound and that the signal probabilities of occurrence are less than half. This was extended to the scenario of nonstationary noise in [67]. The key idea is that in each frequency bin and within a sufficiently short time period, the noise instantaneous power spectrum can be considered approximately constant. Then, it is possible to estimate the variance of complex Gaussian noise from noisy signal observations.

11.6.3 GENERAL PROCESS FOR SPEECH ENHANCEMENT ALGORITHMS

Let $y[n]$ be an N -dimensional vector of recorded data, representing the superposition of the useful signal $s[n]$ and noise $\epsilon[n]$, so that for each sample $n = 1, \dots, N$, we have:

$$y[n] = s[n] + \epsilon[n], \quad (11.6.2)$$

where the noise is assumed to be additive, uncorrelated, and locally stationary. When noise is nonstationary, time-frequency distributions (TFDs) are a natural tool for denoising. The simplest TFD

is the spectrogram, which is the square modulus of the short-time Fourier transform (STFT) (see Section 2.3.1). Applying the discrete-time STFT on both sides of Eq. (11.6.2) yields

$$F_y^w[l, k] = F_s^w[l, k] + F_\epsilon^w[l, k], \quad (11.6.3)$$

where $l = 0, \dots, L - 1$ is the frame index and $k = 0, \dots, K - 1$ is the frequency channel.

11.6.3.1 A priori signal-to-noise ratio estimator

Let us define a general process for denoising using TFD. It is useful to consider a gain function $G[l, k]$ ¹ defined as $G[l, k] = F_s^w[l, k]/F_y^w[l, k]$; this function will depend on SNR_{post} and $\text{SNR}_{\text{prior}}$ defined as

$$\text{SNR}_{\text{post}} = \frac{S_y^w[l, k]}{E\{S_\epsilon^w[l, k]\}}, \quad (11.6.4)$$

$$\text{SNR}_{\text{prior}} = \frac{E\{S_s^w[l, k]\}}{E\{S_\epsilon^w[l, k]\}}, \quad (11.6.5)$$

where $E\{\cdot\}$ denotes the expected value. Generally, the estimation of the $\text{SNR}_{\text{prior}}$ and of the SNR_{post} is realized on each frequency bin independently; however, in practice, the noise power spectral density (PSD) and the $\text{SNR}_{\text{prior}}$ are unknown, as the noisy speech alone is available. Note that in practice, the local PSD is used (over a sliding window where the assumptions are respected) as an estimate of the noise spectrogram $S_\epsilon^w[l, k]$. There are well-known methods to estimate $\text{SNR}_{\text{prior}}$ such as the decision-directed (DD) estimator, for example:

$$\widehat{\text{SNR}}_{\text{prior}}^{\text{DD}}[l, k] = \alpha \widehat{\text{SNR}}_{\text{prior}}^{\text{DD}}[l - 1, k] + (1 - \alpha) \max(\text{SNR}_{\text{post}} - 1, 0). \quad (11.6.6)$$

Where α ($0 \leq \alpha \leq 1$) is a weighting factor. There are other estimators of $\text{SNR}_{\text{prior}}$ (e.g., [68, 69]).

11.6.3.2 Attenuation rules

Methods that account for attenuation rules or gain functions $G[l, k]$ include the Wiener estimate, the minimum mean-squared error log-spectral amplitude (MMSE LSA) estimator [70], and the maximum *a posteriori* estimator [71]. The expressions of these estimators are given below:

$$G_{\text{Wien}}[l, k] = \frac{\text{SNR}_{\text{post}}[l, k] - 1}{\text{SNR}_{\text{post}}[l, k]} \quad (11.6.7)$$

$$G_{\text{MMSE LSA}}[l, k] = \frac{\text{SNR}_{\text{prior}}[l, k] - 1}{1 + \text{SNR}_{\text{prior}}[l, k]} e^{0.5 \int_v^\infty \frac{e^{-t}}{t} dt}, \quad (11.6.8)$$

where $v = \frac{\text{SNR}_{\text{prior}}}{1 + \text{SNR}_{\text{prior}}} \text{SNR}_{\text{post}}$

$$G_{\text{MAP}}[l, k] = u + \sqrt{u^2 + \frac{v}{2\text{SNR}_{\text{post}}[l, k]}}, \quad (11.6.9)$$

¹This notation G like “gain” is used here for convenience; although related, it should not be confused with the kernel G defined in Chapters 2 and 3. The tenuous link is due to the relationship between WVD and spectrogram.

where $u = 0.5 - \frac{v}{4\sqrt{\text{SNR}_{\text{post}}\text{SNR}_{\text{prior}}}}$.

The above functions result from different assumptions. The best combination between $\text{SNR}_{\text{prior}}$ estimator and attenuation rule is data- and application-dependent.

11.6.3.3 General speech enhancement process

This algorithm for speech enhancement is described by the following six steps:

1. Compute the STFT of the noisy signal
2. Estimate the noise power
3. Estimate SNR_{post} and $\text{SNR}_{\text{prior}}$
4. Estimate $G(\text{SNR}_{\text{post}}, \text{SNR}_{\text{prior}})$
5. Apply the gain function G to estimate the useful signal STFT $F_{\tilde{s}}^w[l, k] = F_y^w[l, k]G[l, k]$
6. Synthesize the useful signal $\tilde{s}[n]$ from $F_{\tilde{s}}^w$ using the inverse STFT

11.6.4 AN ADVANCED SONOGRAM USING MEL FILTERS FOR SPEECH ENHANCEMENT: THE DENOISED HEARINGOGRAM

11.6.4.1 Principle

The essence of this algorithm [72] is to combine a sonogram based on the Mel filterbank that accounts for human perception with a wavelet shrinkage method [65]. Let us present the denoising principle that allows the construction of the denoised Hearingogram in the following five steps.

1. Data filtering with the Mel filterbank
2. Threshold values estimation (one set of thresholds per y_{h_m})
3. Wavelet coefficients thresholding
4. Multiresolution analysis from thresholded coefficients to get useful signal approximation \tilde{s}_{h_m}
5. Denoised Hearingogram construction taking for each m value the square magnitude of \tilde{s}_{h_m}

The first step is the computation of y_{h_m} for $m = 1, \dots, M$, where M is the number of filters and y_{h_m} represents the convolution product between the m th Mel filter h_m and the observation y defined in Eq. (11.6.2), such as

$$y_{h_m} = y * h_m. \quad (11.6.10)$$

The aim of this method is to estimate s_{h_m} (denoted \tilde{s}_{h_m}) where s_{h_m} is such that

$$y_{h_m} = s_{h_m} + \epsilon_{h_m}. \quad (11.6.11)$$

The second step consists of defining thresholds for the wavelet coefficients. To do this, let us consider Gaussian unitary power white-noise samples, obtained from random simulations, filtered by the Mel filterbank. Since each Mel filter is band-pass, filtered white noise is colored according to each filter bandwidth. Then, multiresolution analysis is applied on each filtered white noise, allowing us to obtain a set of wavelet coefficients ζ^p depending on the m th Mel filter and the p th coarse level. Finally, the standard deviation of ζ_m^p , $\sigma_{\zeta}^m[p]$, is estimated. The next step is to estimate the noise power σ_{ϵ_m} in

y_{h_m} . This step can be achieved using either the MAD or the extended d -dimensional amplitude trimmed estimator (E-DATE) defined above. Then the universal threshold, which is a simple entropy measure solely dependent on the number of samples N in y_{h_m} , provides the threshold applied to the wavelet coefficients. It is given by [65]

$$\lambda_m[p] = \sigma_{\epsilon_m} \sigma_{\zeta}^m[p] \sqrt{2 \ln(N)}. \quad (11.6.12)$$

Assuming that the analyzed signal contains noise, the universal threshold must be further adjusted. As the entropy in a noisy signal is higher than in the signal without noise, the threshold should be multiplied by σ_{ϵ_m} , giving the threshold value $\lambda_m[p]$, presented in Eq. (11.6.12). So for the m th Mel filter, we have a nonconstant threshold depending on the p th coarse level and σ_{ϵ_m} . Applying multiresolution analysis [73] on y_{h_m} yields the wavelet coefficients. Let ω_m^p be the wavelet coefficients associated with the multiresolution analysis of y_{h_m} for the p th coarse level. The useful signal is handled by large wavelet coefficients, whereas the noise is distributed across small wavelet coefficients [74]. For this reason, y_{h_m} denoising can be performed using a thresholding step on the wavelet coefficients. The hard-thresholding is used to perform this step; it is defined by

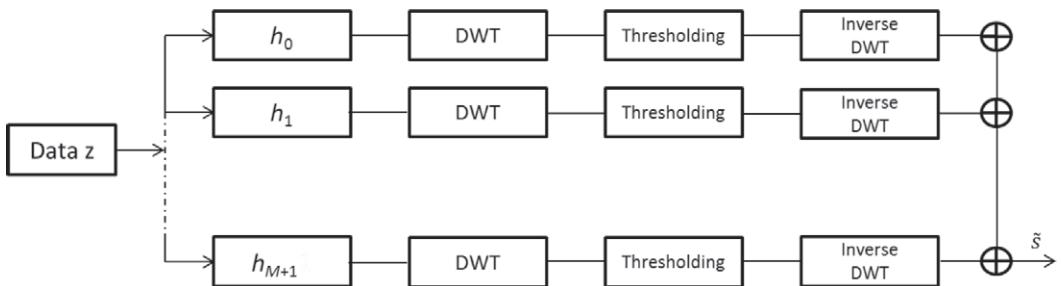
$$\omega_m^p[k] = \begin{cases} 0, & \text{if } |\omega_m^p[k]| < \alpha \lambda_m[p], \\ \omega_m^p[k], & \text{otherwise} \end{cases} \quad (11.6.13)$$

with k running from 1 to $\frac{N}{2^p}$. The constant α is a parameter used to adjust the threshold; the smaller the SNR gain and the higher α , the stronger the useful signal degradation becomes. Then, multiresolution synthesis is used on the thresholded coefficients to provide \tilde{s}_{h_m} . Finally, the stack of the squared magnitudes of the \tilde{s}_{h_m} constitutes the denoised Hearingogram.

11.6.4.2 Algorithm using the denoised Hearingogram

Figure 11.6.1 shows the flowgraph of the denoised Hearingogram algorithm. This algorithm can be summarized as follows:

1. Initialize an N -dimensional vector \tilde{s} to zero
2. For $m = 0, \dots, M + 1$:
 - compute the impulse response h_m
 - determine y_{h_m} by taking the product convolution between y and h_m , i.e., $y_{h_m} = y * h_m$
 - apply multiresolution of y_{h_m} using the DWT
 - apply wavelet coefficients thresholding
 - construct \tilde{s}_{h_m} by applying an inverse DWT
 - perform an iterative construction of \tilde{s} such that: $\tilde{s} = \tilde{s} + \tilde{s}_{h_m}$
3. End for loop

**FIGURE 11.6.1**

Flowchart of the denoising process in five stages: (1) Hearingogram construction; (2) multiresolution analysis; (3) thresholding; (4) inverse multiresolution; (5) synthesis.

11.6.5 DIRECTIONAL NOISE REDUCTION

11.6.5.1 Principle

In order to enhance the speech signal, this algorithm [75] applies different masks on the STFT, with different sizes and shapes. Each masked STFT yields a new TFD, on which a hard-thresholding is applied. Then, each TFD is merged, using the maximum operator in order to create a gain function to denoise the TFD and finally synthesize the signal from this denoised TFD.

11.6.5.2 Proof of concept and description

The noise is assumed locally centered and wide-sense stationary (WSS) Gaussian. In the (t, f) domain, for each frequency channel k , the noise Fourier coefficients $F_\epsilon^w[l, k]$ are circular symmetric complex Gaussian random variables, independent of $F_s^w[l, k]$. The normality is justified, as in the case of the denoised Hearingogram, by the central limit theorem applied to each Fourier coefficient. Each mask is constructed to match a special pattern in the (t, f) domain. The masks were designed using a weighted average, where the weights correspond to a normalized 2D or 3D Gaussian window. Unfortunately, the presence of a high signal component can bias the result of the convolution. To avoid this problem, the TFD is transformed on a time-frequency probability representation (TFPR). First, a scaling step is applied on the TFD coefficients to whiten $F_\epsilon^w[l, k]$, so for each channel frequency, the following scale operation is applied:

$$S_{y_{sc}}^w[l, k] = 2 \frac{S_y^w[l, k]}{E\{S_\epsilon^w[k]\}}, \quad (11.6.14)$$

where $E\{S_\epsilon^w[k]\}$ is estimated with the MAD or the E-DATE defined earlier. Then,

$$S_{y_{sc}}^w[l, k] = S_{s_{sc}}^w[l, k] + S_{\epsilon_{sc}}^w[l, k]. \quad (11.6.15)$$

With this step we can say that the scaled spectrogram (using Eq. 11.6.14) of the noise $S_{\epsilon_{sc}}^w[l, k]$ follows a chi-squared law with two degrees of freedom, such that

$$S_{\epsilon_{sc}}^w \hookrightarrow \chi_2^2. \quad (11.6.16)$$

The TFPR is such that each component is the value of the CDF of $S_{\epsilon_{sc}}^w$ applied to the elements of the TFD

$$\text{TFPR}[l, k] = \int_0^{S_{\epsilon_{sc}}^w[l, k]} f_{S_{\epsilon_{sc}}^w}(x) dx, \quad (11.6.17)$$

where $f_{S_{\epsilon_{sc}}^w}$ denotes the PDF of $S_{\epsilon_{sc}}^w$. Given $S_{\epsilon_{sc}}^w$ is the chi-squared distributed with two degrees of freedom, $\text{TFPR}[l, k]$ can be analytically derived as

$$\text{TFPR}[l, k] = \int_0^{S_{\epsilon_{sc}}^w[l, k]} e^{-\frac{x}{2}} dx \quad (11.6.18)$$

$$= 1 - e^{-\frac{S_{\epsilon_{sc}}^w[l, k]}{2}}. \quad (11.6.19)$$

Thus, TFPR contains values between 0 and 1; when the value is close to 0, it is considered as a noise coefficient and the converse. Now it is possible to compute the threshold for each TFPR obtained after the filtering step (done by applying a mask); this threshold depends on the probability of the false alarm (PFA) defined by the user. This PFA defines the trade off between the noise reduction and the signal degradation. At this step, I TFPR_i have been computed, merging all these TFPRs. By using the maximum operator we can find a new gain function G, and the useful signal TFD will be approximated by

$$F_{\tilde{s}}^w[l, k] = F_y^w[l, k]G[l, k]. \quad (11.6.20)$$

Finally, the time-domain signal is synthesized by applying the inverse STFT on $F_{\tilde{s}}^w$. The denoising process is then described by the following steps:

1. Calculate the STFT of the data, $F_y^w[l, k]$
2. Calculate the scale of the STFT, $S_{\epsilon_{sc}}^w[l, k] = 2 \frac{S_y^w[l, k]}{E[S_y^w[l, k]]}$
3. Compute $\text{TFPR}_Y[l, k] = 1 - e^{-\frac{S_{\epsilon_{sc}}^w[l, k]}{2}}$
4. Create the masks M_i
5. Calculate the convolution product between each mask and TFPR:

$$\text{TFPR}_Y^i = \text{TFPR}_Y * M_i$$

6. Threshold of each TFPR_Y^i

7. Estimate the gain function G such that

$$G[l, k] = \underset{i}{\operatorname{argmax}}(\text{TFPR}_Y^1[l, k] \dots \text{TFPR}_Y^i[l, k] \dots \text{TFPR}_Y^I[l, k])$$

8. Apply the gain function on F_y^w : $F_{\tilde{s}}^w[l, k] = F_y^w[l, k]G[l, k]$

9. Apply the inverse STFT: $\tilde{s} = \text{STFT}^{-1}(F_{\tilde{s}}^w)$

11.6.6 FINDINGS, DISCUSSIONS, AND INTERPRETATIONS

11.6.6.1 The experiment

Experiments were made using real speech signals that represent one Institute of Electrical and Electronics Engineers (IEEE) sentence corrupted by three nonstationary colored noises (pub noise, car noise, and kids' noise) from the NOIZEUS database [76].

On each signal, the denoising process is applied and an objective comparison is made. In both the denoised Hearingogram and the directional noise reduction algorithms, the MAD is used on each frequency bin to estimate the noise-local PSD, whereas the E-DATE was used for the reference algorithm as in [67]. The denoised Hearingogram is computed, with a 200 Mel filter, a 4th-order Daubechies wavelet and with α equal to 1; the MMSE LSA attenuation rule associated with the DD *a priori* SNR estimator described is tuned such that $\alpha = 0.98$ for the DD estimator, and 54 masks are used in the directional denoising method.

11.6.6.2 Results

The results of the two methods on real speech signals are presented in Figs. 11.6.2 and 11.6.3, in the time and time-frequency domains, respectively. The noisy signal and the clean signal are at the top of each figure, whereas the results of denoising using the directional denoising mask and the denoised Hearingogram are at the bottom of each figure. Note that for the two denoising methods, the background noise is strongly reduced; however, in Fig. 11.6.3, in the case of the denoised Hearingogram method, one can notice that the signal is more deteriorated than in the case of the directional denoising method.

For an objective comparison, let us define two measures of denoising efficiency:

- The SNR defined by

$$\text{SNR} = 10 \log_{10} \frac{\sum_{n=0}^{N-1} s^2[n]}{\sum_{n=0}^{N-1} (s[n] - \tilde{s}[n])^2}.$$

- The normalized mean-square error (NMSE) in dB defined by

$$\text{NMSE} = 10 \log_{10} \frac{1}{N} \sum_{n=0}^{N-1} (s[n] - \tilde{s}[n])^2,$$

where, in the two previous expressions, N is the number of signal samples.

11.6.6.3 Performance improvements

Table 11.6.1 indicates that the directional denoising achieves higher performance than the denoised Hearingogram and the LSA DD, resulting in a significant improvement for speech enhancement. In addition, the performance of the denoised Hearingogram is as good as the MMSE LSA DD but with the following advantages: (1) the noise estimation is online and is not an input of the algorithm; (2) it requires no tuning of parameters (the parameters used in this experiment are valid for all signals), whereas the parameter α in the MMSE LSA DD method is data-dependent.

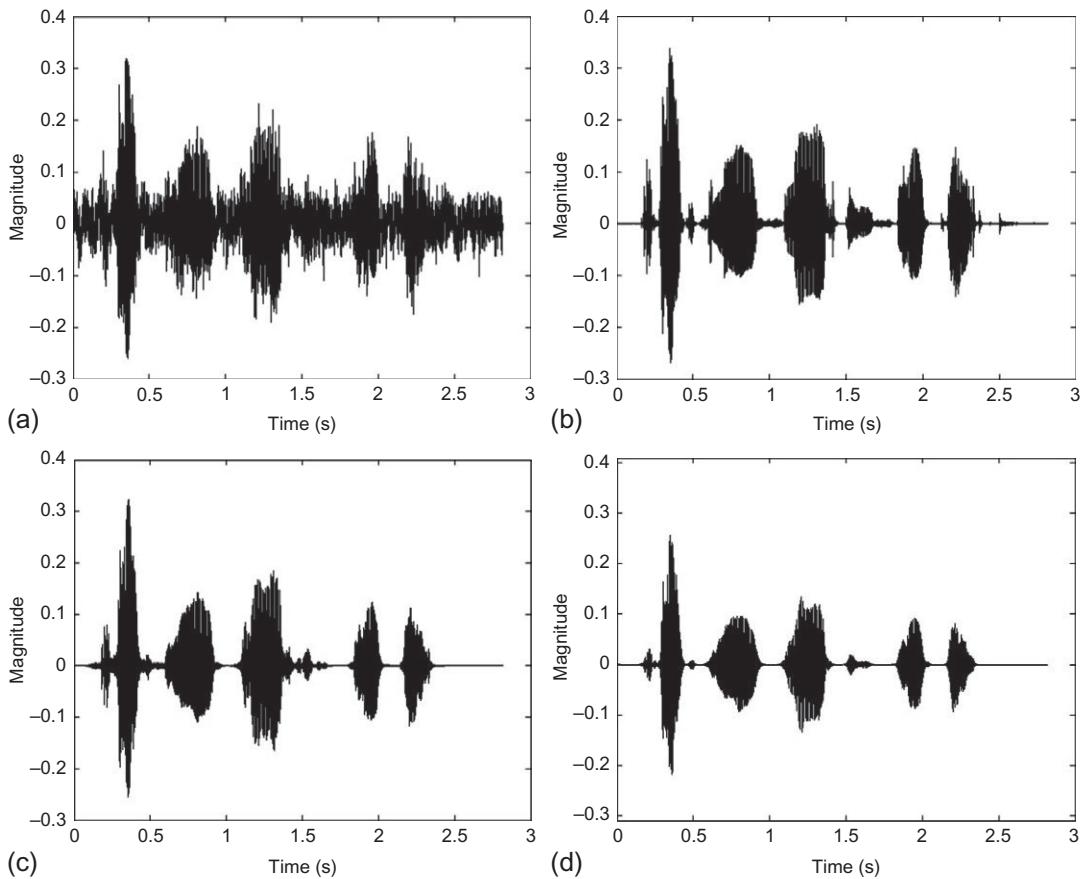
**FIGURE 11.6.2**

Illustration of the two denoising methods for a speech signal in the time domain: (a) represents the noisy observation; (b) is the clean observation; (c) is the denoised noisy observation using directional denoising; (d) is the denoised noisy observation using the denoised Hearingogram.

11.6.7 SUMMARY AND CONCLUSIONS

This section describes and compares methods to denoise audio signals. For these methods, the (t,f) information is used to perform the denoising process. The first one is based on the Hearingogram, which is a sonogram based on Mel filterbanks, coupled with a wavelet shrinkage method. The second one is based on the use of multidirectional masks and detection theory. The general reference speech enhancement algorithm uses the MMSE LSA as attenuation rule associated with DD *a priori* SNR estimator. The denoised Hearingogram performance is similar to the reference method, and the directional denoising shows the best performance resulting in an improvement of 3 dB SNR and 2 dB NMSE compared with the algorithm in [67]. The audio recordings are available as supplementary materials.

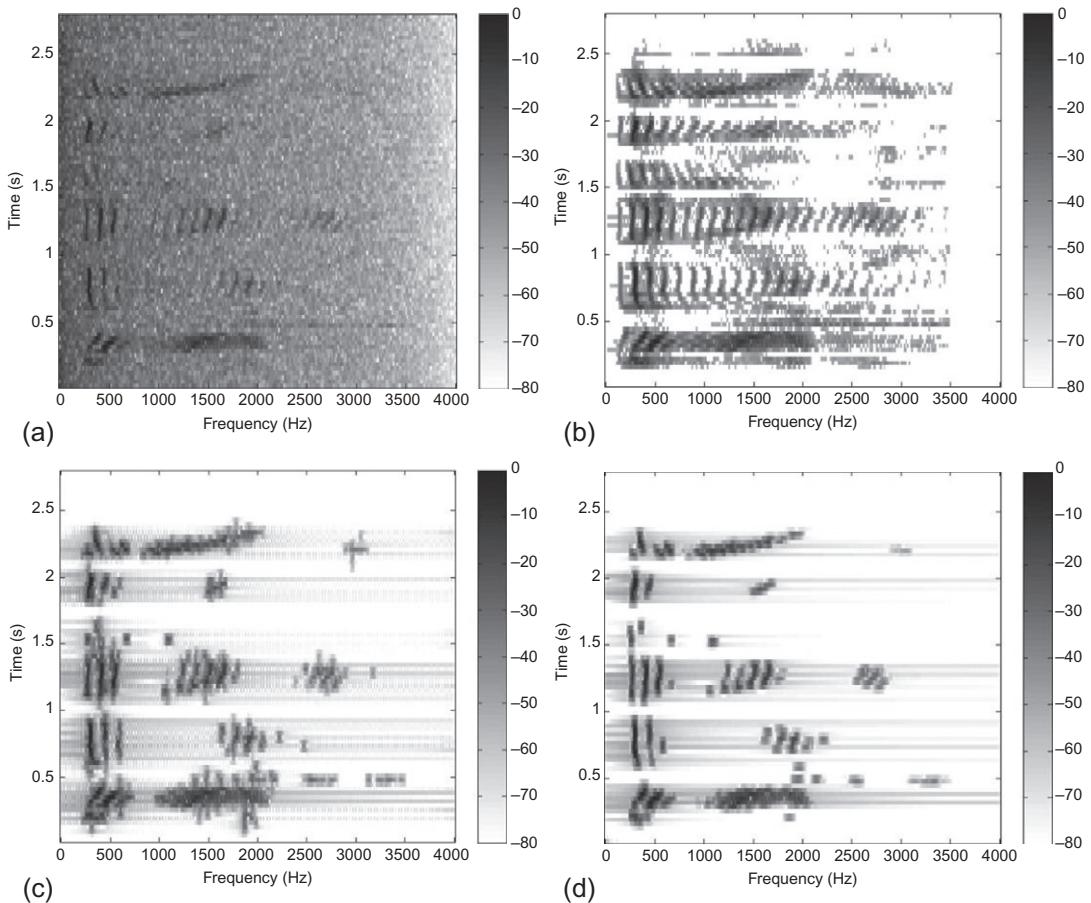
**FIGURE 11.6.3**

Illustration of the two denoising methods for a speech signal using spectrograms: (a) represents the noisy observation; (b) is the clean observation; (c) is the denoised noisy observation using directional denoising; (d) is the denoised noisy observation using the denoised Hearingogram.

Table 11.6.1: Objective Comparison Between MMSE LSA DD, Denoised Hearingogram, (DH) and the Directional Mask Denoising (DMS)

| | Noisy | MMSE LSA DD | DH | DMS |
|-----------|-------|-------------|------|-------|
| SNR (dB) | 5.11 | 9 | 8.79 | 11.87 |
| NMSE (dB) | -32 | -37 | -37 | -39.2 |

REFERENCES

- [1] G. Matz, F. Hlawatsch, “Linear time-frequency filters: on-line algorithms and applications”, in: A. Papandreou-Suppappa (Ed.), Applications in Time-Frequency Signal Processing, CRC Press, Boca Raton, FL, 2002, pp. 205-271 (Chapter 6).
- [2] F. Hlawatsch, Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation, Kluwer, Boston, 1998.
- [3] W. Kozek, “Time-frequency signal processing based on the Wigner-Weyl framework”, *Signal Process.* 29 (1) (1992) 77-92.
- [4] R. Bourdier, J.F. Allard, K. Trumpp, “Effective frequency response and signal replica generation for filtering algorithms using multiplicative modifications of the STFT”, *Signal Process.* 15 (1988) 193-201.
- [5] I. Daubechies, “Time-frequency localization operators: a geometric phase space approach”, *IEEE Trans. Inform. Theory* 34 (1988) 605-612.
- [6] W. Kozek, H.G. Feichtinger, J. Scharinger, “Matched multiwindow methods for the estimation and filtering of nonstationary processes”, in: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS 96), vol. 2, 1996, pp. 509-512.
- [7] S.H. Nawab, T.F. Quatieri, “Short-time Fourier transform”, in: J.S. Lim, A.V. Oppenheim (Eds.), Advanced Topics in Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1988, pp. 289-337 (Chapter 6).
- [8] S. Farkash, S. Raz, “Linear systems in Gabor time-frequency space”, *IEEE Trans. Signal Process.* 42 (3) (1994) 611-617.
- [9] H.G. Feichtinger, T. Strohmer (Eds.), Gabor Analysis and Algorithms: Theory and Applications, Birkhäuser, Berlin/Boston, 1998.
- [10] H.G. Feichtinger, K. Nowak, “A first survey of Gabor multipliers”, in: H.G. Feichtinger, T. Strohmer (Eds.), Advances in Gabor Analysis, Birkhäuser, Boston, 2003, pp. 99-128 (Chapter 5).
- [11] M. Dörfler, B. Torrésani, “Representation of operators in the time-frequency domain and generalized Gabor multipliers”, *J. Fourier Anal. Appl.* 16 (2) (2010) 261-293.
- [12] G. Matz, F. Hlawatsch, “Time-frequency projection filters: online implementation, subspace tracking, and application to interference excision”, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’02), 2002, pp. 1213-1216.
- [13] Y.Y. Zeevi, M. Zibulski, M. Porat, “Multi-window Gabor schemes in signal and image representations”, in: H.G. Feichtinger, T. Strohmer (Eds.), Gabor Analysis and Algorithms: Theory and Applications, Birkhäuser, Berlin/Boston, 1998, pp. 381-407 (Chapter 12).
- [14] P. Balazs, D. Bayer, A. Rahimi, “Multipliers for continuous frames in Hilbert spaces”, *J. Phys. A Math. Theor.* 45 (24) (2012) <http://dx.doi.org/10.1088/1751-8113/45/24/244023>, article ID 244023, 20 pp.
- [15] F. Hlawatsch, G. Matz, “Time-frequency methods for non-stationary statistical signal processing”, in: F. Hlawatsch, F. Auger (Eds.), Time-Frequency Analysis: Concepts and Methods, ISTE/Wiley, London, UK, 2008, pp. 279-320 (Chapter 10).
- [16] S. Qian, “Time-varying filter via Gabor expansion”, in: B. Boashash (Ed.), Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, first ed., Elsevier, Oxford, 2003, pp. 476-480 (Chapter 11, Article 11.2).
- [17] D. Gabor, “Theory of communication”, *J. IEE* 93(III) (26) (1946) 429-457.
- [18] S. Qian, D. Chen, Joint Time-Frequency Analysis: Methods & Applications, Prentice-Hall, Upper Saddle River, NJ, 1996.
- [19] X.-G. Xia, S. Qian, “Convergence of an iterative time-variant filtering based on discrete Gabor transform”, *IEEE Trans. Signal Process.* 47 (10) (1999) 2894-2899.
- [20] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.

- [21] G. Azemi, B. Senadji, B. Boashash, "Ricean K-factor estimation in mobile communication systems", *IEEE Commun. Lett.* 8 (10) (2004) 617-619, ISSN 1089-7798, <http://dx.doi.org/10.1109/LCOMM.2004.835344>.
- [22] D. Iskander, A. Zoubir, B. Boashash, "A method for estimating the parameters of the K distribution", *IEEE Trans. Signal Process.* 47 (4) (1999) 1147-1151, ISSN 1053-587X, <http://dx.doi.org/10.1109/78.752614>.
- [23] C. Breining, P. Dreiseitel, E. Haensler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, J. Tilp, "Acoustic echo control: an application of very-high-order adaptive filters", *IEEE Signal Process. Mag.* 16 (1999) 42-69.
- [24] S.V. Vaseghi, Advanced Signal Processing and Digital Noise Reduction, Wiley and Teubner, New York, 1996.
- [25] J.S. Lim, A.V. Oppenheim, "Enhancement and bandwidth compression of noisy speech", *Proc. IEEE* 67 (12) (1979) 1586-1604.
- [26] J. Yang, "Frequency domain noise suppression approaches in mobile telephone systems", in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'93), vol. 2, 1993, pp. 363-366.
- [27] G. Matz, F. Hlawatsch, W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes", *IEEE Trans. Signal Process.* 45 (6) (1997) 1520-1534.
- [28] L. Stanković, S. Stanković, I. Djurović, "Space/spatial frequency based filtering", *IEEE Trans. Signal Process.* 48 (8) (2000) 2343-2352.
- [29] H.L.L. Van Trees, Detection, Estimation, and Modulation Theory, vol. I: "Detection, Estimation, and Linear Modulation Theory", Wiley, New York, 1968, reprinted 2001.
- [30] F. Hlawatsch, G. Matz, H. Kirchauer, W. Kozek, "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation", *IEEE Trans. Signal Process.* 48 (5) (2000) 1417-1432.
- [31] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.
- [32] L. Stanković, "On the time-frequency analysis based filtering", *Ann. Telecommun.* 55 (5/6) (2000) 216-225.
- [33] P. Flandrin, W. Martin, "The Wigner-Ville spectrum of nonstationary random signals", in: W. Mecklenbräuker, F. Hlawatsch (Eds.), The Wigner Distribution—Theory and Applications in Signal Processing, Elsevier, Amsterdam, 1997, pp. 211-267.
- [34] S. Stanković, I. Orović, N. Žarić, "Robust speech watermarking procedure in the time-frequency domain", *EURASIP J. Adv. Signal Process.* 2008 (12) (2008) <http://dx.doi.org/10.1155/2008/519206>, article ID 519206, 9 pp.
- [35] I. Orović, S. Stanković, "Time-frequency-based speech regions characterization and eigenvalue decomposition applied to speech watermarking", *EURASIP J. Adv. Signal Process.* 2010 (2010) <http://dx.doi.org/10.1155/2010/572748>, article ID 572748, 10 pp.
- [36] S. Stanković, I. Orović, N. Žarić, "An application of multidimensional time-frequency analysis as a base for the unified watermarking approach", *IEEE Trans. Image Process.* 19 (3) (2010) 736-745.
- [37] I. Orović, S. Stanković, T. Chau, C.M. Steele, E. Sejdić, "Time-frequency analysis and Hermite projection method applied to swallowing accelerometry signals", *EURASIP J. Adv. Signal Process.* 2010 (2010) <http://dx.doi.org/10.1155/2010/323125>, article ID 323125, 7 pp.
- [38] L. Cohen, "Introduction: a primer on time-frequency analysis", in: B. Boashash (Ed.), Time-Frequency Signal Analysis: Methods and Applications, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 3-42 (Chapter 1).
- [39] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [40] S. Stanković, "About time-variant filtering of speech signals with time-frequency distributions for hands-free telephone systems", *Signal Process.* 80 (9) (2000) 1777-1785.
- [41] L. Stanković, "A method for time-frequency analysis", *IEEE Trans. Signal Process.* 42 (1) (1994) 225-229.

- [42] S. Stanković, J. Tilp, "Time-varying filtering of speech signals using linear prediction", *Electron. Lett.* 36 (8) (2000) 763-764.
- [43] M.J. Arnold, M. Roessgen, B. Boashash, "Filtering real signals through frequency modulation and peak detection in the time-frequency plane", in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'94), vol. 3, 1994, pp. 345-348.
- [44] M.J. Arnold, B. Boashash, "Time-frequency peak filtering: analysis and implementation details", in: Proceedings of the IEEE-SP International Symposium on Time-Frequency & Time-Scale Analysis, 1994, pp. 256-259.
- [45] M. Roessgen, B. Boashash, "Time-frequency peak filtering applied to FSK signals", in: Proceedings of the IEEE-SP International Symposium on Time-Frequency & Time-Scale Analysis, 1994, pp. 516-519.
- [46] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications", *Proc. IEEE* 80 (4) (1992) 519-568.
- [47] B. Boashash, "Time-frequency signal analysis", in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [48] A. Papoulis, "Random modulation: a review", *IEEE Trans. Acoust. Speech Signal Process.* 31 (1) (1983) 96-105.
- [49] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, third ed., McGraw-Hill, New York, 1991.
- [50] B. Boashash, M. Mesbah, "Signal enhancement by time-frequency peak filtering", *IEEE Trans. Signal Process.* 52 (4) (2004) 929-937.
- [51] B. Boashash, P.J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.* 42 (1) (1994) 216-220.
- [52] H. Lin, Y. Li, B. Yang, H. Ma, "Random denoising and signal nonlinearity approach by time-frequency peak filtering using weighted frequency reassignment", *Geophysics* 78 (6) (2013) 229-237.
- [53] Y. Liu, Y. Li, P. Nie, Q. Zeng, "Spatiotemporal time-frequency peak filtering method for seismic random noise reduction", *IEEE Geosci. Remote Sens. Lett.* 10 (4) (2013) 756-760, ISSN 1545-598X, <http://dx.doi.org/10.1109/LGRS.2012.2221676>.
- [54] Y. Liu, B. Dang, Y. Li, H. Lin, "Local spatiotemporal time-frequency peak filtering method for seismic random noise reduction", *J. Appl. Geophys.* 111 (2014) 76-85, ISSN 0926-9851, <http://dx.doi.org/10.1016/j.jappgeo.2014.09.018>, URL <http://www.sciencedirect.com/science/article/pii/S092698511400281X>.
- [55] I. Djurović, L. Stanković, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment", *Signal Process.* 84 (2004) 631-643.
- [56] H. Hassanpour, "A time-frequency approach for noise reduction", *Digital Signal Process.* 18 (5) (2008) 728-738.
- [57] J.M. O'Toole, M. Mesbah, B. Boashash, "Accurate and efficient implementation of the time-frequency matched filter", *IET Signal Process.* 4 (4) (2010) 428-437.
- [58] G. Strang, *Introduction to Linear Algebra*, fourth ed., Wellesley-Cambridge Press, Wellesley, MA, 2009.
- [59] H. Hassanpour, A. Zehtabian, S.J. Sadati, "Time domain signal enhancement based on an optimized singular vector denoising algorithm", *Digital Signal Process.* 22 (5) (2012) 786-794.
- [60] R.W. Schafer, "What is a Savitzky-Golay filter", *IEEE Signal Process. Mag.* 28 (4) (2011) 111-117.
- [61] G.F. Boudreaux-Bartels, T.W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 34 (3) (1986) 442-451.
- [62] J. Chen, J. Benesty, Y. Huang, S. Doclo, "New insights into the noise reduction Wiener filter", *IEEE Trans. Audio Speech Lang. Process.* 14 (4) (2006) 1218-1234.
- [63] H. Hassanpour, M. Mesbah, B. Boashash, "Time-frequency based newborn EEG seizure detection using low and high frequency signatures", *Physiol. Meas.* 25 (4) (2004) 935-944.

- [64] A.P. Reilly, B. Boashash, "Comparison of time-frequency signal analysis techniques with application to speech recognition", in: Advanced Signal Processing Algorithms, Architectures, and Implementations III, vol. 1770, 1992, pp. 339-350.
- [65] D.L. Donoho, J.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage", *Biometrika* 81 (3) (1994) 425-455.
- [66] D. Pastor, F. Socheleau, "Robust estimation of noise standard deviation in presence of signals with unknown distributions and occurrences", *IEEE Trans. Signal Process.* 60 (4) (2012) 1545-1555, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2012.2184534>.
- [67] V.-K. Mai, D. Pastor, A. Aissa-El-Bey, R. Le-Bidan, "Robust estimation of non-stationary noise power spectrum for speech enhancement", *IEEE/ACM Trans. Audio Speech Lang. Process.* 23 (4) (2015) 670-682, ISSN 2329-9290, <http://dx.doi.org/10.1109/TASLP.2015.2401426>.
- [68] C. Plapous, C. Marro, P. Scalart, "Improved signal-to-noise ratio estimation for speech enhancement", *IEEE Trans. Audio Speech Lang. Process.* 14 (6) (2006) 2098-2108, ISSN 1558-7916, <http://dx.doi.org/10.1109/TASL.2006.872621>.
- [69] I. Cohen, "Speech enhancement using a noncausal a priori SNR estimator", *IEEE Signal Process. Lett.* 11 (9) (2004) 725-728.
- [70] Y. Ephraim, D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator", *IEEE Trans. Acoust. Speech Signal Process.* 33 (2) (1985) 443-445.
- [71] T. Lotter, P. Vary, "Speech enhancement by map spectral amplitude estimation using a super-Gaussian speech model", *EURASIP J. Appl. Signal Process.* 2005 (2005) 1110-1126, ISSN 1110-8657, <http://dx.doi.org/10.1155/ASP.2005.1110>.
- [72] S. Ouelha, "Représentation et reconnaissance des signaux acoustiques sous-marins", Ph.D. thesis, Université de Toulon et du Var, 2014.
- [73] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (7) (1989) 674-693, ISSN 0162-8828, <http://dx.doi.org/10.1109/34.192463>.
- [74] D.L. Donoho, I.M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage", *J. Am. Stat. Assoc.* 90 (432) (1995) 1200-1224.
- [75] U. Moreaud, P. Courmontagne, S. Ouelha, F. Chaillan, J.-R. Mesquida, "Underwater acoustic signal denoising using multi-directional masks on time-frequency representation", in: OCEANS 2014-TAIPEI, IEEE, 2014, pp. 1-8.
- [76] P.C. Loizou, *Speech Enhancement: Theory and Practice*, CRC Press, Boca Raton, FL, 2013.

This page intentionally left blank

DETECTION, CLASSIFICATION, AND ESTIMATION IN THE (t, f) DOMAIN

12

INTRODUCTION AND OVERVIEW

Several studies involving real-life applications have shown that methods for the detection, estimation, and classification of nonstationary signals can be enhanced by utilizing the time-frequency $((t,f))$ characteristics of such signals. Such (t,f) formulations are described in this chapter and include (t,f) matched filtering for detection and extraction of (t,f) features for classification. The topic is covered in six sections with appropriate internal cross-referencing to this and other chapters.

The structure of (t,f) methods is suitable for designing and implementing optimal detectors. Several approaches exist, such as decomposition of TFDs into sets of spectrograms (Section 12.1). For both analysis and classification, a successful (t,f) methodology requires matching of TFDs with the structure of the signal. This can be achieved by a matching pursuit algorithm using (t,f) atoms adapted to the analyzed signals (Section 12.2). We can perform system identification by exciting linear systems with a linear FM signal and relating TFDs of the input and output using (t,f) filtering techniques (Section 12.3). Methods for (t,f) signal estimation and detection can be carried out using time-varying Wiener filters (Section 12.4). The last two sections present advanced formulations and methods for (t,f) matched filtering (Section 12.5) and the formulation of (t,f) features for classification (Section 12.6), both of which are applied to a serious medical problem as an illustration of the performance gained.

12.1 OPTIMAL TIME-FREQUENCY DETECTORS⁰

12.1.1 TIME-FREQUENCY DETECTION

In addition to *exploratory signal analysis*, the structure of time-frequency (t, f) representations (TFRs) (also called (t, f) distributions or TFDs in other parts of the book) can also be fruitfully exploited for designing and implementing optimal detectors for nonstationary signals in noisy environments. Perhaps the simplest and best known examples are the short-time Fourier transform (STFT) and the cross-ambiguity function in radar/sonar processing [1].

This section highlights the key ideas behind the role of TFRs in optimal detection. Essentially, TFRs are useful for detecting nonstationary stochastic signals that exhibit certain unknown parameters, called *nuisance* parameters, such as unknown time or frequency shifts. The characteristics of the signal are matched via proper choice of the TFR kernel and the nuisance parameters are dealt with by exploiting the *covariance* of TFRs to time and frequency shifts. Suboptimal, low complexity (t, f) detectors can also deliver competitive performance in scenarios in which optimal detectors are prohibitively complex.

12.1.2 TIME-FREQUENCY REPRESENTATIONS

The most widely used *linear* TFR is the STFT that correlates the signal of interest, $r(t)$, with (t, f) shifted copies of a prototype (window) signal $g(t)$

$$F_r^g(t, f) = \int r(\tau)g^*(\tau - t)e^{-j2\pi f\tau} d\tau. \quad (12.1.1)$$

An important *quadratic* TFR is the Wigner distribution (WD) or Wigner-Ville distribution (WVD) that essentially correlates the signal with (t, f) shifted copies of a time-reversed version of itself:

$$W_r(t, f) = \int r(t + \tau/2)r^*(t - \tau/2)e^{-j2\pi f\tau} d\tau. \quad (12.1.2)$$

A class of quadratic TFRs or TFDs, defined in Chapter 3, Eq. (3.2.13), can be generated from the WD (or WVD) via (t, f) filtering with a two-dimensional kernel

$$\rho_r(t, f; \gamma) = \int \int W_r(u, v)\gamma(u - t, v - f) du dv, \quad (12.1.3)$$

where ρ is characterized by different choices of the kernel $\gamma(t, f)$. A characteristic property of quadratic TFRs is that they are *covariant* to (t, f) shifts in the signal:

$$r(t) \longrightarrow r(t - t_o)e^{j2\pi f_o t} \implies \rho_r(t, f) \longrightarrow \rho_r(t - t_o, f - f_o). \quad (12.1.4)$$

The magnitude squared STFT (i.e., the spectrogram) is a quadratic TFR with kernel given by the WD of the window function: $|F_r^g(t, f)|^2 = \rho_r(t, f; \gamma = W_g)$.

⁰Author: A.M. Sayeed, University of Wisconsin-Madison, Madison, WI, USA (akbar@engr.wisc.edu). Reviewers: P. Flandrin and A. Papandreou-Suppappola.

An important relation in the context of detection is *Weyl correspondence* that relates a quadratic form to the WD

$$\langle \mathbf{L}x, x \rangle = \int \int x^*(t_1) L(t_1, t_2) x(t_2) dt_1 dt_2 = \int \int W_x(t, f) \text{WS}_L(t, f) dt df, \quad (12.1.5)$$

where \mathbf{L} is a linear operator defined by the kernel $L(t_1, t_2)$ and $\text{WS}_L(t, f)$ denotes the Weyl symbol of the operator defined as [2]

$$\text{WS}_L(t, f) = \int L(t + \tau/2, t - \tau/2) e^{-j2\pi f\tau} d\tau. \quad (12.1.6)$$

If $L(t_1, t_2) = R_s(t_1, t_2) = \mathcal{E}[s(t_1)s^*(t_2)]$, the correlation function of a random signal $s(t)$, then the Weyl symbol of $R_s(t_1, t_2)$ is also the Wigner-Ville spectrum (WVS) of $s(t)$, defined in [Chapter 2](#) as

$$\text{WVS}_s(t, f) = \mathcal{E}[W_s(t, f)] = \int R_s(t + \tau/2, t - \tau/2) e^{-j2\pi f\tau} d\tau = \text{WS}_{R_s}(t, f). \quad (12.1.7)$$

The WVS is a generalization of the notion of power spectral density to the case of nonstationary processes as discussed in [Chapter 2](#). The kernel $\gamma(t, f)$ can be interpreted as the WS of an operator L_γ by inverting Eq. (12.1.6); this results in

$$L_\gamma(t_1, t_2) = \int \gamma\left(\frac{t_1 + t_2}{2}, f\right) e^{j2\pi f(t_1 - t_2)} df. \quad (12.1.8)$$

If $\gamma(t, f)$ generates a real-valued TFR then L_γ is a Hermitian operator and under certain conditions (e.g., if it is a Hilbert-Schmidt operator) it admits an eigendecomposition $L_\gamma(t_1, t_2) = \sum_k \lambda_k u_k(t_1) u_k^*(t_2)$ where $\{\lambda_k\}$ are the real-valued eigenvalues and $\{u_k(t)\}$ are the corresponding orthonormal eigenfunctions. Using this eigendecomposition, the kernel can be expressed as a weighted sum of WDs, $\gamma(t, f) = \sum_k \lambda_k W_{u_k}(t, f)$, and the corresponding TFR can be expressed as a weighted sum of spectrograms

$$\rho_r(t, f; \gamma) = \sum_k \lambda_k \int \int W_r(u, v) W_{u_k}(u - t, v - f) du dv = \sum_k \lambda_k |F_r^{u_k}(t, f)|^2. \quad (12.1.9)$$

This decomposition in terms of spectrograms (discussed in [Section 6.4](#)) provides an intimate connection between linear and quadratic TFRs and is very useful for the efficient implementation of (t, f) -based detectors.

12.1.3 TIME-FREQUENCY DETECTION FRAMEWORK

Let's motivate (t, f) -based detection by considering the classical binary hypothesis testing problem encountered in radar. Let $v(t)$ denote the complex baseband transmitted signal of duration T_o and bandwidth B_o . The received signal $r(t)$ is processed over an observation interval $[T_i, T_f]$ to detect the presence of a target. The two hypothesis on $r(t)$ are

$$\begin{aligned} H_0: r(t) &= \epsilon(t), \quad t \in [T_i, T_f], \\ H_1: r(t) &= as(t; t_o, f_o) + \epsilon(t), \quad t \in [T_i, T_f]. \end{aligned} \quad (12.1.10)$$

Under the null hypothesis, H_0 , the signal contains only the zero-mean complex additive white Gaussian noise (AWGN), $\epsilon(t)$, of variance σ_ϵ^2 . Under the active hypothesis, H_1 , a (t, f) shifted version of the

transmitted signal, $s(t; t_o, f_o) = v(t - t_o)e^{j2\pi f_o t}$, is received in the presence of noise. In Eq. (12.1.10), a denotes an unknown complex gain parameter. The unknown (t, f) parameters, (t_o, f_o) , represent the delay and Doppler shift of the received signal relative to the transmitted signal and correspond to the unknown distance and velocity of the target. Let $\mathcal{S} = [T_{\min}, T_{\max}] \times [-F_{\max}, F_{\max}]$ denote the possible range of values for the nuisance parameters (t_o, f_o) . In this case, the required observation interval is $[T_i, T_f] = [T_{\min}, T_o + T_{\max}]$. For any given value of $(t_o, f_o) \in \mathcal{S}$, the optimal decision statistic, derived from the likelihood ratio [1,3], is the noncoherent matched filter statistic

$$\eta^{(t_o, f_o)} = |\langle r, s(t_o, f_o) \rangle|^2 = \left| \int_{T_i}^{T_f} r(t) v^*(t - t_o) e^{-j2\pi f_o t} dt \right|^2 = |F_r^v(t_o, f_o)|^2, \quad (12.1.11)$$

which correlates the received signal with the reference signal $s(t; t_o, f_o)$ and can be computed via a spectrogram with $v(t)$ as the window. The unknown nature of the nuisance parameters is usually accounted for via a generalized likelihood ratio test (GLRT) corresponding to the decision statistic

$$\eta(r) = \max_{(t, f) \in \mathcal{S}} \eta^{(t, f)}(r) = \max_{(t, f) \in \mathcal{S}} |F_r^v(t, f)|^2, \quad (12.1.12)$$

which represents the peak of the spectrogram over \mathcal{S} . The statistic $\eta(r)$ is compared to a threshold to declare the presence (H_1) or absence (H_0) of a target. Thus, the optimal detector correlates the received signal with all possible (t, f) shifted versions of the transmitted signal and picks the largest correlation as the decision statistic. The location of the peak is in fact a maximum likelihood (ML) estimate of the unknown parameters: $(\hat{t}_o, \hat{f}_o)_{\text{ML}} = \arg \max_{(t, f) \in \mathcal{S}} \eta^{(t, f)}$. The GLRT detector (Eq. (12.1.12)) is

the statistic (Eq. (12.1.11)) computed at $(\hat{t}_o, \hat{f}_o)_{\text{ML}}$. The detector performance is governed by the shape of the magnitude squared of the ambiguity function (AF) $A_v(\tau, \nu)$ of the transmitted signal $v(t)$ near the origin [1]. The AF is defined as

$$A_v(\tau, \nu) = \int v(t + \tau/2) v^*(t - \tau/2) e^{-j2\pi \nu t} dt. \quad (12.1.13)$$

Note that $|A_v(\tau, \nu)|^2 = |F_v^v(\tau, \nu)|^2 = \iint W_v(u, v) W_v(u - \tau, v - \nu) du dv$. Ideally, $A_v(\tau, \nu)$ should approximate a two-dimensional Dirac delta function for perfect resolution in delay and Doppler. Waveforms with large time-bandwidth products, such as those derived from pseudorandom codes used in spread-spectrum communications [4], can be used to approximate this 2D delta function.

The simple detector (Eq. (12.1.12)) is based on a rank-1 quadratic¹ TFR. Higher-rank quadratic TFRs can realize detectors for a much richer class of nonstationary stochastic signals. The next subsection describes this quadratic (t, f) detection framework. In turn, in many cases low-rank approximations of (t, f) detectors, implemented via a bank of a few spectrograms, suffice.

12.1.3.1 Quadratic time-frequency detectors

Quadratic TFRs can be exploited for detection scenarios of the form (Eq. (12.1.10)), where the signal $s(t; t_o, f_o)$ is a (t, f) shifted version of some underlying stochastic signal whose WVS is relatively localized in time and frequency. This represents a rich class of detection scenarios for modeling the inherent randomness of signals as well as measurement uncertainties that manifest themselves as

¹The rank of a TFR is the number of significant nonzero eigenvalues of $\gamma(t, f)$ in Eq. (12.1.9).

unknown (t, f) shifts. For example, radar applications in which the transmitted signal $v(t)$ encounters a randomly time-varying scatterer or channel may be modeled this way [1]. Similarly, signals involved in machine fault diagnostics may exhibit random characteristics along with timing jitters and frequency offsets due to physical mechanisms. The idea behind quadratic (t, f) detectors is to capture the signal statistics via the kernel $\gamma(t, f)$ and the nuisance parameters via the (t, f) covariance property (Eq. (12.1.4)).

Suppose that the signal $s(t; t_o, f_o)$ in Eq. (12.1.10) is a (t, f) shifted version of a zero-mean Gaussian signal $w(t)$; that is, $s(t; t_o, f_o) = w(t - t_o)e^{j2\pi f_o t}$ and $R_{s(t_o, f_o)}(t_1, t_2) = R_w(t_1 - t_o, t_2 - t_o)e^{j2\pi f_o(t_1 - t_2)}$. In this case, for any given (t_o, f_o) , the optimum decision statistic is [3, 5, 6]

$$\eta^{(t_o, f_o)}(r) = \langle \hat{\mathbf{R}}_{s(t_o, f_o)} r, r \rangle = \langle W_r, \text{WS}_{\hat{\mathbf{R}}_{s(t_o, f_o)}} \rangle, \quad (12.1.14)$$

where $\hat{\mathbf{R}}_{s(t_o, f_o)} = (\mathbf{R}_{s(t_o, f_o)} + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{R}_{s(t_o, f_o)}$ is an operator defined by the signal correlation function and the second equality follows from Eq. (12.1.5). The simpler choice $\hat{\mathbf{R}}_{s(t_o, f_o)} = \mathbf{R}_{s(t_o, f_o)}$ in Eq. (12.1.14) yields the *locally optimum*² detector [3].

The unknown nuisance parameters are again accounted for by exploiting the covariance property (Eq. (12.1.4)) which implies that $\text{WS}_{R_{s(t_o, f_o)}}(t, f) = \text{WS}_{R_w}(t - t_o, f - f_o)$ and $\text{WS}_{\hat{\mathbf{R}}_{s(t_o, f_o)}}(t, f) = \text{WS}_{\hat{\mathbf{R}}_w}(t - t_o, f - f_o)$, where $\hat{\mathbf{R}}_w = (\mathbf{R}_w + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{R}_w$ [6].³ Consequently, the quadratic decision statistic in Eq. (12.1.14) can be expressed as

$$\begin{aligned} \eta^{(t_o, f_o)} &= \int \int W_r(u, v) \text{WS}_{\hat{\mathbf{R}}_w}(u - t_o, v - f_o) du dv \\ &= \rho_r(t_o, f_o; \gamma = \text{WS}_{\hat{\mathbf{R}}_w}). \end{aligned} \quad (12.1.15)$$

Similarly, the locally optimal detector corresponds to $\hat{\mathbf{R}}_w = \mathbf{R}_w$ in Eq. (12.1.15). Thus, the decision statistic for different values of (t_o, f_o) can be computed via a TFR with kernel $\text{WS}_{\hat{\mathbf{R}}_w}$. The overall GLRT detector for unknown (t_o, f_o) is given by

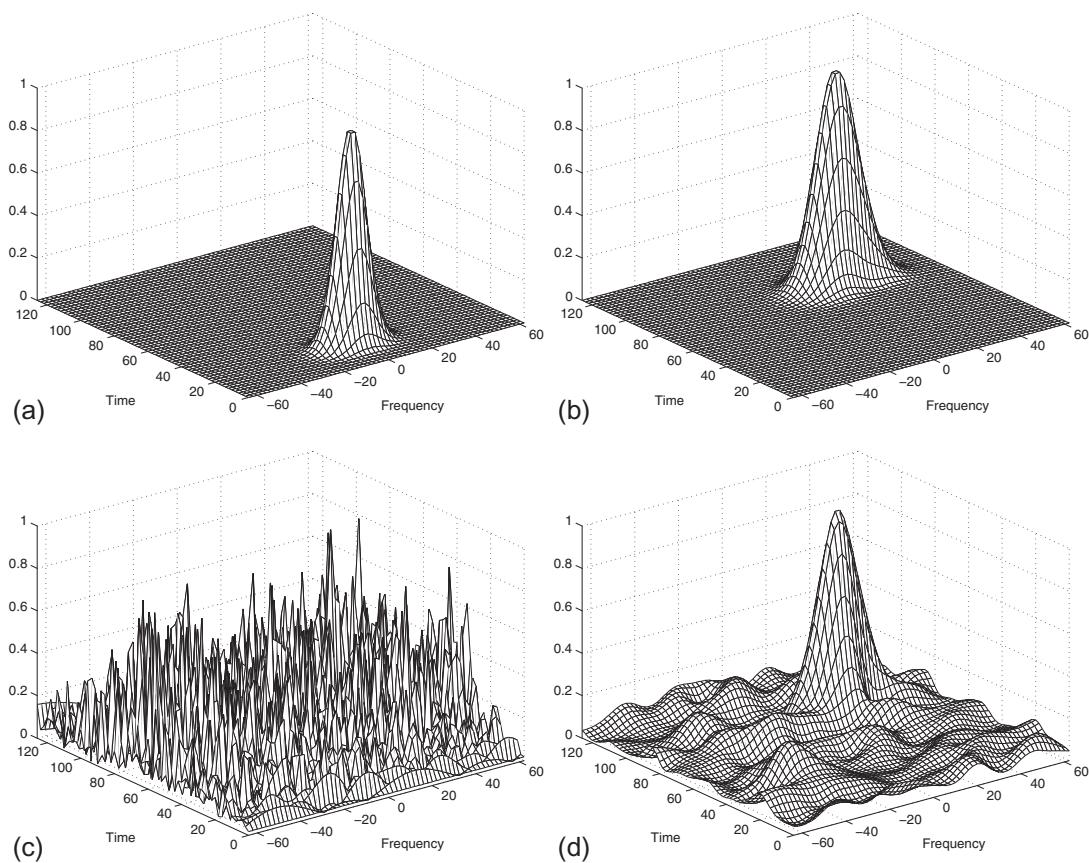
$$\begin{aligned} \eta(r) &= \max_{(t, f) \in \mathcal{S}} \rho_r(t, f; \text{WS}_{\hat{\mathbf{R}}_w}) \\ &= \max_{(t, f) \in \mathcal{S}} \int \int W_r(u, v) \text{WS}_{\hat{\mathbf{R}}_w}(u - t, v - f) du dv, \end{aligned} \quad (12.1.16)$$

where $\mathcal{S} = [T_{\min}, T_{\max}] \times [-F_{\max}, F_{\max}]$. If $\text{WVS}_w(t, f) = \mathcal{E}[W_w(t, f)]$ is supported on $(t, f) \in [0, T_w] \times [-B_w, B_w]$, the required observation interval is $[T_i, T_f] = [T_{\min}, T_{\max} + T_w]$ and $W_r(t, f)$ in Eq. (12.1.16) is computed over the range $[T_{\min}, T_{\max} + T_w] \times [-B_w - F_{\max}, B_w + F_{\max}]$.

The detector (Eq. (12.1.16)) has the intuitively appealing interpretation of *nonstationary spectral correlation*: the WD of the observed waveform is correlated with all possible (t, f) shifted versions of $\text{WS}_{\hat{\mathbf{R}}_w}$ which is the WVS of $w(t)$ in the case of the locally optimum detector and the WVS of a modified version of $w(t)$ in the case of the optimal detector. This interpretation of optimum (t, f) detection is illustrated in Fig. 12.1.1. The location of the maximum in Eq. (12.1.16) is the ML estimate of (t_o, f_o) .

²The first term in the expansion of the likelihood ratio as a function of signal-to-noise ratio (SNR) [3]. This decision statistic is useful in low-SNR scenarios.

³We note that both $\hat{\mathbf{R}}_w$ and \mathbf{R}_w have the same eigenfunctions; the eigenvalues of $\hat{\mathbf{R}}_w$ are obtained from those of \mathbf{R}_w by adding σ_ϵ^2 to each.

**FIGURE 12.1.1**

Nonstationary spectral correlation underlying (t, f) detection. (a) The WVS, $\text{WVS}_w(t, f)$, of the underlying stochastic signal $w(t)$. (b) The idealized TFR, $\rho_r(t, f; \gamma = \text{WVS}_w)$, in the absence of noise; that is, $r(t) = s(t; t_0, f_0)$. (c) The WD of the noisy observed signal $r(t)$. (d) The TFR, $\rho_r(t, f; \gamma = \text{WVS}_w)$, of the noisy signal. The TFR in (d) is computed by smoothing the WD in (c) with the WVS in (a) and consists of the idealized TFR in (b) and additive noisy components. The peak of the TFR in (d) is the decision statistic (compared to a threshold). The location of the peak represents ML estimates of the unknown delay and Doppler shift of the target.

Estimator performance is related to the shape of $F(\tau, \nu) = \int \int \text{WVS}_w(u, v) \text{WVS}_w(u - \tau, v - \nu) du dv$ near the origin.

The correlation function $\hat{R}_w(t_1, t_2)$ is related to $R_w(t_1, t_2)$ by

$$\hat{R}_w(t_1, t_2) = \sum_k \frac{\lambda_k}{\lambda_k + \sigma_\epsilon^2} u_k(t_1) u_k^*(t_2),$$

where $\{\lambda_k\}$ and $\{u_k\}$ denote, respectively, eigenvalues and eigenfunctions, respectively, of $R_w(t_1, t_2)$. Analogous to Eq. (12.1.9), the decision statistic (Eq. (12.1.15)) can be expressed as a weighted sum of spectrograms

$$\eta^{(t_o, f_o)}(r) = \sum_k \hat{\lambda}_k |F_r^{u_k}(t_o, f_o)|^2, \quad (12.1.17)$$

where $\hat{\lambda}_k = \frac{\lambda_k}{\lambda_k + \sigma_e^2}$ in the case of optimum detection and $\hat{\lambda}_k = \lambda_k$ in the case of locally optimum detection. The importance of the above decomposition stems from the fact that many signals encountered in practice exhibit few dominant eigenvalues. This implies that the quadratic (t, f) detector (Eq. (12.1.16)) can be efficiently implemented via a bank of a few spectrograms.

12.1.3.2 An illustrative example

We now illustrate the (t, f) detection framework in the context of an important application in which the transmitted signal $v(t)$ travels over a *randomly time-varying* channel. Such a channel model is appropriate in many applications, including radar [1], wireless communications [4], and underwater acoustic communication [1]. In this case, the stochastic signal $s(t; t_o, f_o)$ is related to $v(t)$ as

$$\begin{aligned} s(t; t_o, f_o) &= \int_0^{T_m} h(t, \tau) v(t - \tau - t_o) e^{j2\pi f_o t} d\tau \\ &= \int_0^{T_m} \int_{-B_d}^{B_d} H(v, \tau) v(t - \tau - t_o) e^{j2\pi(v+f_o)t} dv d\tau, \end{aligned} \quad (12.1.18)$$

where $h(t, \tau)$ is the time-varying channel impulse response and the equivalent representation, $H(v, \tau) = \int h(t, \tau) e^{-j2\pi vt} dt$, is called the *delay-Doppler spreading function* [4]. In Eq. (12.1.18), the received signal is modeled as a linear combination (with stochastic weights) of a continuum of (t, f) shifted copies of $v(t)$. The parameters T_m and B_d are the *multipath* and *Doppler* spreads of the channel, respectively. The parameters (t_o, f_o) represent the *global* delay and Doppler shift encountered during propagation, and T_m and B_d represent the *local* (t, f) spreading around (t_o, f_o) produced by scattering. Each $(v, \tau) \in [-B_d, B_d] \times [0, T_m]$ represents a particular infinitesimal scatterer with stochastic gain $H(v, \tau)$. Under the wide-sense stationary uncorrelated scattering (WSSUS) model, $\{H(v, \tau)\}$ are modeled as a collection of uncorrelated Gaussian random variables [4].

The underlying Gaussian signal $w(t)$ in this case is characterized by

$$R_w(t_1, t_2) = \int_0^{T_m} \int_{-B_d}^{B_d} M(v, \tau) v(t_1 - \tau) v^*(t_2 - \tau) e^{j2\pi v(t_1 - t_2)} dv d\tau, \quad (12.1.19)$$

where $M(v, \tau) = \mathcal{E}[|H(v, \tau)|^2]$ is the channel *scattering function* which quantifies the relative power contributed by each scatterer. Note that $\text{WVS}_w(t, f)$ is supported on $[0, T_o + T_m] \times [-B_o - B_d, B_o + B_d]$. We focus on the locally optimum detector for simplicity which uses $\hat{R}_w = R_w$ in Eq. (12.1.16). Using the Weyl symbol in Eq. (12.1.6), it can be shown that the detection statistic (Eq. (12.1.15)) takes the form

$$\begin{aligned} \eta^{(t_o, f_o)}(r) &= \rho_r(t_o, f_o; \gamma = \text{WVS}_w) \\ &= \int \int W_r(u, v) \text{WVS}_w(u - t_o, v - f_o) du dv \end{aligned} \quad (12.1.20)$$

$$\begin{aligned}
&= \int_0^{T_m} \int_{-B_d}^{B_d} M(v, \tau) \left[\int \int W_r(u, v) W_v(u - t_o - \tau, v - f_o - v) du dv \right] dv d\tau \\
&= \int_{-T_m}^{T_m} \int_{-B_d}^{B_d} M(v, \tau) |F_r^v(t_o + \tau, f_o + v)|^2 dv d\tau.
\end{aligned} \tag{12.1.21}$$

Thus, the test for any given (t_o, f_o) can be computed by smoothing $W_r(t, f)$ in the neighborhood of (t_o, f_o) with $\text{WVS}_w(t, f)$ or by smoothing $|F_r^v(t, f)|^2$ in the neighborhood of (t_o, f_o) with $M(v, \tau)$. The overall GLRT detector (Eq. (12.1.16)) can be realized by (1) computing the $W_r(t, f)$ for $(t, f) \in [T_{\min}, T_{\max} + T_o + T_m] \times [-F_{\max} - B_o - B_d, F_{\max} + B_o + B_d]$, or by computing $|F_r^v(t, f)|^2$ for $(t, f) \in [T_{\min}, T_{\max} + T_m] \times [-F_{\max} - B_d, F_{\max} + B_d]$, (2) computing $\eta^{(t, f)}$ over \mathcal{S} by smoothing $W_r(t, f)$ with $\text{WVS}_w(t, f)$ as in Eq. (12.1.20) or by smoothing the spectrogram with $M(v, \tau)$ as in Eq. (12.1.21), and (3) comparing the maximum of $\{\eta^{(t, f)} : (t, f) \in \mathcal{S}\}$ to a threshold.

While the detector (Eq. (12.1.21)) is in a fairly simple form, the low-complexity implementation in terms of the eigenfunctions of $R_w(t_1, t_2)$ takes an equally simple form. Due to the finite duration and bandwidth of $v(t)$, the decision statistic (Eq. (12.1.21)) admits the following approximate sampled representation:

$$\eta^{(t_o, f_o)}(r) \approx \sum_{l=0}^{\lceil T_m B_o \rceil} \sum_{m=-\lceil B_d T_o \rceil}^{\lceil B_d T_o \rceil} \hat{M}(m/T_o, l/B_o) |F_r^v(t_o + l/B_o, f_o + m/T_o)|^2, \tag{12.1.22}$$

where $\hat{M}(v, \tau)$ represents a smooth version of $M(v, \tau)$ and the number of terms in the summation represents the delays and Doppler shifts that are *resolvable* at the receiver, which is also the number of dominant eigenvalues of $R_w(t_1, t_2)$. Note that the number of dominant eigenvalues is approximately $2T_m B_d T_o B_o$ which is typically relatively small since most practical channels are underspread ($2T_m B_d \ll 1$) [4].

12.1.4 EXTENSIONS

We now briefly discuss two extensions of the quadratic (t, f) detection framework described above. First, if the nuisance parameters are modeled as *random* with known probability density function, $p(t, f)$, the locally⁴ optimum Bayesian test statistic can be realized in the (t, f) domain as

$$\begin{aligned}
\eta_{\text{Bayes}}(r) &= \int_{\mathcal{S}} \rho_r(t, f; \text{WVS}_w) p(t, f) dt df \\
&= \int \int W_r(t, f) \widehat{\text{WVS}}_w(t, f) dt df,
\end{aligned} \tag{12.1.23}$$

where $\widehat{\text{WVS}}_w(t, f) = \int \int p(t, f) \text{WVS}_w(u - t, v - f) du dv$. The Bayesian detector computes a test statistic based on prior expectation of the likelihood of different values of the nuisance parameters. The (t, f) formulation of a quadratic detector in Eq. (12.1.14) has an interesting interpretation in view of the form of the Bayesian detector (Eq. (12.1.23)) and the *quasi-energy density* interpretation of the WD: the quadratic detector (Eq. (12.1.14)) weights the WD of the noisy received signal with our prior expectation of the distribution of signal energy in the (t, f) plane. This fact can also be exploited in

⁴The optimum Bayesian detector essentially corresponds to replacing $\rho_r(t, f; \text{WS}_{R_w})$ with $e^{-\rho_r(t, f; \text{WS}_{R_w})}$ in Eq. (12.1.23).

the design of rank-1 detectors when the underlying reference signal can only be inferred from noisy measurements [7]. In such case, $|\langle r, s \rangle|^2 = \langle W_r, W_s \rangle$, and if $W_s(t,f)$ is mostly concentrated along a curve, for example $W_s(t,f) = \delta(f - v(t))$, it can be estimated more accurately (as opposed to $s(t)$ directly) due to higher SNR along the curve in the (t,f) plane. A matched filter for the underlying signal can then be inferred via the instantaneous frequency or the spectral delay defined by the curve. Such weighted averaging of the TFR can also be exploited (in a suboptimal manner) to design detectors that suppress unwanted signals whose (t,f) support does not completely overlap the support of the desired signal.

Another generalization is to go beyond TFRs to joint signal representations in terms of variables other than time and frequency—time-scale representations (TSRs), for example [8]. The GLRT detection framework described above is best suited to representations that possess some covariance property analogous to Eq. (12.1.4). For example, TSRs are covariant to time shifts and scale changes. In general, such covariance properties are imposed via a parameterized family of unity operators [9]. Each family of unity operators defines a class of signal representations and these parameters of the unitary operators provide a model for nuisance parameters in GLRT detection scenarios. In essence, each family of unity operators defines a one-to-one correspondence between a particular class of detection scenarios and corresponding joint signal representations that serve as optimal detectors for such scenarios.

12.1.5 SUMMARY AND CONCLUSIONS

This section presents basic (t,f) formulations of optimum quadratic detection, along with a discussion of linear detectors and Bayesian detection. The GLRT formulation emphasized in this section exploits the degrees of freedom in a TFR (kernel and covariance properties) [6,9]. The same general framework can be extended to design data-driven detectors and classifiers that infer the optimal processor directly from data [10]. While these formulations emphasize *optimum* detection, a fruitful research direction is the (suboptimal) combination of exploratory data analysis with decision theory: exploiting the degrees of freedom in the TFR to isolate essential signal features that can serve as inputs to pattern classifiers [11]. See also companion Sections 12.5 and 12.6, which represent, respectively, a general approach for (t,f) matched filtering (Section 12.5) [12] and (t,f) features for classification (Section 12.6).

12.2 TIME-FREQUENCY SIGNAL ANALYSIS AND CLASSIFICATION USING MATCHING PURSUITS⁰

12.2.1 SIGNAL TIME-FREQUENCY STRUCTURES

For an appropriate time-frequency (t,f) analysis of nonstationary signals, it is important to match a (t,f) representation (TFR) with (t,f) structure of a signal for successful (t,f) analysis and classification (see Section 1.1.5). In general, TFRs are ideally matched to one or two (t,f) structures based on the properties they satisfy. However, some signals have multiple components with distinctively

⁰Authors: **A. Papandreou-Suppappola**, Arizona State University, Tempe, AZ, USA (papandreou@asu.edu) and **S.B. Suppappola**, Cirrus Logic Inc., Mesa, AZ, USA (seth.suppappola@cirrus.com). Reviewers: F. Auger and X. Xia.

different instantaneous frequency (IF). For example, in sonar acoustics, a received signal (e.g., a linear frequency-modulated (FM) signal) may be interfered by whistles from a group of dolphins (e.g., with hyperbolic IF). The Wigner distribution (WD) can accurately represent signals with constant or linear IF whereas the Altes Q-distribution can accurately represent signals with hyperbolic (dispersive) IF. However, neither one of the two quadratic TFRs can ideally analyze all the components of a signal consisting of both linear and hyperbolic IF as in the aforementioned example. The (t,f) analysis and classification of multicomponent signals is further complicated by the presence of cross terms when quadratic TFRs are used (see Ref. [13] and Section 4.2).

TFRs that can analyze signals with multiple (t,f) structures include the spectrogram [13], reassigned TFRs (see Ref. [14] and Section 7.2), and various adaptive TFRs [15]. The IF of various signals was estimated using TFRs [16]. Although they work well in many applications, such TFRs do not automatically yield the IF of a signal in closed form and do not always provide a well-concentrated representation without cross terms for analyzing signals with nonlinear (t,f) structures. Many natural or synthetic signals have different linear and nonlinear IF, so it is advantageous for an adaptive TFR to exactly match and classify their signal components.

12.2.2 MATCHING PURSUITS FOR ANALYSIS AND CLASSIFICATION

The matching pursuit (MP) iterative algorithm in Ref. [17] decomposes a signal into a linear expansion of waveforms selected from a redundant and complete dictionary . It uses successive approximations of the signal with orthogonal projections on dictionary elements. The dictionary consists of a basic Gaussian atom that is shifted and scaled. A TFR (called the *modified Wigner distribution* in Ref. [17]) is obtained as a weighted sum of the WD of each selected element. This TFR is free of cross terms and preserves signal energy, (t,f) shifts, and scale changes on the analysis signal. In order to analyze linear FM signals more efficiently with fewer waveforms, rotated Gaussian atoms were included in the dictionary in Ref. [18]. On the other hand, a wave-based dictionary consisting of wavefronts, resonances, and linear FM signals was used to process scattering data in Ref. [19].

Although the MP algorithm in Ref. [17] works well for many signals, it uses many Gaussian atoms to represent a signal component with nonlinear (t,f) characteristics. In addition, the modified WD is not well-concentrated along the nonlinear IF of some signal components. In order to exactly match linear or nonlinear FM signals, a MP is used with a dictionary of waveforms that includes complex sinusoids with linear or nonlinear phase functions such as logarithmic and power functions. The aim of the adaptive (modified) MP (AMMP) is to be able to analyze and correctly classify multicomponent signals where each component has a specific (and often different) monotonic phase function that can be written in closed form [20]. The advantage of using a dictionary that is matched to the analysis data is that only a small number of elements are used to decompose the signal, and the algorithm gives fast and parsimonious results. At each iteration, the algorithm *adaptively* chooses the best dictionary element according to some condition, identifies its (t,f) structure, and computes a specific TFR of the element. This TFR is adaptively chosen so that it provides a well-concentrated representation (without interference terms) of the selected element. The resulting TFR is formed as a weighted linear superposition of these matched TFRs. In this respect, it is similar to the TFR obtained in Ref. [21]. The computation of the TFR is only necessary for signal analysis. The algorithm can also be used to identify and classify each signal component. This can be useful, for example, in the classification of multiple received signals in a detection application. Thus, as will be shown, the AMMP decreases the

number of algorithm iterations, improves the (t, f) concentration of different multiple nonlinear FMs, and correctly classifies the IF of each signal component.

12.2.2.1 Adaptive matching pursuit and TFR

The AMMP is based on the algorithm in Ref. [17] but with major differences [20]. First, the AMMP uses more than one type of basic atom in the dictionary. Particularly, the dictionary consists of a large class of different basic atoms which have the form of a nonlinear FM signal (Section 5.6)

$$g(t; \xi, \lambda) = \sqrt{|\nu(t)|} e^{j2\pi \lambda \xi(\frac{t}{t_r})}, \quad (12.2.1)$$

which is uniquely specified by its FM rate λ and its monotonic phase function $\xi(b)$. Note that $\nu(t) = \frac{d}{dt} \xi(\frac{t}{t_r})$ is the IF of the signal in Eq. (12.2.1), and $t_r > 0$ is a reference time. The dictionary may consist of only one type of FM signal with fixed $\xi(b)$ in Eq. (12.2.1) or a linear combination of them including sinusoids with $\xi(b) = b$, linear FM signals with $\xi(b) = \text{sgn}(b)|b|^2$ (where $\text{sgn}(b)$ is -1 for $b < 0$, 0 for $b = 0$, and 1 for $b > 0$), hyperbolic FM signals with $\xi(b) = \ln b$, power FM signals with $\xi(b) = \text{sgn}(b)|b|^k$, and exponential FM signals with $\xi(b) = e^b$. The dictionary is formed by transforming the FM signal in Eq. (12.2.1) as

$$\begin{aligned} g(t; \xi, \lambda, \underline{\theta}) &= \left(\mathcal{S}_\tau \mathcal{C}_a \mathcal{G}_c^{(\xi)} g(\xi, \lambda) \right)(t) = \sqrt{|a|} g(a(t - \tau); \xi, \lambda) e^{j2\pi c \xi(a(\frac{t-\tau}{t_r}))} \\ &= \sqrt{|a \nu(a(t - \tau))|} e^{j2\pi (c + \lambda) \xi(a(\frac{t-\tau}{t_r}))}, \end{aligned} \quad (12.2.2)$$

with the parameter vector $\underline{\theta} = [c, a, \tau] \in (\Theta = \mathbb{R}^3)$. The dictionary is normally formed with a linear (i.e., grid) spacing of the parameters in $\underline{\theta}$, although this is not a requirement. The unitary operators $\mathcal{G}_c^{(\xi)}$, \mathcal{C}_a , and \mathcal{S}_τ result in an IF shift c , scale change a , and constant time shift τ , respectively, of the FM signal. Specifically, the operators transform a signal $x(t)$ as $(\mathcal{G}_c^{(\xi)} x)(t) = x(t) e^{j2\pi c \xi(t/t_r)}$ (where $\xi(b)$ is the same function as in Eq. (12.2.1)), $(\mathcal{C}_a x)(t) = \sqrt{|a|} x(a t)$, and $(\mathcal{S}_\tau x)(t) = x(t - \tau)$. In Eq. (12.2.2), the transformation $\mathcal{G}_c^{(\xi)}$ results in a constant shift (from λ to $c + \lambda$) of the FM rate of the nonlinear FM signal instead of the constant frequency shift used in Ref. [17]. Hence, λ is not a parameter of optimization and is instead used to specify a base FM rate in Eq. (12.2.1). Shifts in the FM rate are effected by varying the c parameter. Thus, the signals considered may be wideband as well as dispersive, as a shift of the IF is a better matched transformation (covering the entire (t, f) plane) than a constant frequency shift. With appropriate normalization, the energy of $g(t; \xi, \lambda, \underline{\theta})$ is restricted to unity for every $\underline{\theta}$ in order to ensure energy preservation when $\xi(b)$ is fixed [17, 20]. The iterative procedure of the AMMP first projects the analysis signal $x(t) = (R_0 x)(t)$ onto each element of the dictionary, and selects¹ $g(t; \xi_0, \lambda, \underline{\theta}_0)$ based on the condition² $|\langle x, g(\xi_0, \lambda, \underline{\theta}_0) \rangle| \geq |\langle x, g(\xi, \lambda, \underline{\theta}) \rangle|, \forall \underline{\theta} \in \Theta$ and for all possible phase functions $\xi(b)$ of the dictionary elements. This ensures that the signal component with the highest energy is separated first by being matched to the element $g(t; \xi_0, \lambda, \underline{\theta}_0)$ that yields the largest correlation with the signal. This results in

$$x(t) = \beta_0 g(t; \xi_0, \lambda, \underline{\theta}_0) + (R_1 x)(t) \quad (12.2.3)$$

¹Note that a subscript n in the parameters R_n , $\xi_n(b)$, $\underline{\theta}_n$, τ_n , and c_n and a superscript n in a TFR $T^n(t, f)$ indicate the algorithm parameters at the $(n + 1)$ th iteration.

²The inner product is defined as $\langle x, g \rangle = \int_{-\infty}^{\infty} x(t) g^*(t) dt$.

with the expansion coefficient $\beta_0 = \langle x, g(\xi_0, \lambda, \underline{\theta}_0) \rangle$. The function $\xi_0(b)$ is the phase function of the first selected element. For example, if this element is a hyperbolic FM signal, then $\xi_0(b) = \ln b$.

The second major difference of the AMMP algorithm from the one proposed in Ref. [17] is that it does not compute the WD of selected elements to form the modified WD. Instead, it adaptively uses the information that the first selected waveform has phase function $\xi_0(b)$ to compute its IF shift covariant *dispersively warped Wigner distribution* (DWD) (Section 5.6). The DWD is a warped version of the WD, $W_y(t, f) = \int_{-\infty}^{\infty} y(t + \tau/2) y^*(t - \tau/2) e^{-j2\pi f \tau} d\tau$, with the warping based on a monotonic and (possibly) nonlinear parameter function $\zeta(b)$. In particular,

$$W_x^{(\xi)}(t, f) = W_y \left(t_r \zeta \left(\frac{t}{t_r} \right), \frac{f}{t_r \mu(t)} \right), \quad (12.2.4)$$

where $\mu(t) = \frac{d}{dt} \zeta \left(\frac{t}{t_r} \right)$, and $y(t) = (\mathcal{W}_\zeta x)(t) = t_r |\mu \left(t_r \zeta^{-1} \left(\frac{t}{t_r} \right) \right)|^{-1/2} x \left(t_r \zeta^{-1} \left(\frac{t}{t_r} \right) \right)$ (Section 5.6) is the warped signal. Note that a specific DWD is obtained simply by fixing its parameter function $\zeta(b)$. By matching $\zeta(b)$ in Eq. (12.2.4) to be equal to the phase function $\xi_0(b)$ in Eq. (12.2.3) (i.e., if $\zeta(b) = \xi_0(b)$), an *adaptive representation for multiple structures* (ARMUS) TFR, at this first iteration, is simply given as $T_x^0(t, f) = |\beta_0|^2 W_{g(\xi_0, \underline{\theta}_0)}^{(\xi_0)}(t, f)$. At the second iteration, the residual function $(R_1 x)(t)$ is obtained by solving Eq. (12.2.3), and it is decomposed in a similar manner as the signal $x(t)$. At the $(n+1)$ th iteration, the condition

$$|\langle R_n x, g(\xi_n, \underline{\theta}_n) \rangle| \geq |\langle R_n x, g(\xi, \underline{\theta}) \rangle|, \quad \forall \underline{\theta} \in \Theta \quad (12.2.5)$$

is used to decompose the n th residual function $(R_n x)(t) = \beta_n g(t; \xi_n, \underline{\theta}_n) + (R_{n+1} x)(t)$, where the expansion coefficient is

$$\beta_n = \langle R_n x, g(\xi_n, \underline{\theta}_n) \rangle. \quad (12.2.6)$$

The DWD of $(R_n x)(t)$ is also obtained adaptively to match the (t, f) structure of the n th residual function by letting $\zeta(b) = \xi_n(b)$ in Eq. (12.2.4).

After N iterations, the MP algorithm results in the signal decomposition

$$x(t) = \sum_{n=0}^{N-1} \beta_n g(t; \xi_n, \lambda, \underline{\theta}_n) + (R_N x)(t). \quad (12.2.7)$$

As the dictionary is complete [20], any signal can be represented as in Eq. (12.2.7) with $N = \infty$ (yielding $(R_N x)(t) = 0$) [17]. In fact, when the signal components match the (t, f) structure of the dictionary elements, the algorithm converges quickly. A maximum number of iterations and an acceptable small residue energy compared to the data energy are used as stopping criteria [17]. The resulting ARMUS of the signal at the N th iteration is the weighted sum of the corresponding DWD of each selected dictionary element with the weights $|\beta_n|^2$ defined in Eq. (12.2.6), i.e.,

$$\text{ARMUS}_x(t, f) = T_x^{N-1}(t, f) = \sum_{n=0}^{N-1} |\beta_n|^2 W_{g(\xi_n, \underline{\theta}_n)}^{(\xi_n)}(t, f). \quad (12.2.8)$$

Note that the same DWD (with fixed $\xi(b)$) may be used in Eq. (12.2.8) for components with the same (t, f) structure but different FM rate.

It is important to note that the success of the algorithm depends highly on the choice of dictionary elements. As a result, preprocessing the data might be necessary for improved performance and to avoid a mismatch between the analysis data and the dictionary elements used. Although this analysis technique follows directly from the MP in Ref. [17] with some simple modifications, it is very powerful for analysis and classification once the dictionary is matched to the data.

12.2.2.2 Properties of the algorithm and the TFR

An important property of the MP in Eq. (12.2.7) is its covariance to certain signal changes. Consider the decomposed signal $x(t) = \sum_{n=0}^{\infty} \beta_n g(t; \xi, \lambda, \underline{\theta}_n)$ in Eq. (12.2.7) with $N = \infty$ and with identical (t, f) structure dictionary elements (i.e., let $\xi_n(b) = \xi(b), \forall n$). If the FM rate of a nonlinear FM signal $x(t)$ is shifted by a constant amount to form $y(t) = (G_u^{(\xi)} x)(t) = x(t) e^{j2\pi u \xi(t/t_r)}$, then its MP is simply given as $y(t) = \sum_{n=0}^{\infty} \beta_n g(t; \xi, \lambda, \hat{\underline{\theta}}_n)$. Note that the expansion coefficients β_n are not affected by this signal change. The parameter vector changes to $\hat{\underline{\theta}}_n = [c_n + u, a_n, \tau_n]$ indicating that the time shifts τ_n and the scale changes a_n remain the same, whereas the dictionary elements undergo a constant shift in their FM rate from $(\lambda + c_n)$ to $(\lambda + c_n + u)$. Note that if $\xi(b)$ is a power or a logarithmic function, then the corresponding AMMP is also covariant to scale changes [20].

The ARMUS TFR in Eq. (12.2.8) also satisfies various desirable properties. By simply combining the DWDs of each selected dictionary element, no cross terms are introduced in the TFR. The TFR also preserves the (t, f) structure of each analysis signal component and provides a highly concentrated representation of each component as it does not apply any smoothing. Specifically, the DWD with parameter $\xi(b)$ of a nonlinear FM signal with phase function $\xi(b)$ results in the highly concentrated representation $DWD_{g(\xi, \lambda)}(t, f; \xi) = |\nu(t)| \delta(f - \lambda \nu(t))$ (Section 5.6) with $\nu(t) = \frac{d}{dt} \xi(t/t_r)$. If a particular application uses signal components with only one type of (t, f) structure, then we should form the dictionary using the corresponding nonlinear FM signal with matched IF. In such cases, the ARMUS satisfies other desirable signal properties such as preservation of signal energy and changes in the analysis signal's FM rate [20]. The TFR also preserves scale changes for dictionary elements that are either hyperbolic or power FM signals.

12.2.2.3 Implementation issues

Since the algorithm parameters vary in order to form the dictionary elements for the MP, the computation is intensive. But a dictionary with elements which approximately span the data in (t, f) structure can be formed with some preprocessing. Thus, the algorithm iterates more rapidly. Additional speedup is possible if the matched DWD of each dictionary element is computed ahead of time. Since the last operation on the basic atom in Eq. (12.2.2) is time shifting, the inner products in the AMMP condition in Eq. (12.2.5) can be computed as a cross-correlation instead of introducing another layer of dictionary elements over all possible time shifts. This increases the computational speed since correlations can be implemented by fast Fourier transform (FFT). The memory consumption by the dictionary can also be reduced since additional dictionary elements are not needed for every time shift. Moreover, since the elements do not change and the residual data is constant during a given AMMP algorithm iteration, additional speedup could be achieved by precomputing and storing the FFTs of these sequences.

If the signal components are well-separated in time, the algorithm can simply find the time support and phase function of each selected element and then use the information to analyze the actual data

(instead of the selected waveforms) with its matched DWD. This greatly reduces computation for just a few DWDs. If classification is needed without analysis, there is no need to compute the TFR. The AMMP provides the IF and the FM rate of each component. It also provides other relevant parameters depending on how adapted the dictionary is to a particular application.

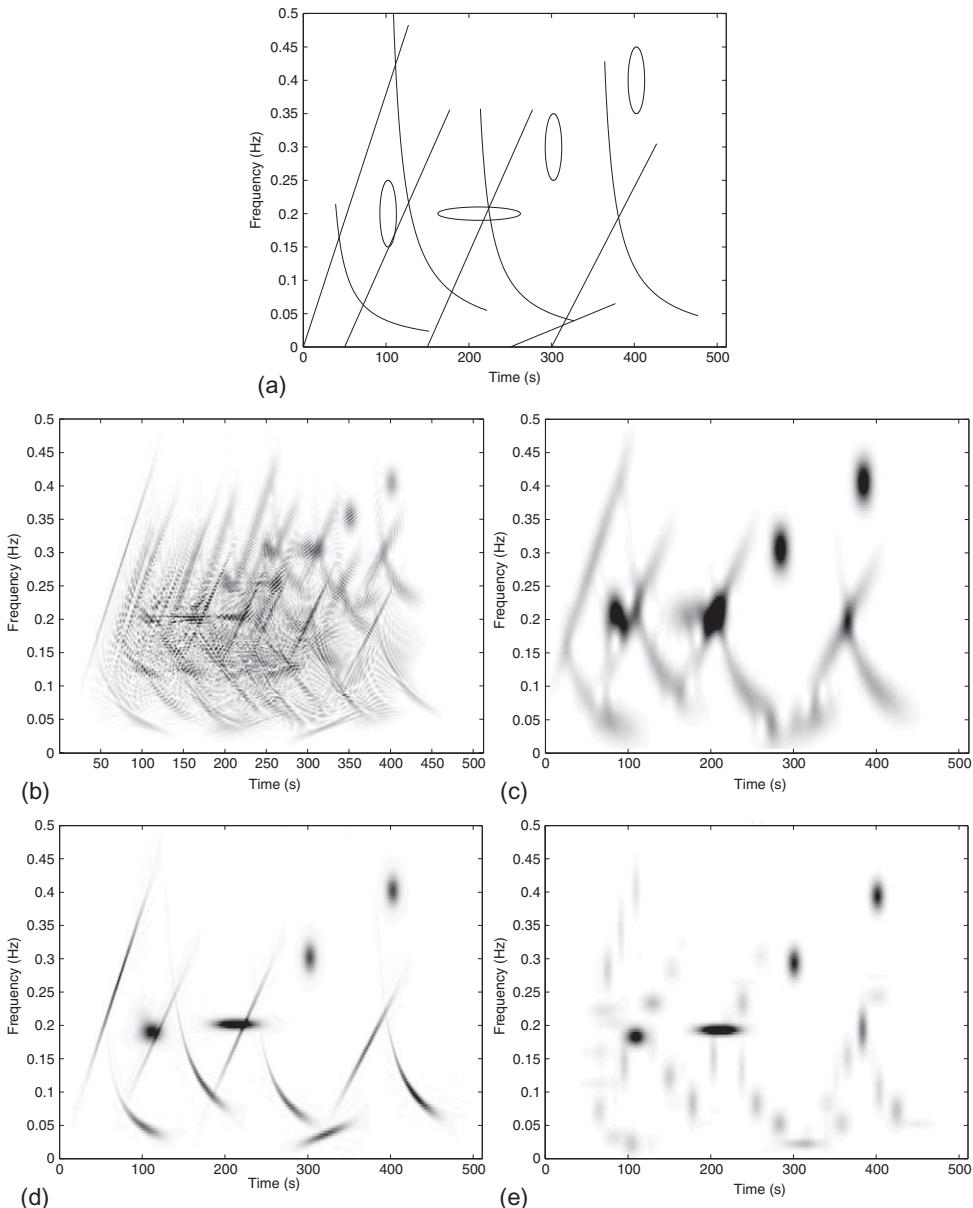
12.2.3 SIMULATION EXAMPLE USING A SYNTHETIC MULTICOMPONENT SIGNAL

The signal used in the simulation is 512 points long and consists of 13 components: 4 windowed hyperbolic FM signals, 5 windowed linear FM signals, and 4 Gaussian waveforms, all with varying parameters. The “ideal” (t,f) representation shown in Fig. 12.2.1(a) is obtained by plotting the IF of each of the hyperbolic and linear FM signals (hyperbolae and lines with positive slope) and the characteristic (t,f) curve of each Gaussian component (ellipses). The WD in Fig. 12.2.1(b) suffers from cross terms and makes it difficult to identify the true (t,f) structure of each component. On the other hand, the spectrogram in Fig. 12.2.1(c) suffers from loss of resolution due to smoothing that complicates the identification of the exact number of signal terms. Furthermore, the spectrogram cannot provide parameter values for any signal component.

The dictionary is formed using hyperbolic FM signals, linear FM signals, and Gaussian waveforms. Note that the Gaussian elements are not in the form of Eq. (12.2.1). However, they were included in the dictionary for a fair comparison with the MP in Ref. [17] and due to the *a priori* knowledge of their presence in the data. After 20 iterations, the data is approximated well as demonstrated by overlaying the signal with its expansion. The ARMUS TFR in Fig. 12.2.1(d) provides a highly concentrated representation for all 13 components without outer cross terms or loss of resolution. This is because it adaptively computes the Altes Q-distribution for selected elements with hyperbolic (t,f) characteristics and the WD for selected elements with linear or constant (t,f) characteristics. Note that the mild spreading of the signal components and some inner interference terms is due to the fact that the data was windowed for processing. As the algorithm iterated 20 times, the ARMUS in Fig. 12.2.1(d) is a TFR of all 20 elements that the algorithm selected. However, only the 13 elements matching the actual signal components are visible. This is because the energy term of the signal residues was very small for 7 of the 20 iterations. If (t,f) analysis is not required, the ARMUS TFR computation step is removed from the algorithm, and the results are used only for classification. In this example, the components are classified to three signal types (hyperbolic FM, linear FM, and Gaussian). Furthermore, the values of the FM rate of each hyperbolic and linear waveform as well as the variance of each Gaussian component are directly obtained and found to be very similar to the parameters of the actual data.

For further comparison, the MP with only Gaussian dictionary elements was used to decompose the signal, and then analyze it using the modified WD [17] as shown in Fig. 12.2.1(e). After 275 iterations, the TFR does not provide a concentrated representation that can easily identify the (t,f) structure of the linear and hyperbolic FM components even though no cross terms are present. Also, the algorithm does not provide a closed form estimate of the IF and FM rate of the signal components for classification. On the other hand, as it is well-matched to Gaussians, the four Gaussian signal components are highly concentrated.

For comparison, the speed of the various methods can be contrasted, when performed on a Pentium III based PC running MATLABTM under Linux. Since the Wigner distribution and the spectrogram use no form of MP iteration, they provide essentially instant results requiring only a few seconds to

**FIGURE 12.2.1**

(a) A linear combination of the ideal (t, f) structure of each component of a signal consisting of four windowed hyperbolic FM signals, five windowed linear FM signals, and four windowed Gaussian signals. Note that the (t, f) structure of a Gaussian waveform is represented as an ellipse delineating the approximate time and frequency support of the signal component. The signal is analyzed using (b) the Wigner distribution, (c) the spectrogram, (d) the ARMUS TFR, and (e) the modified WD in Ref. [17].

generate a plot. The ARMUS representation required approximately 5 minutes to compute, whereas the modified Wigner distribution consumed approximately 45 minutes. The slow processing speed of the modified Wigner distribution is primarily due to the high number of terms (i.e., iterations) required to approximate the analysis signal since Gaussian dictionary elements are not well matched to all of the analysis signal components. In fact, it takes several Gaussians to represent a single hyperbolic FM signal. The ARMUS representation, however, also contains hyperbolic and linear FM dictionary elements and is consequently better matched to the analysis signal.

The AMMP method is used for real data classification in Ref. [22] and applied for structural health monitoring of composite structures in Ref. [23].

12.2.4 SUMMARY AND CONCLUSIONS

In this section, we have presented a method of analyzing complex multicomponent time-frequency signal structures without the usual trade-off of (t, f) resolution versus cross terms. The iterative approach is based on the MP of Ref. [17] but extended to include non-Gaussian signal types. Additionally, we generate a TFR which is free of cross terms by summing TFRs for the decomposed signal components. The TFR for each individual component is chosen to be ideally suited for the particular signal type extracted by the algorithm. Finally, since the TFR both represents and parameterizes individual components, this algorithm lends itself nicely as a feature extraction tool for signal classification purposes.

12.3 SYSTEM IDENTIFICATION USING TIME-FREQUENCY FILTERING⁰

12.3.1 PROBLEM DESCRIPTION

A discrete time system identification problem can be stated as follows:

$$y[n] = \sum_k q[n - k]x[k] + \epsilon[n], \quad (12.3.1)$$

where $x[n]$ is a transmitted signal, $q[n]$ is the impulse response of a linear time invariant (LTI) system, $\epsilon[n]$ is an additive noise, and $y[n]$ is the received signal. The problem is to identify the LTI system transfer function $Q(f)$, i.e., the Fourier transform, of $q[n]$ given the input and the output signals $x[n]$ and $y[n]$.

The conventional method for solving the above problem is the least-squares solution method that is equal to the cross-spectral method in stationary cases, i.e., the system transfer function $Q(f)$ can be estimated by (see, e.g., Refs. [24,25])

$$Q(f) = \frac{S_{xy}(f)}{S_{xx}(f)}, \quad (12.3.2)$$

where $S_{xy}(f)$ is the cross-spectrum of $x[n]$ and $y[n]$, and $S_{xx}(f)$ is the auto-spectrum of $x[n]$. When the additive noise $\epsilon[n]$ in Eq. (12.3.1) is a zero-mean Gaussian process and statistically independent of the

⁰Author: X.-G. Xia, University of Delaware, Newark, DE, USA (xxia@ee.udel.edu). Reviewers: P.M. Oliveira and S. Qian.

input signal $x[n]$, the estimate in Eq. (12.3.2) is asymptotically unbiased but the performance is limited by the noise variance or the signal-to-noise ratio (SNR). When the SNR is low, the performance of the estimate in Eq. (12.3.2) is poor as we will see later. Since the auto-spectrum of the input signal $x[n]$ is in the denominator in the estimate equation (12.3.2), the input signal is, in general, chosen as a pseudorandom signal with flat spectrum. With these types of input signals, noise reduction techniques before system identification do not apply.

In the following, we introduce a different technique [26] for the system identification problem. The main idea is as follows. Instead of pseudorandom signal $x[n]$, chirp type signals are transmitted as training signals, which have wideband characteristics in the frequency domain but are concentrated in the joint time-frequency $((t,f))$ domain. The (t,f) concentration property usually holds after passing through an LTI system (this will be seen later). Since a joint (t,f) distribution usually spreads noise and localizes signals, in particular chirps, the receiver may use a (t,f) analysis technique to map the received signal $y[n]$ from the time domain into the joint (t,f) domain. In this way, the SNR can be significantly increased in the joint (t,f) domain [27]. Furthermore, (t,f) filtering can be used in the (t,f) plane to reduce the noise and the SNR in the time domain can be increased and therefore the system identification after denoising can be improved. Some applications of this approach can be found in Refs. [28–31].

12.3.2 TIME-FREQUENCY FILTERING

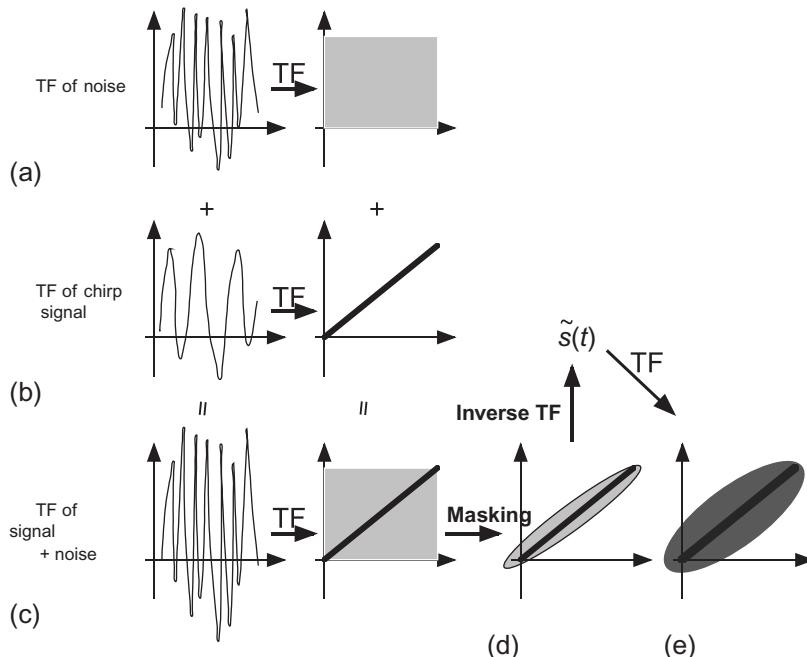
The (t,f) filtering process (described in Chapter 11) includes three major components: (t,f) analysis or mapping, using a tool such as a (t,f) distribution, from the time domain to the joint (t,f) domain, a masking or filtering in the joint (t,f) plane, and a (t,f) synthesis from the joint (t,f) plane back to the time domain, as illustrated in Fig. 12.3.1 for a chirp signal.

The (t,f) mapping in (t,f) filtering may not be one to one as the Fourier transform based filtering in the Fourier domain. In other words, not every signal $S[k, l]$ defined in the joint (t,f) plane corresponds to a time domain signal $s[n]$ such that the (t,f) mapping of the time domain signal $s[n]$ is exactly equal to $S[k, l]$. This causes problems in filtering in the (t,f) domain, since the filtered signal in the (t,f) domain may not correspond to any time domain signal as shown in Fig. 12.3.1(a)–(e). An intuitive solution for this problem is to take the least-squares error (LSE) solution in the time domain (see, for example, Ref. [32]). The LSE, however, usually does not have a desired (t,f) characteristics in the (t,f) domain. When a signal is very long, the computational load for the LSE solution is significantly higher because of the inverse matrix computation. Based on these observations, an iterative algorithm based on the discrete Gabor transform (DGT) was proposed in Ref. [33]; see Section 11.2. Conditions on the convergence, properties of the limit signals, and the relationship between the LSE solutions and solutions from the iterative algorithms were obtained in Ref. [34], where a significant improvement over the LSE solution was also shown. In the remainder of this section, we adopt the DGT as the (t,f) analysis. Other approaches can be found in Ref. [35] (see also Sections 11.2 and 11.6).

12.3.2.1 Discrete Gabor transform

Let a signal $s[n]$, a synthesis window function $h[n]$ and an analysis window function $\gamma[n]$ be all periodic with same period L . Then we have,

$$s[k] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} C_{m,n} h_{m,n}[k], \quad (12.3.3)$$

**FIGURE 12.3.1**

Graphical illustration of the (t, f) filtering process.

$$C_{m,n} = \sum_{k=0}^{L-1} s[k] \gamma_{m,n}^*[k], \quad (12.3.4)$$

$$h_{m,n}[k] = h[k - m\Delta M] W_L^{n\Delta N k}, \quad (12.3.5)$$

$$\gamma_{m,n}[k] = \gamma[k - m\Delta M] W_L^{n\Delta N k}, \quad (12.3.6)$$

where $W_L = \exp(j2\pi/L)$, $j = \sqrt{-1}$, and, ΔM and ΔN are the time and the frequency sampling interval lengths, and M and N are the numbers of sampling points in the time and the frequency domains, respectively, $M \cdot \Delta M = N \cdot \Delta N = L$, $MN \geq L$ (or $\Delta M \Delta N \leq L$). The coefficients $C_{m,n}$ are called the **discrete Gabor transform** (DGT) of the signal $s[k]$ and the representation equation (12.3.3) is called the **inverse discrete Gabor transform** (IDGT) of the coefficients $C_{m,n}$. Let H and G be the following $\Delta M \Delta N$ by L and MN by L matrices, respectively: the element at the $(mN+n)$ th row and the k th column in H is

$$h[k - m\Delta M] W_L^{n\Delta N k}, \quad 0 \leq m \leq M-1, 0 \leq n \leq N-1, 0 \leq k \leq L-1,$$

and the element at the k th row and the $(mN + n)$ th column in G is

$$\gamma_{m,n}^*[k] = \gamma^*[k - m\Delta M] W_L^{-n\Delta N k}, \quad 0 \leq m \leq M - 1, 0 \leq n \leq N - 1, 0 \leq k \leq L - 1.$$

Then, the DGT and IDGT can be rewritten as

$$\mathbf{C} = G\mathbf{s} \quad \text{and} \quad \mathbf{s} = H\mathbf{C}, \quad (12.3.7)$$

where the following condition is needed

$$HG = I, \quad (12.3.8)$$

where I is the $L \times L$ identity matrix. (For more details, see Ref. [33] and Sections 6.3 and 11.2.)

12.3.2.2 Iterative (t, f) filtering algorithm

Let D denote the mask in the (t, f) plane, i.e., an MN by MN diagonal matrix with diagonal elements either 0 or 1. Let \mathbf{s} be a signal with length L in the time domain. The first step in the time-variant filtering is to mask the DGT of \mathbf{s} : $\mathbf{C}_1 = DG\mathbf{s}$, where D masks a desired domain in the (t, f) plane. Since the DGT G is a redundant transformation, the IDGT of \mathbf{C}_1 , $H\mathbf{C}_1$, may not fall in the mask. In other words, in general,

$$H\mathbf{C}_1 \neq DGH\mathbf{C}_1, \quad (12.3.9)$$

which is illustrated in Fig. 12.3.1(e). An intuitive method to reduce the difference between the right and the left hand sides of Eq. (12.3.9) is to mask the right hand side of Eq. (12.3.9) again and repeat the procedure, which leads to the following iterative algorithm:

$$\mathbf{s}_0 = \mathbf{s}, \quad (12.3.10)$$

$$\mathbf{C}_{l+1} = DG\mathbf{s}_l, \quad (12.3.11)$$

$$\mathbf{s}_{l+1} = H\mathbf{C}_{l+1}, \quad (12.3.12)$$

$$l = 0, 1, 2, \dots$$

For more details, see Section 11.2 and Ref. [34].

12.3.3 DENOISING FOR RECEIVED SIGNALS

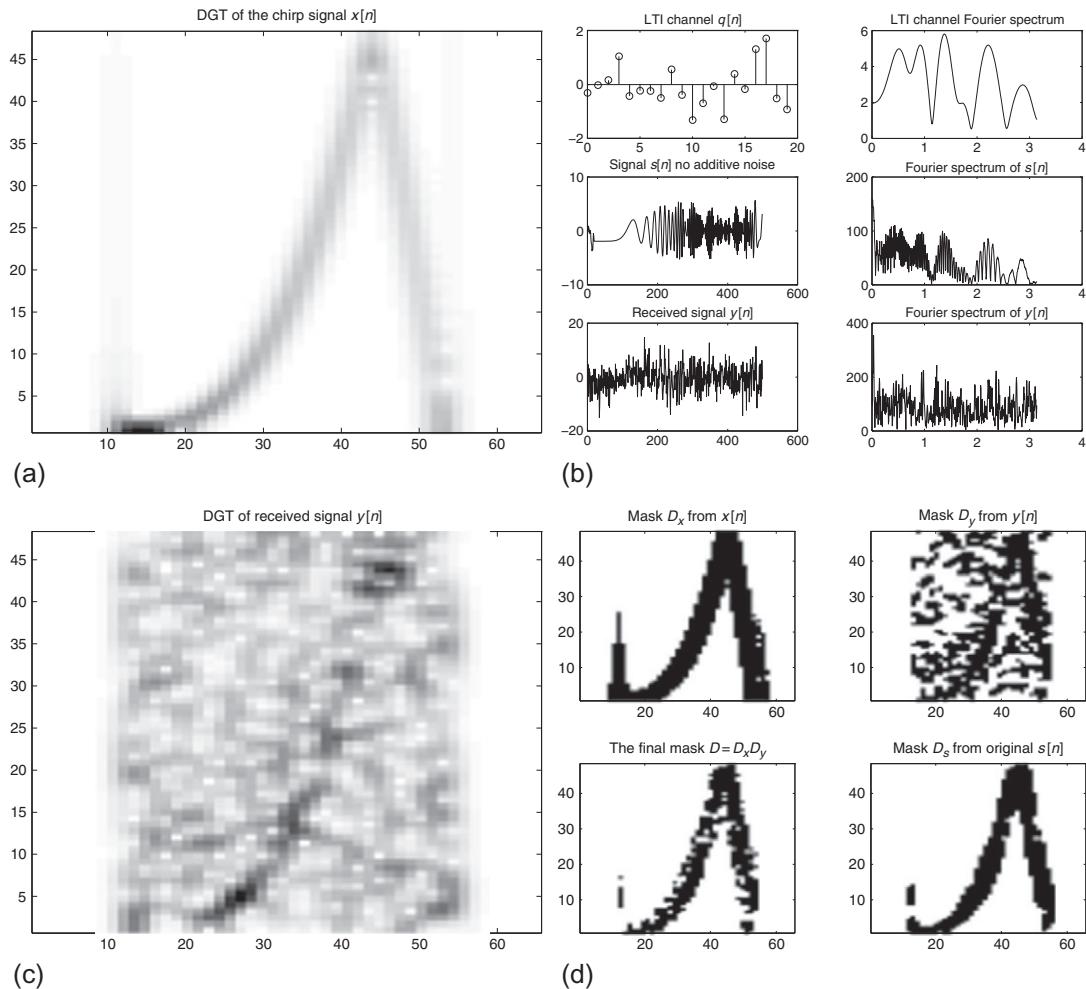
We first describe some parameters. The training signal $x[n]$ is

$$x[n] = \cos \left(2\pi \left[\frac{n+15}{150} \right]^4 \right), \quad n = 0, 1, \dots, 499. \quad (12.3.13)$$

The synthesis and analysis window functions $h[n]$ and $\gamma[n]$ are the orthogonal-like functions from Refs. [33,34] and Section 11.2. Figure 12.3.2(a) shows the DGT of $x[n]$.

We use 20 tap LTI systems in our numerical examples. The system model is

$$y[n] = \sum_{k=0}^{N-1} q[k]x[n-k] + \epsilon[n], \quad (12.3.14)$$

**FIGURE 12.3.2**

- (a) The discrete Gabor transform of signal $x[n]$. (b) An example of LTI channel $q[n]$, signal $s[n]$ and received signal $y[n]$, and their Fourier spectrum, where the SNR = -4.5 dB for the additive white Gaussian noise.
 (c) The discrete Gabor transform of a received signal $y[n]$ with SNR = -4.5 dB.
 (d) Example of masks D_x from $x[n]$, D_y from $y[n]$, the final mask $D = D_x D_y$, and the true mask D_s from $s[n]$.

where $N = 20$, $\epsilon[n]$ is additive white Gaussian noise, independent of $x[n]$, and

$$s[n] = \sum_{k=0}^{N-1} q[k]x[n-k] \quad (12.3.15)$$

is considered as the signal. The original SNR for the received signal is calculated by

$$10\log_{10} \left(\frac{\sum_{n=0}^{499} |s[n]|^2}{\sum_{n=0}^{499} |\epsilon[n]|^2} \right).$$

In the following, we randomly generate the channel $q[n]$. As an example, a random channel, the noiseless signal $s[n]$, the received time waveform $y[n]$ with SNR = -4.5 dB, and their Fourier spectrum are shown in Fig. 12.3.2(b). The DGT of the received signal $y[n]$ with -4.5 dB SNR is shown in Fig. 12.3.2(c). In Fig. 12.3.2(c), the chirp pattern is still visible in the joint time and frequency plane although that is impossible in the time or the frequency domain alone in Fig. 12.3.2(b).

12.3.3.1 Mask D design

The pattern in the DGT domain of the above $s[n]$ in Eq. (12.3.15) is similar to the one for the signal $x[n]$ in Fig. 12.3.2(a). It is true for this and some other examples. The reason can be found in the following analytic argument.

Assume the chirp signal $x[n] = \exp(jcn^r)$ for some constants $r \geq 2$ and $c \neq 0$. Then,

$$s[n] = \sum_k q[k]x[n-k] = x[n] \sum_k q[k] \exp\left(jc \sum_{l=0}^{r-1} c_l n^l k^{r-l}\right),$$

which is dominated by the original chirp $x[n]$ for finite tap LTI systems $q[k]$. It is because the highest chirp order of $s[n]$, r , and the corresponding chirp rate are the same as those of $x[n]$, while the chirp order for the above multiplier of $x[n]$ in $s[n]$

$$\sum_k q[k] \exp\left(jc \sum_{l=0}^{r-1} c_l n^l k^{r-l}\right)$$

is only $r-1$. As a special case, when $r=2$, $s[n] = x[n]\tilde{Q}(\frac{cn}{\pi})$, where $\tilde{Q}(f)$ is the Fourier transform of the signal $q[n]x[n]$: $\tilde{Q}(f) = \sum_k q[k]x[k] \exp(-j2\pi fk)$. When the channel $q[n]$ has only a finite tap, the function $\tilde{Q}(f)$ is usually a smooth signal.

Since the transmitted signal $x[n]$ is known to both transmitter and the receiver, by the above property its pattern in the DGT domain may help in designing a mask in the DGT domain for filtering noise, which motivates the following design algorithm.

A Mask Design Procedure:

Step 1. Implement the DGT, obtaining the coefficients $C_{m,n}$ of the transmitted signal $x[k]$.

Step 2. Threshold the DGT coefficients $C_{m,n}$ and take a mask D_x from $C_{m,n}$:

$$D_x(m, n) = \begin{cases} 1, & \text{if } |C_{m,n}| > t_0, \\ 0, & \text{otherwise,} \end{cases}$$

where t_0 is a predesigned positive number that is called a *thresholding constant*.

Step 3. Implement Steps 1-2 for the received signal $y[k]$ and design a mask D_y with thresholding constant t_1 from the DGT coefficients of $y[n]$ with another predesigned constant $t_1 > 0$.

Step 4. The final mask is the product of D_x and D_y : $D = D_x D_y$.

Since the DGT of the signal $x[n]$ usually dominates the DGT of the signal $s[n]$, the pattern in the DGT domain of the signal $s[n]$ is usually in a close neighborhood of the pattern in the DGT domain of $x[n]$. Therefore, the mask D_x is usually designed so that it covers a relatively larger area, i.e., the thresholding constant t_0 in Step 2 is usually chosen not too large. Since the received signal $y[n]$ is from a noisy channel, the resolution of its DGT pattern may be reduced and therefore the thresholding constant t_1 in Step 3 is usually chosen not too small. Otherwise, the mask D_y will cover too much area which is not desirable. Let us see an example. The mask D_x from $x[n]$, the mask D_y from $y[n]$, their product $D = D_x D_y$, and the mask D_s from the true signal $s[n]$ are shown in Fig. 12.3.2(d), respectively. The SNR in this case is $\text{SNR} = -1.4 \text{ dB}$. The thresholding constants in Steps 1-3 are $t_0 = 0.12 \cdot \max(\text{DGT}(x))$ and $t_1 = 0.15 \cdot \max(\text{DGT}(y))$. It should be pointed out that the above mask design procedure may be improved by using more sophisticated designs. Possible improvements are (i) to find the optimal thresholding constants t_0 and t_1 by training a large number of signals and systems; (ii) to use more sophisticated statistical detection method in the DGT domain for the received signal $y[n]$ instead of a simple thresholding in Step 3; (iii) to smooth the mask $D = D_x D_y$ since the true mask D_s is usually smooth due to the nature of a chirp signal but D_y from the noisy signal $y[n]$ may not be smooth. Some morphological operations, such as dilation, may be used for smoothing the mask D .

12.3.3.2 Denoising experiments

We now implement the denoising method using the iterative (t,f) filtering algorithm with three masking techniques: using the mask $D = D_x$ from the transmitted signal; using the mask $D = D_y D_x$ as designed by Steps 1-4; using the true mask $D = D_s$. We run 100 tests in terms of different LTI systems $q[n]$ (randomly generated) and different additive white Gaussian noises $\epsilon[n]$ for each masking method and take their mean SNR. Nine iterative steps are used in the iterative algorithm Eqs. (12.3.10)–(12.3.12). Figure 12.3.3(a) shows the curves of the mean SNR versus iterative steps for the three masking methods.

12.3.4 SYSTEM IDENTIFICATION

In this section, we first use the iterative time-variant filter Eqs. (12.3.10)–(12.3.12) developed in the previous sections to reduce the additive white Gaussian noise $\epsilon[n]$ from the received signal $y[n]$. For simplicity of calculation we choose the first masking method, i.e., the mask $D = D_x$, throughout this section. With this mask, two iterations are used in the (t,f) filtering. We then implement the conventional system identification method, as shown in Fig. 12.3.3(b).

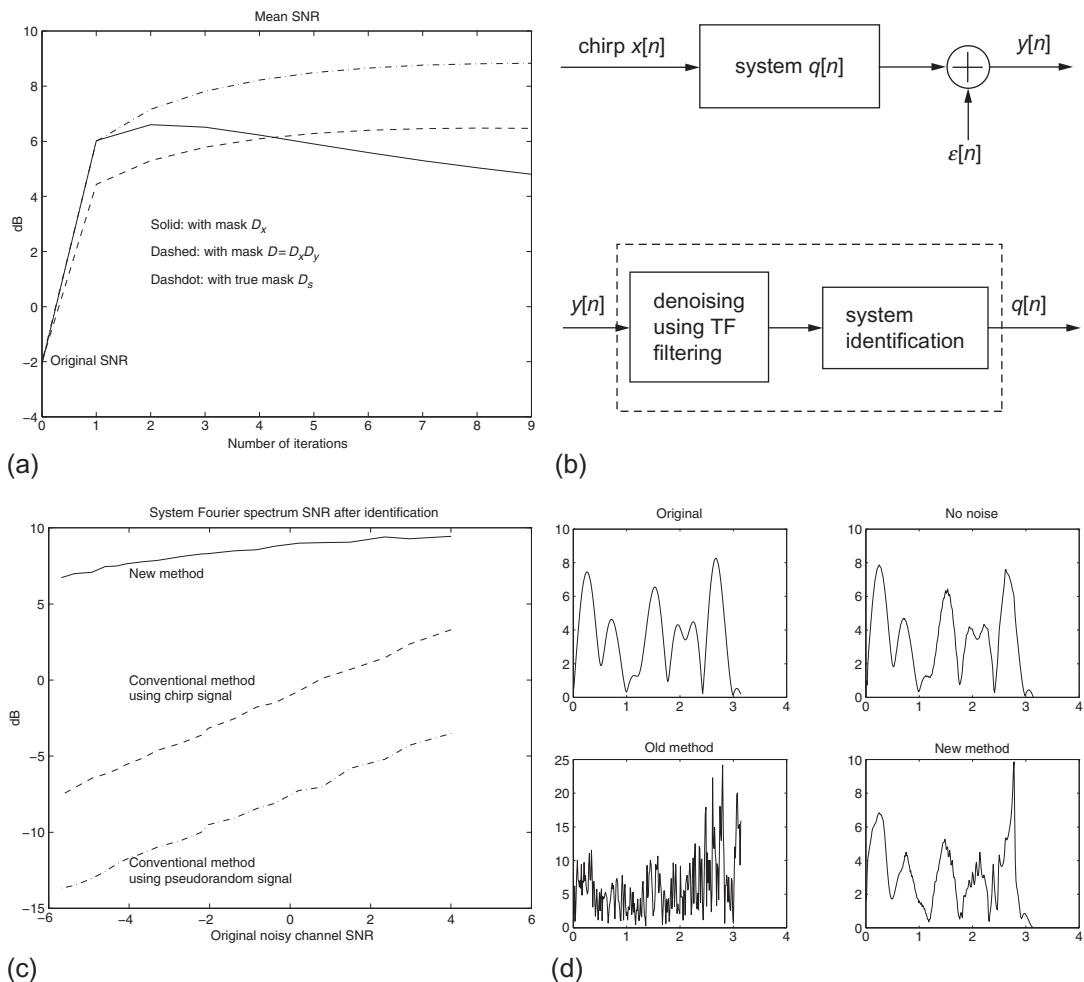
The conventional system identification method used here is the cross-spectral method:

$$Q_{\text{new}}(f) = \frac{S_{\tilde{y}x}(f)}{S_{xx}(f)}, \quad (12.3.16)$$

where $x[n]$ is the chirp signal defined in Eq. (12.3.13) and \tilde{y} is the signal after denoising. It is compared with the conventional method without denoising, i.e.,

$$Q_{\text{old}_1}(f) = \frac{S_{yx}(f)}{S_{xx}(f)}, \quad (12.3.17)$$

where $x[n]$ is also the chirp signal. Since the system identification performance usually depends on the transmitted signal $x[n]$, one might say that it is not fair to compare them using the chirp signal that is pre-

**FIGURE 12.3.3**

(a) The mean SNR curves of the iterative time-variant filtering with the following masks: $D = D_x$, $D = D_x D_y$, and $D = D_s$. (b) System identification method. (c) Comparison of system identification methods: the conventional method using chirp signals; the conventional method using pseudorandom signals; method using chirp signals and time-variant filtering. (d) System identification examples: original spectrum $|Q(f)|$; identified spectrum without additive noise using the chirp signal; the conventional method with additive noise of SNR = -0.4 dB; the described method with additive noise of SNR = -0.4 dB.

ferred here for denoising purposes but which might not be preferred for other methods. For this reason, we also compare the described method with the conventional method using pseudorandom sequences:

$$Q_{\text{old}_2}(f) = \frac{S_{y\tilde{x}}(f)}{S_{\tilde{x}\tilde{x}}(f)}, \quad (12.3.18)$$

where $\tilde{x}[n]$ is a pseudorandom sequence.

Figure 12.3.3(c) shows the improved performance obtained when 200 tests are used for the mean SNR curves for the system spectrum versus the original SNR. Surprisingly, even for the conventional cross spectral method, the chirp signal in Eq. (12.3.13) outperforms pseudorandom signals by approximately 6 dB. In Fig. 12.3.3(d), some identification examples are shown, where the original SNR is -0.4 dB. As a remark, all system identification calculations used in this section are based on the MATLABTM Signal Processing Toolbox.

12.3.5 SUMMARY AND CONCLUSIONS

System identification can be performed by transmitting chirp signals and applying iterative (t,f) filters based on a discrete Gabor transform/expansion. The filters are implemented by projecting the signal back and forth between the time domain and the joint (t,f) domain. This system identification method offers better performance than conventional methods at low signal-to-noise ratios.

12.4 TIME-FREQUENCY METHODS FOR SIGNAL ESTIMATION AND DETECTION⁰

This section uses the generalized Wigner-Ville spectrum (GWVS) and the generalized Weyl symbol (GWS) to develop time-frequency (t,f) techniques for the estimation and detection of underspread nonstationary processes. These (t,f) techniques extend optimal signal estimators (Wiener filters) and optimal signal detectors from the stationary case to underspread nonstationary processes in a conceptually simple and computationally efficient and stable way.

Fundamentals for this section appear in Sections 4.7 and 9.4. The GWVS of a nonstationary random process $x(t)$ with correlation function $r_x(t,t') = E\{x(t)x^*(t')\}$ is defined in Eq. (9.4.2). The GWS of a linear, time-varying (LTV) system \mathbf{H} with kernel (impulse response) $h(t,t')$ is defined in Eq. (9.4.5). For stationary processes, the GWVS reduces to the conventional power spectral density (PSD). Similarly, for a linear, time-invariant (LTI) system, the GWS reduces to the conventional transfer function (frequency response).

A nonstationary process $x(t)$ is called *underspread* if components of $x(t)$ that are sufficiently separated in the (t,f) plane are effectively uncorrelated (see Section 9.4). Two random processes $x(t)$, $y(t)$ are called *jointly underspread* if they satisfy similar underspread conditions [36]. An LTV system \mathbf{H} is called underspread if it produces only small (t,f) displacements of the input signal (see Section 4.7).

12.4.1 NONSTATIONARY SIGNAL ESTIMATION

We first consider the estimation of a nonstationary, zero-mean random signal $s(t)$ from a noise-corrupted observed signal $x(t) = s(t) + n(t)$ by means of an LTV system \mathbf{H} . The signal estimate is thus given by

⁰Authors: **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at) and **G. Matz**, Vienna University of Technology, Vienna, Austria (g.matz@ieee.org). Reviewers: P. Flandrin and S. Carstens-Behrens.

$$\hat{s}(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t, t') x(t') dt'. \quad (12.4.1)$$

The additive noise $n(t)$ is nonstationary and assumed uncorrelated with $s(t)$.

12.4.1.1 The time-varying Wiener filter

The LTV system \mathbf{H} that minimizes the mean-square error (MSE) $E\{\|\hat{s} - s\|^2\}$ is the *time-varying Wiener filter* [3]

$$\mathbf{H}_W = \mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}. \quad (12.4.2)$$

Here, \mathbf{R}_s and \mathbf{R}_n are the correlation operators¹ of signal and noise, respectively. For stationary random processes, \mathbf{H}_W is an LTI system with frequency response [3]

$$H_W(f) = \frac{\mathcal{S}_s(f)}{\mathcal{S}_s(f) + \mathcal{S}_n(f)}, \quad (12.4.3)$$

where $\mathcal{S}_s(f)$ and $\mathcal{S}_n(f)$ denote the PSD of signal and noise, respectively. This frequency-domain expression involves merely a product and a reciprocal of functions (instead of a product and an inverse of operators as in Eq. (12.4.2)) and thus allows a simple design and interpretation of time-invariant Wiener filters.

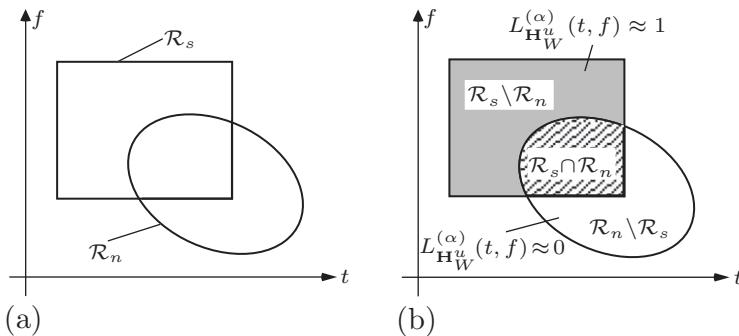
12.4.1.2 (t, f) formulation of time-varying Wiener filters

It is tempting to extend Eq. (12.4.3) to the time-varying Wiener filter \mathbf{H}_W by replacing $H_W(f)$ with the GWS $L_{\mathbf{H}_W}^{(\alpha)}(t, f)$ and $\mathcal{S}_s(f), \mathcal{S}_n(f)$ with the GWVS $\overline{W}_s^{(\alpha)}(t, f), \overline{W}_n^{(\alpha)}(t, f)$. Indeed, for *jointly underspread* processes $s(t)$ and $n(t)$, it can be shown [36] that the time-varying Wiener filter \mathbf{H}_W can be written as the sum of two components: (i) an overspread system component that has negligible effect on the system's performance (MSE) and thus can be disregarded, and (ii) an underspread system component, hereafter denoted as \mathbf{H}_W^u , that allows the approximate (t, f) formulation

$$L_{\mathbf{H}_W^u}^{(\alpha)}(t, f) \approx \frac{\overline{W}_s^{(\alpha)}(t, f)}{\overline{W}_s^{(\alpha)}(t, f) + \overline{W}_n^{(\alpha)}(t, f)}. \quad (12.4.4)$$

Figure 12.4.1 illustrates the simple and intuitively appealing (t, f) interpretation of the time-varying Wiener filter. Let \mathcal{R}_s and \mathcal{R}_n denote the effective support regions of $\overline{W}_s^{(\alpha)}(t, f)$ and $\overline{W}_n^{(\alpha)}(t, f)$, respectively. In the “signal only” (t, f) region $\mathcal{R}_s \setminus \mathcal{R}_n$, Eq. (12.4.4) gives $L_{\mathbf{H}_W^u}^{(\alpha)}(t, f) \approx 1$. Thus, \mathbf{H}_W^u passes all “noise-free” components of $x(t)$ without attenuation or distortion. In the “noise only” (t, f) region $\mathcal{R}_n \setminus \mathcal{R}_s$, Eq. (12.4.4) gives $L_{\mathbf{H}_W^u}^{(\alpha)}(t, f) \approx 0$, i.e., \mathbf{H}_W^u suppresses all components of $x(t)$ located in (t, f) regions where there is no signal. Finally, in the “signal plus noise” (t, f) region $\mathcal{R}_s \cap \mathcal{R}_n$, $|L_{\mathbf{H}_W^u}^{(\alpha)}(t, f)|$ assumes values approximately between 0 and 1. Here, \mathbf{H}_W^u performs an attenuation that depends on the SNR at the respective (t, f) point.

¹The correlation operator \mathbf{R}_x of a nonstationary random process $x(t)$ is the positive (semi-) definite linear operator whose kernel equals the correlation function $r_x(t, t') = E\{x(t)x^*(t')\}$.

**FIGURE 12.4.1**

(t, f) interpretation of the time-varying Wiener filter \mathbf{H}_W for jointly underspread signal and noise processes:
(a) Effective (t, f) support regions of signal and noise, (b) (t, f) pass, stop, and transition regions of the time-varying Wiener filter.

12.4.1.3 (t, f) design of time-varying Wiener filters

Let us define the “ (t, f) pseudo-Wiener filter” $\tilde{\mathbf{H}}_W$ by setting its GWS equal to the right-hand side of Eq. (12.4.4) [36]. Then, for jointly underspread processes $s(t), n(t)$ where Eq. (12.4.4) is a good approximation, the (t, f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ closely approximates (the underspread part of) the optimal Wiener filter \mathbf{H}_W ; furthermore, $\tilde{\mathbf{H}}_W$ will then be nearly independent of the value of α used in Eq. (12.4.4). For processes $s(t), n(t)$ that are *not* jointly underspread, however, $\tilde{\mathbf{H}}_W$ must be expected to perform poorly.

Whereas the (t, f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ is *designed* in the (t, f) domain, the signal estimate $\hat{s}(t)$ can be calculated in the time domain according to Eq. (12.4.1). The impulse response of $\tilde{\mathbf{H}}_W$ is obtained from $L_{\tilde{\mathbf{H}}_W}^{(\alpha)}(t, f)$ according to Eq. (4.7.3). An efficient implementation of the (t, f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ that is based on the multiwindow short-time Fourier transform (STFT) is discussed in Refs. [36, 37]. Compared to the Wiener filter \mathbf{H}_W , the (t, f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ has two practical advantages. Firstly, the prior knowledge required for calculating $\tilde{\mathbf{H}}_W$ is given by the GWVS $\bar{W}_s^{(\alpha)}(t, f)$ and $\bar{W}_n^{(\alpha)}(t, f)$, which are more intuitive and easier to handle than the correlation operators \mathbf{R}_s and \mathbf{R}_n . Secondly, the (t, f) design Eq. (12.4.4) is less computationally intensive and more stable than Eq. (12.4.2) since it requires pointwise (scalar) divisions of functions instead of operator inversions.

12.4.1.4 Robust (t, f) Wiener filters

The performance of the filters \mathbf{H}_W and $\tilde{\mathbf{H}}_W$ is sensitive to deviations of the second-order statistics (correlations or GWVS) from the nominal statistics for which these filters were designed. This motivates the use of *minimax robust Wiener filters* that optimize the worst-case performance (maximum MSE) within specified uncertainty classes of second-order statistics [38, 39].

Consider a partition of the (t, f) plane into K mutually disjoint (t, f) regions $\mathcal{R}_i, i = 1, \dots, K$. Extending the stationary case definition in Ref. [40], we define so-called *p-point uncertainty classes* \mathcal{S} and \mathcal{N} as the sets of all nonnegative (t, f) functions (not necessarily valid GWVS) $\tilde{W}_s(t, f)$ and $\tilde{W}_n(t, f)$ that have prescribed energies s_i and n_i , respectively, within \mathcal{R}_i , i.e., $\iint_{\mathcal{R}_i} \tilde{W}_s(t, f) dt df = s_i$ and

$\iint_{\mathcal{R}_i} \tilde{W}_n(t,f) dt df = n_i$ for $i = 1, \dots, K$. For these uncertainty classes, the GWS of the *minimax robust* (t,f) Wiener filter \mathbf{H}_R is given by [38,39]

$$L_{\mathbf{H}_R}^{(\alpha)}(t,f) = \sum_{i=1}^K \frac{s_i}{s_i + n_i} I_{\mathcal{R}_i}(t,f), \quad (12.4.5)$$

where $I_{\mathcal{R}_i}(t,f)$ is the indicator function of \mathcal{R}_i (i.e., $I_{\mathcal{R}_i}(t,f)$ is 1 for (t,f) inside \mathcal{R}_i and 0 outside \mathcal{R}_i). Note that $L_{\mathbf{H}_R}^{(\alpha)}(t,f)$ is piecewise constant (constant (t,f) weighting within \mathcal{R}_i). The performance of \mathbf{H}_R is approximately independent of the actual second-order statistics as long as they are within \mathcal{S}, \mathcal{N} [38,39]. Signal-adaptive, online implementations of robust time-varying Wiener filters appear in Refs. [39,41].

12.4.1.5 Simulation results

Figure 12.4.2(a) and (b) shows the Wigner-Ville spectra (GWVS with $\alpha = 0$) of jointly underspread signal and noise processes. The Weyl symbols (GWS with $\alpha = 0$) of the corresponding Wiener filter \mathbf{H}_W , of its underspread part \mathbf{H}_W^u , and of the (t,f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ are shown in Fig. 12.4.2(c)–(e). It is verified that the Weyl symbol of $\tilde{\mathbf{H}}_W$ approximates that of \mathbf{H}_W^u . The mean SNR improvement achieved by the (t,f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$ is 6.11 dB, which is almost as good as that of the Wiener filter \mathbf{H}_W (6.14 dB).

To illustrate the performance of the robust (t,f) Wiener filter \mathbf{H}_R , we use $K = 4$ rectangular (t,f) regions \mathcal{R}_i to define p -point uncertainty classes \mathcal{S} and \mathcal{N} as described above. The regional energies s_i and n_i were obtained by integrating the nominal Wigner-Ville spectra in Fig. 12.4.2(a) and (b) over the (t,f) regions \mathcal{R}_i . The Weyl symbol of the robust (t,f) Wiener filter \mathbf{H}_R in Eq. (12.4.5) is shown in Fig. 12.4.2(f). Figure 12.4.3 compares the nominal and worst-case performance of the Wiener filter \mathbf{H}_W (designed for the nominal Wigner-Ville spectra in Fig. 12.4.2(a) and (b)) with the performance of the robust (t,f) Wiener filter \mathbf{H}_R . It is seen that \mathbf{H}_R achieves a substantial performance improvement over \mathbf{H}_W at worst-case operating conditions with only a slight performance loss at nominal operating conditions.

12.4.2 NONSTATIONARY SIGNAL DETECTION

Next, we consider the discrimination of two nonstationary, zero-mean, Gaussian random signals $x_0(t)$ and $x_1(t)$. The hypotheses are $H_0 : x(t) = x_0(t)$ versus $H_1 : x(t) = x_1(t)$.

12.4.2.1 The optimal detector

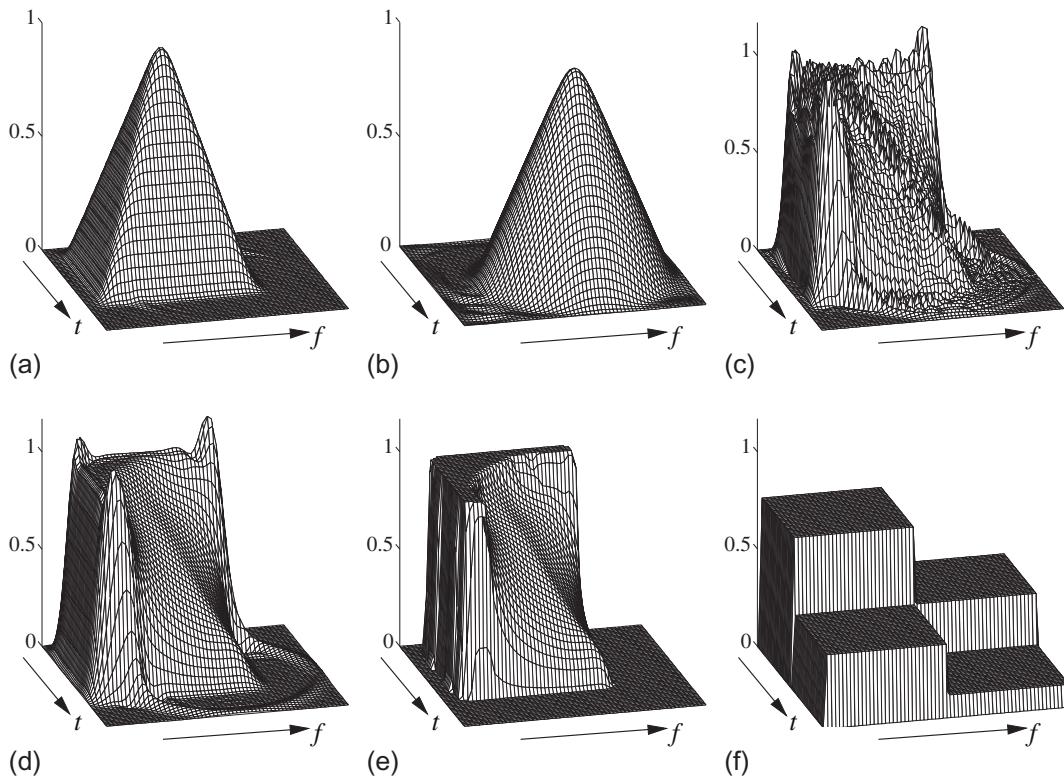
The optimal likelihood ratio detector [3] calculates a quadratic form of the observed signal $x(t)$,

$$\Lambda(x) = \langle \mathbf{H}_L x, x \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_L(t, t') x(t') x^*(t) dt dt', \quad (12.4.6)$$

with the operator (LTV system) \mathbf{H}_L given by

$$\mathbf{H}_L = \mathbf{R}_{x_0}^{-1} - \mathbf{R}_{x_1}^{-1} = \mathbf{R}_{x_0}^{-1} (\mathbf{R}_{x_1} - \mathbf{R}_{x_0}) \mathbf{R}_{x_1}^{-1}. \quad (12.4.7)$$

The test statistic $\Lambda(x)$ is then compared to a threshold to decide whether H_0 or H_1 is in force. For stationary processes, $\Lambda(x)$ can be expressed in terms of the Fourier transform $X(f)$ of $x(t)$ and the PSDs of $x_0(t)$ and $x_1(t)$ as

**FIGURE 12.4.2**

(t, f) representations of signal and noise statistics and of various Wiener-type filters: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) Weyl symbol of Wiener filter \mathbf{H}_W , (d) Weyl symbol of underspread part \mathbf{H}_W^U of \mathbf{H}_W , (e) Weyl symbol of (t, f) pseudo-Wiener filter $\tilde{\mathbf{H}}_W$, (f) Weyl symbol of robust (t, f) Wiener filter \mathbf{H}_R . Time duration = 128 samples; normalized frequency range = $[-1/4, 1/4]$.

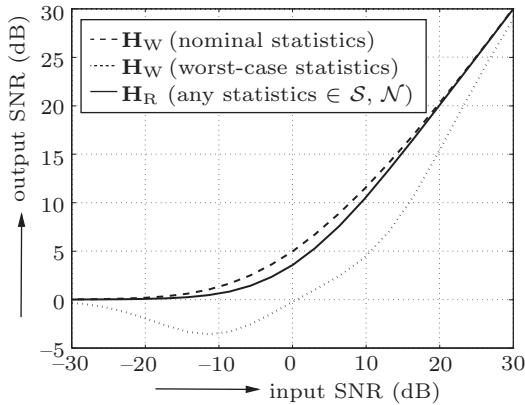
$$\Lambda(x) = \int_{-\infty}^{\infty} |X(f)|^2 H_L(f) df, \quad \text{with } H_L(f) = \frac{\mathcal{S}_{x_1}(f) - \mathcal{S}_{x_0}(f)}{\mathcal{S}_{x_0}(f) \mathcal{S}_{x_1}(f)}. \quad (12.4.8)$$

This frequency-domain expression involves simple products and reciprocals of functions (instead of operator products and inverses as in Eq. (12.4.7)) and thus allows a simple interpretation and design of likelihood ratio detectors in the stationary case.

12.4.2.2 (t, f) formulation of nonstationary detectors

The quadratic test statistic in Eq. (12.4.6) can be rewritten as [5]

$$\Lambda(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x^{(\alpha)}(t, f) L_{\mathbf{H}_L}^{(\alpha)*}(t, f) dt df, \quad (12.4.9)$$

**FIGURE 12.4.3**

Comparison of the performance (output SNR vs. input SNR) of the ordinary Wiener filter \mathbf{H}_W and the robust (t, f) Wiener filter \mathbf{H}_R for various operating conditions.

where,

$$W_x^{(\alpha)}(t, f) = \int_{-\infty}^{\infty} x\left(t + \left(\frac{1}{2} - \alpha\right)\tau\right) x^*\left(t - \left(\frac{1}{2} + \alpha\right)\tau\right) e^{-j2\pi f\tau} d\tau$$

is the generalized Wigner distribution of the observed signal $x(t)$. Thus, $\Lambda(x)$ can be interpreted as a weighted integral of $W_x^{(\alpha)}(t, f)$, with the (t, f) weight function being the conjugate of the GWS of the operator \mathbf{H}_L .

By analogy to [Section 12.4.1](#), a simplified approximate (t, f) formulation of $\Lambda(x)$ exists for *jointly underspread* processes $x_0(t), x_1(t)$. Here, the operator \mathbf{H}_L can be written as the sum of an overspread component whose effect is negligible and an underspread component, denoted \mathbf{H}_L^u , whose GWS can be approximated as [42]

$$L_{\mathbf{H}_L^u}^{(\alpha)}(t, f) \approx \frac{\overline{W}_{x_1}^{(\alpha)}(t, f) - \overline{W}_{x_0}^{(\alpha)}(t, f)}{\overline{W}_{x_0}^{(\alpha)}(t, f) \overline{W}_{x_1}^{(\alpha)}(t, f)}. \quad (12.4.10)$$

Substituting this approximation of $L_{\mathbf{H}_L^u}^{(\alpha)}(t, f)$ for $L_{\mathbf{H}_L}^{(\alpha)}(t, f)$ in Eq. (12.4.9), we obtain the following approximate (t, f) formulation of our test statistic,

$$\Lambda(x) \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x^{(\alpha)}(t, f) \left[\frac{\overline{W}_{x_1}^{(\alpha)}(t, f) - \overline{W}_{x_0}^{(\alpha)}(t, f)}{\overline{W}_{x_0}^{(\alpha)}(t, f) \overline{W}_{x_1}^{(\alpha)}(t, f)} \right]^* dt df. \quad (12.4.11)$$

This extends Eq. (12.4.8) to the underspread nonstationary case and allows a (t, f) interpretation analogous to that of the approximation equation (12.4.4).

12.4.2.3 (t,f) design of nonstationary detectors

Equation (12.4.11) suggests a (t,f) design of nonstationary detectors. By analogy to Eq. (12.4.9), we define the test statistic

$$\tilde{\Lambda}(x) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x^{(\alpha)}(t,f) L_{\tilde{\mathbf{H}}_L}^{(\alpha)*}(t,f) dt df,$$

where the operator (LTV system) $\tilde{\mathbf{H}}_L$ is defined by setting its GWS equal to the right-hand side of Eq. (12.4.10) [42]:

$$L_{\tilde{\mathbf{H}}_L}^{(\alpha)}(t,f) \triangleq \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f) \overline{W}_{x_1}^{(\alpha)}(t,f)}.$$

For jointly underspread processes $x_0(t)$, $x_1(t)$, Eq. (12.4.10) is a good approximation, $\tilde{\mathbf{H}}_L$ will closely approximate (the underspread part of) \mathbf{H}_L , and thus the performance of the (t,f) designed detector $\tilde{\Lambda}(x)$ will be similar to that of the optimal likelihood ratio detector $\Lambda(x)$. For processes $x_0(t)$, $x_1(t)$ that are *not* jointly underspread, however, $\tilde{\Lambda}(x)$ must be expected to perform poorly.

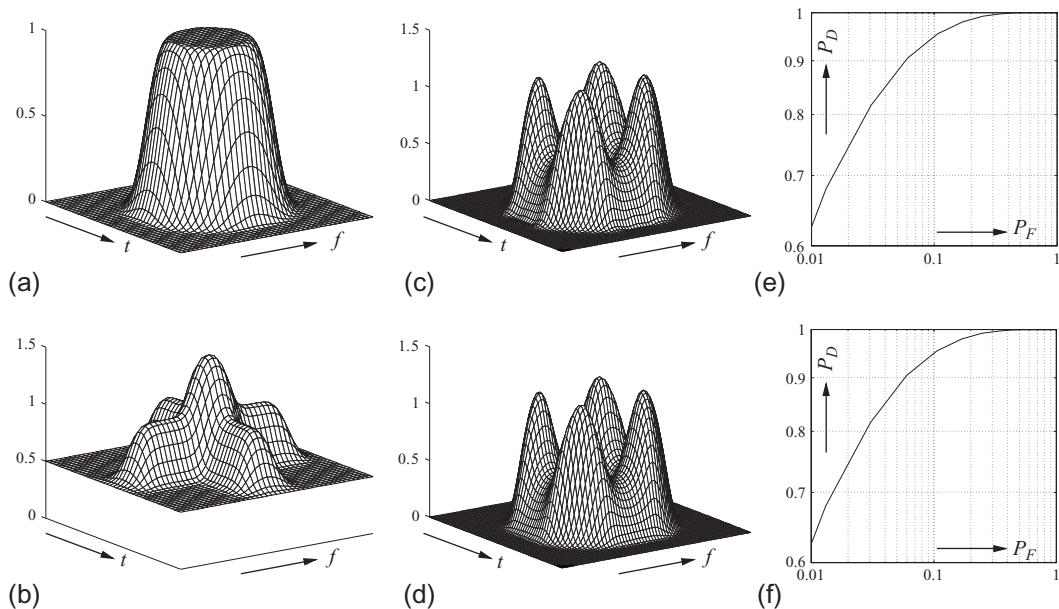
Whereas the detector $\tilde{\Lambda}(x)$ is designed in the (t,f) domain, it can be implemented in the time domain in complete analogy to Eq. (12.4.6). The impulse response $\tilde{h}_L(t,t')$ of $\tilde{\mathbf{H}}_L$ can be obtained from $L_{\tilde{\mathbf{H}}_L}^{(\alpha)}(t,f)$ by an inverse Weyl transformation (cf. Eq. (4.7.3)). An efficient implementation of the (t,f) detector $\tilde{\Lambda}(x)$ that uses the multiwindow STFT is discussed in Ref. [43].

Compared to the likelihood ratio detector $\Lambda(x)$, the (t,f) designed detector $\tilde{\Lambda}(x)$ is practically advantageous because the statistical *a priori* knowledge required for its design is formulated in the intuitively accessible (t,f) domain and because its design is less computationally intensive and more stable since operator inversions are replaced by pointwise divisions of functions. These advantages are analogous to the advantages of the (t,f) pseudo-Wiener filter discussed in Section 12.4.1. For a discussion on minimax robust detectors, see Ref. [44].

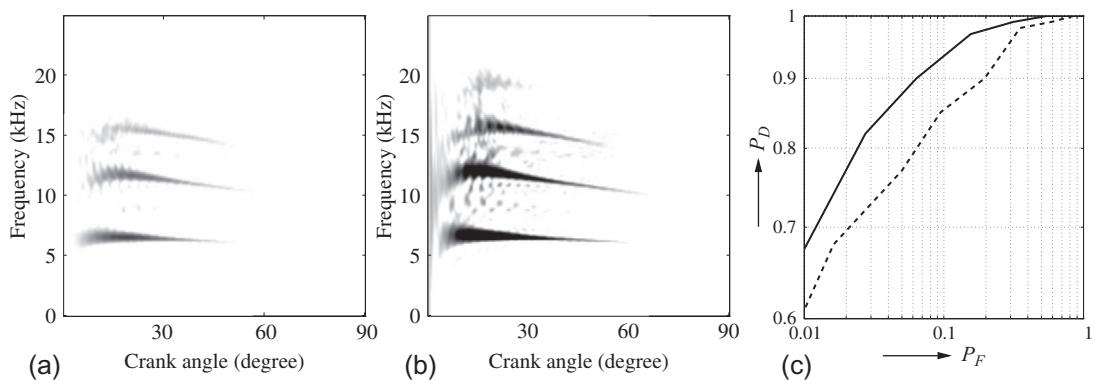
12.4.2.4 Simulation results

Let us first consider $x_0(t) = n(t)$ and $x_1(t) = s(t) + n(t)$, where signal $s(t)$ and noise $n(t)$ are jointly underspread, uncorrelated, zero-mean, Gaussian processes with Wigner-Ville spectra as shown in Fig. 12.4.4(a) and (b). Figure 12.4.4(c) and (d) shows that the Weyl symbols of the optimal operator \mathbf{H}_L and the (t,f) designed operator $\tilde{\mathbf{H}}_L$ are effectively identical, and Fig. 12.4.4(e),(f) shows that the (t,f) designed detector $\tilde{\Lambda}(x)$ performs like the likelihood ratio detector $\Lambda(x)$.

The next example concerns the detection of knocking combustions in car engines (see Section 15.2 and Refs. [43,45,46] for background and details). Here, $x_0(t)$ is the nonknocking signal and $x_1(t)$ is the knocking signal. Estimates of the correlations of $x_0(t)$ and $x_1(t)$ were computed from a set of labeled training data provided by the authors of Section 15.2, and estimates of the Wigner-Ville spectra (shown in Fig. 12.4.5(a) and (b)) were derived according to Eq. (9.4.2). The likelihood ratio detector $\Lambda(x)$ and the (t,f) designed detector $\tilde{\Lambda}(x)$ were constructed using these estimated statistics, and the performance of these detectors was analyzed for a different set of labeled data. Figure 12.4.5(c) shows that the (t,f) designed detector performs significantly better than the theoretically optimal likelihood ratio detector. This is due to numerical problems that occurred in the design of the likelihood ratio detector.

**FIGURE 12.4.4**

Comparison of likelihood ratio detector $\Lambda(x)$ and (t, f) designed detector $\tilde{\Lambda}(x)$: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) Weyl symbol of H_L , (d) Weyl symbol of \tilde{H}_L , (e) receiver operator characteristics (ROC) [3] of $\Lambda(x)$, (f) ROC of $\tilde{\Lambda}(x)$. Time duration = 128 samples; normalized frequency range = $[-1/4, 1/4]$.

**FIGURE 12.4.5**

Detection of knocking combustions: (a) estimated Wigner-Ville spectrum of nonknocking combustion process $x_0(t)$, (b) estimated Wigner-Ville spectrum of knocking combustion process $x_1(t)$ (crank angle is proportional to time; signal length is 186 samples), (c) ROCs of likelihood ratio detector $\Lambda(x)$ (dashed line) and (t, f) designed detector $\tilde{\Lambda}(x)$ (solid line).

Specifically, the estimated correlation matrices² were poorly conditioned. Despite the use of pseudo-inverses, the inversion of these matrices (required for the design of the likelihood ratio detector) could not be stabilized sufficiently. In contrast, the design of the (t,f) detector merely involves a pointwise division of the estimated Wigner-Ville spectra. This is much less affected by numerical problems since divisions by near-to-zero values can easily be stabilized by thresholding.

12.4.3 SUMMARY AND CONCLUSIONS

The generalized Wigner-Ville spectrum (GWVS) provides a natural extension of the PSD to underspread, nonstationary random processes. Similarly, the generalized Weyl symbol (GWS) provides a natural extension of the transfer function (frequency response) to underspread, time-varying linear systems. Using the GWVS and GWS, classical stationary estimators and detectors can be extended to the estimation and detection of underspread, nonstationary random processes in an intuitive manner.

Further discussions of (t,f) methods for nonstationary signal estimation and detection are provided in Sections 8.3, 9.2, 12.1, and 15.2 as well as in Refs. [6,46–49].

12.5 A GENERAL APPROACH TO TIME-FREQUENCY-MATCHED FILTERING⁰

12.5.1 CLASSICAL DETECTION PROBLEM AND ITS SOLUTION

In many applications such as radar, speech, and biomedical signal processing, the presence of a nonstationary signal $s(t)$ that has unknown parameters such as unknown time and/or frequency shifts needs to be detected from a noisy measured signal $x(t)$ of duration T . Following the classical detection problem, the two hypotheses on $x(t)$ are

$$\begin{aligned} H_0 &: x(t) = \epsilon(t), \text{ signal absent,} \\ H_1 &: x(t) = s(t; \Theta) + \epsilon(t), \text{ signal present,} \end{aligned} \quad (12.5.1)$$

where $\epsilon(t)$ is an additive noise and Θ represents unknown parameters of a known deterministic signal $s(t)$ (see also Eq. (12.1.10)). In one such scenario, the signal to detect is a (t,f) shifted version of $s(t)$, that is, $s(t; \Theta) = s(t - t') e^{j2\pi f' t'}$.

The optimal decision strategy for determining the correct hypothesis in the detection problem formulated in Eq. (12.5.1) involves finding the test statistic

$$\eta = \max_{\Theta} \left\{ \int_{(T)} x(t) s^*(t; \Theta) dt \right\} \quad (12.5.2)$$

and comparing it with a predefined threshold value. If there is no unknown parameter, and the signal $s(t)$ is real and the noise $\epsilon(t)$ is a zero-mean white Gaussian process, then the conventional time-domain-matched filter with the test statistic

²In the discrete-time case, correlation operators are replaced by correlation matrices.

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **G. Azemi**, Razi University, Kermanshah, Iran (formerly University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia) (g.azemi@razi.ac.ir). Reviewers: J.M. O'Toole and A.Z. Sha'ameri.

$$\eta = \int_{(T)} x(t) s^*(t) dt \quad (12.5.3)$$

is the optimal detector, where the asterisk denotes complex conjugation.

Based on Moyal's formula, which implies the inner-product invariance property of the WVD (Eq. (3.1.14)), the test statistic of the optimal detector in the (t,f) domain can be expressed as

$$\eta_{\text{QMF}} = |\eta|^2 = \iint_{(T)} W_x(t,f) W_s(t,f) dt df. \quad (12.5.4)$$

The (t,f) -matched filter (TFMF) with the test statistic given in Eq. (12.5.4) is known as the quadrature-matched filter (QMF) [12]. Alternatively, based on the formulation of TFDs presented in the graphical equation (3.2.10) and using the multiplication, time-reversal, and complex conjugation properties of the FT, η_{QMF} can be written in the Doppler-lag (ambiguity) domain as

$$\eta_{\text{QMF}} = \iint A_x(\nu, \tau) A_s(-\nu, -\tau) d\nu d\tau,$$

where $A_s(\nu, \tau)$ is the symmetrical ambiguity function of $s(t)$, defined in Eq. (3.2.14). But from Eq. (3.2.14), it is clear that $A_s(-\nu, -\tau) = A_s^*(\nu, \tau)$, therefore

$$\eta_{\text{QMF}} = \iint A_x(\nu, \tau) A_s^*(\nu, \tau) d\nu d\tau. \quad (12.5.5)$$

12.5.2 GENERAL FORMULATION OF TIME-FREQUENCY-MATCHED FILTERS

Based on Eqs. (12.5.4) and (12.5.5), the test statistic of the general formulation of the WVD-based TFMFs in the (t,f) and ambiguity domains can be written as

$$\eta_{\text{TF}}^{(\text{WV})} = \iint_{(T)} W_x(t,f) \rho_s(t,f) dt df \quad (12.5.6a)$$

$$= \iint A_x(\nu, \tau) A_s^*(\nu, \tau) g^*(\nu, \tau) d\nu d\tau, \quad (12.5.6b)$$

where $g(\nu, \tau)$ is the Doppler-lag kernel of the TFD. The choice of different kernels $g(\nu, \tau)$ in Eq. (12.5.6) results in different test statistics. Table 12.5.1 lists a few known test statistics as special cases of Eq. (12.5.6) with different Doppler-lag kernels for the TFD used in Eq. (12.5.6) [12].

Table 3.3.2 indicates that most commonly used TFDs have real Doppler-lag kernels; therefore, Eq. (12.5.6a) can alternatively be written by swapping s and x as

$$\eta_{\text{TF}}^{(\text{WV})} = \iint_{(T)} W_s(t,f) \rho_x(t,f) dt df, \quad (12.5.7)$$

with

$$\rho_x(t,f) = W_x(t,f) *_{tf} \gamma(t,f), \quad (12.5.8)$$

where $\gamma(t,f)$ is the (t,f) kernel of the TFD.

Table 12.5.1: Test Statistics Resulting from Choices of TFD Doppler-Lag Kernels in Eq. (12.5.6)

| Test Statistic | Doppler-Lag Kernel in Eq. (12.5.6b) |
|---|--|
| $\iint_{(T)} W_x(t, f) W_s(t, f) dt df$ | 1 |
| $\iint_{(T)} W_x(t, f) \rho_s(t - t_0, f - f_0) dt df$ | $g(v, \tau) e^{-j2\pi t_0 v} e^{j2\pi f_0 \tau}$ |
| $\iint_{(T)} \rho_x(t, f) \rho_s(t - t_0, f - f_0) dt df$ | $ g(v, \tau) ^2 e^{-j2\pi t_0 v} e^{j2\pi f_0 \tau}$ |

By replacing $W_x(t, f)$ in Eq. (12.5.6a) with the cross-Wigner-Ville distribution (XWVD) of $x(t)$ and $s(t)$, the test statistic of the general XWVD-based formulation of TFMFs can be defined as

$$\eta_{TF}^{(XWV)} = \iint_{(T)} W_{xs}(t, f) \rho_s(t, f) dt df. \quad (12.5.9)$$

For the special case where $\gamma(t, f) = \delta(t) \delta(f)$, i.e., $\rho_s(t, f) = W_s(t, f)$, we have $|\eta_{TF}^{(XWV)}|^2 = E_s^2 \eta_{QMF}$ where $E_s = \int_{(T)} |s(t)|^2 dt$. This can result in the XWVD-based detection scheme based on Eq. (12.5.9) having better performance in the presence of noise than the WVD-based approach given in Eq. (12.5.6) [50].

12.5.3 TIME-FREQUENCY KERNEL SELECTION FOR OPTIMAL DETECTION

The performance of any solution to the detection problem given in Eq. (12.5.1) can be evaluated using the deflection criterion [1]

$$d = \frac{|E[\eta|H_1] - E[\eta|H_0]|}{(\text{var}[\eta|H_0])^{1/2}}, \quad (12.5.10)$$

where η is the test statistic of the detection method, $\text{var}[\cdot]$ indicates variance and $E[\cdot]$ is the expectation operator. When the signal to be detected, $s(t)$, is deterministic and known, the QMF with the test statistic given in Eq. (12.5.4) is the optimal solution. For this case, the TFMF with the test statistic given in Eq. (12.5.6) is not optimal and [5]

$$d_{TFMF} \leq d_{QMF} = \frac{E_s}{N_0},$$

where E_s is the energy of the reference signal $s(t)$. Note that for $\gamma(t, f) = \delta(t) \delta(f)$, we have $d_{TFMF} = d_{QMF}$. Therefore, the loss in performance in using a TFMF instead of QMF depends on the (t, f) kernel of the TFD.

In most real-life applications, the QMF is not optimal, as the signal to detect is neither deterministic nor completely known. One such case is when $s(t; \Theta) = s(t - t') e^{j2\pi f' t}$, where $s(t)$ is known and deterministic and t' and f' are random variables with joint probability density function (PDF) $p_{t'f'}(t', f')$. For this scenario, the optimal solution is a TFMF with the test statistic given in Eq. (12.5.6), with $\gamma(t, f) = p_{t'f'}(t' = t, f' = f)$ as the (t, f) kernel of the TFD [5]. When the time delay t' and the frequency shift f' are statistically independent, the optimal kernel will be separable, i.e., of the form

$g(v, \tau) = G_1(v) g_2(\tau)$. Procedures for designing such TFDs are presented in Ref. [51] and Section 5.7. For this case, following the approach presented in Ref. [5], we can show that

$$d_{\text{QMF}} \leq d_{\text{TFMF}} = \frac{E_s}{N_0}.$$

Therefore, if the signal $s(t)$ is not known and can only be inferred from noisy measurements, or is randomly perturbed in some way, the QMF never outperforms TFMFs [51–53].

12.5.4 ILLUSTRATION

12.5.4.1 Newborn EEG seizure detection

Electroencephalography (EEG) signals provide objective evidence of the degree and severity of the underlying cause of brain dysfunction and therefore are widely used for diagnosis of abnormal brain activities. In newborns, EEG seizure is one of the major abnormalities correlated with adverse physical and/or neurological outcomes [51]. Newborn EEG seizures are characterized by variations in voltage, duration, frequency content, and waveform shape.

As newborn EEG signals have been shown to be nonstationary multicomponent signals, (t, f) -based approaches have resulted in automatic seizure detectors with better performance than either t domain or f domain methods [12,51,52,54].

12.5.4.2 A TFMF-based approach to newborn EEG seizure detection

Studies have shown that newborn EEG seizures can be modeled as piecewise linear frequency modulation (LFM) signals with harmonics, where the number of LFM pieces depends on the duration of EEG seizure epochs [51,54]. Background patterns, on the other hand, usually exhibit irregular activities with no clear consistent behavior [51]. Figure 12.5.1 illustrates different (t, f) characteristics observed in newborn EEG seizure and background.

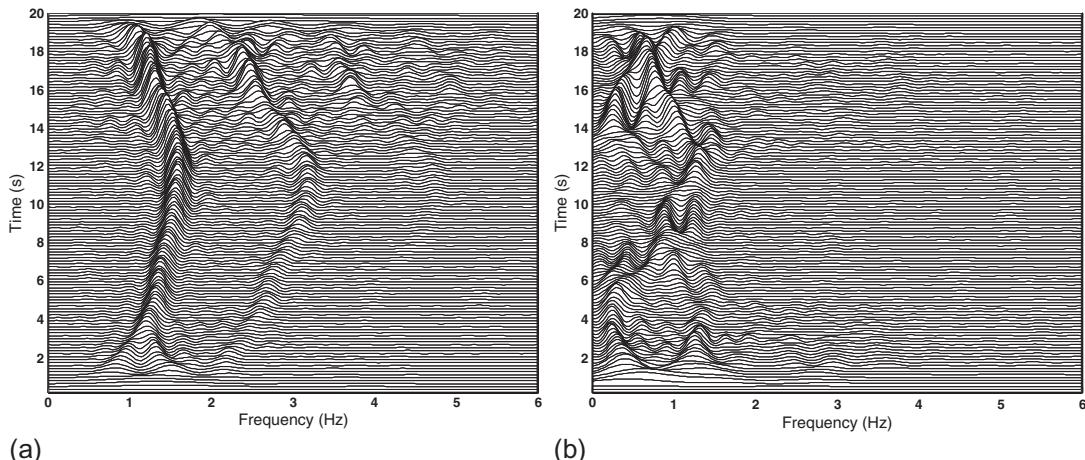
Based on the above findings, the problem of newborn seizure detection can be formulated as Eq. (12.5.1), in which $x(t)$ is the EEG signal and $s(t)$ is a piecewise LFM signal with unknown time delay and/or frequency shift. Since the parameters of the LFM signals (i.e., time duration and slope of each piece of $s(t)$) are not known and may be different for each EEG epoch, one can use a template set $\mathbf{r}(t)$ composed of M piecewise LFM signals, i.e., $\mathbf{r}(t) = \{r_j(t)\}_{j=1}^M$. Also, since the joint p.d.f. of time delay and frequency shift of $s(t)$ is not known, the kernel of the TFD needs to be found adaptively.

12.5.4.3 Implementation

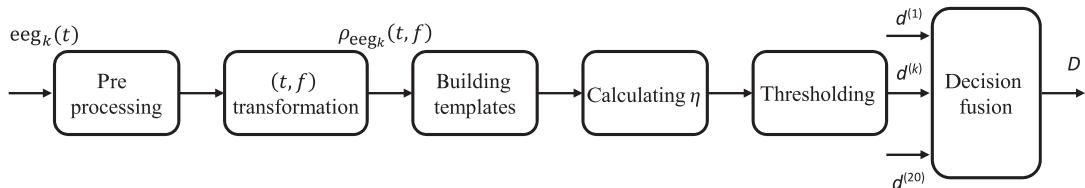
The block diagram in Fig. 12.5.2 illustrates the TFMF-based methodology used for detecting seizures in newborn multichannel EEG signals. An epoch of multichannel EEG is composed of 20 channels, i.e., $\mathbf{eeg}(t) = (\mathbf{eeg}_1(t) \ \mathbf{eeg}_2(t) \ \dots \ \mathbf{eeg}_{20}(t))$ for $0 \leq t \leq T$, with T being the length of the epoch. At the preprocessing stage, the EEG signals are first normalized and filtered in the band 0.5–16 Hz as neonatal EEG seizures have spectral activities mostly below 12 Hz [51]. The filtered signals are then down-sampled at 32 Hz and transformed to the (t, f) domain using Eq. (12.5.8).

Based on the analysis of newborn EEG seizure epochs in Refs. [52,54], a template set composed of only three LFM signals ($M = 3$) is used. It is defined as

$$r_j(t) = e^{-j\pi(2f_0t + \alpha_j(t-T/2)^2)}; \quad 0 \leq t \leq T; j = 1, 2, 3, \quad (12.5.11)$$

**FIGURE 12.5.1**

An example of (t, f) representations of EEG signals acquired from a newborn during seizure (left) and nonseizure activities (right).

**FIGURE 12.5.2**

Methodology for TFMF-based newborn seizure detection using multichannel EEG signals.

where $\alpha_j \in \{-0.05, 0, 0.05\}$, and f_0 is the frequency at which the time slice, at $t = T/2$, of the TFD of the EEG segment under analysis attains its peak

$$f_0 = \arg \max_f \{ \rho_{\text{eeg}_k}(T/2, f) \}. \quad (12.5.12)$$

It is clear that increasing the size of the template set may improve the performance of the detector but with more computations. Note that f_0 may be different for signals acquired from different channels as they may have different (t, f) signatures. With this approach, f_0 is found based on the fundamental component in $\text{eeg}_k(t)$. This is because the energy of the harmonics in $\text{eeg}_k(t)$ is relatively small compared to that of the fundamental component [52].

Based on the results presented in Section 12.5.2, two detection methods based on the general formulation of TFMFs given in Eqs. (12.5.6) and (12.5.9) can be used. The test statistics of these methods are given by

$$\eta_1^{(k)} = \max_{r_j} \iint_{(T)} W_{\text{eeg}_k}(t, f) \rho_{r_j}(t, f) dt df, \quad j = 1, 2, 3, \quad (12.5.13a)$$

$$\eta_2^{(k)} = \max_{r_j} \iint_{(T)} |W_{\text{eeg}_k, r_j}(t, f)| \rho_{r_j}(t, f) dt df, \quad j = 1, 2, 3. \quad (12.5.13b)$$

Equations (12.5.13a) and (12.5.13b) represent the test statistics of the WVD- and XWVD-based TFMFs. Note that as $W_{\text{eeg}_k, r_j}(t, f)$ is not necessarily real, its modulus is used in defining $\eta_2^{(k)}$. By defining f_0 as in Eq. (12.5.12), there is no longer a need for time and frequency shifting of the templates in order to find the test statistic of the matched filter. This reduces the computational load of the EEG seizure detector as needed for fast and real-time implementation. The test statistic for a given EEG segment under analysis is normalized so its value would be in the range [0, 1]. The calculated test statistic is compared with a predefined threshold to find $d^{(k)}$, the binary decision value for the k th channel. Finally, the multichannel decision fusion is achieved by combining the independent decisions of the different EEG channels to form an overall decision as to the presence of a seizure. Here, the overall decision D is found by combining the single-channel results using

$$D = \begin{cases} 1 & \text{if } \sum_{k=1}^{20} d^{(k)} \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

12.5.4.4 Results and discussions

The performance of the above methodology was evaluated using multichannel EEG recordings from five neonates. The signals were marked for seizures by a pediatric neurologist from the Royal Children's Hospital, Brisbane, Australia. A total of 100 seizure segments and 100 nonseizure segments with length $T = 8$ s ($N = 256$ samples after down-sampling at 32 Hz) were extracted randomly from 27 min of artifact-free seizure signals and 39 min of artifact-free nonseizure signals. The procedure illustrated in Fig. 12.5.2 was repeated for different values of thresholds and the *receiver operating characteristic* (ROC) curve was found for each of the two detectors with different (t, f) kernels. The kernels included the WVD, smoothed WVD (SWVD) with a Hamming window $N/4$ samples long, spectrogram (SPEC) with a Hamming window $N/4$ samples long, modified B-distribution (MBD) with $\beta = 0.3$, and extended MBD (EMBD) with $\alpha = 0.1$ and $\beta = 0.9$ (Table 6.1.2). With the WVD kernel, the test statistic η_1 reduces to the test statistic of the QMF given in Eq. (12.5.4). For each detector and each kernel, the ROC curve was found, and the *area under the curve* (AUC) was calculated as the performance criterion (see the pseudocode in Appendix to this section). A perfect classifier would have AUC = 1.

The AUC scores for the WVD- and XWVD-based detectors (with test statistics given in Eqs. (12.5.13a) and (12.5.13b), respectively) with different (t, f) kernels are given in Table 12.5.2. Note that with the WVD as the (t, f) kernel in Eq. (12.5.13a), the WVD-based TFMF becomes the QMF, which has similar performance to the time-domain-matched filter.

The results imply that the XWVD-based TFMF (with highest AUC score of 0.95) outperforms the one based on the WVD (with highest AUC score of 0.89) and the QMF (with AUC score of 0.87).

Table 12.5.2: The AUC Scores of the WVD- and XWVD-Based TFMFs in Detecting Seizures in Multichannel Newborn EEG

| | | TFD Kernel | | | | |
|---------|------------|------------|------|------|------|------|
| | | WVD | SWVD | SPEC | MBD | EMBD |
| Method: | WVD-based | 0.87 | 0.89 | 0.89 | 0.89 | 0.67 |
| | XWVD-based | 0.88 | 0.94 | 0.94 | 0.95 | 0.88 |

Also, among different kernels, the MBD kernel shows higher detection accuracy. These high-detection accuracies show that the minimum-size template set chosen is a good match to the signals in the datasets.

12.5.5 SUMMARY AND CONCLUSIONS

For detection of nonstationary signals with unknown parameters in the presence of noise, the conventional time-domain-matched filter is no longer optimal. For such cases, this section has presented general design of TFMFs based on the WVD and XWVD. The kernel of the TFD is designed and optimized to account for the unknown parameters of the signal to detect. The kernel also gives a (t,f) smoothing, which can attenuate noise components with large lag and/or Doppler values.

The methodology presented here can be extended to TFMF-based classification of different newborn EEG abnormalities [12,51]. It can also be applied to other important areas such as (t,f) -based watermarking.

12.5.6 APPENDIX: PSEUDOCODE FOR THE IMPLEMENTED ALGORITHM

```
% MATLAB pseudo-code for the methodology used
% This code needs the TFSAP toolbox.
% initialization
F_s=256; % original sampling frequency
f_s=32; % sampling frequency after down sampling
L_seg=8; % segment length in sec
TR=1; % TFD time resolution
LWL=127; % length of the lag window
SWL=127; % length of the smoothing window
kernel='WVD'; % choose the  $(t,f)$  kernel
                % other options are: 'SWVD', 'SPEC', 'MBD', and 'EMBD'
method=1; % 1 is for WVD-based and 2 is for XWVD-based
N_seg=200; % 200 is the total number of segments
for n=1:N_seg
    for k=1:20 % 20 is the number of channels
        % normalize the  $k$ th channel, i.e., eeg_k
        eeg_k=eeg_k/max(eeg_k);
```

```

% filter the normalized signal in [0.5 16] Hz band
b=fir1(100,[0.5/(F_s/2) 16/(F_s/2)]);
eeg_k=filtfilt(b,1,eeg_k);
% down sample the signals @ 64 Hz
eeg_k=resample(eeg_k,f_s,F_s);
switch method
    case 1 % WVD-based TFMF
        % find tfd_eeg=ρeegk(t,f) using the quadtdf function in TFSA package
        switch kernel
            case 'WVD' % WVD
                tfd_eeg=quadtdf(eeg_k,LWL,TR,'wvd');
            case 'SWVD' % SWVD with a Hamming window
                tfd_eeg=quadtdf(eeg_k,LWL,TR,'smoothed',SWL,'hamm');
            case 'SPEC' % SPEC with a Hamming window
                tfd_eeg=quadtdf(eeg_k,LWL,TR,'specx',SWL,'hamm');
            case 'MBD' % MBD with β = 0.01
                tfd_eeg=quadtdf(eeg_k,LWL,TR,'mb',0.01);
            case 'EMBD' % EMBD with α = 0.1 and β = 0.9
                tfd_eeg=quadtdf(eeg_k,LWL,TR,'emb',0.1,0.9);
        end
        % find f0
        [max_tmp.f_0]=max(tfd_eeg(:,fix(T2/2)));
        nfft=size(tfd_eeg,1);
        f_0=f_0 * (f_s/2)/nfft;
        % build the templates; r_j is built using the gsig function in TFSA package
        for j=1:3
            % find f1 and f2 based on the slope of the LFM
            r_j=gsig('lin', f1/(f_s/2), f2/(f_s/2), T, 0);
            % find tfd_r_j=ρrj(t,f); codes are similar to above
            out_r_j(j)=sum(sum(tfd_eeg.*tfd_r_j));
            % normalize out_r_j(j)
        end % next j
    case 2 % XWVD-based TFMF
        % codes are similar to above
        % use xwvd(eegk,r_j,LWL,TR) to find the XWVD
    end % over the methods
    eta_method(n,ch)=max(out_r_j);
end % next channel
mask(n)=max(eta_method(n,:));
end % next segment
for ηth = ηmin : ηmax % ηth is the threshold
    for n=1:N_seg
        D(n)=0; % initialization
        if mask(n) < ηth

```

```

D(n)=1;
end
end % next n
compare D and the signal mask and calculate the sensitivity and specificity
end % next  $\eta_{\text{th}}$ 
calculate the AUC

```

12.6 DEFINING TIME-FREQUENCY IMAGE FEATURES BY EXTENSION FROM FREQUENCY DOMAIN OR TIME DOMAIN⁰

The use of time-frequency distributions (TFDs) for the analysis and classification of nonstationary biomedical signals [55] leads to a (t,f) pattern-recognition approach that involves the stages illustrated in Fig. 12.6.1.

TFDs give additional information about nonstationary signals that cannot be extracted from either the time domain or frequency domain representations. This includes the instantaneous frequency. TFDs can be used to classify nonstationary signals by extracting relevant (t,f) features as characteristics—instead of using *all* (t,f) points as features, as this would result in a large number of redundant features, thereby increasing the computational cost and data requirement for training [56]. Features can be extracted by dividing a TFD into a number of tiles and computing the energy in each tile [57], or reducing the dimension of the TFD with methods such as principal component analysis [58], or using IF estimation techniques to extract features such as mean frequency or frequency deviation [55].

This section presents a methodology for defining new (t,f) features by extending time-domain (t -domain) and frequency-domain (f -domain) features to the joint (t,f) domain [55]. As an illustration, the proposed methodology is applied to derive new features for the classification of newborn EEG signals. The performance of the (t,f) features is compared with the corresponding t -domain or f -domain features using a receiver operator characteristic (ROC) analysis. Studies show that these (t,f) features may give better performance compared to the corresponding t -only and f -only features. In addition, several (t,f) image-related features are also described for completeness.

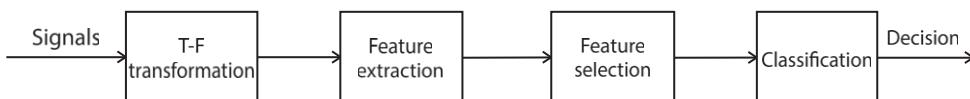


FIGURE 12.6.1

Time-frequency approach to pattern classification.

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **N.A. Khan**, Qatar University, Doha, Qatar (nabeelalikhan@qu.edu.qa). Reviewers: S. Aviyente, G. Azemi, S. Ouelha, and A. Saleem.

12.6.1 EXTENSION OF f -DOMAIN FEATURES TO THE JOINT (t, f) DOMAIN

Typical frequency-domain features such as the mean frequency and spectral entropy can only discriminate among classes of signals with different distributions of the signal energy in the frequency domain; for example, the spectral flatness (which measures the uniformity of signal energy in the frequency domain) can be used to discriminate narrowband signals from wideband signals; however, these features are unable to discriminate among signals with similar signal energy distribution in the frequency domain but completely different (t,f) signatures; e.g., the spectral flatness cannot be used to discriminate white noise from LFM signal, as both signals have wideband characteristics. Such limitations can be overcome by extending the f -domain features to the (t,f) domain as presented in the next section.

12.6.1.1 Extension of spectral flux to (t, f) flux

The spectral flux measures the rate of change of the spectral content of a signal with time [59]. It can be estimated by taking the sum of the absolute difference of the magnitude-squared Fourier transforms of two adjacent segments of a signal, or it can be directly estimated from a TFD using the formula [55]

$$\mathcal{FL}_f = \sum_{n=0}^{N-l} \sum_{k=1}^M \left| \rho_{z_x}[n+l, k] - \rho_{z_x}[n, k] \right|, \quad (12.6.1)$$

where $\rho_{z_x}[n, k]$ represents a TFD (of size $N \times M$) of the analytic associate $z_x[n]$ of a real signal $x[n]$ of length N , and l is the predetermined time-duration between two slices of a TFD, which can assume any integer value between 0 and $N - l$. The spectral flux is a useful measure for distinguishing signals whose spectrum changes slowly from signals whose spectrum changes quickly. It has a lower value for the former “slow” class of signals and a higher value for the latter “fast” class of signals. For example, the energy of EEG seizure signals with LFM characteristics varies slowly with time compared to the EEG background signals, as illustrated in Fig. 12.6.2.

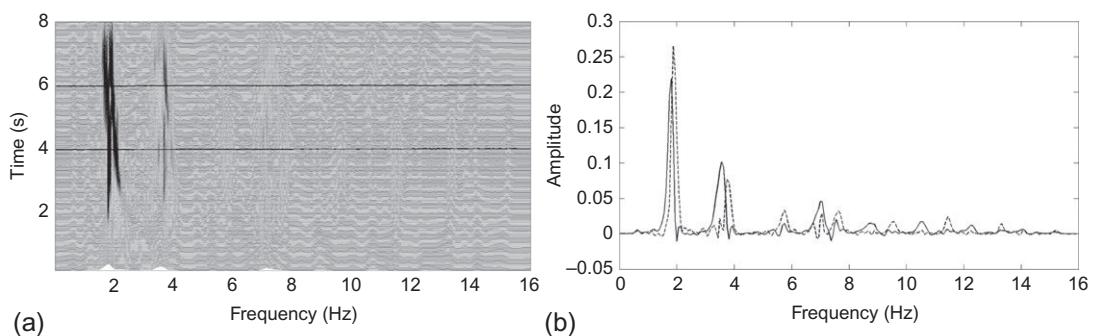


FIGURE 12.6.2

Illustration of the slow variation of the energy of an EEG signal along the frequency axis: (a) TFD (extended modified B distribution) of a seizure signal with piecewise LFM characteristics (black lines indicating two time-slices at $t = 4$ s and $t = 6$ s); (b) plot of the two time slices of the TFD (solid line for $t = 6$ s, dashed line for $t = 4$ s).

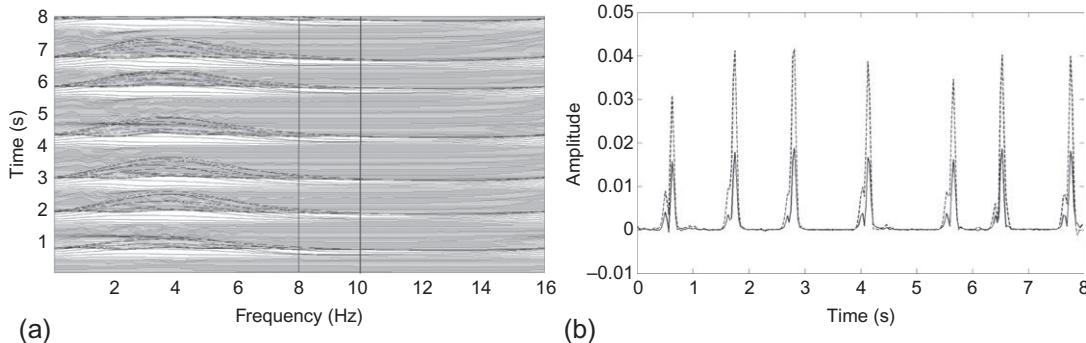
**FIGURE 12.6.3**

Illustration of the slow variation of the energy of an EEG spike signal along the time axis: (a) TFD (extended modified B distribution) of a seizure signal with spike characteristics (black lines indicating two frequency-slices at $f = 8$ Hz and $f = 10$ Hz); (b) plot of the two frequency slices of the TFD (dashed line for $f = 8$ Hz, solid line for $f = 10$ Hz).

Previous studies show that some EEG seizure signals can be modeled as spikes (or a train of impulses) [60]. The (t,f) signature of one such EEG seizure signal is shown in Fig. 12.6.3, which shows that the energy of the signal varies slowly along the frequency-axis due to the spike characteristics. This implies that the spectral flux may not be a good feature for distinguishing EEG seizure signals with spike characteristics from the random background, as the TFDs of spike signals have a sudden variation of the signal energy along the time-axis, resulting in a higher spectral flux than for EEG seizure signals with LFM characteristics. Thus the spectral flux is lower than the random background for EEG seizure signals with LFM characteristics but close to the random background for EEG seizure signals with spike characteristics.

The (t,f) flux extends the spectral flux by measuring the rate of change of signal energy along both the time axis and the frequency axes. It is defined as

$$\mathcal{FL}_{(t,f)} = \sum_{n=1}^{N-l} \sum_{k=1}^{M-m} |\rho_{zx}[n+l, k+m] - \rho_{zx}[n, k]|, \quad (12.6.2)$$

where l and m are predetermined values that depend on the rate of change of signal energy in the (t,f) plane. They can assume any integer values from 0 to $N - 1$ and 0 to $M - 1$, respectively. In this study, l and m are set to 1 to estimate the (t,f) flux. The (t,f) flux accounts for the slow variation in the energy of EEG seizure signals, which takes place either along the frequency axis due to the spike characteristics or along the time-axis due to the LFM characteristics. Therefore, the (t,f) flux usually shows lower values for EEG seizure signals and higher values for EEG background signals.

12.6.1.2 Extension of f domain entropy measures to the joint (t, f) domain

In essence, f -domain entropy measures like the spectral entropy, and spectral flatness (Wiener entropy) estimate the uniformity of signal energy distribution in the frequency domain. High f -domain entropy implies more uniformity in the distribution of the signal energy in the frequency domain, while

low entropy implies less uniformity. These entropy-based measures can be used to discriminate a narrowband signal from a wideband signal, e.g., discriminate between a tone and white noise; but they cannot be used to discriminate two wideband signals, e.g., an LFM signal and noise. In effect, TFDs concentrate energy for the FM signals just as the FT concentrates energy for tones. So, (t,f) extensions of the entropy measures can discriminate between two different classes of wideband signals such that the energy of one class of signals is uniformly distributed in the (t,f) domain, e.g., white noise, while the energy of the second class is concentrated in the (t,f) domain, e.g., an LFM signal. These (t,f) entropy measures can be used to detect the seizure activity in EEG signals, as the energy of EEG seizure signals is usually concentrated in the (t,f) domain along the IFs of the signal components, resulting in low entropy, whereas the energy of the EEG background signal is randomly distributed in the (t,f) domain, resulting in high (t,f) entropy.

- 1. Spectral Flatness (Wiener entropy) to (t,f) Flatness:** The spectral flatness is defined as the geometric mean of the FT of a signal normalized by its arithmetic mean [59]; i.e.,

$$\mathcal{SF}_{(f)} = M \frac{\left(\prod_{k=1}^M |Z_x[k]| \right)^{M^{-1}}}{\sum_{k=1}^M |Z_x[k]|}, \quad (12.6.3)$$

where $Z_x[k]$ is the FT of the analytic associate $z[n]$ of a real signal $x[n]$ and M is the length of $Z_x[k]$. The (t,f) extension of the spectral flatness is obtained by replacing the geometric and arithmetic means of the FT of a signal in Eq. (12.6.3) with the corresponding geometric and arithmetic means of a TFD. The (t,f) flatness is defined as [55]

$$\mathcal{SF}_{(t,f)} = MN \frac{\prod_{n=1}^N \prod_{k=1}^M |\rho_{z_x}[n, k]|^{\frac{1}{NM}}}{\sum_{n=1}^N \sum_{k=1}^M \rho_{z_x}[n, k]}. \quad (12.6.4)$$

In practical implementation, all zeros of a TFD are replaced by very small values.

- 2. Spectral Entropy to (t,f) Entropy:** The spectral entropy is defined as

$$\mathcal{SE}_{(f)} = - \sum_{k=1}^M \mathcal{Z}_x[k] \log_2 \mathcal{Z}_x[k], \quad (12.6.5)$$

where $\mathcal{Z}_x[k] = |Z_x[k]|^2 / \sum_{k=1}^N |Z_x[k]|^2$. The (t,f) Shannon entropy is an extension of the spectral entropy, which can be derived from the spectral entropy by first replacing the FT of a signal with the TFD in Eq. (12.6.5) and then replacing 1D summation by 2D summation. It is defined as

$$\mathcal{SE}_{(t,f)} = - \sum_{n=1}^N \sum_{k=1}^M \frac{\rho_{z_x}[n, k]}{\sum_n \sum_k \rho_{z_x}[n, k]} \log_2 \left(\frac{\rho_{z_x}[n, k]}{\sum_n \sum_k \rho_{z_x}[n, k]} \right). \quad (12.6.6)$$

The above expression is only valid for the TFDs that do not assume negative values, e.g., the spectrogram. The (t,f) Rényi entropy and normalized (t,f) Rényi entropy (TFRNE) are generally used for quadratic TFDs as discussed in Section 7.7. In this section, the normalized (t,f) entropy is

used because it gives superior performance in terms of its ability to discriminate EEG seizure signals from the EEG background. It is defined as

$$\mathcal{RE}_{(t,f)} = \frac{1}{1-\alpha} \log_2 \sum_{n=1}^N \sum_{k=1}^M \left(\frac{\rho_{z_x}[n,k]}{\sum_n \sum_k \rho_{z_x}[n,k]} \right)^\alpha, \quad (12.6.7)$$

where α is an odd integer and $\alpha > 2$. In this study, we have chosen $\alpha = 3$ [61]. In addition to Rényi entropy, other (t,f) concentration measures can also be used (see [Section 7.3](#)).

12.6.2 EXTENSION OF t -DOMAIN FEATURES TO JOINT (t,f) DOMAIN

It is generally assumed that the amplitudes of normal and abnormal signals have different probability distribution functions (PDFs). The PDFs can be characterized by their moments such as mean, variance, and kurtosis, so the moments of the PDF of the signal amplitude can be used as features for classification. The moments can be estimated from the t -domain representation of a signal. This concept can be extended to the joint (t,f) domain, to exploit additional information provided by TFDs, by replacing 1D t -domain moments by the corresponding 2D (t,f) -domain moments as shown below.

1. The **mean** of the t -domain representation is given by

$$m_{(t)} = \frac{1}{N} \sum_n x[n]. \quad (12.6.8)$$

When extended to the (t,f) domain, it becomes

$$m_{(t,f)} = \frac{1}{NM} \sum_n \sum_k \rho_{z_x}[n,k]. \quad (12.6.9)$$

2. The **variance** of the t -domain representation is given by

$$\sigma_{(t)}^2 = \frac{1}{N} \sum_n (x[n] - m_{(t)})^2. \quad (12.6.10)$$

When extended to the (t,f) domain, it becomes

$$\sigma_{(t,f)}^2 = \frac{1}{NM} \sum_n \sum_k (\rho_{z_x}[n,k] - m_{(t,f)})^2. \quad (12.6.11)$$

3. The **skewness** of the t -domain representation is given by

$$\gamma_{(t)} = \frac{1}{N\sigma_{(t)}^3} \sum_n (x[n] - m_{(t)})^3. \quad (12.6.12)$$

When extended to the (t,f) domain, it becomes

$$\gamma_{(t,f)} = \frac{1}{NM\sigma_{(t,f)}^3} \sum_n \sum_k (\rho_{z_x}[n,k] - m_{(t,f)})^3. \quad (12.6.13)$$

4. The **kurtosis** of the t -domain representation is given by

$$k_{(t)} = \frac{1}{N\sigma_{(t)}^4} \sum_n (x[n] - m_{(t)})^4. \quad (12.6.14)$$

When extended to the (t,f) domain, it becomes

$$k_{(t,f)} = \frac{1}{NM\sigma_{(t,f)}^4} \sum_n \sum_k (\rho_{z_x}[n,k] - m_{(t,f)})^4. \quad (12.6.15)$$

5. The **coefficient of variation** of the t -domain representation is given by

$$c_{(t)} = \sigma_{(t)}/m_{(t)}. \quad (12.6.16)$$

When extended to the (t,f) domain it becomes

$$c_{(t,f)} = \sigma_{(t,f)}/m_{(t,f)}. \quad (12.6.17)$$

12.6.3 TIME-FREQUENCY IMAGE FEATURES FOR PATTERN RECOGNITION

One approach to extract features from a given TFD is to interpret it as an image and extract features from the TFD using image-processing techniques. This approach exploits well established image-processing methods to extract highly discriminating features. The image-related features are discussed below.

- Geometric (t,f) features: signal components appear as regions of energy concentration in the (t,f) domain. Image segmentation techniques such as watershed segmentation can be used to locate such regions. Watershed segmentation techniques can be adapted by interpreting image as topographic surface where watershed boundaries separate each energy concentration region [62]. Geometric features are then extracted from these segments to get information regarding the geometry of these segments in the (t,f) plane.

Let us assume that a given TFD $\rho(n, k)$ is segmented into L number of time-frequency regions. For a given (t,f) image, the moments of order (p, q) can be extracted from each segment of the image using the following expression [62].

$$m_{pq}^l = \sum_n \sum_k n^p k^q \rho^l(n, k), \quad (12.6.18)$$

where $p, q = 0, 1, 2, \dots$, $\rho^l(n, k)$ represents the l th segment of the (t,f) image, and m_{pq}^l represents the moment of order (p, q) for l th segment of the (t,f) image. The central moments are expressed as

$$\mu_{pq}^l = \sum_n \sum_k (n - \bar{n}^l)^p (k - \bar{k}^l)^q \rho^l(n, k), \quad (12.6.19)$$

where $\bar{n}^l = \frac{m_{10}^l}{m_{00}^l}$ and $\bar{k}^l = \frac{m_{01}^l}{m_{00}^l}$ are the coordinates of the centroids in the (t,f) plane of the l th segment. The following shape features can be extracted from the moments [62]:

- (t,f) convex hull/area, i.e., μ_{00}^l
- (t,f) perimeter: $(m_{30}^l + m_{12}^l)^2 + (m_{03}^l + m_{21}^l)^2$
- (t,f) compactness: $((m_{30}^l + m_{12}^l)^2 + (m_{03}^l + m_{21}^l)^2)^2 / \mu_{00}^l$

- (t, f) coordinates of the centroid for the segmented region: $\frac{m_{10}^l}{m_{00}^l}$ and $\frac{m_{01}^l}{m_{00}^l}$
Newborn EEG seizure signals appear as large connected regions in the (t, f) domain, while nonseizure signals appear as small disconnected regions of energy concentrations. So, the geometric properties of EEG signals can be used as (t, f) features for detecting seizures.
- Texture (t, f) feature: such features are related to direction and shape of energy distribution in the (t, f) plane. These features can be obtained by first convolving a TFD with a set of convolution masks and then extracting features from the convolved images [63]. The convolution masks are selected to detect certain shapes in the time-frequency images, e.g., $[-1, -2, 0, 1, 2]^T$ can be used to detect edges in the given image.
Let $\rho_j(n, k)$ represent a (t, f) image obtained as a result of convolution with the mask $H_j(n, k)$. It is obtained as

$$\rho_j(n, k) = H_j(n, k) * \rho(n, k). \quad (12.6.20)$$

Statistical (t, f) features such as mean and variance can then be extracted from $\rho_j(n, k)$. Note that Eq. (12.6.20) can be implemented in the ambiguity domain (v, τ) , therefore resulting in a product and two FTs. The operation can be combined with the formulation of high-resolution QTDFs (see Eqs. (2.7.36) and (3.2.13)), given that $\rho_j(n, k)$ can be expressed as

$$\rho_j(n, k) = H_j(n, k) * \gamma(n, k) * W_z(n, k), \quad (12.6.21)$$

where $W_z(n, k)$ is the WVD. If we define $\gamma'(n, k) = H_j(n, k) * \gamma(n, k)$, then the two 2D convolution operations can be reduced to one 2D convolution operation, i.e., $\rho_j(n, k) = \gamma'(n, k) * W_z(n, k)$, thus reducing the computational cost.

- Histogram (t, f) features: the histogram is a graphical representation of the distribution of data. Let us assume that the signal energy in a TFD is quantized to L discrete levels. Then, the histogram of a TFD is the number of samples in the TFD for each energy level. The histogram of TFDs can be estimated using the following expression [64]:

$$h_i = \sum_n \sum_k \delta(\rho(n, k) - i), \quad (12.6.22)$$

where $i = 1, 2, \dots, L$. Statistical quantities such as mean, standard deviation, skewness, and kurtosis describing the shape of the (t, f) histogram can then be used as (t, f) features. Contrast can also be used [65,66].

12.6.4 RESULTS AND DISCUSSION

To evaluate the performance of (t, f) features for detection of seizure activity in newborn EEG signals, the area under the ROC curve (AUC) is used as a performance measure for each feature [67]. The AUC of a feature is the probability that the feature would have a higher value for a randomly chosen positive example than for a randomly chosen negative example [67]. A high AUC (i.e., close to 1) implies that a feature is good in discriminating between two different classes; a low AUC (i.e., close to 0.5) implies that a feature fails to discriminate.

A 20-channel EEG dataset was acquired from 63 neonates at the NICU of the Royal Brisbane and Women's Hospital, Brisbane, Australia. The database contains 26.5 hours of EEG background and 4.3 hours of EEG seizure. A set of 200 single-channel 8s epochs of normal signals, and a set of 200 single-channel 8s epochs of seizure signals are randomly extracted from the database. The length of each epoch is 256 samples at the 32 Hz sampling rate. All the segments are bandpass-filtered to [0.5 Hz, 20 Hz] to remove artifacts and noise. The signals are analyzed using the WVD, the spectrogram (SPEC) with a window length of 71, the extended modified B-distribution (EMBD) with $\alpha = 0.045$ and $\beta = 0.045$, the compact-support-kernel TFD (CSK-TFD) with $C = 0.1$ and $D = 0.1$, and the directional-Gaussian-filter TFD (DGF-TFD) with $a = 3$ and $b = 8$ (see [Section 5.10](#)). The (t,f) features are extracted from these TFDs while the corresponding t -domain and f -domain features are extracted from the t -domain and f -domain representations, respectively.

[Table 12.6.1](#) shows the AUC values for the extracted features. The results indicate two key findings as shown below.

- The (t,f) features have higher AUC compared to the corresponding t -domain or f -domain features in most of the cases
- The performance of the (t,f) features depends on the choice of TFD.

12.6.5 SUMMARY AND CONCLUSIONS

This section demonstrates that the (t,f) pattern-recognition approach offers significant advantages over signal classification using t -domain or f -domain features. In order to demonstrate the significance of the (t,f) approach, a simple methodology for defining new (t,f) features is proposed by extending the t -domain or f -domain features to the joint (t,f) domain. The proposed methodology is applied to define new (t,f) features such as the (t,f) flatness and (t,f) flux for the detection of seizure activity in newborn EEG signals. The performance of the extended (t,f) features is evaluated using AUC analysis.

Table 12.6.1: Result of the ROC Analysis of the (t,f) -Domain, t -Domain, and f -Domain Features for the Detection of Seizures in Newborn EEG Signals

| Feature | TFD | | | | | |
|---------------------|------|------|------|---------|---------|------------------------------|
| | WVD | SPEC | EMBD | CSK-TFD | DGF-TFD | None (f - or t -Domain) |
| Flux | 0.73 | 0.63 | 0.74 | 0.71 | 0.78 | 0.55 |
| Flatness | 0.69 | 0.62 | 0.64 | 0.80 | 0.91 | 0.66 |
| Rényi entropy | 0.61 | 0.72 | 0.68 | 0.72 | 0.66 | 0.59 |
| Mean | 0.53 | 0.54 | 0.53 | 0.51 | 0.54 | 0.51 |
| Variance | 0.53 | 0.67 | 0.62 | 0.54 | 0.62 | 0.51 |
| Skewness | 0.54 | 0.82 | 0.83 | 0.82 | 0.79 | 0.87 |
| Kurtosis | 0.61 | 0.82 | 0.84 | 0.82 | 0.80 | 0.64 |
| Coeff. of variation | 0.51 | 0.67 | 0.61 | 0.67 | 0.61 | 0.50 |

Experimental results indicate that the extended (t,f) features give better performance compared to the corresponding t -domain and f -domain features (with the exception of skewness); e.g., the (t,f) flatness (when extracted from the DGF-TFD) shows an improvement of 24% compared to the spectral flatness. In addition, the classification performance of (t,f) features depends on the choice of TFD and its parameters; e.g., the spectrogram is the best distribution for extracting variance, but the DGF-TFD is the best distribution for extracting the (t,f) flatness. The proposed methodology of extending t -domain and f -domain features to the joint (t,f) domain can be applied to define additional features for other pattern-recognition applications. Other methods for extracting features from TFDs are discussed in Section 16.3.

REFERENCES

- [1] H.L.L. Van Trees, Detection, Estimation, and Modulation Theory, vol. III: “Radar-Sonar Signal Processing and Gaussian Signals in Noise”, Wiley, New York, 1971, reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.
- [2] R.G. Shenoy, T.W. Parks, “The Weyl correspondence and time-frequency analysis”, *IEEE Trans. Signal Process.* 42 (2) (1994) 318-331.
- [3] H.V. Poor, An Introduction to Signal Detection and Estimation, Springer, New York, 1988.
- [4] J.G. Proakis, Digital Communications, third ed., McGraw-Hill, New York, 1995.
- [5] P. Flandrin, “A time-frequency formulation of optimum detection”, *IEEE Trans. Acoust. Speech Signal Process.* 36 (9) (1988) 1377-1384.
- [6] A.M. Sayeed, D.J. Jones, “Optimal detection using bilinear time-frequency and time-scale representations”, *IEEE Trans. Signal Process.* 43 (12) (1995) 2872-2883.
- [7] E. Chassande-Mottin, P. Flandrin, “On the time-frequency detection of chirps”, *Appl. Comput. Harmon. Anal.* 6 (2) (1999) 252-281.
- [8] O. Rioul, P. Flandrin, “Time-scale energy distributions: a general class extending wavelet transforms”, *IEEE Trans. Signal Process.* 40 (7) (1992) 1746-1757.
- [9] A.M. Sayeed, D.L. Jones, “Optimum quadratic detection and estimation using generalized joint signal representations”, *IEEE Trans. Signal Process.* 44 (12) (1996) 3031-3043.
- [10] C. Richard, R. Lengellé, “Data-driven design and complexity control of time-frequency detectors”, *Signal Process.* 77 (1) (1999) 37-48.
- [11] M. Davy, C. Doncarli, G.F. Boudreux-Bartels, “Improved optimization of time-frequency-based signal classifiers”, *IEEE Signal Process. Lett.* 8 (2) (2001) 52-57.
- [12] B. Boashash, G. Azemi, “A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG”, *Digital Signal Process.* 28 (2014) 28-38.
- [13] F. Hlawatsch, G.F. Boudreux-Bartels, “Linear and quadratic time-frequency signal representations”, *IEEE Signal Process. Mag.* 9 (2) (1992) 21-67.
- [14] F. Auger, P. Flandrin, “Improving the readability of time-frequency and time-scale representations by the reassignment method”, *IEEE Trans. Signal Process.* 43 (5) (1995) 1068-1089.
- [15] R.G. Baraniuk, D.L. Jones, “A signal-dependent time-frequency representation: optimal kernel design”, *IEEE Trans. Signal Process.* 41 (4) (1993) 1589-1602.
- [16] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 519-568.
- [17] S.G. Mallat, Z. Zhang, “Matching pursuits with time-frequency dictionaries”, *IEEE Trans. Signal Process.* 41 (12) (1993) 3397-3415.

- [18] A. Bultan, "A four-parameter atomic decomposition of chirplets", *IEEE Trans. Signal Process.* 47 (3) (1999) 731-745.
- [19] M.R. McClure, L. Carin, "Matching pursuits with a wave-based dictionary", *IEEE Trans. Signal Process.* 45 (12) (1997) 2912-2927.
- [20] A. Papandreou-Suppappola, S.B. Suppappola, "Analysis and classification of time-varying signals with multiple time-frequency structures", *IEEE Signal Process. Lett.* 9 (3) (2002) 92-95.
- [21] S. Qian, D. Chen, "Decomposition of the Wigner distribution and time-frequency distribution series", *IEEE Trans. Signal Process.* 42 (10) (1994) 2836-2842.
- [22] S. Pon Varma, A. Papandreou-Suppappola, S.B. Suppappola, "Matching pursuit classification for time-varying acoustic emissions", in: Proceedings of the 35th Asilomar Conference on Signals, Systems, and Computers, 2001, paper TA2-3.
- [23] S. Das, I. Kyriakides, A. Chattopadhyay, A. Papandreou-Suppappola, "Monte Carlo matching pursuit decomposition method for damage quantification in composite structures", *J. Intell. Mater. Syst. Struct.* 20 (6) (2009) 647-658.
- [24] T. Söderström, P. Stoica, *System Identification*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [25] J.C. Ralston, A.M. Zoubir, B. Boashash, "Identification of a class of nonlinear systems under stationary non-Gaussian excitation", *IEEE Trans. Signal Process.* 45 (3) (1997) 719-735, ISSN 1053-587X, <http://dx.doi.org/10.1109/78.558490>.
- [26] X.-G. Xia, "System identification using chirp signals and time-variant filters in the joint time-frequency domain", *IEEE Trans. Signal Process.* 45 (8) (1997) 2072-2084.
- [27] X.-G. Xia, V.C. Chen, "A quantitative SNR analysis for the pseudo Wigner-Ville distribution", *IEEE Trans. Signal Process.* 47 (10) (1999) 2891-2894.
- [28] L. Galleani, "Response of dynamical systems to nonstationary inputs", *IEEE Trans. Signal Process.* 60 (11) (2012) 5775-5786.
- [29] A. Jung, G. Tauböck, F. Hlawatsch, "Compressive spectral estimation for nonstationary random processes", *IEEE Trans. Inform. Theory* 59 (5) (2013) 3117-3138.
- [30] S. Viollet, J. Zeil, "Feed-forward and visual feedback control of head roll orientation in wasps (*Polistes humilis*, *Vespidae, Hymenoptera*)", *J. Exp. Biol.* 216 (7) (2013) 1280-1291.
- [31] A. Yin, L. Zhao, Z. Yang, B. Chen, "Noise reduction method for vibration signals 2D time-frequency distribution using anisotropic diffusion equation", *Math. Methods Appl. Sci.* (2014) <http://dx.doi.org/10.1002/mma.3092>.
- [32] G.F. Boudreax-Bartels, T.W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 34 (3) (1986) 442-451.
- [33] S. Qian, D. Chen, *Joint Time-Frequency Analysis: Methods & Applications*, Prentice-Hall, Upper Saddle River, NJ, 1996.
- [34] X.-G. Xia, S. Qian, "Convergence of an iterative time-variant filtering based on discrete Gabor transform", *IEEE Trans. Signal Process.* 47 (10) (1999) 2894-2899.
- [35] B. Boashash, P. O'Shea, "Detection & classification of underwater transients by time frequency analysis", *J. Electr. Electron. Eng. Aust.* 9 (3) (1989) 63-74.
- [36] F. Hlawatsch, G. Matz, H. Kirchauer, W. Kozek, "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation", *IEEE Trans. Signal Process.* 48 (5) (2000) 1417-1432.
- [37] W. Kozek, H.G. Feichtinger, J. Schäringer, "Matched multiwindow methods for the estimation and filtering of nonstationary processes", in: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS 96), vol. 2, 1996, pp. 509-512.
- [38] G. Matz, F. Hlawatsch, "Minimax robust time-frequency filters for nonstationary signal estimation", in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'99), 1999, pp. 1333-1336.

- [39] G. Matz, F. Hlawatsch, “Minimax robust nonstationary signal estimation based on a p -point uncertainty model”, *J. Franklin Institute* 337 (4) (2000) 403-419.
- [40] S.A. Kassam, H.V. Poor, “Robust techniques for signal processing: a survey”, *Proc. IEEE* 73 (3) (1985) 433-481.
- [41] G. Matz, F. Hlawatsch, A. Raidl, “Signal-adaptive robust time-varying Wiener filters: best subspace selection and statistical analysis”, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’01), 2001, pp. 3945-3948.
- [42] G. Matz, F. Hlawatsch, “Time-frequency formulation and design of optimal detectors”, in: Proceedings of the IEEE-SP International Symposium on Time-Frequency & Time-Scale Analysis, 1996, pp. 213-216.
- [43] G. Matz, F. Hlawatsch, “Time-frequency methods for signal detection with application to the detection of knock in car engines”, in: Proceedings of the Ninth IEEE Workshop on Statistical Signal and Array Processing (SSAP-98), 1998, pp. 196-199.
- [44] G. Matz, A. Raidl, “Robust detection of nonstationary random signals belonging to p -point uncertainty classes”, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’03), vol. 6, 2003, pp. 641-644.
- [45] S. Carstens-Behrens, M. Wagner, J.F. Böhme, “Detection of multiple resonances in noise”, *Int. J. Electron. Commun.* 52 (5) (1998) 285-292.
- [46] G. Matz, F. Hlawatsch, “Time-frequency subspace detectors and application to knock detection”, *Int. J. Electron. Commun.* 53 (6) (1999) 379-385.
- [47] H.A. Khan, L.F. Chaparro, “Nonstationary Wiener filtering based on evolutionary spectral theory”, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’97), vol. 5, 1997, pp. 3677-3680.
- [48] J.A. Sills, E.W. Kamen, “Time-varying matched filters”, *Circuits Syst. Signal Process.* 15 (5) (1996) 609-630.
- [49] F. Hlawatsch, G. Matz, “Time-frequency methods for non-stationary statistical signal processing”, in: F. Hlawatsch, F. Auger (Eds.), *Time-Frequency Analysis: Concepts and Methods*, ISTE/Wiley, London, UK, 2008, pp. 279-320 (Chapter 10).
- [50] B. Boashash, P.J. O’Shea, “A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques”, *IEEE Trans. Acoust. Speech Signal Process.* 38 (11) (1990) 1829-1841.
- [51] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [52] J.M. O’Toole, B. Boashash, “Time-frequency detection of slowly-varying periodic signals with harmonics: methods and performance evaluation”, *EURASIP J. Adv. Signal Process.* 2011 (2011) article ID 193797, 16 pp.
- [53] C. Richard, “Time-frequency-based detection using discrete-time discrete-frequency Wigner distributions”, *IEEE Trans. Signal Process.* 50 (9) (2002) 2170-2176.
- [54] M.S. Khelif, P.B. Colditz, B. Boashash, “Effective implementation of time-frequency matched filter with adapted pre and postprocessing for data-dependent detection of newborn seizures”, *Med. Eng. Phys.* 35 (12) (2013) 1762-1769.
- [55] B. Boashash, G. Azemi, N.A. Khan, “Principles of time-frequency feature extraction for change detection in non-stationary signals: applications to newborn EEG abnormality detection”, *Pattern Recogn.* 48 (3) (2015) 616-627, <http://dx.doi.org/10.1016/j.patcog.2014.08.016>.
- [56] H. Hassanpour, M. Mesbah, B. Boashash, “Time-frequency feature extraction of newborn EEG seizure using SVD-based techniques”, *EURASIP J. Appl. Signal Process.* 16 (2004) 2544-2554.
- [57] A.T. Tzallas, M.G. Tsipouras, D.I. Fotiadis, “Automatic seizure detection based on time-frequency analysis and artificial neural networks”, *Comput. Intell. Neurosci.* 2007 (2007) <http://dx.doi.org/10.1155/2007/80510>, article ID 80510, 13 pp.

- [58] S. Aviyente, E.M. Bernat, S.M. Malone, W.G. Iacono, "Time-frequency data reduction for event related potentials: combining principal component analysis and matching pursuit", *EURASIP J. Adv. Signal Process.* 2010 (2010) <http://dx.doi.org/10.1155/2010/289571>, article ID 289571, 13 pp.
- [59] J. Löfhede, M. Thordstein, N. Löfgren, A. Flisberg, M. Rosa-Zurera, I. Kjellmer, K. Lindecrantz, "Automatic classification of background EEG activity in healthy and sick neonates", *J. Neural Eng.* 7 (1) (2010) article ID 016007, 12 pp.
- [60] H. Hassanpour, M. Mesbah, B. Boashash, "EEG spike detection using time-frequency signal analysis", in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'04), vol. 5, 2004, pp. 421-424.
- [61] V. Sucic, N. Saulig, B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy", *EURASIP J. Adv. Signal Process.* 2011 (1) (2011) <http://dx.doi.org/10.1186/1687-6180-2011-125>, article ID 2011:125, 11 pp.
- [62] B. Boashash, L. Boubchir, G. Azemi, "A methodology for time-frequency image processing applied to the classification of non-stationary multichannel signals using instantaneous frequency descriptors with application to newborn EEG signals", *EURASIP J. Adv. Signal Process.* 2012 (2012) <http://dx.doi.org/10.1186/1687-6180-2012-117>, article ID 117, 21 pp.
- [63] K.-C. Wang, "The feature extraction based on texture image information for emotion sensing in speech", *Sensors* 14 (9) (2014) 16692-16714.
- [64] K. Fu, J. Qu, Y. Chai, Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM", *Biomed. Signal Process. Control* 13 (2014) 15-22.
- [65] A. Saleem, A. Beghdadi, B. Boashash, "Image fusion-based contrast enhancement", *EURASIP J. Image Video Process.* 2012 (10) (2012) 17, <http://dx.doi.org/10.1186/1687-5281-2012-10>.
- [66] M. Bennamoun, B. Boashash, "A structural-description-based vision system for automatic object recognition", *IEEE Trans. Syst. Man Cybern. B Cybern.* 27 (6) (1997) 893-906, ISSN 1083-4419, <http://dx.doi.org/10.1109/3477.650052>.
- [67] T. Fawcett, "An introduction to ROC analysis", *Pattern Recogn. Lett.* 27 (8) (2006) 861-874.

This page intentionally left blank

PART

ENGINEERING APPLICATIONS

V

This page intentionally left blank

TIME-FREQUENCY METHODS IN COMMUNICATIONS

13

INTRODUCTION AND OVERVIEW

The wide range of potential applications of time-frequency (t,f) methods made them an important tool in most fields of science and engineering. A large number of approaches exist, depending on the application considered. Key (t,f) methodologies are presented on specific applications in this Part V of the book and illustrated using selected examples. Telecommunications is one of the key industries where (t,f) methods are already playing an important role. The topic is represented by five sections selected for this chapter, complemented by other sections in other chapters such as [Sections 8.5](#) and [11.3](#).

Due to possible hostile jamming, broadband communication platforms use spread spectrum technology where interference protection is achieved by interference excision. By distributing the signature of received data over a (t,f) region, it is possible to attenuate strong interferences ([Section 13.1](#)). Linear dispersion in wireless communication channels distorts the transmitted signal in both time and frequency. This is accounted for by a (t,f) scattering function. In wireless communication systems with Code Division Multiple Access (CDMA) protocol, fading and multiaccess interference can be dealt with using time-frequency processing. A (t,f) RAKE receiver is described which implements correlations in a (t,f) domain and accounts for both spectral and temporal channel variations resulting from the use of spread spectrum techniques ([Section 13.2](#)). Eigenfunctions of linear systems can be modeled by signals with a (t,f) distribution well localized in the (t,f) plane. The knowledge of the eigenfunctions of time-varying transfer functions allows optimizing the transmission strategy to take advantage of the channel dispersive properties ([Section 13.3](#)). Detection and parameter estimation of chirps in communication systems may be implemented using the fractional Fourier transform ([Section 13.4](#)). The last section focuses on the (t,f) estimation of radio-signal modulation parameters and includes a discussion on cognitive radio, quality of service and communication channels from a (t,f) perspective ([Section 13.5](#)).

13.1 TIME-FREQUENCY INTERFERENCE MITIGATION IN SPREAD SPECTRUM COMMUNICATION SYSTEMS⁰

13.1.1 SPREAD-SPECTRUM SYSTEMS AND INTERFERENCE

In modern broadband communication systems, mitigation of correlated interference is an important aspect of system performance. Benign jamming sources exist ubiquitously in the transmission channel of multiple access systems like cellular telephony and wireless networks. Hostile jamming is certainly a significant issue in military communication systems. For these and many other reasons, broadband communication platforms employ spread spectrum (SS) technology [1], in which

- (1) the signal occupies a bandwidth much in excess of the minimum bandwidth necessary to send the information;
- (2) spreading is accomplished by means of spreading signal, which is often called a code signal that is independent of the data;
- (3) at the receiver, despreading for the recovering the original data is accomplished by the correlation of the received spread signal with a synchronized replica of the spreading signal.

A commonly employed SS technique called direct sequence (DS) is to superimpose a pseudorandom noise (PN) sequence on the data bits. This effectively widens the signal bandwidth by a factor proportional to the ratio of the rate of the PN sequence divided by the data rate. The advantage of this spreading is that the wider bandwidth essentially marginalizes narrowband interference sources, so they have a smaller effect on overall system performance.

At the receiver, the cross correlation with the replica of the PN sequence transfers the information signal back to its original bandwidth while reducing the level of the narrowband interference by spreading it across the bandwidth occupied by the PN sequence. Since the availability of the code at the receiver enables despreading and recovery of data while spreading and suppression of interference, any level of interference rejection can be achieved by using sufficient processing gain. This, however, may entail increasing the bandwidth of the transmitted signal beyond the limits of the available frequency spectrum. Therefore, signal processing techniques have been used in conjunction with the DS spread spectrum receiver to augment the processing gain, permitting greater interference protection without an increase in bandwidth [2].

Typically, interference excision in direct-sequence spread-spectrum (DSSS) systems is performed in the following domains [3].

Frequency domain: The fast Fourier transform (FFT) of the data over the information bit is weighted by appropriate values and then transformed back to the time domain. This is an effective method for stationary narrowband interference. Sidelobes may present a problem in removing the interference without losing some of the signal energy.

Time domain: This includes clipping or gating the high energy regions. It also includes Wiener filtering, adaptive linear predictors, and smoothers. Tracking is highly dependent on the signal-to-noise ratio (SNR) and often fails under rapidly time-varying interference.

⁰Authors: **M.G. Amin**, Villanova University, Villanova, PA, USA (Moeness.Amin@villanova.edu) and **A.R. Lindsey**, Austral Engineering and Software, Inc., Athens, OH, USA (alan.lindsey@australengineering.com). Reviewers: S. Batalama and D.L. Jones.

Time and frequency domains: A transversal filter is designed from the spectral information of the data. Spectral estimation methods combined with open-loop adaptive filtering have been shown to suffer from the same drawbacks as frequency-domain techniques.

Wavelet/Gabor domain: The discrete wavelet transform (DWT) or the Gabor transform is applied to the data, and the coefficients of high energy are removed prior to the inverse transform. The DWT is appropriate for cases of pulse jamming or interference with burst characteristics. The Gabor transform is an effective excision tool only when the interference is consistent with the corresponding tiling of the time-frequency (t, f) plane. The same is true for the wavelet transform [4].

None of the above methods is capable of effectively incorporating the time-varying nature of the interference frequency characteristics. These methods all suffer from their lack of intelligence about the other domain(s) and therefore are limited in their results and their applicability. To illustrate, Fig. 13.1.1 shows that most frequency- and time-domain excisions, in essence, remove all desired signal information over the frequency band ΔF and time duration ΔT . As such, in the case of time-varying interferences, frequency-domain methods ignore the fact that only few frequency bins are contaminated by the jammer at a given time. Dually, time domain excision techniques do not account for the cases where only few time samples are contaminated by the jammer for a given frequency. Applying either method will eliminate the interference but at the cost of unnecessarily reducing the desired signal energy.

The above example clearly demonstrates that nonstationary interferers, which have model parameters that change with time, are particularly troublesome due to the inability of single-domain mitigation algorithms to adequately ameliorate their effects. In this challenging situation, and others like it, joint (t, f) techniques can provide significant performance gains, since the instantaneous frequency (IF), the instantaneous bandwidth, and the energy measurement, in addition to myriad other parameters, are available. The objective is then to estimate the (t, f) signature of the received data using (t, f) distributions (TFDs), attenuating the received signal in those *time-frequency regions* that contain strong interference (see Chapter 3 for details). This is depicted by the region in between the fine lines in Fig. 13.1.1.

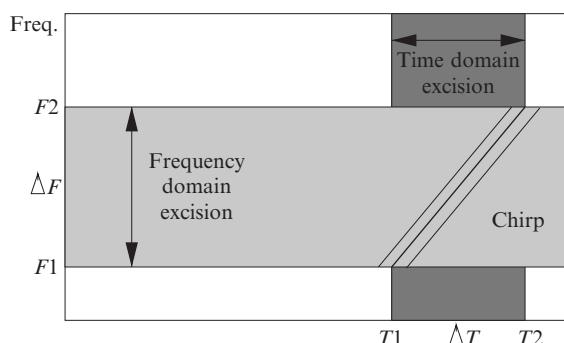


FIGURE 13.1.1

Excision methods for nonstationary signals.

13.1.2 TYPICAL SIGNAL MODEL

The signal model $r(t)$ arriving at a receiver is in the form

$$r(t) = s(t) + n(t) + j(t), \quad (13.1.1)$$

which is composed of the desired spectrally spread signal $s(t)$, the composite additive jamming signal $j(t)$ which may be made up of several different jammers from various sources, and the obligatory uncorrelated thermal noise, $n(t)$, commonly assumed to be white with a Gaussian distribution.

13.1.3 A TIME-FREQUENCY DISTRIBUTION PERSPECTIVE

An application of (t,f) methods in the underlying problem is the design of adaptive notch filters using the IF of the jamming signal [3]. The assumption that $n(t)$ and $s(t)$ are spread across the (t,f) plane and without features. In this model, the correlated features in $j(t)$ rise conspicuously in the (t,f) distribution performed at the input of the receiver, and subsequent IF information allows for the design of a strong notch filter that effectively excises only the portion of the (t,f) spectrum contaminated by the jammer. The process is illustrated in Fig. 13.1.2. The notch filter developed from the TFD can be very short having been shown to be very effective with as few as three or five coefficients.

Let us consider the performance analysis of this system using a general class of multiple-zero FIR excision filters showing the dependence of the bit error rate (BER) on the filter order and its spectral delay [5]. The effect of inaccuracies in the jammer IF on receiver performance is also considered, as a function of the filter notch bandwidth. Closed form approximations for signal to interference plus noise ratio (SINR) at the receiver are given for the various cases. The general form is

$$\text{SINR} = \frac{E^2[D_s]}{\text{Var}(D_s) + \text{Var}(D_j) + \text{Var}(D_n)}, \quad (13.1.2)$$

where the three quantities in the denominator represent the variances of the decision variable, due to the SS signal, the jammer, and the noise respectively. The received signal $r(t)$ in Eq. (13.1.1) is processed with the excision filter and then the result is correlated with the PN sequence $p(t)$ [5]. The parameter D_s

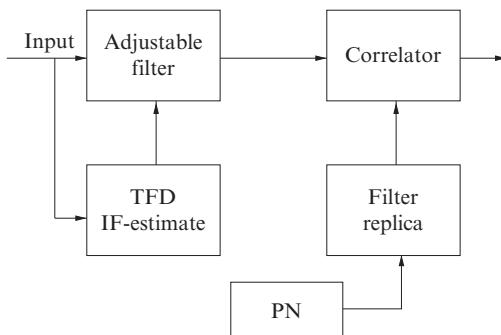


FIGURE 13.1.2

TFD excision system of Amin et al.

is the correlation between $p(t)$ and filtered $s(t)$, D_j is the correlation between $p(t)$ and the filtered $j(t)$, and D_n comes from correlating $p(t)$ and the filtered $n(t)$. The sum of all these three correlators yields the decision variable [5]. $E[\cdot]$ is the expectation operator, and with typical independent and uncorrelated characteristics of jammers and noise, it only leads to nonzero values when applied to the desired signal. If the jammer is fully excised then $\text{Var}(D_j) = 0$. It is clear from the above equation that one of the main goals of the excision process should be reducing the self-noise represented by the term $\text{Var}(D_s)$.

One of the drawbacks to the notch filter approach in Ref. [3] is the infinite notch depth due to the placement of the filter zeros. The effect is a “self-noise” inflicted on the received signal by the action of the filter on the PN sequence underlying the spread information signal. This problem led to the design of an open-loop filter with adjustable notch depth based on the jammer energy. The notch depth is determined by a variable embedded in the filter coefficients chosen as the solution to an optimization problem which maximizes receiver SINR cost function given in Eq. (13.1.2). The TFD is necessary, even for single component signals, because simple IF estimators do not provide energy information [6]. By incorporating a “depth factor” into the analysis and redeveloping all the SINR calculations, the result was significant improvement in SINR, especially at mid-range jammer-to-signal ratios (JSRs), typically around 0-20 dB.

Barbarossa and Scaglione [7] proposed a two-step procedure based on dechirping techniques commonly applied in radar algorithms (see Section 13.3). In the first step, the time varying jammer is converted to a fixed frequency sinusoid eliminated by time invariant filters. The process is reversed in the second step and the jammer-free signal is multiplied by the jammer (t,f) signature to restore the DSSS signal and noise characteristics which have been strongly impacted in the first phase. Comparison of this technique with time-varying excision filters is yet to be conducted in terms of computational complexity and robustness to IF estimation errors.

Synthesis of the interfering signal from information available at the receiver follows by direct subtraction [8]. In many situations, it is possible to make assumptions about certain key parameters which describe fully the nature and composition of a jammer, and then utilize signal processing to extract these parameters from the received signal. Once the parameters are determined, a replica of the jammer can be derived and subtracted from the incoming signal to produce an essentially jammer free channel. However, for this process to work, a jammer of constant modulus and polynomial phase is required. These parameters are extracted at the receiver via TFD where the optimally matched signal in a least squares sense is constructed. The jamming signal is projected to the constant modulus and then phase-matched. The result, especially in high SNR environments, is a signal matching the jammer in amplitude, frequency profile and phase. The last step of generating the difference signal, which is theoretically also the desired signal after the subtraction of the jammer, is straightforward.

To overcome the drawbacks of the least squares synthesis methods, a projection filter approach in which the nonstationary interference subspace is constructed from its (t,f) signature [9]. Since the signal space at the receiver is not specifically mandated, it can be rotated such that a single jammer becomes one of the basis functions. In this way, the jammer subspace is one dimensional and its orthogonal subspace is jammer-free. A projection of the received signal onto the orthogonal subspace accomplishes interference excision with a minimal message degradation. The projection filtering methods compare favorably over the previous notch filtering systems.

13.1.4 EXAMPLE

At this point, in order to further illustrate these excision methods, the work in Ref. [9] will be detailed since it includes comparisons between the two most prominent techniques based on TFDs currently being studied—notch filtering and projection filtering. The signal model is, as expected, given by Eq. (13.1.1), and the major theme of the work is to annihilate interference via projection of the received signal onto a “jammer-free” subspace generated from the estimated jammer characteristics. Figure 13.1.3, reprinted from Ref. [9], clearly illustrates the trade-offs between projection and notch filtering based on the JSR. In the legend, the variable a represents the adaptation parameter for the notch filtering scheme and N represents the block size, in samples, for a 128 sample bit duration in the projection method. Thus, $N = 128$ means no block processing and $N = 2$ corresponds to 64 blocks per bit being processed for projection. Since the projection and nonadaptive notch filter techniques are assumed to completely annihilate the jammer, their performance is decoupled from the jammer power and therefore correctly indicate constant SINR across the graph. The dashed line representing the notch filter with $a = 0$ is really indicating no filtering at all, since the adaptation parameter controls the depth of the notch.

It is evident from Fig. 13.1.3 that without adaptation a crossover point occurs around 2 dB, where filtering with an infinitely deep notch is advantageous. Thus when jammer power exceeds this point, presumably a user would flip a switch to turn on the excision subsystem. However, with adaptation this process happens automatically, while giving superior performance in the midrange. For the projection technique, the block size determines receiver performance conspicuously (*ceteris paribus*). Most important to note, however, is the superior performance of projection over all methods when the block size is equal to the bit duration, i.e., no block processing. It is feasible that computational complexity may warrant a trade-off between SINR and block size, in which case a hybrid implementation may be

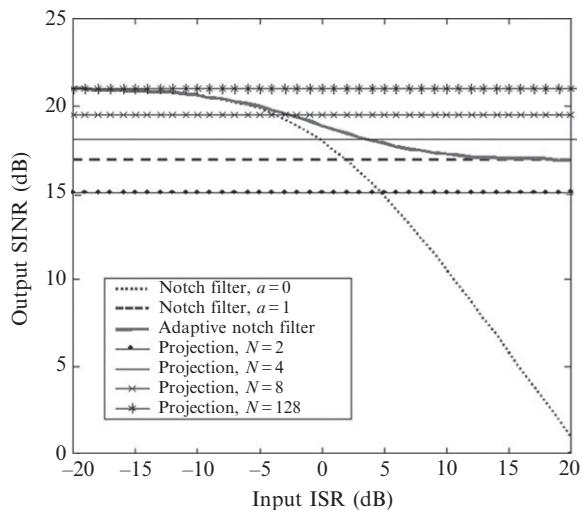


FIGURE 13.1.3

Comparison between projection and notch filtering excision methods.

of benefit—one that automatically switches between adaptive notch filtering and projection depending on the desired SINR. In any case, this example illustrates the parameters involved in the design of modern excision filters for nonstationary jammers.

13.1.5 SUMMARY AND CONCLUSIONS

The prime objective of interference suppression in broadband signal platforms is to cancel the interference with minimum distortions of the desired signal. Time-frequency signal representation provides the mechanism to achieve that objective for a large class of nonstationary interference signals. The signal localization in the (t,f) domain allows signal processing, acting on information of the instantaneous frequency and bandwidth, to play an effective role in enhancing the receiver performance and improving the bit error rates over existing techniques that deal with only the time-domain or the frequency-domain.

We have presented the problem of interference excision from a quadratic (t,f) perspective. Joint (t,f) and time-scale linear transforms have also been successfully applied to suppress nonstationary jammers. Further citations in this area can be found in Refs. [2,10].

Probably the most ubiquitous and consistently apropos application for the (t,f) perspective on nonstationary interference mitigation discussed in this chapter is the global positioning system (GPS) which utilizes DSSS signaling, is globally ubiquitous, and is inherently sensitive to interference because of the extremely low-energy signals. Leading work in this area can be found in Refs. [11–13].

13.2 COMMUNICATION OVER LINEAR DISPERSIVE CHANNELS: A TIME-FREQUENCY PERSPECTIVE⁰

13.2.1 LINEAR DISPERSIVE CHANNELS

Linear dispersive channels encountered in communications disperse the transmitted signal in both time and frequency. To mitigate the effects of dispersion, such as signal fading, appropriate signaling and reception can be used. Accurate modeling of channel characteristics is critical in this regard.

The effect of the channel on the transmitted signal of bandwidth B and duration T depends on the (BT) product of the signaling waveforms. Narrowband signaling schemes with $BT \approx 1$ typically suffer from significant intersymbol interference (ISI) requiring sequence (Viterbi) decoding at the receiver. Linear equalizers constitute a suboptimal low-complexity solution for narrowband systems. In contrast, wideband signaling schemes with $BT \gg 1$ can significantly mitigate ISI and side-step the requirement for equalization by exploiting the dispersion effects for improved performance. Time-frequency (t,f) methods provide useful insights in the analysis and design of such wideband communication systems. This section highlights the key aspects of this (t,f) perspective on wideband communication over linear dispersive channels, in the context of spread-spectrum code division multiple access (CDMA) systems. These general ideas hold in other applications as well.

⁰Author: A.M. Sayeed, University of Wisconsin–Madison, Madison, WI, USA (akbar@engr.wisc.edu). Reviewers: S. Barbarossa and M.G. Amin.

13.2.2 TIME-FREQUENCY MODEL FOR DISPERSIVE CHANNELS

Let $x(t)$ denote the transmitted complex baseband signal. The received signal after passing through a linear dispersive channel can be generally expressed as [14]

$$r(t) = s(t) + \epsilon(t) = \int_0^{T_m} h(t, \tau)x(t - \tau) d\tau + \epsilon(t), \quad (13.2.1)$$

where $h(t, \tau)$ denotes the time-varying channel impulse response, $s(t)$ is the complex baseband signal at the output of the channel, and $\epsilon(t)$ is additive white Gaussian noise (AWGN) with power spectral density σ_ϵ^2 . The maximum delay produced by the channel, T_m , is called the *multipath spread* of the channel. The following equivalent channel representation is particularly relevant from a (t, f) perspective

$$s(t) = \int_0^{T_m} \int_{-B_d}^{B_d} H(v, \tau)x(t - \tau)e^{j2\pi vt} dv d\tau, \quad (13.2.2)$$

where $H(v, \tau) = \int h(t, \tau)e^{-j2\pi vt} dt$ is the (t, f) spreading function of the channel [14], τ corresponds to temporal (multipath) spreading produced by the channel and v corresponds to the spectral (Doppler) spreading produced by the channel. The maximum Doppler frequency produced by the channel, B_d , is called the *Doppler spread* of the channel. Equation (13.2.2) states that the output signal $s(t)$ is a linear combination of time- and frequency-shifted versions of the transmitted signal $x(t)$.

13.2.2.1 Canonical (t, f) channel model

The channel model (13.2.2) being a continuum of scatterers is difficult to incorporate in receiver design. The finite duration T and (essentially) finite one-sided bandwidth B of the signaling waveform $x(t)$ can be exploited to derive a canonical (t, f) channel model that greatly facilitates system design as [14,15]

$$s(t) \approx \sum_{l=0}^L \sum_{m=-M}^M h_{ml} x_{ml}(t) = \sum_{l=0}^L \sum_{m=-M}^M \tilde{H}\left(\frac{m}{T}, \frac{l}{B}\right) x\left(t - \frac{l}{B}\right) e^{j\frac{2\pi mt}{T}}, \quad (13.2.3)$$

where $L = \lceil BT_m \rceil$, $M = \lceil B_d T \rceil$, and $\{h_{ml}\}$ are samples of a smoothed version, $\tilde{H}(v, \tau)$, of the spreading function. The model (13.2.3) is a canonical uniform (t, f) (multipath-Doppler) sampling of the channel induced by the finite duration and bandwidth of $x(t)$. The sampling resolution is $\Delta\tau = 1/B$ in time (multipath) and $\Delta v = 1/T$ in frequency (Doppler). The canonical representation (13.2.3) states that the output signal $s(t)$ belongs to a subspace spanned by the basis functions

$$x_{ml}(t) = x\left(t - \frac{l}{B}\right) e^{j\frac{2\pi mt}{T}}, \quad l = 0, \dots, L, m = -M, \dots, M \quad (13.2.4)$$

that are discretely (t, f) shifted versions of the transmitted signal. The dimension of this (active) subspace is $N_a = (L + 1)(2M + 1) \approx 2BT_m B_d + T_m B + 2TB_d + 1$ which is proportional to the BT of the signaling waveform and the channel spread factor $B_d T_m$. In practice, typical channels are *underspread*: $B_d T_m \ll 1$. While (13.2.4) provides a *fixed* basis for representing the received signal, an alternative signal representation in terms of approximate eigenfunctions of underspread channels is developed in Section 13.3. However, the eigenfunctions are not fixed and need to be computed for each channel realization. Underspread channels imply that the output signal subspace is 1D ($N_a = 1$) for narrowband ($BT = 1$) systems, whereas it is $N_a > 1$ dimensional for wideband ($BT \gg 1$) systems. As

seen later, this increase in dimension due to wideband signaling plays an important role in improved communication over dispersive channels. Note that once the signaling waveform $x(t)$ and the channel spreads T_m and B_d are known at the receiver, all information about the channel is captured *linearly* via $\{h_{ml}\}$ in (13.2.3).

13.2.2.2 Statistical channel characterization

Communication system performance depends on statistical channel characteristics. A widely used model is the wide-sense stationary uncorrelated scattering (WSSUS) model in which $\{H(v, \tau)\}$ are modeled as uncorrelated Gaussian random variables [14]: $\mathcal{E}[H(v, \tau)H^*(v', \tau')] = \Psi(v, \tau)\delta(\tau - \tau')\delta(v - v')$ where $\Psi(v, \tau)$ is the (t, f) scattering function which quantifies the relative channel power at different delays and Doppler shifts. The support of $\Psi(v, \tau)$ is limited to the channel spreads: $[0, T_m] \times [-B_d, B_d]$. Under the WSSUS assumption, the channel coefficients $\{h_{ml}\}$ in the canonical (t, f) model (13.2.3) are approximately uncorrelated if $\Psi(v, \tau)$ is sufficiently smooth. For simplicity, we focus on Rayleigh fading WSSUS channels in which $\{h_{ml}\}$ are zero mean, uncorrelated Gaussian random variables. (See detailed characterization of randomly time-varying channels in Section 9.5.)

13.2.3 COMMUNICATION OVER DISPERSIVE CHANNELS

Signals with duration T and bandwidth B form a vector space of dimension $N_o \approx BT$. Spread-spectrum waveforms used in CDMA systems take the form $q(t) = \sum_{n=0}^{N-1} c[n]v(t-nT_c)$ where $c[n]$ is the length- N spreading code associated with the waveform $q(t)$, T_c is the chip duration, and $N = T/T_c$ is the processing gain. For CDMA signals, $B \approx 1/T_c$ and $N \approx BT \approx N_o$. Given a particular spread-spectrum waveform $q(t)$, a complete basis for the signal space can be generated via distinct waveforms of the form $\{q_{ml}(t)\}$ as defined in (13.2.4). The two most significant factors affecting CDMA system performance are signal *fading* and *multiaccess interference* (MAI). Fading is due to the destructive combination of various (t, f) shifted signal copies and manifests itself as large fluctuations in the received signal power. MAI is caused by the multiple users simultaneously communicating over the channel and can drown the signal of the desired user.

We first discuss (t, f) processing in single-user CDMA systems, highlighting the concept of diversity to combat fading. We then discuss the issue of interference suppression in multiuser systems using (t, f) subspaces. For simplicity, we focus on coherent receivers with binary phase shift keying (BPSK). We assume that $T_m \ll T$; that is, the ISI is negligible. Thus, symbol by symbol detection suffices.

13.2.3.1 Time-frequency RAKE receiver

Consider a single symbol transmission in a single-user system; that is, $x(t) = bq(t)$, where $q(t)$ is the wideband signaling waveform and $b \in \{-1, 1\}$ is the transmitted bit. Equation (13.2.3) leads to the following front-end (t, f) correlation at the receiver

$$z_{ml} = \langle r, q_{ml} \rangle = \int_0^T r(t)q_{ml}^*(t) dt = b \sum_{l'=0}^L \sum_{m'=-M}^M \langle q_{m'l'}, q_{ml} \rangle h_{m'l'} + \langle \epsilon, q_{ml} \rangle. \quad (13.2.5)$$

The correlator outputs $\{z_{ml}\}$ form the N_a -dimensional representation of the noisy received signal with respect to the basis $\{q_{ml}(t)\}$. They are samples of the narrowband *cross-ambiguity function*, $A_{r,q}(v, \tau)$, between the received waveform $r(t)$ and the signaling waveform $q(t)$; that is, $z_{ml} = A_{r,q} \left(\frac{m}{T}, \frac{l}{B} \right)$ where

$A_{r,q}(v, \tau) = \int r(t)q^*(t - \tau)e^{-j2\pi v t} dt$ (Chapter 5, Eq. (5.12.24)). The correlator outputs can be expressed in vector form as $\mathbf{z} = b\mathbf{Q}\mathbf{h} + \mathbf{w}$ where \mathbf{h} is the vector of channel coefficients, $\mathbf{w} \sim \mathcal{N}(0, \sigma_e^2 \mathbf{Q})$, and \mathbf{Q} is an $N_a \times N_a$ matrix of correlation between the different basis waveforms consisting of the entries

$$\langle q_{m'l'}, q_{ml} \rangle = e^{j\frac{2\pi(m-m')l'}{BT}} A_{q,q} \left(\frac{m-m'}{T}, \frac{l-l'}{B} \right). \quad (13.2.6)$$

For typical spread-spectrum CDMA waveforms, and with chip-rate multipath sampling ($B = 1/T_c$ in (13.2.3)), the basis $\{q_{ml}(t)\}$ can be approximately orthogonal resulting in $\mathbf{Q} \approx \mathbf{I}$; that is, $\langle q_{m'l'}, q_{ml} \rangle \approx \delta_{l-l'} \delta_{m-m'}$. This implies that the *sampled* auto-ambiguity function of $q(t)$ in (13.2.6) is approximately a 2D delta function.

Given estimates of the channel coefficients $\{\hat{h}_{ml}\}$, the (t,f) correlator outputs $\{z_{ml}\}$ are coherently combined at the receiver to yield the final bit estimate

$$\begin{aligned} \hat{b} &= \text{sign} \left\{ \text{real} \left[\sum_{l=0}^L \sum_{m=-M}^M \hat{h}_{ml}^* z_{ml} \right] \right\} \\ &= \text{sign} \left\{ \text{real} \left[\hat{\mathbf{h}}^H \mathbf{z} \right] \right\} = \text{sign} \left\{ \text{real} \left[\mathbf{b} \hat{\mathbf{h}}^H \mathbf{Q} \mathbf{h} + \hat{\mathbf{h}}^H \mathbf{w} \right] \right\}, \end{aligned} \quad (13.2.7)$$

where the superscript H denotes complex conjugate transpose. The receiver (13.2.7) is a (t,f) generalization of the RAKE receiver used in conventional CDMA systems [15]. The (t,f) RAKE receiver performs joint multipath-Doppler combining to account for spectral and temporal channel variations within a symbol in contrast to the multipath combining in a conventional RAKE receiver. Conditioned on the channel coefficients \mathbf{h} , and with perfect estimates ($\hat{\mathbf{h}} = \mathbf{h}$), the bit-error-probability (BEP) of the receiver is given by $P_{e,\text{fading}}(\mathbf{h}) = Q(\text{SNR}(\mathbf{h}))$, where $\text{SNR}(\mathbf{h}) = 2\mathbf{h}^H \mathbf{Q} \mathbf{h} / \sigma_e^2$ and $Q(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$. The unconditioned BEP is given by averaging over the statistics of the channel coefficients: $P_{e,\text{fading}} = \mathcal{E}[P_{e,\text{fading}}(\mathbf{h})] = \mathcal{E}[Q(\text{SNR}(\mathbf{h}))]$.

13.2.3.2 Combating fading with time-frequency processing

The performance of the receiver (13.2.7) operating over an AWGN channel with the same *average* SNR as the fading channel is given by $P_{e,\text{AWGN}} = Q(\mathcal{E}[\text{SNR}(\mathbf{h})]) = Q(2 \text{trace}(\mathbf{Q} \mathbf{R}_h) / \sigma_e^2)$, where $\text{trace}(\cdot)$ denotes the trace of a matrix and $\mathbf{R}_h = \mathcal{E}[\mathbf{h} \mathbf{h}^H]$ is the channel correlation matrix. We note that $P_{e,\text{fading}} \geq P_{e,\text{AWGN}}$ and the loss in performance under fading due to the fluctuations in $\text{SNR}(\mathbf{h})$ is quite significant. For example, achieving a BEP of 10^{-3} over a Rayleigh fading channel requires about 18 dB additional SNR compared to an AWGN channel with the same average SNR.

Diversity signaling and reception is a technique that combats the effect of fading by transmitting the signal over multiple *independent* fading channels while keeping the total transmitted power constant. For coherent reception, as the number of diversity channels increases, the performance of the diversity reception over a fading channel approaches that of an AWGN channel at the same average SNR.

The advantage of wideband signaling over WSSUS fading channels is that the different delayed and Doppler shifted copies of the signaling waveform in (13.2.3) serve as N_a *virtual* diversity channels [15]. In essence, the receiver is able to resolve N_a weakly correlated (t,f) shifted copies ($\{q_{ml}(t)\}$) of the transmitted waveform that are linearly independent and serve as virtual diversity channels carrying the same information bit. Using the Karhunen-Loëve expansion, we can write $\text{SNR}(\mathbf{h}) = \sum_{n=1}^{N_a} \lambda_n \chi_n^2$, where $\{\chi_n^2\}$ are independent chi-squared random variables each with 2 degrees of freedom

and $\{\lambda_n\}$ are the eigenvalues of $2\mathbf{Q}\mathbf{R}_h/\sigma_e^2$. Keeping the transmitted power constant, the law of large numbers dictates that $\text{SNR}(\mathbf{h}) \xrightarrow{N_a \rightarrow \infty} \mathcal{E}[\text{SNR}(\mathbf{h})] = \text{trace}(2\mathbf{R}_h\mathbf{Q}/\sigma_e^2)$ and thus $P_{e,\text{fading}} \xrightarrow{N_a \rightarrow \infty} P_{e,\text{AWGN}}$. Thus, as the number of (t,f) signal components increases in (13.2.3), the coherent receiver (13.2.7) completely combats the loss in performance due to fading. For given channel spread factor, $B_d T_m$, the larger the BT of the signaling waveform, the higher the level of diversity. This advantage of wideband signaling is exploited in CDMA wireless communication systems by using direct-sequence spread-spectrum waveforms (in underwater acoustical communications by using of linearly frequency-modulated waveforms). In particular, Doppler diversity can be exploited via time-selective signaling, thus increasing the effective symbol duration [15,16].

13.2.3.3 Interference suppression via time-frequency subspaces

In CDMA systems, interference stems from multiple users simultaneously using the channel. In underwater acoustics, reverberation is the major source of interference. In either case, the (t,f) channel model (13.2.3) can be fruitfully exploited for effective interference suppression. The key idea is the notion of (t,f) subspaces associated with the desired signal and interference [17,18].

Consider a CDMA communication system with K users. The k th user is assigned a signaling waveform $q_k(t)$. According to the model (13.2.3), each user is associated with a canonical basis $\{q_{k,ml}(t)\}$. For simplicity, we assume that all users are transmitting synchronously. The signal at any receiver can be represented as

$$r(t) \approx \sum_{k=1}^K s_k(t) + n(t) = \sum_{k=1}^K b_k \sum_{l=0}^{L_k} \sum_{m=-M_k}^{M_k} h_{k,ml} q_{k,ml}(t) + n(t). \quad (13.2.8)$$

Suppose that we are interested in decoding the signal of the first user. We consider *decentralized* detection in which only the code of the desired user is known at the receiver [18]. In this case, bit detection is accomplished by projecting onto the waveforms $\{q_{1,ml}\}$ associated with the first user. For a discussion of centralized detection which uses information about all users, see Ref. [17].

Recall that a subset of $\{q_{1,ml}\}$ provides a complete basis for the signal space. The key idea from the viewpoint of interference suppression is that of *active* and *inactive* subspaces, \mathcal{S}_a and \mathcal{S}_{ia} , respectively. The active subspace, \mathcal{S}_a , is defined as the span of the basis elements that *lie within the multipath-Doppler channel spread* and form the diversity channels: $\mathcal{S}_a = \text{span}\{q_{1,ml} : l = 0, 1, \dots, L_1, m = -M_1, \dots, M_1\}$. The inactive subspace, \mathcal{S}_{ia} , is spanned by elements that *lie outside the channel spread*: $\mathcal{S}_{ia} = \text{span}\{q_{1,ml} : l \notin \{0, 1, \dots, L_1\}, m \notin \{-M_1, \dots, M_1\}\}$.

The N_a -dimensional vector of correlator outputs corresponding to \mathcal{S}_a is

$$\mathbf{z}_a = b_1 \mathbf{Q}_{11} \mathbf{h}_1 + \sum_{k=2}^K b_k \mathbf{Q}_{1k} \mathbf{h}_k + \mathbf{w}_a = b_1 s_1 + \mathbf{i}_a + \mathbf{w}_a, \quad (13.2.9)$$

where the first term is the signal of interest (as in the single-user case), the second term is the interference corrupting the active subspace, and the third term is background noise. The matrix \mathbf{Q}_{1k} denotes the correlation between the *active* basis waveforms of first and k th user. Let $N_{ia} \leq N_o - N_a$

denote the dimension of the inactive subspace. The vector of correlator outputs corresponding to \mathcal{S}_{ia} is given by

$$\mathbf{z}_{ia} = \sum_{k=2}^K b_k \tilde{\mathbf{Q}}_{1k} \mathbf{h}_k + \mathbf{w}_{ia} = \mathbf{i}_{ia} + \mathbf{w}_{ia}, \quad (13.2.10)$$

where $\tilde{\mathbf{Q}}_{1k}$ represents the correlation between the *inactive* basis waveforms of the first user and the *active* basis waveforms of the k th user. Note that \mathbf{z}_{ia} does not contain the signal of interest—it only contains interference (\mathbf{i}_{ia}) and noise (\mathbf{w}_{ia}). By the very nature of spread-spectrum signals, \mathbf{i}_a is correlated with \mathbf{i}_{ia} and the basic idea is to use \mathbf{i}_{ia} to suppress \mathbf{i}_a corrupting the active subspace. After interference suppression, the “cleaned-up” version of \mathbf{z}_a can be coherently combined as in the single-user case (see (13.2.7)) to decode the bits of the desired user. The dimension of \mathcal{S}_{ia} controls the receiver complexity and the level of MAI suppression.

A linear receiver may be used to accomplish both interference suppression and diversity exploitation in the multiuser case. The bit estimate for b_1 is given by

$$\hat{b}_1 = \text{sign}\{\text{real}[\mathbf{g}^H \mathbf{z}]\} = \text{sign}\{\text{real}[\mathbf{g}_a^H \mathbf{z}_a + \mathbf{g}_{ia}^H \mathbf{z}_{ia}]\}, \quad (13.2.11)$$

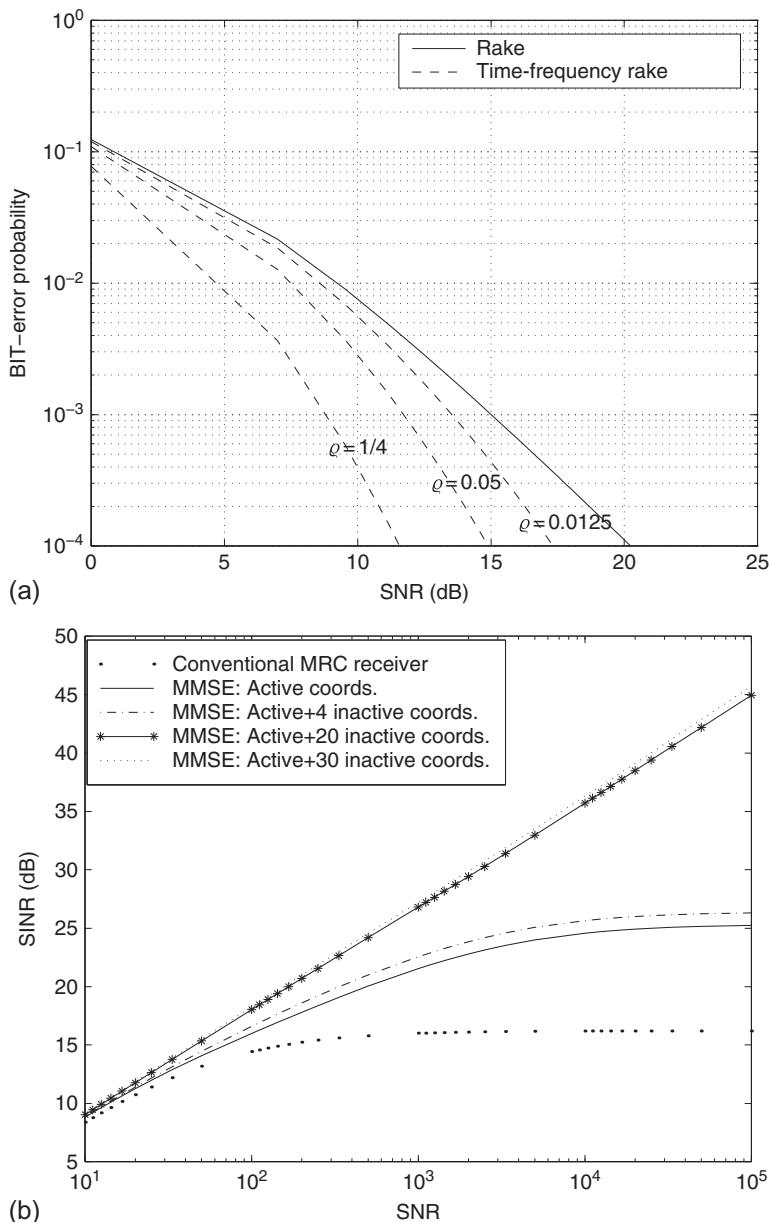
where the $N_a + N_{ia}$ dimensional filter \mathbf{g} processes the active and inactive correlator outputs. If \mathbf{g} is chosen to minimize the mean-squared-error (MSE) between \hat{b}_1 and b_1 , the linear filter admits an intuitively appealing decomposition: $\mathbf{g}_a^H \mathbf{z}_a + \mathbf{g}_{ia}^H \mathbf{z}_{ia} = \mathbf{g}_a^H [\mathbf{z}_a - \mathbf{C}^H \mathbf{z}_{ia}]$. In this decomposition, the matrix filter \mathbf{C} forms a minimum MSE (MMSE) estimate of \mathbf{z}_a from \mathbf{z}_{ia} , thereby suppressing the interference corrupting the active subspace. The filter \mathbf{g}_a then further suppresses any residual interference and exploits multipath-Doppler diversity to combat fading. The processing of N_a dimensional active correlator outputs, \mathbf{z}_a , is necessary for maximal diversity exploitation. The inactive subspace serves the sole purpose of interference suppression. Its dimension can be progressively increased to suppress any desired number of interfering users up to $N_o - N_a$. Thus, the notion of active/inactive (t,f) subspaces provides a natural framework for interference suppression and controlling the complexity of the multiuser (t,f) receiver.

13.2.3.4 An illustrative example

Recall that the key parameters are the signal space dimension, $N_o \approx BT$, which is proportional to the spreading gain $N = T/T_c$, and the products BT_m and BT_d that control the level of multipath and Doppler diversity, respectively. We consider a system with a spreading gain of $N = 64$ and a multipath spread of $T_m = 2T_c$.

Consider first a single-user system to illustrate the effects of multipath-Doppler diversity. Assume chip-rate sampling ($B = 1/T_c$) at the receiver so that $L = 2$, and suppose $0 < BT_d < 0.5$ so that $M = 1$. This results in an $N_a = 9$ dimensional active subspace. Assume uniform power in all the multipath channel components. Let $\varrho \in (0, 1/3)$ denote the ratio of the power in each of the $m = \pm 1$ Doppler components relative to the total power in the three ($m = -1, 0, 1$) components. As $BT_d \rightarrow 0.5$ (faster fading), $\varrho \rightarrow 1/3$ to yield maximum Doppler diversity.¹ Figure 13.2.1(a) shows the BEP of the

¹For a uniform Doppler power spectrum, the values $\varrho = 0.0125, 0.05, 0.25$ are achieved at $BT_d \approx 0.2, 0.4, 0.6$, respectively [15].

**FIGURE 13.2.1**

(a) Joint multipath-Doppler diversity gains achieved by a single-user coherent (t, f) RAKE receiver as a function of ϱ (performance of conventional RAKE corresponds to $\varrho = B T_d = 0$). (b) Monotonic increase in the output SINR of a multiuser (t, f) RAKE receiver with increasing dimension of the inactive (t, f) subspace. For $N_{ia} < 19$, the SINR saturates, whereas for $N_{ia} \geq 19$, interference is suppressed so that output SINR increases linearly with input SNR as in single-user systems.

coherence (t,f) RAKE receiver (13.2.7) as a function of SNR for different values of ϱ . It is evident that the (t,f) RAKE receiver yields significant gains over RAKE as ϱ increases. We note that while typical values of BT_d are relatively small in practice,² simple signaling schemes with longer symbol duration may be used to increase BT_d to achieve significant Doppler diversity [15].

To illustrate the role of active/inactive subspaces in interference suppression, we consider a slow-fading system ($BT_d = 0$) employing four-fold oversampling ($B = 4/T_c$), resulting in signal space dimension $N_o = 4N = 256$ and $N_a = 9$. There are 27 interfering users at the same power as the desired user. The signal-to-interference-and-noise-ratio (SINR) at the output of the desired user's receiver is an appropriate performance metric in the multiuser case. Figure 13.2.1(b) shows the SINR of the MMSE receiver as a function of the transmitted SNR of the desired user for different values for the inactive subspace dimension, N_{ia} . Figure 13.2.1(b) shows that the receiver goes from an interference-limited regime to a noise-limited regime as N_{ia} increases, with the optimal cut-off value around $N_{ia} = 28 - N_a = 19$. Thus, the dimension of the inactive subspace can be progressively increased to achieve a desired level of interference suppression with commensurate receiver complexity.

13.2.4 SUMMARY AND CONCLUSIONS

The (t,f) channel model (13.2.3) [14] of WSSUS channels is exploited to address several key questions relating to reliable communication over dispersive channels, including capacity and error exponents [15]. Results relating to the use of the model in wireless communications show that in terms of diversity and interference suppression, the insights offered by the (t,f) perspective improve efficient transceiver design. Similar ideas can be fruitfully exploited in the design of efficient transmission schemes for reaping the capacity of linear dispersive channels, including the design of orthogonal frequency division multiplexing (OFDM) schemes for high rate wireless communication, and underwater acoustical communications.

13.3 EIGENFUNCTIONS OF UNDERSPREAD LINEAR TIME-VARYING SYSTEMS⁰

In spite of its fundamental role and importance for wireless communications, there is no general theory on the eigendecomposition of kernels of linear time-varying (LTV) systems. A notable exception is the theory developed in Ref. [19, Chapters 1 and 2] (see also the references therein), where the authors show that the kernel of an underspread LTV system is approximately normal and its eigenfunctions can be well approximated by a set of time- and frequency-shifted replicas of a given pulse waveform, which is well localized in time and frequency. Furthermore, the eigenvalues can assume all values within a given interval. This model is shown to provide a small error in average sense, for WSSUS channels. In this section, rather than looking for models valid in average sense, we concentrate on the

² $BT_d \approx 0.013$, at carrier frequency 1.8 GHz, data rate 10 kHz, and maximum speed 50 mph.

⁰Authors: **S. Barbarossa**, Sapienza University of Rome, Rome, Italy (sergio.barbarossa@uniroma1.it) and **M. Tsitsvero**, Sapienza University of Rome, Rome, Italy (tsitsvero@diet.uniroma1.it). Reviewers: G. Matz and A.M. Sayeed.

eigen decomposition of deterministic channels (realizations of a random channel) and, building on the Wentzel-Kramers-Brillouin (WKB) method developed for slowly varying operators [20], we show that (i) the left and right singular functions can be computed exactly for systems whose delay-Doppler spread function is perfectly concentrated on a straight line; (ii) the singular functions are composed of signal components whose instantaneous frequency (IF) can be very well approximated by the contour level of the system time-varying transfer function; and (iii) the singular values coming from our model are inherently discrete for the channels having a square integrable impulse response. The theory is then validated by numerical results on multipath channels and time-frequency limiting channels.

13.3.1 CHANNEL CHARACTERIZATIONS

The input/output relationship of a continuous-time LTV system can be written

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t, \tau) x(t - \tau) d\tau = \int_{-\infty}^{\infty} k(t, \tau) x(\tau) d\tau, \quad (13.3.1)$$

where $h(t, \tau)$ is the channel response at time t to an impulse sent at time $t - \tau$, while $k(t, \tau)$ is the channel response at time t to an impulse sent at time τ . We will refer to $h(t, \tau)$ as the *time-varying impulse response* (TVIR) and to $k(t, \tau)$ as the *TVIR kernel* [14]. Any LTV channel may be equivalently characterized by the *Doppler-delay spreading function*

$$S(v, \tau) := \mathcal{F}_{t \rightarrow v} \{h(t, \tau)\} = \int h(t, \tau) e^{-j2\pi v t} dt \quad (13.3.2)$$

or by the *time-varying transfer function* (TVTF)

$$H(t, f) := \mathcal{F}_{\tau \rightarrow f} \{h(t, \tau)\} = \int h(t, \tau) e^{-j2\pi f \tau} d\tau, \quad (13.3.3)$$

where the limits of integration are from $-\infty$ to ∞ unless otherwise specified.

The corresponding adjoint channel, denoted as \mathbf{H}^* , has the input/output relation

$$y(\tau) = (\mathbf{H}^*x)(\tau) = \int h^*(t, t - \tau) x(t) dt. \quad (13.3.4)$$

13.3.2 CHANNEL DECOMPOSITION

In general, LTV systems are not Hermitian (i.e., $\mathbf{H} \neq \mathbf{H}^*$), so that they may not admit canonical eigendecomposition over a set of orthonormal eigenfunctions with real eigenvalues. However, they may be properly characterized by introducing the *left and right singular functions* with corresponding *real singular values*.

Definition 13.3.1. The functions $u_i(t)$, $v_i(t)$ are *left and right singular functions* of the system \mathbf{H} with corresponding *singular value* $\sigma_i \in \mathbb{R}^+$ if they satisfy

$$\sigma_i u_i(t) = (\mathbf{H}v_i)(t) \text{ and } \sigma_i v_i(t) = (\mathbf{H}^*u_i)(t). \quad (13.3.5)$$

It follows that $u_i(t)$ and $v_i(t)$ are the eigenfunctions of the composite Hermitian systems $\mathbf{H}\mathbf{H}^*$ and $\mathbf{H}^*\mathbf{H}$, respectively, with corresponding eigenvalues σ_i^2 .

If the system impulse response is square-integrable, i.e., if $\int \int |h(t, \tau)|^2 dt d\tau < \infty$, then there exists [21] a sequence (finite or infinite) of positive decreasing numbers $\sigma_1 \geq \dots \geq \sigma_i \geq \dots > 0$, which are the singular values of this system, and two sets of corresponding orthonormal left and right singular functions $u_i(t)$ and $v_i(t)$.

Note however that $h(t, \tau)$ is *non-square-integrable* in at least a few important cases, including linear time-invariant (LTI) channels, in which $h(t, \tau)$ is constant along t , and linear frequency-invariant (LFI) channels, in which $h(t, \tau) = m(t) \delta(\tau)$.

13.3.3 EXACT AND APPROXIMATE SOLUTIONS

13.3.3.1 Spreading function concentrated over a straight line

Consider a channel with spreading function $S(v, \tau)$ whose support is concentrated over a straight line in the (v, τ) plane, so that $S(v, \tau) = g(\tau) \delta(v - \mu\tau - f_0)$. An example could be a multipath channel with two paths. By (13.3.2), the TVIR is

$$h(t, \tau) = g(\tau) e^{j2\pi\mu\tau t} e^{j2\pi f_0 t}. \quad (13.3.6)$$

In this case, the singular functions can be expressed in closed form.

Theorem 13.3.2. *If a system \mathbf{H} has spreading function $S(v, \tau) = g(\tau) \delta(v - \mu\tau - f_0)$, where μ and f_0 are real parameters and $g(\tau)$ is a function that admits a Fourier transform $G(f)$, the left and right singular functions and singular values of \mathbf{H} are*

$$v_i(t) = e^{j2\pi f_i t} e^{j\pi\mu t^2}, \quad (13.3.7)$$

$$u_i(t) = e^{j2\pi f_0 t} e^{j\arg K(f_i, \mu)} v_i(t), \quad (13.3.8)$$

$$\sigma_i = |K(f_i, \mu)|, \quad (13.3.9)$$

where $K(f_i, \mu) = \int g(\tau) e^{-j2\pi f_i \tau} e^{j\pi\mu\tau^2} d\tau$. (Proof: Get (13.3.6) into (13.3.1).)

Remark 1. From (13.3.3), the TVTF of the system in (13.3.6) is

$$H(t, f) = e^{j2\pi f_0 t} G(f - \mu t), \quad (13.3.10)$$

while the IFs of the left and right singular functions are, respectively, $f^l(t) = f_i + \mu t$ and $f^r(t) = f_0 + f_i + \mu t$. Combining this result with (13.3.10), we conclude that the IFs of the left and right singular functions correspond to regions in the (t, f) plane where the magnitude (absolute value) of the TVTF is constant.

13.3.3.2 Spreading function with limited support

Now suppose the spreading function is mainly concentrated over a limited support in the Doppler-delay domain. For the composite systems $\mathbf{H}\mathbf{H}^*$ we can write

$$\sigma_i^2 u_i(t) = \int \mathcal{K}(t, \theta) u_i(\theta) d\theta, \quad (13.3.11)$$

where the kernel $\mathcal{K}(t, \theta)$ is defined as

$$\mathcal{K}(t, \theta) = \int h(t, t - \tau) h^*(\theta, \theta - \tau) d\tau. \quad (13.3.12)$$

Here we cannot express the singular functions in closed form. But, using the WKB method in Appendix (p. 768), we can find the closed-form *approximate* solution

$$u_i(t) = \sum_m A_{i,m}(t) e^{j\phi_{i,m}(t)}, \quad (13.3.13)$$

where $\phi_{i,m}(t)$ is such that

$$\left| H\left(t, \frac{\dot{\phi}_{i,m}(t)}{2\pi}\right) \right|^2 = \sigma_i^2, \quad (13.3.14)$$

and

$$A_{i,m}(t) = \left| \frac{\partial}{\partial f} |H(t,f)|^2 \right|^{-1/2} \Big|_{f=\dot{\phi}_{i,m}(t)/(2\pi)}. \quad (13.3.15)$$

Remark 2. The IF of each component of the left singular function is still given, as in the previous section, by the curves in the (t,f) plane where the TVTF's magnitude is constant. Furthermore, the instantaneous bandwidth (IB) $B(t) = |A'(t)/A(t)|$ of each component depends on the derivative of the TVTF evaluated along the IF curves (see definition of IB in Section 1.3.4).

Remark 3. By (13.3.15), the amplitude $A_{i,m}(t)$ diverges at points where the partial derivative is zero. Near these points—which are called **turning points** by analogy with the WKB solution of the Schrödinger equation—the solution (13.3.14) and (13.3.15) is no longer valid and the problem must be analyzed separately.

13.3.3.3 Area rule

From (13.3.14), it is apparent that the singular values are bounded by the minimum and maximum values of $|H(t,f)|$. However, not all values are possible. The only admissible values for σ_i^2 are those satisfying the *area rule*. To clarify this rule, Fig. 13.3.1 shows an example of contour line of $|H(t,f)|^2$. In this case, there are six turning points. The area rule states that the level σ_i^2 must be such that the area of the region enclosed by the contour line (gray region) must be equal to $n_i + 1/2$, where n_i is a nonnegative integer. Furthermore, the number of components in Eq. (13.3.13) is equal to the number of turning points.

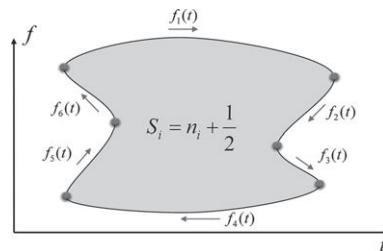


FIGURE 13.3.1

Example of a bubble with six turning points.

In physics this rule is called EBK (Einstein-Brillouin-Keller) quantization [20]. Here it gives a natural criterion for choosing the right values of σ_i . Note that the phase increments by $\pi/2$ at convex turning points and decrements by $\pi/2$ at concave turning points [20] for the traversal shown in Fig. 13.3.1.

13.3.4 NUMERICAL VALIDATION

As the analytic models have specific structures in the (t,f) domain, a natural tool for analyzing the results is a (t,f) representation with good (t,f) localization: the smoothed pseudo Wigner-Ville distribution (SPWVD) with reassignment [22].

Given a continuous-time channel with impulse response $h(v, \tau)$, the corresponding discrete-time equivalent impulse response, assuming band limitation in $[-1/2T_s, -1/2T_s]$ and sampling step T_s , is

$$h[n, n - k] := \int \int h(\theta, \tau) \operatorname{sinc}\left(\frac{nT_s - \theta}{T_s}\right) \operatorname{sinc}\left(\frac{\theta - \tau - kT_s}{T_s}\right) d\tau d\theta. \quad (13.3.16)$$

Using (13.3.16) the discrete-time counterpart of (13.3.1) may be rewritten as $\mathbf{y} = \mathbf{Hx}$, where \mathbf{H} is the channel matrix. The left and right singular vectors $\mathbf{u}_i, \mathbf{v}_i$ of \mathbf{H} can be obtained by computing the singular value decomposition (SVD) of \mathbf{H} : $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^*$, where the columns of \mathbf{U} and \mathbf{V} are the left and right singular vectors $\mathbf{u}_i, \mathbf{v}_i$ associated to the singular values σ_i contained in the diagonal matrix Σ .

We check the validity of our theoretical expressions as follows: (1) build the channel matrix \mathbf{H} ; (2) compute its SVD; (3) compute the reassigned SPWVD (RSPWVD) of the singular vectors associated to some singular value σ_i [22]; and (4) compare the results with those obtained from (13.3.14) and (13.3.15).

We consider now a few numerical examples to validate our theoretical findings.

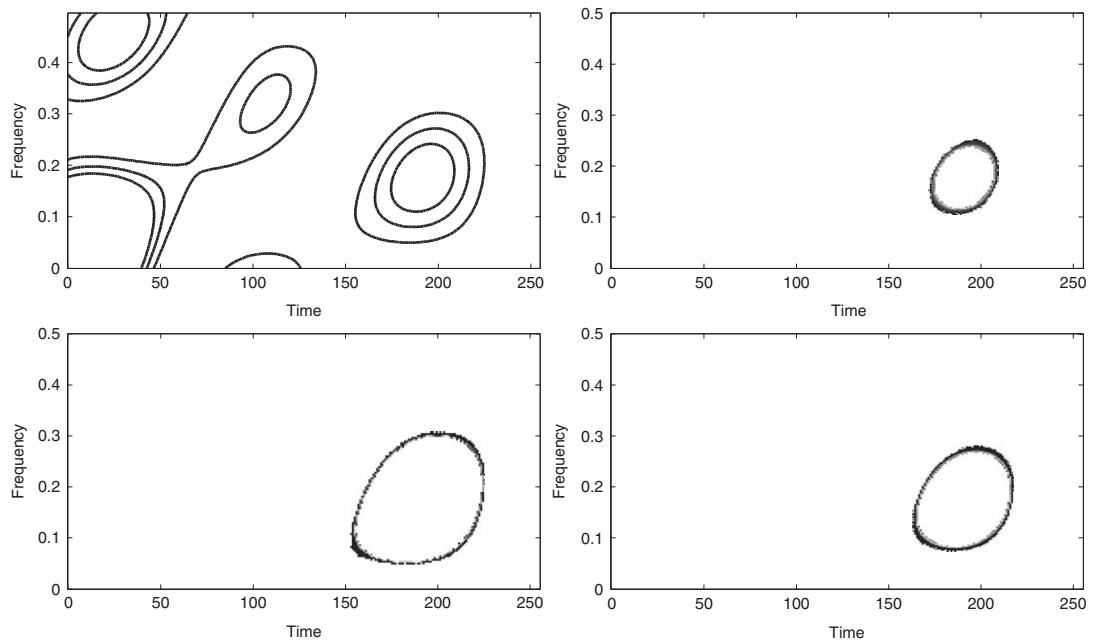
As a test case, we consider a multipath channel with Q paths, whose TVIR is

$$h(t, \tau) = \sum_{q=0}^{Q-1} h_q e^{j2\pi f_q t} \delta(\tau - \tau_q), \quad (13.3.17)$$

so that the q th path is attenuated by h_q , delayed by τ_q , and Doppler-shifted by f_q . Following the Rayleigh fading model, we treat the h_q as independent identically distributed (i.i.d.) complex Gaussian random variables having zero mean and unit variance. The shifts τ_q and f_q are modeled as uniformly i.i.d. random variables within intervals $[0, \Delta\tau]$ and $[-\Delta f/2, \Delta f/2]$, respectively. By (13.3.3), the TVTF is

$$H(t, f) = \sum_{q=0}^{Q-1} h_q e^{j2\pi f_q t} e^{-j2\pi \tau_q f}. \quad (13.3.18)$$

Simulation results use $Q = 10$, $\Delta\tau = 4T_s$, and $\Delta f = 4/(NT_s)$, with $N = 256$. In Fig. 13.3.2, three contour levels (corresponding to σ_{25} , σ_{40} , and σ_{72}) of a realization of $|H(t, f)|$ are given alongside the (t, f) distributions (RSPWVD) of the corresponding singular vectors. We observe a very good agreement between the numerical results and the behaviors predicted by the analytic models. Next, we compare the amplitudes for numerically computed singular vectors \mathbf{u}_{40} with singular functions

**FIGURE 13.3.2**

Contour lines of $|H(t, f)|$ for the multipath case corresponding to three different levels (top left), and TFDs of singular vectors u_{25} (top right), u_{40} (bottom right), and u_{72} (bottom left).

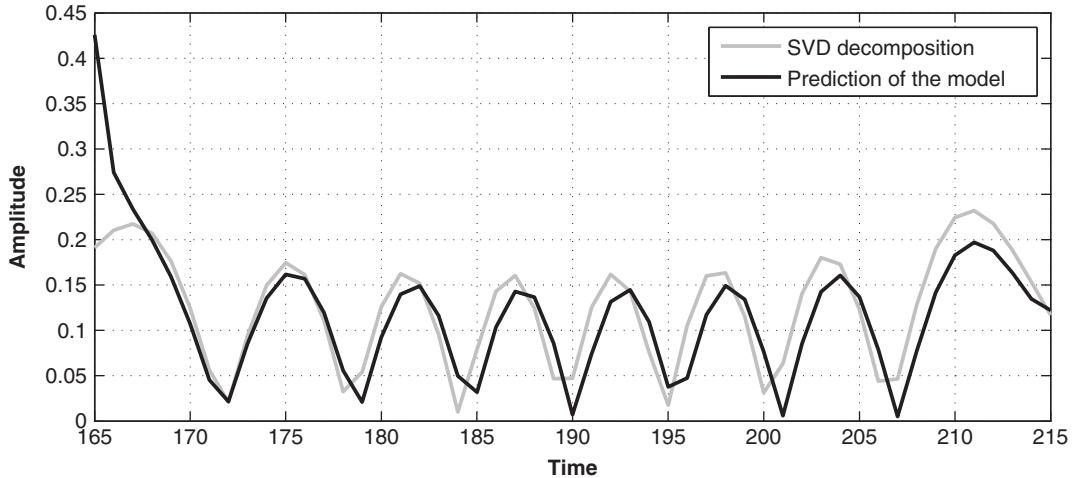
obtained from formula (13.3.15) (see Fig. 13.3.3). The slight departure between the numerical and theoretical results is due to the numerical computation of the instantaneous phase from its derivative, using trapezoidal-rule integration.

As a further example, we consider the prolate spheroidal functions, known to be singular functions of the system approximately limited in both time and frequency [23–25]. Figure 13.3.4 shows the contour plots of the RSPWVDs of three prolate spheroidal sequences corresponding to three different singular values (decreasing from top to bottom) of a given time- and bandwidth-limiting operator. It is interesting to see how these waveforms fill the (t, f) region with nested curves rather than with rectangular tiles.

13.3.5 OPTIMAL WAVEFORMS FOR LTV DIGITAL COMMUNICATION CHANNELS

Section 13.2 shows how to convert the channel dispersiveness, possibly in both t and f , into a source of diversity to enhance SNR at the receiver. Then, if the transmitter can predict the channel TVTF, at least within the next transmission time slot, the transmission strategy can be optimized to exploit the diversity.

For a channel with approximately finite impulse response (FIR) of order L , we can parse the input sequence in blocks of K symbols each and insert null guard intervals of length L between successive

**FIGURE 13.3.3**

Comparison of the amplitude of u_{40} with the prediction of the model.

blocks to avoid inter-block interference. If the symbol rate is $1/T_s$, the time taken to send each block is KT_s . For the i th block, we consider the channel $h_i(t, \tau)$ obtained by windowing $h(t, \tau)$ in time to retain the interval $[iKT_s, (i+1)KT_s]$ and in frequency to retain the band $[-1/(2T_s), 1/(2T_s)]$. The optimal strategy for transmitting a symbol set $s_i[k] := s[iK + k]$, $k = 0, \dots, K - 1$, in additive white Gaussian noise (AWGN), is to send the signal [21, Section 8]

$$x_i(t) = \sum_{k=0}^{K-1} c_{i,k} s_i[k] v_{i,k}(t), \quad (13.3.19)$$

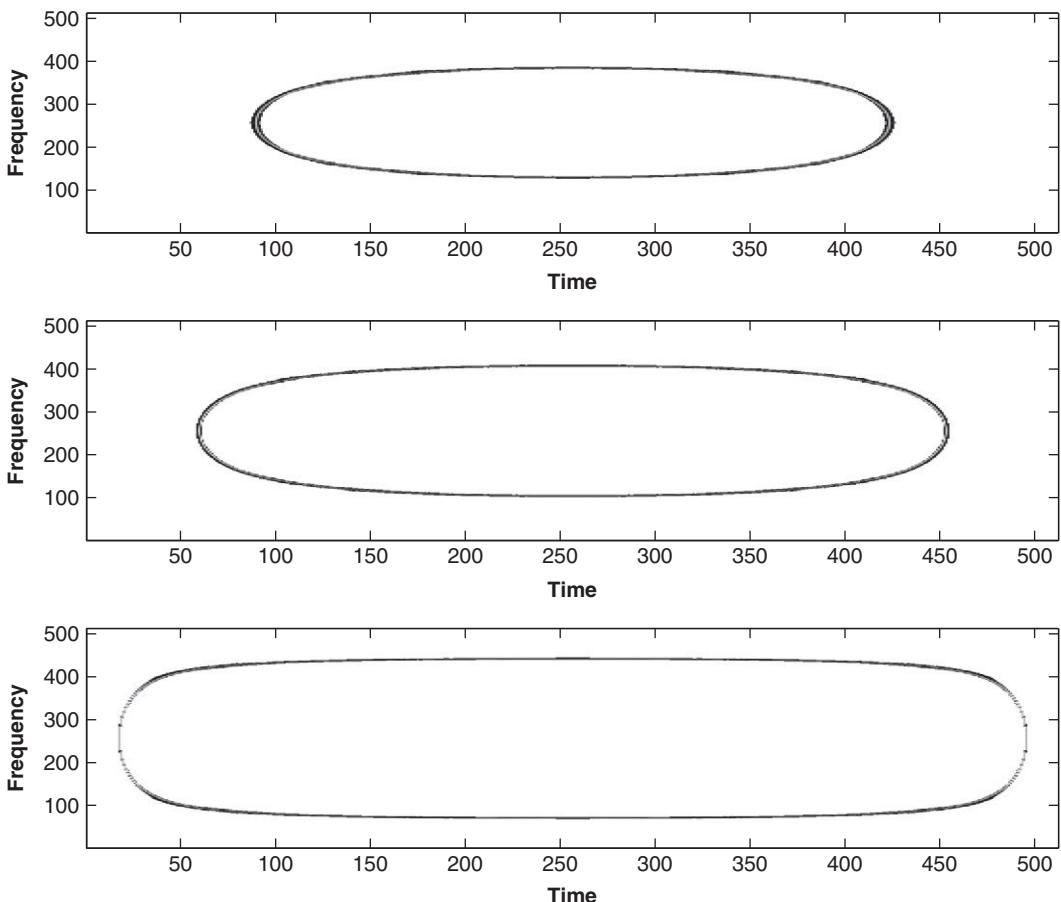
where $v_{i,k}(t)$ is the right singular function belonging to the k th eigenvalue of the channel response $h_i(t, \tau)$ in the i th transmission interval, and $c_{i,k}$ are coefficients used to allocate the available power among the transmitted symbols according to optimization criteria [26]. From (13.3.5), the received signal is

$$y_i(t) = \int_{-\infty}^{\infty} h_i(t, \tau) x_i(t - \tau) d\tau + w(t) = \sum_k c_{i,k} \sigma_{i,k} s_i[k] u_{i,k}(t) + w(t), \quad (13.3.20)$$

where $u_{i,k}(t)$ is the left singular function belonging to the k th singular value of $h_i(t, \tau)$, and $w(t)$ is AWGN. As the $u_{i,k}(t)$ are orthonormal, the transmitted symbols can be estimated as the scalar products of $y(t)$ with the left singular functions, i.e.,

$$\hat{s}_i[m] = \frac{1}{\sigma_{i,m} c_{i,m}} \int_{-\infty}^{\infty} y(t) u_{i,m}^*(t) dt = s_i[m] + w_i[m], \quad (13.3.21)$$

where the noise samples $w_i[m] := \int_{-\infty}^{\infty} w(t) u_{i,m}^*(t) dt$ are i.i.d. Gaussian random variables. Thus the initial LTV channel, possibly dispersive in both t and f , is converted into a set of parallel independent

**FIGURE 13.3.4**

(t, f) distributions of prolate spheroidal functions.

nondispersive subchannels, with no intersymbol interference, with symbol-by-symbol decision by maximum likelihood detector.

Most current transmission schemes are simple examples of the general framework in this section. For example, in a flat fading multiplicative channel, the eigenfunctions are Dirac pulses, so the optimal strategy is time-division multiplexing. By duality, the optimal strategy for LTI channels is orthogonal frequency-division multiplexing (OFDM). In the general underspread channel, the optimal strategy would consist in sending symbols through channel-dependent waveforms that fill the assigned (t, f) frame according to the nested bubble-like structures illustrated in this section. The validity of this approach is limited by the need for a sufficiently accurate short-term prediction of the channel variation, based on past observations.

13.3.6 SUMMARY AND CONCLUSIONS

The analytic model for the eigenfunctions of an underspread linear operator shows that the energy of the eigenfunctions is mainly concentrated along iso-level curves of the transfer function. The model, although approximate, is well supported by numerical results. It provides a general framework for interpreting data transmission schemes and it may guide the selection of the optimal waveforms for transmission over time-varying channels (for more details see, e.g., [26]).

13.3.7 APPENDIX

To facilitate the analysis of slowly varying systems, the kernel $\mathcal{K}(t, \theta)$ may be rewritten in terms of the difference $t - \theta$ and the mean $(t + \theta)/2$ as

$$\mathcal{K}(t, \theta) = K\left(t - \theta, \frac{\epsilon}{2}(t + \theta)\right), \quad (13.3.22)$$

where $\epsilon \ll 1$ represents the small dependence on $(t + \theta)/2$ (for a time-invariant system, $\epsilon = 0$). This suggests [20] the following form for the solution of (13.3.11):

$$u_i(t; \epsilon) = \sum_k A_{i,k}(\epsilon t; \epsilon) e^{j\phi_{i,k}(\epsilon t; \epsilon)/\epsilon}. \quad (13.3.23)$$

To clarify the meaning of (13.3.22) and (13.3.23), consider the following Taylor expansion of the amplitude and phase in (13.3.23) around an arbitrary point $t = t_0$:

$$u_i(t; \epsilon) = e^{j \sum_{n=0}^{\infty} (-1)^n \epsilon^{n-1} (t-t_0)^n \phi_i^{(n)}(\epsilon t_0)/n!} \sum_{n=0}^{\infty} (-1)^n \epsilon^n A_i^{(n)}(\epsilon t_0) (t-t_0)^n / n!. \quad (13.3.24)$$

From (13.3.24) it is clear that a solution of the form (13.3.23) can be expanded in powers of ϵ , by taking different orders for the phase and the amplitude.

Solutions of the form (13.3.23) are usually called Wentzel-Kramers-Brillouin (WKB) solutions. The method is usually applied to *differential* operators, but [20] proved it applicable to *integral* operators. What we have here is a particular case of the well-developed perturbation theory. Hereinafter we assume that series expansions in ϵ converge, although this is subject to additional analysis and corrections.

If we put $\epsilon \approx 0$, we come to the “unperturbed” solution in the form of approximate complex exponentials with constant amplitude. As the “unperturbed” kernel (13.3.22) will depend only on the difference $t - \theta$, this case is generally equivalent to that of LTI systems. On the other hand, $\epsilon = 1$ leads to the true solution of (13.3.11). So the solution is first examined using power series of the small parameter ϵ , and then the results are extended to the solution of (13.3.11).

With a change of variables $q = \epsilon t$, $u = t - \theta$ in (13.3.22), and considering (without loss of generality) a single term in the sum in (13.3.23), we can rewrite (13.3.11) as

$$\sigma_i^2 A_{i,m}(q; \epsilon) = \int K(u, q - \epsilon u/2) A_{i,m}(q - \epsilon u; \epsilon) e^{j[\phi_{i,m}(q - \epsilon u) - \phi_{i,m}(q)]/\epsilon} du. \quad (13.3.25)$$

As $\epsilon \rightarrow 0$, we come to

$$\int K(u, q) e^{-ju\dot{\phi}_{i,m}(q)} du = \sigma_i^2. \quad (13.3.26)$$

Alternatively, by defining $\tilde{K}(p, q) = \int K(u, q) e^{-j2\pi up} du$, [20] as the Wigner transform of $K(u, q)$, we can rewrite (13.3.26) as

$$\tilde{K}(\dot{\phi}_{i,m}(q)/(2\pi), q) = \sigma_i^2. \quad (13.3.27)$$

This implies that the instantaneous phase $\phi_{i,m}(q)$ must be chosen to satisfy (13.3.27). Operator $\mathbf{H}\mathbf{H}^*$ is Hermitian by definition, so that $\tilde{K}(p, q)$ is real.

Now after making a Taylor expansion of each term in (13.3.25) around $\epsilon = 0$, taking into account (13.3.26) and considering only terms $\propto \epsilon$, we can write

$$\int \left[\frac{u}{2} \frac{\partial K(u, q)}{\partial q} A_{i,m}(q) + u K(u, q) \frac{dA_{i,m}(q)}{dq} - j A_{i,m}(q) K(u, q) \frac{u^2}{2} \frac{d^2 \phi_{i,m}(q)}{dq^2} \right] e^{-ju\dot{\phi}_{i,m}(q)} du = 0. \quad (13.3.28)$$

This may be understood as the condition to null the term proportional to ϵ in the expansion of (13.3.25). To simplify the notation further, we introduce the functions $\tilde{K}_{i,j}(p, q) = \frac{\partial^{i+j} \tilde{K}(p, q)}{\partial p^i \partial q^j}$. Noting that

$$j\tilde{K}_{1,1} \left(\frac{\dot{\phi}_{i,m}(q)}{2\pi}, q \right) = 2\pi \int u \frac{\partial K(u, q)}{\partial q} e^{-ju\dot{\phi}_{i,m}(q)} du - j2\pi \frac{d^2 \phi_{i,m}(q)}{dq^2} \int u^2 K(u, q) e^{-ju\dot{\phi}_{i,m}(q)} du, \quad (13.3.29)$$

we can restate (13.3.28) as

$$\frac{d}{dq} \left[A_{i,m}^2(q) \tilde{K}_{1,0} \left(\frac{\dot{\phi}_{i,m}(q)}{2\pi}, q \right) \right] = 0, \quad (13.3.30)$$

which shows that the amplitude must be a solution of (13.3.30).

Next, we establish the relationship between the function $\tilde{K}(p, q)$ and the TVTF $H(t, f)$. It would be useful to represent $\tilde{K}(p, q)$ in terms of the TVIR $h(t, \tau)$. Combining (13.3.12), (13.3.22), and $\tilde{K}(p, q)$, we easily see that

$$\tilde{K}(p, q) = \int \int h \left(\frac{u}{2} + \frac{q}{\epsilon}, \frac{u}{2} + \frac{q}{\epsilon} - \tau \right) h^* \left(-\frac{u}{2} + \frac{q}{\epsilon}, -\frac{u}{2} + \frac{q}{\epsilon} - \tau \right) e^{-jup} du d\tau. \quad (13.3.31)$$

Hence, because $h(t, \tau)$ is the inverse Fourier transform ($f \rightarrow \tau$) of $H(t, f)$,

$$\tilde{K}(p, q) = \int \int H \left(\frac{u}{2} + \frac{q}{\epsilon}, v \right) H^* \left(-\frac{u}{2} + \frac{q}{\epsilon}, v \right) e^{-j2\pi(p-v)u} du dv. \quad (13.3.32)$$

To find the equation of the curve $\tilde{K}(p, q) = \text{constant}$, letting $p(q) = \dot{\phi}_{i,m}(q)/(2\pi)$ for simplicity, and assuming $H \left(\pm \frac{u}{2} + \frac{q}{\epsilon}, v \right) \approx H \left(\frac{q}{\epsilon}, v \right)$, $\tilde{K}(p(q), q)$ becomes

$$\tilde{K}(p(q), q) \approx \int |H \left(\frac{q}{\epsilon}, v \right)|^2 \int e^{-j2\pi[p(q)-v]u} du dv = |H \left(\frac{q}{\epsilon}, p(q) \right)|^2. \quad (13.3.33)$$

If we put $\epsilon = 1$, then, instead of Eqs. (13.3.23), (13.3.27), and (13.3.30), we can write Eqs. (13.3.13) to (13.3.15).

13.4 FRACTIONAL AUTOCORRELATION FOR DETECTION IN COMMUNICATIONS⁰

Linear frequency modulated (FM) signals, also called chirps, are commonly used in communications systems as they allow better resilience against interferences. In this section, a method based on fractional autocorrelation is proposed for detection and parameter estimation of such chirps. To that end, after a brief introduction of the fractional Fourier transform (FrFT), formulations of fractional convolution and correlation are presented. Then, a detection statistic based on fractional autocorrelation is proposed and its performance is demonstrated through simulations.

13.4.1 FRACTIONAL FOURIER TRANSFORM (FrFT)

The FrFT is a linear, energy-preserving signal transformation that generalizes the conventional Fourier transform (FT) via an angle parameter ϕ [27]. For $\phi = \frac{\pi}{2}$, the FrFT simplifies to the conventional FT. The conventional FT of a time-domain signal $s(t)$ can be interpreted as a counterclockwise rotation by $\frac{\pi}{2}$ radians in the time-frequency (t,f) plane. The FrFT of $s(t)$ generalizes this rotation property by defining rotations associated with angles other than $\frac{\pi}{2}$. It is defined [27] as

$$\begin{aligned} (\mathbb{F}^\phi s)(r) &= S^\phi(r) \\ &= \begin{cases} \sqrt{1-j\cot\phi} e^{j\pi r^2 \cot\phi} \int s(t) e^{j\pi t^2 \cot\phi} e^{-j2\pi tr \csc\phi} dt, & \phi \neq n\pi, \\ s(r), & \text{for } \phi = 2n\pi; \text{ and } s(-r), & \text{for } \phi = (2n+1)\pi, \end{cases} \end{aligned} \quad (13.4.1)$$

where \mathbb{F}^ϕ is the FrFT operator associated with angle ϕ . $S^\phi(r)$ denotes the fractional Fourier transformed signal and n is an integer. For $\phi = \frac{\pi}{2}$, the conventional FT is obtained. Similarly, for $\phi = 0$, one obtains the time-domain representation $s(t)$ [27]. Furthermore, for $\phi = \pi$, the FrFT reduces to a time reversal operation [27],

$$(\mathbb{F}^\pi s)(t) = s(-t). \quad (13.4.2)$$

One key property of the FrFT relates it to the WD. The WD of the time-domain signal $s(t)$ (i.e., the WVD if $s(t)$ is analytic) is defined (in Section 2.1.3) as

$$W_s(t,f) = \int s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau. \quad (13.4.3)$$

The relationship between the FrFT and the WD can thus be stated [27] as

$$W_{S^\phi}(t,f) = W_s(t \cos\phi - f \sin\phi, t \sin\phi + f \cos\phi), \quad (13.4.4)$$

where $W_{S^\phi}(t,f)$ denotes the WD of the FrFT signal, $S^\phi(r)$, whereas $W_s(t,f)$ is the WD of the original time-domain signal, $s(t)$. According to (13.4.4), the WD of the fractional Fourier transformed signal $S^\phi(r)$ can be obtained by rotating the WD of the original signal $s(t)$ in the clockwise direction by ϕ .

⁰Authors: **O. Akay**, Dokuz Eylül University, Buca/Izmir, Turkey (olcay.akay@deu.edu.tr) and **G.F. Boudreault-Bartels**, University of Rhode Island, Kingston, RI, USA (boud@ele.uri.edu). Reviewers: M.J. Bastiaans and X.-G. Xia.

13.4.2 FRACTIONAL CONVOLUTION AND CORRELATION

Definitions of fractional convolution and correlation operations are obtained using the unitary fractional-shift operator \mathbf{R}_ρ^ϕ , which is defined [28] as

$$(\mathbf{R}_\rho^\phi s)(t) = s(t - \rho \cos \phi) e^{-j2\pi \frac{\rho^2}{2} \cos \phi \sin \phi + j2\pi t \rho \sin \phi}, \quad (13.4.5)$$

where ρ is the fractional lag variable. For $\phi = 0$, the unitary fractional-shift operator, \mathbf{R}_ρ^0 , simplifies to the well-known unitary time-shift operator, $(\mathbf{R}_t^0 s)(t) = (\mathbf{T}_\tau s)(t) = s(t - \tau)$, with τ denoting the time lag. Analogously, as the special case for $\phi = \frac{\pi}{2}$, the unitary frequency-shift operator is obtained, $(\mathbf{R}_{\frac{\pi}{2}}^{\frac{\pi}{2}} s)(t) = (\mathbf{F}_v s)(t) = e^{j2\pi v t} s(t)$, with v denoting the frequency lag.

The fractional convolution of two signals $s(t)$ and $h(t)$ is calculated [29] by evaluating the inner product¹ of the first signal $s(t)$ with the axis-reversed, complex-conjugated, and fractionally shifted version of the second signal $h(t)$,

$$(s *_{\phi} h)(r) = \langle s, \mathbf{R}_r^\phi \mathbb{F}^\pi h^* \rangle = \langle s, \mathbf{R}_r^\phi \tilde{h} \rangle = e^{j2\pi \frac{r^2}{2} \cos \phi \sin \phi} \int s(\beta) h(r \cos \phi - \beta) e^{-j2\pi \beta r \sin \phi} d\beta. \quad (13.4.6)$$

In (13.4.6), $*_{\phi}$ means fractional convolution associated with angle ϕ , $\tilde{h}(t) = h^*(-t)$, and \mathbb{F}^π is the axis-reversal operator in (13.4.2). For $\phi = 0$, the fractional convolution simplifies to linear time invariant (LTI) convolution given as

$$(s *_0 h)(t) = \langle s, \mathbf{T}_t \mathbb{F}^\pi h^* \rangle = \langle s, \mathbf{T}_t \tilde{h} \rangle = \int s(\beta) h(t - \beta) d\beta. \quad (13.4.7)$$

Similarly, for $\phi = \frac{\pi}{2}$, fractional convolution in (13.4.6) reduces to the linear frequency-invariant (LFrI) convolution which is obtained [29] by

$$(s *_{\frac{\pi}{2}} h)(f) = \langle s, \mathbf{F}_f \mathbb{F}^\pi h^* \rangle = \langle s, \mathbf{F}_f \tilde{h} \rangle = \int s(\beta) h(-\beta) e^{-j2\pi f \beta} d\beta = \int S^{\frac{\pi}{2}}(\beta) H^{\frac{\pi}{2}}[-(f - \beta)] d\beta, \quad (13.4.8)$$

where $S^{\frac{\pi}{2}}(f)$ and $H^{\frac{\pi}{2}}(f)$ denote the conventional FTs of $s(t)$ and $h(t)$, respectively.

Analogously, the fractional cross-correlation, denoted by \star_{ϕ} , of $s(t)$ and $h(t)$ is calculated [29] using the fractional-shift operator, \mathbf{R}_ρ^ϕ , within an inner product,

$$(s \star_{\phi} h)(\rho) = \langle s, \mathbf{R}_\rho^\phi h \rangle = e^{j2\pi \frac{\rho^2}{2} \cos \phi \sin \phi} \int s(\beta) h^*(\beta - \rho \cos \phi) e^{-j2\pi \beta \rho \sin \phi} d\beta. \quad (13.4.9)$$

The fractional autocorrelation at angle ϕ is similarly defined [29] by replacing the second function $h(t)$ in (13.4.9) with the function $s(t)$,

$$(s \star_{\phi} s)(\rho) = e^{j2\pi \frac{\rho^2}{2} \cos \phi \sin \phi} \int s(\beta) s^*(\beta - \rho \cos \phi) e^{-j2\pi \beta \rho \sin \phi} d\beta. \quad (13.4.10)$$

For the special cases of $\phi = 0$ and $\phi = \frac{\pi}{2}$, fractional correlation operations in (13.4.9) and (13.4.10) reduce to the corresponding LTI and LFrI correlation operations, respectively. Table 13.4.1 displays the

¹The inner product of two functions $g(t)$ and $h(t)$ is defined as $\langle g, h \rangle = \int g(t) h^*(t) dt$.

Table 13.4.1: Definitions of Fractional Convolution and Correlation at Angle ϕ

| | |
|-------------|--|
| Convolution | $(s *_{\phi} h)(r) = e^{j2\pi \frac{r^2}{2} \cos \phi \sin \phi} \int s(\beta) h(r \cos \phi - \beta) e^{-j2\pi \beta r \sin \phi} d\beta$ |
| Cross-corr. | $(s \star_{\phi} h)(\rho) = e^{j2\pi \frac{\rho^2}{2} \cos \phi \sin \phi} \int s(\beta) h^*(\beta - \rho \cos \phi) e^{-j2\pi \beta \rho \sin \phi} d\beta$ |
| Autocorr. | $(s \star_{\phi} s)(\rho) = e^{j2\pi \frac{\rho^2}{2} \cos \phi \sin \phi} \int s(\beta) s^*(\beta - \rho \cos \phi) e^{-j2\pi \beta \rho \sin \phi} d\beta$ |

Table 13.4.2: Linear Time Invariant (LTI) and Frequency Invariant (LFRI) Convolution and Correlation

| | LTI | LFRI |
|-------------|--|--|
| Convolution | $(s *_0 h)(t) = \int s(\beta) h(t - \beta) d\beta$ | $(s *_{\frac{\pi}{2}} h)(f) = \int S^{\frac{\pi}{2}}(\beta) H^{\frac{\pi}{2}}[-(f - \beta)] d\beta$ |
| Cross-corr. | $(s \star_0 h)(\tau) = \int s(\beta) h^*(\beta - \tau) d\beta$ | $(s \star_{\frac{\pi}{2}} h)(v) = \int S^{\frac{\pi}{2}}(\beta) [H^{\frac{\pi}{2}}(\beta - v)]^* d\beta$ |
| Autocorr. | $(s \star_0 s)(\tau) = \int s(\beta) s^*(\beta - \tau) d\beta$ | $(s \star_{\frac{\pi}{2}} s)(v) = \int S^{\frac{\pi}{2}}(\beta) [S^{\frac{\pi}{2}}(\beta - v)]^* d\beta$ |

definitions of fractional convolution and correlation operations. Definitions of LTI and LFRI convolution and correlation operations are listed in [Table 13.4.2](#).

13.4.2.1 Alternative formulations of fractional convolution and correlation

Alternative formulations for fractional convolution and correlation allow one to gain additional insight about these fractional operations. Furthermore, they suggest efficient ways for computer implementation. An alternative formulation of fractional convolution is [29]

$$(s *_{\phi} h)(r) = \int S^{\phi}(\beta) H^{-\phi}(r - \beta) d\beta = (S^{\phi} *_0 H^{-\phi})(r), \quad (13.4.11)$$

where $S^{\phi}(r) = (\mathbb{F}^{\phi}s)(r)$ and $H^{-\phi}(r) = (\mathbb{F}^{-\phi}h)(r)$. Equation (13.4.11) indicates that the fractional convolution at angle ϕ can equivalently be calculated by a conventional LTI convolution of $S^{\phi}(r)$ and $H^{-\phi}(r)$. Another alternative formulation is obtained by applying the convolution property of the conventional FT [29] to the right-hand side of (13.4.11) and is given by

$$(s *_{\phi} h)(r) = \left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ S^{\frac{\pi}{2}+\phi}(u) H^{\frac{\pi}{2}-\phi}(u) \right\} \right) (r). \quad (13.4.12)$$

According to this formulation, the fractional convolution associated with angle ϕ is computed by multiplying the FrFT at angle $\frac{\pi}{2} + \phi$ of the first signal, $S^{\frac{\pi}{2}+\phi}(u)$, with the FrFT at angle $\frac{\pi}{2} - \phi$ of the second signal, $H^{\frac{\pi}{2}-\phi}(u)$, and then taking a conventional inverse FT. This formulation suggests a way for efficient computer implementation of fractional convolution using the fast Fourier transform (FFT) along with the fast FrFT algorithm [30].

An alternative formula for fractional autocorrelation in (13.4.10) is given as [29]

$$(s \star_{\phi} s)(\rho) = \int S^{\phi}(\beta) [S^{\phi}(\beta - \rho)]^* d\beta = (S^{\phi} \star_0 S^{\phi})(\rho). \quad (13.4.13)$$

Thus, as can be seen from (13.4.13), the fractional autocorrelation can also be calculated via the LTI autocorrelation of $S^\phi(r)$. By applying the autocorrelation property of the conventional FT to the right-hand side of (13.4.13), another alternative formulation of fractional autocorrelation is obtained as

$$(s \star_\phi s)(\rho) = \left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ \left| S^{\frac{\pi}{2} + \phi}(u) \right|^2 \right\} \right) (\rho). \quad (13.4.14)$$

In (13.4.14), one FrFT with angle $\phi + \frac{\pi}{2}$ and one conventional inverse FT are used to compute fractional autocorrelation. This formulation suggests a fast discrete-time approximation of fractional autocorrelation. The discrete FrFT algorithm proposed in Ref. [30] has a computational load of $\mathcal{O}(N \log N)$ for a discrete-time signal of length N . Thus, utilizing the discrete FrFT and FFT algorithms together, a discrete-time approximation of fractional autocorrelation via (13.4.14) can be computed efficiently with a computational load of $\mathcal{O}(2N \log N + N)$ [29]. If one wants to calculate the fractional autocorrelation for M different angles, then the total computational load is $\mathcal{O}[M(2N \log N + N)]$. Taking the FT of both sides of (13.4.14) yields

$$\left(\mathbb{F}^{\frac{\pi}{2}} \left\{ (s \star_\phi s)(\rho) \right\} \right) (u) = |S^{\frac{\pi}{2} + \phi}(u)|^2. \quad (13.4.15)$$

Equation (13.4.15) can be thought as the fractional autocorrelation theorem. For $\phi = 0$, it simplifies to the autocorrelation theorem of the conventional FT [29],

$$\left(\mathbb{F}^{\frac{\pi}{2}} \left\{ (s \star_0 s)(\tau) \right\} \right) (f) = |S^{\frac{\pi}{2}}(f)|^2. \quad (13.4.16)$$

The alternative formulations of fractional cross-correlation are similarly obtained by replacing the second $s(t)$ in (13.4.13) and (13.4.14) with $h(t)$. All alternative formulations of fractional convolution and correlation operations reported in this section are listed in the second and third columns of Table 13.4.3.

Table 13.4.3: Alternative Formulations of Fractional Convolution and Correlation Associated with the Fractional Domain at Angle ϕ

| Fractional | Alternative I | Alternative II | Alternative III |
|---|--|--|--|
| Convolution $(s *_\phi h)(r) =$ | $\int S^\phi(\beta) H^{-\phi}(r - \beta) d\beta$ | $\left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ S^{\frac{\pi}{2} + \phi}(u) H^{\frac{\pi}{2} - \phi}(u) \right\} \right) (r)$ | $AF_{s,h}(r \cos \phi, r \sin \phi)$ |
| Cross-corr. $(s \star_\phi h)(\rho) =$ | $\int S^\phi(\beta) [H^\phi(\beta - \rho)]^* d\beta$ | $\left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ \left[S^{\frac{\pi}{2} + \phi}(u) \right] \left[H^{\frac{\pi}{2} + \phi}(u) \right]^* \right\} \right) (\rho)$ | $AF_{s,h}(\rho \cos \phi, \rho \sin \phi)$ |
| Autocorr. $(s \star_\phi s)(\rho) =$ | $\int S^\phi(\beta) [S^\phi(\beta - \rho)]^* d\beta$ | $\left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ \left S^{\frac{\pi}{2} + \phi}(u) \right ^2 \right\} \right) (\rho)$ | $AF_s(\rho \cos \phi, \rho \sin \phi)$ |

The cross-AF, $AF_{s,h}(\tau, v) = \int s(t + \frac{\tau}{2}) h^*(t - \frac{v}{2}) e^{-j2\pi vt} dt$, simplifies to the auto AF in (13.4.17) when $h(t) = s(t)$. Also, $h(t) = h^*(-t)$.

13.4.3 FRACTIONAL AUTOCORRELATION AND AMBIGUITY FUNCTION

One of the fundamental properties of the fractional autocorrelation is its relation to the ambiguity function (AF) in Eq. (3.2.14). The AF of $s(t)$ can be defined as

$$\text{AF}_s(\tau, \nu) = \langle \mathbf{F}_{-\frac{\nu}{2}} \mathbf{T}_{-\frac{\tau}{2}} s, \mathbf{F}_{\frac{\nu}{2}} \mathbf{T}_{\frac{\tau}{2}} s \rangle = \int s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{-j2\pi \nu t} dt, \quad (13.4.17)$$

where $(\mathbf{T}_\tau s)(t) = s(t - \tau)$ is the unitary time-shift and $(\mathbf{F}_\nu s)(t) = e^{j2\pi \nu t} s(t)$ is the unitary frequency-shift operator. A well-known property of the AF is that one can recover the LTI and LFrl autocorrelations in the fourth row of Table 13.4.2 by taking horizontal ($\phi = 0$) and vertical ($\phi = \frac{\pi}{2}$) cross sections of the AF, respectively:

$$(s \star_0 s)(\tau) = \text{AF}_s(\tau, 0) \quad \text{and} \quad (s \star_{\frac{\pi}{2}} s)(\nu) = \text{AF}_s(0, \nu). \quad (13.4.18)$$

The relationship between the AF and fractional autocorrelation is given [29] as

$$(s \star_\phi s)(\rho) = \text{AF}_s(\rho \cos \phi, \rho \sin \phi). \quad (13.4.19)$$

This relationship generalizes the properties given in (13.4.18) for time ($\phi = 0$) and frequency ($\phi = \frac{\pi}{2}$) domain autocorrelations into other orientations of the AF. One can see from (13.4.19) that fractional autocorrelation at angle ϕ can also be recovered by taking a radial slice of the AF at angle ϕ in the ambiguity plane. Similarly, fractional convolution and fractional cross-correlation can be related to radial slices of the cross-AF as can be seen in the fourth column of Table 13.4.3.

13.4.4 DETECTION AND CHIRP RATE PARAMETER ESTIMATION OF CHIRPS

This section describes a detection statistic for detection of chirp signals and estimation of their chirp rates.

A continuous-time chirp can be formulated as $e^{j2\pi(\hat{\zeta} + \frac{\hat{m}}{2}t)t}$, where \hat{m} and $\hat{\zeta}$ are the unknown chirp rate and initial frequency parameters, respectively. The AF of all chirps has a linear region of support that passes through the origin of the ambiguity plane [31]. Since fractional autocorrelation corresponds to radial slices of the AF as in (13.4.19), a detection statistic based on fractional autocorrelation can be used for detection of chirps. By expressing the integral of the AF magnitude along the radial line with angle ϕ , the following detection statistic is derived [29]:

$$L(m) = \int |(s \star_{\arctan(m)} s)(\rho)| d\rho. \quad (13.4.20)$$

Equation (13.4.20) states that integration of the modulus of fractional autocorrelation with angle $\phi = \arctan(m)$ of the received signal $s(t)$ can be used for detection of chirp signals [29]. If the detection statistic, $L(m)$, exceeds a threshold for a certain chirp rate value, \hat{m} , then the detection statistic determines that a linear FM signal, with that particular chirp rate, is present in the received signal. The fast fractional autocorrelation algorithm via (13.4.14) does not require the calculation of the AF, and hence is computationally efficient. In fact, if the number of chirp rates M , for which the detection statistic is calculated, is much less than the number of signal samples N ($M \ll N$), then the computational saving provided by (13.4.20) over any detection statistic requiring calculation of the AF is considerable.

13.4.4.1 Simulation examples

In this section, we test the detection statistic in (13.4.20) for detection and chirp rate parameter estimation of monocomponent and multicomponent chirp signals.

Example 13.4.1. A discrete-time monocomponent chirp signal with chirp rate $\hat{m} = 0.35$ is corrupted with complex additive noise, $w[k]$, generated by two independent, zero-mean, Gaussian random processes of equal variance:

$$s[k] = e^{j\hat{m}\frac{\pi}{1024}k^2} + w[k], \quad k = 1, 2, \dots, 512. \quad (13.4.21)$$

The normalized² detection statistic is calculated using (13.4.20) for different chirp rate values $m_l = 0.1 + (\frac{0.4}{200})l$, $l = 0, \dots, 199$. The experiment was realized for the noise-free case and for additive complex noise with signal-to-noise ratios (SNRs) of -6 and -9 dB. Figure 13.4.1 shows how the

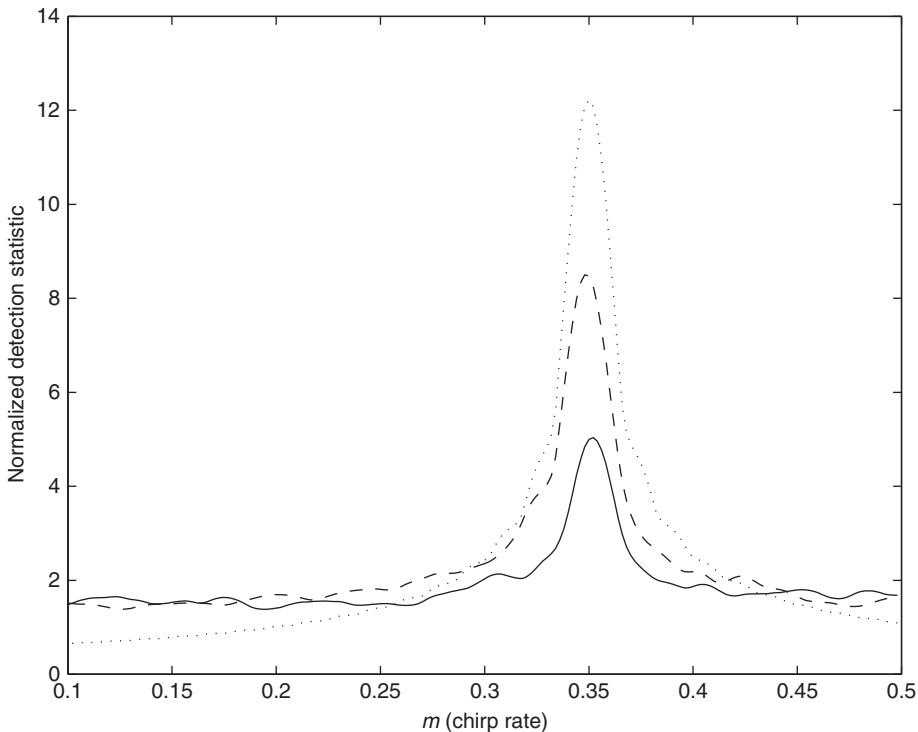


FIGURE 13.4.1

Normalized (see footnote 2) detection statistic; dotted line: noise-free; dashed line: SNR = -6 dB; solid line: SNR = -9 dB.

² We normalize the detection statistic by the area under the received signal magnitude, i.e., $\frac{L(m)}{\sum_{k=1}^N |s[k]|}$, so that the detection threshold level is less affected by the noise level.

normalized detection statistic behaves in different levels of noise. It can be seen in Fig. 13.4.1 that the algorithm is able to detect the monocomponent chirp signal with the correct chirp rate parameter value, $\hat{m} = 0.35$, even in significant levels of noise.

Example 13.4.2. A discrete-time, multicomponent linear FM signal with initial frequency and chirp rate parameter values $\hat{\xi}_0 = 6/1024$, $\hat{\xi}_1 = 3/1024$, $\hat{\xi}_2 = 2/1024$, $\hat{\xi}_3 = 1/1024$, $\hat{m}_0 = 0.10$, $\hat{m}_1 = 0.20$, $\hat{m}_2 = 0.22$, and $\hat{m}_3 = 0.45$ is corrupted with complex additive noise, $w[k]$, generated by two independent, zero-mean, Gaussian random processes of equal variance:

$$s[k] = \sum_{i=0}^3 e^{j[2\pi\hat{\xi}_i + \hat{m}_i \frac{\pi}{4096} k]k} + w[k], \quad k = 1, 2, \dots, 2048. \quad (13.4.22)$$

The normalized detection statistic is calculated via (13.4.20) with respect to different chirp rates $m_l = (\frac{0.5}{200})l$, $l = 0, \dots, 199$. The experiment is done for the noise-free case and for additive complex noise with SNRs of -5 and -10 dB. Figure 13.4.2 shows that all four chirps are detected with their correct chirp rates. Two closely placed chirps with $\hat{m}_1 = 0.20$ and $\hat{m}_2 = 0.22$ are also resolved well.

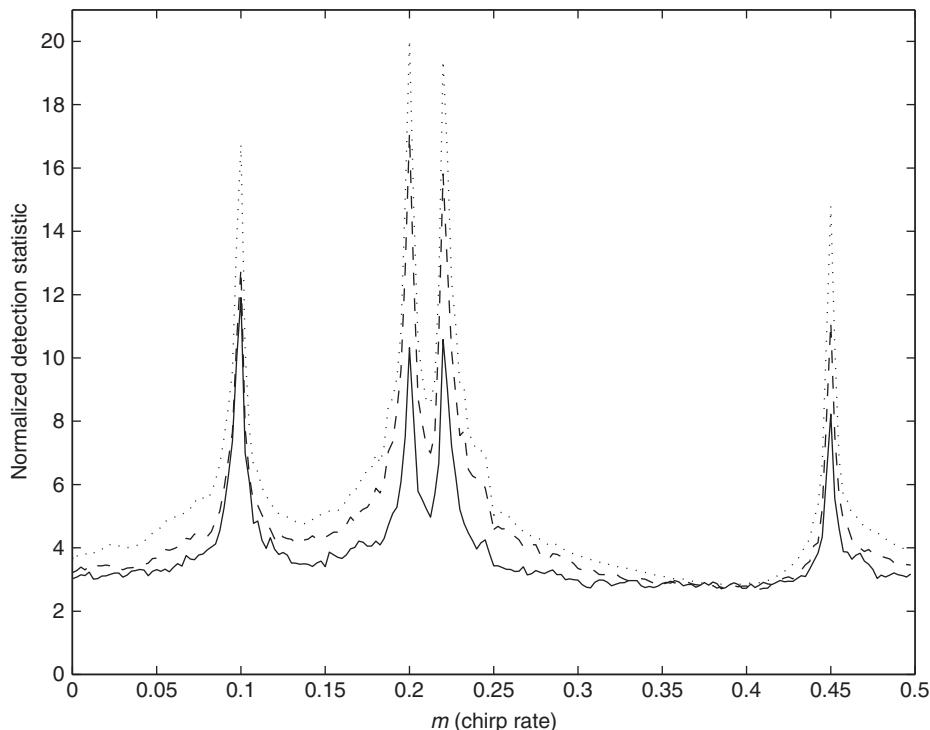


FIGURE 13.4.2

Normalized (see footnote 2) detection statistic; dotted line: noise-free; dashed line: SNR = -5 dB; solid line: SNR = -10 dB.

13.4.5 SUMMARY AND CONCLUSIONS

This section describes a computationally efficient method based on fractional autocorrelation for detection and chirp rate parameter estimation of chirp signals that are utilized frequently in various applications of communications. Formulations of the fractional convolution and correlation operations associated with the fractional domains of the time-frequency plane are also presented. Those fractional domains are defined by the fractional Fourier transform (FrFT). The effectiveness of the proposed method is illustrated with simulation examples. Alternative presentations of the FrFT are given in Sections 4.8 and 5.8.

13.5 TIME-FREQUENCY ESTIMATION OF RADIO-SIGNAL MODULATION PARAMETERS⁰

13.5.1 RADIO-MONITORING PROBLEM

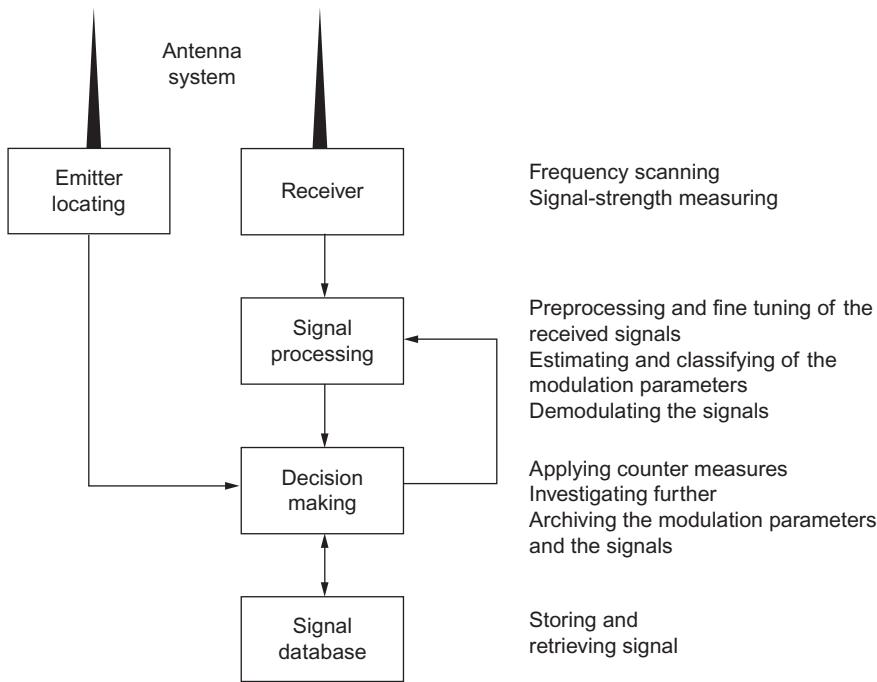
Radio monitoring for military and regulatory organizations involves the scanning of a radio-frequency band, measurement of signal strength, estimation of modulation parameters and classification, geolocating of emitters, and threat identification applicable to both radar and digital communication [32, p. 112]. Figure 13.5.1 shows the block diagram for a typical radio-monitoring system. Radio identification-based sensing in cognitive radio performs the radio-monitoring tasks by identifying the transmission used [33]. Frequency-shift keying (FSK) signals are chosen because they are widely used in digital communication and radar. This section deals with the stage “signal processing” in Fig. 13.5.1 and focuses on classifying FSK signals based on their modulation parameters. Besides digital communication, FSK signals are also used for Costas-coded signaling for pulse compression radar [34, p. 8.1]. In radio monitoring, FSK signal characteristics such as modulation parameters, symbol rate, and frequency deviation are assumed unknown and analysis methods must be applied to estimate them. In this section, a time-frequency (t, f) approach is used for this purpose via instantaneous frequency (IF) estimation. For the other common modulation techniques such as phase-shift keying (PSK) or multiple carrier modulation, the (t, f) approach is still applicable to estimate the frequency content, but the instantaneous information-bearing phase is not represented.

13.5.1.1 Radio communications signal model

An FSK signal is characterized by a frequency transition corresponding to change in the symbol representation. The signal is modeled as a sum of N short-duration complex exponential signals defined as [35]

$$z(t) = A \sum_{k=0}^{N-1} e^{j2\pi f_k(t-kT_b)} \text{rect}\left(\frac{t - (k + \frac{1}{2})T_b}{T_b}\right), \quad (13.5.1)$$

⁰Authors: **A.Z. Sha’ameri**, Universiti Teknologi Malaysia, Johor, Malaysia (zuri@fke.utm.my, ahmad.zuri@gmail.com) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: G. Azemi and J.L. Tan.

**FIGURE 13.5.1**

Block diagram of a radio-monitoring system [32].

where k represents the order of the information-bearing symbol, A is a constant amplitude, f_k is the frequency, T_b is the symbol duration (where “ b ” stands for “bit”) and $\text{rect}(\cdot)$ represents the rectangular or box function. The actual signal parameters are described in Table 13.5.1 and the variation in the frequency f_k depends on the information-bearing symbol (f_k for binary frequency-shift keying (BFSK) can either be equal to f_{b1} or f_{b0} and for 4-frequency-shift keying (4FSK) can be equal to any combination of frequency f_{b0}, f_{b1}, f_{b2} , or f_{b3}).

We assume that received radio-frequency (RF) signals are down-converted to the *intermediate* frequency band and sampled at the Nyquist rate to obtain the discrete-time representation. Unless stated otherwise, “frequency” means *intermediate* frequency of the received signal, normalized to the sampling rate 1 Hz.

Noise affects the IF estimates. The noisy received signal can be modeled as

$$y(t) = z(t) + \epsilon(t), \quad (13.5.2)$$

where $z(t)$ is the signal defined in Eq. (13.5.1) and $\epsilon(t)$ is a complex-valued additive white Gaussian noise (AWGN) process of independent and identically distributed real and imaginary parts with total variance σ_ϵ^2 and zero mean.

Table 13.5.1: Parameters of the Binary and 4FSK Signals (Sampling Rate 1 Hz)

| Signal | Frequency (Hz) Within a Symbol Duration T_b | Symbol Duration (s) |
|--------|---|---------------------|
| BFSK | $f_{b1} = 0.1875$ for symbol “1” $f_{b0} = 0.25$ for symbol “0” | 80 |
| 4FSK | $0.125 \leq f_k \leq 0.3125$, $\delta f = 0.0625$ $f_{b0} = 0.125$ for symbol “00” $f_{b1} = 0.1875$ for symbol “01” $f_{b2} = 0.25$ for symbol “11” $f_{b3} = 0.3125$ for symbol “10” | 80 |

The frequency deviation (minimal difference between permitted frequencies) is 0.0625 Hz in both cases.

13.5.1.2 Relevance of time-frequency analysis

As defined in Eq. (13.5.1), an FSK signal has frequency components that vary with the transmitted symbol. As shown in Part I, a quadratic time-frequency distribution (QTFD) allows better estimation of the true signal characteristics (see Chapter 3). The IF estimated from the peak of the (t,f) representation is important for FSK signals since the information is carried within the frequency.

Two general approaches for classifying digitally modulated signals such as FSK are decision theoretic [36] and pattern recognition [37]. The decision theoretic approach based on detection theory requires prior knowledge of the possible class of signals. The pattern recognition approach requires a signal representation from which the modulation parameters are estimated and used to classify the signals. If the signal is unknown, the signal representation can be used to analyze the signal and estimate its modulation parameters. Time-frequency analysis can provide signal representation in the pattern-recognition approach as shown in Ref. [38]. With efficient estimation of the IF from the (t,f) representation [39], accurate estimation of modulation parameters results in a high classification rate.

An example of (t,f) representation of a nonstationary signal is given in Fig. 2.3.2, p. 79, which shows the spectrogram of a linear FM signal for three different window lengths. Obtaining an accurate (t,f) representation from the spectrogram, the most common TFD, involves a compromise between time resolution and frequency resolution (Section 2.3.2).

Another method, the Wigner-Ville distribution (WVD), defined in Eq. (2.1.17), p. 69, is formed by the Fourier transform (FT) of the instantaneous autocorrelation function (IAF). The cross-term problem in the WVD is described for FSK signals in Ref. [35] and for similar signals, such as frequency-hopping signals, in Ref. [40].

13.5.2 HIGH-RESOLUTION (t,f) DESIGN

The presence of cross-terms in the WVD causes difficulty in identifying the true signal characteristics from the (t,f) representation; the methodology for designing a separable kernel function that preserves the auto-terms while suppressing the cross-terms, presented in Section 5.7, is applied to this problem below.

13.5.2.1 FSK signal representation in (v, τ) domain: Kernel filter selection

From Eq. (2.7.38), the symmetrical ambiguity function of the signal $z(t)$ is

$$A_z(v, \tau) = \int_{-\infty}^{\infty} W_z(t, f) e^{-j2\pi vt} e^{j2\pi f\tau} dt df, \quad (13.5.3)$$

where $W_z(t, f)$ is the WVD of the signal $z(t)$. Figure 13.5.2 shows the ambiguity function for a BFSK signal as defined in Table 13.5.1. The auto-terms are close to the origin, in the Doppler range $-0.02 \leq v \leq 0.02$ Hz and the lag range $-30 \leq \tau \leq 30$ samples (sampling rate 1 Hz). A 2D filtering in the ambiguity domain, based on these ranges, mostly preserves the auto-terms but attenuates the cross-terms. To implement the filter, we define the cutoff frequency in Doppler and the width of the window in lag. It is informative to note that the parameters the 2D filter relates to the BFSK signal parameters; the Doppler range is less than the deviation frequency of 0.0625, while the lag range is close to the symbol duration of 80 samples.

The extended modified B-distribution (EMBD) is the TFD selected, as its parameters can be adjusted for accurate (t, f) representation. The TFD is developed to preserve high (t, f) resolution and attenuate the cross-terms. Compared to other TFDs, the EMBD has shown superior performance over other methods in analyzing real-life biomedical signals. The EMBD [41] has a (t, f) kernel of the *separable* form $\gamma(t, f) = g_1(t) G_2(f)$, whose factors are

$$g_1(t) = \frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} t dt}; \quad G_2(f) = \frac{\cosh^{-2\alpha} f}{\int_{-\infty}^{\infty} \cosh^{-2\alpha} f df}. \quad (13.5.4)$$

The parameters α and β are adjusted to apply the appropriate amount of smoothing in frequency and time, respectively.

13.5.2.2 Time-frequency kernel selection

This section discusses the significance of selecting the right kernel parameters for the EMBD to get the optimal resolution for FSK signals. The selection of the kernel parameters α and β is made by considering the location of the auto-terms and cross-terms in the ambiguity domain of the BFSK signal as shown in Fig. 13.5.2. Figure 13.5.3(a) and (b) shows the kernel function for kernel parameters $(\alpha=0.1, \beta=0.5)$ and $(\alpha=0.5, \beta=0.1)$, respectively. The selection of the kernel function with parameters $(\alpha=0.1, \beta=0.5)$ covers most of the auto-terms that lie along the lag axis ($v=0$) and attenuates the cross-terms of the BFSK signal shown in Fig. 13.5.2. The resulting TFD in Fig. 13.5.3(c) is a precise representation of the true BFSK signal characteristics described in Table 13.5.1. The signal frequencies are correctly represented at 0.1875 and 0.25 Hz, and the duration of the component at 0.1875 Hz is 80 samples. Artifacts are present in the (t, f) representation but are acceptable as they do not disrupt the interpretation of the signal characteristics. The kernel function with parameters $(\alpha=0.5, \beta=0.1)$, shown in Fig. 13.5.3(b), preserves the auto-terms; but in Fig. 13.5.3(d); the wider Doppler range (reduced time smoothing) gives an inner cross-term between the two symbols at 0.25 Hz (time range 150-250 samples). Thus the EMBD provides an accurate (t, f) representation, provided the kernel parameters α and β are optimized for the signal type.

The BFSK and 4FSK signals are different, but both signals have the same symbol duration of 80 samples and frequency deviation of 0.0625 Hz. For the 4FSK signal described in Table 13.5.1, the choice of the kernel parameters $(\alpha=0.1, \beta=0.5)$ made for BFSK does not have an effect on the accuracy

of the (t, f) representation shown in Fig. 13.5.4. Thus, the kernel parameters ($\alpha=0.1$, $\beta=0.5$) when applied to the EMBD give an accurate (t, f) representation for both signals.

13.5.2.3 Comparison with adaptive Wigner-Ville distribution (AWVD)

Since signals in practice are unknown, the kernel parameters have to be adaptively estimated to obtain accurate (t, f) representation. If the kernel parameters are estimated in the time-lag (t, τ) domain, then the AWVD can be defined as [35]

$$\rho_z(t, f) = \int_{-\infty}^{\infty} g_1(t) * K_z(t, \tau) g_2(\tau) e^{-j2\pi f\tau} d\tau, \quad (13.5.5)$$

where $g_1(t)$ is the time window, $g_2(\tau)$ is the lag window, and $K_z(t, \tau)$ is the IAF of the signal $z(t)$

$$K_z(t, \tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right). \quad (13.5.6)$$

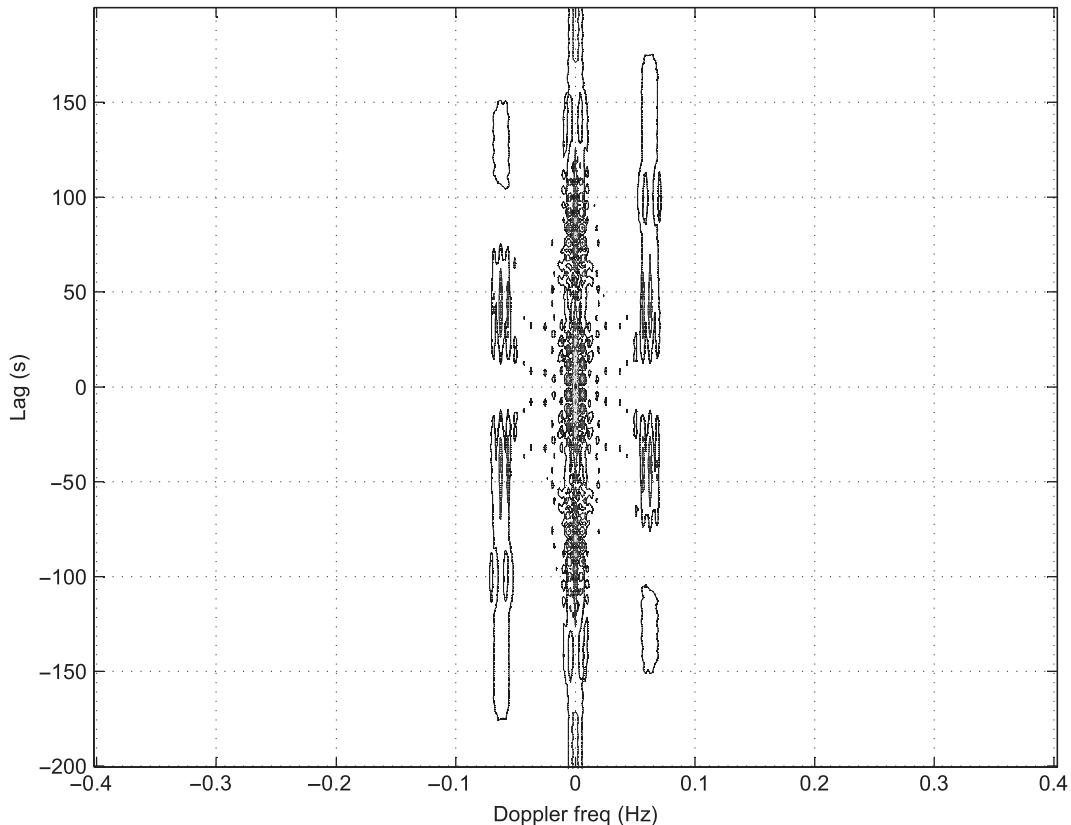
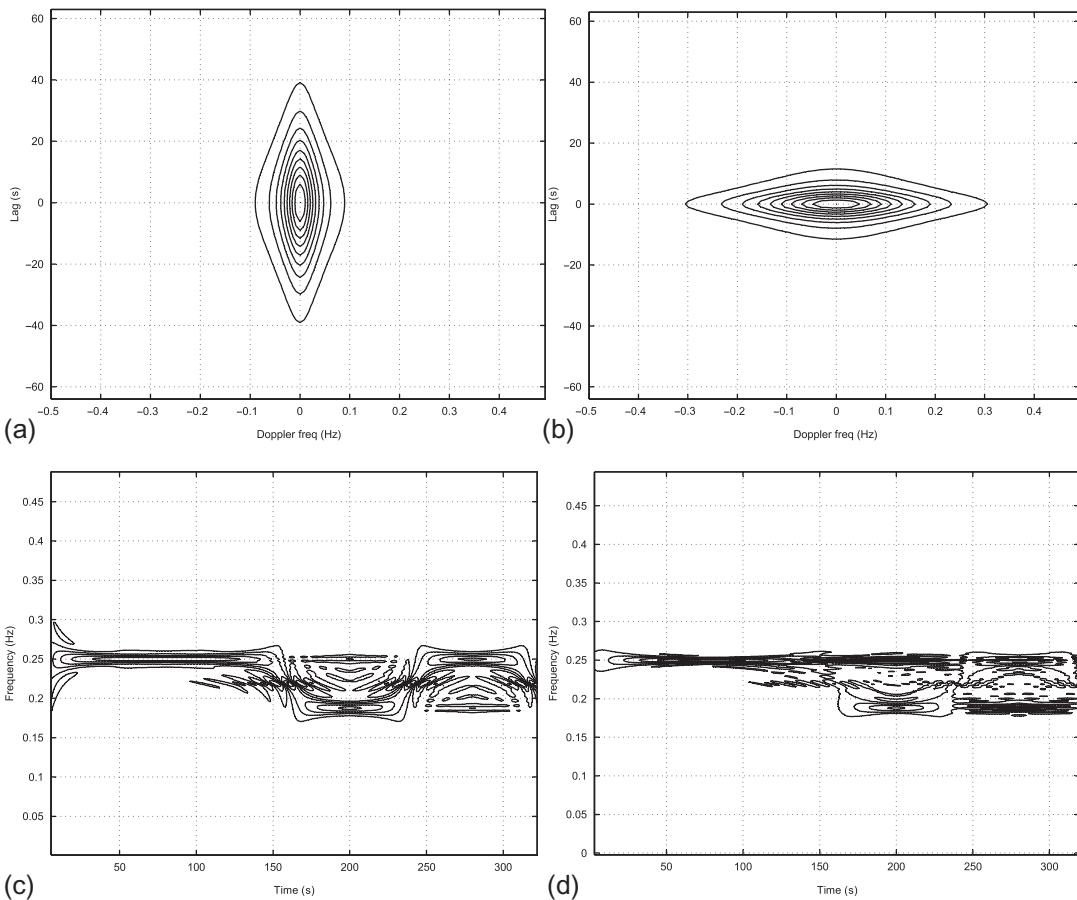


FIGURE 13.5.2

Ambiguity function for the BFSK signal in the Doppler-lag (ν, τ) plane; auto-terms are near the origin; cross-terms are located at $|\nu| > 0.02$ Hz or lag $|\tau| > 30$ samples (sampling rate 1 Hz).

**FIGURE 13.5.3**

Analysis of the BFSK signal in Table 13.5.1 using the EMBD; (a) and (b) Doppler-lag kernels; (c) and (d) resulting (t, f) representations. The choice of α and β determines the accuracy of the (t, f) representation. For $\alpha=0.5$ and $\beta=0.1$, inner cross-terms obscure the frequency transitions: (a) Doppler lag kernel ($\alpha=0.1$, $\beta=0.5$); (b) Doppler lag kernel ($\alpha=0.5$, $\beta=0.1$); (c) EMBD ($\alpha=0.1$, $\beta=0.5$); (d) EMBD ($\alpha=0.5$, $\beta=0.1$).

To reduce cross-terms, both the time window and the lag window are adapted according to the signal characteristics. To achieve an accurate (t, f) representation for FSK signals, the Doppler presentation $G_1(v)$ of the time window, and the lag window $g_2(\tau)$, should cover the Doppler-lag domain as shown in Fig. 13.5.2. Design of the AWVD is influenced by analysis of the IAF, which in this case shows that the auto-terms have zero Doppler (i.e., they are on the time axis) while the cross-terms have nonzero Doppler [35]. The auto-terms are preserved and the cross-terms attenuated by applying a time window at the deviation frequency and lag window at the symbol duration [35]. Adaptive procedures are introduced to determine the parameters of the time window and lag window. Figure 13.5.5 describes the

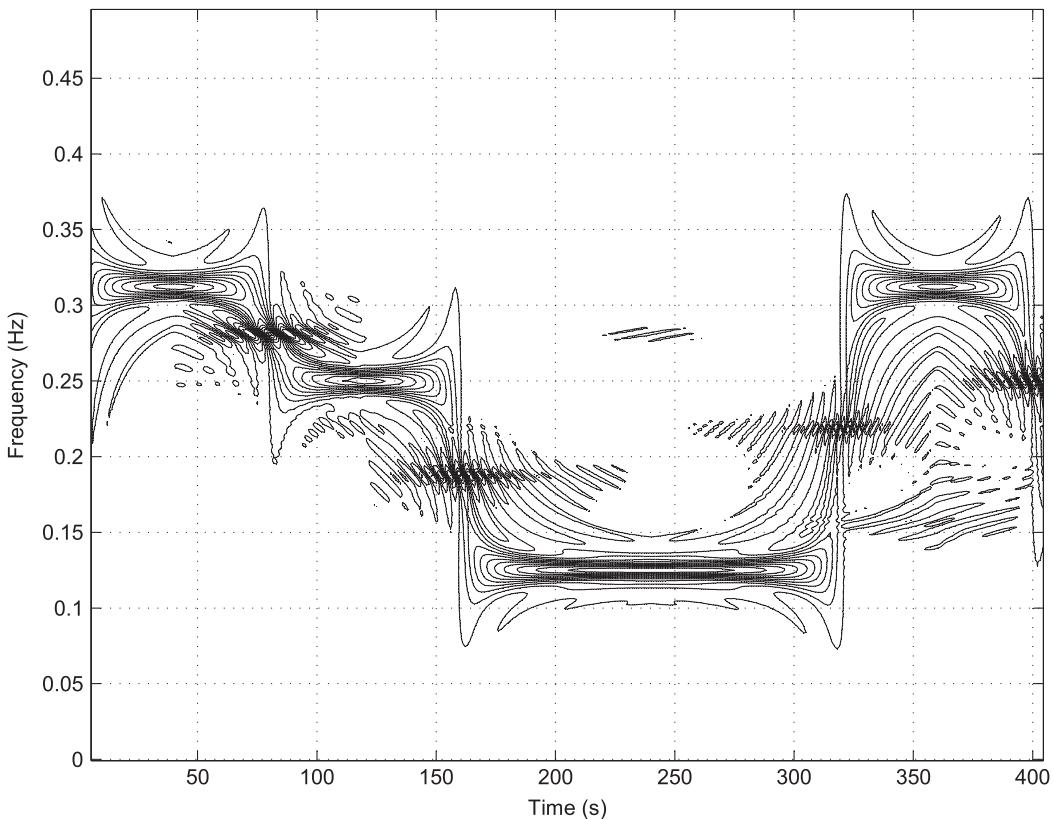


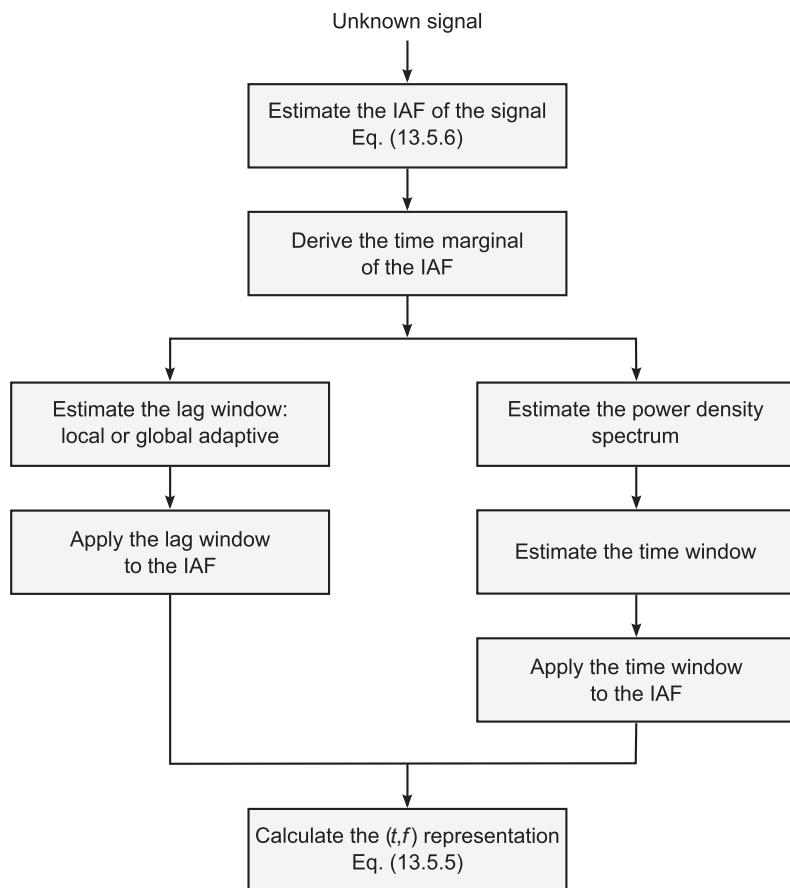
FIGURE 13.5.4

EMBD ($\alpha=0.1$, $\beta=0.5$) of a 4FSK signal. The four possible symbols have frequencies 0.125, 0.1875, 0.25, and 0.3125 Hz. The symbol duration is 80 samples.

procedure to implement the AWVD. The first step is to estimate the IAF and derive the autocorrelation function from the time marginal. From the power spectrum derived from the autocorrelation function, the deviation frequency f_{dev} is estimated from the peak of the power spectrum, which is then used to estimate the time window according to the following relationship $T_1 \geq 3/(2f_{\text{dev}})$ [35].

There are two methods introduced to estimate the lag window width: a global adaptive procedure (by smoothing the autocorrelation function of the signal) [35] and a local adaptive procedure (by applying the local lag correlation function (LLAC) at every time instant over the IAF) [39]. The global adaptive method uses the smoothed modulus of the autocorrelation to estimate the lag window, which is defined as

$$|\mathbb{K}_z(t, \tau)| = h(\tau) * (K_z(\tau)K_z^*(\tau)), \quad (13.5.7)$$

**FIGURE 13.5.5**

Implementation of the AWVD; the two branches are for the lag window (locally or globally adaptive) and the time window. The windows are applied to the IAF and the (t, f) representation is then calculated.

where $h(\tau)$ is the smoothing function. The locally weighted regression scatterplot smoother (LOWESS) smoothing function [35] is shown to be effective in attenuating the out-of-phase terms (see Eq. (13.5.7)). Since the information-bearing sequence carried by the FSK signal is random, the result from Eq. (13.5.7) is a triangular function whose width is approximately the symbol duration, which is used as the lag window width.

The LLAC is used to estimate the lag window for every time instant. The LLAC [39] is a function of time and lag, and it can be defined as

$$R_{KK}(t, \zeta) = \int_{-T}^T w_a(\tau)^2 K_z(t, \tau) K_z(t, \tau - \zeta) d\tau, \quad (13.5.8)$$

where $w_a(\tau)$ is the analysis window, τ is the lag instant, T is the signal duration, and ζ is the lag running variable. The possible range for the normalized LLAC amplitude is

$$0 \leq |R_{KK}(t, \zeta)| \leq 1, \quad (13.5.9)$$

where a value of 1 shows high correlation and vice versa. For a given time instant, the LLAC scans in lag according to the running variable ζ . A drop in the LLAC value of 50% represents a change in the signal characteristics in lag, and the value ζ at this instant is the estimate of the lag window width. The process is then repeated for all time instants. From the estimated time and lag window, the (t, f) representation is calculated from Eq. (13.5.5). In general, both adaptive methods are applicable to FSK signals and the complexity for both adaptive procedures is about the same. The complexity of the adaptive procedure is contributed by the LOWESS smoothing function used in the global adaptive method, while the local adaptive method requires the estimation of the lag window at every time instant. For the cross-TFD (XTFD), only the local adaptive method is applicable, which is due to characteristics of the cross-bilinear product [39].

Figure 13.5.6 shows the (t, f) representation for the BFSK and 4FSK signals using the AWVD. Representations similar to the EMBD ($\alpha=0.1$, $\beta=0.5$) are observed for both signals. For the BFSK signal, frequencies of 0.1875 and 0.25 Hz are represented, and the time interval for the signal component at 0.1875 Hz is correctly represented at 80 samples. For the 4FSK signal, frequencies of 0.125, 0.25, and 0.3125 Hz are represented correctly at the time interval of 80 samples. The main advantage of the AWVD is the adaptive procedure to estimate the kernel parameters (time window and lag window) for an accurate (t, f) representation with minimal artifacts.

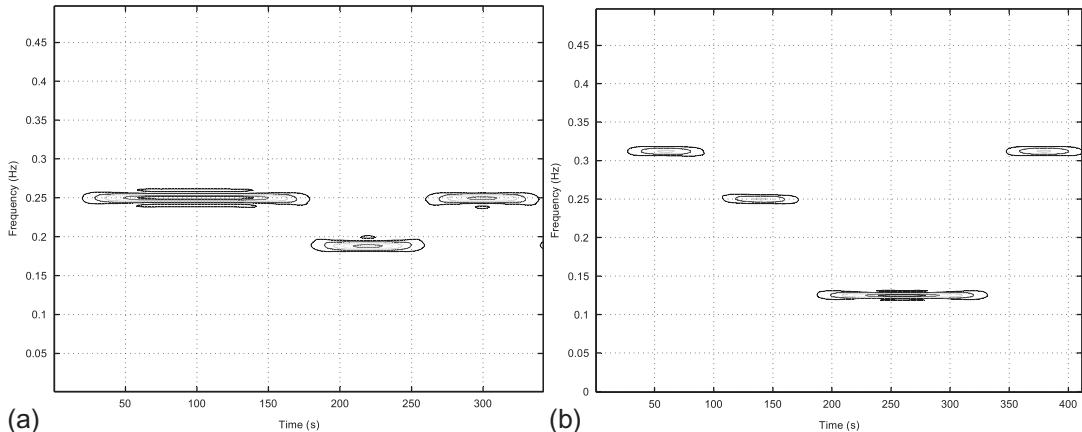


FIGURE 13.5.6

AWVD representations of the BFSK and 4FSK signals; the frequencies and durations of the components are correctly represented according to the signal definitions in Table 13.5.1: (a) BFSK signal; (b) 4FSK signal.

13.5.3 ILLUSTRATION OF TWO EXAMPLES WITH RESULTS AND DISCUSSIONS

Once the TFD is obtained, the IF is estimated as the preliminary step before the signal is further classified according to the actual parameters, such as modulation type, frequencies, and symbol duration. An accurate (t, f) representation ensures an efficient estimation of the IF and a subsequent high signal-classification rate.

13.5.3.1 Stage 1: IF estimation of FSK signals

The IF $f_i(t)$ is estimated from the peak of the TFD as

$$\hat{f}_i(t) = \arg \max_f [\rho_z(t, f)], \quad 0 \leq t \leq T \quad (13.5.10)$$

(cf. Section 10.2). The performance of the IF estimator is benchmarked with the Cramér-Rao lower bound (CRLB), which sets the theoretical lower limit to the variance of any unbiased parameter estimator for a signal in AWGN [39]. Monte Carlo simulations are conducted based on 100 realizations for the signal-to-noise ratio (SNR) range (-3 to 12 dB). The IFs are estimated for BFSK and 4FSK signals from the TFD peaks according to Eq. (13.5.10). The TFDs compared include the spectrogram, EMBD, and AWVD, and their results are shown in Fig. 13.5.7. For the spectrogram, a window length of 64 samples is used to provide the best compromise between time and frequency resolution of the signals described in the section. In general, the EMBD, and AWVD gave the best performance as IF estimators, being close to the CRLB for SNRs above 1 dB. The poor performance of the spectrogram as an IF estimator is due to the smearing in frequency and similar to results reported in Ref. [42].

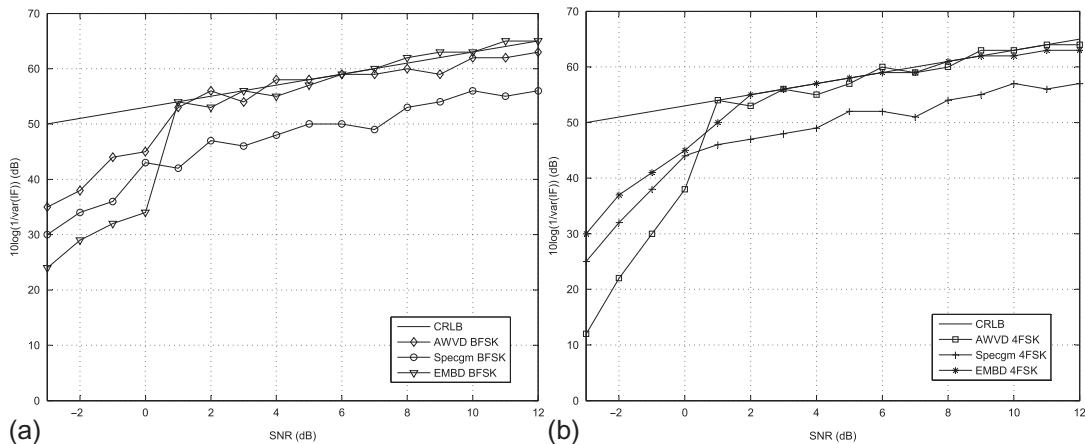
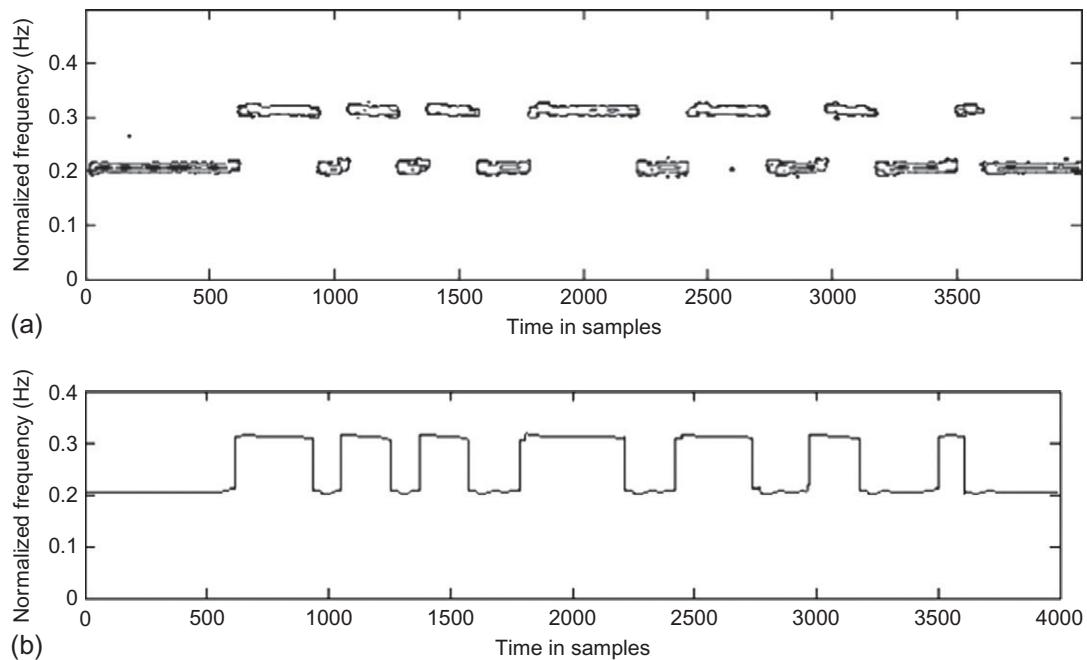


FIGURE 13.5.7

Variance in IF estimation for BFSK and 4FSK signals using the AWVD, spectrogram, and EMBD; the EMBD and AWVD are efficient estimators of IF, meeting the CRLB for SNR above 1 dB: (a) BFSK signals; (b) 4FSK signals.

**FIGURE 13.5.8**

Analysis of captured BFSK signal: (a) accurate (t, f) representation; (b) estimated IF plot, describing true signal characteristics.

13.5.3.2 Stage 2: Analysis and classification of monitored FSK signals

The selected TFD methods developed are validated using real signals monitored in the Straits of Malacca and the South China Sea for maritime communications. The received signal is BFSK-transmitted in the HF band, captured and recorded in October 2012 using the Perseus software-defined radio (SDR). The carrier frequency was at 6.42 MHz with an SNR of 14 dB. Analysis is performed at the baseband after conversion from the HF band. A sampling frequency of 8000 Hz is used to convert the signal to a discrete-time representation for analysis. In general, all the TFDs except the spectrogram show similar results. The IF is estimated from the peak of the (t, f) representation using Eq. (13.5.10). Figure 13.5.8 shows the (t, f) representation and IF using the AWVD. It is observed that the signal has two frequencies, 0.2061 and 0.3125 Hz, with a symbol duration of 110 samples as derived from the minimum time interval in the observed IF. Based on the sampling frequency, the signal frequencies are 1650 and 2500 Hz, with symbol duration 13.8 ms. The analysis performed on the captured signals shows that the AWVD is suitable for use in a noncooperative environment such as spectrum monitoring. The adaptation of the kernel parameters is performed automatically without any prior information on the true signal characteristics, ensuring an accurate (t, f) representation. Any fixed-kernel TFD, such as the EMBD, is not suitable, since the kernel parameters must be first determined from the signal characteristics in the ambiguity domain to obtain an accurate (t, f) representation.

13.5.4 TIME-FREQUENCY ANALYSIS FOR COGNITIVE RADIO APPLICATIONS

The approach known as “cognitive” radio improves radio bandwidth utilization through dynamic bandwidth allocation by allowing secondary users to occupy the frequencies allocated to the primary users when they are not utilized. The process in cognitive radio that identifies the unutilized frequencies or “spectrum holes” allocated to the primary users is spectrum sensing [43]. Spectrum monitoring by the military and regulatory organizations differs from spectrum sensing (by looking for the utilized frequencies). Due to the temporal variation of the unutilized frequencies, time-frequency methods are applicable in spectrum sensing.

Studies on spectrum sensing include the use of QTDFs [43,44], filter banks [43,45,46], wavelet transform, and cyclostationary processing [47]. The spectrogram is used as a spectrum-sensing tool in the European trunked radio system (TETRA) [43]. Unlike energy detection methods, the spectrogram can account for signal nonstationarities and allows basic time and frequency localization information (Section 2.3.2). The WVD provided a joint representation of signals in time and frequency, which was then used to generate an adaptive time-frequency mask to isolate the different spectrum users [44]. Although the WVD gave the best performance, the main issue highlighted is the high computational complexity to implement the method. To optimize the time and energy allocation, simultaneous multiband detection methods are described in Ref. [43]. The wideband sensing-based spectrum-sharing provides a higher throughput compared to the wideband opportunistic spectrum access. A multistage coefficient decimation filter bank [43] has a significant complexity reduction compared to the filter bank. An efficient hardware implementation of the spectrum-sensing method is presented and the TFD is obtained in real-time for a frequency modulation (FM) signal. A channelized wideband receiver architecture is described in Ref. [45] for wideband spectrum-sensing. The RF signal is down-converted to IF, which allows the use of low-end and low cost analog-to-digital converters (ADCs). Filter-bank-based sensing based on uniform discrete Fourier transform (DFT) and modified DFT filter banks with root-Nyquist filters are described in Ref. [46]. The methods are robust to colored noise and the detection process does not require prior knowledge of the received signal. The fast wavelet transform (FWT) is used due to the ability to decompose the sensing process according to a stopping criteria. The selection of the stopping criterion is exploited such that the complexity of using the FWT is lower compared to the fast Fourier transform (FFT). An optimized cyclostationary detection method is described in Ref. [47]. The performance of the algorithm is evaluated by considering the impairments in frequency and sampling clock offsets. The algorithm performs better compared to conventional methods given a constrained sensing time.

Among spectrum-sensing methods, filter-bank-based methods are popular compared to quadratic TDFs, wavelet transform and cyclostationary method. The main reason is simplicity of concept and the reduced computational complexity, since it is important to implement spectrum-sensing at near real-time conditions.

13.5.5 SUMMARY AND CONCLUSIONS

The estimation and classification of signal modulation parameters is a key stage in the emerging area of spectrum-sensing for cognitive radio. It is an important requirement of radio monitoring in applications involving the military, law enforcement and regulatory organizations. Simulation results for the SNR range -3 to 12 dB show that two (t,f) methods, the EMBD and AWVD, are efficient

IF estimators for SNR greater than 1 dB, allowing the direct estimation of the modulation parameters. The verification performed by analyzing and classifying a real monitored HF communication signal shows that an accurate TFD is obtained, and the efficient IF estimation allows precise estimation of the modulation parameters and the subsequent demodulation of FSK signals with a minimal symbol error rate.

TFDs can also be applied to amplitude-shift keying (ASK) signals and biomedical signals such as heart sounds and murmurs. For PSK signals, XTFDs can be considered, since they allow phase recovery [38]. Applications in spread-spectrum communications [40] and low-probability-of-intercept radar [38] highlight the importance of TFDs in the estimation of signal parameters.

REFERENCES

- [1] M.K. Simon, J.K. Omura, R.A. Scholtz, B.K. Levitt, "Spread Spectrum Communications", 3 vols., Computer Science Press, Rockville, MD, 1985.
- [2] M.G. Amin, A.N. Akansu, "Time-frequency for interference excision in spread-spectrum communications", in: G.B. Giannakis (Ed.), Highlights of Signal Processing for Communications *IEEE Signal Process. Mag.* 16 (2), 1999.
- [3] M.G. Amin, "Interference mitigation in spread-spectrum communication systems using time-frequency distributions", *IEEE Trans. Signal Process.* 45 (1) (1997) 90-102.
- [4] L. Boubchir, B. Boashash, "Wavelet denoising based on the MAP estimation using the BKF prior with application to images and EEG signals", *IEEE Trans. Signal Process.* 61 (8) (2013) 1880-1894.
- [5] C. Wang, M.G. Amin, "Performance analysis of instantaneous frequency based interference excision techniques in spread spectrum communications", *IEEE Trans. Signal Process.* 46 (1) (1998) 70-83.
- [6] M.G. Amin, C. Wang, A.R. Lindsey, "Optimum interference excision in spread-spectrum communications using open-loop adaptive filters", *IEEE Trans. Signal Process.* 47 (7) (1999) 1966-1976.
- [7] S. Barbarossa, A. Scaglione, "Adaptive time-varying cancellations of wideband interferences in spread-spectrum communications based on time-frequency distributions", *IEEE Trans. Signal Process.* 47 (4) (1999) 957-965.
- [8] S. Lach, M.G. Amin, A.R. Lindsey, "Broadband nonstationary interference excision in spread-spectrum communications using time-frequency synthesis techniques", *IEEE J. Select. Areas Commun.* 17 (4) (1999) 704-714.
- [9] Y. Zhang, M. Amin, "Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques", *IEEE Trans. Signal Process.* 49 (12) (2001) 3005-3014, ISSN 1053-587X, <http://dx.doi.org/10.1109/78.969509>.
- [10] G.J. Saulnier, M.J. Medley, P.K. Das, "Wavelets and filter banks in spread spectrum communication systems", in: A.N. Akansu, M.J.T. Smith (Eds.), Subband and Wavelet Transforms: Design and Applications, Kluwer, Norwell, MA, 1996, pp. 309-346 (Chapter 10).
- [11] M.G. Amin, L. Zhao, A.R. Lindsey, "Subspace array processing for the suppression of FM jamming in GPS receivers", *IEEE Trans. Aerospace Elect. Syst.* 40 (1) (2004) 80-92.
- [12] M. Villanti, R. Pedone, G.E. Corazza, R. Crescimbeni, "Joint time-frequency domain interference mitigation for Galileo L1 band receivers", in: Proc. Ninth Internat. Symp. on Spread Spectrum Techniques and Applications (ISSSTA-06), 2006, pp. 238-242.
- [13] D. Borio, L. Camoriano, S. Savasta, L. Lo Presti, "Time-frequency excision for GNSS applications", *IEEE Syst. J.* 2 (1) (2008) 27-37.
- [14] P.A. Bello, "Characterization of randomly time-variant linear channels", *IEEE Trans. Commun. Syst.* 11 (4) (1963) 360-393.

- [15] A.M. Sayeed, B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications", *IEEE Trans. Commun.* 47 (1999) 123-132.
- [16] S. Bhashyam, A.M. Sayeed, B. Aazhang, "Time-selective signaling and reception for communication over multipath fading channels", *IEEE Trans. Commun.* 48 (2000) 83-94.
- [17] A.M. Sayeed, A. Sendonaris, B. Aazhang, "Multiuser detection in fast fading multipath environments", *IEEE J. Select. Areas Commun.* 16 (1998) 1691-1701.
- [18] T.A. Kadous, A.M. Sayeed, "Decentralized multiuser detection for time-varying multipath channels", *IEEE Trans. Commun.* 48 (2000) 1840-1852.
- [19] F. Hlawatsch, G. Matz (Eds.), "Wireless Communications over Rapidly Time-Varying Channels", Academic Press, Oxford, UK, 2011.
- [20] B.W. Knight, L. Sirovich, "The eigenfunction problem in higher dimensions: exact results", *Proc. Natl. Acad. Sci. USA* 83 (3) (1986) 527-530.
- [21] R.G. Gallager, "Information Theory and Reliable Communication", Wiley, New York, 1968.
- [22] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: an overview", *IEEE Signal Process. Mag.* 30 (6) (2013) 32-41.
- [23] D. Slepian, H.O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—I", *Bell Syst. Tech. J.* 40 (1) (1961) 43-63.
- [24] H.J. Landau, H.O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—II", *Bell Syst. Tech. J.* 40 (1) (1961) 65-84.
- [25] H.J. Landau, H.O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—III: the dimension of the space of essentially time- and band-limited signals", *Bell Syst. Tech. J.* 41 (4) (1962) 1295-1336.
- [26] S. Barbarossa, A. Scaglione, "Time-varying fading channels", in: G.B. Giannakis, Y. Hua, P. Stoica, L. Tong (Eds.), *Signal Processing Advances in Wireless and Mobile Communications*, vol. 2: Trends in Single- and Multi-User Systems, Prentice-Hall, Upper Saddle River, NJ, 2001 (Chapter 1).
- [27] L.B. Almeida, "The fractional Fourier transform and time-frequency representations", *IEEE Trans. Signal Process.* 42 (11) (1994) 3084-3091.
- [28] O. Akay, G.F. Boudreaux-Bartels, "Unitary and Hermitian fractional operators and their relation to the fractional Fourier transform", *IEEE Signal Process. Lett.* 5 (12) (1998) 312-314.
- [29] O. Akay, G.F. Boudreaux-Bartels, "Fractional convolution and correlation via operator methods and an application to detection of linear FM signals", *IEEE Trans. Signal Process.* 49 (5) (2001) 979-993.
- [30] H.M. Ozaktas, O. Arikan, M.A. Kutay, G. Bozdagi, "Digital computation of the fractional Fourier transform", *IEEE Trans. Signal Process.* 44 (9) (1996) 2141-2150.
- [31] B. Ristic, B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform", *IEEE Trans. Signal Process.* 41 (5) (1993) 1996-2008.
- [32] D.L. Adamy, "EW103: Tactical Battlefield Communications Electronic Warfare", Artech House, Boston, 2009.
- [33] T. Yücek, H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications", *IEEE Commun. Surveys Tutorials* 11 (1) (2009) 116-130.
- [34] M.R. Ducoff, B.W. Tietjen, "Pulse Compression Radar", in: M.I. Skolnik (Ed.), *Radar Handbook*, third ed., McGraw-Hill, New York, 2008 (Chapter 8).
- [35] J.L. Tan, A.Z. Sha'ameri, "Adaptive optimal kernel smooth-windowed Wigner-Ville distribution for digital communication signal", *EURASIP J. Adv. Signal Process.* 2008 (2008) <http://dx.doi.org/10.1155/2008/408341> (article ID 408341, 17 pp.).
- [36] L. Haring, Y. Chen, A. Czylwik, "An automatic modulation classification methods for wireless OFDM systems in TDD mode", *IEEE Trans. Commun.* 58 (9) (2010) 2480-2485.
- [37] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection", *IEEE Circuits Syst. Mag.* 9 (2) (2009) 27-45.

- [38] Y.M. Chee, A.Z. Sha'ameri, B. Boashash, "Efficient phase estimation for the classification of digitally phase modulated signals using the cross-WVD: a performance evaluation and comparison with the S-transform", *EURASIP J. Adv. Signal Process.* 2012 (2012) <http://dx.doi.org/10.1186/1687-6180-2012-65>, article ID 65, 22 pp.
- [39] Y.M. Chee, A.Z. Sha'ameri, M.M.A. Zabidi, "IF estimation of FSK Signals using adaptive smoothed windowed cross Wigner-Ville distribution", *Signal Process.* 100 (2014) 71-84.
- [40] T.-C. Chen, "Joint signal parameter estimation of frequency hopping communications", *IET Commun.* 6 (4) (2012) 381-389.
- [41] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [42] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 2: algorithms and applications", *Proc. IEEE* 80 (4) (1992) 540-568.
- [43] W. Wellisch, A. Barreto, "Spectrum sensing of TETRA systems through time-frequency analysis", *IEEE Latin Amer. Trans. (Rev. IEEE Amer. Lat.)* 11 (3) (2013) 913-919, ISSN 1548-0992, <http://dx.doi.org/10.1109/TLA.2013.6568833>.
- [44] M. Biagi, S. Rinauro, S. Colonnese, G. Scarano, R. Cusani, "WiVCoRA: Wigner-Ville cognitive radio access for secondary nodes", *IEEE Trans. Vehicul. Tech.* 63 (9) (2014) 4248-4264, ISSN 0018-9545, <http://dx.doi.org/10.1109/TVT.2014.2313653>.
- [45] R. Gómez-García, J.P. Magalhães, J.-M. Muñoz-Ferreras, M. Sánchez-Renedo, "Filtering stages for white space cognitive/software-defined radio receivers", in: N.B. Carvalho, A. Cidronali, R. Gómez-García (Eds.), *White Space Communication Technologies*, Cambridge University Press, Cambridge, UK, 2014, pp. 143-147 (Chapter 5).
- [46] K. Maliatsos, A. Adamis, A. Kanatas, "Elaborate analysis and design of filter-bank-based sensing for wideband cognitive radios", *EURASIP J. Adv. Signal Process.* 2014 (82) (2014) <http://dx.doi.org/10.1186/1687-6180-2014-82>.
- [47] E. Rebeiz, P. Urriza, D. Cabric, "Optimizing wideband cyclostationary spectrum sensing under receiver impairments", *IEEE Trans. Signal Process.* 61 (15) (2013) 3931-3943, ISSN 1053-587X, <http://dx.doi.org/10.1109/TSP.2013.2262680>.

This page intentionally left blank

TIME-FREQUENCY METHODS IN RADAR, SONAR, AND ACOUSTICS

14

INTRODUCTION AND OVERVIEW

The fields of radar and sonar are traditionally key application areas and testing grounds for advances in signal processing. Time-frequency (t,f) methodologies have made significant inroads in these fields; their usefulness is demonstrated in seven sections with appropriate internal cross-referencing to this and other chapters.

This chapter begins by considering a baseband Doppler radar return from a helicopter target as an example of a persistent nonstationary signal. A linear (t,f) representation provides a high resolution suitable for preserving the full dynamic range of such complicated signals (Section 14.1). It is then shown that the synthetic aperture principle allows the combination of range resolution, achieved by the use of linear FM signals, with cross-range. For long observation intervals, the phase cannot be assumed to be a linear function of time; then (t,f) based imaging can obtain improvements in focus of the synthetic-aperture image (Section 14.2). In another illustration, it is shown that when a propeller-driven aircraft or a helicopter passes overhead, it produces a Doppler effect which allows the estimation of flight parameters by using IF estimation and the (t,f) interference patterns formed in TFDs (Section 14.3). An example is then presented which shows that to track a theater ballistic missiles launch, the WVD can be used effectively. Its peak provides a direct estimate of the instantaneous Doppler law giving the accelerating target dynamics (Section 14.4). It is then shown that in sonar, there is a clear rationale for using (t,f) processing of returns to provide useful information about targets such as ships (Section 14.5). The last two sections focus on the application of sparse (t,f) distributions to geophysics acoustics (Section 14.6) and a brief tutorial review of (t,f) audio processing for speech and underwater acoustics applications, indicating that high-resolution TFDs can result in much improved performance (Section 14.7).

14.1 TIME-FREQUENCY ANALYSIS OF HELICOPTER DOPPLER RADAR DATA⁰

14.1.1 DYNAMIC RANGE CONSIDERATIONS IN TF ANALYSIS

The classical technique for characterizing the time-evolution behavior of sampled nonstationary signals has been the short-time Fourier transform (STFT), a time-frequency (t,f) analysis technique that uses *linear* operations on sampled data. With momentary nonstationary signals of short duration, one often wishes to improve the (t,f) resolution achievable with the STFT. Time-frequency distributions (TFDs) that use *quadratic* operations on sampled data, such as the Wigner-Ville distribution (WVD), are often able to achieve this improvement. For finite duration data records, this improvement in practice can be shown graphically to be approximately a factor of 2 in the time or frequency sharpening of the TFD response. However, this improvement is achieved in the presence of additional cross-term artifacts and with the sacrifice of detectable dynamic range (DNR). The cross terms (generated by multicomponent signals—more than one signal present at an instant of time), additive noise, and analog-to-digital converter quantization effects often obscure relevant signal components, particularly much weaker signal components (see [Section 4.2](#)). Numerous filtering and smoothing approaches have been developed to mitigate these cross-term artifacts, but this is always achieved at the expense of degrading the (t,f) sharpness of quadratic TFDs. Thus, quadratic TFDs tend to work best in monocomponent signal situations.

The DNR between strongest and weakest signal components in practice can be quite significant. For example, modern radar systems, sonar systems, and medical Doppler ultrasound cardiovascular imaging systems can have DNRs that exceed 70, 50, and 100 dB, respectively. Signals from these systems almost always have multiple components. Thus, quadratic TFDs may not be good candidates for analyzing signals from such systems due to obscuring cross-term artifacts and dynamic range degradation. As this section illustrates, it is still possible to achieve at least a factor of 2 improvement in the TFD response sharpness, without incurring the degradation caused by cross-term artifacts, by exploiting a pair of special two-dimensional (2D) data arrays: the *windowed data function* (WDF) and the *complex WDF transform* (CWT). Application of variants of high-resolution 2D spectral analysis techniques to the CWT can generate the desired high-resolution TFD capability sought in quadratic TFDs but without the cross-term artifacts.

14.1.2 CLASSICAL LINEAR AND QUADRATIC TFDs

The baseline technique for comparing all (t,f) analysis methods is the STFT linear TFD. If $x(t)$ is the signal to be analyzed, define the short-time windowed signal

$$x_h(t, \tau) = x(\tau) h^*(\tau - t), \quad (14.1.1)$$

in which $h(\tau)$ is the analysis window centered about $\tau = 0$, $*$ denotes complex conjugation, and t represents the center time of the analysis interval on $x(t)$. The Fourier transform (FT) of the short-time windowed signal $x_h(t, \tau)$ is therefore

⁰Author: S.L. Marple Jr., Georgia Tech Research Institute, Atlanta, GA, USA (lmarple@orincon.com). Reviewers: J.M. Morris, A. Farina, and E. Tagluk.

$$F_x^h(t, f) = \mathcal{F}_{\tau \rightarrow f} \{x_h(t, \tau)\} = \int_{-\infty}^{\infty} x_h(t, \tau) \exp(-j2\pi f \tau) d\tau, \quad (14.1.2)$$

which is the classical STFT. Its squared magnitude is the localized spectrum or *spectrogram*,

$$S_x^h(t, f) = |F_x^h(t, f)|^2, \quad (14.1.3)$$

which is then plotted as the 2D TFD gram. Trading off (t, f) concentration vs. resolution is achieved by the selection of the analysis window shape and duration.

The classical quadratic TFD method is the Wigner-Ville distribution $W_x(t, f)$ [1]. Most alternatives to the WVD are typically derivatives in which weighting or smoothing kernels are added to the functional definition of the quadratic TFD to better handle finite duration signal records or to suppress cross-term artifacts (see [Chapter 3](#) and [Section 5.7](#)). As explained in [Section 3.2.1](#), the WVD is linked to three companion quadratic functions: the *instantaneous autocorrelation function* (IAF) $K_x(t, \tau)$, the *spectral correlation function* (SCF) $k_x(v, f)$, and the *complex ambiguity function* (CAF) $A_x(v, \tau)$. The WVD can alternatively be computed as (1) the one-dimensional forward transform of the IAF, (2) the one-dimensional inverse transform of the SCF, or (3) the double transform of the CAF, i.e.,

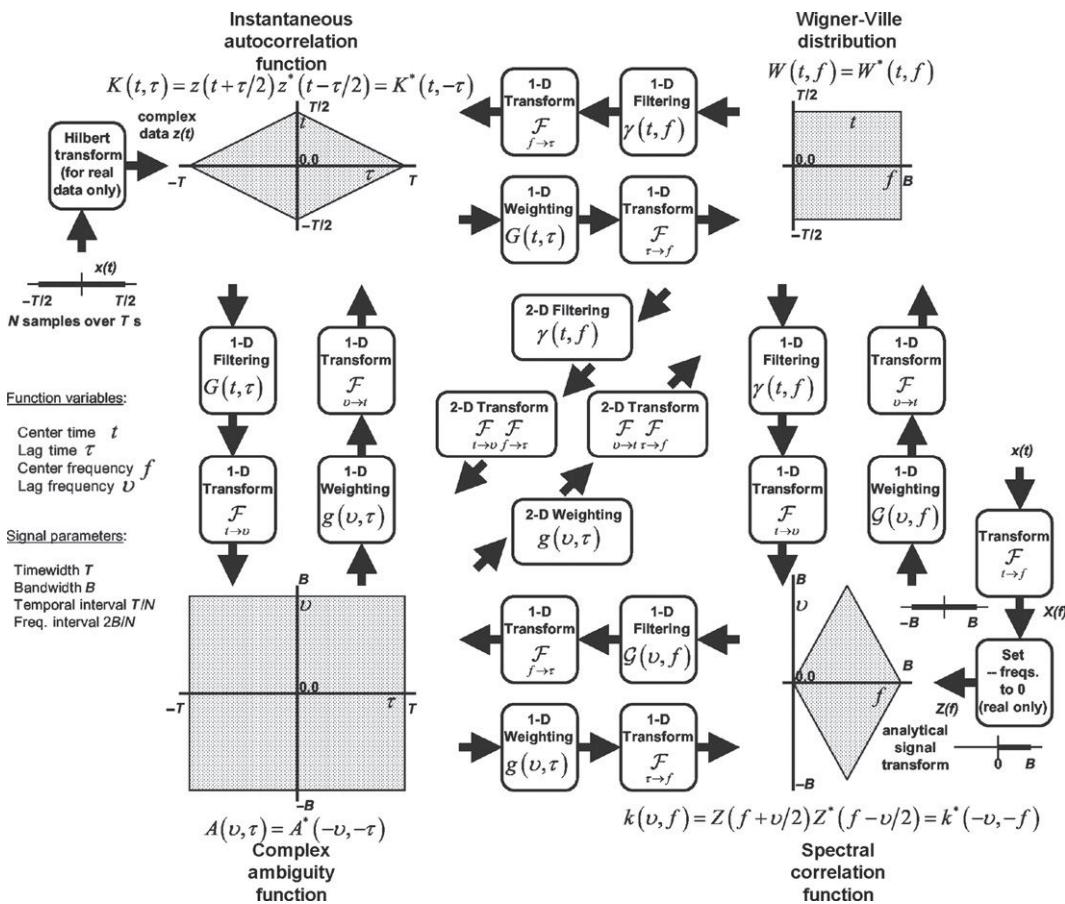
$$W_x(t, f) = \mathcal{F}_{\tau \rightarrow f} \{K_x(t, \tau)\} = \int_{-\infty}^{\infty} x(t + \tau/2) x^*(t - \tau/2) e^{-j2\pi f \tau} d\tau \quad (14.1.4)$$

$$= \mathcal{F}_{v \rightarrow t}^{-1} \{k_x(v, f)\} = \int_{-\infty}^{\infty} X(f + v/2) X^*(f - v/2) e^{+j2\pi vt} dv \quad (14.1.5)$$

$$= \mathcal{F}_{v \rightarrow t}^{-1} \mathcal{F}_{\tau \rightarrow f} \{A_x(v, \tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_x(v, \tau) e^{-j2\pi[f\tau - vt]} dv d\tau, \quad (14.1.6)$$

in which $X(f)$ is the Fourier transform (FT) of $x(t)$, t is the “localized” center time, τ is the relative separation (shift) time displacement, f is the “localized” center frequency, and v is the relative separation (shift) frequency displacement. More details on these relationships may be found in Part I. [Figure 14.1.1](#) depicts the key mathematical definitions and relationships among these four quadratic functions. [Figure 14.1.1](#) also graphically depicts the support regions of the quadratic functions for a finite-duration signal of N samples, $x(n\Delta T) = x[n]$, at ΔT -second intervals. Note that the temporal and spectral correlation functions have diamond-shaped support regions with time-shift and frequency-shift axes, respectively, that are twice the original signal time-width and its transform band-width, due to the correlation processes used to form the IAF and SCF. Having twice the temporal or spectral extents will yield, after transformation, WVD responses that are twice as sharp relative to the original signal’s time extent or bandlimited frequency extent. Note that the temporal and frequency resolutions of the WVD are greatest at $t = 0$ and $f = 0$ and decrease as one advances to the outer time and frequency edges of the IAF’s or SCF’s diamond-shaped region of support.

In order to mitigate the deleterious effects of cross terms, a number of optional smoothing and filtering operations, marked in [Fig. 14.1.1](#) as kernels $g, G, \gamma, \mathcal{G}$, have been developed to suppress the cross-term effects (see [Sections 3.3.4, 5.2, 5.7](#), and [6.1](#) for more details). However, suppressing the cross terms will also degrade the (t, f) concentration; therefore, all kernel designs involve a tradeoff between minimizing cross-term artifacts and maximizing (t, f) resolution.

**FIGURE 14.1.1**

Flow diagram and key relationships and support domains among the four quadratic representations: Wigner-Ville distribution, complex ambiguity function, instantaneous autocorrelation function, and spectral correlation function (see [Section 3.2](#), p. 110, for a detailed discussion).

14.1.3 ALTERNATIVE HIGH-RESOLUTION LINEAR TFD

Let us revisit the 2D STFT and develop another perspective for creating the 2D STFT spectrogram from a 1D data record. Based on this perspective, we will create an alternative 2D matrix array using linear transform operations on data samples from which an opportunity to apply high-resolution 2D spectral analysis procedures is generated. If one assumes a finite-width analysis window of M samples for a case of finite-duration data of N samples ($M < N$) and further assumes the analysis window center time is stepped through the data one sample at a time (multiple sample steps are also possible but not discussed here), then the operations to be performed on the data start with the data matrix in which the

$$\mathbf{x}(t, \tau) = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & x[1] \\ \vdots & \ddots & & \ddots & x[1] & x[2] \\ 0 & & 0 & \ddots & & \vdots \\ \vdots & & \ddots & x[1] & & x[(M-1)/2] \\ 0 & & \ddots & & & \vdots \\ x[1] & & & & & x[N] \\ \vdots & & & & \ddots & 0 \\ x[N-(M-1)/2] & & & x[N] & \ddots & \vdots \\ \vdots & & & 0 & & 0 \\ x[N-1] & x[N] & \ddots & & \ddots & \vdots \\ x[N] & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \quad (14.1.7)$$

analysis center time $t = m\Delta T$ corresponds to the row index m and the data samples within an analysis window interval are all the elements along a selected row, indexed by their column time $\tau = n\Delta T$. Note that the above data matrix has $N + M - 1$ rows corresponding to the center times ranging from $t = -(M+1)/2$ to $N+(M+1)/2$, which requires zero fill when the analysis window runs off the ends of the available data. Other ranges along the vertical axis can also be defined such that zero fill is not required.

Noting the analysis window values as $h[1]$ to $h[M]$, the resultant *windowed data function* (WDF) is obtained element-by-element as $x_h(t, \tau) = x(\tau)h^*(\tau - t)$, to form the WDF matrix $H(t, \tau)$ shown in Fig. 14.1.2. The STFT is then formed by taking the one-dimensional FT of each row of the WDF as shown in Fig. 14.1.2, i.e.,

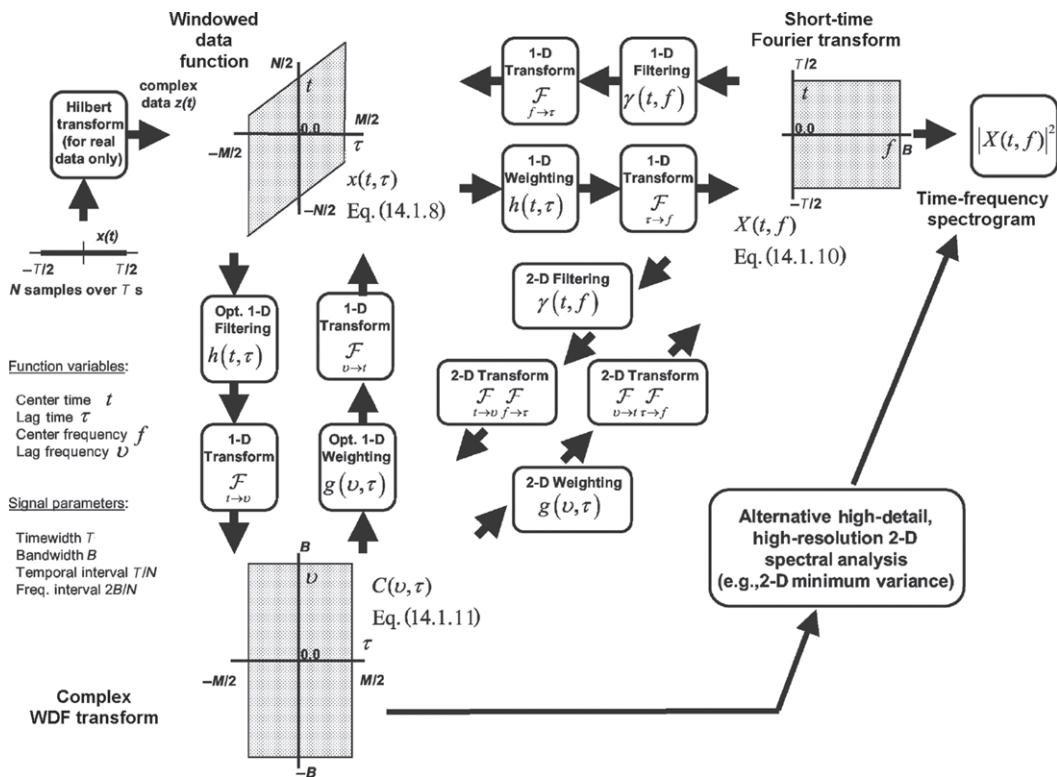
$$F_x^h(t, f) = \mathcal{F}_{\tau \rightarrow f} \{H(t, \tau)\}. \quad (14.1.8)$$

Although counterintuitive, one could instead take the FT of each column (rather than each row) of the WDF to produce the *complex WDF transform* (CWT)

$$C(v, \tau) = \mathcal{F}_{t \rightarrow v} \{H(t, \tau)\} \quad (14.1.9)$$

as shown in Fig. 14.1.2. It may seem computationally disadvantageous to form the CWT, but the motivation is found in the similarity of the CWT to the CAF of Fig. 14.1.1. A double transform of the CWT yields the STFT as shown in Fig. 14.1.2, just as a double transform of the CAF yields the WVD in Fig. 14.1.1. By computing the CWT, we now open the possibility of creating a TFD gram via an alternative high-definition 2D spectral analysis technique in lieu of the spectrogram. Because no cross terms are created in the formation of the CWT, there will be no cross-term artifacts introduced into the 2D spectral analysis results.

Examples of high-definition 2D spectral estimation procedures that may be applied here include the 2D autoregressive [2] and 2D minimum variance techniques [3]. The 2D data set that is used as the input to these methods is the CWT array. The final spectral analysis formula for generating the t -vs- f analysis gram will be a modification of the normal 2D spectral analysis formula in order to account

**FIGURE 14.1.2**

Flow diagram of key relationships among windowed data function (WDF), complex WDF transform (CWT), short-time Fourier transform (STFT), and short-time Fourier transform spectrogram. Insertion of alternative 2D spectral technique possibilities is indicated (see also Section 3.2, p. 110).

for the t -vs- f units of the CWT (rather than time-vs-time). For example, the 2D autoregressive spectral formula is modified to have the structure

$$S_{\text{TFAR}}(t, f) = \frac{1}{\left| \sum_m \sum_n a_{v, \tau}[m, n] \exp(-j2\pi[-m\Delta F t + n\Delta T f]) \right|^2} \quad (14.1.10)$$

in which the two-dimensional autoregressive parameters $a_{v, \tau}[m, n]$ are estimated from the CWT array elements, ΔT is the time-increment along the CWT rows, ΔF is the frequency-increment along the CWT columns, t represents the center time of the analysis evaluation, and f represents the center frequency of the analysis evaluation. Note the opposite signs of the time and frequency variables in the above equation, which contrast with the usual 2D formulae that have the same sign. The modified 2D spectral procedures are more appropriately termed the 2D *time-vs-frequency autoregressive* (TFAR) and the 2D *time-vs-frequency minimum variance* (TFMV) methods for TFD gram creation from the

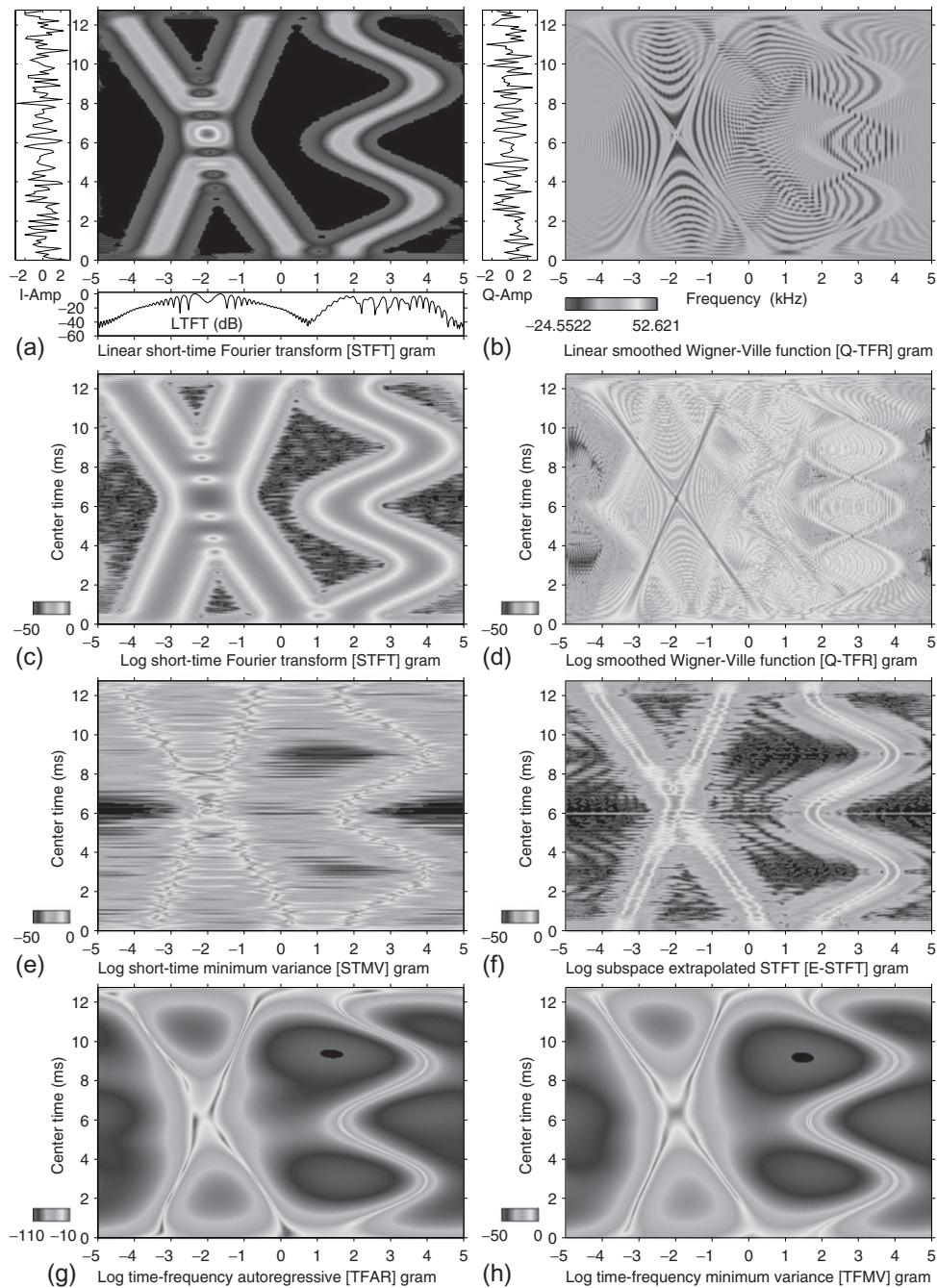
CWT array. Fast computational algorithms for 2D autoregressive parameter estimation may be found in Marple [2,4] and for the 2D minimum variance spectral function in Jakobsson et al. [3].

14.1.4 APPLICATION TO SIMULATED AND ACTUAL DATA

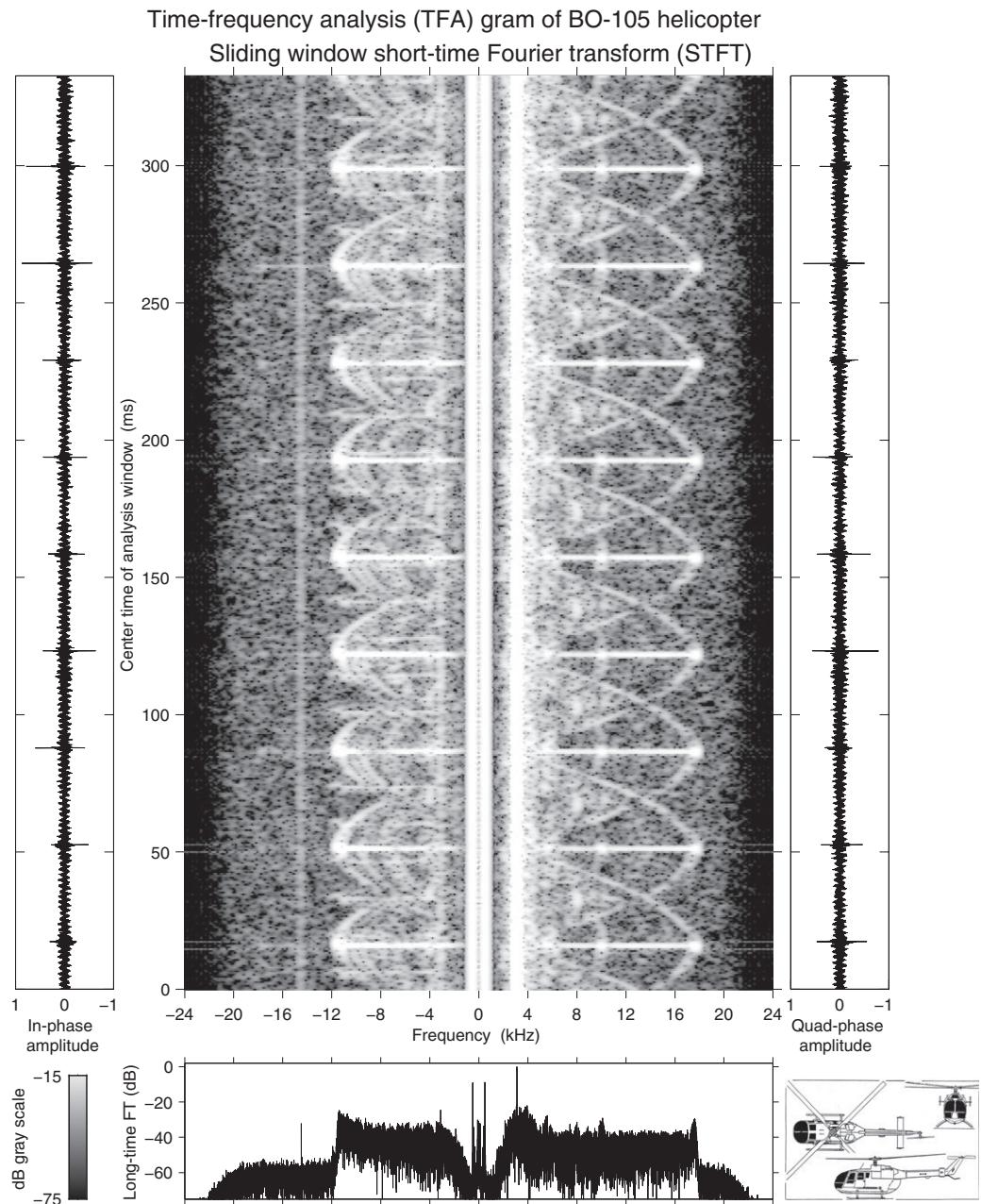
To illustrate the benefit of the 2D high-resolution approach to (t,f) analysis, consider the various TFDs shown in Fig. 14.1.3 for a simulated complex-valued waveform (with in-phase I and quadrature-phase Q components) of 256 samples consisting of two crossing chirp (linear-FM) signals and a sinusoidal-FM signal. The traditional stationary processing approach would be to form the assumed stationary correlation estimate using the entire 256-point data record, followed by Fourier transformation of this estimated correlation to create the long-term Fourier transform (LTFT) spectrum shown at the bottom of Fig. 14.1.3(a), where a logarithmic scale in dB is used to capture the full dynamic range of the spectral response. The I and Q temporal plots of the signal are drawn on the left- and right-hand sides of Fig. 14.1.3(a), respectively. Note that the LTFT spectral shapes of the crossed chirps and the sinusoidal FM component are essentially the same, so that the two signals are almost indistinguishable if the LTFT alone is used for classification. The three signals are distinguishable when the STFT spectrogram is used, as illustrated by Fig. 14.1.3(a) using a color-coded linear plot and in Fig. 14.1.3(c) and a log plot of the top 50 dB. The nonstationary frequency-varying character of the three signal components is now apparent. The Wigner-Ville TFD estimate is plotted with linear units in Fig. 14.1.3(b) and with log units in Fig. 14.1.3(d). The promise of the WVD to increase sharpness in the t -vs- f localization is apparent but comes at the price of introducing significant cross-term artifacts, which are emphasized by the logarithmic gram of Fig. 14.1.3(d). Four alternatives to STFT and WVD characterizations of the nonstationary behavior of the three signals are illustrated in Figs. 14.1.3(e–h). These include the TFAR and TFMV. These last two alternatives produce results as sharp as those obtained with the WVD, without the cross-term artifacts.

Consider the three TFD grams of Figs. 14.1.4–14.1.6 produced from a 330-ms record of baseband Doppler radar returns from a helicopter target. In contrast with the momentary nonstationary signals that characterize much of the (t,f) literature, this is a *persistent* nonstationary signal. Such signals are often called *micro-Doppler* signals [5]. We shall use a sliding analysis window in this case, computing a single line of the TFD with each increment of the window. The helicopter target illuminated by the X-band CW radar is a two-engine Eurocopter Deutschland BO-105, as depicted in the inset of Fig. 14.1.4. The radar signal consists of samples of the complex-valued (I/Q) baseband waveform (carrier frequency demodulated down to 0 Hz); the temporal waveform samples are plotted as vertical displays on the left- and right-hand sides of Figs. 14.1.4–14.1.6. As the helicopter is in motion, the radar signal had Doppler components due to the forward motion of the fuselage, the main rotor and tail rotor rotations, and multipath components between the fuselage and rotors. An LTFT spectral estimate of the entire 330 ms data record produced the spectrum plotted beneath each TFD in Figs. 14.1.4–14.1.6. There is at least 70 dB of dynamic range between the strongest and weakest components in this radar signal.

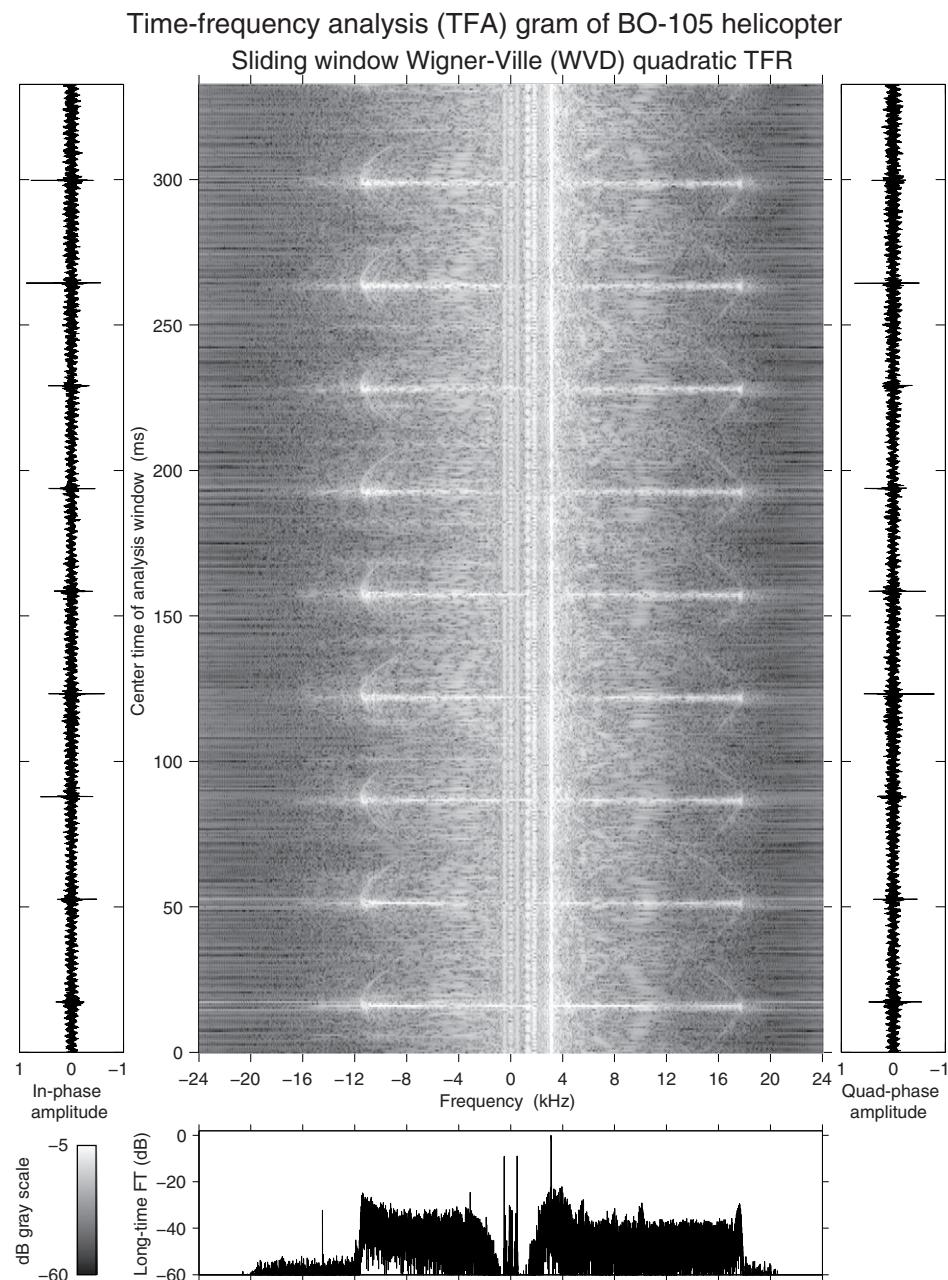
Figure 14.1.4 is the result of a classical STFT analysis of the data record which has been adjusted to use sliding 128-sample Hamming analysis windows to best capture the most rapidly varying Doppler components (main rotor) of the radar signal. Note that positive frequencies represent Doppler components moving toward the radar and negative frequencies represent Doppler components moving away from the radar in this baseband signal. Also note the gray coding assignments of the displayed logarithmic dynamic range; approximately the top 70 dB of each TFR gram is displayed. The dominant

**FIGURE 14.1.3**

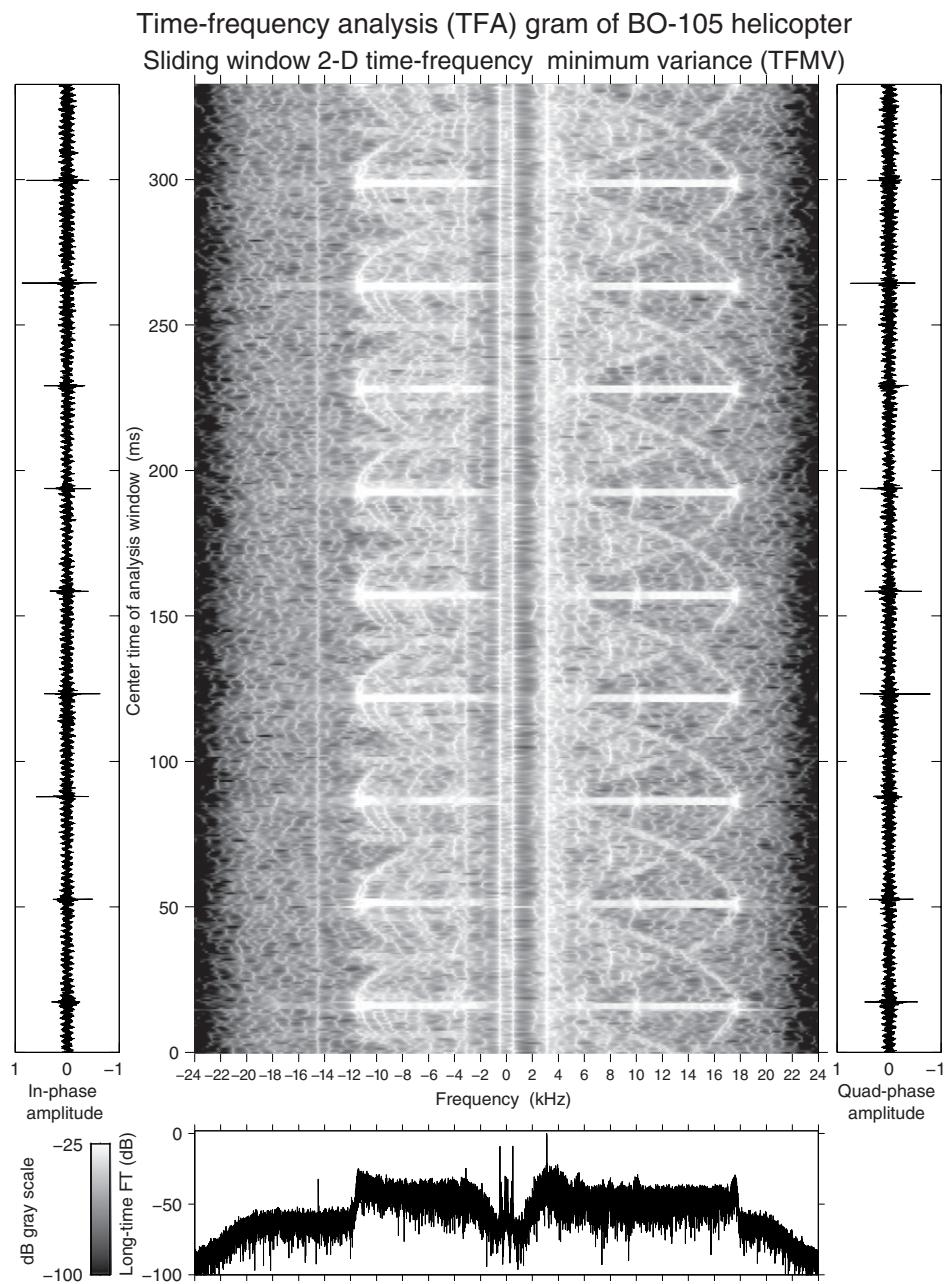
Comparison of various (t, f) analysis techniques to extract the 2D localized nonstationary features of a signal with three components: two criss-crossing chirps and one FM modulated waveform.

**FIGURE 14.1.4**

Classical STFT time-frequency analysis of 330 ms of complex (I/Q) baseband Doppler radar signal of a helicopter target. *Lower right inset:* Profile of BO-105 helicopter target.

**FIGURE 14.1.5**

Time-vs-frequency gram of Doppler radar signal using the Wigner-Ville distribution with sliding window and complex ambiguity function weighting using an exponential kernel.

**FIGURE 14.1.6**

Time-vs-frequency gram of Doppler radar signal using the two-dimensional minimum variance method with sliding window.

signal component is the +2.8 kHz line that represents the Doppler component of the helicopter fuselage due to the constant-velocity motion of the helicopter toward the radar. Other features that can be seen in the STFT gram of Fig. 14.1.4 include the sinusoidal patterns between –12 and +18 kHz due to the Doppler pattern of the four main rotor blades and the periodic broadband horizontal “flashes” corresponding to the periodic alignment of the main or tail rotors to maximally reflect the radar beam. Spectral lines at ± 500 Hz also appear, due to a 500 Hz modulated time tone that was added to the data tape.

Using the same sliding 128-sample analysis window increments as the STFT, the WVD gram shown in Fig. 14.1.5 was produced. An exponential smoothing filter was employed using an exponential parameter factor of 0.3 to mitigate the cross-term artifacts. Although there is an improved sharpness relative to the STFT spectrogram, the cross-term mitigation is insufficient as these terms are still much stronger than the weaker Doppler components that we are attempting to detect and to track. Figure 14.1.6 is the TFD gram resulting from use of the TFMV technique.

14.1.5 SUMMARY AND CONCLUSIONS

The alternative high-resolution methods presented in this section are clearly capable of preserving the full dynamic range of features in complicated signals being analyzed while achieving sharp time-frequency responses comparable to those found in the WVD gram. These attributes are achieved without the artificial introduction of cross-term artifacts typical of some of the standard quadratic TFD methods. The more advanced reduced interference (t,f) methods described in Section 3.3.4 and [6] should be explored for a definitive comparison and conclusion.

14.2 TIME-FREQUENCY MOTION COMPENSATION ALGORITHMS FOR ISAR IMAGING⁰

The synthetic aperture principle is now a well-established concept and has given rise to several applications in high-resolution microwave imaging. Among the most impressive applications is the imaging of Venus’s surface provided by the synthetic aperture radar (SAR) aboard the spacecraft *Magellan* (see, e.g., www.jpl.nasa.gov/magellan). Venus’s surface could not be observed by optical telescopes because of the thick layers of clouds covering the planet. It was thus necessary to use lower frequency waves to penetrate the clouds, but this would have come at the expense of resolution. The radar aboard the *Magellan* orbiter made it possible to send and receive electromagnetic waves through Venus’s atmosphere and still maintain good resolution: the range resolution was achieved by transmitting a wideband impulsive linear FM (chirp) signal, whereas the cross-range resolution was obtained by exploiting the synthetic aperture principle.

The so-called **synthetic aperture** is formed by gathering the echoes collected by the radar platform during the interval in which the radar antenna beam insists on a specific site and processing them coherently. The instantaneous phases of the echoes are compensated as to mimic the presence of

⁰Author: S. Barbarossa, Sapienza University of Rome, Rome, Italy (sergio.barbarossa@uniroma1.it). Reviewers: G. Matz and A.M. Sayeed.

a real antenna gathering the same echoes. If the phase compensation is perfect, we can consider the compensated signal as gathered by a synthetic aperture whose along-track length is equal to the distance traveled by the radar during the time interval within which each point on the ground is illuminated by the radar antenna. Clearly the relative motion between radar and observed scene, or *target*, is fundamental to synthesizing an aperture wider than the real aperture. (In this section, the *target* means the object of the radar imaging.)

The synthesis of the wide aperture requires an accurate knowledge of the temporal evolution of the phase shift induced by the relative radar/target motion. In formulas, if the radar sends an unmodulated sinusoid of frequency f_0 , the echo from a point-like target whose distance from the radar is $r(t)$, assumes the form

$$y(t) = a(t) e^{j2\pi f_0(t-2r(t)/c)} = a(t) e^{j2\pi f_0 t} e^{-j4\pi r(t)/\lambda}, \quad (14.2.1)$$

where c is the speed of light, $\lambda = c/f_0$ is the transmission wavelength, and $a(t)$ is the amplitude modulation.

In the imaging of stationary scenes observed by a radar carried on a spacecraft or an aircraft, the variation of $r(t)$ is entirely due to the radar and an initial coarse estimation of $r(t)$ comes from the navigational system of the vehicle carrying the radar. However, in applications where the observed target is also moving, as in the so-called Inverse SAR (ISAR), the distance $r(t)$ is not known at the radar site. Furthermore, in cases where the target is not moving, the accuracy provided by the navigational instruments may be insufficient, especially for high-resolution imaging. In all these cases, it is necessary to estimate the phase history $\phi(t) := 4\pi r(t)/\lambda$ directly from the received data [7]. This requires the presence of prominent scatterers on the target, whose radar echoes are sufficiently stronger than the background to allow a reliable estimate. The situation is complicated, however, when more dominant scatterers occupy the same range cell, as is likely in many practical circumstances. In fact, if no *a priori* knowledge of the relative motion is available, it is safer not to use a high range resolution at the beginning, simply to avoid the range migration problem. But the more we decrease the range resolution, the higher is the probability of observing more dominant scatterers in the same range cell. In these cases, as proposed in Ref. [8], an iterative procedure can be followed, based on the following steps: (i) start using low range resolution data, for example by smoothing along the range direction, to avoid the range resolution problem; (ii) estimate the instantaneous phase of the dominant scatterer; (iii) estimate the relative motion law $r(t)$ of the dominant scatterer from the instantaneous phase; (iv) use the estimated motion law to compensate for the range migration occurring in the high range resolution data (hence the dominant scatterer automatically becomes the image center); (v) analyze the full range resolution data after compensation of the range migration.

Clearly step (i) increases the probability of having more dominant scatterers in the same range cell and, if the observed signal is really multicomponent, the analysis becomes more complicated because we must separate the components corresponding to each echo before estimating their instantaneous phases. The separation is simple if the components have linear phase. In such a case, in fact, an FFT-based approach may be sufficient. However, especially when long observation intervals are used to obtain high cross-range resolution, the phase cannot be assumed to be a linear function of time. Time-frequency representations of the observed signals thus provide an important analysis tool in such cases. In this section, we illustrate a method for separating the signal components and estimating their instantaneous phases using the so-called smoothed pseudo Wigner-Ville distribution with reassignment

(RSPWVD), described in Section 7.2 and originally proposed in Ref. [9], followed by a parametric estimation method. Intuitively speaking, the time-frequency analysis is used to get an initial idea about the kind of received signals. Based on this preliminary analysis, we can assume a parametric model and then use parametric estimation methods to improve the estimation accuracy. One possible method for combining TFDs and parametric modeling for retrieving the parameters of multicomponent signals, with instantaneous frequencies (IFs) modeled as sinusoids or hyperbolas, was suggested in Ref. [10]; but clearly many alternatives can be followed to optimize the performance of the estimation method.

14.2.1 ECHO FROM A ROTATING RIGID BODY

We assume that the target is a rigid body in the far field of the radar antenna and is characterized by a certain number of dominant scatterers. For a transmission frequency $f_o = c/\lambda$, the echo from the k th scatterer is

$$A_k e^{j2\pi f_o(t-2|\mathbf{r}_o - \mathbf{r}'_k(t)|/c)}, \quad k = 0, \dots, K-1, \quad (14.2.2)$$

where \mathbf{r}_o is the vector indicating the radar position, $\mathbf{r}'_k(t)$ indicates the k th scatterer, and K is the number of scatterers. Under the far field hypothesis, the echo can be approximated as

$$A_k e^{j2\pi f_o t} e^{-j4\pi R_o/\lambda} e^{j4\pi \hat{\mathbf{r}}_o \cdot \mathbf{r}'_k(t)/\lambda}, \quad (14.2.3)$$

where $R_o = |\mathbf{r}_o|$. The motion of a rigid body can always be expressed as the composition of the translation of one of its points plus the rotation of the body around that point. Every imaging or classification procedure must apply some kind of motion compensation and, in general, the translational motion is compensated first. This operation is performed by multiplying the radar echo by a reference signal matched to the echo from one dominant scatterer, and resampling the data in range to remove any range migration of the scatterer assumed as a reference. Taking as a reference the echo from the 0th scatterer in Eq. (14.2.3) (setting $k = 0$ and $A_0 = 1$), the signals after compensation take the form

$$A_k e^{j4\pi \hat{\mathbf{r}}_o \cdot \mathbf{q}_k(t)/\lambda}, \quad k = 1, \dots, K-1, \quad (14.2.4)$$

where $\mathbf{q}_k(t) := \mathbf{r}'_k(t) - \mathbf{r}'_0(t)$. Under the rigid-body constraint, the vectors $\mathbf{q}_k(t)$ can only rotate and the rotation matrix is the same for all points belonging to the target. The differential equation characterizing the rotation of the generic vector $\mathbf{q}_k(t)$ is

$$\frac{d\mathbf{q}_k(t)}{dt} = \boldsymbol{\omega}(t) \times \mathbf{q}_k(t), \quad (14.2.5)$$

where $\boldsymbol{\omega}(t)$ is the vector containing the instantaneous pitch, roll and yaw pulsations ($\omega_p(t), \omega_r(t), \omega_y(t)$), and \times denotes the vector (cross) product. Assuming a constant pulsation, i.e., $\boldsymbol{\omega}(t) = \boldsymbol{\omega}$, and indicating by Ω its modulus, i.e., $\Omega = \sqrt{\omega_p^2 + \omega_r^2 + \omega_y^2}$, and given an initial position $\mathbf{q}_k(0) = \mathbf{q}_k$ of the vector at time $t_0 = 0$, the solution of Eq. (14.2.5) is

$$\mathbf{q}_k(t) = \mathbf{a}_k + \mathbf{b}_k \cos(\Omega t) + \frac{\mathbf{c}_k}{\Omega} \sin(\Omega t), \quad (14.2.6)$$

where the vectors \mathbf{a}_k , \mathbf{b}_k , \mathbf{c}_k are

$$\mathbf{a}_k = \frac{(\boldsymbol{\omega} \cdot \mathbf{q}_k)}{\Omega^2} \boldsymbol{\omega}, \quad \mathbf{b}_k = \mathbf{q}_k - \frac{(\boldsymbol{\omega} \cdot \mathbf{q}_k)}{\Omega^2} \boldsymbol{\omega}, \quad \mathbf{c}_k = \boldsymbol{\omega} \times \mathbf{q}_k. \quad (14.2.7)$$

Hence the echo from the generic k th scatterer is

$$A_k e^{j4\pi \hat{\mathbf{r}}_o \cdot \mathbf{q}_k(t)/\lambda} = A_k e^{j4\pi(m_k \cos(\Omega t + \phi_k) + \alpha_k)/\lambda}, \quad (14.2.8)$$

where

$$m_k = \sqrt{(\hat{\mathbf{r}}_o \cdot \mathbf{b}_k)^2 + \frac{(\hat{\mathbf{r}}_o \cdot \mathbf{c}_k)^2}{\Omega^2}}, \quad \phi_k = \arctan\left(\frac{\hat{\mathbf{r}}_o \cdot \mathbf{c}_k}{\Omega \hat{\mathbf{r}}_o \cdot \mathbf{b}_k}\right), \quad \alpha_k = \hat{\mathbf{r}}_o \cdot \mathbf{a}_k. \quad (14.2.9)$$

It is important to notice that the instantaneous phase of each echo contains a constant term plus a sinusoidal contribution having the *same frequency* for all the scatterers but different amplitudes and initial phases. This is a consequence of the rigid-body constraint.

14.2.2 SIGNAL ANALYSIS BASED ON TIME-FREQUENCY REPRESENTATIONS

The general model for the frequency modulation induced by the relative radar-target motion can always be decomposed into the sum of a slow component, well approximated by a low-order polynomial, plus a possible fast component having a sinusoidal behavior. The echo can then be modeled as

$$x(t) = A e^{j \sum_{m=0}^M a_m t^m} e^{j \beta \cos(\Omega t + \Phi)} + w(t), \quad (14.2.10)$$

where $w(t)$ is additive noise. The slow component is mainly due to the translation; the fast component depends on the rotation. If the sinusoidal component has a period ($T = 2\pi/\Omega$) much longer than the duration (T_0) of the observation interval, it can also be approximated as a low-order polynomial, so that the overall signal can be modeled as a polynomial-phase signal (PPS). Conversely, if T is much less than T_0 , it is better to estimate the parameters of the sinusoidal component without any polynomial modeling.

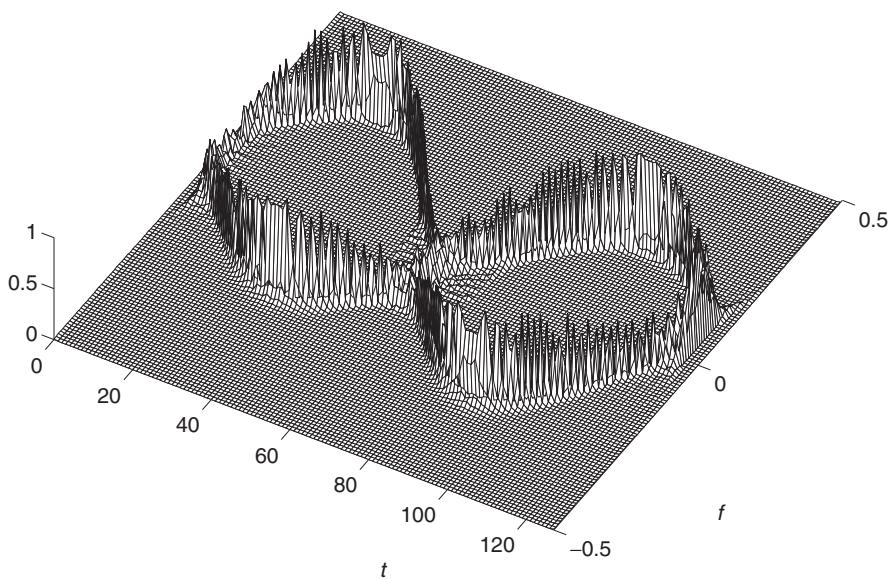
As the velocity of variation is not known *a priori*, it has to be estimated from the data using, at least initially, a nonparametric approach. We used, as a preliminary tool, the reassigned smoothed pseudo Wigner-Ville distribution (RSPWVD) for its good localization and low cross-terms (see [Section 7.2](#)).

As an example, in [Fig. 14.2.1](#) we report the RSPWVD of the echo from a rotating object with two dominant scatterers having the same backscattering coefficients. The two components are clearly visible and show the same period.

We can use the RSPWVD to extract an initial estimate of the period Ω and then exploit the parametric modeling, as shown in the following section. One possibility is to extract the peaks of the TFD, as in Ref. [11] and [Sections 10.3](#) and [10.4](#), and then estimate the modulating frequency of the IF.

14.2.3 PARAMETRIC ESTIMATION OF INSTANTANEOUS PHASES

With the aid of the echo modeling illustrated in [Section 14.2.1](#), we can improve the performance of the estimation method over that obtainable from the simple use of the RSPWVD. In principle, considering signals expressed as in Eq. (14.2.8), we could smooth the WVD by integrating it over all possible sinusoids in the time-frequency plane, as a function of the sinusoidal parameters, using, for example, a

**FIGURE 14.2.1**

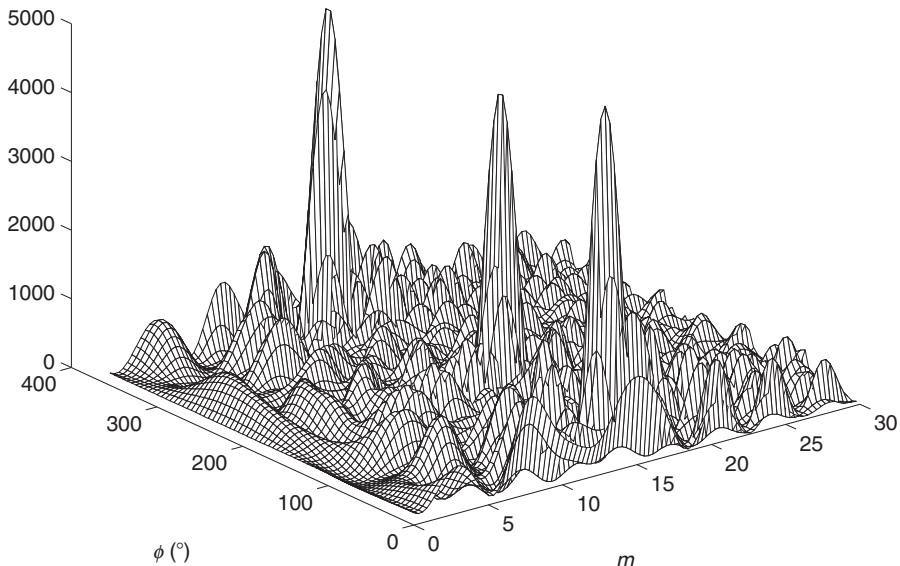
Reassigned smoothed pseudo-Wigner-Ville distribution of the sum of two echoes from a rotating object.

generalized Hough transform as in Ref. [10]. However this operation would be quite troublesome from the computational point of view. In fact such an operation is almost equivalent to computing the square modulus of the scalar product of the received signal and the signal model (14.2.8), with $A_k = 1$. Hence the detection and estimation of FM signals satisfying Eq. (14.2.8) can be carried out by searching for the peaks of the function

$$P(m, \Omega, \phi) = \left| \sum_{n=0}^{N-1} x(n) e^{-j4\pi m \cos(\Omega n + \phi)/\lambda} \right|^2, \quad (14.2.11)$$

where N is the number of samples: if a peak exceeds a suitable threshold, a sinusoidal FM signal is detected and its parameters are estimated as the coordinates of the peak. Of course the straightforward application of Eq. (14.2.11) is also quite troublesome from the computational point of view. Observing, however, that all signal components have the same pulsation Ω , we can estimate Ω from the RSPWVD and then use the estimated value $\hat{\Omega}$ in Eq. (14.2.11) to compute the 2D function

$$P(m, \phi) = \left| \sum_{n=0}^{N-1} x(n) e^{-j4\pi m \cos(\hat{\Omega} n + \phi)/\lambda} \right|^2. \quad (14.2.12)$$

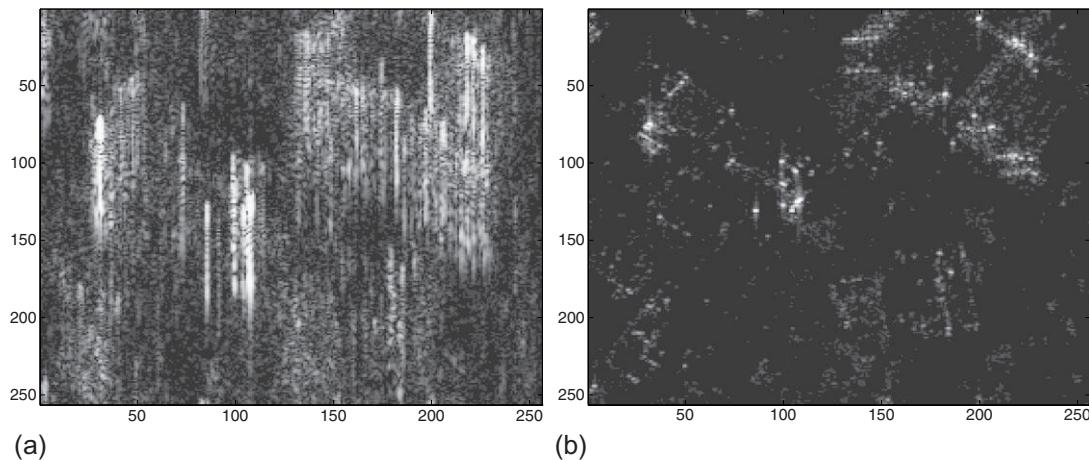
**FIGURE 14.2.2**

$P(m, \phi)$ of a three-component sinusoidal FM signal.

An example is shown in Fig. 14.2.2 for the case where three sinusoidal FM signals occupy the same range cell. Figure 14.2.2 shows the function $P(m, \phi)$ given in Eq. (14.2.12). We can clearly observe the presence of three peaks. Comparing $P(m, \phi)$ with a suitable threshold, we detect the presence of dominant scatterers and estimate, at the same time, their modulation parameters m_k and ϕ_k as the coordinates of the peaks which exceed the threshold. A further simplification of the proposed procedure consists in analyzing time intervals smaller than the rotation period. In these subintervals, the instantaneous phases can be approximated by polynomials (the first terms of their Taylor series expansions). In this case we can use specific algorithms devised for the detection and parameter estimation of multicomponent polynomial phase signals embedded in noise, based on the high order ambiguity function [12].

An example of application to real SAR images is given in Fig. 14.2.3, where we show a SAR image defocused by a third-order polynomial (left side) and the image focused by estimating the polynomial parameters and then compensating the instantaneous phase (right side).

The time-frequency analysis coupled with the parametric method is especially important when the illuminated spot contains moving targets. In such a case, the conventional SAR processing would provide a focused image of the stationary background and a defocused image of the moving object. Indeed, the only way to discriminate moving targets echoes from the background is to analyze their instantaneous frequencies: a different motion law must give rise to a different frequency modulation. Some examples of applications of this idea were shown in Ref. [13].

**FIGURE 14.2.3**

Example of imaging: (a) image defocused with a third-order polynomial instantaneous phase; (b) focused image.

14.2.4 SUMMARY AND CONCLUSIONS

Estimating the frequency modulations of the radar echoes is a fundamental step in the formation of high-resolution images. The tool is especially important in the presence of targets moving on a stationary background, where the only possibility for discriminating moving from fixed targets involves the analysis of their instantaneous frequencies [14]. As the observed signal often comes from multiple scatterers, the availability of TFDs with good localization and low cross-terms is particularly important [15]. The other key idea is that, given a time-frequency representation of the signal of interest, we may derive a parametric model of the signal and then use parametric estimation techniques to improve the estimation accuracy. Other methods for (t, f) post-processing are described in Section 7.5.

14.3 FLIGHT PARAMETER ESTIMATION USING DOPPLER AND LLOYD'S MIRROR EFFECTS⁰

14.3.1 ACOUSTICAL DOPPLER EFFECT

In 1842, Doppler predicted the change in the frequency of a wave observed at a receiver whenever the source or receiver is moving relative to the other or to the wave propagation medium. The effect was first verified for sound waves by Buys-Ballot in 1845 from experiments conducted on a moving train. Nowadays, more striking examples of the acoustical Doppler effect can be found in the time-frequency

⁰Authors: **B.G. Ferguson**, Defence Science and Technology Organisation, Pyrmont, NSW, Australia (Brian.Ferguson@dsto.defence.gov.au) and **K.W. Lo**, Defence Science and Technology Organisation, Pyrmont, NSW, Australia (Kam.Lo@dsto.defence.gov.au). Reviewers: G.J. Frazer and B. Boashash.

distributions (TFDs) from acoustic sensors, whether in air or under water, when a propeller-driven aircraft or a helicopter passes overhead. The dominant feature in the acoustic spectrum of a turboprop aircraft is the spectral line corresponding to the propeller blade-passage frequency. In the case of a helicopter, the dominant line corresponds to the main (or tail) rotor blade-passage frequency. This line of constant frequency represents the signal emitted by the source. During an aircraft transit, the instantaneous frequency (IF) of this signal, when received by a stationary sensor, is observed to change with time due to the acoustical Doppler effect.

14.3.1.1 Time-frequency model: Microphone in air

Consider an acoustic source moving along a straight line at a constant subsonic speed v so that at time τ_c it is at the closest point of approach (CPA) to a stationary acoustic sensor at a separation distance (or slant range) of R_c . Suppose that the source emits an acoustic tone of constant frequency f_0 and that the isospeed of sound propagation in air is c_a . The IF of the signal received by the sensor at time t is given by Ref. [16]

$$f_a(t) = \alpha + \beta p(t; \tau_c, s), \quad (14.3.1)$$

where

$$\alpha = f_0 c_a^2 / (c_a^2 - v^2), \quad (14.3.2)$$

$$\beta = -f_0 c_a v / (c_a^2 - v^2), \quad (14.3.3)$$

$$s = \frac{R_c \sqrt{c_a^2 - v^2}}{v c_a}, \quad (14.3.4)$$

$$p(t; \tau_c, s) = \frac{t - \tau_c}{\sqrt{s^2 + (t - \tau_c)^2}}. \quad (14.3.5)$$

Examples of time-frequency (t, f_a) curves computed by Eqs. (14.3.1)–(14.3.5), with NLS-optimized parameters, are shown as solid lines in Fig. 14.3.1.

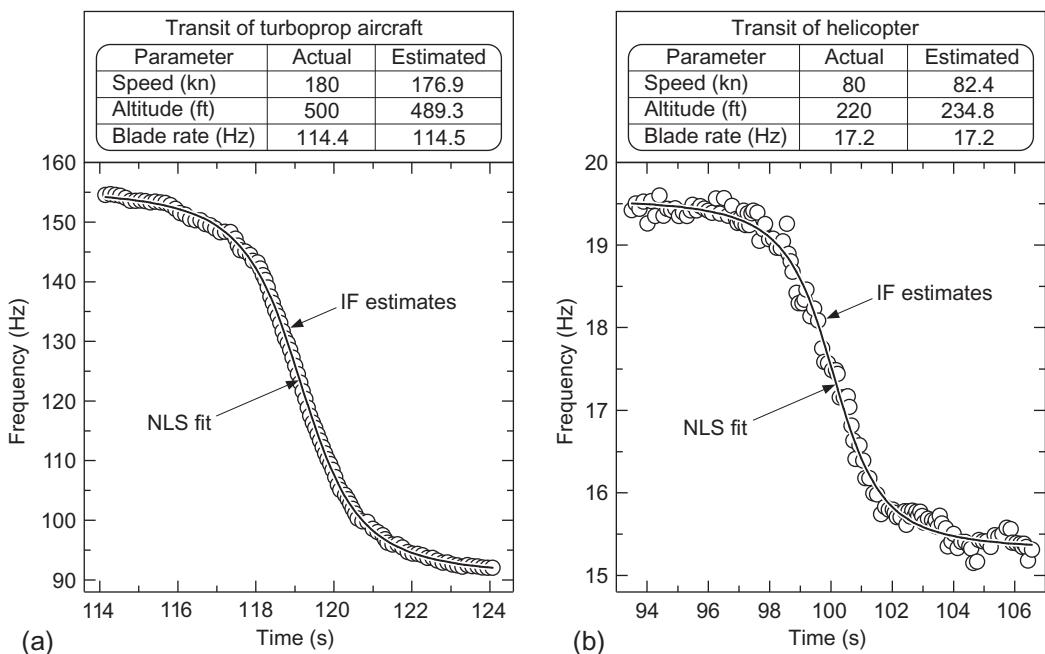
14.3.1.2 Time-frequency model: Hydrophone in water

Now consider an airborne source traveling with a constant subsonic speed v at a constant altitude h in a constant direction so that its flight path passes directly over an acoustic sensor located at a depth d below the sea surface. Due to the propagation delay, the signal emitted by the source at time τ (the *source* time) arrives at the sensor at a later time t (the *sensor* time), which is given by Refs. [17,18]

$$t = \tau + \frac{l_a(\tau)}{c_a} + \frac{l_w(\tau)}{c_w} = \tau + \frac{h}{c_a \cos \phi_a(\tau)} + \frac{d}{c_w \sqrt{1 - (c_w/c_a)^2 \sin^2 \phi_a(\tau)}}, \quad (14.3.6)$$

where c_a and c_w are the respective isospeeds for sound propagation in air and water, $\phi_a(\tau)$ is the angle of incidence (in air) at source time τ , and $l_a(\tau)$ and $l_w(\tau)$ are the respective travel distances in air and water. Assume the source emits an acoustic tone with a constant frequency f_0 . The IF of the signal received by the sensor at time t is given by Refs. [17,18]

$$f_w(t) = f_0 \frac{dt}{d\tau} = \frac{f_0}{1 + v \sin \phi_w(\tau)/c_w} = \frac{f_0}{1 + v \sin \phi_a(\tau)/c_a} \quad (14.3.7)$$

**FIGURE 14.3.1**

- (a) Variation with time of the instantaneous frequency estimates (circles) of the signal received by a microphone during the transit of a turboprop aircraft and the nonlinear least squares (NLS) fit (solid line) to the observations. The actual and estimated values of the source parameters are tabulated at the top of the diagram.
 (b) Similar to (a), but for the transit of a helicopter.

where $\phi_w(\tau)$ is the angle of refraction (in water) at source time τ . Note that Eqs. (14.3.6) and (14.3.7) constitute a set of parametric equations in $\phi_a(\tau)$. Given the source parameter set $\{f_0, v, h, \tau_c\}$ and sensor depth d , there is no explicit analytical expression for the IF f_w as a function of the sensor time t . However, a theoretical curve of f_w as a function of t can be obtained indirectly by first specifying the angle of incidence $\phi_a(\tau)$ and then calculating the corresponding values for f_w and t . Only acoustic ray paths for which the angle of incidence is less than the critical angle of incidence ϕ_c will have (some) acoustic energy transmitted across the air-sea interface into the underwater medium; i.e., $|\phi_a(\tau)|$ must be less than the critical angle of incidence ϕ_c , which is equal to $\sin^{-1}(c_a/c_w)$. Ray paths with angles of incidence exceeding the critical angle have all their acoustic energy reflected from the sea surface. Examples of time-frequency (t, f_w) curves computed by Eqs. (14.3.6) and (14.3.7), with NLS-optimized parameters, are shown as solid lines in Fig. 14.3.2.

14.3.2 ACOUSTICAL LLOYD'S MIRROR EFFECT

When a broadband acoustic source (like a jet aircraft) is in motion, the TFD of the output of a sensor located above the ground displays a pattern of interference fringes known as the acoustical Lloyd's

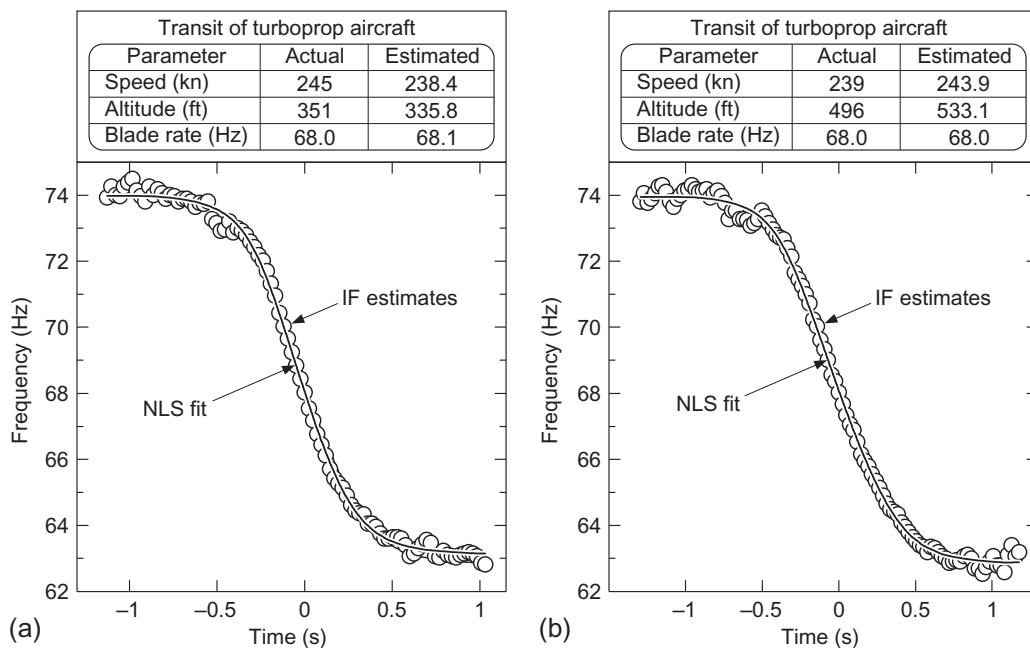


FIGURE 14.3.2

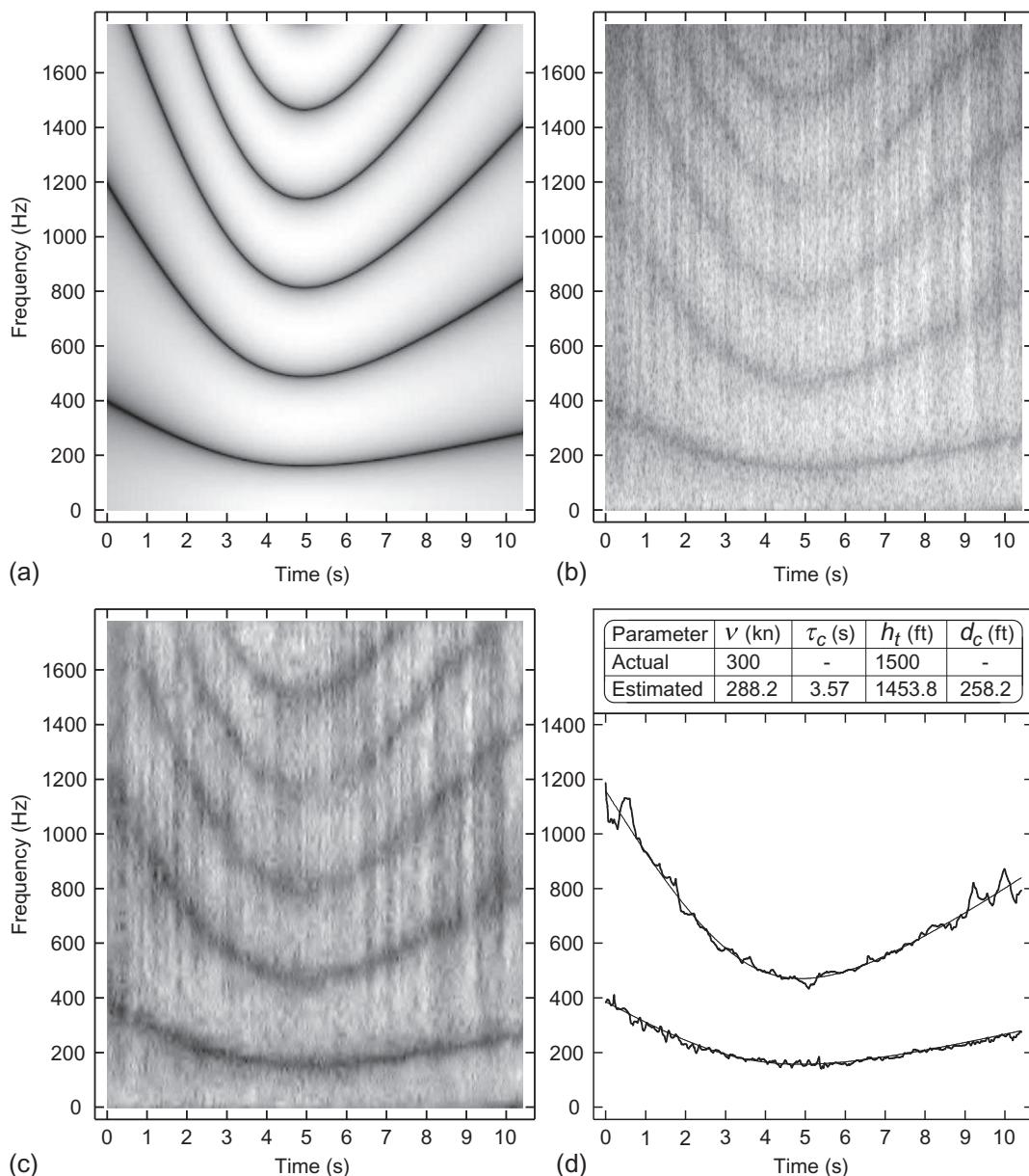
(a) Variation with time of the instantaneous frequency estimates (circles) of the signal received by a hydrophone during the transit of a turboprop aircraft and the nonlinear least squares fit (solid line) to the observations. The actual and source parameters are tabled at the top of the diagram. (b) Similar to (a), but for another transit of the turboprop aircraft.

mirror effect. Figure 14.3.3(a) shows a simulated TFD of the acoustic energy at the sensor output during the transit of an airborne source of broadband sound in an ideal environment. The resultant sound field at the sensor is the sum of the direct and ground-reflected sound fields. The reflected path is longer than the direct path between the source and the sensor, and the path length difference results in a phase difference between the two fields. At certain frequencies, the phase difference is of the correct amount to cause destructive interference (or cancelation) of the two sound fields.

14.3.2.1 Destructive-interference frequency model

Consider an airborne source traveling in a straight line with a constant subsonic speed v at a constant altitude h_t over a hard ground. The acoustic sensor is located at a height h_r above the ground. The source is at the CPA to the sensor at time τ_c , with the ground range at CPA being d_c . Suppose the isospeed of sound propagation in air is c_a . The temporal variation of the n th order destructive-interference frequency as a function of the four parameters $\{v_r, v_t, \tau_c, \gamma\}$ is modeled by Ref. [19]

$$f_n(t) = \frac{2n-1}{4} \frac{c_r^2}{c_r^2 - v_r^2} \left[\sqrt{\gamma^2(c_r^2 - v_r^2) + c_r^2 v_t^2(t - \tau_c)^2} - v_r v_t(t - \tau_c) \right], \quad (14.3.8)$$

**FIGURE 14.3.3**

(a) Simulated time-frequency distribution of a microphone's output for the transit of an airborne broadband acoustic source in an ideal environment. (b) Spectrogram of real data recorded from a microphone's output during a jet aircraft transit. (c) Image enhancement of (b) through preprocessing. (d) Predicted paths (thinner lines) of first and second destructive-interference curves using NLS method superimposed on observed paths (thicker lines) from (c).

where $v_r = v/h_r$, $v_t = v/h_t$, $\gamma = \sqrt{1 + (d_c/h_t)^2}$, and $c_r = c_a/h_r$. The four parameters $\{v_r, v_t, \tau_c, \gamma\}$ are uniquely related to the four flight parameters $\{v, \tau_c, h_t, d_c\}$. Examples of destructive-interference frequency curves $\{f_n(t) : n = 1, 2\}$ computed by Eq. (14.3.8), with NLS-optimized parameters, are shown as thinner lines in Fig. 14.3.3(d).

14.3.3 TIME-FREQUENCY SIGNAL ANALYSIS

Various techniques for estimating the IF of a nonstationary signal have been considered in the literature (see Ref. [20] and Chapter 10). A common approach is to locate the peak of a TFD of the signal. In this section, an IF estimate corresponds to the propeller blade rate of a turboprop aircraft (or main rotor blade rate of a helicopter) as observed at an acoustic sensor. Two examples of using the peaks of the spectrogram (short-time Fourier transform) to estimate the temporal variation of the IF (denoted by circles) of the signal received by a microphone located just above the ground during the respective transits of a turboprop aircraft and a helicopter are shown in Fig. 14.3.1(a) and (b). Two examples of using the peaks of the Wigner-Ville distribution (Section 2.1.4) to estimate the temporal variation of the IF (denoted by circles) of the signal received by a hydrophone located under water during the respective transits of a turboprop aircraft are shown in Fig. 14.3.2(a) and (b).

Similarly, the temporal variation of the destructive-interference frequency of the signal received by an acoustic sensor located above the ground during the transit of an airborne source of broadband sound can be extracted by locating the minima of a TFD of the sensor's output. Figure 14.3.3(d) shows an example of using the minima of the spectrogram to estimate the temporal variation of the first and second destructive-interference frequencies (denoted by thicker lines) of the signal received by a microphone located above the ground during the transit of a jet aircraft.

14.3.4 SOURCE PARAMETER ESTIMATION: AN INVERSE TIME-FREQUENCY PROBLEM

The inverse time-frequency problem is to estimate the source or flight parameters from the variation with time of the received signal's IF or destructive-interference frequency. The problem is solved by a nonlinear least-squares (NLS) method.

14.3.4.1 Narrowband source in level flight with constant velocity: Microphone in air

The source parameters $\{f_0, v, \tau_c, R_c\}$, or equivalently $\{\alpha, \beta, \tau_c, s\}$, are estimated by minimizing the sum of the squared deviations of the noisy IF estimates from their predicted values [16]. Specifically, the NLS estimates of $\{\alpha, \beta, \tau_c, s\}$ are given by

$$\{\hat{\alpha}, \hat{\beta}, \hat{\tau}_c, \hat{s}\} = \arg \left\{ \min_{\{\alpha', \beta', \tau'_c, s'\}} \sum_{k=1}^K [\alpha' + \beta' p(t_k; \tau'_c, s') - g(t_k)]^2 \right\}, \quad (14.3.9)$$

where $g(t_k)$ is the IF estimate at sensor time $t = t_k$ and K is the number of IF estimates. The four-dimensional minimization in Eq. (14.3.9) can be reduced to a two-dimensional maximization [16]:

$$\{\hat{\tau}_c, \hat{s}\} = \arg \left\{ \max_{\{\tau'_c, s'\}} \frac{|\sum_{k=1}^K [g(t_k) - \bar{g}] p(t_k)|^2}{\sum_{k=1}^K [p(t_k) - \bar{p}]^2} \right\}, \quad (14.3.10)$$

$$\hat{\beta} = \frac{\sum_{k=1}^K [g(t_k) - \bar{g}] \hat{p}(t_k)}{\sum_{k=1}^K [\hat{p}(t_k) - \bar{p}]^2}, \quad (14.3.11)$$

$$\hat{\alpha} = \bar{g} - \hat{\beta} \bar{p}, \quad (14.3.12)$$

where $\bar{g} = \frac{1}{K} \sum_k g(t_k)$, $p(t_k) = p(t_k; \tau'_c, s')$, $\bar{p} = \frac{1}{K} \sum_k p(t_k)$, $\hat{p}(t_k) = p(t_k; \hat{\tau}_c, \hat{s})$, and $\bar{p} = \frac{1}{K} \sum_k \hat{p}(t_k)$. Solving Eqs. (14.3.2) and (14.3.3) using the estimated values for α and β gives the estimates of the source speed v and source frequency f_0 as

$$\hat{v} = -(\hat{\beta}/\hat{\alpha}) c_a, \quad (14.3.13)$$

$$\hat{f}_0 = \hat{\alpha}(1 - \hat{v}^2/c_a^2). \quad (14.3.14)$$

From Eq. (14.3.4), the estimate of the CPA slant range R_c is given by

$$\hat{R}_c = \hat{s} \hat{v} c_a / \sqrt{c_a^2 - \hat{v}^2}. \quad (14.3.15)$$

The maximization in Eq. (14.3.10) is performed using the quasi-Newton method where the initial estimates of τ_c and s are given by the method described in Ref. [16]. The results of applying the source parameter estimation method to experimental data (represented by the circles) are shown at the top of Fig. 14.3.1(a) and (b). The estimates closely match the actual values of the aircraft's speed, altitude, and propeller or main rotor blade rate.

14.3.4.2 Narrowband source in level flight with constant velocity: Hydrophone in water

Define the cost function

$$P_f(f'_0, v', h', \tau'_c) = \sum_k [g(t_k) - f'_0 z_k(v', h', \tau'_c)]^2, \quad (14.3.16)$$

where $g(t_k)$ is the IF estimate at sensor time t_k , $\{f'_0, v', h', \tau'_c\}$ are the hypothesized source parameters, and $f'_0 z_k(v', h', \tau'_c) = f_w(t_k)$ is the modeled frequency at sensor time t_k , which is computed using the approximate method found in Ref. [18]. Estimates of the source parameters, $\{\hat{f}_0, \hat{v}, \hat{h}, \hat{\tau}_c\}$, are obtained by finding the set of values of $\{f'_0, v', h', \tau'_c\}$ that minimizes P_f . It can be shown that

$$\{\hat{v}, \hat{h}, \hat{\tau}_c\} = \arg \left\{ \max_{\{v', h', \tau'_c\}} \frac{[\sum_k g(t_k) z_k(v', h', \tau'_c)]^2}{\sum_k z_k^2(v', h', \tau'_c)}, 0 < v' < c_a, h' > 0 \right\}, \quad (14.3.17)$$

$$\hat{f}_0 = \frac{\sum_k g(t_k) z_k(\hat{v}, \hat{h}, \hat{\tau}_c)}{\sum_k z_k^2(\hat{v}, \hat{h}, \hat{\tau}_c)}. \quad (14.3.18)$$

Numerical procedures for maximizing Eq. (14.3.17) use standard (constrained) optimization methods. The initial estimates for $\{v, h, \tau_c\}$ are calculated using the procedure detailed in Ref. [18]. The results of applying the source parameter estimation method to experimental data (represented by the circles) are shown at the top of Fig. 14.3.2(a) and (b). The estimates closely match the actual values of the aircraft's speed, altitude, and propeller blade rate.

14.3.4.3 Broadband source in level flight with constant velocity: Microphone in air

In a TFD of the sensor output, a destructive-interference curve is a sequence of connected points on the time-frequency plane with (locally) minimum amplitudes. The positions of these points, corresponding to the temporal variation of a particular order destructive-interference frequency, define the *path* (or shape) of the destructive-interference curve and their (logarithmic) magnitudes define the *intensity* of the curve. If the TFD is treated as an image, destructive-interference curves can be identified as dark fringes in the image—see Fig. 14.3.3(a). However, in practice, noise and background variations obscure the exact paths of these curves. Figure 14.3.3(b) shows the spectrogram of a set of real data recorded from the output of a microphone during the transit of a jet aircraft. It is thus necessary that the time-frequency image be preprocessed to enhance the appearance of the destructive-interference curves before extracting the flight parameters from the image. The image shown in Fig. 14.3.3(b) is first normalized to remove background variations, and then wavelet de-noising is applied to the normalized image to suppress noise [21]. The result is an enhanced image (Fig. 14.3.3(c)) with the noise suppressed, yet the destructive-interference curves are not degraded in appearance.

Define the parameter vector $\mathbf{z} = [v_r, v_t, \tau_c, \gamma]^T$. A cost function that measures the difference between the observed and predicted paths of the destructive-interference curves is given by

$$P(\mathbf{z}) = \sum_{n=1}^N \sum_{k=1}^{K_n} [g_n(t_{nk}) - f_n(t_{nk}, \mathbf{z})]^2, \quad (14.3.19)$$

where $g_n(t)$ is the observed value of the n th order destructive-interference frequency at time t , which can be obtained by selecting the correct (local) minimum in the enhanced image, and $f_n(t, \mathbf{z})$ is the corresponding predicted value using Eq. (14.3.8). Minimizing $P(\mathbf{z})$ produces the NLS estimate $\hat{\mathbf{z}} = [\hat{v}_r, \hat{v}_t, \hat{\tau}_c, \hat{\gamma}]^T$ of \mathbf{z} . The speed, altitude, and CPA ground range of the source are then estimated as $\hat{v} = h_r \hat{v}_r$, $\hat{h}_t = \hat{v}/\hat{v}_t$, and $\hat{d}_c = |\hat{h}_t \sqrt{\hat{\gamma}^2 - 1}|$. The cost function is minimized using the Gauss-Newton method and the required initial estimate of \mathbf{z} is obtained by the procedure detailed in Ref. [21].

The raw time-frequency image (Fig. 14.3.3(b)) is too noisy for flight parameter estimation. Indeed, direct application of the NLS method to this raw image produces very poor flight parameter estimates, necessitating the use of the enhanced image (Fig. 14.3.3(c)). Figure 14.3.3(d) shows the paths (thicker lines) of the first and second destructive-interference curves extracted from the enhanced image, which are then used in the NLS method to estimate the flight parameters. Only the first and second curves are used for flight parameter estimation, i.e., $n \leq 2$ in Eq. (14.3.19), because the higher order ($n > 2$) destructive-interference curves are too noisy. The estimated and actual values of the flight parameters are shown at the top of Fig. 14.3.3(d). The estimates of the speed and altitude are in good agreement with the actual values.

14.3.5 SUMMARY AND CONCLUSIONS

The acoustical Doppler effect enables the flight parameters and blade-passage frequencies of turboprop and rotary-wing aircraft to be estimated from the time-frequency signal analysis of acoustic sensor data. Also, during the transit of a broadband acoustic source, the acoustical Lloyd's mirror effect manifests itself as an interference pattern in the time-frequency distribution of the output of an acoustic sensor positioned above the ground. The flight parameters of jet aircraft can be estimated from the destructive interference fringes.

14.4 HIGH-FREQUENCY RADAR MEASUREMENTS OF A BALLISTIC MISSILE USING THE WVD⁰

This section shows how the instantaneous frequency concentration property of the WVD can be successfully applied in a real application. As Doppler measurements are considered, the notion of IF is replaced by that of “instantaneous Doppler” (ID).

14.4.1 EXPERIMENT DESCRIPTION

Four surrogate theater ballistic missiles (TBM) were launched from a temporary launch site in northern West Australia with the objective of testing a variety of TBM launch detection sensors. The missiles comprised two stages and used Terrier and improved Orion motors together with a passive payload designed so the flight vehicle had the same in-flight length as a representative TBM. The Terrier first stage burnt for approximately 5 s following launch and the Orion second stage ignited at approximately 18 s following launch and burnt for approximately 24 s.

One of the sensors was a high frequency (HF) line-of-sight radar. The radar was positioned several tens of kilometers from the launch site and operated at a carrier frequency of either (approximately) 8 or 25 MHz. The data described here is from measurements made at 25 MHz and is from the fourth surrogate TBM launch. The radar operated bistatically with the radar transmitter and receiver spaced approximately 20 km apart to avoid the transmitter signal overloading the receiver via a direct line-of-sight or ground wave propagation path. This ensured that there were no occluded regions in the range coverage. The transmitter floodlit the region of interest. The receiver used an ‘L’-shaped array of 32 elements, with 16 elements on each arm of the ‘L’ and each arm approximately 500 m long. A 32-channel receiving system with analog de-ramping and digital output fed data to a real-time radar signal processor. Multiple directive beams covering the region of interest were formed using digital beamforming and both target elevation and azimuth could be determined. For a full description of the radar, see Ref. [22].

14.4.2 SIGNAL DESCRIPTION

The radar used a linear frequency modulated (LFM) continuous waveform with a sweep (or waveform) repetition frequency (WRF) of 50 Hz. A set of coherent measurements were collected, each of 256 sweeps or 5.12-s duration. For hardware limitation reasons, there was a short inter-dwell gap of approximately 12 sweeps ($12 \cdot \frac{1}{50} s$), where no data was recorded between coherent measurement intervals. The radar signal processor pulse compressed each sweep using stretch processing then formed 20 digital beams (10 for each arm of the “L” array). Doppler analysis of the 256 sweeps in each successive coherent processing interval was performed for each range cell in each beam direction.

The selection of WRF (50 Hz) meant that Doppler measurements of the very high-velocity target were ambiguous for most of the flight. The long coherent integration time (CIT), relative to

⁰Authors: **G.J. Frazer**, Defence Science and Technology Organisation, Edinburgh, SA, Australia (gordon.frazer@dsto.defence.gov.au) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: B. Ristic and M. Turley (DSTO).

target kinematics, increased radar sensitivity although the target acceleration reduced the coherent processing gain and limited the accuracy of velocity measurements. The radar used existing software for range, beamforming, and Doppler processing, which assumed constant-velocity targets. Rather than modifying the software to support accelerating targets, post-event acceleration analysis was used to mitigate the accelerating-target-mismatch loss.

Received radar data was displayed in real-time during the experiment and also recorded on tape for subsequent analysis. The data for off-line analysis was range processed (sweep compressed) and beamformed but not Doppler processed. It was organized into a sequence of dwells of data where each dwell contained 2 sets of 10 formed beams (1 set for each arm of the “L” array) with a group of 20 range cells of interest and the complex signal corresponding to each of the 256 sweeps in a CIT for each range and beam. Sequences of dwells were collected into a single file.

A typical range-Doppler map seen by the operator is shown in Fig. 14.4.1. It corresponds to one beam in the direction of the target and 20 range cells stacked vertically. For each range cell the Doppler spectrum is estimated from the appropriate range samples of each of the 256 pulse compressed sweeps, with one spectrum per range cell. Although not shown, the operator actually sees 20 such figures per dwell where each one corresponds to a different beam direction.

This selected example covers the time interval from launch time plus 16 s to launch time plus 21 s ($T+16$ s : $T+21$ s). The range-Doppler map shows the accelerating target smeared in Doppler (from -22 to 0 Hz) in range cells 6 and 7. The target velocity is such that this is an ambiguous Doppler

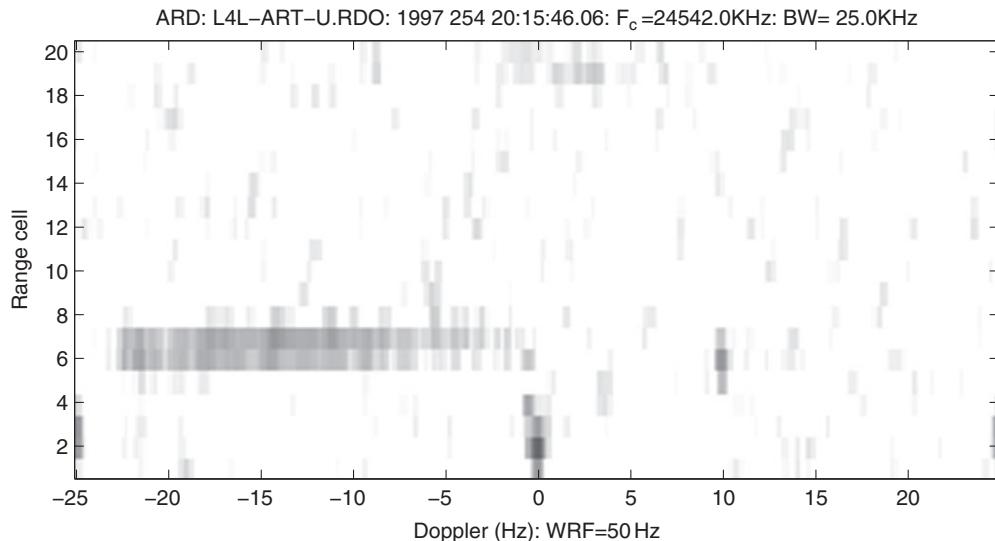


FIGURE 14.4.1

Range-Doppler map showing the accelerating target smeared in Doppler (from -22 to 0 Hz) in range cells 6 and 7. The direct wave from the transmitter and ground clutter are visible surrounding 0 Hz Doppler and centered in range cell 2. The coasting spent first stage of the two stage TBM can be seen at range cell 6 with 10-Hz Doppler.

measurement. The coasting spent first stage of the two stage TBM can be seen at range cell 6 with 10 Hz Doppler. The direct wave from the transmitter as well as ground clutter is visible around 0 Hz Doppler and centered in range cell 2. An injected calibration signal can be seen in range cells 2 and 3 at -25 and $+25$ Hz. Receding targets appear to incorrectly have positive Doppler; however, this is a frequency inversion artifact of the hardware design of the radar.

It is difficult to estimate the true time-varying velocity of the target since it changes significantly during the radar CIT. The aim is therefore to estimate the instantaneous Doppler law and hence the time-varying velocity of the target throughout all dwells which contain the target during the powered segment of flight.

14.4.3 SIGNAL MODEL

Let $z_{k,r,b,a,d}$ be the complex sample output from the radar signal processor prior to Doppler processing corresponding to the k th sweep of the r th range cell and b th beam for the a th arm of the ‘L’-shaped receive array, for the d th dwell. In the data $k \in [0, \dots, 255]$, $r \in [1, \dots, 20]$, $b \in [1, \dots, 10]$, $a \in [1, 2]$, and $d \in [1, \dots, 10]$. Now let the complex signal corresponding to the coherent radar return from all sweeps for the r th range cell, b th beam, and a th arm in the d th dwell be $z_{r,b,d,a}(k)$. This is the signal used for conventional Doppler processing; an example of the results for a set of range cells is shown in Fig. 14.4.1.

Consider now a signal model for the radar return from a particular range and beam cell in a dwell which contained the target and where now for clarity and simplicity we drop the notational dependence on r, b, d, a . Preliminary analysis using a spectrogram [23] suggested that the following discrete-time signal model would be at least piece-wise appropriate:

$$z(t) = A e^{j2\pi \left[f_0 t + \frac{\beta}{2} t^2 \right]} + c(t) + n(t) \quad (14.4.1)$$

for the discrete time $\{t : 0, 1/\text{WRF}, \dots, K/\text{WRF}\}$ and where A is the complex amplitude, f_0, β are the linear FM parameters, $c(t)$ represents clutter, and $n(t)$ represents noise. $T_{\text{CIT}} = \frac{K+1}{\text{WRF}}$ is the radar coherent integration time (CIT). For a bistatic HF line-of-sight radar, $c(t)$ includes contributions such as direct signal from the transmitter, range sidelobes from the direct signal, additional targets (e.g., a booster stage in a multistage rocket), and meteor and ionospheric scatter. In general, both clutter and noise are unknown, but we assume the relative energy satisfies

$$\frac{|A|^2 T_{\text{CIT}}}{\left[\sum_0^{T_{\text{CIT}}} |c(t)|^2 + \sum_0^{T_{\text{CIT}}} |n(t)|^2 \right]} \gg 1. \quad (14.4.2)$$

We thereby consider $z(t)$ as deterministic with unknown parameters, unknown clutter and background noise, with a high signal to clutter-plus-noise energy ratio.

The aim is to estimate the instantaneous Doppler law needed to establish the accelerating target dynamics. At a given radar carrier frequency the instantaneous Doppler as parameterized by f_0 and β can be converted to instantaneous target velocity. In some cases the instantaneous Doppler law of the target is not linear, but it is piece-wise linear only; so, we can adopt the model in Eq. (14.4.1) and apply our analysis over appropriate subsets of $\{t : 0 \leq t < T_{\text{CIT}}\}$.

14.4.4 INSTANTANEOUS DOPPLER ESTIMATION

Several methods can be applied to determine the instantaneous Doppler law (or f_0 and β) of the TBM target from the measured signal $z(t)$. Some are based on various time-frequency distributions (TFDs) [20,23–25]. Others are based on some estimation procedure for the parameters f_0 and β or the Doppler law, e.g., a maximum-likelihood criteria estimator for f_0 and β , or a least-squares criteria polynomial-phase-law estimator for the instantaneous Doppler law.

We have selected a procedure based on the Wigner-Ville distribution (WVD) of the signal $z(t)$ for several reasons. Firstly, the clutter term $c(t)$ in Eq. (14.4.1) is unknown and at least early in the flight will include components due to the Terrier booster and range sidelobes of the direct signal from the transmitter. A simple phase-difference based estimator failed in these cases and, more disturbingly, gave no diagnosis as to the source of the error. On the other hand the WVD provided a clear visualization of the time-frequency law of the signal and since $c(t)$ was not too complicated, the WVD display did not become overwhelmed by cross-terms in the WVD. Secondly, the data is in batch form (a sequence of dwells) with comparatively few samples in each dwell (256), which simplifies selection of the analysis window size for the WVD. There are small temporal gaps between dwells so that each dwell must be examined in isolation. Finally, our analysis is off-line with no requirement to be fully automatic. It was quite permissible to manually inspect and interpret the WVD of any signal of interest. For details on the WVD see Chapters 2 and 3.

14.4.5 RESULTS

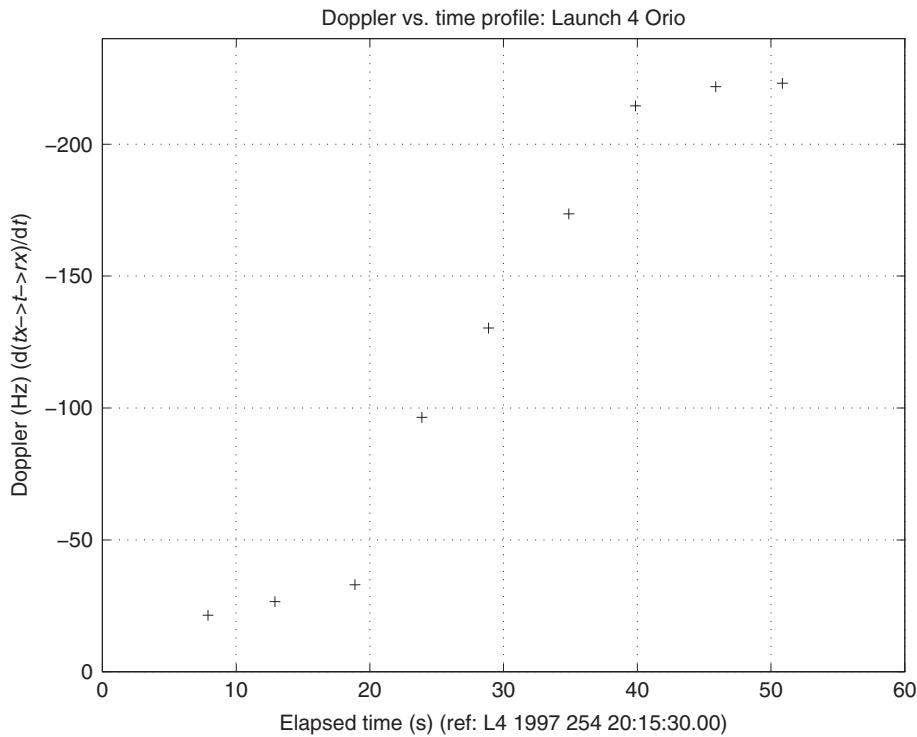
Standard analysis of the measured data from launch 4 consisted of extracting sequences of peaks corresponding to the target from beam-range-Doppler maps for the sequence of dwells covering the interval of powered flight. Each dwell comprised $(256 + 12) \cdot \frac{1}{\text{WRF}} = 5.36$ s elapsed time and a $256 \cdot \frac{1}{\text{WRF}} = 5.12$ s CIT. A single beam example of one of these maps from the dwell collected during second-stage ignition is shown in Fig. 14.4.1, as discussed previously. Figure 14.4.2 shows the extracted Doppler measurements for the full period of powered flight, one per dwell, with a manual correction for the discrete ambiguously measured Doppler caused by the selection of such a low WRF with respect to the target velocity. Similar results were obtained for target azimuth, elevation, range, and signal and noise energy.

Figure 14.4.2 shows that our Doppler measurements have poor temporal resolution. As we have only nine measurements for the period of powered flight, any attempt to predict the full trajectory following motor burnout, including impact location, will be less accurate than desired. Many of the peaks are very approximate because the target accelerated during the CIT and therefore had no single velocity or Doppler.

Using the approximate signal model in Eq. (14.4.1) and known localization properties of the WVD, we can improve on the results of our standard analysis. Our aims are to determine the target instantaneous-Doppler law, to help distinguish between accelerating and transient targets or scatterers, and to determine the instantaneous received target energy law and the processing loss caused by target acceleration.

14.4.5.1 Instantaneous Doppler law

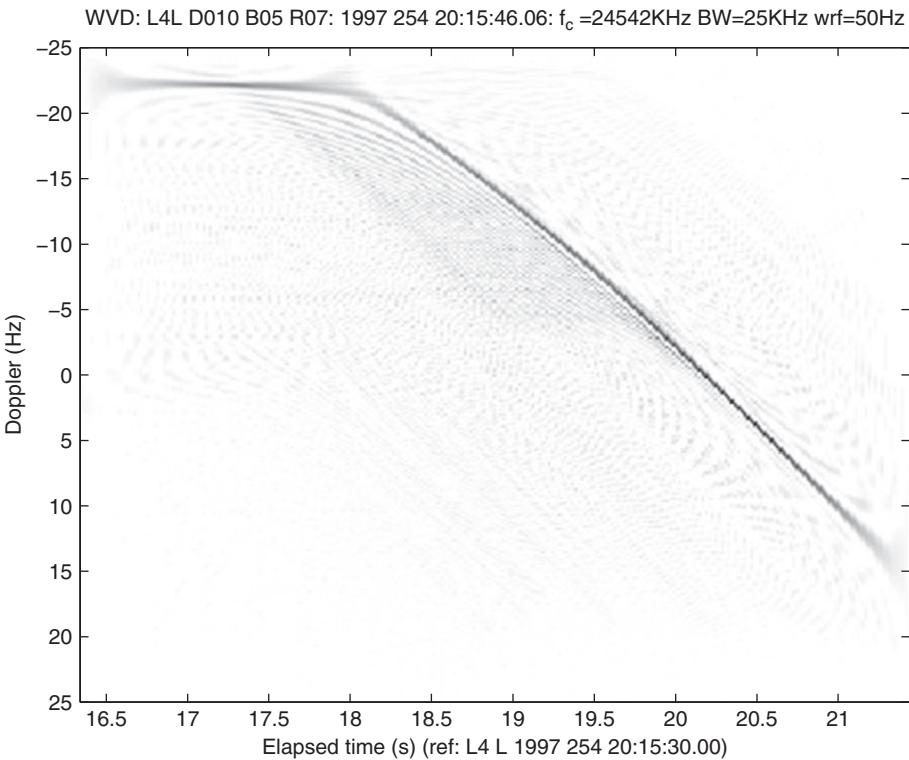
Figure 14.4.3 shows the instantaneous Doppler law estimated from the peak of the WVD. Dwells prior to and later than this measurement interval have also been analyzed using the same procedure.

**FIGURE 14.4.2**

The result of target Doppler law extraction based on conventional processing. One target Doppler measurement is determined for each dwell (5.12 s).

The sequence of Doppler estimates, approximately one per sweep, is shown in Fig. 14.4.4. There is an estimate every 20 ms excluding the inter-dwell gap intervals and some missing estimates at the beginning and end of each CIT. It covers the period of powered flight and is an interval of approximately 10 CITs or some 50 s. Contrasting Figs. 14.4.2 and 14.4.4 shows the benefit of determining accurate instantaneous Doppler law estimates as compared with conventional Doppler processing. The WVD based procedure has a temporal resolution of approximately the sweep duration, as compared with the CIT for conventional Doppler processing (20 ms vs. 5.12 s). Note that the target second-stage motor ignited at approximately 18 s and that an accurate time of ignition can be determined from the instantaneous Doppler law.

In Fig. 14.4.3, cross-terms seen in the WVD are generated between signal segments that precede and follow second-stage ignition. The cross-terms did not impede estimation of the instantaneous Doppler law. Other results showed that smoothing of the instantaneous Doppler estimates using polynomial models can reduce estimate variance and allow Doppler law prediction into the inter-dwell intervals [26].

**FIGURE 14.4.3**

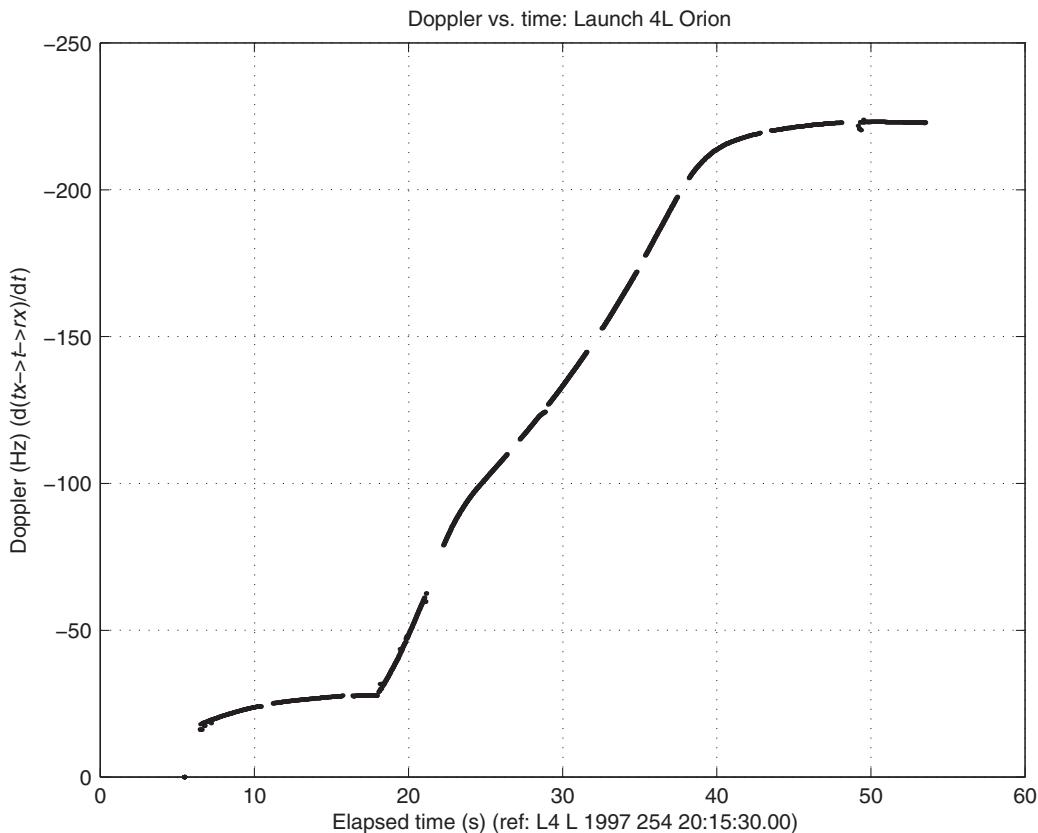
WVD of the signal containing the accelerating target from range cell 7 in Fig. 14.4.1. Second-stage ignition occurred at 18s.

14.4.5.2 Accelerating target vs. transient

The aim here is to improve discrimination between the accelerating and the transient scatterers. Figures 14.4.1 and 14.4.5 show the range-Doppler (RD) maps generated for two separate beam steer directions measured during the same 5.12 s dwell. In Fig. 14.4.1, the accelerating target is visible as a large smear in Doppler at range cell 7. In Fig. 14.4.5, a transient meteor scatterer is visible at range cell 19, which is also smeared in the Doppler domain. Figures 14.4.3 and 14.4.6 show the WVD computed from the signal corresponding to the mentioned range cells. The instantaneous Doppler law of the accelerating target is visible and so is the transient behavior of the meteor scatterer. The WVD provides a display whose features readily allows the two types of scatterers to be distinguished.

14.4.5.3 Instantaneous energy law

Knowledge of the instantaneous Doppler law can also be used to construct a demodulation reference signal, $s(t)$. This signal has unit amplitude and its IF law is the conjugate of the estimated instantaneous

**FIGURE 14.4.4**

Doppler vs. elapsed time. Computed using the WVD. The gaps are due to missing sweeps during the radar inter-dwell gap and some lost instantaneous Doppler detections at either end of individual radar CIT. The plot is a sequence of point measurements, one per radar sweep, and not a continuous line, as it appears at this scale.

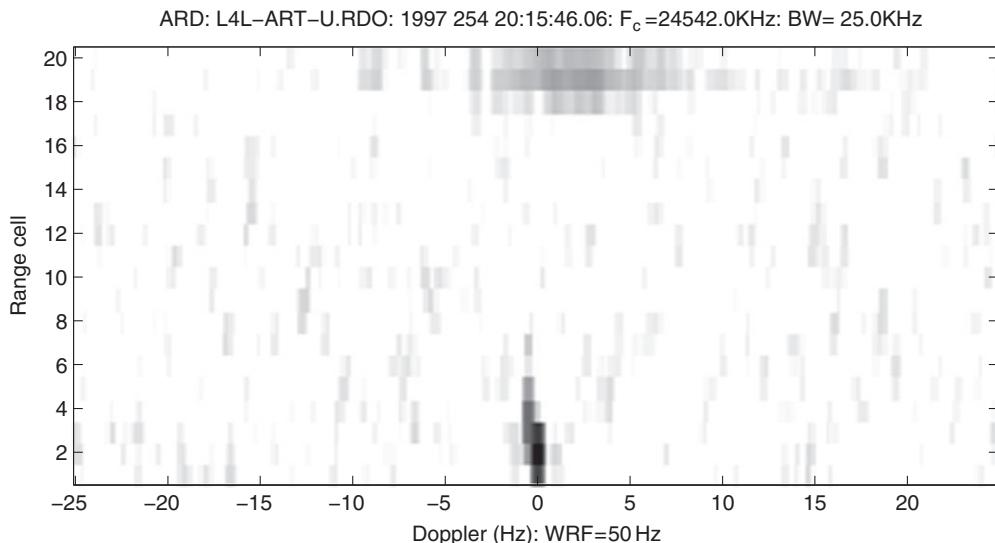
Doppler law of the target, i.e., the IF law of $z(t)$, so that $s(t)$ can be used to demodulate $z(t)$ giving the approximately constant Doppler frequency signal $z_d(t)$:

$$z_d(t) = z(t) \cdot s(t). \quad (14.4.3)$$

The instantaneous energy of the demodulated signal is

$$\hat{p}(t) = \mathbf{G}|z_d(t)|^2, \quad (14.4.4)$$

where \mathbf{G} is some zero-phase smoothing operator. The instantaneous energy is shown in Fig. 14.4.7, which shows three different levels of local smoothing (i.e., different \mathbf{G}). We observe a significant

**FIGURE 14.4.5**

Range-Doppler map showing the transient meteor scatterer in range cell 19. It is difficult to distinguish between this smeared feature and the smeared accelerating target shown in [Fig. 14.4.1](#).

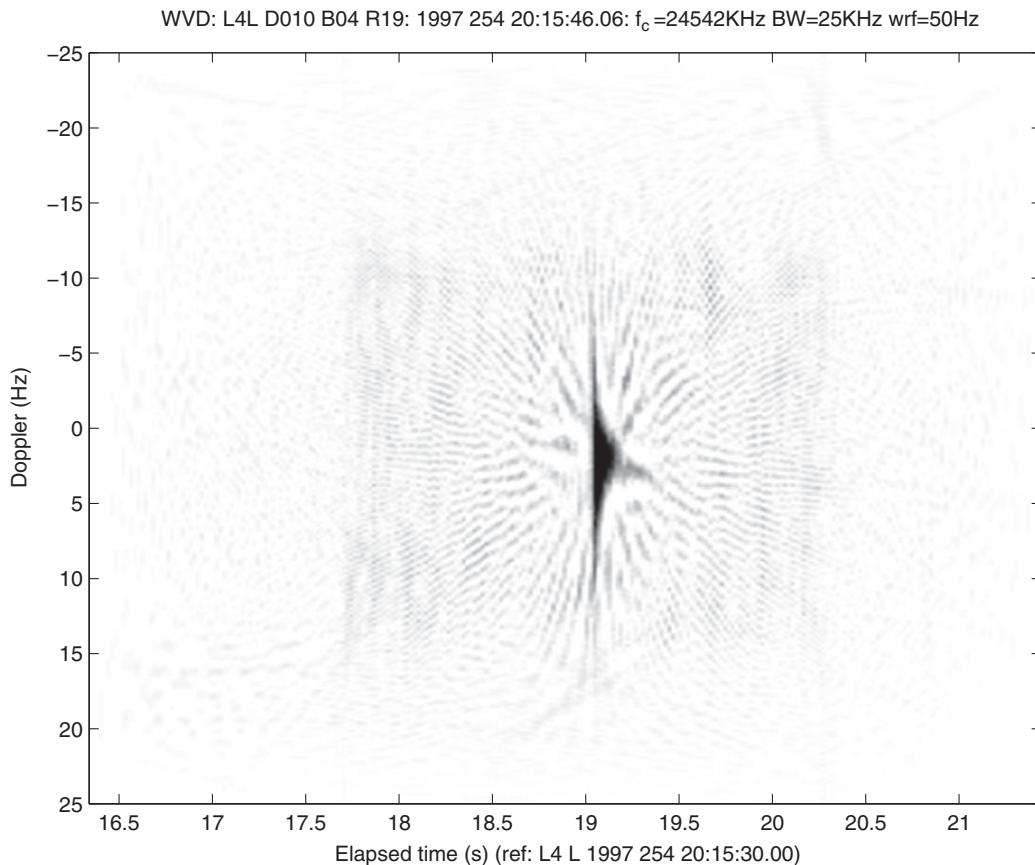
unexplained reduction in the target return energy immediately following second-stage ignition between 18 and 18.5 s.

14.4.5.4 Processing loss due to target acceleration

The processing loss due to target acceleration compared with a similar target of constant velocity can be determined. One contrasts standard Doppler processing applied to the signal $z(t)$ and to the demodulated version $z_d(t)$. [Figure 14.4.8](#) shows that the processing loss is approximately 10 dB for this particular CIT. This is the worst case because the dwell considered covers the period immediately before and after second-stage ignition, including the time of greatest acceleration.

14.4.6 SUMMARY AND CONCLUSIONS

The Wigner-Ville distribution has been applied to HF line-of-sight radar measurements of a surrogate TBM launch. A WVD-based procedure has been used to estimate the Doppler law of the target at a temporal resolution of approximately 20 ms. In contrast, standard processing had a temporal resolution of 5.12 s and contained error due to target acceleration during the radar CIT. The WVD-based display helped distinguish between the accelerating target and interfering transient meteor scatterers. It has also allowed determination of the instantaneous energy law of the target, identifying a still unexplained reduction in received signal level immediately following stage-two ignition. Finally, analysis based on the procedure using the WVD has provided an estimate of the processing loss when standard Doppler processing is applied for this particular accelerating target.

**FIGURE 14.4.6**

WVD of the signal containing the transient meteor scatterer from range cell 19 in Fig. 14.4.5.

14.5 TIME-FREQUENCY SONAR PROCESSING⁰

Sonar (for “sound navigation and ranging”) and Radar (for “radio detection and ranging”) were originally used for finding the distances and velocities of targets. They have evolved to scanning systems that are capable of localizing, identifying, and classifying targets. Other applications that use sound under water include tracking and locating shoals of fish, identifying and tracking whales, sea-bottom surveying for mining, estimating depth or turbulence or temperature variations of the ocean, tracking aircraft flying above the ocean, vision systems for underwater robotic vehicles, etc. These tasks

⁰Authors: **V. Chandran**, Queensland University of Technology, Brisbane, QLD, Australia (v.chandran@qut.edu.au) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: P. J. O’Shea, K. Abed-Meraim, A. Boudraa, and S. Ouelha.

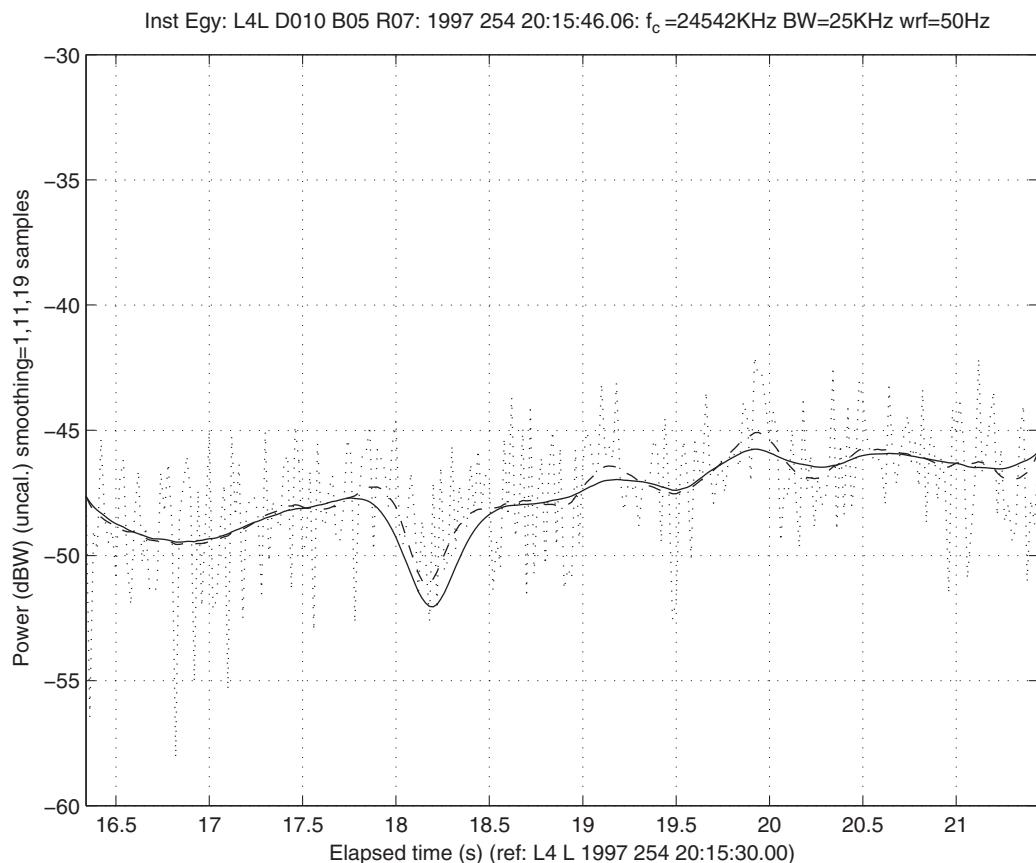


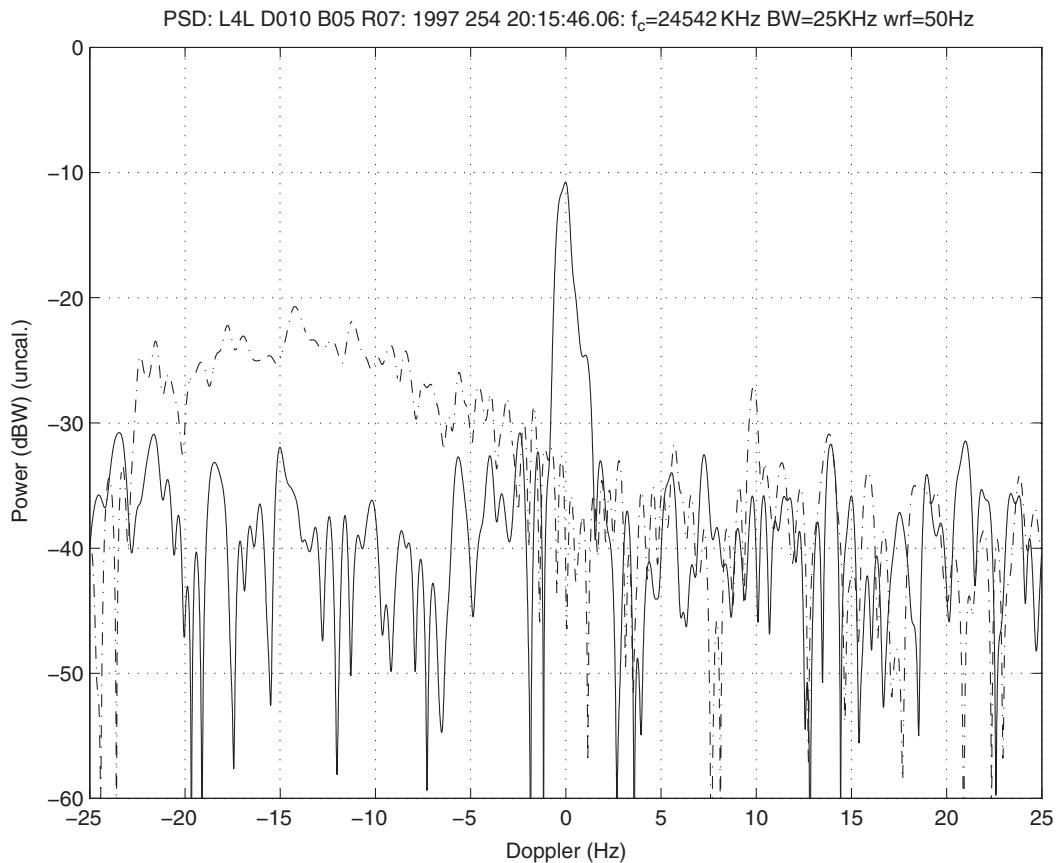
FIGURE 14.4.7

Instantaneous energy law vs. elapsed time for the interval 16.5–21.5 s after launch. The three curves correspond to 1 (..), 11 (- -), and 19 (-) sample zero-phase moving-average smoothing.

require the processing of a received signal to extract the relevant information. Some of these systems are active, that is, they transmit a sound signal and receive the reflected or backscattered signal. Others are passive and only receive sound emanating from other sources. The transmitters and receivers of sound in water are called hydrophones. Often arrays of hydrophones are employed to provide narrow beamwidths and increased sensitivity.

14.5.1 PRINCIPLES OF SONAR PROCESSING

Sonar can be active or passive. In **active sonar**, the transmitted signal has a known shape. The basic problem is to determine the delay in receiving the reflected or scattered signal. This delay is proportional to the target distance (i.e., range) and inversely proportional to the velocity of sound in water (about

**FIGURE 14.4.8**

The Doppler spectrum computed over the full CIT for the original signal (—) and for the demodulated signal (---). The processing loss caused by assuming a constant-velocity target is approximately 10 dB in this case.

1500 m/s). The wavelength of a 15 kHz component is therefore about 10 cm. To obtain better resolution, ultrasound at frequencies above 20 kHz is frequently employed in active sonar. The signal travels through water on its way to the target and back and is modified by the channel response. The received signal could be a superposition of scattered signals from several points on the target, reflections from other objects in the environment (referred to as multipath), and noise.

14.5.1.1 Active sonar modeling

To model active sonar in general, both object and channel are considered linear and time-varying. The channel impulse response $c(\vec{r}, t)$ and the insonified object's impulse response $o(\vec{r}, t)$ are functions

of space and time. The received signal $s_{\text{rec}}(\vec{r}; t)$ is related to the transmitted signal $s_{\text{tr}}(t)$ through the superposition integral

$$s_{\text{rec}}(\vec{r}; t) = \int_{-\infty}^{\infty} c(\vec{r}; t, \tau) v(\vec{r}; t, \tau) d\tau + n(t), \quad (14.5.1)$$

where $\vec{r} = [x, y, z]$ is the spatial position vector, t is the time, $n(t)$ is the noise, and

$$v(\vec{r}; t) = \int_{-\infty}^{\infty} u(\vec{r}; t, \tau) o(\vec{r}; t, \tau) d\tau, \quad (14.5.2)$$

where

$$u(\vec{r}; t) = \int_{-\infty}^{\infty} c(\vec{r}; t, \tau) s_{\text{tr}}(\tau) d\tau. \quad (14.5.3)$$

The spatial and temporal coherence of the ocean medium determine the largest aperture that can be used for image formation and the longest waveform duration that can be used in sonar systems. It is usually assumed that the coherence time of the channel is much greater than the time taken by the signal to travel to the target and back. Then the phase shift remains nearly constant over one pulse period. The signal bandwidth is also assumed much smaller than the coherence bandwidth of the channel, making the channel wideband and frequency nonselective.

If the channel impulse response does not change over the duration of the signal, the system can be considered to be linear and time-invariant, and the superposition integrals are replaced by convolution integrals.

$$s_{\text{rec}}(t) = \left[\left(s_{\text{tr}}(t) * \underset{t}{c}(t) \right) * \underset{t}{o}(t) \right] * \underset{t}{c}(t) + n(t), \quad (14.5.4)$$

where $*$ represents convolution in time. The channel impulse response consists of delayed impulses caused by various paths of propagation, such as echoes from the target and the sea-bottom. A typical channel impulse response that accounts for multiple reflections (but ignores point scatterers) is the discrete summation of impulses

$$c(t) = \sum_{i=1}^{M} c_i \delta(t - \tau_i), \quad (14.5.5)$$

where the c_i represent the attenuation suffered by each reflection term. Usually, in active sonar, the return from the target of interest is the strongest. The ideal condition is when there is only one reflection, i.e., when $M = 1$. In practice, there are not only multiple reflections from other objects in the environment (such as fish or debris) but also backscattering from the ocean floor. If the lags for which the returned signal is analyzed are restricted to those corresponding to reasonable expected distances of the target, bottom backscatter and reflections from cluttering objects outside this “ballpark” can be ignored. The problem then essentially becomes one of detecting and locating the principal echo in additive noise that may be nonstationary owing to clutter.

Time-frequency analysis plays an important role in analyzing the received signal and estimating the range of the echo from the target because

- the transmitted signal is often frequency-modulated;
- there is a Doppler shift if the target is moving;

- time-delay is the parameter of most interest;
- noise levels are high and signals are highly nonstationary, so that classical methods such as correlation or matched filters do not perform well;
- (t,f) displays are physically meaningful for human observers to interpret.

14.5.1.2 Passive sonar modeling

In **passive sonar** (or acoustic) systems, there is no transmitted pulse; the sound emanates from the target itself. However, there may still be multiple paths by which the sound travels to any hydrophone. In fact, the interference between the direct path and the bottom-reflected path can be used to advantage in determining the distance and velocity of the source, and is particularly useful in detecting and tracking ships off the coast; it leaves parabolic patterns in the (t,f) plane as shown in Ref. [27]. If the channel is modeled as

$$c(t) = \delta(t) + c_1 \delta(t - \tau), \quad (14.5.6)$$

then the magnitude-squared frequency response is $|C(f)|^2 = 1 + c_1^2 + 2c_1 \cos(2\pi f \tau)$. If h is the bottom depth, $h-k$ the hydrophone depth, and d the distance from the hydrophone to the ship along the surface, then it can be inferred from the geometry [27] that the direct signal travels a distance $(d^2 + (h-k)^2)^{1/2}$ and the bottom-reflected signal travels a distance $(d^2 + (h+k)^2)^{1/2}$. The time delay between them is $\frac{1}{c} \left[(d^2 + (h+k)^2)^{-1/2} - (d^2 + (h-k)^2)^{-1/2} \right]$, where c is the velocity of sound in water. Because d is often far greater than $h-k$ or $h+k$, the square roots can be approximated by first-order Taylor expansions to show that $\tau \approx \frac{2hk}{dc}$. Since the ship is in motion, τ will vary with time. If a is the minimum range of the ship and v its velocity of passage, and if the time origin is at the moment of closest approach, then $d^2 = a^2 + (vt)^2$. Again, when vt is much larger than a , this becomes $d \approx vt + \frac{a^2}{2vt}$, and the first interference maximum occurs at the frequency

$$f_{\max} = \frac{1}{\tau} \approx \frac{\left[vt + a^2/(2vt) \right]}{2hk}. \quad (14.5.7)$$

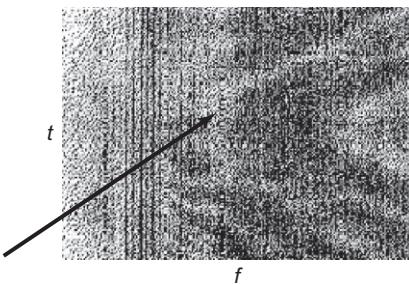
Therefore, a (t,f) display will show parabolic patterns (see Fig. 14.5.1). Even if there is only flow noise from the vessel, these patterns can show up in long-time history averaged time-frequency representations. From these patterns, the speed of the vessel and its range can be estimated.

The passive sonar problem can be considered as one of blind deconvolution where the emitted signal parameters and the channel parameters need to be simultaneously estimated using methods such as those presented in Section 8.4. Time-frequency methods are able to overcome the ill-conditioning typical of single sensor deterministic deconvolution techniques. Passive sonar can also be viewed as a transient signal detection problem in nonstationary noise when the signal-to-noise ratio is poor.

14.5.2 STANDARD METHODS USED IN SONAR

14.5.2.1 Matched filter

The optimal detector for a signal of known shape, $s(t)$, in additive Gaussian noise is the matched filter. If the received signal is $s_{\text{rec}}(t)$ and there is no Doppler shift, the matched filter can be formulated as a correlation, i.e.,

**FIGURE 14.5.1**

Spectrogram of ship noise. Parabolic patterns are visible toward the right of the plot. The time (t) varies from 0 to 12 min and the frequency (f) from 0 to 280 Hz. (From Ref. [27].)

$$\eta(\tau) = \int_{-\infty}^{\infty} s_{\text{rec}}(t) s^*(t - \tau) dt. \quad (14.5.8)$$

The greatest correlation between the delayed transmitted signal and the received signal will occur at the value of τ which corresponds to the delay of the principal echo. Hence the ideal matched filter output exhibits a peak at this delay. The matched filter output is thresholded to make a detection decision.

With increasing noise, dispersion in the medium and uncertainty in the transmitted signal shape, it becomes more and more difficult to obtain a discernible peak in the output at the correct location. When there is relative velocity between the source and receiver and a corresponding Doppler frequency v , the matched filter can be used for range-Doppler processing as

$$\eta(v, \tau) = \int_{-\infty}^{\infty} s_{\text{rec}}(t) s^*(t - \tau) e^{-j2\pi vt} dt, \quad (14.5.9)$$

where the delay τ corresponds to the range parameter. Equation (14.5.9) is also known as the Woodward ambiguity function. A symmetric formulation known as the Sussman ambiguity function is discussed in [Chapters 2 and 3](#) and [Section 5.1](#).

14.5.2.2 Hypothesis tests

Hypotheses tests, based on energy in sliding window signal segments, spectral density correlations, and other statistics, are employed to decide whether a segment around a particular location resulted from only noise, or signal (echo of the transmitted pulse in active sonar or signature of some acoustic source of interest in passive sonar) plus noise. The alternative hypotheses for the detection of an unknown signal $s(t)$ in additive noise $n(t)$ are therefore

$$H_0 : s_{\text{rec}}(t) = n(t), \quad (14.5.10)$$

$$H_1 : s_{\text{rec}}(t) = s(t) + n(t). \quad (14.5.11)$$

A test (or decision) statistic η is computed from the received signal $s_{\text{rec}}(t)$ or one of its transformed representations. If the value of the statistic exceeds a certain threshold, the hypothesis is accepted. A measure of goodness of the test statistic (η) is the signal-to-noise ratio (SNR) [28]:

$$\text{SNR} = \frac{|E(\eta/H_1) - E(\eta/H_0)|}{\sqrt{\frac{1}{2} [\text{var}(\eta/H_1) + \text{var}(\eta/H_0)]}}, \quad (14.5.12)$$

where $E(\dots)$ denotes the expected value and $\text{var}(\dots)$ the variance. This SNR can be used to compare test statistics. If the energy of the signal $s(t)$ is A , and if the noise is additive, Gaussian and white with power spectral density N_0 , then the SNR of the matched filter output as a decision statistic is given by $\sqrt{A/N_0}$. This approach is extended by defining (t,f) matched filters in [Section 12.5](#) and [\[6,29\]](#).

For classification of the signal into one of several known categories rather than detection of one category, the transient event would need to be segmented and a set of features extracted from the segment. Instead of thresholding a likelihood, a comparison of likelihoods can be used (e.g., the maximum-likelihood Bayes classifier).

Time-frequency methods are also used in hypothesis testing frameworks using a test statistic and in classification frameworks using a set of features extracted from the TFDs of the signal. The two-dimensional nature of the representation offers some potential for the selection of discriminating features that are also robust to noise (for more details see [Section 12.6](#)).

14.5.3 TIME-FREQUENCY APPROACH TO SONAR

14.5.3.1 Quadratic and adaptive TFDs in sonar: Basic approaches

For sonar signal detection with TFDs energy in the (t,f) plane is used as a test statistic. Let us consider the WVD defined in [Section 2.1](#) as

$$W(t,f) = \int_{-\infty}^{\infty} z_{\text{rec}}(t+\tau/2) z_{\text{rec}}^*(t-\tau/2) e^{-j2\pi f\tau} d\tau, \quad (14.5.13)$$

where $z_{\text{rec}}(t)$ is the analytic associate of $s_{\text{rec}}(t)$. The noise performance [\[28\]](#) is

$$\text{SNR} = \sqrt{\frac{A}{N_0}} \frac{1}{\sqrt{1+N_0/A}}. \quad (14.5.14)$$

Thus the WVD is known to be suboptimal and its performance degrades in high noise, due to its cross-terms. However, it provides the potential for time-varying filtering, and the effect of noise on the test statistic can be reduced by windowing the distribution. It is thus possible to obtain adaptive filtering prior to statistic computation or feature selection and use a high-resolution reduced interference TFD as defined in [Chapter 3](#).

If the reference signal waveform is known, we can apply the cross WVD

$$W_{z_{\text{rec}},z}(t,f) = \int_{-\infty}^{\infty} z_{\text{rec}}(t+\tau/2) z^*(t-\tau/2) e^{-j2\pi f\tau} d\tau \quad (14.5.15)$$

(cf. [Section 3.1.2.3](#)). It is equivalent to the matched filter approach in this case but with the added advantage of a 2D (t,f) plane for adaptation and filtering.

If the transient signal is monocomponent and of known instantaneous frequency (IF), a two-dimensional window in the (t,f) plane that preserves points in a neighborhood of the IF, while filtering out all others, can be used to provide a test statistic with improved SNR. If the transient signal is multicomponent, one can compute and add the cross WVDs of the received signal with

the separate components of the transient signal, thereby eliminating the effect of cross-terms while retaining auto-terms in the (t,f) plane. The IF laws in this case would need to have been identified and are assumed known [28]. Otherwise, they can be estimated iteratively.

Other quadratic TFDs such as the **cone-shaped kernel** (t,f) representation (CK-TFR) or ZAM distribution, given in [Section 3.3](#), can attenuate the cross-terms and hence may be better suited for (t,f) displays of active sonar returns [30].

The **adaptive optimal kernel** (AOK), described in [Section 5.3](#), uses a radially Gaussian signal-dependent kernel. The shape of the kernel is changed to optimally smooth the TFD as discussed in [Chapter 3](#) and [Section 5.3](#). The optimal kernel is one that maximizes a quality measure for the TFD, subject to constraints that force the kernel to suppress cross-terms while keeping a good (t,f) resolution.

Performance comparisons of the AOK (t,f) representation with the STFT, WVD, and ZAM kernels on real multicomponent active-sonar signals are shown in Ref. [31]. [Figure 14.5.2](#) illustrates various TFDs of a man-made underwater acoustic signal comprising three chirps in noise. The TFD plots, here, show that the horizontal axis is time and the vertical axis is frequency. The chirps show up in the AOK plot as three line-like features decreasing in frequency with time. They overlap in time but not significantly in the (t,f) plane of the AOK TFD. Hence the chirps can be detected and delays estimated from this representation. Notice that the Wigner and ZAM distributions suffer more severely from cross-term artifacts while the resolution of the STFT is too poor to bring out the relevant features.

14.5.3.2 Advanced approaches using high-resolution TFDs

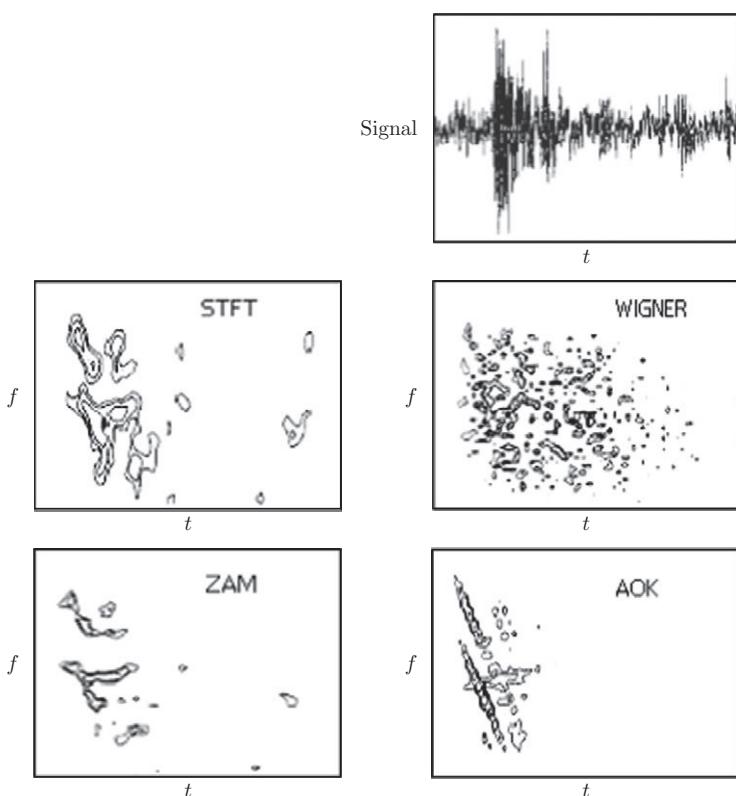
The above methods can be improved by using high-resolution TFDs such as those defined in [Chapter 3](#) and [6], leading to advanced (t,f) detection methods as described in [Sections 12.5](#) and [12.6](#). Another approach in passive sonar uses TFDs designed using human physiology and human perceptions considerations [32]. Such physiology based TFDs are essentially filter bank methods [33].

14.5.3.3 Gabor expansion and wavelets in sonar

A **cross-term deleted Wigner representation** (CDWR), based on Gabor expansions (see [Section 11.2](#)), is applied in Ref. [34] to sonar detection and classification, on data collected by the US Navy. The test data includes hyperbolic FM chirp signals and continuous-wave signals for two events. The decision statistic is a cross-correlation between the cross-CDWR of the received and replica signals and the auto-CDWR of the replica signal. This method achieves better detection accuracy than the square-law detector used by the US Navy. It is tailored to helicopter-deployed arrays of hydrophones, which are often close to the target but have wide beamwidths and consequently poor SNRs.

The **wavelet transform** and its squared magnitude (the scalogram) are compared with the WVD for automatic classification of active sonar data in Ref. [35]. Classification accuracy of up to 92% is achieved at -4 dB SNR on synthetic data of acoustic scattering from high-aspect-ratio solids (length-to-width ratios of 5 and 10). The wavelet transform often outperforms the WVD partly because it does not suffer from cross-terms. The scalogram, however, does suffer from such effects because it is a quadratic representation.

Wavelet-based features and artificial neural networks were used to classify underwater mammal sounds in Ref. [36]. Accuracy is shown to be significantly greater than for the use of the same number of Fourier energy and frequency features. The improvement is largely due to the wavelets' nonuniform tiling of the (t,f) plane and also due to BCM feature extraction that performs unsupervised nonlinear dimension reduction and supervised discriminant pursuit.

**FIGURE 14.5.2**

Time trace (top) and various TFRs (labeled) of a sonar signal comprising three chirps. The AOK representation resolves features (nearly linear with negative slopes, showing frequencies decreasing with time) corresponding to the three chirps. The frequency sweep would typically be about 10-20 kHz and the transients would last tens of milliseconds. (Adapted from [31].)

14.5.4 PRONY AND HIGHER-ORDER SPECTRAL METHODS IN SONAR

There are other “indirect” methods that can provide (t,f) representations (primarily by means of short-time or sliding-window analysis) although they are adapted from non-time-frequency techniques first proposed for, e.g., linear and higher-order spectral analysis. They are not members of the quadratic class of (t,f) representations but are sufficiently important to warrant description in this section.

14.5.4.1 Prony model method

The modified Prony method of detection of short-duration transient signals in noise is based on complex exponential signal modeling. It employs forward and backward predictions using high prediction orders and eliminates noise zeroes from the forward and backward characteristic polynomials by using the fact that zeroes of white, stationary noise do not exhibit conjugate reciprocal relationships in the z -plane

as do true signal zeroes. Singular-value decomposition is used to provide further noise enhancement. After these steps, an enhanced short-duration signal spectrum for the analysis window is estimated. A sliding window ensures that a (t,f) representation is obtained from which features can be extracted for detection and/or classification. The performance of this technique is compared with the STFT and the Wigner distribution on underwater acoustical data (tuna and sperm-whale clicks) in Ref. [37]. The transient events lasted 10–30 ms. Significantly improved features were extracted by the Prony technique.

14.5.4.2 Higher-order spectral methods

Gaussian noise immunity, translation invariance, and other useful properties of higher order spectra are also used in obtaining robust (t,f) representations and in the feature-extraction stage after a representation. Higher-order spectra (HOS) are Fourier representations of cumulants or moments of a stationary random process. They are functions of more than one frequency. The bispectrum is a function of two frequencies and is the FT of the third-order cumulant, which is a function of two lag variables. Unlike the power spectrum (which is the second order spectrum), HOS retain phase information. They are zero for Gaussian processes and can be used to detect non-Gaussian processes and identify nonlinear systems. For transient detection, as required in sonar, the stationarity assumption needs to be relaxed. Further, the noise is often nonstationary as well and can contribute to nonzero higher-order spectral contributions. Time-varying versions of higher-order spectra include the third-order Wigner distribution called the Wigner bispectrum [38], and a class of generalized Wigner-Ville distributions based on polynomial phase laws [31], which is applied to underwater acoustic signals.

A sonar image processing technique based on HOS feature extraction, presented in Ref. [39], is also indirectly related to (t,f) representation. This method applies to sonar scans that result in 2D images with azimuth and range as independent variables. Sonar0, Sonar1 and Sonar3 sea-mine databases from the US Naval Surface Warfare Center, Coastal Systems Station, Florida (references in Ref. [39]) are used. The three databases together contain 122 sea-mines in the training set and 123 in the testing set. The processing stages comprise adaptive Wiener filtering, sliding-window matched filtering, and adaptive thresholding. These steps account for nonstationary background noise. In this method, the thresholding after matched filtering is used for removing outlying noise rather than for detection. The resulting image is then used to extract features—the principal ones being bispectral and trispectral features (references in Ref. [39]) that are designed to be invariant to scaling and translation. Features that are robust to scaling make the task of the classifier easier. A statistical multistage classifier is used. The method achieves close to 90% accuracy with about 10% false alarms for detection of sea-mines. Since HOS are evaluated over overlapping windows, this method is an indirect “space-frequency” representation approach applied to 2D input. It is capable of being trained to achieve robustness to intra-class variations and some channel variations. Other features for pattern recognition using invariants defined from HOS are described in Ref. [40].

14.5.5 DISPERSION AND ANGLE FREQUENCY REPRESENTATION

Different propagating modes through the ocean (viewed as a waveguide) can exhibit different dispersion (group velocity vs. frequency) curves. Different frequencies travel with different velocities within the same mode and between different modes of propagation. As a result, it is possible for a linear FM up-ramp pulse to be converted into a down-ramp pulse of a different frequency law and increased

duration, in extreme cases. This effect is studied in Ref. [41] along with a method for deconvolution of the channel response. Such deconvolution is sometimes necessary before feature extraction to reduce classification or parameter estimation errors.

Dispersion can also be beneficial in classifying ocean bottoms. Bottom backscattering is a function of the type of surface (such as sand, pebble, clay, rock) and of the frequency of sound and incident angle. Wideband sonar (typically 20–140 kHz [42]) is used to characterize the frequency dependence of bottom backscatter and estimate the backscattering coefficient or the impulse response of the bottom. Conventional methods are not adequate for this. The directivity of the transducer and propagation attenuation cannot be ignored and must be deconvolved from the echo response. An angle frequency representation is then obtained which exhibits features similar to (t,f) representations that can be used for classification. When the scale of the roughness is much smaller than the depth, it affects only the phase and not appreciably the magnitude of the response at any frequency. However, there is an incident angle dependence which can be used for classification. A sand bottom shows a decrease in the backscattering coefficient as the angle increases, whereas, for surfaces with greater roughness such as the pebble bottom, the backscattering coefficient is nearly independent of the angle.

14.5.6 SUMMARY AND CONCLUSIONS

Time-frequency analysis may be applied with considerable success to active and passive underwater sonar data. Although multipath interference is usually a problem in sonar (and radar) applications, one of the first applications of (t,f) analysis to underwater sonar actually exploited multipath propagation for detecting and tracking ships. Reduced-interference quadratic TFDs allow (t,f) characteristics of transient signals to be displayed and interpreted without difficulty even at low SNR and even for multicomponent signals. Filtering in the (t,f) plane allows for robust feature selection and decision statistics of better quality.

Automated classification leads to multi-class problems with more than a few classes with signatures that often show significant intra-class variation due to changing channel conditions such as multipath and dispersion.

Two- and three-dimensional scans make the classification task easier through the exploitation of spatial correlations of the returned signal or the matched filtered output. However, they introduce additional difficulties with real-time processing. The concept of first expanding a signal in dimensionality to an information-rich space, and then selectively reducing the dimensionality through projections, slices, or segmentation, is a powerful one and is being exploited in many pattern recognition problems. TFDs are one example of such an approach. They can reveal the nature of the evolution of correlations in a transient signal and provide information on the appropriate subspaces to extract features from. Practical exploitation of this in an automated detection, classification, and tracking system requires its interfacing with other powerful concepts from other areas such as feature extraction, feature selection, adaptive information processing, classifier selection, and classifier fusion, as detailed in [Section 12.6](#).

14.6 SPARSE TIME-FREQUENCY DISTRIBUTIONS APPLIED TO GEOPHYSICS⁰

Time-frequency (t,f) tools are particularly successful in analyzing nonstationary signals (Chapter 1), but the window function limits the resolution of the transform (Fig. 2.3.1). Many attempts have been made to improve the resolution of the short-term Fourier transform (STFT) by optimizing the window function either parametrically [43] or nonparametrically (Section 5.3); however, for most practical signals, such as seismic signals, the desired time-frequency distribution (TFD) cannot be reached by just optimizing the window parameters; other tools such as regularization also need to be employed.

For a linear TFD, the transform is a linear mapping between the time and the (t,F) domains, whereby the signal is decomposed into a series of windowed Fourier atoms. A common way of computing the transform (series) coefficients is projecting the signal onto the basis elements using the adjoint operator. But because the range of the adjoint operator is just a small subset of its domain, the original TFD may be inaccessible to it. This section considers this problem from the viewpoint of linear algebra and inverse theory. As the TFD for a given signal is nonunique with a large degree of freedom, the (t,f) analysis, instead of using the adjoint operator, is formulated as a least-squares (LS) problem, allowing additional constraints on the coefficients in the (t,f) domain. The sparse regularization proposed here considerably increases the performance (resolution, concentration, etc.) of (t,f) representations and allows development of an invertible transform with adjustable (t,f) resolution.

14.6.1 LINEAR DISCRETE TIME-FREQUENCY DISTRIBUTIONS

The joint discrete (t,f) (DFT) pair of the N -length discrete signal $\mathbf{x} \in \ell^2(\mathbb{Z})$ (the space of finite energy discrete signals) can be extended to their nonstationary forms

$$\mathbf{x}(n) = \sum_{n'=0}^{N-1} \sum_{k=0}^{N-1} F_x^w(k, n') w(n - n') e^{j2\pi nk/N}, \quad (14.6.1)$$

$$X(k) = \frac{1}{N} \sum_{k'=0}^{N-1} \sum_{n=0}^{N-1} F_X^w(k', n) w(k - k') e^{-j2\pi nk/N}, \quad (14.6.2)$$

by modifying the time basis $\{e^{-j2\pi nk/N}\}_{n=0}^{N-1}$ and frequency basis $\{e^{j2\pi nk/N}\}_{k=0}^{N-1}$ ($n, k = 0, \dots, N-1$), to the respective (t,f) bases $\{w(k - k') e^{-j2\pi nk/N}\}_{k',n=0}^{N-1}$ and $\{w(n - n') e^{j2\pi nk/N}\}_{n',k=0}^{N-1}$, where w is the smooth window function (close to one near the origin and decaying toward zero at the edges) and both F_x^w and F_X^w are TFDs of the signal \mathbf{x} . Equation (14.6.1) is the classical short-term Fourier transform (STFT) (cf. Section 2.3.1) and Eq. (14.6.2) is the S-transform (Section 2.7.6.2 and [44]). The main difference of the formulation equation (14.6.2) over Eq. (14.6.1) is that it is performed in the frequency domain and thus the window width can be tuned as a function of frequency. In the original S-transform, the window is Gaussian, with variance inversely proportional to frequency [44].

⁰Author: **A. Gholami**, Institute of Geophysics, University of Tehran, Iran (agholami@ut.ac.ir). Reviewers: I. Djurović, E. Sejdić, and S. Ventosa.

The main problem, however, is determination of the coefficients \mathbf{F}_x^w or \mathbf{F}_X^w . Let $\mathcal{F} \in \mathbb{C}^{N^2 \times N^2}$ be a block diagonal matrix with the DFT matrix as its diagonal blocks. Let $\mathcal{W} \in \mathbb{R}^{N \times N^2}$ be a row block matrix constructed from the window functions, so that its i th block is a diagonal matrix constructed from the window corresponding to the i th time/frequency sample. Then Eqs. (14.6.1) and (14.6.2) can be written in matrix-vector forms as $\mathbf{x} = \mathcal{W}\mathcal{F}^H \text{vec}(\mathbf{F}_y^w)$ and $\mathbf{X} = \mathcal{W}\mathcal{F} \text{vec}(\mathbf{F}_X^w)$, respectively, where the vec operator vectorizes a matrix by stacking its columns one under the other. Therefore, the linear TFR in its general form reads as

$$\rho = \Omega \text{vec}(\mathbf{F}_\rho^w) \quad (14.6.3)$$

(see [45] for more details), where ρ is an N -length vector (real or complex), $\Omega = \mathcal{W}\mathcal{F}$ is the inverse TFR operator, and \mathbf{F}_ρ^w is the TFR of ρ . The matrix Ω , being of size $N \times N^2$, has no inverse; when it is of full rank, it has a nullspace of dimension $N(N - 1)$, so that the solution \mathbf{F}_ρ^w to Eq. (14.6.3) has $N(N - 1)$ degrees of freedom, making it highly nonunique. Regularization tools can be employed to determine the desired TF map. The minimal-norm criterion leads to the closed-form solution $\text{vec}(\mathbf{F}_\rho^w) = \Omega^H(\Omega\Omega^H)^{-1}\rho$, which exists whenever Ω has full rank (H denotes the Hermitian transpose). From the viewpoint of frame theory, Ω^H constructs a dirty frame, and $\Omega\Omega^H = \mathcal{W}\mathcal{F}\mathcal{F}^H\mathcal{W}^H = \mathcal{W}\mathcal{W}^H \approx \mathbf{I}_N$. The conjugate inverse Ω^H can also be constructed from a dual window function \tilde{w} so as to make $\Omega\Omega^H = \mathbf{I}_N$. In this case, the minimal-norm solution is simply $\text{vec}(\mathbf{F}_\rho^w) = \Omega^H\rho = \mathcal{F}^H\mathcal{W}^T\rho$; however, despite this simplicity, the range of Ω^H is just a small subset of the solution space. Thus it may not be possible to determine the desired TFD for a given signal via Ω^H , whatever the chosen window may be. Figure 14.6.1 shows the relation between the signal space \mathbb{R}^N , its (t, f) space \mathbb{C}^{N^2} , Ω , and Ω^H . It can be shown that $\mathcal{R}(\Omega^H) \perp \mathcal{N}(\Omega)$ and $\mathcal{R}(\Omega^H) + \mathcal{N}(\Omega) = \mathbb{C}^{N^2}$, where \mathcal{R} and \mathcal{N} denote the range and null space, respectively; that means any possible TFD of the signal ρ can be written uniquely as

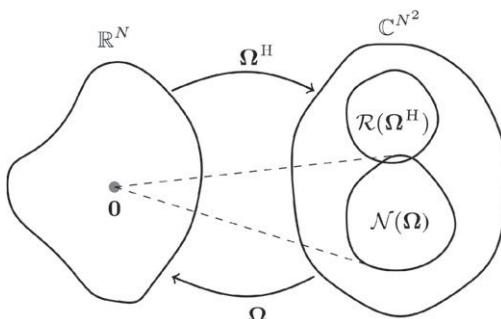


FIGURE 14.6.1

The relation between the signal space \mathbb{R}^N , (t, f) space \mathbb{C}^{N^2} , Ω , and Ω^H . Note that only a small subset of the solution space is accessible to the adjoint operator Ω^H and the TF space includes some nullspace, which is mapped to null vector in the signal space by Ω . Furthermore, $\mathcal{R}(\Omega^H) \perp \mathcal{N}(\Omega)$ and $\mathcal{R}(\Omega^H) + \mathcal{N}(\Omega) = \mathbb{C}^{N^2}$. Therefore, getting help from the elements in the nullspace is a way of improving the time-frequency representation (TFR) performance.

$\underline{\mathbf{F}}_{\rho}^w = \underline{\mathbf{F}}_{\rho}^w + \underline{\underline{\mathbf{F}}}_{\rho}^w$, where $\underline{\mathbf{F}}_{\rho}^w \in \mathcal{R}(\Omega^H)$ and $\underline{\underline{\mathbf{F}}}_{\rho}^w \in \mathcal{N}(\Omega)$. This is not a shortcoming but a feature, allowing the desired TFD to be created using appropriate regularizers.

14.6.2 SPARSITY INDUCED TIME-FREQUENCY DISTRIBUTIONS

The TFD generated by Ω^H has minimal energy (ℓ_2 -norm). This criterion leads to low resolution as it dislikes energy concentration and prefers solutions with many small nonzero coefficients. High-resolution TFDs can be obtained using other norms (measures) that favor energy concentration (sparsity). Generally, the resolution of the TFD of a given signal is governed by the chosen window and optimization procedure. Thus we have the optimization problem

$$\begin{aligned} & \underset{\mathcal{W}, \mathbf{F}_{\rho}^w}{\text{minimize}} \quad \varphi(\mathbf{F}_{\rho}^w) \\ & \text{subject to} \quad \boldsymbol{\rho} = \boldsymbol{\Omega} \text{vec}(\mathbf{F}_{\rho}^w), \end{aligned} \quad (14.6.4)$$

where φ is a sparsity-inducing functional. Because Eq. (14.6.4) is nonlinear in the window function, its solution is difficult and may be computationally expensive. One approach is to optimize over \mathcal{W} and \mathbf{F}_{ρ}^w separately, as follows.

14.6.2.1 Optimization over \mathcal{W}

For simplicity, to optimize Eq. (14.6.4) over the window function, we solve for \mathbf{F}_{ρ}^w with $\varphi(\mathbf{x}) = \|\mathbf{x}\|_2^2$ to obtain the closed form $\mathbf{F}_{\rho}^w = \mathcal{F}^H \mathcal{W}^T \boldsymbol{\rho}$. This reduces Eq. (14.6.4) to

$$\underset{\mathcal{W}}{\text{minimize}} \quad \varphi(\mathcal{F}^H \mathcal{W}^T \boldsymbol{\rho}), \quad (14.6.5)$$

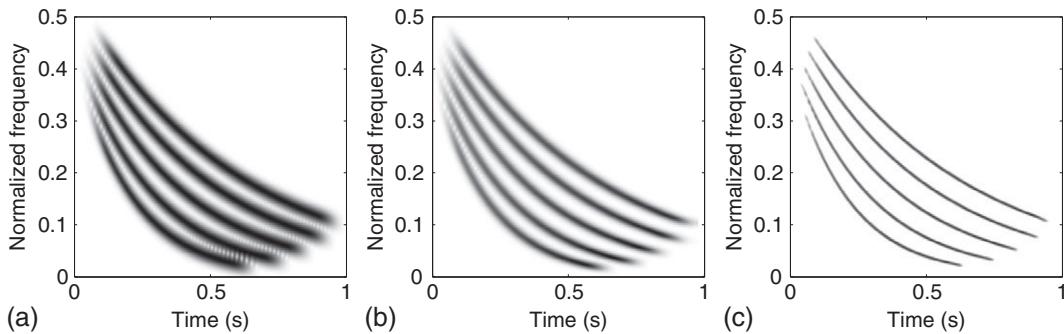
which has been widely used for adaptive determination of the window for conventional TFRs. Usually a known window type is chosen and its width is determined by Eq. (14.6.5) either globally [45] or instantaneously for each time/frequency sample [43,46]. Various sparsity functionals have been proposed (Section 7.3). One straightforward algorithm for solving Eq. (14.6.5) is to compute the function for a range of parameter values and choose the best value (Section 5.3). Gradient-based algorithms may also be used for the optimization (Section 7.3).

14.6.2.2 Optimization over \mathbf{F}_{ρ}^w

Having the optimized \mathcal{W} resulting from Eq. (14.6.5), the sparsity of \mathbf{F}_{ρ}^w can further be increased by employing sparse regularizers. The following convex problem has been proposed for this purpose, which also considers uncertainties in the signal [45]

$$\begin{aligned} & \underset{\mathbf{F}_{\rho}^w}{\text{minimize}} \quad \|\mathbf{F}_{\rho}^w\|_1 + \frac{\gamma}{2} \|\mathbf{F}_{\rho}^w\|_2^2 \\ & \text{subject to} \quad \|\boldsymbol{\rho} - \boldsymbol{\Omega} \text{vec}(\mathbf{F}_{\rho}^w)\|_2 \leq \varepsilon, \end{aligned} \quad (14.6.6)$$

where $\varepsilon \geq 0$ is a user-specified bound for the error in the signal, $\|\cdot\|_p$ is the entrywise ℓ_p -norm, and $\gamma \geq 0$ controls the amount of sparsity induced to the solution. Approaching γ to 0 leads to the minimal ℓ_1 -norm TFD, which can be over sparse with discontinuous structure [47]; and as $\gamma \rightarrow \infty$ the resulting

**FIGURE 14.6.2**

TFDs of a signal ρ simulated as a sum of five down-sweep chirps. Cases (a) and (b) are obtained by the conjugate inverse operator for (a) fixed window and (b) instantaneously optimized windows. The adaptive sparse distribution (c) is the solution of Eq. (14.6.6) for the windows used in (b).

TFD converges to a nonsparse one with minimum energy. A nonzero γ prevents the solution from being oversparse and allows adjustment of its resolution. Equation (14.6.6) provides an invertible transform to produce a desirable TFD. But it is nonlinear and its solution cannot be obtained in a closed form. Thus iterative algorithms must be employed. This is the cost we must pay for increasing the resolution; however, the problem is convex and can be solved by convex programming. A fairly efficient algorithm in Ref. [45] solves it with computational complexity ($N^2 \log(N)$) using the concepts of the alternating split Bregman technique [48]; however, other algorithms such as matching pursuit can also be employed to construct sparse TFDs (see Section 7.6).

Example 14.6.1. A signal ρ was simulated by superimposing five down-sweep chirps. Its TFD was computed by using the conjugate inverse operator and the sparse regularization (14.6.6). The results of the conjugate inverse operator, $\mathcal{F}^H \mathcal{W}^T \rho$, for a fixed Gaussian window and instantaneously optimized Gaussian windows are presented in Fig. 14.6.2(a) and (b), respectively. The optimization was performed by Eq. (14.6.5) for φ as the fourth power of the ratio of ℓ_4 and ℓ_2 norms (Section 5.3). Figure 14.6.2(c) shows the sparse distribution by Eq. (14.6.6) for $\varepsilon = 10^{-2}$, $\gamma = 0.1$, and window matrix \mathcal{W} used for Fig. 14.6.2(b). It can be seen that the sparsity constraint increased the resolution and concentration of the TFD significantly.

14.6.2.3 Sparsity for quadratic TFDs

Quadratic TFDs (see Chapters 2 and 3) have been developed for improving the TFR of nonstationary signals; however, they suffer from the cross-terms which, by masking the signal components, can make interpretation of the resulting TFR very difficult. Various methods have been introduced for reducing the effect of cross-terms (see Sections 3.1.2.3 and 4.2). The concept of sparsity can also be extended to the quadratic TFDs for suppressing cross-terms. Note that any quadratic TFD can be considered as a smoothed version of the Wigner-Ville distribution; i.e., it can be formulated as two-dimensional convolution of the signal WVD with a smoothing kernel (Eq. (3.2.13)). Therefore, the cross-terms of the WVD can be reduced by sparse deconvolution of quadratic TFDs using an appropriate kernel.

14.6.3 GEOPHYSICAL APPLICATIONS

Seismic exploration is a common method for prospecting resources of oil, natural gas, and minerals. Usually a short-duration seismic impulse is generated at or near the surface, propagated into the earth, reflected where geological properties change (boundaries), and then recorded by a reflection seismogram positioned at the surface. During propagation, the frequency content of the outgoing pulse is subjected to strong variations due to the subsurface absorption effects, the layers' thickness, and lateral or vertical variations in the pertophysical properties of the material such as velocity, porosity, density, pore-fluid, etc. Reflections are received at different times depending on the depth of the reflecting layers and the velocity of the subsurface, to form a nonstationary seismic trace. The recorded trace is then analyzed and interpreted to extract information about different properties of the subsurface. Due to the nonstationary nature of seismic signals, TFD tools gained significant attention for analysis of coherent ground roll noise attenuation [45,49], low-frequency anomalies detection [45,46,50], thin beds detection [46], and nonstationary deconvolution [51].

14.6.3.1 *Ground roll elimination*

Surface seismic data are subject to some random and coherent noise, which needs to be attenuated prior to further processing and interpretation. Ground roll, a source-generated surface wave, is the main type of coherent noise in land data, which severely obscures reflection events. It disperses during propagation and thus appears in a cone region on a time-vs.-offset display of the records (Fig. 14.6.3(a)). Due to the low-frequency and low-velocity characteristic of ground roll, a common approach for its filtering is to mask-related coefficients in the frequency-vs.-wave number (f - k) domain. The method has been extended by means of (t,f) tools in order to make the rejection dependent on time or position [49]. The sparse TFD method has been used to remove the ground roll noise from the field record shown in Fig. 14.6.3(a) (see Ref. [45] for more details). The record consists of 60 traces, each having 1000 time samples. The source has been at coordinate (0,0). The cone-shaped ground rolls are clearly seen in the

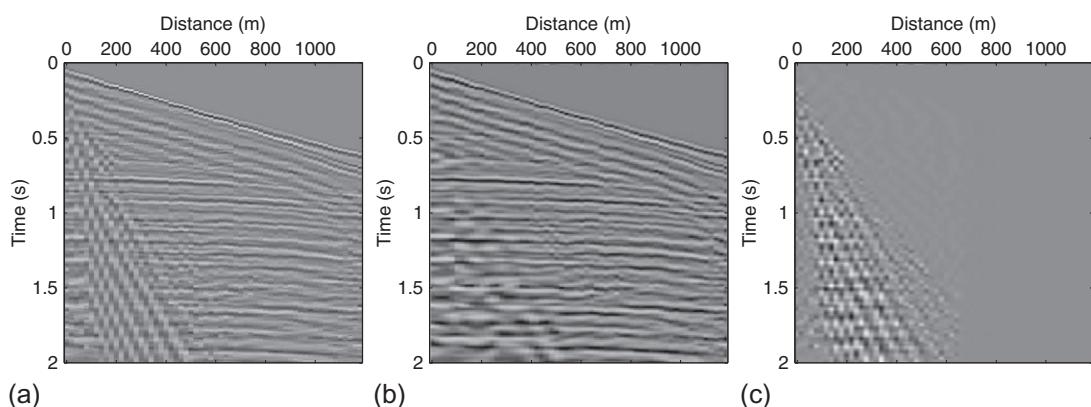


FIGURE 14.6.3

(a) Noisy field seismic signal; (b) filtered signal obtained by rejecting noise coefficients in the f - k - x domain using the sparse TFD method; (c) difference between (a) and (b).

traces near the source. The sparse TFD was applied along the distance direction with $\gamma = 0.1$, $\varepsilon = 10^{-8}$. Then, the DFT was applied in the time direction, followed by simply muting ground roll energy in the $f\text{-}k$ domain generated for each spacial coordinate. The inverse transformation of the resulting $f\text{-}k\text{-}x$ cube produced the section shown in Fig. 14.6.3(b) whose difference with the original record is shown in Fig. 14.6.3(c). As seen, the sparse TFD, due to its localized property and its invertibility, successfully eliminated such coherent noises from the signal while preserving useful reflection events.

14.6.3.2 Detection of hydrocarbon reservoirs

Time-frequency tools are also useful for locating reservoirs with seismic signals. Generally, the quality factor Q in gas reservoirs is low, so that high frequencies of a seismic pulse are attenuated more severely when propagating through a reservoir than when propagating through the surrounding media. Q is not the only factor that causes high-frequency loss; others include processing effects and high-amplitude multiple reflections from tops and bottoms of reservoirs [52]. It is also believed that *low-frequency shadows* occur for reflectors on seismic sections beneath gas reservoirs [45]. Another way is first to generate monofrequency sections around the dominant frequency corresponding to near- and far-offset data, and then to consider high-amplitude anomalies in the difference section as the location of laminated gas reservoirs. Figure 14.6.4 (top) shows field seismic sections corresponding to near- and far-offset traces. The difference amplitude sections corresponding to monofrequency 17 Hz sections, generated by the conjugate inverse operator Ω^H with fixed window, Ω^H with instantaneously optimized window

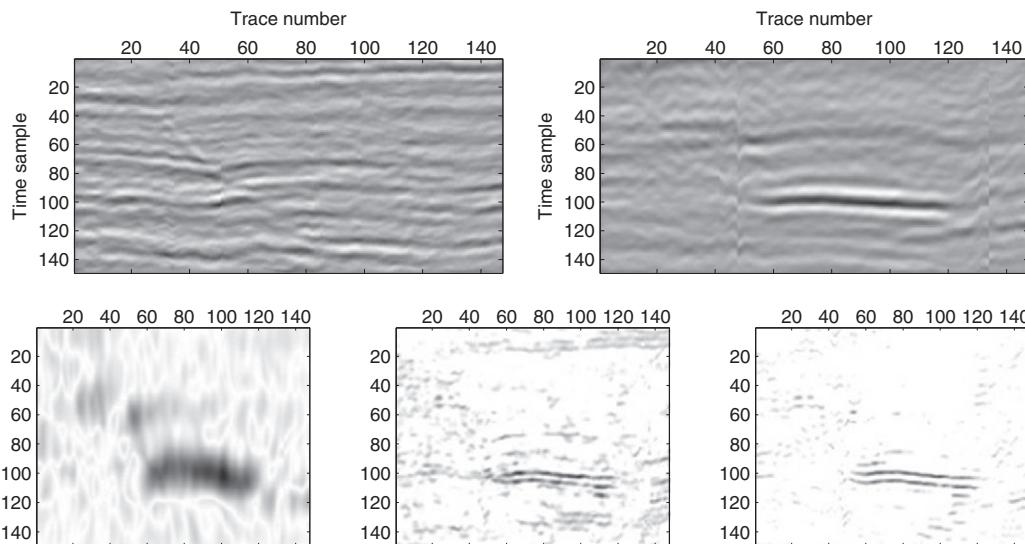


FIGURE 14.6.4

Top: Near offset (left) and far offset (right) field seismic data. *Bottom, left to right:* Difference amplitude sections corresponding to monofrequency 17 Hz generated by Ω^H with fixed window, Ω^H with instantaneously optimized window (14.6.5), and sparse TFD (14.6.6).

windows (14.6.5), and the sparse regularization (14.6.6), are also shown in the same figure (bottom). Observe that the sparse TFD, due to its high resolution and concentration, successfully resolved the thin-bedded reservoir in the section [46].

14.6.4 SUMMARY AND CONCLUSIONS

Linear discrete TFDs were presented using the concepts of linear algebra and inverse theory. The TFD for a given signal is nonunique with a large degree of freedom. The conventional conjugate inverse operator, despite its simplicity, provides a low-resolution distribution satisfying the minimal-energy constraint. But the range of the conjugate inverse operator is just a small subset of the solution space, so that the desired TFD may be inaccessible to it. Therefore, high performance (resolution, concentration, etc.) in a TFD is accessible only by appropriate regularization over the window function (see Sections 5.3 and 7.3) and (t,f) plane. The algorithm presented here can generate a very high-resolution TFD with high concentration for a given signal. But it is still suboptimal and could be further improved by simultaneous optimization of the window function and TFD (14.6.4).

14.7 AUDIO SPEECH AND UNDERWATER SIGNALS TIME-FREQUENCY CHARACTERISTIC ENHANCEMENT⁰

As audio signals are nonstationary, clarity of representation is essential for their TFDs, so that their relevant features can be read, selected, and extracted. In addition, real audio signals are affected by noise, which is troublesome for signal interpretation; for this reason, a signal-to-noise ratio (SNR) enhancement step is necessary. Desired information includes the distribution of the signal over the time-frequency (t,f) plane and the instantaneous frequency (IF) characterizing the variation of its frequency content over time. A general scheme to identify audio signals is described in Figure 14.7.1.

The objective of this section is to show the improvements of (t,f) analysis of speech and underwater acoustic signals using high-resolution (t,f) methods compared to standard methods like the spectrogram.

14.7.1 EXAMPLES OF AUDIO SIGNALS AND THEIR (t, f) CHARACTERISTICS

We will present in the next two sections the key characteristics for speech signals and underwater signals.

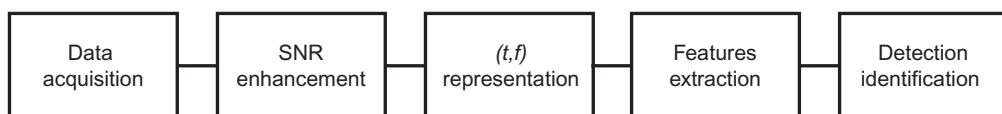


FIGURE 14.7.1

General scheme for audio signals identification

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Ouelha**, Qatar University, Qatar (samir.ouelha@qu.edu.qa). Reviewers: S. Ghaemmaghami and V. Chandran

14.7.1.1 Speech signals

Most speech signals are nonstationary processes with multiple components that may vary in time and frequency. Classical stationary methods are unable to represent these variations accurately, whereas (t,f) representations allow a more precise description of nonstationary signals. There are two useful acoustic features in a voiced-speech signal: fundamental frequency (pitch) and formant. The fundamental frequency is usually the lowest frequency component of the signal; it represents the vibration frequency of the vocal cords during sound production. The formant is a concentration of acoustic energy around a particular frequency in the speech wave; each formant corresponds to a resonance on the vocal tract. A model used to represent voiced speech signals is defined by Ref. [53]

$$s(t) = \sum_{k=1}^N s_k(t) \quad (14.7.1)$$

with $s_k(t) = a_k(t) \cos(2\pi[f_{ck}t + \int_0^t f_k(\tau) d\tau] + \theta)$, where f_{ck} is the central frequency, $f_k(t)$ is the frequency modulation, and $a_k(t)$ is time-varying amplitude. The instantaneous frequency (IF) of $s_k(t)$ is defined as $f_{ik}(t) = f_{ck} + f_k(t)$; this model is a sum of N amplitude modulation and frequency modulation (AM-FM) signals, where each signal represents one formant [54]. Fundamental and formant frequencies, represented by major peaks in the spectrum, convey important information about speech. In fact, a large amount of phonetic information is conveyed by the corresponding parts of voiced speech signals. Accordingly, formant detection and tracking are important in extracting speech features and in recognizing its evolutionary behavior. The representation of a voiced speech signal by the formant amplitude envelope and instantaneous frequency is rich, because it reveals both the spectral structure and the excitation timing information of different formant bands. Although there is basically an infinite number of formants in speech signals, four formants are sufficient to represent the vocal-tract characteristics [54]. Consequently, there is no need to detect and follow all speech frequency components in most applications. Traditionally, formant trackers use Linear Prediction or are based on the STFT [54]. Such formant trackers work well when the formants change slowly and gradually, but in the case of abrupt changes in formant trajectories, poor results are obtained. Generally, two spectrograms with different window lengths are used. In wideband spectrogram (small window length), the temporal resolution is good and one can see the periodicity of pitch, but we sacrifice the localization of the formant structures. The narrowband spectrogram improves the latter in terms of frequency resolution, but the time resolution is degraded. There is another difficulty associated with this type of methods; if two adjacent formants are too close in the (t,f) plane, this phenomenon may produce a single peak in the spectrum during certain intervals, leading to missing one of the existing formants. There are two possible ways to overcome the limitations of conventional spectrograms. The first approach is to use high-resolution (t,f) methods (see Chapters 2 and 3), while the second approach is to use some postprocessing methods like reassignment (see Section 7.5).

14.7.1.2 Underwater acoustic signals

An underwater acoustic signal is obtained using devices, such as a SONAR device that can supply the audio record of the acoustic field activity in a given direction. Underwater acoustic signals are used with many applications, such as acoustic communications, detection and localization of surface and subsurface objects, depth sounders, sub-bottom profilers, and other applications. The identification and recognition of signals have become an important issue. Conventionally, human experts perform

identification by both listening to and visual inspection of (t, f) representations; however, the increasing complexity of sonar arrays means that manual classification is often impractical; there is, therefore, a real need for reliable automatic classification. Underwater signals are generally a mixture of several sounds, which include biological sounds (dolphins, shrimps, whales, etc.) with environmental sounds (rain, ice cracking) and man-made sounds (torpedoes, surface ships), which makes the problem of recognition very difficult, because of the signal shape variability. The second problem is that for the same source, the signature in the (t, f) domain may not be the same because of the marine environment. An approach is to use (t, f) representations inspired by human physiology for processing signals. In Ref. [55], the authors developed a specific process for underwater acoustic signal representations, with human perceptions considered. This approach allows one to obtain a denoised (t, f) representation with a frequency resolution based on human perception. This transformation is a sonograph (see Section 2.4) based on the use of Mel filterbanks, coupled with a wavelet shrinkage method. Mel filterbanks are composed of triangular filters spaced according to Mel's scale, defined by the following formula for a given frequency f in Hz:

$$\text{mel}(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right).$$

Thus Mel's frequency scale is linear below 1 kHz and logarithmic above. If we consider M Mel filters, $H_{\text{Mel}}(f; m)$, each one centered around a frequency f_m with $m = 1, 2, \dots, M$, and has a bandwidth $B(m)$ defined as follows:

$$B(m) = f_{m+1} - f_{m-1}, \quad \forall m = 2, 3, \dots, M-1.$$

The center frequency f_m is computed from its corresponding center frequency on Mel's scale using the following inverse formula:

$$f_m = 700 \left(10^{\frac{\text{mel}(f_m)}{2595}} - 1 \right)$$

with

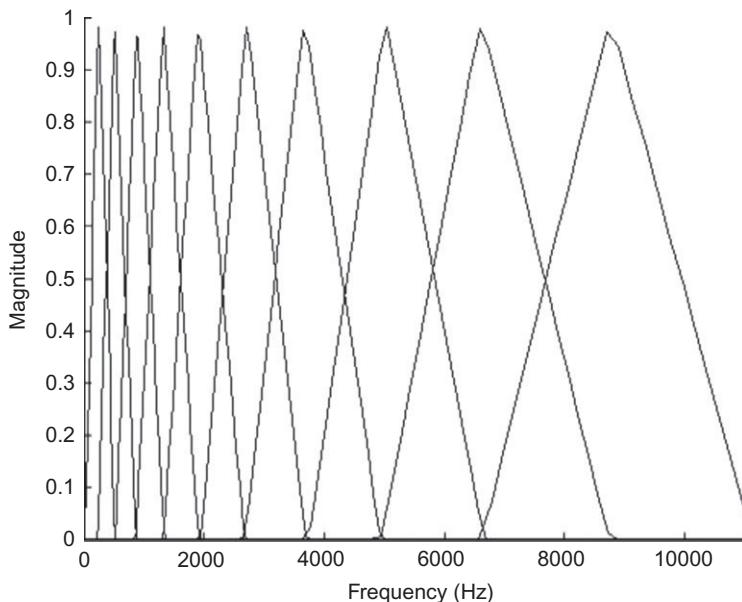
$$\text{mel}(f_m) = \frac{m}{M+1} (\text{mel}(f_{\max}) - \text{mel}(f_{\min})),$$

where f_{\max} and f_{\min} correspond, respectively, to the highest and lowest frequency components of the studied signal (generally $f_{\min} = 0$ and $f_{\max} = \frac{F_s}{2}$, where F_s is the sampling frequency). The sonograph equation defined in Section 2.4 corresponds to a uniform filterbank; however, in the case of Mel filterbanks, the impulse response of each filter depends on the analyzed frequency. The sonograph based on the use of Mel filterbank can be described by

$$S_s^H(t, f) = \left(\frac{f_m}{f} \right) \left| \int S(v) H \left(v \left(\frac{f_m}{f} \right) \right) e^{j2\pi vt} dv \right|^2.$$

Figure 14.7.2 shows an example of a Mel filterbank composed of 10 filters.

The advantage of this method is to use a bank of filters with different resolutions, which is better adapted to the perceptual interpretation of audio signals, because it is well known that the human ear

**FIGURE 14.7.2**

An example of a Mel filterbank composed of 10 filters, $f_{\min} = 0$ and $f_{\max} = 11,025$ Hz.

acts like a filterbank, concentrated only on certain frequency components. That is why Mel filters are nonuniformly spaced on the frequency axis, with more filters on the low frequency regions and less on high frequency regions.

The key part of automatic recognition of underwater acoustic signals involves features extraction; various approaches have been described, including time waveform analysis, frequency spectrum analysis, wavelet transformation, and nonlinear signal processing (Section 12.6). Features based on IF are useful for classification, especially for the differentiation between mechanical and biological signals [15]. We can also use (t,f) image-processing techniques to compute morphological features to characterize the shape, position, and direction of each (t,f) component. This approach processes the (t,f) plane of a given signal considered as an image (see Section 12.6). In the context of underwater signals, texture features such as statistics of spatial grayscale variations of the TFDs, entropy, energy, and other features based on gray level co-occurrence matrix are useful for classification; these features are described in Section 12.6 and in Ref. [6].

14.7.2 PREPROCESSING OF AUDIO SIGNALS FOR SNR ENHANCEMENT

It is desired to remove noise so as to enhance the useful signal before recognition in order to better represent the (t,f) features that characterize the signal and improve the overall performance of the classification system. In developing a noise reduction technique, keeping the structure of the original signal is one of the main concerns. In real audio signals, noise is nonstationary; therefore, the signal

enhancement process becomes more challenging. In such cases, (t,f) distributions can be used in noise attenuation along with keeping the structure of the original signal unchanged. A number of methods has been proposed for denoising audio signals, including

- spectral subtractive algorithms and Wiener filtering [56]; these algorithms are simple to implement but suffer for musical noise distortion;
- statistical model-based algorithms for speech enhancement, including the maximum likelihood estimator, minimum mean square error (MMSE), and maximum *a posteriori* estimators [57];
- subspace algorithms (see Section 11.5) based on the principle that the vector space of the noisy signal can be decomposed into signal and noise subspaces;
- independent component analysis, in which the components of the desired representation have minimal stochastical dependence in order to capture the essential structure of the data;
- (t,f) image processing techniques for the removal of noise from a signal; for example, in Ref. [58], an adaptive block thresholding algorithm is proposed to remove noise from audio signals in (t,f) representations.

The performance of these methods is application-dependent. In the case of underwater signals, it is affected by ocean condition and ambient noise during transmission. The sources of ambient noise are both natural and human-made, with different sources exhibiting different directional and spectral characteristics. Therefore, before recognizing the received acoustic signals, it is necessary to remove the noise as much as possible so as to preserve the important signal features.

14.7.3 RESULTS FOR SPEECH AND UNDERWATER ACOUSTIC SIGNALS

The selection of a suitable (t,f) method for representing audio signals is the first key step for any classification system. A suitable (t,f) method is one that is capable of highlighting the signal nonstationary features that best discriminate among different classes under consideration. The best TFD to use for audio signals needs to be adapted to its characteristics; when the signals have an IF parallel to the time axis in the (t,f) domain, lag-independent TFDs, like the MBD, are suitable. In real underwater signals, there is a train of impulses called echolocation signals, for which the IF is parallel to the frequency axis in the (t,f) domain; Doppler-independent TFDs, like the windowed WVD, are suited for this type of signal. Finally, for other cases, we can use separable kernels, like the B-distribution, the EMBD and CKD (see Section 3.3 and [15], because these methods employ smoothing filters along both time and frequency axes).

14.7.3.1 Results of simulated signals

Let us compare the performance of some (t,f) methods in terms of their ability to accurately concentrate the energy of signal components along the corresponding IF curves. Let us then model a speech signal by a mixture of two sinusoidal FM signals, expressed as

$$s(t) = \begin{cases} s_1(t) + s_2(t), & \text{if } 0 \leq t \leq T, \\ 0, & \text{otherwise,} \end{cases}$$

where the signal components have the characteristics shown below:

$$\Phi(t) = 2\pi \frac{f_1 + f_0}{2}(t - t_0) + \frac{f_1 - f_0}{2}T \left(\sin\left(\frac{2\pi t}{T} + \phi\right) - \sin(\phi) \right)$$

with

$$\phi = \arccos\left(\frac{(2f_{\text{ref}} - (f_1 + f_0))}{f_1 - f_0}\right),$$

$$s_1(t) = e^{j\Phi(t)},$$

where $f_0 = 0.1, f_1 = 0.2, T = 256, t_0 = 128, f_{\text{ref}} = 0.1$

$$s_2(t) = e^{j\Phi(t)},$$

where $f_0 = 0.18, f_1 = 0.28, T = 256, t_0 = 128, f_{\text{ref}} = 0.18$.

The IFs of the signal components (see Fig. 14.7.3) are described by

$$f(t) = f_{\text{ref}} + \frac{f_1 - f_0}{2} \cos\left(\frac{2\pi t}{T} + \phi\right).$$

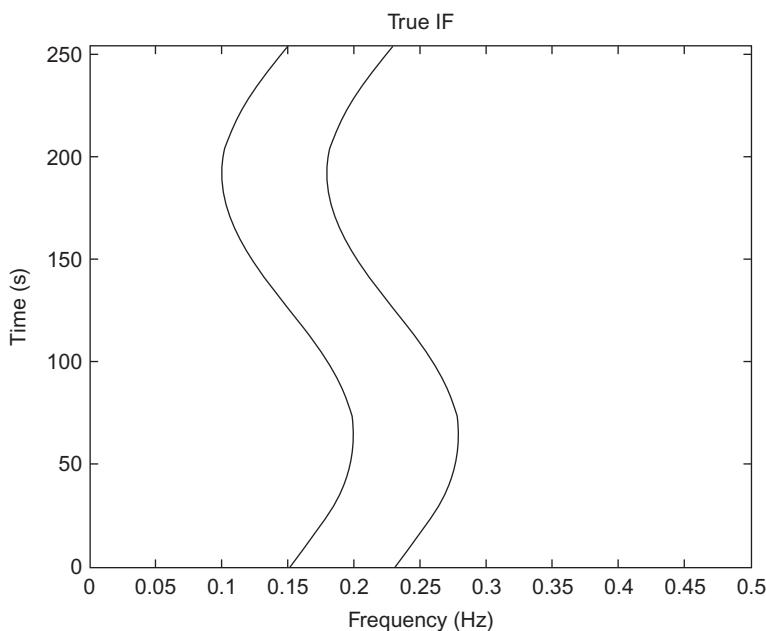


FIGURE 14.7.3

IF law of signal $s(t)$.

This multicomponent signal is sampled at 1 Hz. The signal is analyzed using compact kernel distribution (CKD), multidirectional distribution (MDD) [15], and the spectrogram with two different window lengths (wideband and narrowband) (see Fig. 14.7.4). The parameters of the TFDs have been optimized based on visual inspection to maximize their energy concentration and resolution properties. White Gaussian noise is added to the signal. The IF of the signal is estimated using the connected component-linking algorithm (see Section 10.6), for SNR ranging from -3 to 15 dB. The mean square error (MSE), between the estimated IF and the original IF is used to evaluate the accuracy of IF estimates. The MSE is estimated using 200 Monte Carlo simulations. The following observations can be made regarding the results of IF estimates that are shown in Fig. 14.7.5:

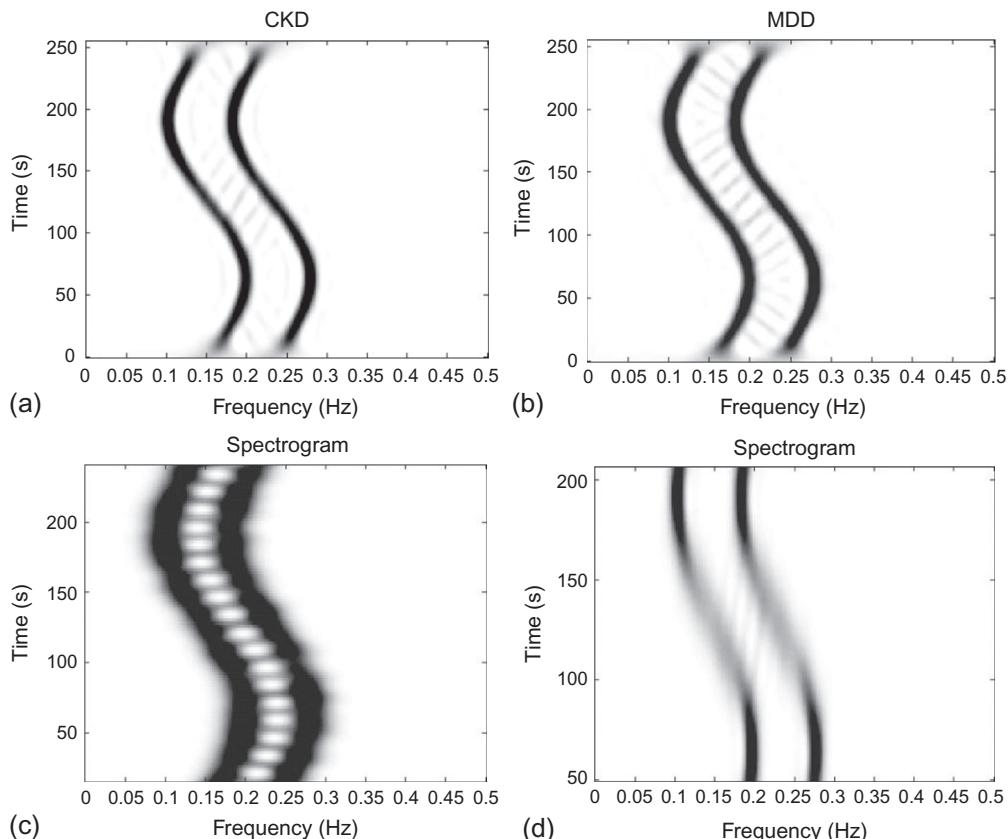
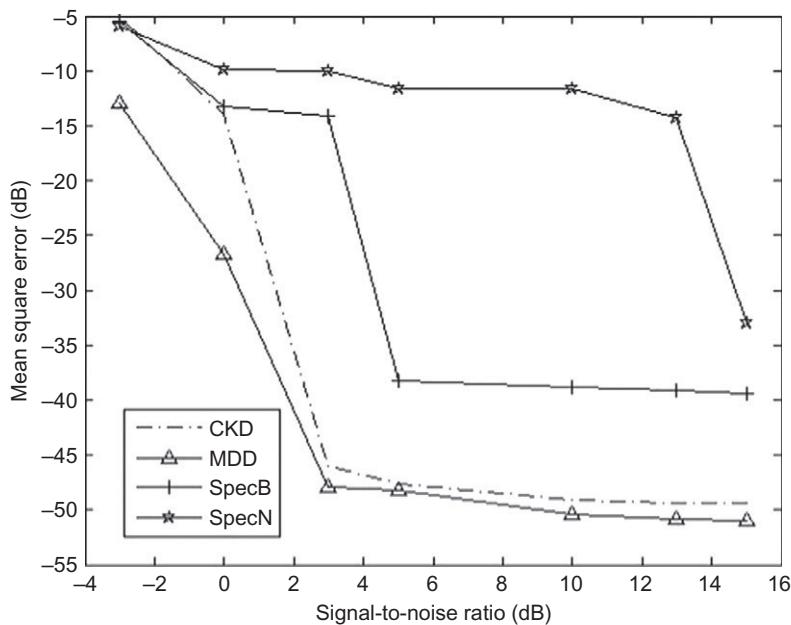


FIGURE 14.7.4

Comparison of (t, f) methods for signal $s(t)$ composed of two sinusoidal FM components: (a) the CKD ($c = 1, D = 0.12, E = 0.075$); (b) the MDD($\theta_1 = 60^\circ, \theta_2 = 120^\circ$); (c) spectrogram (Hanning window length = 31); (d) spectrogram (Hanning window length = 99).

**FIGURE 14.7.5**

Mean square error of the estimates for the multicomponent signal.

- The MDD is the best-performing TFD and allows a good separation of the close components at high SNR; the CKD also gives good results, but not as good as the MDD. The MDD kernel performs better, because it can follow the directions of the auto-terms.
- The spectrogram gives poor results in the two cases. In one case, the components are well separated, but the IF law is not well estimated, whereas in the second case, the variations of IF law are well estimated, but the components are not separated, because of the poor energy concentration.
- All methods fail to give accurate estimates of IF laws for a low SNR (from 0 dB). This shows the importance of using a preprocessing SNR enhancement step.

14.7.3.2 Results of real speech signals

This experiment compares the CKD and the spectrogram of a real speech signal that is a voiced female speech with a sampling frequency equal to 16 kHz. We are focused only on the band [0;1000 Hz] in order to illustrate the benefit of high-resolution (t,f) method. The noise reduction method described in Ref. [57] is applied.

Figure 14.7.6 shows the benefit of the CKD compared to the spectrogram. The CKD in Fig. 14.7.6(b) provides better results compared to the spectrogram in Fig. 14.7.6(a) shown in terms of energy concentration and resolution. A high-resolution TFD is required when the signal has close components in the (t,f) domain and/or in the case of high variations in the IF law of its components, which is the case with many real signals.

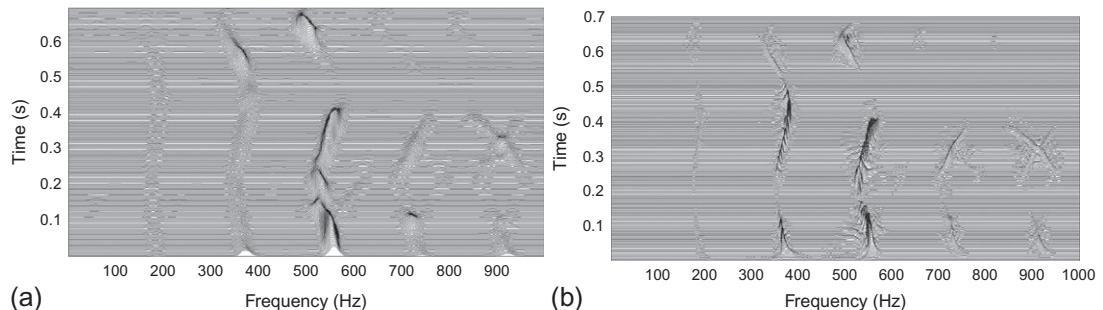


FIGURE 14.7.6

Comparison of (t, f) methods for a denoised real voiced speech signal: (a) spectrogram (Hanning window length = 255); (b) the CKD ($c = 0.1$, $D = 0.12$, $E = 0.03$), scale in dB with application of a threshold max-30 dB.

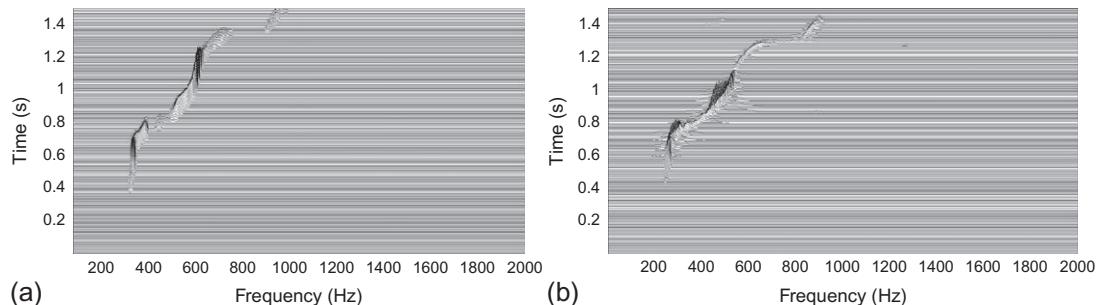


FIGURE 14.7.7

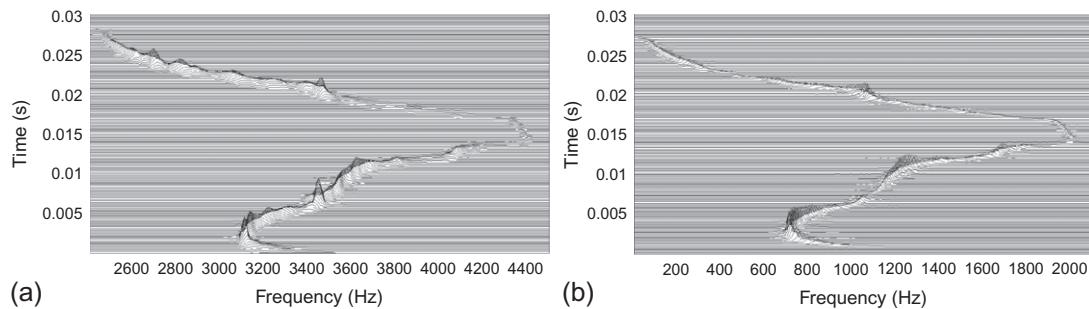
Comparison of (t, f) methods for a humpback whale vocalization signal: (a) spectrogram (Hanning window length = 255); (b) the EMBD ($\alpha = 0.01$, $\beta = 0.01$), scale in dB with application of a threshold max-30 dB.

14.7.3.3 Results of real underwater acoustic signals

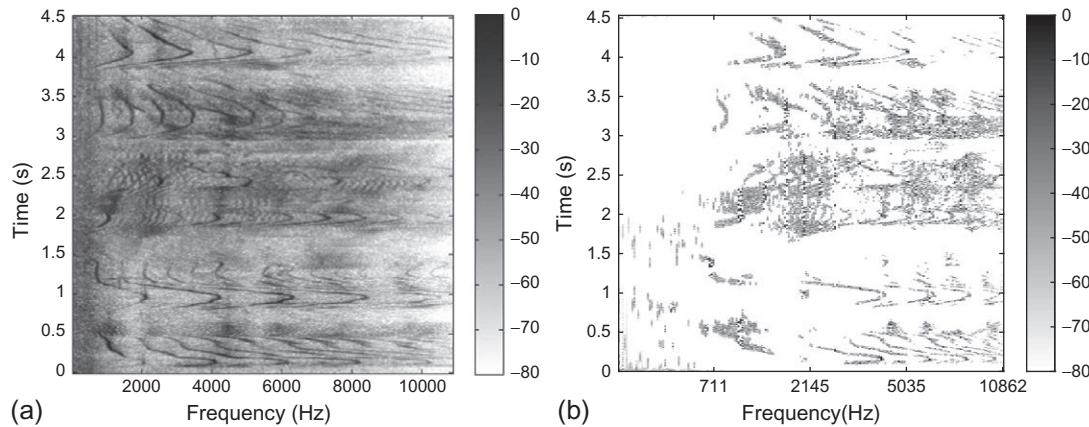
To assess the benefit of high-resolution (t, f) methods for underwater signals, we compare the EMBD with the spectrogram. The first signal is a humpback whale song and the second signal is a beluga whale song.

In both cases, we can see a better (t, f) representation when we use high-resolution (t, f) methods. The EMBD in Fig. 14.7.7(b) and the CKD in Fig. 14.7.8(b) give high energy concentration compared to the spectrogram (see Figs. 14.7.7(a) and 14.7.8(a)). For the second signal, the CKD gives good energy concentration because of the flexibility to adjust both shape and size of the kernel smoothing independently.

Now, let us compare two (t, f) representations of a signal representing a dolphin communication. If we look closer at Fig. 14.7.9(a), at around $t = 2$ s and around $f = 2$ kHz, we notice the richness of this record, showing a strong spectral entanglement. In the case of spectrogram, the noise is making automatic identification difficult. Whereas on the denoised Hearingogram [55] (see Fig. 14.7.9(b)),

**FIGURE 14.7.8**

Comparison of (t, f) methods for Beluga whale signal: (a) spectrogram (Hanning window length = 255); (b) the CKD ($c = 1, D = 0.012, E = 0.03$), scale in dB with application of a threshold max-30 dB

**FIGURE 14.7.9**

Comparison between spectrogram (linear f scale) (a) and sonograph based on Mel filterbank (nonlinear f scale); (b) denoised Hearingogram applied on signal of dolphin communications, scale in dB with application of a threshold max-80 dB.

only patterns of interest are preserved. These patterns show a strong spectral entanglement. Thus the interpretation of the signal is easier and allows us to conclude that it represents a communication among several dolphins. As the denoised Hearingogram is based on human physiology, we can extract auditory features from this domain.

14.7.4 SUMMARY AND CONCLUSION

This section links important audio signal characteristics to their (t, f) features. Underwater and speech signals are considered together because conventionally, human experts perform identification of

underwater signal by both listening and visual inspection of (t,f) representation. The various examples presented for both speech and underwater signal indicate that high-resolution (t,f) methods can improve the performance of audio signal analysis. They allow accurate IF estimation and the extraction of features, which are more precise, like formant and fundamental frequencies for the speech signal. An SNR enhancement step is often necessary, especially for underwater signals, which generally have a low SNR. Other useful features, like morphological or statistical features from the (t,f) domain, could be used to improve classification performance. Moreover, the superior performance of experienced human listeners in sound-event classification suggests that some advantage can be gained from a perceptually motivated approach, which involves the use of auditory features [32]. Thus a combination of perceptual features and (t,f) features such as IF and other image-processing methods can be used to characterize audio signals. The high-resolution (t,f) methods discussed above need to be implemented with care, noting that

- These methods have longer computation time and memory storage issue compared to the spectrogram. So, for the analysis of audio signals, we recommend a segmentation step using the spectrogram and then applying high-resolution (t,f) methods on each segment (see [Section 6.6](#) for algorithms to reduce computation and memory).
- For high-resolution methods, the number of parameters to optimize is larger compared to the spectrogram. The parameters can be selected using some energy concentration/resolution measures (see [Section 7.3](#)). Another approach is to use adaptive methods (see [Section 5.1](#)) where kernel shape is automatically optimized based on signal characteristics. For example, the adaptive optimal kernel (AOK), described in [Section 5.3](#), has been used successfully on man-made underwater acoustic (see example in [Section 14.5](#)).

Finally, we described the benefit of integrating an SNR enhancement stage in a (t,f) representation. This representation is based on auditory model (Mel filterbank), so it allows us to directly extract some perceptual features.

REFERENCES

- [1] T.A.C.M. Claasen, W.F.G. Mecklenbräuker, “The Wigner distribution—a tool for time-frequency signal analysis”, *Philips J. Res.* 35 (1980) 217-250 (Part 1), 276-300 (Part 2), 372-389 (Part 3).
- [2] S.L. Marple Jr., “Two-dimensional lattice linear prediction parameter estimation method and fast algorithm”, *IEEE Signal Process. Lett.* 7 (6) (2000) 164-168.
- [3] A. Jakobsson, S.L. Marple Jr., P. Stoica, “Computationally efficient two-dimensional Capon spectrum analysis”, *IEEE Trans. Signal Process.* 48 (9) (2000) 2651-2661.
- [4] S.L. Marple Jr., *Digital Spectral Analysis with Applications*, Prentice-Hall, Upper Saddle River, NJ, 1987.
- [5] V.C. Chen, *The Micro-Doppler Effect in Radar*, Artech House, Norwood, MA, 2011.
- [6] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [7] W.G. Carrara, R.S. Goodman, R.M. Majewski, *Spotlight Synthetic Aperture Radar*, Artech House, Norwood, MA, 1995.
- [8] S. Barbarossa, “Detection and imaging of moving objects with synthetic aperture radar—part 1: optimal detection and parameter estimation theory”, *IEEE Proc. F Radar Signal Process.* 139 (1) (1992) 79-88.

- [9] F. Auger, P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Process.* 43 (5) (1995) 1068-1089.
- [10] S. Barbarossa, O. Lemoine, "Analysis of nonlinear FM signals by pattern recognition of their time-frequency representation", *IEEE Signal Process. Lett.* 3 (4) (1996) 112-115.
- [11] B. Barkat, B. Boashash, "Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: statistical performance in the presence of additive Gaussian noise", *IEEE Trans. Signal Process.* 47 (9) (1999) 2480-2490.
- [12] S. Barbarossa, A. Scaglione, G.B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling", *IEEE Trans. Signal Process.* 46 (3) (1998) 691-708.
- [13] S. Barbarossa, A. Scaglione, "Autofocusing of SAR images based on the product high-order ambiguity function", *IEEE Proc. Radar Sonar Navigat.* 145 (5) (1998) 269-273.
- [14] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals; part 2: algorithms and applications", *Proc. IEEE* 80 (4) (1992) 519-568.
- [15] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digit. Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [16] B.G. Ferguson, B.G. Quinn, "Application of the short-time Fourier transform and the Wigner-Ville distribution to the acoustic localization of aircraft", *J. Acoust. Soc. Amer.* 96 (1994) 821-827.
- [17] B.G. Ferguson, "Time-frequency signal analysis of hydrophone data", *IEEE J. Oceanic Eng.* 21 (4) (1996) 537-544.
- [18] B.G. Ferguson, K.W. Lo, "Transiting aircraft parameter estimation using underwater acoustic sensor data", *IEEE J. Oceanic Eng.* 24 (4) (1999) 424-435.
- [19] K.W. Lo, B.G. Ferguson, "Passive estimation of aircraft motion parameters using destructive interference between direct and ground-reflected sound waves", in: *Proc. Information Decision and Control 99*, IEEE, 1999, pp. 171-176.
- [20] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 2: algorithms and applications", *Proc. IEEE* 80 (4) (1992) 540-568.
- [21] K.W. Lo, S.W. Perry, B.G. Ferguson, "Aircraft flight parameter estimation using acoustical Lloyd's mirror effect", *IEEE Trans. Aerospace Elect. Syst.* 38 (1) (2002) 137-151.
- [22] G.J. Frazer, "Project DUNDEE results: high frequency line-of-sight radar measurements and analysis", Tech. Rep. DSTO-TR-0734, Defence Science and Technology Organisation, 1998, aR 010-659.
- [23] B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992.
- [24] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals", *Proc. IEEE* 80 (4) (1992) 520-538.
- [25] I. Djurović, L. Stanković, "Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions", *IEEE Signal Process. Lett.* 7 (11) (2000) 317-319.
- [26] B. Boashash, B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise", *Signal Process.* 67 (1) (1998) 1-23.
- [27] J.G. Lourens, "Passive sonar detection of ships with spectro-grams", in: *Proc. 3rd South African IEEE Conference on Theoretical and Practical Work in Communications and Signal Processing*, 1990, pp. 147-151.
- [28] B. Boashash, P.J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 38 (11) (1990) 1829-1841.
- [29] B. Boashash, G. Azemi, "A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG", *Digit. Signal Process.* 28 (2014) 28-38.

- [30] W.L.J. Fox, J.C. Luby, J.W. Pitton, P.J. Loughlin, L.E. Atlas, "Sonar and radar range-Doppler processing using a cone-shaped kernel time-frequency representation", in: Proc. 24th Asilomar Conf. on Signals, Systems, and Computers, vol. 2, 1990, pp. 1079-1083.
- [31] B. Boashash, G. Frazer, "Time-varying higher-order spectra, generalised Wigner-Ville distribution and the analysis of underwater acoustic data", in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'92), vol. 5, 1992, pp. 193-196.
- [32] S. Tucker, G.J. Brown, "Classification of transient sonar sounds using perceptually motivated features", *IEEE J. Ocean. Eng.* 30 (3) (2005) 588-600, <http://dx.doi.org/10.1109/JOE.2005.850910>.
- [33] P. Courmontagne, S. Ouelha, U. Moreaud, F. Chaillan, "A blind denoising process with applications to underwater acoustic signals", in: IEEE Proc. Oceans, IEEE, San Diego, 2013, pp. 1-7.
- [34] S. Kadambe, T. Adali, "Application of cross-term deleted Wigner representation (CDWR) for sonar target detection/classification", in: Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers, vol. 1, 1998, pp. 822-826.
- [35] F. Lari, A. Zakhori, "Automatic classification of active sonar data using time-frequency transforms", in: Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis, 1992, pp. 21-24.
- [36] Q.Q. Huynh, L.N. Cooper, N. Intrator, H. Shouval, "Classification of underwater mammals using feature extraction based on time-frequency analysis and BCM theory", *IEEE Trans. Signal Process.* 46 (5) (1998) 1202-1207.
- [37] S.L. Marple Jr., T. Brotherton, "Detection and classification of short duration underwater acoustic signals by Prony's method", in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'91), vol. 2, 1991, pp. 1309-1312.
- [38] N.L. Gerr, "Introducing a third-order Wigner distribution", *Proc. IEEE* 76 (3) (1988) 290-292.
- [39] V. Chandran, S. Elgar, A. Nguyen, "Detection of mines in acoustic images using higher order spectral features", *IEEE J. Ocean. Eng.* 27 (3) (2002) 610-618.
- [40] V. Chandran, B. Carswell, B. Boashash, S. Elgar, "Pattern recognition using invariants defined from higher order spectra: 2-D image inputs", *IEEE Trans. Image Process.* 6 (5) (1997) 703-712, ISSN 1057-7149, <http://dx.doi.org/10.1109/83.568927>.
- [41] Z.-H. Michalopoulou, "Underwater transient signal processing: marine mammal identification, localization, and source signal deconvolution", in: Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97), vol. 1, 1997, pp. 503-506.
- [42] N. Ma, D. Vray, P. Delachartre, G. Gimenez, "Sea-bottom backscattering modeling with a wideband constant beamwidth sonar at normal incidence", in: Proc. IEEE Ultrasonics Symposium, vol. 2, 1995, pp. 1077-1080.
- [43] I. Djurović, E. Sejdić, J. Jiang, "Frequency-based window width optimization for S-transform", *Internat. J. Electr. Commun.* 62 (4) (2008) 245-250.
- [44] R.G. Stockwell, L. Mansinha, R.P. Lowe, "Localization of the complex spectrum: the S transform", *IEEE Trans. Signal Process.* 44 (4) (1996) 998-1001.
- [45] A. Gholami, "Sparse time-frequency decomposition and some applications", *IEEE Trans. Geosci. Remote Sens.* 51 (6) (2013) 3598-3604.
- [46] H. Sattari, A. Gholami, H.R. Siahkoohi, "Seismic data analysis by adaptive sparse time-frequency decomposition", *Geophysics* 78 (5) (2013) V207-V217.
- [47] P. Flandrin, P. Borgnat, "Time-frequency energy distributions meet compressed sensing", *IEEE Trans. Signal Process.* 58 (6) (2010) 2974-2982.
- [48] T. Goldstein, S. Osher, "The split Bregman method for L1 regularized problems", *SIAM J. Imaging Sci.* 2 (2) (2009) 323-343.
- [49] R. Askari, H.R. Siahkoohi, "Ground roll attenuation using the S and x-f-k transforms", *Geophys. Prospect.* 56 (1) (2008) 105-114.
- [50] G. Liu, S. Fomel, X. Chen, "Time-frequency analysis of seismic data using local attributes", *Geophysics* 76 (6) (2011) P23-P34.

- [51] G.F. Margrave, M.P. Lamoureux, D.C. Henley, "Gabor deconvolution: estimating reflectivity by nonstationary deconvolution of seismic data", *Geophysics* 76 (3) (2011) W15-W30.
- [52] D. Ebrom, "The low-frequency gas shadow on seismic sections", *Leading Edge* 23 (8) (2004) 772-772.
- [53] G. Degottex, P. Lachantin, A. Roebel, X. Rodet, "Mixed source model and its adapted vocal tract filter estimate for voice transformation and synthesis", *Speech Commun.* 55 (2) (2013) 278-294.
- [54] S. Ghaemmaghami, M. Deriche, B. Boashash, "Formant detection through instantaneous-frequency estimation using recursive least square algorithm", in: Proc. Fourth Internat. Symp. on Signal Processing and its Applications (ISSPA'96), vol. 1, IEEE, 1996, pp. 81-84.
- [55] P. Courmontagne, S. Ouelha, F. Chaillan, "A new time-frequency representation for underwater acoustic signals: the denoised hearingogram", in: OCEANS-Bergen, 2013 MTS/IEEE, IEEE, 2013, pp. 1-6.
- [56] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction", *IEEE Trans. Acoust. Speech Signal Process.* 27 (2) (1979) 113-120.
- [57] I. Cohen, "Relaxed statistical model for speech enhancement and a priori SNR estimation", *IEEE Trans. Speech Audio Process.* 13 (5) (2005) 870-881.
- [58] G. Yu, S. Mallat, E. Bacry, "Audio denoising by time-frequency block thresholding", *IEEE Trans. Signal Process.* 56 (5) (2008) 1830-1839.

TIME-FREQUENCY DIAGNOSIS, CONDITION MONITORING, AND FAULT DETECTION

15

INTRODUCTION AND OVERVIEW

Time-frequency (t,f) applications are now so widespread that they cannot be comprehensively covered in one volume. For this reason, this chapter aims to further illustrate the (t,f) approach by selecting a few key generic applications of diagnosis and monitoring. The topic is represented by seven sections covering a wide range of diverse applications.

One key application is electrical power quality as it is often severely affected by transient disturbances. It is necessary to detect and assess their effect on voltage and current stability. This is achieved by a time-localized frequency analysis where the instantaneous frequency (IF) allows assessing disturbance propagation (Section 15.1). In the automotive industry, the treatment and prevention of knock is a major problem for internal combustion engines as car spark knocks caused by an abnormal combustion may lead to engine damage. The Wigner-Ville distribution is used to optimize the position for placement of knock sensors (Section 15.2). Other applications involve signals that have dispersive spectral delays governed by a power law, such as dispersive propagation of a shock wave in a steel beam and cetacean mammal whistles. A power class of TFDs suitable for such applications is formulated and a methodology is described (Section 15.3). In applications of image processing, image quality may be assessed using a 2D-WVD based measure correlated with subjective human evaluations. It is shown that this SNR measure based on the WVD outperforms conventional SNR measures (Section 15.4). Some general principles of (t,f) diagnosis are then reviewed for medical applications with focus on heart sound abnormality diagnosis (Section 15.5). For machine condition monitoring, a task crucial to the competitiveness of a wide range of industries, the tasks of detecting and diagnosing faults in machines, is made easier using (t,f) approaches such as the WVD, wavelets, and wavelet packets (Section 15.6). The last section presents a specific example of condition monitoring of assets using (t,f) methods that focus on the prevention of steel beam damage (Section 15.7).

15.1 TIME-FREQUENCY ANALYSIS OF ELECTRIC POWER DISTURBANCES⁰

With the advent of various power electronic devices in modern power systems, power quality is becoming an important issue for both customers and service providers. Ideally the voltage and current signals in power systems are supposed to be constant in amplitude and frequency. However, the voltage and current frequently suffer distortions in amplitude, frequency, and phase due to various sources of disturbances [1]. The effects of periodic distortion or harmonics have been treated with Fourier series and are characterized by a total harmonic distortion (THD) index [1]. Transient disturbances, which occur within less than one cycle, are of great interest since they also degrade power quality. Specifically, the detection, estimation, classification, and assessment of transient disturbance signals have become an important aspect in power quality analysis. To overcome the inability of Fourier analysis to provide time-localized frequency information for the transient disturbance signals, advanced signal processing techniques such as the wavelet transform [2] and quadratic time-frequency distributions (TFDs) [3] have been introduced to provide time-localized frequency analysis of the disturbances.

In this section, we focus on applications of time-frequency (t, f) analysis to transient power quality events. A power system is mainly divided into distribution and transmission levels. We will provide one example from the distribution level for power quality assessment and the other from the transmission level for the propagation of disturbance waves.

15.1.1 TIME-FREQUENCY ANALYSIS: REDUCED INTERFERENCE DISTRIBUTION

As explained in [Chapter 3](#), various types of TFDs may be expressed in terms of the following quadratic (or bilinear) distribution function with a kernel $g(v, \tau)$ [4]:

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j2\pi v(t-u)} g(v, \tau) z\left(u + \frac{\tau}{2}\right) z^*\left(u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} dv du d\tau. \quad (15.1.1)$$

This $\rho_z(t, f)$ is the TFD of $z(t)$. The arguments regarding the kernel selection and suppression of interference effects are also applicable to the case of the power quality analysis. As the disturbance signal is characterized by the presence of multiple frequency components over a short time, interference is also problematic and a high resolution TFD is required. Among the various types of TFDs, the reduced interference distribution (RID) [4] has been shown to exhibit the most suitable properties for the analysis of power quality disturbance events (see Ref. [3] and [Section 5.2](#)).

The advantage of the RID can be found in the motivation and philosophy of the RID. RID is a more general definition of the TFD kernel that satisfies the following criteria: realness, time/frequency shift invariance, time/frequency marginal properties, instantaneous frequency (IF)/group delay (GD) availability, and time/frequency support. Various definitions of the RID kernel are available as a two dimensional low pass filter and the requirements listed above. Among the various types of the RID kernels, we employ the binomial distribution kernel provided in Ref. [4].

In the beginning of transient power quality signal analysis, the wavelet transform has been mainly utilized [2], because wavelet analysis provides time-localized properties. However, the TFD, especially

⁰Authors: **E.J. Powers**, The University of Texas at Austin, Austin, TX, USA (ejpowers@mail.utexas.edu), **Y.-J. Shin**, Yonsei University, Seoul, Korea (yongjune@yonsei.ac.kr), and **W. Mack Grady**, Baylor University, Waco, TX, USA (Mack_Grady@baylor.edu). Reviewers: W. J. Williams and G. J. Frazer.

the RID, is a potentially more useful distribution [5] because it allows one not only to visualize the transient signal information but also to extract transient parameters useful in assessing the severity of various transient power quality events.

15.1.2 POWER QUALITY ASSESSMENT VIA TIME-FREQUENCY ANALYSIS

In this section, we present the application of (t,f) analysis to the assessment of power quality [3]. A capacitor switching disturbance waveform with its corresponding RID is provided in Fig. 15.1.1. The objective of capacitor switching in power systems is to correct the power factor and/or mitigate the effects of harmonics associated with nonlinearities. However, the switching capacitor event also generates undesirable disturbance waveforms as shown at the top of Fig. 15.1.1. At the bottom of

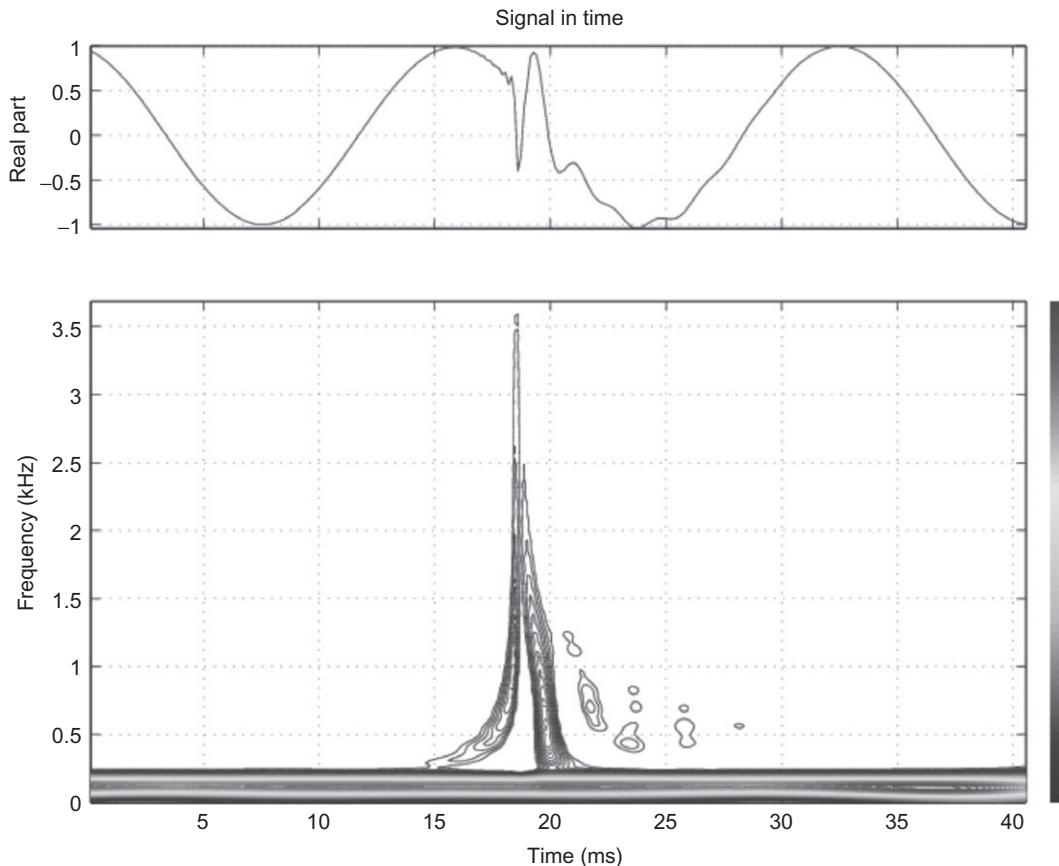


FIGURE 15.1.1

Capacitor switching disturbance time series (top) and its reduced interference distribution (bottom). From Ref. [6], © 1999 IEEE.

[Fig. 15.1.1](#) the corresponding RID is provided. The time-varying frequency content of the capacitor switching disturbance is clearly seen. Note that frequencies up to 3.5 kHz appear to be associated with this particular event. This voltage signal was recorded in the field and is provided through the courtesy of EPRI (Electric Power Research Institute).

Besides the visualization of the time-varying spectral characteristics of the disturbance provided by the binomial RID, one can characterize the frequency variation of the disturbance in terms of instantaneous frequency.

For a TFD $\rho_z(t, f)$ satisfying the IF property (see [Section 3.1.1](#)), the instantaneous frequency (IF) may be expressed as

$$f_i(t) = \frac{\int_{-\infty}^{\infty} f \cdot \rho_z(t, f) df}{\int_{-\infty}^{\infty} \rho_z(t, f) df}. \quad (15.1.2)$$

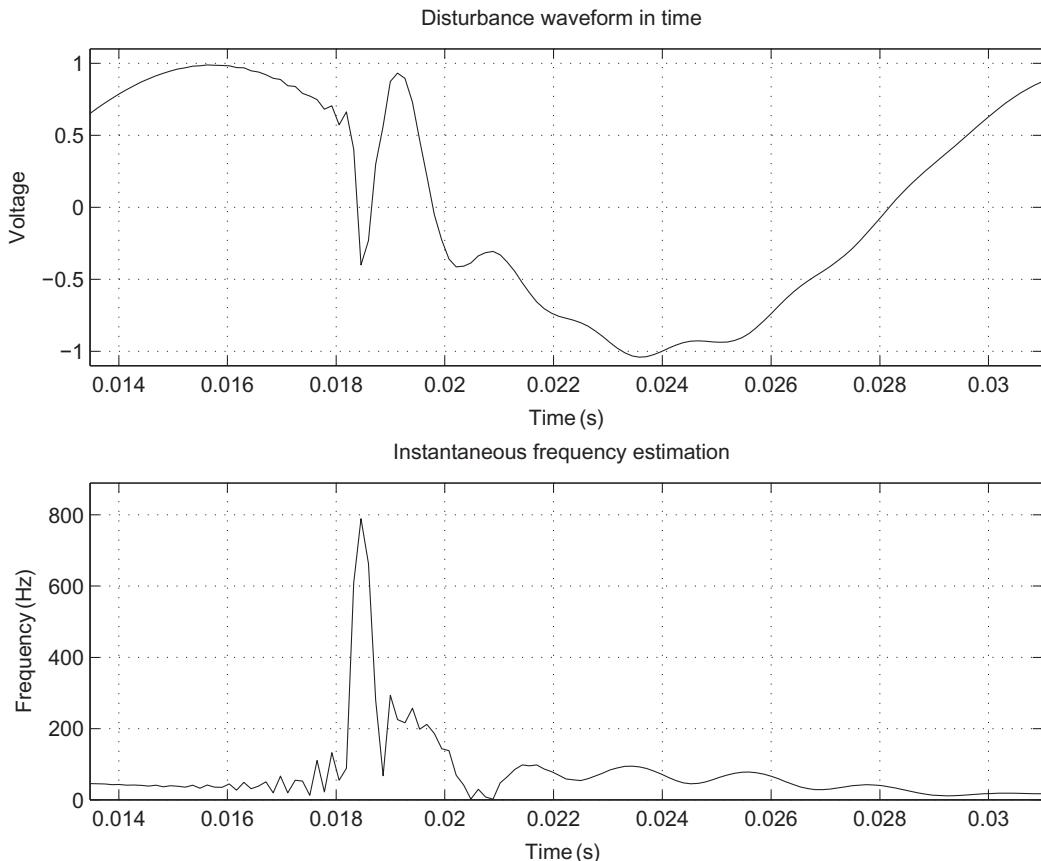
The instantaneous frequency is a normalized first order frequency moment for the TFD and corresponds to the “average” frequency for given time t , where each frequency is weighted by the relative energy associated with that frequency at time t . Note that all types of TFDs can be used for IF estimation. But, for a reasonable estimate of the instantaneous frequency, a kernel must meet the requirements that $g(v, \tau = 0) = 1$ and that $\frac{\partial g(v, \tau)}{\partial \tau}|_{\tau=0} = 0$, which are satisfied by the binomial RID kernel used in this section (see [Tables 3.3.1](#) and [6.1.1](#)).

As a result, the instantaneous frequency of the disturbance provides a quantitative assessment of the transient frequency distortion. The same capacitor switching disturbance waveform (zoomed for the disturbance portion) used in [Fig. 15.1.1](#) and its corresponding instantaneous frequency are plotted in [Fig. 15.1.2](#). The peak of the instantaneous frequency is estimated to be around 800 Hz. During the disturbance (between times 16 and 28 ms), the disturbance is reflected in the change of the instantaneous frequency. After the transient disturbance the instantaneous frequency returns to 60 Hz.

15.1.3 APPLICATION OF INSTANTANEOUS FREQUENCY FOR DISTURBANCE PROPAGATION

The time of arrival of disturbance signals on high-voltage transmission lines is of great interest for relay and fault localization. Traditional fault localization in a transmission line network is based on a fault-study using voltage and current measurements. The traditional methodology is subject to inaccurate results, because the calculation depends on the rough assumption of the fault impedance and the type of fault. Power system monitoring systems employ GPS (global positioning systems) receivers to provide time synchronized data. GPS synchronized data enables one to solve the fault location problem based on time-of-arrival of the disturbance waveforms. The propagation properties of high voltage transmission lines have been carefully treated and shown to be dispersive [7]. To treat the time synchronized disturbance data, an accurate estimation of the arrival time is critical. In this section, an application example is provided to show how the instantaneous frequency can be utilized for the arrival time estimation.

In [Fig. 15.1.3](#) a simulation circuit diagram is provided. For a long transmission line (345 kV), there occurs a typical line-to-ground fault which is 84.6 km away from “SEND” and 394.3 km away

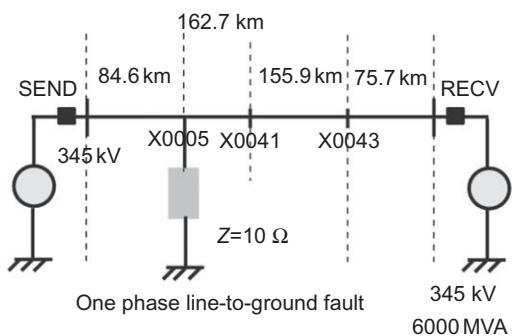
**FIGURE 15.1.2**

Capacitor switching disturbance and corresponding instantaneous frequency based on the RID. From Ref. [6], © 1999 IEEE.

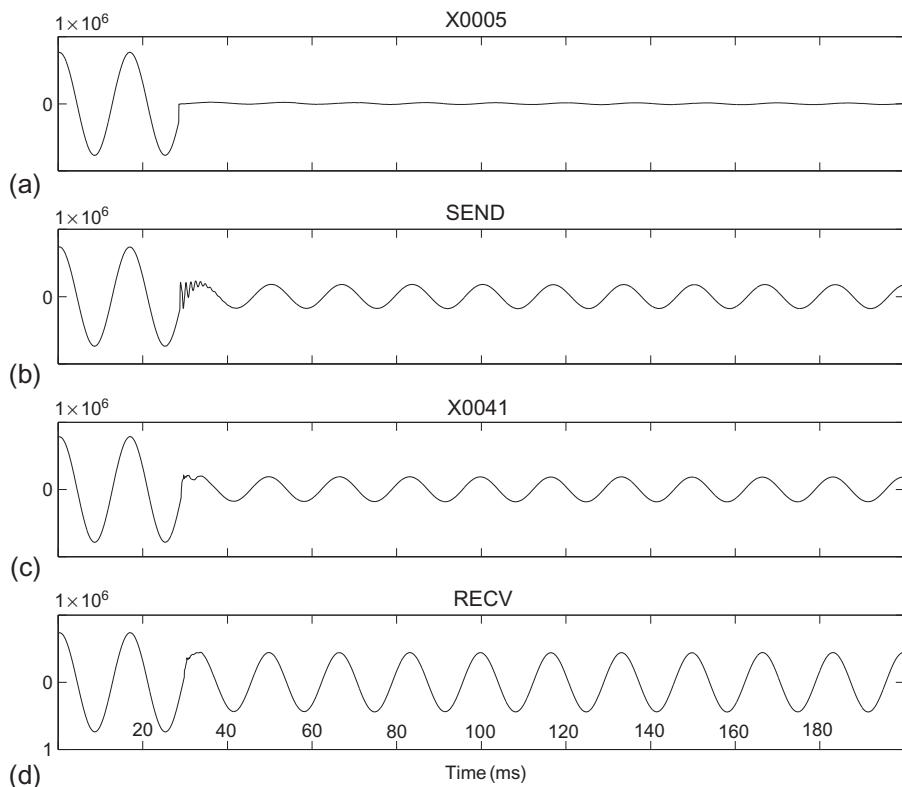
from “RECV” as indicated in Fig. 15.1.3. For this transmission line configuration, EMTP (Electro-Magnetic Transient Program) simulates the voltage and current disturbances. The corresponding voltage waveforms at individual buses (X0005, SEND, X0041, RECV) are provided in Fig. 15.1.4.

As the transmission line is characterized by frequency-dependent attenuation and dispersion, different frequencies suffer different amounts of attenuation and also propagate with different phase and group velocities; consequently, the waveforms observed at different buses appear “distorted” or “dispersed” from the original waveform. Therefore, it is difficult to assign time-of-arrival for “distorted” signals. There are perhaps many ways to determine time-of-arrival; however, in this section we focus on one, namely instantaneous frequency.

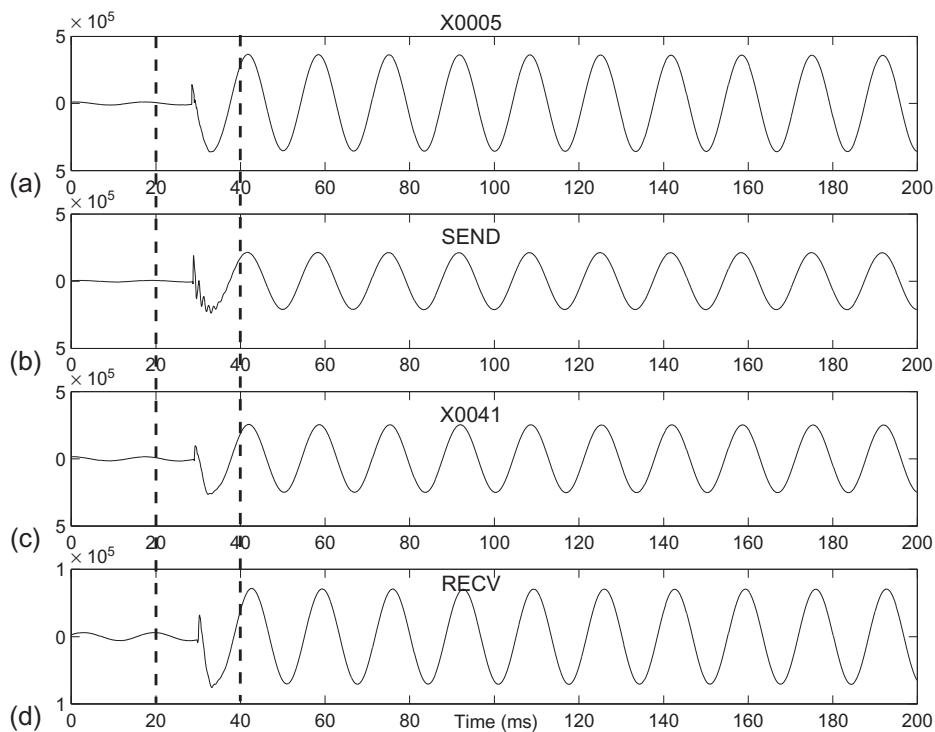
The corresponding zero sequence mode disturbance voltage is provided in Fig. 15.1.5. The zero sequence mode is a summation of the individual three-phase waveforms and is ideally zero for a

**FIGURE 15.1.3**

EMTP simulation circuit configuration. From Ref. [7], © 2000 IEEE.

**FIGURE 15.1.4**

Disturbance voltage waveforms recorded at individual buses. From Ref. [7], © 2000 IEEE.

**FIGURE 15.1.5**

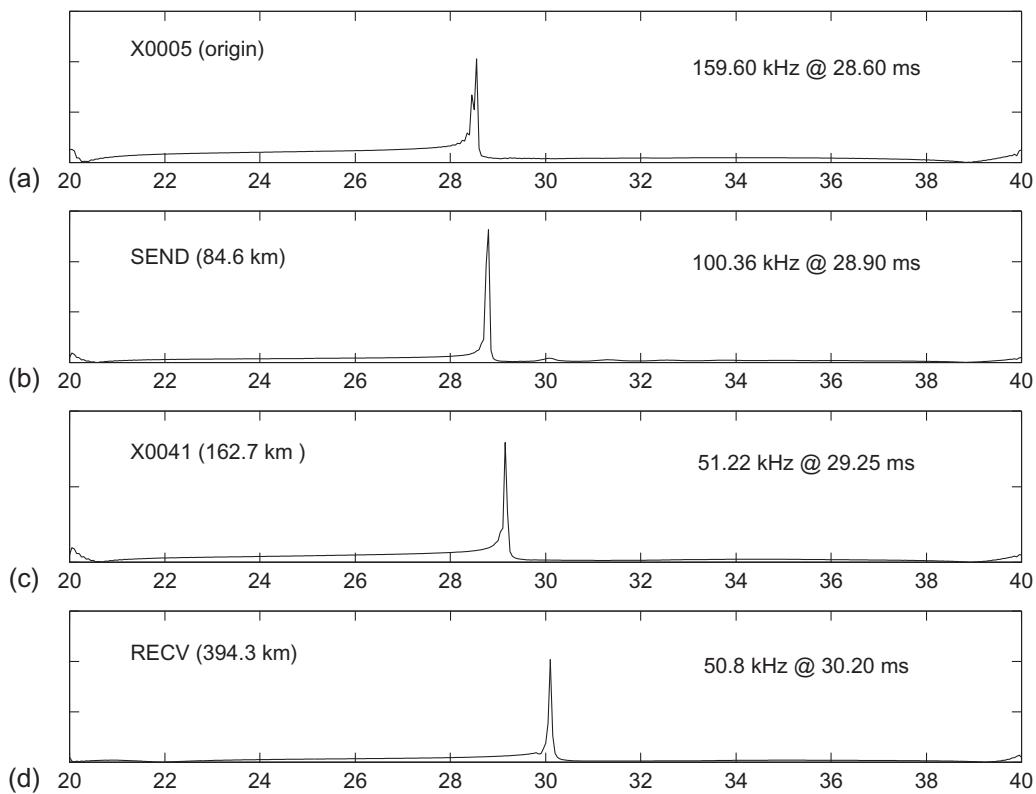
Zero sequence disturbance voltage waveforms recorded at individual buses. From Ref. [7], © 2000 IEEE.

balanced three-phase system. Thus it is very sensitive to a fault on any of the three phases as shown in Fig. 15.1.5. The reduced interference distribution has been calculated for the disturbance waveforms in zero sequence mode in order to generate the instantaneous frequency of the zero sequence disturbance signals. The instantaneous frequency, its peak value, and time of arrival of the disturbance at various observation points are provided in Fig. 15.1.6. Note that the time axis in Fig. 15.1.6 is zoomed to within a 20-40 ms interval as indicated in Fig. 15.1.5. The time of arrival t^{arrival} has been assigned as follows:

$$t^{\text{arrival}} = \arg\{\max_t[f_i(t)]\}. \quad (15.1.3)$$

As the frequency bandwidth of the disturbance is broad since the disturbance is transient, the assignment of the arrival time via the peak instantaneous frequency is a reasonable approximation.

The arrival times and peak values of the instantaneous frequency are presented in Table 15.1.1. To convert times to distance we utilize the results of the analysis presented in Ref. [7], where it was shown that for a range of peak frequencies appearing in Table 15.1.1, the corresponding zero-sequence group velocity is $v_g = (2.6 \pm 0.2) \times 10^8$ m/s. The corresponding estimates of distance are compared to the true distance in Table 15.1.1. Note the range of estimated distance agrees quite well with the known true

**FIGURE 15.1.6**

Instantaneous frequency estimation of the disturbance voltage waveforms in zero sequence. From Ref. [7], © 2000 IEEE.

Table 15.1.1: Summary of the Zero Sequence Disturbance via RID

| Bus Name (unit) | Arrival Time (ms) | Peak IF (kHz) | True Distance (km) | Estimated Distance (km) |
|----------------------------|------------------------------|--------------------------|-------------------------------|------------------------------------|
| X0005 | 28.60 | 159.60 | 0.0 | N/A |
| SEND | 28.90 | 100.36 | 84.6 | 72–84 |
| X0041 | 29.25 | 51.22 | 162.7 | 156–182 |
| RECV | 30.20 | 50.80 | 394.3 | 384–448 |

distance. Note, also, that the peak instantaneous frequency is lower for the larger propagation distance. This is due to the fact that higher frequencies associated with the disturbance suffer greater attenuation than lower frequencies. Ongoing work in voltage-only distance localization involves refinement of the instantaneous frequency approach and consideration of the use of group delay.

15.1.4 SUMMARY AND CONCLUSIONS

We have given two examples of (t,f) analysis applied to power system disturbances. RID-based (t,f) analysis successfully provides a useful characterization of power system disturbance signals. We have also demonstrated that the concept of instantaneous frequency is quite useful in providing a simplified picture of the (t,f) features of a transient disturbance and in determining the arrival time of a disturbance observed in the zero-sequence mode of a three-phase system. These days, GPS-based time synchronization technology is being adopted for fault location of electric power systems discussed in this section.

15.2 COMBUSTION DIAGNOSIS BY TF ANALYSIS OF CAR ENGINE SIGNALS⁰

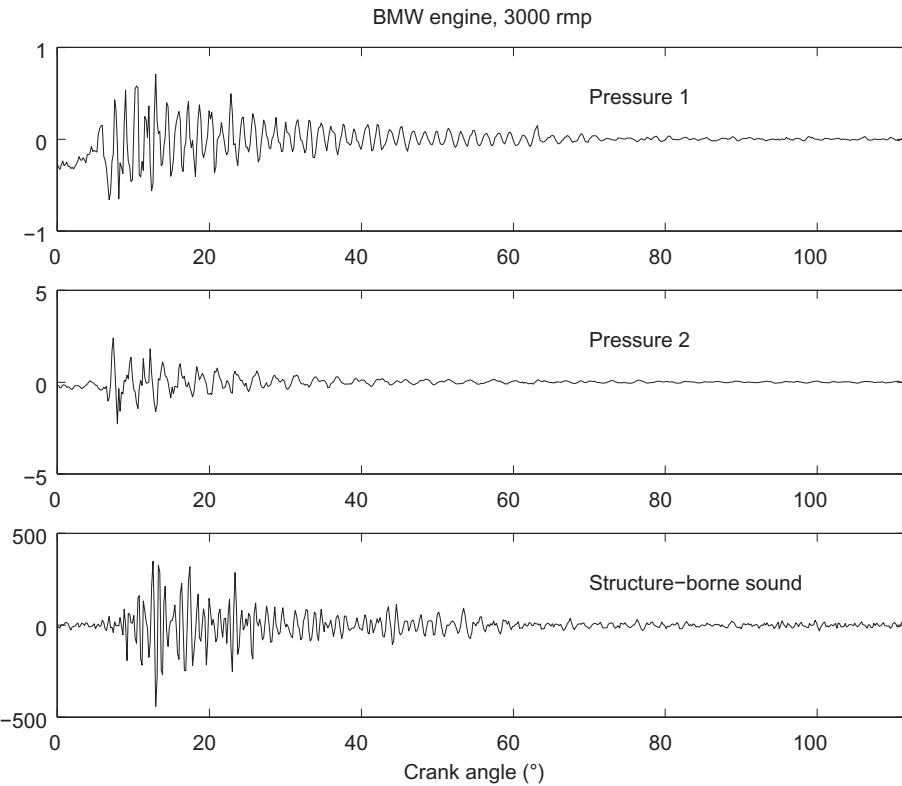
15.2.1 KNOCKING COMBUSTIONS

A permanent challenge for car manufacturers is to increase efficiency, reduce pollution, and prolong life of internal combustion engines. A restriction to these aims is the occurrence of knock. Knock is an undesired spontaneous auto-ignition of the end gas causing a sharp increase of pressure and temperature [8]. Generally, rare knock has no effect to engine performance but frequent or very strong knock can damage the engine. Knock excites combustion chamber acoustic resonances that can be measured by special pressure sensors. But these are too expensive for use in serial vehicles. Nowadays, acceleration sensors mounted on the engine housing measure structure-borne sound as a distorted version of pressure to detect knock. Time-frequency analysis can help to understand the nature of pressure and structure-borne sound signals and to improve knock detection. The following applications for combustion diagnosis make use of the high frequency part of the signals which means frequencies above 3 kHz. Therefore, all signals were high pass filtered. [Figure 15.2.1](#) shows two pressure signals and a structure-borne sound signal of a BMW engine recorded simultaneously. The sensor of pressure 1 is mounted in the spark plug and the sensor of pressure 2 in the cylinder head.

15.2.2 SIGNAL MODELS

Considering homogeneous gas distribution, the frequencies of the resonances depend on the speed of sound and on the combustion chamber geometry. The latter can be approximated by an ideal cylinder [9] or more accurately using finite element methods (FEM) [10]. [Figure 15.2.2](#) shows the instantaneous frequencies of a BMW engine estimated by FEM simulations [11]. But they yield information neither about the time instant when knock occurs nor about the resonance amplitudes. The time instant of excitation is random, usually between 5° and 15° crank angle after top dead center of the piston. The amplitudes are damped due to heat and friction losses and to increasing combustion chamber volume. Previous considerations motivate modeling of the undisturbed pressure signal in the time domain by Ref. [12].

⁰Authors: **J.F. Böhme**, Ruhr-Universität Bochum, Bochum, Germany (Johann.Boehme@rub.de) and **S. Carstens-Behrens**, University of Applied Sciences Koblenz, Remagen, Germany (carstens-behrens@rheinahrcampus.de). Reviewers: L.J. Stanković and G. Matz.

**FIGURE 15.2.1**

Example of time signals of a knocking combustion.

$$X(t) = \sum_{p=1}^P A_p e^{-d_p(t-t_0)} \cos \left(2\pi \int_{t_0}^t f_p(\alpha(\tau)) d\tau + \Phi_p \right) u(t-t_0). \quad (15.2.1)$$

The pressure signal is a superposition of P resonances; index p refers to the p th component. A_p and Φ_p are random initial amplitude and phase, $e^{-d_p t}$ describes the damping of the oscillation, t_0 is a suitable time instant after excitation, $u(t)$ is the step function, and $f_p(\alpha(t))$ is the instantaneous frequency at crank angle $\alpha(t)$. The engine housing transfer function is modeled as a linear, time-varying function $h(z(t), \tau)$, whereas time variance stems only from piston position $z(t)$ which is a function of crank angle. Merging the pressure model and the engine housing transfer function yields the model of the undisturbed structure-borne sound signal

$$Y(t) = \int_{-\infty}^{\infty} h(z(t), \tau) X(t - \tau) d\tau \approx \sum_{p=1}^P A_p e^{-d_p(t-t_0)} \operatorname{Re} \left\{ H_p(z(t), 2\pi f_p(\alpha(t))) e^{j2\pi \int_{t_0}^t f_p(\alpha(\tau)) d\tau + \Phi_p} \right\}, \quad (15.2.2)$$

where

$$H_p(z(t), f) = \int_{-\infty}^{\infty} h(z(t), \tau) e^{d_p \tau} e^{-j2\pi f \tau} d\tau. \quad (15.2.3)$$

The approximation is valid if the instantaneous frequencies $f_p(\alpha(t))$ do not change fast. Thus, the structure-borne sound signal is a complex amplitude modulated version of the pressure signal.

15.2.3 SIGNAL ANALYSIS USING WIGNER-VILLE SPECTRUM

Pressure and structure-borne sound signals consist of a superposition of amplitude and frequency modulated components. Therefore, identification of the components in the time domain or frequency domain alone is difficult: a suitable time-frequency representation is required. The Wigner-Ville spectrum (WVS) yields a high time-frequency resolution. For a stochastic process $S(t)$ the WVS is defined by

$$W_S(t, f) = \int_{-\infty}^{\infty} r_S \left(t + \frac{1}{2}\tau, t - \frac{1}{2}\tau \right) e^{-j2\pi f \tau} d\tau, \quad (15.2.4)$$

where $r_S(t, t') = E[S(t)S^*(t')]$ is the correlation function of $S(t)$ (see Sections 2.2 and 9.4). In the case of constant speed and load pressure the structure-borne sound signals are cyclo-stationary with low stochastic dependency from combustion to combustion [12]. Exploiting this fact, the correlation function can be estimated by

$$\hat{r}_S(t, t') = \frac{1}{L} \sum_{l=1}^L s_l(t) s_l(t'), \quad (15.2.5)$$

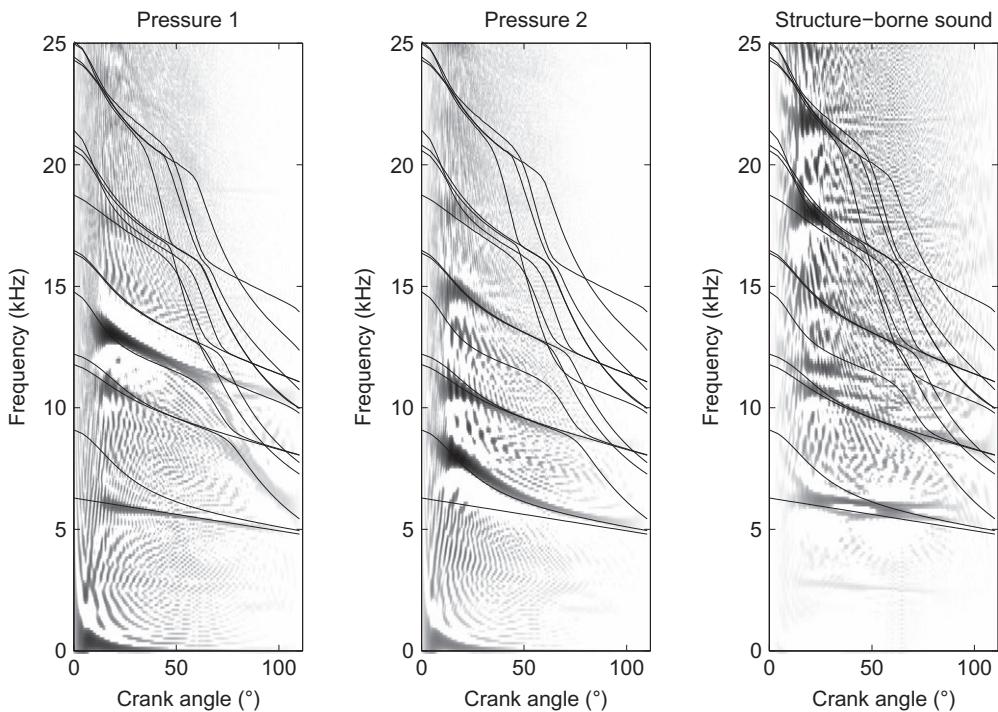
where $s_l(t)$ is observed pressure or structure-borne sound of the l th combustion, $l = 1, \dots, L$, with $s_l(t) = 0$ if t is not in the observation interval $[0, T]$, and $t = 0$ corresponds to top dead center for each combustion cycle. Substituting the estimated correlation function in Eq. (15.2.4) yields the estimated WVS of pressure or structure-borne sound, respectively,

$$\hat{W}_S(t, f) = \int_{-\infty}^{\infty} \hat{r}_S \left(t + \frac{1}{2}\tau, t - \frac{1}{2}\tau \right) e^{-j2\pi f \tau} d\tau. \quad (15.2.6)$$

This WVS estimate has three advantages: it has a high time-frequency resolution, cross-terms are reduced effectively, and the algorithm is fast compared to other approaches that first estimate the WVS of each combustion and then take the mean over all combustions.

15.2.3.1 Model verification

Figure 15.2.2 shows the estimated WVS of the signals of two differently located pressure sensors and a structure-borne sound signal of a BMW engine. In order to enhance weak signal components, signal amplitudes are scaled logarithmically, so we have to be careful in interpreting gray scales as amplitudes. In addition, Fig. 15.2.2 shows the instantaneous frequencies estimated by FEM simulations. Obviously, the pressure sensors observe different resonances due to their positions and the pressure nodes' and anti-nodes' locations. The frequencies of the visible pressure resonances coincide well with the frequencies estimated by FEM simulations. The structure-borne sound signal shows similar behavior but there are some strange components at approximately 6.5 kHz and between 10 and 12 kHz. They may be

**FIGURE 15.2.2**

Frequencies estimated by finite elements (solid lines) and estimated WVS of pressure and structure-borne sound signals of a BMW engine at 3000 rpm based on 96 realizations measured simultaneously.

a consequence of the engine housing transfer function, additional noise, or of insufficiently reduced cross-terms of the estimated WVS. As mentioned before, we have to be careful in interpreting the gray scale amplitudes. Nevertheless, we can recognize for pressure and structure-borne sound signals that knock occurs around 10° crank angle. Furthermore, signal power (and amplitude) is maximal at excitation of knock and decreases with increasing crank angle. In general, the results of the estimated WVS justify the decomposition of pressure and structure-borne sound signals into frequency and amplitude modulated components.

15.2.3.2 Knock sensor position optimization

The number of resonances being observed by a sensor depends strongly on its position. This holds for pressure as well as for structure-borne sound. Therefore, it is very important to find a good acceleration sensor position. This can be done at test bed engines by optical methods like laser holography or by assessing the position of several acceleration sensors with statistical analysis [13]. A very simple, intuitive, and fast approach is to use the estimated WVS. Since there are fast algorithms for its implementation, it can be calculated at the test bed while the engine is running. The operator compares the WVS of the pressure signal with those of the acceleration sensor signals. Using pressure

as the reference, structure-borne sound signals that contain the same resonances are more suitable for knock detection than signals with fewer or different resonances. As a spin-off, observing the sensor signals by the WVS is very useful for verifying measurement setup, a very important task at test bed measurements.

15.2.3.3 Nonequidistant sampling

Knock detection is usually based on estimated resonance energies of structure-borne sound signals. As previous considerations have shown, the frequencies vary with crank angle. Optimum estimation of resonance energies can therefore be achieved by time-varying filtering. Alternatively, optimized nonequidistant sampling in the time domain leads to resonances with approximately constant frequencies [14]. Using Wigner-Ville analysis, parameters can be found to optimally compensate the frequency modulation in a crank angle interval with significant resonance energy. Then, ordinary band-pass filtering is sufficient to estimate the resonance energies well.

15.2.4 SIGNAL ANALYSIS USING S-METHOD

The estimated WVS is a useful tool for analyzing a large number of combustions at once. But if the number of combustions is not large enough, cross-terms disturb the estimated spectrum. In many cases, we are interested in analyzing single combustions. There are a couple of good WVS estimators reducing or avoiding cross-terms; see, e.g., Refs. [15,16]. Among these, the S-method has been proven to be an appropriate tool to represent single knock signals in time-frequency domain [17]. Using the short-time Fourier transform of the signal $s(t)$

$$F_s^w(t,f) = \int_{-\infty}^{\infty} s(t+\tau) w(\tau) e^{-j2\pi f\tau} d\tau, \quad (15.2.7)$$

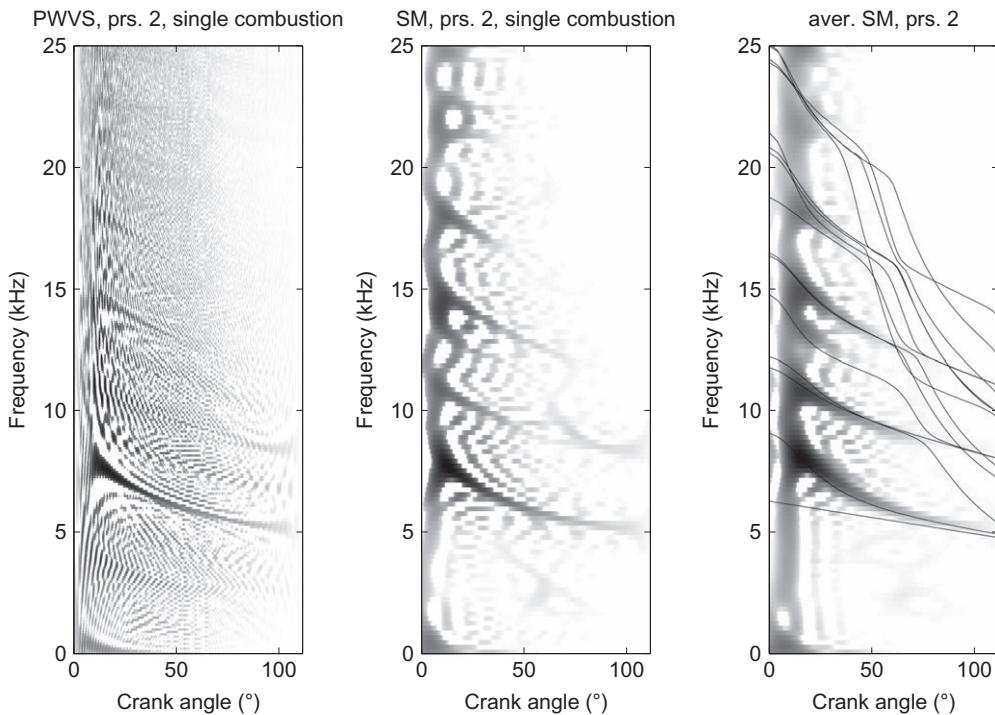
where $w(\tau)$ is a lag window, the S-method can be defined as

$$\text{SM}(t,f) = 2 \int_{-\infty}^{\infty} P(\theta) F_s^w(t,f+\theta) F_s^{w*}(t,f-\theta) d\theta \quad (15.2.8)$$

with an appropriate window $P(\theta)$. The width of $P(\theta)$ significantly affects the behavior of the S-method with regard to reduction of cross-terms. For a detailed discussion of the S-method see Section 6.2. Exploiting fast fourier transform algorithms, the S-method can be implemented very efficiently [18], so that real-time processing is possible on a fast DSP.

15.2.4.1 Observation of single combustions

Figure 15.2.3 shows the estimated WVS and the S-method of a single knocking combustion. The estimated WVS seems to contain mainly one component. There are some more but it is very difficult to distinguish between auto-terms and cross-terms. The time-frequency concentration of the S-method is worse but it is easy to identify at least four components. Therefore, the S-method is much more suitable to observe single combustions than the estimated WVS. Normally, the WVS is estimated by averaging over a large number of realizations. Analogously, the S-method of a large number of single knocking combustions was averaged in Fig. 15.2.3 (right). Comparing this representation to the corresponding WVS in Fig. 15.2.2 (middle), we recognize the worse resolution of the averaged

**FIGURE 15.2.3**

Left: estimated WVS of a single knocking combustion of a BMW engine at 3000 rpm observed by pressure sensor 2; middle: S-method of the same signal; right: S-method averaged over 96 realizations.

S-method. But the amplitude and frequency modulation is observable. In contrast to the WVS, the averaged S-method is able to resolve the component that starts at 15 kHz at 0° crank angle which is the strongest component in pressure 1; see Fig. 15.2.2 (left). In the WVS this component is covered by cross-terms. Apparently, the number of realizations was too small to reduce cross-terms sufficiently.

15.2.4.2 Power and energy estimation

As demonstrated before, the S-method can be used to investigate single combustions. As shown in Ref. [17], it is possible to track the resonances and to estimate the power $\hat{A}_p^2(t)$ of component p at each time instant t by integrating along the frequency axis in the component's band $B_p(t)$:

$$\hat{A}_p^2(t) = \int_{B_p(t)} \text{SM}(t,f) df. \quad (15.2.9)$$

In the simplest case, the band width $B_p(t)$ is constant. But considering knock signals, the band width decreases with crank angle; see Fig. 15.2.2. So the signal-to-noise ratio can be improved by adapting

the band width. In [Section 6.2](#) an adaption algorithm is proposed. Finally, integration of $\hat{A}_p^2(t)$ along the observation interval $[0, T]$ yields the estimated resonance energy \hat{E}_p of component p

$$\hat{E}_p = \int_0^T \hat{A}_p^2(t) dt. \quad (15.2.10)$$

15.2.4.3 Knock detection

The main application of estimating resonance energies of internal combustion engines is the detection of knock. A knock detector in today's cars based on structural-borne sound usually estimates the signal energy in a wide band, e.g., 5-12 kHz. The energy is compared to a threshold which yields the knock detection. This is a simple, fast, and successfully applied method. But it can be improved by taking into consideration that signal energy is the sum of the energies of single components. Since the engine housing transfer function may attenuate or amplify the components differently, we get a wrong idea of the energy distribution inside the cylinder which is used as reference at test bed engines. Among other time-frequency detectors like those presented in Refs. [19,20] or [Section 12.4](#), the S-method is a useful tool for estimating the resonance energies of single components and is therefore advantageous for detecting knock; see, e.g., Ref. [21].

15.2.5 SUMMARY AND CONCLUSIONS

On the basis of the WVS and the S-method, it was demonstrated that time-frequency representations are suitable and very useful tools for the task of combustion diagnosis of internal combustion engines. They allow an intuitive insight into relevant signal parameters and properties like amplitude and frequency modulation. Therefore, they simplify and improve investigations during the development and application of new engines. Furthermore, they help to improve signal processing for higher engine efficiency.

15.3 POWER CLASS TIME-FREQUENCY REPRESENTATIONS AND THEIR APPLICATIONS⁰

15.3.1 POWER CLASS QUADRATIC TIME-FREQUENCY REPRESENTATIONS

Various classes of quadratic time-frequency representations (QTFRs) are best suited for analyzing signals with certain types of time-frequency (t,f) geometries. For example, when a signal has constant (t,f) characteristics, the quadratic class as defined on p. 113 is most appropriate. The aforementioned QTFR classification is based on the grouping together of all QTFRs that satisfy the same two signal transformation covariance properties (see Refs. [22,23] and [Sections 4.3](#) and [5.6](#)). Specifically, the quadratic class [24–26] consists of QTFRs that are covariant to constant (nondispersive) time shifts and

⁰Authors: **A. Papandreou-Suppappola**, Arizona State University, Tempe, AZ, USA (papandreou@asu.edu), **F. Hlawatsch**, Vienna University of Technology, Vienna, Austria (franz.hlawatsch@tuwien.ac.at), and **G.F. Boudreux-Bartels**, University of Rhode Island, Kingston, RI, USA (boud@ele.uri.edu). Reviewers: J. Bertrand and J.P. Ovarlez.

frequency shifts of the signal whereas the affine class (see Refs. [25,27–29] and Section 7.1) consists of QTFRs that are covariant to scale changes (dilations) and constant time shifts. Furthermore, the hyperbolic class [30,31] consists of QTFRs that are covariant to scale changes and hyperbolic dispersive time shifts and are best suited to analyze signals with hyperbolic (nonlinear) group delay. When the analysis signal has a group delay that is a power function of frequency, the aforementioned QTFRs do not provide an adequate representation as they do not match power (t,f) characteristics. Thus, *power class* QTFRs can successfully localize signals along their power law group delay functions [32,33].

The importance of power QTFRs is that many applications involve signals with dispersive group delays governed by some power parameter κ . Examples include the dispersive propagation of a shock wave in a steel beam ($\kappa = 1/2$), trans-ionospheric chirps measured by satellites ($\kappa = -1$), acoustical waves reflected from a spherical shell immersed in water, various cetacean mammal whistles, and signal solutions of the diffusion equation ($\kappa = 1/2$) (e.g., waves propagating along uniform distributed RC transmission lines) [32]. Power laws can also be used to roughly approximate other, more complex, group delays.

15.3.1.1 Localized signal analysis application

The type of signals mentioned above that constitute the family of *power impulses* are defined in the frequency domain as

$$I_c^{(\kappa)}(f) \stackrel{\Delta}{=} \sqrt{|\tau_\kappa(f)|} e^{-j2\pi c \Lambda_\kappa\left(\frac{f}{f_r}\right)} = \sqrt{(|\kappa|/f_r) |f/f_r|^{\kappa-1}} e^{-j2\pi c \operatorname{sgn}(f)|f/f_r|^\kappa} \quad (15.3.1)$$

with monotonic phase spectrum $\Lambda_\kappa(f/f_r) = \operatorname{sgn}(f) |f/f_r|^\kappa$ and power group delay $\tau_g(f) = c \tau_\kappa(f) = c \frac{|\kappa|}{f_r} |f/f_r|^{\kappa-1} = c \frac{d}{df} \Lambda_\kappa(f/f_r)$ with $f \in \mathbb{R}$. Here, $\operatorname{sgn}(f)$ provides the sign (± 1) of f , and $f_r > 0$ is a fixed reference frequency. For successful analysis, an ideal QTFR T must be localized along the group delay (also called spectral delay in Chapter 1) $\tau_g(f)$ of the power impulse in Eq. (15.3.1). In particular,

$$T_{I_c^{(\kappa)}}(t,f) = |\tau_\kappa(f)| \delta(t - c \tau_\kappa(f)) \quad (15.3.2)$$

is desirable in applications where information about the signal analyzed could be obtained from the localized curve $t = c \tau_\kappa(f)$ in the (t,f) plane. For example, the localization could be used in applications such as signal classification or estimation of the parameter c in Eq. (15.3.1). As shown next, some power class QTFRs ideally provide the localized representation in Eq. (15.3.2) for analyzing power impulses as well as other signals with power group delay.

The power law (t,f) structure can also be observed in applications where a system can cause a time shift that varies dispersively in frequency to an input signal with Fourier transform (FT) $X(f)$ yielding the output signal $Y(f) = e^{-j2\pi c \Lambda_\kappa(f/f_r)} X(f)$. Thus, power QTFRs could be used in applications where a signal with constant group delay passes through a system with power dispersive (t,f) characteristics that transforms the signal's constant group delay to a power group delay [32]. E.g. the ocean is a medium (system) that could cause power dispersive changes to an underwater communications message and its echoes. These changes could be accounted for at the receiver when matching processing tools, like power QTFRs, are used for detection.

15.3.1.2 The power of QTFR classes

Following the covariance-based classification method, for a signal $x(t)$ with FT $X(f)$, we define all κ th power class QTFRs, $T_X^{(\kappa)}(t, f)$, to satisfy two properties [32,33]. The first is covariance to scale changes on $x(t)$, i.e.,

$$T_{\mathcal{C}_a X}^{(\kappa)}(t, f) = T_X^{(\kappa)}(at, f/a), \quad (15.3.3)$$

where the scaling operator \mathcal{C}_a is defined as $(\mathcal{C}_a X)(f) = X(f/a)/\sqrt{|a|}$. The second property is covariance to power time shifts on $x(t)$ that correspond to frequency-dependent shifts, $\tau_\kappa(f)$, in the signal's group (or spectral) delay. Specifically,

$$T_{\mathcal{D}_c^{(\kappa)} X}^{(\kappa)}(t, f) = T_X^{(\kappa)}(t - c \tau_\kappa(f), f) = T_X^{(\kappa)}\left(t - c(\kappa/f_r)|f/f_r|^{\kappa-1}, f\right) \quad (15.3.4)$$

for $f \in \mathbb{R}$. The effect of the power time shift operator $\mathcal{D}_c^{(\kappa)}$ is given by Ref. [32]

$$(\mathcal{D}_c^{(\kappa)} X)(f) = e^{-j2\pi c \Lambda_\kappa(f/f_r)} X(f) = e^{-j2\pi c \operatorname{sgn}(f)|f/f_r|^\kappa} X(f). \quad (15.3.5)$$

Here, $c \in \mathbb{R}$ and $\kappa \in \mathbb{R}$ ($\kappa \neq 0$) is the power parameter associated with each power class. The κ th power function $\Lambda_\kappa(b) = \operatorname{sgn}(b)|b|^\kappa$, $b \in \mathbb{R}$, corresponds¹ to a transformation of the phase spectrum of the signal as shown in Eq. (15.3.5). The frequency-dependent time shift $\tau_\kappa(f)$ in Eq. (15.3.4) corresponds to the derivative of the power function, i.e., $\tau_\kappa(f) = \frac{d}{df} \Lambda_\kappa(f/f_r) = \frac{\kappa}{f_r} |f/f_r|^{\kappa-1}$. Figure 15.3.1 depicts $\tau_\kappa(f)$ in the (t, f) plane for various choices of the power parameter κ .

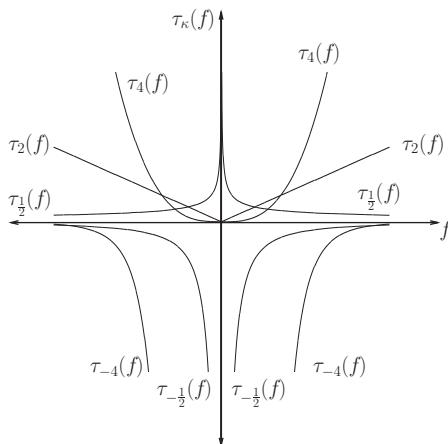


FIGURE 15.3.1

The power group delay $\tau_\kappa(f)$ for various choices of the power parameter κ .

¹ The definition $\Lambda_\kappa(b) = \operatorname{sgn}(b)|b|^\kappa$ vs. $\Lambda_\kappa(b) = b^\kappa$ extends the power function to $b < 0$ so that $\Lambda_\kappa(b)$ is an odd, strictly monotonic function constituting a one-to-one mapping from \mathbb{R} to \mathbb{R} .

The importance of the κ th power class QTFRs is directly linked to the two covariances in Eqs. (15.3.3) and (15.3.4). The power time shift covariance in Eq. (15.3.4) is useful in analyzing signals passing through dispersive systems with power law group delay signals localized along power law curves in the (t, f) plane. On the other hand, the scale covariance in Eq. (15.3.3) is important for multiresolution analysis.

15.3.1.3 Power class formulation

Any QTFR of the κ th power class can be expressed as [32]

$$T_X^{(\kappa)}(t, f) = \frac{1}{|f|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma_T(f_1/f, f_2/f) \cdot e^{j2\pi(tf/\kappa)[\Lambda_\kappa(f_1/f) - \Lambda_\kappa(f_2/f)]} X(f_1) X^*(f_2) df_1 df_2, \quad (15.3.6)$$

where the 2-D kernel $\Gamma_T(b_1, b_2)$ uniquely characterizes the QTFR. Specific choices of $\Gamma_T(b_1, b_2)$ define specific QTFRs $T^{(\kappa)}$ in the κ th power class. When $\kappa = 1$ we obtain the affine class (see Section 7.1), which is an important special case of the power classes corresponding to the constant (nondispersive) time shift $\tau_1(f) \equiv 1/f_r$.

The κ th power class QTFR in Eq. (15.3.6) can also be obtained via a unitary warping operation (see Sections 4.5 and 5.6 and Refs. [32–34]). Specifically, if $T_X^{(A)}(t, f)$ is a QTFR of the affine class, then the corresponding κ th power class QTFR, $T_X^{(\kappa)}(t, f)$, can be obtained by warping the affine class QTFR, $T_X^{(A)}(t, f)$, according to Ref. [32]

$$T_X^{(\kappa)}(t, f) = T_{\mathcal{U}_\kappa X}^{(A)} \left(\frac{t}{f_r \tau_\kappa(f)}, f_r \Lambda_\kappa(f/f_r) \right) = T_{\mathcal{U}_\kappa X}^{(A)} \left(\frac{t}{\kappa |f/f_r|^{\kappa-1}}, f_r \operatorname{sgn}(f) |f/f_r|^\kappa \right). \quad (15.3.7)$$

Here, the unitary² frequency axis warping operator [32,33] \mathcal{U}_κ is given by

$$(\mathcal{U}_\kappa X)(f) = \frac{X(f_r \Lambda_\kappa^{-1}(f/f_r))}{\sqrt{f_r |\tau_\kappa(f_r \Lambda_\kappa^{-1}(f/f_r))|}} = \frac{1}{\sqrt{|\kappa|} |f/f_r|^{\frac{\kappa-1}{2\kappa}}} X(f_r \operatorname{sgn}(f) |f/f_r|^{\frac{1}{\kappa}}),$$

where the inverse function $\Lambda_\kappa^{-1}(b)$ satisfies $\Lambda_\kappa^{-1}(\Lambda_\kappa(b)) = \Lambda_\kappa(\Lambda_\kappa^{-1}(b)) = b$. The QTFRs of the affine class are defined as [25]

$$T_X^{(A)}(t, f) = \frac{1}{|f|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma_T(f_1/f, f_2/f) e^{j2\pi t(f_1-f_2)} X(f_1) X^*(f_2) df_1 df_2,$$

where $\Gamma_T(b_1, b_2)$ is a 2-D kernel characterizing the affine class QTFR (cf. (15.3.6)). Note that $T_X^{(A)}(t, f) = T_X^{(\kappa)}(t, f)|_{\kappa=1}$. The unitary warping relation in Eq. (15.3.7) preserves certain desirable characteristics of the affine class while transforming other ones to match the dispersive nature of the signals to be analyzed by power class QTFRs. For example, whereas both classes preserve scale changes of the signal, only the affine class preserves constant (nondispersive) time shifts. On the other hand, the warping in Eq. (15.3.7) transforms constant time shifts to power dispersive time shifts in the power class, and thus the constant time shift covariance of the affine class is transformed into the power time shift covariance of the power classes. The warping also provides an efficient method for computing power class QTFRs when algorithms for computing affine class QTFRs are available [33].

²Unitarity of the operator \mathcal{U}_κ implies that \mathcal{U}_κ preserves inner products. Specifically, the operator satisfies $\int_{-\infty}^{\infty} (\mathcal{U}_\kappa X)(f) (\mathcal{U}_\kappa X)^*(f) df = \int_{-\infty}^{\infty} X(f) X^*(f) df$.

15.3.1.4 Class members

Specific QTFRs of the power classes satisfy various desirable properties in addition to the covariance properties in Eqs. (15.3.3)–(15.3.4) satisfied by *all* members of the power classes. Some power class QTFRs of particular importance include the *power Wigner distribution*, the *powergram*, the *smoothed pseudo power Wigner distribution* [32], and the *Bertrand P_κ -distributions* (see Section 7.1 and Refs. [28,32]). All these QTFRs have counterparts in the affine class by virtue of the power warping relation in Eq. (15.3.7). For example, the power Wigner distribution,

$$\begin{aligned} W_X^{(\kappa)}(t,f) &= \left| \frac{f}{\kappa} \right| \int_{-\infty}^{\infty} X\left(f\Lambda_\kappa^{-1}\left(1 + \frac{\beta}{2}\right)\right) X^*\left(f\Lambda_\kappa^{-1}\left(1 - \frac{\beta}{2}\right)\right) e^{j2\pi\frac{f}{\kappa}\beta} \frac{d\beta}{|1 - \frac{\beta^2}{4}|^{(\kappa-1)/(2\kappa)}} \\ &= W_{U_\kappa X}\left(t/\left(f_r \tau_\kappa(f)\right), f_r \Lambda_\kappa(f/f_r)\right) \end{aligned}$$

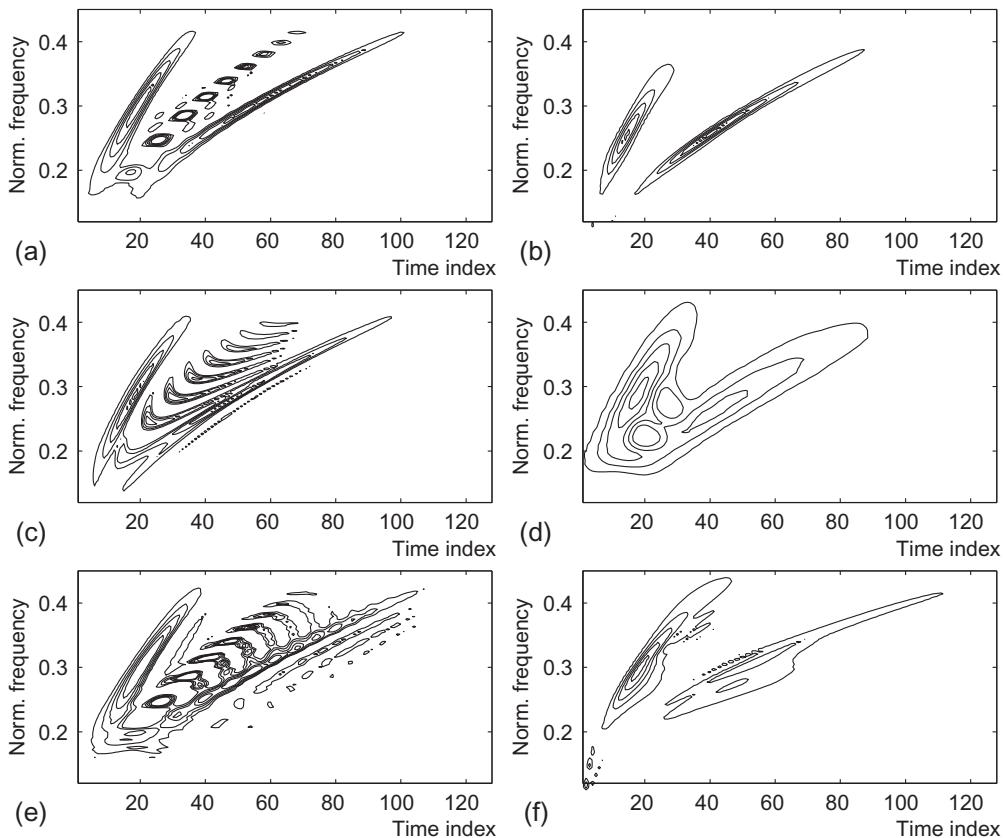
(cf. (15.3.7)), is the power warped version of the well-known Wigner distribution (WD), $W_X(t,f) = \int_{-\infty}^{\infty} X(f + \frac{v}{2}) X^*(f - \frac{v}{2}) e^{j2\pi tv} dv$ defined in Chapter 2, Eq. (2.1.32) [22,24,25]. As such, it satisfies many properties such as a specific set of marginal properties and the perfect (t,f) localization property for power impulses in Eq. (15.3.2). Just like the WD, the power WD contains oscillatory and partly negative cross-terms when multicomponent signals are analyzed (see Section 4.2 and Ref. [33]). In order to suppress cross-terms, a specific type of smoothing can be applied that is matched to the power (t,f) geometry. The powergram and the smoothed pseudo power WD apply such a smoothing to the power WD, at the expense of the loss of some properties (such as the marginal properties) and the loss of (t,f) resolution. The Bertrand P_κ -distributions (see Section 7.1 and Ref. [28]) are also perfectly localized for power impulses.

15.3.2 POWER CLASS APPLICATIONS

15.3.2.1 Synthetic data analysis example

The discrete implementation of power QTFRs (outlined in Refs. [32,33]) is applied to analyze a two-component signal consisting of two power impulses with power parameter $\kappa_{\text{signal}} = 3$. For computational purposes, the impulses are windowed in the frequency domain. Figure 15.3.2(a) and (b) shows the results obtained with the power WD and a smoothed pseudo power WD with a very short smoothing window. Both QTFRs have power parameter $\kappa = 3$, matched to the power impulse parameter κ_{signal} . The power WD in Fig. 15.3.2(a) has very good (t,f) concentration but large cross-terms [33] which are effectively suppressed in the smoothed pseudo power WD in Fig. 15.3.2(b) with hardly any loss of (t,f) concentration. Also shown (in Fig. 15.3.2(c) and (d)) are the results obtained with the WD and an affine-smoothed pseudo WD, both members of the affine class [25] (i.e., both QTFRs have power parameter $\kappa = 1 \neq \kappa_{\text{signal}}$). The WD in Fig. 15.3.2(c) is not matched to the power impulses, displaying complicated cross-terms. The affine-smoothed pseudo WD in Fig. 15.3.2(d) does not suppress all the cross-terms and has a larger loss of (t,f) concentration than does the smoothed pseudo power WD in Fig. 15.3.2(b). Although all QTFRs in Fig. 15.3.2 are scale covariant, the results of the two power QTFRs with $\kappa = 3$ in Fig. 15.3.2(a) and (b) are better than those of the two affine QTFRs with $\kappa = 1$ in Fig. 15.3.2(c) and (d) because the former two are optimally matched to the $\kappa_{\text{signal}} = 3$ power law group delays of the power impulse signal components.

In order to further demonstrate the effect of mismatch in the signal parameter κ_{signal} and the QTFR power parameter κ , Fig. 15.3.2(e) and (f) shows the results obtained when analyzing the above signal using the power WD and a smoothed pseudo power WD with QTFR power parameter $\kappa = 4$.

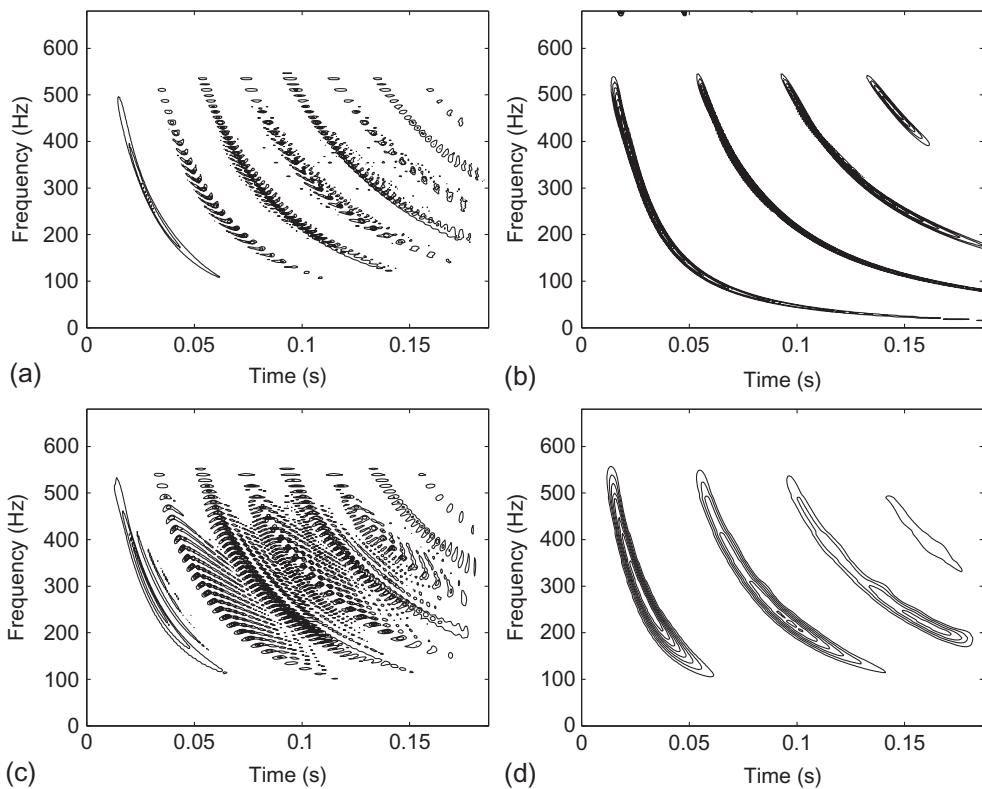
**FIGURE 15.3.2**

Power class analysis of a two-component analytic signal consisting of the sum of two windowed power impulses with signal power parameter $\kappa_{\text{signal}} = 3$. (a) Power WD with $\kappa = 3$, (b) smoothed pseudo power WD with $\kappa = 3$, (c) WD ($\kappa = 1$), (d) affine-smoothed pseudo WD ($\kappa = 1$), (e) power WD with $\kappa = 4$, and (f) smoothed pseudo power WD with $\kappa = 4$.

In Fig. 15.3.2(e) and (f) the power parameter of the power class QTFRs, $\kappa = 4$, is different from that of the signal, $\kappa_{\text{signal}} = 3$. The smoothed pseudo power WD in Fig. 15.3.2(f) has better cross term suppression and better (t,f) concentration along the true group (spectral) delay than the affine-smoothed pseudo WD in Fig. 15.3.2(d) since the power parameter mismatch in Fig. 15.3.2(f) is smaller than in Fig. 15.3.2(d).

15.3.2.2 Real data analysis example

Next, we demonstrate the use of power class QTFRs for analyzing real data with dispersive (t,f) structure. Figure 15.3.3 shows two power class QTFRs with $\kappa = 0.35$ and two affine ($\kappa = 1$) QTFRs of the measured impulse response of a steel beam with rectangular cross section (data collected at

**FIGURE 15.3.3**

Power class analysis of a bandpass-filtered segment of the measured impulse response of a steel beam (sampling frequency 4096 Hz): (a) power WD with $\kappa = 0.35$; (b) smoothed pseudo power WD with $\kappa = 0.35$; (c) WD ($\kappa = 1$); and (d) affine-smoothed pseudo WD ($\kappa = 1$).

Cambridge University) [35]. The impulse response was obtained by lightly tapping one end of the steel beam in the direction orthogonal to the flat side of the beam. Bending waves travel along the beam until they are reflected at the free end. They return to the point of impact, are reflected again, etc., thereby producing a series of echoes with increasing dispersion. The QTFRs in Fig. 15.3.3 display a bandpass-filtered segment of the measured impulse response. The smoothed pseudo power WD with $\kappa = 0.35$ in Fig. 15.3.3(b) shows better resolution and/or cross term suppression than the other three QTFRs depicted. The specific value of $\kappa = 0.35$ was chosen empirically to match the (t,f) curvature of the primary reflection. Other applications can be found in Ref. [36].

15.3.3 SUMMARY AND CONCLUSIONS

This section presented QTFR classes specifically matched to signals and systems with power law group (or spectral) delay characteristics. These *power class* QTFRs preserve scale changes and power law

frequency-dependent time shifts of the signal under analysis. Thus, these QTFRs are potentially useful in applications where a propagation medium causes power dispersive time shifts. The implementation of power class QTFRs can be based on a warping transformation that relates the κ th power class with the affine class. Successful application of power class QTFRs presupposes sufficient *a priori* knowledge about the signal to aid in choosing the appropriate power parameter κ .

15.4 THE 2D WIGNER-VILLE DISTRIBUTION APPLIED TO IMAGE DISTORTION ESTIMATION⁰

In this section the 2D Wigner-Ville distribution (WVD) and its properties are reviewed and analyzed in the joint space/spatial-frequency domain. The advantages and limitations of the 2D-WVD and practical implementation aspects are discussed in the context of image distortion analysis. A WVD-based image quality metric relates to basic characteristics of the human visual system (HVS), namely spatial frequency sensitivity and nonlinear visual sensitivity to energy distribution. Results show that for some images and common distortions, the WVD-based metric is consistent with subjective human evaluation and offers good performance but at the cost of high computational complexity. WVD-based perceptual masking models can make the metric more competitive with other state-of-the-art metrics.

15.4.1 IMAGE QUALITY AND SPATIAL/SPATIAL-FREQUENCY REPRESENTATIONS

The WVD (see [Section 2.1.4](#)) was introduced as a tool for analyzing 1D nonstationary signals. Subsequently, WVDs were successfully applied in capturing essential nonstationary image structures [37–39], using properties of joint spatial/spatial-frequency representations of images [40]. This section presents the WVD from the point of view of image analysis with illustration on a specific application, namely computing an image quality measure for analyzing and tracking image distortions. This simple WVD-based distortion measure does not take into account the masking effect but still yields good results. The qualitative correlation of the WVD-based measure with subjective human evaluation is illustrated through experimental results. This measure could be used in various applications such as image segmentation, bilinear problems in computer vision, and for image coder optimization.

15.4.2 CONTINUOUS 2D WIGNER-VILLE DISTRIBUTION

Given an image $p(x, y)$, where x and y are the spatial coordinates, the associated 2D-WVD is given by

$$W_p(x, y; u, v) = \int_{\mathbb{R}^2} p\left(x + \frac{\alpha}{2}, y + \frac{\beta}{2}\right) p^*\left(x - \frac{\alpha}{2}, y - \frac{\beta}{2}\right) e^{-j2\pi(\alpha u + \beta v)} d\alpha d\beta, \quad (15.4.1)$$

where u and v are spatial frequencies, and the asterisk denotes complex conjugation. Note that the function $W_p(x, y; u, v)$ is called the “Wigner distribution” of p , while the “Wigner distribution”

⁰Authors: **A. Beghdadi**, Université Paris 13, Villetaneuse, France (azeddine.beghdadi@univ-paris13.fr), **R. Iordache**, GE Healthcare, 78530 Buc, France (razvan.iordache@ge.com), and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: H. Wechsler, K. Abed-Meraim, and A. Saleem.

of an analytic image associated to p is referred to as the “Wigner-Ville distribution” (WVD) of p (see Chapter 2). For simplicity, in this section only the “WVD” appellation is used, as defined in Eq. (15.4.1).

This 4D function provides a joint spatial/spatial-frequency representation of the image. A 2D spatial-frequency spectrum is assigned to each point (x, y) of the image. The WVD can also be written in terms of $P(u, v)$, the 2D FT of $p(x, y)$, as

$$W_p(x, y; u, v) = \int_{\mathbb{R}^2} P\left(u + \frac{\eta}{2}, y + \frac{\xi}{2}\right) P^*\left(u - \frac{\eta}{2}, y - \frac{\xi}{2}\right) e^{j2\pi(\eta x + \xi y)} d\eta d\xi. \quad (15.4.2)$$

The image signal can be reconstructed up to a sign ambiguity from its WVD:

$$p(x, y) p^*(0, 0) = \int_{\mathbb{R}^2} W_p\left(\frac{x}{2}, \frac{y}{2}; u, v\right) e^{j2\pi(xu+yv)} du dv \quad (15.4.3)$$

(cf. Eq. (3.1.13), p. 105). Among the various properties of 2D WVD, the most important one for image processing applications is that it is always a real-valued function and, at the same time, contains the phase information. Other properties include translation, modulation, scaling, and convolution. For details, see Ref. [41].

The WVD localizes the spatial/spatial-frequency structures of $p(x, y)$. For instance, if the signal energy is concentrated in space around (x_0, y_0) and in frequency around (u_0, v_0) , then the energy of $W_p(x, y; u, v)$ is centered at $(x_0, y_0; u_0, v_0)$ and has the same spatial and frequency spread as $p(x, y)$. Actually, it can be shown that the spatial-frequency support of $W_p(x, y; u, v)$ is included in the support interval of $P(u, v)$, for all (x, y) , and that the spatial support of $W_p(x, y; u, v)$ is included in the support interval of $p(x, y)$ for all (u, v) .

The WVD is often considered as the image energy distribution in the joint spatial/spatial-frequency domain. Indeed, the local power at a given position (x_0, y_0) is equal to the marginal integral of the 2D WVD over the spatial-frequency domain at the considered spatial position:

$$|P(x_0, y_0)|^2 = \int_{\mathbb{R}^2} W_p(x_0, y_0; u, v) du dv. \quad (15.4.4)$$

Similarly, the spectral energy density at spatial frequency (u_0, v_0) is given by

$$|P(u_0, v_0)|^2 = \int_{\mathbb{R}^2} W_p(x, y; u_0, v_0) dx dy. \quad (15.4.5)$$

These properties allow the use of 2D WVD in image analysis applications where the spatial/spatial-frequency features of images are of interest.

The main drawback which limits the use of WVD is the cross-term interference (see Section 4.2 for the one-dimensional case) and its computational complexity. Indeed, due to its bilinearity property, the WVD of the sum of two images p_1 and p_2 contains an interference term corresponding to the cross-WVD term:

$$W_{p_1+p_2}(x, y; u, v) = W_{p_1}(x, y; u, v) + W_{p_2}(x, y; u, v) + 2 \operatorname{Re} \{W_{p_1, p_2}(x, y; u, v)\}. \quad (15.4.6)$$

This interference term W_{p_1, p_2} is given by

$$W_{p_1, p_2}(x, y; u, v) = \int_{\mathbb{R}^2} p_1\left(x + \frac{\alpha}{2}, y + \frac{\beta}{2}\right) p_2^*\left(x - \frac{\alpha}{2}, y - \frac{\beta}{2}\right) e^{-j2\pi(\alpha u + \beta v)} d\alpha d\beta. \quad (15.4.7)$$

These undesirable artifacts are highly oscillatory and their power could be larger than that of the useful signal. They can be classified [38] into two categories: (a) interference terms due to the interaction between the conjugate symmetric spectral components of real images, and (b) interference terms due to the interaction between any two spectral components of the image (real or complex) at different spatial frequencies. The WVD of any real image suffers from the first type of artifacts, due to its conjugate symmetric spectrum. If the real image is multicomponent, as in most applications, both types of artifacts are present, making the spatial/spatial-frequency representation difficult to interpret. The following section presents a solution to eliminate the first type of interference and to reduce the second.

15.4.2.1 Analytic image

If the spectrum of an image contains only positive (or only negative) frequency components, there is no interference between the positive and negative frequencies in its WVD. The analytic image has such a spectral property; therefore, a cleaner spatial/spatial-frequency representation of a real image, $p(x, y)$, may be obtained by computing the WVD of an analytic image $z(x, y)$ associated with it.

The analytic image is an extension of 1D analytic signal to 2D. Unlike the 1D analytic signal, the analytic image is not unique. For a reliable spatial/spatial-frequency representation of the real image, the analytic image is chosen so that

- (a) the useful information from the 2D WVD of the real signal is found in the 2D WVD of the analytic image, and
- (b) the 2D WVD of the analytic image minimizes the interference effect.

As in 1D case where the analytic image is formed by suppressing the negative frequencies, the 2D analytic image may be constructed by setting to zero one half-plane of the spatial-frequency domain. By introducing a reference direction $(\cos \theta, \sin \theta)$, a pair of frequency (u, v) is called positive if $u \cos \theta + v \sin \theta > 0$ and is called negative if $u \cos \theta + v \sin \theta < 0$.

The Fourier transform with respect to direction $(\cos \theta, \sin \theta)$, of the analytic image corresponding to the real image p , is given by Ref. [42]

$$Z^{(\theta)}(u, v) = P(u, v)(1 + \text{sign}(u \cos \theta + v \sin \theta)). \quad (15.4.8)$$

The choice of θ depends on knowledge of spectral characteristics of the real image.

Unfortunately, the associated WVD $W_{Z^{(\theta)}}$ contains interference cross-terms between the spectral components of the two quadrants in the positive frequency half-plane. To eliminate these cross-terms a second definition of the analytic image seems appropriate, as a complex signal with single-quadrant spectrum [43]. To avoid any loss of information, two analytic images are associated to the real image. The single-quadrant spectra $Z_1(u, v)$ and $Z_2(u, v)$, whose 2D IFTs are the analytic images $z_1(x, y)$ and $z_2(x, y)$ corresponding to the real image p , are defined as

$$Z_1(u, v) = P(u, v)(1 + \text{sign}(u))(1 + \text{sign}(v)), \quad (15.4.9)$$

$$Z_2(u, v) = P(u, v)(1 + \text{sign}(u))(1 - \text{sign}(v)). \quad (15.4.10)$$

Using z_1 and z_2 , the interference between frequencies from different quadrants is eliminated for any spectral distribution of the real signal. W_{z_1} reflects the spatial/spatial-frequency distribution of p for $uv > 0$, and W_{z_2} for $uv < 0$.

15.4.2.2 Continuous 2D windowed Wigner-Ville distribution

As most images are of finite support it is appropriate to apply the WVD to windowed images. As in the 1D case [44], to compute the windowed WVD at a spatial location (x_0, y_0) , we use a windowed version of p with the window w centered on (x_0, y_0) :

$$p_w(x, y; x_0, y_0) = p(x, y) w(x - x_0, y - y_0). \quad (15.4.11)$$

Then the windowed WVD at (x_0, y_0) is defined as the WVD of p_w

$$\begin{aligned} W_p^{(w)}(x, y; u, v) &= \int_{\mathbb{R}^2} w\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) w^*\left(-\frac{\alpha}{2}, -\frac{\beta}{2}\right) \\ &\quad \cdot p\left(x + \frac{\alpha}{2}, y + \frac{\beta}{2}\right) p^*\left(x - \frac{\alpha}{2}, y - \frac{\beta}{2}\right) e^{-j2\pi(\alpha u + \beta v)} d\alpha d\beta. \end{aligned} \quad (15.4.12)$$

at $x = x_0, y = y_0$. The 2D windowed WVD can be shown to represent the convolution with respect to the spatial frequencies of W_p and W_w :

$$W_p^{(w)}(x, y; u, v) = \int_{\mathbb{R}^2} W_p(x, y; \gamma, \sigma) W_w(x, y; u - \gamma, v - \sigma) d\gamma d\sigma. \quad (15.4.13)$$

The effect of windowing is to smear the WVD in the frequency plane only, so that the frequency resolution is decreased but the spatial resolution is unchanged.

15.4.3 DISCRETE 2D WIGNER-VILLE DISTRIBUTION

For simplicity, we consider square shaped images and analysis windows. Let $p(m, n)$ be the discrete image obtained by sampling $p(x, y)$, adopting the convention that the sampling period is normalized to unity in both directions. The associated windowed WVD (15.4.13) is given by

$$W_p^{(w)}(m, n; u, v) = 4 \sum_{r=-L}^L \sum_{s=-L}^L K(m, n; r, s) e^{-j2\pi(ru+sv)}, \quad (15.4.14)$$

where

$$K(m, n; r, s) = w(r, s) w^*(-r, -s) p(m + r, n + s) p^*(m - r, n - s) \quad (15.4.15)$$

and $w(r, s) = 0$ for $|r|, |s| > L \in \mathbb{Z}^+$.

The properties of the discrete-space WVD are similar to the continuous WVD, except for the periodicity in the frequency variables, which is one-half the sampling frequency in each direction:

$$W_p^{(w)}(m, n; u, v) = W_p^{(w)}\left(m, n; u + \frac{k}{2}, v + \frac{l}{2}\right) \quad \forall m, n, k, l \in \mathbb{Z}. \quad (15.4.16)$$

Therefore, if $p(x, y)$ is a real image, it should be sampled at twice the Nyquist rate to avoid aliasing effects in $W_p^{(w)}(m, n; u, v)$.

There are three practical solutions to avoid aliasing, when the original continuous image is not available:

- (a) to filter out the frequency regions that cause aliasing (lowpass filtering);
- (b) to up-sample the image by a factor of 2 in both directions and then appropriately band-limit the image using an anti-aliasing filter [37]; and

- (c) to compute the WVD of one or more analytic images associated with the real image, which eliminates or at least reduces aliasing [38,41].

To have a discrete spatial-frequency representation of the WVD, the frequency variables should be sampled. The 2D discrete windowed WVD of an image $p(m, n)$, for $(m, n) \in \mathbb{Z}^2$, is defined as

$$W_p^{(w)}(m, n; u_a, v_b) = 4 \sum_{r=-L}^L \sum_{s=-L}^L K(m, n; r, s) W_4^{ra+sb}, \quad (15.4.17)$$

where $N = (2L+2)$, $W_4 = e^{-j4\pi/N}$, and the normalized spatial frequency pair is $(u_a, v_b) = (a/N, b/N)$. By making a periodic extension of the kernel $K(m, n, r, s)$, for fixed (m, n) , Eq. (15.4.17) can be transformed to match the standard form of a 2D DFT, except that the twiddle factor is W_4 instead of W_2 (see Ref. [44] for more details for 1D case; the 2D construction is a direct extension). Thus standard FFT algorithms can be used to calculate the discrete $W_p^{(w)}$. The additional power of 2 represents a scaling along the frequency axes and can be neglected in calculations.

As real scene images have rich frequency content, interference cross-terms may mask the useful components' contributions. A common method to reduce the interference in image analysis applications is to smooth the 2D discrete windowed WVD in the spatial domain using a smoothing window $h(m, n)$. This comes at a cost of a reduction in the spatial resolution. The result is the 2D discrete smoothed windowed WVD [37]:

$$W_p(m, n; u_a, v_b) = 4 \sum_{k=-M}^M \sum_{l=-M}^M h(k, l) W_p^{(w)}(m + k, n + l; u_a, v_b). \quad (15.4.18)$$

In the case of a symmetric spatial-shift window, i.e., $w(r, s) = w(-r, -s)$, the WVD can be defined by

$$W_p(m, n; u_a, v_b) = 4 \sum_{r=-L}^L \sum_{s=-L}^L \left[|w(r, s)|^2 W_4^{ra+sb} \sum_{k=-M}^M \sum_{l=-M}^M h(k, l) p(m + k + r, n + l + s) p^*(m + k - r, n + l - s) \right]. \quad (15.4.19)$$

An important aspect to consider when using WVD is the choice of the two windows, $w(r, s)$ and $h(k, l)$. The size of the first window, $w(r, s)$, is dictated by the resolution required in the spatial-frequency domain. The spectral shape of the window should be an approximation of the delta function that optimizes the compromise between the central lobe width and the side lobes height. A window that complies with these demands is the 2D extension of Kaiser window, used in Ref. [37].

The role of the second window, $h(k, l)$, is to allow spatial averaging. Its size determines the degree of smoothing. The larger the size of the window, the lower is the spatial resolution. The common choice for this window is the rectangular window.

15.4.3.1 Choice of analytic image in computing discrete WVDs

In the discrete case, the analytic image(s) must eliminate aliasing effect and preserve all the information of the real image. Because no *single* analytic image satisfies both requirements, either one analytic image is used and some aliasing is allowed, or more analytic images are employed which satisfy two

restrictions: (a) the real image can be perfectly reconstructed from the analytic images, and (b) each analytic image is alias free with respect to the WVD.

The discrete version of the half-plane analytic image (15.4.8) was used for texture segmentation in Ref. [38]. While preserving all the information of the real image, it fails to produce a substantial reduction in aliasing, as it does not prevent frequency fold-over along one direction.

One way to avoid aliasing is to use two analytic images, obtained by splitting the region of support of the half-plane analytic image into two equal-area subregions. This is the procedure presented in Ref. [41], where one analytic image contains the low positive frequencies and the other the high positive frequencies. Although this method requires the computation of two WVDs, no aliasing artifacts appear and the WVD of the analytic images can be combined to produce a spatial/spatial-frequency representation of the real image having the same frequency resolution and support as the original real image. This approach was successfully applied in texture analysis and segmentation in Ref. [39].

The energy of most natural images is concentrated at low frequencies, in which case the low-frequency analytic image contains most of the energy of the real image, and therefore its WVD exhibits strong cross-term interferences. A more balanced energy sharing is done when discrete single-quadrant analytic images are used; these are the discrete version of Eqs. (15.4.9) and (15.4.10). As the positive-frequency domain is split along v -axis, the inter-quadrant cross-terms are eliminated.

A full-domain WVD $FW_p(m, n; u_a, v_b)$ of the real image $p(m, n)$ can be constructed from $W_{z_1}(m, n; u_a, v_b)$ and $W_{z_2}(m, n; u_a, v_b)$ [41]. In the spatial-frequency domain, the full-domain WVD is, by definition, of periodicity 1 and symmetric with respect to the origin, as the WVD of a real image. It is fully specified by

$$FW_p(m, n; u_a, v_b) = \begin{cases} W_{z_1}(m, n; u_a, v_b), & 0 \leq u_a < \frac{1}{2}, 0 \leq v_b \leq \frac{1}{2}, \\ W_{z_2}(m, n; u_a, v_b), & 0 \leq u_a < \frac{1}{2}, 0 > v_b \geq -\frac{1}{2}, \end{cases} \quad (15.4.20)$$

$$FW_p(m, n; u_a, 0) = W_{z_1}(m, n; u_a, 0) + W_{z_2}(m, n; u_a, 0), \quad 0 \leq u_a < \frac{1}{2}, \quad (15.4.21)$$

$$FW_p(m, n; u_a, v_b) = FW_p(m, n; -u_a, -v_b), \quad 0 > u_a, v_b \geq -\frac{1}{2}, \quad (15.4.22)$$

$$FW_p(m, n; u_a + k, v_b + l) = FW_p(m, n; -u_a, -v_b) \quad \forall k, l, a, b \in \mathbb{Z}. \quad (15.4.23)$$

In conclusion, the full-domain WVD provides information unaltered by aliasing artifacts on the spatial/spatial-frequency distribution of the real image over the entire frequency spectrum. A potential drawback of these approaches is that additional sharp filtering boundaries may introduce ringing effects.

15.4.4 AN IMAGE QUALITY METRIC BASED ON THE 2D WVD

15.4.4.1 *Image quality assessment*

Many approaches have been proposed for objective evaluation of image quality [45]. Perceptual approaches are more promising than purely signal-based approaches. However, the use of perceptual vision models is complex and depends on many unpredictable parameters. Therefore, one must find a compromise by searching for simple models, which take into account only the most relevant properties of the human visual system (HVS) [46], such as its frequency and directional selectivity and its nonlinear response to visual stimuli.

The frequency-selectivity of the HVS can be modeled by using wavelet-based decomposition [47] or other directional subband decompositions such as cortex transforms. Image distortion measures based on these properties are given in Ref. [45]. It is generally acknowledged that the early visual processing stages involve the creation of a joint spatial/spatial-frequency representation of visual stimuli [47]. It has also been observed that localized and structured distortions are more annoying than random and evenly distributed distortions. These observations motivate the use of the WVD as a tool for analyzing image distortions.

The simple WVD-based distortion measure does not take into account the masking effect. This factor needs to be introduced in future work to make the metric more consistent with human visual perception. The qualitative correlation of the WVD-based measure with subjective human evaluation is illustrated through experimental results. This metric could be used for image quality assessment, or for image coding optimization, or for bilinear problems encountered in computer vision, or for image segmentation.

Structured distortions affecting an image, which are more annoying than the unstructured distortions, are usually highly concentrated in the spatial/spatial-frequency domain. For two distortions with the same energy, i.e., the same signal-to-noise ratio (SNR), the more disturbing is the one having a peaked energy distribution in the spatial/spatial-frequency plane.

In terms of the effect on the WVD, the noise added to an image not only influences the coefficients in the positions where the noise has nonzero WVD coefficients, but also induces cross-interference terms (see Eq. (15.4.16)). The stronger the noise WVD coefficients, the more important the differences between the WVD of the noisy image and the original image.

The WVD-based SNR of a distorted version $g(m, n)$ of the original discrete image $p(m, n)$ is defined as

$$\text{SNR}_W = 10 \log_{10} \frac{\sum_m \sum_n \max_{a,b} (|FW_p(m, n; u_a, v_b)|)}{\sum_m \sum_n \max_{a,b} (|FW_p(m, n; u_a, v_b) - FW_g(m, n; u_a, v_b)|)}. \quad (15.4.24)$$

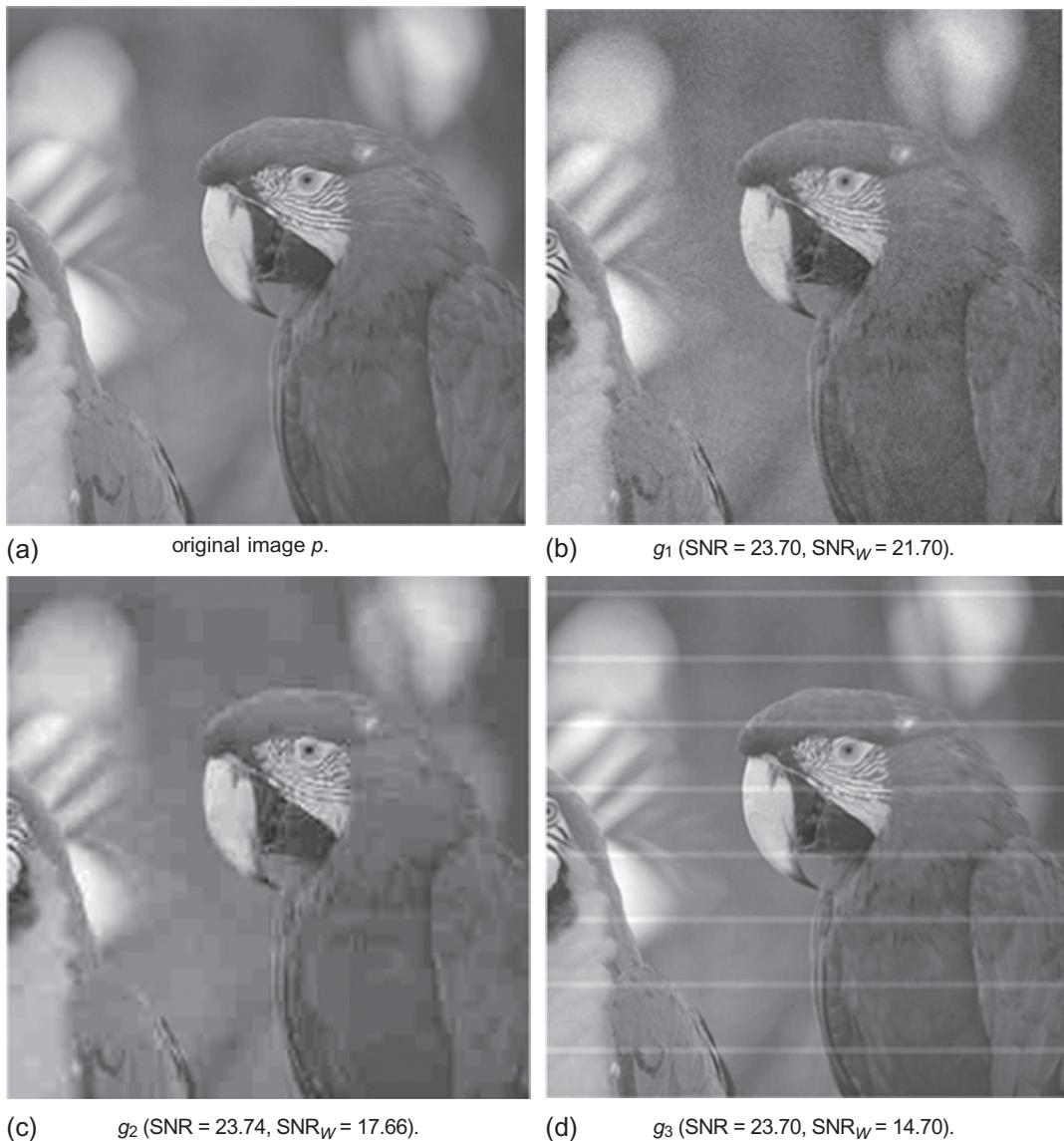
For each position (m, n) , only the highest-energy WVD component is retained, as if the contribution of the other components are masked by it. Of course the masking mechanisms are much more complex, but this coarse approximation leads to results which are more correlated to the HVS perception than SNR.

Let η_1 and η_2 be two degradations having the same energy. Suppose η_1 is additive white Gaussian noise and η_2 is an interference pattern. While the energy of the noise is evenly spread in the spatial/spatial-frequency plane, the energy of the structured degradation is concentrated in the frequency band of the interference. Thus the WVD of η_2 contains terms which have absolute values larger than any term of WVD of η_1 , as the two degradations have the same energy. These peak terms induce larger local differences between WVD of $g_2 = p + \eta_2$ and the WVD of p , which are captured by “max” operation in the denominator of Eq. (15.4.24) and lead to a smaller SNR_W for g_2 .

15.4.4.2 Results and discussion

To show the performance of the proposed image-distortion measure compared to the signal-to-noise ratio (SNR), three types of degradation are considered: white noise, an interference pattern, and JPEG coding (Fig. 15.4.1). They induce nearly the same SNR, whereas visual comparison clearly reveals different perceptual quality.

The WVD-based distortion measure correlates with subjective quality evaluation done by five nonexpert observers, who prefer the white-noise-distorted image to the interference-perturbed image

**FIGURE 15.4.1**

Distorted versions of 256×256 pixel parrot image: (a) original image p ; (b) g_1 is obtained by adding white Gaussian noise to p ; (c) g_2 is a JPEG reconstruction of p , with a quality factor of 88; (d) g_3 is the result of imposing a grid-like interference over p . The SNRs are given in dB.

and to the JPEG-coded image. The reason is that for random degradation the noise has the same effect in the entire spatial-frequency plane. Therefore, the maximum spectral difference at almost any spatial position is lower than the just-noticeable perceptual difference. On the other hand, when the distortion is localized (as interference patterns or distortion induced by JPEG coding), the maximum spectral difference corresponding to an important proportion of the pixels has a significant value, much larger than the just-noticeable perceptual difference.

15.4.5 SUMMARY AND CONCLUSIONS

The advantages and usefulness of the 2D WVD in image analysis are demonstrated by considering an application, namely distortion evaluation using the spatial/spatial-frequency representation of images. The efficiency of this measure is validated through experiments and informal visual quality-assessment tests. The results are consistent with basic knowledge of the HVS (the joint spatial/spatial-frequency representation and nonlinear inhibition models), showing that this simple image-distortion measure correlates with the perceptual evaluation. Further details on the use of the WVD for image analysis and processing can be found in Refs. [48,49].

15.5 EXTRACTING TIME-FREQUENCY FEATURES FROM PCG SIGNALS FOR MEDICAL DIAGNOSIS⁰

This section presents an application of high-resolution time-frequency $((t,f))$ techniques to extract useful information for medical diagnosis [50,51] with illustration to phonocardiogram (PCG) signals also known as heart sounds or murmurs [52].

15.5.1 TIME-FREQUENCY CHANGE DETECTION IN BIOMEDICAL SIGNALS

Past studies have shown that (t,f) concepts and pattern recognition methods can be merged to analyze, process, and classify nonstationary data [50,51,53]. This (t,f) pattern recognition approach leads to an improved methodology for detecting abnormalities or faults in a wide range of applications, including condition monitoring, fault detection, and a more general detection of system changes. The approach relies on the definition of (t,f) features [51] which are obtained by translation from time-domain-only or frequency-domain-only features (e.g., signal-related features); or they can be purely (t,f) domain features with no relation to time-domain or frequency domain (e.g., image-related features) [50]. These (t,f) features that characterize the nonstationary nature and the multicomponent characteristic of signals are extracted directly from the (t,f) distributions (TFDs) of signals. The signal-related features, derived by extending features to the (t,f) domain, include the instantaneous frequency (IF), singular-value decomposition, and energy-based features. The image-related features are extracted from the (t,f) representation, considered an image, using image-processing techniques. These combined signal and image features allow us to extract more information from a signal than could be possible using only

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **A. Djebbari**, Université Abou Bekr Belkaïd, Tlemcen, Algeria (a.djebbari@umontreal.ca). Reviewers: F. Berekci-Reguig, R. Mazhar, and S.H. Sheikh-Salleh.

1-D signal-processing methods. Studies have shown that such (t,f) features improve the performance in classification systems based on multiple support vector machine (SVM) systems (Section 12.6 and [50]). Further improvements can be made using an improved design of high-resolution TFDs that can enhance classification performance by defining data-dependent kernels [50].

Change detection within the (t,f) domain can be applied to biomedical signals such as PCG signals (i.e., the acoustic recording of heart sounds and murmurs on the chest of patients). Despite the prevalence of Doppler echocardiography (DE) for noninvasive evaluation of blood flow through the heart's chambers and valves, several studies reported DE inaccuracies in measuring intracardiac pressure [54]. The DE can overestimate or underestimate the pulmonary artery pressure (PAP) for both normal and abnormal intracardiac hemodynamics due to the poor Doppler imaging of the transtricuspid regurgitant blood flow. Furthermore, for right ventricular and PAP measurement in patients with pulmonary hypertension, DE is inaccurate when compared to the right-sided heart catheterization (RHC). The PCG signal has been shown to be a reliable tool in estimating the PAP in comparison to DE [55]. Furthermore, PCG signals of patients with valvular bioprostheses can be used as a noninvasive early degeneration prevention tool by using (t,f) analysis to detect any suspected pathological change [52,56].

15.5.2 PRINCIPLES OF DIAGNOSIS IN BIOMEDICAL AND PCG SIGNALS

15.5.2.1 Characteristics of PCG signals

Auscultation of the cardiac sounds of patients is a common clinical method used by physicians to investigate abnormalities of blood flow in and out of the beating heart. This acoustic monitoring is noninvasive but requires a long experience to accurately recognize each heart sound and murmur, especially for complex cardiovascular pathologies. Automatic change detection can be used as an objective diagnosis method, hence removing the subjectivity factor and dependence on the availability of experts.

During a cardiac cycle, the human heart generates up to four groups of sounds denoted as S1, S2, S3, and S4. The first two, S1 and S2, are the higher intensity, primary heart sounds that respectively mark the onset of systole and diastole of the cardiac cycle. The S1 heart sound occurs at the closure of the mitral (M1) and tricuspid (T1) valves. The S2 heart sound corresponds to the closure of the aortic (A2) and pulmonary (P2) valves.

The S3 and S4 heart sounds are of lower intensity than the primary heart sounds. They may occur in healthy or in diseased state. S3 occurs after S2, in the beginning of diastole, due to reverberations produced by in-rushing blood from the atria. The S4 sound (atrial gallop) is produced by blood being forced into a stiff ventricle.

The multidegree-of-freedom theory (MFT) considers the heart as a system that transforms the biochemical energy of the blood into mechanical, acoustical energy and heat [52]. This theory stipulates that the heart is a multidegree-of-freedom viscoelastic system assumed to generate heart sounds as a result of vibrations of its internal structures as well as turbulence of the blood mass. The produced heart sounds are therefore multicomponent signals; they are assumed to be the direct result of pressure fluctuations around the valves, leaflets, cordae tendinae, and myocardium, as well as turbulence in the intracardiac blood mass. Thus, PCG signals are formed by vibrations of intracardiac structures that are of different resonance modes. Studies have found that heart sounds are mainly formed by frequency-modulated (FM) chirps [52]. According to the MFT, any vibrations below 60 Hz are due to

the intracardiac structures motions as a whole owing to the incompressibility of blood mass. Above 60 Hz, intracardiac structures (valve leaflets, cordae tendinae, etc.) can vibrate according to their own resonance modes.

15.5.2.2 Segmentation of PCG signals in systole and diastole

PCG signals can be segmented in systole and diastole phases by detecting the R-peak and the end of the T-wave of the simultaneously recorded electrocardiogram ECG signal. Detected systole and diastole segments are adjusted in the time domain through cross-correlation to avoid phase cancellation of valvular heart sounds [52]. PCG signals were recorded at the LGB-IRCM (Laboratoire de génie biomédical, Institut de recherches cliniques de Montréal) from the cardiac valve database; they were used to study systolic murmurs of bioprostheses implanted in the aortic orifice for patients with aortic stenosis (AS) [52]. This database is composed of simultaneously recorded PCG and ECG signals of 172 patients with a prosthetic valve implanted in aortic or mitral orifices. Each recording includes around 30 cardiac cycles. Figure 15.5.1 illustrates three cardiac cycles of a PCG/ECG recording of a normal subject from the LGB-IRCM cardiac valve database.

15.5.3 TIME-FREQUENCY ANALYSIS OF PCG SIGNALS

DE measurement of the PAP is often inaccurate [54]. The PCG signal is a noninvasive alternative to ensure an accurate measurement of the PAP [55]. Moreover, the PCG signal can reveal early valvular pathologies and prevent degeneration of intracardiac valvular prostheses. Due to their nonstationary multicomponent nature, PCG signals require (t,f) methods to detect pathological patterns occurrence such as heart murmurs during systole or diastole of cardiac phases.

The Wigner-Ville distribution (WVD) and its smoothed versions can yield high-resolution (t,f) representations, allowing us to conveniently quantify heart murmurs as they take into account both time and frequency characteristics jointly. Indeed, as formulated in Section 3.2, the smoothed pseudo WVD (SPWVD) (also called smoothed windowed WVD) with its separable kernel was used to

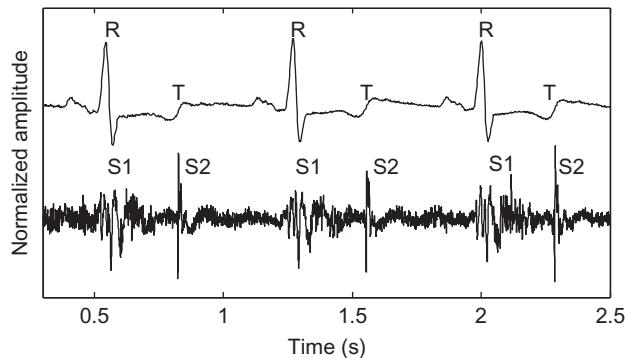


FIGURE 15.5.1

LGB-IRCM cardiac valve database: Normal ECG and PCG signals of the data file 10001 over three cardiac cycles.

detect valvular components of heart sounds within the (t,f) plane in Refs. [52,56]. But despite the improvement obtained by the SPWVD in reducing cross-terms, auto-terms are widened in the (t,f) domain. Therefore, a better trade-off should be considered between cross-terms reduction and (t,f) resolution. An improved separable (t,f) kernel, the extended modified B-distribution (EMBD) defined in Ref. [51] was shown to achieve better analysis of heart sounds and murmurs at a higher (t,f) resolution [53].

In this study, normal and abnormal PCG signals of the LGB-IRCM cardiac valve database are analyzed with the EMBD to show its ability to better detect abnormal changes in pathological conditions.

15.5.3.1 Extended modified B-distribution (EMBD)

The EMBD is a member of the general class of quadratic (t,f) distributions (QTFDs) with a separable (t,f) kernel. Such QTFDs are formulated as (see Eqs. (2.2.15) and (3.2.13) in [Chapters 2](#) and [3](#), respectively)

$$\rho_z(t,f) = \gamma(t,f) \underset{(t,f)}{\ast\ast} W_z(t,f), \quad (15.5.1)$$

where $W_z(t,f)$ denotes the WVD of the analytic version $z(t)$ of the analyzed signal. The (t,f) kernel $\gamma(t,f)$ defining the selected QTFD is associated with a Doppler-lag kernel $g(\nu, \tau)$ through direct and inverse FTs as

$$\gamma(t,f) = \mathcal{F}_{t \leftarrow \nu}^{-1} \left\{ \mathcal{F}_{\tau \rightarrow f} \{g(\nu, \tau)\} \right\}. \quad (15.5.2)$$

By replacing the expression of the (t,f) kernel of Eq. (15.5.2) in Eq. (15.5.1), QTFDs can be written as an inverse and direct FT combination of the symmetric ambiguity function (SAF) $A_z(\nu, \tau)$ smoothed or filtered by a Doppler-lag kernel $g(\nu, \tau)$ as

$$\rho_z(t,f) = \mathcal{F}_{t \leftarrow \nu}^{-1} \left\{ \mathcal{F}_{\tau \rightarrow f} \{g(\nu, \tau) A_z(\nu, \tau)\} \right\}. \quad (15.5.3)$$

The SAF $A_z(\nu, \tau)$ is defined as the FT of the IAF $K_z(t, \tau)$ as

$$A_z(\nu, \tau) = \mathcal{F}_{t \rightarrow \nu} \{K_z(t, \tau)\} = \int_{-\infty}^{+\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\nu t} dt. \quad (15.5.4)$$

QTFDs with a separable kernel have the following kernels:

$$\gamma(t,f) = g_1(t) \cdot G_2(f) \stackrel{2\text{DFT}}{\Rightarrow} g(\nu, \tau) = G_1(\nu) \cdot g_2(\tau), \quad (15.5.5)$$

with $g_1(t) = \mathcal{F}_{t \leftarrow \nu}^{-1} \{G_1(\nu)\}$ and $G_2(f) = \mathcal{F}_{\tau \rightarrow f} \{g_2(\tau)\}$.

The modified B-distribution (MBD) is a quadratic TFD with a lag-independent kernel. Its Doppler-lag kernel $g(\nu, \tau) = G_1(\nu)$ with $g_2(\tau) = 1$ is a time-only kernel with a Doppler low-pass filter, allowing a smoothing of the TFD in the time direction. The separable time-lag kernel $G(t, \tau)$ of the MBD is given in [Chapters 2](#) and [3](#) (now renamed $g_1(t)$) and expressed as

$$g_1(t) = \frac{\cosh^{-2\beta} t}{\int_{-\infty}^{+\infty} \cosh^{-2\beta} \xi d\xi}. \quad (15.5.6)$$

The denominator term ensures normalization of the kernel. Lower values of the smoothing control-parameter β ensures adequate smoothing for cross-terms.

The EMBD is an improvement of the MBD; it extends the above kernel to the frequency domain [53] using the $G_2(f)$ window defined as

$$G_2(f) = \frac{\cosh^{-2\alpha} f}{\int_{-\infty}^{+\infty} \cosh^{-2\alpha} \xi \, d\xi}. \quad (15.5.7)$$

This formulation is similar to g_1 , which is given in Eq. (15.5.6), where α and β are respectively smoothing parameters in the time and frequency domains. The full kernel of the EMBD is therefore given by Refs. [50,51]

$$g(\nu, \tau) = \frac{|\Gamma(\beta + j\pi\nu)|^2}{\Gamma^2(\beta)} \frac{|\Gamma(\alpha + j\pi\nu)|^2}{\Gamma^2(\alpha)}. \quad (15.5.8)$$

Studies have shown that this 2D (t,f) smoothing formulation ensures a more flexible (t,f) analysis by selecting appropriate values for α and β .

15.5.3.2 Time-frequency change detection in abnormal PCG signals

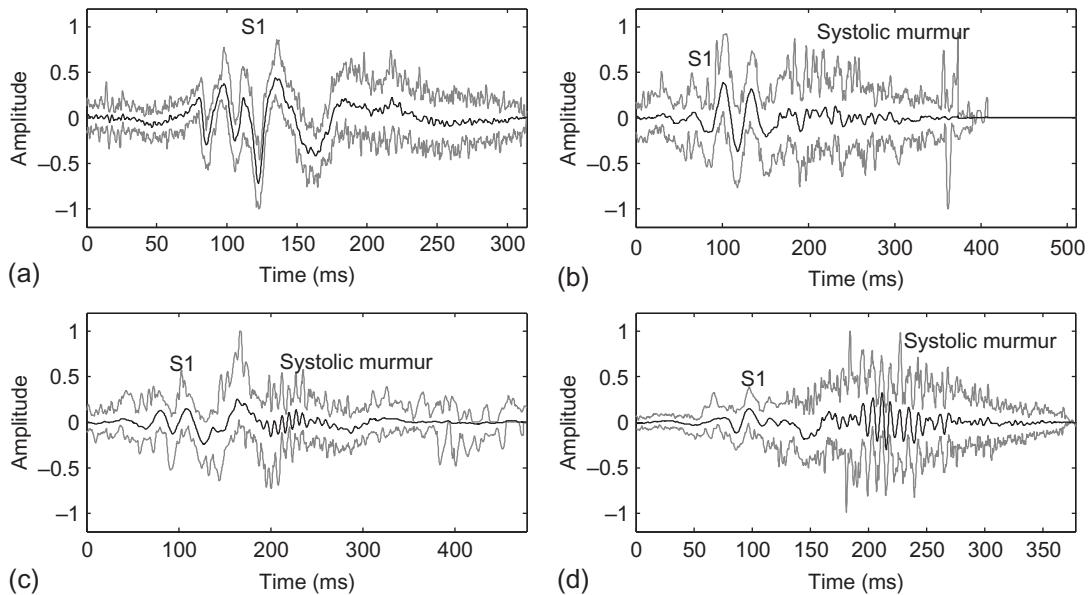
Valvular prostheses are maintained through regular medical consultations in a cardiology department to detect their degeneration into either stenosis, regurgitation, or a combination of both malfunctions. Such stenosis and regurgitation through intracardiac valves yields an additional murmur recorded within PCG signals during the systole or diastole phase as a result of blood ejection across the obstructed or regurgitant valvular orifice. This murmur is the abnormal change to be detected within PCG signals using (t,f) techniques resulting in a useful noninvasive diagnosis methodology. The nonstationary content of heart sounds and murmurs is conveniently handled by QTFDs such as the EMBD.

This study considers PCG signals from the LGB-IRCM cardiac valve database of patients diagnosed with AS following aortic valve replacement (AVR) with bioprostheses. The AS consists of a reduction of the effective orifice area (EOA) of the aortic valve. Blood passing through the obstructed aortic valve generates a systolic murmur, which occurs after the S1 heart sound. This murmur is of higher frequency content in comparison to S1 with an amplitude varying according to stenosis severity. Three severity levels at mild, moderate, and severe AS are selected to show the performance of the EMBD in detecting systolic heart murmurs within the (t,f) plane at different severity levels.

The first patient (data file 10078) has a St. Jude Medical (SJM) bioprosthesis with a mild AS; the second patient (data file 20008) has a Carpentier-Edwards bioprosthesis with a moderate AS, and the third patient (data file 10019) has a noncommon bioprosthesis with a severe AS. Since all three patients were diagnosed with AS, a systolic murmur occurs at different intensities according to severity as illustrated in averaged and cross-correlation-adjusted systole segment in Fig. 15.5.2. Both of these PCG signals were recorded from the aortic auscultation area to get the adequate intensity for systolic murmurs.

15.5.3.2.1 Ambiguity functions

For verification and validation, Fig. 15.5.3 shows the symmetric ambiguity functions of averaged and time-adjusted systole phases of normal and abnormal PCG signals. In all cases, auto-terms are concentrated at the origin of the Doppler-lag domain, whereas cross-terms are scattered away from the

**FIGURE 15.5.2**

LGB-IRCM cardiac valve database: systole phases of (a) normal subject; (b) mild AS (St. Jude bioprosthetic); (c) moderate AS (Carpentier-Edwards bioprosthetic); (d) severe AS (noncommon bioprosthetic type). The gray waveforms represent the envelope of adjusted systole phases of the overall cardiac cycles, whereas the black waveforms represent the average of cross-correlation-adjusted systole phases of the segmented PCG signals.

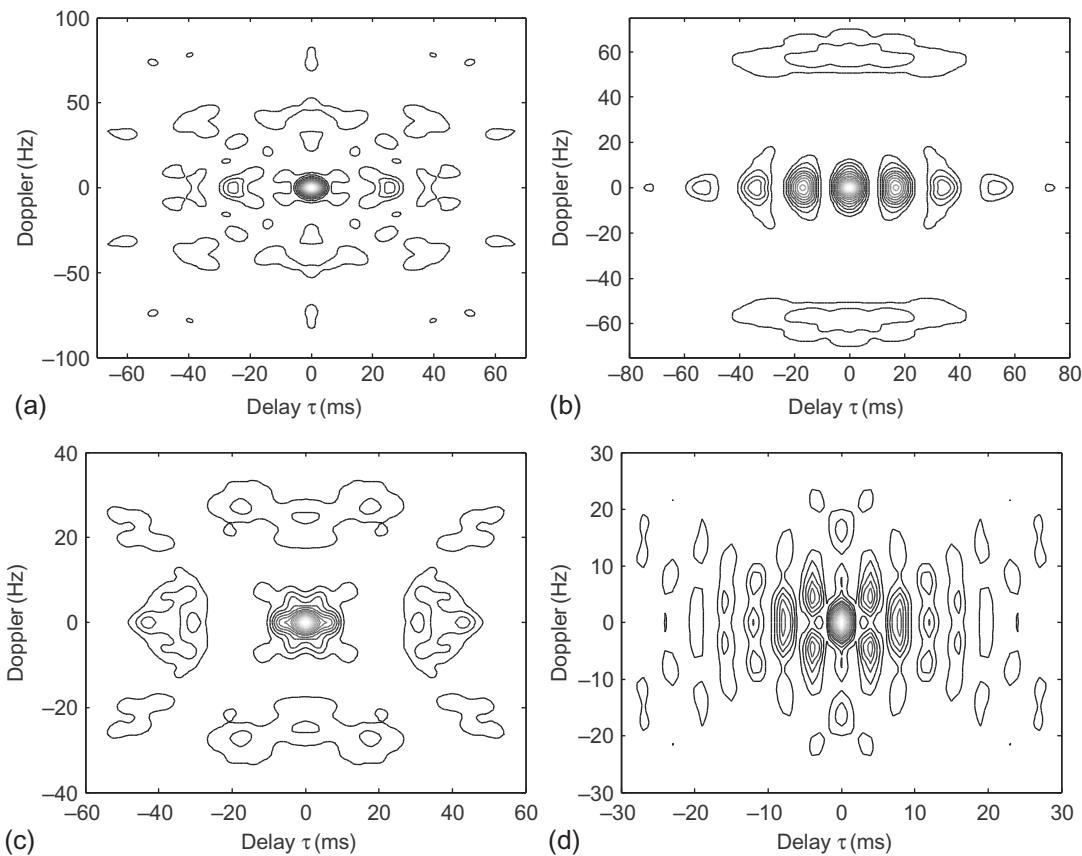
origin. The EMBD uses g_1 and g_2 as smoothing functions in the time and frequency directions over the (t,f) domain to obtain the clear TFD plots shown in Fig. 15.5.4.

15.5.3.3 Example: (t,f) detection of aortic stenosis systolic murmurs

Let us consider the (t,f) detection of heart murmurs for the overall abnormal PCG signals of the LGB-IRCM cardiac valve database; TFDs of averaged and cross-correlation adjusted systole phases for normal and AS PCG signals are shown in Fig. 15.5.4 for comparison purposes. The segmentation process used ECG signals as a reference; this involves detecting the R-peak and the end of the T-wave to delimit the systole and diastole phases within a cardiac cycle [57].

Figure 15.5.4 shows time-frequency representations (TFRs) obtained using the EMBD with smoothing parameters $\alpha = \beta = 0.01$. The TFD in Fig. 15.5.4(a) represents averaged and adjusted systole phases of the normal PCG signal (data file: 10001) and highlights the hyperbolic chirp behavior of valvular components of the S1 heart sound, namely M1 and T1 of the mitral and tricuspid valves, respectively. These valvular closure components are formed by hyperbolic FM laws within the (t,f) plane. The M1 valvular component occurs at higher frequency in comparison to the T1 component.

In Fig. 15.5.4(c) and (d), the systolic murmur is detected at around 130 Hz for both moderate and severe AS at their respective severity levels, whereas the mild AS (Fig. 15.5.4(b)) does not show any noticeable power (in this studied case). The results indicate that the detected change within the (t,f)

**FIGURE 15.5.3**

LGB-IRCM cardiac valve database: ambiguity functions of averaged systole phases of (a) normal heart sounds; (b) mild AS (St. Jude bioprosthetic); (c) moderate AS (Carpentier-Edwards bioprosthetic); (d) severe AS (noncommon bioprosthetic type).

plane corresponds to an additional abnormal heart murmur generated by deteriorated bioprosthetic valves for AS at mild, moderate, and severe AS. These (t,f) changes can be quantified and used as a reliable diagnosis tool in preventing their degeneration. So, there is a need for a more detailed study that can justify and assess in detail the actual improvements made by using a (t,f) approach rather than a conventional spectral approach, the criteria being improved diagnostic compared to the labeled database.

The results also indicate that the EMBD provides a high (t,f) resolution which is reliable in detecting the valvular split within S1 and S2 heart sounds between their respective valvular components. Indeed, the FM behavior is clearly illustrated in Fig. 15.5.4 for TFDs of both normal and abnormal PCG signals.

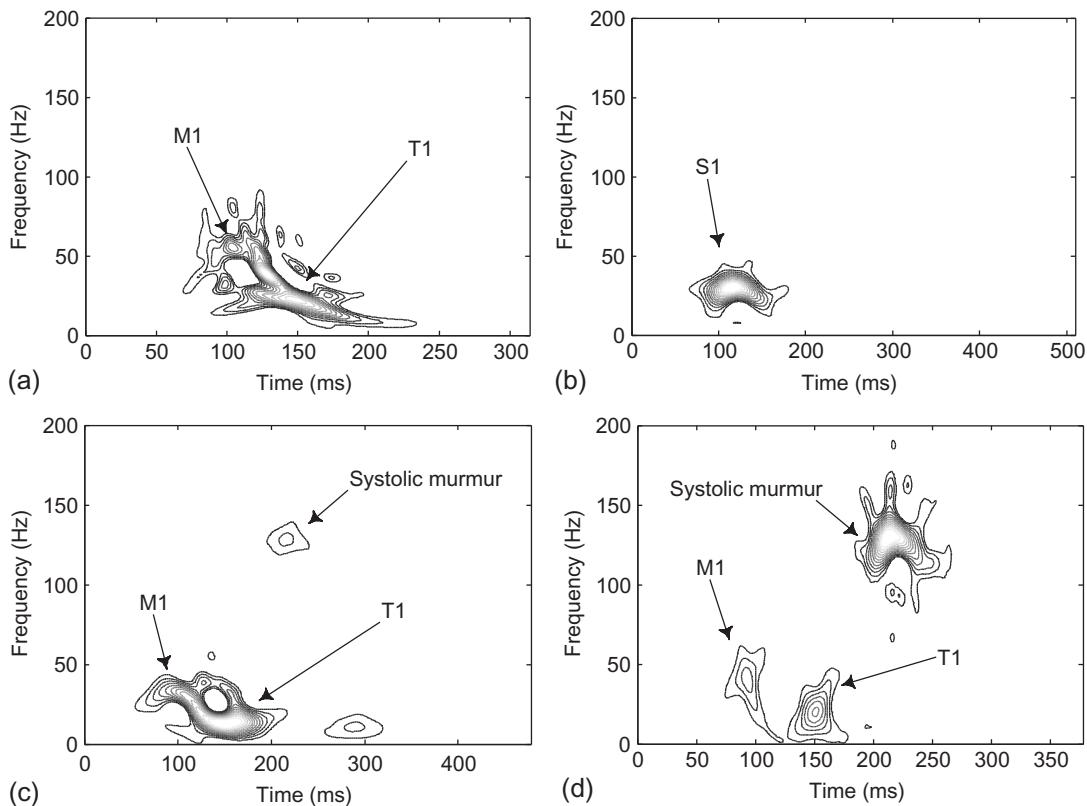


FIGURE 15.5.4

LGB-IRCM cardiac valve database: EMBDs ($\alpha = 0.01$ and $\beta = 0.01$) of averaged systole phases of (a) normal heart sounds; (b) mild AS (St. Jude bioprosthetic); (c) moderate AS (Carpentier-Edwards bioprosthetic); (d) severe AS (noncommon bioprosthetic type).

15.5.4 SUMMARY AND CONCLUSIONS

Degeneration of the valvular prostheses is commonly detected through regular medical consultations. Early detection of such malfunction is important for proper medical/surgical management to preserve cardiac function. This study validates the use of quadratic TFDs to study early change detection in PCG signals. As an improved separable kernel (t,f) distribution, the EMBD shows a good performance in the detection of heart murmurs within the (t,f) plane for all abnormal PCG signals of the studied database. As an illustration, systolic heart murmurs at mild, moderate, and severe AS were conveniently detected and classified by the EMBD at their respective severity levels.

Detected systolic murmurs generated by bioprosthetic aortic valves have high frequency in comparison to S1 heart sound. A resonant mode of ejected blood flow is detected at around 130 Hz for both AS cases at different power levels correlated with the stenosis severity. The early change detection in PCG signals as heart murmurs accurately localized within the (t,f) region can therefore

be a significant tool toward prevention of valvular prostheses degeneration. Further improvements can be made using data-dependent TFDs designed following the methodology described in [Chapter 3](#) and [Section 13.5](#).

The (t,f) analysis achieved on the PCG signals of the LGB-IRCM cardiac valve database confirmed the multidegree of freedom genesis theory of heart sounds and murmurs, which stipulates that heart sounds are mainly caused by pressure fluctuations around the valves, leaflets, and therefore are mainly formed by chirps. Such (t,f) analysis can be also applied to other valvular abnormalities such as stenosis or regurgitation affecting the four intracardiac valves.

15.6 DIAGNOSIS AND CONDITION MONITORING USING TIME-FREQUENCY PATTERN RECOGNITION AND MACHINE LEARNING⁰

15.6.1 CONDITION MONITORING VS. FAULT DIAGNOSIS

Machine-condition monitoring is the process of checking whether a machine has abnormal symptoms. Fault diagnosis, on the other hand, means deciding the nature and the cause of the fault by examining the symptoms obtained by processing the collected data [58]. This section aims to provide a time-frequency (t,f) methodology for implementing techniques suitable for the efficient interpretation of the symptoms relevant to machine-condition monitoring. It provides two examples and relevant references. It focuses on one-dimensional time-domain signals, but the methodology presented can be extended to 2D signals and images.

15.6.2 MACHINE-CONDITION MONITORING

In modern manufacturing, the quest for automation, speed, and flexibility has resulted in the construction of machines performing extremely complex processes with their performance dependent on the trouble-free operation of all components. When a fault occurs, it is critical to detect it, isolate the causes, and take appropriate maintenance action at an early stage. Consequently, techniques were developed to monitor condition-relevant parameters.

Predictive maintenance through condition monitoring and diagnosis is intended to enhance product quality, improve worker safety, and reduce the costs of maintenance. This is achieved by (1) allowing an early detection of potentially catastrophic faults and (2) allowing the implementation of condition-based maintenance rather than periodic or failure-based maintenance.

An efficient condition-monitoring technique should be able to provide a timely warning and predict faults at early stages by obtaining information about the machine in the form of primary data. Digital signal processing (DSP) methods then allow us to extract the critical information from these data and correlate them to the condition of the machine. Effectiveness depends on the goodness of the matching between the DSP techniques and the characteristics of the monitored signals. If the signal characteristics are nonstationary, then (t,f) approaches become indispensable to reach maximum efficiency.

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **S. Ouelha**, Qatar University, Doha, Qatar (samir.ouelha@qu.edu.qa). Reviewers: D.M. Onchis and V. Sucic.

Systems for condition monitoring and diagnosis can be either offline or online. In an offline (periodic) monitoring system, the monitored signal is measured at preselected time intervals. In an online (permanent) monitoring system, the signal is continuously measured and compared with a reference level.

Whereas human operators tend to rely on a combination of sight and sound to perform condition monitoring, automatic techniques may rely on direct or indirect measurements. Direct methods use sensing techniques that directly measure the extent of the deterioration, such as tool wear. Indirect methods rely on sensing machine parameters such as forces, acoustic emission, temperature, vibration, current, voltage, torque, strain, or image characteristics of the tools being monitored. In methods using indirect measurements, features indicative of condition are extracted and correlated to assess the nature, extent, and location of the fault [59].

15.6.3 THREE STAGES OF MACHINE-CONDITION MONITORING AND DIAGNOSIS

Machine-condition monitoring can be posed as a pattern-recognition problem with three key stages: first, data acquisition; second, feature extraction and selection; and finally, decision-making. Data are acquired using transducers and normally recorded in either analog or digital form on magnetic tape or computer disk.

15.6.3.1 Feature extraction and selection

It is usually not feasible to automatically determine the machine condition using the recorded raw data. Instead, the objective is, with expert assistance, to identify different classes and perform a supervised classification; then, assign a label to each instance of training data. Then for each example, several features are defined (feature extraction). Therefore, some transformation or processing is required, e.g., mapping the original data from the time domain to another domain, such as the frequency or (t, f) domain, where the differences between the normal and abnormal behaviors can be magnified. It is in this new domain that features are extracted. It is important to select an optimal subset of features as it is most likely that using all the extracted features would not result in the best classification rate. Indeed, having more features does not result in better classification; when the number of examples in the training data is too small compared to the number of features, most classification procedures cannot find good classification boundaries. This is the curse of dimensionality [60]. In fact, often one may introduce “noise” to the classifier by adding irrelevant features. A feature that is irrelevant or redundant with another may penalize the process. Manual selection of features is a difficult task because two irrelevant features when taken individually can be very relevant when operated together [60]. We can generalize this; in some cases, the k worst features (those that got the lowest score in a given performance criterion) can be better when combined than the k best features. This is the phenomenon of interpertinence.

15.6.3.2 Feature selection methods

Good feature selection yields data compression and performance enhancement. This requires smart search algorithms, as an exhaustive exploration of all subsets has exponential complexity. Feature selection methods include filters, wrappers, and embedded techniques.

15.6.3.3 Filter methods

Here, the performance criterion used to assess the pertinence of one feature is a measure based on the properties of the training data, without taking into account the classifier performance for the selected features. Three popular filter methods are described below:

- 1. Fisher criterion:** This criterion measures the separability between data clusters as [61]

$$F(i) = \frac{\sum_{c=1}^C (n_c(\mu_c^d - \mu^d)^2)}{\sum_{c=1}^C (n_c(\sigma_c^d)^2)}, \quad (15.6.1)$$

where n_c is the number of elements in class c , μ_c^d and σ_c^d represent respectively the mean and the standard deviation of the d th feature in the class c , and μ^d is the overall mean of the d th feature.

- 2. Minimum redundancy maximum relevance (MRMR) [62]:** This criterion uses mutual information so that both relevancy and nonredundancy of features are considered simultaneously. The mutual information between two random variables (RV) X and Y is defined by

$$I(X, Y) = \sum_{x,y} p_{X,Y}(x, y) \log \frac{p_{X,Y}(x, y)}{p_X(x)p_Y(y)}, \quad (15.6.2)$$

where p_X is the probability density function (pdf) of X and $p_{X,Y}$ is the joint pdf of X and Y . The relevance of a feature set S for class c is the average value of all mutual information values between the individual feature f_d and class c , i.e.,

$$D(S, c) = \frac{1}{\text{card}(S)} \sum_{f_d \in S} I(f_d, c), \quad (15.6.3)$$

where $\text{card}(S)$ represents the number of elements in the set S . The redundancy of features in set S is the average value of mutual information values between the feature f_d and the feature f_j :

$$R(S) = \frac{1}{(\text{card}(S))^2} \sum_{f_d, f_j \in S} I(f_d, f_j). \quad (15.6.4)$$

Thus, the MRMR criterion is a combination of the above two measures, i.e.,

$$\text{MRMR} = \max_S (D(S, c) - R(S)). \quad (15.6.5)$$

- 3. Maximal marginal diversity (MMD):** this criterion for feature selection is based on the Kullback-Liebler divergence (KLdiv), defined as

$$\text{KLdiv}(p_X || p_Y) = \sum_{x \in S} p_X(x) \log \frac{p_X(x)}{p_Y(y)}. \quad (15.6.6)$$

The RVs X and Y take the same set of choices S . KLdiv is a measure of relative entropy between the two pdfs, but it is not a distance, because it is not symmetrical as $\text{KLdiv}(p_X || p_Y) \neq \text{KLdiv}(p_Y || p_X)$. Its symmetrical extension is the Kullback-Liebler distance, defined by

$$\text{dKL}(p_X, p_Y) = \text{KLdiv}(p_X || p_Y) + \text{KLdiv}(p_Y || p_X). \quad (15.6.7)$$

Let us assume the data contains a number C of clusters (or classes) G_1, \dots, G_C . Let L_n be the n th line of X_{train} (matrix that represents the training data with N rows and D columns). Thus, the prior probability of occurrence of the c th class $P(G_c)$ is estimated by $\hat{P}(G_c) = \text{card}(L_n \in G_c)/N$, where N is the number of examples in the training data. Let us define the event \bar{G} as the empirical mean of all C clusters, and let C_d be the d th column of X_{train} . The MMD criterion, for the d th feature, is defined in the case of C clusters G_1, \dots, G_C by

$$J(C_d) = \sum_{c=1}^C \hat{P}(G_c) d\text{KL}(p(G_c, C_d), p(\bar{G}, C_d)), \quad (15.6.8)$$

where $p(\bar{G}, C_d)$ is the joint pdf of the mean cluster \bar{G} and C_d . The selected best solution for the problem (of selecting the most discriminating features) is to choose the axes that maximize the MMD.

There are other criteria for assessing feature performance. The AUC is popular; it represents the area under the receiver operating characteristics (ROCs). The AUC of a feature is the probability that the feature would have a higher value for a randomly chosen positive example than for a randomly chosen negative example [50].

15.6.3.4 Wrapper methods

The name of the wrapper method refers to the fact that this feature selection algorithm takes into account the selected classifier. Feature selection by wrapping generally gives better results than the filtering approach. Given a cost function associated with the classification performance, each subset is evaluated, then updated iteratively for as long as necessary to obtain, after convergence, the best subset according to the chosen cost function (see Fig. 15.6.1). Consequently, with this method, the selected subset of features that is optimal for the classifier used may not give a good performance for another classifier. For wrapper methods, computation costs are very high compared with filter methods (mainly due to the cost function computation). Generally, the cost function is the classifier performance that involves computationally expensive steps such as repeated training and testing using

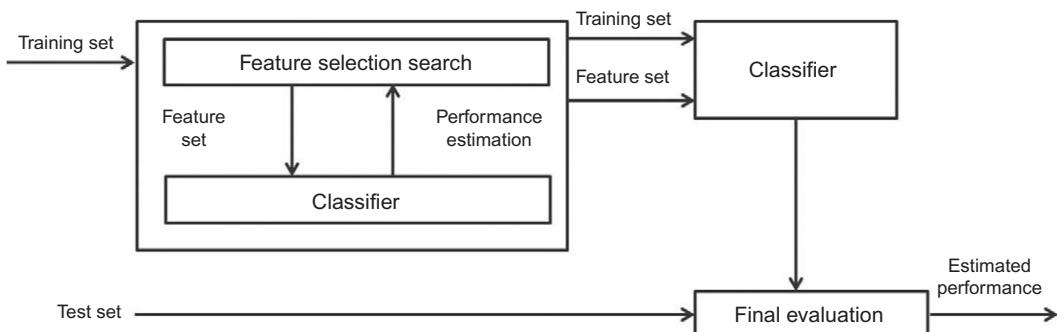


FIGURE 15.6.1

Principle of wrapper for feature selection.

the cross-validation approach. The problem complexity is high, so it is not possible to use an exhaustive search strategy (an NP-complete problem, where “NP” means “nondeterministic polynomial-time”). Methods or random search heuristics can be used such as sequential forward feature selection (SFFS) [63]. The search operation can be enhanced by increasing the size of the initial set of features. Wrapper methods can strongly enhance the processing performance, providing a small subset of features. Embedded methods can also be used; they incorporate the feature selection in the classifier learning process [60].

15.6.3.5 Classification

Final decision making is done by automatically correlating the feature set (obtained from the previous stage) to the machine conditions. Through supervised learning, an operator instructs the computer of possible patterns in the feature sets and relates them to the machine conditions. Typical automatic methods include fuzzy logic, decision trees, artificial neural networks, and support vector machines (SVMs). The latter is a modern computational learning method based on statistical learning theory. Compared to other classifiers, it makes no assumption on the shape of the pdf of the data. Its basic principle is to place a linear boundary between two classes and orient it in such a way that the margin is maximized (i.e., the distance between the boundary and the nearest data point in each class is maximal). The nearest data points are used to define the margin and are known as support vectors. The boundary can be expressed with the use of the support vectors as [64]

$$(\mathbf{w}^T \mathbf{x}) + b = 0, \quad \mathbf{w} \in \mathbb{R}^D, \quad b \in \mathbb{R}, \quad (15.6.9)$$

where \mathbf{x} is the input vector of dimension D , the vector \mathbf{w} and the scalar b determine the position of the boundary. Taking into account the noise with positive slack variables ξ_i and the error penalty E , the optimal hyperplane separating the data is obtained by the following optimization [64]:

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|\mathbf{w}\|^2 + E \sum_{i=1}^D \xi_i, \\ & \text{subject to} && \begin{cases} y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, \quad i = 1, \dots, N. \end{cases} \end{aligned} \quad (15.6.10)$$

Often, real data are nonlinearly separable, so the linear boundary in the input space cannot separate the two classes properly. Nonlinear mapping can then be used to generate the classification features from the original data. The data to be classified is mapped by using a transformation $\varphi(\mathbf{x})$ onto a high-dimensional feature space, where the data can be linearly classified or separated. A kernel function K is used to perform the transformation; common kernel functions are linear functions, polynomial functions, radial basis functions, and sigmoid functions, but any function that satisfies Mercer’s theorem [64] can be used. Based on the input data vectors that describe the possible faults in machines, SVM can recognize these patterns. Usually, each fault produces special features that can be considered patterns.

15.6.4 STANDARD DSP METHODS FOR FEATURE EXTRACTION

15.6.4.1 Time-domain methods

The basic t -domain approach for fault detection uses measurements of the power of the monitored signal. The method exploits the principle that as the machine’s condition deteriorates, the vibration energy is expected to increase. Another approach uses statistical parameter analysis for fault detection

by considering the monitored signal as a random variable and then using higher-order statistical moments and cumulants as features.

15.6.4.2 Frequency-domain methods

The basic spectral analysis approach exploits the principle that the spectrum of the monitored signal changes when faults occur. The type and extent of the change depends on the nature of the fault in the machine being monitored. The condition of the machine is assessed through monitoring the change in the spectrum or a number of discriminating features extracted from the spectrum of the measured signal. These features may be some specific frequency components that depend on the type of machine and the nature of the fault. They are compared to references established when the machine was known to work properly under similar conditions, and an appropriate decision is taken when the feature vector deviates from the reference by more than a predetermined threshold.

15.6.5 NONSTATIONARY SIGNALS IN MACHINES

15.6.5.1 Limitations of classical methods

Traditional time-domain and spectral analysis techniques have two main shortcomings: (1) they do not provide information about the time-dependence of the frequency content of nonstationary signals, and vice versa; and (2) the Fourier transform (FT) does not conveniently represent a signal that has nonperiodic components, such as a transient signal.

In the case of rotating machines, the presence of certain frequency components has been shown to be indicative of faults; however, since some of these frequencies depend on the rotational speed, it is not possible using spectral analysis to determine these frequencies when the bearing runs at variable rotational speed. Studies have stressed the importance of machine monitoring during the transient states—such as start-up, shutdown, and acceleration periods—because some machine failures happen during these types of transition periods. Transient signals can be a good source of information about machine condition that is not available during steady states. For these reasons, FT-based methods may be unsuitable for machine monitoring in the above-mentioned circumstances [65].

15.6.5.2 The need for (t, f) methods in condition monitoring

The most frequently used (t, f) methods in the area of machine condition monitoring and diagnosis are quadratic time-frequency distributions (QTFDs) and time-scale analysis (mainly wavelet transforms (WTs)). These methods represent the signals in a wider (t, f) space that allow easier and more precise discrimination between fault and normal machine conditions. Previous studies defined a procedure based on the Wigner-Ville distribution (WVD) for robust detection and classification of helicopter gearbox faults [66]. The principle followed is that different faults produced different patterns in the (t, f) plane. The WVD-based patterns of vibration and acoustic signals were also used to detect faults in a number of machines and machine components such as engines [67] and gearboxes [68]. Other TFDs such as higher-order Wigner-Ville moment distributions [69] and reduced-interference distributions (RIDs) [70] were also used. Most of these methods were visually based detection and classification techniques that were meant to show the effectiveness of the respective TFDs for early fault detection. Other methods went further and used automatic feature extractors in an overall classification process. Some of the features extracted are amplitude values of the contour plots [68] and singular values of the TFD [70].

Another group of techniques collectively referred to as time-scale techniques were shown to be suitable for the detection and localization of transient events; in particular, WTs (both continuous and discrete) have been adopted in condition monitoring and diagnosis of turning and drilling machines [58], gears or gear trains [71], and bearings [72]. As in the case of the TFDs, some of the methods are used with feature extractors whose output is fed to a detector/classifier.

Parts I and II present detailed background on the two most widely used (t,f) methodologies, i.e., QTFDs and WTs (see [Chapters 2](#) and [3](#)). These two classes of representations are related through the short-time Fourier transform (STFT) and the Gabor transform ([Sections 2.3](#) and [2.7](#)). QTFDs are suitable for large bandwidth duration (BT) signals ([Chapter 1](#)), while time-scale methods give best results when used with low BT and transient signals. A QTFD estimate is obtained from the quadratic class of TFDs [51], expressed in [Section 3.2.2](#) as $\rho_z(t,f) = W_z(t,f) * \gamma(t,f)$, where $\gamma(t,f)$ is a two-dimensional kernel window that is application-dependent, $W_z(t,f)$ is the WVD, and $*\gamma$ indicates a double convolution. The kernel window, $\gamma(t,f)$, characterizes a particular TFD and is generally adapted to the data so as to obtain the best possible (t,f) resolution [70] (see [Section 3.3](#) for more details) or the most sparse TFD. In classification applications, it may be advantageous to design a kernel γ that specifically maximizes the differences between classes.

15.6.6 EXAMPLE 1: GEARBOX FAULT DETECTION USING THE WVD

This section presents an example to illustrate the (t,f) methodology for condition monitoring and diagnosis.¹ The application is the detection of a broken tooth in a spur gear. The chosen TFD is the WVD. The system considered is composed of a 24-tooth input gear driven by an electric motor and meshing with 16 teeth of a pinion whose rotational frequency is 37.5 Hz. The applied load was 70% of the maximum load. The study simulated five fault types, each involving the partial or total removal of one or more teeth. The faults included the removal of 25%, 50%, 75%, and 100% of the face-width at a given radius, plus the same defect with 100% advancement on two pinion teeth. The acceleration vibration signal obtained from the above-mentioned system was low-pass filtered and sampled at a rate of 6.4 kHz.

15.6.6.1 Feature extraction and selection

The vibration signal is synchronously averaged in order to remove any periodic events not exactly synchronous with the gear of interest and to reduce the effects of noise and vibration sources other than that of the gear. The averaged signal is then transformed to the (t,f) domain using the pseudo-WVD (discrete WVD) with a Hamming window. The negative values of the WVD are set to zero and the resulting TFD is normalized. The results are displayed in the form of contour plots. To enhance the sidebands around the meshing frequencies, the residual signal is obtained by removing the meshing harmonics using a bandstop filter. The extracted features are the amplitude values of the contour plots (see [Fig. 15.6.2](#)). To reduce the dimension of the feature vector, WVD cross-sections at and around a chosen meshing frequency are selected.

¹Another example is presented in [Section 15.7](#).

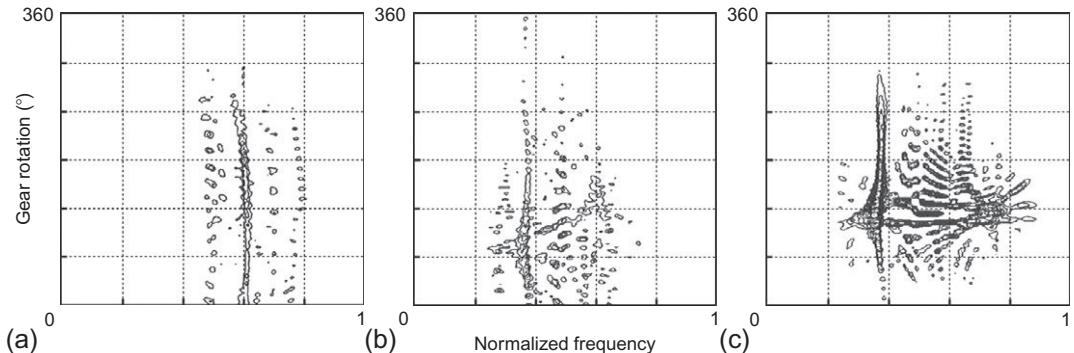


FIGURE 15.6.2

Weighted WVD of the residual signal: (a) normal condition of the spur gear; (b) one broken tooth with 50% fault advancement; (c) one broken tooth with 100% fault advancement.

15.6.6.2 Decision making

Two main classification approaches can be considered: statistical and neural pattern recognition. In the first approach, to assign the feature vector from the last stage to one of the C classes considered, the Mahalanobis distance can be selected as the similarity measure; it is given by

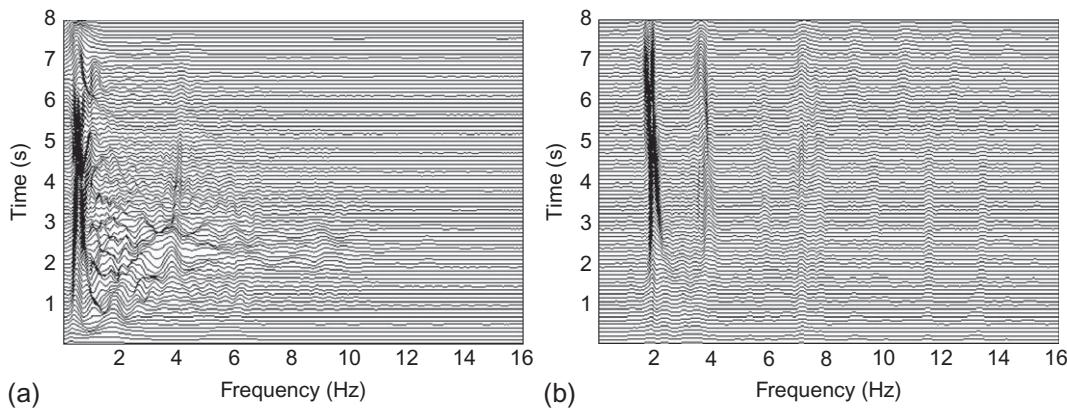
$$d = \left[\sum_{k=1}^K (\vec{y} - \vec{\mu}_k)^T \Sigma_k^{-1} (\vec{y} - \vec{\mu}_k) \right]^{1/2}, \quad (15.6.11)$$

where \vec{y} is the feature vector and $\vec{\mu}_k$ and Σ_k are respectively the mean vector and covariance matrix representing the k th class. The study considered only two classes, namely normal (no fault) and abnormal (fault), and used only one template representing the normal condition. The classification approach can also use a neural network trained in a supervised mode using the back-propagation algorithm [68].

Fault diagnosis of rotating machinery can also be performed using time-scale methods such as the WT. The difference with the procedure described above is that features can be extracted from the coefficients of the WT or other time-scale methods [72].

15.6.7 EXAMPLE 2: DETECTION OF EEG ABNORMALITIES IN THE BRAIN

The above (t,f) methodology for condition monitoring and diagnosis is illustrated on the detection of brain electroencephalogram (EEG) abnormalities. The EEG signal recordings are described in Ref. [51]. The EEG was filtered with an analog bandpass filter with cutoff frequencies at 0.5 and 70 Hz, sampled at 256 Hz, then resampled to 32 Hz, with digital anti-aliasing filters, as the significant energy in newborn EEG ($>95\%$) does not exceed the alpha band (8-12 Hz).

**FIGURE 15.6.3**

(t, f) features of EEG signals using the EMBD: (a) nonseizure; (b) seizure.

15.6.7.1 Feature extraction and selection

The feature extraction step includes statistical, image processing, and morphological methods (Section 12.6) applied to a (t, f) image (see, e.g., Fig. 15.6.3). At this point, the problem has been divided into several biclass problems; the feature selection must be done for each biclass problem, as SVMs were designed for such cases (see Fig. 15.6.4). A configuration that looks for the best set of features for each decision is preferred, as this optimizes the classification locally rather than globally. Knowing that for this step, the calculation time is not important, the wrapper method is used for feature selection. The SFFS algorithm is selected to carry out the search algorithm; it generates the configuration that has the highest rate of correct classifications using the method of leave-one-out. In the case of supervised learning, the choice of the database and labeling examples has a significant impact on the results, and the database must have been formed with expert assistance.

15.6.7.2 Decision making

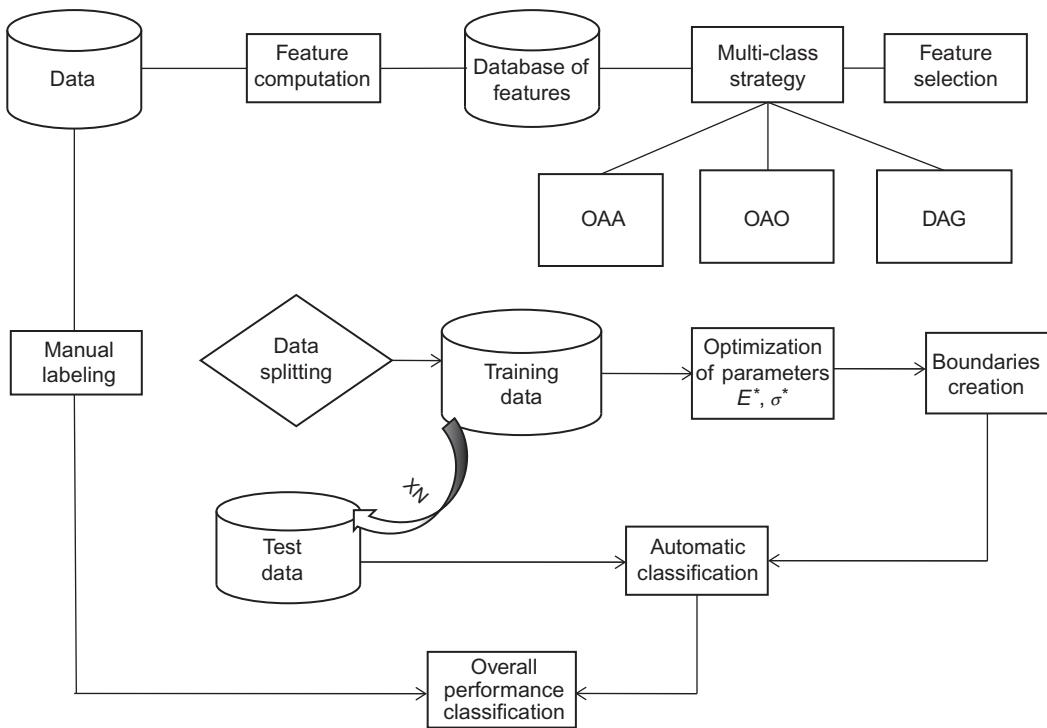
15.6.7.2.1 Kernel choice

In the implementation of SVM, the kernel must be adapted to the data. We selected the Gaussian radial basis function (RBF) kernel for the two reasons: (1) the number of parameters induced by this kernel is small compared to other common kernels like polynomial kernels; and (2) numerical difficulties are reduced (indeed $0 < K(\vec{x}_i, \vec{x}_j) \leq 1$ for a Gaussian RBF kernel while for a polynomial kernel, values can be infinite or zero; however, when the number of features is large, typically greater than 500, it may be better to use a linear kernel, since the size of the input space is quite large and thus the probability of finding a linear boundary is greater [64].

15.6.7.2.2 Multiclass classification strategy

Three strategies are considered:

1. One-against-all (OAA): This simplest strategy consists of using a binary classifier for each class. Once class c is chosen, the second class is the union of all others. Thus, the c th binary classifier

**FIGURE 15.6.4**

General scheme for decision-making for abnormalities detection (OAA, one-against-all; OAO, one-against-one; DAG, direct acyclic graph).

separates the examples of class c from all other elements of other classes. To assign an example to one class, it is presented to C binary classifiers, and then outputs (i.e., soft decisions) from all the classifiers are combined. The class assigned to the example is the one for which the classifier returns the maximum distance from the margin. As this method involves training with a very unbalanced number of examples among classes, it causes practical difficulties.

2. One-against-one (OAO): This approach develops a classifier for each pair of possible classes. Each classifier is indexed by the pair (c, l) with $1 \leq c, l \leq C$. For each pair (c, l) a classifier is constructed to separate class c from class l . This method constructs $\binom{C}{2}!$ different classifiers; then the decision for a new example is obtained using majority voting. One problem is that indeterminations may occur when several classifiers get equal votes.
3. Direct acyclic graph (DAG): Here, a decision tree is constructed; each node of this tree is a binary decision. This approach starts from the root node and evaluates the binary decision. Then it moves to either left or right depending on the output value, and the process ends with the leaves (the end nodes). We can construct this tree using the measure of generalization ability: $E_{cl} = N_{sv}/N$, where N is the number of examples of the c th and the l th class and N_{sv} is the number of support vectors used for the biclass problem between class c and class l .

15.6.7.2.3 Parameter optimization

This discussion focuses on SVMs as an illustration of the overall approach. In the case of multiclass problems, this step must be performed for every biclass problem, as each problem is different; we must optimize each two-class problem to obtain an overall good performance for the multiclass classification problem. A grid search strategy is used to select σ and E as, in the case of two parameters, the complexity is reasonable. It is essential to seek the optimal couple simultaneously because we cannot seek both independently, as parameter E and parameter σ are strongly linked. At the mesh, we will assign E logarithmically spaced values between 10^{-6} and 10^6 and for the value of σ , values between 0.1 and 20 also distributed logarithmically. The couple that results in the maximal rate of correct classifications, estimated using the cross-validation, is selected. The complete decision-making process is illustrated in Fig. 15.6.4.

15.6.8 SUMMARY AND CONCLUSIONS

Time-frequency signal analysis and processing methods can bring significant improvements in performance in the area of machine-condition monitoring and diagnosis. (t,f) methods can be used for detection, classification, and monitoring the progression of the faults and wear with time. This can enable more precise prediction and prevention of catastrophic failures. (t,f) techniques can be used either as visual indicators of the presence of faults or as the basis of feature extractors in a fully automated pattern-recognition process.

Sections 11.2, 15.2, and 15.7 describe three other time-frequency approaches to machine-condition monitoring in more detail.

15.7 CONDITION MONITORING OF ASSETS USING TIME-FREQUENCY METHODS⁰

15.7.1 TIME-FREQUENCY METHODS FOR CONDITION MONITORING

The efficient condition monitoring of an asset requires quantifiable indicators that can characterize the dynamic behavior of the asset in both the undamaged and damaged states. One of the most used indicators in noninvasive damage detection is the change in natural frequency occurring with the alteration of the asset's geometrical and mechanical characteristics. But since the changes in the natural frequencies are small even for severe damage, we need elaborate procedures that can compute the signal's frequency spectrum with very fine resolution. Following the line of Section 15.6, this section overviews some practical methods of improving the reading of the classical time-frequency (t,f) spectrum, so the frequency content can be more precisely determined and used in condition monitoring of assets.

The focus will then be on the application of these methods to the identification of the natural frequencies of cantilever beams for condition monitoring and damage detection. Since vibrations of

⁰Authors: D.M. Onchis, Eftimie Murgu University of Resita, Resita, and West University of Timisoara, Timisoara, Romania (d.onchis@uem.ro, donchis@hpc.uvt.ro). Reviewers: R. Yan, M. Gaianu, D. Frunzaverde, and G.A. Velasco.

the cantilever beam following a mechanical excitation are of damped harmonic type, it is natural to try to accurately identify the frequencies contained in the oscillating waves.

For simplicity, we will use a steel cantilever beam having the following geometrical characteristics: length $l = 1000$ mm, width $b = 50$ mm, height $h = 5$ mm, hence (for the undamaged state) cross-sectional area $A = 250 \times 10^{-6}$ m² and moment of inertia $I = 520.833 \times 10^{-12}$ m⁴. The mechanical properties are mass density $\rho = 7850$ kg/m³, Young's modulus $E = 2.0 \times 10^{11}$ N/m², and Poisson's ratio $\mu = 0.3$. This beam is used as a reference. For other dimensions (l, b, h) or mechanical properties (ρ, E, μ), the problem can be solved by considering the scale influence. For recording the signal, we place an accelerometer on the free end of the unloaded beam.

Damage was simulated at 1000 positions on the beam using the finite-element method (FEM). Real damage was created at several exposed locations on the beam for direct measurement. The change in natural frequency due to damage depends on the forces acting on the beam and the dimensions of the damage (cross-section reduction). Changes in ambient temperature could also induce small changes in frequency. In the numerical experiments, we computed the first 10 natural frequencies, which are usually enough for the requirements of damage detection [73].

15.7.2 IMPROVING THE READABILITY OF TIME-FREQUENCY SPECTRA

We compare two classical (t, f) methods, namely the linear short-time Fourier transform (STFT) and the quadratic Wigner-Ville distribution (WVD), with a novel adaptive (t, f) representation, namely the nonstationary Gabor transform (NSGT). The WVD and the STFT are defined in Eqs. (2.1.17) and (2.3.2) (pp. 69 and 76). For the practical implementation, the transforms will be applied in a finite discrete setting using the Hilbert space \mathbb{C}^L . We consider a signal s of length L , with window functions $\{\varphi_k\}_k$ and frequency shift parameters $\{b_k\}_k$. Under these considerations, the direct discrete versions of the STFT and the WVD can be immediately obtained.

It remains to define the NSGT, which is a variant of the STFT. The finite discrete NSGT [74] is given by

$$c_{n,k} = \sum_{i=0}^{L-1} s(i) \overline{\varphi_n(i)} e^{-j2\pi i n b_k / L}. \quad (15.7.1)$$

Unlike the fixed window of the STFT, the window φ_n varies over time n . But it still allows use of the fast Fourier transform (FFT). Unlike standard Gabor analysis, which uses time translation to generate atoms, nonstationary Gabor frames allow different windows in different time positions. For each position, atoms are built by regular frequency modulation; the frequencies, hence the atoms, are indexed by k .

The *frequency-side* finite discrete NSGT is given by

$$c_{n,k} = \sum_{i=0}^{L-1} \hat{s}(i) \overline{\widehat{\varphi_k}(i)} e^{j2\pi i n a_k / L} = \sum_{l=0}^{L-1} s(l) \overline{\varphi_k(l - n a_k)}. \quad (15.7.2)$$

In this case, the window φ_k varies over the frequency index k , generating a transform that is frequency-adaptive like wavelet transforms, or nonuniform linear time-invariant (LTI) filterbanks. The fast implementation requires a full FFT of s .

At first glance, both the WVD and the STFT offer the same advantage in frequency resolution. In principle, the NSGT can have a similar resolution with the appropriate choice of bins per octave but that is at the expense of some of its desirable properties, such as emphasis on specific frequency ranges. The advantage of the WVD over the other two methods is the fact that the signal is correlated with itself, thus giving the most accurate localization of the natural frequencies. This is also a drawback since it causes many prominent cross-terms to appear. In order to determine the relevant frequencies from the WVD, the method presented here uses prior knowledge from the STFT and NSGT. This can be done in two different ways.

The first approach determines the frequencies that are most significant in the STFT and NSGT and determines a range in which all the natural frequencies may lie. This range comprises the frequency bins with the largest contribution and their closest neighbors (within a distance that is variable depending on the bin value). By searching for high contributions in the WVD only within this predetermined range, we exclude the possibility of reading cross-terms of any great intensity except when cross-terms overlap auto-terms. This approach is heavily dependent on the choice of the neighborhoods around the key frequency bins determined in the STFT/NSGT. Also, small correlation effects might appear in the STFT between the signal and the window function. Although unlikely, this might lead to an unwanted extension of the range in which the WVD is investigated and to inconclusive results.

The second approach is a progressive refinement of the search range going through all three representation methods, starting with the NSGT. First an appropriate choice of window functions must be made so that only natural frequencies are detected, even with an imprecise resolution. The NSGT is ideal for this, because of the advantage of allocating a variable number of bins to each octave. Having determined this initial set of peak frequencies, the STFT is investigated in the areas of interest and a further refinement is made. This intermediate step is necessary in order to approach the natural frequencies and account for the change in window length. Equivalently, this could be done by taking an STFT with variable windows, but using both the NSGT and the STFT ensures a smooth transition. The final step is identifying the peak frequencies in the WVD that lie within the refined area of interest.

One should note that in the end, both approaches yield the same results in the optimal case, but the second algorithm minimizes the risk of having too wide a search range in the WVD. Both paths use the high resolution and accuracy of the WVD to determine peak frequencies within a certain range and employ the NSGT and the STFT in order to determine where the natural frequencies lie up to a certain degree of precision. This algorithm makes use of the advantages of the three spectral representation methods, while avoiding each of their drawbacks, so that in the end, it reaches better results than any of the three methods individually.

15.7.3 ASSESSMENT OF THE METHODS ON CANTILEVER BEAMS

In this section, we employ the above methods to accurately detect the natural frequencies of the cantilever beams for condition monitoring purposes. But before applying the algorithms, we overview the maximum possible ranges of variation of the natural frequencies due to damage.

Using the equation of motion for a prismatic beam [75], the 10 natural frequencies of the undamaged beam were analytically calculated; they are given in the first column in [Table 15.7.1](#). We then performed FEM analysis for 1000 positions of the damage to determine the frequency intervals of interest; the results are included in the same table. The third verification step consisted of direct measurements on the damaged beam on the locations where the amplitudes for the 10 vibration modes have maximum

Table 15.7.1: Comparison of the First 10 Natural Frequencies (in Hz) Obtained Analytically for the Undamaged Beam, and by FEM Analysis and Measurements for the Extreme Damage Case

| Analytical | FEM | | Measurements | |
|------------|--------|---------|--------------|---------|
| | Low | Up | Low | Up |
| 5.12 | 4.20 | 5.30 | 4.27 | 5.40 |
| 23.55 | 21.32 | 25.30 | 21.27 | 25.40 |
| 62.47 | 56.7 | 69.65 | 54.27 | 65.40 |
| 136.19 | 133.36 | 147.34 | 135.40 | 148.34 |
| 232.45 | 225.53 | 257.34 | 224.40 | 255.40 |
| 342.43 | 322.56 | 367.43 | 320.40 | 360.40 |
| 485.39 | 471.28 | 495.01 | 475.40 | 495.40 |
| 650.25 | 640.13 | 656.25 | 645.40 | 655.40 |
| 826.4 | 811.33 | 831.28 | 815.40 | 835.40 |
| 1035.4 | 1024.1 | 1067.82 | 1020.40 | 1055.40 |

values or are null or where we encountered a point of inflection. These results are also joined in the same table. Using the overall information, we have taken an extra-safe margin of 5% on both sides (“Low” and “Up”) in order to create the prospective frequency variation intervals.

To obtain the frequencies from the refined $(t-f)$ spectrum of the WVD, STFT, and NSGT, one can apply a peak-peaking technique on the respective spectrum in conjunction with the application of one of the two methods proposed in [Section 15.7.2](#).

Finally, [Table 15.7.2](#) presents a benchmarking between the frequencies obtained from the analytical equation of the beam, the gradually improved $(t-f)$ procedure, and the wavelet-based detection.

15.7.4 CONDITION MONITORING USING THE EMD METRIC

For condition monitoring purposes, we compare the extended (t,f) signatures for the set of 10 natural frequencies using the *earth mover’s distance* (EMD) metric [\[76,77\]](#). The extended signature is a signature in which the minimal number of empty bins has been added to assure that, given a pair of signatures to be compared, the number of bins is the same. In the extended (t,f) signatures, to separate the damage effect from the ambient effect, we have also incorporated the temperature variation during the probable functioning conditions. For each Celsius degree, we have computed the natural frequency for the first 10 vibration modes over a series of repeated measurements.

The EMD is a method of evaluating the dissimilarity between two multidimensional distributions in some feature space where a distance measure between single features, called the *ground distance*, is given. The EMD metric is suitable for condition monitoring because it can accurately measure the dissimilarities on signatures of sets with different frequency values for the first 10 vibration modes.

Table 15.7.2: Comparison of the First 10 Natural Frequencies Obtained Using the Analytical Equation, the (t,f) Method, and the Wavelet Method

| Analytical | Time-Frequency | Wavelet-Based |
|------------|----------------|---------------|
| 5.12 | 5.07 | 4.87 |
| 23.55 | 22.32 | 21.32 |
| 62.47 | 61.7 | 59.65 |
| 136.19 | 133.36 | 137.34 |
| 232.45 | 235.53 | 237.34 |
| 342.43 | 343.56 | 347.43 |
| 485.39 | 481.28 | 485.01 |
| 650.25 | 645.13 | 646.25 |
| 826.4 | 824.33 | 811.28 |
| 1035.4 | 1034.1 | 1057.82 |

The EMD was originally introduced by Rubner et al. [76]. Pele and Werman [78] introduced an EMD algorithm with thresholded ground distances and showed that the EMD is indeed a metric for all signatures.

Let $P = (p_1, w_{p_1}), \dots, (p_m, w_{p_m})$ be the first signature with m clusters, where p_i is the representative cluster and w_{p_i} is the weight of the cluster; and let $Q = (q_1, w_{q_1}), \dots, (q_n, w_{q_n})$ be the second signature with n clusters. Let $D = [d_{ij}]$ be the ground distance matrix, where d_{ij} is the ground distance between clusters p_i and q_j (and j is now a counter). The $\widehat{\text{EMD}}$ could be defined as

$$\widehat{\text{EMD}}_\alpha(P, Q) = \left(\min_{f_{ij}} \sum_{i,j} f_{ij} d_{ij} \right) + \left| \sum_i w_{p_i} - \sum_j w_{q_j} \right| \alpha \max_{i,j} d_{ij} \quad (15.7.3)$$

such that the EMD constraints are satisfied.

The advantage of this definition is that the number of edges in the flow network is reduced by a factor of magnitude, making it much faster than the original algorithm proposed by Rubner. In addition, the EMDs with thresholded distances have many other desirable properties—including correspondence with the way humans perceive distances, and robustness to outlier noise and quantization effects—while retaining the property-of-being metrics. Pele and Werman [78] proved that $\widehat{\text{EMD}}$ is a metric for any two signatures if the ground distance is a metric and $\alpha > 0.5$. The metric property is necessary, because it enables fast computational algorithms.

The problem we are dealing with at this point is the following: having determined the first 10 natural frequencies from the time-frequency spectrum of a possibly damaged beam, $F_d = f_{d,1}, f_{d,2}, \dots, f_{d,10}$, one must determine with the best degree of precision if the beam is damaged. Using the signatures of repeated measurements by taking the averages, the temperature changes t_1, t_2, \dots, t_n can be simulated as in Ref. [73]. The signatures of the first 10 natural frequencies for the temperature t_j can be determined and stored in a database as $f_{ij,1}, f_{ij,2}, \dots, f_{ij,10}$, where $i = 1, \dots, m$ and $j = 1, \dots, n$. In our study, we have

considered the simulated temperature changes in $n = 200$ fractions of Celsius degrees, ranging in the interval $[-25.8\%, +38.9\%]$. The number of repeated measurement tests is $m = 10$.

The $\widehat{\text{EMD}}$ metric of Pele and Werman [78] can be used to determine which entry in the database best resembles the measurements on the damaged beam and to what degree. But to obtain reliable and precise results, the data must be normalized.

First, we consider the natural frequencies, of which the first was measured between 3.109 and 4.102 Hz, with a standard deviation of 0.157, while the tenth was measured between 945.801 and 1041.226 Hz, with a standard deviation of 18.511. The increase in standard deviation places greater priority on the higher frequencies. Because equal priority is clearly desired, the data set was given a set of weights w_1, w_2, \dots, w_{10} , where $w_i = \frac{1}{\sigma_i}$ for $i = 1, \dots, 10$, and σ_i represents the standard deviation of the i th frequency. After this normalization, the standard deviation of each natural frequency was 1, giving equal importance to each frequency.

Second, in order to have a very precise observation of the damage, we would like to penalize as little as possible the difference between the total mass of two signatures in Pele and Werman's $\widehat{\text{EMD}}_\alpha$, by manipulating the coefficient α . But this method has the limitation that it remains a metric as long as $\alpha > 0.5$. So in order to improve the precision of the damage observation, the total mass of each set of frequencies was also normalized.

After this process, the k th natural frequency from a signal signature measured from a beam at the temperature t_j and with the repetitions D_i becomes

$$f'_{ij,k} = \frac{f_{ij,k}}{\sigma_k} \frac{M}{M_{ij}}, \quad (15.7.4)$$

where σ_k is the standard deviation of the k th frequency in the original data table, M_{ij} is the total mass of the frequency set for D_i and t_j , and M is the average total mass of a frequency set. Note that M_{ij} and M are computed for the intermediate data table, namely after the standard deviation normalization with the weight set $\{w_k\}_{k=1,\dots,10}$. F_d is normalized in a similar fashion:

$$f'_{d,k} = \frac{f_{d,k}}{\sigma_k} \frac{M}{M_D}, \quad (15.7.5)$$

where M_D is the total mass of F_d . By comparing the extended $(t - f)$ signature for a signal with the signatures from the database, the possibility of ambient effects on the small frequency variation is eliminated, and the conclusion that a beam is damaged or not can be immediately drawn.

15.7.5 SUMMARY AND CONCLUSIONS

We present in this section the condition monitoring of beams based on a set of 10 precomputed natural frequencies. We survey ways to improve the natural frequencies' readability using complementary information extracted from the spectra of the STFT, WVD, and NSGT. The results show that we are able to improve the accuracy of natural frequency identification using the proposed methods. Using the $\widehat{\text{EMD}}$ metric applied on the set of (t, f) extended signature frequencies, we are able to accurately determine whether a beam is damaged or not.

REFERENCES

- [1] J. Arrillaga, M.H.J. Bollen, N.R. Watson, "Power quality following deregulation", *Proc. IEEE* 88 (2) (2000) 246-261.
- [2] S. Santoso, E.J. Powers, W.M. Grady, P. Hofmann, "Power quality assessment via wavelet transform analysis", *IEEE Trans. Power Deliv.* 11 (2) (1996) 924-930.
- [3] Y.-J. Shin, E.J. Powers, W.M. Grady, A. Arapostathis, "Power quality indices for transient disturbances", *IEEE Trans. Power Deliv.* 21 (1) (2006) 253-261.
- [4] W.J. Williams, J. Jeong, "Reduced interference time-frequency distributions", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 74-97 (Chapter 3).
- [5] L. Cohen, "Wavelet moments and time-frequency analysis", in: *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, Soc. of Photo-optical Instrumentation Engineers, 1999, pp. 434-445.
- [6] Y.-J. Shin, A.C. Parsons, E.J. Powers, W.M. Grady, "Time-frequency analysis of power system disturbance signals for power quality", in: *Proc. IEEE Power Engineering Soc. Summer Meeting*, vol. 1, 1999, pp. 402-407.
- [7] Y.-J. Shin, E.J. Powers, W.M. Grady, S.C. Bhatt, "Effects of dispersion on disturbance propagation on high voltage transmission lines", in: *Proc. IEEE Power Engineering Soc. Summer Meeting*, vol. 2, 2000, pp. 851-854.
- [8] J.B. Heywood, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York, 1988.
- [9] R. Hickling, D.A. Feldmaier, F.H.K. Chen, J.S. Morel, "Cavity resonances in engine combustion chambers and some applications", *J. Acoust. Soc. Amer.* 73 (4) (1983) 1170-1178.
- [10] D. Scholl, C. Davis, S. Russ, T. Barash, "The volume acoustic modes of spark-ignited internal combustion chambers", in: *SI Engine Combustion*, no. 1315 in SAE Special Publications, Soc. of Automotive Engineers, 1998, pp. 1-8, *Proc. SAE Internat. Congress and Exposition*, Detroit, MI, 23-26 February 1998, SAE paper no. 980893.
- [11] S. Carstens-Behrens, A. Podvoiski, "Estimation of combustion chamber resonances for improved knock detection in spark ignition engines", in: *Proc. 18th CAD-FEM Users' Meeting*, Internationale FEM-Technologietage, 2000.
- [12] J.F. Böhme, D. König, "Statistical processing of car engine signals for combustion diagnosis", in: *Proc. Seventh IEEE Workshop on Statistical Signal and Array Processing (SSAP-94)*, 1994, pp. 369-374.
- [13] A.M. Zoubir, B. Boashash, "The bootstrap and its application in signal processing", *IEEE Signal Process. Mag.* 15 (1) (1998) 56-76, ISSN 1053-5888, <http://dx.doi.org/10.1109/79.647043>.
- [14] D. König, "Application of time-frequency analysis for optimum nonequidistant sampling of automotive signals captured at knock", in: *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'96)*, vol. 5, 1996, pp. 2746-2749.
- [15] H.-I. Choi, W.J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE Trans. Acoust. Speech Signal Process.* 37 (6) (1989) 862-871.
- [16] Y. Zhao, L.E. Atlas, R.J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals", *IEEE Trans. Acoust. Speech Signal Process.* 38 (7) (1990) 1084-1091.
- [17] L. Stanković, J.F. Böhme, "Time-frequency analysis of multiple resonances in combustion engine signals", *Signal Process.* 79 (1) (1999) 15-28.
- [18] L. Stanković, "A method for time-frequency analysis", *IEEE Trans. Signal Process.* 42 (1) (1994) 225-229.
- [19] G. Matz, F. Hlawatsch, "Time-frequency subspace detectors and application to knock detection", *Internat. J. Elect. Commun.* 53 (6) (1999) 379-385.

- [20] B. Samimy, G. Rizzoni, "Mechanical signature analysis using time-frequency signal processing: application to internal combustion engine knock detection", *Proc. IEEE* 84 (9) (1996) 1130-1343.
- [21] S. Carstens-Behrens, M. Wagner, J.F. Böhme, "Detection of multiple resonances in noise", *Internat. J. Elect. Commun.* 52 (5) (1998) 285-292.
- [22] F. Hlawatsch, G.F. Boudreux-Bartels, "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.* 9 (2) (1992) 21-67.
- [23] A. Papandreou-Suppappola, F. Hlawatsch, G.F. Boudreux-Bartels, "Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: a unified framework for the affine, hyperbolic, and power classes", *Digit. Signal Process.* 8 (1) (1998) 3-48.
- [24] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [25] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, 1999 (translated by J. Stöckler from the French editions, *Temps-fréquence*, Hermès, Paris, 1993, 1998).
- [26] F. Auger, E. Chassande-Mottin, "Quadratic time-frequency analysis I: Cohen's class", in: F. Hlawatsch, F. Auger (Eds.), *Time-Frequency Analysis: Concepts and Methods*, ISTE/Wiley, London, UK, 2008, pp. 131-163 (Chapter 5).
- [27] O. Rioul, P. Flandrin, "Time-scale energy distributions: a general class extending wavelet transforms", *IEEE Trans. Signal Process.* 40 (7) (1992) 1746-1757.
- [28] J. Bertrand, P. Bertrand, "A class of affine wigner functions with extended covariance properties", *J. Math. Phys.* 33 (7) (1992) 2515-2527.
- [29] P. Gonçalves, J.-P. Ovarlez, R.G. Baraniuk, "Quadratic time-frequency analysis III: the affine class and other covariant classes", in: F. Hlawatsch, F. Auger (Eds.), *Time-Frequency Analysis: Concepts and Methods*, ISTE/Wiley, London, UK, 2008, pp. 193-226 (Chapter 7).
- [30] A. Papandreou, F. Hlawatsch, G.F. Boudreux-Bartels, "The hyperbolic class of quadratic time-frequency representations—part I: constant- Q warping, the hyperbolic paradigm, properties, and members", *IEEE Trans. Signal Process.* 41 (12) (1993) 3425-3444.
- [31] F. Hlawatsch, A. Papandreou-Suppappola, G.F. Boudreux-Bartels, "The hyperbolic class of quadratic time-frequency representations—part II: subclasses, intersection with the affine and power classes, regularity, and unitarity", *IEEE Trans. Signal Process.* 45 (2) (1997) 303-315.
- [32] F. Hlawatsch, A. Papandreou-Suppappola, G.F. Boudreux-Bartels, "The power classes—quadratic time-frequency representations with scale covariance and dispersive time-shift covariance", *IEEE Trans. Signal Process.* 47 (11) (1999) 3067-3083.
- [33] A. Papandreou-Suppappola, F. Hlawatsch, G.F. Boudreux-Bartels, "Power class time-frequency representations: interference geometry, smoothing, and implementation", in: Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis, 1996, pp. 193-196.
- [34] R.G. Baraniuk, D.L. Jones, "Unitary equivalence: a new twist on signal processing", *IEEE Trans. Signal Process.* 43 (10) (1995) 2269-2282.
- [35] D.E. Newland, "Time-frequency and time-scale analysis by harmonic wavelets", in: A. Prochazka, J. Uhlir, P.J.W. Rayner, N.G. Kingsbury (Eds.), *Signal Analysis and Prediction*, Birkhäuser, Boston, 1998 (Chapter 1).
- [36] J. Bonnel, G. Le Touzé, B. Nicolas, J.I. Mars, "Physics-based time-frequency representations for underwater acoustics: power class utilization with waveguide-invariant approximation", *IEEE Signal Process. Mag.* 30 (6) (2013) 120-129.
- [37] T.R. Reed, H. Wechsler, "Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations", *IEEE Trans. Pattern Anal. Machine Intell.* 12 (1) (1990) 1-12.
- [38] Y.M. Zhu, R. Goutte, M. Amiel, "On the use of two-dimensional Wigner-Ville distribution for texture segmentation", *Signal Process.* 30 (3) (1993) 329-353.
- [39] G. Cristóbal, J. Hormigo, "Texture segmentation through eigen-analysis of the Pseudo-Wigner distribution", *Pattern Recognit. Lett.* 20 (3) (1999) 337-345.

- [40] S. Stanković, I. Orovčić, M. Chabert, B. Mobasseri, “Image watermarking based on the space/spatial-frequency analysis and Hermite functions expansion”, *J. Elect. Imaging* 22 (1) (2013) <http://dx.doi.org/10.1117/1.JEI.22.1.013014>, article ID 013014.
- [41] J. Hormigo, G. Cristóbal, “High resolution spectral analysis of images using the pseudo-Wigner distribution”, *IEEE Trans. Signal Processing* 46 (6) (1998) 1757-1763.
- [42] T. Bülow, G. Sommer, “A novel approach to the 2D analytic signal”, in: F. Solina, A. Leonardis (Eds.), Proc. 8th Internat. Conf. on Computer Analysis of Images & Patterns (CAIP’99), Ljubljana, Slovenia, 1-3 September 1999, no. 1689 in Lecture Notes in Computer Science, Springer, 1999, pp. 25-32.
- [43] S.L. Hahn, “Multidimensional complex signals with single-orthant spectra”, *Proc. IEEE* 80 (8) (1992) 1287-1300.
- [44] B. Boashash, P.J. Black, “An efficient real-time implementation of the Wigner-Ville distribution”, *IEEE Trans. Acoust. Speech Signal Process.* 35 (1987) 1611-1618.
- [45] D.M. Chandler, “Seven challenges in image quality assessment: past, present, and future research”, *ISRN Signal Process.* 2013 (2013) <http://dx.doi.org/10.1155/2013/905685>, article ID 905685, 53 pp.
- [46] A. Beghdadi, M.-C. Larabi, A. Bouzerdoum, K.M. Iftekharudin, “A survey of perceptual image processing methods”, *Signal Process. Image Commun.* 28 (8) (2013) 811-831.
- [47] M.A. García-Pérez, V. Sierra-Vázquez, “Visual processing in the joint spatial/spatial-frequency domain”, in: E. Peli (Ed.), Vision Models for Target Detection and Recognition: In Memory of Arthur Menendez, World Scientific Publishing, Singapore, 1995, pp. 16-62 (Chapter 2).
- [48] H. Suzuki, F. Kobayashi, “A method of two-dimensional spectral analysis using the Wigner distribution”, *Elect. Commun. Jpn. III Fundament. Elect. Sci.* 75 (1) (1992) 101-110.
- [49] R. Redondo, S. Fischer, F. Šroubek, G. Cristóbal, “A 2D Wigner distribution-based multisize windows technique for image fusion”, *J. Visual Commun. Image Represent.* 19 (1) (2008) 12-19.
- [50] B. Boashash, N.A. Khan, T. Ben-Jabeur, “Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review”, *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [51] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [52] A. Djebbari, F. Bereksi-Reguig, “Detection of the valvular split within the second heart sound using the reassigned smoothed pseudo Wigner-Ville distribution”, *BioMed. Eng. OnLine* 12 (1) (2013) 37.
- [53] I. Kamarulafizam, S.-H. Salleh, A.R. Harris, K. Yusoff, M. Hamed, A.A. Jamil, A.S. Rahmani, “Heart sound feature representation using extended modified B-distribution”, *Trans. Jpn. Soc. Med. Biol. Eng.* 51 (Suppl.) (2013) R-204.
- [54] G.K. Groh, P.T. Levy, M.R. Holland, J.J. Murphy, T.J. Sekarski, C.L. Myers, D.P. Hartman, R.D. Roiger, G.K. Singh, “Doppler echocardiography inaccurately estimates right ventricular pressure in children with elevated right heart pressure”, *J. Amer. Soc. Echocardiog.* 27 (2) (2014) 163-171.
- [55] J. Xu, L. Durand, P. Pibarot, “A new, simple, and accurate method for non-invasive estimation of pulmonary arterial pressure”, *Heart Br. Card. Soc.* 88 (1) (2002) 76-80.
- [56] S. Barma, B.W. Chen, K.L. Man, J.F. Wang, “Quantitative measurement of split of the second heart sound (S2)”, *IEEE/ACM Trans. Comput. Biol. Bioinformat.* 99 (2014) 1-1, <http://dx.doi.org/10.1109/TCBB.2014.2351804>.
- [57] A. Illanes-Manríquez, Q. Zhang, “An algorithm for robust detection of QRS onset and offset in ECG signals”, *Comput. Cardiol.* 2008 (2008) 857-860.
- [58] R. Du, “Engineering monitoring and diagnosis using wavelet transforms”, in: C.T. Leondes (Ed.), Computer-Aided Design, Engineering, and Manufacturing: Systems Techniques and Applications—Vol. I: Systems Techniques and Computational Methods, CRC Press, Boca Raton, FL, 2000 (Chapter 8).

- [59] D.E. Dimla Snr., "Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods", *Internat. J. Machine Tools Manufact.* 40 (8) (2000) 1073-1098.
- [60] I. Guyon, A. Elisseeff, "An introduction to variable and feature selection", *J. Mach. Learn. Res.* 3 (2003) 1157-1182.
- [61] T.S. Furey, N. Cristianini, N. Duffy, D.W. Bednarski, M. Schummer, D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data", *Bioinformatics* 16 (10) (2000) 906-914.
- [62] H. Peng, F. Long, C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy", *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (8) (2005) 1226-1238, ISSN 0162-8828, <http://dx.doi.org/10.1109/TPAMI.2005.159>.
- [63] C.-C. Chang, T.-Y. Lin, "Linear feature extraction by integrating pairwise and global discriminatory information via sequential forward floating selection and kernel {QR} factorization with column pivoting", *Pattern Recognit.* 41 (4) (2008) 1373-1383, ISSN 0031-3203, <http://dx.doi.org/10.1016/j.patcog.2007.09.008>.
- [64] C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining Knowl. Discov.* 2 (2) (1998) 121-167, ISSN 1384-5810, <http://dx.doi.org/10.1023/A:1009715923555>.
- [65] R. Burnett, J.F. Watson, S. Elder, "The application of modern signal processing techniques for use in rotor fault detection and location within three-phase induction motors", *Signal Process.* 49 (1) (1996) 57-70.
- [66] B.D. Forrester, "Time-frequency analysis in machine fault detection", in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 406-423 (Chapter 18).
- [67] B. Boashash, P.J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques", *IEEE Trans. Acoust. Speech Signal Process.* 38 (11) (1990) 1829-1841.
- [68] W.J. Staszewski, K. Worden, G.R. Tomlinson, "Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition", *J. Mech. Syst. Signal Process.* 11 (5) (1997) 673-692.
- [69] S.K. Lee, P.R. White, "Higher-order time-frequency analysis and its application to fault detection in rotating machinery", *J. Mech. Syst. Signal Process.* 11 (4) (1997) 637-650.
- [70] S. Gu, J. Ni, J. Yuan, "Non-stationary signal analysis and transient machining process condition monitoring", *Internat. J. Mach. Tools Manufact.* 42 (1) (2002) 41-51.
- [71] H. Zheng, Z. Li, X. Chen, "Gear fault diagnosis based on continuous wavelet transform", *J. Mech. Syst. Signal Process.* 16 (2-3) (2002) 447-457.
- [72] S. Prabhakar, A.S. Sekhar, A.R. Mohanty, "Detection and monitoring of cracks in a rotor-bearing system using wavelet transforms", *J. Mech. Syst. Signal Process.* 15 (2) (2001) 447-450.
- [73] D.M. Onchis, "Observing damaged beams through their time-frequency extended signatures", *Signal Process.* 96 (2014) 16-20.
- [74] G.A. Velasco, N. Holighaus, M. Dörfler, T. Grill, "Constructing an invertible constant-Q transform with non-stationary Gabor frames", in: Proc. 14th Int. Conference on Digital Audio Effects (DAFx-11), 2011, 7 pp.
- [75] S.M. Han, H. Benaroya, T. Wei, "Dynamics of transversely vibrating beams using four engineering theories", *J. Sound Vibrat.* 225 (5) (1999) 935-988.
- [76] Y. Rubner, C. Tomasi, L.J. Guibas, "A metric for distributions with applications to image databases", in: Sixth International Conference on Computer Vision, 1998, pp. 59-66.
- [77] Y. Rubner, C. Tomasi, L.J. Guibas, "The earth mover's distance as a metric for image retrieval", *Internat. J. Comput. Vision* 40 (2) (2000) 99-121.
- [78] O. Pele, M. Werman, "Fast and robust Earth Mover's Distances", in: Proc. IEEE 12th International Conference on Computer Vision, 2009, pp. 460-467.
- [79] F. Hlawatsch, F. Auger (Eds.), *Time-Frequency Analysis: Concepts and Methods*, ISTE/Wiley, London, UK, 2008.

This page intentionally left blank

TIME-FREQUENCY METHODOLOGIES IN NEUROSCIENCES

16

INTRODUCTION AND OVERVIEW

This chapter presents a number of time-frequency (t,f) techniques that can provide advanced solutions to several problems in neurosciences with focus on the monitoring of brain abnormalities using the EEG (t,f) characteristics as a diagnosis and prognosis tool. The methods presented illustrate the improved performance obtained by using a time-frequency approach to process EEG data, including a focus on detecting abnormalities in sick newborns in a Neonatal Intensive Care Unit (NICU) as well as mental health issues in elderlies.

The material includes methods for the assessment of newborn EEG and ECG abnormalities using a time-frequency identification approach (Section 16.1); TF modeling of nonstationary signals with illustration on newborn EEGs (Section 16.2); time-frequency features for nonstationary signal classification with illustration on newborn EEG burst-suppression detection (Section 16.3); time-varying analysis of brain networks using the EEG for the analysis and detection of Alzheimer disease (Section 16.4); EEG time-frequency analysis and noise reduction using empirical mode decomposition (Section 16.5). Finally the chapter concludes with a discussion on the perspectives of using advanced (t,f) methods for medical diagnosis and prognosis in other areas of neurosciences (Section 16.6).

16.1 TIME-FREQUENCY DIAGNOSIS OF ABNORMALITIES IN NEWBORN PHYSIOLOGICAL SIGNALS⁰

A key objective of clinical neuroscience is the study of brain activity using the electroencephalogram (EEG) and other signals to understand the characteristics of normal activity and detect any deviation resulting in an abnormal activity. The seriousness of the abnormality can then be assessed in terms of risk to the patient.

Typically, multimodal and multichannel data are recorded in the clinic including EEG and heart data. The EEG represents electrical activity on the scalp that is recorded using multiple sensors resulting in several channels of data (ranging from 20 to 512). Heart data can include electrocardiogram (ECG), which represents the electrical activity of the heart, and heart rate variability (HRV), both of which are of interest as autonomic control of the heart by the brain manifests in these signals.

The signal processing engineering approach analyzes the signals with the view of extracting signal information in the form of characteristics that can then be linked to known or unknown conditions of the brain. This approach is used for adults, children, and newborns.

In the specific case of the newborn, the perinatal period has a high risk of morbidity and mortality, due to cerebral injury during birth. Minimizing this risk is a key clinical problem requiring practical solutions. The EEG is a high-value noninvasive tool for screening, diagnosis, and monitoring in the clinic. Given the simultaneous recording of other signals, a multimodal multichannel approach can be established to maximize the use of available physiological information. Innovative signal-processing methods can produce high quality information from the multimodal multichannel data, resulting in discriminating features that can be used to reduce neurodevelopmental disability and its high financial costs to society.

This section illustrates the application of time-frequency (t,f) methods to newborn abnormality diagnosis using the physiological signals routinely recorded in the Neonatal Intensive Care Unit (NICU). Three specific applications are outlined: heart rate estimation from the ECG, seizure detection in the EEG, and classifying the abnormality of the EEG.

16.1.1 THE CLINICAL CONTEXT OF THE NEWBORN

16.1.1.1 Monitoring of the Newborn at Risk

The monitoring of physiological signals in the NICU is a critical technological aspect of clinical care. These signals represent parameters related to the physiological function of the circulatory, neurological, and respiratory systems. They are measured from an array of different sensors attached to the newborn. Consistent and continual deviation from a nominal range (homeostasis) implies a reduction or absence of autonomic control that may result from an underlying pathology. This may require prompt intervention by the clinician. The detection of such abnormal situations is therefore critical to the health of the patient.

Such signals are typically displayed on vital-sign monitors, and deviations from a normal range are highlighted using alarms. In some situations, signs of abnormal function in a physiological signal

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa) and **N.J. Stevenson**, Helsinki University Central Hospital, 00029 HUS, Finland (n.stevenson@ucc.ie). Reviewers: J. M. O'Toole and L. Rankine.

are not manifested by a simple absolute increase or decrease in the signal amplitude, for example, the detection of seizures using the EEG or sepsis using the ECG. Modern biomedical signal analyses attempt to extract these signs of abnormality using more sophisticated algorithms, taking into account several key parameters of signals best estimated using (t,f) representations. The advantage of such advanced (t,f) methods is that they can extract more information from these signals (in the form of parameters, such as the instantaneous frequency (IF) or other spectral features) in order to assist medical decision making. Such methods can also aid in ensuring that the signals generated by the sensing equipment are accurate and not artefactual.

Physiological signals commonly monitored in the NICU are described below, followed by three specific examples where (t,f) methods were used successfully.

16.1.1.2 Physiological Signals in the NICU

The recorded signals are stored, processed and then displayed for interpretation by the clinician. Common physiological parameters are given below.

- EEG (voltage): This recording of the electrical activity of the brain is made using several electrodes positioned on the scalp over brain lobes.
- ECG (voltage): This is a separate but simultaneous multichannel recording of the electrical activity of the heart made using several electrodes placed on the body. The morphology and timing of waveform landmarks generated during a cardiac cycle (a heartbeat) provide important information on the heart function. The timing of the beats provides an estimate of the heart rate (HR) in beats per minute which is also diagnostically useful.
- Blood pressure (mmHg): This measure of the hydraulic pressure within the circulatory system is a key component for estimating the cardiac output. It is defined with two parameters: the systolic (during a beat) and diastolic (between beat) pressure, which can be summarized as a mean arterial pressure.
- Oxygen saturation (%): To monitor the percentage of oxygen in the blood, peripheral oxygenation saturation can be measured with an optical sensor on the hand or foot. Near-infrared spectroscopy can measure tissue oxygenation in the brain using a sensor positioned on the scalp.
- Respiration rate (breaths per minute): This is a measure of the movement of the lungs corresponding to breathing (inhale followed by exhale). It can be measured with sensors that respond to movement or bioimpedance.

16.1.2 NEWBORN SIGNAL ANALYSIS: THE (t, f) APPROACH

The above-mentioned physiological signals are inherently time-varying (i.e., nonstationary). Traditional stationary methods of analysis attempt to account for this time-variation by segmenting signals into short-duration epochs and applying stationary methods. The use of (t,f) signal processing eliminates the need for this assumption (which is commonly flawed) and permits the use of epoch lengths that result in the highest quality analysis [1,2].

The analysis of such newborn physiological signals requires the design of a joint (t,f) representation of signal energy to extract more precise and pertinent information for the task at hand, be it HR estimation, seizure detection or abnormality classification. A critical stage to any analysis is, therefore, the selection of a (t,f) distribution (TFD) that results in an optimal analysis and precise diagnosis.

16.1.2.1 TFD Selection

A desirable TFD should have the typical requirements of high (t,f) resolution and reduced cross-terms. It may also be required to satisfy additional properties such as marginal/energy conservation and nonnegativity or directional energy concentration. These properties are related to the window and kernel used to form the TFD in the time-frequency (t,f) , Doppler-lag (ν, τ) , time-lag (t, τ) or Doppler-frequency (ν, f) domains [1,3].

For example, the performance and characteristics of several TFDs were compared to find an optimal representation of newborn EEG data in the (t,f) domain in Ref. [1]. TFD has been applied to epochs of real newborn EEGs for various data window lengths and TFD parameter values. The performances were compared both visually, and quantitatively using objective criteria (described in Section 7.4). Based on these criteria, the B-distribution (BD) was selected as a suitable TFD for newborn EEG signals, although more advanced but more complex TFDs exist (see Sections 5.9 and 12.5). The (t,f) analysis was performed using the *Time-Frequency Signal Analysis package* (qspace.qu.edu.qa/handle/10576/10904), described in Chapter 17.

The BD is defined in terms of its time-lag kernel (see Section 2.7.5) as

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{|\tau|}{\cosh^2(t-u)} \right)^{\beta} z(u + \tau/2) z^*(u - \tau/2) e^{-j2\pi f \tau} du d\tau. \quad (16.1.1)$$

The parameter β ($0 < \beta \leq 1$) controls the sharpness of the cutoff of the two-dimensional filter in the ambiguity domain. The modified B-distribution (MBD) uses a slightly different kernel [1,3] obtained by making the following change in Eq. (16.1.1) (see Chapters 2 and 3 for more details)

$$\left(\frac{|\tau|}{\cosh^2(t)} \right)^{\beta} \rightarrow \left(\frac{\cosh^{-2\beta}(t)}{\int_{-\infty}^{\infty} \cosh^{-2\beta}(\xi) d\xi} \right). \quad (16.1.2)$$

The advantage of the MBD is due to the shape of the smoothing window that ensures more smoothing in the time-domain rather than the frequency domain for cross-term attenuation. This results in a representation that accentuates continuous multi-component signals with slowly varying frequency content; such signals were shown to be a good model for newborn EEG seizures [1,16].

16.1.2.2 Time-Frequency Patterns of Physiological Signals

This section illustrates the array of (t,f) patterns present in one key modality of newborn physiological signal, the EEG. Newborn EEG can be considered as a colored-noise process with distinct periods of amplitude modulation (AM) that denote different states of brain activity. This background signal is often interspersed with several distinct patterns that relate to maturational characteristics of the brain or disturbances in the normal function [2,4]. Several examples of (t,f) patterns are shown in Fig. 16.1.1 and detailed below.

1. Normal patterns:

Active sleep: This is the most common type of normal healthy newborns. It consists of combined theta (4-8 Hz) and delta (0.5-4 Hz) waveforms with intermixed irregular activity [5]. This pattern is represented in the (t,f) domain as a colored-noise process with a spectral coloring following an inverse power law. It has been modeled as a fractional Brownian process [4].

Quiet sleep: This is the second-most common type of normal healthy newborns. There are two activities that define quiet sleep: high-voltage slow-wave and tracé alternant [5]. Tracé alternant is

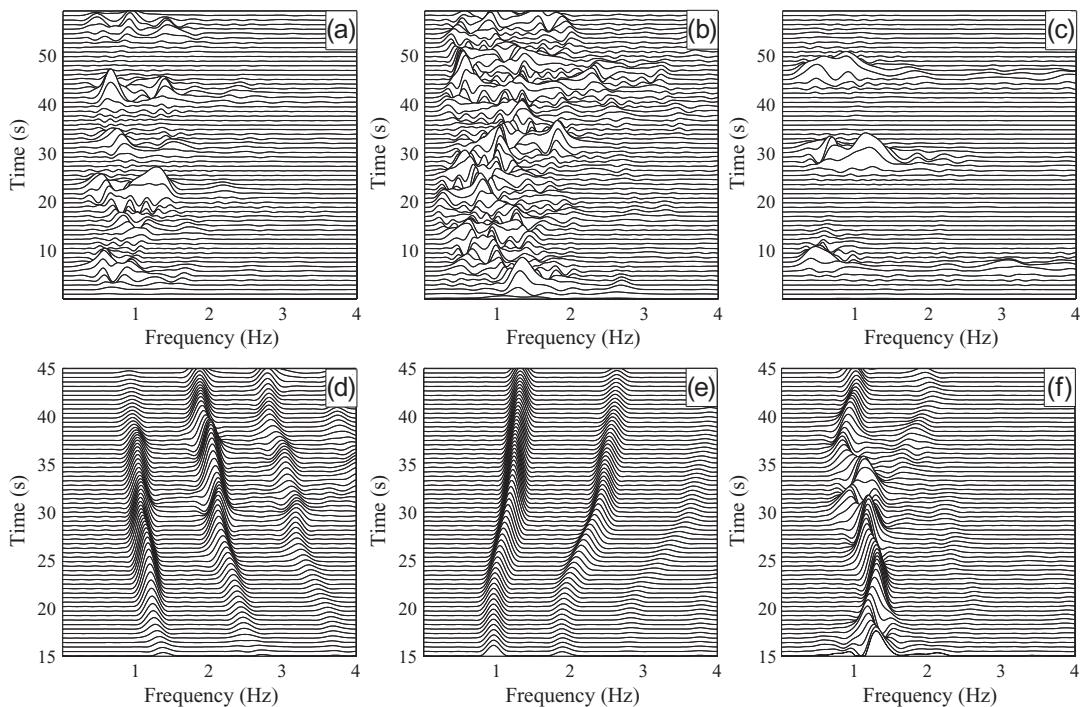


FIGURE 16.1.1

Time-frequency patterns of newborn EEG: (a) tracé alternant in quiet sleep; (b) active sleep; (c) burst suppression; (d) seizure with decreasing linear FM; (e) seizure with increasing linear FM; (f) seizure with complex, piecewise linear FM.

the predominant activity and can be defined as a colored-noise process modulated by a time-varying sinusoid. This is represented in the (t,f) domain as a periodic concentration of high-energy activity, followed by low-energy activity. An analogous pattern exists in premature newborns. This pattern is called tracé discontinu and the AM is more significant with long periods of inactivity in the EEG.

2. Abnormal patterns:

Seizure: This is a common abnormality seen in the EEG of critically ill newborns. They manifest on the EEG as *a clear ictal event characterized by the appearance of sudden, repetitive, evolving, stereotyped waveforms that have a definite beginning, middle and end, and last for a minimum of 10 s* [6]. Newborn seizures can be modeled as multicomponent with piecewise linear IF laws with AM [1]. They appear in the (t,f) domain as a piecewise main ridge (component) describing the IF, i.e., a frequency modulation (FM) with slowly time-varying amplitude alongside several subharmonic ridges.

Burst suppression: One can observe short-period signals with a high energy lasting for few seconds and usually occurring at frequencies below 4 Hz. These features are characteristic of the *burst-suppression* pattern. This is an abnormal pattern in the EEG; it is defined as a burst of

high-voltage activity lasting 1-10 s and composed of various waveforms—e.g., delta (0-4 Hz) and theta (4-8 Hz) with superimposed and intermixed spikes, sharp waves, and faster activity—followed by a marked attenuation [5].

16.1.3 DETECTION OF SEIZURES IN NEWBORN EEG

Newborn seizure is a common emergency that is routinely treated with an array of anticonvulsant medications in the NICU. Seizures do not always have overt clinical manifestations, with as many as 70% exhibiting no clinical signs. The interpretation of the EEG is a difficult task and requires a level of expertise that is not readily available around the clock. Automated analyzes of the EEG offer a potential practical solution to the problem of seizure detection. Newborn seizures emerge from an EEG background that exhibits stochastic properties with spectral coloring, AM and long-term changes in states. The EEG is further contaminated by several artifacts of biological or external origin, such as fine/gross muscle movement or interference from power supplies. The repetitive and evolving characteristics of seizure emerging from a nonseizure background contribute to the finding that (t,f) analysis can be used to design a detection technique [1].

The *matched filter* is a basic method for detection. It essentially reduces to a correlator receiver whose output is compared to a threshold [7] (see [Section 12.5](#)). In the (t,f) approach, the correlator receiver is replaced by a 2D correlation between the TFDs $\rho(t,f)$ of $x(t)$ and $s(t)$ (see also [Section 12.5](#)).

A correlator receiver usually requires knowledge of a reference signal (e.g., its wave shape or related information, such as (t,f) characteristics or features) and the noise statistics. For the (t,f) -based matched detector applied to seizure detection, the reference TFD $\rho_{\text{ref}}(t,f)$ can be typically based on a piecewise linear FM model with parameters that are estimated directly from the EEG epoch under analysis [1]. The chosen correlator statistic $\eta(z)$ is the two-dimensional cross-correlation between the TFD of the EEG analytic associate $z(t)$ and the reference TFD ρ_{ref} [7]:

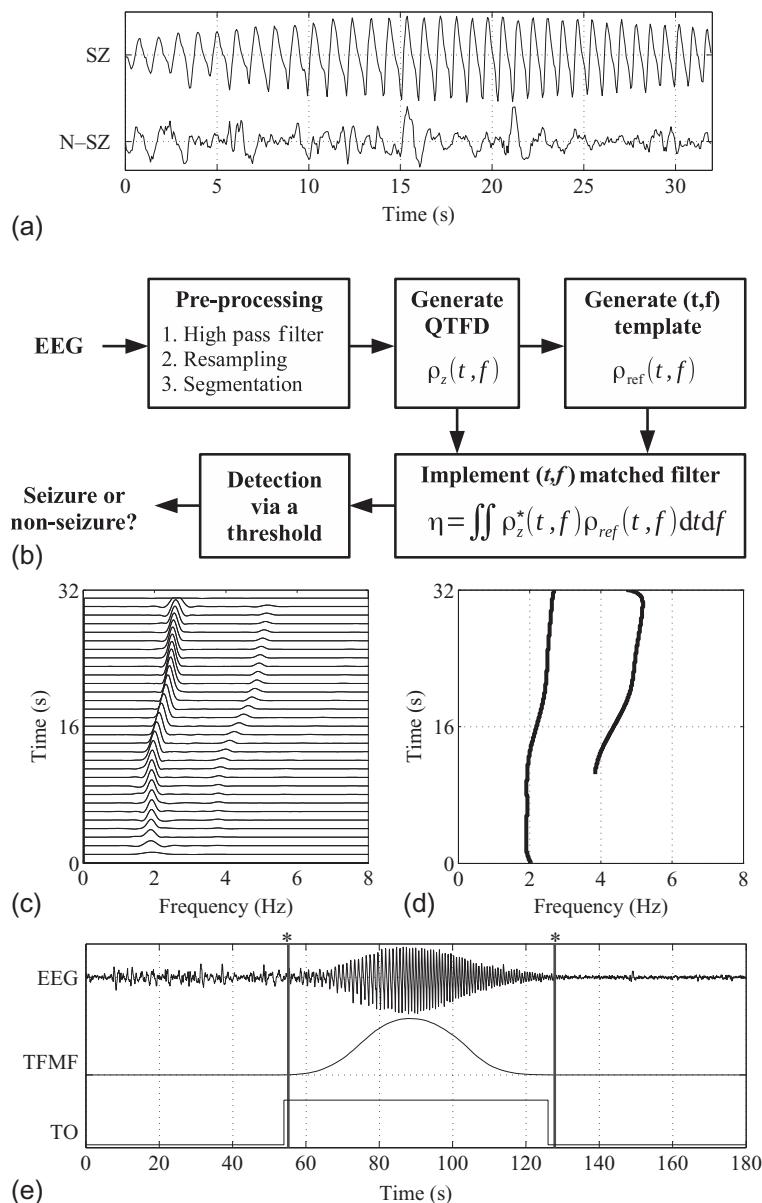
$$\eta(z) = \iint \rho_{\text{ref}}(t,f) \rho_z^*(t,f) dt df, \quad (16.1.3)$$

where z is the analytic signal corresponding to the EEG signal under consideration. The template parameters are estimated from the EEG epochs. The detection of seizures uses the (t,f) matched filter shown in [Fig. 16.1.2](#), and it results in a good performance when applied to seizure detection in real-world newborn EEG [2,7]. An example of a (t,f) -matched filter process is shown in [Fig. 16.1.2](#).

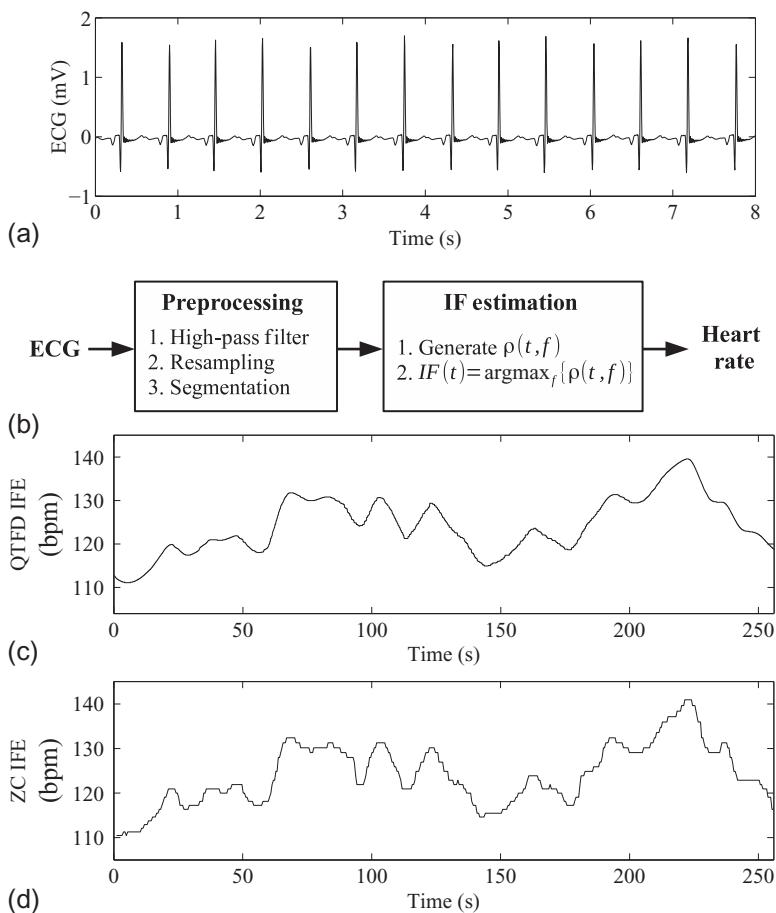
16.1.4 ESTIMATING THE HEART RATE FROM ECG

The heart rate is another important clinical indicator in the NICU. Terms such as tachycardia (high heart rate) and bradycardia (low heart rate) are commonly used in diagnosis. Furthermore, the oscillations of HR, called HRV, reflect the modulation by the autonomic nervous system, suggesting a relationship between HRV and activity in the brain stem. This intrinsic link between heart rate and neurological function has been used to detect seizures and assess the severity of hypoxic ischemic encephalopathy (HIE) in the NICU setting.

The estimation of the heart rate from the ECG signal is fundamentally an IF estimation problem [8]. This is traditionally performed by post-processing (interpolating) the output of peak detection algorithms, a method that is similar in concept to a zero-crossing IF estimation algorithm (see

**FIGURE 16.1.2**

Implementation of a (t, f) -matched filter (TFMF) for automated seizure detection on 180 s of newborn EEG: (a) 32 s epochs of seizure (SZ) and nonseizure (N-SZ); (b) seizure-detection algorithm for a single channel of EEG (a multi-channel detector can be implemented by postprocessing the single channel outputs); (c) quadratic TFD of seizure (MBD, $\beta = 0.01$); (d) adaptive template estimate using multicomponent IF estimation algorithm; (e) the EEG, corresponding TFMF output and TFMF output after thresholding (TO) (asterisks denote the start and end of seizure as marked by the human expert).

**FIGURE 16.1.3**

The estimation of HR from the ECG: (a) a short duration epoch of the ECG signal; (b) the HR estimation algorithm based on the analysis of a quadratic TFD; (c) the output of the HR estimation algorithm based on the analysis of a TFD; (d) the output of an alternate algorithm based on zero-crossings analysis (IFE denotes instantaneous frequency estimate which in this case is equivalent to a HR estimate).

[Chapter 10](#) for several IF estimation methods). As the ECG signal is typically recorded with a high signal-to-noise ratio, and is multicomponent with a harmonic relationship between components, the TFD peak IF estimate $f_i(t) = \text{argmax}_f\{\rho_z(t,f)\}$ is effective. The estimation of HR from ECG using this IF estimate is shown in [Fig. 16.1.3](#).

The peak detection methods of IF estimation result in a nonuniformly sampled signal (as the temporal locations of the peaks are variable). As a result, the IF estimate must be interpolated, typically with a zero-order hold or a cubic spline, in order to generate a uniformly sampled signal. Methods based on (t,f) analysis overcome this limitation and do not require interpolation.

16.1.5 CLASSIFYING ABNORMALITY IN NEWBORN EEG

Hypoxic-ischemic encephalopathy (HIE) is a condition that develops in newborns when the oxygen and blood supply to the brain is interrupted for a period of time. The degree of HIE is commonly classified into three grades: mild, moderate and severe, reflecting the level of injury. Therapeutic hypothermia is a treatment for HIE that must be applied within 6 h of birth for efficacy. The treatment can have complications, so it is only recommended for newborns with moderate to severe HIE. Determining which newborns require treatment at an early stage is critical.

The EEG can be used for assessment of brain activity and for grading HIE severity with high accuracy. This method normally requires specific expertise for practical implementation in a NICU setting. Computer-assisted methods of analysis offer a potential automatic solution. For example, a computer can be used to estimate the severity of the HIE by classifying hourly epochs of multichannel EEG within 6 h of birth.

Time-frequency methods can be applied to this task [9] using a time-varying model of newborn EEG background. In this case, the EEG is modeled as a colored noise process with AM,

$$\text{EEG}(t) = a(t)x(t) = a(t)T_{\text{filter}}\{y(t)\}, \quad (16.1.4)$$

where $a(t)$ is the AM, the transformation $T_{\text{filter}}\{\cdot\}$ is a linear or nonlinear filter and $y(t)$ is white Gaussian noise. The frequency response of the transformation can be approximated with a power-law response of $1/f^{\alpha}$ that varies over time, resulting in nonstationary behavior. By constraining the frequency content of $a(t)$ to be much less than $x(t)$ so there is no spectral overlap between $a(t)$ and $x(t)$, and assuming information on the underlying signal transformation $T_{\text{filter}}\{\cdot\}$ can be estimated from the frequency-domain representation of $x(t)$ (via estimates of the IF, $f_i(t)$), the interpretation of $a(t)$ and $T_{\text{filter}}\{\cdot\}$ aligns with the visual interpretation of the EEG made by the neurophysiologist of amplitude and frequency content, respectively.

The AM $a(t)$ and IF $f_i(t)$ can be estimated as follows:

$$|a(t)|^2 = \int_{-\infty}^{\infty} \rho_z(t,f) \, df, \quad (16.1.5)$$

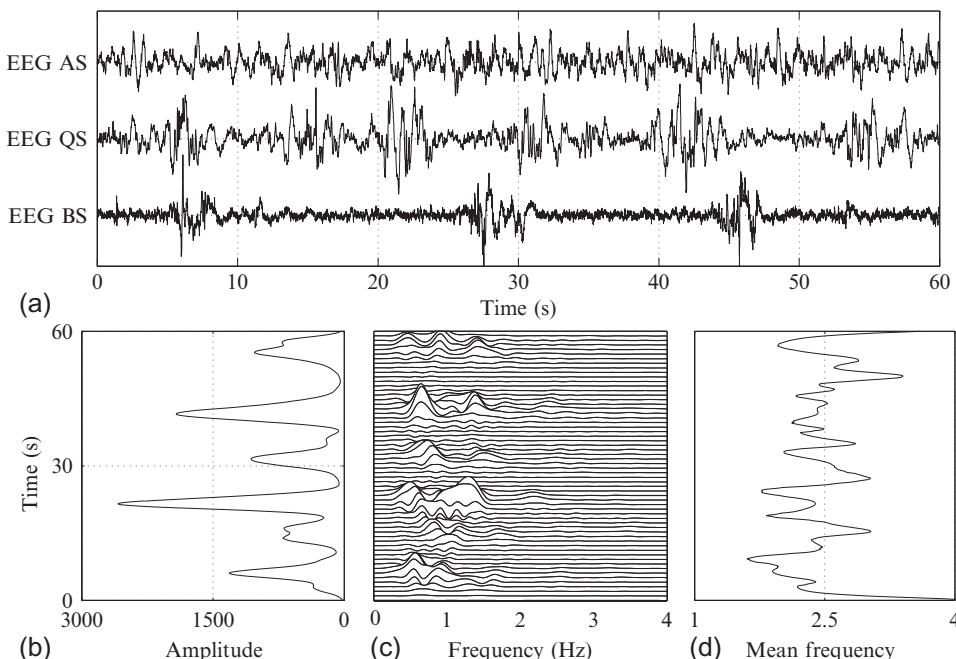
$$f_i(t) = \frac{\int_{-\infty}^{\infty} f \rho_z(t,f) \, df}{\int_{-\infty}^{\infty} \rho_z(t,f) \, df}. \quad (16.1.6)$$

If the kernel used to generate the TFD is a delta function (resulting in the Wigner-Ville distribution (WVD)), then the mean frequency equals the IF.

Summary statistics of these AM and FM parameters provide a useful feature set to classify the newborn EEG in several grades of abnormality [9]. Examples of the EEG, as well as the AM and IF parameters, are shown in Fig. 16.1.4.

16.1.6 SUMMARY AND CONCLUSIONS

One of the key problems in clinical neuroscience is detecting and interpreting abnormalities in the brain of people, and especially critically ill newborns. In such situations, the acquisition of physiological signals allows the possibility of an automated analysis of these signals, which can provide useful information for the clinician. Physiological signals routinely collected in a clinical setting include

**FIGURE 16.1.4**

Normal and abnormal EEG signals, AM and IF used for computer-assisted classification: (a) 60s epochs of active sleep (AS), quiet sleep (QS), both normal patterns and burst suppression (BS), abnormal pattern; (b) AM estimated from the TFD of QS in (c); (d) IF estimated from the TFD of QS in (c).

the ECG, EEG, blood pressure, oxygen saturation and respiration. These signals have nonstationary characteristics that justify a (t,f) approach. The successful implementation of a dedicated (t,f) processing of physiological signals has the potential to improve the quality and type of information extracted, which can assist the clinician in diagnosis. More details can be found in Sections 16.2, 16.3, and 16.6 and in Refs. [10,11].

16.2 TIME-FREQUENCY MODELING OF NONSTATIONARY SIGNALS FOR NEWBORN EEG⁰

Accurate signal models can be useful in many applications, such as filtering, feature extraction and compression, by helping provide insights into the underlying signal-generating system. Considering

⁰Authors: **L.J. Rankine**, Business Engineers Australia, QLD (luke@beaustralia.com.au), **N.J. Stevenson**, Helsinki University Central Hospital, 00029 HUS, Finland (n.stevenson@ucc.ie), and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: G. Azemi and J. O'Toole.

the signals' nonstationary behavior as demonstrated in Chapters 2 and 1, it is useful to attempt to model these signals by taking into account these time-varying characteristics. The development of such a model can be separated into four distinct stages: time-frequency (t,f) analysis, (t,f) model selection, model fitting, and model evaluation. This section reviews the process of (t,f) modeling using the general multicomponent time-varying signal model outlined by Boashash [12]. This signal model is applicable to a number of real-life signals with nonstationary characteristics, such as the example of newborn EEG seizure signals that illustrate this section.

16.2.1 THE TIME-FREQUENCY MODELING PROCESS

16.2.1.1 Time-Frequency Analysis

To apply the time-varying multicomponent signal model of Boashash [12], we must first establish and characterize the time-varying nature of the signal. This can be done by observing the distribution of signal energy in the joint (t,f) domain—a qualitative process undertaken by the signal analyst. Examples of (t,f) patterns that may characterize a nonstationary signal include: mono-component or multicomponent, linear or nonlinear instantaneous frequency (IF) laws or frequency modulation (FM), amplitude modulation (AM), and discontinuities/transients.

Methods of generating (t,f) representations include the short-time Fourier transform (STFT), quadratic (t,f) distributions (QTFDs), wavelets, atomic decomposition using (t,f) dictionaries and the empirical mode decomposition. Of these, the reduced-interference class of QTFDs has been shown to be particularly useful in analyzing and characterizing nonstationary signals that contain multiple components with time-varying IF laws [13]. In practice, it is recommended that several different TFDs are used during the initial analysis before a final choice is made.

16.2.1.2 Model Selection

A general (t,f) model that has been used to describe real signals is [12]

$$s(t) = \sum_{k=1}^K a_k(t) \cos \left(2\pi \int_0^t f_k(\tau) d\tau + \theta_k \right) = \sum_{k=1}^K a_k(t) \cdot \cos \theta(t) \quad (16.2.1)$$

(see also Chapter 1, p. 52), where t is time, K is the number of components, and the k th component has an AM of $a_k(t)$ and an IF law described by $f_k(t)$. Time-varying frequency content is modeled by $f_k(t)$, time-varying amplitude by $a_k(t)$, and time-varying morphology or wave shape as varying proportionalities between amplitudes of components. The variable θ_k is the starting frequency of the instantaneous phase function. Please note, there is a direct relationship between the instantaneous phase and the IF given by $\theta(t) = 2\pi \int_0^t f_k(\tau) d\tau + \theta_k$ (see Chapter 1, p. 50). The use of a linear sum of two functions can result in an infinite number of choices of K , $a_k(t)$ and $f_k(t)$ that represent $s(t)$ in general, but the use of the analytic associate of the signal limits these choices by ensuring that the frequency content of the AM of a component is significantly lower than $\cos \left(2\pi \int_0^t f_k(\tau) d\tau + \theta_k \right)$ (see Section 1.2.3, p. 44). This definition can be further simplified in some applications by assuming a fixed harmonic relationship between time-varying components, that is, $f_k(t) = kf_1(t)$.

An example of a function that can be used to model both the amplitude and frequency modulation is the polynomial function

$$g(t) = r_0 + r_1 t + r_2 t^2 + \cdots + r_n t^n, \quad (16.2.2)$$

whose parameters are the order (degree) n and the coefficients $[r_0, r_1, \dots, r_n]$.

16.2.1.3 Model Fitting

Once a model is determined, its parameters need to be extracted from the (t, f) representation of signals from the class of interest. The advantage of the multi-component time-varying models is that signal components are well represented by QTDFs (see [Section 3.3.1](#), p. 119). This synergy between model and (t, f) representation simplifies the process of model parameterization. Automated methods of parameterization can be based on techniques developed for the estimation of multicomponent IF laws in a signal [2,13] (see [Section 10.6](#)). Polynomial functions can then be fitted to the extracted FM and AM functions using a least-squares minimization. For example, in the case of a nonstationary signal with polynomial modulation of amplitude and frequency as shown in [Eq. \(16.2.2\)](#), the parameters n and $[r_0, r_1, \dots, r_n]$ for each AM and FM component would be required. Initially, the IF laws would be identified; these IF laws can then be used to estimate the AM directly from the QTDF. Methods for estimating the IF law are provided in [Section 10.6](#).

In the majority of real signals, the model parameters are random variables. To simulate a signal accurately through the random choice of model parameters, we need to find an accurate description of the probability density function of the model parameters; however, it must be clearly noted that not all model parameters are independent; joint distributions of some parameters must be estimated [14]. For example, the IF components of a multicomponent signal may be highly dependent on the main signal component. Only through the accurate representation of the random distribution of model parameters can accurate signal simulation be achieved.

16.2.1.4 Model Evaluation

In order to assess the model and evaluate its performance, it is important that signal examples generated using the model are compared to examples of the real signal. There is an array of measures that can be used for this comparison, such as root-mean-square error (RMSE), normalized RMSE and linear correlation coefficient [15]. These measures can be applied in the time, frequency or joint time-frequency domain; however, another consideration to be made in evaluating a model is the number of model parameters; it is important in practice to minimize the number of model parameters while at the same time providing an accurate model.

16.2.2 EXAMPLE OF TIME-FREQUENCY MODELING ON SYNTHETIC DATA

To demonstrate the process of (t, f) modeling a synthetic example is presented. A two-component signal with a linear-IF component and quadratic-IF component is used with polynomial AM laws for each component. It is defined as

$$s[n] = a_1[n] \cos(\theta_1[n]) + a_2[n] \cos(\theta_2[n]), \quad (16.2.3)$$

where

$$a_1[n] = 3.80 \times 10^{-5} (n/f_s)^3 - 3.64 \times 10^{-3} (n/f_s)^2 + 8.36 \times 10^{-2} (n/f_s) + 1,$$

$$a_2[n] = -4.41 \times 10^{-4} (n/f_s)^2 + 3.29 \times 10^{-2} (n/f_s) + 0.6,$$

$$\theta_1[n] = \frac{2\pi}{128} (n/f_s)^2 + \pi (n/f_s),$$

$$\theta_2[n] = \frac{2\pi}{3096} (n/f_s)^3 - \frac{2\pi}{32} (n/f_s)^2 + 6\pi (n/f_s)$$

with

$$n = [0, 1, \dots, 512], f_s = 8 \text{ Hz}; \text{ the IF is the derivative of } \theta.$$

The top signal in Fig. 16.2.1(a) is the time-domain representation of the original two-component signal. The (t, f) representation using a smoothed WVD (2D Hamming window of length 31) is shown in Fig. 16.2.1(b), which indicates that there are two clear signal components. The AM and IF of each component can be extracted from the TFD using the IF estimation method outlined in Ref. [13]. The

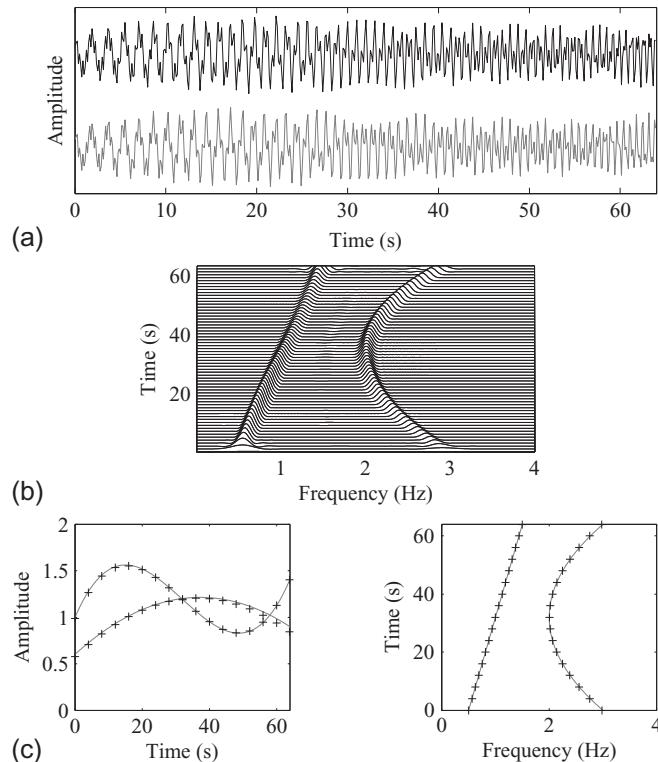


FIGURE 16.2.1

An example of the model-fitting process for a two-component time-varying signal: (a) the original two-component signal (top) and the modeled signal (bottom); (b) the representation of the original signal energy in the (t, f) domain; (c) crosses show the AM and FM estimates generated by fitting polynomials to the amplitude and frequency estimated directly from the (t, f) representation; continuous curves show the AM and FM laws used to generate the original signal in (a).

AM and IF can then be fitted to a polynomial function using a least-squares method. Figure 16.2.1(b) shows the actual AM and IF (solid lines) compared to the original polynomial (crosses). The bottom signal in Fig. 16.2.1(a) is the simulated signal based on these parameter estimates. The simulated signal and the original signal are highly correlated in the joint- (t, f) domain (the linear correlation coefficient is 0.998).

A database of signals from a class of interest permits the estimation of the distributions of the modeled signal parameters. Random selection of model parameters from these distributions results in the simulation of a variety of signals from the signal class of interest.

16.2.3 APPLICATION TO NEWBORN EEG MODELING

16.2.3.1 Newborn EEG Signal Characteristics

The newborn EEG is a multichannel stochastic signal embedded with deterministic patterns that reveal important information about cortical function. The EEG signal is of the order of tens of microvolts and has a frequency range centered in the delta band (0.5–4 Hz) with little energy above 13 Hz. The EEG from a normal, healthy term neonate exhibits continuous activity that is synchronous and symmetric across hemispheres with distinct state changes corresponding to sleep-wake cycling interspersed with several normal patterns [5]. Disruption to cortical activity is, therefore, manifested on the EEG and can be used to diagnose the underlying neonatal encephalopathy.

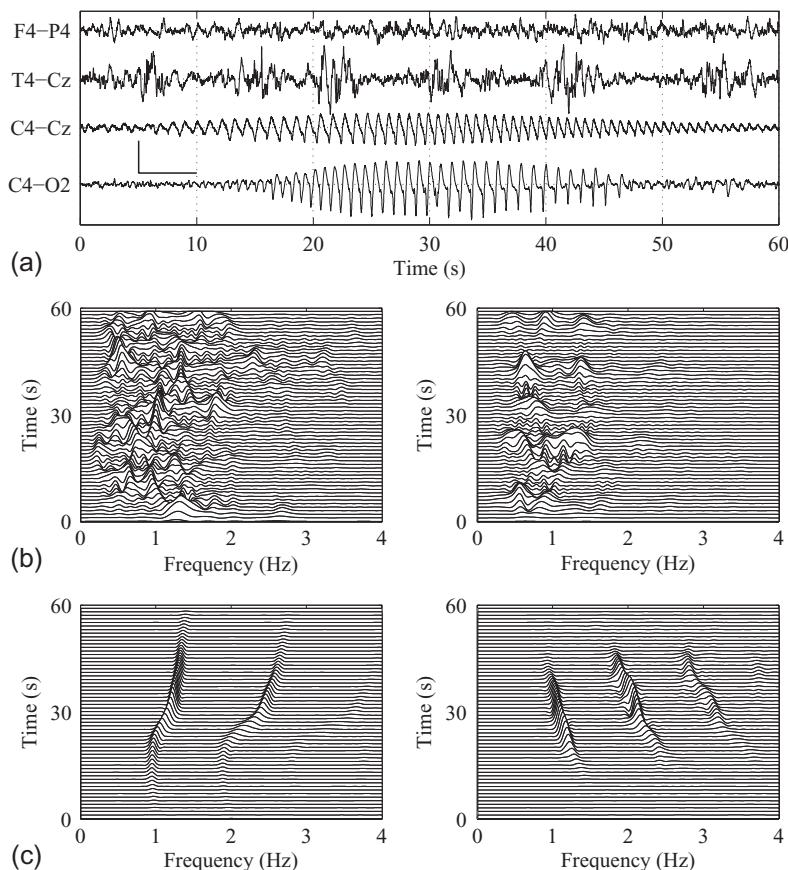
The primary use of the EEG in the neonatal intensive care unit (NICU) is as an aid to the detection of seizures. Seizures are seen in a range of neonatal encephalopathy with an incidence of 1 to 3 per 1000 (admission rates to the NICU are approximately 70 per 1000 births) [5]. The neonatal brain is prone to seizures, as it is more excitatory than the adult brain [2]. Any disruption to normal brain function caused by an injury is therefore more likely to trigger a seizure in the newborn. Seizures are generated when large populations of colocated neurons synchronize and discharge with near-periodic repetition. Individually, seizures can last from 10 s to many hours (a condition known as *status epilepticus*). Collectively, they can last for several days after the primary injury. As mentioned in the previous section, seizures manifest on the EEG as a clear ictal event characterized by the appearance of sudden, repetitive, evolving, stereotyped waveforms that have a definite beginning, middle and end, and last for a minimum of 10 s [6]. These seizures emerge from a background pattern that displays both stochastic and chaotic characteristics with notable modulations of amplitude apparent over minutes and hours [9,16]. Example epochs of newborn EEG seizure, active sleep (nonseizure) and quiet sleep (nonseizure) are shown in Fig. 16.2.2 in the time and (t, f) domain.

16.2.3.2 Time-Frequency Modeling of Newborn EEG Seizure

Studies show that newborn EEG seizure is a nonstationary signal with moderate rates of change in the frequency content and multiple harmonically related components (see Fig. 16.2.2(c)) [1]. The characteristics exhibited by newborn EEG seizure in the (t, f) domain are mathematically definable and can be modeled with a limited parameter set using the general signal model in Eq. (16.2.1). Subsequent studies have modeled the AM with a polynomial function (a maximum order of 8) and the FM with a piecewise linear function ($M = 3$ pieces) with a limitation on the number of harmonics (K) to 5 [16].

In [16], the AM function of the k th component is defined as

$$a_k(t) = \phi(t; R_k, P, V_n). \quad (16.2.4)$$

**FIGURE 16.2.2**

Newborn EEG signals: (a) signal energy in the time domain for (from top to bottom) nonseizure (active sleep and quiet sleep) and seizure (two examples); the L-shaped scale indicator is 5 s wide and 100 μ V high; (b) signal energy in the joint- (t, f) domain for newborn EEG nonseizure in active sleep (left) and quiet sleep (right); (c) signal energy in the joint- (t, f) domain for newborn EEG seizure on C4-Cz (left) and C4-O2 (right).

The gain factor, R_k , is referred to as the harmonic ratio (i.e., the ratio between the average amplitude of the harmonic and the fundamental). This infers that $R_1 = 1$. The component amplitude modulation function is determined from a cubic spline interpolation of P randomly assigned turning points with amplitudes given by $a_k = R_k(0.67 + V_n)$ where the mean of $V_n = 0.33$. The locations of the turning points are found according to $q = \frac{N(p+X)}{p}$ where $p = [0, \dots, P - 1]$ and X is a stationary random process, uniformly distributed between 0 and 1. The boundary conditions of the cubic spline fit are set to have a derivative of zero.

The FM function of the k th component was defined as

$$f_k(t) = \sum_{m=1}^M F_m(t; \zeta_m, C_m) \operatorname{rect}\left(\frac{t - 0.5(B_{m+1} - B_m)}{B_{m+1} - B_m}\right), \quad (16.2.5)$$

where

$$F_m(t; \zeta_m, C_m) = \zeta_m t + C_m, \quad (16.2.6)$$

and

$$C_m = \begin{cases} f_{\text{st}} & \text{for } m = 1, \\ F_{m-1}(B_m; \zeta_{m-1}, C_{m-1}) - \zeta_m B_m & \text{for } m \geq 2. \end{cases} \quad (16.2.7)$$

Here, f_{st} is the start frequency, $\zeta = [\zeta_1, \zeta_2, \dots, \zeta_m]$ are the gradients of the IF law in Hz/s for each piece, $B = [B_1 = 0, B_2, \dots, B_M, B_{M+1} = N]$ are the turning points, N is the discrete length of the seizure, “rect” is the unit rectangular pulse function, and C_m are the alignment intercepts that ensure $f_k(t)$ is continuous.

In order to assess this model of the newborn EEG seizure, the model was fitted to real newborn EEG seizure epochs. The output from the model using the estimated parameters was compared with real signals using the linear correlation coefficient evaluated in the time, frequency and joint-(t, f) domains. The results of the analysis are shown in [Table 16.2.1](#).

The results indicated that the model accurately represented the time-frequency characteristics of the real newborn EEG signal. The model was then used as the basis of a newborn EEG seizure simulator. The distributions of the parameters in the parameter vector were estimated on a large dataset of newborn EEG signals. The source code for this newborn EEG simulator can be obtained from supplementary material available on this book’s website.

16.2.4 SUMMARY AND CONCLUSIONS

Signals with time-varying frequency content require time-varying models. This section describes four stages of (t, f) modeling including (t, f) analysis, time-varying model selection, model fitting, and model evaluation. This method of (t, f) modeling is then applied to newborn EEG seizure signals to illustrate the modeling process from initial (t, f) analysis, through model evaluation, to signal simulation.

Table 16.2.1: A Comparison of Real and Simulated Epochs (via a Fit of the (t, f) Model in the Time, Frequency and Joint-(t, f) Domains

| Domain | Time | Frequency | Time-Frequency |
|--------|---------------|---------------|----------------|
| R | 0.345 (0.176) | 0.799 (0.093) | 0.901 (0.056) |

A 1D and 2D linear correlation coefficient, R , was used to measure the comparison. The results are presented as a mean and a standard deviation (in parentheses) across 204 epochs, each containing 12.8 s of newborn EEG seizure from 12 neonates.

16.3 TIME-FREQUENCY FEATURES FOR NONSTATIONARY SIGNAL CLASSIFICATION WITH ILLUSTRATION ON NEWBORN EEG BURST-SUPPRESSION DETECTION⁰

16.3.1 CLASSIFICATION OF NONSTATIONARY SIGNALS USING INFORMATION EXTRACTED FROM TFDs

As discussed in [Section 12.6](#), the design of high-performance classification systems requires signal transformations that provide discriminative and effective features that are easier to extract and result in an enhanced classification. For applications involving nonstationary signals, features extracted from the signal's (t,f) representation can result in reliable and accurate classification systems, as they can leverage the nonstationary characteristics of the signal using the extra information available in the (t,f) domain. Such information cannot be obtained directly from the signal's representation in the time or frequency domain, as detailed in [Section 12.6](#).

For classification of nonstationary signals, TFDs allow the extraction of important features of the signals and provide greater insight into the nature of the information they carry [1]. Such information includes the number of signal components, their instantaneous frequencies (IFs) and instantaneous amplitudes (IAs), and the regions of energy concentration in the (t,f) domain (see [Chapters 1](#) and [2](#)). This section presents a selection of (t,f) -based features suitable for the classification of nonstationary signals. The features are used to detect abnormal brain activities observed in newborn electroencephalography (EEG) signals. In [Section 12.6](#), a methodology for extending time-domain and frequency-domain features to the joint- (t,f) domain is introduced, and new (t,f) -based features are defined. This section complements and supplements the material described in [Section 12.6](#).

16.3.2 TIME-FREQUENCY FEATURES

For a given real N -point discrete-time signal $x[n]$, the TFD $\rho_{z_x}[n, k]$ is calculated using Eq. [\(6.1.26\)](#):

$$\rho_{z_x}[n, k] = \text{DFT}_{m \rightarrow k} \left\{ G[n, m] * \left(z_x[n+m] z_x^*[n-m] \right) \right\}, \quad (16.3.1)$$

where $z_x[n]$ is the analytic associate of $x[n]$. The TFD in [\(16.3.1\)](#) can be represented as an $N \times M$ matrix ρ_{z_x} , from which the following (t,f) features can be extracted:

- IF-based features** are derived from statistics of the signal's IF, such as its mean and deviation. The IF-based features can be extracted from the signal's IF estimated from the first moment of its TFD (Eq. [\(3.1.8\)](#)) or from the IFs of the components of the signal. See [Chapter 10](#) for a thorough review of the methods and algorithms for IF estimation.
- IA-based features** are derived from the statistics of the signal IA. For a monocomponent signal, the envelope of $z[n]$, the analytic associate of $x[n]$, can be used as an estimate of the signal IA (see [Section 6.5.2](#)). For multicomponent signals, the components need to be separated before the IA of each component is estimated ([Section 10.6](#) presents two methods).

⁰Authors: **G. Azemi**, Razi University, Kermanshah, Iran (g.azemi@razi.ac.ir) and **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: E. Sejdic and N. A. Khan.

- 3. Matrix decomposition-based features** are derived from a decomposition of the TFD matrix ρ_{zx} .

Two matrix decomposition methods are commonly used, namely *singular value decomposition* (SVD) and *nonnegative matrix factorization* (NMF) as described below.

- (a) **The SVD** divides the TFD matrix ρ_{zx} into two subspaces of the form [10]:

$$\rho_{zx} = USV^H, \quad (16.3.2)$$

where U is a real matrix, S is an $N \times M$ diagonal matrix with nonnegative real numbers (S_i , $i = 1, 2, \dots, N$) on the diagonal, and V^H (the conjugate transpose of V) is a real unitary matrix. The diagonal entries of S are known as the singular values of ρ_{zx} . Features are then extracted from the singular values of the matrix ρ_{zx} . Such SVD-based features include the maximum and variance of the singular values. Another possible feature that is based on the SVD is the (t,f) complexity measure, i.e., $CM(x)$, which is derived from the Shannon entropy of the singular values of ρ_{zx} and represents the magnitude and the number of the nonzero singular values of a TFD [1]. It is given by:

$$CM(x) = - \sum_{i=1}^N \bar{S}_i \log_2 \bar{S}_i \quad (16.3.3)$$

where \bar{S}_i is the i th normalized singular value, i.e., $\bar{S}_i = \frac{S_i}{\sum_i S_i}$.

- (b) **The NMF** is another matrix decomposition method. It has the advantage of preserving the nonnegativity of the entries, which is important for meaningful physical interpretation in some circumstances. The NMF factorizes the TFD matrix ρ_{zx} into two nonnegative matrices as [17]

$$\rho_{zx} \approx W_{N \times R} H_{R \times M} = \sum_{r=1}^R w_r h_r, \quad (16.3.4)$$

where the columns of W and H^T are known as base and coefficient vectors, respectively, and R ($R \ll \min(M, N)$) is the decomposition parameter, which is usually application-dependent.

Note that the decomposition of ρ_{zx} using the NMF requires the matrix to be nonnegative [17]. The base vectors w_r , $r = 1, \dots, R$, can be interpreted as characteristic frequency structures, whereas the coefficient vectors h_r , $r = 1, \dots, R$, can be regarded as the temporal location of these structures [18]. Based on the decomposition in Eq. (16.3.4), the sparsity of the base and coefficient vectors are defined respectively as

$$S_{w_r} = \frac{1}{\sqrt{N}} \left(\sqrt{N} - \frac{\sum_{l=1}^N w_r(l)}{\sqrt{\sum_{l=1}^N w_r^2(l)}} \right), \quad r = 1, \dots, R, \quad (16.3.5)$$

$$S_{h_r} = \frac{1}{\sqrt{M}} \left(\sqrt{M} - \frac{\sum_{l=1}^M h_r(l)}{\sqrt{\sum_{l=1}^M h_r^2(l)}} \right), \quad r = 1, \dots, R, \quad (16.3.6)$$

and their statistics can be considered as NMF-based features.

- 4. TFD concentration measures** can also be used for the classification of nonstationary signals.

Different measures were presented in Section 7.3 and their properties were discussed. One such measure is:

$$M_p(x) = \left(\sum_{n=1}^N \sum_{k=1}^M |\rho_{z_x}[n, k]|^{1/p} \right)^p \quad (16.3.7)$$

with $p > 1$. The advantage of this measure is that it is not sensitive to small values in $\rho_{z_x}[n, k]$. Note that signals with power distributed all over the (t, f) plane have a larger $M_p(x)$, while those with power concentrated in certain areas are characterized by a smaller $M_p(x)$.

- 5. The (t, f) Rényi entropy** provides a measure of complexity of a nonstationary signal in the (t, f) plane and is given in Eq. (7.3.3) as

$$R_\alpha(x) = \frac{1}{1 - \alpha} \log \left(\sum_{n=1}^N \sum_{k=1}^M \frac{\rho_{z_x}^\alpha[n, k]}{\sum_n \sum_k \rho_{z_x}[n, k]} \right) \quad (16.3.8)$$

with $\alpha > 2$. The advantage of (t, f) Rényi entropy over other expressions of entropy is that it can also be applied to all TFDs that may assume negative values (see Section 7.3). The Rényi entropy can also be used to define another feature, which is the estimate of the number of signal components in a multicomponent signal [19]

$$n_\alpha(x) = 2^{R_\alpha(x)} - 2^{R_\alpha(\text{ref})}, \quad (16.3.9)$$

where $\text{ref}[n]$ is a reference stationary cosine signal (see Section 7.7).

- 6. (t, f) image-based features** are image descriptors extracted from the TFD matrix ρ_{z_x} considered as an image. First a segmentation technique, e.g., water-shed, is applied on ρ_{z_x} to detect regions where most signal information appears, and a binary-segmented image $\rho_{z_x}\text{seg}[n, k]$ is generated [10]. Morphometric features are then extracted from the *moments* of $\rho_{z_x}\text{seg}[n, k]$, defined as

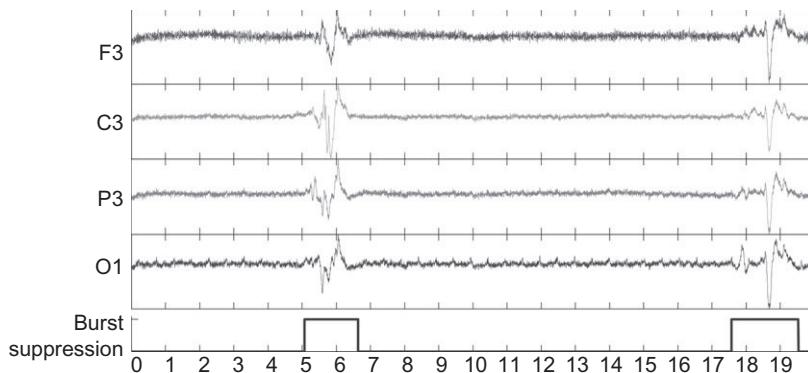
$$m_{pq} = \sum_{n=1}^N \sum_{k=1}^M n^p k^q \rho_{z_x}\text{seg}[n, k], \quad p, q = 0, 1, 2, \dots$$

Such features include [1,10]:

- (a) the (t, f) area: $A(x) = m_{00}$;
- (b) the (t, f) compactness: $C(x) = P^2/A(x)$ where $P = (m_{30} + m_{12})^2 + (m_{03} + m_{21})^2$;
- (c) the (t, f) coordinates of the centroid for the segmented region: $\text{CC}_X(x) = m_{10}/m_{00}$ and $\text{CC}_Y(x) = m_{01}/m_{00}$;
- (d) the (t, f) rectangularity: $R(x) = (m_{20} - m_{02})^2 + 4m_{11}^2$;
- (e) the (t, f) aspect ratio: $\text{AR}(x) = m_{20} - m_{02}$.

16.3.3 ILLUSTRATION ON REAL NEWBORN EEG DATA

As an illustrative application, let us consider detection of burst-suppression (B-S) patterns in newborn EEG signals. This discontinuity-based abnormality is characterized by a burst of high-voltage activity, followed by periods of marked attenuation, i.e., $< 5 \mu\text{V}$ (Fig. 16.3.1). The presence of B-S patterns in newborn EEG is proven to be associated with a high probability of poor neurodevelopmental outcomes [20].

**FIGURE 16.3.1**

EEG of a newborn showing burst-suppression patterns. *The waveform at the bottom shows the binary mask prepared based on the clinical observation of a pediatric neurologist.*

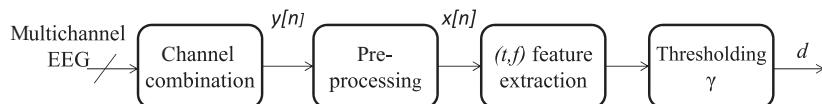
16.3.3.1 Burst-Suppression Detection in Newborn EEG

Detecting B-S patterns in EEG signals by visual inspection is laborious and time-consuming, and requires skilled interpreters who could be prone to subjective judgment and error. Automated detection of such patterns has been achieved using time-domain and spectral features extracted from EEG signals. Here, we evaluate the performance of an arbitrary subset of the (t, f) features presented in [Section 16.3.2](#) in detecting B-S patterns in continuous multichannel newborn EEG signals. The subset includes nine features F_i , $i = 1, 2, \dots, 9$, as follows: mean(IF), std(IA), CM(x), std(S_{hr}), mean(S_{wr}), $M_2(x)$, $R_3(x)$, $A(x)$, and AR(x), respectively.

For this purpose, a database was built, composed of continuous multichannel EEG signals with 20 channels collected from three newborns admitted to the Royal Brisbane and Women's Hospital, Brisbane, Australia. The signals were recorded using bipolar montage, according to the 10-20 standard, at sampling rate $f_s=256$ Hz. B-S patterns were marked in the acquired EEG signals by a pediatric neurologist.

16.3.3.2 Methodology

[Figure 16.3.2](#) shows the methodology used for labeling an epoch of multichannel EEG, i.e., $\mathbf{eeg}[n] = (\text{eeg}_1[n] \text{ eeg}_2[n] \dots \text{ eeg}_{20}[n])$, of length $T = Nf_s = 2$ s. We use 115 burst epochs and 115 suppression epochs extracted from artifact-free EEG signals.

**FIGURE 16.3.2**

The methodology used for classifying an epoch of multichannel EEG as burst or suppression using (t, f) features.

The signals are first combined using

$$y[n] = \sqrt{\frac{1}{20} \sum_{i=1}^{20} \text{eeg}_i^2[n]}, \quad (16.3.10)$$

to improve their SNR and enhance the signature of the abnormality in the (t,f) domain. The signal $y[n]$ is then bandpass-filtered in the (0.5–16) Hz band and down-sampled at 64 Hz (to reduce computational load). The resulting signal $x[n]$ is transformed to the (t,f) domain and its TFD, i.e., $\rho_{z_x}[n, k]$, is estimated. From $\rho_{z_x}[n, k]$, the nine (t,f) -based features F_i , $i = 1, \dots, 9$, are extracted. The values of a particular feature are compared for all segments with a predefined threshold γ to find a binary decision signal, which is compared with the mask provided by an expert neurologist; and the sensitivity and specificity of the feature are calculated. This procedure is repeated for different values of γ to find the receiver operating characteristic (ROC) curve for the feature. The area under the curve (AUC) is calculated as the performance criterion (see the pseudocode in the appendix to this section). Note that a classifier with AUC=1 is perfect, whereas a classifier with AUC=0.5 is as good as a random guess.

16.3.3.3 Results and Discussions

The methodology presented in the previous section was applied to the features extracted from several TFDs. They include the WVD, smoothed-WVD (SWVD) with a Hamming window $N/2$ samples long, spectrogram (SPEC) with a Hamming window $N/2$ samples long, modified B-distribution (MBD) with $\beta = 0.01$, and extended MBD (EMBD) with $\alpha = 0.9$ and $\beta = 0.01$ (see Table 6.1.2). The IF of $x[n]$ was estimated using the first-order moment of its TFD. Also, the NMF approximation described in Ref. [21] with parameter $R = 16$ was used. Before implementing the NMF algorithm, a zero was assigned to the negative values in $\rho_{z_x}[n, k]$. Table 16.3.1 shows the results of performance analyses of the subset of the features presented in Section 16.3.2 extracted from different TFDs.

The results show that, among the selected features for detecting B-S patterns in newborn EEG, features F_2 , F_5 , and F_7 (i.e., std(IA), mean(S_{w_r}), and $R_3(x)$, respectively) outperform the others. The AUC scores in Table 16.3.1 also imply that the performance of a (t,f) feature in detecting B-S patterns in newborn EEG depends on the type of TFD used for representing the signal $x[n]$ in the (t,f) domain.

Table 16.3.1: The Results of the ROC Analysis of a Subset of the (t,f) Features Presented in Section 16.3.2 Extracted from WVD and MBD in Detecting B-S Patterns in Multichannel Newborn EEG

| TFD | (t,f) Features | | | | | | | | |
|------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| | F_1 | F_2 | F_3 | F_4 | F_5 | F_6 | F_7 | F_8 | F_9 |
| WVD | 0.55 | 0.96 | 0.89 | 0.71 | 0.63 | 0.66 | 0.98 | 0.99 | 0.72 |
| SWVD | 0.58 | 0.96 | 0.61 | 0.68 | 0.91 | 0.66 | 0.99 | 0.79 | 0.68 |
| SPEC | 0.67 | 0.96 | 0.51 | 0.64 | 0.96 | 0.62 | 0.99 | 0.65 | 0.87 |
| MBD | 0.58 | 0.96 | 0.53 | 0.85 | 0.86 | 0.55 | 0.99 | 0.79 | 0.76 |
| EMBD | 0.88 | 0.96 | 0.92 | 0.50 | 0.99 | 0.78 | 0.99 | 0.99 | 0.53 |

This methodology was also applied to the detection of another newborn EEG abnormality, i.e., the seizure, as reported in Ref. [1].

16.3.4 SUMMARY AND CONCLUSIONS

This section shows that (t,f) -based features exploit the additional information provided by the signal variations in terms of nonstationarities observed in the (t,f) domain and therefore can lead to better classification of nonstationary signals. Other (t,f) -based features include those presented in Section 12.6. The choice of features is application-dependent, and a thorough data analysis is needed to choose the ones that best represent different classes in the signal under analysis.

For a particular application, the performance of a (t,f) feature depends on the type of TFD used to represent the signal in the (t,f) domain, and it can be improved by optimizing the parameters of the TFD kernel. The computational load of a (t,f) -based classifier can be significantly reduced by using more computationally efficient algorithms for implementing TFDs, such as the one presented in Ref. [1]. The approach presented here for detecting B-S patterns in newborn EEG can be customized and used for the detection and classification of other newborn EEG abnormalities (e.g., seizures) and artifacts (see [1] and the other sections in this chapter).

16.3.5 APPENDIX: PSEUDOCODE FOR THE IMPLEMENTED ALGORITHM

```
% MATLAB pseudo-code used to find the AUC for a given feature
% initialization
fs = 256; % sampling frequency
Nseg = 230; % total number of segments
for n = 1:Nseg
    % combine the channels
    y = sqrt(mean(eeg.^2,1)); % eeg is a multichannel segment
    % filter the signal in [0.5 16] Hz band
    b = fir1(100,[0.5/(fs/2) 16/(fs/2)]);
    y = filtfilt(b,1,y);
    % down sample the signals @ 64 Hz
    x = resample(y,64,fs);
    % calculate the TFD of x using quadtfdf function
    % find the value of the (t,f) feature, i.e., F(n)
end % next segment
% ROC analysis
i = 0;
for γ = γmin : γmax % γ is the threshold
    i = i+1;
    d = F<=γ;
    % ref_mask is the binary mask provided by the neurologist
    TP(i) = sum(ref_mask.*d); % True Positive Rate
    FP(i) = sum((1-ref_mask).*d); % False Positive Rate
    TN(i) = sum((1-ref_mask).*(1-d)); % True Negative Rate
```

```

FN(i) = sum(ref_mask.*(1-d)); % False Negative Rate
Sen(i) = TP(i)/(TP(i)+FN(i)); % Sensitivity
Spe(i) = TN(i)/(TN(i)+FP(i)); % Specificity
end % next γ
% calculate the AUC
auc = trapz(1-Spe,Sen);

```

16.4 TIME-VARYING ANALYSIS OF BRAIN NETWORKS⁰

Typical neurophysiological recordings of brain activity, such as electroencephalography (EEG) or magnetoencephalography (MEG), are nonstationary signals. Time-frequency tools are therefore suitable for analysis of such signals. Over the years, time-frequency (t,f) analysis has found numerous applications in computational neuroscience from the diagnosis of schizophrenia in adults [22] to understanding autism-spectrum disorders in a pediatric population [23] or seizure detection in newborns [24–26]. With advances in computational resources and various imaging modalities, (t,f) analysis is poised to become one of the main tools in the area of computational neuroscience.

Time-frequency analysis is especially critical in the analysis of brain networks, which represent the human brain as a complex network of nodes and edges. Nodes can represent individual neurons or groups of neurons, or groups of *voxels* (volume pixels) obtained from magnetic-resonance imaging (MRI), while edges correspond to the relationships between these regions [27]. Applications that emphasize the fact that the properties of these networks can be used as diagnostic biomarkers for various diseases include Alzheimer's [28], stroke [29], and other brain disorders [30].

16.4.1 BRAIN NETWORKS

Brain networks typically denote networks of interconnected distinct units in the brain [27]. The units can represent individual neurons or anatomically distinct brain regions. In general, the networks may be classified as structural, functional or effective. *Structural* networks describe physical connections between nodes (e.g., neurons). *Functional* networks describe statistical dependence between nodes, regardless of whether the nodes are physically connected; the statistical dependence can be quantified using various approaches, from simple correlation to more advanced (t,f)-based approaches. *Effective* networks usually correspond to direct or indirect networks in which the relationships are determined through causality [27].

Brain networks are usually represented as graphs, which in turn can be represented as matrices, in which the element m_{ij} represents the connection between the i th and j th nodes. To form these matrices, usually called *connectivity matrices*, we begin with neuroimaging data obtained via, for example, MRI, EEG, MEG, or near-infrared spectroscopy. The imaging procedure determines whether the connectivity is structural or functional. When connectivity matrices have been established, graph-theoretical tools are used to understand the network properties [27].

⁰Author: E. Sejdic, University of Pittsburgh, Pittsburgh, PA, USA (esejdic@ieee.org), S. Aviyente, Michigan State University, East Lansing, MI, USA (avyente@egr.msu.edu), and B. Boashash, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa). Reviewers: G. Azemi and A. Omidvarnia.

16.4.2 TIME-FREQUENCY ANALYSIS FOR BRAIN NETWORKS

The brain and its various functions represent a nonstationary system. The strength and the number of interconnections between brain regions continuously change, enabling a number of uses of (t,f) analysis [31]. Here, we outline the main three cases.

16.4.2.1 Establishing Connectivity Matrices via Time-Frequency Analysis

Cross-correlation is typically utilized to assess the relationship between different brain regions; however, a simple correlation coefficient at the zeroth lag is limited in its ability to assess the interactions of these signals, as it can only quantify the linear relationships and capture mostly the amplitude-based relationships. Time-frequency tools become important in this case, and publications typically examine phase synchrony among neuroimaging signals representing different nodes. Synchrony measurements relate two signals' temporal structures without considering their amplitude. Hence, any two signals are denoted synchronous as long as their rhythms (i.e., temporal structures) are similar. The degree of synchrony between two signals is usually assessed via estimation of instantaneous phase around a particular frequency, which is usually accomplished via the Hilbert transform or (t,f) methods. Given that neuroimaging signals are predominantly nonstationary, (t,f) methods are more suitable for the calculation of synchrony between two signals [32]. Traditionally, a Morlet wavelet-based method was used [33], but here we present a recently proposed method based on a quadratic (t,f) representation [32].

Let us consider a (t,f) representation of a signal based on the reduced-interference Rihaczek (t,f) distribution (RID-Rihaczek), defined as

$$\rho_z(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-v^2\tau^2/\sigma} e^{-j\pi v\tau} A_z(v,\tau) e^{-j2\pi(vt-f\tau)} dv d\tau, \quad (16.4.1)$$

where $A_z(v,\tau)$ is the ambiguity function as defined in previous chapters (e.g., Eq. (3.2.4), p. 111), $e^{-j\pi v\tau}$ is the kernel for the Rihaczek distribution, and $e^{-v^2\tau^2/\sigma}$ is the exponential kernel used to reduce the effect of the cross-terms. Other kernels can be used as long as they remove the cross-terms in the Rihaczek amplitude spectrum. The next step is to estimate the time-varying phase in the (t,f) plane. This is accomplished via the equation

$$\Phi(t,f) = \arg \{ \rho_z(t,f) \}, \quad (16.4.2)$$

where we use the convention that $\arg\{w\} = 0$ if $w = 0$. Using the definition of the time-varying phase spectrum, the phase difference between two signals, $z_1(t)$ and $z_2(t)$, can be similarly defined as

$$\Phi_{12}(t,f) = \arg \{ \rho_{z_1}(t,f) \rho_{z_2}^*(t,f) \}. \quad (16.4.3)$$

Again, if the expression in braces is zero, the “arg” is taken as zero.

Based on the estimate of a time-varying phase spectrum as shown above, a synchrony measure still needs to be defined. In this section, the phase-locking value (PLV) is considered. The PLV between two signals, averaged across realizations/trials, can be defined as [32]

$$\text{PLV}(t,f) = \frac{1}{\text{TR}} \left| \sum_{k=1}^{\text{TR}} \exp \left[j\Phi_{12}^k(t,f) \right] \right|, \quad (16.4.4)$$

where TR is the number of trials/realizations, and $\Phi_{12}^k(t, f)$ is the time-varying phase estimate between two electrodes/voxels for the k th trial. The PLV measures the intertrial/interrealization variations of phase differences at time (t) and frequency (f). A PLV close to 1 indicates small phase difference across trials/realizations. It should be noted that PLV is applied to neuroimaging signals between pairs of channels/voxels; i.e., it is a bivariate measure, not a multivariate measure. Most neuroimaging experiments have multiple trials, but for a single trial, a so-called single-trial PLV is calculated, denoting the consistency of the phase across time. Lastly, the described phase synchrony measure assesses the instantaneous phase differences between signals in the (t, f) domain [34].

16.4.2.2 Understanding the Time-Varying Nature of Brain Networks

The next challenge in the analysis of brain networks lies in the analysis of connectivity matrices, which change over time. This is expected in many cases, especially when a person is performing repeated operations over a certain time span. A second possibility is the case when it is expected that activated brain nodes will change over time due to neurological disease or aging; however, to utilize (t, f) analysis for the examination of the time-varying nature of these brain networks, one needs to use spectral graph theory.

Let us consider undirected weighted graphs $\mathfrak{G} = \{\nu, \xi, W\}$, where ν is a finite set of vertices, ξ is a set of edges, and W is a weighted connectivity matrix. Let N be the cardinality of ν . A signal $z : \nu \rightarrow \mathbb{R}^N$ defined on the vertices of the graph may be represented as a vector $z \in \mathbb{R}^N$, whose n th component represents the signal value at the n th vertex in ν . Briefly, the classical Fourier transform (FT) expands a signal in terms of complex exponentials that can be considered to be eigenfunctions of the one-dimensional Laplace operator: $-\Delta(\exp(j2\pi ft)) = (2\pi f)^2 \exp(j2\pi ft)$. Using the same analogy, the graph FT of a function $z(n)$ can be defined in terms of the eigenvectors $\chi_l(n)$ of the connectivity matrix Laplacian [35]

$$Z(\lambda_l) := \langle z, \chi_l \rangle = \sum_{n=1}^N z(n) \chi_l^*(n), \quad (16.4.5)$$

with the inverse graph FT being equal to

$$z(n) := \sum_{l=1}^{N-1} Z(\lambda_l) \chi_l(i). \quad (16.4.6)$$

The graph Laplacian eigenvalues and eigenvectors can be interpreted as follows: the Laplacian eigenvector χ_0 is constant and equal to $1/\sqrt{L}$ at each vertex. The graph Laplacian eigenvectors associated with low eigenvalues λ_l vary slowly across the graph; i.e., if two vertices are connected by an edge with a large weight, the values of the eigenvector at those locations are likely to be similar. The eigenvectors associated with larger eigenvalues oscillate more rapidly and are more likely to have dissimilar values on vertices connected by an edge with high weight [35].

Using the definition of the graph FT, the windowed graph FT of a function $z(n) \in \mathbb{R}$ can be given as [36]

$$F_z^g(n, k) = \langle z, g_{i,k}(n) \rangle, \quad (16.4.7)$$

where a windowed graph Fourier atom $g_{i,k}(n)$ is given as

$$g_{i,k}(n) = N \chi_k(n) \sum_{l=0}^{N-1} \hat{g}(\lambda_l) \chi_l^*(i) \chi_l(n), \quad (16.4.8)$$

where $\hat{g}(\lambda)$ is a window definition in the graph spectral domain.

16.4.2.3 Time-Frequency Features of Time-Varying Networks

Time-varying networks can be analyzed using vertex-frequency analysis, the graph analog of (t,f) analysis of one- or two-dimensional signals. As we obtain a two-dimensional representation similar to a (t,f) representation, we can utilize some of the tools developed for analysis of (t,f) representations. In particular, we can calculate (t,f) features, such as [1],

- the (t,f) energy concentration (TFEC)

$$\text{TFEC} = \left(\sum_{n=1}^N \sum_{k=1}^M |S_x^g(n, k)|^{1/2} \right)^2, \quad (16.4.9)$$

- the (t,f) Rényi entropy (TFRE)

$$\text{TFRE} = \frac{1}{1-\alpha} \log_2 \left[\sum_{n=1}^N \sum_{k=1}^M \left(\frac{S_x^g(n, k)}{\sum_{n=1}^N \sum_{k=1}^M S_x^g(n, k)} \right)^\alpha \right], \quad (16.4.10)$$

- the (t,f) (or time-varying) spectral flatness (TFSF)

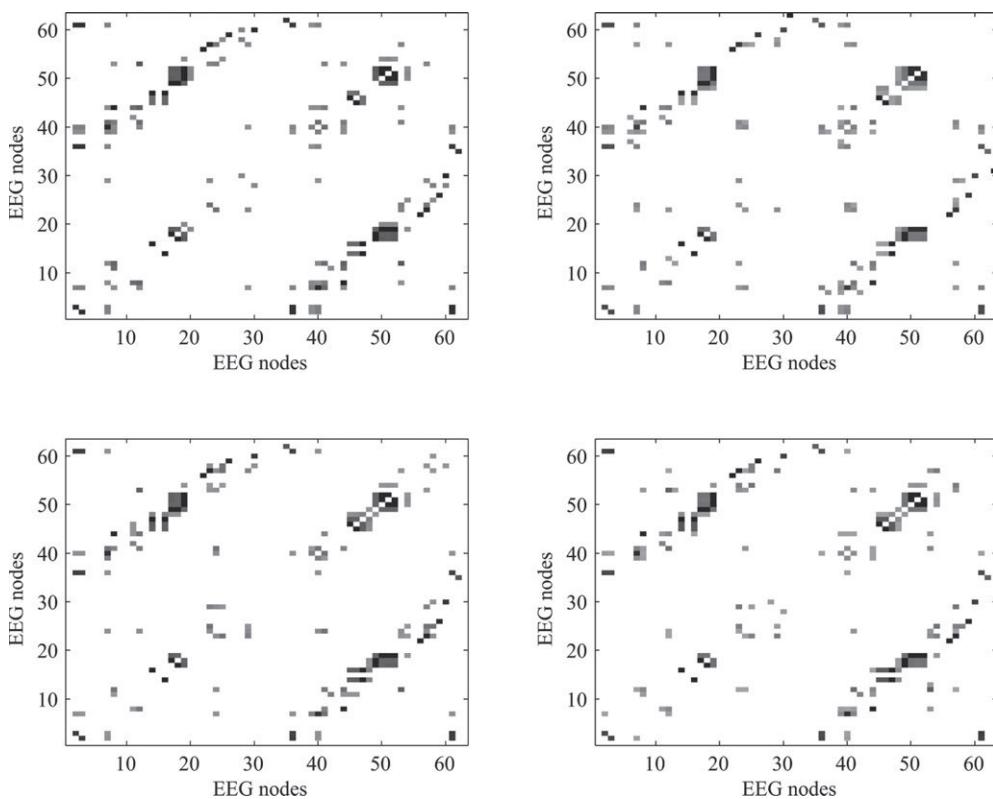
$$\text{TFSF} = -\frac{1}{NM} \sum_{n=1}^N \sum_{k=1}^M S_x^g(n, k) \log_2 S_x^g(n, k), \quad (16.4.11)$$

where $S_x^g(n, k) = |F_x^g(n, k)|^2$ is the spectrogram of the considered signal, and $\alpha = 3$ is a typical value for TFRE. There are many more (t,f) features that can be utilized in the analysis (see Section 12.6). Furthermore, if a (t,f) distribution can be considered as an image (e.g., if the distribution is nonnegative), then features used in image processing can be used here as well.

16.4.3 ILLUSTRATIVE EXAMPLE

As an illustration of the proposed approach, let us consider a sample EEG recording of a healthy adult performing four saliva swallows in the head-neutral position. The sample signals were collected from 64 electrodes positioned according to the 10-20 international electrode system. Positioning was accomplished using the “actiCAP” active electrodes (Brain Products GmbH, Germany), and the signals were amplified using the “actiCHamp” amplifier (Brain Products). The P1 electrode was chosen as the reference; i.e., EEG potentials are relative to P1. During all data collections, the electrode impedance was below 15 kΩ. The “PyCorder” acquisition software (Brain Products) provided a 10 kHz sampling frequency, and it was also used for saving collected data on a computer hard drive.

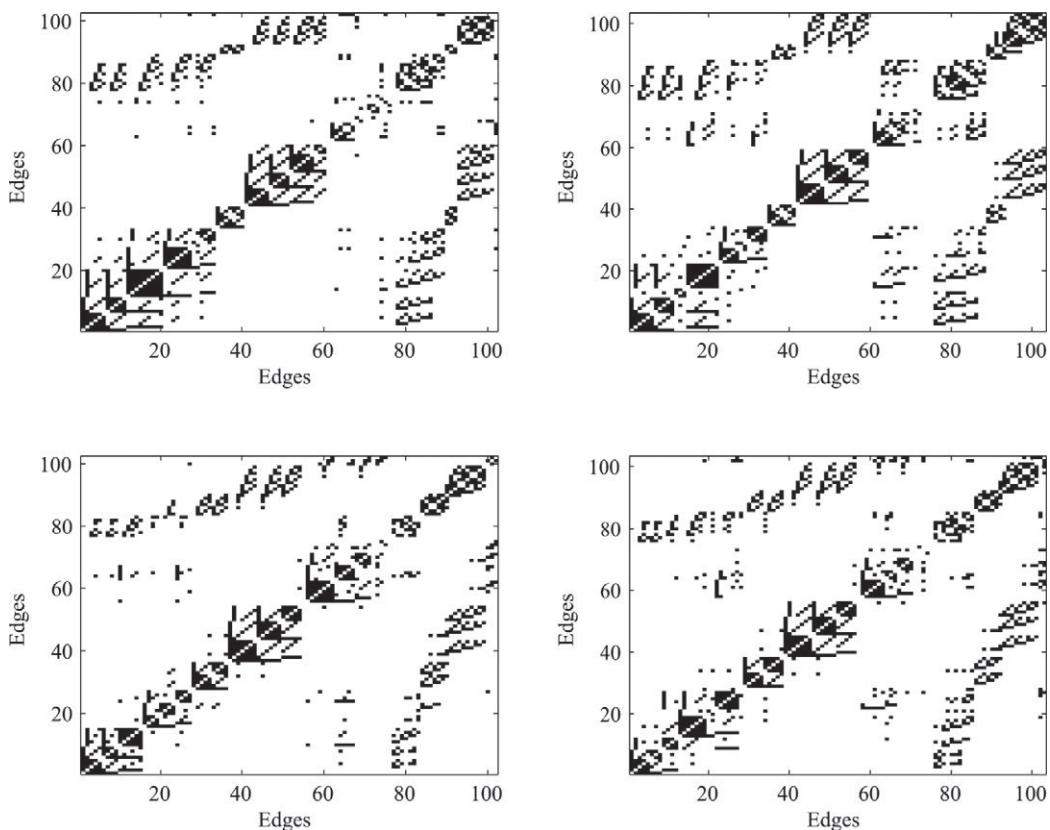
To establish brain connectivity matrices, we utilized Eqs. (16.4.3) and (16.4.4) to calculate bivariate phase differences and corresponding PLV values for all EEG channels. The computed connectivity matrices are not sparse, as many spurious connections also appear. Therefore, to concentrate on the

**FIGURE 16.4.1**

Brain networks for the sample of four swallows.

most relevant connections, we thresholded connectivity matrices to keep only the strongest 5% of connections. The established brain connectivity matrices for four swallows are shown in Fig. 16.4.1. Even though the swallows are from the same healthy participant, differences in the established networks can be observed. It is obvious that even though most of the functional connections are present for all swallows, there are swallow-by-swallow differences even in healthy individuals, which need to be further quantified. While typical network metrics such as clustering coefficients or global efficiency measures can be calculated, those metrics do not reveal any information about the time-varying nature of these brain networks; even though the values of these network *metrics* can be different on a swallow-by-swallow basis, we do not have any details on how the networks themselves have changed. Therefore, it would be beneficial to apply the vertex-frequency analysis in order to understand the time-varying changes in the swallowing networks.

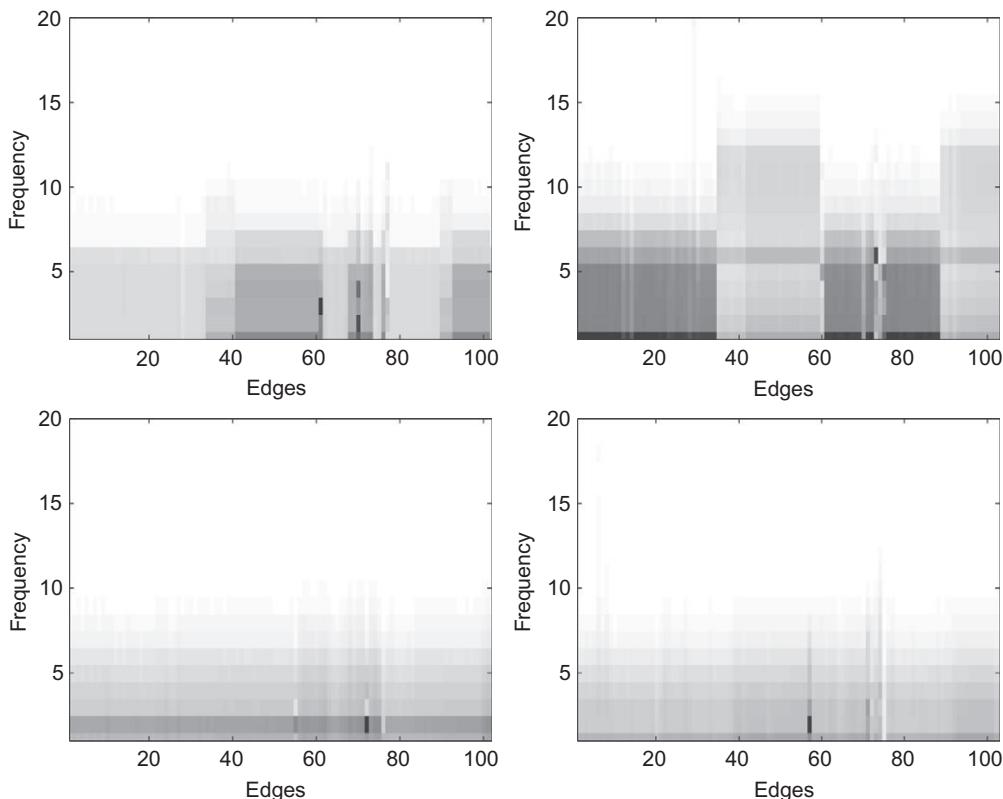
For the vertex-frequency analysis to be applicable to these networks, time-varying data points need to be on *vertices*. Hence, as the changes are in the *edges* (connections), one needs to convert the network representations into line graphs by converting edges to vertices that are connected by unity edges. Such

**FIGURE 16.4.2**

Line graphs of brain networks for the four swallows.

conversions yield edge-connectivity matrices as shown in Fig. 16.4.2. These edge-edge connectivity matrices are then used to calculate edge-frequency representations via Eqs. (16.4.7) and (16.4.8). These representations, as shown in Fig. 16.4.3, depict spectrograms of the brain networks shown in Fig. 16.4.1. The obtained representations clearly depict the time-varying nature of the swallowing networks, as the frequency content of the networks changes on a swallow-by-swallow basis.

Next, we would like to quantify these changes by examining the three (t,f) features outlined in the previous section. Our results for the four sample swallows, calculated from the network spectrograms, are summarized in Table 16.4.1. As these features clearly capture the differences in the edge-frequency representations, we can state that features based on (t,f) representations (or, more specifically, edge-frequency representations) are suitable for understanding differences in brain networks during the four swallowing actions. The presented results are just sample results that demonstrate the applicability of typical (t,f) measures in the network analysis. Future works should investigate the relationships between typical network metrics and (t,f) features. Those investigations should provide us with more understanding about the utility of the (t,f) network analysis.

**FIGURE 16.4.3**

Edge-frequency representations of the four swallows.

Table 16.4.1: Values of the Three (t,f) Features

| | SW1 | SW2 | SW3 | SW4 |
|------|-------|-------|-------|-------|
| TFEC | 0.562 | 0.754 | 0.396 | 0.494 |
| TFRE | 8.958 | 9.394 | 8.886 | 8.924 |
| TFSF | 7274 | 5206 | 3532 | 4612 |

16.4.4 SUMMARY AND CONCLUSIONS

The applications of (t,f) analysis in modern computational neuroscience include the establishment and postanalysis of brain networks. A network connection matrix can be established using (t,f) measures of synchrony between the signals at the nodes, and analyzed using features derived from (t,f) theory. The illustrative study of a healthy human subject performing the same action several times shows that changes in features of edge-frequency representations can reveal changes in brain networks.

16.5 TIME-FREQUENCY ANALYSIS AND EEG NOISE REDUCTION USING EMPIRICAL MODE DECOMPOSITION⁰

16.5.1 EMPIRICAL MODE DECOMPOSITION

Empirical-mode decomposition (EMD) is a fully adaptive and data-driven signal analysis method for multicomponent signals that assumes no *a priori* information about the linearity or stationarity of the signal [37]. It decomposes the signal into a set of intrinsic mode functions (IMFs) such that their sum is equal to the original signal. IMFs must satisfy two conditions: (i) the number of extrema and the number of zero crossings should be equal or differ by one at most; and (ii) the mean envelope determined by the upper and lower envelopes should be zero. The crucial part of an EMD algorithm is the *sifting* process, which is summarized as follows [38]:

- (a) compute the signal upper envelope $M(t)$, and the lower envelope $m(t)$, by interpolating local maxima and minima using cubic spline interpolation;
- (b) compute the local mean $\mu(t) = [M(t) + m(t)] / 2$ of the signal $x(t)$;
- (c) find the detail by subtracting $\mu(t)$ from $x(t)$ as $h(t) = x(t) - \mu(t)$;
- (d) if $h(t)$ satisfies the aforementioned IMF conditions, keep it as an IMF $\varphi(t)$, and allocate the new value of $x(t)$ as $x(t) - h(t)$;
- (e) repeat from (a) until $x(t)$ becomes a constant or monotonic “residue”, say $r(t)$.

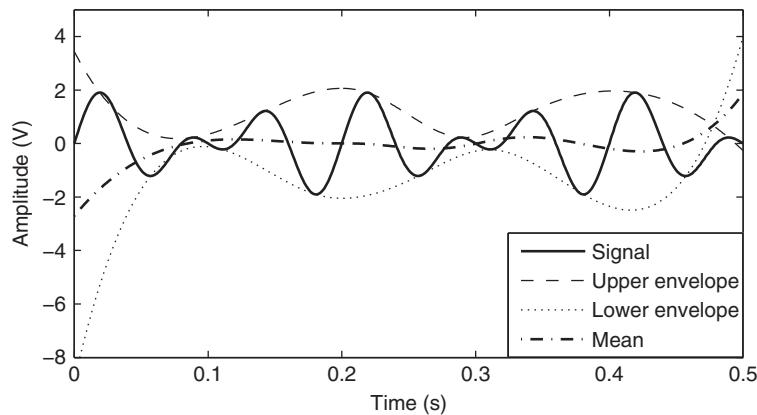
Once the decomposition is finalized, the original signal $x(t)$ is represented as

$$x(t) = \sum_{i=1}^L \varphi_i(t) + r(t), \quad (16.5.1)$$

where L is the total number of extracted IMFs. The steps of one iteration of the sifting process are shown graphically in Fig. 16.5.1.

If the mean signal in Fig. 16.5.1 satisfies the IMF conditions, it is used as an IMF, and subtracted from the signal. Otherwise, it is only removed from the signal and the iterations continue until the residue is extracted. Since the residue is a monotonic function, no more IMFs can be decomposed. Hence, the decomposition of $x(t)$ is obtained as L components $\varphi_i(t)$, and a single residue $r(t)$. The zero mean condition of the IMFs is not valid for the residue, which may be an alternative approach to analyzing direct current (DC) components in the signal. From this point of view, each component of a multicomponent signal may be estimated and analyzed separately to provide richer information without selection of any basis function (mother wavelet, harmonics, etc.). Thus EMD may be used to characterize the local (t, f) content of any signal and for IF estimation, classification, and denoising [39–41]. Shortcomings of the EMD algorithm include end effects (divergence of the envelopes; see Fig. 16.5.1) and mode mixing. Attempted improvements focus on the interpolation and the stopping criterion [37,38]. The original standard-deviation-based stopping criterion, which aims to ensure that IMFs retain sufficient physical sense of amplitude and frequency modulation, is defined by [37]

⁰Authors: A. Akan, Istanbul University, Avclar, Istanbul, Turkey (akan@istanbul.edu.tr) and A. Mert, Piri Reis University, Tuzla, Istanbul, Turkey (amert@pirireis.edu.tr). Reviewers: N. Stevenson, E. Sejdic, and I. Shafii.

**FIGURE 16.5.1**

Graphical representation of *sifting* process.

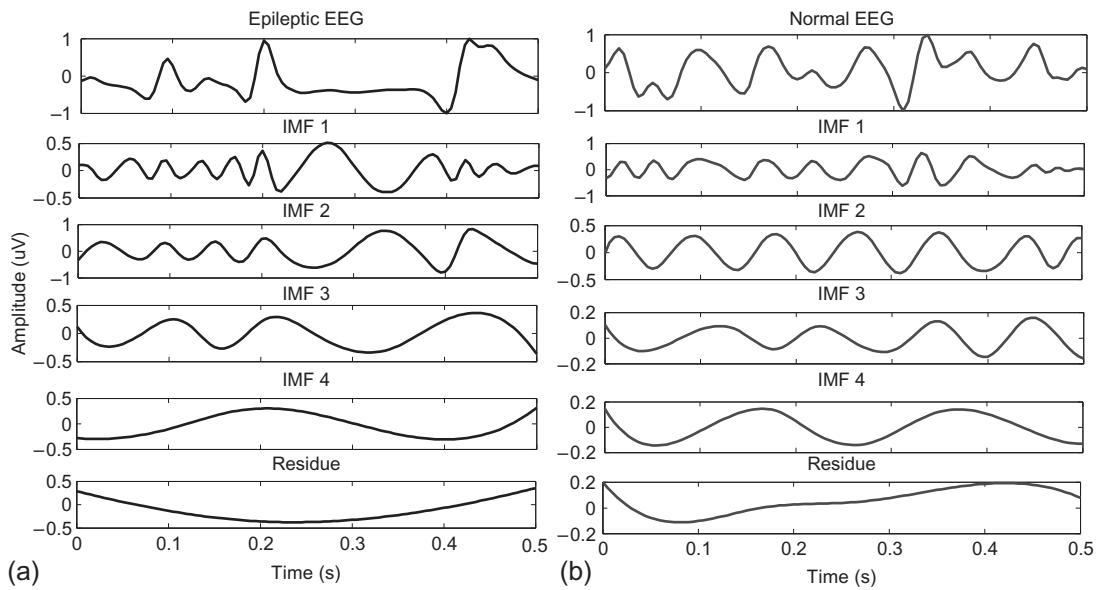
$$\text{SD} = \sum_{t=0}^T \left[\frac{|h_{(k-1)}(t) - h_k(t)|^2}{h_{(k-1)}^2(t)} \right], \quad (16.5.2)$$

where SD is the standard deviation and k is the iteration number of the sifting process. SD may be chosen between 0.2 and 0.3 for meaningful results [37]. Another EMD algorithm [38] uses an alternative stopping criterion by predefined resolutions

$$q_{\text{resol}} = 10 \log \left(\sigma_{x(t)}^2 / \sigma_{h(t)}^2 \right) \quad \text{and} \quad q_{\text{resid}} = 10 \log \left(\sigma_{x(t)}^2 / \sigma_{r(t)}^2 \right), \quad (16.5.3)$$

where q_{resol} and q_{resid} refer to the ratios of signal energy to IMF and residue energy, respectively. By selecting these parameters, an oscillation may be assigned as an IMF, or the residue. Studies [42] to enhance the sifting algorithm focus on properly decomposing a multicomponent signal into mono-component oscillations. This is called the “mode-mixing” problem, because it is complicated by oscillations of one component interfering with other IMFs. In some cases, the EMD algorithm cannot extract a component but distributes its oscillations to other IMFs. This problem generally arises when signal components have (i) disparate amplitudes at *some* times, and (ii) close statistical properties. Under these conditions, an envelope estimation based on the detection of extrema of the discretized signal may lead to “mismatched points”—one detected extremum may be due to one component, and the next detected extremum may be due to another component. Hence, an IMF may contain distributed parts of other components, obtained by detecting extrema of different components. To combat the mode-mixing problem of EMD, improved envelope estimation approaches have been presented using advanced interpolation algorithms [43].

To illustrate the EMD algorithm on real signals, a set of IMFs extracted from single-channel epileptic and normal electroencephalography (EEG) recordings are given in Fig. 16.5.2. The publicly available recordings (<http://eeganalysis.web.auth.gr/dataen.htm#EpilepticEEG>) with 200 Hz sampling are filtered by a lowpass filter ($f_c = 40$ Hz) and normalized to the range ± 1 . The first IMFs seem

**FIGURE 16.5.2**

IMFs of epileptic (a) and normal (b) EEG.

random and nonstationary, while the others are similar to a sinusoidal waveform. The time-frequency (t, f) characteristics of epileptic and normal EEG IMFs are different, justifying the EMD approach.

16.5.2 HILBERT-HUANG TRANSFORM

The basic idea of the Hilbert-Huang transform (HHT) is to combine EMD and the Hilbert transform (HT) [37]. This method suggests an alternative approach to analyzing a multicomponent signal via the instantaneous frequency (IF) and instantaneous amplitude (IA) domains (Section 1.3.1) after decomposing it into mono-component IMFs. Then each IMF, $\varphi_A(t)$, is computed as

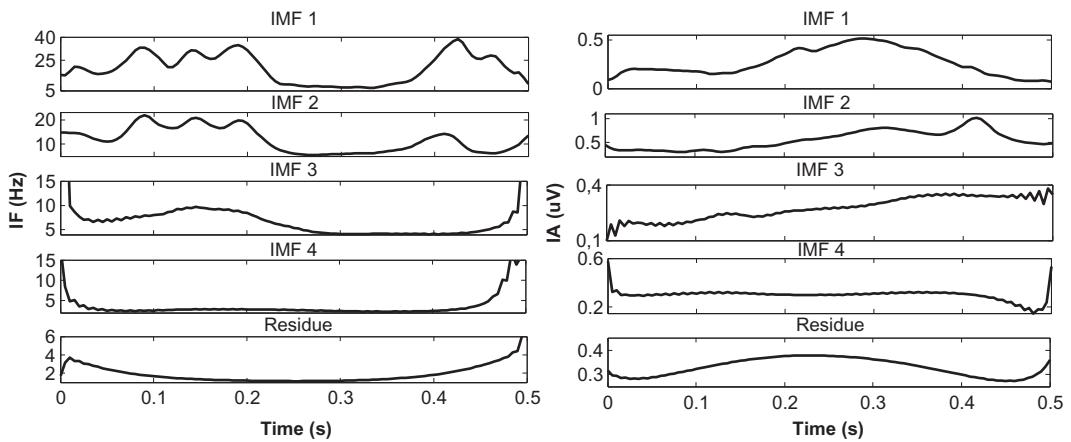
$$\varphi_{A_i}(t) = \varphi_i(t) + j\varphi_{H_i}(t) = a_i(t)e^{j\phi_i(t)}, \quad i = 1, 2, \dots, L, \quad (16.5.4)$$

where $\varphi_{H_i}(t)$ is the HT of $\varphi_i(t)$, defined as

$$\varphi_{H_i}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\varphi_i(\tau)}{t - \tau} d\tau, \quad (16.5.5)$$

where the integral is to be understood as the Cauchy principal value (Eq. (1.2.16)). The IA, $a_i(t)$, and the instantaneous phase, $\phi_i(t)$, are given by

$$a_i(t) = \sqrt{\varphi_i^2(t) + \varphi_{H_i}^2(t)} ; \quad \phi_i(t) = \arctan \left[\varphi_{H_i}(t) / \varphi_i(t) \right]. \quad (16.5.6)$$

**FIGURE 16.5.3**

IF and IA of the epileptic EEG signal.

After unwrapping $\phi_i(t)$, the IF of the IMF is defined as

$$f_i(t) = \frac{1}{2\pi} \phi'_i(t). \quad (16.5.7)$$

To demonstrate the Hilbert-Huang spectrum (HHS), we apply the HHT to the epileptic EEG signal in Fig. 16.5.2(a), and show its IA and IF spectra in Fig. 16.5.3. Notice in Fig. 16.5.3 that the IF fluctuations of the IMFs decrease from top to bottom, while IA envelopes do not have specific trends. Finally, the HHS of a multi-component signal provides more detailed information than a spectral analysis in distinguishing different oscillations contained in the signal.

16.5.3 EMD-BASED EEG DENOISING

Denoising means estimating a noise-free signal $\tilde{x}(t)$ from a noisy observation

$$x(t) = \tilde{x}(t) + \epsilon(t), \quad (16.5.8)$$

where the corrupting noise $\epsilon(t)$ has known or unknown variance $\sigma_{\epsilon(t)}^2$. The performance measure for the denoising process is the signal-to-noise ratio (SNR)

$$\text{SNR} = 10 \log \left(\sigma_{\tilde{x}(t)}^2 / \sigma_{\epsilon(t)}^2 \right). \quad (16.5.9)$$

Once the IMFs of $x(t)$ are obtained, some of them need to be detected as the components of the noise $\epsilon(t)$, and the remaining ones as components of the noise-free signal $\tilde{x}(t)$. Hence, any EMD-based denoising approach focuses on identifying the IMFs corresponding to noisy oscillations, and then estimating the original signal by combining the rest of the IMFs. From this point of view, a reliable and robust metric is required to test whether the IMFs belong to noise or not. Models based on clustering, white Gaussian noise (WGN) [44], and fractional Gaussian noise (fGn) [45] can be successfully applied

to determine the noisy oscillations; however, these are models referenced to the signal itself, and based on observing discrepancies of the IMF energies that are not applicable in practical situations.

The performance of the EMD-based denoising also depends on the ability of the sifting algorithm to solve the mode-mixing problem. If the EMD cannot properly decompose a multicomponent signal (or noisy signal $x(t)$) into its components, IMFs can contain noise and the estimation of the desired signal can be distorted. Hence, the excluded IMFs in the reconstruction will impair the denoising performance. In the following section, we focus on determining the purely noisy oscillations or IMFs that contain mostly noise, using reliable metrics.

16.5.3.1 HHT-Based Clustering For EEG Denoising

HHT-based clustering investigates the similarity of the IMFs in the IF domain. The corrupted signal is decomposed by the EMD, and then the HHS is computed and analyzed by forming a hierarchy of clusters [46]. Another method is to use the Euclidean distances between IF spectra of IMFs, defined as

$$d_{ij}(\text{IF}_i, \text{IF}_j) = \|\text{IF}_i(t) - \text{IF}_j(t)\| ; i = 1, 2, \dots, L-1; j = i+1, i+2, \dots, L \quad (16.5.10)$$

where $\|\cdot\|$ denotes the Euclidean norm. Hence, we define the similarity vector

$$D = [d_{12}, d_{13}, \dots, d_{1L}, d_{23}, d_{24}, \dots, d_{2L}, \dots, d_{(L-1)L}] . \quad (16.5.11)$$

The steps of the proposed method can be described as follows:

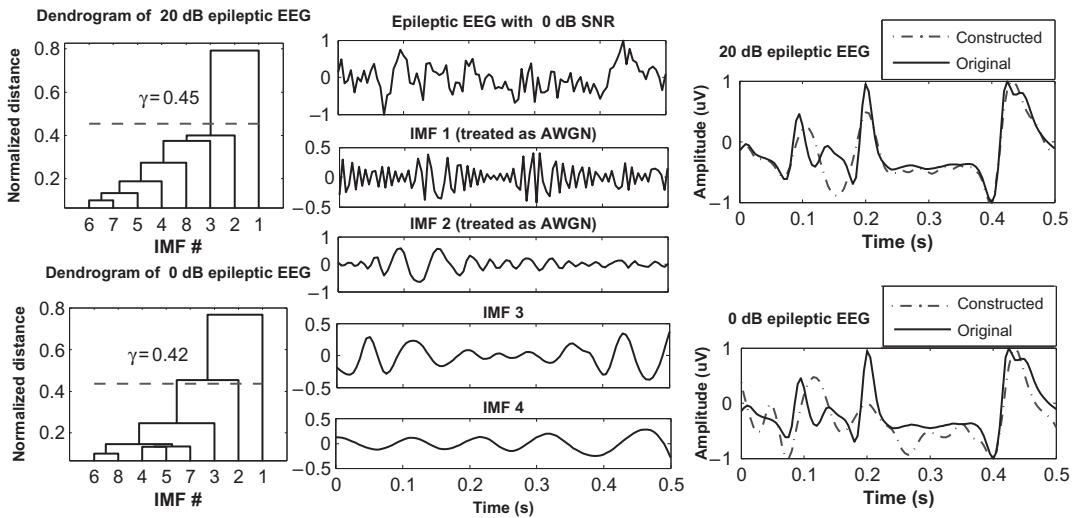
- (a) Let the noisy signal be $x(t) = \tilde{x}(t) + \epsilon(t)$, where $\tilde{x}(t)$ is the noise-free signal to be estimated, and $\epsilon(t)$ is additive white Gaussian noise (AWGN).
- (b) Decompose $x(t)$ into $\varphi_i(t)$, $i = 1, 2, \dots, L$, where L is the number of IMFs.
- (c) Compute the similarity vector D from the IF spectra $\text{IF}_i(t)$ of the IMFs calculated by the HHT, and then normalize them to the range $[0, 1]$.
- (d) Determine a threshold γ as the average component of the similarity vector D , i.e., $\gamma = \frac{1}{M} \sum_{i=1}^M D_i$, where $M = L(L-1)/2$ denotes the length of D .
- (e) After single-linkage, plot the dendrogram to observe the result. The estimated version of $\tilde{x}(t)$ is the sum of the IMFs with lower similarity than the threshold; i.e., $\tilde{x}(t) = \sum_{\{i:D_i < \gamma\}} \varphi_i(t)$.

Thus, irrelevant IMFs may be identified and excluded in the reconstruction phase.

Example 16.5.1. The HHT-based denoising is tested on the single-channel epileptic EEG signal shown in Fig. 16.5.2, with HHS shown in Fig. 16.5.3. Noisy signals with 20 and 0 dB AWGN are applied to the above algorithm, and the obtained results are shown in Fig. 16.5.4. The first IMF for 20 dB, and the first and the second IMFs for 0 dB cases are thresholded and treated as the components of the AWGN. Thus, denoised versions of the epileptic EEG are calculated as $\sum_{i=2}^8 \varphi_i$ and $\sum_{i=3}^8 \varphi_i$, resulting in SNRs of 22.62 and 7.58 dB, respectively. Note that for both cases, the number of IMFs in the decomposition is $L = 8$. The method is also tested on a normal EEG, and the SNRs are computed as 24.06 and 12.05 dB. As observed in the examples, white-noise-dominant IMFs are successfully identified; however, interference of noise into informative IMFs remains as an artifact of the EMD algorithm [43].

16.5.3.2 Detrended Fluctuation-Based IMF Thresholding

In the case of nonstationary signals, detrended fluctuation analysis (DFA) is considered to be more suitable and reliable than the “Hurst” analysis for investigating the long-range dependence and

**FIGURE 16.5.4**

Results of the proposed denoising applied to epileptic EEG with 0 and 20 dB noise.

autocorrelation. The advantages of DFA are that it permits the analysis of self-similarity embedded in a nonstationary time series, and it avoids spurious detection [47,48]. The principle of the DFA algorithm is to determine how the average root-mean square (RMS) fluctuation of the signal changes around the estimated local trend. The integrated time series $y(k)$ is computed as the first step of the DFA after removing the mean μ_x :

$$y(k) = \sum_{i=1}^k [x(i) - \mu_x]. \quad (16.5.12)$$

The sequence $y(k)$ is partitioned into n -sample-long segments, called boxes, for DFA. The local trend for a length- n box μ_{y_n} is estimated by least-squares linear fitting. Finally the RMS fluctuation F_n is computed by subtracting μ_{y_n} from the integrated series $y_n(k)$:

$$F_n = \left(\frac{1}{N} \sum_{k=1}^N [y_n(k) - \mu_{y_n}]^2 \right)^{1/2}. \quad (16.5.13)$$

F_n is calculated for $1 \leq n \leq N$. The slope α of the plot of $\log F_n$ vs. $\log n$ is called the “scaling exponent” and it presents a power-law behavior as $F_n \propto n^\alpha$. The main reason why α is reliable is that DFA yields fixed scores for uncorrelated white noise ($\alpha = 0.5$), pink noise ($\alpha = 1.0$) and Brownian noise ($\alpha = 1.5$); if $0.5 \leq \alpha \leq 1$, this is an indication of temporal correlations. In the case where $\alpha > 1$, the signal is nonstationary. Where $\alpha < 0.5$, it is anticorrelated; i.e., large fluctuations are likely to be followed by small ones. In other words, the slope may be used as an indicator of the roughness [49]. The denoising algorithm is based on the advanced ability of DFA to identify noisy IMFs. Similarly to the algorithm of HHT-based clustering, IMFs $\varphi_i(t)$ are analyzed by the DFA, and the

resulting slopes α_i are used for thresholding. The denoised estimate of the original signal is computed by $\tilde{x}(t) = \sum_{\{i: \alpha_i > \gamma\}} \varphi_i(t)$.

Example 16.5.2. The threshold γ is chosen depending on the autocorrelation property of the noise. When the noise is AWGN, the threshold is set to $\gamma = 0.7$, corresponding to white noise ($\alpha = 0.5$), plus a 0.2 confidence interval to cope with the mode-mixing problem of the EMD. To test the DFA-based denoising, the method is applied to the epileptic EEG with 1024 samples, corrupted by 20 and 0 dB AWGN. Table 16.5.1 shows the denoising performance with a comparison to the wavelet hard and soft thresholding (WHT and WST). The DFA-EMD denoising is capable of suppressing AWGN better than WHT and WST for low SNR levels (0-10 dB), but has close performance at high levels (20-30 dB).

Example 16.5.3. Here, we consider removing electro-oculogram (EOG) artifacts from an EEG signal segment. The threshold is set to $\gamma = 0.4$ due to the anticorrelated nature of EOGs. The reconstruction results of Examples 16.5.2 and 16.5.3 are shown in Fig. 16.5.5. Performance of the EOG artifact removal in Example 16.5.3 is illustrated in Fig. 16.5.6. IMFs with $\alpha < 0.4$, namely $\varphi_i(t)$ for $i = 1, 4, \dots, 11$, are identified as the oscillations of the EOG. On the right, the estimations using the selected IMFs are compared with the original EOG signal, showing agreement.

Table 16.5.1: The Performance Results of the Proposed Denoising

| SNR (dB) | DFA-EMD (dB) | WHT (dB) | WST (dB) |
|----------|--------------|----------|----------|
| 0 | 5.390 | 0.382 | 0.436 |
| 10 | 13.047 | 10.380 | 10.517 |
| 20 | 19.583 | 20.385 | 20.716 |
| 30 | 30.682 | 30.384 | 30.650 |

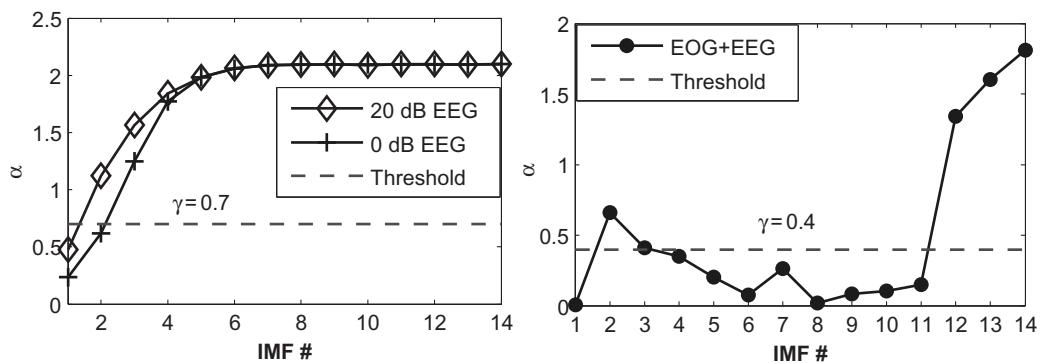
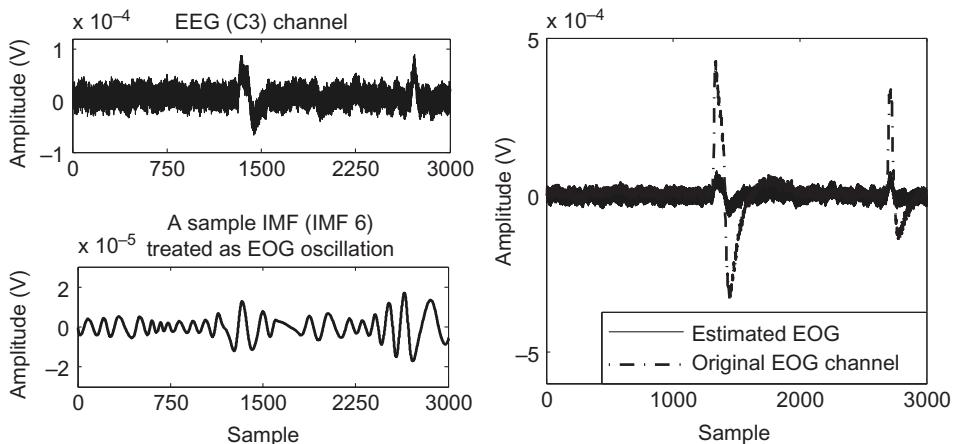


FIGURE 16.5.5

Proposed DFA thresholding for AWGN- and EOG-corrupted EEG signals.

**FIGURE 16.5.6**

EOG removal using the DFA-EMD method.

16.5.4 CLASSIFICATION OF NORMAL AND EPILEPTIC EEGS USING TIME-DOMAIN PROPERTIES BY DFA AND EMD

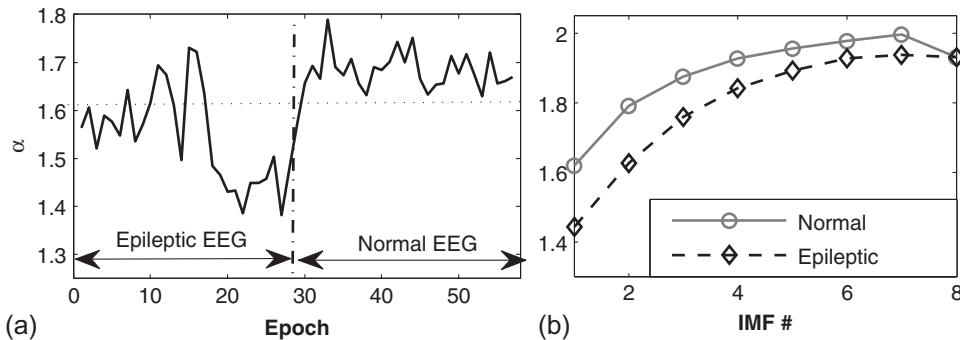
Epileptic and normal EEG recordings are both nonlinear and nonstationary signals; however an epileptic seizure makes the epileptic EEG less nonstationary than normal activity (see Section 16.1) [50,51]. As mentioned in the previous section, the slope α of the DFA characterizes a time series as nonstationary when $\alpha > 1$. Thus, the slope can be used as an indicator of epileptic seizures in the EEG signal.

Example 16.5.4. First, DFA slopes are computed from the EEG recordings directly. Figure 16.5.7(a) shows the results for 58 epochs composed of 29 epileptic and 29 normal EEG epochs with two second record lengths. Note here that the DFA slope calculated from the signal is not capable of distinguishing seizures successfully, due to overlapping scores. We therefore employ EMD together with DFA to enhance the ability to characterize nonstationarities in the EEG signal. The steps of the proposed classification algorithm are given as follows:

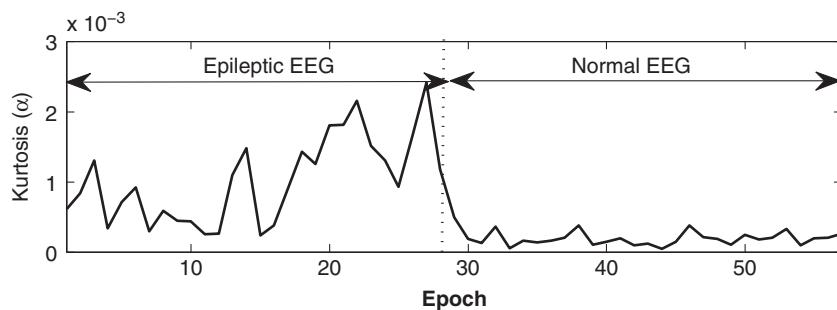
- apply EMD to a given EEG signal $x(t)$ and obtain the IMFs $\varphi_i(t)$, $1 \leq i \leq L$, where L is the number of IMFs;
- compute the DFA slopes α_i of the IMFs;
- deploy the pattern of DFA slopes directly or characterize their statistical properties using kurtosis

$$k = \frac{\frac{1}{L} \sum_{i=1}^L (\alpha_i - \bar{\alpha})^4}{\left[\frac{1}{L} \sum_{i=1}^L (\alpha_i - \bar{\alpha})^2 \right]^2}.$$

Figure 16.5.7(b) shows the averaged DFA slopes of the IMFs for 29 epileptic and 29 normal EEG epochs calculated by the above method. It is observed that IMFs of the epileptic EEGs have lower DFA slopes α than the normal EEG. It is also shown that the components of the epileptic signal are

**FIGURE 16.5.7**

DFA slopes of epileptic and normal EEG (a) with their IMFs (b).

**FIGURE 16.5.8**

Kurtosis of IMFs' DFA slopes for epileptic and normal EEGs.

more suitable for the extraction of classification features in the t domain. Furthermore, kurtosis of the DFA slope pattern is a reliable statistical parameter for distinguishing epileptic from normal EEGs. The kurtosis of the DFA slope for 58 EEG epochs is shown in Fig. 16.5.8. The envelopes of the DFA slopes in Fig. 16.5.7 and their kurtosis in Fig. 16.5.8 could then be integrated into a seizure-detection algorithm.

16.5.5 SUMMARY AND CONCLUSIONS

EMD is a nonlinear, nonstationary, data-driven method that extracts signal components called IMFs. EMD can be used for identifying or testing IMFs. Denoising can be performed by hierarchical clustering based on the HHT, and by using metrics based on DFA to identify irrelevant noisy IMFs. Analysis of IMFs in the time domain using DFA slopes can improve classification of multicomponent signals for which DFA alone is not sufficient, such as epileptic and normal EEGs.

16.6 TIME-FREQUENCY METHODOLOGIES FOR ASSESSMENT OF BIOSIGNALS IN NEUROSCIENCES⁰

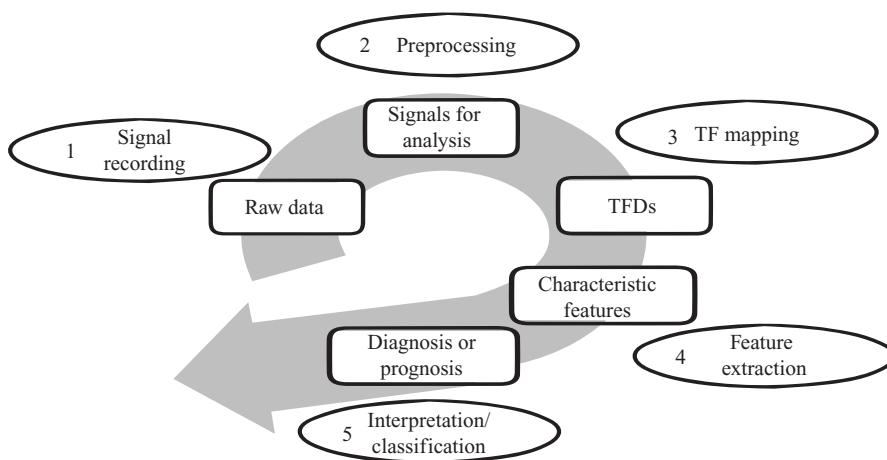
Bio-signals originating from the nervous system represent physiological phenomena that require interpretation for information gathering and decision making related to diagnosis or prognosis [52]. Despite the high level of apparent randomness and variability in neuronal data, advances in computational neuroscience, biomedical signal processing and bioinformatics indicate that with the right representation and analysis methods, biosignals can convey significant information that reflects normal activities in healthy nervous systems [53]; this may then suggest dysfunction in various pathological scenarios, which can be detected with change detection methods [10,54]. Effective analysis of biosignals is therefore crucial for accurate assessment, diagnosis and prognosis of neurological disorders and diseases. Such biosignals can be acquired through two broad approaches: noninvasive recording without penetrating the body, and invasive recording through incision, needle electrodes etc. Noninvasive modalities have a significant role in clinical applications due to their ease of use. Such recordings cover a wide range of bio-signals including, e.g., electroencephalogram (EEG), magnetoencephalogram (MEG), electrocardiogram (ECG), electromyogram (EMG), functional magnetic-resonance imaging (fMRI), and heart-rate variability (HRV) measurements.

16.6.1 TIME-FREQUENCY ANALYSIS ASSESSMENT OF THE NERVOUS SYSTEM

16.6.1.1 *Nonstationarities in Bio-signals*

Time-frequency (t,f) methods offer an elegant and precise approach to exploring the temporal-spectral aspects of biosignals by taking their nonstationary nature into account, thus maximizing the extraction of clinically relevant information. For instance, in newborns, the instantaneous heart rate (derived from newborn ECG) increases as perinatal hypoxia-ischemia (HI) takes place but then gradually reduces or oscillates if the HI is prolonged [55]. The duration that the heart rate (HR) is sustained during an HI insult is associated with the neurological outcome in experiments conducted with newborn piglets using precise HRV measurements made with (t,f) methods [55]. The results indicate that the instantaneous evaluation of the neural responses (and resulting biosignals) is more informative than stationary measurements that project the (t,f) information of the neural states along the time span of analysis. For example, the (t,f) analysis of HRV signals was able to assess the dynamic response of autonomic cardiovascular regulation to perinatal hypoxia [56]. More generally, the (t,f) representation of biosignals offers a more advanced and extensive platform to extract clinically related features that can be used alone or fused to establish classifiers for diagnostic and prognostic purposes. Such (t,f) feature extraction methods have shown superior performance compared to time-only and/or frequency-only features (see Sections 12.6 and 16.3) [10].

⁰Authors: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa), **S. Dong**, University of Queensland Centre for Clinical Research, Herston, QLD, Australia (s.dong@uq.edu.au), **A. Omidvarnia**, The University of Melbourne, Melbourne, Australia (a.omidvarnia@brain.org.au), **P. Zarjam**, Kermanshah Azad University, Iran, and University of New South Wales, Sydney, Australia (p.zarjam@unsw.edu.au), **J.M. O'Toole**, University College Cork, Corcaigh Ireland (j.otoole@ieee.org), and **P. Colditz**, University of Queensland Centre for Clinical Research, Brisbane, Australia (p.colditz@uq.edu.au). Reviewer: G. Azemi.

**FIGURE 16.6.1**

General methodology for using (t, f) methods in biosignal analysis; the arrow indicates the process of how recorded raw data are transformed into clinically useful information by (t, f) methods.

16.6.1.2 General methodology for (t, f) bio-signal analysis

A typical (t, f) approach for studying biosignals can be described by five stages illustrated in Fig. 16.6.1: (1) data acquisition; (2) preprocessing of the recorded data; (3) mapping into the (t, f) domain; (4) exploring the relevant (t, f) features and finally; (5) clinical interpretation and classification of the features. The major preprocessing steps performed need to meet the criteria of the (t, f) analysis (see Section 3.3.1). In the case of HRV analysis, the signals are generated from raw ECG signals through a preprocessing process including R-peak detection, outlier removal, resampling and detrending [56]. The next step is to transform the signals into the (t, f) domain using a suitable (t, f) distribution (TFD) (see Chapters 2 and 3 for more details about the required TFD properties). This step is followed by extracting features relevant to the study with acceptable biological plausibility or assumptions. These features, either standalone or along with other clinical measurements, can be used to interpret the underlying physiological or pathophysiological processes and/or to derive effective discriminants for clinical classification.

Selected examples of applied neuroscience studies are developed below to illustrate the use of the general methodology of (t, f) biosignal analysis. Such applications in general follow the five analysis stages of Fig. 16.6.1, where the (t, f) method either plays a crucial role in the analysis or shows superior performance.

16.6.2 AUTONOMIC NERVOUS SYSTEM RESPONSE TO PERINATAL HYPOXIA

16.6.2.1 The problem

The autonomic nervous system, with elements within both the central and peripheral nervous systems, regulates the internal organs such as the heart via two components: the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS). The SNS accelerates the HR and increases the

contractility of myocytes so as to enhance the cardiac output. The PNS (vagal pathway) works in an antagonistic way to reduce the HR and cardiac output. The global effect of the autonomic regulation on the cardiovascular function is the combination of contributions from both SNS and PNS. Under hypoxic conditions, the sympathetic branch is activated, whereas the parasympathetic branch is suppressed [56]. The related HRV signals include two major components: the high-frequency (HF) component and the low-frequency (LF) component. The HF component reflects the activity of PNS, whereas the LF component provides a measure of sympathetic regulation. The spectral power of the components is used as a quantification of each autonomic branch [55]. A study quantified the dynamic sympathetic response to hypoxia using the instantaneous power of the LF component in a newborn piglet experiment [55,56]. This measurement is then used to investigate whether the SNS is a determinant in the physiological response to hypoxia and subsequent neurodevelopmental outcome in neonates who are at risk of HI brain injury [56].

16.6.2.2 Principle

A quantification of HRV components is obtained using the TFD of HRV epochs and the estimate of instantaneous component power. The TFD representation provides a dynamic measure of HRV component power that traditional time or frequency domain methods are not capable of doing. For a given HRV epoch $x(t)$, its analytic signal $z(t)$ is formed using the Hilbert transform (HT) (Section 1.2.3) and then mapped into the (t,f) domain using a HRV-adapted TFD denoted CS-MBD (see Section 3.3.4.2). As the CS-MBD can attain negative values which is undesirable for interpreting the TFD values as instantaneous power, an option is to use a nonnegative TFD (e.g., spectrogram; see Section 3.4), but an objective comparison has shown that the compact support modified B-distribution (CS-MBD) outperforms the spectrogram in HRV signal representation in terms of resolution [56]. As the study also indicates, the negative values can be neglected as shown in Fig. 16.6.2(c); therefore, the CS-MBD has been chosen in the analysis for this application.

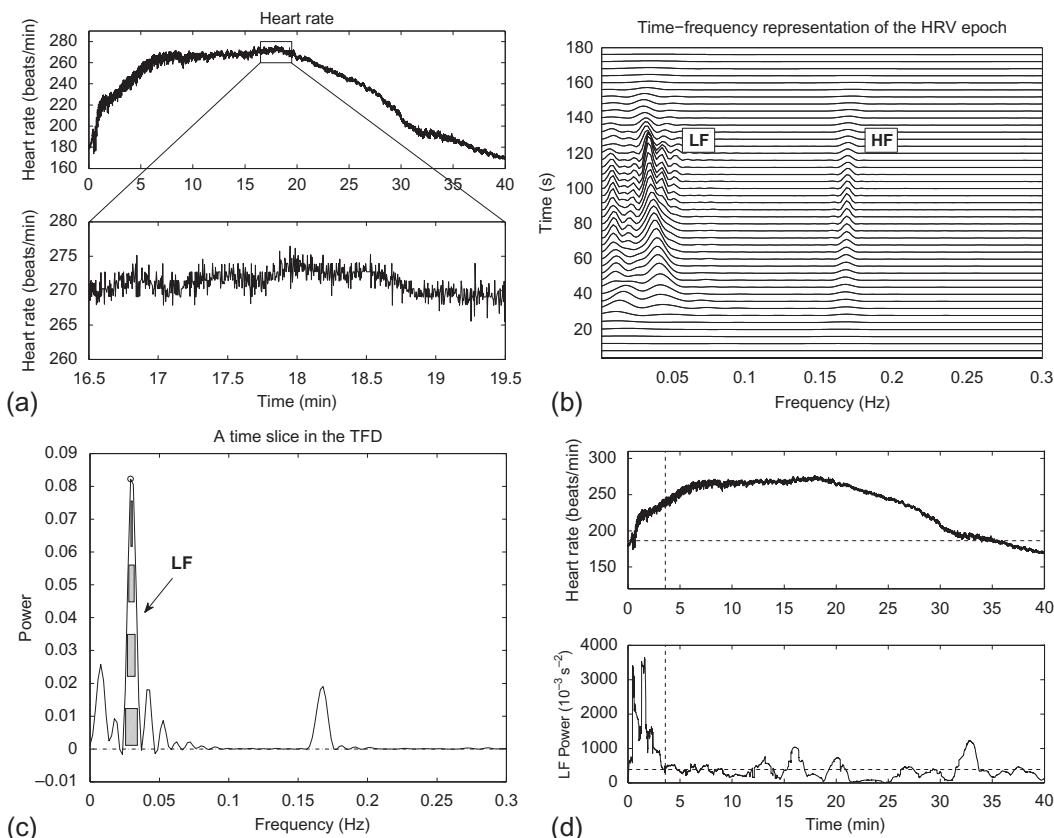
The IF of HRV components is then estimated from the TFD (Section 10.6). The power of the components at each instant is calculated by integration of the area between the main lobe centered on the estimated IF and the zero axis. The evolution of HRV power is used as a continuous quantification of the sympathetic and parasympathetic regulation of the cardiovascular system [55].

16.6.2.3 Methods

In several experiments, anesthesia was induced in nine full-term one-day-old piglets that were intubated and ventilated using a neonatal ventilator at a rate of 30 breaths per minute (BPM). Hypoxia was induced by reducing the ventilation rate to 10 BPM and the fraction of inspired oxygen (FiO_2) to 0.1 for 40 min. A 10-min epoch prior to hypoxia was used to compute the HRV baseline [55,56].

The HRV signals were derived via HT and thresholding to locate the R-peaks in the ECG. The instantaneous HR was then computed as the inverse of the time interval between consecutive R-peaks.

The 40-min HRV signals during hypoxia were divided into 3-min epochs with 50% overlap. Each epoch was mapped into the (t,f) domain using the CS-MBD. The 40-min TFD was then obtained by connecting 25–75% of the consecutive TFD epochs in the time direction. This overlap approach effectively preserves the accuracy of the power estimation by attenuating the distortion introduced by the finite length of epochs. Next, the IF of the LF component is estimated.

**FIGURE 16.6.2**

Time-frequency analysis of HRV signals in a newborn piglet exposed to hypoxia: (a) is an HR series (and a 3-min epoch for analysis) from a piglet exposed to hypoxia; (b) displays the TFD of the epoch where the LF and HF components appear clearly; (c) plots a time slice in the TFD where the circle indicates the IF of the LF component and the shaded area represents the power; (d) shows the HR (top) and concurrent LF spectral power (bottom) over 40 min. Note that in (d), the horizontal dashed lines represent the respective baseline level, and the vertical dashed line indicates the time that the elevated LF power reduces to the baseline level.

16.6.2.4 Results

Figure 16.6.2 shows the nonstationary HRV HF and LF components. The figures show that the LF power that reflects the sympathetic response to hypoxia increases at the beginning of hypoxia and then declines to the baseline level within 5 min. On the other hand, the HR remains elevated and then maintains its maximum level until about 20 min. The initial increase in LF power can be interpreted as the sympathetic reaction to hypoxia, which accelerates the HR and strengthens the cardiac contractility in order to maintain the cardiac output to compensate for the reduction of blood oxygen content. The LF power reduction, however, appears prior to the HR reaching maximum; this indicates that the sympathetic pathway is not yet fully functional at this age or at least is not the dominant component of the response. Either interpretation has important implications for treatment of the newborn following hypoxia ischemia [55].

16.6.3 SEIZURE DETECTION USING COMBINED EEG AND HRV

16.6.3.1 The problem

The use of combined scalp EEG and HRV signals for the detection of seizure in both newborns and adults is justified, because scalp EEG is a potentially effective representation of brain abnormalities, and HRV can describe changes that manifest in the autonomic nervous system as seizure activity [57].

16.6.3.2 Principle

The application of seizure detection using EEG and HRV signals follows the methodology described in [Section 16.6.2](#) and [Fig. 16.6.1](#). EEG and HRV signals contribute complementary information that is sensitive and specific to seizure events. Thus, characteristic features can be extracted, selected, and fed into classifiers for classification training.

16.6.3.3 Methods

In an experiment, EEG and ECG signals are recorded simultaneously. The features from both types of signals can be extracted either separately or jointly [57]. A selection of possible features for seizure detection is presented in [Table 16.6.1](#). Joint features can be designed by taking into account the characteristics of interaction between the central and autonomic nervous systems; e.g., one particular feature, the wavelet coherence, quantifies the synchronization between EEG and HRV signals in the analysis of epileptic seizure activity [58].

A feature-conditioning process can help to minimize the overlap between elements of the feature space by selecting the most discriminative features and decreasing the feature-space dimension to reduce the computation load. Two feature-selection approaches can be described [10,57]. The first approach is based on the selection of those features (or a combination of features) that leads to the highest classification rate among all feature sets ([Section 15.6](#)) [57]. In the second approach, the features are chosen through a classification-independent procedure based on the feature-discriminative capability; the area under the curve (AUC) of the receiver operating characteristic (ROC) analysis is one commonly used criterion [10].

The selected features are then fed into a classifier that consists of one of several possible machine-learning algorithms [57].

Table 16.6.1: EEG and HRV Features for Seizure Detection Combining EEG and HRV Modalities (see list in [Section 12.6](#))

| Domain | EEG | HRV |
|---------|--|--|
| t | Statistical features: (maximum, minimum, mean, variance, standard deviation, skewness, kurtosis, percentiles [57,59]); mobility and complexity [59]; entropy [57,59]; nonlinear energy [57]; line length [57] | Statistical features: (maximum, minimum, mean, variance, standard deviation, skewness, kurtosis, percentiles [57,59]); mobility and complexity [59]; derivative of HRV [57] |
| f | Dominant spectral peak [57]; spectral peak bandwidth [57]; power ratio [57]; autoregressive models [57,59]; spectral band power [57,59]; spectral entropy [57]; spectral edge and power; decorrelation time [59] | Power spectral density [57]; spectral power of VLF, LF, and HF component [59]; LF/HF power ratio [59]; spectral entropy [57] (see also Sections 15.6 and [1,3,10]) |
| (t,f) | Wavelet coefficients [57]; IF, band and amplitude [1]; TFD entropy; (t,f) energy [1]; singular TFD values [1]; (t,f) variance, skewness, kurtosis, spectral flux, etc. | Wavelet coefficients [59]; IF, band and amplitude [1]; TFD entropy [1] (see also Section 12.6 and [3,10]) |

16.6.3.4 Results

The (t,f) methodology described here has been applied in a number of studies and has shown promising performance in the detection of seizures in the newborn and adult [57,59]. In the study of adult epileptic seizure detection [59], high sensitivity and specificity (94.5% and 91.2%, respectively) have been reported for a classification method using both EEG and HRV features, which outperforms the method that uses either type of signal alone. In addition, the detection performance generally improves when (t,f) features are used (Section 12.6) [10].

16.6.4 MEASURING PHASE SYNCHRONY FOR EEG ABNORMALITY ASSESSMENT

16.6.4.1 The problem

Multivariate phase-synchrony analysis techniques can be used for exploring the time-varying interactions between scalp EEG channels [11].

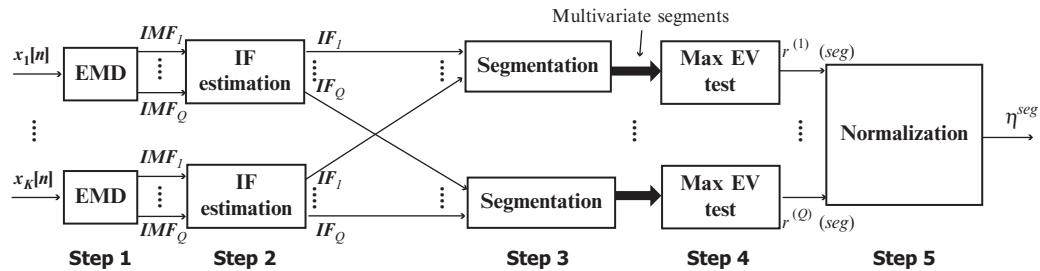
A generalized phase-synchrony measure can mitigate the limitations of classical pair-wise asynchrony assessment in multichannel neonatal EEG, as a bivariate measure is blind to the global disruption of synchrony among abnormal EEG channels. This approach uses the concept of cointegration between the corresponding instantaneous phase signals through multivariate autoregressive (MVAR) modeling [11].

16.6.4.2 Principle

Suppose $x[n]$ and $y[n]$ are two periodic signals, phase-locked of order $P_x : P_y$ where both P_x and P_y are positive integer numbers. Let the two signals evolve temporally from a similar point on the time axis. If the phase-locking ratio (P_x/P_y) is rational, the two signals will cross each other periodically at the same initial common value with a time period associated with the least common multiple (LCM) between P_x and P_y . Therefore, the rational phase-locking order can be intuitively interpreted with a physical meaning for periodic signals. In contrast, the two periodic signals never arrive at the same point with passing time if the phase locking ratio is irrational. The situation of phase synchrony in nonperiodic signals is not as straightforward as the periodic ones. It becomes even more difficult for nonstationary signals that, by definition, cannot be periodic. In this case, the issue can be clarified by using the concept of frequency flows in the (t,f) domain. Phase synchrony is equivalent to frequency synchrony due to the direct relationship between the instantaneous phase (IP) and the IF [11]; e.g., if two signals have similar IF laws during a time interval, they are phase-locked of order 1:1 over that time period. Consequently, a linear relationship between two IF laws with rational gain (P_x/P_y) implies phase-locking of order $P_x : P_y$. This classical definition of phase synchrony cannot explain generalized phase synchrony (GePS) where the linear relationships between phase signals are irrational. An alternative interpretation is therefore provided for GePS based on the concept of cointegration; for a multichannel nonstationary signal, if there is a linear relationship between the IF laws of a subset of channels during a reasonably long time period, they are said to be generally phase-synchronized over that time period [11].

16.6.4.3 Methods

For a nonstationary K -dimensional signal $\mathbf{x}[n]$, the procedure of GePS assessment is shown in Fig. 16.6.3 and can be described by the following steps [11]:

**FIGURE 16.6.3**

Block diagram of the GePS assessment approach. The narrow arrows represent the univariate signals, while the large bold arrows show the multivariate signals.

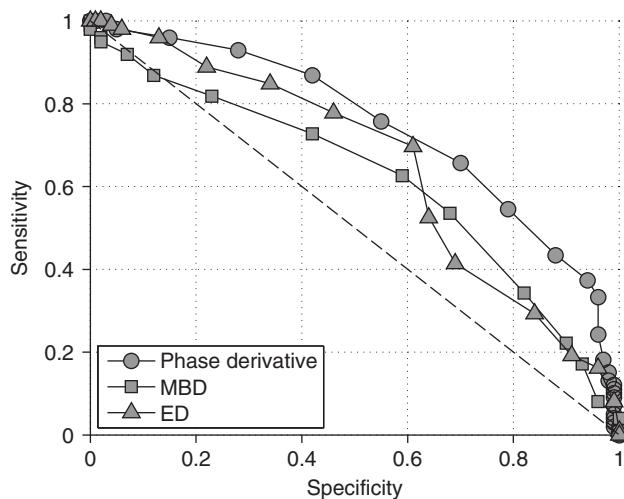
1. Each channel $x_i[n](i = 1, 2, \dots, K)$ is decomposed into Q intrinsic mode functions (IMFs) ($i = 1, 2, \dots, Q$) using the empirical mode decomposition (EMD) (Section 4.12). The parameter Q can be estimated by an EMD stopping criterion such as $Q = \min Q_i(i = 1, 2, \dots, K)$, where Q_i is the number of IMFs in the i th channel. The same value is then used for all channels.
2. The IFs $f_i^q[n]$ are then estimated as derivatives of the unwrapped phase signals $\varphi_i^q[n]$, or by using other IF estimates.
3. The IFs of all channels at each decomposition level, i.e., $\mathbf{f}^q = [f_1^q f_2^q \dots f_K^q]$ are divided into nonoverlapping temporal segments. Their minimum window length is determined based on the requirement of the MVAR model parameter estimation. As a rule of thumb, the length should be significantly larger than Kp^2 (p is the MVAR model order in the maximum eigenvalues test [11]).
4. The maximum eigenvalues test is applied on each multivariate segment at 99% confidence level and the linear relationships between IFs are extracted [11].
5. The phase synchrony measure for each segment is defined as the normalized number of the cointegrating relationships r^q over the IMF components [11]:

$$\eta^{\text{seg}} = \frac{1}{QK} \sum_{q=1}^Q r^q. \quad (16.6.1)$$

This measure varies between 0 and 1 where 0 means no cointegrating relationship within IF laws and 1 implies complete generalized phase-locking. Extraction of the GePS measure from all segments over the entire multichannel signal leads to a feature-time series representing the time-varying generalized phase synchrony within channels. The segment-based GePS values can also be used to test the statistical difference between distinct conditions (e.g., seizure vs. nonseizure).

16.6.4.4 Results

Multichannel seizure EEG recordings of five newborns with hypoxic-ischemic encephalopathy (HIE) were collected from 14 monopolar channels using the 10-20 standard, and 8 channels (F3, F4, C3, C4, P3, P4, O1 and O2) out of 14 were considered for further analysis [52]. The arrangement of electrodes was chosen as the combination of four symmetric electrode locations from left and right hemispheres in order to facilitate interhemispheric phase-synchrony assessment [11]. In total, 100

**FIGURE 16.6.4**

ROC curves of the IF estimation methods used in this section to extract GePS within the seizure/nonseizure EEG segments. The methods include: derivative of instantaneous phase, modified B-distribution (MBD), and Exponential distribution (ED). The diagonal dashed line in the figure represents the chance level.

nonoverlapping ictal segments (with seizures) and 100 nonoverlapping interictal segments (with no marked seizures) were extracted randomly from 27 min of artifact-free seizure signals and 39 min of artifact-free nonseizure signals obtained from all subjects. Then, for both seizure and nonseizure groups, (t,f) -based IF estimation quantified GePS in newborn EEG [11].

The ROC IF curves on the seizure/nonseizure EEG segments indicated that the sensitivity is much higher than the false alarm (1-specificity) (see Fig. 16.6.4). It indicates an elevated level of GePS within the seizure segments in contrast to the nonseizure segments [11], therefore validating the use of the (t,f) approach. See Section 16.4 for another application of phase synchrony.

16.6.5 MEASURING COGNITIVE WORKLOAD BY ASSESSING EEG SIGNALS

16.6.5.1 *The problem*

There is an ongoing interest for cognitive monitoring in the fields of cognitive neuroscience, biomedical engineering, human-computer-interactions, brain-computer interfaces and psychology. The objective is to gather information about the user's cognitive state, such as the mental or cognitive activities, memory workload and task engagement. These states indicate how hard a user's cognitive system is working to solve a problem or use an interface, and can be employed to support the user's goals or adapt an interface [60]. In cognitive monitoring, measuring the imposed load on the working memory during a cognitive process, i.e., cognitive load, is of high importance. This is to avoid the user's mental overload and maintain efficiency and productivity in critical or high mental load workplaces, such as medical or emergency departments, air traffic control and military operations [60]. Among different techniques

available to measure cognitive load, monitoring the brain activity using neuro-imaging techniques such as EEG has been recognized as the most sensitive and consistent reflector of working memory load [61].

16.6.5.2 Principle

The wavelet transform (WT) provides a (t,f) representation of the given signal with a good time-resolution at high frequencies and good frequency resolution at low frequencies, so-called multiresolution property [57,59] (see Section 2.7.6). The characteristic feature suitable for classifying cognitive load levels can be extracted from the wavelet coefficients. One such feature is the entropy, which is a measure of regularity or order, and in the case of EEG signals, indicates the degree of synchrony of the neural groups participating in different neural responses.

16.6.5.3 Methods

Let $v = (v_1 \ v_2 \ \cdots \ v_M)^T$ denote the wavelet coefficients of an EEG segment in a particular scale, and p_i is defined as $p_i = \frac{|v_i|}{\sum_{i=1}^M |v_i|}$ [60]. Note that $\sum_{i=1}^M p_i = 1$. The following entropies can be calculated as features for capturing the EEG signal's regularity and order variations [60]:

1. Shannon entropy: $H_S = -\sum_{i=1}^M p_i \log(p_i)$,
2. Tsallis entropy: $H_T = \frac{1}{q-1} \sum_{i=1}^M (p_i - p_i^q); \ 0 < q < 1$,
3. Escort-Tsallis entropy: $H_{ET} = \frac{1}{q-1} \left[1 - \left(\sum_{i=1}^M (p_i)^{q-1} \right)^{-q} \right]$.

16.6.5.4 Results

In an experiment, an “addition” task with seven levels of difficulty was designed. The task levels of difficulty varied from one-digit addition (very easy) to multidigit addition (extremely difficult). The EEG recordings were conducted under controlled conditions in an electrically isolated lab, at the ATP Laboratory of National ICT Australia in Sydney, for 12 participants. The above-mentioned entropy-based features were extracted from the artifact-free EEG segments (of 5-s length) from selected EEG channels (i.e., Fp1, AF3, F7, F3, FC1, FC5, FC6, FC2, F4, F8, AF4, Fp2, and Fz) of each participant. The extracted features were then fed into a multilayer neural network classifier, and a leave-one-out cross-validation technique was used for training and testing data. The results showed a consistent decline in the medians of the entropy-based features as the task load increased, indicating that the degree of the signal disorder decreases as the working memory load imposed increases. Classification accuracies of 98.35%, 98.44%, and 92.21% were obtained using H_S , H_T , and H_{ET} , respectively. The assessment of the strength of phase synchrony among the EEG channels revealed that as the load level increases, the level of synchrony of the cell groups involved in neural responses increases [60]. See Section 16.4 for another application of phase synchrony.

16.6.6 ASSESSING BRAIN INJURY BY IF ESTIMATION OF PLEDs

16.6.6.1 The problem

Periodic lateralized epileptiform discharges (PLEDs) are sharp EEG transients that can occur after brain injury; treatment options for patients with PLEDs are unclear. The periodicity of PLEDs is an important prognostic feature and can aid treatment decisions. Although they are called periodic, the period between discharges can significantly vary over time and they are in fact nonstationary in nature.

Estimating this time-varying periodicity, or the IF, is difficult because of the relatively high-amplitude background EEG activity.

16.6.6.2 Principle

PLED signals consist of a series of spike-like waveforms with a time-varying periodicity embedded in background EEG activity. These are harmonic signals with a slowly varying IF law and are thus similar to seizure and voiced-speech signals [62]. Standard IF methods ([Section 10.6](#)) fail to adequately estimate the IF of PLEDs for two reasons. First, the relatively high-amplitude background EEG activity masks the spike-like PLEDs. Second, the varying morphology of PLED waveforms changes the spectral envelope, thus making it difficult to determine the fundamental frequency of a harmonic signal. These two issues are dealt with as follows:

1. The nonlinear energy operator is applied to enhance these spike-like waveforms. This is achieved either by taking the envelope of the derivative, or using the Teager-Kaiser operator ([Section 4.10](#)) on the (discrete) signal $s[n]$,

$$y[n] = s^2[n] - s[n+1]s[n-1]. \quad (16.6.2)$$

2. A homomorphic filtering then suppresses the spectral envelope from the TFD of $z[n]$, the analytic associate of $y[n]$. Using the log function, the TFD $\rho_z[n, k]$ is transformed to the modified pseudo-time-lag domain as: $R_z^m[n, m] = \mathcal{F}^{-1}_{m \leftarrow k} \{\log[\rho_z[n, k]]\}$, where m indicates that $R_z[n, m]$ is modified as shown. Before taking the log transform of the TFD, any TFD values less than or equal to zero are set to a small number (approximately equal to machine precision). In this domain, the spectral modulation is suppressed using a low-pass filter $l(\tau)$ and then transforming back to the (t, f) domain:

$$\eta_z[n, k] = \exp \left[\mathcal{F}_{m \rightarrow k} \{R_z^m[n, m]l[m]\} \right]. \quad (16.6.3)$$

This filtering approach removes the spectral envelope making it easier for the component-linking method to estimate the IF ([Section 10.6](#)).

16.6.6.3 Methods

First, Eq. (16.6.2) is applied to enhance the peaks of the PLED waveforms. Second, the signal is transformed to the (t, f) domain using a quadratic TFD with a separable kernel ([Section 5.7](#)), using Hanning-Hamming windows [1] to suppress cross-term energy [62]. Third, the spectral envelope in the TFD is suppressed using Eq. (16.6.3). And lastly, the IF is estimated from this modified TFD [1].

To test the performance, the method was compared with the IF component, linking method on 33 segments of EEG containing PLEDs that were extracted from the EEG records of eight patients at the Cruces University Hospital, Barakaldo, Spain. The start of each PLED waveform was annotated by two clinical neurophysiologists, and the IF was calculated by inverting this time-varying periodicity.

16.6.6.4 Results

Enhancing the spike-like waveforms and suppressing the spectral envelope of the TFD significantly improve the IF estimation of PLEDs. This has significant prognostic information to aid physicians to optimize neuroprotective treatments [1].

16.6.7 SUMMARY AND CONCLUSIONS

The results obtained in the studies described above suggest that: (1) (t,f) methods are capable of characterizing the dynamic activity of the underlying neurological system; (2) characteristic (t,f) features exhibit improved performance in the identification of abnormalities in the underlying neurological systems.

Hence, (t,f) techniques should play an increasing role in applications including clinical diagnosis and prognosis of neurological diseases. In addition, spatial or multi-channel (t,f) distributions for neonatal EEG signals can assist in accurately locating the source of abnormalities in the brain by accounting for the nonstationary and multicomponent characteristics of signals collected from multisensors (see [Section 8.1](#) and [63]). An STFD matrix extracted from multichannel EEG signals can be decomposed by the multiple signal classification (MUSIC) algorithm [64] over preselected regions of interest in the (t,f) domain. Alternatively, (t,f) blind-source separation methods can be utilized to estimate the temporo-spectral independent components of the recorded EEG signals [1].

Extending the STFT approach to more efficient TFDs (as discussed in [Chapters 2](#) and [3](#)) in neurosciences should result in valuable improvements and allow further insights into healthy or damaged brain functions.

REFERENCES

- [1] B. Boashash, G. Azemi, J.M. O'Toole, "Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Proc. Mag.* 30 (6) (2013) 108-119.
- [2] N.J. Stevenson, J.M. O'Toole, L.J. Rankine, G.B. Boylan, B. Boashash, "A nonparametric feature for neonatal EEG seizure detection based on a representation of pseudo-periodicity", *Med. Eng. Phys.* 34 (4) (2012) 437-446.
- [3] B. Boashash, N.A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review", *Digital Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [4] N.J. Stevenson, M. Mesbah, B. Boashash, "Modelling newborn EEG background using a time-varying fractional Brownian process", in: Proceedings of the 15th European Signal Processing Conference, vol. 1, 2007, pp. 1246-1250.
- [5] M.S. Scher, "Electroencephalography of the newborn: normal and abnormal features", in: E. Niedermeyer, F.H. Lopes da Silva (Eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, fifth ed., Lippincott Williams & Wilkins, Philadelphia, 2005, pp. 937-989 (Chapter 49).
- [6] R.R. Clancy, A. Legido, "The exact ictal and interictal duration of electroencephalographic neonatal seizures", *Epilepsia* 28 (5) (1987) 537-541.
- [7] B. Boashash, G. Azemi, "A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG", *Digit. Signal Process.* 28 (2014) 28-38.
- [8] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals; Part 2: Algorithms and applications", *Proc. IEEE* 80 (4) (1992) 519-568.
- [9] N.J. Stevenson, I. Korotchikova, A. Temko, G. Lightbody, W.P. Marnane, G.B. Boylan, "An automated system for grading EEG abnormality in term neonates with hypoxic-ischaemic encephalopathy", *Ann. Biomed. Eng.* 41 (4) (2013) 775-785.

- [10] B. Boashash, G. Azemi, N.A. Khan, “Principles of time-frequency feature extraction for change detection in nonstationary signals: applications to newborn EEG abnormality detection”, *Pattern Recognit.* 48 (3) (2015) 616-627, <http://dx.doi.org/10.1016/j.patcog.2014.08.016>.
- [11] A. Omidvarnia, G. Azemi, P.B. Colditz, B. Boashash, “A time-frequency based approach for generalized phase synchrony assessment in nonstationary multivariate signals”, *Digital Signal Process.* 23 (3) (2013) 780-790.
- [12] B. Boashash, “Time-frequency signal analysis”, in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 418-517 (Chapter 9).
- [13] L.J. Rankine, M. Mesbah, B. Boashash, “IF estimation for multicomponent signals using image processing techniques in the time-frequency domain”, *Signal Process.* 87 (6) (2007) 1234-1250.
- [14] N.J. Stevenson, M. Mesbah, G.B. Boylan, P.B. Colditz, B. Boashash, “A nonlinear model of newborn EEG with nonstationary inputs”, *Ann. Biomed. Eng.* 38 (9) (2010) 3010-3021.
- [15] M. Basseville, “Distance measures for signal processing and pattern recognition”, *Signal Process.* 18 (4) (1989) 349-369.
- [16] L.J. Rankine, N.J. Stevenson, M. Mesbah, B. Boashash, “A nonstationary model of newborn EEG”, *IEEE Trans. Biomed. Eng.* 54 (1) (2007) 19-28.
- [17] M.W. Berry, M. Browne, A.N. Langville, V.P. Pauca, R.J. Plemmons, “Algorithms and applications for approximate nonnegative matrix factorization”, *Comput. Stat. Data Anal.* 52 (1) (2007) 155-173.
- [18] B. Ghoraani, S. Krishnan, “Time-frequency matrix feature extraction and classification of environmental audio signals”, *IEEE Trans. Audio Speech Lang. Process.* 19 (7) (2011) 2197-2209.
- [19] V. Sucic, N. Saulig, B. Boashash, “Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy”, *EURASIP J. Adv. Signal Process.* 2011 (1) (2011) <http://dx.doi.org/10.1186/1687-6180-2011-125>, article ID 2011:125, 11 pp.
- [20] L.M. Douglass, J.Y. Wu, N.P. Rosman, C.E. Stafstrom, “Burst suppression electroencephalogram pattern in the newborn: predicting the outcome”, *J. Child Neurol.* 17 (6) (2002) 403-408.
- [21] H. Kim, H. Park, “Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method”, *SIAM J. Matrix Anal. Appl.* 30 (2) (2008) 713-730.
- [22] M.-E. Lynall, D.S. Bassett, R. Kerwin, P.J. McKenna, M. Kitzbichler, U. Muller, E. Bullmore, “Functional connectivity and brain networks in schizophrenia”, *J. Neurosci.* 30 (28) (2010) 9477-9487.
- [23] J.M. Peters, M. Taquet, C. Vega, S.S. Jeste, I.S. Fernández, J. Tan, C.A. Nelson III, M. Sahin, S.K. Warfield, “Brain functional networks in syndromic and nonsyndromic autism: a graph theoretical study of EEG connectivity”, *BMC Med.* 11 (2013) <http://dx.doi.org/10.1186/1741-7015-11-54>, 16 pp.
- [24] P. Zarjam, M. Mesbah, B. Boashash, “Seizure detection in newborns’ EEG signals using time-scale analysis”, *GESTS Int. Trans. Comput. Sci. Eng.* 27 (1) (2006) 7-14.
- [25] P. Zarjam, G. Azemi, M. Mesbah, B. Boashash, “Newborn EEG seizure detection using t-f distance measure”, *GESTS Int. Trans. Comput. Sci. Eng.* 23 (1) (2005) 163-171.
- [26] B. Boashash, M. Mesbah, “Using DSP to detect seizures in newborns”, *Elect. Syst. Software* 1 (3) (2003) 34-37, ISSN 1479-8336, URL http://digital-library.theiet.org/content/journals/10.1049/ess_20030605.
- [27] M. Rubinov, O. Sporns, “Complex network measures of brain connectivity: uses and interpretations”, *NeuroImage* 52 (3) (2010) 1059-1069.
- [28] C.J. Stam, W. de Haan, A. Daffertshofer, B.F. Jones, I. Manshanden, A.M. van Cappellen van Walsum, T. Montez, J.P.A. Verbunt, J.C. de Munck, B.W. van Dijk, H.W. Berendse, P. Scheltens, “Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer’s disease”, *Brain* 132 (1) (2009) 213-224.
- [29] C. Grefkes, G.R. Fink, “Reorganization of cerebral networks after stroke: new insights from neuroimaging with connectivity approaches”, *Brain* 134 (5) (2011) 1264-1276.
- [30] P.J. Uhlhaas, W. Singer, “Abnormal neural oscillations and synchrony in schizophrenia”, *Nat. Rev. Neurosci.* 11 (2) (2010) 100-113.

- [31] A. Omidvarnia, G. Azemi, B. Boashash, J. O'Toole, P. Colditz, S. Vanhatalo, "Measuring time-varying information flow in scalp EEG signals: orthogonalized partial directed coherence", *IEEE Trans. Biomed. Eng.* 61 (3) (2014) 680-693, ISSN 0018-9294, <http://dx.doi.org/10.1109/TBME.2013.2286394>.
- [32] S. Aviyente, E.M. Bernat, W.S. Evans, S.R. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain", *Human Brain Mapping* 32 (1) (2011) 80-93.
- [33] J.-P. Lachaux, E. Rodriguez, J. Martinerie, F.J. Varela, "Measuring phase synchrony in brain signals", *Human Brain Mapping* 8 (4) (1999) 194-208.
- [34] S. Aviyente, A.Y. Mutlu, "A time-frequency-based approach to phase and phase synchrony estimation", *IEEE Trans. Signal Process.* 59 (7) (2011) 3086-3098.
- [35] D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, "The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains", *IEEE Signal Process. Mag.* 30 (3) (2013) 83-98.
- [36] D.I. Shuman, B. Ricaud, P. Vandergheynst, "Vertex-frequency analysis on graphs", arXiv:1307.5708 (2013).
- [37] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis", *Proc. R. Soc. A* 454 (1971) (1998) 903-995.
- [38] R.T. Rato, M.D. Ortigueira, A.G. Batista, "On the HHT, its problems, and some solutions", *J. Mech. Syst. Signal Process.* 22 (6) (2008) 1374-1394.
- [39] R. Gupta, A. Kumar, R. Bahl, "Estimation of instantaneous frequencies using iterative empirical mode decomposition", *Signal Image Video Process.* (2012) <http://dx.doi.org/10.1007/s11760-012-0305-5>, 14 pp.
- [40] L. Ai, J. Wang, R. Yao, "Classification of parkinsonian and essential tremor using empirical mode decomposition and support vector machine", *Digit. Signal Process.* 21 (4) (2011) 543-550.
- [41] Z. Wu, N.E. Huang, "Statistical significance test of intrinsic mode functions", in: N.E. Huang, S.S.P. Shen (Eds.), *Hilbert-Huang Transform and Its Applications*, World Scientific, Singapore, 2005, pp. 107-127 (Chapter 5).
- [42] J. Fleureau, J.-C. Nunes, A. Kachenoura, L. Albera, L. Senhadji, "Turning tangent empirical mode decomposition: a framework for mono- and multivariate signals", *IEEE Trans. Signal Process.* 59 (3) (2011) 1309-1316.
- [43] X. Hu, S. Peng, W.-L. Hwang, "EMD revisited: a new understanding of the envelope and resolving the mode-mixing problem in AM-FM signals", *IEEE Trans. Signal Process.* 60 (3) (2012) 1075-1086.
- [44] Z. Wu, N.E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method", *Proc. R. Soc. A* 460 (2004) 1597-1611.
- [45] P. Flandrin, G. Rilling, P. Gonçalvès, "Empirical mode decomposition as a filter bank", *IEEE Signal Process. Lett.* 11 (2) (2004) 112-114.
- [46] J.F. Cui, H.S. Chae, "Applying agglomerative hierarchical clustering algorithms to component identification for legacy systems", *Informat. Software Technol.* 53 (6) (2011) 601-614.
- [47] C.K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series", *Chaos: Interdiscipl. J. Nonlinear Sci.* 5 (1) (1995) 82-87.
- [48] A. Schaefer, J.S. Brach, S. Perera, E. Sejdić, "A comparative analysis of spectral exponent estimation techniques for $1/f^\beta$ processes with applications to the analysis of stride interval time series", *J. Neurosci. Meth.* 222 (2014) 118-130.
- [49] M. Jospin, P. Caminal, E.W. Jensen, H. Litvan, M. Vallverdú, M.M. R.F. Struys, H.E.M. Vereecke, D.T. Kaplan, "Detrended fluctuation analysis of EEG as a measure of depth of anesthesia", *IEEE Trans. Biomed. Eng.* 54 (5) (2007) 840-846.
- [50] B. Boashash, M. Mesbah, "A time-frequency approach for newborn seizure detection", *IEEE Eng. Med. Biol. Mag.* 20 (5) (2001) 54-64.
- [51] P. Celka, B. Boashash, P. Colditz, "Preprocessing and time-frequency analysis of newborn EEG seizures", *IEEE Eng. Med. Biol. Mag.* 20 (5) (2001) 30-39.

- [52] E. Kaniusas, Biomedical Signals and Sensors I: Linking Physiological Phenomena and Biosignals, Springer-Verlag Berlin, Heidelberg, 2012, pp. 1-26 (Chapter: Fundamentals of biosignals).
- [53] R.B. Stein, E.R. Gossen, K.E. Jones, “Neuronal variability: noise or part of the signal?” *Nat. Rev. Neurosci.* 6 (5) (2005) 389-397.
- [54] M. Roessgen, A. Zoubir, B. Boashash, “Seizure detection of newborn EEG using a model-based approach”, *IEEE Trans. Biomed. Eng.* 45 (6) (1998) 673-685, ISSN 0018-9294, <http://dx.doi.org/10.1109/10.678601>.
- [55] S. Dong, B. Boashash, G. Azemi, B.E. Lingwood, P.B. Colditz, “Automated detection of perinatal hypoxia using time-frequency-based heart rate variability features”, *Med. Biol. Eng. Comput.* 52 (2) (2014) 183-191, ISSN 0140-0118, <http://dx.doi.org/10.1007/s11517-013-1129-3>.
- [56] S. Dong, G. Azemi, B. Boashash, “Improved characterization of HRV signals based on instantaneous frequency features estimated from quadratic time-frequency distributions with data-adapted kernels”, *Biomed. Signal Process. Control* 10 (2014) 153-165, ISSN 1746-8094, <http://dx.doi.org/10.1016/j.bspc.2013.11.008>.
- [57] I. Mporas, V. Tsirka, E.I. Zacharaki, M. Koutroumanidis, M. Richardson, V. Megalooikonomou, “Seizure detection using EEG and ECG signals for computer-based monitoring, analysis and management of epileptic patients”, *Expert Syst. Appl.* 42 (6) (2015) 3227-3233, ISSN 0957-4174, <http://dx.doi.org/10.1016/j.eswa.2014.12.009>.
- [58] D. Piper, K. Schiecke, L. Leistritz, B. Pester, F. Benninger, M. Feucht, M. Ungureanu, R. Strungaru, H. Witte, “Synchronization analysis between heart rate variability and EEG activity before, during, and after epileptic seizure”, *Biomed. Eng. (Biomed. Technik)* 59 (4) (2014) 343-355.
- [59] M. Valderrama, C. Alvarado, S. Nikolopoulos, J. Martinerie, C. Adam, V. Navarro, M.L.V. Quyen, “Identifying an increased risk of epileptic seizures using a multi-feature EEG-ECG classification”, *Biomed. Signal Process. Control* 7 (3) (2012) 237-244, ISSN 1746-8094, <http://dx.doi.org/10.1016/j.bspc.2011.05.005>.
- [60] P. Zarjam, J. Epps, F. Chen, N.H. Lovell, “Estimating cognitive workload using wavelet entropy-based features during an arithmetic task”, *Comput. Biol. Med.* 43 (12) (2013) 2186-2195, ISSN 0010-4825, <http://dx.doi.org/10.1016/j.combiomed.2013.08.021>.
- [61] P. Antonenko, F. Paas, R. Grabner, T. van Gog, “Using electroencephalography to measure cognitive load”, *Educat. Psychol. Rev.* 22 (4) (2010) 425-438, ISSN 1040-726X, <http://dx.doi.org/10.1007/s10648-010-9130-y>.
- [62] J.M. O’Toole, B. Boashash, “Time-frequency detection of slowly-varying periodic signals with harmonics: methods and performance evaluation”, *EURASIP J. Adv. Signal Process.* 2011 (2011) article ID 193797, 16 pp.
- [63] F. Miwakeichi, E. Martínez-Montes, P.A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, Y. Yamaguchi, “Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis”, *NeuroImage* 22 (3) (2004) 1035-1045, ISSN 1053-8119, <http://dx.doi.org/10.1016/j.neuroimage.2004.03.039>.
- [64] J.C. Mosher, S. Baillet, R.M. Leahy, “EEG source localization and imaging using multiple signal classification approaches”, *J. Clin. Neurophysiol.* 16 (3) (1999) 225-238.

GETTING STARTED WITH A PRACTICAL AND EFFICIENT TIME-FREQUENCY TOOLBOX TFSAP-7.0⁰

17

INTRODUCTION AND OVERVIEW

The focus of this chapter is the practical application of the time-frequency (t,f) algorithms described in the book to simulated and real signals, using an advanced flexible platform for Time-Frequency Signal Analysis and Processing (TFSAP). This platform, known as the (t,f) toolbox TFSAP 7.0, provides two options to the user: either to use the available interactive graphical user interface (GUI), or to use online commands in the user's own code. Updates and improvements made to this TFSAP (t,f) toolbox are inline with the latest research done in the (t,f) field. The TFSAP toolbox has proved since its inception to be an efficient and popular tool for analyzing nonstationary signals. Previous versions provided users with an array of tools to analyze nonstationary signals accurately and efficiently [1]. The updated TFSAP (t,f) toolbox extends the functionality of previous versions. It provides additional options such as generating new (t,f) distributions, and synthesizing a signal from a (t,f) image or specifications. Several demo scripts are included in the current version to demonstrate the functionality of the toolbox and to coach new users to use the TFSAP toolbox for advanced signal processing applications dealing with nonstationarities. The version current at the time of publication is renamed TFSAP 7.0 so as to coincide with the second edition of this book. It can be downloaded for free as a service to the community. The toolbox is intended to work in a MATLAB/Octave environment and it includes several databases of real signals, including EEG and other medical data, as an encouragement to foster further research in biomedicine and related fields.

⁰Author: **B. Boashash**, Qatar University, Doha, Qatar; University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia (boualem@qu.edu.qa).

17.1 INTRODUCTION AND (t, f) TOOLBOX DESCRIPTION

This current updated version TFSAP 7.0 extends and enhances previous TFSA versions continually updated since the first public release in 1987 at the inaugural ISSPA conference in Brisbane, Australia.

In terms of terminology, TFSA refers to the name of the previous versions of the (t, f) toolbox; TFSAP refers to the versions distributed at the time of publication in 2015 and after. Version 7.0 is simply the 1st upgrade made in 2015. This version is provided for free as supplementary material along with several real data sets.

17.1.1 FEATURES AND CONTENTS

17.1.1.1 Standard Original Features of the (t, f) Toolbox

The previous versions of this toolbox¹ broadly addressed the issues of nonstationary signal and noise generation, quadratic time-frequency distributions (QTFDs), multilinear (t, f) analysis, instantaneous frequency (IF) estimation, and tools for data visualization and plotting [1]. For QTFDs, the algorithms included were the Wigner-Ville distribution (WVD), smoothed Wigner-Ville distribution (SWVD), cross-Wigner-Ville distribution (XWVD), spectrogram, Rihaczek-Margenau, exponential distribution (also known as Choi-Williams distribution), and Born-Jordan and Zhao-Atlas-Marks distributions. Multilinear (t, f) analysis is also available with the polynomial WVD of 4th- and 6th-order kernels. In addition, time-scale analysis includes algorithms for wavelet coefficients computation [2]. For IF estimation, a substantial suite of algorithms was also included. These are: the general phase difference, weighted phase difference, adaptive, zero-crossing, least square polynomial fit, peak of spectrogram, peak of WVD [3], and peak of polynomial WVD IF estimators (Section 5.4).

To facilitate the use of TFSA tools and utilities, a GUI was developed in MATLAB with computationally intensive routines written in C for efficient implementation. For MATLAB and C interfacing, MEX file formatting is used. The previous package of TFSA was written to run only on 32-bit Windows or Linux operating system platforms.

A comprehensive user guide and tutorial was also provided as part of the (t, f) toolbox; this is briefly described in the next sections. This tutorial and user guide was designed to give an overview of the features and capabilities of the toolbox as well as provide a step-by-step procedure from installation to usage. The version (TFSAP 7.0) extends this to 64-bit Windows and 64-bit Linux platforms.

17.1.1.2 Features Introduced in the Updated (t, f) Toolbox

Although (t, f) -based signal analysis is a mature research area, a continual improvement is ongoing for the development of more accurate and efficient algorithms/approaches. In particular, an intensive research effort has been done for designing and implementing high-resolution QTFDs as they have proven to be very effective and allowed improvements in terms of resolution and accuracy in the analysis and characterization of nonstationary and/or multicomponent signals. Although the

¹To the best of the author's knowledge, the core of the earliest (t, f) toolbox was developed by the author during the course of his earlier research in the period 1978-1982, while working at Elf-Aquitaine, France, then upgraded during his stay in INSA, Lyon, France and finally completed it as an interactive Fortran package at the University of Queensland, Australia, in the mid-1990s with the assistance of several PhD students. It was first released publicly as part of the conference ISSPA-1987 in Brisbane, Australia, and then continually upgraded since then as a MATLAB toolbox. Several hundred educational and research institutions obtained this toolbox over the years, with or without source code. (See user manual and supplementary material for more details.)

quadratic class of time-frequency distributions (TFDs) offers good temporal and spectral resolutions simultaneously, these transforms suffer from cross-terms when analyzing multicomponent signals. These cross-terms that result because of the quadratic nature of the transforms may be useful in some applications, such as classification, as they provide extra detail for the recognition algorithm; however, in some other applications, these cross-terms need to be reduced by carefully choosing a smoothing function. This is achieved at the expense of resolution.

These quadratic TFDs were considered in earlier chapters as smoothed versions of the WVD. The choice of a particular TFD depends on the specific application at hand and the properties that are desirable for that application. As a result, many new members of the quadratic class of TFDs have been designed, and the list is still increasing. So there was a need to update the TFSA package. In addition to the standard (t, f) methods listed in the previous section, the new version TFSAP 7.0 includes the following advanced TFDs:

1. B-distribution (BD) and modified B-distribution (MBD) (see [Sections 2.7.5](#) and [5.7](#) of the book),
2. Extended modified B-distribution (EMBD) (see [Section 2.7.5](#) and Ref. [4]),
3. Compact support kernel TFD (CSK) (see [Section 3.3.4](#) and Ref. [4]),
4. Extended compact support kernel TFD (CKD) (see [Section 3.3.4.2](#)),
5. Multidirectional TFD (MDD) (see [Sections 3.3.4.3](#) and [5.9](#) and Ref. [5]),
6. Enhanced spectrogram (or S-method) (see [Sections 2.3.2](#) and [3.3.4.1](#))

and several others (see the menu of TFSAP 7.0 for full details).

More importantly, this (t, f) toolbox allows the user to define new data-dependent TFDs suitable for the application considered (see details in [Section 17.5](#)).

17.1.2 TEST SIGNALS UPDATE

Researchers use (t, f) methods as an effective tool in many new applications, such as in the fields of biomedical, speech, radar, and sonar signal processing. This presented a greater need to improve the TFSA toolbox to include several test signals with the above-mentioned quadratic TFD-based tools to help researchers perform advanced (t, f) -based analysis on test data relevant to their needs.

More particularly, in the biomedical field, (t, f) methods have shown great potential in a number of applications involving, e.g., electroencephalogram (EEG), heart rate variability (HRV) signals, and fetal movements (FetMov) to mention just a few [4] (see [Chapter 16](#) for details). EEG, HRV, and FetMov abnormalities analysis using TFDs is an important research area, and these datasets are therefore desirable in a toolbox. TFSAP 7.0 contains several such data sets.

Synthesis of a time-domain signal from a given TFD or (t, f) image also has important applications in signal processing. By using the signal-synthesis property, any signal can be reconstructed from TFDs or from a (t, f) image that has a minimum of desirable properties (see [Sections 11.2](#) and [11.5](#)), allowing applications such as denoising, (t, f) filtering, and signal enhancement (see [Section 11.6](#), p.678).

17.1.3 STANDARD FORMULATIONS AND THEIR IMPLEMENTATIONS

A few basic definitions and formulas are reproduced here from earlier chapters for convenience and completeness so that the reader can find in a nutshell what is implemented in the (t, f) toolbox.

Given a real signal $s(t)$, we first form its analytic associate $z(t)$ as follows (see Eq. [\(1.2.7\)](#), p.44):

$$z(t) = s(t) + j\mathcal{H}\{s(t)\}, \quad (17.1.1)$$

where \mathcal{H} is the Hilbert transform (see [Section 1.2.3](#)). We then define the QTFD as per Eq. (3.2.13), p.113, i.e., as:

$$\rho_z(t,f) = \underset{tf}{W_z(t,f)} \ast \gamma(t,f), \quad (17.1.2)$$

where $\gamma(t,f)$ is a 2-D smoothing kernel and $W_z(t,f)$ is the WVD defined in [Chapter 2](#) as (see Eq. (2.1.17), p.69):

$$W_z(t,f) = \int_{-\infty}^{\infty} z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) e^{-j2\pi\tau f} d\tau = \mathcal{F}_{\tau \rightarrow f} \{z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2})\}. \quad (17.1.3)$$

The equivalent discrete-time QTFD of Eq. (17.1.2) is given in [Section 6.1](#) and Refs. [4,6] as:

$$\rho[n,k] = W[n,k] \underset{n}{\circledast} \underset{k}{\circledast} \gamma[n,k] |_{k=0,1,\dots,N-1} \quad (17.1.4)$$

with

$$W[n,k] = \sum_{m=0}^{2N-1} K[n,m] e^{-j2\pi mk/2N} = \text{DFT}_{m \rightarrow k} \{K[n,m]\}, \quad (17.1.5)$$

where \circledast is the circular convolution; N , n , k , and m are the signal length, discrete time, discrete frequency, and discrete lag, respectively; $W[n,k]$ is the discrete WVD, and $K[n,m]$ represents a discrete-time version of $K(t,\tau) = z(t + \tau/2) z^*(t - \tau/2)$ as detailed in [Chapters 2 and 6](#).

17.2 TECHNICAL CONTENT OF THE (t, f) TOOLBOX TFSAP 7.0

17.2.1 LIST OF UPDATED FUNCTIONS IN VERSION 7.0

In order to update the features of the previous versions of TFSA toolbox, the following options are included in the 2015 version renamed TFSAP 7.0¹:

- 1. Generation of test signals and noise such as:** linear frequency modulation (FM), quadratic FM, cubic FM, stepped FM, sinusoidal FM, hyperbolic FM and Gaussian and uniform noise; others to appear in future versions.
- 2. Generation of demo signals (`signal1`, `whale1`, `bat1`, `eeg1`, `hrv1`):** `signal1` is an arbitrary test signal; `whale1` is a whale signal; `bat1` is a bat signal; `eeg1` is an EEG signal; and `hrv1` is an HRV signal. Other signals such as `FetMov`, `bird`, and `speech` are also included in this and future upgrades.
- 3. Generation of TFDs such as:** WVD, SWVD, spectrogram, Rihaczek-Margenau-Hill distribution, Windowed-Rihaczek-Margenau-Hill distribution, exponential distribution, B-distribution, MBD, EMBD, CSK-based distributions, extended CSK-based distributions, Born-Jordan distribution, Zhao-Atlas-Marks distribution, XWVD, polynomial Wigner-Ville distribution (order-6 kernel), polynomial Wigner-Ville distribution (order-4 kernel), and ambiguity function (AF); S-method; MDD; others to appear in future upgrades.

¹Next updates will be uploaded on a regular basis on several websites, including the author's webpage, Elsevier book website, The University of Queensland espase, Qatar University Qspace, and other relevant websites associated with the author.

4. **Generation of time-scale signal representations:** Wavelet transforms (WTs) and scalogram; others to be added in future upgrades.
5. **IF estimation algorithms:** Finite-phase difference including: first-order (FFD), second-order (CFD), fourth-order and sixth-order, weighted-phase difference, zero crossing, adaptive least mean square (LMS), adaptive Recursive Least Squares, least squares polynomial coefficients, peak of the spectrogram, peak of the WVD, and peak of the polynomial WVD; other options to be added in future updates.
6. **Signal synthesis using the following TFDs:** Short-time fourier transform (STFT), spectrogram, and WVD.
7. **Other digital signal processing (DSP) tools:** Analytic signal calculation, power spectrum, and previously mentioned demo signals (e.g., signal1, whale1, bat1, eeg1 and hrv1).
8. **Data visualization routines including:** plot, image, pseudocolor, waterfall, mesh, surf, contour, and tfsapl.
9. **The current version of TFSAP toolbox runs on the following four different operating system platforms:** 32-bit Windows, 64-bit Windows, 32-bit Linux, and 64-bit Linux; others to appear in future upgrades.

17.2.2 TFSAP DEVELOPMENT PROCESS

TFSAP has been developed in MATLAB and C as an extension of a previous Fortran and C toolbox that were themselves an extension of an APL version.² The GUI of TFSAP is developed in MATLAB, while the computationally intensive scripts are written in C for efficient implementation. The MATLAB and C code communicate with each other using MEX file formatting.

Adding a new function to TFSAP requires the coding of source files in C to perform the required function. Then TFSAP GUI files are edited to provide access to the new function. For example, to add the EMBD in TFSAP requires editing the source code of the quadratic function in C, as the EMBD is a member of the quadratic class of TFDs. Afterward, the source code of the quadratic TFD analysis GUI is updated to incorporate the functionality of the EMBD in the GUI of the quadratic (t,f) analysis. This task is performed by the (t,f) toolbox maintainers.

17.3 USER'S GUIDE FOR (t, f) TOOLBOX TFSAP-7.0

This section takes a step-by-step approach to the understanding and use of this (t,f) toolbox. It also has an adjunct purpose in that it familiarizes the user with TFSAP 7.0 for MATLAB. TFSAP 7.0 has been written to run on a platform supporting MATLAB. A version supporting OCTAVE is under development, and it is planned to be part of the next version, TFSAP 7.1.

17.3.1 INSTALLATION

The MATLAB (t,f) toolbox TFSAP 7.0 can be downloaded from several links, including www.time-frequency.net, the Elsevier book website, <http://espace.library.uq.edu.au/view/uq:211321>, and the author's personal page. The corresponding code is described in Ref. [9] and Chapter 6.

²The main code of the APL version appears in the Appendix in the 1982 PhD thesis of the author. The first working version was published by the author in an ELF-Aquitaine company report dated September 1978 [7] and reported in Ref. [8].

The versions for Linux and MS-Windows are distributed pre-compiled. To install them, just extract the zipped file and follow the instructions in the README text file. To run TFSAP 7.0, you need to have MATLAB already installed. To run the TFSAP 7.0 GUI, a minimum system requirement of 32 megabytes RAM on PCs is needed to avoid memory problems.

17.3.2 RUNNING TFSAP 7.0

To run TFSAP 7.0 on MS-Windows, all you need to do is to click the MATLAB icon on the screen and then type: `tfsa7`

at the MATLAB prompt.¹ If you are running TFSAP 7.0 on UNIX workstations, invoke MATLAB from the shell by typing: `matlab`

(followed by the “Enter” key), and then type: `tfsa7`

at the MATLAB prompt. The Main Menu of the GUI will then appear as shown in Fig. 17.3.1. The GUI of TFSAP 7.0 contains the pop-up menus and interface fields which control the various functions and tools available in the toolbox. The GUI provides an alternative to typing TFSAP 7.0 commands on the MATLAB command line, to make it easy for beginners.

It contains four buttons. These are:

- Welcome: Clicking this button gives general information on TFSAP-7.0.
- Copyright: The license agreement screen appears by clicking this button.

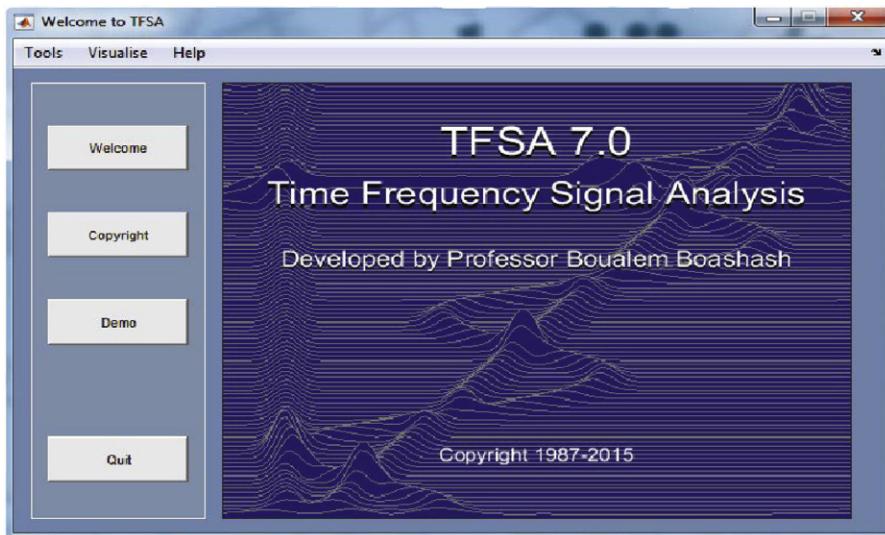


FIGURE 17.3.1

Main menu of GUI of TFSAP 7.0 toolbox

¹If you use an earlier version, e.g., TFSA 5.5 or 6.0, then the number seven changes to, e.g., five for version 5.5 or six for version 6.0 (in this case, consult the manual accompanying the software).

- Demo: Clicking this button gives a demonstration of TFSAP-7.0.
- Quit: Clicking this button aborts the TFSAP session.

The TFSAP 7.0 menu bar is composed of three major menu items: Tools, Visualize, and Help. Any menu item can be opened by placing the mouse's pointer on the corresponding item and then clicking the primary button of the mouse. You can access any menu sub-item by clicking with the primary button of the mouse on the corresponding menu item and while holding this button down, dragging the mouse across the submenu bar until encountering the required subitem, then releasing the mouse button.

17.3.3 DATA ACCESSES AND SIGNAL GENERATION GUI

[Figure 17.3.2](#) shows the main GUI through which different types of nonstationary signals can be generated. [Figure 17.3.2](#) shows an example of generating a linear FM signal with 128 samples, starting and ending frequencies of 0.1 and 0.4 Hz. There is also an option to generate the real or analytic form of a test signal. The complex analytic signal (see [Chapter 1](#)) can be generated by adding to the time-domain signal an imaginary part using the Hilbert transform. In TFSAP, the analytic signal is generated using an equivalent frequency domain method (see [Chapters 1 and 6](#)).

The different types of test signals that can be generated from TFSAP 7.0 are listed in [Section 17.2.1](#). There is also an option to generate noise only or a test signal with noise. An illustration to generate a quadratic FM noisy signal is shown in [Fig. 17.3.3\(a\)](#). A Gaussian noise of signal-to-noise ratio (SNR) 20 dB is added with the signal. The length of the signal is 128 samples with starting frequency of 0.1 Hz and ending frequency of 0.4 Hz. As previously mentioned, TFSAP 7.0 also includes several real-world nonstationary signals. These can be generated by selecting Demo Signal as signal type. Then, examples of demo signals will appear in the Demo Signal tab as shown in [Fig. 17.3.3\(b\)](#).

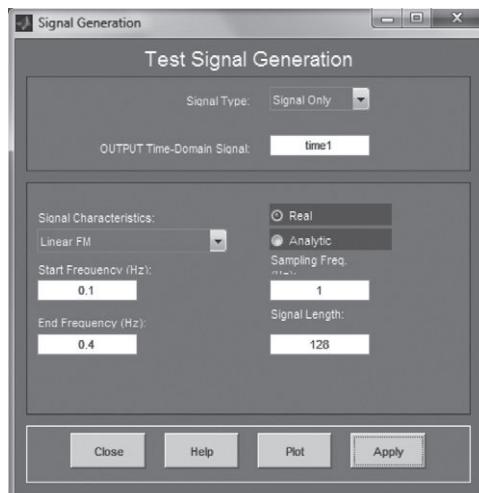
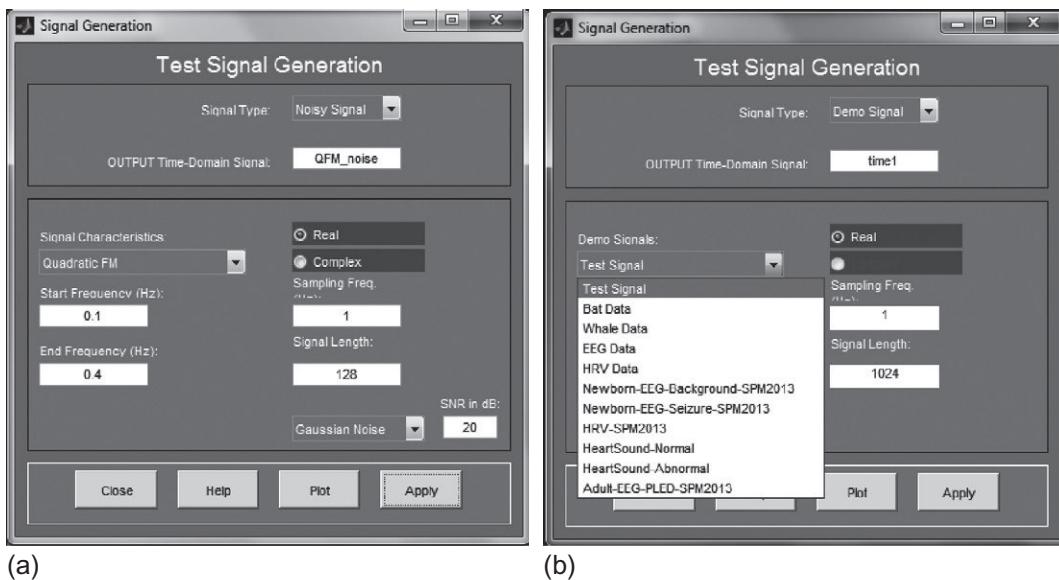


FIGURE 17.3.2

Signal generation GUI.

**FIGURE 17.3.3**

Generating test signals: (a) a quadratic FM signal with noise; (b) a demo signal.

All references to “signal” or “array” names in TFSAP 7.0 refer to variables in the MATLAB workspace. Values are read from and written to this space, and can be accessed normally from the MATLAB command line.

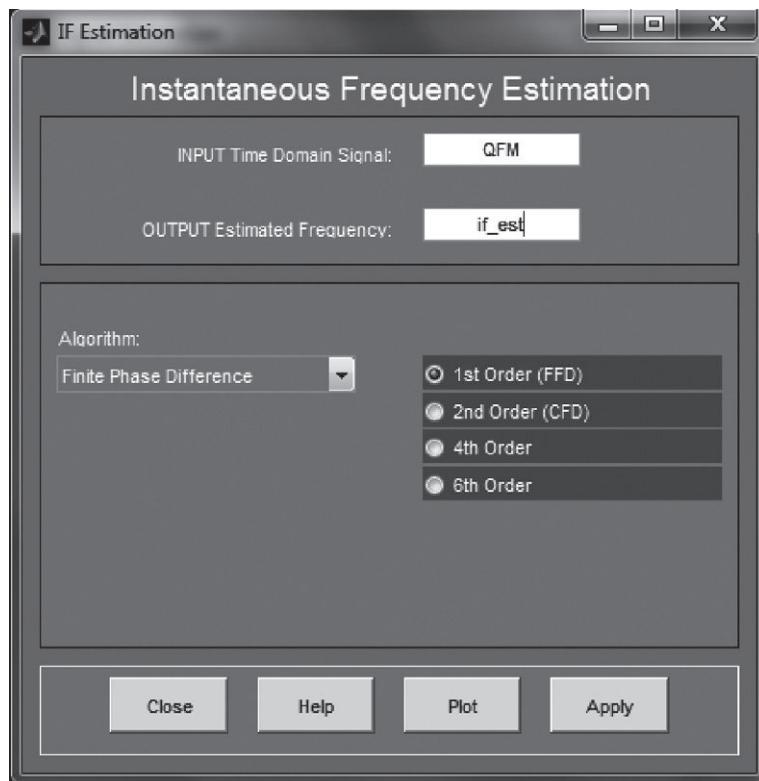
Thus two signals can be added together using the following procedure:

1. Generate time1 from the GUI Signal Generation popup.
2. Generate time2 from the GUI Signal Generation popup.
3. In the MATLAB command-line window, type: `time3 = time1 + time2`
4. Then, time3, the sum of the two signals, appears in the workspace and it can now be accessed from the GUI.

17.3.4 IF ESTIMATION GUI

As described in the previous chapters, IF estimation is an important technique in (t,f) analysis (see [Chapter 1](#) and Ref. [3]). There are several algorithms that can be used to compute IF, e.g., differentiation schemes, adaptive techniques, peaks of TDS, and zero-crossings (for more details, see Chapter 10 and Ref. [10]). The detailed list of functions for IF estimation is given in [Section 17.2.1](#). For any signal, the user needs to perform the following tasks for IF estimation:

- Generate a time-domain signal as explained in [Section 17.3.3](#).
- Open the IF estimation GUI.
- Enter the name of the signal in the Input Time-Domain Signal textbox

**FIGURE 17.3.4**

GUI for IF estimation.

- Enter the name of the output in the Output Estimated Frequency textbox
- Select any IF estimation method from the list of methods shown in the algorithm tab and then click the Apply button to generate the IF.

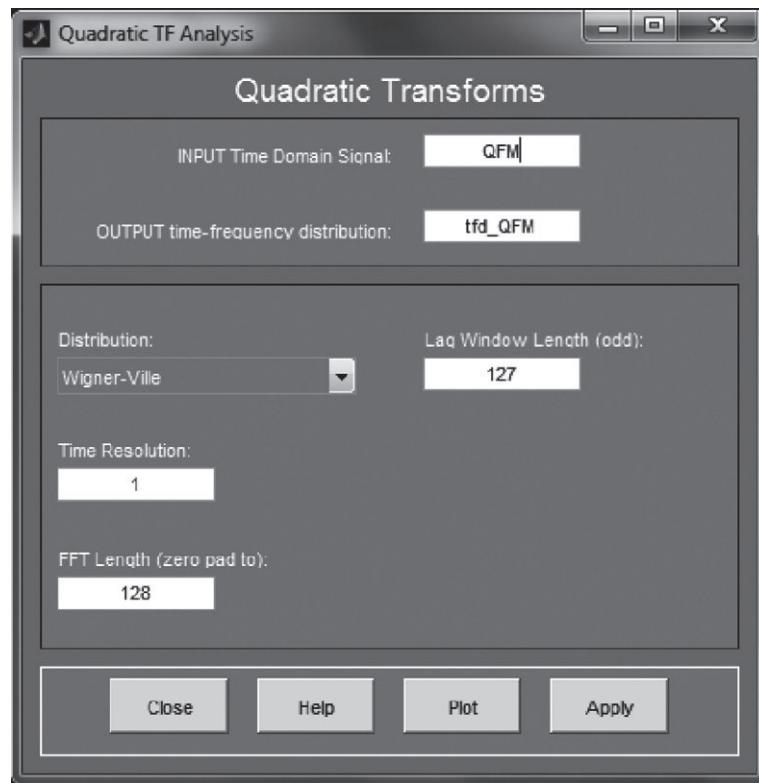
Figure 17.3.4 shows an example of the IF estimation GUI for a quadratic FM signal.

17.3.5 QUADRATIC TFD ANALYSIS GUI

17.3.5.1 Using an Existing Time-Frequency Method

In TFSAP 7.0, a GUI is included that gives a menu with the option to generate different TFDs for any signal. For analyzing the quadratic TFD of any signal, the user needs to perform the following tasks:

- Generate a time-domain signal as explained in [Section 17.3.3](#).
- Open the quadratic TF analysis GUI.
- Enter the name of the signal in the Input Time Domain Signal textbox
- Enter the name of the output in the Output TFD textbox

**FIGURE 17.3.5**

GUI for quadratic TF analysis.

- Select any quadratic TFD from the list shown in the distribution tab
- Set the FFT length and lag window length and press apply to generate quadratic TFD (see Chapter 6 for more details).

Figure 17.3.5 shows an example of estimating the WVD of a quadratic FM signal using the quadratic (t,f) analysis GUI.

17.3.5.2 User-Defined TFD

If the user wants to define and test a new kernel, he can do so following the procedure described below:

1. Define the kernel $\gamma[n,k]$ in the MATLAB line command as per Eq. (17.1.4).
2. Generate the WVD $W_z[n,k]$ of the signal z as described above.
3. Perform a double discrete convolution product between $W_z[n,k]$ and $\gamma[n,k]$ (in n and k) using the MATLAB command `conv2` (see Eq. (17.1.2)). Then, the result of the above operation is the TFD obtained using the kernel $\gamma[n,k]$ defined in step 1.
4. Proceed to plot the above result or perform any other desired operation.

For more details, see Sections 17.1.3 and 17.5.2.

17.3.6 SIGNAL SYNTHESIS GUI

A time-domain signal $z[n]$ can be synthesized from a given TFD $\rho_z[n, k]$ under some conditions (see Sections 2.3.1, 3.1.1, and 11.2) or from any (t, f) image. TFSAP 7.0 includes a GUI to help users explore and use the different features of signal synthesis in applications such as (t, f) filtering and signal enhancement. The main GUI for signal synthesis is shown in Fig. 17.3.6. To reconstruct a discrete time-domain signal $z[n]$ from a given discrete (t, f) image $\rho_z[n, k]$, the user needs to perform the following tasks:

- Generate a desired (t, f) image (e.g., the modified TFD of a particular signal).
- Open the signal synthesis GUI.
- Enter the name of the discrete (t, f) image in the Input TFD textbox.
- Enter the name of the output in the Output time domain signal textbox.
- Select any synthesis method along with the lag window length and then click Apply to generate a time-domain signal.

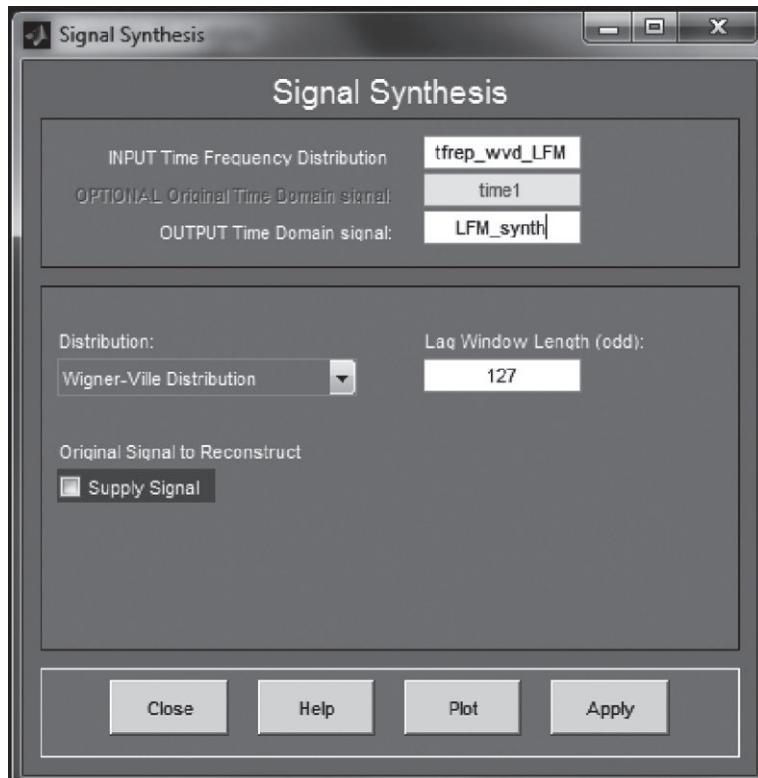


FIGURE 17.3.6

GUI for signal synthesis (note that the input here is similar to the output of the quadratic transform shown in Fig. 17.3.5).

17.3.7 TFSAP-MATLAB FLEXIBLE INTERFACE

17.3.7.1 Compatibility with Other MATLAB Functions

The GUI is an interface to the TFSAP 7.0 functions. All normal MATLAB features remain in place. For example, titles, labels, figure properties and variables may all be changed using the MATLAB command-line interface. Figures can be printed using the MATLAB's print command, as normal.

17.3.7.2 MATLAB Command-Line Interface

TFSAP 7.0 functions can be accessed in the same manner as the standard MATLAB functions (see Sections 17.4.3.3, 17.4.4, and 17.5). For a list of functions in TFSAP 7.0 type: `help tfsa7`

Help on any of the functions can be obtained by typing: `help <functionname>`. A detailed list of functions that TFSAP 7.0 supports is given in Section 17.2.1.

17.4 A BRIEF (t, f) TUTORIAL USING THE (t, f) TOOLBOX

This section illustrates the usability of the updated TFSAP 7.0 toolbox for real-life applications. It is a concise version of a tutorial that is part of the (t, f) toolbox.

17.4.1 RESULTS FOR SIMULATED SIGNALS

17.4.1.1 Results for a Quadratic FM Signal

A comparison of TFDs for a quadratic FM signal is shown in Fig. 17.4.1. The frequency of quadratic FM signal (length = 128 samples) ranges from 0.1 to 0.4 Hz. Figure 17.4.1(a) shows the WVD; Fig. 17.4.1(b) shows the exponential distribution with sigma 20; Fig. 17.4.1(c) shows the spectrogram with a smooth “hamm” window and window length of 15 samples; and Fig. 17.4.1(d-f) shows the B-distribution with $\beta = 0.1$ for a quadratic FM signal, then the MBD, followed by the EMBD. All the TFDs are generated with lag window length of 15 samples, time-resolution of 1, and length of FFT 128 samples.

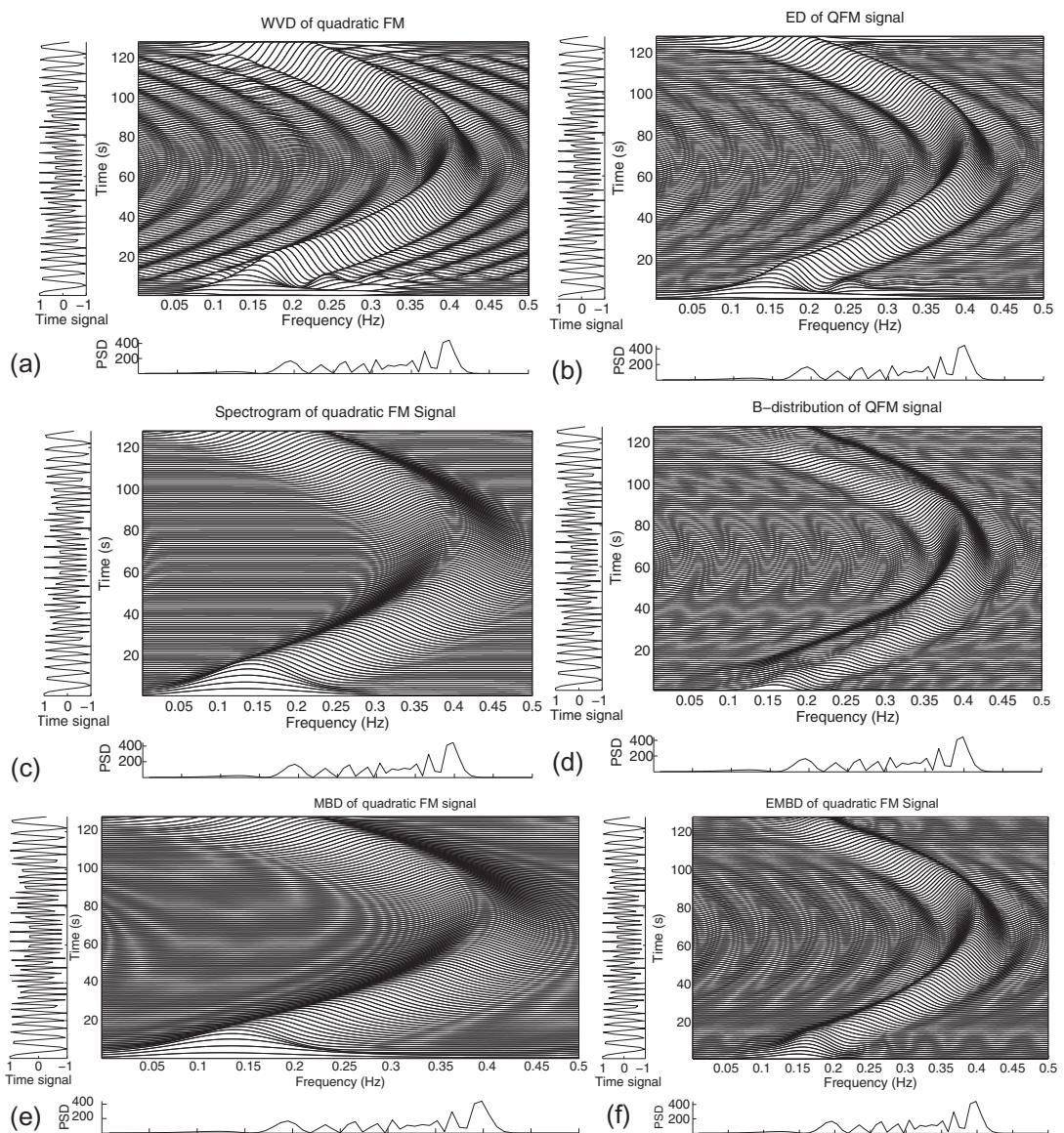
17.4.1.2 Results for IF Estimation of Quadratic FM Signals

Figure 17.4.2 illustrates the IF estimation of a Quadratic FM signal using three different methods; i.e., zero-crossing, phase differentiation, and peaks of WVD, all available from TFSAP 7.0. For more advanced high-resolution TFDs, such as the MBD or EMBD, the peak extraction method can yield an IF estimation with different properties. The user is encouraged to perform this comparison.

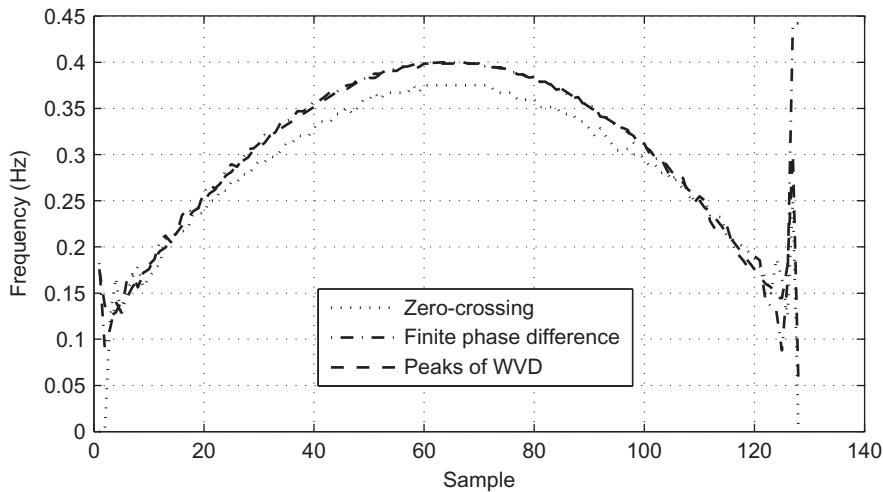
17.4.2 RESULTS FOR REAL SIGNALS

17.4.2.1 Results for a Bat Signal

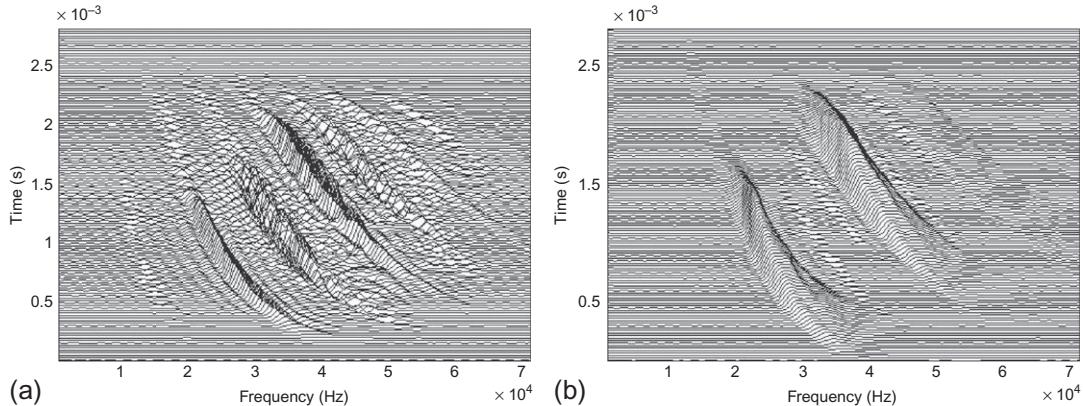
Several real demo signals are included in the toolbox, i.e., bat signal, whale signal, EEG and HRV signals. A comparison using two different quadratic TFDs on a bat signal sampled at 142 kHz is shown in Fig. 17.4.3. Figure 17.4.3(a) shows the WVD of the bat signal with a lag window length of 63 samples and a time-resolution of 3. Figure 17.4.3(b) shows the modified B-distribution of the bat signal with a lag window length of 63 samples, a time-resolution of 3, and $\beta = 0.01$. From Fig. 17.4.3, it can be observed that the bat signal is multicomponent with a decreasing nonlinear FM. It can also be observed that the MBD performs much better in suppressing the cross-terms. (Note that increasing the value of the time resolution results in a less detailed figure, but with improved visual clarity as the

**FIGURE 17.4.1**

Comparison of standard quadratic TFDs for a quadratic FM signal: (a) Wigner-Ville distribution; (b) exponential distribution; (c) spectrogram; (d) B-distribution; (e) MBD; (f) EMBD.

**FIGURE 17.4.2**

IF estimation of a quadratic FM signal. (Note that the results of the phase differentiation and peaks of WVD methods are very close in this example.)

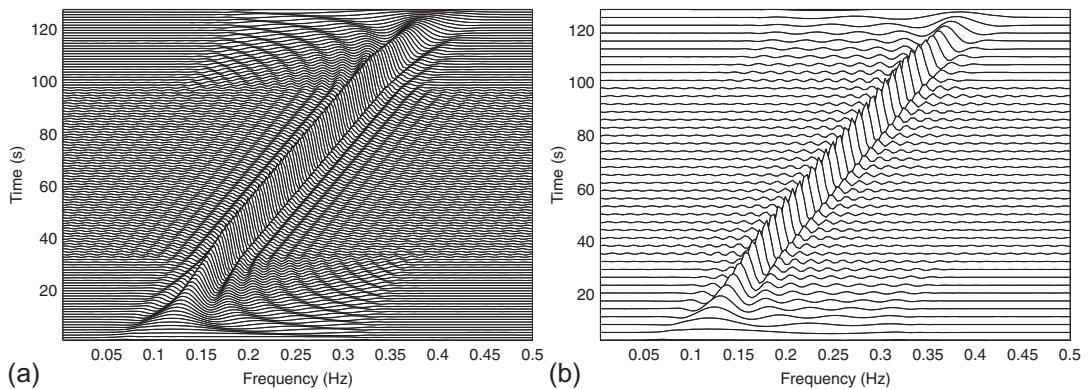
**FIGURE 17.4.3**

Comparison of WVD and MBD for bat echolocation signal: (a) WVD of a bat signal; (b) MBD of a bat signal.

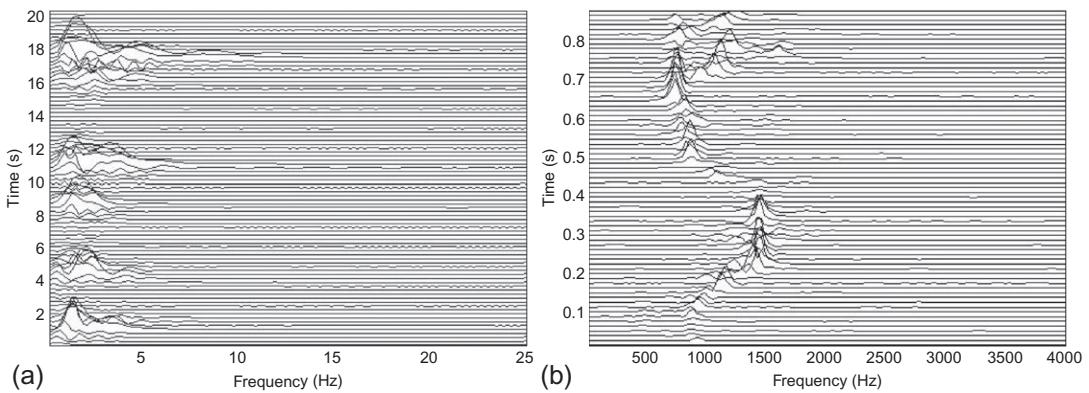
number of time lines decreases. This is demonstrated in Fig. 17.4.4 where a time resolution of 1 was used in Fig. 17.4.4(a) and a time resolution of 3 was used in Fig. 17.4.4(b) for the WVD of a linear FM signal.)

17.4.2.2 Results for an EEG Signal

EEG signal analysis based on quadratic TFDs is illustrated in Fig. 17.4.5. Graph (a) shows the EMBD of an EEG signal. Figure 17.4.5 shows that the EMBD enhances the ridge structure of multicomponent EEG signal, and it significantly reduces the cross-terms [4].

**FIGURE 17.4.4**

WVD of a linear FM signal: (a) time resolution = 1; (b) time resolution = 3.

**FIGURE 17.4.5**

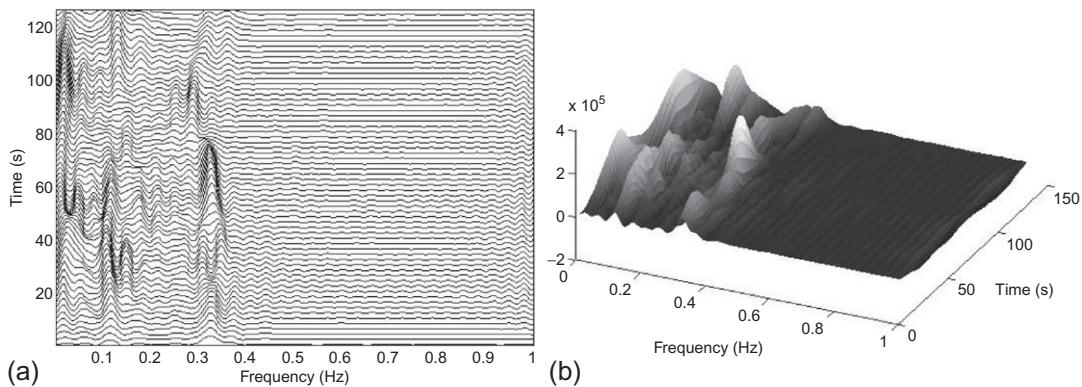
Time-frequency representations of real signals: (a) EMBD of EEG signal; (b) EMBD of signal whale1.

17.4.2.3 Result for a Whale Signal

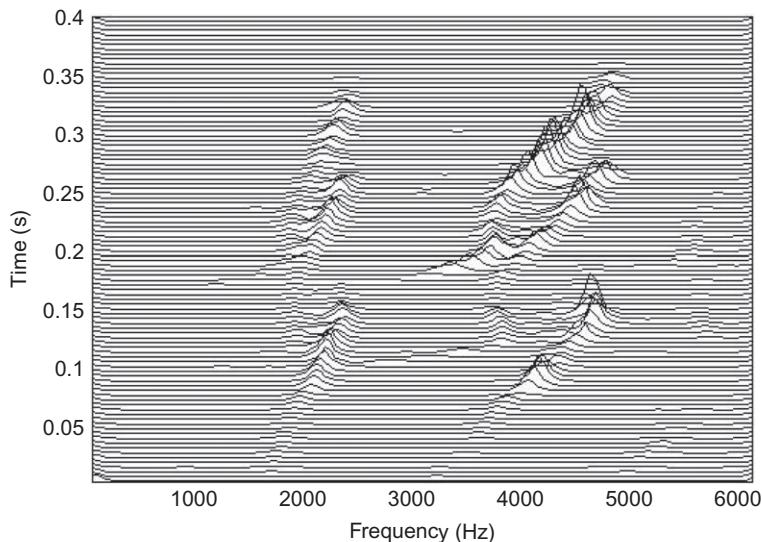
Several whale signals are also included in the TFSAP 7.0 package. The signal whale1 contains 7000 data points and was collected at a sampling rate of 8 kHz. Figure 17.4.5(b) shows the EMBD of the signal whale1. A second whale signal, whale2, presented in Chapter 2, is also included in this toolbox (see Figure 2.1.1).

17.4.2.4 Result for an HRV Signal

HRV signals are formed from electrocardiogram (ECG) signals [4]. An HRV signal with length of 256 samples and sampling frequency 2 Hz is included in the TFSAP 7.0 package. Figure 17.4.6(a) shows the MBD of the HRV signal using the tfsapl function, and Fig. 17.4.6(b) shows the same TFD as in Fig. 17.4.6(b) by using the surf plotting function.

**FIGURE 17.4.6**

MBD of HRV epoch from a healthy adult: (a) graph obtained using `tfsapl` function; (b) graph is plotted using `surf` function with the same MBD parameters as (a): $\beta = 0.02$ and the lag window length is 63. From the figure, three components and their variations can be observed.

**FIGURE 17.4.7**

EMBD of a bird signal.

17.4.2.5 Result for a Bird Signal

This signal is a bird song recorded in Australia; it is shown on the book's cover of the first edition. The signal is sampled at 24417 Hz and is 9767 samples long. This signal, called `bird1`, is included in TFSAP 7.0. [Figure 17.4.7](#) shows the EMBD of the bird signal, obtained with $\alpha = 0.05$, $\beta = 0.5$, a lag window length of 127 samples, a FFT length of 128 samples, and a time resolution of 50. The signal is

downsampled by a factor 2 because after 6 kHz, the signal energy is insignificant. [Section 6.6](#) shows a TFD of the same signal using a separable Hamming/Hanning kernel.

17.4.3 RESULTS FOR SIGNAL SYNTHESIS AND (t, f) FILTERING

17.4.3.1 Signal Reconstruction and Denoising Using the WVD

17.4.3.1.1 Signal reconstruction

An example of signal synthesis using the WVD is shown in [Fig. 17.4.8](#): (a) shows a linear FM signal with starting and ending frequency of 0.1 and 0.4 Hz, respectively with signal length of 127 samples; (b) shows the corresponding WVD of the signal; and (c) shows the synthesized time-domain

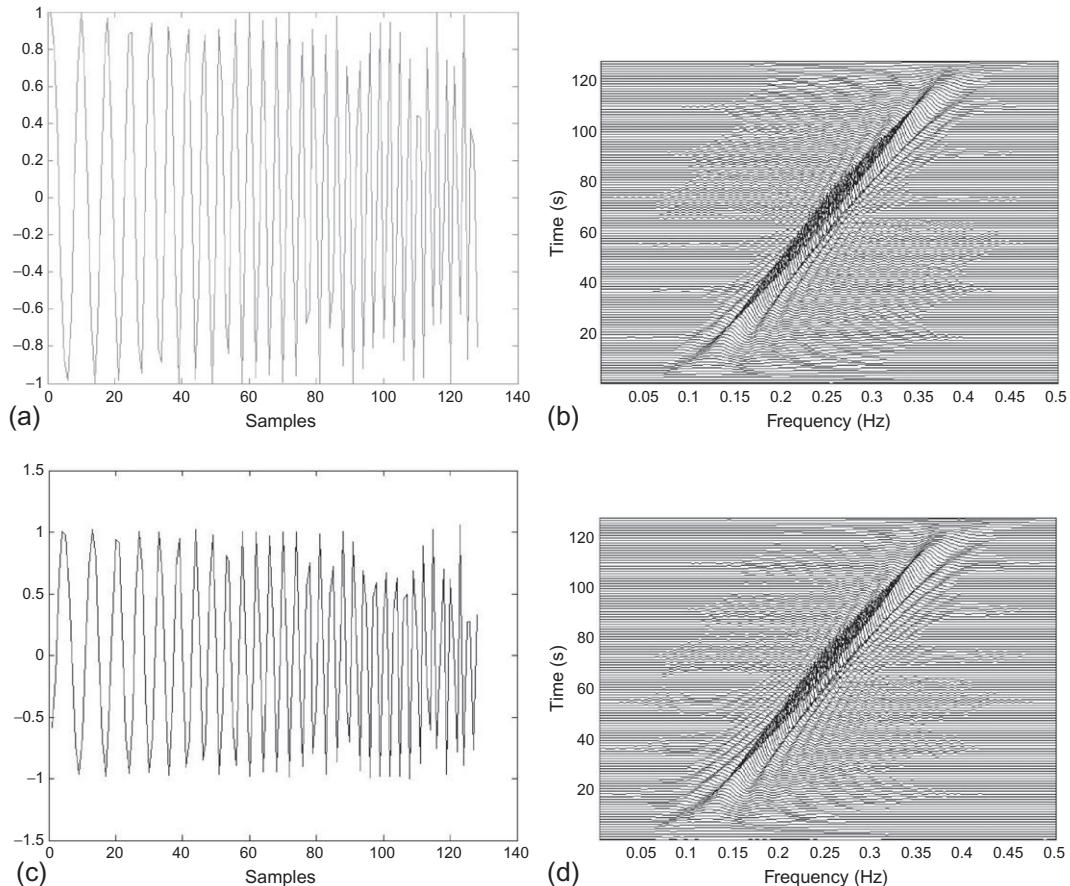


FIGURE 17.4.8

Example of signal reconstruction using the WVD method: (a) signal; (b) WVD of signal; (c) synthesized signal; (d) WVD of synthesized signal.

signal using the WVD method. Graph (d) shows the WVD of synthesized time-domain signal. The results indicate that the reconstructed and original are similar.

17.4.3.1.2 Denoising

[Figure 17.4.9](#) shows the (t,f) filtering of a noisy FM signal with starting and ending frequency of 0.1 and 0.4 Hz, respectively and having a Gaussian noise of 20 dB. This is followed by a signal reconstruction stage. In this example, the following steps are performed to remove noise from this signal:

1. The IF of the signal is estimated by detecting peak locations in the WVD of the noisy signal.
2. The IF estimates are then used to create a binary image such that the value is one along the IF points and zero elsewhere.
3. The binary image is smoothed by convolving it with a 1D Gaussian filter.
4. The smoothed image is then used to mask the WVD.
5. The masked WVD is then “inversed” to yield a reconstructed (t,f) filtered time-domain signal.

[Figure 17.4.9\(a\)](#) shows the time domain representation of the noisy signal. [Figure 17.4.9\(b\)](#) shows the WVD of the noisy signal. [Figure 17.4.9\(c\)](#) shows the filtered TFD of the noisy signal. [Figure 17.4.9\(d\)](#) shows the time-domain representation of the filtered TFD; it shows that (t,f) filtering has significantly reduced the effect of noise and the reconstructed signal closely matches the original one as shown in [Fig. 17.4.8\(a\)](#). In order to show the effect of filtering, the WVD of a synthesized filtered signal is again computed as shown in [Fig. 17.4.9\(e\)](#). By comparing the results of [Fig. 17.4.9\(b\)](#) with [Fig. 17.4.9\(e\)](#), it can be observed that the noise of the synthesized signal is significantly reduced and the WVD of the synthesized signal is closer to the actual WVD of the original signal.

17.4.3.2 Using the STFT and Spectrogram

[Section 2.3.1](#) shows that a signal can be reconstructed exactly from its STFT. The following steps may be performed to remove noise from a signal using the STFT:

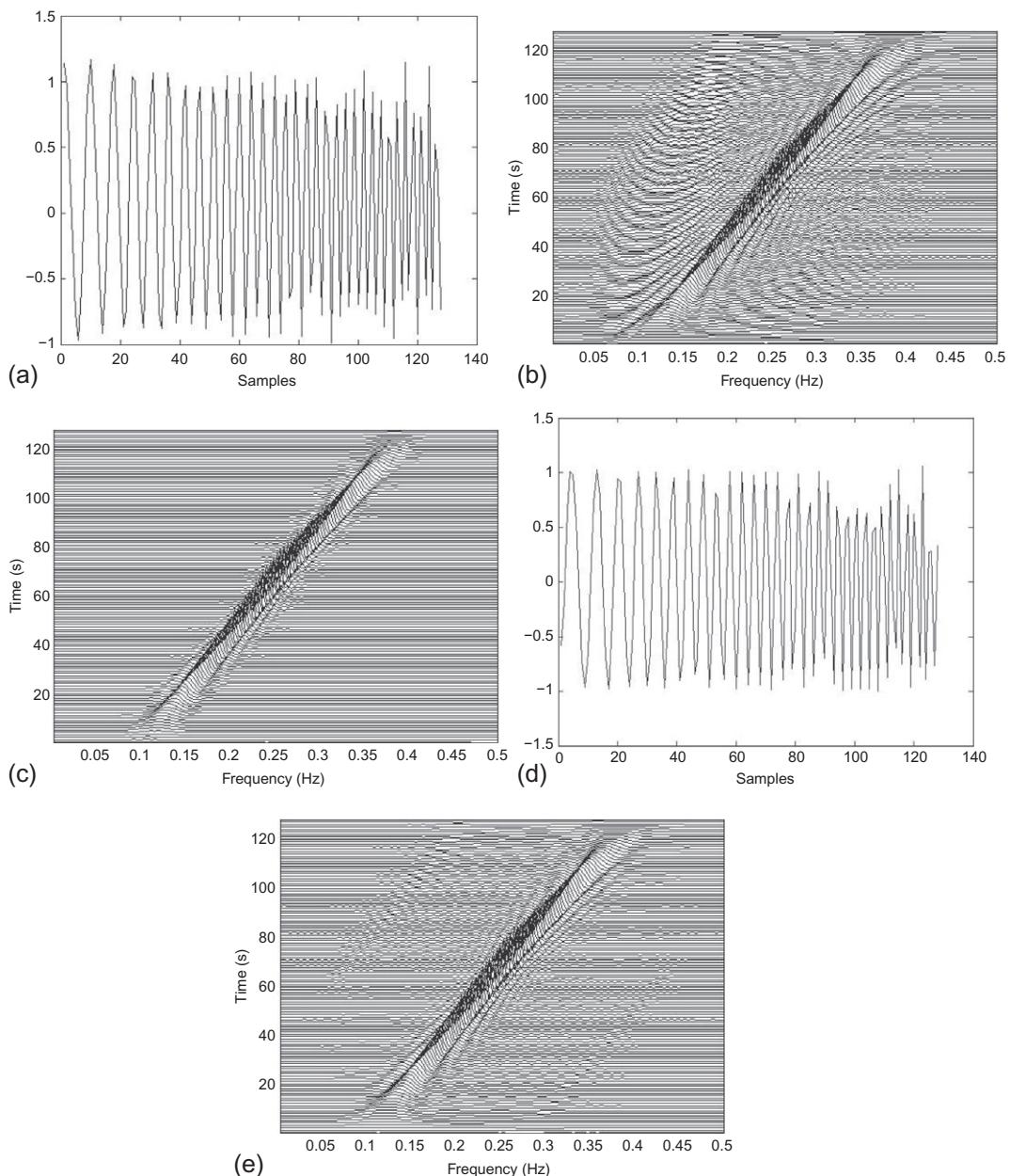
1. compute the STFT of the time-domain signal,
2. apply a mask on the STFT (element-by-element multiplication between the STFT coefficients and the mask),
3. the masked STFT is “inversed” to reconstruct a filtered time-domain signal.

[Sections 11.2](#) and [11.6](#) show several examples of (t,f) signal denoising using the STFT. In particular, [Section 11.2](#) describes in detail how to efficiently implement the inverse STFT, taking into account issues such as window overlap.

17.4.3.3 Time-Frequency Synthesis of a Signal from an Arbitrary Image

Let us present an example of signal synthesis from an arbitrary image describing some desired (t,f) specifications. The procedure involves the following steps:

1. Form the desired (t,f) image.
 2. Import the image in the MATLAB environment, i.e.,
- ```
tfid = double(imread('signall.jpg','jpg'));
```

**FIGURE 17.4.9**

Signal synthesis of a noisy signal using WVD-based TF filtering: (a) signal + noise; (b) WVD of noisy signal; (c) filtered TFD; (d) synthesized signal; (e) WVD of synthesized signal.

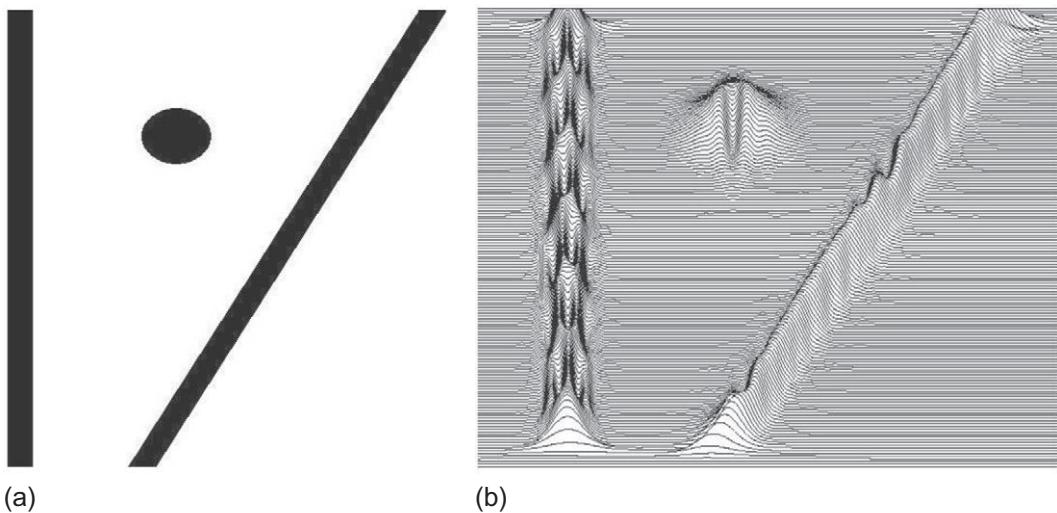
**FIGURE 17.4.10**

Illustration of signal synthesis from an image: (a) image to synthesize; (b) spectrogram of the reconstructed time-domain signal.

3. Transform the image in gray level (note that this step is necessary only for rgb images):  
`tfd = rgb2gray(tfd);`
4. Transpose the image:  
`tfd = tfd.';`
5. Follow the GUI procedure described in [Section 17.3.6](#); e.g., choose the “Modified Spectrogram” technique.
6. Use the following parameters: “Smoothing Window Length”=127 and “Iterative Tolerance Level”=0.45.

[Figure 17.4.10](#) shows the desired  $(t,f)$  image used in this example and the spectrogram of the synthesized time-domain signal. The parameters used to generate the TFD are: a window length of 127 samples, a time resolution of 3, a Hamming smoothing window, and a 512 FFT length.

#### 17.4.4 DUAL USE OF $(t, f)$ TOOLBOX (GUI OR COMMAND LINE)

The toolbox can be used in two different ways: (1) using the GUI as described before; or (2) using the command line. TFSAP 7.0 functions can be accessed from MATLAB in the same manner as the standard MATLAB functions. For a list of functions in TFSAP 7.0, type:

`help <functionname>.`

Further information can be found typing:

`type <functionname>.`

The next section provides more details and an example.

---

## 17.5 EXPANDING AND PERSONALIZING THE TFSAP TOOLBOX

### 17.5.1 PRINCIPLE AND ILLUSTRATION

Using the MATLAB command line, it is possible to use most TFSAP functions in one's personal code. For example, let us present how one can use the user-defined kernel option defined in [Section 17.3.5.2](#). In this example, we use a separable Hamming/Hanning kernel (see [Sections 3.2.5.1](#) and [5.7](#)):

1. Generate a signal  $z$ , e.g., a quadratic FM signal in the range 0.1 and 0.4 Hz:  

$$z = \text{gsig}(\text{'quad'}, 0.1, 0.4, 128, 1);$$
2. Define the kernel  $\gamma[n, k]$ ; in this example, we use an average filter:  

$$\text{gamma} = \text{hann}(11) * \text{hamming}(23)';$$
3. Generate  $W_z[n, k]$  of 128 point signal using a time resolution of 3:  

$$\text{tf} = \text{wvd}(z, 127, 3);$$
4. Calculate the double convolution product between  $W_z[n, k]$  and  $\gamma[n, k]$  to get the TFD associated to the user-defined kernel:  

$$\text{tfnew} = \text{conv2}(\text{tf}, \text{gamma}, \text{'same'})$$
5. Plot the TFD using the `tsaps1` function:  

$$\text{tsaps1}(z, \text{tf});$$

A list of all these command lines are available in the reference guide provided in the  $(t, f)$  toolbox (available as supplementary materials).

### 17.5.2 EXERCISE

Redo the same procedure presented above for the EMBD; define the kernel  $\gamma[n, k]$  corresponding to this TFD.

---

## 17.6 SUMMARY AND CONCLUSIONS

This current TFSAP toolbox 7.0 provides an excellent opportunity for students and researchers to start their journey in the field of nonstationary signal analysis. It allows the users to exploit the wide range of  $(t, f)$  tools in different signal processing application. The version TFSAP 7.0 updates and extends the functionality of previous versions of the TFSAP toolbox. The author intends to continue to update it with his team in order to reflect new ideas and algorithms such as Ref. [11], and to improve functionality. This version TFSAP 7.0 is intended also to be available in Octave and to include more biomedical data for testing and research. It should become available on “[www.time-frequency.net](#)” the author's personal webpage, on the website of this book (under supplementary material), on The University of Queensland espace, on Qatar University Qspace, and on other associated websites.

---

## REFERENCES

- [1] B. Boashash, “Efficient software implementation for the upgrade of a time-frequency signal analysis package”, in: Proc. Eighth Australian Joint Artificial Intelligence Conference (A.I.'95), 1995, pp. 26-31.

- [2] I. Daubechies, “The wavelet transform: a method for time-frequency localization”, in: S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing*, vol. 1, Prentice-Hall, Englewood Cliffs, NJ, 1991, pp. 366-417 (Chapter 8).
- [3] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals”, *Proc. IEEE* 80 (4) (1992) 520-538.
- [4] B. Boashash, G. Azemi, J.M. O’Toole, “Time-frequency processing of nonstationary signals: advanced TFD design to aid diagnosis with highlights from medical applications”, *IEEE Signal Process. Mag.* 30 (6) (2013) 108-119.
- [5] B. Boashash, N.A. Khan, T. Ben-Jabeur, “Time-frequency features for pattern recognition using high resolution TFDs: a tutorial review”, *Digit. Signal Process.* 40 (2015) 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [6] A.P. Reilly, G.J. Frazer, B. Boashash, “Analytic signal generation—tips and traps”, *IEEE Trans. Signal Process.* 42 (11) (1994) 3241-3245.
- [7] B. Bouachache, “Representation temps-fréquence”, Tech. Rep. 373/78, Soc. Nat. ELF-Aquitaine, Pau, France, 1978, 56 pp.
- [8] B. Bouachache, B. Escudie, J.M. Komatitsch, “Sur la possibilité d’utiliser la représentation conjointe en temps et fréquence dans l’analyse des signaux modulés en fréquence émis en vibrosismiques”, in: Seventh GRETSI Symp. on Signal Processing and its Applications, 1979, pp. 121/1-121/6.
- [9] B. Boashash, A.P. Reilly, “Algorithms for time-frequency signal analysis”, in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman-Cheshire/Wiley, Melbourne/New York, 1992, pp. 163-181 (Chapter 7).
- [10] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications”, *Proc. IEEE* 80 (4) (1992) 540-568.
- [11] J.M. O’Toole, B. Boashash, “Fast and memory-efficient algorithms for computing quadratic time-frequency distributions”, *Appl. Comput. Harmo. Anal.* 35 (2) (2013) 350-358.

# Additional Notes and Further Reading

A final literature search, made just before going to press (late 2015), indicated that the field of Time Frequency Signal Analysis and Processing (TFSAP) was lively and expanding fast at the time of publication. Accordingly, in order to make this volume as comprehensive as possible, we conclude with a survey of related literature that appeared too recently to be cited in the relevant preceding chapters.

To show the emphases of recent research, statistics were compiled using only journal papers published in the last three years with keywords “Time-Frequency” or “Instantaneous Frequency”, as recorded in two databases: (1) IEEE *Xplore* and (2) *ScienceDirect*, which cover most of the important journals in the target research areas. About 400 papers were captured in the search. The following table classifies these papers based on the main themes or application areas (overlap may occur).

The table shows that in the last three years, besides theoretical areas, the top five application areas of TFSAP were: (1) *Biomedical Engineering*, (2) *Industrial Monitoring/Diagnosis*, (3) *Acoustics/Audio/Speech*, (4) *Instrumentation* and (5) *Communication systems*.

The following bibliography is an alphabetical listing of a selected subset of sources that appeared while this book was in production and before going to press. It is intended as a snapshot, dated late 2015, of the complementary research activity in TFSAP and its applications that is not directly referenced in the earlier chapters.

A consolidated bibliographic database compiled from all references in this book appears in the folder “Supplementary material” of the book’s companion website.

| Research Areas                                                                                                                            | Fraction                     | Related Chapters |
|-------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|------------------|
| TFSAP Theory, including:<br>- <i>Time-Frequency Methods</i><br>- <i>Instantaneous Frequency Methods</i>                                   | 30%<br>(20%)<br>(10%)        | 1–12             |
| Biomedical Engineering, including:<br>- <i>Neuroscience/EEG/MRI</i><br>- <i>ECG/HRV</i><br>- <i>Medical Engineering</i>                   | 20%<br>(10%)<br>(5%)<br>(5%) | 16               |
| Industrial Monitoring, including:<br>- <i>Fault Diagnosis of Materials/Structures/Machines</i><br>- <i>Vibrations/Structural Dynamics</i> | 15%<br>(10%)<br>(5%)         | 15               |
| Acoustics/Audio/Speech                                                                                                                    | 10%                          | 8, 11, 14        |
| Instrumentation/Measurements                                                                                                              | 10%                          | 15, 17           |
| Communication Systems (Optical/MIMO/Channel Estimation/Coding/Synchronization/Broadcasting)                                               | 10%                          | 13               |
| Satellite/GNSS/GPS/Navigation/Radar/Sonar                                                                                                 | 5%                           | 8, 14            |
| Electrical Power Lines/Systems                                                                                                            | 5%                           | 15               |
| Geoscience/Geophysics/Remote Sensing                                                                                                      | 5%                           | 14               |
| Electromagnetic/Microwave/Radio Frequency                                                                                                 | 5%                           | 13               |

- [1] A. S. Al-Fahoum, A. A. Al-Fraihat, "Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains", *ISRN Neuroscience*, vol. 2014, Article ID 730218, 7 pages, 2014, <http://dx.doi.org/10.1155/2014/730218>.
- [2] S. Aldirmaz, L. Durak-Ata, A. Akan, L. F. Chaparro, "Signal-adaptive discrete evolutionary transform as a sparse time-frequency representation", *Digital Signal Processing*, vol. 23 (5), September 2013, pp. 1747-1755, <http://dx.doi.org/10.1016/j.dsp.2013.05.010>.
- [3] O. A. Alkishriwo, A. Akan, L. F. Chaparro, "Intrinsic mode chirp decomposition of non-stationary signals", *IET Signal Processing*, vol. 8 (3), pp. 267-276, May 2014, <http://dx.doi.org/10.1049/iet-spr.2013.0396>.
- [4] J. P. Amezquita-Sanchez, H. Adeli, "A new music-empirical wavelet transform methodology for time-frequency analysis of noisy nonlinear and non-stationary signals", *Digital Signal Processing*, vol. 45, October 2015, pp. 55-68, <http://dx.doi.org/10.1016/j.dsp.2015.06.013>.
- [5] S. Barma, B. Chen, W. Ji, F. Jiang, J.-F. Wang, "Measurement of duration, energy of instantaneous frequencies, and splits of subcomponents of the second heart sound", *IEEE Transactions on Instrumentation and Measurement*, vol. 64 (7), pp. 1958-1967, 2015, <http://dx.doi.org/10.1109/TIM.2014.2383071>.
- [6] H. Bleton, E. Sejdić, "A cerebral blood flow evaluation during cognitive tasks following a cervical spinal cord injury: A case study using transcranial Doppler recordings", *Cognitive Neurodynamics*, vol. 9 (6), pp. 615-626, December 2015, <http://dx.doi.org/10.1007/s11571-015-9355-z>.
- [7] B. Boashash, G. Azemi, N. A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: Applications to newborn EEG abnormality detection", *Pattern Recognition*, vol. 48 (3), March 2015, pp. 616-627, <http://dx.doi.org/10.1016/j.patcog.2014.08.016>.
- [8] B. Boashash, G. Azemi, and J. M. O'Toole, "Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications", *IEEE Signal Processing Magazine*, vol. 30, pp. 108-119, November 2013.
- [9] B. Boashash, N. A. Khan, T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review", *Digital Signal Processing*, 40, 2015, 1-30, <http://dx.doi.org/10.1016/j.dsp.2014.12.015>.
- [10] G. Chen, J. Chen, G. M. Dong, "Chirplet Wigner-Ville distribution for time-frequency representation and its application", *Mechanical Systems and Signal Processing*, vol. 41 (1-2), December 2013, pp. 1-13, <http://dx.doi.org/10.1016/j.ymssp.2013.08.010>.
- [11] Q. Chen, L. Li, Y. Wang, "Amplitudes of mono-component signals and the generalized sampling functions", *Signal Processing*, vol. 94, January 2014, pp. 255-263, <http://dx.doi.org/10.1016/j.sigpro.2013.06.034>.
- [12] C. K. Chui, H. N. Mhaskar, "Signal decomposition and analysis via extraction of frequencies", *Applied and Computational Harmonic Analysis*, available online 13 January 2015, <http://dx.doi.org/10.1016/j.acha.2015.01.003>.
- [13] Y. Ding, I. W. Selesnick, "Sparsity-based correction of exponential artifacts", *Signal Processing*, vol. 120, March 2016, pp. 236-248, <http://dx.doi.org/10.1016/j.sigpro.2015.09.017>.
- [14] R. G. Everitt, C. Andrieu, M. Davy, "Online Bayesian inference in some time-frequency representations of non-stationary processes", *IEEE Transactions on Signal Processing*, vol. 61 (22), pp. 5755-5766, November 2013, <http://dx.doi.org/10.1109/TSP.2013.2280128>.
- [15] Z. Feng, M. Liang, F. Chu, "Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples", *Mechanical Systems and Signal Processing*, vol. 38 (1), July 2013, pp. 165-205, <http://dx.doi.org/10.1016/j.ymssp.2013.01.017>.
- [16] S. K. Hadjidimitriou, L. J. Hadjileontiadis, "EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings", *IEEE Transactions on Affective Computing*, vol. 4 (2), pp. 161-172, April-June 2013, <http://dx.doi.org/10.1109/T-AFFC.2013.6>.
- [17] T. Y. Hou, Z. Shi, "Data-driven time-frequency analysis", *Applied and Computational Harmonic Analysis*, vol. 35 (2), September 2013, pp. 284-308, <http://dx.doi.org/10.1016/j.acha.2012.10.001>.

- [18] T. Y. Hou, Z. Shi, P. Tavallali, “Convergence of a data-driven time-frequency analysis method”, *Applied and Computational Harmonic Analysis*, vol. 37 (2), September 2014, pp. 235-270, <http://dx.doi.org/10.1016/j.acha.2013.12.004>.
- [19] C. Huang, L. Yang, L. Yang, “ $\epsilon$ -mono-component: Its characterization and construction”, *IEEE Transactions on Signal Processing*, vol. 63 (1), pp. 234-243, 2015, <http://dx.doi.org/10.1109/TSP.2014.2370950>.
- [20] D. Iatsenko, P. V. E. McClintock, A. Stefanovska, “Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms”, *Digital Signal Processing*, vol. 42, July 2015, pp. 1-26, <http://dx.doi.org/10.1016/j.dsp.2015.03.004>.
- [21] G. Kim, J. Lee, Y. Kim, H. Oh, “Sparse Bayesian representation in time-frequency domain”, *Journal of Statistical Planning and Inference*, vol. 166, November 2015, pp. 126-137, <http://dx.doi.org/10.1016/j.jspi.2015.02.008>.
- [22] Y. Liu, Y. Li, H. Lin, H. Ma, “An amplitude-preserved time-frequency peak filtering based on empirical mode decomposition for seismic random noise reduction”, *IEEE Geoscience and Remote Sensing Letters*, vol. 11 (5), May 2014, pp. 896-900, <http://dx.doi.org/10.1109/LGRS.2013.2281202>.
- [23] M. Lozano, J. A. Fiz, R. Jané, “Performance evaluation of the Hilbert-Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization”, *Signal Processing*, vol. 120, March 2016, pp. 99-116, <http://dx.doi.org/10.1016/j.sigpro.2015.09.005>.
- [24] J. Ma, J. Jiang, “Analysis and design of modified window shapes for S-transform to improve time-frequency localization”, *Mechanical Systems and Signal Processing*, vols. 58-59, June 2015, pp. 271-284, <http://dx.doi.org/10.1016/j.ymssp.2014.12.015>.
- [25] D. Malnar, V. Sucic, Z. Car, “Optimizing the reference signal in the cross Wigner-Ville distribution based instantaneous frequency estimation method”, *Procedia Engineering*, vol. 100, 2015, pp. 1657-1664, <http://dx.doi.org/10.1016/j.proeng.2015.01.540>.
- [26] D. P. Mandic, N. U. Rehman, Z. Wu, N. E. Huang, “Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis”, *IEEE Signal Processing Magazine*, vol. 30 (6), November 2013, pp. 74-86, <http://dx.doi.org/10.1109/MSP.2013.2267931>.
- [27] X. Mao, B. Zhang, Z. Ye, “Profile recovery for LAMOST fiber spectral data with low SNR via improved time-frequency peak filtering”, *Optik - International Journal for Light and Electron Optics*, available online 17 October 2015, <http://dx.doi.org/10.1016/j.ijleo.2015.10.035>.
- [28] A. Moukadem, Z. Bouguila, D. O. Abdeslam, A. Dieterlen, “A new optimized Stockwell transform applied on synthetic and real non-stationary signals”, *Digital Signal Processing*, vol. 46, November 2015, pp. 226-238, <http://dx.doi.org/10.1016/j.dsp.2015.07.003>.
- [29] A. Moukadem, S. Schmidt, A. Dieterlen, “High order statistics and time-frequency domain to classify heart sounds for subjects under cardiac stress test”, *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 157825, 15 pages, 2015, <http://dx.doi.org/10.1155/2015/157825>.
- [30] S. B. Nagaraj, N. J. Stevenson, W. P. Marnane, G. B. Boylan, G. Lightbody, “Neonatal seizure detection using atomic decomposition with a novel dictionary”, *IEEE Transactions on Biomedical Engineering*, vol. 61 (11), pp. 2724-2732, November 2014, <http://dx.doi.org/10.1109/TBME.2014.2326921>.
- [31] I. Orović, S. Stanković, T. Thayaparan, “Time-frequency-based instantaneous frequency estimation of sparse signals from incomplete set of samples”, *IET Signal Processing*, vol. 8 (3), pp. 239-245, May 2014, <http://dx.doi.org/10.1049/iet-spr.2013.0354>.
- [32] B. A. Reyes, S. Charleston-Villalobos, R. González-Camarena, T. Aljama-Corrales, “Assessment of time-frequency representation techniques for thoracic sounds analysis”, *Computer Methods and Programs in Biomedicine*, vol. 114 (3), May 2014, pp. 276-290, <http://dx.doi.org/10.1016/j.cmpb.2014.02.016>.
- [33] F. Riaz, A. Hassan, S. Rehman, I. Niazi, K. Dremstrup, “EMD based temporal and spectral features for the classification of EEG signals using supervised learning”, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, available online 08 June 2015, <http://dx.doi.org/10.1109/TNSRE.2015.2441835>.

- [34] J. Rodrigues, A. Andrade, "Causal inference in neuronal time-series using adaptive decomposition", *Journal of Neuroscience Methods*, vol. 245, April 2015, pp. 73-90, <http://dx.doi.org/10.1016/j.jneumeth.2015.02.013>.
- [35] K. Schiecke, M. Wacker, F. Benninger, M. Feucht, L. Leistritz, H. Witte, "Matching pursuit-based time-variant bispectral analysis and its application to biomedical signals", *IEEE Transactions on Biomedical Engineering*, vol. 62 (8), pp. 1937-1948, 2015, <http://dx.doi.org/10.1109/TBME.2015.2407573>.
- [36] N. K. Sharma, T. V. Sreenivas, "Event-triggered sampling using signal extrema for instantaneous amplitude and instantaneous frequency estimation", *Signal Processing*, vol. 116, November 2015, pp. 43-54, <http://dx.doi.org/10.1016/j.sigpro.2015.03.025>.
- [37] R. Sharma, R. B. Pachori, U. R. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals", *Entropy*, vol. 17 (2), pp. 669-691, 2015, <http://dx.doi.org/10.3390/e17020669>.
- [38] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, M. Daković, "Instantaneous frequency in time-frequency analysis: Enhanced concepts and performance of estimation algorithms", *Digital Signal Processing*, vol. 35, December 2014, pp. 1-13, <http://dx.doi.org/10.1016/j.dsp.2014.09.008>.
- [39] L. Stanković, I. Orović, S. Stanković, M. Amin, "Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency", *IEEE Transactions on Signal Processing*, vol. 61 (18), pp. 4562-4572, September 2013, <http://dx.doi.org/10.1109/TSP.2013.2271752>.
- [40] L. Stanković, S. Stanković, M. Daković, "From the STFT to the Wigner Distribution [Lecture Notes]", *IEEE Signal Processing Magazine*, vol. 31 (3), pp. 163-174, May 2014, <http://dx.doi.org/10.1109/MSP.2014.2301791>.
- [41] J. Urbanek, T. Barszcz, J. Antoni, "A two-step procedure for estimation of instantaneous rotational speed with large fluctuations", *Mechanical Systems and Signal Processing*, vol. 38 (1), July 2013, pp. 96-102, <http://dx.doi.org/10.1016/j.ymssp.2012.05.009>.
- [42] M. Wacker, H. Witte, "Time-frequency techniques in biomedical signal analysis: A tutorial review of similarities and differences", *Methods of Information in Medicine*, vol. 52 (4), May 2013, pp. 235-270, <http://dx.doi.org/10.3414/ME12-01-0083>.
- [43] P. Wang, J. Gao, "Extraction of instantaneous frequency from seismic data via the generalized Morse wavelets", *Journal of Applied Geophysics*, vol. 93, June 2013, pp. 83-92, <http://dx.doi.org/10.1016/j.jappgeo.2013.04.003>.
- [44] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, Z. He, "Matching demodulation transform and synchrosqueezing in time-frequency analysis", *IEEE Transactions on Signal Processing*, vol. 62 (1), pp. 69-84, January 2014, <http://dx.doi.org/10.1109/TSP.2013.2276393>.
- [45] S. Wang, X. Chen, Y. Wang, G. Cai, B. Ding, X. Zhang, "Nonlinear squeezing time-frequency transform for weak signal detection", *Signal Processing*, vol. 113, August 2015, pp. 195-210, <http://dx.doi.org/10.1016/j.sigpro.2015.01.022>.
- [46] W. D. Wellisch, A. N. Barreto, "Spectrum sensing of TETRA systems through time-frequency analysis", *IEEE (Revista IEEE America Latina) Latin America Transactions*, vol. 11 (3), pp. 913-919, May 2013, <http://dx.doi.org/10.1109/TLA.2013.6568833>.
- [47] Y. Xia, W. Zhou, C. Li, Q. Yuan, S. Geng, "Seizure detection approach using S-transform and singular value decomposition", *Epilepsy & Behavior*, vol. 52 (Part A), November 2015, pp. 187-193, <http://dx.doi.org/10.1016/j.yebeh.2015.07.043>.
- [48] L. Xu, Y. Yang, "Parameter estimation of underwater moving sources by using matched Wigner transform", *Applied Acoustics*, vol. 101, January 2016, pp. 5-14, <http://dx.doi.org/10.1016/j.apacoust.2015.07.020>.
- [49] Y. Yang, X. Dong, Z. Peng, W. Zhang, G. Meng, "Component extraction for non-stationary multi-component signal using parameterized de-chirping and band-pass filter", *IEEE Signal Processing Letters*, vol. 22 (9), pp. 1373-1377, 2015, <http://dx.doi.org/10.1109/LSP.2014.2377038>.

- [50] Y. Yang, Z. Peng, X. Dong, W. Zhang, G. Meng, "Application of parameterized time-frequency analysis on multicomponent frequency modulated signals", *IEEE Transactions on Instrumentation and Measurement*, vol. 63 (12), December 2014, pp. 3169-3180, <http://dx.doi.org/10.1109/TIM.2014.2313961>.
- [51] Y. Yang, W. Zhang, Z. Peng, G. Meng, "Multicomponent signal analysis based on polynomial chirplet transform", *IEEE Transactions on Industrial Electronics*, vol. 60 (9), pp. 3948-3956, September 2013, <http://dx.doi.org/10.1109/TIE.2012.2206331>.
- [52] H. Zhang, G. Bi, W. Yang, S. G. Razul, C. M. S. See, "IF estimation of FM signals based on time-frequency image", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51 (1), pp. 326-343, 2015, <http://dx.doi.org/10.1109/TAES.2014.130554>.
- [53] C. Zhang, Y. Li, H.-b. Lin, B.-j. Yang, "Seismic random noise attenuation and signal-preserving by multiple directional time-frequency peak filtering", *Comptes Rendus Geoscience*, vol. 347 (1), January-February 2015, pp. 2-12, <http://dx.doi.org/10.1016/j.crte.2014.10.003>.

This page intentionally left blank

# Index

## Symbols

1/0 kernel, 254–255  
10-neighborhood, 420

## A

a priori SNR, 685  
a-posteriori, 680  
abnormality diagnosis, 916  
ACS, *see* almost-cyclostationary  
active sleep, *see* electroencephalogram (EEG), sleep patterns  
AD-MUSIC, *see* MUSIC  
adaptive (modified) matching pursuit, 702–706, 708  
adaptive cone-shaped kernel, 256, 833  
adaptive notch filter, 750–753  
adaptive optimal kernel (AOK), 255–256, 299, 301, 304–306, 833–834  
- with chirp signals, 833, 834  
adaptive quadratic TFDs, 254–256  
adaptive representation for multiple structures (ARMUS), 704–708  
adaptive spectrogram, 184, 185, 222, 252–254  
adaptive STFT, 252–254  
adaptive TFDs, 252–256  
- directional, 299–307  
- fractional spectrogram, 299, 302, 304–306  
- of bat sound, 253  
- with 1/0 kernel, 254–255  
- with adaptive cone-shaped kernel, 256  
- with adaptive optimal kernel (AOK), 255–256, 299, 301, 304–306  
- with local time-frequency-varying adaptation, 301–303, 306  
- with radially Gaussian kernel, 255, 442–444  
- with reduced interference, 256  
adaptive Wigner-Ville distribution (AWVD), 781–789  
- for IF estimation, 786–789  
additive noise, 6, 10, 476, 477, 664  
- and polynomial FM signals, 597, 604  
- converted to multiplicative, 665  
- IF estimation in, 588, 591  
- smearing the WVD, 665  
adjoint operator, 837, 838  
adjoint system, 190  
admissibility condition, 143  
affine class, 147, 161, 166, 266, 269–273, 872, 874–878  
- with reassignment, 398  
affine class of TFDs, 142

affine group, 159, 388, 394, 395  
- defined, 388  
affine transformation, 388  
affine Wigner class, 148  
affine Wigner distribution, 148  
affine Wigner function, 390–395  
- unitarity of, 394–395  
affine WVD, *see* affine Wigner function  
affine-smoothed pseudo-Wigner distribution, 875–877  
aliasing  
- in analytic associate, 375  
- in discrete quadratic TFDs, 336–337, 339–340  
- in discrete WVD, 333–334  
- in images, 881–883  
almost-cyclostationary (ACS) signals, 562–569  
- generalized, 566–568  
- generalized expected ambiguity function (GEAF), 564, 566  
- Wigner distribution, 564–565, 568  
alternative Dunn's index (ADI), 441  
Altes, *see* Q-distribution  
AM-FM model, 212–213, 215  
- demodulation, 213, 216–217  
- multicomponent, 925–930  
AM-FM signals, 844  
ambiguity domain, 114, 298, *see* Doppler-lag domain  
ambiguity function, 90, 105, 111, 114, 238–243, 290, 321  
- and compressive sensing, 426–428  
- and fractional autocorrelation, 774  
- and signal design, 242  
- and TFDs, 239  
- and the STFT, 240  
- and the wavelet transform, 240  
- and uncertainty, 241–242  
- as time-frequency correlation, 240, 241  
- computation of, 243  
- filtered/generalized, *see* filtered ambiguity function  
- generalized, 557  
- higher-order, 809  
- in radar, 114, 238  
- in sonar, 238, 831  
- narrowband, 238, 240–243  
- properties of, 240–243  
- Sussman (symmetrical), 114, 239, 245  
- time-localized, 256  
- weighted, *see* filtered ambiguity function  
- wideband, 239, 243  
- Woodward, 239, 831  
ambiguity surface, 240–242

amplitude spectrum, 17  
 amplitude-shift keying (ASK), 789  
 analytic associate, 44, 45  
   - and aliasing in the WVD, 375  
   - and oversampling, 365  
   - computation of, 365, 375  
   - used to reduce cross-terms, 153  
 analytic image, 879–880  
   - and aliasing, 882  
   - choice of, 882–883  
 analytic signal, 11, 44, 58–60, *see also* analytic associate  
 anechoic model, 504  
 ANFIS, *see* adaptive network-based fuzzy inference system  
 angle of arrival, *see* direction-of-arrival...  
 ANN, *see* neural network  
 AOK, *see* adaptive optimal kernel  
 aortic stenosis, 888, 891  
 area rule, 763–764  
 area-under-the-curve (AUC) score, 224, 729–730, 738–739, 935  
 ARMUS, *see* adaptive representation for multiple structures  
 array processing, 454  
 artifacts, 68, 106  
 artificial neural network, *see* neural network  
 ASK (amplitude-shift keying), 789  
 asymptotic mean squared error  
   - of PWVD-based IF estimate, 600–601  
 asymptotic signal, 48  
 asymptotic-phase signals, 60  
 atmospheric turbulence, 604  
 atomic decomposition, 142–145  
 atoms, 142  
 attenuation rules, 680  
 AUC, *see* area-under-the-curve  
 audio signals, 843, 847  
 audio signals identification, 843  
 auditory model, 853  
 auto-term energy, 295  
 auto-terms, 108  
   - as superpositions of coherent cross-terms, 156  
   - in STFDs, 511, 513  
   - location of, 115  
   - orientation of, 295  
   - selection of, 495  
   - support, 291  
 autocorrelation function, 11–13, 74, 867  
 automatic recognition, 846  
 autonomic nervous system, 954, 957  
 average frequency, 51  
 averaged spatial TFD matrix, 464  
 AWVD, *see* adaptive Wigner-Ville distribution  
 axis transformation, 176, 178

**B**

B-distribution, *see also* binomial distribution  
 B-distribution (BD), 14, 91–93, 122–124, 279, 592, 918  
   - computation of, 368  
   - defined using time-lag kernel, 918  
   - discrete, 340, 341, 367  
   - for component extraction, 496  
   - for component separation, 501  
 B-S (EEG pattern), *see* burst-suppression  
 backscattering, *see* scattering  
 band-limited signal, 46, 73  
 bandpass signal, 54  
 bandwidth, 46, 48  
   - effective, 46  
   - essential, 47  
   - finite, 19  
   - instantaneous, 3, 17  
   - relaxed measures of, 46  
   - significant, 47  
 bandwidth-duration (*BT*) product, 14, 47, 48, 53, 200, 243, 696, 753, 754, 757  
 bat signal, 978  
 bat sound, 221, 222, 617–619  
   - adaptive TFDs of, 253  
   - comparison of TFDs of, 442–444  
   - Gabor spectrograms of, 221  
   - spectrogram and MBD of, 589  
 Bayesian detector, 700  
   - optimal, 700  
 Bayesian regularized LMB, *see* Levenberg-Marquardt  
   backpropagation  
 BD, *see* B-distribution  
 beamforming, 461, 502  
 beamforming pattern, 507  
 Bedrosian's theorem, 60  
 Bernoulli filter, 629  
 Bertrand distribution, 272–273, 533, 875  
 BFSK, *see* binary frequency-shift keying  
 bias  
   - of IF estimation, 580–586, 589–591, 593, 594, 596, 664–667  
   - of WVD, 522  
 bilinear class TFDs, 142  
 bilinear time-frequency representations, *see* quadratic TFDs  
 bilinear transformation, 106  
 binary  $(t, f)$  image, 420  
 binary frequency-shift keying (BFSK), 779–787  
 binary image, 615–616  
 binary masking approach, 505  
 binomial distribution, 250  
   - kernel decomposition, 362  
   - of electric power disturbances, 858, 860

- biological sounds, 845
  - bioprostheses, 892, 893
  - biorthogonality, 352, 357
  - biosignals, 953
  - bird song, 982
    - noisy miner, 414, 417, 418
  - BJD, *see* Born-Jordan distribution
  - blind source separation, 459
    - underdetermined, *see* underdetermined BSS
  - blind source separation (BSS), 467, 471, 474–476, 483–490, 503, 611–615, 830
    - based on STFD, 484–489, 491–501, 511
    - conditions permitting, 483
    - underdetermined, 487
  - blood pressure, 917
  - blurring, *see* smoothing
  - borehole acoustics, 314–316
    - flexural mode, 314–316
  - Born-Jordan distribution (BJD), 91–93, 121, 123, 124, 247
    - computation of, 368
    - discrete, 340, 341, 367
    - kernel decomposition, 364
    - minimum kernel energy property, 526
  - brain abnormalities, 957
  - brain injury, 955, 961
  - brain network
    - defined, 937
    - effective, 937
    - functional, 937, 941
    - representation of, 937
    - structural, 937
    - time-varying, 939, 941–943
  - Bregman technique, 840
  - BSS, *see* blind source separation
  - BSS-STRE, *see* short-term Rényi entropy (STRE)
  - BT product, *see* bandwidth-duration (BT) product
  - burst-suppression (B-S) EEG pattern, 223, 919, 933–936
  - butterfly function, 121, *see* cone-shaped kernel
- C**
- canonical pair, 58–59
    - characterization of, 59–60
  - cantilever beam, 904–906
  - Cauchy principal value, 45
  - CDMA (code-division multiple-access)..., *see* spread-spectrum...
  - CDWR, *see* cross-term deleted Wigner representation
  - central finite-difference (CFD) approx., 67, 258, 264
  - centroid, *see* local centroid
  - CFD..., *see* central finite-difference...
  - channel
    - linear dispersive, 753–760
    - multiplicative, 767
    - overspread, 558
    - random, 554
    - tetherless underwater, 539
    - time-varying, 554
    - underspread, 558, 754
    - wireless, 533
  - WSSUS, *see* WSSUS channel, 555
  - characteristic function, 557, 665
  - Chebyshev polynomial, 262
  - chirp, 32, *see also* frequency modulation; linear FM; quadratic FM; hyperbolic FM
    - and adaptive optimal kernel, 833, 834
    - detection by fractional autocorrelation, 774–777
    - sweep rate estimation of, 774–777
  - Choi-Williams distribution (CWD), *see* Exponential distribution (ED)
  - CIT, *see* coherent integration time
  - CKD, *see* Compact-support-Kernel Distribution
  - classification rate, 895
  - clustering
    - of components, 496–497, 947–949
    - of vectors, 439–441, 492–498, 501
  - code-division multiple-access..., *see* spread-spectrum...
  - cognitive load, 960
  - cognitive monitoring, 960
  - cognitive radio, 788
  - Cohen's class, 113n, *see* quadratic class; quadratic TFDs
  - coherence length, 478
  - coherent integration time (CIT), 818–825
  - coherent radar, 820
  - cointegration, 958
  - Compact support kernel (CSK), 125, 296
  - Compact-support-Kernel Distribution (CKD), 125, 126, 128, 297, 617–619, 739, 847, 850
  - complex argument TFD, 265
  - complex envelope, 59
  - complex lag TFD, 265
  - complex time lag, 265
  - complex WDF transform (CWT), 794, 797–799
    - cross-terms avoided by, 797
  - complex-lag distributions, 265
  - complex-lag TFD, 266
  - complex-time form, *see* polynomial WVDs
  - complexity estimation, 431
  - component
    - linking of, 615–618
  - component clustering, 496–497, 947–949
  - component extraction, 496–501, 611–615, 617–619, 944–948, 952
  - component separation, 500, 611–615, 617–619

- components
  - crossing, 434, 435
  - local number of, 434–437
  - number of, 430, 432–437
  - overlapping, 434, 435
  - real-life signal, 433
  - resolution of, 410
- compressive sensing, 424–429
  - and concentration measures, 403, 408
- computation, *see also entries for particular distributions, kernels, etc.*
  - latency of, 366
  - of analytic associate, 365, 375
  - of discrete kernels, 367–369
  - of discrete quadratic TFDs, 364–383
  - of discrete WVD, 365
  - of modified B-distribution, 370–372
  - of spectrogram, 372–373
  - of windowed DWVD, 365
  - real-time, 366
  - throughput of, 366
- concentration
  - and probabilities, 401
  - by analogy with duration, 403–404
  - enhancement by neural networks, 438, 441–444
  - measured by Rényi entropy, 402–403, 441–442, 444
  - measured by ratio of norms, 401–402, 441–442, 444
  - normalized measures, 403
  - of S-method, 405–407
  - of spectrogram, 405
  - optimization of, 406–408
  - time-frequency energy (TFEC), 940, 943
- condition monitoring, *see also machine condition monitoring*
  - of cantilever beam, 904–909
  - using natural frequencies, 904–909
- cone-shaped kernel, 121, 247
  - adaptive, 256, 833
- conformity
  - index of, 388, 390, 392–395
- connectivity criterion, 420
- connectivity matrix, 937–939
  - construction of, 938–942
- constant IF, 55
- constant Q transform, 97
- continuous wavelet transform (CWT), *see* wavelet transform
- convolution invariance, 105
- convolutive mixing, 458
- convolutive mixtures, 503
  - separation of, 487–490
- convolutive model, 503
- correlation operator, 547, 717
- covariance, 145
  - in the group domain, 160
  - in the time-frequency domain, 164
  - of bilinear/quadratic TFRs, 164
  - of linear TFRs, 164
- covariance matrix, 457
- covariance property, 158, 266, 871, 873
- covariance theory of time-frequency analysis, 158
- CPF, *see* cubic phase function
- Cramer representation, 563
- Cramer-Rao bound, 242, 473, 624, 664, 786
  - and QML method, 625
- CRB, *see* Cramer-Rao bound
- CRLB, *see* Cramer-Rao bound
- CRLS (cascade recursive least squares), *see* neural network
- cross modified S-transform (XMST), 312–317
  - co- and quadrature, 313, 315
- cross polynomial WVD (XPWVD), 599
  - iterative IF estimation by, 599
- cross Teager-Kaiser (CTK) energy operator, 213–215
  - and cross WVD, 214–215
  - and cross-correlation, 214
- cross-ambiguity function, 131, 292, 755
- cross-correlation
  - 2-D, 920
- cross-correlation function, 291
- cross-MST, *see* cross modified S-transform
- cross-range resolution, 804, 805
- cross-Rihaczek-Margenau-Hill (XRMH) distribution, 313, 314
  - co- and quadrature, 313, 315
- cross-spectrogram, 358, 361, 364
- cross-term deleted Wigner representation (CDWR), 833, *see also* Gabor expansion
- cross-term energy, 293, 295
- cross-terms, 68, 106, 107, 151–158, 288, 833
  - amplitude, 410
  - and localization, 155–157
  - avoided by CWT, 797
  - avoided by TFAR and TFMV, 799
  - away from origin in Doppler-lag domain, 153–154, 245
  - geometry of, 298
  - in polynomial TFDs, 264
  - in STFDs, 474–476, 511, 513
  - in the energy spectrum, 151
  - in the spectrogram, 151
  - in the WVD, 151–153
  - location of, 115
  - reduced by using analytic associate, 153
  - related to XWWDs, 108
  - statistical, 551
  - support, 291, 293

- suppressed by LI kernels, 278
  - suppression of, 110, 114
  - cross-TFD, 455, 479
  - cross-validation, 898
  - cross-Wigner-Ville distribution (XWVD), 219
    - co- and quadrature, 313, 315
    - defined, 108
    - for matched detection, 726, 729–730
    - in IF estimation, 576
  - CSK..., *see* Compact-support-Kernel...
  - CTK, *see* cross Teager-Kaiser...
  - cubic FM signals, 264
  - cubic phase function (CPF), 621, 622
    - and higher-order PPSs, 622
    - noise and cross-terms, 622
  - cubic spline, 929
  - cumulant, 665
  - CWD, *see* Choi-Williams distribution
  - CWT, *see* complex WDF transform; wavelet transform
  - cyclostationarity, 562, 569
    - and quadratic TFRs, 564–565
    - of engine signals, 867
  - cyclostationary detection, 788
  - cyclostationary random process, 557, 562–569
- 
- D**
  - data compression, 895
  - data visualization, 971
  - data-clustering, 214
  - data-dependent TFDs, 969
  - DCT, *see* discrete cosine transform
  - de-blurring, 438–439
  - decay time, 79
  - dechirping, 71, 258
    - in polynomial WVDs, 597–598
  - decimation
    - of TFDs, 378–383
  - decision making, 4, 5
  - decision tree, 903
  - deflection, 478
  - degrees of freedom, 20
  - delay spread, 558
  - delay-Doppler spread, 189, 699, 761, 762
    - straight-line, 762
  - delay/Doppler scenario, 477
  - demo signals, 970
  - denoised Hearingogram, 681
  - denoising, 422, 643, 680, 711–714
    - by detrended fluctuation analysis, 949–952
    - by empirical mode decomposition, 944, 947–952
    - experiments, 714
    - mask design for, 713–714
    - with the Hilbert-Huang transform, 948–952
  - DEO, *see* differential energy operator
  - detectable dynamic range (DNR), 794
  - detection, 719, *see also* test statistic
    - as hypothesis testing, 477, 695–696, 831–832
    - Bayesian, 700
    - by fractional Fourier transform, 199
    - by quadratic TFDs, 696–701
    - decentralized, 757
    - of chirps, 774–777
    - of knock, *see* knock detection
    - of lines, 497
    - optimal, 694–701, 726
    - quadratic, 476–483
  - deterministic signal, 6, 8
  - detrended fluctuation analysis (DFA), 948–952
  - DFA, *see* detrended fluctuation analysis
  - DGF, *see* directional Gaussian filter
  - DGT, *see* discrete Gabor transform
  - DI..., *see* Doppler-independent...
  - diastolic pressure, 917
  - dictionary, time-frequency, 426–429
  - differential energy operator, 213
  - differential equations, 200–207
  - dilation invariance, 390, 398
  - dipole sources, 459
  - Dirac delta function, 11
  - Direct acyclic graph, 903
  - direct-sequence spread-spectrum communication, 748
  - direction finding, 464
  - direction-of-arrival (DOA) estimation, 456, 467, 469–471, 473, 476–480, 482, 483
  - directional denoising, 685
  - directional filtering, *see* smoothing, directional
  - directional Gaussian filter
    - double derivative, 303–307, 739–740
  - directional smoothing, *see* smoothing, directional
  - discontinuities
    - well handled by wavelet transform, 95
  - discrete Fourier transform, 648
  - discrete Gabor transform (DGT), 709–711
    - inverse, 710
  - discrete PWVDs, 599
  - discrete SM, 345–347
    - examples, 346–347
  - discrete spectrogram, 338–369, 372–373
    - latency of, 369
    - sampling for, 338–339
  - discrete TFDs, 332–342
    - aliasing in, 333–334, 336–337, 339–340
    - computation of, 364–383

- discrete TFDs (*Continued*)
  - decimated, 378–383
  - derived from standard kernels, 340–341, 367
  - Doppler-frequency form, 337–338
  - general quadratic, 335–338
  - IF property for, 339
  - properties related to kernel, 339–340
  - real-time computation of, 366
  - sampling for, 335–337
  - with Doppler-independent kernels, 340, 341, 367
- discrete TFPF
  - bias-variance tradeoff, 667
  - iterative algorithm, 668
  - signal scaling, 667
  - window length selection, 667
- discrete wavelet, 145
- discrete wavelet packet analysis (DWPA), 96
- discrete wavelet transform (DWT), 95, 96, 447, 448,
  - see also* scalogram
  - in interference excision, 749
- discrete WVD, 332–334, 340, 341, 367
  - 2D, 881–883
  - aliasing in, 333–334
  - as discrete quadratic TFD, 337
  - computation of, 365
  - sampling for, 332–334
  - windowed, *see* windowed DWVD
- discrete WWVD, *see* windowed DWVD
- discrete-domain frames, 445–448
- dispersion
  - in borehole acoustics, 314–316
  - time shift, 268
  - time-frequency characteristics, 267
- dispersive
  - IF shift, 531, 534
  - spreading function, 535–537
  - system, 267, 531
  - transformation, 531, 872, 874
- displacement function, 161
- DNR, *see* detectable dynamic range
- DOA..., *see* direction-of-arrival...
- dolphin click
  - binomial distribution of, 362
  - Exponential distribution of, 251
- dominant frequency, 51
- dominant scatterers, 805
- dominant time, 51
- Doppler ( $v$ ), 72
  - in radar echo from helicopter, 799–804
- Doppler ambiguity, 130
- Doppler analysis, 818
- Doppler Echocardiography (DE), 887
- Doppler effect
  - acoustical, 810–812
  - hydrophone in water, 811–813
  - microphone in air, 811–812
  - vs. Doppler ( $v$ ), 72n
- Doppler processing, 820
- Doppler shift, 158, 188
- Doppler spectrum, 819
- Doppler spread, 554, 558, 754
- Doppler tolerance, 243
- Doppler-delay spread, 189, 761, 762
  - straight-line, 762
- Doppler-frequency kernel, 112, 116
- Doppler-independent kernel, 91, 93, 107, 118, 119, 121, 123, 124, 276–277, 283, *see also* windowed WVD
  - and TFD properties, 277–278
  - discrete, 340, 341, 367, 375, 377, 379–380
  - giving windowed WVD, 277
  - smoothing effect of, 276
- Doppler-invariant signal, 271
- Doppler-lag domain, 114, 115
  - and compressive sensing, 426–428
  - and cross-terms, 153, 154, 245
- Doppler-lag kernel, 112, 114, 115
- double convolution, 76, 276
- DSF, *see* spreading function, dispersive
- DSSS..., *see* direct-sequence spread-spectrum...
- dual function, 80, 143
- dual operator, 163
- Dunn's index (DI), 441
- duration, 46, 48
  - effective, 47
  - essential, 47
  - finite, 19
  - relaxed measures of, 46
  - significant, 47
- duration-bandwidth product, *see* bandwidth-duration ( $BT$ ) product
- DWD, *see* Wigner distribution, dispersively-warped
- DWS, *see* Weyl symbol, dispersive
- DWT, *see* discrete wavelet transform
- DWVD, *see* discrete WVD
- dyadic grid wavelet, 145
- dyadic sampling grid, 95
- dynamic bandwidth, 56
- dynamic bandwidth allocation, 788
- dynamic signal, 215–216

## E

- echolocation, 10
- echolocation signals, 847
- EEG, *see* electroencephalogram

- EEG abnormalities, 462  
 EEG Seizure signals, 297  
 EEG signal, 980  
 effective analysis window, 75  
 effective bandwidth, 46  
 effective duration, 47  
 effective lag window, 73  
 effective network, *see* brain network, effective  
 eigenvalue decomposition, 461  
 Einstein-Brillouin-Keller (EBK) quantization, 764  
 elbow criterion, 439–440  
 electrocardiogram (ECG), 917  
 electroencephalogram (EEG), 917, 935
  - analyzed by B-distribution, 918
  - and connectivity matrix, 937, 940
  - background patterns, 733–735, 920–923, 928
  - epileptic, 945–948, 950–952
  - for seizure detection, 677, 678, 727–730, 736–739, 928
  - multicomponent nature of, 947
  - nonstationarity of, 946, 951
  - seizure patterns, 303, 305–307, 727, 728, 733–735, 928–930, 936
  - simulation of, 930
  - sleep patterns, 928, 929
  - time-frequency peak filtering of, 669, 670
- elementary messages, 95  
 elementary signals, 142  
 EMBD, *see* extended modified B-distribution  
 EMD, *see* empirical mode decomposition  
 Empirical Mode Decomposition (EMD), 212–213, 225, 227, 944–952  
 energy
  - concentration about IF law, 42, 132, 410
  - of signal, 34
- energy atom, 219
  - oscillation of, 219
- energy concentration, 421  
 energy density, 83, 99
  - 2D, 432
  - quasi-, 700
- energy distributions, 142, 145  
 energy flow, 20  
 energy gradient, 83, 99  
 energy property, 104  
 energy separation algorithm (ESA), 212–213, 217  
 energy spectral density, 11  
 energy spectrum, 35, 70
  - cross-terms in, 151
- engine pressure signal, 406, 407  
 engine sound, 652
  - Gabor coefficients of, 652, 653
- enhanced  $(t,f)$  images, 128  
 ensemble average, 12, 13, 75  
 ensemble view, 170  
 entropy, 431
  - counting property of, 430, 432–437
  - Leipnik measure, 404
  - Rényi, 402–404, 410, 432–434, 436, 437, 612, 613, 735, 736, 739, 933
  - Shannon, 401n, 404, 432, 735, 932
  - short-term, 433, 436
  - spectral, 735
  - Wiener, *see* flatness, spectral
  - Zakai parameter, 404
- entropy estimation, 437  
 entropy measure, 682  
 entropy-based features, 961  
 epilepsy, 945–952  
 epileptic seizure activity, 957  
 ergodicity, 13  
 ESA, *see* energy separation algorithm  
 Escort-Tsallis entropy, 961  
 ESF, *see* spreading function, exponential  
 essential bandwidth, 47  
 essential duration, 47  
 estimation, 716  
 evolutionary spectrum, 548
  - generalized, 548
  - transitory, 548
- evolutive spectrum, 74  
 EWS, *see* Weyl symbol, exponential  
 expected ambiguity function
  - generalized, 549
- exponential class, 269, 271–273  
 Exponential distribution (ED), 91, 92, 123, 244–246
  - computation of, 368, 376, 378–379
  - discrete, 340, 341, 367
  - invariance properties of, 248
  - of dolphin click, 251
  - sensitivity to noise, 530
- exponential T-distribution, 594  
 extended displacement function, 162  
 extended modified B-distribution (EMBD), 91, 92, 123, 128, 289, 299, 304–306, 729, 730, 739, 780–789, 847, 851, 889
  - for IF estimation, 618, 619, 786–789, 935
  - kernel selection, 780–781
  - of FSK signals, 780–781, 783, 786
- extended signatures, 907, 909

- extended-DATE, 679
- extraction
  - of components, 496–501, 611–615, 617–619, 944–948, 952
- F**
- fading
  - in CDMA communication, 755, 758, 760
  - in wireless communication, 604
  - Rayleigh, 755, 756
  - reduced by time-frequency processing, 756–757
- family of atoms, 143
- fast Fourier transform (FFT)
  - computational cost, 374, 377
- fast wavelet transform, 788
- fault diagnosis, 894, 898
  - in frequency domain, 899
- fault location
  - in electric power networks, 860–863
- feature extraction, 895, 931–935
  - for machine condition monitoring, 899, 900, 904
- feature selection, 895, 931, 935
- features, 18
  - time-frequency, 128
- features (time-frequency)
  - analogous to  $f$ -domain features, 733–736
  - analogous to  $t$ -domain features, 736–737
  - of time-varying networks, 940–943
- features extraction, 846
- FEM, *see* finite-element method
- FFT, *see* fast Fourier transform
- fifth order derivatives, 265
- filter, *see* time-frequency filter
- filter bank, 81
  - perfect reconstruction, 447, 448
- filter methods, 896
- filtered ambiguity function, 113, 114, 239, 245, 254
  - for separable kernel, 275
- filtered STFT, 645
- filtered TFD, 984
- filtered WVD, 74, 119
- finite bandwidth, 46
- finite duration, 46
- finite-element method (FEM), 865, 867, 905–907
- Fisher criterion, 896
- Fisher information, 171
- Fisher information matrix, 242
- flatness
  - spectral, 733–735, 739
  - time-frequency, 735, 739, 740, 940, 943
- flux
  - spectral, 733–734, 739
  - time-frequency, 128, 734, 739
- FM-like signals
  - underdetermined BSS for, 491–501
- FM..., *see* frequency modulation...
- Form I (of polynomial WVD), 261
- Form II (of polynomial WVD), 261
- Form III (of polynomial WVD), 261
- formant, 844
- formant detection, 844
- formant trackers, 844
- Fourier series, 858
- Fourier theorem, 8
- Fourier transform (FT), 8, 35
- Fourier transform pair, 11, 12
- Fourier-Stieltjes integral, 563
- fractional autocorrelation theorem, 773
- fractional convolution, 771–773
  - alternative formulations, 772–773
  - special cases of, 771–772
- fractional correlation, 771–773
  - alternative formulations, 772–773
  - and ambiguity function, 774
  - and chirp detection, 774–777
  - auto-, 771
  - computational complexity of, 773
  - cross-, 771
  - special cases of, 771–772
- fractional Fourier transform, 284, 770
  - and Hermite-Gauss functions, 195
  - applications of, 199
  - computational complexity of, 200
  - defined, 194–195
  - global moments of, 197–198
  - invariance properties of, 197–198
  - local moments of, 198–199
  - moments, 197–199
  - of common functions, 196
  - of linear FM signal, 199
  - Parseval's relation for, 197
  - properties, 195
  - rotation property, 195, 284, 770
  - special cases of, 284
- fractional power spectrum
  - defined, 196
- fractional spectrogram
  - adaptive, 299, 302, 304–306
- fractional-powers form, *see* polynomial WVDs
- frame, 446
  - cascaded, 447–448
  - defined, 445
  - discrete-domain, *see* discrete-domain frames
  - dual, 446
- frame operator, 446

- frames, 905  
 frequency bin, 679  
 frequency covariance, 104, *see* frequency-shift invariance  
 frequency deviation, 732  
 frequency flows, 958  
 frequency marginal, 104  
 frequency modulation, *see also* chirp; linear FM; quadratic FM; hyperbolic FM
  - law, 32, 51
  - of multiple components, 702
  - rate (nonlinear), 703
 frequency resolution, 77  
 frequency support, 105, 121, 247
  - discrete, 339
  - strong, 84
 frequency window, 73, 74, 115  
 frequency-domain, 8  
 frequency-invariant system, 186  
 frequency-invariant WSSUS channel, 556  
 frequency-shift invariance, 121, 248, 266
  - and positivity, 318
  - of robust WVD, 543
  - preserved by reassignment, 398
 frequency-shift invariance (FI), 104  
 frequency-shift keying (FSK), 8, 777–789
  - classification, 777, 786–789
  - IF estimation, 777
  - in ambiguity domain, 780, 782, 787
  - parameters, 777–779, 786–789
  - time-frequency representation, 777, 779–789
 FRFT, *see* fractional Fourier transform  
 FSK, *see* frequency-shift keying  
 FT, *see* Fourier transform  
 Fubini's theorem, 74  
 functional network, *see* brain network, functional
- G**
- Gabor coefficient, 218, 653–656
  - of engine sound, 652, 653
 Gabor expansion, 218, 833
  - critical sampling in, 653
  - discrete, 653–655
  - dual window functions, 653–655
  - in time-varying filtering, 652–656
  - oversampling in, 351, 354–355, 653
  - product form, 353
  - sum-of-products form, 354
  - with non-orthogonal sampling, 351, 356–358
  - with rectangular lattice, 352
 Gabor filter, 143, 641  
 Gabor frames
  - non-stationary, 905
 Gabor spectrogram, 217–224
  - concept of, 218
  - convergence to WVD, 219–221
  - defined, 218–219
  - of bat sound, 221
 Gabor transform, 80, 143, 352, 448, 641
  - discrete, *see* discrete Gabor transform
  - in interference excision, 749
  - multi-window, 446
  - non-stationary, 905–906
  - oversampling in, 354–355
  - product form, 353
  - sum-of-products form, 354
 Gabor transform pair
  - orthogonal-like, 654, 656
 Gabor transform, non-stationary, *see* non-stationary Gabor transform  
 Gabor wavelet, 143  
 GACS, *see* generalized almost-cyclostationary  
 GAF, *see* generalized ambiguity function  
 gain function, 680  
 gas reservoir detection, 842–843  
 Gaussian atoms, 223  
 Gaussian process
  - complex circular, 608
 Gaussian signal, 48, 52, 697
  - as basis function, 218
  - in derivation of Gabor spectrogram, 218
 GEAF, *see* generalized expected ambiguity function  
 generalized almost-cyclostationary (GACS) signals, 566–568  
 generalized ambiguity function (GAF), 113, *see also* filtered ambiguity function  
 generalized expected ambiguity function (GEAF), 564, 566  
 generalized inner-product invariance, 131  
 generalized input/output method, 204–205  
 generalized likelihood ratio test (GLRT), 478, 482, 696, 697, 700, 701, *see also* likelihood ratio detector  
 generalized marginals, 256, 284–288, 389  
 generalized Radon transform, 390  
 generalized spreading function, 640  
 generalized Wigner distribution (GWD), 835  
 generalized-marginal property, 286  
 generalized-marginal TFDs, 286–288
  - kernel examples, 288
  - kernel properties, 288
 geophysical applications, 837, 841–843  
 Givens rotation, 514–515  
 global energy property, 104  
 Global Positioning System, 753  
 GLRT, *see* generalized likelihood ratio test  
 GPS, *see* Global Positioning System

- graph Fourier transform, 939
  - inverse, 939
  - windowed, 939
 graphical user interface (GUI), 968
 ground roll, in seismic exploration, 841–842
 group, 159
 group delay, 14, 55, 105, 162, 864, 872, *see also* spectral delay; time delay
  - constant, 872
  - dispersive, 873
  - hyperbolic, 390
  - local, 397
  - nonlinear, 873
 group delay property
  - discrete, 339
 group velocity, 861, 863
 GWD, *see* generalized Wigner distribution
- H**
- Haar window, 358, 361
 Hadamard product, 510
 HAF, *see* higher-order ambiguity function
 Hamming window, 279
 Hankel form, 344
 Hanning window, 279, 405, 523, 529
 harmonic
  - decomposition, 10
 harmonic components, 10
 harmonic oscillator, 210–211
  - energy of, 210
 harmonizable random process, 563, 566–567
 HCPF-HAF (hybrid cubic phase function / higher-order ambiguity function), *see* hybrid CPF-HAF
 Hearingogram, 851
 heart rate, 917, 920, 953
 heart rate variability, 953
 heart rate variability measurements, 953
 heart sounds, 887, 890
 Heisenberg uncertainty relation, 79, 182–183, 396, *see also* uncertainty
  - applied to marginals, 169
  - in quantum mechanics, 20n
  - misleading as limit on resolution, 168–169
 helicopter
  - radar echo from, 799–804
  - sound from, 811, 815
 Hermite function expansions, 183
 Hermite normal form, 356
 Hermite-Gauss functions, 194
 Hermitian operator, 175, 177
 Hermitian symmetry, 10, 43
 Hessian, 171
 HHS, *see* Hilbert-Huang spectrum
 HHT, *see* Hilbert-Huang transform
 high-resolution ( $t, f$ ) methods, 422
 high-resolution QTFD, 289, 968
 high-resolution TFD, 126, 295
 higher order derivatives, 265
 higher-order ambiguity function (HAF), 602, 621, 622, 624, 809
  - for multiple component analysis, 603–604, 622
  - for single component analysis, 603, 622
 higher-order cumulants, 610
 higher-order IAF, 259
 higher-order moment, 610
 higher-order polynomial FM signals, 263
 higher-order spectra, 610, 835
  - bispectrum, 835
  - time-varying, 609, 610
  - trispectrum, 487, 835
 higher-order statistics, 263
 Hilbert space, 388, 445
 Hilbert spectrum, 227
 Hilbert transform, 11, 44, 45, 946
 Hilbert-Huang spectrum (HHS), 946–948
 Hilbert-Huang transform (HHT), 946–950, 952
 histogram of TFDs, 738
 homomorphic filtering, 962
 Hough Transform, 808
 HSF, *see* spreading function, hyperbolic
 Huber  $M$ -estimates, 539, 540, 546
 human perception, 681, 845
 human visual system (HVS), 878, 883, 884, 886
 HVS, *see* human visual system
 HWS, *see* Weyl symbol, hyperbolic
 hyperbolic
  - class, 167, 269, 271–274, 872
  - FM signal, 106, 243, 390, 392–394, 538, 671, 702–708, 833
  - frequency-shift operator, 166
  - time-frequency structure, 536
  - wavelet transform, 167
 hyperbolic T-distribution (HTD), 592, *see modified*
 B-distribution
 hypoxia-ischemia, 953
- I**
- IA, *see* instantaneous amplitude
 IAF, *see* instantaneous autocorrelation function
 IDGT, *see* discrete Gabor transform (inverse)
 IF estimation algorithms, 971
  - zero-crossing, 978
 IF..., *see* instantaneous frequency...
 IFT, *see* inverse Fourier transform

- image
    - binary, 615–616
  - image distortion, 878, 884–886
  - image processing, 419, 737
  - image quality, 878
    - as signal-to-noise ratio, 884, 885
    - time-frequency measures of, 884–885
  - image segmentation, 737
  - IMF, *see* intrinsic mode function
  - impulse noise, 539, 540, 545, 546
    - and TFDs, 539
  - index of (dimensional) conformity, 388, 390, 392–395
  - index of significance, 389, 390, 395
  - information
    - hidden, 14
    - spectral, 170–174
  - inner artifacts, 106, 120, 132, 155
    - suppressed by DI kernels, 278
    - suppressed by polynomial WVDs, 262
  - inner product, 10
  - inner-product invariance, 105, 389
  - innovations system, 547
  - instantaneous amplitude, 40, 50, 931, 934, 935
  - instantaneous autocorrelation function (IAF), 70, 86, 110
    - higher-order/polynomial, 259
    - spatial, 508
  - instantaneous bandwidth, 41, 57, 410
  - instantaneous cross-correlation function, 108, 214
  - instantaneous Doppler, 818, 820–825
  - instantaneous frequency (IF), 3, 6, 12, 17, 39, 41, 49, 52, 56, 58, 132, 162
    - adaptive estimation of, 580–586, 588–596
    - and Teager-Kaiser (TK) energy operator, 212, 215–217
    - density of, 400
    - dispersive, 702
    - dispersive change of, 531
    - encoding signal as, 663–664
    - estimated by matched spectrogram, 577–580
    - estimated by maxima of MBD, 592–596, 616–617
    - estimated by maxima of quadratic TFDs, 589–592, 594, 613–619
    - estimated by maxima of spectrogram, 588–589, 595
    - estimated by maxima of TFDs, 580, 581
    - estimated by polynomial WVD, 259–260, 600–601
    - estimated for noisy signals, 581
    - estimated for random signals, 576–580
    - hyperbolic, 702
    - hyperbolic shifts, 535
    - lag window affecting estimate of, 590–591
    - local, 397
    - localization by WVS, 604
    - mean, 55
  - multicomponent, *see* multiple IFs
  - nonlinear, 702, 703
  - of EEG signal, 934, 935
  - of electric power disturbance, 860–865
  - of engine noise, 867
  - of engine pressure signal, 865–867
  - PD estimators of, 620, 623, 624
  - power-function, 533, 536, 538
  - property, 104, 121, 259–260, 339
  - unbiased estimation of, 132
  - uniqueness of, 58
  - visualization of, 119
- instantaneous frequency curves, 421
- instantaneous frequency estimation, 920
- instantaneous heart rate, 953
- instantaneous mixing, 457
- instantaneous mixtures
  - separation of, 483–487, 489
- instantaneous phase, 50, 51, 56
  - and phase synchrony, 307
  - parametric estimation of, 807, 926
- instantaneous power, 11, 34, 70
- integer-powers form, *see* polynomial WVDs
- interference
  - traded against localization, 156–157
- interference terms, 43
- intermediate frequency (IF)
  - definitions of, 576
- interpolation, *see* oversampling
- intrinsic mode function (IMF), 225, 226, 944
- inverse Fourier transform (IFT), 35
- inverse Gabor transform, 80
- inverse STFT, 77, 507, 645, 984
- inverse synthetic aperture radar (ISAR), 805
- inverse time-frequency problems, 815–817
- invertibility
  - of TFD, 677
  - of WVD, 105
- invertible signal, 52
- ionospheric scatter, 820
- IP, *see* instantaneous phase
- ISAR, *see* inverse synthetic aperture radar
- iterated projections distribution, 256

## J

- Jacobi diagonalization, 485, 489, 515
- JAD, *see* joint anti-diagonalization; joint off-diagonalization
- Jakes Doppler power profile, 557
- jammer, 749, 751, 753
- Janssen's interference formula, 156
- JBDF, *see* joint block-diagonalization
- JD, *see* joint diagonalization

jet aircraft  
   - sound from, 812, 814  
 JOD, *see* joint off-diagonalization; joint anti-diagonalization  
 joint  $(t,f)$  density, 12  
 joint anti-diagonalization (JAD), 485–487, *see also* joint off-diagonalization  
 joint block-diagonalization (JBD), 488–489  
 joint diagonalization (JD), 485, 511  
 joint distributions  
   - axis transformation approach, 176  
   - by linked signal/axis transformations, 177–179  
   - construction of, 176  
   - signal transformation approach, 176–177  
 joint off-diagonalization (JOD), 511–512, 514–515, *see also* joint anti-diagonalization  
 Joint Photographic Experts Group (JPEG) coding, 884–886  
 joint spatial/spatial-frequency..., *see*  
   spatial/spatial-frequency...

**K**

Kaiser window  
   - 2D extension of, 882  
 Karhunen-Loëve Transform (KLT), 359  
 kernel, 99, 104, *see also* Doppler-frequency kernel;  
   Doppler-lag kernel; kernel filter; time-frequency kernel; time-lag kernel  
   - and variance of TFD, 527–528  
   - as filter, 115n, 117  
   - as in “signal kernel”, 66–67  
   - computation of, 367–369  
   - decomposition into spectrogram kernels, 358–364, 695  
   - discrete, 340, 341, 367–369  
   - Doppler-independent, 91, 93, 107, 118, 123, 276–277, 283, 340, 341, 367  
   - energy of, 525–526  
   - for optimal detection, 726–727  
   - for selected discrete quadratic TFDs, 341, 367  
   - generalized-marginal, 288  
   - lag-independent, 93, 118, 119, 277, 283  
   - non-separable, 295  
   - optimizing parameter of, 413, 414, 417  
   - separable, 93, 117, 154, 274–283, 398, 726  
   - singularities in, 368–369  
   - time-only, *see* lag-independent kernel  
   - transformations of, 112  
 kernel filter, 115, 117, *see* kernel  
 Kirkwood-Rihaczek distribution, 84, *see* Rihaczek distribution  
 Klauder wavelet, 395  
 KLT, *see* Karhunen-Loëve Transform (KLT)  
 knife-edge ridge, 66  
 knock detection, 722, 865–867

- by observing single combustions, 869–870  
   - by power and energy estimation, 870–871  
   - by S-method, 869–871  
   - position of sensors for, 868–869  
   - sampling for, 869  
   - signal models for, 865–868  
 Kullback-Liebler distance, 896  
 Kullback-Liebler divergence, 896

**L**

L-Wigner distribution (LWD), 348–349, 621  
   - optimal lag window for, 582  
 lag-independent kernel, 93, 118, 119, 122, 277  
   - and TFD properties, 277–279  
   - discrete, 375, 377, 380  
   - smoothing effect of, 277  
 Laplacian distribution, 540  
 latency, 366  
   - of discrete spectrogram, 369  
 LD, *see* Levin distribution  
 Lead Field Matrix, 458  
 leakage, *see* spectral dispersion  
 least squares (LS) method, 649  
 Leipnik entropy measure, 404  
 Levenberg-Marquardt backpropagation (LMB), 438, 440  
 Levin distribution (LD), 87, 92, 93, 123, 124  
   - defined, 84–85  
   - discrete, 340, 341, 367  
 LFM, *see* linear FM  
 LI..., *see* lag-independent...  
 Lie bracket, 212, 214  
 Lie group, 160  
 likelihood ratio, 629  
 likelihood ratio detector, 719, *see also* generalized likelihood ratio test  
   - time-frequency design, 722  
   - time-frequency formulation, 720  
 line detection, 497  
 linear FM signal, 13, 34, 37, 65  
   - as test signal, 32, 676, 677  
   - defined, 32  
   - finite-duration, 71  
   - for frequency-hopped CDMA, 492  
   - IF of, 32  
   - in multiplicative noise, 604–605  
   - optimal window duration for, 79  
   - piecewise, 289  
   - TFDs of, 91  
   - Wigner-Ville spectrum of, 604–607  
 linear frequency modulated (LFM) waveform, 818  
 linear frequency sweep, 32  
 linear mixtures, 6

- blind separation of, 483–490
  - linear phase, 54
  - linear time-frequency representations, 110, 158, 837–839
    - in PPS parameter estimation, 620–626
  - linear time-varying channel, 554, 761, 764
    - optimal waveforms for, 765–768
  - linear time-varying system, 185, 548, 552, 554, 638, 716, 760–761, 765–767
    - nonlinear frequency shift in, 531
  - Lloyd’s mirror effect, 812–815
    - nodal frequencies, 813–815
  - LMB, *see* Levenberg-Marquardt backpropagation
  - LNN, *see* neural network, localized
  - local cross time-frequency product, 184
  - local energy, 119
  - local entropy, 433
  - local ergodicity, 75, 610
  - local frequency bandwidth, 184
  - local time bandwidth, 184
  - localization, 10, 32
    - and cross-terms, 155–157
    - as by-product of interference, 155, 156
    - traded against interference, 156–157
  - localized spectrum, *see* short-time Fourier transform
  - Loéve bifrequency spectrum, 563, 567, 569
  - logon, 79
  - loss function, 539–541, 543, 546
  - LTV..., *see* linear time-varying...
  - LWD, *see* L-Wigner distribution
- M**
- M*-STFT, *see* robust
  - machine condition monitoring, 894–895, 899, 900, 904
    - data acquisition for, 900
    - during transient states, 899–900
  - magnitude spectrum, 35, 38
  - Mahalanobis distance, 901
  - MAI, *see* multiaccess interference
  - mainlobe amplitude, 410
  - mainlobe bandwidth, 410
  - man-made sounds, 845
  - Margenau-Hill distribution, 85, *see* Levin distribution
  - marginal median, 542
  - marginal properties, 70, 84, 169, 547, 548
    - and the kernel, 285–286
    - discrete, 339
    - violated by spectrogram, 120
  - marginals
    - generalized, 256, 284–288, 389
  - Marinovich-Altes distribution, 178
  - Markov chain, 628, 629
  - masked STFT, 984
  - masked WVD, 110
  - matched detector, *see* time-frequency matched detector
  - matched filter, 920, *see* time-frequency matched detector
  - matched spectrogram, 576, 580
    - computation of, 576–580
    - convergence, 578–580
    - convergence of computation, 578–580
    - in IF estimation, 576–580
    - window of, 576, 579, 580
  - matching pursuit, 222
    - adaptive (modified), 702–706, 708
    - adaptive algorithm, 704
    - concept, 702
    - dictionary of waveforms, 427–429, 702
    - iterative algorithm, 702
  - MATLAB<sup>TM</sup> code, 341, 369, 370, 373, 508, 510, 512, 513, 515
  - matrix TFD, 479
  - Maximal marginal diversity, 896
  - maximum eigenvalues test, 959
  - maximum likelihood (ML), 540
    - quasi (QML), 621, 623–626
  - MBD, *see* modified B-distribution; and extended modified B-distribution
  - MDD, *see* Multi-directional Distribution
  - MDK, *see* Multi-directional kernel
  - MDPSS, *see* prolate spheroidal sequence, modulated discrete
  - mean IF, 55
  - mean instantaneous intensity, 546
  - mean phase coherence, 308
  - mean squared error (MSE)
    - minimum (MMSE), 758, 760
    - of IF estimation, 581, 584, 624, 625, 926
    - of noise-reduced signal, 676
    - of PWVD-based IF estimate, 600–601
  - mean value
    - of quadratic TFD, 524
  - measure of complexity, 430
  - Median absolute deviation (MAD), 679
  - median WVD, 544
  - Mel filterbank, 681, 845
  - Mellin transform, 177, 390, 536
  - Mellin-Fourier duality, 239
  - memory load, 961
  - mental load, 960
  - mental overload, 960
  - Mexican hat wavelet, 95
  - MIMO, 454, *see* multiple-input multiple-output
  - minimax Huber  $M$ -estimates, 539, 540, 546
  - minimax robust Wiener filter, 718
  - minimum description length (MDL), 495

- minimum-energy TFD, *see* time-frequency distributions, minimum-energy
- MISO, 454
- missile tracking, 818–825
  - and instantaneous energy, 823–825, 827
- mixing matrix, 456, 467, 483–484
- ML, *see* maximum likelihood
- MMD criterion, 897
- mobile radio channel, 554
- modified B-distribution (MBD), 91, 92, 122–124, 279, 414, 417, 729, 730, 918
  - computation of, 370–372
  - discrete, 279, 340, 341, 367
  - IF estimation by, 592, 613, 614, 616, 617, 935
  - lag-independent kernel of, 279, 592
  - of bat sound, 589
  - properties of, 592
  - versus spectrogram, 590
- modified S-transform (MST)
  - cross-, *see* cross modified S-transform
  - resolution of, 312
- modulation invariance, 105
- modulation parameter estimation, 777–779, 807–809
- monocomponent FM signal, 42
- monocomponent linear FM signal, 132
- monocomponent signal, 6, 15, 49, 410
  - imprecision of terminology, 155
  - optimal window duration for, 78
- Monte Carlo method, 473, 607, 609, 627
- Moore-Penrose pseudo-inverse matrix, 649
- Morlet wavelet, 95, 310, 938
- mother wavelet, 95
- motion compensation, 806
- moving targets
  - discriminated from fixed, 809
- Moyal's formula, 131, 147, 240, 394, 725
- MRMR, 896
- MSE, *see* mean squared error
- MST, *see* modified S-transform
- Multi-directional Distribution, 126, 295, 849
  - kernel, 297
- Multi-directional kernel, 126, 288, 296
- Multi-directional TFD, *see* Multi-directional Distribution
- multi-path propagation, 458
- multi-resolution property, 961
- multi-sensor QTDFs, 454, 467
- multiaccess interference (MAI), 755, 758
- multichannel EEG, 454, 462
- multichannel TFDs, 454, 463
- multicomponent FM signal, 42
- multicomponent signal, 6, 15, 16, 19, 49, 410, 418, 491
  - imprecision of terminology, 155
- instantaneous frequencies of, 42, 588–592, 594–596, 611–619
- time-frequency peak filtering of, 669
- WVD of, 107
- multipath propagation, 478, 554
- multipath spread, 754
- multiple IFs, 42, 611–619
- multiple sensors, 454
- multiple-input multiple-output (MIMO), 200, 205–209, 487, 489
- multiplicative noise, 6, 8, 10, 43, 263, 604, 608
  - analyzed as additive, 529
- multiplicative system, 767
- Multiresolution analysis, 681
- multivariate phase synchrony, 958
- multiwindow Gabor filter, 641
- multiwindow STFT, 718, 722
  - filter based on, 640
- murmurs
  - heart, 887, 890, 892, 893
- MUSIC, 467, 470–473, 476
  - ambiguity-domain (AD-MUSIC), 471–473
  - time-frequency (t-f MUSIC), 471, 473–474
- MUSIC algorithm, 462, 464, 504
- musical notation
  - as time-frequency representation, 33
- mutual information, 896

## N

- natural frequencies, 904–909
- negative frequencies, 10, 43
- network analysis (time-frequency), 937–943
- neural network, 833
  - in machine condition monitoring, 901
  - in time-frequency analysis, 438–444
  - localized (LNN), 439–440, 442
- neural responses, 953
- neural-network-based TFD (NTFD), 442, 444
- neurological disorders, 953
- neuroprotective treatment, 962
- NLS, *see* nonlinear least-squares
- NMF, *see* non-negative matrix factorization
- noise
  - additive, *see* additive noise
  - analytic, 525–526
  - complex white, 524–525, 527–529
  - impulse, *see* impulse noise
  - in quadratic TFDs, 524–526, 920
  - in TFDs, 522–531
  - in windowed WD, 522–523
  - multiplicative, *see* multiplicative noise
  - nonstationary white, 524–525, 528–529

- real, 526
  - removal of, *see* denoising
  - stationary colored, 524–525, 529
  - stationary white, 524, 527–528
  - thresholding, 494, 500
  - noise reduction, 846
    - using TFDs, 670–677
  - noise subspace, 461
  - noise subspace subtraction, 672–673
  - noisy signals
    - polynomial WVD of, 598
    - quadratic TFDs of, 526–529
  - non-linear FM, 418
  - non-negativity, 120, *see* “positivity”
  - non-orthogonal sampling, 355–356
  - non-stationary Gabor frames, 905
  - non-stationary Gabor transform (NSGT), 905–907, 909
    - frequency-side, 905
  - nonlinear FM signal, 257, 265
    - as basic atom, 703
    - examples of, 257
    - WVD of, 106
  - nonlinear frequency shift, 531
  - nonlinear least-squares (NLS) method, 812–815, 817
  - nonnegative matrix factorization (NMF), 932, 935
  - nonstationary interference, 749, 751, 753
  - nonstationary random process, 531, 532, 537, 538, 546, 716
    - almost-cyclostationary (ACS), 562–569
    - cyclostationary, 557, 562–569
    - overspread, 550, 551
    - underspread, 550, 551, 716, 717, 721
    - white, 547–549
  - nonstationary signals, 4, 6, 13, 33
    - detection, 719
    - estimation, 716
  - nonstationary spectral correlation, 697, 698
  - normalized mean squared error (NMSE), 651
  - notation
    - standardization of, xxxiii
  - NSGT, *see* non-stationary Gabor transform
  - NTFD, *see* neural-network-based TFD
  - nuisance parameters, 694, 696, 697
    - and covariance properties, 694, 697
    - random, 700
  - Nyquist term, 365
- O**
- observation time
    - optimal, 170–175
  - OFDM, *see* orthogonal frequency division multiplexing
  - OKM, *see* optimal-radially-Gaussian kernel method (OKM)
  - One-against all, 902
  - One-against-one, 903
  - operator
    - integral, 266, 532, 873
  - optimal detection, *see* detection, optimal
  - optimal detector, 476, 479
  - optimal resolution, 295
  - optimal signal detector, *see* likelihood ratio detector, 719
  - optimal signal estimator, *see* Wiener filter, 717
  - optimal-radially-Gaussian kernel method (OKM), 441–444
  - orthogonal frequency division multiplexing (OFDM), 560, 561, 760, 767
  - orthogonal wavelet, 143
  - orthogonal-like functions, 711
  - orthogonal-like Gabor transform pair, 654, 656
  - orthogonality
    - quasi-, 493
    - time-frequency, 493–495, 501
  - orthonormal wavelet basis, 145
  - outer artifacts, 107, 120, 132, 155
    - suppressed by LI kernels, 278
  - oversampling
    - rational, 354–355, 358
  - overspread process, 550, 551
  - overspread system, 189, 192, 717, 721
  - overspread WSSUS channel, 558
  - oxygen saturation, 917
- P**
- p-point uncertainty classes, 718
  - Page distribution, 89, 92, 93, 123, 124
    - defined, 82–83
    - discrete, 340, 341, 367
  - parameters
    - of signals, *see* signal models; signal parameters
  - parasympathetic nervous system, 954, 955
  - Parseval’s relation, 240, 659
  - Parseval’s theorem, 11
  - partial displacement operator, 160
  - particle filter, 627
  - pattern recognition
    - in machine condition monitoring, 900, 901
    - time-frequency approach, 732, 739
  - PCA, *see* principal component analysis
  - PD (phase difference) IF estimators, 620, 623, 624
  - peak detection and tracking, 497, 500–501
  - perceptual features, 853
  - performance enhancement, 895
  - periodic auxiliary function, 654
  - Periodic lateralized epileptiform discharge, 961, 962
  - permutation problem, 507
  - PFFs, *see* product-function frames

- PHAF (product higher-order ambiguity function), *see* product HAF
- phase coupling, 307, 314
- phase delay, 55
- phase differentiation, *see* PD
- phase information, 96
- phase lag
  - of Hilbert transformer, 44
- phase locking value (PLV), 308, 938, 940
- phase retrieval, 214
- phase shift keying (PSK), 755
- phase spectrum, 35, 38
  - time-varying, 938
- phase synchrony, 307–317, 938–939, 943
  - narrowband, 307–308
- phase-locking, 958
- phase-shift keying (PSK)
  - and cross time-frequency distributions, 789
- Phonocardiogram (PCG), 886
- piecewise LFM signals, 289
- pitch, 33, 844
- PLV, *see* phase locking value
- poles
  - time-frequency, 201–203, 205
- polynomial derivatives, 266
- polynomial FM signal, 257, 926–928, *see* polynomial-phase signal
- polynomial IAF, 259
- polynomial phase estimation, 601–602
- polynomial phase transform (PPT), 264
- polynomial TFDs, *see also* polynomial WVDs, 257–988
- polynomial WVDs, 257–264, 608–610
  - coefficients of, 260–262
  - cross, *see* cross polynomial WVD
  - cross-terms in, 264
  - derivation of, 257–259
  - discrete, 599
  - for multiple component analysis, 603
  - for single component analysis, 603
  - general definition of, 259
  - IF estimation by, 589, 597–601, 621
  - IF property of, 259–260
  - in time-frequency peak filtering, 667
  - of multicomponent signals, 264
  - of noisy signals, 598
  - optimal lag window for, 582
  - reducing to ordinary WVD, 261
  - S-method for, 349–350
  - signal kernel of, 257, 597–598
  - sixth-order, 262, 609
  - suppressing inner artifacts, 262
  - symmetry about IF, 259–260
  - with complex time, 261
  - with fractional powers, 261
  - with integer powers, 261–262
- polynomial-phase signal (PPS), 257, 263, 264, 608, 620, 926–928
  - estimation of phase, *see* polynomial phase estimation
  - in additive noise, 596–604, 620–624, 626
  - in model of radar echo, 807, 809
  - matched spectrogram of, 580
- positive frequencies, 10
- positivity, 106, 120, 128–321
  - and shift invariance, 318
  - incompatible with IF and group delay properties, 321
  - necessary condition for, 318–319
  - of sonograph, 129
  - of spectrogram, 128
  - sufficient condition for, 320–321
- post-processing, 419
  - $(t,f)$ , 128
- power class, 872–877
  - formulation of, 874
- power quality, 858–859
  - time-frequency assessment of, 859–860
- power spectra, 13
- power spectral density (PSD), 11–12, 74, 546
  - frequency resolution in, 410
- power time shift property, 873
- power Wigner distribution, 875
- power-function
  - time-frequency representation, 537
  - transformation, 536
- PPS, *see* polynomial-phase signal
- PRFB, *see* filter bank, perfect reconstruction
- principal component analysis (PCA), 359, 732
- prior probability, 897
- probability of “birth”, 628
- probability of “survival”, 628
- product HAF (PHAF), 621, 622, 624
  - noise and cross-terms in, 622
- product kernel, 117
- product transform, 448
- product-function frames (PFFs), 448
  - for non-periodic spaces, 446
  - for periodic spaces, 446
- prognostic feature, 961
- prognostic information, 962
- projective group representation, 159
- projectogram, 360
- prolate spheroidal sequence, 765, 767
  - modulated discrete (MDPSS), 428, 429
- Prony method, 834–835
- propagation matrix, 456

- PSD, *see* power spectral density  
 pseudo-Wigner distribution
  - affine-smoothed, 875–877
 pseudo-WVD, *see* windowed WVD  
 PSF, *see* spreading function, power  
 PSK, *see* phase-shift keying  
 pulmonary artery pressure (PAP), 887  
 PWD (pseudo-Wigner distribution), *see* windowed WVD  
 PWS, *see* Weyl symbol, power  
 PWVDs, *see* polynomial WVDs

## Q

- Q-distribution, 272, 274, 536, 702, 706
  - and the ambiguity function, 239
  - smoothed pseudo, 273
 QMF, *see* quadrature matched filter  
 QML, *see* maximum likelihood, quasi  
 QTFRs (quadratic time-frequency representations), *see*
  - quadratic TFDs
 quadratic class, 113, 114, 158, 161, 166, 244, 245, 266, 694, 871  
 quadratic FM signal
  - IF estimation of, 609, 926–928
 quadratic TFDs, 112, 142, 158, 245, 968
  - adaptive, 254–256
  - affine class of, 266, 269–273, 872, 874–878
  - computation of, 364–383
  - decimated, 378–383
  - design of kernels for, 117, 246–248
  - desirable properties of, 119–121, 132, 591–592
  - detection by, 696–701
  - discrete, 332–342
  - exponential class of, 269, 271–273
  - exponentially warped, 272
  - generalized-marginal, 286–288
  - hyperbolic class of, 269, 273, 872
  - IF estimation by, 589–592, 927
  - in machine condition monitoring, 899
  - kernel decomposition, 358–361
  - noise in, 524–526, 920
  - of noisy signals, 526–529
  - power-class, 269, 271–272, 871–878
  - power-warped, 272
  - properties related to kernel, 121–122, 277–279, 317–321, 339–340
  - real-time computation of, 366
  - S-method in, 343–345
  - subclass  $\mathcal{P}$  of, 120
  - subclass  $\mathcal{P}'$  of, 120
  - table of, with properties, 124
  - time/frequency covariance of, 871
  - variance of, 524–526, 531

- with Doppler-independent kernels, 91, 93, 107, 118, 119, 123, 124, 276–277, 340, 341, 367
  - with lag-independent kernels, 93, 118, 119, 277
  - with reassignment, 397–398
  - with separable kernels, 93, 117–119, 154, 274–283, 398
- quadrature matched filter (QMF), 725, 729  
 quadrature signal, 44  
 quasi-energy density, 700  
 quasi-orthogonality
  - time-frequency, 493
 quiet sleep, *see* electroencephalogram (EEG), sleep patterns

## R

- R-peak detection, 954  
 radar
  - and the ambiguity function, 114, 238
  - fluctuating target in, 604
  - HF line-of-sight, 818–820, 824, 825
 radar antenna, 806  
 radially Gaussian kernel, 255, 442–444  
 radio monitoring, 777, 787–789
  - block diagram, 778
 Radon transform
  - generalized, 390
 Radon-Wigner distribution
  - rotation of, 284
 Radon-Wigner transform, 196  
 RAKE receiver, 755–756, 759–760  
 random processes, 14, 75  
 random set theory, 627  
 random signal, 12  
 random time-frequency shift, 557  
 random time-varying channel, 554  
 random walk model, 628  
 range ambiguity, 130  
 range-Doppler map, 819  
 range-velocity transformation, 238  
 rational oversampling, 354–355, 358  
 Rayleigh fading, 755, 756  
 RC, *see* reconstruction collection  
 RD, *see* Rihaczek distribution  
 real-time computation, 366  
 real-time filtering
  - of speech, 657, 659
 realness, 104, 121
  - computational exploitation of, 366
  - of polynomial WD, 259
  - of robust TFDs, 544

- realness (*Continued*)
  - of robust WVD, 543
- reassignment, 419
- reassignment method, 419
- reassignment points
  - asymptotic, 400
- reassignment principle, 396–401
  - and bilinearity, 398
  - and time/frequency shift invariance, 398
  - supervised reassignment, 400
- receiver operating characteristic (ROC), 224, 729, 732, 738, 739, 935–937
- reconstruction collection, 446
- reduced interference, 120
- reduced-interference distribution, 110, 132, 275
  - adaptive, 256
  - compared, 250
  - computation of, 358, 359, 374–383
  - design of, 132
  - design of kernels for, 246–248
  - discrete kernels for, 249–250, 380–381
  - in Doppler-lag domain, 244–245, 374–381
  - of dolphin click, 251
  - of electric power disturbances, 858–860, 864, 865
  - optimization of, 249
  - support properties, 247
- redundancy of features, 896
- redundant features, 732
- redundant representation, 143
- relative entropy, 896
- relaxation time, 56, 79
- relevance of a feature, 896
- Rényi entropy, 402–404, 410, 432–434, 436, 437, 612, 613, 735, 736, 739, 933, 940, 943
  - counting property of, 434
  - local, 433
  - time-frequency, 735, 736, 940, 943
- reproducing kernel, 240
- reservoir detection, 842–843
- residual spectrogram, 541, 545
- resolution, 120
  - enhancement by neural networks, 438, 441–444
  - evaluation of, 413
  - of components, 410
  - time-frequency, 113
- resolved components, 410
- respiration rate, 917
- reverberation, 458
- RID, *see* reduced-interference distribution
- Rihaczek distribution (RD), 87, 92, 123, 124
  - defined, 83–84
  - discrete, 340, 341, 367
  - reduced-interference, 938
- Rihaczek spectrum, 547
- ringing, 255, 883
- RMSE (root-mean-square error), *see* mean squared error
- road detection algorithm, 497, 501
- road network tracking algorithm
  - in underdetermined BSS, 497, 499, 501
- robust spectrogram, 539–542, 545
- robust STFT, 539–542
  - iterative realization, 541–542
  - vector filter realization, 542
- robust Wigner distribution, 543–545
  - properties of, 543–544
- ROC, *see* receiver operating characteristic
- rotation property, 284
- running energy spectrum, 82
- running spectrum, 20, 82
- running transform, 82
- RWT, *see* Radon-Wigner transform

## S

- S-method, 124, 126, 342–351, 419, 420
  - affine form, 345
  - and L-Wigner distribution, 348–349
  - and polynomial WVDs, 349–350
  - basic form, 343–344
  - concentration of, 405–407
  - cross-terms in, 343
  - discrete realization, 345–347
  - examples, 346–347, 350
  - forms in quadratic TFDs, 343–345
  - fractional domain form, 345
  - in spectral subtraction filtering, 661
  - related to STFT, 869
  - sensitivity to noise, 530
  - time direction form, 344
- S-transform, 96, 97, 310–311, 837
  - modified, 311–312
  - resolution of, 312
- sampling
  - for discrete quadratic TFDs, 335–337
  - for discrete spectrogram, 338–339
  - for discrete WVD, 332–334
  - for windowed DWVD, 334–335
  - ideal, 332
  - with compression, *see* compressive sensing 424
- SAR, *see* synthetic aperture radar
- Savitzky-Golay (SG) filter, 212, 215, 216, 673–674, 677
- SC, *see* spectrally correlated
- scalar product, *see* inner product
- scale covariance, 248, 272
- scale covariant, 142

- scale modulation, 446, 448
- scale operator, 177
- scaling property, 873
- scalogram, 147
  - cross-terms in, 151, 833
  - related to wavelet transform, 151, 399, 833
  - with reassignment, 400
- scalp EEG channels, 958
- scattering, 699, 827–829, 833, 836
  - by rotating rigid body, 806–807
  - in missile tracking, 820, 821, 823, 825, 826
- scattering function, 555, 699, 755
- SDS..., *see* spectral delay shift...
- seismic exploration, 837, 841–843
- seizure, 957
- seizures, 919
  - matched detection of, 920
  - neonatal vs. adult, 928
  - piecewise-linear FM, 727, 928–930
  - simulation of, 930
  - time-frequency detection of, 920
  - time-frequency patterns of, 303, 305–307, 727, 728, 928–930
- self-adjoint operator, 175
  - eigensystem of, 360, 362
- separable kernel, 93, 117–119, 121, 132, 154, 274–283
  - and TFD properties, 277–279
  - design examples, 279–280
  - for optimal detection, 726
  - general theory of, 275–276
  - numerical examples, 280–283, 380–381
- separate convolutions in  $t$  and  $f$ , 117
- separation
  - of components, 500, 611, 617
- separation algorithm, 507
- separation filter, 507
- separation measure, 412
- separation performance, 506
- sequential Bayesian estimation, 627
- SG, *see* Savitzky-Golay...
- Shannon entropy, 401n, 403–404, 432, 932, 961
  - time-frequency, 735
- shift covariant, 142
- short-term Rényi entropy, 612, 615, 618, 619
- short-time ambiguity function, 256
- short-time Fourier transform (STFT), 124, 142, 143, 158, 161, 166, 639, 645, 653, 680, 694
  - adaptive, 252–254
  - and QML method, 623–624
  - and S-method, 869
  - and the ambiguity function, 240
  - defined, 76
- in realization of higher-order TFDs, 348–350
- in realization of quadratic TFDs, 343–345
- in speech filtering, 659
- inverse, 77
- multiwindow, 718, 722
- robust, 539–542
- SIAF, *see* spatial instantaneous autocorrelation function
- sidelobe amplitude, 410
- Siebert's self-translation property, 241
- signal characteristics, 18
- signal characterization, 4
- signal classification, 701, 705, 931–932, 934–936
- signal decomposition, 229
  - by matching pursuit algorithm, 427–429, 702, 704, 706
- signal denoising, 422, 984
- signal detection, 719
- signal enhancement, 678, 847
- signal estimation, 716
- signal features, 14
- signal formulations, 43
- signal kernel, 66–67
  - for polynomial WVD, 257, 597–598
- signal mixtures, 503
- signal models, 43, 924–930
  - evaluation, 926
  - fitting, 926
  - parameters of, 43, 926–930
  - selection, 925–926
- signal moments
  - extended to two dimensions, 181–182
  - in one dimension, 180–181
- signal parameters, 414, 926–930
- signal power, 13
- signal reconstruction, 984
- signal scattering, 478
- signal spread, 80
- signal structure, 437
- signal synthesis, 649, 675–676, 969, 971, 983, 984
- signal transformation, 176–178
- signal-to-noise ratio, 14
- signatures
  - extended, 907, 909
- significance
  - index of, 389, 390, 395
- significant bandwidth, 47
- significant duration, 47
- SIMO, 454, *see* single-input multiple-output
- sinc function, 13
- single-input multiple-output (SIMO), 489
- single-input single-output (SISO), 200–205
- singular function, 766

- singular value, 766
  - derivative of, 673
  - normalized, 673
- singular-value decomposition (SVD), 670–674, 677, 835, 932
- sinusoidal FM signal, 77, 676
  - defined, 32
  - IF of, 32
- SISO, 454, *see* single-input single-output
- SM, *see* S-method
- smearing, 77n, *see* smoothing
- smoothed pseudo Wigner-Ville distribution, 154, 398
  - computation of, 376–377, 379–381
  - has separable kernel, 275, 283
  - reassigned (RSPWVD), 398, 764, 805, 807
- smoothed SIAF, 510
- smoothed WVD, 90, 414
- smoothing
  - as convolution, 87
  - directional, 299–307
  - versus “squeezing”, 398
- snake-based ridge extraction, 421
- SNR, *see* signal-to-noise ratio
- sonar, 702, 827–830
  - and the ambiguity function, 238, 831
- sonogram, *see* sonograph
- sonograph, 81, 845
  - defined, 81
  - positivity of, 129
  - related to spectrogram, 81, 129
- source localization, 458, 504
- source location, 456, 502
- source parameter estimation, 815–817
  - broadband, 817
  - narrowband, in air, 815–816
  - narrowband, in water, 816
  - wavelet denoising for, 817
- source separation, 459
- Space-Time-Frequency Processing, 462
- sparse decomposition, 425, 501, 837, 839
- spatial coherence, 478
- spatial instantaneous autocorrelation function (SIAF), 508
  - smoothed, 510
- spatial TFD (STFD), 454, 467, 484, 508–516
  - auto-term points, 493
  - auto-terms in, 511, 513
  - cross-terms in, 474–476, 511, 513
  - for blind source separation, 484–489, 491–501
  - matrices, 488–489
  - quadratic, 467, 510
  - structure under linear model, 511
- spatial/spatial-frequency representations, 878–880, 883, 884, 886
- spectral analysis, 15
- spectral autocorrelation function, 73
- Spectral Characteristics, 35
- spectral complexity, 171
- spectral correlation function, 72, 111
- spectral delay, 3, 14, 17, 38, 52, 55, 56
  - exponential, 272
  - hyperbolic, 271
  - nonlinear, 267
  - power-function, 272
- spectral delay shift (SDS) covariance, 266–273
- spectral disjointness, 45
- spectral entropy, 735
- spectral flatness, *see* flatness, spectral
- spectral gradient, 83
- spectral information, 170–174
  - for nonstationary signals, 172–174
  - for stationary signals, 170–172
- spectral spread, 17
- spectral subtraction
  - SM-based, 661
  - spectrogram-based, 660, 661
- spectral window, 20
- spectrally correlated (SC) signals, 567–569
  - jointly, 568–569
- spectrogram, 78, 86, 92, 93, 123, 126, 147, 694
  - adaptive, 184, 185, 222, 252–254
  - approximate decomposition into, 362–364, 699
  - as special case of S-method, 343
  - auto-, 358
  - computation of, 372–373
  - concentration of, 405
  - cross-, 358, 361, 364
  - cross-terms in, 151
  - decomposition into, 358, 695, 699
  - defined, 77
  - discrete, 338–369, 372–373
  - Doppler-lag kernel of, 128
  - enhanced, 123, 124
  - IF estimation by, 588–589, 625, 626
  - in spectral subtraction filtering, 660
  - in speech filtering, 659–660
  - in time-varying Wiener filter, 659
  - limitations of, 218, 244
  - marginals violated by, 120
  - matched, *see* matched spectrogram
  - of bat sound, 442–444, 589
  - optimal window duration for, 78, 582
  - positivity of, 128
  - related to sonograph, 81, 129

- residual, 541, 545
- robust, 539–542
- sensitivity to noise, 530
- subsumed by quadratic TFDs, 132
- time-lag kernel of, 86–87, 128
- variance of, 531
- with reassignment, 396–397, 441–443
- spectrum**
  - phase, 14, 17
- spectrum sensing**, 788
- speech**, 490
  - convolutive mixture, 490
  - recovered, 490
  - time-varying filtering of, 656–663
- speech enhancement**, 643, 679
- speech features**, 844
- spike**, 297
- spline**, 212
  - cubic, 929
- spread function**, *see* delay-Doppler spread
- spread-spectrum communication**
  - code-division multiple-access (CDMA), 492, 753, 755–760
  - described, 748
  - direct-sequence, 748
  - interference mitigation in, 748–753, 757–760
  - reduction of fading in, 756–757
- spreading function**, 754
  - dispersive, 535–537
  - generalized, 187, 549, 554, 640
  - narrowband, 533
  - wideband, 533
- SPWVD**, *see* smoothed pseudo Wigner-Ville distribution
- SS...**, *see* spread-spectrum...
- ST**, *see* S-transform
- state vector**, 628
- state-space approach**, 627
- stationarity**, 13
- statistical ( $t,f$ ) features**, 738
- statistical cross-terms**, 551
- statistical learning theory**, 898
- steering vectors**, 456, 467
- stenosis**, 890
- STFD**, *see* spatial TFD
- STFD matrix**, 455
- STFT**, *see* short-time Fourier transform
  - vector form of, 647
- STFT filter**, 639
- STRE**, *see* short-term Rényi entropy (STRE)
- stretch processing**, 818
- strong frequency support**, 84
- strong time support**, 84
- structural network**, *see* brain network, structural
- subspace noise filtering**, 670–677
- supervised reassignment**, 400
- Support Vector Machines**, 898
- support vectors**, 898
- Sussman ambiguity function**, *see* ambiguity function
- SVD**, *see* singular-value decomposition
- swallowing accelerometry**, 428, 431
- symmetrical ambiguity function**, *see* ambiguity function
- sympathetic nervous system**, 954, 955
- synchrony (phase)**, 307–317, 938–939, 943
  - narrowband, 307–308
- synchrosqueezed plane**, 400
- Synchrosqueezing**, 419
- synthesis (of signal)**, 675–676
- synthesized filtered signal**, 984
- synthetic aperture principle**, 804
- synthetic aperture radar (SAR)**, 804
  - basic principles, 804–805
  - images, 809
  - inverse (ISAR), 805
- synthetic aperture radar (SAR) processing**, 809
- system**
  - adjoint, 190
  - approx. eigenfunctions/eigenvalues of, 191
  - halfband, 643
  - innovations, 547
  - linear frequency-invariant, 186
  - linear time-invariant, 185, 762, 767, 768
  - linear time-varying, 185, 548, 552, 554, 638, 716, 760–761, 765–767
  - multiple-input multiple-output (MIMO), 200, 205–209, 487
  - overspread, 189, 192, 717, 721
  - random time-varying, 554
  - single-input single-output (SISO), 200–205
  - underspread, 188, 190, 552, 638, 716, 760–768
  - wideband, 538
- system analysis**, 4
- system identification**, 708–716
  - using chirp signal, 709
  - using pseudo-random signal, 708–709
- systolic murmurs**, 890, 891, 893
- systolic pressure**, 917

## T

- t-f MUSIC**, *see* MUSIC
- target accelerated**, 821
- Taylor series expansion**, 264
- TBM (theater ballistic missile)**, *see* missile tracking
- TBP (time-bandwidth product)**, *see* bandwidth-duration (*BT*) product

- Teager-Kaiser operator, 962  
 Teager-Huang transform (THT), 213  
 Teager-Kaiser (TK) energy operator, 209–217
 - and dynamic signal, 215–216  
 - and instantaneous frequency (IF), 212, 215–217  
 - cross-, 213–215  
 - derivation, 210  
 - discrete, 211  
 Teager-Kaiser energy operators
 - and time-frequency distributions, 214–216  
 test statistic, 477, 479, 482, 696–700, 719, 726, 728–729,  
*see also* detection  
 - optimal, 478, 480, 482, 483, 696, 724–726  
 texture ( $t, f$ ) features, 738  
 texture features, 846  
 TFAR..., *see* time-vs-frequency autoregressive...  
 TFDs, *see* time-frequency distributions  
 TFEC, *see* concentration, time-frequency energy  
 TFMF (time-frequency matched filter), *see* time-frequency matched detector  
 TFMV, *see* time-vs-frequency minimum-variance...  
 TFPF, *see* time-frequency peak filtering; discrete TFPF  
 TFRE, *see* Rényi entropy, time-frequency  
 TFRs, *see* time-frequency representations; time-frequency distributions  
 TFSA package, 262, 341, 373  
 TFSAP Toolbox, 967  
 TFSP, *see* time-frequency signal processing  
 theater ballistic missiles (TBM), 818  
 third order phase derivative, 265  
 thresholding constant, 713–714  
 tiling ( $t, f$ ) diagram, 98  
 time average, 75  
 time covariance, 104, *see* time-shift invariance  
 time delay, *see* spectral delay
 - critique of terminology, 52n  
 time delay estimation, 214  
 time marginal, 104  
 time of arrival
 - of electric power disturbances, 860–865  
 time resolution, 78  
 time support, 105, 121, 247
 - discrete, 339  
 - strong, 84  
 time window, 73, 74, 115  
 time-advanced signal, 366  
 time-bandwidth product (TBP), *see* bandwidth-duration  
 (BT) product  
 time-frequency
 - toolbox, 41  
 time-frequency “squeezing”, 398, 400  
 time-frequency analysis
 - matching, 269  
 - time-frequency centroid, 738  
 time-frequency characteristics, 920  
 time-frequency compactness, 737  
 time-frequency complexity, 19  
 time-frequency concentration, 19  
 time-frequency concentration measures, 736  
 time-frequency Convex Hull, 737  
 time-frequency correlation function, 549, 555  
 time-frequency correlation spread, 550  
 time-frequency dictionary, 426–429  
 time-frequency directions, 291  
 time-frequency displacement operator, 159  
 time-frequency distributions, 4, 14, 17, 32
 - adaptive, 252–256, 704  
 - and ambiguity functions, 239  
 - and Teager-Kaiser energy operators, 214–216  
 - comparison of, 408, 414, 417  
 - concentration of, 401–408, 438, 441–444, 932, *see also* concentration  
 - desirable characteristics of, 39  
 - detection by, 695–701, 832–833  
 - discrete, 332–342  
 - generalized-marginal, 286–288  
 - high-resolution linear, 796–804  
 - IF estimation by, 580–586, 588–596  
 - interpretation of, 20  
 - minimum-energy, 839, 840, 843  
 - of helicopter radar data, 794–804  
 - optimal, 413, 414  
 - optimizing performance of, 413, 414  
 - performance criteria for, 410  
 - performance measure for, 413, 414  
 - performance of, 408  
 - polynomial, *see* polynomial TFDs; polynomial WVDs  
 - poorly-performed, 413  
 - power-class, 269, 271–272, 871–878  
 - quadratic, *see* quadratic TFDs  
 - robust, 539–546, *see* robust spectrogram; robust STFT;  
 robust Wigner distribution; median WD  
 - SDS-covariant, 266–273  
 - selection of, 408, 413, 414, 417  
 - sparse, 403, 425–429, 837, 839–843  
 - well-performed, 413  
 time-frequency domain, 18, 115, 464, 895  
 time-frequency entropy measures, 735  
 time-frequency extensions, 735  
 time-frequency features, 732  
 time-frequency filter, 638
 - adaptive notch, 750–753  
 - applications of, 643  
 - discrete-time formulation of, 642

- explicit design of, 638
- for noisy speech, 656–663
- Gabor, 641
- Gabor expansion based, 652–656
- implicit design of, 638
- iterative algorithm for, 711
- multiwindow Gabor, 641
- multiwindow STFT, 640
- projection, 639, 751–753
- SM-based, 660–663
- spectral subtraction, 657, 659–663, 672–673
- spectrogram-based, 663
- STFT, 639
- system identification using, 708–716
- Weyl, 639
- Zadeh, 639
- time-frequency filtering, 418, 651, 984
- time-frequency flux, 19
- time-frequency image, 737, 967, 986
- time-frequency image enhancement, 128, 420
- time-frequency kernel, 90, 112, 121
- time-frequency law, 17
- time-frequency matched detector, 724–730, 920
  - area-under-the-curve (AUC), 729–730
  - for EEG seizures, 727–730
- time-frequency matched filter (TFMF), *see* time-frequency matched detector
- time-frequency measures
  - application of, 183–184
  - generation of, 182
  - properties & interpretation of, 182–183
- time-frequency pattern recognition, 732
- time-frequency peak filtering (TFPF), 422, 663–670
  - bias-variance tradeoff, 667
  - concept, 663
  - definitions, 663–664
  - discrete signal scaling, 667
  - discrete window length selection, 667
  - iterative algorithm, 668
  - of EEG signal, 669, 670
  - of multicomponent signal, 668–670
  - principles, 664
  - properties, 664–666
- time-frequency peak filtering (TFPF), 422
- time-frequency perimeter, 737
- time-frequency phase difference (TFPD), 309, 310
- time-frequency poles, 201–203, 205
- time-frequency post-processing, 419
- time-frequency probability representation, 683
- time-frequency projection filter, 639, 751–753
- time-frequency reassignment, *see* reassignment principle
- time-frequency relationships, 10
- time-frequency representation
  - sparse, 501
- time-frequency representations, 33, *see also* time-frequency distributions
  - linear, 110, 158, 621, 624, 626
  - linear discrete, 837–839
- time-frequency resolution, 113
- time-frequency scaling operator, 159, 166
- time-frequency shift, 15
- time-frequency shift operator, 158, 187
- Time-Frequency Signal Analysis* package, 262, 341, 373
- Time-Frequency Signal Analysis* toolbox, 373
- time-frequency signal detector, 722
- time-frequency signal estimator, 718
- time-frequency signal processing (TFSP), 3, 39
- time-frequency signatures, 24, 463, 733
- time-frequency specifications, 18, 984
- time-frequency transfer function, 190
- time-frequency window, 20
- time-frequency-space, 480
- time-invariant system, 185
- time-invariant WSSUS channel, 556
- time-lag kernel, 86, 91, 110, 112, 115, 694
  - of B-distribution, 918
  - of spectrogram, 87, 128, 694
- time-limited signal, 46, 55, 73
- time-Mellin distributions, 178
- time-only kernel, 277, *see* lag-independent kernel
- time-scale, 21
- time-scale analysis, 968
  - in machine condition monitoring, 899
- time-scale distributions, 145, 178, 701
- time-scale methods, 97
- time-scale representation, 143, 148
- time-shift invariance, 104, 121, 248, 266, 398
  - and positivity, 318
  - of robust WVD, 543
  - preserved by reassignment, 398
- time-varying amplitude, 43
- time-varying channel, 554
- time-varying components, 43
- time-varying frequency, 6
- time-varying higher-order spectra, 263, 609, 610
- time-varying impulse response, 531, 761–764, 769
- time-varying phase spectrum, 938
- time-varying power spectrum, 546, 551, 553
- time-varying spectrum, 74, 76
- time-varying system, 185, 531, 538, 548, 552, 554, 638, 716
- time-varying transfer function, 532, 638, 761–765, 769
- time-varying Wiener filter, 660, 662, 663, 717
  - SM-based, 661
  - spectrogram-based, 659, 661

time-vs-frequency autoregressive (TFAR) method, 798, 799  
 time-vs-frequency minimum-variance (TFMV) method, 798, 799, 804  
 TK, *see* Teager-Kaiser  
 Toeplitz factorization, 344  
 tomographic methods, 389–391  
 topographic surface, 737  
 total harmonic distortion (THD), 858  
 Tracé alternant, 918  
 tracé discontinu, 919  
 transformation, *see* axis transformation; signal transformation  
 transient detection, 214  
 transient spectrum, 200, 207–209  
 transient states  
   - of machines, 899  
 transitional probability matrix, 628  
 Tsallis entropy, 961  
 tuning width, 79  
 turboprop aircraft  
   - sound from, 811, 815  
 turbulence, 604  
 TV-HOS, *see* time-varying higher-order spectra  
 TVIR, *see* time-varying impulse response  
 TVTS, *see* time-varying transfer function  
 two-dimensional filtering, 115, 117

**U**

ULA, *see* uniform linear array  
 uncertainty, *see also* Heisenberg...  
   - and the ambiguity function, 241–242  
   - in time-frequency plane, 168–170  
   - measures of, in  $(t,f)$  plane, 170  
 uncertainty class, 718  
 uncertainty relation, 20  
   - Heisenberg, 182–183  
 underdetermined blind source separation, *see* underdetermined BSS  
 underdetermined BSS (UBSS), 505  
   - for FM-like signals, 491–501  
   - peak detection and tracking, 500–501  
   - road network tracking, 497, 499, 501  
   - vector clustering, 493–496  
 underdetermined case, 504  
 undersampling, *see* decimation  
 underspread approximations, 190, 551, 558, 717, 721  
 underspread process, 550, 551, 716  
 underspread system, 188, 190, 552, 638, 716  
   - eigenfunctions of, 760–768  
   - linear, 760–768  
 underspread WSSUS channel, 558, 700  
   - approximate eigenfunctions and eigenvalues, 558

  - sampling approximation, 560  
 underwater acoustic, 844  
 underwater acoustic channel, 554  
 uniform linear array (ULA), 476  
 unitarity  
   - of affine Wigner function, 394–395  
 unitary equivalence, 176–177  
 unitary equivalence principle, 168  
 unitary group representation, 159  
 unitary transform, 11  
 unitary transformation, 105, 268, 389, 486  
 unresolved components, 410  
 user-defined kernel, 987

**V**

validation indices, 441  
 valvular prostheses, 893, 894  
 variance  
   - mean value of, 528  
   - of IF estimation, 580, 581, 583, 584, 586, 591–594, 596, 606–607  
   - of quadratic TFD, 524–526, 531  
   - of spectrogram, 531  
   - of windowed WVD, 523, 530  
 vector clustering, 439–441, 492, 493, 497, 498, 501  
   - distance measure, 496  
 vector filter, 542  
 vector median, 542  
 Vibroseis signal, 32  
 vital signs monitors, 916  
 Viterbi algorithm, 631  
 vocal tract, 844  
 Volterra filter, 211  
 volume invariance  
   - of ambiguity function, 240

**W**

warp, 177–179  
   - dispersive, 268, 874  
   - operator, 177, 178, 268, 534, 874  
   - power, 875  
   - signal, 270, 535, 874  
   - time-frequency representation, 270, 535, 704, 874  
   - transformation, 177, 534, 874, 878  
   - unitary, 874  
 warped Wigner distribution, 178  
 watershed segmentation, 737  
 wave packets, 142  
 wavelet, 143  
   - dilation, 144  
   - translation, 144  
 wavelet basis, 95

- wavelet coherence, 957
- wavelet packet transform (WPT), 447, 448
- wavelet shrinkage, 681
- wavelet transform, 95, 142, 143, 161, 166, 310, 399, 446, 833, 961, *see also* scalogram
  - and discontinuities, 95
  - and phase, 310
  - and the ambiguity function, 240
  - and transient events, 95, 900
  - discrete, *see* discrete wavelet transform
  - in machine condition monitoring, 899
  - of electric power disturbances, 858
  - vs. S-transform, 310–311
- wavelets, 92, 394–395
  - as basis functions, 92, 95
  - Klauder, 395
  - Mexican hat, 95
  - Morlet, 95, 310, 938
  - mother, 95
  - theory of, 80
- WDF, *see* windowed data function
- Wentzel-Kramers-Brillouin (WKB) method, 761, 763, 768
- Weyl correspondence, 479, 695
- Weyl filter, 639
  - generalized, 639
- Weyl spectrum, 548
- Weyl symbol, 186, 638, 695, 699, 719
  - discrete-time, 642
  - dispersive, 534–537
  - generalized, 186, 547, 548, 553, 554, 638, 716
  - narrowband, 531, 537, 538
  - wideband, 533
- Weyl-Heisenberg group, 389
- whale signal, 981
- whale song, 184, 185
- white nonstationary process, 547–549
- whitened STFD matrix, 484, 485
- whitening, 484–485, 488, 511–512
- wide-sense stationary (WSS), 563, 564
- wide-sense stationary uncorrelated scattering, *see also* WSSUS channel
- wide-sense stationary uncorrelated scattering (WSSUS), 555, 699, 755, 756, 760
- Wiener attenuation rule, 651
- Wiener entropy, *see* flatness, spectral
- Wiener filter, 657, 658, 676–677, 748
  - robust, 718
  - time-frequency design, 718
  - time-frequency formulation, 717
  - time-varying, 659–663, 717
- Wiener-Khintchine theorem, 12, 74, 457, 462, 549, 564
  - cyclic, 564
- Wigner bispectrum, 835
- Wigner distribution, 67, 875
  - artifacts in, 69
  - dispersively-warped, 270, 704
  - generalized, 187, 547, 721
  - nonlinearity of, 68
  - power, 875–877
- Wigner-Ville distribution (WVD), 69, 87, 92, 93, 111, 123, 124, 132, 694
  - 2D, 878, 886
  - 2D continuous, 878–879
  - 2D discrete, 881–883
  - as special case of polynomial WVD, 261
  - as special case of S-method, 343
  - cross-terms in, 151–153
  - discrete, *see* discrete WVD
  - discrete windowed, *see* windowed DWVD
  - holographic nature of, 156
  - in derivation of Gabor spectrogram, 219
  - in gearbox fault detection, 900–901
  - in missile tracking, 821–825
  - in terms of the spectrum, 72
  - limitations of, 106, 218, 244
  - local centroid of, 397
  - masked, 110
  - median, 544
  - of linear FM signal, 71
  - polynomial, *see* polynomial WVDs
  - properties of, 104–106, 218
  - pseudo-, *see* windowed WVD
  - robust, 543–545
  - rotation of, 195, 284
  - satisfies *all* generalized marginals, 286
  - smoothing of, 118, 119, 414
  - spectrogram decomposition, 360–361
  - windowed, *see* windowed WVD
- Wigner-Ville spectrum, 75, 532, 547, 657, 665, 695, 697, 719, 722, 867
  - generalized, 546, 557, 716
  - IF localization by, 604
  - in additive noise, 606–607
  - in multiplicative noise, 604–606
  - knock detection by, 868, 869
  - of engine signal, 867, 868
  - of single combustion, 869–870
  - optimality of, 604–605
- Wigner-Ville trispectrum, 263, 487
- window
  - Hamming, 279
  - Hanning, 279, 405, 523, 529
- windowed data function (WDF), 794, 797–798

windowed DWVD, 334–335, 340, 341, 367

- as discrete quadratic TFD, 337
- computation of, 365
- sampling for, 334–335

windowed graph Fourier transform, 939

windowed Levin distribution, 85, 88, 92, 123

- discrete, 340, 341, 367

windowed Rihaczek distribution, 85, 92, 123

- discrete, 340, 341, 367

windowed WVD, 74, 91–93, 107, 119, 123, 124, 473,

*see also* Doppler-independent kernel

- 2D continuous, 881
- adaptive lag window for, 580, 583, 584, 586, 587, 782–785
- also called pseudo-WVD, 334
- discrete, *see* windowed DWVD
- kernel of, 277
- noise in, 522–523
- optimal lag window for, 523, 580–587
- sensitivity to noise, 530
- smoothed, 154, 398
- variance of, 523, 530

windowing process, 647

WKB, *see* Wentzel-Kramers-Brillouin

Woodward ambiguity function, *see* ambiguity function

WPT, *see* wavelet packet transform

wrapper methods, 897

WSS, *see* wide-sense stationary

WSSUS, *see* wide-sense stationary uncorrelated scattering

WSSUS channel, 555, 755, 756, 760

- delay spread, 558
- Doppler spread, 558
- frequency correlation function, 556

- frequency-invariant, 556

- overspread, 558
- random time-frequency shift, 557
- scattering function, 555, 755
- statistical input-output relations, 557
- time correlation function, 556
- time-frequency correlation function, 555
- time-invariant, 556
- underspread, 558

WT, *see* wavelet transform

## X

Xie-Beni (XB) index, 441

XMST, *see* cross modified S-transform

XPWVD, *see* cross polynomial WVD

XRMH, *see* cross-Rihaczek-Margenau-Hill...

XXWVD, *see* cross-Wigner-Ville distribution

## Z

Zak transform, 353, 354, 357

- and Fourier transform, 353
- defined, 353

Zakai entropy parameter, 404

ZAM..., *see* Zhao-Atlas-Marks...

zero-padding, 333–335, 338, 365

- in Doppler domain, 376
- in lag domain, 377, 379
- of analytic associate, 375

Zhao-Atlas-Marks (ZAM) distribution, 91–93, 121, 123,

184, 185, 247

- computation of, 368
- discrete, 340, 341, 367
- limitations of, 256

# Time-Frequency Signal Analysis and Processing

## A Comprehensive Reference

Edited by Boualem Boashash

A unique reference on Time-Frequency Signal Analysis and Processing (TFSAP), now updated with a strong emphasis on developing algorithms for advanced applications.

*Time-Frequency Signal Analysis and Processing* is a collection of theory, techniques, and algorithms used for the analysis and processing of nonstationary signals, as found in a wide range of applications, including telecommunications, radar, and biomedical engineering. This book gives the university researcher and R&D engineer insights into how to use Time-Frequency Signal Analysis and Processing (TFSAP) methods to develop and implement the engineering application systems they require.

### New to this Edition

- New sections on efficient and fast algorithms; a “Getting Started” chapter that enables readers to start using the algorithms on simulated and real examples with the TFSAP toolbox, compare the results with the ones presented in the book, and then insert or adapt the algorithms into their own applications
- Two new chapters and twenty-three new sections, including updated references
- New topics including: efficient algorithms for optimal TFDs, the enhanced spectrogram, time-frequency modeling, more mathematical foundations, the relationships between QTDFs and Wavelet Transforms, new advanced applications such as cognitive radio, watermarking, noise reduction in the time-frequency domain, algorithms for time-frequency image processing, and a new chapter on time-frequency applications in neuroscience. Relevant code is provided as Supplementary Material

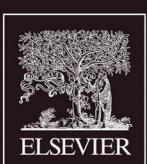
### Key Features

- A comprehensive tutorial introduction to TFSAP, accessible to anyone who has taken a first course in signals
- Key advances in theory, methodology, and algorithms are concisely presented by some of the leading authorities on the respective topics
- Applications written by leading researchers showing how to use TFSAP methods

### About the Editor

**Dr. Boualem Boashash** is a scholar, professor, and senior academic with experience in five leading Universities in France and Australia, mainly at Queensland Universities. He has published over 500 technical publications, including over 100 journal publications that cover engineering, applied mathematics, and biomedicine. He pioneered the field of time-frequency signal processing for which he published the most comprehensive book and most powerful software package. Among many initiatives, he founded ISSPA, which has been a leading conference since 1985. After a previous appointment in the UAE as Dean of Engineering, he moved to Qatar University in 2009 as a research professor. His work has been cited over 11,000 times.

Communications Engineering



ACADEMIC PRESS

An imprint of Elsevier  
[store.elsevier.com](http://store.elsevier.com)

ISBN 978-0-12-398499-9



9 780123 984999