1997 IEEE Signal Processing Workshop on Higher-Order Statistics

19981001 041

AQF99-01-0001

July 21–23, 1997
Banff, Alberta, Canada

Sponsored by
IEEE Signal Processing Society

IEEE COMPUTER SOCIETY

IEEE

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources,
gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1997	Final 7/1/97 - 12/31/97

4. TITLE AND SUBTITLE
IEEE 1997 Signal Processing Workshop on Higher Order Statistics

5. FUNDING NUMBERS

6. AUTHOR(S)
Keh-Shin Lii (as PI), collection of papers with different authors

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Riverside
Riverside, CA 92521

8. PERFORMING ORGANIZATION REPORT NUMBER
ONR N00014-97-1-0733

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
4520 Executive Dr. Suite 300
San Diego, CA 92121-3019

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The workshop is co-sponsored by IEEE Signal Processing Society

12a. DISTRIBUTION / AVAILABILITY STATEMENT
unclassified, unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The content of the Proceedings includes higher order statistics with applications, nonstationary signals, nonlinear systems, multidimensional signals and system identification and deconvalution, alpha-stable distributions, array processing, source separation, detection, estimation, modeling and analysis of nonGaussian processes.

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	unlimited

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI-Std Z39-18
298-102

UNIVERSITY OF CALIFORNIA, RIVERSIDE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF STATISTICS
FAX: (909) 787-3286

RIVERSIDE, CALIFORNIA 92521-0138

September 21, 1998

Dr. Neil Gerr
Office of Naval Research
800 N. Quincy Street
Ballston Tower One
Arlington, VA 22217-5660

RE: N00014-97-1-0733

Dear Dr. Gerr:

This is the final report on ONR grant N00014-97-1-0733 which partially supported the 1997 IEEE Signal Processing Workshop July 21-23, 1997. The main outcome of the conference is the publication of the Proceedings which includes 93 papers. Four copies of the Proceedings are enclosed with this report. One copy of the Proceedings is being sent to the Defense Technical Information Center (DTIC). The content of the Proceeding is unclassified and is for unrestricted release. A copy of this letter as well as SF298 are being sent to Ms. Jan Meyers, the Grant Administrator in San Diego. The support of this endeavor from ONR is gratefully acknowledged.

Sincerely,

Keh-Shin Lii

cc: Ms. Jan Meyers
Grant Administrator
Office of Naval Research
Regional Office, San Diego
4520 Executive Dr., Suite 300
San Diego, CA 92121-3019

SPW-HOS'97

Student Paper Awards

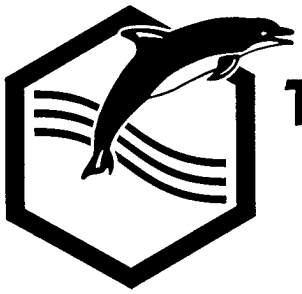
Sponsored by EURASIP

Student paper awards were provided by The European Signal Processing Organization (EURASIP).

EURASIP was founded in1980 and its aim is to further all aspects of signal processing research and education. The awards are seen by EURASIP as a suitable mechanism for encouraging and recognizing young researchers.

HOS'97 Icebreaker

Sponsored by Thomson Marconi, France



THOMSON MARCONI
SONAR

Proceedings of the

IEEE SIGNAL PROCESSING WORKSHOP
on
HIGHER-ORDER STATISTICS

July 21-23, 1997

Banff, Alberta, Canada

Sponsored by

The IEEE Signal Processing Society

With Support from

U.S. Office of Naval Research

IEEE COMPUTER SOCIETY

Los Alamitos, California

Washington • Brussels • Tokyo

Copyright © 1997 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Computer Society Order Number PR08005
ISBN 0-8186-8005-9
ISBN 0-8186-8006-7 (case)
ISBN 0-8186-8007-5 (microfiche)
IEEE Order Plan Catalog Number 97TB100151
Library of Congress Number 97-72013

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1-714-821-8380
Fax: + 1-714-821-4641
E-mail: cs.books@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1-908-981-1393
Fax: + 1-908-981-9667
mis.custserv@computer.org

IEEE Computer Society
13, Avenue de l'Aquilon
B-1200 Brussels
BELGIUM
Tel: + 32-2-770-2198
Fax: + 32-2-770-8505
euro.ofc@computer.org

IEEE Computer Society
Ooshima Building
2-19-1 Minami-Aoyama
Minato-ku, Tokyo 107
JAPAN
Tel: + 81-3-3408-3118
Fax: + 81-3-3408-3553
tokyo.ofc@computer.org

Editorial production by Deborah Plummer

Cover art production by Joe Daigle/Studio Productions

Printed in the United States of America by Sony Electronic Publishing Services

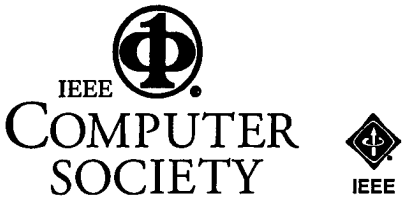


Table of Contents

IEEE Signal Processing Workshop on Higher-Order Statistics

Conference Organization xii

Plenary: High-order Cyclostationary Signal Analysis - An Overview
Georgios Giannakis, *University of Virginia, USA*

MAP: Applications of HOS
Chair: Steve McLaughlin, *University of Edinburgh, Scotland*

Autoregressive Modeling of Lung Sounds Using Higher-Order Statistics: Estimation of Source and Transmission 4
L. Hadjileontiadis and S. Panas

Multipath Time Delay Estimation using Higher Order Statistics 9
Y.-C. Liang, A. Leyman and B.-H. Soong

Discrimination of Local Seismic Events in Panama by Means of Higher-Order Statistics 14
L. Persson and J. Boutet

Performance Evaluation of Time-Delay Estimation in Non-Gaussian Conditions 20
G. Shor and H. Messer

Time-varying Third-order Cumulant Spectra and Its Application to the Analysis and Diagnosis of Phonocardiogram 24
M. Shen and F. Shen

The Analysis and Classification of Phonocardiogram Based on Higher-order Spectra 29
M. Shen and L. Sun

Coherent Interference Excision Using Higher Order Spectra 34
I. Jouny and A. Low

The Application of Non-linear Volterra Type Filters to Television Images 39
W. Collis, M. Weston and P. White

Higher Order Statistics for Detection and Classification of Faulty Fanbelts Using Acoustical Analysis. 43
G. Gelle, M. Colas and G. Delaunay

Classification Of Linear Modulations By A Combination Of Different Orders Cyclic Cumulants 47
P. Marchand, C. Le Martret and J.-L. Lacoume

Tracking Nonlinear Systems Using Higher Order Moments 52
K. Turner, F. Faruqi and C. Brown

Isolated Word Recognition Using High-Order Statistics and Time-Delay Neural Networks 57
M. Ashouri

MAL: Applications of HOS

Chair: Gaetano Scarano, *University of Rome, Italy*

Impact of Blind versus Non-Blind Channel Estimation on the BER Performance of GSM Receivers 62
D. Boss, T. Petermann and K.-D. Kammeyer

Inverter Fed Induction Machine Condition Monitoring Using the Bispectrum 67
N. Arthur and J. Penman

Higher-Order Statistics for Tissue Characterization from Ultrasound Images 72
U. Abeyratne and A. Petropulu

Signal-Dependent Film Grain Noise Removal and Generation Based on Higher-Order Statistics 77
J. Yan and D. Hatzinakos

Blind Identification Methods Applied to Electricité de France's Civil Works and Power Plants Monitoring 82
G. D'Urso, P. Prieur and C. Vincent

MPL: Nonstationary Signals, Nonlinear Systems

Chair: Chong-Yung Chi, *National Tsing Hua University, Taiwan*

Time-Frequency Representations and Their Structure 88
B. Friedlander and L. Scharf

Hybrid FM-Polynomial Phase Signal Modeling: Parameter Estimation and Performance Analysis 93
F. Gini and G. Giannakis

Higher-order Statistics and Extreme Waves 98
E. Powers, I.-S. Park, S. Im, S. Mehta and E.-J. Yi

Exploring Lag Diversity in the High-order Ambiguity Function for Polynomial Phase Signals 103
G. Zhou and Y. Wang

MPP: Nonstationary/Nonlinear/Multidimensional Signals & Systems

Chair: Athina Petropulu, *Drexel University, USA*

Déblurring Two-Tone Images by a Joint Estimation Approach Using Higher-Order Statistics 108
T.-H. Li and K.-S. Lii

Linear Algebraic Approaches for (Almost) Periodic Moving Average System Identification 112
Y.-C. Liang, A. Leyman and X.-D. Zhang

Fast Estimation of Wiener Kernels of Nonlinear Systems in the Frequency Domain 117
M. Shcherbakov

Second-Order Statistics versus HOS for the Estimation of Harmonics in Additive and Multiplicative Noise 122
M. Ghogho

Design of a Volterra Series-Based Nonlinear Compensator 127
J. Kim, K. Cho, Y. Kim, J. Chung and S. Nam

Nonlinearly Constrained Optimisation Using a
Penalty-Transformation Method for Volterra Parameter Estimation 132
T. Stathaki

An Iterative Mixed Norm Image Restoration Algorithm 137
M.-C. Hong, T. Stathaki and A. Katsaggelos

On Identifying Volterra Transfer Functions of Cubically
Nonlinear Systems Using Minimally Sampled Data 142
C.-H. Tseng

Higher-Order Statistics of Natural Images and their
Exploitation by Operators Selective to Intrinsic Dimensionality 147
G. Krieger, C. Zetzsche and E. Barth

A Normalized Block LMS Algorithm for Frequency-Domain
Volterra Filters 152
S. Im

Windows and Volterra Transfer Function Estimation 157
H. Yoo and E. Powers

Performance Analysis of Volterra Kernel Estimators with
Gaussian Inputs 162
A. Redfern and G. Zhou

Stochastic Resonance in a Discrete Time Nonlinear
SETAR(1,2,0,0) Model 166
S. Zozor and P.-O. Amblard

Texture Classification Using Third Order Correlation Tools 171
C. Coroyer, D. Declercq and P. Duvaut

On the Identifiability of Bilinear Stochastic Systems 176
C. Bourin and P. Bondon

Plenary: Signal Processing with Alpha-Stable Distributions: Current and Future Trends
Chrysostomos L. Nikias, *University of Southern California, USA*

TAP: System Identification & Deconvolution
Chair: Ed Powers, *University of Texas-Austin, USA*

Blind Channel Identification Based on Higher-Order Cumulant
Fitting Using Genetic Algorithms 184
S. Chen, Y. Wu and S. McLaughlin

Blind System Identification Using Fourth Order Spectral
Analysis of Complex Signals 189
C. Huet and J. Le Roux

Fractionally-Spaced Signal Reconstruction Based on Maximum Likelihood 194
B. Porat and B. Friedlander

A Class of Blind Deconvolution and Equalization Algorithms
for Nonminimum Phase Multi-Input Multi-Output Systems 199
C.-C. Feng and C.-Y. Chi

On-line Algorithms for Blind Deconvolution of Multichannel
Linear Time-Invariant Systems 204
Y. Inouye and T. Sato

Fourth-Order HOS Deconvolution in Wideband Communication Applications with Lattice Geometric Motion 209
J. Solinsky

Higher-Order Statistics-Based Deconvolution of Ultrasonic Nondestructive Testing Signals 214
A. Yamani, M. Bettayeb and L. Ghouti

Combination of HOS Based Blind Equalization Algorithms for Use in Mobile Communications 219
A. Nandi and C. Schmidt

System Reconstruction from Selected HOS Regions 224
H. Pozidis and A. Petropulu

Blind System Identification in an Impulsive Environment 229
J. Yin and A. Petropulu

Non-Causal ARMA Model Identification by Maximizing the Kurtosis 234
J.-L. Vuattoux and E. Le Carpentier

Identification of Multivariable Stochastic Linear Systems Using Integrated Polyspectrum Given Noisy Input-Output Data 239
J. Tugnait

On the Optimality of Bussgang and Super Exponential Blind Deconvolution Methods 244
G. Scarano and G. Jacovitti

Recursive Estimation Algorithm for FIR Systems Using the 3rd and 4th Order Cumulants 248
H. Kim, B. Jeon, T. Yang and K.-M. Sung

TAL: Alpha-Stable Distributions
Chair: Brian Sadler, *Army Research Laboratories, USA*

Zero-Order Statistics: A Signal Processing Framework for Very Impulsive Processes 254
J. Gonzalez, D. Griffith and G. Arce

Comparison between Asymmetric Generalized Gaussian (AGG) and Symmetric-alpha-Stable (S-/alpha/-S) Noise Models for Signal Estimation in Non Gaussian Environments 259
A. Tesei, R. Bozzano and C. Regazzoni

No Evidence of Stable Distributions in Radar Clutter 264
J. Ilow and H. Leung

On the Modeling of Network Traffic and Fast Simulation of Rare Events using alpha-Stable Self-Similar Processes 268
A. Karasaridis and D. Hatzinakos

TDE, DOA and Related Parameter Estimation Problems in Impulsive Noise 273
A. Swami and B. Sadler

TPL: Array Processing
Chair: Ben Friedlander, *University of California-Davis, USA*

Which Cumulants Should be Selected for Steering Vector Estimation? 280
T. Kaiser and J. Mendel

Improving the Threshold Performance of Higher-Order
Direction Finding Methods via Pseudorandomly Generated Estimator Banks 285
A. Gershman and J. Böhme

MUSIC's and Cramér-Rao Bound in Fourth-Order Cumulant Domain 290
H. Wu, Z. Bao and K. Yang

Adaptive Separation of an Unknown Number of Sources 295
Z. Malouche and O. Macchi

Near-Field Localization Using Inverse Filter Criteria-Based
Blind Separation And Cumulant Matching 300
A. Govindaraju and J. Tugnait

TPP: Source Separation & Array Processing

Chair: Joel LeRoux, *University of Nice, France*

Covariance of Finite-Sample Cumulants in Array-Processing 306
T. Kaiser and J. Mendel

DOA Estimation Using Blind Separation of Sources 311
M. Hirari and M. Hayakawa

Dimensionality Reduction in Higher-Order-Only ICA 316
L. De Lathauwer, B. De Moor and J. Vandewalle

Blind Source Separation with Noisy Sources 321
C. Servière

Blind Separation of Convolutive Mixtures: A Gauss-Newton Algorithm 326
S. Cruces and L. Castedo

Identifiability of the Superimposed Signals Using
Third-Order Cumulants 331
D. Korze and D. Zazula

Minimum Entropy Approach for Multisensor Data Fusion 336
Y. Zhou and H. Leung

Blind Separation of Sources Applied to Convolutive Mixtures
in Shallow Water 340
M. Gaeta, F. Briolle and P. Esparcieux

Separation of Sinusoidal Sources 344
C. Servière, V. Capdevielle and J.-L. Lacoume

A Linear Feedforward Neural Network with Lateral Feedback
Connections for Blind Source Separation 349
S. Choi and A. Cichocki

Fast Wideband Source Separation Based on Higher-Order Statistics 354
S. Bourennane, M. Frikel and A. Bendjama

Computational Savings Using Nonlinear Statistics in DOA Estimation 359
P. Campisi and G. Scarano

Behaviour of Higher Order Blind Source Separation Methods
in the Presence of Cyclostationary Correlated Multipaths 363
P. Chevalier, V. Capdevielle and P. Comon

Towards non Symmetrical Optimal Source Separation Contrasts 368
E. Moreau

Narrow Band Source Separation in Wide Band Context
Applications to Array Signal Processing 373
J. Galy, C. Adnet and E. Chaumette

Plenary: Cumulant Tensors

Pierre Comon, *EURECOM, France*

WAP: Detection, Estimation, Modeling & Analysis of Non-Gaussian Signals

Chair: Karl-Dirk Kammeyer, *University of Bremen, Germany*

On Representation For Nongaussian ARMA Processes 380
K. Chandrasekhar and S. Joshi

Factorizability of Complex Signals Higher (Even) Order Spectra : a Necessary and Sufficient Condition 385
J. Le Roux and C. Huet

The Estimation of Stable Distribution Parameters 390
S. Bates and S. McLaughlin

Fourier Series Based Nonminimum Phase Model for Second- and Higher-Order Statistical Signal Processing 395
C.-Y. Chi

Variability in Higher Order Statistics of Measured Shallow-Water Shipping Noise 400
L. Pflug, G. Ioup, J. Ioup and P. Jackson

Higher Order Moment Detection of Transients in Measured Shallow-Water Noise 405
L. Pflug, G. Ioup and J. Ioup

Detection and Classification of Spectrally Equivalent Processes: A Parametric Approach 410
M. Coulon, J.-Y. Tourneret and M. Ghogho

Robust Time-Frequency Representations for Signals in alpha-Stable Noise Using Fractional Lower-Order Statistics 415
D. Griffith, J. Gonzalez and G. Arce

Application of the Positive Alpha-Stable Distribution 420
R. Pierce

Asymptotic Distribution of the Hermite Normality Test 425
D. Declercq and P. Duvaut

High-Order Properties of M-Band Wavelet Packet Decompositions 430
D. Leporini and J.-C. Pesquet

Bussgang Gaussianity Test for Stationary Series 434
G. Giunta, G. Jacovitti and G. Scarano

Testing Multivariate Gaussianity with the Characteristic Function 438
A. Zoubir, C. Brown and B. Boashash

WAL: Modeling, Statistical Tests

Chair: Ananthram Swami, *Malgudi Systems, USA*

Testing the Gaussian Assumption for Self-similar Teletraffic models 444
S. Bates and S. McLaughlin

Sampling Jitter Detection using Higher-Order Statistics 448
J. Tourneret, A. Ferrari and B. Lacaze

Playing with Long Range Dependence and HOS 453
P. Amblard, J. Brossier and J. Lacoume

Higher and Lower-Order Properties of the Wavelet Decomposition of Self-Similar Processes 458
B. Pesquet-Popescu and P. Larzabal

Detection of a Common Non-Gaussian Signal in Two Sensors Using the Bootstrap 463
H.-T. Ong, D. Iskander and A. Zoubir

Author Index **469**

Organizers and Co-Chairs

Keh-Shin Lii
Department of Statistics
University of California - Riverside
Riverside, CA 92521 USA
Tel: (909) 787-3836
Fax: (909) 787-3286
E-mail: ksl@ucrstat.ucr.edu
URL: http://cnas.ucr.edu/~stat/lii.html

Jerry M. Mendel
Department of Electrical Eng. - Systems
University of Southern California
Los Angeles, CA 90089-2564 USA
Tel: (213) 740-4445
Fax: (213) 740-4651
E-mail: mendel@sipi.usc.edu
URL: http://sipi.usc.edu/~mendel

Technical Program Co-Chairs

Dimitrios Hatzinakos
Dept. of Electrical & Computer Engineering
University of Toronto
Toronto, Ontario, Canada M5S 1A4
Tel: (416) 978-1613
Fax: (416) 978-4425
E-mail: dimitris@comm.toronto.edu
URL: http://www.comm.toronto.edu/
~dimitris/dimitris.html

Jitendra K. Tugnait
Department of Electrical Engineering
Auburn University
Auburn, Alabama 36849 USA
Tel: (334) 844-1846
Fax: (334) 844-1809
E-mail: tugnait@eng.auburn.edu
URL: http://www.eng.auburn.edu/~tugnait/

Registration/Treasurer

Keh-Shin Lii
University of California, Riverside, USA

Publicity

Dimitrios Hatzinakos
University of Toronto, Canada

Publication

Athina P. Petropulu
Electrical and Computer Engineering Dept.
Drexel University
Philadelphia, PA 19104 USA
Tel: (215) 895-2358
Fax: (215) 895-1695
E-mail: athina@cbis.ece.drexel.edu

Local Arrangements

Brent Peterson
Department of Electrical Engineering
University of Calgary
Calgary, Alberta Canada T2N 1N4
Tel: (403) 289-3140
E-mail: b.peterson@ieee.org

International Representatives

Stephen McLaughlin, *UK*
Pierre Comon, *France*
Yujiro Inouye, *Japan*
Vito Cappellini, *Italy*

Technical Committee

Gaetano Scarano, *Italy*
Javier R. Fonollosa, *Spain*
Anathram Swami, *USA*
Jean-Francois Cardoso, *France*
Chong-Yung Chi, *Taiwan*
Abdelhak M. Zoubir, *Australia*

PLENARY

High-order Cyclostationary Signal Analysis: An overview

Dr. Georgios Giannakis, *Univ. of Virginia (USA)*

Processes encountered in statistical signal processing, communications, and time series analysis applications, are often assumed stationary. Due to the varying nature of physical phenomena and certain manmade operations however, time-invariance and the related notion of stationarity are often violated in practice. In this talk, I will focus on cyclostationary processes which are characterized by the periodicity they exhibit in their second- and/or higher-order statistical descriptors. Periodicity is omnipresent in real life problems entailing phenomena and operations of repetitive nature: communications, geophysical and atmospheric sciences (hydrology, oceanography, meteorology, climatology), rotating machinery, econometrics, and biological systems. Background material will deal with polyspectral representations, sample estimation of cyclic statistics, testing a time series for second- and high-order cyclostationarity, and noise suppression by joint exploitation of cyclostationarity and non-Gaussianity. The diversity offered by periodic variations will be emphasized in the context of blind identification of time-invariant and periodically varying systems and separation of cyclostationary signals on the basis of their cycles. Specific applications will include time delay estimation, harmonic retrieval in the presence of multiplicative noise, modeling with systematically missing observations, polynomial phase signals for modeling motion induced variations, equalization of random channels, and generalized differential encoding of communication

MAP: Applications of HOS

Chair: Steve McLaughlin
University of Edinburgh, Scotland

Autoregressive Modeling of Lung Sounds Using Higher-Order Statistics: Estimation of Source and Transmission

Leontios J. Hadjileontiadis and Stavros M. Panas

Department of Electrical & Computer Engineering, Division of Medical Signal and Image Processing, Aristotle University of Thessaloniki, GR-540 06 Thessaloniki, Greece.
(Tel/Fax: +3031-9963-03/12, e-mail: leontios@ccf.auth.gr, panas@vergina.eng.auth.gr)

Abstract

The use of higher-order statistics in an autoregressive modeling of lung sounds is presented, resulting in a characterization of their source and transmission. The lung sound source in the airway is estimated using the prediction error of an all-pole filter based on higher-order statistics (AR-HOS), while the acoustic transmission through the lung parenchyma and chest wall is modeled by the transfer function of the same AR-HOS filter. The parametric bispectrum, using the estimated a_i coefficients of the AR-HOS model, is also calculated to elucidate the frequency characteristics of the modeled system. The implementation of this approach on pre-classified lung sound segments in known disease conditions, selected from teaching tapes, was examined. Experiments have shown that a reliable and consistent with current knowledge estimation of lung sound characteristics can be achieved using this method, even in the presence of additive Gaussian noise.

1. Introduction

Pulmonary diagnosis is often based on acoustical pulmonary signals analysis, since the generated acoustical energy, produced by air flow during inspiration and expiration, is highly correlated with pulmonary dysfunction. This dysfunction is caused by some anatomical or physiological changes in the pulmonary system and it is characterized by the changes in the acoustical properties of the various parts or organs involved [1]. The description of the influence of a disease in the production or/and transmission characteristics of lung sounds is of great importance, since it can give insight to the causes of the production mechanism of the disease.

In this paper, a characterization of source and transmission of lung sounds by means of autoregressive modeling based on higher-order statistics (AR-HOS) is presented. The model performance is evaluated through pre-classified lung sounds included in teaching tapes, recorded from patients with various kinds of lung diseases. Results from the applied AR-HOS model proved the establishment of an efficient and objective modeling approach of lung sounds.

2. AR-HOS modeling of lung sounds

In this section, the description of the proposed AR-HOS modeling of lung sounds is presented.

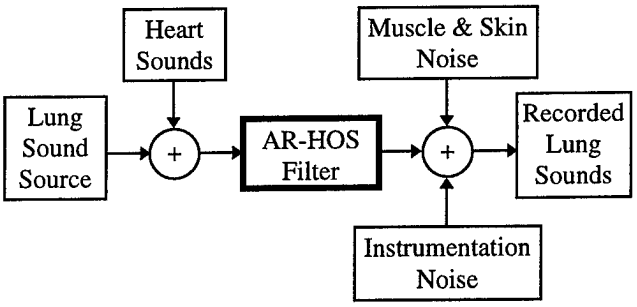


Fig. 1. Schematic diagram of the AR-HOS model employed for lung sounds analysis.

According to Fig. 1, lung sounds, originated inside the airways of the lung, are modeled as the input to an all-pole filter (AR-HOS), which describes the transmission of lung sounds through the parenchyma and chest wall structure. The output of this filter is considered to be the lung sounds recorded at the chest wall. These sounds also contain heart sounds interference, the reduction of which has been addressed recently by the authors using fourth-order statistics in [2] and other pending publications. Muscle and skin noise along with instrumentation noise are modeled as an additive Gaussian noise.

Using the model hypothesized above, given a signal sequence of lung sounds at the chest wall, an autoregressive analysis can be applied to compute the model parameters, and therefore the source and transmission filter characteristics can separately be estimated [3]. This becomes clear from the following analysis.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2.1 Modeling of transmission of lung sounds

The AR-HOS prediction filter performs autoregressive prediction based on third-order statistics (AR-TOS). The equation describing the autoregressive model is:

$$y_n + \sum_{i=1}^{p} a_i y_{n-i} = w_n \quad , \quad a_0 = 1 , \qquad (1)$$

where, y_n represents a p^{th} order AR process of N samples *(n=0,...,N-1)*, a_i are the coefficients of the AR-model, and w_n are i.i.d., non-Gaussian, third-order stationary, zero-mean, with $E\{w_n^3\}=\beta\neq 0$ and y_n independent of w_l for $n<l$. Since w_n is third-order stationary, y_n is also third-order stationary, assuming it is a stable AR model. For the model of Eqn. (1) we can write the *cumulant-based 'normal' equations* [4]:

$$\sum_{i=0}^{p} a_i R(\tau_1 - i, \tau_2) = 0 , \tau_1 = 1, \ldots, p \ \& \ \tau_2 = -p, \ldots, 0 , \qquad (2)$$

where, $R(\tau_1,\tau_2)$ is the third order cumulant sequence (TOS) of the AR process. In practice we use sample estimates of the cumulants. Eqn. (2) yields consistent estimates of the AR parameters maintaining the orthogonality of the prediction error sequence to an instrumental process derived from the data [5]. The calculation of the order p of the AR-TOS model is reduced to a rank determination problem. According to Giannakis and Mendel [6], the rank of a matrix $\boldsymbol{C}_e$ formed by exact third-order cumulants is equal to p, even when only $\overline{p}$, the upper bound of p, is known. A subjective rule to select the 'effective' AR order p is to find the largest drop among two successive normalized singular values of $\hat{C}_e$. Estimations of the transfer function $H(\omega)$ of the system and the normalized parametric bispectrum C_3^y can be derived using the following equations, respectively [7]:

$$\hat{H}(\omega) = \frac{1}{1 + \sum_{i=1}^{p} \hat{a}_i e^{-j\omega i}} \quad , \quad |\omega| \le \pi , \qquad (3)$$

$$\frac{1}{\hat{\beta}} \hat{C}_3^y(\omega_1, \omega_2) = \hat{H}(\omega_1)\hat{H}(\omega_2)\hat{H}^*(\omega_1 + \omega_2) , \qquad (4)$$

where, * denotes the complex conjugate part.

The most profound motivations behind the use of third-order statistics in the estimation of a_i are [7]: i) *suppression of Gaussian noise,* since third-order statistics of Gaussian signals are identically zero, and ii) *preservation of the true phase character of the signal,* unlike the second-order statistics (autocorrelation). Hence, when the analysis waveform consists of a non-Gaussian signal in additive symmetric noise (e.g. Gaussian), the parameter estimation of the original signal using third-order statistics takes place in a high signal-to-noise ratio (SNR) domain and the whole parametric presentation of the process is more accurate and reliable [4].

2.2 Modeling of source of lung sounds

The model used for the lung sounds originated inside the airways of the lung is depicted in Fig. 2. According to Fig. 2, the lung sound source consists of three kinds of noise sequences [11]. The first sequence (periodic impulse) describes the wheezes and ronchus sources, since they have characteristic distinct pitches associated with them, and they are believed to be produced by periodic oscillations of the air and airway walls [8]-[11]. The second sequence (random intermittent impulse) describes the crackle sources since they believed to be produced by sudden opening/closing of airways or bubbling of air through extraneous liquids in the airways, both phenomena associated with sudden intermittent bursts of sounds energy [8]-[11]. Finally, the third sequence (white non Gaussian noise) describes the breath sound sources since they are believed to be produced by turbulent flow in a large range of airway dimensions [8]-[11]. An additive combination of these sequences can produce almost all kinds of lung sounds, resulting in a complete description of the lung sound sources.

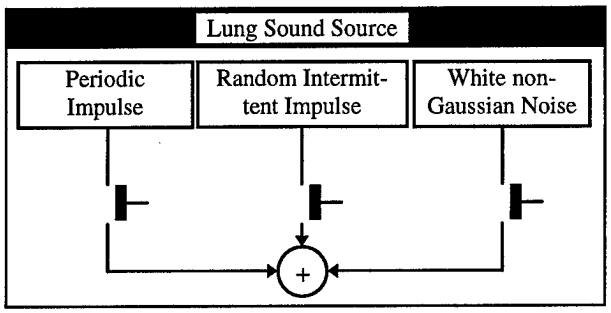


Fig. 2. The lung sound source model.

The estimation of the AR-TOS model input (lung sound source) can be derived from the prediction error of Eqn. (1), by means of inverse filtering.

3. Implementation

Pre-classified signals, each *25s* in duration, representing typical diseases affecting the airways, parenchyma, and chest wall, were selected from teaching tapes [8]-[10]. After antialiasing filtering, the signals were digitized with a 12-Bit A/D converter at a sampling rate of *2.5KHz*. The signals (sections of *N=1024, 2048, 4096*) were subjected to AR-HOS analysis and third-order statistics were estimated by averaging the cumulants of successive subsections (*M=8*). The value of order p ranged from *2* to *11*.

4. Results and Discussion

The results obtained by means of AR-TOS modeling of lung sounds are presented in Table I.

Table I. Results of AR-TOS modeling of lung sounds.

Description of segment	Associated Phenomenon	Estimated Source	Transmission Filter	p
Bronchial (insp.)	Normal Breathing	White NGN	LPF (0-550Hz)	2
Bronchial (insp.)	Pneumonia (early)	White NGN	LPF (0-700Hz)	4
Vesicular (insp.)	Normal Breathing	White NGN	LPF (0-500Hz)	4
Broncho-vesicular (insp.)	Normal Breathing	White NGN	LPF (0-450Hz)	2
Tracheal (insp.)	Normal Breathing	White NGN	LPF (0-700Hz)	6
Wheeze (expir.)	Asthma	Periodic Impulse (T=0.004s)	LPF (0-550Hz) d.p. at 250 500Hz	10
Early Inspiratory Crackles (insp.)	Airway Obstruction	Random Intermittent Impulses + White NGN	LPF (0-600Hz)	4
Late Inspiratory Crackles (insp.)	Airway Restriction	Random Intermittent Impulses + White NGN	LPF (0-600Hz)	11
Fine Crackles (insp.)	Pulmonary Fibrosis	Random Intermittent Impulses + White NGN	BPF (350-700Hz) c.f. 530Hz	2
Coarse Crackles (insp.)	Chronic Bronchitis	Random Intermittent Impulses + White NGN	BPF (200-400Hz) c.f. 300Hz	2
Squawks (insp.)	Allergic Alveolitis and Interstitial fibrosis	Random Intermittent Impulses + White NGN	BPF (180-480Hz) c.f. 380Hz	2
Stridors (insp.)	Croup	Periodic Impulse (T=0.006s)	LPF with d.p. at 380 760Hz	4
Pleural Friction Rub (insp.)	Inflammatory Pleural	Random Intermittent Impulses + White NGN	LPF (0-400Hz)	2

NGN: Non Gaussian Noise; c.f.: central frequency; L/BPF: Low/Band Pass Filter; d.p.: distinct peaks.

Representative results of AR-TOS modeling of lung sounds (normal bronchial sounds, wheezes, fine crackles and squawks) are shown in Figs 3, 4, 5, and 6, respectively. The lack of energy noted in the low-frequency range (0-100Hz) is attributed to analog highpass filters used prior to the recording of these sounds to avoid heart sounds, muscle, and skin noise [8]-[10].

Fig. 3 shows the results of AR-TOS modeling of an inspiratory segment of normal bronchial breath sounds [Fig. 3(a) and (c)]. The source [Fig. 3(b) and (d)] was found to be random white noise. The order p of the AR-TOS model was found equal to *2* [Fig. 3(e)]. The estimated filter response [Fig. 3(f)] was noted to be essentially the same over the breath segment [Fig. 3(g)], indicating a primarily stationary filter response over the inspiratory phase. The typical response [Fig. 3(f)] was found to have a lowpass characteristic, with a frequency band range around 0-550Hz. According to bispectrum [Fig. 3(h)], the power of low-frequencies was found to be rather low, while the central frequency was located around 250Hz.

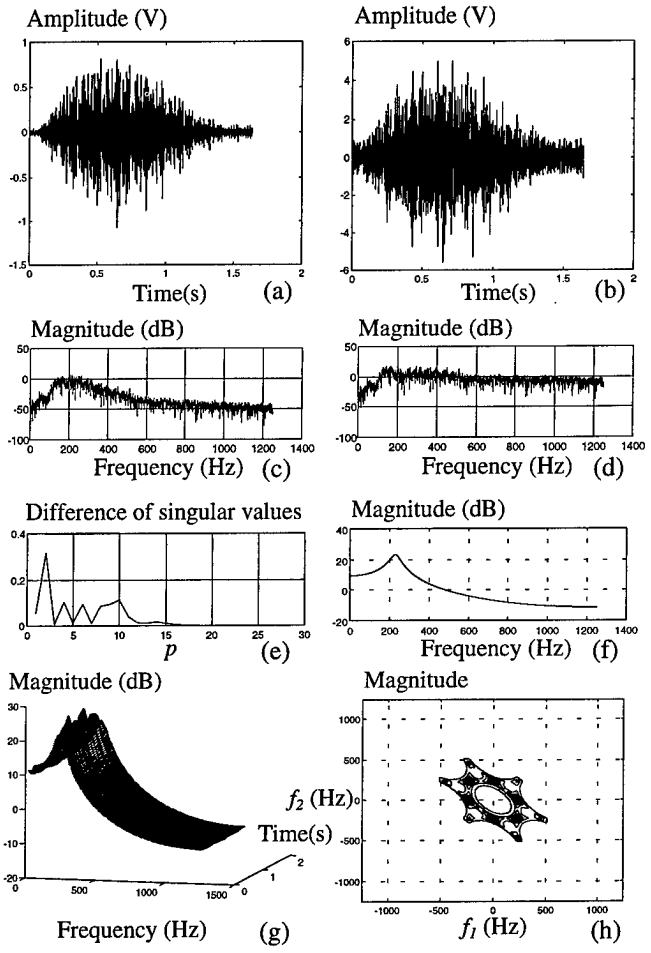


Fig. 3. (a) A time section of 1.63s of normal bronchial sounds (inspiration). (b) The estimated lung sound source. (c) The estimated power spectrum of the recorded sound. (d) The estimated power spectrum of the estimated source. (e) The estimation of *p*. (f) The estimated filter frequency response. (g) The estimated filter frequency response vs. time. (h) The estimated magnitude of the parametric bispectrum of recorded sounds.

Fig. 4 shows the results of AR-TOS modeling of an expiratory asthmatic segment [Fig. 4(a) and (c)]. The estimated source [Fig. 4(b1,b2) and (d)] was found to be a periodic train of impulses (with period T≈0.004s). The order p of the AR-TOS model was found equal to *10* [Fig. 4(e)]. The estimated filter response [Fig. 4(f)] was noted to be essentially the same over the breath segment [Fig. 4(g)], indicating a primarily stationary filter response over the expiratory phase. The typical response [Fig. 4(f)] was found to have a lowpass characteristic, with a frequency band range around 0-550Hz, and two distinct resonance peaks at 250Hz and 500Hz. According to bispectrum [Fig. 4(h)], the power of the signal is located in the resonance frequencies, result which is consistent with the production mechanism of wheezes, i.e. distinct resonance in the transmission path (airway walls) [8]-[10].

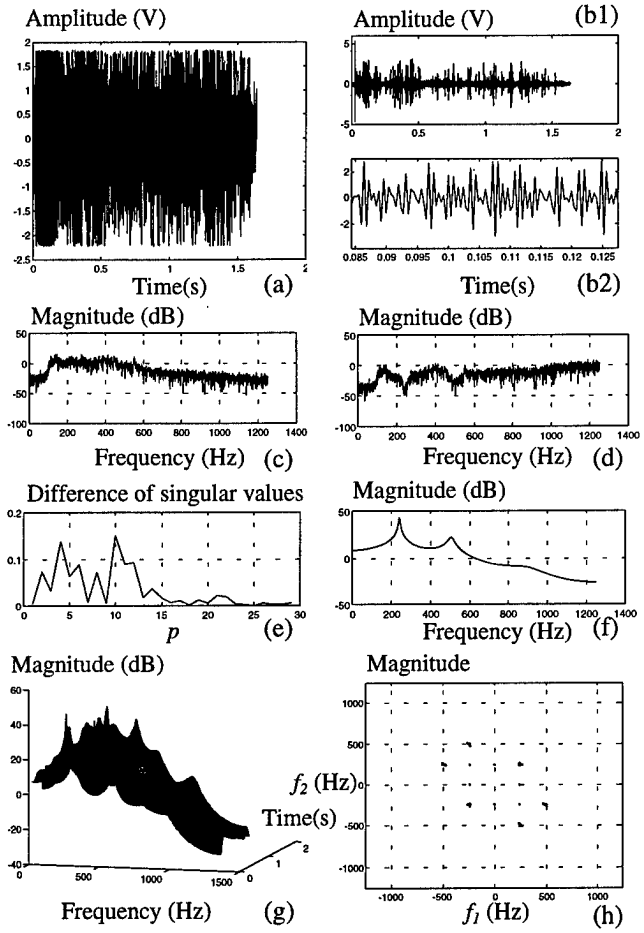


Fig. 4. (a) A time section of 1.63s of asthmatic wheezes (expiration). (b)-(h) The same as (b)-(h) in Fig. 2.

In case of the inspiratory segment of fine crackles [Fig. 5(a) and (c)], the estimated source waveform [Fig. 5(b) and (d)] contains impulsive bursts, corresponding to fine crackles, and white NGN. This could be explained by the associated with fine crackles phenomenon of explosive reopening of small airways that had closed during the previous expiration. The order p of the AR-TOS model was found equal to *2* [Fig. 5(e)]. The estimated filter response [Fig. 5(f)] was noted to be essentially the same over the breath segment [Fig. 5(g)], indicating a primarily stationary filter response over the inspiratory phase. The typical response [Fig. 5(f)] was found to have a bandpass characteristic, with a frequency band range around 350-700Hz, with central frequency at 530Hz. According to bispectrum [Fig. 5(h)], the power of the signal is shifted to higher frequencies. These results are consistent with the understanding of pulmonary fibrosis, since the associated abnormal airway closure that precedes the 'crackling' reopening is due to increased lung stiffness, which probably causes the transmission of higher frequencies [8]-[10].

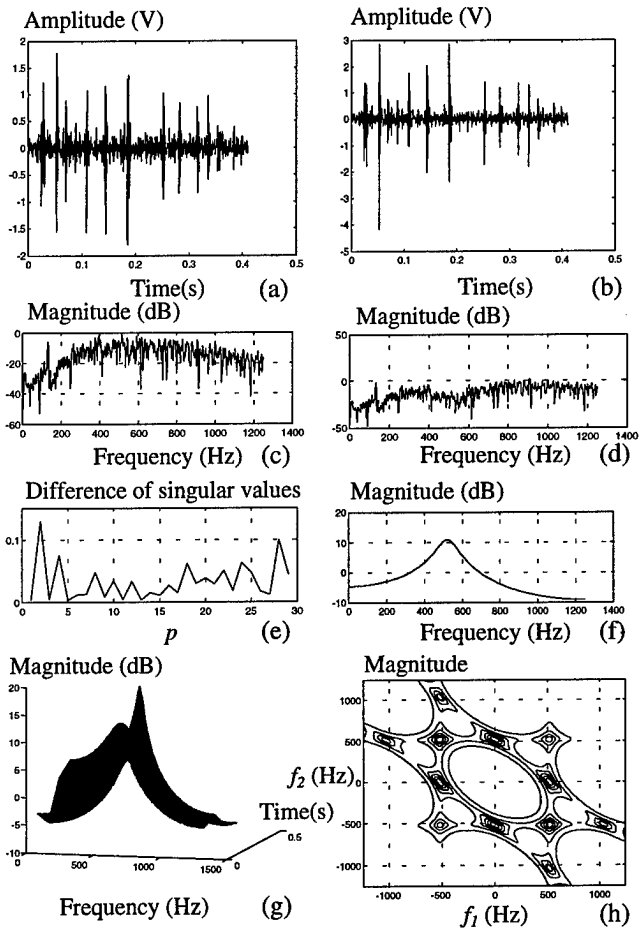


Fig. 5. (a) A time section of 0.4096s of fine crackles (inspiration). (b)-(h) The same as (b)-(h) in Fig. 2.

In case of squawks [Fig. 6(a) and (c)], the estimated source [Fig. 6(b) and (d)] combines impulsive bursts followed by short, almost exponential, decaying periodic train of impulses. This result is consistent with the accepted

theory that, squawks are produced by the explosive opening and decaying fluttering of an unstable airway [8]. The order p of the AR-TOS model was found equal to 2 [Fig. 6(e)]. The estimated filter response [Fig. 6(f)] was noted to be essentially the same over the breath segment [Fig. 6(g)], indicating a primarily stationary filter response over the inspiratory phase. The typical response [Fig. 6(f)] was found to have a bandpass characteristic, with a frequency band range around 180-480Hz, with central frequency at 380Hz. According to bispectrum [Fig. 6(h)], the power of the signal is shifted to lower frequencies, with smooth characteristics.

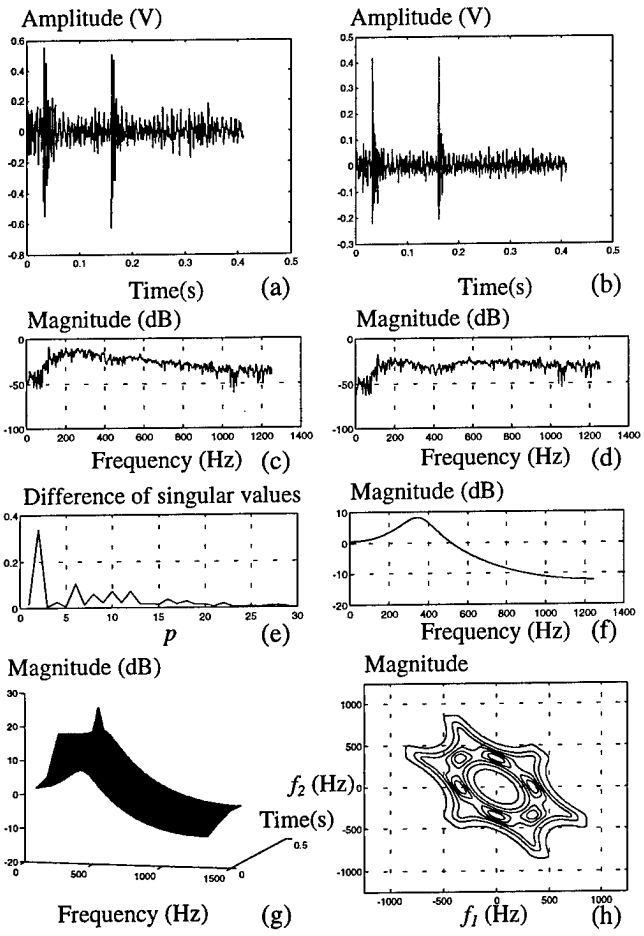


Fig. 6. (a) A time section of 0.4098s squawks (inspiration). (b)-(h) The same as (b)-(h) in Fig. 2.

The robustness of the AR-HOS model to symmetric noise has been tested through the same signals, contaminated with additive colored Gaussian Noise (0, 0.03). Figs. 7(a) and (b) depict the estimated parametric bispectrum of AR-HOS modeling of noisy fine crackles and squawks, from where it can be noticed that, the main frequency characteristics remain unchanged, compared to Figs 5(h) and 6(h), respectively. From the derived results it is evident that the AR-HOS modeling of lung sounds characterizes their source and transmission efficiently, providing useful diagnostic information to the clinicians.

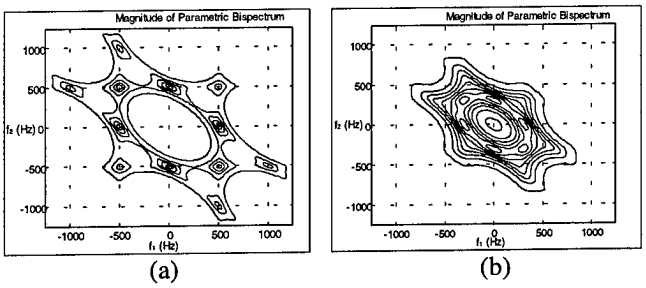


Fig. 7. Estimated parametric bispectrum of noisy (a) Fine crackles, (b) Squawks.

5. Summary

An autoregressive analysis of lung sounds based on higher-order statistics (HOS) was presented in this work. Experiments proved the attribute of HOS in estimating source and transmission characteristics of lung sounds efficiently, even in the presence of additive symmetric noise.

References

[1] A. Cohen, "Signal Processing Methods for Upper Airway and Pulmonary Dysfunction Diagnosis," *IEEE Eng. in Medicine & Biology Mag.*, pp. 72-75, March 1990.

[2] L. J. Hadjileontiadis, and S. M. Panas, "Adaptive Reduction of Heart Sounds from Lung Sounds using Fourth-Order Statistics," *IEEE Trans. Biomed. Eng.*, in press (July 1997).

[3] V. K. Iyer, P. A. Ramamoorthy, and Y. Ploysongsang, "Autoregressive Modeling of Lung Sounds: Characterization of Source and Transmission," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 11, pp. 1133-1137, Nov. 1989.

[4] A. Swami, J. M. Mendel, and Ch. L. Nikias, *Higher-Order Spectral Analysis Toolbox*, The MathWorks, USA, 1995.

[5] A. Swami, *System Identification Using Cumulants*, Ph.D. Dissertation, Univ. of South. California, pp. 107-108, 1988.

[6] G. B. Giannakis, and J. M. Mendel, "Cumulant-Based Order Determination of Non-Gaussian ARMA Models," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 8, pp. 1411-1423, August 1990.

[7] Ch. L. Nikias and A. M. Petropulu *Higher-Order Spectra Analysis: A nonlinear signal processing framework*, New Jersey: PTR Prentice-Hall, Inc., USA, 1993.

[8] S. S. Kraman, *Lung Sounds: An Introduction to the Interpretation of the Auscultatory Finding*, American College of Chest Phys., Northbrook, USA, workbook & tape, 1993.

[9] S. Lehrer, *Understanding lung sounds*, 2nd ed., W.B. Saunders Co., Philadelphia, Pa. 19106, workbook & tape, 1993.

[10] A G. Tilkian, and M. B. Conover, *Understanding Heart Sounds and Murmurs with an Introduction to Lung Sounds*, 3rd ed., W.B. Saunders Co., Philadelphia, Pa. 19106, workbook & tape, 1993.

[11] L. J. Hadjileontiadis, *Analysis and Processing of Lung Sounds using Higher-Order Statistics-Spectra and Wavelet Transform*, Ph.D. Dissertation, A.U.T., Thessaloniki, Greece, chap. 6, pp. 139-175, May 1997, (in Greek).

Multipath Time Delay Estimation using Higher Order Statistics

Ying-Chang Liang, A. Rahim Leyman and Boon-Hee Soong
School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798
Email: eycliang(earleyman, ebhsoong)@ntuvax.ntu.ac.sg

Abstract

This paper addresses the problem of time delay estimation (TDE) in spatially correlated noises. Two cases are considered: (i) TDE of non-Gaussian signal in spatially correlated Gaussian noises; and (ii) TDE of Gaussian signal in spatially correlated non-Gaussian noises. For the first case, a new approach based upon the use of higher order statistics of the measurements is proposed; for the second case, a hybrid approach by using higher and second order statistics of the measurements is suggested. Simulation examples are presented to illustrate the effectiveness of these approaches.

1. Introduction

Time delay estimation (TDE) between the received signals at two spatially separated sensors has found applications in many areas such as radar, sonar, biomedicine and geophysics. The traditional methods [1] are based on shifting one measurement with respect to the other to compare the similarities between the two records and then choosing the shift at which best match occurs. In theory, these methods yield consistent time delay estimate when the noises are uncorrelated or their cross-correlation functions are known; however, they may yield ambiguous results when the noise processes are spatially correlated.

Recently, higher order statistics have found wide applications in signal processing field. In those practical settings where the noise processes are jointly Gaussian, possibly correlated and the signal is non-Gaussian, several TDE approaches have been proposed based on higher order cumulants of the measurements [2,3]. Among these approaches, the time domain approach of [2] may fail to work when the cumulant matrix constructed is ill conditioned. On the other hand, although Tugnait's approaches [3] do not suffer from this limitation, they are applicable only to fourth order cumulant case. Motivated by the above observations, the first objective of this paper is to study the problem of multipath TDE of non-Gaussian signals in spatially correlated Gaussian noises, referred to as NSGN case, and to propose new normal equations which yield unique solution and are applicable to any higher than second order case;

On the other hand, colored non-Gaussian noise environments are known to exist widely in practical applications. For example, Hinich *et al.* [4] have shown that the ship radiated noise is non-Gaussian with nonzero bispectrum. Some types of underwater noises are definitely non-Gaussian [5]. Moreover, the Gaussianity of signal has been assumed and tested in conventional TDE methods [1]. Therefore, another obvious problem needed to be solved is that of estimating multipath time delays of Gaussian signal in unknown spatially correlated non-Gaussian noises, referred to as GSNN case, which is the second objective of this paper. We propose a hybrid approach for solving this problem by using second and higher order statistics of the measurements.

2. TDE in NSGN Case

In this section, we study the problem of TDE of non-Gaussian signal in spatially correlated Gaussian noises.

2.1. Model Assumptions

Let $x(n)$ and $y(n)$ denote two sensor measurements satisfying

$$x(n) = s(n) + w_1(n) \tag{1}$$

$$y(n) = \sum_{j=1}^{M} A_j s(n - D_j) + w_2(n) \tag{2}$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

where $s(n)$ is an unknown signal, M is the multipath number, A_j and $s(n-D_j)$ are the amplitude attenuation and delayed (or advanced) version of the jth path reflection, respectively.

ASa1) The signal $s(n)$ is a zero mean, non-Gaussian linear random process, given by

$$s(n) = \sum_k h(k)e(n-k) \tag{3}$$

where $\sum_{i=-\infty}^{\infty} |h(i)| < \infty$, $H(0) \triangleq \sum_{i=-\infty}^{\infty} h(i) \neq 0$, and $e(n)$ is zero mean, independently and identically distributed (i.i.d.), non-Gaussian process with kth ($k > 2$) order cumulant denoted by γ_{ke} ($\gamma_{ke} \neq 0$);

ASa2) The noises $w_1(n)$ and $w_2(n)$ are zero mean *possibly spatially correlated* stationary Gaussian processes;

ASa3) $s(n)$ is statistically independent of $w_1(n)$ and $w_2(n)$.

Our task is to estimate the multipath time delays, $D_j, j = 1, \cdots, M$, from $x(n)$ and $y(n)$ only.

2.2. New Normal Equations

Let $c_{k,x}(\tau_1, \cdots, \tau_{k-1})$ denote the kth order cumulant of $x(n)$

$$\begin{aligned} & c_{k,x}(\tau_1, \cdots, \tau_{k-1}) \\ \triangleq\ & cum\{x(n), x(n+\tau_1), \cdots, x(n+\tau_{k-1})\} \end{aligned} \tag{4}$$

and $c_{k,xy}(\tau_1, \cdots, \tau_{k-1})$ the kth order cross cumulant of $x(n)$ and $y(n)$

$$\begin{aligned} & c_{k,xy}(\tau_1, \cdots, \tau_{k-1}) \\ \triangleq\ & cum(x(n), \cdots, x(n+\tau_{k-2}), y(n+\tau_{k-1})) \end{aligned} \tag{5}$$

where $cum(\cdot)$ denote the cumulant operator. Under ASa1), a simple relationship between the higher- and lower-order cumulants of $s(n)$, which is called the projection property of cumulants, is given by, for $k > l$,

$$\begin{aligned} & c_{l,s}(\tau_1, \cdots, \tau_{l-1}) \\ =\ & \alpha_{lk} \sum_{\tau_l=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} c_{k,s}(\tau_1, \cdots, \tau_{k-1}) \end{aligned} \tag{6}$$

where $\alpha_{lk} = \gamma_{le}/\gamma_{ke}[H(0)]^{l-k}$. Especially, when $l = 2$ and $k = 3, 4$, we have

$$c_{2,s}(\tau_1) = \alpha_{23} \sum_{\tau_2=-\infty}^{\infty} c_{3,s}(\tau_1, \tau_2) \tag{7}$$

and

$$c_{2,s}(\tau_1) = \alpha_{24} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} c_{4,s}(\tau_1, \tau_2, \tau_3), \tag{8}$$

respectively, where $\alpha_{23} = \gamma_{2e}/[\gamma_{3e}H(0)]$, and $\alpha_{24} = \gamma_{2e}/[\gamma_{4e}(H(0))^2]$. In practical computation, "∞" and $c_{3,s}(\cdot), c_{4,s}(\cdot)$ in (7) and (8) are, respectively, replaced by a positive integer L and $\hat{c}_{3,s}(\cdot), \hat{c}_{4,s}(\cdot)$, which are computed from N data samples. It is known that the estimate $\hat{c}_{2,s}(\tau_1)$ is an asymptotically unbiased and strongly consistent estimate of the true autocorrelation of $s(n)$ (within a scale factor), when $L = O(N^{1/(4+\delta)}), \delta > 0$.

It is well known that the delayed signal, $s(n-D_j)$, can be expressed approximately as

$$s(n-D_j) = \sum_{i=-P}^{P} sinc(i-D_j)s(n-i) \tag{9}$$

where $sinc(v) := sin(\pi v)/(\pi v)$, and $P \geq max(D_1, \cdots, D_M)$. From (1) and (2), we have

$$\begin{aligned} & c_{k,xy}(\tau_1, \cdots, \tau_{k-1}) \\ =\ & \sum_{i=-P}^{P} \sum_{j=1}^{M} A_j sinc(i-D_j) \\ & \times c_{k,x}(\tau_1, \cdots, \tau_{k-2}, \tau_{k-1}-i) \end{aligned} \tag{10}$$

because of the Gaussianarity of the noises. By letting

$$r_x(\tau_{k-1}) = \sum_{\tau_1=-L}^{L} \cdots \sum_{\tau_{k-2}=-L}^{L} c_{k,x}(\tau_1, \cdots, \tau_{k-1}) \tag{11}$$

and

$$r_{xy}(\tau_{k-1}) = \sum_{\tau_1=-L}^{L} \cdots \sum_{\tau_{k-2}=-L}^{L} c_{k,xy}(\tau_1, \cdots, \tau_{k-1}) \tag{12}$$

it follows that from (10) and (6)

$$r_{xy}(\tau) = \sum_{i=-P}^{P} \sum_{j=1}^{M} A_j sinc(i-D_j) r_x(\tau - i) \tag{13}$$

or

$$r_{xy}(\tau) = \sum_{i=-P}^{P} a_i r_x(\tau - i) \tag{14}$$

where

$$a_i = \sum_{j=1}^{M} A_j sinc(i-D_j) \tag{15}$$

Eqn. (14) forms the new normal equations for time delay estimation using higher order statistics. Concatenating (14) for $\tau = -P, \cdots, P$, we obtain the following matrix equation

$$\mathbf{Ax} = \mathbf{b} \tag{16}$$

where

$$\mathbf{A} = \{r_x(i-j)\}_{i,j=1}^{2P+1} \tag{17}$$

$$\mathbf{x} = [a_{-P}, \cdots, a_P]^T \tag{18}$$

$$\mathbf{b} = [r_{xy}(-P), \cdots, r_{xy}(P)]^T \tag{19}$$

Note that from (11), (16) and (6), $\mathbf{A}$ is equal to a nonsingular autocorrelation matrix (within some non-zero constant scale), which guarantees the uniqueness of the solution to the new normal equations. If D_j's are all integer, then $a_i = A_j$, for $i = D_j, j = 1, 2, \cdots, M$ and $a_i = 0$ for others; otherwise, interpolation computations based upon (15) can be employed for determining D_j's.

2.3. Algorithm 1

Algorithm 1 for TDE in NSGN case is now summarized as follows:
(1) estimate $\hat{c}_{k,x}(\cdot)$ and $\hat{c}_{k,xy}(\cdot)$ from $x(n)$ and $y(n)$, $n = 0, 1, \cdots, N$;
(2) estimate $\hat{r}_x(\cdot)$ and $\hat{r}_{xy}(\cdot)$ using (11) and (12);
(3) estimate a_i, $i = -P, \cdots, P$ by solving the matrix equation of (15);
(4) choose the time delays D_j's and the attenuations A_j's from the definition of a_i's.

3. TDE in GSNN Case

In this section, we study the problem of TDE of Gaussian signal in spatially correlated non-Gaussian noises.

3.1. Model Assumptions

The measurement models are the same as (1) and (2), while the conditions on the models are as follows:
ASb1) The signal $s(n)$ is a zeromean, stationary Gaussian process.
ASb2) The noises $w_1(n)$ and $w_2(n)$ are spatially correlated non-Gaussian processes with non-zero kth ($k > 2$) cumulants, and $w_2(n)$ can be generated by passing $w_1(n)$ through a linear finite dimensional asymptotically stable system with impulse response g_i, i.e.,

$$w_2(n) = \sum_{i=-Q}^{Q} g_i w_1(n-i); \tag{20}$$

in addition, $\sum_{i=-Q}^{Q} g_i z^{-i} \neq 0$ for $|z| > 1$.
ASb3) $s(n)$ is independent of $w_1(n)$ and $w_2(n)$.

3.2. Theoretical Results

From the above assumptions, the measurement models (1) and (2) can be rewritten as

$$x(n) = s(n) + \sum_{i=-Q}^{Q} g_i w(n-i), \tag{21}$$

$$y(n) = \sum_{i=-P}^{P} a_i s(n-i) + w(n). \tag{22}$$

The problem is to estimate a_i's, thus the time delays D_j's from $x(n)$ and $y(n)$ only. Note this problem is similar to that of multichannel signal separation in [6], the differences lie in the probability distributions of the two components of $s(n)$ and $w(n)$, where in [6] $s(n)$ and $w(n)$ are all non-Gaussian, and in this section $s(n)$ is Gaussian while $w(n)$ is non-Gaussian. It is pointed out that the separation approaches developed in [6] are not applicable to our case since the coefficient matrix to the related normal eauations are zero and obviously singular; however, the hybrid approach developed in this section is also applicable for multichannel signal separation.
Note that the signal $s(n)$ is Gaussian and the noises parts $w_1(n)$ and $w_2(n)$ are non-Gaussian, we may use the method developed in the previous section to estimate g_i's first. Comparing (1) and (2) with (21) and (22), we have

$$r_{yx}(\tau) = \sum_{i=-Q}^{Q} g_i r_y(\tau - i) \tag{23}$$

where $r_{yx}(\cdot)$ and $r_y(\cdot)$ are similar to $r_{xy}(\cdot)$ and $r_x(\cdot)$ as in (11) and (12), respectively. Concatenating (23) for $\tau = -Q, \cdots, Q$, we may obtain the unique solution of g_i's since the corresponding coefficient matrix is nonsingular. Once g_i's are available, one may use the cross-correlation based approach to estimate a_i's. To this end, we prefilter $y(n)$ in (22) as

$$\begin{aligned} \tilde{y}(n) &= \sum_{j=-Q}^{Q} g_j y(n-j) \\ &= \sum_{i=-P}^{P} \sum_{j=-Q}^{Q} a_i g_j s(n-i-j) \\ &\quad + \sum_{j=-Q}^{Q} g_j w(n-j) \end{aligned} \tag{24}$$

Combination of (21) and (24) yields

$$\hat{y}(n) = x(n) - \tilde{y}(n)$$

$$= s(n) - \sum_{i=-P}^{P} \sum_{j=-Q}^{Q} a_i g_j s(n-i-j)$$

$$= \sum_{i=-P-Q}^{P+Q} b_i s(n-i) \qquad (25)$$

where

$$b_i = \delta(i) - a_i * g_i, \; i = -P-Q, \cdots, P+Q \qquad (26)$$

and "$*$" denotes convolution computation.
From (21) and (25), it is easy to show that

$$r_{\hat{y}}(\tau) = \sum_{i=-P-Q}^{P+Q} b_i r_{x\hat{y}}(\tau + i) \qquad (27)$$

Using the above normal equations, we can estimate b_i's, and thus estimate a_i's from (26).

3.3. Algorithm 2

Algorithm 2 for TDE in GSNN case is now summarized as follows:
(1) estimate $\hat{c}_{k,y}(\cdot)$ and $\hat{c}_{k,yx}(\cdot)$, thus $\hat{r}_y(\cdot)$ and $\hat{r}_{yx}(\cdot)$ from $x(n)$ and $y(n)$, $n = 0, 1, \cdots, N$;
(2) estimate g_i, $i = -Q, \cdots, Q$ by solving the normal equation of (23);
(3) use (24) and (25) to obtain $\hat{y}(n)$, compute $\hat{r}_{\hat{y}}(\cdot)$ and $\hat{r}_{x\hat{y}}(\cdot)$, and then estimate b_i's using (27);
(4) estimate a_i's from (26), and thus D_j's and A_j's.

4. Simulations and Conclusions

Two simulation examples are presented to illustrate the effectiveness of the new algorithms of this paper. In the simulations, we choose $M = 2$, $D_1 = 2$, $D_2 = 5$, $A_1 = 0.8$, $A_2 = 0.6$, and $N = 2000$, and define $SNR = 10log_{10}\{E[s^2(n)/E[w_1^2(n)]\}$.
Example 1(*NSGN case*): In (1) and (2), the signal and noises are generated by

$$s(n) = u(n) + u(n-1) \qquad (28)$$

$$w_1(n) = w_2(n) = e(n) + 0.8e(n-1) \qquad (29)$$

respectively. Here, $u(n)$ and $e(n)$ are i.i.d. exponential and Gaussian processes, respectively. Figs.1-3 show the estimated a_i's, $i = -10, \cdots, 10$, obtained by traditional approach [1], Tugnait's approach [4], and our *Algorithm* 1 ($SNR = 2.5$dB, 20 Monte Carlo runs).
Example 2(*GSNNcase*): The signal $s(n)$ and and noise $w_2(n)$ are the same as in (28) and (29), while $w_1(n) = w_2(n) - 0.8w_2(n-1)$, and $u(n)$ and $e(n)$ are i.i.d. Gaussian and exponential processes, respectively. The corresponding results are plotted in Figs.4-6 ($SNR = -2.5$dB, 20 Monte Carlo runs).
From the results shown in Figs.1-6, we have the following observations.
(1) The tradictional approach always produce three peaks, with two of them due to the signal delays, while the other due to the correlatedness of the noises. However, the true time delays cannot be correctly picked out from the ambiguous peaks.
(2) For the NSGN case, although the Tugnait's approach can give the correct time delays, the estimation results are with higher variances as compared with those obtained by the *Algorithm* 1. For the GSNN case, the Tugnait's approach fail to produce the true time delays of the Gaussian signals.
(3) On the other hand, our *Algorithm* 1 produces correct time delay estimates in the NSGN case, while *Algorithm* 2 works well in the GSNN case. It is pointed out that although the SNR is relatively low in *Example* 2, the *Algorithm* 2 performs quite well since it is finally implemented in the domain of second order statistics, which is powerful in suppressing the noises.

References

[1] G. C. Carter, "Time delay estimation for passive sonar signal processing," IEEE Trans. Acoust., Speech, Signal Processing, 1981, vol. 29, No. 6, pp.463-470.

[2] C. L. Nikias and R. Pan, "Time delay estimation in unknown Gaussian spatially correlated noise processing", IEEE Trans. Acoust., Speech, Signal Processing, vol.36, No. 11, 1988, pp.1706-1714.

[3] J. K. Tugnait, "On time delay estimation with unknown spatially correlated Gaussian noise using fourth-order cumulants and cross cumulants," IEEE Trans. Signal Processing, vol.39, No. 6, 1991, pp.1258-1267.

[4] M. J. Hinich, D. Marandino and E. J. Sullivan, "Bispectrum of ship-radiated noise," J. Acoust. Soc. Amer., vol.85, pp.1512-1517, Apr. 1989.

[5] S. A. Kassam, Signal Detection in Non-Gaussian Noise, New York: Springer-Verlag, 1988.

[6] D. Yellin and E. Weinstein, "Criteria for multichannel signal separation", IEEE Trans. Signal Processing, vol.42, No.8, 1994, pp.2158-2167.

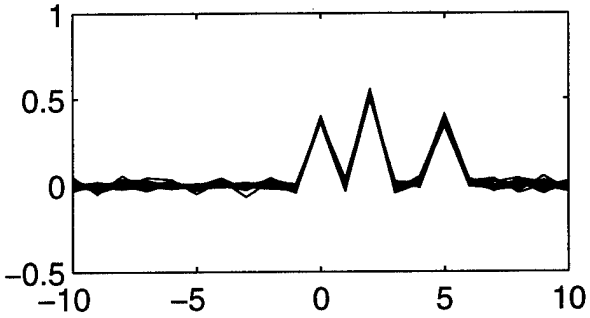


Figure 1. The estimated a_is obtained via traditional approach ($N = 2000$, $SNR = 2.5$dB, 20 Monte Carlo runs)

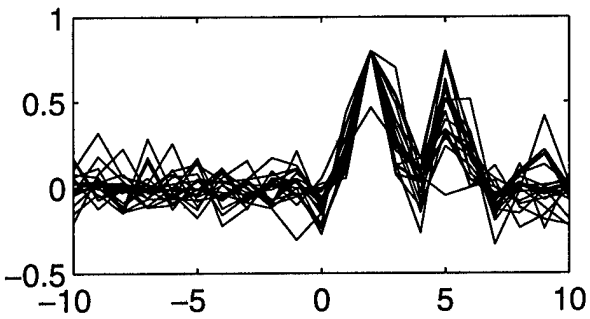


Figure 2. The estimated a_is obtained via Tugnait's approach ($N = 2000$, $SNR = 2.5$dB, 20 Monte Carlo runs)

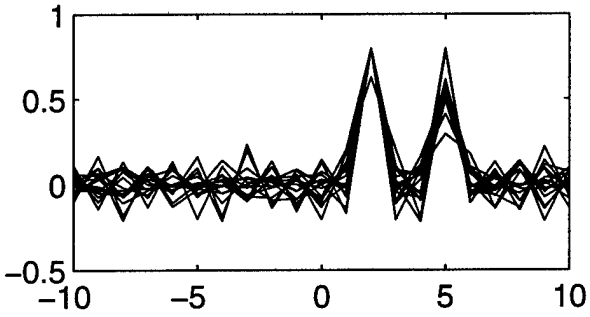


Figure 3. The estimated a_is obtained via *Algorithm* 1 ($N = 2000$, $SNR = 2.5$dB, 20 Monte Carlo runs)

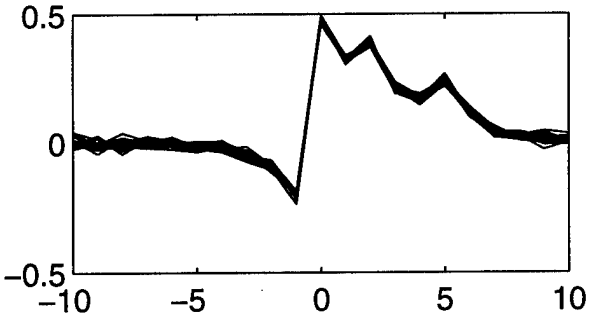


Figure 4. The estimated a_is obtained via traditional approach ($N = 2000$, $SNR = -2.5$dB, 20 Monte Carlo runs)

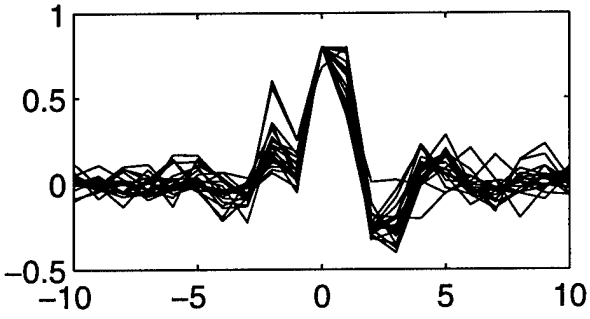


Figure 5. The estimated a_is obtained via Tugnait's approach ($N = 2000$, $SNR = -2.5$dB, 20 Monte Carlo runs)

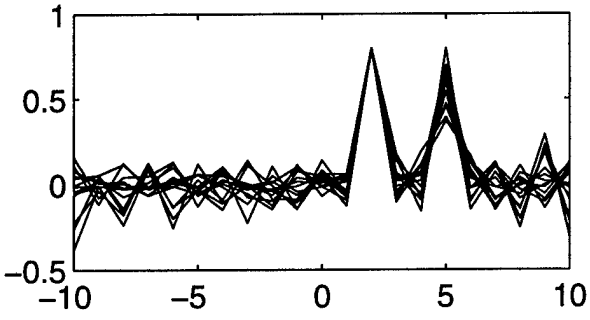


Figure 6. The estimated a_is obtained via *Algorithm* 2 ($N = 2000$, $SNR = -2.5$dB, 20 Monte Carlo runs)

Discrimination of Local Seismic Events in Panama by Means of Higher-Order Statistics

Leif Persson[1] and Jaime Toral Boutet[2]

1) National Defense Research Establishment FOA, S-172 90 Stockholm, Sweden. Tel: +46 8 706 3615, Fax: +46 8 7063543, email:leifp@sto.foa.se. Also at Uppsala University, Seismological Department, Box 2101, S-750 02 Uppsala, Sweden. Tel: +4618 18 1472, Fax: +46 18 181471, email: leif.persson@seismo.uu.se.

2) University of Panama, IGC, Panama city, Panama. Tel: +507 263 7703, Fax: +507 263 7671. email:igc2@ancon.up.ac.pa

Keywords: discrimination, higher-order cross-correlation's, master-event correlation, local seismic events.

Abstract

The analysis of seismicity need to be restricted to earthquakes. The estimated seismic energy release and the spatial distribution of seismic events will be erroneous and biased if the data catalogue includes artificial events such as explosions. Therefore, it is necessary to separate earthquakes and explosions prior to the compilation of the seismicity in a particular region. In the canal area of Panama, the events are equally distributed with signal energy contents in the same range, making localization and magnitude estimation not effective as discriminators. In this study we use a method based on master-event correlation's for filtering out the explosions in the seismological catalogue. We use a library of time series with the a priori knowledge of the source type. The method is based on comparison of unknown events with the library by second, third and fourth-order cross-correlations. Studied time series are divided into three parts, the first phases of the P and S-wave and a window from the preceding noise. Twenty known and twenty-one unknown events are tested with master-event correlation in second, third and fourth-order domain. The second-order based method discriminate correctly 40 % of the known events while the third and fourth-order based correlations succeed in 75 % respectively 80 % of the known cases.

1. Introduction

We examine the problem of discrimination between local chemical explosions and earthquakes in the canal area of Panama by second, third and fourth-order master-event cross-correlations. Regional and local seismic sources exhibit considerably large contents in both higher-order spectra and statistics, e.g. [1] and [2]. Chemical explosions generated in a pre-stressed and fracturated region as around the canal area of Panama in Central America are similar in the seismic energy contents as shallow earthquakes making conventional discrimination methods difficult. Also, localization as a method for identifying sources are not effective because the small distinction in the spatial distribution of the events. The additional information in the third and fourth-order correlation domain's are useful for the possibilities of a reduction of the misclassfication.

2. Seismic data

The number of man-made seismic events in the region of central Panama are high due to the activity of expanding the water depth and the size of the canal. Charges as large as 1800 kg of explosives are used in a daily basis. During a week several hundred events are made at the bottom of the canal and on land. In the area it is also a large number of events caused by road works, quarries and military activity both on land and in the water. Some of the explosions are reported to the seismological unit in Panama city but many events are not. In Figure 1 a typical pattern of the seismicity of a four year period is displayed. The estimation of many seismological quantities such as seismic hazard, seismicity and energy release distribution are impossible with a data base polluted by artificial events. Therefore, it is important on a routine basis to discriminate between natural and man-made events. Table I and II present some parameter's for the used events in the master-event library and Table III is some parameter's for the unknown events. The events are all in the same magnitude-energy range and in a distance range from 6 km to 74 km from the recording station UPA.

0-8186-8005-9/97 $10.00 © 1997 IEEE

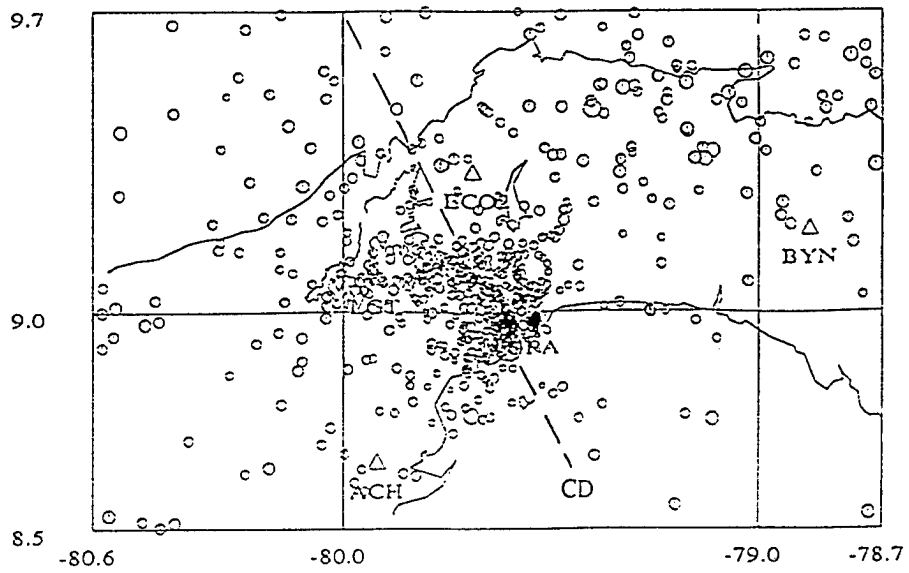


Figure 1. Seismicity map including earthquakes and man-made events in the central of Panama from the period of May 1992 to May 1996. The size of the circles are related to magnitude of the events. CD means Canal discontinuity, and UPA, ECO2, LGT, BYN and ACH are seismological stations.

No.	Year M/D	Date Origin Time	Epicenter Lat. °N	Long. °W	Mag.	Dist. fUPA km	Az. deg.
	1994						
1	May 20	23:01:07	9.10	79.76	2.8	28	298
2	Sep 28	00:33:03	8.79	79.82	2.8	39	237
	1995						
3	Mar 01	07:36:06	8.93	79.61	2.0	10	232
4	May 12	14:31:56	9.05	79.65	2.9	15	301
5	May 19	16:43:25	9.15	79.79	2.9	30	303
6	May 23	02:12:20	9.36	79.98	2.8	64	289
7	Jun. 06	21:51:07	9.40	79.01	3.3	74	52
	1996						
8	Jan 26	12:47:08	9.07	79.02	3.2	57	80
9	May 02	10:38:00	9.00	79.26	3.3	30	85
10	May 06	18:03:11	9.16	79.65	3.1	23	320
11	May 11	07:57:28	9.00	79.26	3.1	30	83
12	May 16	16:20:07	9.08	79.69	3.2	20	305

Table I. Earthquakes used in the master event library. Epicenter - Lat., Long., means estimated localization. Mag. means the duration magnitude and, distance and azimuth means position relative to the UPA station.

No.	Date Year M/D	Origin Time	Ripp no.	Epicenter Lat. °N	Long. °W	Charge kg	Mag.	Dist. from UPA km	Az. deg.
	1995								
13	Feb. 06	11:09:18	3	9.07	79.69	158.7	2.6	20	302
		:09:19				158.7			
		:09:20				158.7			
14	Feb. 07	11:25:58	2	9.16	79.71	158.8	2.7	28	303
		:26:00				158.7			
15	Feb. 10	12:09:19	3	9.07	79.73	108.8	2.7	24	298
		:09:19				108.8			
		:09:20				108.8			
16	Feb. 14	13:08:12	3	9.07	79.68	79.4	2.7	19	303
		:08:19				79.4			
		:08:22				79.4			
17	Feb. 17	11:23:59	3	9.10	79.70	174.6	3.0	20	304
		:26:47				174.6			
		:26:57				174.6			
18	Feb. 24	14:17:17	3	9.13	79.70	158.7	2.8	24	308
		:17:18				174.8			
		:17:29				174.8			
19	Jun. 16	16:00:41	7	9.08	79.72	254.0	2.8	23	300
		:00:42				254.0			
		:00:44				254.0			
		:00:45				254.0			
		:00:47				254.0			
		:00:49				254.0			
		:00:50				254.0			
20	Sep. 15	17:17:12	1	9.01	79.69	269.8	2.8	17	283

Table II. Explosions used in the master event library. Epicenter - Lat., Long., means estimated localization of the events. Ripple means the number of charges. Mag. means the duration magnitude and, distance and azimuth means position relative to the UPA station.

No.	Date Year M/D	Origin Time	Epicenter Lat. °N	Long. °W	Mag	Dist. UPA km	Az. UPA deg.
	1994						
p1	Dec 29	22:08:57	9.07	79.66	2.5	17	306
	1996						
p2	Jan 05	19:18:29	9.04	79.56	2.2	6	351
p3	Jan 06	16:36:30	9.02	79.63	2.3	12	286
p4	Jan 08	21:38:30	9.10	79.58	1.9	14	345
p5	Jan 12	19:43:46	9.00	79.60	2.2	7	284
p6	Jan 12	19:59:57	9.00	79.60	1.6	8	280
p7	Jan 19	17:19:49	9.16	79.54	2.7	19	352
p8	Jan 24	18:55:55	9.02	79.57	1.9	6	320
p9	Jan 25	11:40:58	9.35	79.96	3.1	62	306
p10	Jan 25	19:56:27	9.13	79.65	2.0	20	325
p11	Jan 26	11:03:25	9.13	79.78	2.8	33	288
p12	Jan 26	19:12:47	8.86	79.70	1.9	23	234
p13	Feb. 02	13:10:37	9.12	79.67	2.7	21	305
p14	Feb. 02	16:31:08	9.18	79.63	3.0	24	333
p15	Feb. 02	16:44:08	9.06	79.63	2.1	14	307
p16	Feb. 02	18:16:03	9.06	79.60	2.0	12	320
p17	Feb. 05	11:28:12	9.03	79.80	2.6	30	280
p18	Feb. 05	17:41:48	9.23	79.86	2.7	46	300
p19	Feb. 07	19:44:53	9.06	79.62	1.9	13	311
p20	Feb. 08	17:10:15	9.07	79.64	2.2	16	311
p21	Feb. 12	11:58:12	9.14	79.71	2.8	26	305

Table III. Unknown events used in the analysis. Epicenter - Lat., Long., means estimated localization of the events. Mag. means the duration magnitude and, distance and azimuth means position relative to the UPA station.

3. Master-event correlation analysis

We performed master-event correlation analysis by using a library of known earthquakes and explosions to selected unknown events in the second, third and fourth-order cross-correlation domain's. The measures are all possible non-redundant cross-correlation combinations for each data window between the unknown events and the library events. The data windows are matched along the time axis and the maximum correlation value with the main part of the information is used for the analysis. Some authors e.g. [3] use an interval determined by the maximum time between coherent phases. The complexity in both P and S-phases in our time series restrict us to use only the lag for maximum correlation. The conventional second-order cross-correlation between the time series of the unknown event x(t) and the time series of one of the known master events y(t) is defined as,

$$m^2_{xy}(\tau) = \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)y(t+\tau). \qquad (1)$$

The size N and the position of the data window α need to be matched to the particular phase of interest. In this study, the time series of the master and the unknown event are divided into three data windows; 1) preceding noise only, 2) the first arrived phase, normally the P-wave and 3) the second large phase, normally the S-wave. We perform the analysis at lag τ with the maximum correlation value. The analysis of the noise window serve as a reference. Similar to the different properties in the bispectrum and the trispectrum compared to power spectrum the third and fourth-order cross-correlations measure other similarities than the amplitudes, [4]. Also, other possible combinations of correlation's occur between the known master events y(t) and the tested event x(t) and is useful in the discrimination analysis. The third-order cross-correlation's are defined as,

$$m^3_{xxy}(\tau_1,\tau_2) = \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)x(t+\tau_1)y(t+\tau_2) \quad (2)$$

and

$$m^3_{xyy}(\tau_1,\tau_2) = \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)y(t+\tau_1)y(t+\tau_2). \quad (3)$$

The fourth-order cross-correlation can be defined in three different non-redundant combinations as,

$$m^4_{xxxy}(\tau_1,\tau_2,\tau_3)$$

$$= \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)x(t+\tau_1)x(t+\tau_2)y(t+\tau_3), \qquad (4)$$

$$m^4_{xxyy}(\tau_1,\tau_2,\tau_3)$$

$$= \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)x(t+\tau_1)y(t+\tau_2)y(t+\tau_3), \qquad (5)$$

and

$$m^4_{xyyy}(\tau_1,\tau_2,\tau_3)$$

$$= \frac{1}{N}\sum_{t=\alpha}^{N+\alpha} x(t)y(t+\tau_1)y(t+\tau_2)y(t+\tau_3). \qquad (6)$$

The cross-correlations in equation (1) to (6) of the unknown events x(t) are performed to all the library events y(t) and gives values in clusters. The cluster's are compared to both the earthquake and the explosion 'master' cluster composed of library events only. The discrimination space for the second-order cross-correlation values in equation (1) is spanned by both the correlation of the first and the second data window. This is not necessary for the third-order correlation values spanned by equation (2) and (3) and the fourth-order estimates, equation (4) to (6). The shortest squared Mahalanobis distance of the unknown events define if the particular event is an explosion or an earthquake. If the clusters overlap with 50 % or more the method failed and the cross-correlation values are not used in the analysis.

4. Results

The procedure begins with defining the position α and size (N) of data window for the noise, P and S-wave. Independent P and S-wave parts (vertical and horizontal components) are averaged into groups, numerically quantified in clusters, and displayed in diagrams, e.g. Figure 2. The analyses are demonstrated by the cross-correlation technique described above with a) second-order (m2xy), b) third-order (m3xyy) and c) fourth-order (m4xxyy, m4xxxy and m4xyyy). The measures used are defined in equations (1) to (6). The cross-values between the master events, earthquakes (open circles), explosions (crosses) and the unknown events (star) are measured and the shortest squared Mahalanobis distances for each group define the discrimination of the event. The first analysis is to test each of the master events to the library. The second-order based method discriminate correctly 40 % of the known events while the third and fourth-order based correlation's succeed in 75 % respectively 80 % of the cases. The results of the analysis by using 8 explosions and 12 earthquakes as known master events in a library and 21 unknown events are compiled in Table IV. For the unknown events the analysis of second, third and fourth-order measures succeed to declare 11 explosions and 8 earthquakes. Two of the events disagree between the vertical P and the horizontal S-waves.

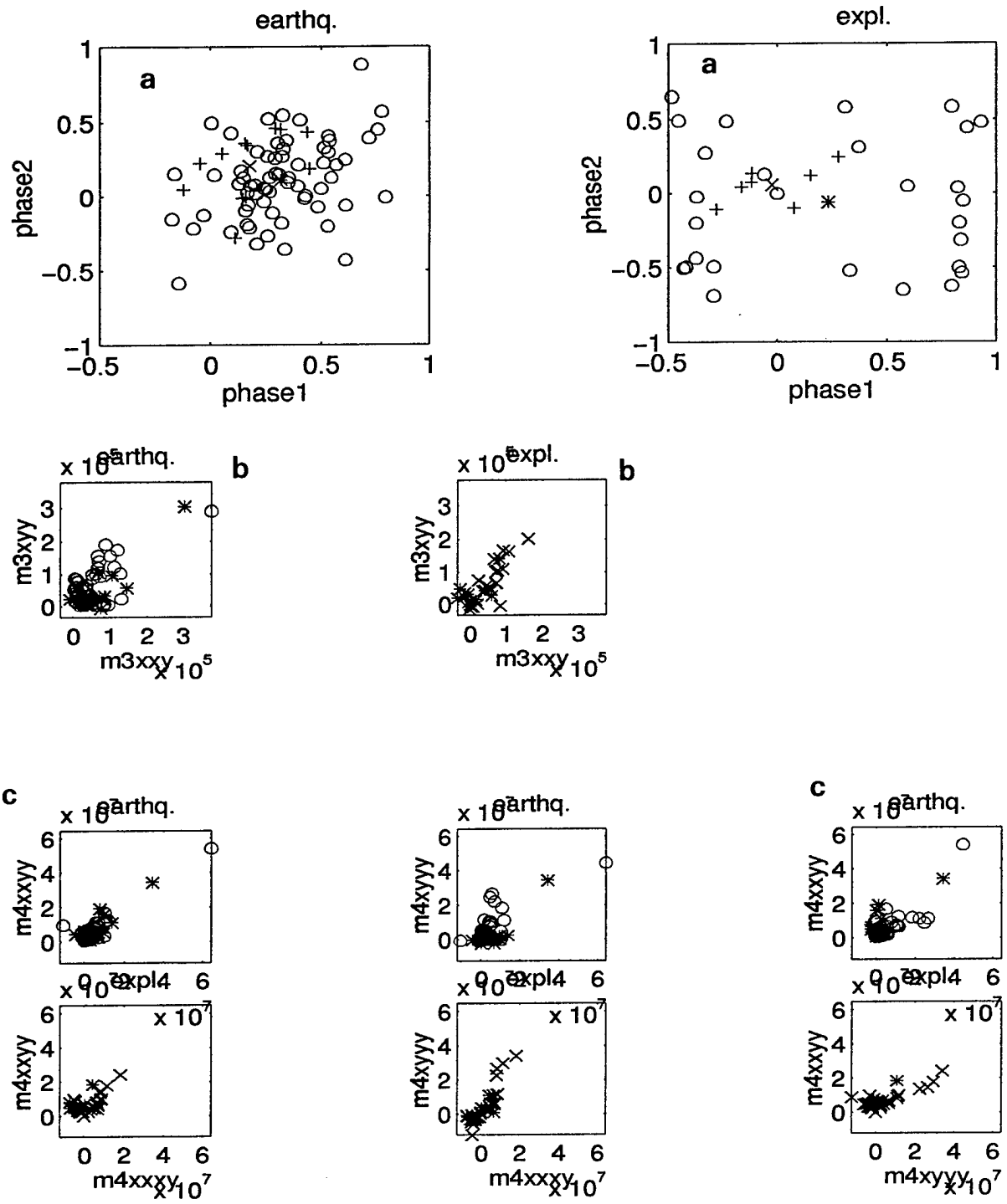


Figure 2. Examples of analysis of an unknown event (star) compared with the master events, earthquakes (open circles) and explosions (crosses) by **a**) second order, **b**) third order and **c**) fourth-order cross-correlation measures. The unknown event is p1 in Table IV and is estimated to be an earthquake.

No	P+S-m2	P m3	P-m4	S-m3	S-m4	Res. P	S
p1	ex	ex	ex	ex	eq	ex	ex
p2	eq	eq	eq	eq	ex	eq	eq
p3	ex	ex	ex	ex	ex	ex	ex
p4	eq	eq	eq	eq	ex	eq	eq
p5	eq	eq	eq	eq	eq	eq	eq
p6	ex	ex	ex	ex	ex	ex	ex
p7	eq	ex	eq	ex	ex	eq	ex
p8	ex	ex	ex	ex	ex	ex	ex
p9	ex	ex	ex	ex	ex	ex	ex
p10	ex	ex	ex	ex	ex	ex	ex
p11	eq	ex	ex	ex	ex	ex	ex
p12	ex	ex	ex	ex	ex	ex	ex
p13	eq	eq	eq	eq	eq	eq	eq
p14	eq	eq	eq	ex	eq	eq	eq
p15	eq	eq	eq	eq	eq	eq	eq
p16	ex	ex	ex	ex	ex	ex	ex
p17	eq	eq	eq	ex	eq	eq	eq
p18	ex	ex	ex	ex	ex	ex	ex
p19	eq	ex	eq	ex	ex	eq	ex
p20	ex	ex	ex	ex	ex	ex	ex
p21	eq	eq	eq	ex	eq	eq	eq

Table IV. Analysis of unknown events used in the analysis. Abbreviation ex. means explosion and eq. earthquake.

5. Discussion and Conclusions

At present, access to digitized waveforms permits fast evaluation of signal source parameters applicable in the analysis of weak and low energy seismicity (like natural fault-ruptures or man-made events). In the classification of seismic sources, we use the quantitative appearance of seismograms in the higher-order cross-correlation domain, instead of conventional statistical processing of catalogues, contrast in magnitude units, or polarity signs and envelope and energy contents in the time series.

We describe a method to register and discriminate industrial explosions in the canal region of Panama against low-magnitude local earthquakes. The higher-order based master-event cross-correlation method is introduced and its potential for new applications in observational seismology is examined. The study comprises of 21 unknown events compared to a library of 8 explosions and 12 earthquakes, inside the canal area of Panama. Discriminations by second-, third- and fourth-order succeed in 11 explosions and 8 earthquakes. Four of the events disagree between the vertical P and the horizontal S-waves. One explanation for this might be wrong phase identification, multiple events or the library is to small to cover all the waveform characteristics. Another possible source for errors is the position and size of the data window and the matching of the compared phases in each window. Results suggest that there are no influence in the size of the events (magnitudes), and the method seems to be insensible to azimuth or distance fluctuation without any corrections of the incoming waves. From a pure seismological point of view, the method has a robust physical meaning using real waveform records.

The seismic catalogue at UPA contains a high number of presumed explosions. The seismicity show that after semi-automatic identification and removing of the hidden explosions, the hypothesis that assumes an active single lineament of the Canal Discontinuity fault becomes rather weak. This method can assist seismologists in the region to improve their identification of seismological data. Therefore, the higher-order based master-event correlation technique is suggested to be used before any investigation or conclusion about local seismicity, and is recommended for future routine analysis at other seismological observatories.

6. References

[1] Persson L., 1993, 'Bispectrum analysis of the coda records from local earthquakes and mine explosions', *Computers & Geosciences*, vol. 19, no 2, pp. 243-247.

[2] Persson L., 1994, 'Detection of seismic phases by higher-order statistics', In *Proc. of Norsig/IEEE*, Ålesund, Norway.

[3] M. Joswig and Schulte-Theis H., 1993, 'Master-event correlation of weak local earthquakes by dynamic waveform matching', Geophys. J. Int., vol 113, pp. 562-574.

[4] Lii K. S. and Helland K. N., 1981, 'Cross-bispectrum, computation and variance estimation', ACM Trans. On Math. Soft. Vol. 7, no 3, pp. 284-294.

PERFORMANCE EVALUATION OF TIME-DELAY ESTIMATION IN NON-GAUSSIAN CONDITIONS

Gadi Shor and Hagit Messer

Department of Electrical Engineering - Systems, Tel Aviv University
Tel-Aviv 69978, Israel
Tel. +972-36408119, Fax. +972-36407095
E-mail: messer@eng.tau.ac.il

ABSTRACT

In recent years, time delay estimation TDE processors for non-Gaussian conditions were developed and it has been shown that exploiting the non-Gaussianity of the processes can lead to improved performance, relative to the Gaussian case. However, while for the later standard tools (as the Cramer-Rao bound) can easily be applied to predict the achievable performance, for the general, non-Gaussian case using such tools rarely leads to analytical solutions. In this paper we present a close-form expression for a figure of merit FOM for any TDE processor under any statistical model. It is then evaluated for different TDE processors and it is confirmed (by simulations) that the proposed FOM accurately predicts the performance of these processors for reasonable values of time-bandwidth product.

1. INTRODUCTION

Time delay estimation (TDE) is a problem of applications in many areas as: radar and sonar, communication, bio-medical engineering, etc. This problem has been deeply studied for the case where the additive noise process is Gaussian [1]. For this case, the optimal processor and the achievable performance are well understood. There are applications, however, where the noise and/or the signal are non-Gaussian. In recent years, TDE processors for non-Gaussian conditions were developed. In general, exploiting the non-Gaussianity of the processes can lead to improved performance, relative to the Gaussian case. However, while for the Gaussian case standard tools (as the Cramer-Rao bound) can easily be applied to predict the achievable performance, for the general, non-Gaussian case using such tools rarely leads to analytical solutions.

In this paper we present a close-form expression for a figure of merit (FOM) for any TDE processor under any statistical model. It has been developed under a simplified version of the TDE problem, namely: a discrete, *iid* model. For this case, simulation results show very good agreement with the theoretical results. However, as we will show in the sequel, TDE processors for the *iid* case can be used with non-*iid* signals with performance degradation which is proportional to the amount of correlation in the processes.

Consider two discrete signals:

$$\begin{aligned} x_1(i) &= s(i) + n_1(i) \quad ; \quad 1 \le i \le N \\ x_2(i) &= s(i-D) + n_2(i) \quad ; \quad 1 \le i \le N \end{aligned} \tag{1}$$

D is an unknown parameter which can get one out of $K+1$ possible values, $0 \le D \le K \le N$. We assume that $n_1(i)$, $n_2(i)$ and $s(i)$ are mutually independent, *i.i.d* processes. We denote the probability density function of $s(i)$ by f_s and that of $n_j(i)$, $j = 1, 2$, by f_n for $i = 1, ..., N$. Our aim is to estimate D given the two vectors, $\underline{x}_1 = [x_1(1), ..., x_1(N)]^T$ and $\underline{x}_2 = [x_2(1), ..., x_2(N)]^T$.

Any TDE processor looks for the best matching between $x_2(i)$ and the delayed signal, $x_1(i-d)$ for $0 \le d \le K$. Because of the *i.i.d* assumption, such a processor is of the general form: $max\{z_d\}_{d=0}^{K}$, where

$$z_d = \sum_{j=1}^{N} g(x_1(j-d), x_2(j)) \tag{2}$$

$g(\cdot,\cdot)$ is a weighting function which operates on pairs of terms from the two sequences. Different processors (e.g., linear or non-linear correlators, minimum norm processors, maximum likelihood, etc.) differ by the function g. However, based on the assumption that the statistics of the two noise processes are the same, it is reasonable to assume that $g(y,z) = g(z,y)$.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. THE PROPOSED *FOM*

Let $z_0 = z(d = D)$, $z_r = z(d = d_r \neq D)$ and $v_r = z_0 - z_r$. Then the probability of correct decision is:

$$P_c = Pr\{v_1 > 0, v_2 > 0, ..., v_K > 0\} = Pr\{\underline{v} > \underline{0}\} \quad (3)$$

z_i is the sum of N independent random variables with finite and equal second moments. Thus $\underline{v}$ is asymptotically normal for large N. Under the stated assumptions all v_i have the same means and variances: $E(v_r) = E(z_0) - E(z_r) \equiv \eta_0 - \eta_1$, $Var(v_r) = Var(z_0 - z_r) \equiv \sigma^2$. v_r and v_s also have the same correlation coefficient $\rho > 0$ for all $r \neq s$. It follows that

$$E(\underline{v}) = (\eta_0 - \eta_1)\underline{1} \quad ; \quad cov(\underline{v}) = \sigma^2[\rho \underline{1}\underline{1}^T + (1-\rho)I] \quad (4)$$

where $\underline{1}$ is a K-dimensional vector whose elements are 1 and I is the $K \times K$ identity matrix. Let a and b_r be independent Gaussian random variables

$$a = N(\eta_0 - \eta_1, \rho\sigma^2) \quad , \quad b_r = N(0, (1-\rho)\sigma^2) \quad (5)$$

then the their sum $a + b_r$ has the statistics of v_r specified by (4) so one can think of v_r as the sum of two components:$v_r = a + b_r$. Because of the independence of b_r, (3) can now be written in the form:

$$P_c = \int_{-\infty}^{\infty} [P_r(b_1 > -a|a)]^K f_a(a) da. \quad (6)$$

Using (5) one then obtains

$$P_c = \int_{-\infty}^{\infty} \left[\frac{1 + erf(\frac{u}{\sqrt{2}})}{2}\right]^K \frac{1}{\sqrt{2\pi\frac{\rho}{1-\rho}}} e^{-\frac{[\frac{0}{\sqrt{1}} - \frac{1}{\sqrt{1}}]^2}{2\frac{1}{1}}} du \quad (7)$$

where $u = a/\sqrt{(1-\rho)\sigma^2}$ and $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$. Here we suggest the approximation: $[w(u)]^K \approx w(\frac{u - c_1(K)}{c_2(K)})$ where $w(u) = \frac{1 + erf(\frac{u}{\sqrt{2}})}{2}$ and $c_1(K)$ and $c_2(K)$ are predetermined numbers. Under this approximation:

$$P_c \approx \frac{1 + erf(\frac{\frac{0}{1} - c_1(K)\sqrt{1-\rho}}{\sqrt{\rho + c_2^2(K)(1-\rho)}})}{2}. \quad (8)$$

The probability of correct decision, or any quantity which is proportional to it, is a reasonable figure of merit. In particular, under our assumptions, $Pr\{\hat{D} = d|D\} = P_c$ if $d = D$ and $Pr\{\hat{D} = d|D\} = \frac{1-P_c}{K}$ if $d \neq D$, so P_c uniquely determines the mean square error (mse) of any reasonable estimator of D:

$$\begin{aligned} mse\{\hat{D}\} &= \sum_{i=0}^{K} (d_i - D)^2 Pr\{\hat{D} = d_i|D\} \\ &= \sum_{i=1}^{K} (d_i - D)^2 \frac{1 - P_c}{K} = (1 - P_c)B(K) \end{aligned} \quad (9)$$

where $B(K)$ is a measure of the distance between the true delay and the alternatives and is independent of the estimator.

Since the error-function is a monotonic increasing function of its argument, a proper figure of merit for any processor can be this argument, given by (see (8)):

$$FOM = \frac{\frac{\eta_0 - \eta_1}{\sigma} - c_1(K)\sqrt{1-\rho}}{\sqrt{\rho + c_2^2(K)(1-\rho)}} \quad (10)$$

Note that the FOM has units of (the square root of) the signal-to-noise ratio ($\sqrt{SNR}$), as the common FOM used in binary detection problems [2].

It can be shown, that any given TDE processor of the form of (2) and under the stated statistical assumptions, the FOM (10) becomes:

$$FOM = \frac{\sqrt{N\frac{R_0}{R_1}} - c_1(K)}{\sqrt{\frac{R_2}{R_1} + c_2^2(K)}} \quad (11)$$

where R_0, R_1 and R_2 are evaluated from the 2nd order moments of the variables:
$g_{ij} = g(s(i) + n_1(i), s(j) + n_2(j))$, $i, j = 1, 2, 3$:

$$\begin{aligned} R_0 &= [E(g_{11}) - E(g_{12})]^2 \\ R_1 &= E(g_{12}^2) - 2E(g_{12}g_{13}) + E^2(g_{12}) \\ R_2 &= 4E(g_{12}g_{13}) - 4E^2(g_{12}) + E(g_{11}^2) - E^2(g_{11}) \\ &- 4E(g_{11}g_{12}) + 4E(g_{11})E(g_{12}) \end{aligned} \quad (12)$$

Note that the dependence of the FOM on N and on K is explicit in (11). The specific TDE processor and the statistic of the signal and the noise processes affect the FOM via R_0, R_1 and R_2 which are independent of N and K.

The (approximated) probability of a correct decision for this problem can also be derived using (8), which is related to the mse by (9).

2.1. Example

We study time delay estimation of a Gaussian signal in non-Gaussian noise. The signal $s(i)$ is of zero mean and variance σ_s^2 and the noise processes are mixed-Gaussian in the sense that their samples are picked from two zero-mean Gaussian populations of variances σ_1 and σ_2 with equal probability. That is:

$$f_n(x) = \frac{1}{2\sqrt{2\pi\sigma_1^2}} \exp^{-\frac{2}{2\frac{2}{1}}} + \frac{1}{2\sqrt{2\pi\sigma_2^2}} \exp^{-\frac{2}{2\frac{2}{2}}} \quad (13)$$

We consider three discrete TDE processor: the well-known linear correlator [3], the maximum-likelihood ML estimator for this problem [6], and the $minorm_p$

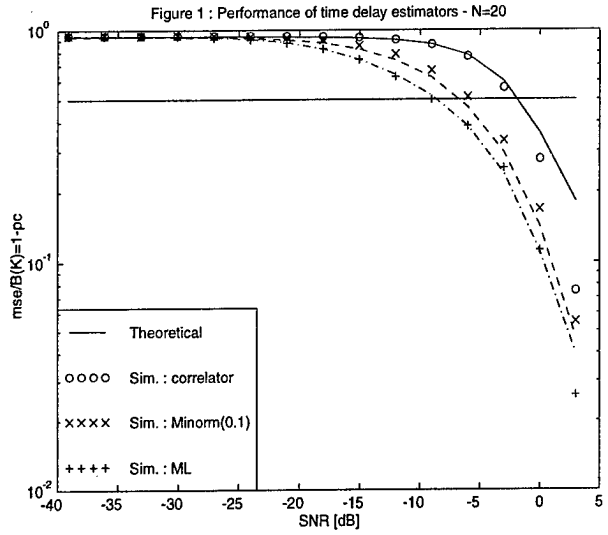


Figure 1: **The mean square error of the estimate of the unknown delay $D = 0$ for a Gaussian signal in a zero-mean, symmetric mixed-Gaussian noise ($\delta = 10$) as a function of the signal-to-noise ratio. Simulation of 3 processors are compared with the theoretical results. $N = 20$, 5000 Monte-Carlo runs.**

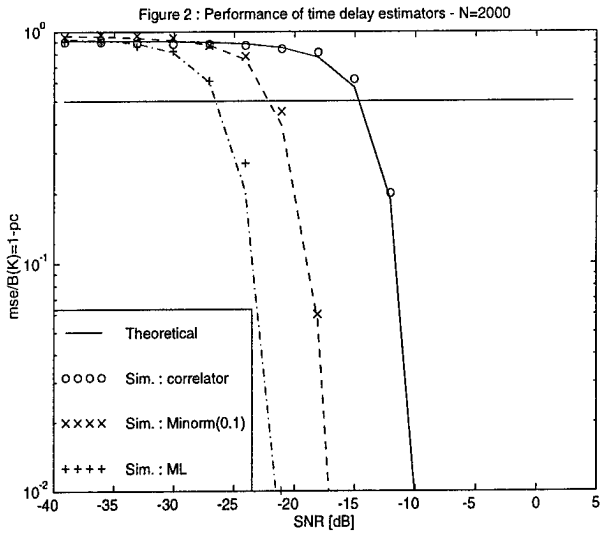


Figure 2: **The mean square error of the estimate of the unknown delay $D = 0$ for a Gaussian signal in a zero-mean, symmetric mixed-Gaussian noise ($\delta = 10$) as a function of the signal-to-noise ratio. Simulation of 3 processors are compared with the theoretical results. $N = 2000$, 100 Monte-Carlo runs.**

processor [4,5], which minimizes the norm-p ($p \neq 2$) of the difference between $x_2(i)$ and $x_1(i-d)$[1], $d = 0, 1, ..., K$. Figs. 1 and 2 present the theoretical performance of the three processors in comparison with simulation results. The theoretical performance is evaluate using (9), where η_0, η_1, σ and ρ were derived from the first and the second order moments of the statistics of a given processor, which are evaluated directly from the simulated data by averaging over the correspondence quantities. The performance of the simulations is evaluated by averaging over the TDE results of a given processor using 5000 Monte-Carlo runs for the case of $N = 20$ (Fig. 1) and 100 runs for $N = 2000$ (Fig. 2). The performance is depicted as a function of the signal to noise ratio (SNR): $\frac{2\sigma^2}{\sigma_1^2+\sigma_2^2}$ for noise with $\delta = \frac{\sigma_1}{\sigma_2} = 10$. We consider two cases: $N = 20$ and $N = 2000$. In both cases, $K = 16$. The two cases represent the extreme cases where the threshold is places in high/low SNR.

The simulation results show very good agreement with the theoretical performance. Note, however, that since they theoretical FOM is based on the central limit theorem, it describes better the performance for $N = 2000$ than for $N = 20$.

[1]Note that the linear correlator is also a $minorm_2$ processor.

3. DISCUSSION

Although the results of Section 2.1 are promising, the basic assumptions seem restrictive. First, the *iid* assumption on both the signal and the noise processes can rarely be justified when practical applications are considered. Second, we assume that the unknown delay is confined to multiples of the sampling integral. The two assumptions are related. To study the effect of correlation in the noise and/or the signal processes on the results, we have run simulations of estimators designed under the *iid* assumption when the actual distributions deviate from the *iid* pattern. Mixed Gaussian correlated samples were generated by the following procedure: Each of 3 independent sets of *iid* Gaussian samples was passed through a low-pass filter of the form.

$$H(z) = \frac{z}{z-p} \tag{14}$$

The constant p controls the bandwidth of the filter. If the resulting sequences are designated as h_1, h_2 and h_3, the desired mixed Gaussian correlated sequence u_i is given by

$$u_i = h_{1i}(y_i - 0.5) + h_{2i}(y_i + 0.5). \tag{15}$$

Here y_i is generated by hard-limiting h_{3i} at levels

of +0.5 and -0.5. Thus

$$y_i = \begin{cases} 0.5 & h_{3i} > 0 \\ -0.5 & h_{3i} < 0 \end{cases} \quad (16)$$

To make the effect of correlation explicit, Figs. 3-4 show curves for the linear correlator and for the maximum likelihood estimator working with: a) *iid* signal and noise b) correlated signal and noise, $p = 0.8$, c) correlated signal and noise, $p = 0.9$. Simulations were

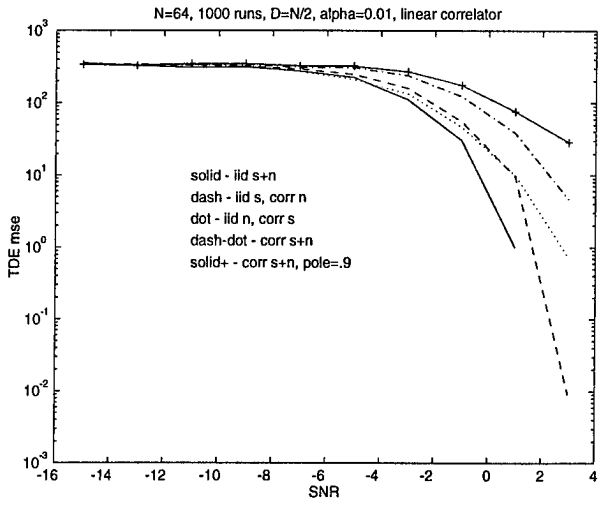


Figure 3: **Simulation results of the linear correlator: *iid* signal and noise (solid line), *iid* signal and correlated noise (dash line), *iid* noise and correlated signal (dot), correlated signal and noise ($p = 0.8$) (dash-dot) and more correlated signal and noise ($p = 0.9$) (solid +). The mean square error of the time delay estimates is depicted as a function of the signal-to-noise ratio. Number of data points: $N = 64$, number of Monte-Carlo runs: 1000, true delay: $D = N/2$, Gaussian signal and Mixed-Gaussian noise with $\delta = 10$.**

also carried out for *iid* signal/correlated noise and correlated signal/*iid* noise. The results of the various experiments show that the performance with non-*iid* signal and/or noise processes shows degradation relative to the *iid* case, which is proportional to the amount of correlation in the noise. Since the proposed FOM describes the performance for the *iid* case, it can serve as a lower bound on the performance for any, non-*iid*, case. Note, that if the actual delay is not an integral multiply of the sampling time, the estimation error under a correct decision is not zero, but a constant proportional to the fraction of the residual delay, so (3) is no longer correct. However, for comparing the performance of two estimators it is not important whether

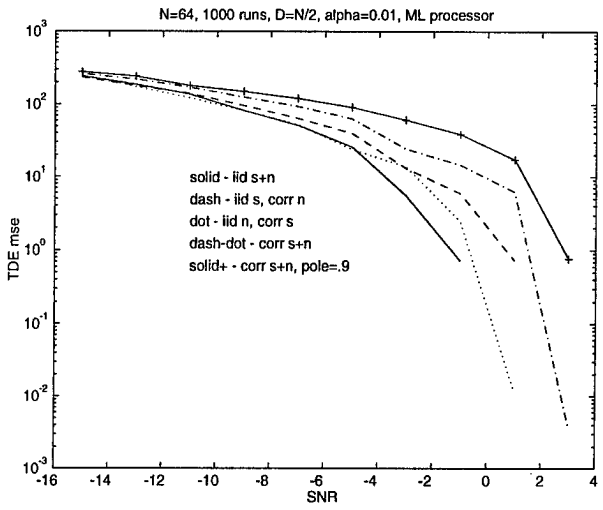


Figure 4: **As Fig. 3, for the ML processor.**

the true delay is a integral number of the sampling period or not.

4. REFERENCES

1. G.C. Carter, Editor: ***Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test, and Evaluation Engineers***, IEEE Press, pp. 134-137, 1993.

2. C.W. Helstrom: ***Elements of Signal Detection and Estimation***, Prentice Hall, 1995.

3. C.H. Knapp, G.C. Carter: "The Generalized Correlation Method for Estimation of Time Delay", IEEE Trans. ASSP-24, No. 4, pp. 320-327, August 1976.

4. H. Messer, Y. Anu: "A Novel Algorithm for Time Delay Estimation with Improved Performance for Non-Gaussian Signals", ATHOS Workshop on System Identification and High Order Statistics, Campus of Sophia-Antipolis, France, Sept. 20-21, 1993.

5. H. Messer, P. M. Schultheiss: "On Time-Delay Estimation in Non-Gaussian Noise", IEEE Proc. of the 7th SSAP Workshop, Quibec, Canada, June 1994.

6. P.M. Schultheiss, H. Messer, G. Shor: "Maximum Likelihood Time Delay Estimation in non-Gaussian Noise", submitted to the IEEE Trans. on Signal Processing.

Time-varying Third-order Cumulant Spectra and Its Application to the Analysis and Diagnosis of Phonocardiogram

Minfen Shen
Scientific Research Dept.
Shantou University
Guangdong, 515063
China
mfshen@mailserv.stu.edu.cn

Fenglin Shen
Dept. of Electronic Engineering
University of Science & Technology of China
Hefei, Anhui, 2300
China
flshen@ustc.edu.cn

Abstract

Time-varying third-order comulant spectra for analyzing phonocardiographic signals has been proposed as an effective tool to detect and quantify the temporal quadratic nonlinear interactions. The cumulant-based Wigner bispectra (CWB) is applied to investigate the nonstationarity and non-Gaussianility of both actual normal and clinical phonocardiogram. Significant time-varying bispectral structure are found and discussed. It is expected to use the Wigner bispectra in the understanding of the heart sound mechanism and the improvement of the assistant diagnosis of some heart diseases.

1. Introduction

The diagnostic use of heart sound has a long history in cardiology. However, it still dose not enable the analyst to obtain both qualitative and quantitative characteristics of the phonocardiographic signals. People still have an insufficient understanding of heartbeat sound mechanisms and the inherent complexity of heart sounds. Abnormal phonocardiogram may contain, in addition to the first and second sounds, murmurs and aberrations caused by different pathological conditions of the cardiovascular system [1]. A physician's capability to diagnose the heart sounds is quite limited because the human ear is poorly suited for the cardiac auscultation.

With the rapid development of electronic instrumentation and the understanding of the genesis of heart sounds, many characteristics and other features of PCG signals can be analyzed more accurately by emplying digital signal processing techniques [2] [3]. It has been shown that the measurement and processing of heart sound signals are clinically significant. In the past 20 years, a variety of methods are used to analyzed the PCG signals, such as the sound spectrograph, wavelet detection, power spectral estimation, linear model-based techniques and so on [4] [5] [6]. However, many approaches are based on the analysis of power spectrum, or second-order statistics, of PCG signals. To extract the useful features and more information involved in heart sounds, a newly emerging technique, higher-order statistical analysis, is proposed in this paper to provide new insights into the nature of the phonocardiogram. Bispectral analysis has been developed in response to a need to examine square nonlinear interaction relationships in the heart sound signals. On the other hand, since heart sound exhibit marked changes with time and frequencies, time-varying techniques are necessary to be considered for analyzing the PCG. The earlier time-frequency analysis for phonocardiographic signals was based on the sound spectrograph. Recently, several kinds of windowed Wigner-Ville distributions and wavelet transformations

0-8186-8005-9/97 $10.00 © 1997 IEEE

were applied to detect the temporal power spectra with high resolution [7] [8]. However, second order time-frequency representations are not sufficient in the study of non-Gaussianility of the nonstationary PCG signals. To investigate the nonlinear nature of different non-stationary and non-Gaussian PCG signals, as was done for the stationary case, it is important to extend the quadratic time-frequency distribution to the higher order time-frequency representation [9] [10]. For this purpose, the time-varying higher-dimensional spectra in this paper is proposed for the analysis and diagnosis of phonocardiogram. Several kinds of real PCG data are analyzed based on third-order cumulant-based Wigner higher-order spectra. Significant differences of the transition patterns are analyzed between the normal and pathological phonocardiographic signals in the time-varying bifrequency domain. The aim is to get a new insight into the detection of square nonlinear quadratic interactions of the nonstationary phonocardiographic signals in terms of time varying third-order cumulant spectra.

2. Wigner higher-order spectra

Let $x(t)$ be a complex deterministic signal. The kth-dimensional local autocorrelation of $x(t)$ is defined as [11] [12]

$$R_k(t,\tau_1,\tau_2,\cdots\tau_k) = \\ = x^*(t-t_0)\prod_{i=1}^{k} x(t+\tau_i-t_0) \tag{1}$$

where t_0 is an arbitrary time delay. It is given by the following expresion [13]

$$t_0 = \sum_{i=1}^{k} \tau_i/(k+1) \tag{2}$$

The kth order Wigner higher-order moment spectra (WHOS) is defined as the k-dimensional Fourier transformation of the local autocorrelation

$R_k(t,\tau_1,\tau_2,\cdots\tau_k)$:

$$W_{kx}(t,f_1,f_2,\cdots,f_k) = \\ = FT_k\{R_k(t,\tau_1,\tau_2,\cdots,\tau_k)\} \tag{3}$$

The special case of Wigner higher-order moment spectra for $k=2$ is called as Wigner third-order moment spectra or bispectrum (WB) which is widely used in many practical applications. Therefore, we focus on the analysis of the third-order cumulant spectra. The Wigner bispectra of the signal $x(t)$ is expressed as

$$WB_x(t,f_1,f_2) = \int_{\tau_1}\int_{\tau_2} x^*(t-\tau_1/3-\tau_2/3)\cdot \\ \cdot x(t+2\tau_1/3-\tau_2/3)x(t+2\tau_2/3-\tau_1/3)\cdot \\ \cdot\exp[-j2\pi(f_1\tau_1+f_2\tau_2)d\tau_1 d\tau_2 \tag{4}$$

The definition of the Wigner higher-order moment spectra is only suitable for deterministic signals. For many kinds of biomedical signals and other kinds of engineering problems, they should be considered as non-stationary random process. It is assumed that $\{x(t)\}$ is a non-stationary complex random process with zero-mean value. The time-varying third-order cumulant is denoted by

$$C_{3x}(x) = cum(t,\tau_1,\tau_2) \tag{5}$$

Then the cumulant-based Wigner bispectrum (CWB) is defined as

$$CWB_x(t,\tau_1,\tau_2) = \iint cum(t,\tau_1,\tau_2)\cdot \\ \cdot\exp[-j2\pi(f_1\tau_1+f_2\tau_2)]d\tau_1 d\tau_2 \tag{6}$$

Two main properties of CWB are summaried as following

(1) If complex random processes $\{x(t)\}$ and $\{y(t)\}$ are statistically independent and $z(t) = x(t)+y(t)$, then we have

$$CWB_z(t,\tau_1,\tau_2) =$$

$$= CWB_x(t,\tau_1,\tau_2) + CWB_y(t,\tau_1,\tau_2) \quad (7)$$

(2) If complex random process $\{x(t)\}$ is jointly Gaussian with zero mean and independent and equally correlated real and imaginary parts, then

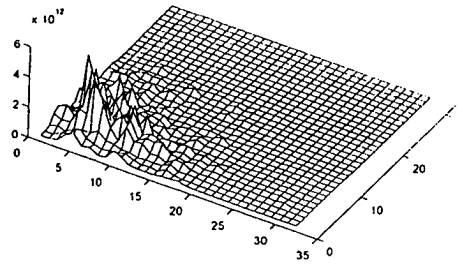
$$CWB_x(t,\tau_1,\tau_2) = 0 \quad (8)$$

Therefore, CWB can preserve many important properties of third-order statistics of a stationary signal, for example, elimination of additive Gaussian noise with unknown covariance, and at the same time are able to characterise the time variations of frequency contents of the signals.

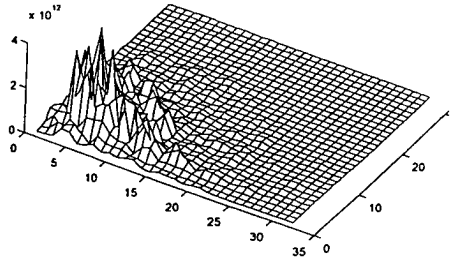
3. Experimental results and discussion

A personal computer-based system for heart sound measurement and data acquisition is implemented. ECG lead is considered to collect ECG as a timing signal. Normal and clinical PCG signal were digitally recorded in Beijing Fuwai Hospital. In order to test the nonlinear changes in high frequency, the sampling rate used is 1250 and 2500 samples/s. The length of each record of PCG signal is 20 second. To investigate the time-varying quadratic phase coupling of the PCG signals, the cumulant-based Wigner bispectral (CWB) technique is employed to analyze the time-bifrequency distribution of two cases of phonocardiographic signals, normal and abnormal with murmurs. The cumulant-based Wigner bispectra of the heart sound signal is estimated by using FFT technique with 2048 points of heart sound in discrete time and bifrequency form.

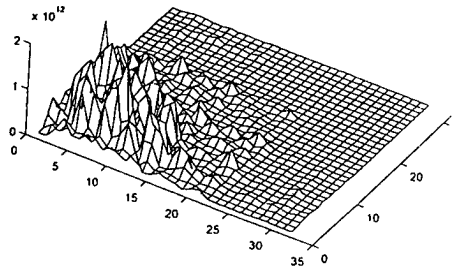
The magnitude of time-varying bispectrum for both actual normal and pathological PCG signals at (a) n_1=1024, (b) n_2=2048 and (c) n_3=3072 are shown in Fig.1 and Fig.2, respectively. It can be seen that the quadratic phase coupling is significantly changing with time in time-bifrequency domain. The time-varying bispectral structure of two kinds of phonocardiographic signals are clearly different. The frequencies of the PCG signals that occur quare phase



(a)

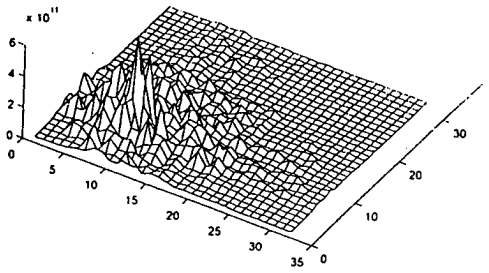


(b)

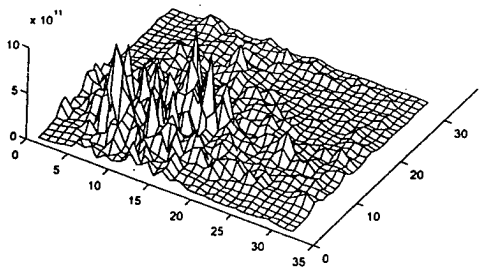


(c)

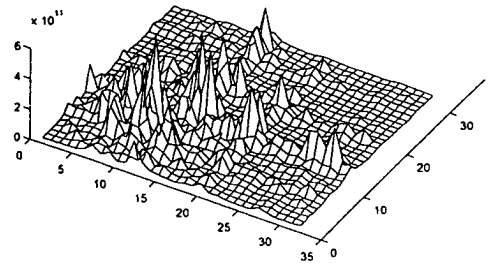
Fig. 1 The time-varying third-order cumulant spectra of the normal phonocardiogram



(a)



(b)



(c)

Fig. 2 The time-varying third-order cumulant spectra of the pathological phonocardiogram with murmurs

coupling in bifrequency domain are time-varying. Some bispectral peaks are observed for the normal heart sound signal in the relatively low bifrequency band with the frequencies from 100 Hz to 300 Hz. In the pathological PCG signal with murmurs case, several bispectral peak also exist, but they slowly change with time and occur in higher bifrequency domain with the frequencies changing from 120 Hz to 500 Hz. This result is expected since the frequency contents of murmurs are higher than the first and second heart sound. It is indicated that the pathological cardiac vibration characteristics has changed. In practical application, heart sounds collected are unavoided to be corrupted by additive Gaussian noise. The affect of backgroud Gaussian noise to the phonocardiographic signals was analyzed with several SNR, defined as $10\log\sigma_s^2/\sigma_n^2$. We observe that the CWB technique has high immunity to additive Gaussian noise with arbitrary covariance, while the power spectral analysis or time-frequency representation are limited for elimination of background Gaussian noise.

4. Conclusion remarks

We present a new method, higher order time-frequency distribution, for analyzing the nonstationity and the non-Gaussianility of phonocardiographic signals. In this paper, it can be seen that the cumulant-based Wigner higher-order spectra is applicable for the analysis of practical PCG signals. Both actual normal and pathological heart sounds are digitally collected and processed by employing the advanced signal processing technique: time-varying third-order cumulant spectra. The cumulant-based Wigner bispectra (CWB) is found to be a successful tool for the analysis of heart sound due to the fact that the PCG signals are characterized by changes in bifrequency domain as time progresses. We found that the significant different patterns of time-varying bispectral structure exist between the PCG signals of normal persons and the patients in terms of CWB method. The CWB reveals more details about the temporal square nonlinear information of the heart sounds. New nonstationary and non-Gaussian signal processing techniques provide a powerful tool for PCG signals, particularly in regard to the nonlinear relationship of the clinical heart sounds. This may

be applied in the improvement of the diagnostic techniques of some heart diseases in cost-effective approaches and extended to the analysis of other biomedical signals.

5. References

[1] P. H. Stein. Frequency spectra of the first heart sound and of the aortic component of the second heart sound in patients with degenerated porcine bioprodthetic valves. *Am. J. Cardiol.*, 53 (557), 1984.

[2] A. Iwata. et al. Automatic classification of the phonocardiogram. *Automediaca,* Vol. 3, 1980.

[3] J. H. Tae. et al. Pole-zero modeling and classification of phonocardiograms. *IEEE Trans. on Biomedical Engineering, BME -30 (2),* February 1983.

[4] M. R. Rangaraj. et al. Phonocardiogram signal analysis: a review. *Critical Reviews in Biomedical Engineering,* Vol. 15. Issue 3, 1988.

[5] M. S. Obaidat. Phonocardiogram signal analysis: techniques and perfomance comparison. *Journal of Medical Engineering & Technology,* Vol. 17(6), November 1993.

[6] A. P. Yoganathan. Use of fast Fourier transformation for frequency analysis of the first heart sound in normal man. *Medical Biology Engineering,* Vol. 14(69), 1976.

[7] J. C. Wood. and D. T. Barry. time-frequency analysis of the first heart sound. *IEEE Engineering in Medicine and Biology,* March/April 1995.

[8] H. I. Choi and W. J. Williams. Improved time-frequency representation of multicomponent signals using exponential kernals, *IEEE Trans. on Acoust Speech and Signal Processing.,* Vol. 37: 862-871, 1989.

[9] B. Boashash and G. Frazer. Time-varying higher order spectra, generalized Wigner-Ville distribution and the analysis of underwater acoustic data, *Proceedings of 92' ICASSP,* Vol. V, 1992.

[10] J. R. Fonollosa and C. L. Niakias. Analysis of finite-energe signal using higher order moment and spectra-based time-frequency distributions", *Signal Processing,* Vol..36, 1994.

[11] J. R. Fonollosa and et al. Wigner higher order moment spectra: definition, properties, computation and application to transient signal analysis, *IEEE Trans. on Signal Processing* , Vol. 41, 1993.

[12] P. O. Amblard and J. L. Lacoume. A deductive construction of third order time-frequency distributions, *Signal Processing,* Vol. 36 (3), 1994.

The Analysis and Classification of Phonocardiogram Based on Higher-order Spectra

Minfen Shen
Scientific Research Dept.
Shantou University
Shantou, Guangdong 515063, China
mfshen@mailserv.stu.edu.cn

Lisha Sun
Dept. of Electronic Engr.
Shantou University
Shantou, Guangdong 515063, China

Abstract

This paper investigates the application of non-Gaussian AR model and the parametric bispectral estimation in analyzing normal and pathological heart sound signals. The non-Gaussian AR model of PCG signals (phonocardiogram) is used to detect quadratic nonlinear interactions and to classify the two patterns of phonocardiograms in terms of the parametric bispectral estimate. The bispectral cross-correlation is proposed to the order determination of the model. Several real PCG data are implemented to show that the quadratic nonlinearity exist in both normal and clinical heart sounds. It was found that parametric bispectral techniques are effective and useful tools in analyzing PCG and other biomedical signals, such as EMG, ECG and EEG.

1. Introduction

Cardiac auscultation has become a basic clinical tool to analyze heart sounds for a long time. However, quantitative phonocardiogram (PCG) and advanced processing technique have lagged behind. The understanding of heart sound mechanisms and the inherent complexity of PCG signals is still limited. With the development of digital signal processing techniques, many research results have shown that the measurement and processing of PCG signals are clinically significant because the signals involve plenty of useful information relating to the different states of the heart. A variety of methods, such as the sound spectrograph, envelope distribution, FFT technique, time-frequency representation and model-based methods, are widely used to detect the information of PCG signals for clinical purpose. Both time and frequency domain analysis are developed for the PCG signals analysis [1] [2] [3]. Parametric methods, such as Burg and Marple algorithm, are widely applied for the analysis and classification of PCG signals since the parametric methods give a better estimation of spectral features of the signals [4] [5]. However, many methods based on power spectrum or correlation for PCG signals processing start with the assumption of linearity, Gaussianility and minimum phase systems [6]. In other words, all these methods generally depend on the second order statistics. PCG signals often do not comply with these assumption in practical applications. To understand the exact feature and extract more information involved in phonocardiographic signals, this contribution proposes the higher-order spectra (HOS) for analysis and classification of heart sounds. HOS can reveal more information than power spectrum can. Cumulant or polyspectra, provides both amplitude and phase information of the signals [7] [8]. Furthermore, HOS provides a measurement of Gaussianility since both cumulant and polyspectra are identically zero for any stationary Gaussian processes [9]. Consequently, HOS-based analysis significantly increases the SNR when the signals are contaminated by additive Gaussian noise with unknown covariance.

The purpose of this contribution are three folds: to discuss non-Gaussian AR model of PCG signals, to

0-8186-8005-9/97 $10.00 © 1997 IEEE

determine the order of non-Gaussian AR model for PCG signals, and to analyze the classification of two patterns of phonocardiogram by bispectrum.

2. Method

2.1 Non-Gaussian AR model of PCG signals

It is assumed that the phonocardiographic signals are non-Gaussian and third-order stationary. Let a PCG sequence $\{s(k)\}$ be modelled by a non-Gaussian AR model as follows [10]:

$$s(n)+\sum_{k=1}^{p} a_k s(n-k) = w(n) \qquad (1)$$

where $w(n)$ is a white non-Gaussian i.i.d. process with zero mean and third-order stationary. $E[w^2(n)] = Q$ and $E[w^3(n) = \beta \neq 0$. The system given in (1) is assumed to be stable and $s(n)$ is statistically independent of $w(n)$ for $k \leq n$. It is followed from (1) that [11]

$$C_{3s}(-m,-n) = \beta \cdot \delta(m,n) - \qquad (2)$$

$$-\sum_{k=1}^{p} a_k C_{3s}(k-m,k-n); \quad m,n \geq 0$$

where $C_{3s}(m,n)$ is the third-order cumulant of sequences $s(k)$. The parameters $\{a_k\}$ can be decided by solving expresion (2) along the line $m = n = 0,1,\cdots,p$.

2.2 The order determination of non-Gaussian AR model

Like most cases for order selection in parametric spectral estimation, the model order in (2) must be considered. The criteria like FPE, AIC or CAT for order determination is not valid for the model given in expresion (1) since all of these criterias are developed under the assumption of second-order statistics [12]. Some methods for order decision of ARMA models were reported, such as the SVD approach [13] and information theory criteria [14]. Third-order statistics are found to be approapriate for order selection of non-Gaussian ARMA model in which the non-Gaussian input is contaminated by additive Gaussian noise with unknown covariance. For the model in (2), an approach for AR model order selection is proposed to the parametric bispectral estimation of the PCG signals [15]. The bispectral cross correlation (BCC) is defined as :

$$BBC = \sum_{\omega_1}\sum_{\omega_2} \left|D_s(\omega_1,\omega_2)\right| \cdot \left|B_{sp}(\omega_1,\omega_2)\right| \qquad (3)$$

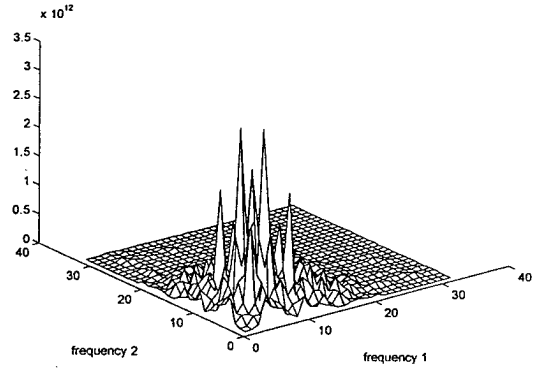
where $D_s(\omega_1,\omega_2)$ represents the bispectral estimate by using the conventional method, and $B_{sp}(\omega_1,\omega_2)$ is the parametric bispectral estimate with order p. Note that the energy content of each AR estimation should be normalized when using (3). The BCC changes with the AR model order p. The optimal order is decided when the maximum value of BCC occurs.

2.3 Pattern classification based on the AR model

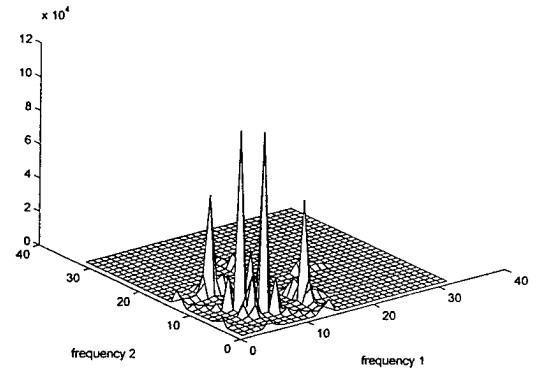
Since HOS retain both amplitude and phase information of the signal, the features of the HOS can be extracted for the purpose of clinical pattern classification. Moreover, features derived from HOS of the signals have high immunity to additive Gaussian noise. The parameters of non-Gaussian AR model are proposed to form the features vector for the pattern classification of PCG signals. The classification of normal and patient phonocardiogram is analyzed. A scheme for obtaining minimal set of parameters $\{a_k\}$ from equation (2) is selected as choosing features vectors. Both linear and nonlinear classifier for the different PCG signals are performed and tested. A group of normal and clinical phonocardiographic signals are employed for the two classes of pattern classification. The results with different order of non-Gaussian AR model are provided to demonstrated that the perfomance of patterns classification in terms of third-order statistics is effective and applicable. The perfomance of classification with different noisy background are investigated to exhibited the high SNR performance when parametric bispectral technique is applied to the pattern classification of the PCG signals.

3. Results and discussion

A computer-based system for PCG signals detection and data acquisition is carried out, in which the ECG was collected as a timing signal. 10 normal and 12 patient PCG signals were digitally recorded from Fuwai Hospital in Beijing and stored as data files for postprocessing. The analog PCG signals were converted to digital format through an A/D converter at a sampling rate of 1250 Hz and 2500 Hz. In order to understand the bispectral structure of PCG signals, both FFT method and parametric method are used for bispectral estimation of PCG. The amplitude bispectra in terms of conventional method and non-Gaussian AR meodel are shown in Fig.1 (a) and (b), respectively.

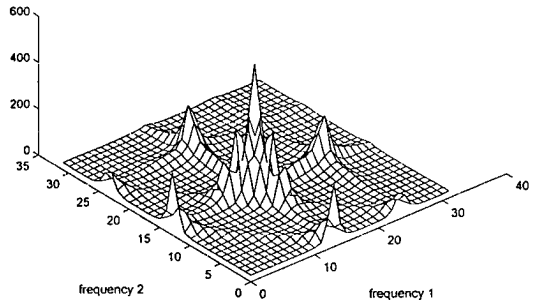


(a)

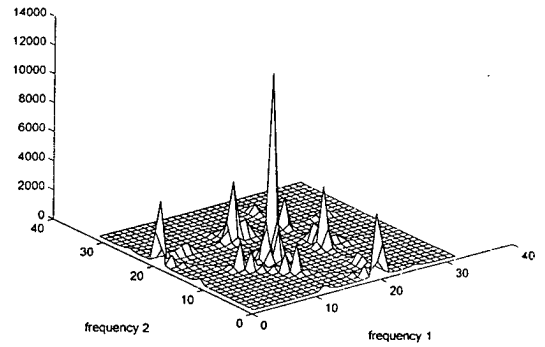


(b)

Fig. 1 The amplitude bispectra of normal PCG by FFT and parametric model



(a)



(b)

Fig.2 The bispectra of clinical PCG with MI and MS

By comparing the bispectral structure, it is found that parametric approach offers much better resolution of bispectral estimation. Fig.2 (a) and (b) give the parametric bispectral estimation obtained from two records of clinical PCG signals with heart disease of MI and MS. The results clearly indicate that different bispectral structure exist between the normal and some abnormal phonocardiograms. Consequently, PCG signals should be treated as a non-Gaussian process. For the normal phonocardiogram, there exist four bispectral peaks with the frequency from 110 Hz to 180 Hz in bifrequency domain. For the pathological phonocardiogram, four bispectral peaks also exist in the bifrequency domain, but the bispectral structure is

quite different from that of normal PCG. The bispectral peaks clearly move to high frequency area in bifrequency domain, with the frequency ranging from 180 Hz to 500 Hz since the clinical reasons and the murmurs. The bifrequencies corresponding to the significant levels of detected square nonlinearity are summarized in Table 1.

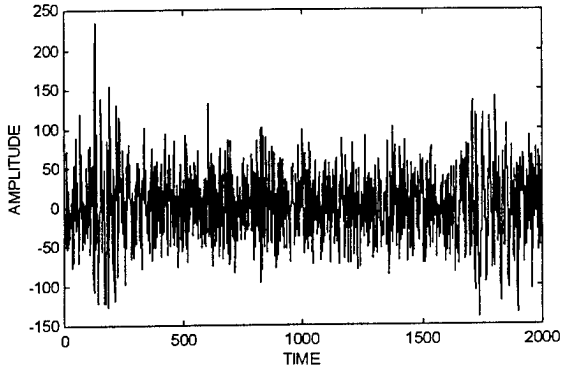
Table 1. Frequencies occuring bispecral peaks for three kinds of PCG signals in bifrequency domain.

	$B_1(f_1,f_2)$	$B_2(f_1,f_2)$	$B_3(f_1,f_2)$	$B_4(f_1,f_2)$
normal	120, 120	180, 120	280, 120	180, 100
clinical 1	280, 280	500, 280	280, 220	280, 180
clinical 2	270, 270	270, 160	470, 340	0, 270

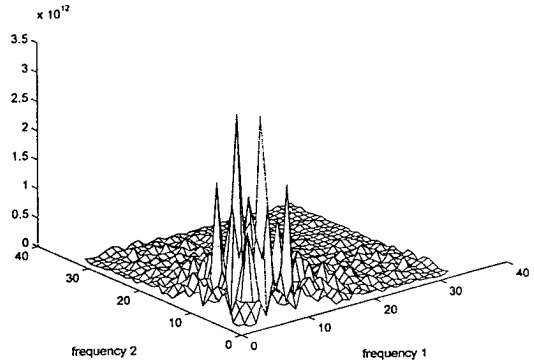
The background noise was tested for PCG bispectral estimation with different SNR. As shown in Fig.3 (a), one record of typical normal phonocardiographic signals contaminated with Gaussian white noise is illustrated. Fig.3 (b) shows the amplitude of parametric bispectral estimate. It can be seen that the bispectral analysis has high immunity to the background noise by comparing Fig.1 with Fig.3.

The bispectral cross correlation (BCC) is employed to the optimal order determination for the normal PCG non-Gaussian AR model. The bispectral estimation with different model orders p was implemented to estimate the BCC values. We observe that the maximum BCC value occurs when order p is equal to 13. Moreover, it can be seen that the BCC values change randomly as model order increases. We found that the optimal orders for non-Gaussian AR model is between 10 and 15 for a normal PCG sequence.

The model parameters stand for the third-order statistics of the signals. They are selected to form the features vectors for two classes of pattern classification of phonocardiogram, normal and clinical phonocardiograph with murmurs. Nonlinear method is applied to differiate the normal and the pathological phonocardiogram. These particular feature vectors, standing for the third-order statistics of PCG data, have provided a good classification accuracy even when the parameters with orders p are less than optimal order 13. A group of real normal and pathological PCG records are employed for two classes of pattern classification based on parametric bispectral estimate. Some experiment of classification of noisy patterns are also investigated to demonstrate the high SNR performance when third-order cumulant is used to the pattern classification of PCG.



(a)



(b)

Fig.3 The PCG signal contaminated with additive Gaussian noise and the bispectrum

4. Conclusion

The higher-order spectra was used in this paper to investigate the heart sound signals. Non-Gaussian AR model was fitted for the bispectral estimation and pattern classification of phonocardiogram. Parametric bispectral estimate of PCG were employed to detect and quantify the presence of quadratic phase coupling occurring in the different clinical phonocardiograms. Furthermore, The parameters of the non-Gaussian AR model were applied as the features for two classes of pattern clssification. Several real PCG data were

analyzed in terms of bispectrum. The results demonstrate that a PCG should be treated as a non-Gaussian process. Different bispectral structure exist in both normal and pathological heart sounds. The information of the bispectrum was used as the main features for the two pattern classification of the PCG signals. It was found that parametric bispectral techniques are useful tools in analyzing the quadratic nonlinear interactions of biomedical signals, such as PCG, ECG and EEG. Polyspectra may be an effective and useful tool for understanding the basic heart sound mechanism and improving diagnostic sensitivity via heart sounds.

References

[1] M. S. Obaidat. Phonocardiogram signal analysis: techniques and perfomance comparison. *Journal of Medical Engineering & Technology,* Vol. 17(6), November 1993.

[2] J. H. Tae. et al. Pole-zero modeling and classification of phonocardiograms. *IEEE Trans. on Biomedical Engineering, BME-30(2),* February 1983.

[3] M. R. Rangaraj. et al. Phonocardiogram signal analysis: a review. *Critical Reviews in Biomedical Engineering,* Vol. 15. Issue 3, 1988.

[4] S. L. Marple. A new autoregressive spectrum analysis algorithm. *IEEE Trans. on ASSP.* Vol. 28. 1980.

[5] M. Yokoi. Clinical evaluation on 5 years experience of automated phonocardiographic analysis. *Japan Heart Journal,* Vol. 18. 1977.

[6] M. L. Mendel Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *Proceedings of The IEEE,* Vol. 79, No. 3, March. 1991.

[7] M. M. Jerry and et al. Editorial applications of higher order statistics, *IEE Proceedings-F,* Vol. 140, No. 6. December 1993.

[8] M. R. Raghuveer and C. L. Nikias. Bispectrum estimation: a parametric approach. *IEEE Trans. on ASSP,* Vol. ASSP-33, October. 1985.

[9] R. Subba. *An Introduction to Bispectral Analysis and Bilinear Time Series Models,* New York: Spring-Verg, 1984.

[10] C. L. Nikias and M.R. Raghuveer. Bispectrum estimation: a digital signal processing framwork, *Proceedings of the IEEE,* Vol. 75, No. 7, July 1987.

[11] C. L. Nikias. *Higher-order spectra analysis, A nonlinear signal processing framework* PTR Prentice Hall, Eaglewood Cliff, New Jersey 1993

[12] K. T. Jitendra. Linear model validation and order selection using higher order statistics, *IEEE Trans. on Signal Processing,* Vol. 42, No. 7, July 1994

[13] G. Giannakis. et al. Cumulant-based order determination of non-Gaussian ARMA models, *IEEE Trans. on ASSP,* Vol. 38(8), Aug 1990.

[14] G. Giannakis and S. Shamsunder. Information theoretic criteria for non-Gaussian ARMA order determination and parameter estimation. *Proc. of ICASSP'93,* Vol. 4, USA. 1993

[15] J. Noonan and ea al. AR model order selection based on bispectral cross correlation, *IEEE Trans. on Signal Processing* , Vol. 39, No. 6. June 1991

COHERENT INTERFERENCE EXCISION USING HIGHER ORDER SPECTRA

Ismail Jouny and Aichen Low

Department of Electrical Engineering
Lafayette College
Easton, PA 18042

ABSTRACT

The bispectrum of a direct sequence spread spectrum communication signal contaminated with coupled multitone jamming provides significant information about the jammer frequencies and thus permits its excision using a bank of notch filters. Jammers that can be modeled as an autoregressive process can also be examined in the bispectral domain and mitigated using a linear FIR filter with temporally changing coefficients. The results indicate that utilizing bispectral analysis in DSS spread spectrum communications is an attractive alternative to conventional excision techniques and in some scenarios the only excision option available.

Keywords: Excision, bispectrum, spread spectrum.

1. INTRODUCTION

Interference mitigation in military communication systems can be accomplished using array processing in which a jammer or an undesired signal is spatially nulled in favor of the desired signal. Alternatively the desired narrowband signal may be spread spectrally prior to transmission according to a preset psuedo noise code and then despread upon reception. Simultaneously any channel interference is going to be spread at the receiver to a much lower signal level and then excised using the appropriate filtering technique. Spreading the signal prior to transmission is often accomplished by modulating the desired narrowband signal using a spreading sequence such as a psuedo noise code word sequence, or by modulating the desired signal in a psuedo random manner (frequency hopping).

Spread spectrum concept also permits code division multiple access communications (CDMA) in which many signals share the same time and frequency range but are multiplexed according to a combination of psuedo random sequences. CDMA has many civilian and military applications particularly in wireless cellular technology. Clearly excision of deliberate jamming as well as multipath interference is a problem that requires considerable attention. Excision can be accomplished using a filter whose coefficients can be estimated in a minimum mean square sense assuming prior knowledge of signal statistics, or using excision in the transform domain, the time-frequency domain, or the time-scale domain [3, 6]. Interference excision can also be accomplished using adaptive processing and learning techniques, closed loop estimation and locking methods, or using open loop filtering. The appropriate excision system depends on the nature of the jammer and its statistics and stationarity. The temporal characteristics of the interference are also important factors in determining the proper excision operation. In this paper we consider two specific type of jammers with inherent temporal changes, namely multitone jammers with implicit coupling, and jammers that can be modeled as the output of a non-Gaussian autoregressive system.

2. MULTITONE COUPLED JAMMING

Direct sequence spread-spectrum (DSS) communication systems are vulnerable to line-of-sight as well as multipath components of an interference source, in addition to the multipath components of the desired DSS communication signal. While the bispectrum of the desired communications signal is theoretically zero, and that of the correlated components of the multipath interference is not, it is argued that the bispectral information of a received communications signal can be used to identify the instantaneous frequencies of the correlated jammer components and can thus be utilized in placing a more accurate and effective spectral excision filter.

In this paper we extract bispectral information for a class of jammers with multipath components. The higher order spectrum of the desired signal is assumed known a priori. A bi-frequency excision filter is then applied to the spread-spectrum signal to remove the correlated spectral components of the interference. Time varying excision filters are of course needed to track the instantaneous frequencies of the correlated interference sources. The bispectrum is thus computed for rela-

0-8186-8005-9/97 $10.00 © 1997 IEEE

tively short data segments to maintain the stationarity of the jammer but long enough segments to obtain low variance bispectral estimates.

The desired direct sequence spread spectrum signal associated with the $n-th$ information bit $\{b_n\}$ of the $m-th$ user is

$$b_{nm}(t) = \sum_{k=1}^{L} p_{nm}(k) q(t - k\tau_c) \tag{1}$$

where L denote the number of PN chips per information bit, and $p_{nm}(k)$ represents the $k-th$ bit of the $m-th$ PN code and $q(t)$ is the chip pulse and τ_c is the chip duration. The desired DSS transmitted signal associated with the $m-th$ user is thus

$$s_m(t) = \sum_{n} I_n b_{nm}(t - nT_b) \tag{2}$$

where $T_b = L\tau_c$ represent the bit duration and I_n is the $n-th$ information bit. A code division multiple access signal is then expressed as

$$s(t) = \sum_{m} s_m(t) \tag{3}$$

The transmitted signal is corrupted with additive jamming and additive white noise and can be expressed as

$$r(t) = s(t) + j(t) + n(t) \tag{4}$$

where $n(t)$ is white zero-mean Gaussian noise. The multitone jammer is modeled as

$$j(t) = \sum_{l=1}^{P} \sum_{k=1}^{3} \alpha_{kl} \cos(\lambda_{kl} + \theta_{kl}) \tag{5}$$

where $\lambda_{3l} = \lambda_{2l} + \lambda_{1l}$ which indicates a coherent coupling between the jammers multitone components. The amplitude of these components are denoted by α_{kl}, and the angles θ_{kl} are uniformly distributed random phases over $(0, 2\pi)$ where $\theta_{3l} = \theta_{1l} + \theta_{2l}$. The number of coupled multitone signals is denoted by P.

The above jammer model is used in the simulation phase of this study, but the concept of excision using bispectral features can be implemented with any jammer that exhibits definable bispectral features and assuming that the bispectrum of the desired signal is theoretically zero.

The envisioned receiver estimates the bispectrum of the received signal which is the two-dimensional Fourier transform of the third order cumulant of the received signal $c_r(m,n)$

$$\begin{aligned} c_r(m,n) &= E\{r(t)r(t+m)r(t+n)\} \qquad (6) \\ B_r(\omega_1,\omega_2) &= \sum_{m}\sum_{n} C_r(m,n) e^{-j(\omega_1 m + \omega_2 n)} \end{aligned}$$

where $E\{\}$ denotes the statistical expectation. The additive white Gaussian noise is independent of the psuedo-noise like information signal, thus

$$B_r(\omega_1,\omega_2) = B_j(\omega_1,\omega_2) \tag{7}$$

and the bispectrum of the received signal is entirely due to the jammer signal and its multipath components and any other coupled components.

The bispectrum of the received signal has twelve symmetry regions in the bi-frequency plane. It is sufficient to examine the estimated bispectrum in one of these twelve symmetry areas where a peak search algorithm is implemented to locate those bispectral components that are above a certain level (the bispectrum of white Gaussian noise in our case). The search algorithm identifies peaks within a window of 8 or 12 bispectral pixels. The algorithm covers the entire bispectral domain area defined by $\omega_1 + \omega_2 < \pi$ which is sufficient assuming a properly sampled received signal. The frequency pair (ω_1, ω_2) associated with each bispectral peak indicates the presence of three jammer tones at frequencies ω_1, ω_2 and $\omega_1 + \omega_2$. The bispectrum may also identify a jammer tone who is not coupled with other tones, such a bispectral peak appears on one of the axes in the bifrequency plane. In addition to the bispectrum the receiver may also examine the power spectrum of the incoming signal as well as its time-frequency or time-scale signatures. Thus the proposed excision system could operate either independently or in conjunction with spectral based excision system. The bispectrum based excision system clearly identifies all jammer components that may have been missed using spectral analysis either because of their low level or because of the presence of a strong degree of coupling between jammer components. A nonparametric estimation of the bispectrum results in a frequency resolution that is equivalent to that of the spectrum and is limited by the entropy of the received signal. A nonparametric bispectral estimation technique could improve the quality of the estimates of the jammer frequencies. The nonparametric estimation of the jammer frequencies from their bispectral signatures is not included in this study.

Excision is accomplished using a cascade of notch filters or a multinotch filter tuned to the estimated jammer frequencies. Three types of filters were examined in this study: FIR linear phase filters

$$H\left(e^{j\omega}\right) = \sum_{k=1}^{M} b_k e^{-j\omega k} \tag{8}$$

where the symmetric coefficients are obtained such that the filter response at the nulling frequency λ_n, $H\left(e^{j\lambda}\right) =$

0. Three or five tap finite impulse response filters similar to those proposed in [1, 6] were utilized in this study to perform the appropriate excision. The results were inferior simply because such wide notch filters perform poorly in a cascade environment. Alternatively a multinotch FIR filter was designed and the coefficients were estimated in a least mean square sense to fit the constraints of the multinotch filter. The resulting filter exhibited high ripple error and did not produce sufficient jammer excision and resulted in significant signal distortion. Successful nulling of multitone jamming is accomplished using a cascade of IIR notch filters or using a multinotch IIR filter. A IIR filter is given as

$$H\left(e^{j\omega}\right) = \frac{e^{j2\omega} - 2e^{j\omega}\cos(\lambda_n) + 1}{e^{j2\omega} - 2\left(1 - \frac{B}{2}\right)e^{j\omega}\cos(\lambda_n) + (1 - B)} \tag{9}$$

where B is the notch bandwidth normalized with respect to the sampling period. A multinotch filter can also be used for excision of multitone signals

$$H\left(e^{j\omega}\right) = \frac{1}{1 + \sum_{n=1}^{P} \frac{B}{2}\left[U(e^{j(\omega-\lambda\ )}) + U(e^{(\omega+\lambda\ )})\right]} \tag{10}$$

where $U(x) = 1/(x-1)$, and B_n denotes the notch filter bandwidth at the frequency λ_n. While an IIR filter is easily implemented using today's high speed digital signal processors, phase nonlinearity which leads to signal distortion remains a concern. Furthermore, IIR filters who span segments of temporally changing frequency content are less desirable in a highly time-dependent changing jamming environment.

There are several factors that affect the performance of the proposed bispectral based excision system. First, the stationarity of the signal, the noise, and the jammer, and in particular the third order stationarity of the jammer. While the assumption of a stationary signal and a stationary additive noise is a valid one, the assumption of a perfect third order stationary multitone jammer is not absolute. The proposed excision system mitigates any jammer that exhibits genuine bispectral features at frequency pairs in the bispectral domain. It is therefore necessary to assume that the received signal is third order stationary within a short observation window (in the order of 50 or 60 information bits) or in a few nano seconds time period. This assumption over such a short period of time is arguably a valid one. Secondly, the issue of temporal change in the frequencies of the multitone coupled jammer. It is assumed in this study that the jammer frequencies do change temporally over very short periods of time (in the order of few nano seconds). The bispectral search algorithm is expected to detect any temporal change in the tones of the jammer. However, within each observation window which corresponds to few nano seconds of digital information, the jammer tones are assumed fixed and the signal is block filtered using the above excision systems. The bandwidth of the excision filter is chosen wide enough such that any temporal change in the jammer tones which is undetected by the bispectral search algorithm remains within the filter 3-db notch region. The third factor is the bandwidth of the notch filter and how it may change from one jammer tone to another. The notch filter bandwidth does effect the performance of the proposed excision system. A choice of B that is consistent with the search algorithm frequency resolution limits seem to be the best compromise. It is desirable, in the presence of coupling, that the notch at the third jammer tone be twice as wide as that of the frequency resolution simply because it is the sum of two estimated frequencies from the bifrequency plane. The difference between the excision performance of either of the above filters (cascade or multi-notch) seems to be negligible. Figure 3 shows the effect of applying bispectral based excision (bit error rate being zero for an 80 bit sequence) (top). Note that considerable bit error is experienced when the jammer frequencies change by about 1% of their original value (middle), and the bottom figure shows the effects of ignoring the jammer coupling component (at frequency $\lambda_1 + \lambda_2$) which is only identifiable in the bispectral domain. Clearly the information obtained from the bispectrum of the incoming signal is important for a satisfactory jammer excision.

The performance of the proposed multitone jamming excision system changes significantly when the transmitted signal is binary phase shift keying (BPSK) modulated and depends on the carrier frequency of the BPSK system. The effect is primarily due to the multinotch excision filter used and the separation between the jammer tones and the carrier. As the carrier frequency approaches either of the jammer frequencies, the bit error rate increases significantly, which indicates that in a BPSK mode the presence of multitone coupled jamming could have severe effect on the performance of a DS spread spectrum communications system.

3. NON-GAUSSIAN AR JAMMING

The second class of jamming sources considered in this study are those that can be modeled as MA, AR, or ARMA processes driven by non-Gaussian noise. It is well known that the ARMA parameters of a non-Gaussian random process cannot be accurately estimated using second order statistics or the power spectral density. The third order cumulant and the bispectrum are used to estimate the ARMA parameters

of a non-Gaussian jammer interfering with a pseudo random communication signal with zero bispectrum. An excision filter based on the estimate of the jammer ARMA parameters is then used to mitigate the effects of the jammer. The simulations performed in this study focus on the autoregressive case where the jammer is synthesized using

$$j(n) = \sum_{k=1}^{m} a_k j(n-k) + w(n) \qquad (11)$$

where $\{w(n)\}$ is a non-Gaussian white signal with constant bispectrum. It is noted in [1] that as a consequence of sampling at the chip rate, the desired signal $s(n)$ is uncorrelated with the received signal $r(n-l)$ for $l = 1, 2, \ldots, m$ provided that m is less than the length of the PN sequence. Therefore the jammer(without the signal $s(n)$) can be predicted from the received signal using

$$j(n) = \sum_{k=1}^{m} a_k r(n-k) \qquad (12)$$

Which indicates that the jammer AR parameters can be estimated from the bispectrum or the third order cumulant of the received signal after which the incoming signal is whitened using a transversal filter with coefficients equal to those of the estimated AR parameters given as

$$H\left(e^{j\omega}\right) = 1 - \sum_{k=1}^{m} a_k e^{-j\omega k} \qquad (13)$$

Clearly, a better estimate of the jammer AR parameters can be attained by using higher order statistics along with second order statistics. The are several algorithms proposed in the literature for estimating the parameters of a non-Gaussian AR processes. The method used in this study relies on using cumulant slices of second and third order [9]. An adaptive AR parameter estimation technique that relies on the third order statistics of the incoming signal is more desirable in scenarios of strong temporal dependence of the AR parameters.

4. CONCLUSIONS

Higher order spectral analysis in general and the bispectrum in particular provide a unique platform for identifying mulltitone jammers with inherent coupling or correlation. The bispectral frequency pair of each jammer component is estimated and an excision is applied at the estimated frequencies as well as their sum. Autoregressive non-Gaussian jammers are another class of interference that benefit from cumulant based excision where the jammer AR signal is whitened using a transversal reciprocal filter. The results show the significance of bispectral based excision techniques in direct sequence spread spectrum communications and they also show the unique ability of bispectral technique to mitigate the effects of two potential classes of jammers.

5. REFERENCES

[1] J. Ketchum and J. Proakis,"Adaptive algorithms for estimating and suppressing narrow band interference in PN spread spectrum systems",*IEEE Transactions on Communications*, Vol. 30, No. 5, pp. 913-924, May 1982.

[2] S. Tyler and M. Amin, "Mitigating interference in direct sequence spread spectrum systems," *Rome Laboratory Technical Journal*, 1st Volume, June pp. 35-45, 1995.

[3] M. Medley, G. Saulinier, and P. Das, "Applications of the wavelet transform in spread spectrum communications systems," *SPIE Wavelet Applications*, Orlando, Fl, Vol. 2242, pp. 54-68, April 1994.

[4] J. R. Glover Jr., "Adaptive Noise Cancelling Applied to Sinusoidal Interferences," *IEEE transactions on Acoustics, Speech, and Signal Processing*, Vol. 25, No. 6, pp. 484-491, December 1977.

[5] M. Medley, G. saulnier, and P. Das, " The application of wavelet-domain adaptive filtering to spread spectrum communications," *SPIE, Wavelet Applications*, Orlando, FL., April 1996.

[6] M. Amin, "Interference mitigation in spread spectrum communication systems using time-frequency distributions," *IEEE transactions on Signal Processing*, Vol. 45, No. 1, pp. 90-101, January 1997.

[7] E. Masry, "Closed-form analytical results for the rejection narrow-band interference in PN spread-spectrum systems-Part I: Linear prediction filters," *IEEE Transactions on Communications*, Vol. 32, No. 8, pp. 888-896, August 1984.

[8] M. Tazbey and A. Akansu, "Adaptive subband transforms in time-frequency excisers dor DSSS communications systems," *IEEE Transactions on Signal Processing*, Vol. 43, No.11, pp. 2776-2782, November 1995.

[9] C. Nikias and M. Raghuveer,"Bispectrum Estimation: A Digital Signal Processing Framework," *Proceedings of IEEE*, Vol. 75, No. 7, July 1987.

[10] L. Milstein and P. Das, "An analysis of a real-time transform domain filtering digital communication system: Part I: Narrowband Interference rejection," *IEEE Transactions on Communications*, Vol. 28, No. 6, pp. 816-824, June 1980.

[11] R. Ziemer and R. Peterson, *Digital Communications and Spread-Spectrum Systems*, Macmillan, New York, 1985.

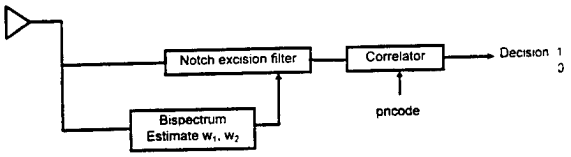


Figure 1: Bispectral based coupled multitone jammer excision system

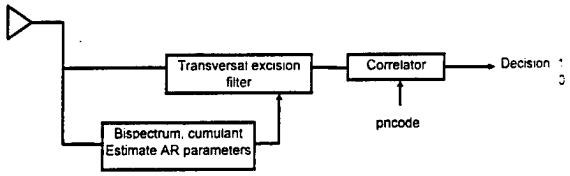


Figure 2: Cumulant based non-Gaussian AR jammer excision system

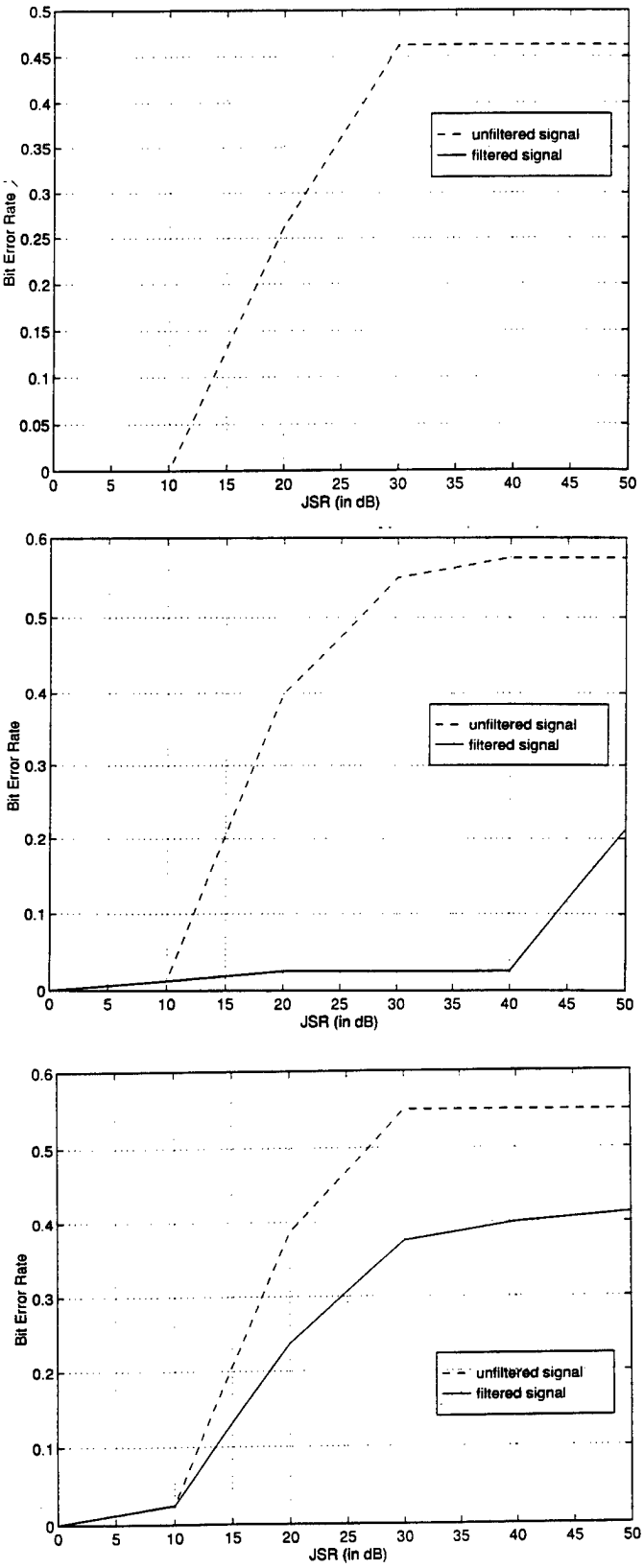


Figure 3: Performance in bit error rate versus jammer-to-signal ratio showing complete jammer excision (top), the effects of 1% temporal change in jammer frequencies (middle), and the effect of not nulling the third coupling component (bottom).

The Application of Non-linear Volterra Type Filters to Television Images

W.B.Collis [†‡], M. Weston [‡], P.R.White [†]

[†] *Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, UK.*

[‡] *Snell & Wilcox, Durford Mill, Petersfield, Hampshire, GU31 5AZ, UK.*

email: bc@isvr.soton.ac.uk

Abstract

Linear filter theory based on Wiener filtering is well understood and used widely in many fields of image and signal processing. However, the use of linear filters is generally associated with implicit approximations. Therefore, in this work a series of non-linear filters is developed based on the concepts of Volterra series and these are applied to image interpolation problems. More explicitly, the aim is to interpolate one field of a frame of a television picture to form an estimate of the second field. This is known as de-interlacing and is useful in many areas of video processing, for example standards conversion. Conventional de-interlacing systems use a fixed linear combination of the pixels in the aperture. In this paper we consider the extension of these methods to allow estimators based on non-linear combinations of pixel values.

1. Introduction

Realistic images have non-Gaussian statistics; hence a linear interpolator is not guaranteed to be optimal and there may be a performance gain to be exploited in employing a non-linear interpolation scheme [4]. Such a non-linear interpolation scheme can exploit any of the plethora of non-linear models available. However we concentrate on a Volterra type of model [1] in which only polynomial combinations of pixel values in the aperture are formed.

The de-interlacing problem can be summarised as: Given the knowledge of one field of a frame of a television picture, is it possible to reconstruct the other field? This is done by assuming knowledge of both fields of a particular frame and then calculating the best (linear/non-linear) filter that maps between them. This filter can then be applied to the first field from other frames in order to reconstruct the whole frame.

2. Non-linear filtering

For a linear system the relationship between the input x(t), the output y(t) and the system's impulse response h(t) is given by the convolution integral. For de-interlacing, if the process of mapping between one field and the other is a linear process it would be possible with the knowledge of the two fields to calculate the impulse response or filter, h(t), that when applied to field, f_1, gave exactly field f_2. In practice this filter would have to be of finite length and so it would not exactly give field f_2. However, in reality this mapping process is non-linear. This will be due to a number of reasons, for example, the field will contain aliased elements, and aliasing is a non-linear process. Hence it is likely that some form of non-linear operation will yield a better estimate of f_2 from f_1.

3. The Volterra Model

In time series analysis the Volterra model is thought of as an extension of the linear convolution integral. Rather than calculating the convolution of the linear impulse response of a system with its input, one considers an infinite sum of higher order impulse responses convolved with interactions of the input. A Volterra model, truncated at the second order, contains only linear and quadratic filters and so is only able to model systems which contain quadratic type non-linear elements. Such models are useful when identifying non-linear systems which skew the input. A Volterra model truncated at the third order, contains linear, quadratic, and cubic filters. These models allow the identification of non-linear systems which symmetrically distort the input signal. Higher order Volterra models allow even greater generalisations. There are two penalties associated with the use of higher order models; firstly the computational requirements increase dramatically, secondly the data lengths required to obtain accurate parameter estimates also increases. In discrete form the Volterra model can be expressed as,

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$y(n) = \sum_{k_1=0}^{L} h_1(k_1)x(n-k_1) + \sum_{k_1=0}^{L}\sum_{k_2=0}^{L} h_2(k_1,k_2)x(n-k_1)x(n-k_2) + \sum_{k_1=0}^{L}\sum_{k_2=0}^{L}\sum_{k_3=0}^{L} h_3(k_1,k_2,k_3)x(n-k_1)x(n-k_2)x(n-k_3) + \vdots + \sum_{k_1=0}^{L}\cdots\sum_{k_M=0}^{L} h_M(k_1,\cdots,k_M)x(n-k_1)\cdots x(n-k_M)$$

$h_M(k_1,...k_M)$ is called the M^{th} order Volterra kernel or Volterra filter. Note that a linear system is just a first order Volterra system and hence only has a first order filter $h_1(k)$. To fully characterise a non-linear system using the Volterra series it is necessary to calculate an infinite number of higher order filters. In practice this is not possible and so it is necessary to truncate the Volterra series at some order. In this work we will consider Volterra filters which are truncated at the second order, third order and fifth order.

4. Optimisation of filters

Consider a system as in figure 1. The object is to design an N point digital finite impulse response filter, h, to modify the input, x(n), in such a way as to minimise the mean square error, e(n), between the filter output and the desired signal, y(n). In the case of de-interlacing, x(n) is field f_1 and y(n) is field f_2. The aim is to create a filter, h, that when operating upon f_1 gives an optimal estimate of f_2 in the sense that the mean squared error is minimised.

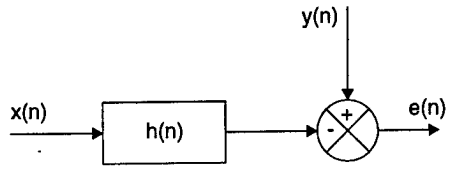


Figure 1: Linear filtering

The filter impulse response which minimises the sum of the squared errors of data of length L, is given by the solution of the over-determined (assuming L>N) system of equations

$$Xh = y \text{ where,}$$

$$X = \begin{bmatrix} x(L) & x(L-1) & \cdots & x(L-N+1) \\ x(L-1) & x(L-2) & \cdots & x(L-N) \\ \vdots & \vdots & & \vdots \\ x(2) & x(1) & \cdots & 0 \\ x(1) & 0 & \cdots & 0 \end{bmatrix} \quad y = \begin{bmatrix} y(L) \\ y(L-1) \\ \vdots \\ y(2) \\ y(1) \end{bmatrix}$$

the least squares solution of which is [2],

$$h = (X^TX)^{-1}X^Ty.$$

Note X^TX and X^TY are usually much smaller than X. Hence, it is much more efficient to compute X^TX and X^TY directly from x(n) and y(n) rather than to form X.

The same method can be extended to obtain an estimate of the M^{th} order non-linear filter. For the sake of this explanation, a second order Volterra system which contains only linear and quadratic filters will be considered, although it is conceptually easy to extend the method to any arbitrary order.

The extension of this to a more general Volterra model is in principle simply a matter of modifying the data matrix X. Below we show the data matrix for a second order Volterra model, in which a constant (DC) term has also been included. A symmetric form for the Volterra kernel has been assumed so this matrix has dimension $L \times \left(N + \frac{N(N+1)}{2}\right)$.

$$X = \begin{bmatrix} 1 & x(L) & x(L-1) & \cdots & x(L-N+1) & x(L)^2 & x(L)x(L-1) & \cdots & x(L)x(L-N+1) & x(L-1)^2 & \cdots & x(L-N+1)^2 \\ 1 & x(L-1) & x(L-2) & \ldots & x(L-N) & x(L-1)^2 & x(L-1)x(L-2) & \cdots & x(L-1)x(L-N) & x(L-2)^2 & \cdots & x(L-N)^2 \\ 1 & x(L-2) & x(L-3) & \ldots & x(L-N-1) & x(L-2)^2 & x(L-2)x(L-3) & \ldots & x(L-1)x(L-N-1) & x(L-3)^2 & \cdots & x(L-N-1)^2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & x(3) & x(2) & \cdots & 0 & x(3)^2 & x(3)x(2) & \cdots & 0 & x(2)^2 & \cdots & 0 \\ 1 & x(2) & x(1) & \cdots & 0 & x(2)^2 & x(2)x(1) & \cdots & 0 & x(1)^2 & \cdots & 0 \\ 1 & x(1) & 0 & \cdots & 0 & x(1)^2 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

The optimal filter, in the least squares sense, can then be estimated by computing $(X^TX)^{-1}X^Ty$. The filter will contain three separate components, the DC term, the standard linear coefficients which should be multiplied by single pixel values, and the quadratic coefficients which will be multiplied by products of pixel values. A six point cubic Volterra filter is shown as an example in figure 3. The first point is the DC term, the next six points are the linear coefficients which can be seen, typically, to have an approximate sin(x)/x structure. The next 21 points are the quadratic coefficients, and the final 56 points are the cubic coefficients.

5. Choice of filter order and aperture

A Volterra model truncated at the third order is used predominantly in this work. This will contains linear, quadratic, and cubic filters and so is able to model systems which contain both quadratic and cubic non-linear elements. These generate both skewed and symmetric distortions of the probability density function. Higher order models can be used and are shown to give improved results but the size of the filter and the computation required in its estimation rise exponentially and there are rapidly diminishing returns. For example, the fifth order, six pixel cubic Volterra filter does perform better than the third order six pixel filter but there are over five times as many terms.

For the linear case a simple four point vertical filter, see figure 2, gives an error of 18.61. Increasing the number of taps does not significantly reduce the mean squared error, (the error for a eight point vertical filter is 18.58, and for a 36 point filter is 18.55), nor does utilising pixels in the horizontal direction significantly reduce the error.

For the non-linear case the choice of aperture has much more dramatic results. A two-dimensional aperture does give a significant improvement over a one-dimensional one. This is thought to be due to the ability of the non-linear filter to deal with sloping edges and lines and hence the filter needs gradient information. However the number of filter coefficients rises rapidly with the number of pixels. Due to computational constraints the maximum size for a cubic filter is 20 pixels and for a fifth order filter is 6 pixels.

As the number of pixels available is limited it is important to choose the correct shape of aperture. A good compromise are apertures that contain four vertical pixels and then a number of horizontal pixels. The apertures used for the 4,6,8 and 20 pixel filters are shown in figure 2 (× denotes the pixels used in field, f_1, to estimate the pixel denoted by • in field, f_2). It is thought that the horizontal pixels help to cope with the near horizontal lines and edges that often cause problems due to jagging in a de-interlaced picture.

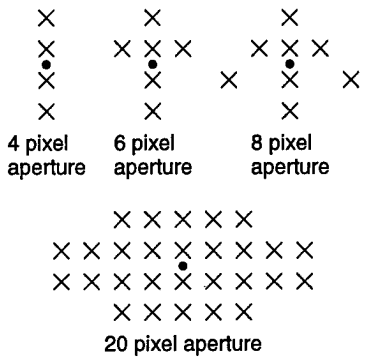


Figure 2: Choice of aperture for non-linear filters

6. Results for Volterra filters

Table 1 shows the mean squared error between the estimated field and actual field for the picture 'girl' for a series of different filters. It can be seen that in all cases the non-linear filters perform better than standard linear filters. The higher order non-linear filters produce very good results, with smooth edges on both curves and straight lines and little jagging. They only differ from the original field in regions of fine detail where it would be unexpected for any filter to work as they cannot create information that is not in the input field.

Filter type	Mean square error	Number of coefficients in filter
4 pixel linear filter	18.61	4
8 pixel linear filter	18.58	8
36 pixel linear filter	18.55	36
4 pixel cubic filter	16.14	35
6 pixel cubic filter	15.67	84
8 pixel cubic filter	15.21	165
20 pixel cubic filter	13.50	1770
6 pixel fifth order filter	14.88	462

Table 1: Mean squared errors for various filters used on picture 'girl'

Figures 4 and 5 show a small section from a de-interlaced version of the picture 'girl,' estimated with an optimum four point filter, and with an optimum twenty pixel cubic Volterra filter, respectively. These can be compared with the original picture, shown in figure 6. It can be seen how the non-linear filter produces much smoother edges and curves than its linear counterpart with

reduced jagging. This is most evident around the inner curve of the table leg.

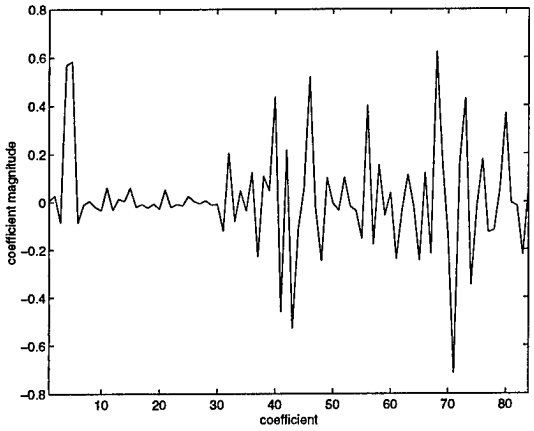


Figure 3: Filter coefficients for a six pixel cubic Volterra filter

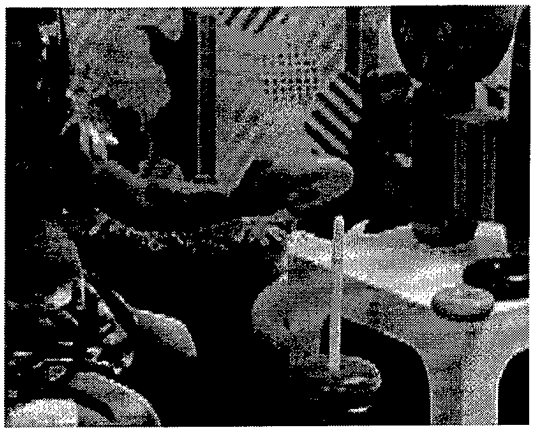


Figure 4: Section of 'girl' estimated with a 4 point linear filter

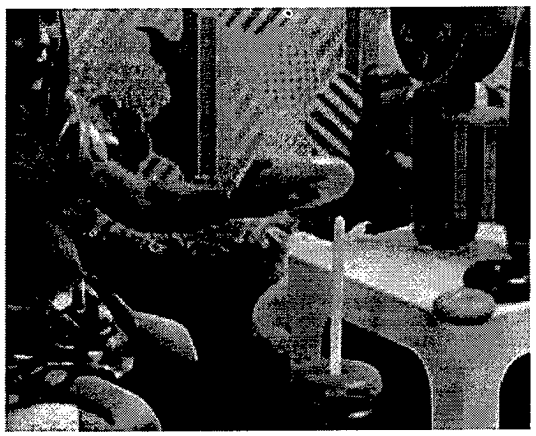


Figure 5: Section of 'girl' estimated with a 20 point cubic Volterra filter

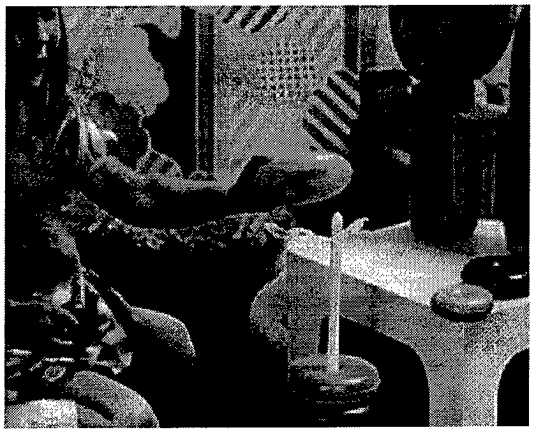


Figure 6: Section of 'girl' from original field

7. Conclusions

- Non-linear Volterra type filters can give dramatically improved performance over conventional linear predictors when used for de-interlacing.
- Volterra filters are generally more complex than their linear equivalents, both to design and implement.
- So far these techniques have only been applied to de-interlacing, but there are many other areas in this field where we can expect non-linear techniques to give improved performance.

8. References

[1] W.B.Collis, 'Higher order spectra and their application to non-linear mechanical systems', PhD Thesis, University of Southampton, (1996)

[2] S.Haykin, '**Adaptive filter theory**', Prentice Hall, (1986)

[3] M.Schetzen, '**The Volterra and Wiener theories of non-linear systems**', John Wiley & Sons, (1980)

[4] G.L.Sicuranza, 'Quadratic filters for signal-processing', Proceedings of the IEEE, Vol. 80, No. 8, pp. 1263-1285, (1992)

Higher Order Statistics for Detection and Classification of Faulty Fanbelts Using Acoustical Analysis.

G.Gelle, M.Colas and G.Delaunay.
Laboratoire d'application de la microélectronique, Université de Reims
BP 1039 ; 51687 Reims Cedex 2
France
e-mail : guillaume.gelle@univ-reims.fr

Abstract

Higher Order Statistics (HOS) are well suited to solving detection and classification problems because they can suppress gaussian noise and preserve some of the non-gaussian information [1] [2]. This paper describes the use of these methods for acoustic quality control of manufactured goods on a production line, and specifically the detection of faulty fanbelts on the drying block of a washing machine. Two HOS based methods were used in this paper. The first is based on the properties of the bispectrum in the outer triangle and particularly of the normalized bispectrum also called skewness function.The second uses the third order cumulant of a Matched Filter output. This combination has the advantages of matched filtering plus the properties of higher than second order statistics making this algorithm more robust than the conventional matched filter. The method was used to classify real signals from fanbelts suffering from specific known defects.

1.Introduction

For detecting defaults in manufactured goods in a production line, acoustical quality control is a very interesting tool due to its simple handling. Sounds frequently reveal the defects in an engine and can even identify the fault. This study was designed to detect non destructively faulty fanbelts on the drying block of a washing machine at the end of the production line. The fanbelt defect appears like a shock on the acoustical signature and changes the shape of its probality density function. With this point of view, we suggest to use Higher Order Statistics . In fact, HOS with their ability to suppress gaussian noise and to preserve some of non-gaussian information seem to be more interesting for analysing measurements made in a noisy environment.

The first section describes the use of a bispectrum based detector. This detector uses the properties of the bispectrum which are : for a stationary and unaliased signal, the bispectrum will be zero in the Outer Triangle (OT) [3]. When the process is nonstationary, Hinich and Wolinsky [4] have shown that the bispectrum is usually not zero. A transient cause the observed signal to becomes nonstationary in the window where the shock is present.

The second section examines another detection scheme using HOS. The proposed algorithm was performed by Giannakis and Tsatsanis [5], and combines the matched filtering operations with properties of HOS.

The last section examine two different fanbelts classes which are called "click" and "beating". The beat is caused by defective belt tension due to its pull. The second faulty class is due to the fanbelt hitting the drying block of the washing machine, producing a sharp sound in the acoustical signature.

To classify the different classes we use the work performed by Giannakis and Tsatsannis and the results were compared to those obtained with the conventionnal matched filter.

2.The detection hypothesis testing

Let us consider the following simple binary hypothesis testing problem :

$$\begin{cases} H_0 : X(i) = N(i) \\ H_1 : X(i) = S(i) + N(i) \end{cases}$$

for $i = 0,1....N$.

where $S(i)$ is a known reference signal and $N(i)$ is a sample realization from a zero-mean, gaussian noise process with unknown covariance sequence. The detection method must decide between H_0 and H_1 by analysing the received signal $S(i)$.

If $S(i)$ is a non-gaussian signal with non-zero kth-order polyspectrum (or moment spectrum), and if the noise is gaussian, then, their cumulant $c_{KN}(u_1,...,u_{K-1}) = 0$ and their higher order spectra $C_{KN}(f_1,...,f_{K-1}) = 0$ over their corresponding time and frequency domain. The above

0-8186-8005-9/97 $10.00 © 1997 IEEE

hypothesis testing can be reformulated in the higher order domain as :

$$\begin{cases} H_0 : c_{KX}(u_1,\ldots,u_{K-1}) = 0 \\ H_1 : c_{KX}(u_1,\ldots,u_{K-1}) \neq 0 \end{cases}$$

in the time domain for all lags $(u_1,\ldots,u_{K-1})$.
Similarly in the frequency domain we have :

$$\begin{cases} H_0 : C_{KX}(f_1,\ldots,f_{K-1}) = 0 \\ H_1 : C_{KX}(f_1,\ldots,f_{K-1}) \neq 0 \end{cases}$$

on all the principal domain PD_K.
A thorough review of higher order statistics can be found in [4].

3.The bispectrum based detector

Reviews of the bispectrum can be found in [1] and [6]. However, there is need for action to remind the properties of the bispectrum in the principal domain.

In fact, in respect with their symetry properties, it's sufficient to consider the behavior of the bispectrum in the principal domain (figure 1), defined as the triangle $\{0 \le g \le f ; 2f + g \le N\}$.

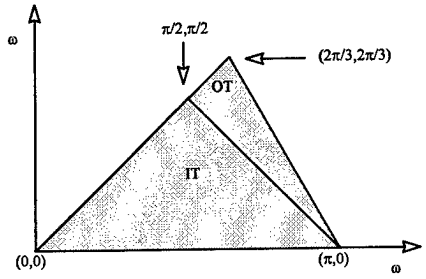


Figure 1 : The Principal domain (in normalized pulsation)

The principal domain is divided into two regions. The first region, satisfying $\{f + g \le N/2\}$ is called the inner triangle IT ($IT = Triangle[(0,0),(\pi,0),(\frac{\pi}{2},\frac{\pi}{2})]$ in figure 1). The remaining region is called the outer triangle OT ($OT = Triangle[(\pi,0),(\frac{\pi}{2},\frac{\pi}{2}),(\frac{2\pi}{3},\frac{2\pi}{3})]$ in figure 2).

For a stationary unaliased signal, the bispectrum is zero in the OT triangle. And more, if the signal is gaussian, the bispectrum will be theoritically zero in all the principal domain. If the process is nonstationary Hinich and Wolinsky [4] proved that the bispectrum will usually be nonzero. A non-random transient signal in additive stationary gaussian noise makes the observed signal nonstationary in the window where the signal is present. See also [3] and [7] for more detail.

For a sampled signal, an unbiased estimate of the discrete bispectrum called the biperiodogram is given by :

$$B(f,g) = \frac{1}{N+1} X(f)X(g)X^*(f+g) \qquad (1)$$

where $X(f) = \sum_{n=0}^{N} x(n)e^{-2\pi ifn/N+1}$ represents the discrete Fourier transform of the N+1 samples $\{x(n); 0 \le n \le N\}$.

The estimate of the biperiodogram is unbiased but its variance is proportional to the product of the Fourier transforms. Furthermore, as N becomes larger, the variance of the estimate increases and the estimate becomes inconsistent.

A consistent estimate can be obtained by dividing the data record into K frames of L samples, with or without overlapping. A time domain window can be applied to each frame and the FFT computed by :

$$X_k(f) = \sum_{n=(k-1)L+1}^{kL} x(n)e^{-2\pi ifn/L}$$

The bispectrum $B_k(f,g)$ of the kth record is computed as in (1). The final estimate is obtained by averaging across the K frames :

$$\hat{B}(f,g) = \frac{1}{K}\sum_{k=1}^{K} B_k(f,g)$$

The method used in this paper is based on the change of amplitude and temporal skew of the signal when a transient appears. A natural measure of this skew is the third order cumulant. Parsons and Williams [8] showed that this measure can be constructed in the bispectral domain as follow :

- for amplitude skew :

$$d_a = \sum_{f,g} \mathrm{Re}(B(f,g))^2$$

and, for temporal skew :

$$d_t = \sum_{f,g} \mathrm{Im}(B(f,g))^2$$

If these measures are summed, and integrated in the OT triangle, we can construct a more robust estimate for testing the deviation to stationarity when the transient signal appears. The output of this bispectral energy detector will be :

$$d = \sum_{OT} |B(f,g)|^2$$

And the detection scheme is :

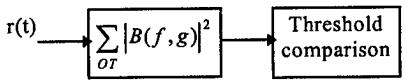


Figure 2 : block diagram of the bispectral energy detector

The hypothesis testing becomes :

- If hypothesis H_0 is true, the measure d will be zero.
- If hypothesis H_1 is true, the signal will be non-stationary and the measure d in OT will be different from zero. The detection of transient signal presence will be realized by thresholding. The performances of this detector are given in simulation in [9].

The performances of the test can be improved by using the normalized bispectrum also call the Skewness function :

$$Bsk(f,g) = \frac{|B(f,g)|^2}{P(f)P(g)P(f+g)}$$

where $P(f)$ is the final estimate of the power spectrum. When the transient appears, the measure d is defined now as :

$$d = \sum_{OT} Bsk(p,q)$$

increases because the signal becomes locally non-stationary.

We can see on the figure 3 the results of the skewness function with (a) and without a fault (b) :

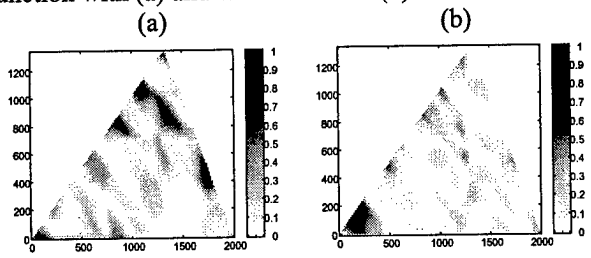


Figure 3 : Skewness function with and without a fault

But this method is not adequate for our application. In fact, the test was adequate if the size of the analysing window was adapted to the transient size. In our case it is not possible to know the duration of the transient in the acoustical signature and the results were very poor. For this reason we prefer to use the algorithm described in the next section.

4.Matched filtering and HOS detector

Let us consider that the reference signal has a finite impulse response $s(i)$, so that, the filter matched to this signal has a finite impulse response $h(i) = s(N-i)$. If the MF is excited by $x(i)$, then the kth-correlation of the MF's output $y(i)$ is given by :

$$y_K(i_1,\ldots,i_{K-1}) = \sum_{j_1,\ldots,j_{K-1}=-N}^{N} x_K(j_1,\ldots,j_{K-1}) s_K(j_1-i_1,\ldots,j_{K-1}-i_{K-1})$$

in a noise free case we obtain :

$$y_K(i_1,\ldots,i_{K-1}) = \sum_{j_1,\ldots,j_{K-1}=-N}^{N} s_K(j_1,\ldots,j_{K-1}) s_K(j_1-i_1,\ldots,j_{K-1}-i_{K-1})$$

and for zero lag :

$$y_K(0,\ldots,0) = \sum_{j_1,\ldots,j_{K-1}=-N}^{N} y_K^2(j_1,\ldots,j_{K-1}) = E_{KS}$$

where E_{KS} is the energy of the signal's kth-order correlation. The Cauchy-Schwartz inequality can be use to show that :

$$|y_K(i_1,\ldots,i_{K-1})| \le |y_K(0,\ldots,0)|$$

therefore, the kth-order correlation peaks at the origin. This is the basis for the detection algorithm proposed by Giannakis and Tsatsanis in [5].

- For the H_0 hypothesis, the kth-correlation of the MF's output leads toward c_{KN} which is zero (for k>2) if the noise is gaussian.

- According to the H_1 hypothesis, the kth-order cumulant of the MF's output leads toward the kth-order cumulant energy c_{KS}^2 of the signal $s(i)$ if the signal has a zero mean and if several independant records are available. (see [5] for details)

The detection rule becomes a comparison of $|c_{KY}(0,\ldots,0)|$, the absolute value of zeroth kth-order cumulant lag of the filtered observation sequence with a threshold. For k=3, the block diagram of the detector is shown in figure 4.

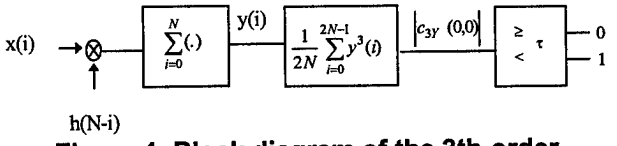


Figure 4 :Block diagram of the 3th-order cumulant based detector

In practice, the true correlation value is not available. This value can be estimated by computing the three order correlation directly, and for zeroth lag we obtain the simple sum :

$$y_K(0,0) = \frac{1}{2N+1}\sum_{i=0}^{2N} y^k(0,0)$$

however, cumulant squared with correlation up to the third order statistics. If k≥4, we need to compute the kth order cumulant of the MF's output. We can also compute a more robust estimate by averaging over several independent segments, insensitivity to gaussian noise being true in the mean .

This detector can be viewed as a cumulant energy detector or in the frequency domain like a polyspectral energy detector. This algorithm is attractive because it requires the computation of a single cumulant lag.

Our application has not one but several reference signals (for all classes of fanbelts). Then, we are interested in generalizing detection algorithm to classification problems. In this case, we have a bank of L MF's with impulse response $\left\{h^{(l)}(i)\right\}_{l=1}^{L}$. The filters are normalized to have a zero mean and equal kth-order cumulant energy E_K. One way to accomplish this is by scaling i.e. :

$$h^{(l)}(i) = \lambda_l s^{(l)}(N-i)$$

Where $\lambda^{(l)}$ are chosen such that $E_K = \lambda_l^{2K} E_{KS_l}$ [5].

The classification was done with the properties of higher order statistics that can be summarized :"if a bank of

equal correlation energy MF's is used to classify an incoming signal, the maximum zeroth kth-order correlation appears at the output of the filter matched to the incoming signal".

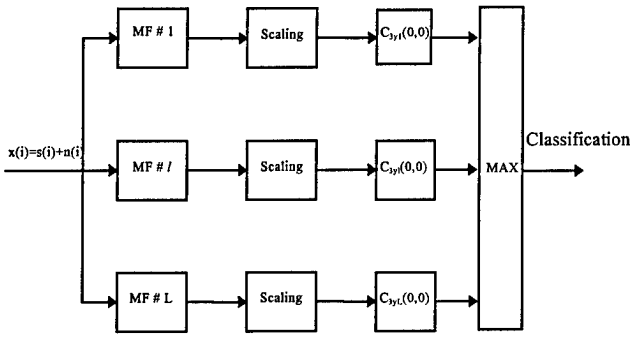


Figure 5 : The third order cumulant-based classifier

5. Experimental results

The reference signatures for the design of each matched filter, were obtained by synchronous averaging over 1500 records made in an anechoid chamber for each class of fanbelts. Fanbelts were assigned unambiguously in a reference class by a human expert. When the decision was not unambiguous, the fanbelt with the greatest ambiguity was excluded from the classification testing.

Test signals were obtained by synchronous averaging over ten records and sent to the filter bank. The estimating cumulant was obtained by averaging many cumulants of MF's output from several independant records.

The results are shown in table 1 for 15 averages of the filter output :

true conditions	Classified conditions		
	'good '	'beat'	'click'
'good'	**86.5%**	0%	13.5%
'beat'	6.5%	**67%**	26.5%
'click'	0%	0%	**100%**

Table 1 : results of the classification procedure with HOS classifiers

and for conventionnal matched filter :

true conditions	Classified conditions		
	'good '	'beat'	'click'
'good '	**69%**	11%	20%
'beat'	14%	**63%**	23%
'click'	6.2%	30%	**63.8%**

Table 2 : results of the classification procedure with conventionnal MF's classifiers

The above table 2 shows that the results for the HOS classifier were much better than those for the MF classifier. This improvement is due to the HOS classifier being less sensitive to ambiant noise. The number of averages in the cumulant estimation also influence detection.

6. Conclusion

This paper, describes the use of higher order statistics to classify fanbelt defects by acoustical analysis. Their use resulted in excellent performance of the third order cumulant detector with regard to the matched filter. Performance can be improved by using larger records and numbers of averages for the cumulant estimation. Other methods are presently being investigated to improve the classification. Thus time-frequency appears to be a promising tool [10].

7.references

[1] : C.L. Nikias and A.P. Petropulu, "Higher-Order Spectra Analysis. A Nonlinear Signal Processing Framework", Prentice Hall Signal Processing Series, 1993.

[2] : L.M.Garth and H.V. Poor, "Detection of Non-Gaussian Signals : A Paradigm for Modern Statistical Signal Processing", in Proceedings of the IEEE, Vol 82, No 7,July 1994.

[3] : M.J.Hinich, "Detecting a transient signal by bispectral analysis", IEEE trans on ASSP, vol 38, no 7, pp1277-1283, July 1990

[4] : M.J.Hinich and M.A. Wolinsky, "A test for aliasing using bispectral analysis", J. Amer., Stat. Assos., Vol 83, No 402, pp. 499-502, 1988.

[5] : G.B. Giannakis and M.K. Tsatsanis, "Signal Detection and Classification using Matched Filtering and Higher-Order Statistics", in IEEE Transactions on ASSP, vol 38, No 7, July 1990

[6] : D.R.Brillinger and M.Rosenblatt, "Asympototic theory of estimates of kth-order spectra," in Spectral Analysis of times series, B.Harris, Ed. New York : Wiley , 1967.

[7] : M.J.Hinich and H.Messer,"on the principal domain of the discrete bispectrum of a stationary Signal", IEEE transaction on signal processing vol 43, n°9,1995.

[8] : A.T.Parsons, M.L.Williams, "The construction of broadband Higher-Order Spectral measures", Proc. Of the IEEE Signal Processing Workshop on Higher Order statistics, pp 85-89, Chamrousse, 1991.

[9] : A.Moineau, G.Gelle, G.Delaunay,"Detection of transient signal in additive gaussian noise : comparison between time-frequency analysis and higher order statistics approaches", proc. of the IASTED International Conference Signal and Image Processing, Nov 11-14, 1996 Orlando,USA, pp 269-273.

[10] : G. Gelle, A. Moineau, G. Delaunay, "Détection et classification de courroies défectueuses par analyse acoustique : comparaison entre une approche temps-fréquence et une approche statistique d'ordre supérieure.", to be published in Colloque GRETSI 97, Grenoble France.

CLASSIFICATION OF LINEAR MODULATIONS BY A COMBINATION OF DIFFERENT ORDERS CYCLIC CUMULANTS

Pierre Marchand
TEAMLOG
"Le Grand Sablon"
4 av. de l'Obiou
38700 La Tronche, France
Pierre.Marchand@cephag.inpg.fr

Christophe Le Martret
Centre d'ÉLectronique de l'Armement (CELAR)
35170 Bruz, France
lemartre@celar.fr

Jean-Louis Lacoume
CEPHAG-ENSIEG, BP 46
BP 46, 38402 Saint Martin d'Hères Cedex, France
Jean-Louis.Lacoume@cephag.inpg.fr

Abstract

The search for discriminating features is a crucial point when a modulation classification task is aimed. This paper introduces new features based on a combination of fourth- and second- order temporal cyclic cumulants. Such a combination enhances the theoretical discrimination that can be achieved by a single stationary cumulant, and moreover, the cyclic parameters become discriminating, whereas it is not the case when they are considered at pure orders. As an application, we propose a process to classify 4-PSK vs 16-QAM modulations. The classification is achieved by estimating the feature for the received signal, and comparing it with theoretical ones by a matched filter technique. Simulations show that though the cyclic parameters are a priori more discriminating than their stationary counterparts, the variance of their estimates may overcome this advantage.

1. Problem statement

The modulation classification process consists in determining the modulation type of an intercepted signal corrupted by noise. This is a challenging problem that has been investigated for several years. Since a general approach of the modulation classification process is very difficult to draw up, each case has to be considered separately, and specific methods must be found according to the classes of modulation we want to discriminate. One of the classical way to tackle with the problem of modulation classification is based on maximum likelihood theory [2]. It consists in processing the log-likelihood function of the signal (or an approximation thereof) and then comparing it to an appropriate threshold. This very general approach can be restricted to some features extracted from the signal, for which the theoretical expressions are perfectly known. For example, QAM modulation classification has been recently achieved [3] using stationary fourth-order cumulants of the signal as discriminating features.

In our paper, we generalize the discriminating feature given in [3], by considering the cyclostationary property of $x(t)$ and enlightening an original way to enhance the discrimination.

The use of (second- or higher-order) cyclostationarity property of time series has already proved to be very helpful in many applications [1][8], especially to characterize telecommunication signals. The use of this property enables access to more information about the analyzed signal and can be very helpful when a modulation classification task is addressed. In our application, we will show that the introduction of the cyclostationary property leads to features that are more discriminating as far as a theoretical point of view is considered. However, since reasonably good estimates of cyclic parameters require much more samples than for stationary parameters, both methods have to be

0-8186-8005-9/97 $10.00 © 1997 IEEE

compared, which will be done in the sequel.

In our algorithm the discriminating feature is obtained by a proper combination of second- and fourth-order cyclic temporal cumulants. This approach is quite new for two reasons. Firstly, we build a new statistic which is neither a cumulant nor a moment, but is choosen in order to be the more discriminating as possible. Secondly, cyclic statistics have not been used yet to classify M-QAM ($M \geq 4$) signals. Moreover, it enforces the idea that several orders must be combined to achieve modulation classification, which has already been pointed in a different manner in [7]. In the sequel, details are given only for the 4-PSK *vs* 16-QAM case (4-PSK can be seen as a 4-QAM), but the method can be extended to more than two classes.

The paper is composed as follows. In section 2, we derive the expressions of fourth- and second order cyclic correlations for digital modulations. In section 3, the basic idea of the classifier is explained, and the general structure of the algorithm is given. Simulations are provided in section 4, where our algorithm is compared to its stationary counterpart.

2. Digital modulations and their cyclic multicorrelations

2.1. Signals of interest

In our study, we are interested in M-QAM modulation classification (*i.e.* M-states Quadrature Amplitude Modulation). The analytic signal representation for these modulations is given by:

$$x(t) = e^{2i\pi f_c t} \sum_k s_k q(t - kT_b - t_0) \qquad (1)$$

where $\{s_k = a_k + i.b_k\}_{k \in Z}$ is a complex-valued, zero-mean, and *i.i.d.* symbol sequence, T_b is the symbol duration, f_c is the carrier frequency, $q(t)$ is the real-valued pulse function, and t_0 is a non-random time shift. In this paper, we do not stationarize the signals, and consequently, time-dependency must be taken into account when expressing the temporal cumulants of (1). In other words, $x(t)$ is modeled as a cyclostationary process.

The received signal, for which we want to determine to type of modulation, can be written as:

$$x(t) = s(t) + b(t) \qquad (2)$$

where $b(t)$ is a stationary white gaussian noise whose power is unknown.

2.2. Cyclic multicorrelations

Let $C_{x,p+q,p}(t;\underline{\tau})$ be the $(p+q)$th-order cumulant-based correlation of the process $x(t)$, defined with p non-conjugated terms and q conjugated terms, as in [4]:

$$\begin{aligned} &C_{x,p+q,p}(t;\underline{\tau}) = \\ &\quad Cum[x(t), x(t+\tau_1), ..., x(t+\tau_{p-1}), \\ &\qquad x^*(t-\tau_p), ..., x^*(t-\tau_{p+q-1})] \end{aligned} \qquad (3)$$

Since $x(t)$ is almost-cyclostationary, there are at most countably many values of α for which the so-called $(p+q)$th-order cyclic correlation, defined by:

$$c^{\alpha}_{x,p+q,p}(\underline{\tau}) = \lim_{T \to +\infty} \frac{1}{T} \int_T C_{x,p+q,p}(t;\underline{\tau}) \exp(-2i\pi\alpha t) dt \qquad (4)$$

is non-zero.

Let us precise the expressions at order four ($p+q=4$). We will consider the definition in which there are as many conjugated as non-conjugated terms ($p=q=2$). Applying (4) to the process (1), it can be readily shown that the modulus of the cyclic tricorrelation (also called fourth-order temporal cumulant [1],[8]) is given by:

$$\begin{aligned} \left|c^{\alpha}_{x,4,2}(\underline{\tau})\right| &= \left|\frac{C_{s,4,2}}{T_b} \int q(t)\, q(t+\tau_1)\, q(t-\tau_2) q(t-\tau_3) \right. \\ &\quad \left. \cdot \exp(-2i\pi\alpha t) dt\right| \end{aligned} \qquad (5)$$

where $C_{s,4,2}$ is the stationary fourth-order cumulant of the random sequence $\{s_k\}$: $C_{s,4,2} = Cum[s_k, s_k, s_k^*, s_k^*]$. It is necessary to consider the modulus, in order to avoid terms depending on t_0 and f_c, which are *a priori* both unknown.

Similarly, at order two, the modulus of the cyclic correlation of the process (1) is:

$$\left|c^{\alpha}_{x,2,1}(\tau)\right| = \left|\frac{C_{s,2,1}}{T_b} \int q(t)\, q(t-\tau) \exp(-2i\pi\alpha t) dt\right| \qquad (6)$$

3. Classification strategy

3.1. The basic idea

It has been shown in [5] that the $c^{\alpha}_{x,p+q,p}(\underline{\tau})$ are useless in a modulation classification task for QAM signals, because they lead to proportional functions whatever p, q, and α. Besides, it has been shown in [3] that the tricorrelations of the stationarized signal are non proportional functions of $\underline{\tau}$, and so make classification possible even if the power of the signal is unknown. Moreover, we pointed out in [5] that cumulant-based

stationary and cyclic tricorrelations correspond to two different arrangements of the same fourth– and second-order statistics. Since it is clear that one of these arrangements makes classification possible, whereas the other does not, it seems natural to search for the combination of fourth- and second-order statistics which is optimal with a view to modulation classification.

Following this principle, our discriminating feature will be:

$$f_x(\tau_1,\tau_2,\tau_3,\tau_4,\lambda) = \left|c^{\alpha}_{x,4,2}(\tau_1,\tau_2,\tau_3)\right| + \lambda \left|c^{\alpha}_{x,2,1}(\tau_4)\right|^2 \quad (7)$$

where the parameter λ should be determined in order to maximize the distance between D_x and D_y if x and y are different modulation types. As for determining the most appropriate cycle frequency, a natural choice is $\alpha = \frac{1}{T_b}$, since the effect of stationary noise is thus avoided, and since energy for higher harmonics of $\frac{1}{T_b}$ is less significative. In order to get a correlation receiver which does not need threshold computation, we normalize (7) and the new feature becomes:

$$F_x(\underline{\tau},\lambda) = \frac{f_x(\underline{\tau},\lambda)}{\sqrt{\int_{\underline{\tau}} f_x^2(\underline{\tau},\lambda)d\underline{\tau}}} \quad (8)$$

where $\underline{\tau} = (\tau_1,\tau_2,\tau_3,\tau_4)$.

3.2. Determination of the practical discriminating feature

Since (8) is a five dimensional pattern (with respect to $\underline{\tau}$ and λ), we propose a feature involving lower complexity by working only with a slice of $F_x(\underline{\tau},\lambda)$. We will keep only a line of the three dimensional tricorrelation $c^{\alpha}_{x,4,2}(\tau_1,\tau_2,\tau_3)$, and we impose that this line contains the origin. In other words, τ_1, τ_2 and τ_3 can be parametrated as:

$$\tau_1 = a\tau, \quad \tau_2 = b\tau, \quad \tau_3 = c\tau. \quad (9)$$

Besides, we impose that the correlation $c^{\alpha}_{x,2,1}(\tau_4)$ be a function of the running parameter τ defined in (9), ans so the resulting feature is given by:

$$G_x(a,b,c,\lambda,\tau) = \frac{g_x(a,b,c,\lambda,\tau)}{\sqrt{\int_\tau g_x^2(a,b,c,\lambda,\tau)d\tau}} \quad (10)$$

where

$$\begin{aligned} g_x(a,b,c,\lambda,\tau) &= f_x(a\tau,b\tau,c\tau,\tau,\lambda) \quad (11)\\ &= \left|c^{\alpha}_{x,4,2}(a\tau,b\tau,c\tau)\right| + \lambda\left|c^{\alpha}_{x,2,1}(\tau)\right|)^2 \end{aligned}$$

Now, the four parameters a, b, c and λ should ideally maximize the distance between G_x and G_y. Since G_x and G_y are normalized, a, b, c and λ should equivalently minimize the correlation coefficent:

$$\rho = \int_\tau G_x(a,b,c,\lambda,\tau)\cdot G_y(a,b,c,\lambda,\tau)d\tau \quad (12)$$

Consequently, given two types of modulation x and y, the first step is to use the theoretical expressions of the cyclic fourth- and second-order correlations (5) and (6) in order to compute (10) and then determine $(a,b,c,\lambda)_{opt}$ such as:

$$(a,b,c,\lambda)_{opt} = \arg\min(\rho) \quad (13)$$

To achieve this step, it is of course necessary to know *a priori* the symbol duration T_b and the pulse function $q(t)$.

3.3. Comparison between theoretical and estimated features

Following [3], the two theoretical expressions $G_x((a,b,c,\lambda)_{opt},\tau)$ and $G_y((a,b,c,\lambda)_{opt},\tau)$ are compared to $g_s((a,b,c,\lambda)_{opt},\tau)$ (which is the feature estimated on the received signal $s(t)$) by a matched filter technique as described in figure 1. The outputs of each of the two matched filters are compared, and the strongest correlation measure provides the recognized modulation.

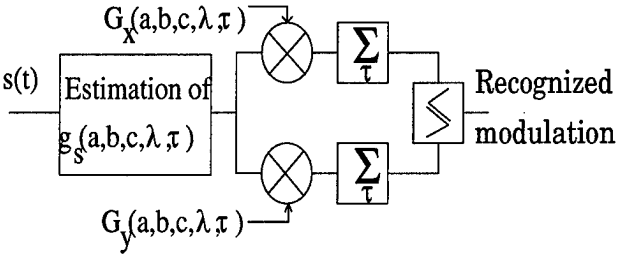


Figure 1. The x *vs.* y classification system

4. Application to 4-PSK *vs.* 16-QAM classification

4.1. Computation of optimal parameters.

We suppose now x =4-PSK and y =16-QAM. The correlation coefficient (12) is easily computable using (10), (5) et (6) with:

$$\text{4-PSK}\begin{cases} C_{s,4,2} = -1 \\ C_{s,2,1} = 1 \end{cases} \qquad \text{16-QAM}\begin{cases} C_{s,4,2} = -0.68 \\ C_{s,2,1} = 1 \end{cases} \quad (14)$$

We also suppose that the pulse function of the modulations is given by: $q(t) = 1$ if $t \in [0, T_b[$ and $q(t) = 0$ elsewhere.

The minimization of (12) is then performed thanks to the SIMPLEX algorithm, which leads to the following optimal parameters:

$$(a, b, c, \lambda)_{opt} = (0.04, 0.99, 0.84, -2.87) \quad (15)$$

and the corresponding correlation coefficient is given by:

$$\rho_{\min} = -0.06. \quad (16)$$

The optimal values for a, b and c given in (15) are not suitable for discrete-time series, since estimation can be performed only for integer lags. Consequently, we are compelled to use sub-optimal parameters, which are the interger values of a, b and c nearest to the optimal ones, *i.e.* $(a, b, c) = (0, 1, 1)$. Once these parameters are fixed, minimization of (12) is performed again with respect to λ (*cf.* figure 2), which leads to:

$$\lambda_{opt} = -2.94 \text{ and } \rho_{\min} = 0.038. \quad (17)$$

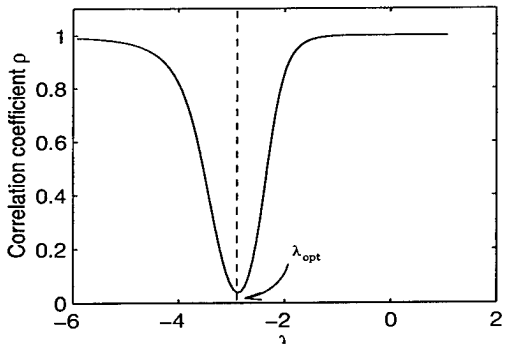


Figure 2. Correlation coefficient

The corresponding discriminating features (10) are shown on figure 3.

4.2. Simulations and discussion

Simulations have been performed on synthetic data in white gaussian noise for different signal to noise ratios S/N (S/N=0 , 5 dB). The number of transmitted symbols N_s varies from 50 to 5000 symbols, with $T_b = 10$ (*i.e.* 500 to 5000 samples). For each couple (S/N, N_s), 500 different signals (different symbol sequences and noise samples) are generated for each modulation. The figures (4-5) give the performance (probability of correct classification (P_{cc}) in %) obtained for different N_s and for a given S/N.

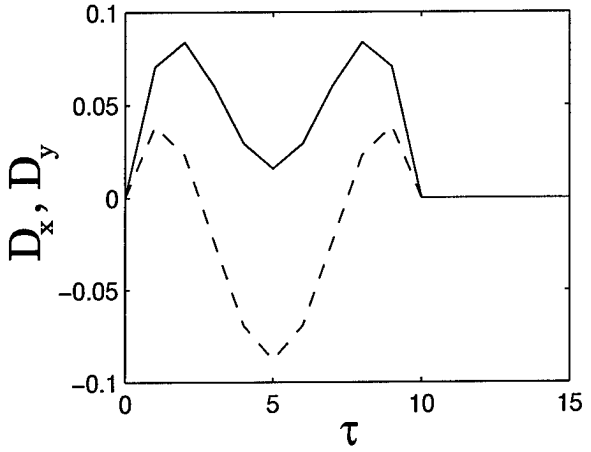


Figure 3. Discriminating features (–): 4-PSK, (- -): 16-QAM

On the same figures, we reported the performances of a classifier build on exactly the same principles discussed in section 3, but where the discriminating features (10) are based on stationary fourth- and second-order moment-based correlation, i.e.:

$$g_{\widetilde{x}}(a, b, c, \lambda, \tau) = m_{\widetilde{x},4,2}(a\tau, b\tau, c\tau) + \lambda \left(m_{\widetilde{x},2,1}(\tau) \right)^2 \quad (18)$$

where $\widetilde{x}(t)$ is a stationarized version of the cyclostationary process $x(t)$. Details about this classifier can be found in [6], but it should be stressed that the minimum correlation coefficient is:

$$\rho_{\min} = 0.54 \quad (19)$$

This correlation coefficient is much higher than the one exhibited in (17), which means that the patterns extracted in the cyclic domain are theoretically more discriminating than the ones obtained with a stationary model. However, figures (4) and (5) show that simulated performances are better if a stationary model is adopted. This can be explained by the following reasons. The performance of the classifier depends not only on the distance between the discriminating features, but also on the variance of the estimates of these features. It is well known that the variance of the estimates is higher in the cyclic domain than in the stationary one, and taking the modulus of the cumulants in (7) renders the cyclic classifier more sensitive to noise (for both variance estimation and thermic noise).

One can remark that the previous performance are immuned to white noise because $\tau = 0$ is avoided in

the stationary case. But in case of coulored noise, one can guess that the cyclic approach will perform better.

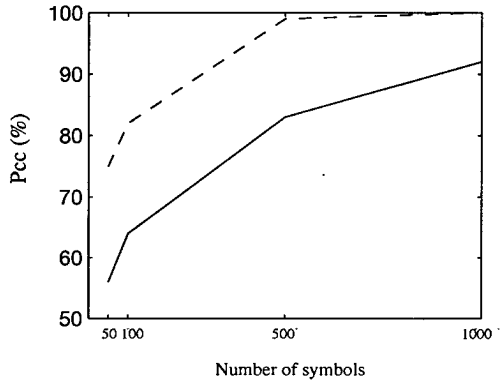


Figure 4. Performance of the classifier, S/N=0dB, (–): cyclic model (- -): stationary model

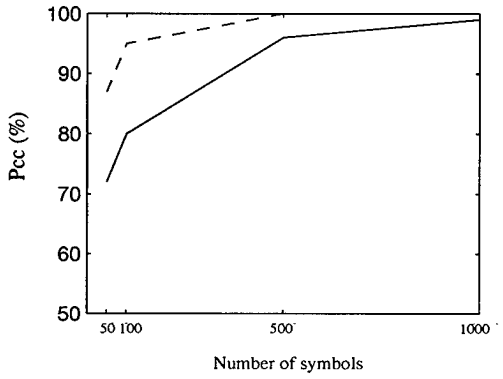


Figure 5. Performance of the classifier, S/N=5dB, (–): cyclic model (- -): stationary model

5. Conclusion

We have presented a new principle of modulation classifier, which is derived from maximum likelihood techniques. It is new relatively to the kind of the feature that is processed to achieve the classification. This feature is an *ad hoc* function which has to be built for each class we want to discriminate. Properly mixing fourth- and second- order statistics can result in features that are more or less discriminating, depending on the model adopted for the stochastic process (stationary or cyclostationary). But it has been proved that the accuracy of the classification is also highly conditioned by the variance of the estimated features.

References

[1] C.M. SPOONER, W.A. GARDNER, "The cumulant theory of cyclostationary time-series, Part I: Foundation; Part II: Development and applications," *IEEE trans. on sig. proc.*, vol. 42, No. 12, Dec. 1994.

[2] C.Y. HUANG, A. POLYDOROS, "Likelihood methods for MPSK modulation classification," IEEE Trans. on Com., vol. 43, No. 2/3/4, Feb./Mar./Apr. 1995.

[3] D. BOITEAU, C. LE MARTRET, "Classification of linear modulations by mean of a fourth order cumulant," EUSIPCO '96, Trieste, Sept. 1996.

[4] P.O. AMBLARD, M. GAETA AND J.L LACOUME, "Statistics for complex variables and signals - Part II: Signals," Signal Processing, vol. 53, No 1, August 1996

[5] P. MARCHAND AND D. BOITEAU, "Higher-order statistics for QAM signals: a comparison between cyclic and stationary representations," EUSIPCO '96, Trieste, Sept. 96.

[6] C. LE MARTRET, D. BOITEAU, "Modulation classification by means of different orders statistical moments" submitted to MILCOM 5-7 Nov 1997, Monterey, California.

[7] C.M. SPOONER, "An overview of recent developments in cyclosationary signal processing," *Proc. of the second workshop on cyclostationary signals*, Monterey, CA, August 1-2, 1994.

[8] A.V. DANDAWATÉ AND G.B. GIANNAKIS, "Statistical tests for presence of cyclostationarity," IEEE Trans. on Sig. Proc., vol. 42, No. 9, Sept. 1994.

Tracking Nonlinear Systems Using Higher Order Moments

K. J. Turner, F. A. Faruqi and C. L. Brown
Signal Processing Research Centre
Queensland University of Technology
GPO Box 2434, Brisbane, Q. 4001, Australia
k.turner@qut.edu.au

Abstract

A sub-optimal nonlinear time-recursive filter is developed which considers an arbitrary number of moments of the conditional density. The filter assumes a quadratic truncation of the system dynamics and measurement functions and retains N moments, thus requiring knowledge of up to $2N+2$ a priori moments and $N+1$ moments of the measurement noise process, which may be non-Gaussian. Prediction and update relations are given for moments of arbitrary order, along with mechanisms which facilitate their closed forms. Numerical examples are given for both scalar and vector systems and show promising results.

1 Introduction

It is seen, in tracking problems and indeed in all stochastic frameworks, that nonlinearity inherently breeds non-Gaussianity and subsequently non-negligible higher order statistics for the conditional density of the state estimate. For this discussion, a generalised sampled continuous model for nonlinear stochastic filtering is considered. Specifically, nonlinear system dynamics and measurement functions are considered along with non-Gaussian disturbances. The state dynamics and measurement functions may be described thus,

$$\frac{dx}{dt} = f(x) + g\,w \quad (1)$$

$$y_k = h(x_k) + v_k \quad (2)$$

where g is a scalar constant and w and v_k are the process and measurement disturbances respectively. Using knowledge of $f(x)$ and $h(x_k)$, a general filter structure is constructed as in Figure 1. The estimate for the state is described by a probability density function $p_X(x_k|Y_k)$ conditioned on the measurement set (Y_k). This conditional density is parameterised or approximated such that relations for the prediction and correction of the state estimate may be formulated.

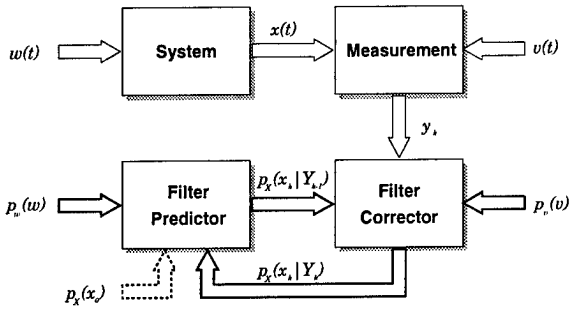


Figure 1. Filter Block Diagram.

The simplified case where the system and measurement equations are linear or linearised and the disturbances are Gaussian has previously been addressed by Kalman [7], and closed form expressions for the filter parameters have been derived. Unfortunately, the general solution to the recursive filtering problem for non-linear/non-Gaussian systems is itself ill-posed, with either parameterisations or computations of the posterior conditional probabilities approaching infinite orders.

What is required is a set of recursive closed-form relationships upon a fixed and finite number of parameters. Subsequently, certain assumptions are made in order to facilitate a tractable solution, at the cost of compromising the optimality of the estimates. These *sub-optimal* filters range widely in the fundamental assumptions and structures that are imposed [1, 3, 5, 6, 9, 10].

Generally, these methods suffer either from large parameter set requirements or considerable analytical complexity when considering vector systems. As a consequence, the literature regarding these more involved nonlinear filtering methods consider only simple examples with scalar systems and measurements [1, 5, 6]. A more recent and notable exception is [2] in which examples containing vector systems and scalar measurement functions are investigated.

It remains to devise a method by which the filter parameterisation and computational requirements may be kept at

0-8186-8005-9/97 $10.00 © 1997 IEEE

a reasonable level when applied to practical problems with state and measurement vectors of non-trivial dimension.

2 Higher order moment based filtering

The approach proposed by the authors embodies the extension of the truncated and Gaussian second-order filters as developed by Jazwinski [6] which considers the effect of higher order moments on the filtering solution.

For these second-order filters, a Gaussian assumption is assumed for the conditional density and noise densities, while the system dynamics and measurement functions are approximated by a Taylor series expansion which retains up to the quadratic terms. These quadratic nonlinearities introduce terms in the estimate and error covariance update relations which include higher-order moments of the conditional density and the noise processes.

Facilitation of closed form relations requires that certain *closure* methods be employed which expand or extrapolate the required higher-order moments in terms of the error covariance. Specifically, the truncated second-order filter assumes that the higher-order moments are negligible and may be set to zero, while assuming a Gaussian extrapolation yields the Gaussian second-order filter. Jazwinski notes that these filters frequently produce negative error variances due to the nature of the assumptions imposed.

Alternatively, consider the relaxation of the Gaussianity assumption through the inclusion of update relations for moments of up to some arbitrary order (N). The quadratic truncation of the dynamics and measurement functions is maintained, as is the assumption of Gaussianity for the process noise (w). Non-Gaussian measurement disturbances are inherently included in this filter structure, by considering the moments of the noise process (v_k).

The filter structure may be sectioned into 2 parts; the *prediction*, in which the evolution or diffusion of the moments between observations is considered, and the *correction* or measurement update. In order to maintain brevity, the formulation for the scalar systems will discussed herein, with the vector formulations being treated briefly [11].

3 Moment Prediction

The diffusion relations for the i th central moment (m_i) may be formulated by considering the derivative of polynomial expansion of its fundamental definition. Proceeding,

$$\frac{d}{dt}(m_i) = \frac{d}{dt}\mathsf{E}\left[(x - \mathsf{E}[x])^i\right] \tag{3}$$

which gives,

$$\frac{d}{dt}(m_i) = \frac{d}{dt}\mathsf{E}[x^i] + \sum_{j=1}^{i}(-1)^j c_{i,j}\mathsf{E}[x]^{(j-1)} \times\left[\frac{d}{dt}\left(\mathsf{E}[x^{i-j}]\right)\mathsf{E}[x] + j\mathsf{E}[x^{i-j}]\frac{d}{dt}\left(\mathsf{E}[x]\right)\right] \tag{4}$$

where $c_{i,j}$ is the coefficient for the j th term of an i th order polynomial. Evaluating the remaining quantities using [8, Eqn. (3.6)] and [6, Lemma 6.1] respectively gives,

$$\mathsf{E}[x^k] = \sum_{l=0}^{k} c_{k,l} m_{(k-l)}\mathsf{E}[x]^l, \tag{5}$$

and,

$$\frac{d}{dt}\left(\mathsf{E}[x^k]\right) = k\left\{\sum_{l=0}^{k-1} c_{k-1,l}\mathsf{E}[x]^l \times\left(f m_{(k-l-1)} + f_x m_{(k-l)} + \frac{1}{2} f_{xx} m_{(k-l+1)}\right)\right\} + \frac{1}{2}k(k-1)g^2 q\sum_{l=0}^{k-2} c_{k-2,l} m_{(k-l-2)}\mathsf{E}[x]^l, \tag{6}$$

where q is the variance of the process noise. Also f, f_x and f_{xx} are the dynamics function and its first and second partial derivatives respectively, evaluated at the current state estimate. From these relations it is evident that in order to predict N moments one requires knowledge of $(N+1)$ moments. Hence, closure methods are employed in order to calculate the $(N+1)$th moment.

4 Moment Correction

One may calculate the posterior or corrected moments (m_i^+) by considering the first order Power Series expansion about the measurement residual $(y - \mathsf{E}[y])$,

$$\mathsf{E}[x]^+ = a_1 + b_1(y - \mathsf{E}[y]) \tag{7}$$

$$m_i^+ = a_i + b_i(y - \mathsf{E}[y]) \tag{8}$$

where the coefficients are expressed as a polynomial expansion in terms of the predicted moments (m^-), the moments (r) of the measurement noise and the first and second partial derivatives of the measurement function.

Utilising [6, Eqn. 9.21], expressions for a_1 and b_1 are evaluated thus,

$$a_1 = \mathsf{E}[x] \tag{9}$$

$$b_1 = \mathsf{E}[(x - \mathsf{E}[x])(y - \mathsf{E}[y])] \times\mathsf{E}[(y - \mathsf{E}[y])^2]^{-1} \tag{10}$$

Also, a_i and b_i are expressed as,

$$a_i = \sum_{j=0}^{i}(-1)^j c_{i,j}(b_1)^j \times \mathsf{E}[(x-\mathsf{E}[x])^{(i-j)}(y-\mathsf{E}[y])^j] \quad (11)$$

$$b_i = \sum_{j=0}^{i}(-1)^j c_{i,j}(b_1)^j \times \mathsf{E}[(x-\mathsf{E}[x])^{(i-j)}(y-\mathsf{E}[y])^{(j+1)}] \times \mathsf{E}[(y-\mathsf{E}[y])^2]^{-1} \quad (12)$$

with all of the expectations conditioned on the set of previous measurements. The remaining expectations are of the form,

$$\mathsf{E}[(x-\mathsf{E}[x])^k(y-\mathsf{E}[y])^l] = \sum_{m=0}^{l} c_{l,m} h_x^{(l-m)} \sum_{n=0}^{m} c_{m,n}(\frac{1}{2}h_{xx})^n r_{(m-n)} \times \sum_{p=0}^{n}(-1)^p c_{n,p} m_{(k+l-m+2n-2p)}^{-}(m_2^{-})^p \quad (13)$$

From these relations it is seen that in order to compute the ith *posterior* moment one requires $(2i+2)$ *a priori* moments. This is by far a more serious problem than that for the case of moment prediction.

5 Closure Methods

It remains to derive a method by which higher-order moments may be generated in terms of lower-order moments. Consider, as an alternative to the truncated and Gaussian closure, a method which utilises a cumulant based expansion. Specifically, if N moments are considered by the filter, the moments of order greater than N are calculated by setting the cumulants of order $N+1$ and greater to zero. For example, consider the expansion of the 5th cumulant (κ_5) in terms of the central moments [8].

$$\kappa_5 = m_5 - 10m_3m_2$$

Setting $\kappa_5 = 0$ and solving yields,

$$m_5 = 10m_3m_2$$

This process is generalised such that the required higher-order moments may be calculated.

6 Vector Systems

For vector systems, a novel tensor notation is utilised such that the symmetric nature of the moments is considered [4],

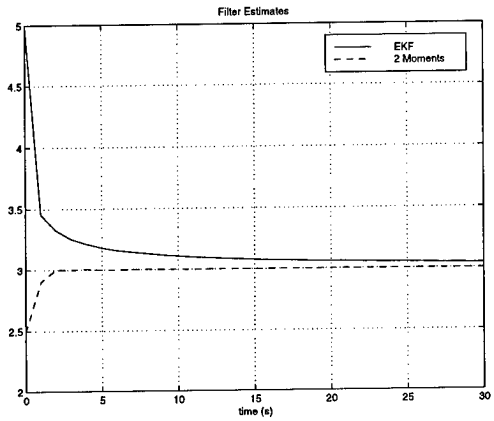


Figure 2. Example 1: A typical realisation.

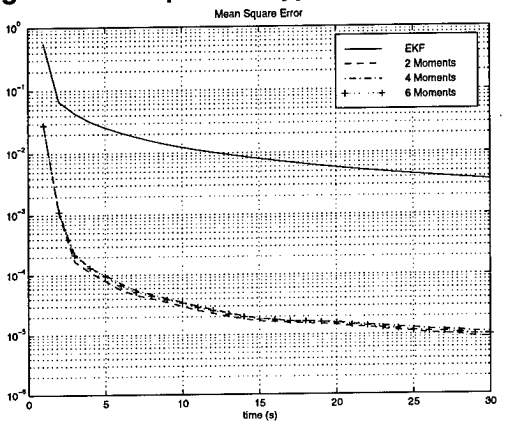


Figure 3. Example 1: MSE for 500 Monte Carlo simulations.

and the quantities are stored in a vector form. For example, consider the third moment of a 2-state system as,

$$\mu_3 = \left[\mathsf{E}[x_1^3]\ 3\mathsf{E}[x_1^2x_2]\ 3\mathsf{E}[x_1x_2^2]\ \mathsf{E}[x_2^3]\right]^T. \quad (14)$$

The algebra surrounding the notation in Equation (14) contains a number of useful properties. One property which is of particular value is that of commutativity, such that the polynomial expansions present in the derivation for the scalar system are also valid for the vector case. Additionally, all quantities manipulated by the filter are standard vectors and matrices, thus avoiding complicated tensor notations and issues.

7 Numerical Analysis

To illustrate the effectiveness of the higher order moment techniques three numerical examples are investigated.

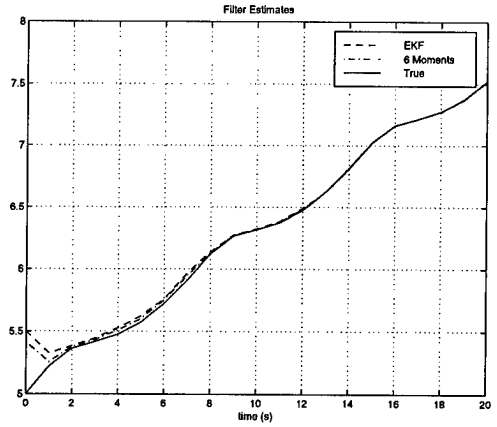


Figure 4. Example 2: Estimates for a typical realisation.

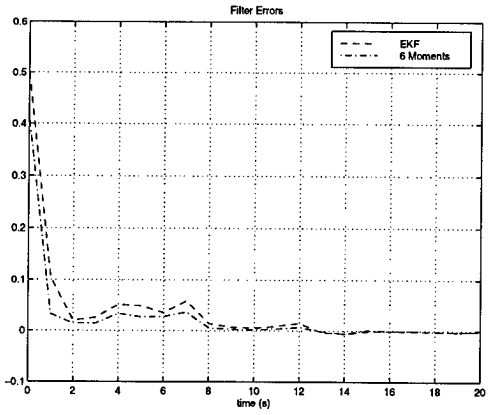


Figure 5. Example 2: Errors for a typical realisation.

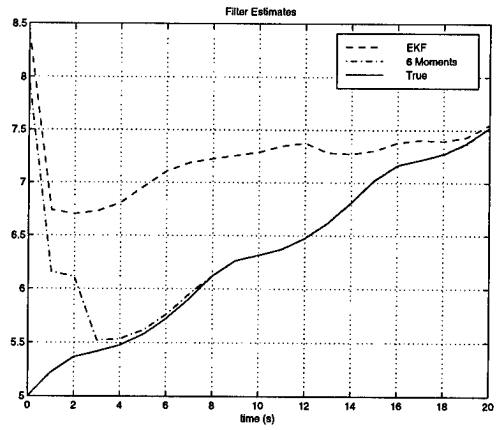


Figure 6. Example 2: Realisation with biased estimates.

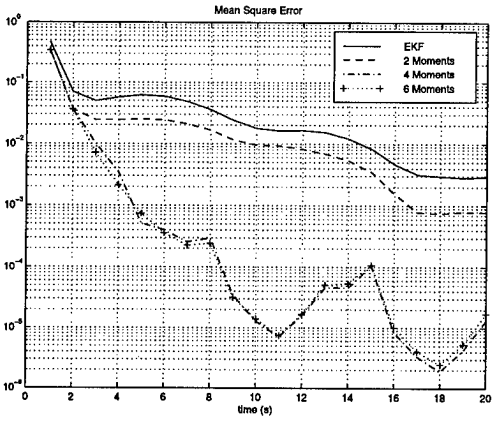


Figure 7. Example 2: MSE for 500 Monte Carlo simulations.

7.1 Example 1

Consider firstly a scalar example in which estimation of a constant state from noisy measurements of its squared value is investigated, such that,

$$\frac{d}{dt}(x) = 0 \qquad \text{and} \qquad y_k = x_k^2 + v_k. \tag{15}$$

An experiment is constructed for which the true state $x_0 = 3$ and the standard deviation of the measurement noise $\sigma_v = 0.1$. The initial filter estimate is chosen $\hat{x_0} \sim \mathcal{N}(3, 4)$ such that the nonlinearity effects are accentuated.

Figure 2 shows a typical realisation of the estimate for the proposed method, for $N = 2$ and the Extended Kalman Filter (EKF). The EKF clearly shows its characteristically biased convergence while a marked improvement is achieved for the proposed method. The improvement is also illustrated by the Monte-Carlo analysis results for 500 realisations in given Figure 3. Results for 4 and 6 moments offer little improvement over the 2 moment filter for this case.

7.2 Example 2

Consider a more interesting scalar system, in which the both the dynamics and the measurement functions contain higher than quadratic nonlinearities. Namely,

$$\frac{dx}{dt} = -\cos^3(ax) + b \tag{16}$$

$$y_k = x_k^3 + v_k \tag{17}$$

where $a = 7$ and $b = 1.5$. The initial estimate is $\sim \mathcal{N}(5, 1)$ and $\sigma_v = 2$. Figures 4 and 5 show the respective estimates and errors for a typical realisation ($N = 6$). While the improvement is marginal for this case, the bias effects as shown in Figure 6 are largely avoided. Monte Carlo analysis in Figure 7 shows that a suitable number of moments is 4,

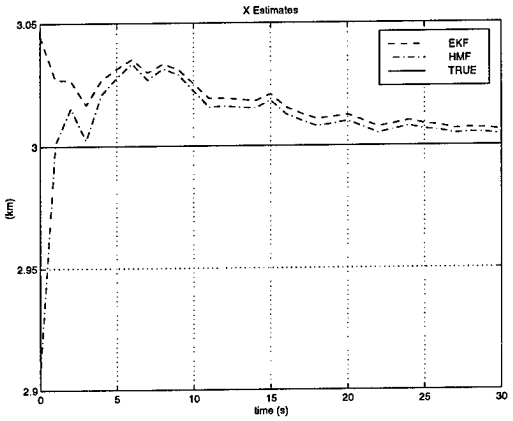


Figure 8. Example 3: Estimates for x.

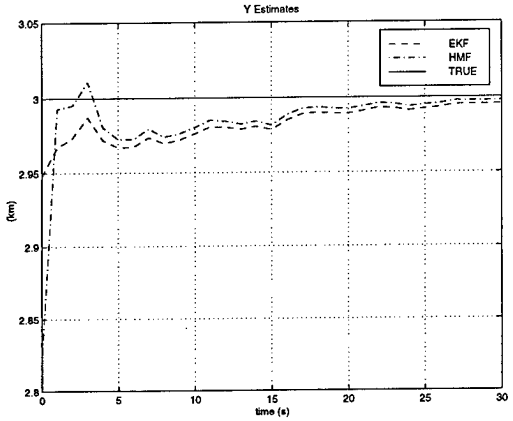


Figure 9. Example 3: Estimates for y.

as the 6 moment filter offers negligible improvement. The oscillations in the MSE for the 4 and 6 moment filters are attributable to the neglected higher-order nonlinearities in the system dynamics.

7.3 Example 3

Consider a simple 2 dimensional vector system with a constant state,

$$\boldsymbol{X} = [x \;\; y]^T \tag{18}$$

where range and bearing angle measurements available for filtering,

$$\boldsymbol{Y_k} = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \tan^{-1}(y_k/x_k) \end{bmatrix} + \begin{bmatrix} v_{1_k} \\ v_{2_k} \end{bmatrix} \tag{19}$$

with standard deviations for the range and bearing measurement noises being 30 metres and 0.015 radians respectively. The initial estimate is chosen such that $\boldsymbol{X}_0 \sim \mathcal{N}([3 \;\; 3]^T, \boldsymbol{I})$. Figures 8 and 9 show some preliminary results for a 4 moment filter for vector systems. These results show improved convergence over the EKF in the estimates for both states and motivates continuing research into vector higher-order filters.

8 Conclusion

A sub-optimal nonlinear time-recursive filter was developed which considers an arbitrary number of moments of the conditional density for the filter estimate. It was seen that under a quadratic truncation of the system dynamics and measurement functions that a filter which retains N moments requires knowledge of up to $2N + 2$ *a priori* moments and $N+1$ moments of the measurement noise process. Prediction and update relations were given for moments of arbitrary order, along with mechanisms which facilitate their closed forms. Numerical examples are given for both scalar and vector systems and show promising results.

References

[1] D. L. Alspach and H. W. Sorenson. Nonlinear bayesian estimation using gaussian sum approximations. *IEEE Transactions on Automatic Control*, AC-17(4), August 1972.

[2] R. L. Bellaire, E. W. Kamen, and S. M. Zabin. A new nonlinear iterated filter with applications to target tracking. *Proceedings of the SPIE*, 2561:240–251, 1995.

[3] F. E. Daum. Beyond kalman filters: practical design of nonlinear filters. *Proceedings of the SPIE*, 2561:252–262, 1996.

[4] F. A. Faruqi. Algebraic structure of quadratic dynamical systems. Technical report, Queensland University of Technology, Brisbane, Australia, 1997.

[5] T. W. Hilands and S. C. A. Thomopoulos. High-order filters for estimation in non-gaussian noise. *Information Sciences*, 80:149–179, 1994.

[6] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, Inc., 1970.

[7] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, pages 35–43, March 1960.

[8] M. G. Kendall and A. Stuart. *Advanced theory of statistics: distribution theory*, volume 1. Griffin Books, London, 3rd edition, 1969.

[9] C. J. Masreliez. Approximate non-Gaussian filtering with linear state and observation relations. *IEEE Trans. Automat. Contr.*, AC-20:107–110, 1975.

[10] S. Maybank. Finite-dimensional filters. *Philosophical Transactions of the Royal Society of London A*, 354:1099–1123, 1996.

[11] K. J. Turner and F. A. Faruqi. Higher-order moment based filtering for vector systems. Technical report, Queensland University of Technology, Brisbane, Australia, 1997.

Isolated Word Recognition Using High-Order Statistics and Time-Delay Neural Networks

M.R. Ashouri
Department of Electrical and Computer Engineering
Isfahan University of Technology
Isfahan 84154, Iran.
e-mail: ashouri@cc.iut.ac.ir

Abstract

In this paper, two isolated word recognition method based on high-order statistics and time-delay neural network (TDNN) for recognition of Farsi spoken digits have been studied. The adopted speech recognition system consists of four modules, namely, preprocessor, end-points' detector, feature extractor and classifier. The first method estimates the AR parameters of speech based on the third- and fourth-order cumulants using high-order Yule-Walker, W-slice and 1-D slice approaches. In the second method, statistical features are extracted from the estimated high-order probability density function (pdf)of theresholded amplitude features. For each pdf estimate, the values of mean, variance, third order moment and entropy are computed. The total number of features for each frame of approximate length of 15 ms is 16.. The adopted TDNN has 16 nodes in its input layer, 10 nodes in its output layer and two hidden layers. The learning rule of the adopted TDNN that is based on the back-propagation rule has been modified to decrease the training time. Computer simulation results obtained from recognizing 10 Farsi digits spoken by different speakers shows that the first method has a better recognition rate while the second method necessitate less computation

1. Introduction

Speech is a form of measures that human beings use to communicate intention and emotion. Therefore speech signal processing techniques including automatic speech recognition has attracted attention as practical application areas widen. As classical methods, Dynamic TimeWarping (DTW) and Hidden Markov Model (HMM) have been used for speech recognition. In recent years, automatic speech recognition using Artificial Neural Networks (ANN) have been studied by many researchers. A number of recognition systems have employed original or modified version of the Time Delay Neural Network (TDNN) as defined by Waibel, and some of these methods outperform the conventional HMM approach [1], [2].

Classical speech analysis techniques for feature extraction are based on second-order statistics and their performance dramatically decreases when noise is present in the signal under analysis.

In the last few years there has been an increasing interest in the application of High-Order Statistics (HOS) in speech recognition. Paliwal [3] made use of HOS in a recognition system and he showed that results remain constant under a great variability of SNR. Results show that speech signals can be characterized not only by its autocorrelation but also by its third- and fourth-order cumulants. Cumulants of order greater than two are zero for white and colored Gaussian noise. Analysis of noisy speech signals based on HOS permits separation of speech from noise.

In this paper two HOS based methods for Farsi isolated digit recognition has been studied. The first method adopts the estimation of AR parameters using HOS techniques. The second method employs features based on the estimates of high order pdf.

2. AR Parameters Estimation Using HOS

From the high-order Yule-Walker algorithm [5] a set of following equations is obtained;

$$\sum_{l=0}^{p} a(l) C_{k,y}(m-l, k_0, 0, ..0) = 0 \qquad (1)$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

To solve these equations, it is necessary to concatenate p+1 slices, k_0=-p.....,0 in (1) because a single slice does not guarantee a full rank system of equations. To improve the stability, two modification has been proposed: to increase the number of slices or to increase the number of equations per slice [4].

The w-slice algorithm [6] is based on the following weighted sum of cumulant slices;

$$C_w(i) = w_2 C_{2,y}(i) + \sum_{j=-L}^{N} w_3(j) C_{3,y}(i,j) + \sum_{j=-1}^{N}\sum_{k=-L}^{N} w_4(j,k) C_{4,y}(i,j,k) + \ldots \quad (2)$$

and it is developed in three steps:
a). choose w_2, w_3 (j)and w_4(i,j) so that:

$$C_w(i) = 0, \quad i=-P,\ldots,-1 \qquad C_w(0)=1 \quad (3)$$

where $P \geq p$, $N \geq 0$ and $L \geq p+M$, where M is the over determination. b). Estimate the first P term of the impulse response from the weighted cumulant $C_w(i)$.

$$h(i) = C_w(i) \quad i=1,\ldots P \quad (4)$$

c) Solve the filter coefficients from the following equation:

$$\sum_{l=0}^{p} a(l) h(i-l) = 0 \quad i=1\ldots P \quad (5)$$

In the 1-D slice algorithm, the AR parameters are obtained from the single cumulant slice $C_{k,y}(m,k_0,\ldots,0)$. An one-dimensional slice does not guarantee a full rank system of equations and for this reason is not reasonable to solve (1) directly since the solution may not be stable and unique. Instead the autocorrelation of the cumulant slices considered [4]. If we multiply (1) by the one-dimensional slice $C_{k,y}(m-l',k_0,0\ldots0)$ and we sum in an interval with m>0, we obtain:

$$\sum_{l=0}^{p} a(l) \sum_{m>0} C_{k,y}(m-l,k_0,0..0) C_{k,y}(m-l',k_0,0..0) = 0 \quad (6)$$

The above equation can be expressed as:

$$\sum_{l=0}^{p} a(l) f_c(l,l',k_0,0\ldots0) = 0 \quad (7)$$

where $f_c(l,l',k_0,0\ldots0) =$

$$\sum_{m>0} C_{k,y}(m-l,k_0,0\ldots0) C_{k,y}(m-l',k_0,0\ldots0) \quad (8)$$

Instead of $f_c(.)$ we approximately use $R_c(.)$ that is the autocorrelation of the causal part of the one slice cumulant which yields;

$$\sum_{l=0}^{p} a(l) R_c(l-l') = 0 \quad l'=1.....p \quad (9)$$

The coefficients obtained from (9) differ from the true AR coefficients but they are very useful for speech recognition tasks. The autocorrelation method that assures a stable solution also has been considered;

$$\sum_{k=0}^{p} a(k) R_{hh}(k-l) = 0 \quad k=1\ldots p \quad (10)$$

3. HOS Based Feature Estimation using PDF

Let X be a feature vector and let x_1, x_2,... x_d be the values that specific features take. If each element x_i of the vector X is quantized and assumes a finite number of different values. For example if the feature used is an amplitude feature with 8 bit resolution (256 levels), then each x_i can be represented as a vector itself as follows;

$$x_i = (x_{i1}, x_{i2}, \ldots\ldots, x_{it}) \quad (11)$$

As X is a multidimensional vector, there are several means, variances and moments (first-, second- and third order moments). For example we can define several means as follow;

$$\mu_{x_i} = \sum_{j=1}^{j=t} x_{ij} \{ \sum_{1k=1}^{1k=t} \sum_{2l=1}^{2l=t} \ldots \sum_{dn=1}^{dn=t} p(x_{1k},\ldots,x_{dn}) \} \quad (12)$$

One problem with the implementation of this model is the high-dimensionality of the estimate which require an

impractical amount of memory[8]. To solve this problem, we build several binary feature vector from our feature vector. In our case we generate four binary vector using four equally spaced theresholds.

Let f(x) be the amplitude of sample x and let T be the value of the threshold. The binary vectors g(x) are generated by applying the condition;

$$g(x) = \begin{matrix} 1, if..f(x) \geq T \\ 0, if..f(x) < T \end{matrix} \qquad (13)$$

Each of these binary vector form a d digit binary number from which we can compute its histogram. From each histogram the mean, variance, third order moment and entropy are computed.

The values calculated are used as descriptors or features of a specific word class. For classification we used neural networks (we used MLP and TDNN). The neural networks are trained with these descriptors. From the histogram the mean, variance, third order moment and entropy are computed.

4. Implementation

To evaluate the performance of the two methods described earlier, a database composed of utterances of 10 speaker was generated. Each speaker uttered 10 repetition of each Farsi digit. Five utterance of all the speaker are used for training and five for testing. Noisy signals are obtained adding white Gaussian noise at SNR of 0, 10, 20 dB.

For the evaluation of the first method in recognition of Farsi spoken digits, the autocorrelation algorithm (or second order Yule-Walker method YW2), the third- and fourth-order Yule-Walker algorithm (YW3 and YW4), the W-slice algorithm(WS3 and WS4) and the 1-D slice algorithm(DS3 and DS4)were examined. 16 features that are cepstrum, delta cepstrum and delta energy are used in the recognition system.

The TDNN architecture consists of 2 hidden layer plus input and output layers. The input layer has 16 nodes with 4 delays that makes the total number of input nodes to be 16*5. The first hidden layer has 10 nodes with 6 delays. The second hidden layer has 10 nodes with 10 delays and the output has 10 nodes with no delay. The configuration of the adopted TDNN is shown in Fig. 1.

The speed of the training of the TDNN has been significantly increased by modification of its learning algorithm. This is achieved by adding a linear function to the sigmoid transfer function of neurons. This prevents the derivative of the sigmoid transfer function become zero. Otherwise in situations when the derivative of the transfer function is zero or very small, the weights will not be updated and the learning time will be increased.

By considering the sigmoid transfer function:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (14)$$

When a linear function is added to the sigmoid function such as :

$$f_1(x) = f(x) + x \qquad (15)$$

$$f_1'(x) = f'(x) + 1 \qquad (16)$$

Thus in this case the derivative of the transfer function will always be nonzero.

The recognition rate of the aforementioned methods are summarized in table 1.

SNR	clean	20 dB	10 dB	0 dB
YW2	97.5	96.2	79.3	33.4
YW3	97.3	95.9	73.5	23.4
YW4	96.3	96.1	75.4	23.6
WS3	96.2	94.3	80.7	44.9
WS4	97.2	95.1	83.6	48.8
DS3	95.9	93.3	87.6	57.8
DS4	96.1	94.4	89.9	63.1

Table 1. Summary of recognition rate of Farsi digits recognition(method 1)

In the second method which exploits high-order pdf based features, each Farsi digit signal , after end-points' detection is divided into equal ,non overlapping frames. The total number of frames for each digit chosen to be 100. We pad enough zeroes at the end of each word so that the number of samples in each frame is an integer.

In the next step we select d equally spaced samples from each frame. In the preliminary investigation d selected to be 10. Although amplitude feature were adopted, other types such as LPC features could be used. For the frames that their length are not multiples of 10, interpolation technique used to evaluate the values of the appropriate amplitude features. Each obtained feature vector is converted to 4 binary vectors using 4 equally spaced threshold levels. Each binary vector is in fact a 10 binary digit number ranging from 0 to 1024. The histogram of each vector is then calculated and from that the first-, second- and third order moments and entropy are computed.

The moments are computed from the following equations:

$$\mu_n(x) = \sum_{i=1}^{L} (x_i - m)^n p(z_i) \qquad (17)$$

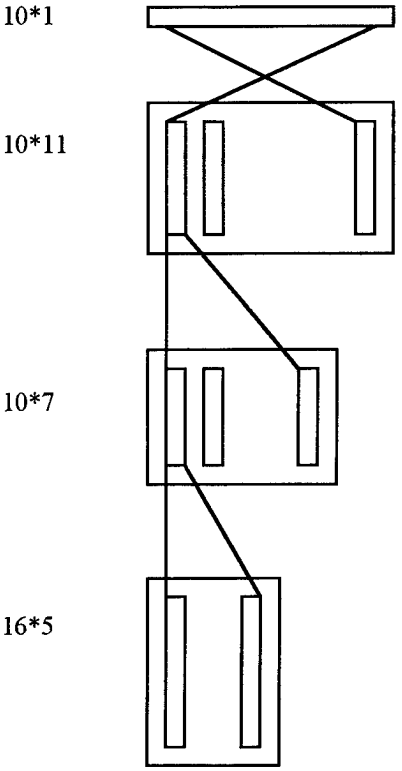


Fig.1. Configuration of the TDNN

Where m is the mean value of z (first-order moment)

$$m = \sum_{i=1}^{L} z_i p(z_i) \qquad (18)$$

The average information or entropy is:

$$H(z) = -\sum_{i=1}^{L} P(a_i) \log P(a_i) \qquad (19)$$

The total number of features obtained for each frame is 16. Before application of these feature vectors to the TDNN, they were compressed by a logarithmic function to limit their dynamic range. The employed speech database and the TDNN are the same as used for the first method.The recognition rate of the second method is not as good as the rate of the first method, and is about ten percent lower. By considering the fact that the amount of computation involved in the second method is much lower, this results is very encouraging. Some modification is going to be implemented to improve the results of the second method.

5. Conclusion

In this paper two methods based on high-order statistics for recognition of Farsi spoken digits have been studied. The first method estimates the AR parameters of speech based on third- and fourth-order cumulants using Yule-Walker, W-slice and 1-D slice methods and compare them with the autocorrelation method. The second method is based on high-order pdf and the first-, second-, third-order moments and entropy. The first method showed better recognition rate while the second method needs much less computation. Farsi spoken digits has more similarity to each other than their English counterparts, thus their recognition rate is lower.

6. Acknowledgment

The author wishes to acknowledge the Research Department of Isfahan University of Technology for supporting the project "Application of Neural Networks in Farsi Speech Recognition".

7. References

[1] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano and K. Lang, "Phoneme Recognition Using Time-Delay Neural Networks", IEEE Trans. on ASSP, vol.37, no.3, Mar 89

[2] H. Hild and A. Waibel,"Multi-speaker/speaker independent Architectures For The Multi-State Time-Delay Neural Networks", Proc. ICASSP-93, pp.255-258.

[3] K.K. Paliwal and M.M. Sondhi, "Recognition of Speech Using Cumulant-Based Linear Prediction Analysis", ICASSP-91, pp. 429-432

[4] A. Moreno, S. Tortola, J. Vidal and J.A.R .Fonollosa, "New HOS-Based Parameter Estimation Methods for Speech Recognition in Noisy Environments" , ICASSP95 pp. 429-432.

[5] A. Swami, J.M. Mendel, "AR Identifiability Using Cumulant Slices", Proc. of the Workshop in Higher Order Spectr al Analysis, pp.13-18. June 1989

[6] J. Vidal and J.A.R. Fonollosa, "Causal AR Modeling Using Linear Combination of Cmulant Slices",Elsevier Science: Signal Processing 36, Aug. 1992.

[7] J.M. Mendel,"Tutorial on Higher Order Statistics(Spectra) in Signal Processing and System Theory: Theoretical Resu lts and Some Applications", Proc.IEEE, vol.97, Mar. 1991

[8] H. Araujo, J. Dias and A.T. Almedia," Parallel Texture A alysis Using High-Order Statistics and Neural Networks. Pro.1994 Int.Symp. on Speech, Image Pro. and Neural Ne 13-16 April 1994, Hong Kong, pp.401-404

MAL: Applications of HOS

Chair: Gaetano Scarano
University of Rome, Italy

Impact of Blind versus Non-Blind Channel Estimation on the BER Performance of GSM Receivers

Dieter Boss, Thorsten Petermann and Karl-Dirk Kammeyer
University of Bremen, FB-1, Department of Telecommunications
P.O. Box 33 04 40, D-28334 Bremen, Germany, Fax: +(49)-421/218-3341, e-mail: boss@comm.uni-bremen.de

Abstract

We investigate in this paper whether the HOS-based blind channel estimation method EVI (EigenVector approach to blind Identification) can compete with the non-blind cross-correlation-based scheme used in state-of-the-art GSM receivers (Global System for Mobile comm.). For blind, non-blind, and ideal estimates of COST-207[2] mobile radio channels, we give simulated bit error rates (BER) after Viterbi detection in terms of the mean signal-to-noise ratio ($\overline{\text{SNR}}$). Averaged over three COST-207 propagation environments, EVI leads to an $\overline{\text{SNR}}$ loss of 1.2dB only, while it saves the 22% overhead in GSM data rate due to the transmission of training sequences. Since just 142 samples are used for channel estimation, we consider this performance outstanding for an approach based on HOS.

1. Linear time-variant GSM channel model

CONSIDER[1] the equivalent baseband representation of a GSM communication system in Figure 1, where source and channel coding are omitted to enhance clarity. In the transmitter, coded information bits $d(k)$ together with reference bits $f(k)$ are assembled into *bursts* of 142 bits, where a value from $\{-1, 1\}$ is taken each bit period $T = 48/13\ \mu\text{s} \approx 3.7\ \mu\text{s}$. Each burst is encoded differentially to facilitate demodulation. Then, it is modulated by *Gaussian Minimum Shift Keying (GMSK)* with $f_{3\,\text{dB}} = 0.3/T$ and transmitted over the multipath radio channel.

In a mobile scenario, the physical multipath radio channel is *time-variant* with a baseband impulse response depending on the time difference τ between the observation and excitation instants as well as the observation time t. We adopt the stochastic *Gaussian Stationary Uncorrelated Scattering (GSUS)* model leading to the following impulse response of the channel including the receive filter [1]

$$h_c(\tau, t) = \frac{1}{\sqrt{N_e}} \sum_{\nu=1}^{N_e} e^{j(2\pi f_{d,\nu} t + \Theta_\nu)} \cdot g_{Rc}(\tau - \tau_\nu), \quad (1)$$

where N_e is the no. of elementary echo paths, $g_{Rc}(\tau)$ denotes the receive filter impulse response, and the subscript in $h_c(\cdot)$ suggests its continuous-time property. 3D sample impulse responses can easily be determined from (1) by independently drawing N_e Doppler frequencies $f_{d,\nu}$, N_e initial phases Θ_ν, and N_e echo delay times τ_ν from random variables with Jakes, uniform, and piecewise exponential probability density functions, respectively. As for the echo delay times τ_ν, we use standard COST-207[2] *Typical Urban (TU)*, *Bad Urban (BU)* and *Hilly Terrain (HT)* profiles.

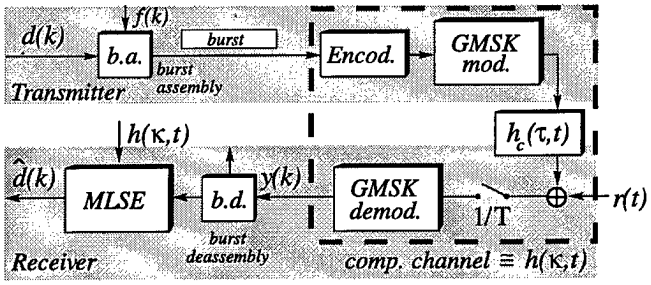


Figure 1: GSM communication system

According to Figure 1, each received burst is corrupted by additive Gaussian noise $r(t)$ which is colored by the 5th order Butterworth receive filter $g_{Rc}(\tau)$ with cut-off frequency at 75 kHz. We use this filter throughout the paper for reasons of adjacent channel suppression (GSM carrier spacing is 200 kHz). Upon symbol-rate sampling, a simple derotation demodulator can be used to obtain $y(k)$.

Maximum Likelihood Sequence Estimation (MLSE) represents the optimum procedure to remove intersymbol interference from a received digital communication signal such as $y(k)$. However, it assumes the "composite channel",

[1]This research is supported by the German NSF (DFG). *Matlab* programs and compressed postscript files of our publications are also available from our WWW server `http://www.comm.uni-bremen.de`.

[2]European Cooperation in the field of Scientific and Technical research.

0-8186-8005-9/97 $10.00 © 1997 IEEE

i.e. the equivalent symbol-rate system between $d(k)$ and $y(k)$ (see the dashed frame in Figure 1), to be *(R1)* linear with finite impulse response which *(R2)* must be known. Moreover, *(R3)* the noise component of $y(k)$ is supposed to be white. Strictly speaking, none of the requirements *(R1)* to *(R3)* is met in typical GSM systems.

However, as for *(R1)*, it can be shown that the composite channel can be approximated by the *linear* model

$$h(\kappa,t) = \begin{cases} j^{-\kappa} \cdot g_0(\kappa T, t) & \text{for } \kappa \in \mathcal{K} \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

with

$$g_0(\tau,t) \triangleq c_0(\tau) * h_c(\tau,t)\,, \qquad (3)$$

where $c_0(\tau)$ represents the Laurent approximation [2] of the (non-linear) GMSK modulator, "$*$" denotes the convolution operator and the factor $j^{-\kappa}$ takes the combined effects of differential encoding and derotation demodulation into account. In order to obtain a (time-variant) FIR model, let $h(\kappa,t)$ be limited to the range $\mathcal{K}$ of "relevant" indices κ.

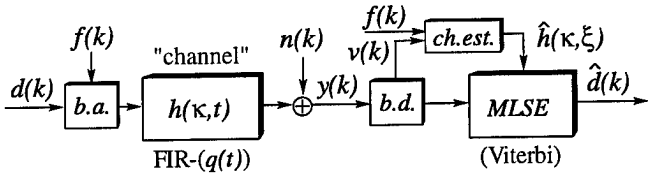


Figure 2: GSM system with linear time-variant channel model

Using the linear model (2) for the composite channel, Figure 1 can be redrawn according to Figure 2. In the sequel, $h(\kappa,t)$ will simply be termed "channel". Note that for any fixed value t_0, the channel slice $h(\kappa,t_0)$ may be mixed-phase. Let $q(t_0)$ denote its effective order, which may vary from slice to slice due to time selective fading. The noise sequence $n(k)$ emerges from $r(t)$ in Figure 1 by symbol-rate sampling and demodulation. Although $n(k)$ is colored due to the Butterworth receive filter, this is tolerated in typical GSM receivers (c.f. requirement *(R3)*).

According to *(R2)*, MLSE requires the knowledge of $h(\kappa,t)$ in order to equalize a block of the demodulated sequence $y(k)$. As this knowledge is not available, the problem of *channel estimation* arises. This is indicated in Fig. 2 by a box termed "ch.est.". Let $\hat{h}(\kappa,\xi)$ denote the estimate which will be used to equalize the ξ-th burst.

In the following section, we focus on two algorithms to calculate estimates $\hat{h}(\kappa,\xi)$. They will be compared in section 3 in terms of their channel estimation quality as well as the resulting bit error performance after MLSE.

2. Blind and non-blind channel estimation

Channel estimation is a particular form of *system identification*. All methods we apply in this paper suppose the system to be mixed-phase, linear and to have a finite impulse response. With time-variance being relatively slow in both GSM-900 and DCS-1800 applications ($f_{d,max}^{-1} = (226\,\text{Hz})^{-1} \approx 4.4\,\text{ms} \gg T$), the channel can also be assumed *piecewise (quasi) time-invariant*, i.e. time-invariant over a certain number Δk of bit periods T

$$h(\kappa,t) \approx h(\kappa,t_0) \quad \text{for} \quad |t-t_0| \le \Delta k T/2\,. \qquad (4)$$

If at most Δk consecutive samples of $y(k)$ are observed at once, $y(k)$ can be considered quasi stationary so that the respective channel estimation algorithm may assume a mixed-phase linear time-invariant FIR system $h(\kappa,t_0)$. Generally, a system can be identified with or without reference data (training sequences).

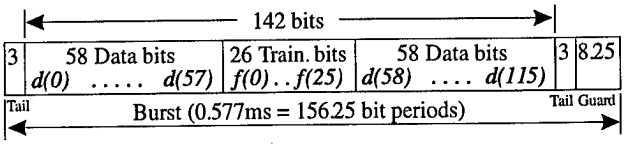


Figure 3: GSM "normal" burst

As an example for a channel estimation approach based on training sequences, consider the cross-correlation (CC) scheme used in state-of-the-art GSM receivers. According to Figure 3, each "normal" burst contains a training sequence $f(k)$ of 26 bits surrounded by two packets of 58 data bits emerging from the information sequence $d(k)$. On the assumption of time-invariance over $\Delta k = 21$ bit periods, channel estimates $\hat{h}(\kappa,\xi)$ can be derived from the received sequence sampled at symbol-rate by using the sample cross-correlation between the demodulated (corrupted) and the stored (ideal) training sequences $v(k)$ and $f(k)$, respectively (see Fig. 2). For any given order $\hat{q}$ of the FIR system to be estimated, the estimate is given by

$$\begin{bmatrix} \hat{h}(0,\xi) \\ \vdots \\ \hat{h}(\hat{q},\xi) \end{bmatrix} = \mathbf{F} \cdot \begin{bmatrix} v(10-\hat{q}) \\ \vdots \\ v(25) \end{bmatrix}, \qquad (5)$$

where $\mathbf{F}$ denotes the $(\hat{q}+1) \times (\hat{q}+16)$ Toeplitz matrix containing the orthogonal part of the training sequence

$$\mathbf{F} \triangleq \begin{bmatrix} f(5) & \cdots & f(20) & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & f(5) & \cdots & f(20) \end{bmatrix}. \qquad (6)$$

N.B.: Although the above channel estimation scheme supposes time-*in*variance of the channel over a period of $\Delta k = 21$ training bits only, the resulting estimate is used for MLSE on the adjacent data fields. As the channel coefficients might already have changed in the data fields, there is an implicit assumption of *quasi time-invariance over one burst* ($\Delta k = 142$) in this concept.

The repeated transmission of training sequences leaves a GSM system with an overhead capacity of $26/116 = 22.4\%$. This capacity could be used for other purposes if the channel was estimated from the received signal only (*blind* system identification). As just one symbol-rate sampled received sequence is available in GSM mobile units, in the downlink, this can only be accomplished by exploiting HOS of the (quasi) stationary demodulated sequence $y(k)$.

Remark: If the sampling period was a fraction of T, or alternatively, the symbol-rate sampled signals received by several antennae were interleaved (in the uplink, e.g.), the resulting demodulated sequence is (quasi) *cyclostationary*. Generally, *Second Order Cyclostationary Statistics* (SOCS) are sufficient to retrieve the channel. However, there are "singular" channel classes which can *not* be identified [3]. We have shown in [4] that singular channels represent a severe limitation to SOCS-based methods because estimation performance from few samples is heavily affected if subchannel zeros are just "close" to each other. As this can *not* be prevented in a mobile environment, SOCS-based algorithms are neglected in this paper.

In summary, we have used the following criteria for the selection of a blind channel estimation algorithm:

(1) **Reliable channel estimates must be obtained from 142 samples of the demodulated sequence $y(k)$ only.**
(2) This should apply to arbitrary channels (even if there are zeros on/close to the complex plane's unit circle).
(3) As the effective channel order is unknown (and time-variant), an order misfit must not represent a problem.
(4) The estimates should be as robust as possible with respect to stationary additive Gaussian noise.

We have selected the EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI) by Kammeyer, Jelonnek, and Boss [5, 6], which is based on 4th (and 2nd) order statistics, because recent simulation results suggested that it meets the above criteria. For linear modulation with raised cosine transmit and receive filtering, we have demonstrated in [7] that EVI can blindly estimate COST-207 channels from 142 samples within a normalized mean square error bound of about 5 per cent (at a constant SNR of 7 dB). EVI's estimation performance was also compared with methods based on SOCS [4] and HOS [7]: On the above conditions, EVI's estimation performance was found to be superior in both cases. Finally, for GSM data transmission over COST-207 mobile channels, we have compared EVI with the optimum non-blind least squares (LS) scheme in terms of the resulting bit error rate (BER) after MLSE. We have demonstrated that EVI entails an average $\overline{\text{SNR}}$ loss of 1.1 dB only [8]. While the channel used for those simulations was time-*in*variant within each burst, we apply true time-variant filtering in this paper and investigate the effects on channel estimation as well as BER performance.

3. Simulation results

Fig. 4 shows the magnitude impulse response $|g_0(\tau,t)|$ according to (3), where $h_c(\tau,t)$ is a sample COST-207 *Hilly Terrain (HT)* channel obtained from (1) with $N_e = 100$. Both time axes are normalized to the GSM bit period $T \approx 3.7\,\mu$s. The max. Doppler frequency is set to $f_{d,max} = 200$ Hz. For GSM-900 and DCS-1800, this corresponds to velocities of the mobile unit in the ranges $225\ldots243$ km/h and $115\ldots126$ km/h, respectively. For Fig. 4, eq. (3) is evaluated over a t range covering 3 min. Doppler periods $T_{d,min} = 1/f_{d,max} = 5$ ms, i.e. 15 ms, or equivalently, 4062 bit periods or 26 burst periods. The surface lines are obtained by sampling $|g_0(\tau,t)|$ four times each bit period on the τ axis and once each burst period ($156.25\,T$, c.f. Fig. 3) on the t axis. Note that the magnitude impulse response $|h(\kappa,t)|$ of the linear model (2), which is the time-variant FIR system to be estimated, can also be seen at $\tau = \kappa T$.

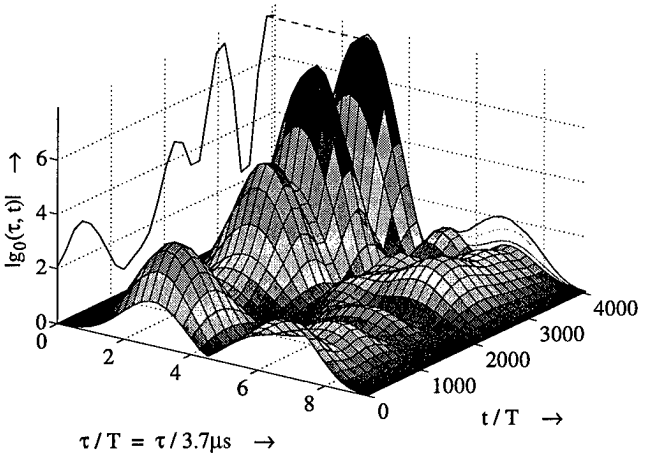


Fig. 4: Linear model of a sample *Hilly Terrain* composite channel

For the simulation of a GSM data transmission link according to Fig. 1, $h_c(\tau,t)$ is required rather than $g_0(\tau,t)$. However, $|h_c(\tau,t)|$ looks pretty much like Fig. 4 because of the bell-shaped form of the Laurent impulse $c_0(\tau)$. The "hills" are just slightly narrower (in direction of τ).

In order to ensure *(i)* meaningful measurements of BER and *(ii)* a satisfactory approximation of a GSUS channel by a sample impulse response $h_c(\tau,t)$, we use a range of t for the simulations which is much wider than the one displayed in Fig. 4. Let D_ξ denote the ξ-th burst containing 156 bits with values from $\{-1,1\}$. For the results given below, a total of 2000 bursts $D_0,\cdots,D_{1999}$ is transmitted, i.e. 232000 information bits. Since at most each 4th burst is sent to/from the same mobile station, this covers a t range of 8000 burst periods, i.e. 4.6s or 924 min. Doppler periods. To take the channel's time-variance within each burst period into account, $h_c(\tau,t)$ is sampled seven times per burst at

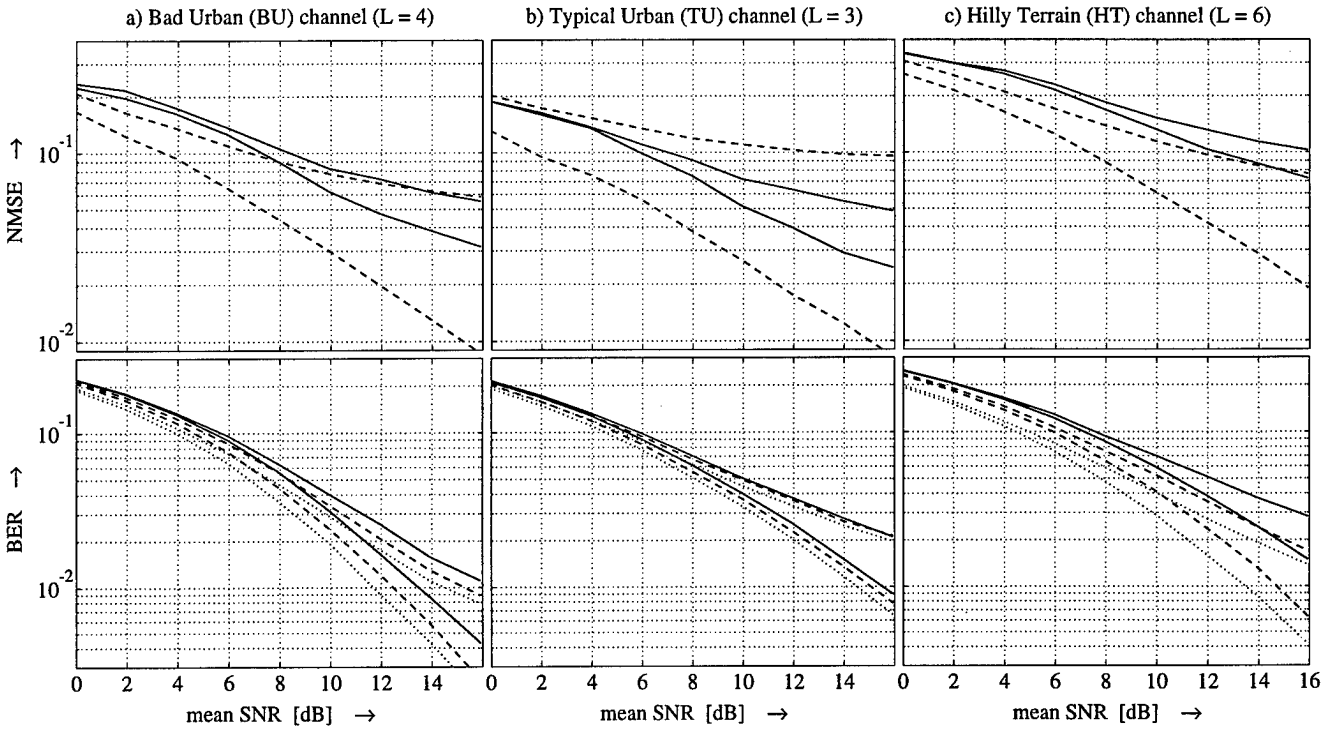


Figure 5: NMSE (above) and BER (below) in terms of $\overline{\text{SNR}}$ for the estimates of three sample COST-207 GSUS channels
solid: blind EVI channel estimation, dashed: non-blind CC channel estimation, dotted: "ideal" non-blind channel estimation

$t_{\xi,\mu} = (4 \cdot 156.25\,T)\,\xi + (26\,T)\,\mu$ with $\mu = 0, \cdots, 6$ and then interpolated linearly.

For the following description of the simulation procedure, assume that a sample impulse response $h_c(\tau, t)$ and a value of the mean signal-to-noise ratio $\overline{\text{SNR}}$ was selected. Referring to Fig. 1, the burst D_ξ is encoded, modulated, and then propagated through the time-variant channel $h_c(\tau, t)$, where $t_{\xi,0} \le t \le t_{\xi,6}$. Gaussian noise $r(t)$, colored by the receive filter, is added according to $\overline{\text{SNR}}$. Upon symbol-rate sampling and demodulation, 148 samples of $y(k)$ are obtained. Then, both the non-blind cross-correlation (CC) and the blind EVI channel estimation schemes are applied to $y(k)$, where the former utilizes the training midamble (see eq. (5)) while the latter uses 142 samples[3] of $y(k)$. Both approaches are given the number $L = \hat{q}+1$ of channel coefficients to be estimated, where L is calculated from the effective length of the sample power delay spectrum of the GSUS channel. Note that the actual mean effective length $\bar{q} + 1$ of the channel $h(\kappa, t)$ may well be shorter due to time selective fading (see Fig. 4 at $t \approx 1200T$, e.g.). The resulting channel estimates $\hat{h}(\kappa, \xi)$ are then passed to the Viterbi detector to obtain the estimated bursts $\hat{D}_\xi$.
In the frame of Monte-Carlo runs, this procedure is executed for 2000 bursts $D_0, \cdots, D_{1999}$ and $\overline{\text{SNR}}$ values in the range from 0 to 16 dB. Finally, BER is calculated from all bursts $\hat{D}_\xi$ and D_ξ transmitted at a given $\overline{\text{SNR}}$.

[3]For EVI, the training sequence was replaced with 26 additional data bits.

For each of the seven true channel slices $h(\kappa, t_{\xi,\mu})$ per burst ($\mu = 0, \cdots, 6$), the *Normalized Mean Square Error (NMSE)* of the associated estimate $\hat{h}(\kappa, \xi)$ is calculated[4]. Averaging over all values of ξ delivers the quality measure

$$\text{NMSE}(\mu) \triangleq \frac{1}{2000} \sum_{\xi=0}^{1999} \frac{\sum_\kappa |\hat{h}(\kappa, \xi) - h(\kappa, t_{\xi,\mu})|^2}{\sum_\kappa |h(\kappa, t_{\xi,\mu})|^2} \quad (7)$$

at seven instants $26T\mu$ per burst. Finally, a weighted average over all values of μ is calculated to obtain an overall measure (denoted NMSE) for each value of $\overline{\text{SNR}}$.

Figure 5: For three sample COST-207 GSUS channels, this figure displays the performance of both the channel estimation algorithms and the Viterbi detector in terms of $\overline{\text{SNR}}$. The upper row of subplots displays the NMSE values on a logarithmic scale, while the bottom row shows the bit error rates (BER). For the Figures 5a, b, and c, a *Bad Urban (BU)*, a *Typical Urban (TU)*, and a *Hilly Terrain (HT)* sample channel is used, where the number of coefficients of each estimate is set to $L = 4$, 3, and 6, respectively. The solid lines refer to EVI, while the dashed lines indicate the performance of the CC approach. All subplots contain two lines of either type, where the bottom one is obtained

[4]As all blind system identification methods can *not* estimate one complex factor, each estimate was multiplied with the optimum constant ($\Rightarrow$ min. Euclidean distance from the true channel) before NMSE was calculated. To ensure fairness of the comparison, this is done for CC's estimates, too.

by suppressing time-variance of the channel within each burst (although this is in accordance with the assumption made by the channel estimation and Viterbi algorithms, it is quite unrealistic and meant for reference only). Note that all curves suffer from the linearity assumption for the GMSK modulator as well as from the order misfit between L and the mean effective length $\bar{q}+1$ of the current channel.

From the bottom CC and EVI lines in the upper subplots, we realize that for EVI, NMSE is 1.3 to three times as high as for CC, where the factor 1.3 holds for low values of $\overline{\text{SNR}}$, while 3 applies above 14 dB. From the upper solid and dashed lines, however, we realize that CC's estimates are heavily affected by time-variance within each burst period while those of EVI degrade moderately so that estimation quality becomes rather similar. This is due to the fact that non-blind approaches have no information about the channel coefficients outside the training midamble, while blind approaches estimate some sort of *mean* channel slice. Note that for the *TU* channel, EVI even outperforms CC.

With the bottom subplots of Fig. 5, dotted lines are added for comparison. They are obtained when using "ideal" non-blind channel estimates for MLSE, i.e. $\hat{h}(\kappa,\xi)=h(\kappa,t_{\xi,3})$ for $\kappa=0,\cdots,L-1$. For the channels with suppressed time-variance within each burst (bottom three lines), we realize that at BER $\approx 2\%$, blind EVI channel estimation requires about 1 dB *(BU)*, 0.4 dB *(TU)*, and 2.2 dB *(HT)* more in $\overline{\text{SNR}}$ than the non-blind CC scheme. If the quasi time-invariance assumption over each burst is violated (upper three lines) BER levels off for high values of $\overline{\text{SNR}}$. For the *BU* channel, again, EVI requires about 1 dB more in $\overline{\text{SNR}}$ than CC. While this loss falls to zero for the *TU* channel, it increases to about 2.5 dB for the sample *HT* channel. Averaging over the three sample channels results in a mean loss of about 1.2 dB.

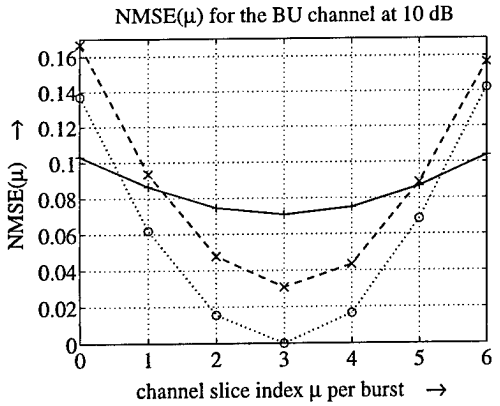


Figure 6: NMSE in terms of the channel slice index μ per burst for the EVI (solid), CC (dashed) and "ideal" (dotted) estimates of the *BU* channel at $\overline{\text{SNR}}=10$ dB

Figure 6: For the *Bad Urban (BU)* simulation results at 10 dB, this figure provides a detailed view of channel estimation quality in terms of the bit index per burst. It shows the evolution of NMSE as a function of the channel slice index μ. Obviously, "ideal" estimates (dotted line) based on the training midamble have a vanishing value of NMSE(μ) in the center of each burst only ($\mu=3$), while it increases as the burst boundaries are approached. At higher values, CC's estimates (dashed line) reveal a similar behavior. Although EVI's estimation quality is inferior at the center, it degrades just slightly towards the beginning and the end of the burst, as can be seen from the solid line.

4. Conclusions

We have demonstrated that for GSM data transmission over mobile comm. channels, the HOS-based blind channel estimation method EVI entails an $\overline{\text{SNR}}$ loss of 1.2 dB only (averaged over three COST-207 propagation environments) compared with the non-blind cross-correlation scheme. As just 142 samples of the demodulated sequence can be used for channel estimation, we consider these results quite remarkable for an algorithm based on HOS.

References

[1] P. Hoeher. A Statistical Discrete-Time Model for the WSSUS Multipath Channel. *IEEE Trans. on Vehicular Technology*, VT-41(4):461–468, April 1992.

[2] P. A. Laurent. Exact and Approximate Construction of Digital Phase Modulations by Superposition of Amplitude Modulated Pulses (AMP). *IEEE Trans. on Communications*, COM-34:150–160, 1989.

[3] J. K. Tugnait. On Blind Identifiability of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics. In *Proc. Global Telecom. Conf.*, pages 2001–2005, Houston, Texas, Dec. 1993.

[4] D. Boss and K. D. Kammeyer. Blind Identification of Mixed-Phase FIR Systems with Application to Mobile Communication Channels. In *Proc. ICASSP-97*, volume 5, pages 3589–3592, Munich, Germany, April 1997.

[5] D. Boss, B. Jelonnek, and K. D. Kammeyer. Eigenvector Algorithm for Blind MA System Identification. *EURASIP Signal Processing*, summer 1997. To appear.

[6] K. D. Kammeyer and B. Jelonnek. A New Fast Algorithm for Blind MA-System Identification based on Higher Order Cumulants. In *Proc. SPIE Advanced Signal Proc.: Algorithms, Architectures and Implementations V*, volume 2296, pages 162–173, San Diego, California, July 1994.

[7] D. Boss and K. D. Kammeyer. Blind GSM Channel Estimation. In *Proc. VTC-97*, Phoenix, Arizona, May 1997.

[8] D. Boss and K. D. Kammeyer. Blind GSM Channel Estimation based on Higher Order Statistics. In *Proc. ICC-97*, Montréal, Canada, June 1997.

Inverter Fed Induction Machine Condition Monitoring Using the Bispectrum

N. Arthur & J. Penman, Department of Engineering, University of Aberdeen, Scotland.

Abstract

This submission proposes the use of the unnormalised bispectrum as a signal processing tool for the diagnosis of inverter fed induction machine fault conditions. Increasingly, induction machines are supplied from non-sinusoidal, variable speed sources which increases the complexity and magnitude of the machine cage vibration. In addition, contamination of the vibration signal from both known and unknown sources makes accurate fault detection more difficult. This paper addresses both issues, experimental results are presented, and it is shown that the unnormalised bispectrum improves on the diagnostic capability of more conventional second order statistical measures.

1. Introduction

The induction machine is the single most common electromechanical energy conversion device available. It represents the backbone of the electrical engineering industry, and is generally a reliable and robust system component. That said, the induction machine does fail, most usually as the result of aging or poor construction. If failure is of a catastrophic nature, danger to plant and personnel may follow. Without a comprehensive mechanism for induction machine Condition Monitoring (CM), incipient machine failure often goes undetected, increasing the risk of catastrophic breakdown. If a CM regime is implemented, the impending failure can be detected, a risk assessment made and remedial action scheduled. For example, an induction machine may be a part of a critical process, the machine might develop a minor fault subsequently detected by an accurate induction machine CM system. If the fault is only likely to deteriorate gradually and not result in catastrophic failure, the machine can continue to be operated and the subsequent deterioration monitored. Maintenance can be planned for the next available shutdown, preserving the criticality of the machine without compromising safety.

Many methods of induction machine CM exist, but in general diagnosis is obtained by evaluation of a second order statistical measure, the power spectral density (psd) of a characteristic machine signal. That signal can be any one of a number of system properties, but is usually the stator electromagnetic vibration, motor phase current, air gap axial flux or emitted noise. [1-4]. Each of these methods has certain advantages and allows induction machine CM to a certain extent. However, no panoptic means of fault detection has been found as yet. It is the authors opinion that this lack of integrity is symptomatic of the processing techniques employed, rather than the signals being monitored.

In recent years there has been an increase in research into Higher Order Statistics (HOS). Initially this research concentrated primarily on the theoretical aspects of HOS, with little recourse to practical applications. Recent work [5-7] has suggested that third order measures conveys more pertinent CM information than the power spectrum alone, consequently more attention is being paid to the application of HOS measures to real life issues. A proportion of the literature has been devoted to the use of HOS as a CM tool, i.e. for the diagnosis of electrical machine faults [1], gearbox faults [8] and structural fault diagnosis [9]. It is contended that the application of HOS techniques to the CM of real electrical systems is timely [10], as rotating machine faults give rise to a non-linear mode of operation and HOS techniques allow single sensor diagnosis of system non-linearities. Further, the application of HOS to machine vibration signals results in the reduction of Gaussian noise. In addition, the retention of all signal phase information is an inherent feature of the calculation of HOS.

To the best of the authors knowledge, none of the previous work has quantitatively compared HOS and the psd for the detection of electrical machine faults. This paper will introduce variable frequency supplies and will compare the abilities of the power spectrum and the unnormalised bispectrum to diagnose inverter fed induction machine fault conditions using contaminated vibration signals.

2. Inverter Drives

A relatively recent advance in the utilization of induction machines is the wide use of variable speed

0-8186-8005-9/97 $10.00 © 1997 IEEE

drives, for example, the Pulse Width Modulated (PWM) inverter. These drives enjoy huge popularity as they allow flexible, variable speed operation of the induction machine. Figure 1 shows a simplistic PWM inverter drive.

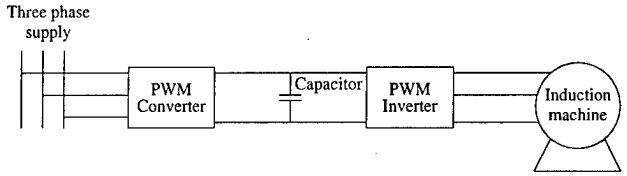


Figure 1. Simple PWM inverter drive.

The PWM converter rectifies the three phase input voltage and controls the magnitude of the voltage in the dc link. The capacitor is primarily for smoothing the dc link voltage, and the PWM inverter is used to control the frequency of the output voltage waveform. Figure 2 shows a typical single phase output voltage waveform from a PWM inverter supply.

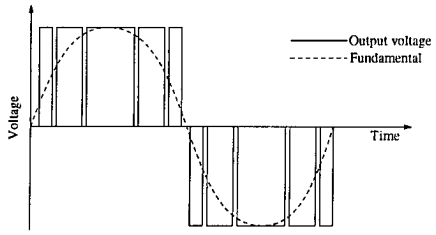


Figure 2. Single phase PWM output.

Clearly, the output voltage is far from sinusoidal and contains many harmonic components. One of the consequences of the use of PWM inverters is a subsequent increase in the magnitude and complexity of machine vibration, due to the high harmonic content of the motor stator voltage [11]. This has implications for the mechanical reliability of inverter fed induction machines and the levels of acoustic noise emitted. A further consideration when using PWM inverters is the derating of the machine because of additional heating effects caused by the non sinusoidal supply harmonics. Any diagnostic system based on the measurement of machine vibration must be able to interpret this complex machine signal and produce definitive and reliable fault diagnosis. For this to happen, important or relevant features identifying each machine condition must be monitored, extracted and diagnosis reached if these features reach prescribed levels.

3. Contamination of the Vibration Signal

In any industrial environment the ambient or background vibration at any location varies temporally as machines operate and loading criteria fluctuate. The ambient vibration also varies spatially as the material properties of foundations and machinery mountings change. This impacts directly on any vibration based CM system which is in use. Clearly, any change in background vibration manifests itself as a change in the vibration of the machine under observation, resulting in a variation of any signal characterising machine performance [12]. Further contamination of the machine vibration signal may originate from interference present in the electrical supply, or from other, unknown sources.

It is clear that any contamination of the induction machine vibration signal will result in degradation of the information contained therein, thus making interpretation of the signal more difficult. It is helpful therefore, to have a CM system as impervious to contaminants as possible, such that important features contained in the machine vibration are retained. It has been shown [13] that the evaluation of a signals HOS results in the elimination of additive Gaussian noise. However, it has also been shown [14] that for data records of fixed length, the noise level is attenuated and not removed completely. Assuming we can define these vibration signal contaminants to be random, and to have a Gaussian distribution and zero mean [9], it is proposed that an induction machine CM system based on HOS, and in particular the unnormalised bispectrum, will result in the reduction of noise in the vibration signal and allow the retention of important fault characteristic features.

4. Definitions

The decomposition of a signal into its spectral components provides a great deal of information. The Fourier Transform (FT) is the most common way of computing the spectral components of a signal. Consider a discrete-time series $x(n)$, $n=0, \pm1, \pm2 \ ...etc$, its FT, $X(k)$, is described by

$$X(k) = \sum_{n=-\infty}^{\infty} x(n) \exp(-j2\pi kn) \qquad (1)$$

The most frequently used tool for signal processing over the past thirty years has been the evaluation of signals' power spectral density (psd). The psd is a measure of the spread of the signal power over a range of frequencies, considering each frequency to be

independent of all others. The psd of a signal can be given in terms of its FT, by

$$S_{xx} = E[X(k)X^*(k)] \quad (2)$$

Where $E[\ ]$ represents the expectation operator and $X^*(\ )$ is the complex conjugate of $X(\ )$. As an extension to the estimation of these second order measures, a product of Fourier Transforms can be formed at different frequencies to produce the third order measure called the bispectrum. The bispectrum considers the degree of coupling between three frequencies, k, l and $k+l$. The bispectrum of $x(n)$ is described by

$$B_x(k,l) = E[X(k)X(l)X^*(k+l)] \quad (3)$$

Clearly, we can see that the bispectrum is a function of two frequency indices and as such is more computationally intensive to estimate than the psd. In the calculation of the bispectrum only bifrequencies in the Inner Triangle (IT) of the principle domain will be considered, in order to minimise the possibility of sampling rate discrepancies degrading the bispectral estimation [15], ensuring best reliability.

5. Laboratory Work

A 3kW, 6-pole, squirrel cage induction machine was used to drive an axial flow screw air compressor via a speed up helical gearbox, the induction machine being fed from a P.W.M. inverter as shown in Figure 3.

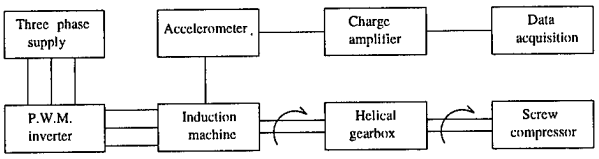


Figure 3. Laboratory apparatus.

The induction machine vibration was measured by a piezoelectric accelerometer mounted radially on the casing of the machine. The ability to replicate common induction machine faults was a central feature of the electromechanical system. In addition to the no fault condition, four single and four combined fault conditions were seeded on the induction machine. The faults implemented were selected to reflect common faults found in induction machines in normal service. After the implementation of each fault, a fifteen minute warm up period was allowed to ensure that the induction machine had reached steady state operation. The induction machine conditions were as detailed in Table 1.

Table 1. Conditions implemented on test machine.

Condition	Description
1	No fault
2	One broken rotor bar
3	Two broken rotor bars
4	Stator short circuit
5	Unbalanced supply
6	Two broken rotor bars, stator short circuit
7	Two broken rotor bars, unbalanced supply
8	Stator short circuit, unbalanced supply
9	Two broken rotor bars, stator short circuit, unbalanced supply

For each of the nine machine states, 60 motor vibration signatures of 20480 points were conditioned, sampled at 100kHz and acquired.

6. Data Processing

Data processing was achieved in two stages; induction machine fault templates were constructed, then the diagnostic capabilities of the psd and the bispectrum were evaluated.

6.1 Fault template construction

The induction machine vibration signatures were split into two groups, half being designated 'template' data and half 'testing' data. The power spectrum and bispectrum of the 270 'template' vibration signatures was computed using a 1024 point fft, Hanning window and 70% overlap. The frequency and bifrequency components of the 'template' power spectra and bispectra were sorted into a descending order of importance, according to the product of the component mean and standard deviation, evaluated over all machine conditions.

The range of each frequency and bifrequency component of the psd and bispectrum was stored for each machine condition, giving a characteristic template for the psd and bispectrum for all machine conditions.

6.2 Diagnostic evaluation

The power spectrum and bispectrum of the 'testing' data with 0, 10, 20 and 30% of normally distributed,

white, additive Gaussian noise was computed, the noise being added to simulate contamination of the vibration signal. The frequency and bifrequency components of the 'testing' data were sorted as previously. These spectra were then compared with all machine condition 'templates'. The machine condition was only diagnosed if all spectral components considered fell within the range of components of a characteristic template. Thè condition thus determined was compared with the machine state in order to evaluate the accuracy of diagnosis. The whole process was repeated with 10, 20 and 30% contamination of the 'template' vibration traces.

7. Results

Figure 4 shows the variation of diagnostic capability of the psd for a range of important frequency components, and with varying degrees of vibration contamination. In this instance the characteristic template was constructed with 0% vibration contamination.

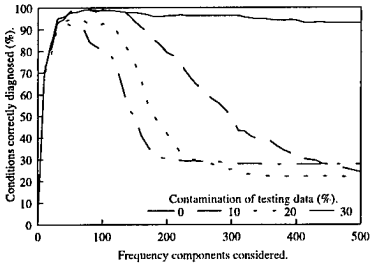


Figure. 4. Psd diagnosis, 0% 'template' contamination.

It can be seen that maximum diagnosis was achieved when the contamination of the 'testing' data was 0%, the same as when the 'template' was created. Further, the diagnosis worsened in proportion to the added contamination. In all cases the diagnosis worsened as the number of frequency components considered increased after maximum diagnosis. Figure 5 shows the diagnostic variation of the psd with 20% vibration contamination.

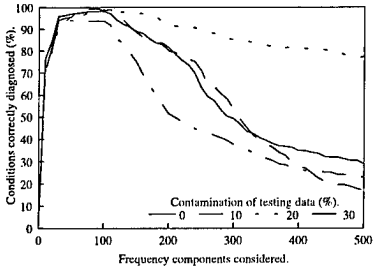


Figure. 5. Psd diagnosis, 20% 'template' contamination.

Figure 5 shows that maximum diagnosis was attained when the contamination of the 'testing' data was 20%. Other values of vibration contamination degraded performance. Similar results were obtained for 10 and 30% contamination of vibration in the 'template' data. Figure 6 shows the diagnostic variance of the bispectrum for 0% contamination of the vibration when creating templates.

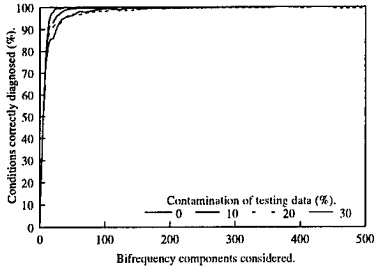


Figure. 6. Bispectrum diagnosis, 0% 'template' contamination.

Figure 6 shows how the maximum diagnosis of the bispectrum is effectively invariant of the vibration contamination in the 'testing' data. Also, the diagnosis is invariant of the number of important bifrequencies considered. Figure 7 details bispectrum variance for 20% vibration contamination in the template data.

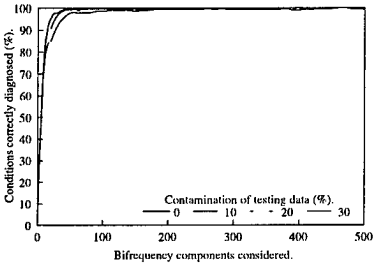


Figure 7. Bispectrum diagnosis, 20% 'template' contamination.

Again it can be seen that the maximum diagnosis is invariant of the contamination of the vibration testing data and the number of bifrequency components considered. These results were consistent for 10 and 30% vibration contamination in the template data, also.

In summary, it is clear that the diagnostic success of the psd is dependent on the degree of vibration contamination contained in both the 'template' and 'testing' data. Acceptable diagnosis is only achieved when the vibration contamination contained in the 'template' data matches that of the 'testing' data. Similarly, degradation of diagnosis is proportional to the

difference between the vibration contamination contained within the 'template' and 'testing' data. Also, it has been shown that the psd diagnosis is dependent on the number of frequency components of the power spectrum which are considered to form the diagnosis, the added noise making accurate diagnosis impossible when the less important components were considered. Clearly, this makes the psd a somewhat limited diagnostic device, with accurate diagnosis dependent on the level of vibration corruption and, on analysing the optimal number of frequency components.

In contrast, it can be seen that the bispectrum provides 100% accurate diagnosis irrespective of the vibration contamination difference between 'template' and 'testing' data. In addition, this diagnosis is maintained independent of the number of bifrequency components considered after maximum diagnosis. Clearly, this use of HOS reduces the effect of the added noise, even when considering less important components of the signal.

8. Conclusions

It is clear from the results presented that the bispectrum offers a valid means of CM inverter-fed induction machines. Accurate machine condition diagnosis was achieved repeatably over a number of different machine faults. It is also evident that the bispectrum provides enhanced fault recognition when compared with the more conventional power spectrum based technique. In addition, the diagnostic capability of the bispectrum based regime is insensitive to vibration contamination and retains conclusive diagnosis regardless of the number of bispectral components considered.

The authors feel this represents an advance in diagnostic technology, with the bispectrum clearly identifying spectral components which uniquely identify various fault conditions. New work is to be concentrated on identifying important components explicitly in an attempt to make the 'fingerprint' approach redundant. This will allow predictive analysis to be made and remove the need for *a priori* data.

Acknowledgments

This work was supported by contract SSDW3/0005 from DERA Winfrith under the MOD Corporate Research Programme. (C) British Crown copyright 1997/DERA, published with permission of the Controller of Her Britannic Majesty's Stationery Office.

References

[1] T.W.S. Chow, G. Fei, "Three phase induction machine asymmetrical faults identification using bispectrum", IEEE Trans. on Energy Conversion, Vol. 10, No. 4, December 1992, pp. 688-693.

[2] R. R. Schoen et al, "An Unsupervised, On-Line System for Induction Motor Fault Detection Using Stator Current Monitoring", IEEE Trans. On Industry Applications, Vol. 31, No. 6, 1995, pp. 1280-1286.

[3] J. Penman *et al*, "Detection and location of interturn short circuits in the stator windings of operating motors", IEEE Trans. On Energy Conversion, Vol. 9, No. 4, December 1994, pp. 652-658.

[4] A. Gaylard, A. Meyer, C. Landy, "Acoustic evaluation of faults in electrical machines", IEE Conf. Proc. On EMD, September 1997, pp. 147-150.

[5] C.L. Nikias, J.M. Mendel, "Signal processing with higher-order spectra", IEEE Signal Processing Magazine, July 1993, pp. 10-36.

[6] T.W.S. Chow, G. Fei, "Three phase induction machine asymmetrical faults identification using bispectrum", IEEE Trans. on Energy Conversion, Vol. 10, No. 4, December 1992, pp. 688-693.

[7] Fackrell et al, "The interpretation of the bispectra of vibration signals-II experimental results and applications", Mechanical Systems and Signal Processing, Vol. 9, No. 3, 1995, pp. 267-274.

[8] J. Sato, K. Sasaki, Y. Nakamura, "Real-time bispectral analysis of gear noise and its application to contactless diagnosis", Journal Acoust. Soc. Wm., Vol. 62, No.2, August 1977, pp. 382-387.

[9] F. Huang, S. Gu, "Application of higher order cumulants to structure fault diagnosis", Proc. SPIE, Vol. 1923, No. 2, 1993, pp.1237-1240.

[10] Fackrell et al, "The interpretation of the bispectra of vibration signals-I experimental results and applications", Mechanical Systems and Signal Processing, Vol. 9, No. 3, 1995, pp. 257-266.

[11] R. Yacamini, S. C. Chang, "Noise and vibration from induction machines fed from harmonic sources", IEEE Trans. on Energy Conversion, Vol. 10, No. 2, June 1995, pp. 286-292.

[12] C. K. Mechefske, "Machine Condition Monitoring:Part 2-The Effects of Noise in The Vibration Signal", Brit. Jour. of NDT, Vol. 35, No. 10, Oct. 1993, pp.574-579.

[13] J.W.A. Fackrell et al, "Practical issues concerning the use of the bicoherence for the detection of Quadratic Phase Coupling", IEEE Sig. Proc./ATHOS Workshop on HOS, Begur, Spain, 1995. pp.310-314.

[14] J.W.A. Fackrell, S. McLaughlin, "Quadratic phase coupling detection using higher order statistics", IEE Colloquium on H.O.S., Savoy Place, London, May 1995.

[15] I. Harfer, H. Messer, "The Bispectrum of Sampled Data: Part I-Detection of the Sampling Jitter", IEEE Trans on Signal Processing, Vol. 41, No. 1, 1993, pp. 296-312.

Higher-Order Statistics for Tissue Characterization from Ultrasound Images *

Udantha R. Abeyratne and Athina P. Petropulu
Electrical and Computer Engineering Department
Drexel University, Philadelphia, PA 19104
Tel. (215) 895-2358 Fax. (215) 895-1695

ABSTRACT

We model tissue as a collection of point scatterers embedded in a uniform media, and show that the higher-order statistics (HOS) of the scatterer spacing distribution can be estimated from digitized RF scan line segments and be used in obtaining tissue signatures. Based on our model for tissue microstructure, we estimate resolvable periodicity and correlations among non-periodic scatterers. Using higher-order statistics of the scattered signal, we define as tissue "color" a quantity that describes the scatterer spatial correlations, show how to estimate it from the higher-order correlations of the digitized RF scan line segments, and investigate its potential as a tissue signature.

1 Introduction

Ultrasound imaging is a widely used technique in the diagnosis of tumors of soft tissues. Currently, the clinical diagnosis is based on visual interpretation of images by radiologists. There have been many attempts at developing objective tissue characterization criteria on the premise that there is much more observer-independent information available from ultrasound than what is currently being used. These are rooted on the fundamental notion that the biological tissues are composed of characteristic structures whose ultrasonic properties often change due to diseases. The goal of tissue characterization is the extraction of signatures that assume distinct values in the presense of normal and diseased states of tissues, such that it is possible to differentiate between them.

Although the exact identities of the physical structures responsible for ultrasound backscattering are generally unknown [17], [3], they can usually be characterized either as macroscopic or microscopic, compared with the wavelength of the interrogating ultrasonic beam. The tissue is often modeled as a collection of point scatterers, embedded in a uniform non-scattering medium. Considering the biological variabilities associated with tissues, the spatial distribution and the scattering strengths (scattering cross-sections) associated with these scatterers are usually described in statistical terms. The statistics of the inter-scatterer spacing distribution are commonly used as tissue signatures.

In organs such as the liver the overall arrangement of the tissues consists of an organized repetition of a basic structural unit. This repeated structure is regarded to be providing resolvable, repeated scattering centers for the propagating ultrasonic pulse [8], [10], [11], [5]. It has been shown that a periodicity can be observed from the liver pulse-echo data, and the estimated periodicity can be linked to the mean-scatterer-spacing of the underlying tissue [8]. It was also shown that the mean scatterer spacing may be used as a feature for tissue characterization of diffuse liver diseases, such as cirrhosis and chronic active hepatitis [8],[11].

The scatterer spatial correlation has been used in the past to model ultrasonic properties of tissue [9],[13],[6],[?] and was shown to lead to promising tissue signatures. In [13], it was shown that using power spectra, deterministic, membrane-like structures as well as structures consisting of random, diffuse structures can be characterized. In [14], effective scatterer sizes, concentrations and acoustic impedances were investigated using power spectra, as potential tissue signatures. The effective scatterer size was reported to be the most important tissue feature sensed with the method of [14].

In this paper, we use a point scatterer model similar to the ones in [11] and [7] to describe tissue microstructure, with the exception that our model takes the correlations existing among both resolvable and non-

*This work was supported by NSF under grant MIP-9553227 and the Whitaker Foundation.

0-8186-8005-9/97 $10.00 © 1997 IEEE

resolvable scatterers into account. This enables us to also consider cases of coherent long-range scatterer distributions which have high variances associated with the inter-scatterer spacing, i.e. almost no periodicity, and, non-periodic echos resulting from correlated, non-periodic components of both diffused and structural scatterers. We assume that RF echoes are non-Gaussian, on the grounds of empirical/theoretical justifications presented in works such as [19], [4], [2], [18] and [20]. Based on our model for tissue microstructure, we develop schemes for the estimation of resolvable periodicity as well as correlations among non-periodic scatterers. Using higher-order statistics of the scattered signal, we define as tissue "color" a quantity that describes the scatterer spatial correlations, show how to evaluate it from the higher-order correlations of the digitized RF scan line segments, and investigate its potential as a tissue signature. We also preliminary evidence that even when there is no significant periodicity in data, we are still able to characterize tissues using signatures based on the higher-order cumulant structure of the scatterer spacing distribution.

2 Theoretical developments

Assuming a narrow ultrasound beam, linear propagation and weak scattering in the tissue medium, we model the ultrasonic RF echo, $y(t)$, by,

$$y(t) = h(t) * f(t) + w(t), \tag{1}$$

where $h(t)$ is the pulse-echo wavelet, $w(t)$ is the observation noise, and "*" denotes convolution. The quantity $f(t)$, which we will call *the tissue response*, represents the underlying tissue microstructure.

Following [3],[11] and [7] we model the structures within tissue that are responsible for backscattered ultrasonic field by point-scatterers organized at different levels. Existing work such as [11] and [7] attempt to characterize liver tissue based on a two component point scatter model: (a) the *diffused component*, which represents [11] randomly positioned scatterers of sufficient concentrations to produce echo signals with circular Gaussian statistics. The positions of individual scatterers are assumed uncorrelated [11]; (b) the *coherent component*, which represents non-randomly distributed scatterers with long-range order [11], [10].

In this paper (see also [1]), we describe liver tissue by a three-component model, which in addition to (a) and (b) above, also takes into account the correlations existing among (i) diffused scatterer locations, (ii) resolvable and unresolvable pseudo-periodic scatterers leading to long range order, and (iii) collection of scatterers leading to short range order, but not strong enough to violate the conditions of weak-scattering or the stationarity of the RF echo in a significant manner.

We model the tissue response $f(t)$ as:

$$f(t) = r_d(t) + r_m(t) + r_r(t), \tag{2}$$

where $r_d(t)$ represents unresolvable diffuse scatterers leading to fully-developed speckle with Gaussian statistics [11]; $r_r(t)$ represents resolvable, coherent scatterers with long-range order (pseudo-periodicity) [11]. The quantity $r_m(t)$ represents the combined effects of unresolvable periodicity from structural scatterers, and, correlated non-periodic components of diffused and structural scatterers (short-range order). The terms $r_m(t)$, $r_r(t)$ (and their contributions to the RF-echo $y(t)$) are non-Gaussian, due to contributions from periodic structures and/or scatterers of short range order. The use of higher-order statistics (HOS) as the tool allows us to properly treat the non-Gaussian components, to which deterministic structures such as large cell assemblies and necrotic areas inside tumors contribute.

The ultrasonic rf echo, $y(t)$, can be expressed as:

$$y(t) = y_d(t) + y_m(t) + y_r(t) + w(t), \tag{3}$$

where $y_d(t) = r_d(t) * h(t)$, $y_m(t) = r_m(t) * h(t)$ and $y_r(t) = r_r(t) * h(t)$ are respectively called [1] the *diffused* (DC), *mixed* (MC) and *resolvable-periodic* (RPC) components of $y(t)$. The observation noise $w(t)$ is assumed zero-mean, Gaussian and uncorrelated with $y_i(t)$, $i = d, m, r$.

Under the assumption that $y_i(t)$, $i = d, m, r$ are zero-mean, mutually uncorrelated, the third-order cumulants $c_3^y(\tau_1, \tau_2)$ [15] of the process $y(t)$ can be written as

$$c_3^y(\tau_1, \tau_2) = c_3^{y_d}(\tau_1, \tau_2) + c_3^{y_m}(\tau_1, \tau_2) + c_3^{y_r}(\tau_1, \tau_2) + c_3^w(\tau_1, \tau_2) \tag{4}$$

where $c_3^{y_d}(\tau_1, \tau_2)$, $c_3^{y_m}(\tau_1, \tau_2)$, $c_3^{y_r}(\tau_1, \tau_2)$, $c_3^w(\tau_1, \tau_2)$ respectively denote the third-order cumulants of $y_d(t)$, $y_m(t)$, $y_r(t)$ and the observation noise $w(t)$.

Since noise and **diffused compoment** are Gaussian processes,

$$c_3^w(\tau_1, \tau_2) = c_3^{y_d}(\tau_1, \tau_2) = 0, \ \forall(\tau_1, \tau_2). \tag{5}$$

The suppression of the Gaussian component of the tissue response is an important advantage because it does not carry any information about tissue.

The **mixed component** $y_m(t)$ of the RF echo $y(t)$ is modeled as:

$$y_m(t) = \left\{ \sum_{p=1}^{N_m} b_p \delta(t - \lambda_p) \right\} * h(t). \tag{6}$$

Let us assume that: (A1) λ_i is uncorrelated to b_j, $i,j = 1,2,\cdots,N_m$; (A2) λ_i, $i = 1,2,\cdots$, form a stationary process; (A3) $y_m(t)$ is a zero-mean process.

Under (A3) the third order cumulants of $y_m(t)$ equal,

$$\begin{aligned} c_3^{y_m}(\tau_1,\tau_2) &= E\{y_m(t)y_m(t+\tau_1)y_m(t+\tau_2)\}, \\ &= M_3^{r_m}(\tau_1,\tau_2) * c_3^h(\tau_1,\tau_2), \end{aligned} \quad (7)$$

with,

$$\begin{aligned} M_3^{r_m}(\tau_1,\tau_2) &= \sum_{p=1}^{N_m}\sum_{q=1}^{N_m}\sum_{s=1}^{N_m} E\{b_p b_q b_s\} \\ \times E\{\delta(t-\lambda_p)\delta(t+\tau_1-\lambda_q)\delta(t+\tau_2-\lambda_s)\}, \end{aligned} \quad (8)$$

where (A1) has been used in obtaining (8).

Let the joint probability density function (arbitrary) for the scatterer location triplets $\lambda_p, \lambda_q, \lambda_s$ be given by $f_m(\lambda_p,\lambda_q,\lambda_s)$. Then from (8) we obtain:

$$\begin{aligned} M_3^{r_m}(\tau_1,\tau_2) &= \sum_{p=1}^{N_m}\sum_{q=1}^{N_m}\sum_{s=1}^{N_m} E\{b_p b_q b_s\} \\ &\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t-\lambda_p)\delta(t+\tau_1-\lambda_q)\delta(t+\tau_2-\lambda_s) \\ &\quad \times f_m(\lambda_p,\lambda_q,\lambda_s)d\lambda_p d\lambda_q d\lambda_s \\ &= \sum_{p=1}^{N_m}\sum_{q=1}^{N_m}\sum_{s=1}^{N_m} E\{b_p b_q b_s\} \\ &\quad \times f_m(t,t+\tau_1,t+\tau_2). \end{aligned} \quad (9)$$

Under the assumption (A2), the probability density function $f_m(t,t+\tau_1,t+\tau_2)$ does not depend on the particular t chosen. Thus, $f_m(t,t+\tau_1,t+\tau_2)$ can be written as $f_m(\tau_1,\tau_2)$. Therefore, from (7) and (9) we obtain:

$$c_3^{y_m}(\tau_1,\tau_2) = K_m f_m(\tau_1,\tau_2) * c_3^h(\tau_1,\tau_2), \quad (10)$$

where K_m is given by,

$$K_m = \sum_{p=1}^{N_m}\sum_{q=1}^{N_m}\sum_{r=s}^{N_m} E\{b_p b_q b_s\}. \quad (11)$$

The **resolvable periodic component** has been shown to be of great significance in tissue characterization[8], [7]. *To be classified as the resolvable periodic component, the scatterer separation should be regular enough, and the repeat distances should be large enough (compared to the length of the pulse-echo impulse response) to be resolvable.* The Gamma distribution has proven to be a particularly useful tool in describing inter-scatterer space distribution [12], [7], because of its versatility in producing scatterer locations ranging from almost periodic to clearly non-periodic.

The general mathematical description of the RPC follows directly from the expressions derived for the case of MC component. The resolvable periodic component, $y_r(t)$, can be written as:

$$y_r(t) = \left\{\sum_{p=1}^{N_r} v_p \delta(t-\theta_p)\right\} * h(t). \quad (12)$$

Again, we assume: (B1) θ_i is uncorrelated to v_j, $i,j = 1,2,\cdots,N_r$; (B2) θ_i, $i = 1,2,\cdots$, froms a stationary process; (B3) $y_r(t)$ is a zero-mean process.

From (12) we can obtain as in the case of MC component,

$$c_3^{y_r}(\tau_1,\tau_2) = K_r f_r(\tau_1,\tau_2) * c_3^h(\tau_1,\tau_2), \quad (13)$$

$$K_r = \sum_{p=1}^{N_r}\sum_{q=1}^{N_r}\sum_{s=1}^{N_r} E\{v_p v_q v_s\}, \quad (14)$$

where the joint probability density function $f_r(\tau_1,\tau_2)$ now describes a (pseudo) periodic phenomenon, i.e. $f_r(\tau_1,\tau_2)$ is a two dimensional bed of spikes, whose separation is equal to the mean-scatterer-spacing.

In the following we derive an expression for $f_r(\tau_1,\tau_2)$, under the constraint that the resolvable periodic scatterers are separated by a constant time interval, T, i.e. the process is strictly periodic and the resolvable periodicity (unknown) is T. Then the periodic component $r_r(t)$ can be written as:

$$r_r(t) = \sum_{p=1}^{N_r} v_p(t)\delta(t-pT), \quad (15)$$

where $v_p(t)$ is the *scatterer strength function*, and is a bounded function defined over the set of real numbers, $\mathcal{R}_{(-\infty,\infty)}$. The Fourier transform $\tilde{R}_r(\omega)$ of $r_r(t)$ can be obtained as,

$$R_r(\omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty} V(\omega - \frac{2\pi}{T}k), \quad (16)$$

where $V_p(\omega)$ is the Fourier transform of $v_p(t)$. Let the region of support of the function $v_p(t)$ be $\{t : |t| < B\}$. As $B \to \infty$, $V(\omega) \to \delta(\omega)$. As $B \to \infty$ the bispectrum [15] $C_3^{r_r}(\omega_1,\omega_2)$ of $r_r(t)$ is given by:

$$\begin{aligned} C_3^{r_r}(\omega_1,\omega_2) &= R_r(\omega_1)R_r(\omega_2)R_r(-\omega_1-\omega_2), \\ &= (\frac{2\pi}{T})^3 \sum_{k_1}\sum_{k_2}\sum_{k_3} \delta(\omega_1 - \frac{2\pi}{T}k_1)\delta(\omega_2 - \frac{2\pi}{T}k_2) \\ &\quad \times \delta(-\omega_2-\omega_1 - \frac{2\pi}{T}k_3), \end{aligned} \quad (17)$$

$$= (\frac{2\pi}{T})^3 \sum_{k_1} \sum_{k_2} \delta(\omega_1 - \frac{2\pi}{T}k_1)\delta(\omega_2 - \frac{2\pi}{T}k_2). \quad (18)$$

For the case $V(\omega) \rightarrow \delta(\omega)$, from (12) and (18) we can obtain the bispectrum $C_3^{y_r}(\omega_1, \omega_2)$ of $y_r(t)$ as:

$$C_3^{y_r}(\omega_1, \omega_2) = (\frac{2\pi}{T})^3 \{C_3^h(\omega_1, \omega_2) \sum_{k_1} \sum_{k_2} \delta(\omega_1 - \frac{2\pi}{T}k_1) \times \delta(\omega_2 - \frac{2\pi}{T}k_2)\}. \quad (19)$$

The third-order cumulant sequence $c_3^{y_r}(\tau_1, \tau_2)$ of the RPC can be obtained from (19) through an inverse Fourier transform, i.e.,

$$c_3^{y_r}(\tau_1, \tau_2) = c_3^h(\tau_1, \tau_2) * \sum_{k_1} \sum_{k_2} \delta(\tau_1 - k_1 T)\delta(\tau_2 - k_2 T). \quad (20)$$

Comparing (14) and (20) we get:

$$f_r(\tau_1, \tau_2) = \sum_{k_1} \sum_{k_2} \delta(\tau_1 - k_1 T)\delta(\tau_2 - k_2 T), \quad (21)$$

Finally, the third-order cumulant of the RF-echo equals:

$$c_3^y(\tau_1, \tau_2) = \{K_m f_m(\tau_1, \tau_2) + K_r f_r(\tau_1, \tau_2)\} * c_3^h(\tau_1, \tau_2),$$
$$= \{K_m f_m(\tau_1, \tau_2) + \sum_{k_1} \sum_{k_2} \delta(\tau_1 - k_1 T)\delta(\tau_2 - k_2 T)\} * c_3^h(\tau_1, \tau_2), \quad (22)$$

when the inter-scatterer spacing of the RPC component is a constant T.

The periodicity can be estimated from (22). The term $K_m f_m(\tau_1, \tau_2) * c_3^h(\tau_1, \tau_2)$ in Eq. (10) can be considered as the third-order cumulant sequence of the signal

$$z(t) = e(t) * u(t), \quad (23)$$

where $e(t)$ is a non-Gaussian noise process whose third order cumulants equal K_m, and $u(t) = g(t) * h(t)$ is a linear time invariant system. We call $g(t)$ the *color of the tissue response*. Based on (10) the third-order cumulant sequence $c_3^z(\tau_1, \tau_2)$ of $z(t)$ can be retrieved from $c_3^y(\tau_1, \tau_2)$. Using methods of system reconstruction from higher-order spectra, we can recover $u(t)$ from $c_3^z(\tau_1, \tau_2)$.

The quantities $u_i(t)$ and $u_j(t)$ obtained from two different RF line segments "i" and "j" can be written as $u_k(t) = h(t) * g_k(t)$, $k = i, j$, where $h(t)$, are invariant between "i" and "j"; $g_i(t)$ and $g_j(t)$ respectively represent the color of the tissue responses at "i" and "j". The quantities $g_k(t)$, $k = i, j$ carry information about the scatterer spacings corresponding to the MC component, and is also free from the effects of $h(t)$, making it a good candidate for a tissue signature. Using $u_1(t)$ and $u_2(t)$ we form a tissue signature $S_{ij}(t)$ as: $S_{ij}(t) = u_i(t) * \{u_j(t)\}^{-1} = h(t) * g_i(t) * \{h(t) * g_j(t)\}^{-1} = g_i(t) * g_j^{-1}(t)$ where $\{\cdot\}^{-1}$ denotes the convolutional inverse.

If the estimations are from two regions with similar ultrasonic properties, $S_{ij}(t)$ will approach a Dirac delta-function. The closeness of S_{ij} to a Dirac delta performs as a measure of the similarity between two different regions.

3 Results

We obtained RF ultrasound data corresponding to clinical images of human livers containing tumors. An example of such an image is shown in Fig. 1(a) where the tumor is visible in the middle part of the image. Using RF-echoes inside and outside the known tumor location, we estimated the periodicity of the scatterer spacing (m). Histogram for m is shown in Fig.1 (b). From Fig. 1 we conclude that m serves as a feature for characterizing the tumor in shown in Fig. 1(a). Figure 2 shows the color of the tissue responses estimated from three different liver images where the periodicity failed as feature. Figs. 2(a), (b) and (c) show the color of the tissue responses estimated inside tumors and Figs. 2(d), (e) and (f) show the corresponding ones estimated outside the tumors. The solid lines indicate the average estimate and the dotted lines indicate the average $\pm$ standard deviation. The color of the tissue response compactly summarizes the correlation structure existing among the scatterers of both short-range and long-range order. It performs as a tissue signature which supplements the periodicity, and can be measured even when the periodicity is not defined.

References

[1] U.R.Abeyratne, A.P.Petropulu and J.M.Reid, "On Modeling the Tissue Response From Ultrasonic B-Scan Images", *IEEE Tr. on Medical Imaging*, vol. 15, no. 4, pp. 479-490, August 1996.

[2] U. R. Abeyratne, A. P. Petropulu, J. M. Reid, "Higher Order Spectra Based Deconvolution of Ultrasound Images", *IEEE Trans. on Ultr., Ferroelec. and Frequency Control*, vol. 42, no.6, pp. 1064-1075, 1995.

[3] J.C. Bamber, "Theoretical Modelling of the Acoustic Scattering Structure of Human Liver", *Acoustic Letters*, Vol.3, No. 6, pp. 114-119, 1979.

[4] J.-F. Chen, J.A. Zagzebski, and E.L. Madsen, "Non-Gaussian Vs. Non-Rayleigh Statistical Properties of Ultrasound Echo Signals", *IEEE Tr. on Ultrasonics, Ferroelec., and Frequency Control*, vol. 41, no. 4, pp. 435-440, 1994.

[5] F.S. Cohen and G. Georgiou, "Detecting and Estimating Structure Regularity of Soft Tissue Organs From Ultrasound Images", *International Conference on Image Processing*, Washington D.C., October, 1995.

[6] F.S. Cohen, "Modeling of Ultrasound Speckle with Application in Flaw Detection in Metals", *IEEE Tr. on Signal Processing*, vol. 40, no. 3, pp. 624-632, 1992.

[7] K.D. Donohue, J.M. Bressler, T. Varghese and N.M. Bilgutay, "Spectral Correlation in Ultrasonic Pulse-echo Signal Processing", *IEEE Tr. on Ultrasonics, Ferroelec., and Frequency Control*, vol. 40, no. 4, pp.330-337, 1993.

[8] L. L. Fellingham and F.G. Sommer, "Ultrasonic Characterization of Tissue Structure in the *In Vivo* Human Liver and Spleen", *IEEE Trans. on Sonics and Ultr.*, vol. SU-31, No-4, pp 418-428, 1984.

[9] J.C. Gore, S. Leeman, C. Metreweli, N.J. Plessner, K. Willson, "Dynamic Autocorrelation Analysis of A-scans *in-vivo*", *Ultrasonic Tissue Characterization II*, M. Linzer (ed.), National Bureau of Standards, Washington, D.C, pp. 275-280, 1979.

[10] J.F. Greenleaf and C.M. Sehgal, **Biologic System Evaluation with Ultrasound**, Springer-Verlag, New York, 1992.

[11] M.F. Insana, R.F. Wagner, B.S. Garra, D.G. Brown and T.H. Shawker, "Analysis of ultrasound image texture via generalized Rician statistics", *Optical Engineering*, Vol. 25, No. 6, pp. 743-748, 1986.

[12] L. Landini and L. Verrazzani, " Spectral characterization of tissue microstructure by ultrasound: A stochastic approach", *IEEE Tr. Ultra.,Ferroelec. and Freq. Control*, vol. 37, pp 448-456, 1990.

[13] F.L. Lizzi, M. Greenebaum, E.J. Feleppa, M. Elbaum, D.J. Coleman, "Theoretical Framework for Spectrum Analysis in Ultrasonic Tissue Characterization", *J. Acoust. Soc. Am.*, 73(4), pp. 1366-1373, 1983.

[14] F.L. Lizzi, M. Ostromogilsky, E.J. Feleppa, M. Rorke, M.M. Yaremko, "Relationship of Ultrasonic Spectral Parameters to Features of Tissue Microstructure", *IEEE Tr. on Ultr. Ferroelec. and Frequency Control*, 33(3), pp 319-329, 1986.

[15] C. L. Nikias and A.P. Petropulu, **Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework**, Prentice Hall Incorporated, Oppenheim Series in Signal Processing, 1993.

[16] A.V.Oppenheim and R.W. Schafer, **Discrete Time Signal Processing**, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1989.

[17] J.M.Reid, "The Measurement of Scattering", in *Tissue Characterization with Ultrasound*, CRC Press, Boca Raton, Florida, pp.81-114, 1986.

[18] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli and B. B. Goldberg, "Use of Non-Rayleigh Statistics for the Identification of Tumors in Ultrasonic B-Scans of the Breast", *IEEE Trans. on Medical Imaging*, Vol. 12, pp 687-692, 1993.

[19] R. H. Tuthill, R. H. Sperry and K. J. Parker, "Deviations From Rayleigh Statistics in Ultrasound Speckle", *Ultrasonic Imaging*, 10, pp 81-89, 1988.

[20] L. Weng, J.M. Reid, P.M. Shanker, K. Soetanto and X.-M. Lu, "Nonuniform phase distribution in ultrasound speckle analysis-Part I: Background and experimental demonstration", *IEEE Tr. Ultrasonics, Ferroelec., and Frequency Control*, Vol. 39, pp 352-359, 1992.

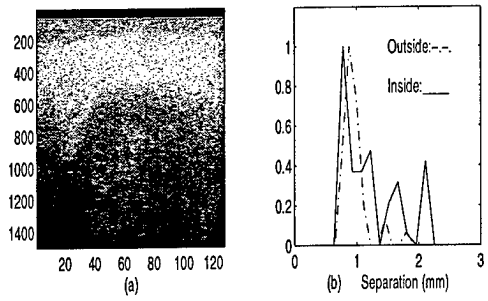


Figure 1

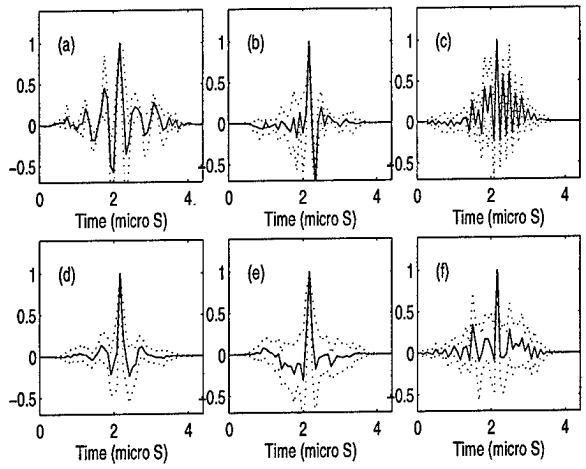


Figure 2

Signal-Dependent Film Grain Noise Removal and Generation Based on Higher-Order Statistics

Jacky Chun Kit Yan and Dimitrios Hatzinakos
University of Toronto, Department of Electrical and Computer Engineering
10 King's College Road, Toronto, Ontario, M5S-3G4, Canada
dimitris@comm.toronto.edu

Abstract

In this paper, we propose a new noise filtering scheme that is based on higher-order statistics (H.O.S.) for photographic images corrupted by signal-dependent film grain noise. In addition, reliable estimation of the noise parameter using H.O.S. is proposed. This parameter estimation technique can be used to generate film grain noise which has applications in motion picture and television productions. Simulation results show that the proposed filter perform better than existing methods which are based on second-order statistics.

1. Introduction

Noise suppression is a common application in image processing. Traditionally the noise process is taken to be additive white Gaussian and statistically independent of the signal. However, in the case of photographic images, it is well known that they contain film grain noise that is signal-dependent, that is, the noise statistics depend on the signal. The noise model describing film grain noise and zero mean Gaussian measurement noise w has the form [1]:

$$r(x,y) = s(x,y) + ks^p(x,y)n(x,y) + w(x,y) \quad (1)$$

where s is the noiseless image measured in density, k is the scanning constant, p is an exponent that depends on film, n is a Gaussian noise with zero mean and unit variance. Because conventional filtering techniques assume a different noise model, they do not perform well in this case.

Existing methods for suppressing signal-dependent film grain noise include modifications of standard techniques designed for additive noise. Examples are the Wiener filter [1] and statistical estimators [2] for the above noise model. These techniques yield better performance than those assuming a false signal-independent noise. However, the above methods assume the parameter k is known *a prior*. Moreover, due to the nonlinearity in the noise model, expressions for the statistical estimators (MMSE and MAP) have a complicated form even in the special case of Gaussian image s, and involve solving a polynomial equation and numerical integration at every pixel.

In this paper, reliable estimation of noise model parameter using higher-order statistics (H.O.S.) is considered. In this respect a new filter (Wiener type) based on H.O.S. is proposed. Also, realistic film grain noise generation, which has applications in television and motion picture productions, becomes possible. Because measurement noise is Gaussian, higher-order statistics of the observed image r would contain contributions from the non-Gaussian image s and signal-dependent noise only, which leads to better parameter estimation. Furthermore, since photographic images are highly non-Gaussian and film grain noise is nonlinearly related to the original image, a lot more information can be extracted from their higher-order statistics [3]. Filtering schemes based on H.O.S. can give better performance.

2. Design of Higher-Order Statistics Based Filter

Assuming the proposed filter $h(x,y)$ be a finite impulse response (FIR) filter with a support region of

$$h(x,y) \neq 0 \quad for \quad a \leq x \leq b, c \leq y \leq d, \quad (2)$$

the filter coefficients $h(x,y)$ can be solved by minimizing a higher-order statistics criterion that is an extension of the mean square error (MSE) criterion used in the correlation based Wiener filter. Let the error signal

0-8186-8005-9/97 $10.00 © 1997 IEEE

$e(x,y)$ be defined as:

$$e(x,y) = s(x,y) - \sum_{i=a}^{b}\sum_{j=c}^{d} h(i,j)r(x-i,y-j). \quad (3)$$

the proposed filter $h(x,y)$ is designed by minimizing the following criterion [4]:

$$J_C^M(h) = \left\{ \sum_{\alpha=-\infty}^{\infty}\sum_{\beta=-\infty}^{\infty} Cum^M(e_{xy}, e_{xy}, r_{x-\alpha,y-\beta}, \ldots, r_{x-\alpha,y-\beta}) \right\}^2 \geq J_C^M(\hat{h}) \quad (4)$$

with $e_{xy} = e(x,y)$ and $\hat{h}(x,y)$ being the optimum filter based on the cumulant based criterion J_C^M.

3. Cumulant-Based Wiener-Hopf Equation

To compute the optimum filter coefficients, we can extend the correlation based Wiener-Hopf equation and the orthogonality principle to higher-order statistics. By using the idea described in [4], let $\tilde{h}(x,y)$ be the filter satisfying the following cumulant based orthogonality condition:

$$\sum_{\alpha=-\infty}^{\infty}\sum_{\beta=-\infty}^{\infty} Cum^M(e_{xy}, r_{x-i,y-j}, r_{x-\alpha,y-\beta}, \ldots, r_{x-\alpha,y-\beta}) = 0 \quad (5)$$

Then it can be shown that $\tilde{h}$ is the optimum filter associated with the criterion J_C^M. To derive the cumulant based Wiener-Hopf equation, we start with the orthogonality condition (Eq. (5)), and substitute expression for $e(x,y)$ into Eq. (5) to give:

$$\Rightarrow C_{sr}(p,q) = \sum_{i=a}^{b}\sum_{j=c}^{d} h(i,j)C_{rr}(p-i,q-j) \quad (6)$$

where C_{sr} and C_{rr} are defined as

$$C_{sr}(i,j) = \sum_{\alpha=-\infty}^{\infty}\sum_{\beta=-\infty}^{\infty} Cum^M(s_{xy}, r_{x-i,y-j}, r_{x-\alpha,y-\beta}, \ldots, r_{x-\alpha,y-\beta}) \quad (7)$$

$$C_{rr}(i,j) = \sum_{\alpha=-\infty}^{\infty}\sum_{\beta=-\infty}^{\infty} Cum^M(r_{xy}, r_{x-i,y-j}, r_{x-\alpha,y-\beta}, \ldots, r_{x-\alpha,y-\beta}) \quad (8)$$

By substituting estimates of C_{rr} and C_{sr} and forming a linear system of equations, filter coefficients $h(i,j)$ can be solved. Note that the above derivation can be applied similarly to moments, i.e., moment based criterion leading to moment based orthogonality condition and moment based Wiener-Hopf equation.

4. Estimation of Higher-Order Statistics

The use of Eq. (6) requires that C_{sr} and C_{rr} are known. In practice, higher-order cumulants are estimated by replacing the expectation operator by sample averaging over the data. Estimating C_{rr} is not difficult since we have access to the observed signal. However, C_{sr} is not easy to obtain unless we know the signal exactly, which is impossible. Thus we need to determine their relationships with the original signal statistics C_{ss}, which is assumed known.

By substituting the noise model (Eq. (1)) into the definitions of C_{rr} and C_{sr}, we have for $M = 3$ and $p = 0.5$:

$$C_{rr}(i,j) = C_{sr}(i,j) + k^2(R_s(i,j) - m_s^2) + k^2\sum_{\alpha}\sum_{\beta}(R_s(\alpha,\beta) - m_s^2)\delta(i,j) \quad (9)$$

and

$$C_{sr}(i,j) = \sum_{\alpha}\sum_{\beta} Cum(s_{xy}, s_{x-i,y-j}, s_{x-\alpha,y-\beta}) + k^2(R_s(i,j) - m_s^2) + E[s(x,y) - m_s]\sum_{\alpha}\sum_{\beta} R_w(i-\alpha, j-\beta) \quad (10)$$

where R_s and m_s are the autocorrelation and the mean of the signal s respectively:

$$R_s(i,j) = E[s(x,y)s(x-i,y-j)] \quad (11)$$

5. Parameter Estimation

The calculations of Eqs. (9) and (10) require that the constant k be known. However, when this information is not available, we must estimate the constant from the observed image statistics and the *a prior* ideal image statistics. In the case of $p = 0.5$, the variance, skewness, and kurtosis of the received image are related to that of the original image by the following equations:

$$\sigma_r^2 = \sigma_s^2 + k^2 E[s] + \sigma_w^2 \quad (12)$$

$$c_3^r = c_3^s + 3k^2\sigma_s^2 \quad (13)$$

$$c_4^r = c_4^s + 6k^2c_3^r - 15k^4\sigma_s^2 \quad (14)$$

The value of k can then be solved by substituting the statistics of the observed image (which can be estimated) and the *a prior* image statistics (σ_s^2, c_s^3, and c_s^4) into any of the above equations. Note from the above equations that the use of higher-order statistics in the presence of Gaussian measurement noise leads to better estimation of k, as cumulants of Gaussian noise are identically zero.

6. Film Grain Noise Generation

As outlined in Introduction, film grain noise generation has applications in television and motion picture productions since digitized film images, video images and computer generated images are routinely combined into one frame. In this process artificial noise is added to the video and computer generated images to match the grain pattern of the film. To generate the right amount of artificial film grain noise, the noise parameter k must be known.

Using the above method to estimate the noise parameter k would require statistics of the corrupted and ideal images. When k is solved, generate and add a noise image to the computer generated image according to Eq. (1), but without the measurement noise w. In cases where the *a prior* ideal image statistics are not known, an approximate solution was proposed in [5] which uses statistics of a filtered image as the ideal image statistics. The observed image r is processed using a sub-optimal filtering technique that requires no knowledge of the noise statistics.

7. Simulation Results

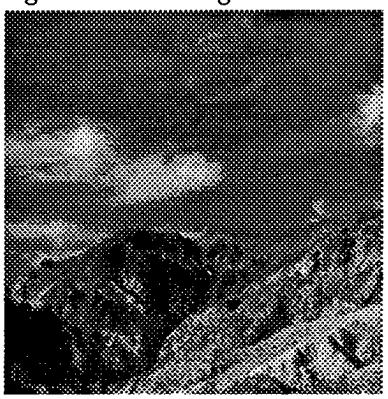
In this section, the above proposed methods in estimation of the parameter k, and signal-dependent film grain noise removal based on H.O.S. are applied. Two test images of size 256 x 256 were used: Lenna and Mountain. Test image 'Mountain' is shown in Fig. 1.

To test the validity of parameter estimation of k, a number of simulations were performed for the following two cases:

- signal-dependent noise only
- mixture of signal-dependent / signal-independent noise

Signal-dependent film grain noise and Gaussian measurement noise are added to the image 'Lenna'. Sample

Figure 1. Test image: 'Mountain'.



cumulants are calculated using the following relationships:

$$c_1^r = m_1^r \quad (15)$$

$$c_2^r = \sigma_r^2 = m_2^r - (m_1^r)^2 \quad (16)$$

$$c_3^r = m_3^r - 3m_1^r m_2^r + 2(m_1^r)^2 \quad (17)$$

$$c_4^r = m_4^r - 4m_1^r m_3^r - 3(m_2^r)^2 + 12(m_1^r)^2 m_2^r - 6(m_1^r)^4 \quad (18)$$

The quantities m_1^r, m_2^r, m_3^r, and m_4^r are estimated from an M x N image using sample averaging. The signal statistics σ_s^2, c_3^s, and c_4^s of the image 'Lenna' are known *a prior* and are used to solve for k. The parameter p was fixed to be 0.5 throughout the experiments since this is typical for a variety of film stocks. Because the variance of the signal-independent noise is assumed unknown, k is solved using Eq. (12) with zero measurement noise variance. A value of $k = 0.1$ was selected for moderate noise corruption. The value of k determines the degree of degradation, as can be seen from the variance of the signal-dependent noise term:

$$\sigma_{noise}^2 = k^2 E[s] \quad (19)$$

Fifty independent runs were performed, with the results summarized below.

Table 1. Estimation of k with true value k=0.1.

	estimated k (mean± standard deviation):		
σ_w^2	2nd order	3rd order	4th order
0	0.0999±0.0015	0.0999±0.0024	0.0998±0.0033
0.05	0.1299±0.0014	0.1003±0.0030	0.1006±0.0041
0.10	0.1939±0.0013	0.1013±0.0039	0.1009±0.0072
0.15	0.2681±0.0015	0.1004±0.0058	0.1007±0.0082
0.20	0.3465±0.0014	0.0990±0.0096	0.0984±0.0159

The advantage of using H.O.S. in estimation is evident from the tables. Second-order statistics results in

high bias and low variance, whereas estimation using H.O.S. has low bias and higher variance. It is clear that estimating k using H.O.S. is better than using second-order statistics in the presence of Gaussian measurement noise.

For noise filtering, again two different cases were investigated:

- signal-dependent noise only
- mixture of signal-dependent/signal-independent noise

A value of $k = 0.1$ was used. For a mixture of noise, variance of measurement noise σ_w^2 was chosen to be 0.005.

The criteria used in evaluating the performance were 1) signal-to-noise ratio (SNR), 2) mean absolute error (MAE), and 3) mean square error (MSE) which is similar to SNR. These are defined below for an image of size M x N:

$$\mathrm{SNR} = 10 \log_{10} \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} s^2(x,y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{s}(x,y) - s(x,y)]^2} \quad (20)$$

$$\mathrm{MAE} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{s}(x,y) - s(x,y)] \quad (21)$$

$$\mathrm{MSE} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{s}(x,y) - s(x,y)]^2 \quad (22)$$

where s and $\hat{s}$ are the ideal and estimated images, respectively. A filter size of 3 x 3 was chosen. Wiener filter designed based on the usual correlation based criterion [1] (designated M_2) and the two proposed higher-order statistics based criteria (third-order moment M_3 and third-order cumulant C_3) are compared.

A number of observations can be made:

- On average higher-order statistics based filters achieved better SNR than the correlation based filter by 1 dB. This improvement may be due to the fact that more information about the image statistics was utilized.
- The cumulant based filter performed as good as the higher-order moment and correlation based filters in the case of signal-dependent noise only. However, for a mixture of noise, the cumulant based filter is not suitable. By examining the cumulant based Wiener-Hopf equation in Eq. (6):

$$C_{sr}(p,q) = \sum_{i=a}^{b} \sum_{j=c}^{d} h(i,j) C_{rr}(p-i, q-j) \quad (23)$$

we have for a mixture of noise:

$$C_{sr} \text{ (} \textit{mixture of noise} \text{)} \approx C_{sr} \text{ (} \textit{film grain noise only} \text{)} \quad (24)$$

and

$$C_{rr} \text{ (} \textit{mixture of noise} \text{)} \approx C_{rr} \text{ (} \textit{film grain noise only} \text{)} \quad (25)$$

thus the cumulant based criterion cannot not recognize the Gaussian measurement noise. The filter performed as if only film grain noise is present.

- Performance of cumulant based filter depends heavily on the properties of the image. For third-order cumulant, if the distribution of an image is close to Gaussian or is symmetric, then it is not appropriate to use the cumulant based filter. Table 5 shows the statistics of the two images. It can be observed that 'Mountain' has the lowest skewness, thus third-order cumulant based filter did not perform well. If fourth-order cumulant is used, the filter should perform well in the case of film grain noise only, as indicated in Table 4.

To test the noise generation procedure, the image 'Lenna' was used for noise generation. Corrupted image was filtered using the method described in [5]. Then k was computed using fourth-order statistics of the two images. Although k can be solved by matching their variances, it was found that variance of the filtered image is lower than that of the corrupted image because edges are blurred to some extent. Thus using variance to obtain k would lead to over-estimation, and the final image would be too noisy. To compare the noise level of the original corrupted and the final image, SNR was used. For the original corrupted image the signal power is the ideal signal power, whereas the signal power of the final image is that of the filtered image. It can be seen from Table 6 that the noise level in two images (noise-added and original corrupted images) are about the same.

8. Conclusions

This paper presents a H.O.S. based filter for filtering images corrupted by signal-dependent film grain noise. In addition, estimation of noise parameter using H.O.S. is proposed and successfully applied in film grain noise generation. Simulation results show that the performance of this filter is better than that of filter based on second-order statistics, and parameter estimation using H.O.S. is more reliable in the presence of Gaussian measurement noise.

Table 2. Test image 'Lenna' with signal-dependent noise only (k=0.1).

	SNR (dB)	MAE	MSE
unfiltered	16.7591	4.6375e-2	3.6820e-3
M_2	19.8458	3.1849e-2	1.8089e-3
M_3	19.9902	3.1212e-2	1.7498e-3
C_3	20.1143	3.0150e-2	1.7005e-3

Table 3. Test image 'Lenna' with mixture of signal-dependent noise (k=0.1) and measurement noise ($\sigma_w^2 = 0.005$).

	SNR (dB)	MAE	MSE
unfiltered	13.0270	7.3938e-2	8.6957e-3
M_2	17.7796	4.1604e-2	2.9110e-3
M_3	17.8464	4.1266e-2	2.8666e-3
C_3	13.7575	6.7819e-2	7.3494e-3

Table 4. Test image 'Mountain' with signal-dependent noise only (k=0.1).

	SNR (dB)	MAE	MSE
unfiltered	15.9772	4.5280e-2	3.4005e-3
M_2	20.9042	2.5644e-2	1.0936e-3
M_3	21.6888	2.1905e-2	9.1283e-4
M_4	21.5968	2.2600e-2	9.3237e-4
C_3	20.7989	2.3356e-2	1.1204e-3
C_4	21.7648	2.2393e-2	8.9398e-4

Table 5. Image statistics of various test images.

	Image Statistics		
Image	mean	variance	skewness
Lenna	3.6363e-1	4.2353e-2	7.2266e-3
Mountain	3.3807e-1	2.0373e-2	1.0025e-3

Table 6. SNR of noise-generated images using different image statistics (original SNR is 16.7620 dB).

statistics	SNR (dB)
second order	16.0929
third order	15.8685
fourth order	16.9721

Figure 2. Corrupted 'Mountain' with $k = 0.1$.

Figure 3. Filtered 'Mountain' using M_2 criterion.

Figure 4. Filtered 'Mountain' using C_4 criterion.

References

[1] J. F. Walkup and R. C. Choens. Image processing in signal-dependent noise. *Opt. Eng.*, 13:258–266, 1974.

[2] J. F. Walkup G. K. Froehlich and R. B. Asher. Optimal estimation in signal-dependent noise. *J. Opt. Soc. Am.*, 68:1665–1672, 1979.

[3] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework.* Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[4] C. C. Feng and C. Y. Chi. Design of wiener filters using a cumulant based mse criterion. *Signal Processing*, 54:23–48, 1996.

[5] J. C. K. Yan. Statistical methods for film grain noise removal and generation. *M. A. Sc. Thesis, University of Toronto*, 1997.

Blind Identification Methods Applied to Electricité de France's Civil Works and Power Plants Monitoring

G. D'URSO, P. PRIEUR, C. VINCENT
Electricité de France/Etudes et Recherches
6, Quai Watier, 78401 Chatou, France
phone : 33 - 1 30 87 77 52 , fax : 33 - 1 30 87 84 34
Guy.Durso, Pascale.Prieur, Claude.Vincent@der.edfgdf.fr

Abstract

In this article, we would like to present results obtained on industrial data with source separation techniques in instantaneous mix. One introduces three applications developed to perform the monitoring of Electricité de France civil works and power plants.

The first application concerns the monitoring of nuclear power plants. Each internal component generates specific vibration modes and "neutron noise" is a combinaison of all modes generated. The aim of this study is to separate such independent vibration modes.

The second application concerns the dams supervision : it consists in separating the various types of motion of a dam according to their physical origin.

The third application concerns non destructive testing on steam generators in nuclear power plants. The aim is to reduce the flattening noise. The classical methods operate only when a noise reference is available. We propose to use the multi-sensor approach with the blind separation methods (the noise reference is not necessary).

Considering the specifications of the signals, we get better performance using a two-order statistics algorithm than a higher-order statistics algorithm.

1. Monitoring of nuclear power plants

1.1 Introduction

Nuclear reactors are composed of many elements which vibrate and become deformed. An efficient monitoring could be achieved if we are able to separate the vibration signature of the thermal shield from that of the core barrel. These signatures are characterized by vibratory modes : « mode 1 » (around 7.5 Hz) related to the swinging of the core barrel and « mode 2 » (around 11.5 Hz) due to the deformation of the thermal shield. In many cases, however, because of its amplitude or its frequency deviation (it can shift from 7.5 to 11.5 Hz), mode 1 totally masks mode 2.

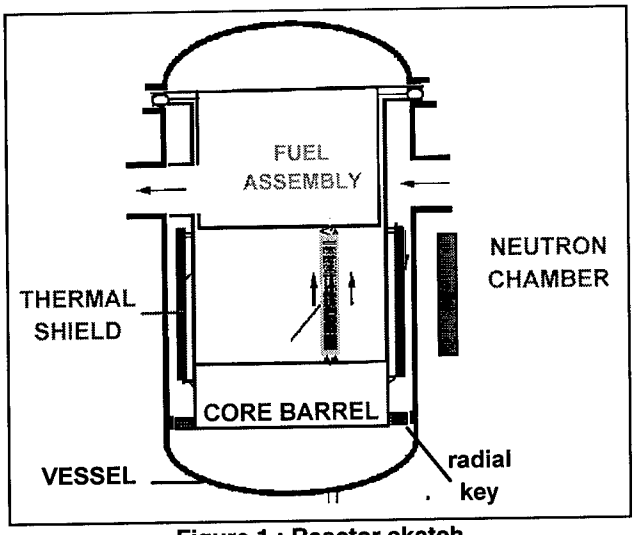


Figure 1 : Reactor sketch

We are in the presence here of several gaussian sources vibratory with different, and therefore independent, physical origins. We have four sensors positioned at 90° angles around the core barrel. Each delivers a theoretically linear combination of various vibratory sources. The medium of propagation here cannot be modeled, due to its complexity.

1.2. The used techniques

Because many vibratory diagnostic methods rely on spectra, the objective is to estimate the spectrum of each source.

We present two methods : interspectral matrices which enable estimation of the spectral density of each source and the SOBI algorithm (Second Order Blind Identification), which performs a blind separation of the temporal signatures of each source.

- Interspectral matrices are used to estimate source spectra. By diagonalizing the matrix at one frequency,

0-8186-8005-9/97 $10.00 © 1997 IEEE

we estimate the power spectral density (PSD) values of the sources (equal to the eigenvalues) [1]. A simple classification of eigenvalues in descending order does not show the evolution in source power. It is not able to deal with the crossing of sources spectra.

We therefore propose an alternative method, based on phase and amplitude continuity. We assume that the eigenvalues and the signal sub-space of each source are continuous in frequency. We then use an algorithm which minimizes the discontinuities of the amplitude of the eigenvalues and the eigenvector phase to find the ordered values of the sources in frequency [6].

- To separate temporal signals, we then use the SOBI algorithm [4]. The method exploits the time coherence of the source signals. The algorithm is based on a 'joint diagonolisation' of correlation matrices. You will find further information in [5].

1.3 Experimental results

These two methods have now been applied to real data. The first (figure 2, above) gives successful results, and the PSDs of the physical phenomena are well estimated. With the second method, the SOBI algorithm gives a good estimation of real signals. The PSDs of the separated signals are consistent with those found using the first method (figure 2,below).

An examination of PSDs reveals several vibratory modes :

- the two highest eigenvalues show a major contribution around 7.5 Hz (core barrel mode 1), and are related to swinging of the core barrel ;
- the two lowest represent one source related to the thermal shield (mode 2, at 11.5 Hz) ;
- an interfering source which most often presents two modes: one around 8 Hz and the second around 10 Hz.

1.4 Conclusion

Used algorithms are sub-optimal because :

- for the first one, the eigendecomposition of the interspectral matrix has no sole solution. The solution is known with exception of a unitary matrix (rotation matrix). In our case, we choose a classic seismic phase constraint [8] [9] : we decide that the first component of all eigenvectors will have a nil phase. But such an approach limits significantly the general nature of this algorithm.
- the second algorithm makes the hypothesis hat an instantaneous linear mixture, even though the mix is certainly convolutive.

However, for gaussian sources, this method restitute mode 2 correctly, and eliminates the "parasite modes" found with traditional methods.

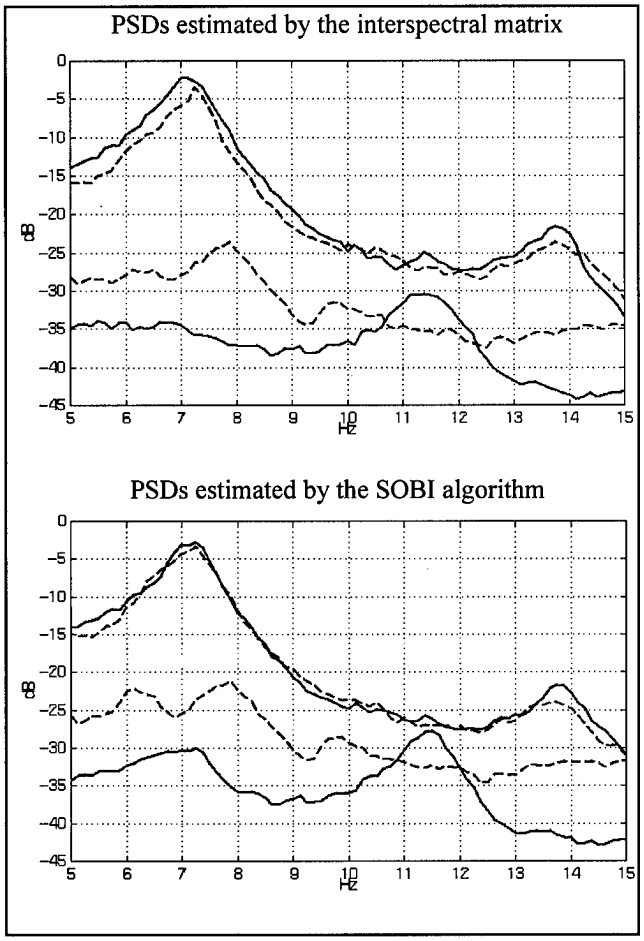


Figure 2: Power spectral density (PSD) values of the vibratory sources

2. Dams supervision

2.1. Introduction

In spite of their apparent immobility, dams move in working conditions. Civil works monitoring allows a best estimation on their condition. And so, we should distinguish between the displacement of the dam resulting from solicitations and these which denote a damage of the work.

Dams are equipped with sensors of radial and tangential displacement, flow sensors, strain sensors .... The registered signals contain, on the one hand information about the ageing of the dam, on the other hand the response of the dam due to external solicitations. There are two kinds of solicitations :

- mechanical solicitations corresponding to the upward thrust generated by the water retained ;

• thermal solicitations due to temperature changes.

Our study is devoted to displacement measures.

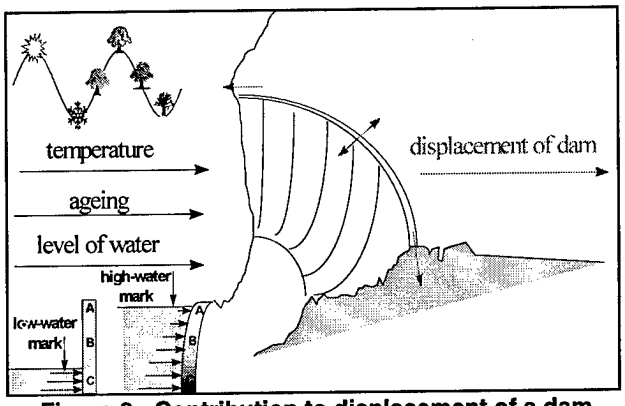


Figure 3 : Contribution to displacement of a dam

We suppose that the displacement of a dam is an additive mix of the responses of the dam to these three phenomena (supposed independent and non gaussian) to which a measure noise is added. The measures of displacement are very few (800 samples for around 20 years in the best of cases) and their sampling is non regular (being able to vary from 3 to 20 days) and non evenly distributed. We can note that the variations of both solicitations are very different. If mechanical solicitation varies every day, the thermal solicitation, for its part, presents mainly yearly and monthly periods.

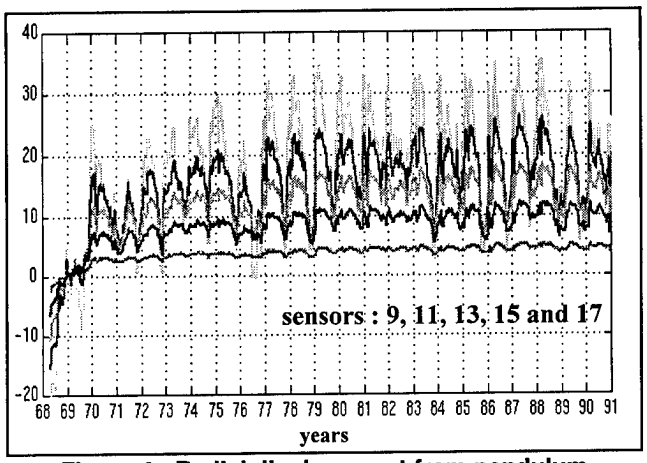


Figure 4 : Radial displacement from pendulum

2.2. Methods used

One uses a source separation algorithm that aims to find the various contributions from signals observed on a network of sensors.

As sources are non gaussian, we could have used an algorithm based on higher-order statistics. But we have very few data records at our disposal (from 150 to 800 from case to case) we chose the Second Order Blind Identification algorithm : SOBI. Actually, a good estimation of correlation requires less samples than estimation of higher order cumulants.

We would like to insist on the fact that the hypothesis of instantaneous mix allows to use classic sources separation methods, even for non regularly sampled signals. SOBI uses the property of independence of sources which technically finds expression in a diagonal structure of the correlation matrices. Estimation of correlation matrices is certainly biased, but their diagonal structure is preserved, so that the algorithm keeps its efficiency.

2.3. Application to displacement of concrete dam

Using five sensors (one above the others straight up the dam) allows to extract three solicitations, which constitutes a good result.

Everyday temperature measurements and of the level of the water retained in reservoir that we got for this dam, will allow to check the results obtained.

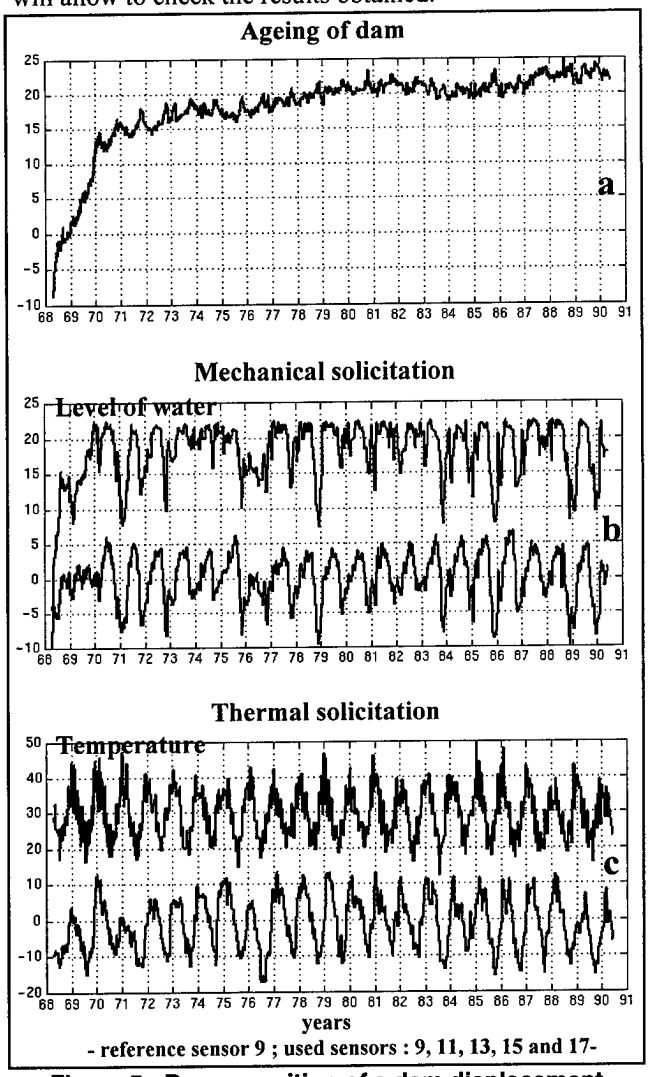


Figure 5 : Decomposition of a dam displacement

Figure 5 presents the different contributions comparatively to the level of the water retained and the temperature.

reference a : displacement due to ageing

reference b : displacement due to mechanical contribution below and level of water retained above

reference c : displacement due to thermal contribution below and temperature above

These curves highlight a cause and effect relationship between constraints (temperature and level of water) and their contributions upon the dam : mechanical and thermal displacement. Besides, the effect of time (ageing) allows to follow the evolution of the dam when it moves downstream.

3. Testing on steam generators

This application concerns the non destructive testing on steam generators of nuclear power plants. The aim is to detect the appearance of cracks superficially and in depth on the tubes.

Because of the manufacturing process, data records contain mainly noise. One proposes to use an array processing approach to restore the signals.

3.1. Introduction

Nuclear power plants steam generators are tested by eddy current probes. These probes are sensitive to every unevenness of the material and detect the potential presence of defects that corrupt the evenness (of the tube).

Now the surface of these tubes (manufactured by cold-rolling) comprises undulations (normal unevenness) which generate a so called flattening noise which can perturb the monitoring, hiding defects. It is essential to eliminate this noise in order to make the monitoring possible.

Several monitoring complex measures work simultaneously. The most advanced of the methods developed up to now -particularly the adaptative filtering- are based on the utilization of noise reference [2]. In practice, even if some measures contain mustly noise, we have no noise reference alone and that sets limits to this class of methods performance.

One proposes to use here sources separation methods. One considers that several types of unevenness (defect, flattening noise) are present in the steam generators and are mixed in data record. These irregularities can be assimilated to sources and we can suppose that they are statistically independent of each other. The array processing, thanks to separation methods, allows to separate the contributions of unevenness and the flattening noise in the monitoring signal.

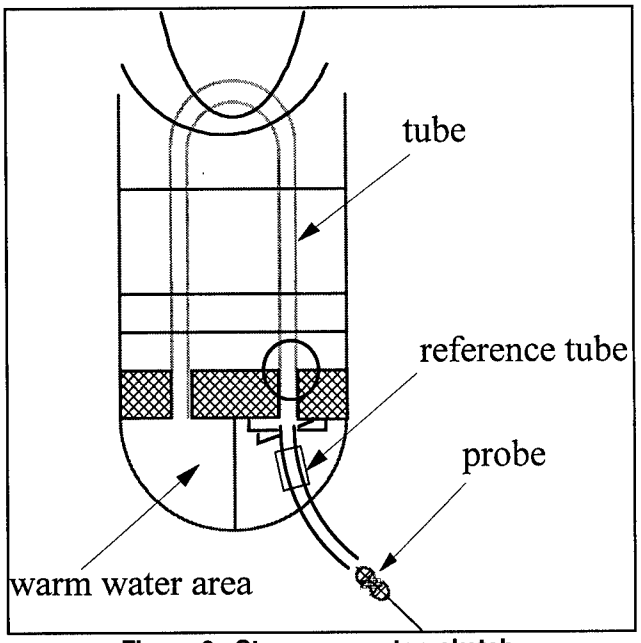


Figure 6 : Steam generator sketch

3.2. Separation algorithm

The test with algorithms based on higher-order statistics [3] [7] and on two-order [4] was successful.

The results presented in the following paragraph have been obtained with Cardoso's algorithm [7].

3.3 Experimental results

The results show an improvement of the signal to noise ratio (SNR). Some hidden defects thus be detected (fig. 7). We used the real part of two complex measures (on the left hand). The treatment separates the effective signal and the noise (on the rigth hand).... The defect, which was hidden in the raw signal is detected.

To sum up, it is possible to eliminate the flattening noise by sources separation. This treatment can plainly improve the ratio signal to noise and makes the small range defects visible.

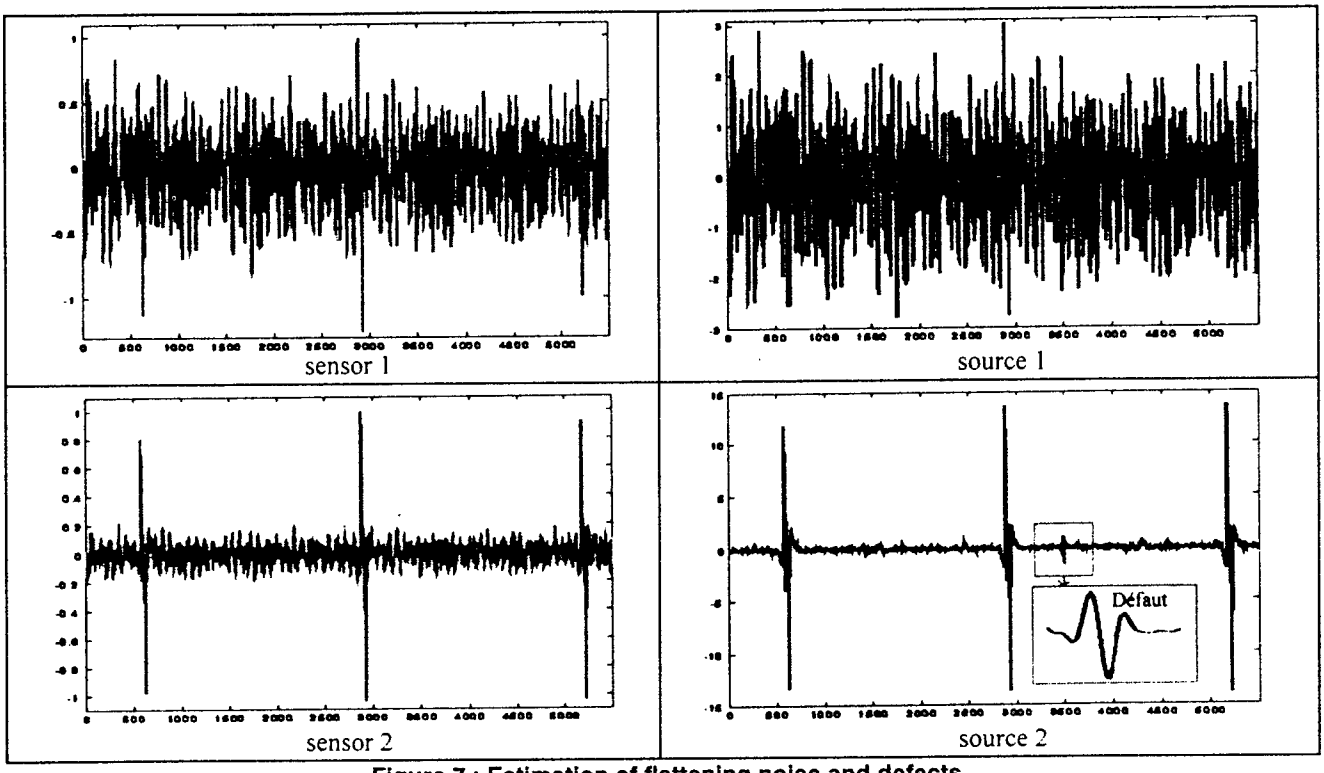


Figure 7 : Estimation of flattening noise and defects

4. Conclusion

Mixes appearing during the non destructive testing of E.D.F.'s civil works are more or less convolutive. However, we tried to apply to them source separation techniques conceived for instantaneous mixes. The test with three different methods yields conclusive results.

The SOBI algorithm was more suitable for the data records than the higher-order statistics methods. Actually, the sources had a temporal coherency and different spectra. Furthermore, SOBI allows to treat gaussian sources (first example : vibratory mode separation) and short signals (a few hundred points for the dam displacement measures). It keeps a good reliability because it is based on two order statistics. The higher-order statistics allow to enlarge the signal's representation space. We find the same idea in the SOBI algorithm, which uses the temporal dimension (sources coherence) to enlarge the representation space.

At the present time, a study about new source separation methods suitable for convolutive mixes is in progress.The selected methods are based on higher-order statistics. We plan to test them and to estimate their performance for our applications.

References

[1] Bienvenu & Kopp, "Optimality of high resolution processing using he eigensystem approach", IEEE Trans. on ASSP, No. 31, pp. 1235-1248, 1983.

[2] Y. Poupeau, "Amélioration du rapport S/B des signaux de contrôle par courants de Foucault", Thèse de doctorat université P. et M. Curie, 1989

[3] Comon, "Independent component analysis and blind identification", Traitement du signal, GRETSI, vol. 7, No. 5, pp. 435-450, 1990.

[4] A. Belouchrani and K. Abed Meraim, "Séparation aveugle au second ordre de sources temporellement corrélées", Proc. GRETSI, Juan-les-Pins, sept. 1993.

[5] A. Belouchrani, K. Abed Meraim, J. F. Cardoso and E. Moulines, "Asymptotic Performance of Second Order Blind Separation", IEEE Proc. ICASSP, Adelaide, Australia, april 1994.

[6] G. d'Urso & L. Cai, "Source separation methods applied to reactor monotoring", ATHOS, worshop on HOS, Edinburgh, 1994.

[7] J. F. Cardoso and A Souloumiac, "An efficient technique for the blind separation of complex sources", in Proc. IEEE SP, worshop on HOS, lake Tahoe, 1993

[8] Glangeaud & Gavin & Mari, "The use of the eigenvectors of the matrix, seismic application", Traitement du signal, vol. 10 No. 2, pp85-104, 1994.

[9] N. Thirion, J. L. Lacoume and J. Mars, "Resolving power of spectral matrix filtering : a discussion on the links steering vectors/eigenvectors", in Proc. IEEE SP, worshop on SSAP, Corfu, june 1996.

MPL: Nonstationary Signals, Nonlinear Systems

Chair: Chong-Yung Chi
National Tsing Hua University, Taiwan

Time-Frequency Representations and Their Structure *

Benjamin Friedlander
Dept. of Elec. & Comp. Eng.
University of California
Davis, CA 95616, USA.

Louis L. Scharf
Dept. Elec. & Comp. Eng
University of Colorado
Boulder, CO 80309, USA.

Abstract

This paper explores the structure of time-frequency representations (TFR) in a discrete time setting. We define a proper TFR to be a function of the signal that has natural time- and frequency-shift properties. We then derive the basic structure of a proper TFR and argue for a quadratic TFR as the simplest form of a proper TFR.

1. Introduction

Time-Frequency Representations (TFRs) have been the subject of renewed interest in recent years. The variety of techniques which have been studied makes it somewhat difficult to provide a complete classification. It is apparent that the largest number of publications focuses on a class of TFRs which are quadratic (or bilinear) functions of the signal, commonly called "the Cohen class" [1]. Various non-quadratic techniques, such as Cohen class TFRs with data dependent kernels (which make the TFR non-quadratic) [2], data-adaptive representations [3, 4], and positive time-frequency distributions (TFDs) [5]–[6] have also been proposed.

When looking at this variety of TFRs the question naturally arises: what is a "proper" time frequency representation? Or should we accept *any* function of time and frequency constructed from the signal, as a valid TFR? If not, then what are the properties which a function should have, in order to qualify it as a TFR, and what is the minimal set of properties required to uniquely determine its form? These questions seem to have received relatively little attention in the literature.

In this paper we attempt to answer these questions by defining a proper TFR to be a function of the signal which has natural time-shift, frequency shift, scaling and non-negativity properties. We are able to show that these properties impose a particular structure on any proper TFR. In effect, we are able to derive a "representation theorem" for proper TFRs, thereby extending the philosophy of [7] from time-invariant spectrum analysis to time-varying spectrum analysis.

The organization of the paper is as follows. We start in section 2 with a general characterization of proper TFRs. In section 3 we describe the time and frequency response of a proper TFR. Section 4 presents the *quadratic* proper TFR. In section 5 we derive the form of a quadratic proper TFR for the case of finite data sequences.

2. The General Form of TFRs

Let $s(t), t \in \mathcal{I}$, where $\mathcal{I}$ is the set of all integers, be a complex discrete time signal. We will denote by s the infinite dimensional vector whose elements are $s(t)$. Without loss of generality we may view this signal as samples of an underlying bandlimited continuous-time signal, sampled at 1 second intervals. We will assume that the signal is sampled at a sufficiently high rate to avoid aliasing

Let $P(t,w)$ be any function of time and frequency, constructed from the signal. The dependence of $P(t,w)$ on the signal is not shown explicitly for notational convenience. The frequency w is assumed for the moment to be continuous – corresponding to the Discrete-Time Fourier Transform. In other words, $P(t,w)$ is defined on a discrete time grid, but a continuous frequency grid, $-\pi \leq w \leq \pi$.

We will say that a function $P(t,w)$ is a proper TFR, if the following properties hold, for *any* signal $s(t)$.

1. Frequency-shift, or modulation property – If $s(t)$ is replaced by $s(t)e^{jw_0 t}$, then $P(t,w)$ becomes $P(t, w-w_0)$, for any frequency w_0.

*This work was supported by the Office of Naval Research under contracts No. N00014-95-1-0912 and N00014-89-V-1070.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Time-shift property – if $s(t)$ is replaced by $s(t - t_0)$, where t_0 is an arbitrary integer, then $P(t, w)$ becomes $P(t - t_0, w)$, for any delay t_0.

3. Scaling property – if $s(t)$ is replaced by $cs(t)$ where c is an complex constant, then $P(t, w)$ becomes $|c|^2 P(t, w)$.

4. Non-negativity property – $P(t, w) \geq 0$.

The first property is tied directly to the basic notion of frequency, and could be used as an alternative way of defining frequency. Without this property the variable w would not have the usual interpretation of frequency. This is the "modulation-invariant" property of [7]. The time-shift property is essential for introducing time dependence into the TFR. Shifting the signal must translate into a similar shift of the time-frequency representation of the signal. Otherwise the variable t would not have the usual meaning of time.

Assumptions 1 and 2 are fundamental properties, and it seems difficult to envision any reasonable definition of a TFR which does not obey them. We note that these properties hold for most of the TFRs discussed in the literature. However, here we propose to use them as the *defining* properties of proper TFRs, rather than derive them for an assumed TFR structure.

Assumptions 3 and 4 are necessary for $P(t, w)$ to be an energy function. Energy is, of course, non-negative, and is the "squared value" of some underlying physical quantity (voltage, current, velocity, etc.). Properties 3 and 4 could be combined into a single statement saying that any energy estimator $P(t, w)$ must be a non-negative function of the signal which scales quadratically under linear signal scaling.

Let us start by assuming that $P(t, w)$ is an arbitrary function of time and frequency, constructed from the signal. We will then examine the consequences of requiring that this function have the properties above. Thus, let

$$P(t, w) = g(t, w, s) \tag{1}$$

where g is a mapping from $\mathcal{C}^N$ to $\mathcal{R}$.

Now let us follow the methodology of [7] and consider what happens when we replace $s(t)$ by $s(t)e^{jw_0 t}$. Then s is replaced by $D(w_0)s$, where $D(w)$ is the frequency shift operator. A matrix representation of $D(w)$ would be an infinite dimensional diagonal matrix whose diagonal elements are $\{e^{jwt}\}$. From property 1 it follows that

$$g(t, w - w_0, s) = g(t, w, D(w_0)s) \tag{2}$$

Setting $w = 0$ and then replacing $-w_0$ by w we get

$$g(t, w, s) = g(t, 0, D(-w)s) \tag{3}$$

Next consider what happens when we replace $s(t)$ by $s(t - t_0)$. Then s is replaced by $Z^{-t_0}s$ where Z^{-t_0} denotes the delay operator. The unit delay operator Z^{-1} shifts s one sampling unit. Using property 2 it follows that

$$g(t - t_0, w, s) = g(t, w, Z^{-t_0}s) \tag{4}$$

Setting $t = 0$ and then replacing $-t_0$ by t we get

$$g(t, w, s) = g(0, w, Z^t s) \tag{5}$$

If we combine multiplication by $e^{jw_0 t}$ and delay by t_0 it is straightforward to show that

$$g(t, w, s) = g(0, 0, D(-w)Z^t s) \tag{6}$$

(using time shift followed by frequency shift), and

$$g(t, w, s) = g(0, 0, Z^t D(-w)s) \tag{7}$$

(using frequency shift followed by time shift). Note that Z^t and $D(-w)$ do not commute:

$$Z^t D(-w)s = e^{-jwt} D(-w)Z^t s \tag{8}$$

However, by property 3, $g(t, w, s)$ is invariant to unitary scaling, so either order of delay and modulation produces the same result:

$$g(t, w, s) = g(0, 0, Z^t D(-w)s) = g(0, 0, D(-w)Z^t s) \tag{9}$$

This is a representation theorem for the most general form of a proper TFR. The dependence of the proper TFR on time, frequency, and the signal, must be in the particular combination appearing above, and $g(0, 0, x)$ must have the quadratic scaling property.

To simplify the notation we will suppress the first two arguments of g from now on. In other words, $g(0, 0, \cdot)$ will be replaced by $g(\cdot)$, and therefore the *representation* for a proper TFR is

$$P(t, w) = g(Z^t D(-w)s) = g(D(-w)Z^t s) \tag{10}$$

where g satisfies the quadratic scaling property

$$g(cx) = |c|^2 g(x) \tag{11}$$

For fixed t and free w, the TFR $P(t, w)$ is a *frequency scan*, and for fixed w and free t it is a *time scan*. Each of these scans has a natural frequency or time-shift property.

3. Time and Frequency Responses of a Proper TFR

It is straightforward to prove a few more facts without specifying the function g, such as

1. If $s = \mathbf{1}$ (where $\mathbf{1}$ is an infinite vector of all 1's) then $P(t,w)$ is a function of w only. In other words, the TFR of a constant is not time-dependent. This follows immediately from the fact that $Z^t\mathbf{1} = \mathbf{1}$, for any t, in which case

$$P_1(t,w) = g(D(-w)Z^t\mathbf{1}) = g(D(-w)\mathbf{1}) = g(\psi(-w)) \tag{12}$$

where $\psi(w)$ is the infinite DTFT sequence $\{e^{jwn}\}$. This might reasonably be called the *frequency response* of the proper TFR.

2. If $s = \delta_n$ (where δ_n is an infinite vector of all zeros, except for a 1 at the n-th location) then $P(t,w)$ is a function of t only. In other words, the TFR of an impulse is not frequency dependent, in which case

$$P_\delta(t,w) = g(Z^tD(-w)\delta_n) = g(Z^t\delta_n) \tag{13}$$

where $Z^t\delta_n$ is the infinite delta sequence δ_{t+n}. This might reasonably be called the *temporal response* of the proper TFR.

3. If $s(t) = e^{j(\beta_1 t+\beta_2 t^2)}$ – a linear chirp – then $P(t,w) = P(0,w-2\beta_2 t)$. In other words, the TFR has a fixed shape, which gets shifted along the frequency axis at the *instantaneous frequency* of the chirp.

Property 3 raises the question of the most general signal s for which an *instantaneous carrier frequency* may be defined for a proper TFR. We shall say $\beta(t)$ is the instantaneous carrier frequency for a proper TFR $P(t,w)$ if

$$P(t,w) = P(0,w-\beta(t)) \tag{14}$$

That is, the TFR has a fixed shape which propagates along the (t,w) line $w = \beta(t)$. The requirement on s is

$$Z^tD(-w)s = cZ^0D(-(w-\beta(t)))s; \quad |c| = 1 \tag{15}$$

or

$$e^{-jw(n+t)}s(n+t) = ce^{-j(w-\beta(t))n}s(n) \tag{16}$$

It is just a few steps to show that the most general sequence with this property is the unit modulus, quadratic phase sequence

$$s(n) = e^{j(\phi_0+\phi_1 n+\phi_2 n^2)} \tag{17}$$

in which case $\beta(t) = 2\phi_2 t$. In summary, the most general instantaneous carrier frequency we can associate with a proper TFR is a *linear* carrier frequency – and it is just the *instantaneous* frequency of a linear chirp!

In summary, every proper TFR has the representation of equations (10) and (11), and the most general carrier frequency that can be determined for such a TFR is the linear carrier frequency of a linear chirp. No other signal produces a proper TFR for which a carrier frequency may be determined.

4. Quadratic TFRs

To get more useful insights it is necessary to specify the functional form of g. To do so, we shall assume that g is quadratic in the data. That is,

$$g(x) = \sum_m\sum_n x^*(m)Q(m,n)x(n); \quad Q(m,n) = Q^*(m,n) \tag{18}$$

where Q is an infinite dimensional non-negative definite kernel. In this case the *quadratic* proper TFR is

$$P(t,w) = \sum_m\sum_n s^*(m+t)e^{jwm}Q(m,n)e^{-jwn}s(n+t) \tag{19}$$

This is a representation theorem for the most general quadratic proper TFR.

A reasonable way to normalize $P(t,w)$ is to require its expected value to be 1 when s is stationary unit variance white noise:

$$E\{P(t,w)\} = 1 = \sum_n Q(n,n) \tag{20}$$

In this case $P(0,0)$ is a reasonable definition of generalized energy of s:

$$P(0,0) = \sum_m\sum_n s^*(m)Q(m,n)s(n) \tag{21}$$

The slices $P(0,w)$ and $P(t,0)$ are, respectively, *frequency scans* and *time scans*:

$$P(0,w) = \sum_m\sum_n s^*(m)Q(m,n)s(n)e^{-jw(n-m)} \tag{22}$$

$$P(t,0) = \sum_m \sum_n s^*(m+t)Q(m,n)s(n+t) \qquad (23)$$

One can go further to constrain Q by imposing restrictions on *marginals*. However, there is no choice of Q which reproduces both $|s(t)|^2$ and $|S(e^{jw})|^2$ for its marginals.

Now suppose the kernel Q is the rank-r non-negative definite kernel

$$Q(m,n) = \sum_{i=1}^{r} f_i(m) f_i(n) \qquad (24)$$

where the sequences $f_i, i = 1,2,\ldots,r$, are linearly independent in $l_2(C)$ and normalized so that

$$\sum_{i=1}^{r} \sum_n |f_i(n)|^2 = 1 \qquad (25)$$

Then the quadratic proper TFR is the following multiwindow STFT

$$P(t,w) = \sum_{i=1}^{r} |\sum_n f_i(n)s(n+t)e^{-jwn}|^2 \qquad (26)$$

This fixed-window — sliding-signal version of the TFR may also be written in the fixed-signal — sliding-window and its filtering forms:

$$\begin{aligned} P(t,w) &= \sum_{i=1}^{r} |\sum_n f_i(n-t)s(n)e^{-jwn}|^2 \\ &= \sum_{i=1}^{r} |\sum_n h_i(t-n)s(n)e^{-jwn}|^2 \quad (27) \\ & \qquad (28) \end{aligned}$$

If the impulse response $h_i(t) = f_i(-t)$ is finitely parametrized as an ARMA impulse response, then the rank-r multiwindow is finitely computable, for each t, on a finite lattice of the Nyquist band $(-\pi,\pi]$. The window design problem for quadratic proper TFRs is to design normalized windows f_i which concentrate energy in (t,w) cells of shape $[-\alpha,\alpha] \times [-\beta\pi,\beta\pi]$, where 2α is the desired temporal resolution and 2β the derived frequency resolution.

If the windows f_i are orthonormal, i.e. $(f_i,f_j) = \delta_{ij}$, then the TFR may be written as

$$P(t,w) = \frac{1}{r}|P_f Z^t D(-w)s|^2 \qquad (29)$$

where P_f is a projection onto the rank-r subspace of l_2 spanned by the windows f_i. This makes the most general rank-r, quadratic, proper TFR, a matched subspace energy detector [9].

We note that some of the TFRs proposed in the literature employ kernels Q which are signal dependent, see e.g. [2]. Strictly speaking, such TFRs do not belong to the class of quadratic TFRs, since they are complicated non-quadratic functions of the signal. Furthermore, these TFRs are proper TFRs *only* if the kernel depends on the signal in the particular functional form discussed earlier, i.e., $Q(Z^t D(-w)s)$ or $Q(D(-w)Z^t s)$. The signal dependent kernels proposed in the literature seem to all have this property.

5. TFRs for Finite Sequences

In many situations the signals for which we wish to compute TFRs will be of finite extent, i.e. $s(t)$ will be non-zero only for $t = 0,\ldots,N-1$. In this section we will specialize the quadratic TFRs for infinite data sequences to the case of finite sequences.

First note that in the finite data case $P(t,w)$ needs to be computed only for $t = 0,\ldots,N-1$. In other words, it is not necessary to consider more time-shifts than there are data samples. In order to do this properly we will define the signal vector as the signal samples, preceded by $N-1$ zeros,

$$[\underbrace{0,\ldots,0}_{N-1 \text{ zeros}}, s(0),\ldots,s(N-1)]^T \qquad (30)$$

This will "leave room" for the necessary number of shifts. This zero-padded vector can be written more compactly as $\mathbf{Fs}$, where $\mathbf{s} = [s(0),s(1),\ldots,s(N-1)]^T$. That is,

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} \\ I_N \end{bmatrix} \qquad (31)$$

where $\mathbf{0}$ is an $(N-1) \times N$ matrix of zeros, $\mathbf{I}_N$ is the $N \times N$ identity matrix, and $\mathbf{Fs}$ is a $(2N-1) \times 1$ vector. In the following we will denote the length of $\mathbf{Fs}$ by M, i.e. $M = 2N-1$.

When dealing with finite sequences it is convenient to use the circular delay operator, which will be defined by $\mathbf{Z}^{-1}$ – the circular down-shift matrix,

$$\mathbf{Z}^{-1} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \qquad (32)$$

The circular "advance" operator is given by $\mathbf{Z}$ – the circular up-shift matrix, which is the transpose of $\mathbf{Z}^{-1}$,

$$\mathbf{Z} = (\mathbf{Z}^{-1})^{-1} = (\mathbf{Z}^{-1})^T \qquad (33)$$

We define the modulation matrix

$$\mathbf{D}(w) = \text{diag}\{\psi(w)\} \tag{34}$$

where $\psi(w)$ is the DTFT vector,

$$\psi(w) = [1, e^{jw}, \ldots, e^{jw(M-1)}]^T \tag{35}$$

Using this notation, it is straightforward to verify that the quadratic TFR introduced earlier can be rewritten as

$$P(t,w) = \mathbf{s}^H \mathbf{F}^H \mathbf{Z}^{-t} \mathbf{D}(w) \mathbf{Q} \mathbf{D}(-w) \mathbf{Z}^t \mathbf{F}\mathbf{s} \tag{36}$$

or equivalently

$$P(t,w) = \mathbf{s}^H \mathbf{F}^H \mathbf{D}(w) \mathbf{Z}^{-t} \mathbf{Q} \mathbf{Z}^t \mathbf{D}(-w) \mathbf{F}\mathbf{s} \tag{37}$$

where $t = 0, \ldots, N-1$. It is easy to show that P(t,w) has the required properties for a proper TFR. The matrix $\mathbf{Q} \geq 0$ is an $M \times M$ block of the infinite dimensional kernel Q defined earlier. Equations (36) and (37) provide a *representation theorem* for quadratic proper TFRs, for the finite data case.

Note that because of the presence of $N-1$ zeros in $\mathbf{Fs}$, there are parts of $\mathbf{Q}$ which never enter in the computation of $P(t,w)$. Specifically, only the values of $\mathbf{Q}$ which are in the band of width M, between the N-th upper diagonal and N-th lower diagonal, enter the calculations. Values outside this band could be arbitrary.

In [10] we relate the form of the TFR in the equation above to the multi-windowed short-time Fourier transform (STFT), and also present an alternative form of the quadratic proper TFR.

6. Conclusions

Starting with the requirement that any "proper" TFR must obey natural time- and frequency-shift properties, have a quadratic scaling property, and be non-negative, we have derived the structure of such time-frequency representations in equation (10) and (11). We have shown that any quadratic proper TFR must have the form specified in equation (19). In the finite data case the quadratic proper TFR must have the form of equations (36) or (37). When the kernel of the TFR is of rank r, the TFR is just a windowed STFT.

The main point we have tried to make is that the TFR structure is a direct consequence of a simple set of defining properties. Thus, it is not necessary to introduce various auxiliary notions such as time-frequency distributions obeying certain marginal properties, in order to come up with quadratic TFRs.

References

[1] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, 1995.

[2] Baraniuk, R.G.; Jones, D.L. "A signal-dependent time-frequency representation: fast algorithm for optimal kernel design," *IEEE Transactions on Signal Processing*, Jan. 1994, vol.42, no.1, pp. 134-46.

[3] Jones, D.L.; Baraniuk, R.G. "An adaptive optimal-kernel time-frequency representation" *IEEE Transactions on Signal Processing*, Oct. 1995, vol.43, no.10, pp. 2361-71.

[4] Jones, D.L.; Baraniuk, R.G. "A simple scheme for adapting time-frequency representations," *IEEE Transactions on Signal Processing*, Dec. 1994, vol.42, no.12, pp. 3530-5.

[5] L. Cohen and T. E. Posch, "Positive Time-Frequency Distribution Functions," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-33, No. 1, pp. 31–37, February 1985.

[6] P. J. Laughlin, J. W. Pitton, and L. E. Atlas, "Construction of Positive Time-Frequency Distributions," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2697–2705, October 1994.

[7] C. T. Mullis and L. L. Scharf, "Quadratic Estimators of the Power Spectrum," pp. 1–57, in *Advances in Spectrum Analysis and Array Processing: Volume I*, Simon Haykin, Ed., Prentice-Hall, 1991.

[8] D. J. Thomson, "Spectrum Estimation and Harmonic Analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[9] L. L. Scharf and B. Friedlander, "Matched Subspace Detection," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146–2157, August 1994.

[10] B. Friedlander and L. L. Scharf, "On the Structure of Time-Frequency Representation," submitted for publications.

HYBRID FM-POLYNOMIAL PHASE SIGNAL MODELING: PARAMETER ESTIMATION AND PERFORMANCE ANALYSIS

[1] *Fulvio Gini* and [2] *Georgios B. Giannakis*

[1]Dept. "Ingr. dell' Informazione", Univ. of Pisa, via Diotisalvi, 2 I-56126 Pisa, ITALY
Tel: (+39) 50-568550 Fax: (+39) 50-568522 E-mail: gini@iet.unipi.it
[2]Dept. of Electrical Engr., Univ. of Virginia, Charlottesville, VA 22903-2442, U.S.A.
Tel: (804) 924-3659 Fax: (804) 924-8818 E-mail: georgios@virginia.edu

ABSTRACT

Parameter estimation for a combination of a polynomial phase signal (PPS) and a frequency modulated (FM) signal is addressed. A novel approach is proposed that allows one to decouple estimation of the FM parameters from that of the PPS parameters, exploiting the properties of the multi-lag High-order Ambiguity Function (ml-HAF). Performance analysis is carried out and Cramer-Rao bounds are compared with simulation results.

1. PROBLEM STATEMENT

The discrete-time model of the hybrid FM-PPS signal of interest is:

$$s(n) = \rho\, e^{j2\pi b \sin(\omega_0 n+\phi_0)}\, e^{j2\pi \sum_{i=0}^{M} a_i n^i/i!}, \qquad (1)$$

where: $\{a_i\}_{i=0}^{M}$ are the PPS parameters, ρ denotes a constant (perhaps unknown) amplitude, and M is the PPS order; constant $b > 0$ is the so-called modulation index, ω_0, ϕ_0 stand for the frequency and initial phase, respectively.

This paper concerns parameter estimation of the hybrid FM-PPS signal in (1) from a finite number of N noisy observations, $\{x(n) = s(n) + v(n)\}_{n=0}^{N-1}$. The additive noise $v(n)$ is assumed stationary, complex white Gaussian, mixing in the sense of [4, p. 25], with zero-mean and variance σ_v^2. Signals described by (1) are encountered in many engineering applications such as radar, SAR, sonar, acoustics, and optics. When the target is moving, the received signal can be modeled as a discrete-time PPS (see e.g., [10]), and the polynomial coefficients $\{a_i\}$ are related to the kinematics of the moving target [8, p. 403]. The High-order Instantaneous Moment (HIM) and its Fourier transform, the High-order Ambiguity Function (HAF), introduced by Peleg and Porat (see [8, Ch. 12]), have provided a simple albeit sub-optimum algorithm for estimating the PPS coefficients recursively.

Signals arising from moving targets however, cannot always be modeled as a pure PPS. For example, sinusoidal FM signals arise from vibrating targets [5], [12], or, rotating parts of the target [3], [7]. When both of these effects (motion and vibration/rotation) are present, the noise-free backscattered signal obeys (1), and the estimation algorithms proposed in [5] and [11] for a pure sinusoidal FM are no more appropriate.

The main contribution of this paper is to show that the HAF offers a good alternative to the computationally intensive maximum likelihood (ML) approach, even when the observed data cannot be modeled as a pure PPS, but a sinusoidal FM component is also present.

The multi-lag HIM is a nonlinear transformation, originally introduced for continuous-time signals in [1], defined recursively as [2]:

$$x_1(n) = x(n),\ x_2(n;\tau_1) = x_1(n+\tau_1)x_1^*(n-\tau_1),\ \ldots,$$
$$x_M(n;\tau_1,\ldots,\tau_{M-1}) = x_{M-1}(n+\tau_{M-1};\tau_1,\ldots,\tau_{M-2}) \times x_{M-1}^*(n-\tau_{M-1};\tau_1,\ldots,\tau_{M-2})\ . \qquad (2)$$

We term x_M the multi-lag HIM (ml-HIM) because it reduces to the HIM for $\tau_1 = \ldots = \tau_{M-1} = \tau$, [8, Ch. 12], [10]. The ml-HAF is defined as the (generalized) Fourier Series of the ml-HIM:

$$X_M(\alpha;\tau) := \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} x_M(n;\tau_1,\ldots,\tau_{M-1}) e^{-j\alpha n}\ , \qquad (3)$$

where $\tau := [\tau_1\ \tau_2\ \ldots\ \tau_{M-1}]$.

2. ESTIMATION ALGORITHMS

Applying the ml-HIM in (2) to the noise-free FM-PPS signal we find:

$$s_M(n;\tau) = \rho^{2^{M-1}} e^{j[\omega_c n+\psi_c+\beta\sin(\omega_0 n+\psi_0)]}\ , \qquad (4)$$

where:

$$\omega_c := 2^M \pi a_M \prod_{m=1}^{M-1} \tau_m\ ,\quad \psi_c := 2^M \pi a_{M-1} \prod_{m=1}^{M-1} \tau_m\ , \qquad (5)$$

$$\beta := 2^M \pi b \prod_{m=1}^{M-1} \sin(\omega_0 \tau_m)\ ,\ \psi_0 := \phi_0 + (M-1)\frac{\pi}{2}\ . \qquad (6)$$

Eq. (4) shows that the M-th order ml-HIM of a sinusoidal FM signal is still a sinusoidal FM signal, with the same vibration frequency ω_0 but with different modulation index, β, that is proportional to the original one, b.

Being (almost) periodic in n, $s_M(n;\tau)$ of (4) can be written as a superposition of equally spaced harmonics $\{\omega_c \pm k\omega_0\}$ with amplitudes $\{J_k(\beta)\}$, the kth-order Bessel functions of the first kind, [9, p. 311]:

$$s_M(n;\tau) = \rho^{2^{M-1}} \sum_{k=-\infty}^{\infty} J_k(\beta) e^{j[(\omega_c+k\omega_0)n+\psi_c+k\psi_0]}\ . \qquad (7)$$

Because $J_k(\beta)$ vanishes for large $|k|$, [9], there exists a large enough integer K such that most of the energy in $\{J_k(\beta)\}$ is contained in the range $k \in [-K, K]$. It has been shown in [11] that the smallest integer K for which $\sum_{k=-K}^{K} J_k^2(\beta) > 0.99$, is $K \approx \beta+1$ for $\beta > 1$, whereas $K = 0$ for $\beta \in [0, 0.14)$. For $\beta \in [0.14, 1]$ K is either 1 or 2. Depending on β, we can

0-8186-8005-9/97 $10.00 © 1997 IEEE

thus approximate the infinite sum in (7) by the partial sum

$$s_M(n;\boldsymbol{\tau}) = \rho^{2^{M-1}} \sum_{k=-K}^{K} J_k(\beta) e^{j[(\omega_c+k\omega_0)n+\psi_c+k\psi_0]} , \quad (8)$$

which consists of multi-component constant amplitude tones with harmonic frequencies. Fourier transforming (8), we obtain the ml-HAF as

$$S_M(\alpha;\boldsymbol{\tau}) = \rho^{2^{M-1}} \sum_{k=-K}^{K} J_k(\beta) e^{j(\psi_c+k\psi_0)} \delta(\alpha-\omega_c-k\omega_0) , \quad (9)$$

where $\delta(\cdot)$ denotes the Kronecker delta function.

The ml-HAF in (9) peaks at $\{\omega_c + k\omega_0\}_{k=-K}^{K}$. Hence, the vibration frequency ω_0 is given by the distance between successive peaks of $|S_M(\alpha;\boldsymbol{\tau})|$. The position of the central peak is related to the highest order PPS coefficient a_M through ω_c [c.f. (5)]. Then, making use of (9) one can obtain β and ψ_0, and subsequently the modulation index b and the phase ϕ_0, from the values at the peaks.

Sample estimates of $S_M(\alpha;\boldsymbol{\tau})$ are computed from $\{x(n)\}_{n=0}^{N-1}$ as

$$\widehat{S}_M(\alpha;\boldsymbol{\tau}) = \frac{1}{N} \sum_{n=0}^{N-1} x_M(n;\boldsymbol{\tau})\, e^{-j\alpha n}, \quad (10)$$

where $x_M(n;\boldsymbol{\tau})$ is the ml-HIM, and $\widehat{S}_M(\alpha;\boldsymbol{\tau}) = X_M(\alpha;\boldsymbol{\tau})$ is the ml-HAF of the data. Under the stated assumptions on $v(n)$, the estimator in (10) is asymptotically unbiased and consistent in the mean square sense (see [10] for a proof). Once $\widehat{S}_M(\alpha;\boldsymbol{\tau})$ is computed, the parameters of interest can be estimated by substituting $\widehat{S}_M(\alpha;\boldsymbol{\tau})$ for $S_M(\alpha;\boldsymbol{\tau})$ in (9).

We can take advantage of the multiple lags and calculate X_M for L different sets of lags, $\boldsymbol{\tau}_l := [\tau_{1,l} \ \ldots \ \tau_{M-1,l}]$, $l = 1, 2, \ldots, L$, having the same product; i.e., $\prod_{m=1}^{M-1} \tau_{m,l_1} = \prod_{m=1}^{M-1} \tau_{m,l_2}, \forall\, l_1, l_2 \in [1, L]$. The product of L ml-HAF amplitudes $\prod_{l=1}^{L} |X_M(\alpha, \boldsymbol{\tau}_l)|$ enhances the desired peaks and reduces noise further than the single ml-HAF [2].

The resulting algorithm is summarized in the following steps.

Step 1: Calculate from the data $x_M(n;\boldsymbol{\tau}_l)$, the ml-HIM of order M, for each set of lags $\boldsymbol{\tau}_l$, $l = 1, 2, \ldots, L$.

Step 2: Estimate the ml-HAF of $s(n)$ using (10).

Step 3: Find the peak location of the product ml-HAF, $\prod_{l=1}^{L} |\widehat{S}_M(\alpha, \boldsymbol{\tau}_l)| = \prod_{l=1}^{L} |X_M(\alpha, \boldsymbol{\tau}_l)|$, corresponding to $\{\omega_c + k\omega_0\}_{k=-K}^{K}$ [see also (9)], and collect these values in a $(2K+1) \times 1$ vector $\widehat{\boldsymbol{p}}$.

Step 4: Define $\mathbf{A} := [\mathbf{k}\ \mathbf{1}]$, $\widehat{\boldsymbol{\omega}} := [\widehat{\omega}_0\ \widehat{\omega}_c]^T$, and

$$\mathbf{k} := [-K \quad -K+1 \quad \ldots \quad K-1 \quad K]^T,$$

where $\mathbf{1}$ is the vector of ones. Solve the overdetermined linear system $\mathbf{A}\widehat{\boldsymbol{\omega}} \overset{LS}{=} \widehat{\boldsymbol{p}}$, to compute $\widehat{\omega}_0$ and $\widehat{\omega}_c$ as: $\widehat{\boldsymbol{\omega}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\widehat{\boldsymbol{p}}$.

Step 5: Estimate a_M as:

$$\widehat{a}_M = \widehat{\omega}_c / (2^M \pi \prod_{m=1}^{M-1} \tau_{m,1}) . \quad (11)$$

Step 6: Estimating b from (10) and (6) requires some care, because (10) is exact only as $N \to \infty$. For finite sample size we find from (1) and (10) that:

$$\widehat{S}_M(\omega_c + k\omega_0; \boldsymbol{\tau}_l) = \rho^{2^{M-1}} \frac{N - 2\sum_{m=1}^{M-1} \tau_{m,l}}{N} \times J_k(\beta_l) e^{j(\psi_c+k\psi_0)} + \textit{noise terms} . \quad (12)$$

where $\beta_l := 2^M \pi b \prod_{m=1}^{M-1} \sin(\omega_0 \tau_{m,l})$. To estimate b based on (10), we first find

$$\widehat{J}_k(\beta_l) = \frac{\widehat{S}_M(\widehat{\omega}_c + k\widehat{\omega}_0; \boldsymbol{\tau}_l)}{\rho^{2^{M-1}}(N - 2\sum_{m=1}^{M-1} \tau_{m,l})/N} , \text{ for } k \in [-K, K] , \quad (13)$$

and estimate β_l as (see also [11]):

$$\widehat{\beta}_l = \frac{2s}{s^2-1} \ln \left[\sum_{k=-K}^{K} \widehat{J}_k(\beta_l)\, s^k \right] . \quad (14)$$

Finally, obtain an estimate of b via (6).

Step 7: Define:

$$\widehat{\boldsymbol{\psi}} := [\widehat{\psi}_0 \ \ \widehat{\psi}_c]^T,\ \widehat{\Psi}_{k,l} := \arg\{\widehat{S}_M(\widehat{\omega}_c + k\widehat{\omega}_0; \boldsymbol{\tau}_l)\},$$
$$\widehat{\boldsymbol{\Psi}}_l := [\widehat{\Psi}_{-K,l} \ \ldots \ \widehat{\Psi}_{K,l}]^T, \widehat{\boldsymbol{\Psi}} := [\widehat{\boldsymbol{\Psi}}_1^T \ \ldots \ \widehat{\boldsymbol{\Psi}}_L^T]^T,$$

and the $(2M+1)L \times 2$ dimensional matrix $\mathbf{B}$ as:

$$\mathbf{B} := [\mathbf{A}^T \ \mathbf{A}^T \ \ldots \ \mathbf{A}^T]^T .$$

Solve the overdetermined linear system $\mathbf{B}\widehat{\boldsymbol{\psi}} \overset{LS}{=} \widehat{\boldsymbol{\Psi}}$ to find $\widehat{\psi}_0$ and $\widehat{\psi}_c$ as: $\widehat{\boldsymbol{\psi}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\widehat{\boldsymbol{\Psi}}$.

Then, $\widehat{\phi}_0$ is obtained from $\widehat{\psi}_0$. Now, having estimated a_M and the FM component we can remove it from the data and proceed with the PPS part using techniques described in [10].

As an example, in Fig. 1a we plot the values of the Bessel coefficients $J_k(\beta)$ versus k, for $b = 6$, $\omega_0 = 0.0491$, $\tau_1 = 129$. From (5)-(6) we find $\omega_c = 0.5629$ and $\beta = 3.6996$. The values of the other parameters can be found in the Example 2, Section 4. As expected, the coefficients become negligible for $|K| > 5$ in (7), confirming that $K \approx \beta + 1$. In Fig. 1b the ml-HAF $X_2(\alpha;\tau_1)$ is plotted for $\alpha \in [0, 1]$, $SNR := \rho^2/\sigma_v^2 = 16\ dB$, and $N = 1,024$. From the location (and value) of the peaks we derive our estimates.

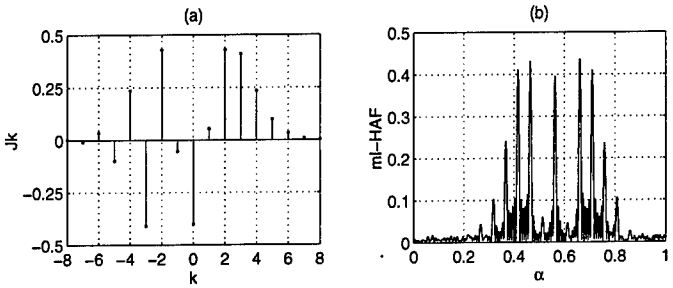


Figure 1. (a) Bessel coefficients $J_k(\beta)$ vs. k, for $\beta = 3.6996$. (b) $X_2(\alpha;\tau_1)$ vs. α, for $SNR = 16\ dB$ and $N = 1,024$.

Narrowband FM. According to Carson's rule the bandwidth of an FM signal is approximately $2\omega_0(1+\beta)$, [9, p. 313]. We will refer to the case discussed so far, for general β, as the wideband FM (WBFM). Expanding $\exp[j\beta\sin(\omega_0 n + \psi_0)]$ in Taylor series we find:

$$e^{j\beta\sin(\omega_0 n+\psi_0)} = \sum_{k=0}^{\infty} \frac{[j\beta\sin(\omega_0 n+\psi_0)]^k}{k!} . \quad (15)$$

If β is small enough, we can truncate the series expansion as follows:

$$e^{j\beta\sin(\omega_0 n+\psi_0)} \approx$$
$$1 + j\beta\sin(\omega_0 n+\psi_0) - \frac{1}{2}\beta^2\sin^2(\omega_0 n+\psi_0)$$
$$= 1 - \frac{\beta^2}{4} + j\beta\sin(\omega_0 n+\psi_0) - \frac{\beta^2}{4}\cos(2\omega_0 n+2\psi_0).$$

A value of $\beta < 1$ is sufficient to guarantee that the second order approximation is valid. We will refer to the case of $\beta < 1$ as narrowband FM (NBFM). Using this approxmation in (8) we obtain:

$$\begin{aligned} S_M(\alpha;\tau) \approx \rho^{2^{M-1}}[&-\frac{\beta^2}{8}\,e^{j(\psi_c-2\psi_0)}\delta(\alpha-\omega_c+2\omega_0) \\ &+ \frac{\beta}{2}\,e^{j(\psi_c-\psi_0)}\delta(\alpha-\omega_c+\omega_0) + (1-\frac{\beta^2}{4})\,e^{j\psi_c}\delta(\alpha-\omega_c) \\ &+ \frac{\beta}{2}\,e^{j(\psi_c+\psi_0)}\delta(\alpha-\omega_c-\omega_0) \\ &- \frac{\beta^2}{8}\,e^{j(\psi_c+2\psi_0)}\delta(\alpha-\omega_c-2\omega_0)], \end{aligned} \quad (16)$$

which makes estimation of β from the amplitude of the peaks easier. Note that the bandwidth of the signal in (16) is $4\omega_0$ as predicted by Carson's rule (with $\beta = 1$). The estimation algorithm is obtained by modifying Step 6 as follows:

Step 6 (NBFM): Exploit (15) to obtain different estimates of β as:

$$\begin{aligned} \widehat{\beta}_{0,l} &= 2\sqrt{1 - \frac{|\widehat{S}_M(\widehat{\omega}_c;\tau_l)|}{\rho^{2^{M-1}}(N-2\sum_{m=1}^{M-1}\tau_{m,l})/N}}, \\ \widehat{\beta}_{1,l} &= \frac{|\widehat{S}_M(\widehat{\omega}_c+\widehat{\omega}_0;\tau_l)| + |\widehat{S}_M(\widehat{\omega}_c-\widehat{\omega}_0;\tau_l)|}{\rho^{2^{M-1}}(N-2\sum_{m=1}^{M-1}\tau_{m,l})/N}, \\ \widehat{\beta}_{2,l} &= 2\sqrt{\frac{|\widehat{S}_M(\widehat{\omega}_c+2\widehat{\omega}_0;\tau_l)| + |\widehat{S}_M(\widehat{\omega}_c-2\widehat{\omega}_0;\tau_l)|}{\rho^{2^{M-1}}(N-2\sum_{m=1}^{M-1}\tau_{m,l})/N}}, \end{aligned} \quad (17)$$

and then take their mean, $\widehat{\beta} = 1/(3L)\sum_{l=1}^{L}\sum_{i=0}^{2}\widehat{\beta}_{i,l}$. Finally, $\widehat{b}$ is derived by inserting $\widehat{\beta}$ in (6). The other steps remain unchanged. It is worth observing that because $\beta = 2^M\pi b\prod_{m=1}^{M-1}\sin(\omega_0\tau_m)$, it is possible to select the lags τ in order to have $\beta < 1$. This allows estimation of β using the second-order approximation (16).

3. PERFORMANCE EVALUATION

Consider the PPS-FM model in (1) and denote the parameter vector as $\boldsymbol{\theta}$. To avoid numerical problems in the inversion of the Fisher information matrix, it is useful to define $\boldsymbol{\theta}$ as follows:

$$\boldsymbol{\theta} := [N^0 a_0 \;\ldots\; N^M a_M \;\; b \;\; N\omega_0 \;\; \phi_0]^T. \quad (18)$$

If $\widehat{\boldsymbol{\theta}}$ is an unbiased estimator of $\boldsymbol{\theta}$, then it must satisfy $\mathrm{cov}(\widehat{\boldsymbol{\theta}}) \geq \mathbf{J}^{-1}$, where $\mathbf{J}$ is the so-called Fisher information matrix (FIM). As a first step towards deriving the CRLBs, let us write the observed data in vector form as $\mathbf{x} = \mathbf{s}(\boldsymbol{\theta}) + \mathbf{v}$, where $[\mathbf{x}]_n := x(n)$, $n = 0, 1, \ldots, N-1$; similar definitions apply to $\mathbf{s}(\boldsymbol{\theta})$ and $\mathbf{v}$. Due to the assumptions on $v(n)$, we have that $\mathbf{x}$ is an $N \times 1$ complex Gaussian circular vector, with mean value $E\{\mathbf{x}\} = \mathbf{s}$ and covariance matrix $E\{\mathbf{x}\mathbf{x}^H\} = \sigma_v^2\mathbf{I}$, where ^{H} denotes conjugate transpose, and $\mathbf{I}$ is an $N \times N$ identity matrix. Calculation of the FIM for a vector of correlated complex Gaussian observations is carried out in [6, Ch. 3], and in our case it produces for the (i,j)th entry:

$$J_{ij} = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial \mathbf{s}^H(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial\theta_j}\right\}. \quad (19)$$

The CRLBs are obtained from the diagonal elements of $\mathbf{J}^{-1}$, which we denote as $CRLB(\theta_i) = [\mathbf{J}^{-1}]_{ii}$. Inserting $s(n)$ of (1) in (19) and taking partial derivatives, we obtain the elements of the FIM. Evaluation of J_{ij} is not difficult but tedious. Skipping the details, we found:

$$J_{a_i,a_m} = \frac{8\pi^2\rho^2}{i!m!\,\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)^{i+m},\; 0 \leq i,m \leq M, \quad (20)$$

$$J_{b,b} = \frac{8\pi^2\rho^2}{\sigma_v^2}\sum_{n=0}^{N-1}\sin^2(\omega_0 n+\phi_0), \quad (21)$$

$$J_{\omega_0,\omega_0} = \frac{8\pi^2\rho^2 b^2}{\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)^2\cos^2(\omega_0 n+\phi_0), \quad (22)$$

$$J_{\phi_0,\phi_0} = \frac{8\pi^2\rho^2 b^2}{\sigma_v^2}\sum_{n=0}^{N-1}\cos^2(\omega_0 n+\phi_0), \quad (23)$$

$$J_{a_i,b} = \frac{8\pi^2\rho^2}{i!\,\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)^{i}\sin(\omega_0 n+\phi_0), \quad (24)$$

$$J_{a_i,\omega_0} = \frac{8\pi^2\rho^2 b}{i!\,\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)^{i+1}\cos(\omega_0 n+\phi_0), \quad (25)$$

$$J_{a_i,\phi_0} = \frac{8\pi^2\rho^2 b}{i!\,\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)^{i}\cos(\omega_0 n+\phi_0), \quad (26)$$

$$J_{b,\omega_0} = \frac{4\pi^2\rho^2 b}{\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)\sin(2\omega_0 n+2\phi_0), \quad (27)$$

$$J_{b,\phi_0} = \frac{4\pi^2\rho^2 b}{\sigma_v^2}\sum_{n=0}^{N-1}\sin(2\omega_0 n+2\phi_0), \quad (28)$$

$$J_{\omega_0,\phi_0} = \frac{8\pi^2\rho^2 b^2}{\sigma_v^2}\sum_{n=0}^{N-1}\left(\frac{n}{N}\right)\cos^2(\omega_0 n+\phi_0). \quad (29)$$

Equations (20)-(29) allow the exact numerical evaluation of the FIM, but it is not clear how the parameters affect the bounds. However, a large sample approximation can be derived by regarding the sums as an Euler approximation of the integral, $\sum_{n=0}^{N-1} f(n) \approx \int_0^N f(t)dt$. With this in mind, we obtain:

$$\begin{aligned} J_{a_i,a_m} &= \frac{8\pi^2\rho^2}{(i+m+1)i!m!\,\sigma_v^2},\; 0 \leq i,m \leq M, \\ J_{b,b} &= \frac{4\pi^2\rho^2}{\sigma_v^2}\left[N - \frac{\sin(2\omega_0 N+2\phi_0)-\sin(2\phi_0)}{2\omega_0}\right], \\ J_{\omega_0,\omega_0} &= \frac{4\pi^2\rho^2 b^2}{\sigma_v^2}\left[\frac{N}{3} + \frac{\sin(2\omega_0 N+2\phi_0)}{2\omega_0}\right], \\ J_{\phi_0,\phi_0} &= \frac{4\pi^2\rho^2 b^2}{\sigma_v^2}\left[N + \frac{\sin(2\omega_0 N+2\phi_0)-\sin(2\phi_0)}{2\omega_0}\right], \end{aligned}$$

$$
\begin{aligned}
J_{a_0,b} &= -\frac{8\pi^2\rho^2}{i!\,\sigma_v^2}\cdot\frac{\cos(\omega_0 N+\phi_0)-\cos(\phi_0)}{\omega_0},\\
J_{a_i,b} &= -\frac{8\pi^2\rho^2}{i!\,\sigma_v^2}\cdot\frac{\cos(\omega_0 N+\phi_0)}{\omega_0},\ 1\le i\le M,\\
J_{a_i,\omega_0} &= \frac{8\pi^2\rho^2 b}{i!\,\sigma_v^2}\cdot\frac{\sin(\omega_0 N+\phi_0)}{\omega_0},\\
J_{a_0,\phi_0} &= \frac{8\pi^2\rho^2 b}{\sigma_v^2}\cdot\frac{\sin(\omega_0 N+\phi_0)-\sin(\phi_0)}{\omega_0},\\
J_{a_i,\phi_0} &= \frac{8\pi^2\rho^2 b}{i!\,\sigma_v^2}\cdot\frac{\sin(\omega_0 N+\phi_0)}{\omega_0},\ 1\le i\le M,\\
J_{b,\omega_0} &= -\frac{4\pi^2\rho^2 b}{\sigma_v^2}\cdot\frac{\cos(2\omega_0 N+2\phi_0)}{2\omega_0},\\
J_{b,\phi_0} &= -\frac{4\pi^2\rho^2 b}{\sigma_v^2}\cdot\frac{\cos(2\omega_0 N+2\phi_0)-\cos(2\phi_0)}{2\omega_0},\\
J_{\omega_0,\phi_0} &= \frac{2\pi^2\rho^2 b^2}{\sigma_v^2}\left[N+\frac{\sin(2\omega_0 N+2\phi_0)}{2\omega_0}\right].
\end{aligned}
$$

In Fig. 2 the CRLBs of the FM and PPS parameters are reported for $SNR = 12\ dB$, $b = 0.05$, $\omega_0 = 0.015$, and $\phi_0 = 0$. The solid and dashed lines refer to the exact and approximate CRLBs, respectively. The dashdotted lines refer to the exact CRLBs when only FM or only PPS parameters are present (let us call them marginal CRLBs). As we see, asymptotic behavior is reached at about $N = 1000$. The effect of coupling beetwen the PPS and the FM parameters is evident by comparing the plots with solid and dashdotted lines. The results show that for $N > 1000$ the coupling tends to be negligible.

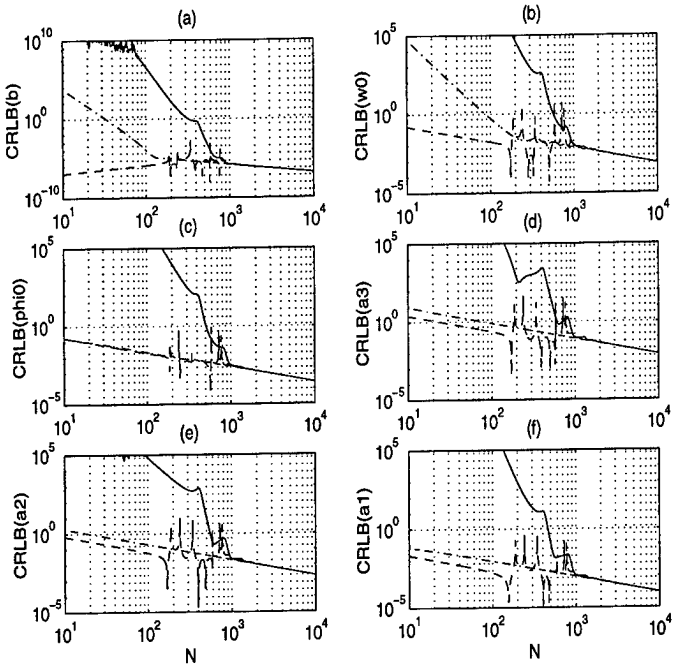


Figure 2. Exact (solid), approximate (dashed) and marginal (dashdotted) CRLBs vs. N.

To illustrate and evaluate the various parameter estimation algorithms discussed so far, we simulated numerically the performance by Monte Carlo experiments and we compared them with the CRLBs.

Example 1: Narrowband FM signal and 3^{rd}-order PPS. We first generated $N = 2,048$ samples according to (1), with parameters: $\rho = 1$, $b = 0.05$, $\omega_0 = 0.015$, $\phi_0 = 0$, $a_0 = 0$, $a_1 = 0.25$, $a_2 = 1.3889\cdot 10^{-3}$, $a_3 = 1.3022\cdot 10^{-5}$. The noise variance σ_v^2 was set to obtain the desired signal-to-noise ratio, defined as $SNR := \rho^2/\sigma_v^2$. The sets of lags in the product ml-HAF were: $(\tau_{1,l},\tau_{2,l}) = (60,60), (72,50), (75,48), (80,45), (90,40)$, and $(100,36)$, so that $\tau_{1,l}\cdot\tau_{2,l} = 3600$, for $l = 1,\cdots,6$; consequently $\omega_c = 1.1782$ and $\beta_l = 0.7711, 0.7555, 0.7476, 0.7319, 0.6923, 0.6445$, respectively. Note that all of them are less than one, so the second-order approximation is valid. All FFT operations were carried out by zero-padding to $N_{zp} = 2^{16}$ points so that the frequency bins were small enough to allow accurate estimation of the peaks' location.

In Fig. 3 we show the mean square errors (mse) of the estimates versus SNR, obtained from 400 independent Monte Carlo runs (in each run we generated $N = 2,048$ samples). The solid lines refer to the approximate CRLBs ($CRLB(\omega_0)$ and $CRLB(a_m)$ were properly scaled by N^2 and N^{2m}, respectively); the dashed lines correspond to the mse obtained assuming known amplitude ρ; and the dashdotted lines refer to the case where ρ is unknown and is estimated using the following algorithm based on fourth order cumulants:

$$
\widehat{\rho} = \left[2\left(\frac{1}{N}\sum_{n=0}^{N-1}|x(n)|^2\right)^2 - \frac{1}{N}\sum_{n=0}^{N-1}|x(n)|^4\right]^{1/4}. \qquad (30)
$$

Unbiasedness and mean square consistency follow from the mixing conditions assumed to be satisfied for $v(n)$; the proof is standard and is not reported here.

In the range of $SNRs$ investigated, no sensible difference was observed, between the two cases (known ρ and unknown ρ), except for the mse of $\widehat{b}$. We also note that the mse's of $\widehat{a}_3$ and $\widehat{a}_2$ are not very close to the bounds. The reason is related to the amount of zero-padding [10]. Similar results have been obtained for the other parameters, but are not reported here due to space limitations. Finally, in Fig. 4 we report the mse's and the exact CRLBs for the estimators of b and ω_0 versus N. As expected, when $CRLB(\omega_0)$ drops below 10^{-10} the mse is no more able to track the bound, due to the limited FFT resolution.

Example 2: Wideband FM signal and 2^{nd}-order PPS. Here, we generated $N = 1,024$ samples of the FM-PPS process with second-order polynomial phase, and parameters given by: $b = 6$, $\omega_0 = 0.0491$, $\phi_0 = 0$, $a_0 = 0.5$, $a_1 = 0.1$, $a2 = 3.4722\cdot 10^{-4}$, and $\tau_1 = 129$, so $\omega_c = 0.5629$ and $\beta = 3.6996$. The modulation index was estimated according to the method reported in (14), proposed in [11]. Note that $\beta > 1$, so we cannot use (17).

Tables I and II show biases and variances of the estimates for the $SNR = 16\ dB$ (first two rows) and $8\ dB$ (last two rows), obtained from 400 independent Monte Carlo runs ($N = 1,024$ samples per each run). As we see from Table II, low SNR increases the variance of b. Even if the estimator (14) guarantees reliable estimates of the modulation index for a pure FM tone observed in additive white Gaussian noise (see [11]), when it is applied to the "data" $x_M(n;\tau_l)$ all the cross-terms between $s(n)$ and $v(n)$, present in the ml-HIM of $x(n)$, contribute to the disturbance term, which consequently can no longer be considered white or Gaussian process. This seems to have deleterious effects on the performance of (14), as shown in Table II.

A possible remedy is to use the estimate ω_0 to select a new set of lags in such a way that $\beta < 1$, and then estimate b as in Step 6. We repeated the simulations, this time with $\tau_1 = 256$, so that $\beta = 0.2435$, and we obtained

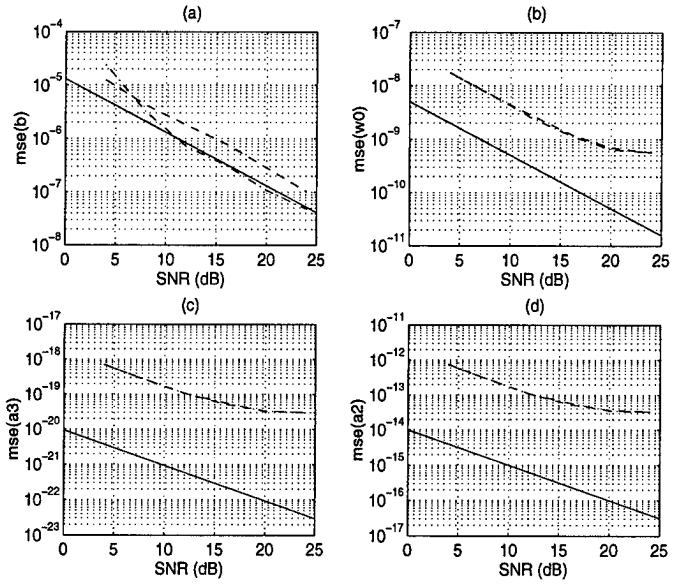


Figure 3. Mean square error for known ρ (dashed), unknown ρ (dashdotted), and CRLBs (solid) vs. SNR.

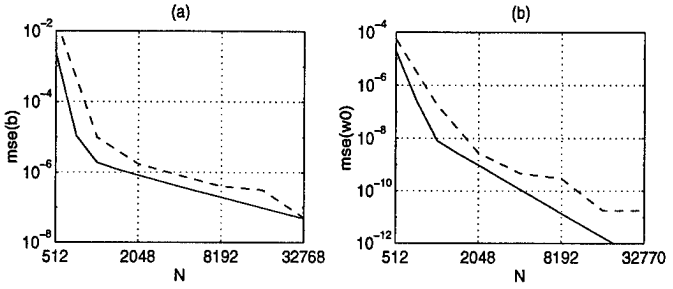


Figure 4. Mean square error (dashed) and CRLBs (solid) vs. N.

bias$(\widehat{b})$ = 0.1384, var$(\widehat{b})$ = 0.6407 for SNR = 16 dB, and bias$(\widehat{b})$ = 0.1469, var$(\widehat{b})$ = 3.0400 for SNR = 8 dB. The accuracy improvement (at least for the low SNR case) is now noticeable.

Table I - PPS parameters ($SNR = 16\ dB$ and $8\ dB$)

	$\widehat{a}_2$	$\widehat{a}_1$	$\widehat{a}_0$
Bias	$-1.8236 \cdot 10^{-8}$	0.0 174	-0.2938
Var	$2.327 \cdot 10^{-15}$	0.0028	0.0376
Bias	$-4.3700 \cdot 10^{-7}$	-0.0226	-0.4019
Var	$5.6254 \cdot 10^{-13}$	0.0276	0.0765

Table II - FM parameters ($SNR = 16\ dB$ and $8\ dB$)

	$\widehat{b}$	$\widehat{\omega}_0$	$\widehat{\psi}_0$
Bias	−1.3579	$-1.9890 \cdot 10^{-5}$	0.0574
Var	0.0460	$2.7497 \cdot 10^{-10}$	$4.7490 \cdot 10^{-5}$
Bias	−4.1559	$-7.6639 \cdot 10^{-5}$	0.0549
Var	$1.1172 \cdot 10^{4}$	$3.1892 \cdot 10^{-8}$	0.0106

4. CONCLUSIONS

Parameter estimation of a class of nonstationary complex signals whose phase can be modeled as a superposition of a polynomial term and a sinusoidal frequency modulated term was addressed. Previously proposed methods for estimating the FM parameters (and in particular the modulation index) do not include PPS components, and, viceversa, the standard approach based on the HAF for PPS parameter estimation cannot work when the FM component is present. The proposed method handles the more general scenario of hybrid FM-PPS signals exploiting the properties of the multi-lag HAF. This approach allows one to decouple estimation of the FM parameters from that of the PPS, with a noticeable reduction in computational complexity. The redundancy offered by the multi-lags was also used to reduce the FM term to a narrowband process, and thereby improves the accuracy of the modulation index estimation. The exact Fisher information matrix for the FM-PPS parameters was also derived along with an asymptotic form of the CRLBs. Computer simulations were carried out to compare the performance of the proposed methods with the relevant CRLBs. The model adopted herein is potentially useful for practical radar/sonar modeling and target classification.

REFERENCES

[1] S. Barbarossa, "Detection and estimation of the instantaneous frequency of polynomial phase signals by multilinear time-frequency representations," *Proc. of IEEE-SP Workshop on Higher Order Statistics*, pp. 168-172, Lake Tahoe, CA, June 1993.

[2] S. Barbarossa, A. Porchia, and A. Scaglione, "Multiplicative multilag higher-order ambiguity function," *Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 3022-3025, Atlanta, GA, May 1996.

[3] M. R. Bell, R. A. Grubbs, "JEM modeling and measurement for radar target identification," *IEEE Trans. on AES*, vol. 29, pp. 73-87, Jan. 1993.

[4] D. R. Brillinger, *Time Series: Data Analysis and Theory*, Holden-day Inc., San Francisco, 1981.

[5] S.-R. Huang, R. M. Lerner, and K. J. Parker, "On estimating the amplitude of harmonic vibration from the Doppler spectrum of reflected signals", *J. of the Acoust. Soc. Am.*, vol. 88, pp. 2702-2712, Dec. 1990.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, NJ, 1993.

[7] S. Palumbo, S. Barbarossa, A. Farina, and M. R. Toma, "Classification techniques of radar signals backscattered by helicopter blades", *Proc. of Int. Symp. on Digital Signal Processing*, London, UK, July, 1996.

[8] B. Porat, *Digital Processing of Random Signals, Theory & Methods*, Prentice Hall, NJ, 1994.

[9] F. G. Stremler, *Introduction to Communication Systems*, third edition, Addison Wesley, 1990.

[10] G. T. Zhou, G. B. Giannakis and A. Swami, "On polynomial phase signals with time-varying amplitudes," *IEEE Trans. on Signal Processing*, vol. 44, pp. 848-861, April 1996.

[11] G. T. Zhou and G. B. Giannakis, "Parameter estimation of FM signals using cyclic statistics," *IEEE-SP Intl. Symposium on Time-Frequency and Time-Scale Analysis*, pp. 248-251, Philadelphia, Pennsylvania, October 25-28, 1994.

[12] J.-E. Wilbur and R. J. McDonald, "Nonlinear analysis of cyclically correlated spectral spreading in modulated signals", *J. of the Acoust. Soc. of America*, vol. 92, pp. 219-230, July 1992.

Acknowledgements: The authors would like to thank Profs. G. Tong Zhou and Alfonso Farina for useful comments. The work in this paper was supported by ONR Grant No. N00014-93-1-0485.

HIGHER-ORDER STATISTICS AND EXTREME WAVES

E.J. Powers, I.-S. Park, S.Im, S. Mehta, and E.-J. Yi
Department of Electrical and Computer Engineering and
Electronics Research Center
The University of Texas at Austin
Austin, Texas 78712-1084

Abstract

A sparse second-order time-domain Volterra model is used to decompose a random (sea) wave train into its first- and second-order components. Extreme waves are shown to result from short-term phase locking of the first- and second-order components. The feasibility of using a wavelet-based bicoherence "spectrum" to detect the strong, but short lived, phase coupling is investigated. The results are encouraging and suggest the wavelet-based bicoherence is a topic worth considering further.

1. Introduction

Extreme waves are very, very large amplitude sea waves that are occasionally generated during severe storm conditions such as hurricanes. They are of immense practical importance, since such waves are capable of capsizing ships and offshore structures used in petroleum production. Despite their great importance, our understanding of such extreme waves is quite poor for two basic reasons. First, they are relatively rare, and secondly they involve nonlinear phenomena. In this paper, we report on the application of a number of higher-order statistics based techniques that have been utilized to analyze and interpret extreme wave experimental data. We make use of both Fourier-based and wavelet-based higher order spectra, and frequency- and time-domain Volterra models, to provide new physical insight into important nonlinear wave interaction phenomena that underlies extreme wave generation. Such experimental insight has heretofore been unavailable.

Time series data of instantaneous wave elevation were collected from a series of probes separated one meter apart in the direction of propagation. The experiments were carried out in a model basin as described in Sec. 2. The generated random waves were unidirectional and were generated in such a way as to model (scale 1:54.5) a 100-year Gulf of Mexico storm. In Sec. 4 we review the use of higher-order spectra to determine a frequency-domain Volterra model which models the first-order (i.e., linear) and second-order (i.e., quadratic) physics which govern wave propagation between the two probes. Since the waves at any point are non-Gaussian due to prior nonlinear wave interactions, determination of the Volterra kernels must take this into account. Thus the wave statistics must be characterized by a hierarchy of higher-order spectra up to 4th order. Furthermore, the model must be orthogonal (here we make use of a modified Gram-Schmidt procedure) to facilitate the decomposition of the experimentally observed wave power spectrum into its constituent first-order and second-order components. Such decomposition indicates that most of the wave power is associated with first-order (i.e., linear) phenomena, while the power associated with second-order effects is considerably smaller and resides in a higher frequency band.

Next, in Sec. 5, we utilize the same experimental time series data to determine a sparse time-domain Volterra model in order to gain insight in the temporal relationship between the first-order and second-order wave components. These components correspond to the output of the linear and quadratic Volterra filters, respectively. The results indicate a second-order component oscillating (the wave spectrum is a JONSWAP spectrum [1] which is moderately narrow band) at roughly twice the frequency of the first-order component. However, when the occasional extreme wave is present, the relationship between first- and second-order components changes dramatically. We observe in this case that the second-order component is large in amplitude and phase-locked to the first-order component in such a way that constructive interference takes place. Such constructive interference lasts only for a cycle or so and gives rise to an extreme wave.

Since the first- and second-order components are phase-locked over a relatively short period of time and since the phase-locking (or phase coupling) is relatively weak over most of the remaining experimental time series record, Fourier- transform-based bicoherence spectra are really not suitable for detecting such short-lived nonlinear interac-

0-8186-8005-9/97 $10.00 © 1997 IEEE

tions. For this reason, we utilize a wavelet-based bicoherence spectra in Sec. 6 in an attempt to detect the short-lived but high degree of phase coupling associated with the generation of extreme waves. The wavelet-based bicoherence results are quite encouraging and suggest this may be a powerful tool with which to detect and quantify short-time-duration nonlinear events. The paper is summarized in Sec. 7.

2. Experimental Setup

The experiments were conducted at the Offshore Technology Research Center's Model Basin located at College Station, Texas. The data utilized in this paper correspond to unidirectional random waves, with a JONSWAP spectrum [1]. Three probes, separated one meter apart, were utilized to record the waves as they propagate past the probes. An example of the wave elevation time series is shown in Fig. 1 and its corresponding spectrum is shown in [2]. Of particular interest is the large-amplitude wave located at approx. 5365 secs in Fig. 1, and the extent to which possible nonlinear interactions between first- and second-order propagating wave components might be responsible for generating such a large-amplitude extreme wave.

3. Decomposition Via Volterra Models

To quantify the relative roles of first- and second-order wave components in the generation of large amplitude extreme waves we must first decompose the observed waves into their first- and second-order components. In Sec. 4 the decomposition is carried out in the frequency domain and in Sec. 5, the time domain. The key idea is to take the data from two probes and construct a second-order (i.e., quadratic) Volterra model (see Fig. 2), where the linear component models the first-order (i.e., linear) wave propagation physics and the quadratic component models second-order (i.e., quadratic) propagation phenomena. To the extent that we have a good model, so that the error is small in Fig. 2, we may regard the outputs of the linear and quadratic Volterra filters as a decomposition of the observed random waves into their first- and second-order components.

4. Frequency-domain Volterra Model

We have previously reported on a second-order frequency-domain orthogonal Volterra-like model of the first- and second- order wave physics [2]. Of particular interest are the plots of the linear and quadratic coherence spectra (see Fig. 4 in [2]). These plots, combined with the autopower spectrum, indicate that there is a small amount of second-order power oscillating at the twice the characteristic frequencies associated with the waves. Since, the second-order power is approximately two-orders of magnitude smaller than the first-order power, one might conclude that second-order effects are not important. However, since determination of the higher-order spectral moments (necessary to compute the Volterra transfer functions) is based on the Fourier transform, no temporal information is available regarding a possible short-term interaction of first- and second-order wave components. Since the large-amplitude extreme wave in Fig. 2 is localized in time, we must utilize other approaches. Thus, in Sec. 5 we consider a time domain Volterra model, and in Sec. 6 a wavelet-based bicoherence.

5. Time-domain Volterra Model

Using wave elevation time series data from the two probes we next construct a sparse discrete time-domain second-order Volterra model described by,

$$y[n] = \sum_{i=0}^{N-1} h_1[i]x[n-i] + \sum_{i=0}^{N-1}\sum_{j=i}^{N-1} h_2[i,j]x[n-i]x[n-j] + e[n], \quad (1)$$

where $h_1[i]$ and $h_2[i,j]$ correspond to the linear (first-order) and quadratic (second-order) Volterra kernels (filter coefficients) respectively. The terms $x[n]$ and $y[n]$ correspond to the wave elevation time series recorded at the upstream and downstream probes, respectively. Orthogonal search techniques are used to identify the most significant linear and quadratic filter coefficients. Of a total of 30 linear, and 465 quadratic coefficients, the orthogonal search identified the 9 most significant linear and 15 most significant quadratic coefficients. Thus we have a sparse model. Next we input the upstream wave elevation into the sparse Volterra model. Of particular interest to us in this paper is the output of the linear and quadratic filters. The relevant time traces for a third probe located one meter further downstream are shown Fig. 3 where we have zoomed in on the region containing the large-amplitude extreme wave. Note that, with the exception of the time surrounding the extreme wave, the second-order component is relatively small in amplitude relative to the first-order component. Furthermore, the linear component rather closely models the observed wave elevation. However, this situation changes significantly in the presence of the extreme wave in that the first-order wave component underpredicts the wave peak, and overpredicts the depth of the following wave trough. Note, however, that the second-order wave component increases in amplitude and adds to the first-order component to correctly predict the amplitude of the extreme wave (constructive interference). Furthermore, since the second-order component

oscillates at roughly twice the frequency of the first-order component, it is out of phase with the first-order component at the following wave trough. As a result, the summation of the out-of-phase first-and second-order components (destructive interference) correctly predicts the relatively shallow trough depth. These results demonstrate that the generation of large-amplitude extreme waves is associated with a short-term nonlinear wave interaction characterized by a high-degree of phase coupling between first- and second-order components. These experimental results agree closely with numerical and analytical studies of such phenomena [3]. Furthermore, these results provide confidence that the orthogonal search technique identified those linear and quadratic filter coefficients which are most important in terms of modeling the linear and nonlinear physics.

6. Wavelet-based Bicoherence

Since the phase locking (or phase coupling) associated with very large-amplitude extreme waves occurs over a relatively short time period, Fourier-based bicoherence analysis fails to detect the short-term nonlinear interaction. For this reason we employ a wavelet-based bicoherence time-scale "spectrum" first put forward in [4]. We utilize the continuous wavelet transform defined by

$$X_w(a,\tau) = \frac{1}{\sqrt{|a|}} \int x(t)\psi^*(\frac{t-\tau}{a})dt \qquad (2)$$

where $x(t)$ is the wave evaluation time series data, and a is the variable scale. The quantity $\psi(t)$ is a Morlet wavelet defined by

$$\psi(t) = e^{i2\pi t}e^{-t^2/2} \qquad (3)$$

which obviously corresponds to a complex exponential multiplied by a Gaussian.

Next, following [5], we define the wavelet-based bispectrum as

$$B^w_{xxx}(a_1,a_2) = \int_T X_w(a,\tau)X^*_w(a_1,\tau)X^*_w(a_2,\tau)d\tau \quad (4)$$

where T is the time duration over which the wavelet-based bispectrum is computed. The scales a, a_1, and a_2 satisfy the scale selection rule

$$a^{-1} = a_1^{-1} + a_2^{-1} \qquad (5)$$

This is analogous to the frequency selection rule ($f = f_1 + f_2$) associated with Fourier-based bispectral analysis of stationary data, and with the frequency selection rule ($f = f_1 + f_2$) associated with frequency mixing resulting from quadratic nonlinear interactions. Since the Morlet wavelet has a well defined peak in the frequency domain, frequency and scale are inversely related (i.e, $f \propto a^{-1}$). The wavelet-based auto-bicoherence is given by analogy to [5, see also [4]]

$$(b^w_{xxx}(a_1,a_2))^2 = \frac{|B^w_{xxx}(a_1,a_2)|^2}{(\int_T |X_w(a_1,\tau)X_w(a_2,\tau)|^2 d\tau)(\int_T |X_w(a,\tau)|^2 d\tau)} \qquad (6)$$

Next the wavelet-based bicoherence was computed for several wave-elevation segments of 150 secs duration in the vicinity of the large-amplitude extreme wave. Plots of the time series data associated with a selected segment, and the corresponding wavelet-based bispectrum and bicoherence are shown in Figs. 4 and 5. Note that Fig. 5 contains the large-amplitude extreme wave shown in Fig. 1. Observe that the peaks in the wavelet-based bispectrum and bicoherence are fairly broad, due both to the random non-narrow bandwidth of the random sea waves and the finite time duration of the analyzing wavelets. Of particular significance is the fact that the maximum vertical scale for the bicoherence is greater in Fig. 5 than in Fig. 4. This suggests, of course, that the wavelet-based bicoherence is sensitive to the short-term but high degree of phase coupling characteristic of the strong nonlinear interaction associated with the generation of the large-amplitude extreme wave in the 150 secs segment starting at 5250 secs.

Next we follow a procedure used in studies of plasma turbulence, whereby we sum the wavelet-based bicoherence over the upper triangle in the bifrequency plane. Low frequencies are excluded because the wavelet-based bicoherence contains several spurious peaks due to the fact that there is virtually no power in the JONSWAP spectrum at these low frequencies. The summed bicoherence for 16 overlapping (50%) segments of 150 sec duration are tabulated in Table 1. The values are normalized to the largest value, associated with the segment extending from 5250 to 5400 secs. This, of course, is the segment containing the largest amplitude wave. All other segments have a lower value of summed bicoherence thereby suggesting that the wavelet-based bicoherence is indeed sensitive to the high-degree of phase coupling characteristic of the short-time strong nonlinear interactions associated with extreme waves.

7. Conclusion

The results on the use of wavelet-based bicoherence "spectra" to detect strong, but short-lived, phase coupling associated with the generation of extreme waves are of a preliminary nature. Nevertheless, the early results are sufficiently encouraging to warrant further investigation of this technique.

Start time [sec]	Normalized summed bicoherence	Start time [sec]	Normalized summed bicoherence
4500	0.5753	5100	0.3264
4575	0.5604	5175	0.3656
4650	0.4504	5250	1.00
4725	0.4045	5325	0.8745
4800	0.3969	5400	0.6807
4875	0.4185	5475	0.8856
4950	0.4813	5550	0.5797
5025	0.4486		

Table 1. Table of the summed bicoherence for 50% overlapping time intervals of 150 sec duration.

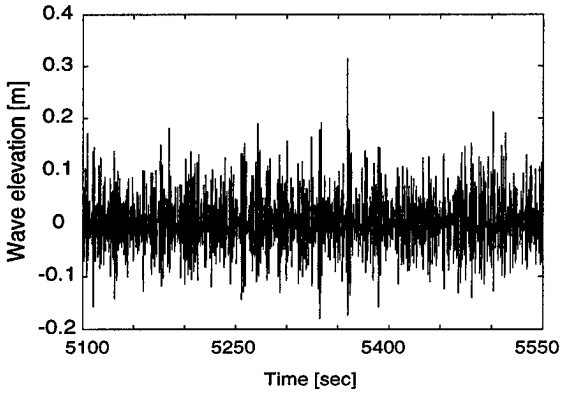


Figure 1. A portion of the wave elevation time trace containing a large-amplitude extreme wave at ~ 5365 secs.

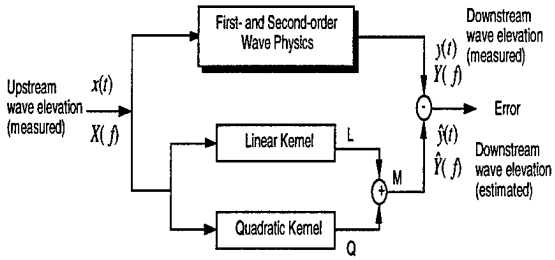


Figure 2. A second-order Volterra model.

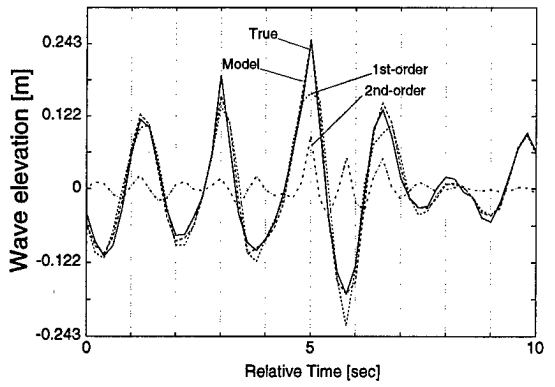


Figure 3. True, that is, observed, wave elevation at downstream probe, wave elevation "predicted" by a sparse second-order Volterra model, and the 1st-order and 2nd-order wave components in the vicinity of the large-amplitude extreme wave.

Acknowledgment

This study was supported by the National Science Foundation Engineering Research Centers Program Grant Number CDR-8721512, through the Offshore Technology Research Center (OTRC). The Volterra modeling and higher-order wavelet-based spectral analysis was supported in part by the Joint Service Electronic Program AFOSR-F-4962-95-C-0045. The authors would also like to thank Dr. Arun Duggal, Mr. Peter Johnson and the OTRC facility staff for their assistance in providing the data.

References

[1] Subrata K. Chakrabarti, "Nonlinear Methods in Offshore Engineering," Elsevier, Amsterdam, 1990.

[2] E.J. Powers, S. Im, A. Duggal, and P. Johnson, "Application of High-Order Statistics to Nonlinear Hydrodynamics," *Proceedings of IEEE Signal Processing ATHOS workshop on Higher-Order Statistics.*, Begur, Spain, June 12-14,1995, pp. 1758–1769.

[3] C. T. Stansberg, "Second-order Effects in Random Wave Modeling," *Proc. Int. Symp. Waves - Physical and Numerical Modeling*, Vol. 2, pp. 798-802, Vancouver, 1994.

[4] B. Ph. van Milligen, C. Hidalgo, and E. Sanchez, *Nonlinear Phenomena and Intermittency in Plasma Turbulence*, Physical Review Letters, 74, pp. 395-398, January 16, 1995.

[5] Y.C. Kim and E.J. Powers, "Digital Bispectral Analysis and Its Applications to Nonlinear Wave Interactions," *IEEE Trans. on Plasma Science*, Vol. PS-7, pp. 120–131, June 1979.

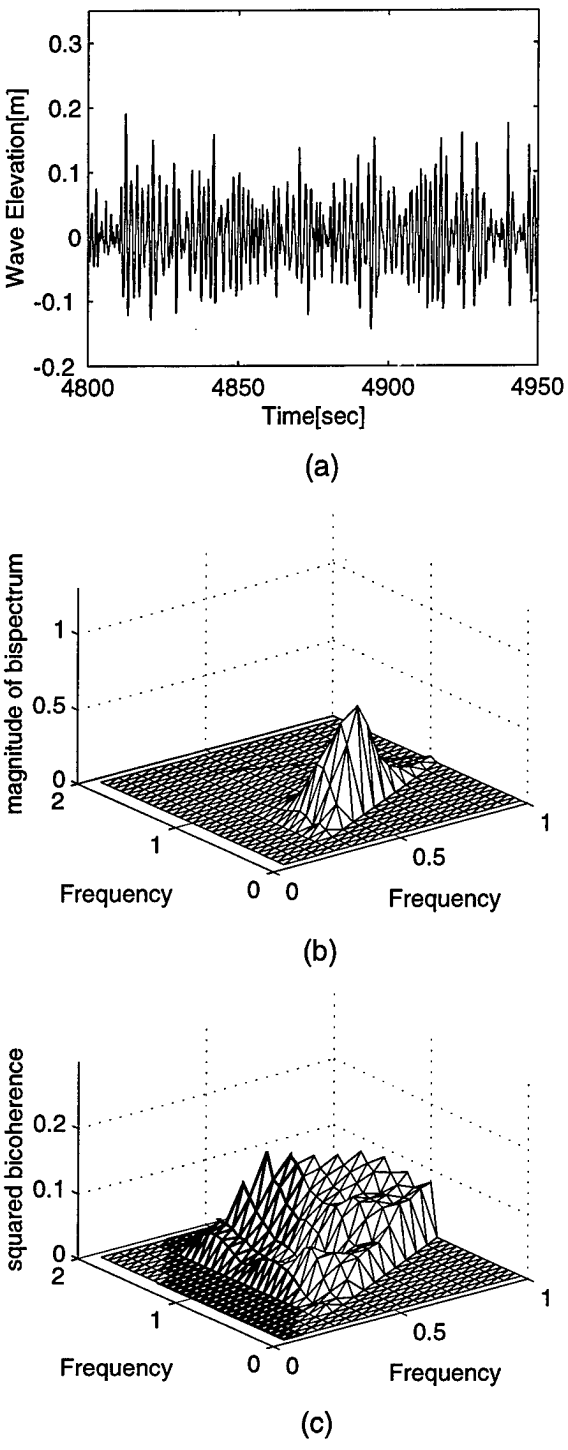


Figure 4. Time-trace, wavelet based bispectrum, and wavelet based bicoherence for the time interval 4800 to 4950 secs.

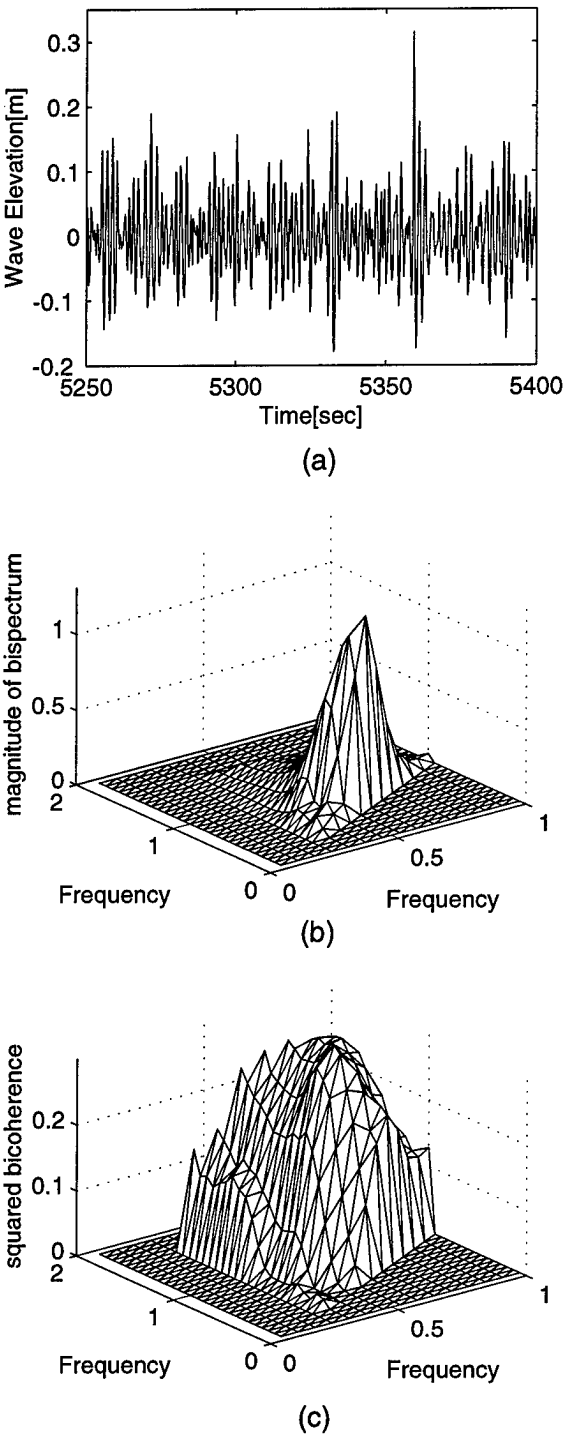


Figure 5. Time-trace, wavelet based bispectrum, and wavelet based bicoherence for the time interval 5250 to 5400 secs,the interval containing the large-amplitude extreme wave.

EXPLORING LAG DIVERSITY IN THE HIGH-ORDER AMBIGUITY FUNCTION FOR POLYNOMIAL PHASE SIGNALS

G. Tong Zhou[1] *Yang Wang*[2]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA.
[2]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA.

ABSTRACT

High-order ambiguity function (HAF) is an effective tool for retrieving coefficients of polynomial phase signals (PPS). The lag choice is dictated by conflicting requirements: a large lag improves estimation accuracy but drastically limits the range of the parameters that can be estimated. By using two (large) co-prime lags and solving linear Diophantine equations using the Euclidean algorithm, we are able to recover the PPS coefficients from aliased peak positions without compromising the dynamic range and the estimation accuracy. Separating components of a multi-component PPS whose phase polynomials have very similar leading coefficients has been a challenging task, but can now be tackled easily with the two-lag approach. Numerical examples are presented to illustrate the effectiveness of our method.

1. INTRODUCTION

Polynomial phase signals (PPS) have been studied extensively in recent years as a means of modeling certain nonstationary processes [2], [4], [9], [10]. One of the primary motivations for studying PPS comes from Doppler radar applications. Although the continuous-time transmitted radar signal does not have polynomial phase, the samples taken at the matched filter output of a pulsed radar system give rise to a discrete-time PPS when the target is moving (see e.g., [7, pp. 58-65]). The polynomial coefficients are then related to the kinetic parameters of the target [7, p. 59]. PPS modeling of SAR images has also appeared in [6], [3].

A discrete-time single component PPS of order M has the form

$$y(t) = A\, e^{j \sum_{m=0}^{M} a_m\, t^m}, \qquad t = 0, 1, \ldots, T-1. \quad (1)$$

Set $y^{(*q)}(t) = y(t)$ for even q and $y^{(*q)}(t) = y^*(t)$ for odd q, where $y^*(t)$ denotes the conjugate of $y(t)$. Peleg and Porat have shown that the *high-order instantaneous moment* (see [5, Ch. 12]),

$$\mathcal{P}_M[y(t);\tau] \triangleq \prod_{q=0}^{M-1} [y^{(*q)}(t - q\tau)]^{\binom{M-1}{q}}, \quad (2)$$

reduces the Mth-order PPS in (1) to a single harmonic at frequency $\tilde{\omega} = M!\, a_M \tau^{M-1}$. The generalized FS coefficient of $\mathcal{P}_M[y(t);\tau]$ defined as

$$P_M[y;\alpha,\tau] \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{P}_M[y(t);\tau]\, e^{-j\alpha t}, \quad (3)$$

is referred to as the *high-order ambiguity function* (HAF). The magnitude of (3) peaks at $\alpha = \tilde{\omega}$. Finding this peak thus enables us to estimate the coefficient a_M. Once a_M is found, we demodulate $y(t)$ by $\exp\{-j a_M t^M\}$ to reduce the order of the PPS by one, and repeat the above procedure until all remaining coefficients of the phase polynomial are found.

To avoid aliasing in estimating a_M, one must limit $\tilde{\omega}$ to within $[-\pi, \pi)$ and hence a_M to within $[-\varepsilon\pi, \varepsilon\pi)$ with $\varepsilon = \tau^{-(M-1)}/M!$. Therefore τ has to be small in order for the dynamic range of a_M to be reasonably large. On the other hand, noise is often present in physical systems and we observe $x(t) = y(t) + v(t)$, where $v(t)$ is assumed to be zero-mean. The coefficient a_M is then estimated using $\mathcal{P}_M[x(t);\tau]$ in place of $\mathcal{P}_M[y(t);\tau]$. It is shown by Peleg that in the presence of noise, the best estimate $\widehat{a}_M$ of a_M, in the sense of minimizing the asymptotic variance of $\widehat{a}_M$, is attained by taking $\tau = T/M$ for $M = 2, 3$, or $\tau = T/(M+2)$ for $4 \le M \le 10$, see [5, p. 398]. Taking such a large τ will drastically limit the dynamic range of a_M. A dilemma is thus reached here regarding our choice of τ.

2. JOINT PARAMETER ESTIMATION USING CO-PRIME LAGS

The above dilemma is easily resolved by using two large lags τ_1 and τ_2 that are co-prime. Using our approach, the *identifiability condition* for a_M (i.e., the range of a_M that can be recovered uniquely) is $[-\pi/M!, \pi/M!)$. This offers a much larger dynamic range than $a_M \in [-\varepsilon\pi, \varepsilon\pi)$, $\varepsilon = \tau^{1-M}/M!$, when only one lag is used. Using two large lags, the peaks are likely to be at aliased positions, but the correct a_M can nonetheless be recovered. Both the theory and simulations confirm that while a much relaxed identification condition is maintained, our estimation accuracy is at least as good as that with one large lag. We wish to point out here that multiple lags are also used in [1] to improve accuracy and resolve identifiability problems with multicomponent PPS.

We now describe our method. Let $y(t)$ be a PPS of order M. Observe that for any given lag τ, the peak location of $P_M[y;\alpha,\tau]$ is at $M!a_M\tau^{M-1}$, in theory. But due to the inherent wrap-around effect of the DFT, we actually observe

0-8186-8005-9/97 $10.00 © 1997 IEEE

the peak location at $b \in [-\pi, \pi)$, with $M! a_M \tau^{M-1} = b + 2\pi k$ for some integer k. Suppose that two co-prime lags τ_1, τ_2 are now used, and from which we obtain two (possibly aliased) peak locations b_1, $b_2 \in [-\pi, \pi)$ from the HAFs. Hence there exist k_1, $k_2 \in \mathbf{Z}$ such that

$$M!\, a_M \tau_1^{M-1} = 2\pi k_1 + b_1, \quad (4)$$

$$M!\, a_M \tau_2^{M-1} = 2\pi k_2 + b_2. \quad (5)$$

Multiplying (4) by τ_2^{M-1} and (5) by τ_1^{M-1} and then subtracting the resulting equations, we obtain

$$k_2 \tau_1^{M-1} - k_1 \tau_2^{M-1} = \frac{b_1 \tau_2^{M-1} - b_2 \tau_1^{M-1}}{2\pi} \triangleq N. \quad (6)$$

Note that N in the above equation is an integer, in theory, but may not be so in the presence of noise, and we round it off to the nearest integer in that case. Next, we show how the condition $M! a_M \in [-\pi, \pi)$ and the co-primeness of τ_1, τ_2 allow us to uniquely solve for k_1 and k_2.

First we find two integers n_1, n_2 such that $n_2 \tau_1^{M-1} - n_1 \tau_2^{M-1} = 1$ using the Euclidean algorithm which is a standard algorithm in number theory, see e.g. [8] or the Appendix. It is well known that the solutions k_1, k_2 of the linear Diophantine equation (6) are completely characterized by

$$k_1 = n_1 N + l\tau_1^{M-1}, \quad k_2 = n_2 N - l\tau_2^{M-1}, \quad (7)$$

where $l \in \mathbf{Z}$. Since k_1 increases by τ_1^{M-1} every time l increases by 1, there exists a unique k_1 such that $2\pi k_1 + b_1 \in [-\pi\tau_1^{M-1}, \pi\tau_1^{M-1})$. Denote this k_1 by k_1^*. It follows from the identifiability condition $M! a_M \in [-\pi, \pi)$ and (4) that

$$M! a_M \tau_1^{M-1} = 2\pi k_1^* + b_1, \text{ and } a_M = \frac{2\pi k_1^* + b_1}{M! \tau_1^{M-1}}. \quad (8)$$

We now summarize the two-lag algorithm for estimating a_M below:

Step 1. Choose co-prime integers τ_1 and τ_2 that are close to the optimal lag: T/M for $M = 2$, 3, or $T/(M+2)$ for $4 \le M \le 10$. Assume that $\tau_1 > \tau_2$. Find the (possibly aliased) peak locations b_1, $b_2 \in [-\pi, \pi)$ from $|P_M[x; \alpha, \tau_1]|$ and $|P_M[x; \alpha, \tau_2]|$, respectively.

Step 2. Find two integers n_1, n_2 satisfying $n_2 \tau_1^{M-1} - n_1 \tau_2^{M-1} = 1$ using the Euclidean algorithm.

Step 3. Solve for the unique $k_1 = n_1 N + l\tau_1^{M-1}$ with $l \in \mathbf{Z}$ such that $2\pi k_1 \in [-\pi\tau_1^{M-1}, \pi\tau_1^{M-1})$, where N is defined in (6). Let

$$a_M = \frac{2\pi k_1 + b_1}{M! \tau_1^{M-1}}.$$

Finally, we remark that there are easy ways to find τ_1 and τ_2 that are co-prime. For example, one may take any two consecutive integers, or take one of them a power of 2 while the other an odd integer.

3. APPLICATION TO MULTI-COMPONENT PPS

The HAF method is equally effective in estimating parameters of a multi-component PPS. Although cross-terms are created in $\mathcal{P}_M[y(t); \tau]$ when $y(t)$ has multiple components, it is shown in [9] that these cross-terms do not generally give rise to peaks in the corresponding HAF $P_M[y; \alpha, \tau]$. Therefore the HAF method works for multi-component PPS. Our two-lag approach thus offers the same advantage in this setting.

In the past, HAF has been problematic for multi-component PPS whose two or more leading phase coefficients are very close. As an added advantage, the two-lag approach solves this problem with ease. To illustrate that, let $y(t)$ be an M-th order two-component PPS and assume that the two leading coefficients $a_{M,1}$, $a_{M,2}$ of the phase polynomials are very close. When τ is small, the peak locations $M! a_{M,1} \tau^{M-1}$ and $M! a_{M,2} \tau^{M-1}$ of $P_M[y; \alpha, \tau]$ are indistinguishable and one may be misled to believe that there is only one component present. However, by choosing a large τ, we magnify the difference between $a_{M,1}$ and $a_{M,2}$ and thus separating the two peaks. The two-lag method then recovers $a_{M,1}$ and $a_{M,2}$.

We wish to point out that there is a minor complication in the actual implementation of the algorithm. Suppose that τ_1, τ_2 are co-prime and for each τ_i we find two (possibly aliased) peak locations $b_{1,i}$, $b_{2,i}$ of the HAF $P_M[y; \alpha, \tau_i]$. It is not immediately clear which peaks from τ_1 correspond to those from τ_2. There are two possible pairings: $\{(b_{1,1}, b_{1,2}), (b_{2,1}, b_{2,2})\}$ or $\{(b_{1,1}, b_{2,2}), (b_{2,1}, b_{1,2})\}$. But notice that according to (6), $[b_{i,1}\tau_2^{M-1} - b_{j,2}\tau_1^{M-1}]/(2\pi)$ must be an integer if $(b_{i,1}, b_{j,2})$ is to be a correct pair. We choose the two pairs so that the above expression is the closest to being integers. Pairing ambiguity is thus resolved. Once the leading coefficients are found, we follow the procedure delineated in [9] to find all remaining parameters.

4. SIMULATIONS

We present here two numerical examples to illustrate the two-lag algorithm.

Example 1. We have $T = 512$ samples of a chirp $x(t) = \exp(ja_2 t^2) + v(t)$, where $a_2 = 0.1$ and $v(t)$ is the zero-mean i.i.d. complex Gaussian noise with variance 0.5. For a given lag τ, we form $P_2[x; \alpha, \tau]$ by taking the normalized (by data length T) DFT of the product process $x^*(t-\tau)x(t)$. Method I uses a single lag $\tau = 1$, and the corresponding $2a_2\tau = 0.2$ is free of the aliasing effect. The bias and standard deviation of the estimate $\widehat{a}_2$ calculated from 100 independent realizations were 4.5219×10^{-6} and 5.7433×10^{-5}, respectively. Method II uses two lags $\tau_1 = T/2 = 256$ and $\tau_2 = 201$, and aliasing occurs for both lags. We recover a_2 using the previously outlined two-lag algorithm, by solving the linear Diophantine equation (6). The bias and the standard deviation of our estimate $\widehat{a}_2$ calculated from the same 100 independent realizations were -1.1679×10^{-7} and 1.5475×10^{-6}, respectively. *Using the two-lag approach, we have gained an order of magnitude improvement in both the bias and*

the standard deviation!

Example 2. We have $T = 512$ samples of

$$x(t) = 1.2\ e^{j(0.4t^2-0.3t)} + e^{j(0.401t^2-t)} + v(t),$$

where $v(t)$ is the zero-mean i.i.d. complex Gaussian noise with variance 0.5. Figure 1 illustrates the magnitude of $P_2[x;\alpha,\tau]$ for different τ's and Table 1 shows the corresponding peak locations. A small lag such as $\tau_1 = 1$ reveals a single peak, but a large lag has the ability to split the peak into two, showing that there are actually two chirps. From the four peak locations obtained using two large co-prime lags τ_2 and τ_3, we recover the chirp coefficients $a_{1,2}$ and $a_{2,2}$. For the realization shown in Figure 1, we obtain $\widehat{a}_{1,2} = 0.40000115150160$, $\widehat{a}_{2,2} = 0.40099542805416$, which are very accurate. We have generated 100 independent realizations, and have found the bias and the standard deviation in $\widehat{a}_{1,2}$ to be -1.6063×10^{-8} and 1.4883×10^{-6}, respectively, and the bias and the standard deviation in $\widehat{a}_{2,2}$ to be -2.0721×10^{-6} and 1.8729×10^{-6}, respectively.

5. CONCLUSIONS

The high-order ambiguity function has been used frequently to estimate parameters of polynomial phase signal (PPS). It is well known that large lags offer better estimation accuracy but limit the range of parameters that can be identified. We demonstrate that by using two (large) co-prime lags and solving linear Diophantine equations by way of the Euclidean algorithm, we can recover the true PPS coefficients from aliased positions. Good estimation accuracy and dynamic range of the PPS coefficients are simultaneously achieved with this method. The two-lag method for multi-component PPS parameter estimation offers the same advantage in accuracy and dynamic range, even when the leading coefficients of the phase polynomials are very close, because we are free to use lag τ's to magnify their difference.

6. APPENDIX: LINEAR DIOPHANTINE EQUATIONS AND THE EUCLIDEAN ALGORITHM

Let $a > b > 0$ be co-prime integers and $N \in \mathbf{Z}$. The linear Diophantine equation

$$ax + by = 1, \tag{9}$$

can be solved easily using the Euclidean algorithm, which finds a pair of integers x_0, y_0 such that $ax_0 + by_0 = 1$. Details on linear Diophantine equations and the Euclidean algorithm can be found in textbooks on elementary number theory, see e.g. [8]. We illustrate the Euclidean algorithm here by way of an example with $a = 256$, $b = 191$.

First, we carry out the following successive quotients until the remainder is 1:

$$\begin{aligned} 256 &= 1 \times 191 + 65, \\ 191 &= 2 \times 65 + 61, \\ 65 &= 1 \times 61 + 4, \\ 61 &= 15 \times 4 + 1. \end{aligned}$$

Reverse the process to obtain

$$\begin{aligned} 1 &= 61 - 15 \times 4 \\ &= 61 - 15 \times (65 - 1 \times 61) \\ &= 16 \times 61 - 15 \times 65 \\ &= 16 \times (191 - 2 \times 65) - 15 \times 65 \\ &= 16 \times 191 - 47 \times 65 \\ &= 16 \times 191 - 47 \times (256 - 1 \times 191) \\ &= -47 \times 256 + 63 \times 191. \end{aligned}$$

Therefore we have found $x_0 = -47$ and $y_0 = 63$ such that $ax_0 + by_0 = 1$. Once x_0 and y_0 are found, the solutions to the linear Diophantine equation (9) are given by

$$x = Nx_0 + bl, \quad y = Ny_0 - al, \quad \forall\, l \in \mathbf{Z}. \tag{10}$$

The Euclidean algorithm converges very fast. In fact, it can be shown that the number of steps is at most 6.65 times the number of digits in b [8, p. 29].

REFERENCES

[1] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Multilag high-order ambiguity function for multicomponent polynomial phase signal modeling," *IEEE Transactions on Signal Processing,* 1997 (to appear); see also [6].

[2] P. M. Djurić and S. M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. on Acoustics, Speech, and Signal Processing,* vol. 38, pp. 2118-2126, 1990.

[3] M.Z. Ikram, K.A. Meraim, and Y. Hua, "Estimating Doppler parameters in SAR imaging for moving targets," *Proc. of the IEEE Nordic Signal Processing Symposium,* pp. 207-210, Espoo, Finland, September 1996.

[4] S. Peleg and B. Porat, "Estimation and classification of signals with polynomial phase," *IEEE Trans. on Information Theory,* vol. 37, pp. 422-430, 1991.

[5] B. Porat, *Digital Processing of Random Signals, Theory & Methods,* Prentice Hall, Englewood Cliffs, NJ, 1994.

[6] A. Porchia, S. Barbarossa, A. Scaglione, and G.B. Giannakis, "Autofocusing techniques for SAR Imaging using the multi-lag high-order ambiguity function," *Proc. of Intl. Conf. on Acoustics, Speech, and Signal Processing,* vol. IV, pp. 2086-2089, Atlanta, GA, May 7-11, 1996.

[7] A.W. Rihaczek, *Principles of High-Resolution Radar,* Peninsula Pub., CA, 1985.

[8] J.H. Silverman, *A Friendly Introduction to Number Theory,* Prentice Hall, 1997.

[9] Y. Wang and G.T. Zhou, "On the use of high order ambiguity function for multicomponent polynomial phase signals," *Signal Processing* (accepted).

[10] G.T. Zhou, G.B. Giannakis, and A. Swami, "On polynomial phase signals with time-varying amplitudes," *IEEE Trans. on Signal Processing,* vol. 44, no. 4, pp. 848-861, April 1996.

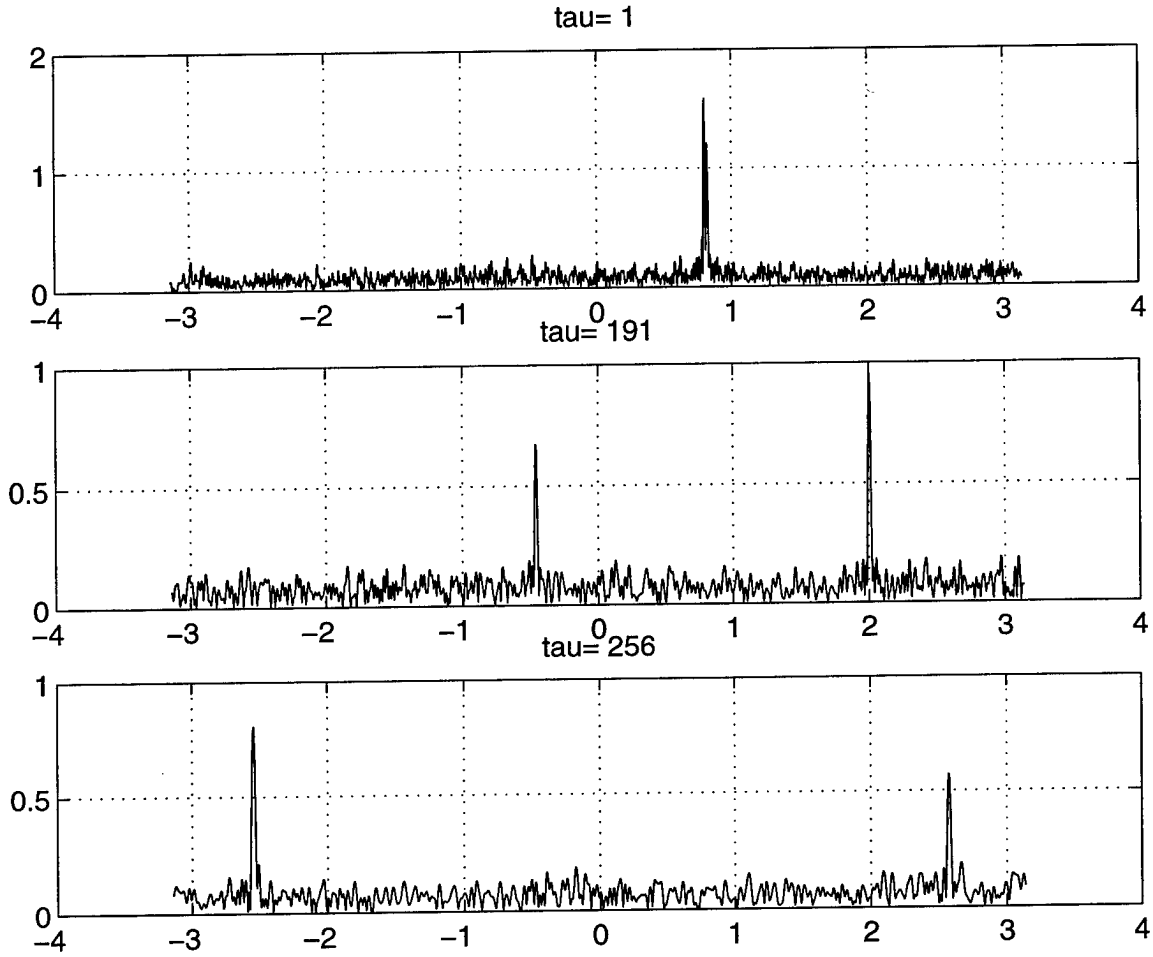


Figure 1. The ambiguity function (HAF of order 2) of a two-component chirp for various τ's.

Theoretical peak locations			Actual peak locations		
$\tau_1 = 1$	$\tau_2 = 191$	$\tau_3 = 256$	$\tau_1 = 1$	$\tau_2 = 191$	$\tau_3 = 256$
0.8000	2.0036	-2.5451	0.8007	2.0040	-2.5433
0.8020	2.3856	-2.0331	0.8007	2.3838	-2.0365

Table 1. Theoretical vs. experimental peak locations in Example 2 (c.f. Figure 1).

MPP: Nonstationary/Nonlinear/ Multidimensional Signals & Systems

Chair: Athina Petropulu
Drexel University, USA

Deblurring Two-Tone Images by A Joint Estimation Approach Using Higher-Order Statistics

Ta-Hsin Li* and Ke-Shin Lii†
*Department of Statistics, Texas A&M University, College Station, TX 77843
†Department of Statistics, University of California, Riverside, CA 92521

Abstract

A method is proposed for the restoration of linearly blurred two-tone images without requiring the knowledge of the blur parameters. The method jointly estimates the original image and the blur parameters based on some statistical parameters at the output of an inverse filter. Unlike some other blind image restoration procedures, the proposed method does not require the estimation or modeling of the statistical properties of the original image, yet can be justified even for non-i.i.d. images.

1. Introduction

In typical automatic recognition and identification systems, the input images are usually assumed to be blur-free. Sometimes, blurred images can be encountered due to the impairment of imaging systems and environment. If the blur is not properly removed, the recognizer's performance can severely deteriorate. For example, if the blurred text image in Fig. 1 is segmented, without deblurring, for automatic character recognition, the result could become unrecognizable even to trained human eyes (see Fig. 2).

In this article we are concerned with the restoration of linearly blurred two-tone images such as texts and bar codes [1]. We propose a restoration method that requires no prior knowledge of the blur parameters or the statistical properties of the original images. The method is shown to be able to handle not only images with *i.i.d.* (independent and identically distributed) pixels but also images with non-*i.i.d.* pixels that are often encountered in reality. Due to the use of higher-order statistics [2], the method is also capable of handling minimum as well as non-minimum phase blur filters. These features are very desirable for image restoration problems.

This work is an extension of some previous works on blind deconvolution (or equalization) of 1-D signals in the digital communications context [4], [5]. One of the major differences is that unlike the digital communications problem where the alphabets of the transmitted digital signal are known *a priori* at the receiver, the two tones of the original image in the image deblurring problem usually are not available. The proposed method in this article is able to take advantage of the two-tone information without requiring knowledge of the specific values of the tones – they are estimated from the blurred image jointly with the blur filter and the original image.

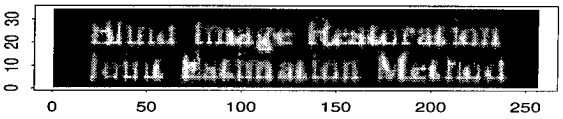


Figure 1. A blurred text image.

2. Deblurring Method

For linear degradations, the blurred image, denoted by $\mathcal{Y} := \{Y(m,n)\}$, can be regarded as the output from a blur filter

$$\mathcal{Y} = \mathcal{B} \otimes \mathcal{X},$$

where $\mathcal{B} := \{b(j,k)\}$ is the point-spread function (PSF) of the blur filter, $\mathcal{X} := \{X(m,n)\}$ is the original image, and $\otimes$ stands for two-dimensional convolution.

To deblur the image $\mathcal{Y}$ when the blur filter $\mathcal{B}$ is not available, we filter $\mathcal{Y}$ with an inverse filter $\mathcal{F} := \{f(j,k)\}$ to obtain $\mathcal{Z} := \{Z(m,n)\}$, where

$$\mathcal{Z} := \mathcal{F} \otimes \mathcal{Y}.$$

In this article, we propose the following *cost function* for the selection of $\mathcal{F}$:

$$J(\mathcal{F}) := \min_{a_2 = 1} \boldsymbol{a}^T H(\mathcal{Z})\, \boldsymbol{a},$$

where $\boldsymbol{a} := [a_0, a_1, a_2]^T$ is a constant vector and $H(\mathcal{Z})$ is a three-by-three matrix

$$H(\mathcal{Z}) := \begin{bmatrix} 1 & \mu_1(\mathcal{Z}) & \mu_2(\mathcal{Z}) \\ \mu_1(\mathcal{Z}) & \mu_2(\mathcal{Z}) & \mu_3(\mathcal{Z}) \\ \mu_2(\mathcal{Z}) & \mu_3(\mathcal{Z}) & \mu_4(\mathcal{Z}) \end{bmatrix},$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

with $\mu_i(\mathcal{Z})$, defined by

$$\mu_i(\mathcal{Z}) := E\{Z^i(m,n)\},$$

being the ith statistical moment of $\mathcal{Z}$.

3. Does the Method Work?

To see why minimizing $J(\mathcal{F})$ leads to a reasonable solution, it is instrumental to note that

$$\boldsymbol{a}^T H(\mathcal{Z})\,\boldsymbol{a} = E\{|\psi(Z^i(m,n))|^2\},$$

where $\psi(Z)$ is a two-degree polynomial defined by

$$\psi(Z) := a_0 + a_1 Z + a_2 Z^2.$$

Let β_1 and β_2 represent the zeros of $\psi(Z)$ so that

$$\psi(Z) = a_2\,(Z-\beta_1)\,(Z-\beta_2).$$

Then, it easy to see that if $\mathcal{F}$ equals the inverse of $\mathcal{B}$ or any of its shifted and/or scaled versions, the output image $\mathcal{Z}$ would coincide with $\mathcal{X}$ up to a shift and/or a scale and thus become a two-tone image. With β_1 and β_2 representing the two tones of $\mathcal{Z}$, one would obtain $\psi(Z(m,n)) = 0$, and therefore $J(\mathcal{F})$ would attain its minimum value zero.

On the other hand, if $J(\mathcal{F}) = 0$ for some $\mathcal{F} \neq 0$, then $\psi(Z(m,n)) = 0$ for some $\beta_1 \neq \beta_2$, so that $\mathcal{Z}$ must be a two-tone image. Therefore, for the justification of the method, it suffices to ensure that the two-tone image $\mathcal{Z}$ so obtained can only be a shifted and/or scaled version of the original two-tone image $\mathcal{X}$. Although one should not expect this for every two-tone image $\mathcal{X}$, it can be shown that it is true for a large class of two-tone images [3].

For example, if the pixels of $\mathcal{X}$ are *i.i.d.*, then $\mathcal{Z}$ must be a shifted and/or scaled version of $\mathcal{X}$ as soon as it becomes a two-tone image as a result of the filtering. To see this, we first note that

$$\mathcal{Z} = \mathcal{T} \otimes \mathcal{X},$$

with $\mathcal{T} := \{t(j,k)\} := \mathcal{F} \otimes \mathcal{B}$ being the combination of the deblurring filter with the blur filter. We proceed with the justification by the method of contradiction. Suppose that $\mathcal{Z}$ is a two-tone image but $\mathcal{T}$ is not a shifted and/or scaled delta function; for simplicity, let us assume that $t_0 := t(0,0) \neq 0$ and $t_1 := t(1,1) \neq 0$. Furthermore, without loss of generality, let us assume that $\mathcal{X}$ takes values in $\{0,1\}$. Then, at any location (m,n), we can write

$$\begin{aligned} Z(m,n) \;&=\; t_0\,X(m,n) \\ &\quad + t_1\,X(m-1,n-1) + R(m,n), \end{aligned}$$

where $R(m,n) := \sum{}' t(j,k)\,X(m-j,n-k)$; the sum $\sum'$ is computed for $(j,k) \neq (0,0),(1,1)$. (For simplicity, let's assume that $\mathcal{T}$ has a finite support.) Because the pixels of $\mathcal{X}$ are independent, it is always possible to find the locations (m_i,n_i), $(i=1,2,3,4)$, such that $R(m_i,n_i) = \text{const.}$ and $\{X(m_i,n_i), X(m_i-1,n_i-1)\} = \{(0,0),(0,1),(1,0),(1,1)\}$. This implies that the image $\mathcal{Z}$ would at least take all the values in $\{r, t_0+r, t_1+r, t_0+t_1+r\}$ and thus contradicts with the assumption that $\mathcal{Z}$ is a two-tone image.

The above argument can be generalized to non-*i.i.d.* images, with the relaxed requirement that any finite collection of the two tone values be visited by the pixels of $\mathcal{X}$ with a positive probability [3]. This requirement is analogous to the ergodicity requirements in the classic statistical theory for consistent parameter estimation. For two-tone images that satisfy this requirement, it can be shown that $\mathcal{Z}$ must be a shifted and/or scaled version of $\mathcal{X}$ once $J(\mathcal{F})$ is minimized with some filter $\mathcal{F} \neq 0$.

4. How to Compute the Minimizer?

The minimization of $J(\mathcal{F})$ is a nonlinear problem and therefore has to rely upon some iterative procedures. In the following, we outline a three-step iterative algorithm (TSIA) for the minimization of $J(\mathcal{F})$. More sophisticated variations of TSIA are possible, especially for Step 3, but we do not address them in this article.

1. With $\mathcal{Z}_k := \mathcal{F}_k \otimes \mathcal{Y}$, compute $\boldsymbol{a}_k$ such that

   $$\boldsymbol{a}_k := \arg\min\{\boldsymbol{a}^T H(\mathcal{Z}_k)\,\boldsymbol{a} :\; a_2 = 1\}.$$

2. Compute $\beta_{1,k}$ and $\beta_{2,k}$ as the zeros of the two-degree polynomial $\psi_k(Z) := a_{0,k} + a_{1,k} Z + Z^2$, where $\boldsymbol{a}_k := [a_{0,k}, a_{1,k}, 1]^T$.

3. Update the filter by a steepest descent method, with step size $\mu > 0$, such that

   $$\mathcal{F}_{k+1} := \mathcal{F}_k - \mu\,\nabla J_k(\mathcal{F}_k),$$

   where $J_k(\mathcal{F}) := \boldsymbol{a}_k^T H(\mathcal{Z})\,\boldsymbol{a}_k$ and $\mathcal{Z} = \mathcal{F} \otimes \mathcal{Y}$.

It is clear that TSIA jointly estimates the inverse blur filter, the original image, and the tones by $\mathcal{F}_k$, $\mathcal{Z}_k$, and $(\beta_{1,k}, \beta_{2,k})$, respectively. Note that in order to avoid the trivial solution of $\mathcal{F} = 0$ it may be necessary to impose a constraint, such that $\sum f(j,k) = 1$, on the filter. This constraint does not affect the viability of the method because the tones are automatically adjusted in TSIA according to the scale of $\mathcal{Z}$.

The computation of $\boldsymbol{a}_k$ in Step 1 is quite straightforward. In fact, since $\boldsymbol{a}^T H(\mathcal{Z}_k)\,\boldsymbol{a}$ is a quadratic function

of a_0 and a_1 under the constraint $a_2 = 1$, it is easy to show that $\boldsymbol{a}_k$ can be uniquely determined by

$$a_{0,k} = (c_{1,k}c_{3,k} - c_{2,k}^2)/(c_{2,k} - c_{1,k}^2)$$

and

$$a_{1,k} = (c_{1,k}c_{2,k} - c_{3,k})/(c_{2,k} - c_{1,k}^2),$$

where $c_{i,k} := E\{Z_k^i(m,n)\}$ is the ith central moment of $\mathcal{Z}_k$, $(i = 1, 2, 3)$.

As a simple variation of TSIA, one may update the tones after the filter is iterated $M > 1$ times. In other words, one may carry out Steps 1 and 2 once every M iterations of Step 3 with fixed tones. Experience shows that this variation reduces the computational complexity but still achieves reasonable results as long as M is not too large.

5. How Does It Work On Data?

To test the proposed method, let us consider the blurred text image shown in Fig. 1. The blur is caused by an autoregressive filter of the form

$$\begin{aligned} Y(m,n) &= \rho_1 Y(m-1,n) + \rho_2 Y(m,n-1) \\ &\quad - \rho_1\rho_2 Y(m-1,n-1) + X(m,n), \end{aligned}$$

where ρ_1 and ρ_2 are the filter parameters. In Fig. 1 we use $\rho_1 = \rho_2 = 0.7$.

For automatic character recognition, the raw images are typically required to be segmented into binary ones before they can be fed into a trained recognizer, such as an artificial neural network. A simple segmentation method is to classify each pixel as being black or white by comparing the pixel with a predetermined threshold. If the blurred image in Fig. 1 is segmented without deblurring, one would obtain a binary image that could be difficult for the recognizer to handle.

Fig. 2(a) shows a segmentation result for the blurred image in Fig. 1. As we can see, the direct segmentation of the blurred image produces a result that is unrecognizable perhaps even by trained human eyes. The difficulty is also reflected objectively by the proportion of misclassified pixels, which equals 16.19% for the 33-by-256 image in Fig. 2(a).

To improve the segmentation results, we apply the proposed deblurring method to the blurred image in Fig. 1. The filter $\mathcal{F}$ is restricted to be of size three-by-three, so that $f(j,k) = 0$ if $|j| \geq 2$ or $|k| \geq 2$. The nonzero coefficients of the initial filter $\mathcal{F}_0 := \{f_0(j,k)\}$ are $f_0(0,0) = 1.0$, $f_0(1,0) = f(0,1) = -0.2$, and $f_0(1,1) = 0.04$. Fig. 2(b) shows the segmentation result based on the output from the initial filter – the image is still difficult to recognize, with the proportion of misclassified pixels equal to 10.58%.

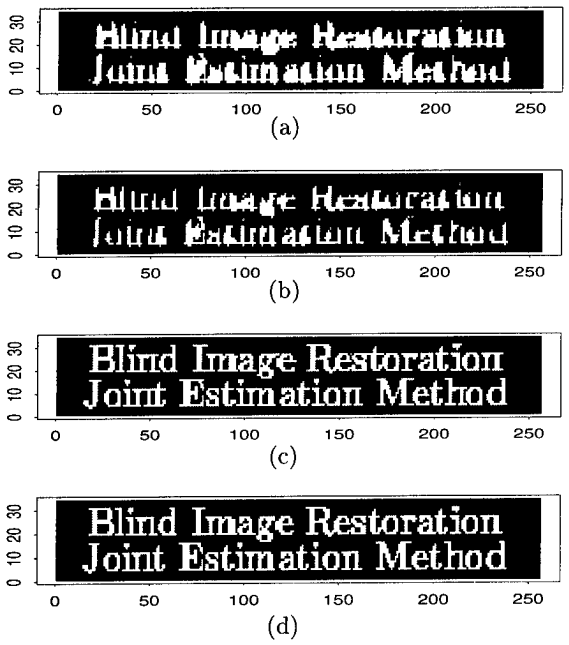


Figure 2. Segmentation results from: (a) the blurred image in Fig. 1; (b) the output of the initial filter; and (c) the deblurred image after 4,000 iterations of TSIA. The original text image is shown in (d).

Initiated with $\mathcal{F}_0$, we update the deblurring filter using the TSIA algorithm – the step size parameter is taken to be $\mu = 2 \times 10^{-10}$ and the tones are estimated once every $M = 30$ iterations of Step 3. Fig. 2(c) shows the segmentation result based on a deblurred image after 4,000 iterations. Compared with Figs. 2(a)-(b), and with the original two-tone (0-255) image in Fig. 1(d), the improvement on recognizability is quite significant, not only visually but also in terms of the proportion of misclassified pixels, which is now reduced to 0.57%.

The convergence behavior of TSIA is shown in Fig. 3 in terms of the cost function $J(\mathcal{F})$. As we can see, the cost function decreases monotonically as the iteration proceeds. The convergence rate is high at the beginning, but the algorithm slows down after a certain number of iterations (roughly 2,000 in this example). This behavior is typical for the steepest-descent-type algorithms [6].

To obtain the segmented images in Fig. 2, the segmentation threshold is determined so that the proportions of black and white pixels after the segmentation equal those in the original image. Fig. 4 and Table 1

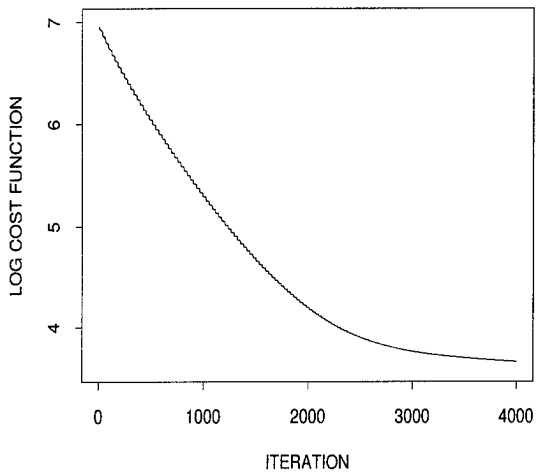


Figure 3. Plot of the cost function $J(\mathcal{F})$ (in logarithmic scale) against the number of iterations. The TSIA parameters are: $\mu = 2 \times 10^{-10}$ and $M = 30$.

Table 1. Percentage of mis-classified pixels

Threshold	Blurred	Initial	Deblurred
63	20.33	19.53	0.86
76	18.34	17.34	0.47
89	16.60	14.73	0.24
102	15.42	12.56	0.16
114	15.11	10.90	0.33
127	15.29	10.36	0.54
140	16.50	11.50	1.06

demonstrate the impact of the threshold on the segmentation results in terms of the classification errors. When the images are blurred, the segmentation error depends crucially on the threshold, as shown by the dashed and dotted lines in Fig. 4. The threshold is much less critical when the blur is properly removed, as shown by the solid line in Fig. 4. In this example, the deblurred image produces dramatically improved segmentation results for all the thresholds.

6. Conclusions

In this article we have proposed a method for the restoration of blurred two-tone images when the blur filter is unknown. The method jointly estimates the blur filter along with the original image and the tone values based solely on the blurred image.

The method shows that it is possible to blindly deblur certain two-tone images of correlated pixels without explicitly estimating or modeling their statistical properties. The text image example demonstrates the potential usefulness of the method for front-end processing in automatic character recognition systems.

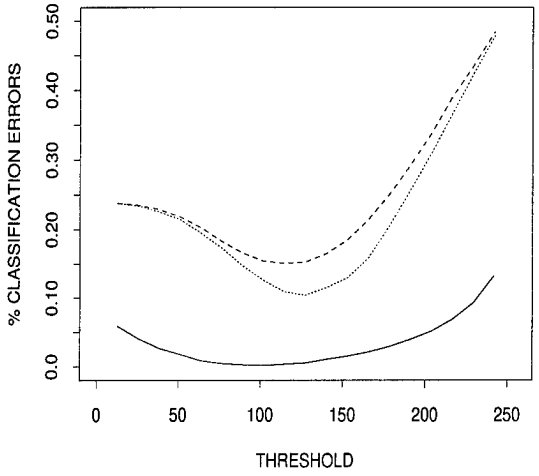


Figure 4. Proportion of mis-classified pixels as a function of the segmentation threshold – dashed line for the blurred image, dotted line for the image from the initial filter, and solid line for the deblurred image after 4,000 iterations.

For the future research, one can certainly extend the results in this article to multi-tone images and explore other possible cost functions and algorithms for the implementation. The bottom line is that if one judiciously takes advantage of the prior two-tone (or multi-tone) information, simple yet powerful deblurring algorithms can be obtained.

References

[1] T. Pavlidis, J. Swartz, and Y. P. Wang, " Information encoding with two-dimensional bar codes," *IEEE Computer*, vol. 25, no. 6, pp. 18–28, 1992.

[2] K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, vol. 10, pp. 1195–1208, 1982.

[3] T. H. Li and K. S. Lii, "A Joint Estimation Approach for Two-Tone Image Deblurring," submitted for publication, 1997.

[4] T. H. Li, "Blind identification and deconvolution of linear systems driven by binary random sequences," *IEEE Trans. Inform. Theory*, vol. 38, no. 1, pp. 26–38, 1992.

[5] T. H. Li, "Estimation and blind deconvolution of autoregressive systems with nonstationary binary inputs," *J. Time Series Anal.*, vol. 14, no. 6, pp. 214–236, 1993.

[6] S. Haykin, *Adaptive Filter Theory*, 3rd Edn., New York: Prentice Hall, 1996.

Linear Algebraic Approaches for (Almost) Periodic Moving Average System Identification

Ying-Chang Liang, A. Rahim Leyman
School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798
eycliang(earleyman)@ntuvax.ntu.ac.sg

Xian-Da Zhang
Department of Automation
Tsinghua University
Beijing 100084, P.R. China
zxd-dau@mail.tsinghua.edu.cn

Abstract

This paper addresses the problem of (almost) periodic moving average (APMA) system identification. Two normal equations are established by using time-varying higher order cumulants of the measurements, from which two new linear algebraic algorithms are presented for parameter estimation. Simulation examples are given to demonstrate the performance of these new approaches.

1. Introduction

During the last decade, the problem of identifying linear time-invariant moving average (LTI-MA) systems via higher order statistics (HOS) has been well studied; see [1] for an excellent overview. However, much less attention has been paid to identifying time-varying systems.

This paper addresses the problem of (almost) periodic moving average (APMA) system identification. APMA system is an important class of time-varying systems which are often encountered in hydrology, meteorological data and telecommunication systems. Recently, Dandawaté and Giannakis [2] introduced a novel approach for modeling APMA processes using higher order cyclic-statistics (HOCS). This approach is based on two main steps: (i) by exploiting the cyclostationarity property of APMA processes, the time-varying higher order cumulants (TV-HOC's) are estimated via the HOCS's computed from one data record; and then (ii) the model parameters are estimated from the TV-HOC's by using linear or nonlinear programming algorithms. This method is conceptually simple, however, as pointed out in [2], the bias and variance of these parameter estimators are considerably higher especially for the linear algorithms. This is expected since the linear algorithms in [2] are time-varying extensions of the so-called closed-form solutions designed for LTI-MA system identification, which would exhibt poor numerical performance without smoothing out the effects of the estimation errors of cumulants [1]. Therefore, new linear algorithms with better numerical performance are highly desirable.

As a matter of fact, linear algebraic methods are appropriate choices for improving the performance of parameter estimates in LTI-MA system identification [1]. It is reasonable to expect that linear algebraic methods outperform the closed-form methods even for the time-varying case. The objective of this paper is to present a linear algebraic approach for APMA system identification. Two new normal equations based on TV-HOC's are established, from which the time-varying system parameters are uniquely recovered (within some constant scale factor). Simulation examples are presented to demonstrate the performance of this paper's approaches.

2. Problem Statements

Throughout this paper, it is assumed that the channel is an APMA system with order-q, and that the measurement output $z(n)$ is described by

$$z(n) = x(n) + w(n) = \sum_{k=0}^{q} b(n;k)\varepsilon(n-k) + w(n) \quad (1)$$

where $\varepsilon(n)$ and $x(n)$ are the driving noise and the "real" system output, respectively; and $w(n)$ is the measurement noise. The impulse response coefficients, $b(n;k)$'s, vary almost periodically (AP) with respect to n at each lag k, thus $x(n)$ is a time-varying, but cyclostationary process [2].

0-8186-8005-9/97 $10.00 © 1997 IEEE

Let $m_{lx}(n;\boldsymbol{\tau})$ and $c_{lx}(n;\boldsymbol{\tau})$ denote the lth-order time-varying moment and cumulant of $x(n)$ at lag $\boldsymbol{\tau} \triangleq (\tau_1,\cdots,\tau_{l-1})$, respectively. The rigorous relation between $c_{lx}(n;\boldsymbol{\tau})$ and $m_{lx}(n;\boldsymbol{\tau})$ can be found in [6]. Specifically, for (time average) zero-mean processes with $l = 2, 3$ and 4, we have

$$c_{2x}(n;\tau) = m_{2x}(n;\tau), \tag{2}$$

$$c_{3x}(n;\tau_1,\tau_2) = m_{3x}(n;\tau_1,\tau_2), \tag{3}$$

and

$$\begin{aligned} & c_{4x}(n;\tau_1,\tau_2,\tau_3) \\ = \; & m_{4x}(n;\tau_1,\tau_2,\tau_3) - m_{2x}(n;\tau_1)m_{2x}(n+\tau_2,\tau_3-\tau_2) \\ & -m_{2x}(n;\tau_2)m_{2x}(n+\tau_3,\tau_1-\tau_3) \\ & -m_{2x}(n;\tau_3)m_{2x}(n+\tau_1,\tau_2-\tau_1). \end{aligned} \tag{4}$$

Since $b(n;k)$ is AP in n, $m_{lx}(n;\boldsymbol{\tau})$ is also AP in n for each $\boldsymbol{\tau}$. Hence, it possesses the following Fourier series representation

$$m_{lx}(n;\boldsymbol{\tau}) = \sum_{\alpha\in A_{lx}} m_{lx}^{(\alpha)}(\boldsymbol{\tau})e^{j\alpha n} \tag{5}$$

where $m_{lx}^{(\alpha)}(n;\boldsymbol{\tau})$ is called the lth-order cyclic-moment at cycle α, and

$$m_{lx}^{(\alpha)}(\boldsymbol{\tau}) = \lim_{T\to\infty}\frac{1}{T}\sum_{n=0}^{T-1} m_{lx}(n;\boldsymbol{\tau})e^{-j\alpha n} \tag{6}$$

and the set of cycles at (6) is defined as $A_{lx} \triangleq \{\alpha : m_{lx}^{(\alpha)}(\boldsymbol{\tau}) \neq 0, 0 \le \alpha < 2\pi\}$. By [2], when $x(n)$ is a mixing process, a consistent and asymptotically normal estimate of $m_{lx}^{(\alpha)}(\boldsymbol{\tau})$ can be obtained via sample average technique based on one data record, and the lth-order time-varying moment, thus cumulant can be computed via (5), and (2)-(4).

In this paper, we concentrate our attention on identifying the time-varying impulse response coefficients from the TV-HOC's of the output measurements. Therefore, we assume that the cycles $\alpha \in A_{lx}$ and the order q are known *a priori*. Otherwise, they can be estimated first by performing the statistical tests for presence of cyclostationarity in the measurements[2, 5]. In addition, the following conditions are assumed to hold.

AS1) The driving noise sequence $\varepsilon(n)$ is unobservable, and is zero-mean, independent and identically distributed (i.i.d.), and non-Gaussian process with some l such that $0 < |\gamma_{l\varepsilon}| < \infty, l > 2$, where $\gamma_{l\varepsilon}$ is the lth order cumulant of $\varepsilon(n)$.

AS2) The measurement noise $w(n)$ is a Gaussian (colored) process independent of $\varepsilon(n)$, and hence of $x(n)$.

3. Two Linear Algebraic Algorithms

It is well known that under condition AS2)

$$c_{lz}(n;\tau_1,\cdots,\tau_{l-1}) = c_{lx}(n;\tau_1,\cdots,\tau_{l-1}), \quad l \ge 3, \tag{7}$$

because of the higher than second-order cumulants of Gaussian processes being identical to zero, and that the lth-order cumulants are related to the impulse response coefficients by [2]

$$\begin{aligned} c_{lz}(n;\tau_1,\cdots,\tau_{l-1}) &= \gamma_{l\varepsilon}\sum_{k=0}^{q} b(n;k)b(n+\tau_1;k+\tau_1) \\ &\quad \cdots b(n+\tau_{l-1},k+\tau_{l-1}). \end{aligned} \tag{8}$$

By using the fact that $b(n;k) \equiv 0$, for $k < 0$ and $k > q$, it follows that

$$\begin{aligned} & c_{lz}(n;\tau_1,\tau_2,\cdots,q,k) \\ = \; & \gamma_{l\varepsilon}b(n;0)b(n+\tau_1;\tau_1)\cdots b(n+q;q) \\ & \times b(n+k;k). \end{aligned} \tag{9}$$

Based on (9), three linear algorithms were developed by Dandawaté and Giannakis [2] for estimating the unknown parameters, $b(n;k)$'s, within some constant scale factor. Here we refer to them as DG-1, DG-2 and DG-3, respectively.

(1) DG-1: The impulse response coefficients are given by

$$b(n;k) = \frac{\gamma_{(l-1)\varepsilon}}{\gamma_{l\varepsilon}}\frac{c_{lz}(n-k;\tau_1,\cdots,\tau_{l-3},q,k)}{c_{(l-1)z}(n-k;\tau_1,\cdots,\tau_{l-3},q)}. \tag{10}$$

(2) DG-2: The impulse response coefficients are given by

$$b(n;k) = \frac{1}{\gamma_{3\varepsilon}^{1/3}}\frac{c_{3z}(n-k;q,k)}{[c_{3z}(n-k;q,q)c_{3z}(n-k;q,0)]^{1/3}}. \tag{11}$$

(3) DG-3: The impulse response coefficients are given by

$$\begin{aligned} b(n+k;k) &= \gamma_{lz}^{-1/l}\left\{\frac{c_{lz}^2(n;0,\cdots,0,0,q)}{c_{lz}(n;0,\cdots,0,q,q)}\right\}^{1/l} \\ &\quad \times\frac{c_{lz}(n;\tau_1,\cdots,\tau_{l-3},q,k)}{c_{lz}(n;\tau_1,\cdots,\tau_{l-3},q,0)}. \end{aligned} \tag{12}$$

Although the above algorithms are based upon simple computations, they belong to the closed-form solutions which do not smooth out the effects of the estimation errors of the TV-HOC's. Now, let us establish two normal equations based on the TV-HOC's of the output measurements, and develop two linear algebraic algorithms for estimating the APMA system parameters.

3.1. Normal Equation based on the Diagonal Slice of TV-HOC's

Let $\tau_1 = \cdots = \tau_{l-1} = \tau$ in (8), we have

$$c_{lz}(n;\tau,\cdots,\tau) = \gamma_{l\varepsilon}\sum_{k=0}^{q} b(n;k)[b(n+\tau;k+\tau)]^{l-1}. \tag{13}$$

On the other hand, from (9), we have

$$\begin{aligned} & c_{lz}(n-k;0,\cdots,0,q,k+\tau) \\ = \; & \gamma_{l\varepsilon}[b(n-k;0)]^{l-2}b(n-k+q;q) \\ & \times b(n-k+\tau;k+\tau) \end{aligned} \tag{14}$$

and

$$\begin{aligned} & c_{lz}(n-k;0,\cdots,0,q,0) \\ = \; & \gamma_{l\varepsilon}[b(n-k;0)]^{l-1}b(n-k+q;q), \end{aligned} \tag{15}$$

which yield

$$\begin{aligned} & b(n+\tau;k+\tau) \\ = \; & b(n-k;0)\frac{c_{lz}(n-k;0,\cdots,0,q,k+\tau)}{c_{lz}(n-k;0,\cdots,0,q,0)} \end{aligned} \tag{16}$$

and

$$b(n-k;0) = \left\{\frac{c_{lz}^2(n-k;0,\cdots,0,q,0)}{\gamma_{l\varepsilon}c_{lz}(n-k;0,\cdots,0,q,q)}\right\}^{1/l} \tag{17}$$

where we assume $b(n-k;0) > 0$ when even l is used. Substitute (16) and (17) into (13), and denote

$$\begin{aligned} & a_l(n+\tau;k+\tau) \\ = \; & \left\{\frac{c_{lz}^2(n-k;0,\cdots,0,q,0)}{c_{lz}(n-k;0,\cdots,0,q,q)}\right\}^{(l-1)/l} \\ & \times\left\{\frac{c_{lz}(n-k;0,\cdots,0,q,k+\tau)}{c_{lz}(n-k;0,\cdots,0,q,0)}\right\}^{l-1}. \end{aligned} \tag{18}$$

Considering that $c_{lz}(n,\tau,\cdots,\tau) \equiv 0$, for $|\tau| > q$, and that

$$a_l(n+\tau;q) \neq 0, \quad \text{for any } \tau \tag{19}$$

$$a_l(n+\tau;k+\tau) \equiv 0, \quad \text{for } k+\tau < 0, \text{ and } k+\tau > q \tag{20}$$

due to (9) and (18), we have

$$\begin{aligned} c_{lz}(n;\tau,\cdots,\tau) \; = \; & \sum_{k=0}^{q}\gamma_{l\varepsilon}^{1/l}b(n;k)a_l(n+\tau;k+\tau), \\ & -q \le \tau \le q. \end{aligned} \tag{21}$$

Equation (21) is referred to as the normal equation based on the diagonal slice of TV-HOC's for time-varying MA(q) system identification.

Concatenating (21) for $\tau = -q,\cdots,0,\cdots,q$, we obtain the following matrix equation

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{22}$$

where $\boldsymbol{A}$ contains the submatrix $\boldsymbol{A_1}$ as

$$\boldsymbol{A_1} = \begin{bmatrix} a_l(n;q) & a_l(n;q-1) & \cdots & a_l(n;0) \\ & a_l(n+1;q) & \cdots & a_l(n+1;1) \\ & & \ddots & \vdots \\ 0 & & & a_l(n+q;q) \end{bmatrix} \tag{23}$$

and

$$\boldsymbol{x} = \gamma_{l\varepsilon}^{1/l}[b(n;q), b(n;q-1),\cdots,b(n;0)]^T. \tag{24}$$

From (23) and (19), it is easy to see that $\boldsymbol{A_1}$ has full rank $q+1$ due to $\det(\boldsymbol{A_1}) = \prod_{i=0}^{q} a_l(n+i;q) \neq 0$, thus the coefficient matrix in (22) has full rank $q+1$. This shows that the matrix equation (22) with $(q+1)$ unknown parameters has a unique least-square (LS) solution given by

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{b}. \tag{25}$$

Solving (25), we obtain the estimates of $\gamma_{l\varepsilon}^{1/l}b(n;k)$'s, which are related to $b(n;k)$'s within a constant scale. We refer to the solution based on (21) as *Algorithm* 1.

3.2. Normal Equation based on All Slices of TV-HOC's

From (16) and (17), $b(n+\tau_i;k+\tau_i)$ can be expressed as

$$b(n+\tau_i;k+\tau_i) = \gamma_{l\varepsilon}^{-1/l}b_l(n+\tau_i;k+\tau_i) \tag{26}$$

where

$$\begin{aligned} & b_l(n+\tau_i;k+\tau_i) \\ = \; & \left\{\frac{c_{lz}^2(n-k;0,\cdots,0,q,0)}{c_{lz}(n-k;0,\cdots,0,q,q)}\right\}^{1/l} \\ & \times\left\{\frac{c_{lz}(n-k;0,\cdots,0,q,k+\tau_i)}{c_{lz}(n-k;0,\cdots,0,q,0)}\right\}. \end{aligned} \tag{27}$$

By substituting (29) and (30) into (8) and denoting

$$e_l(n;k;\tau_1,\cdots,\tau_{l-1}) = \prod_{i=1}^{l-1} b_l(n+\tau_i;k+\tau_i), \tag{28}$$

it follows that

$$\begin{aligned} & c_{lz}(n;\tau_1,\cdots,\tau_{l-1}) \\ = \; & \sum_{k=0}^{q}\gamma_{l\varepsilon}^{1/l}b(n;k)e_l(n;k;\tau_1,\cdots,\tau_{l-1}), \\ & -q \le \tau_i \le q, i = 1,\cdots,l-1. \end{aligned} \tag{29}$$

Equation (29) is referred to as the normal equation based on all slices of TV-HOC's for time-varying MA(q) system identification.

Concatenating (29) for $\tau_i = -q, \cdots, 0, \cdots, q$, and $i = 1, \cdots, l-1$, we obtain an overdetermined set of linear equations, which includes that given by (22). Since the rank of the coefficient matrix in (22) is $q+1$, the coefficient matrix based on (29) is of rank-$(q+1)$ too. This means that equation (29) gives a unique LS solution (within some constant scale). We refer to the solution based on (29) as *Algorithm* 2.

Several remarks are in order.

Remark 1 : Equations (21) and (29) hold for a general time-varying MA(q) model. However, to implement these schemes in practice, it is necessary to estimate the involved TV-HOC's first, which may not be available from only one record of data. Fortunately, for an APMA system, the output TV-HOC's can be computed via sample average technique from only one record of data, as discussed in Section II, which makes it possible to identify the time-varying impulse response coefficients.

Remark 2 : For time-invariant MA(q) system identification, a variety of linear algebra solutions have been developed by using the overparameterized techniques. However, these techniques can not be applied to the time-varying system identification because the number of unknowns may be greater than the number of equations available for such case. On the other hand, our two algorithms are well parameterized in the sense that only $b(n;k)$'s are estimated.

Remark 3 : *Algorithm* 1 exploits the diagonal slice of TV-HOC's only, however, *Algorithm* 2 uses all slices of TV-HOC's. By using the symmetry property of TV-HOC's, it is possible to reduce the number of equations in *Algorithm* 2. In addition, it is widely recognized [1] that additional cumulant slices may be exploited to improve the performance of parameter estimates. Therefore, *Algorithm* 2 is expected to outperform *Algorithm* 1.

4. Simulations

Two simulation examples are presented to illustrate the performance of our new algorithms for TVMA system identification. Here, $b(n;k)$ was chosen to be a periodically time-varying filter with period 2, thus the cyclic frequencies are $\alpha = 0$ and $\alpha = \pi$, and 100 Monte Carlo runs were performed with each consisting of N data samples. In each run, the input $\varepsilon(n)$ was chosen to be i.i.d. exponential with $\gamma_{3\varepsilon} = 1$, and the measurement noise $w(n)$ was a zero-mean Gaussian AR process with poles at $0.8 \pm j0.2$. The signal-to-noise ratio (SNR) was defined as $\mathrm{SNR} \triangleq 10\log_{10}(E\{\varepsilon^2(n)\}/E\{w^2(n)\})$. The accuracy of system identification is assessed by calculating the mean square error (MSE)

$$\mathrm{MSE} \triangleq 10\log_{10}\frac{\sum_{n=0}^{1}\sum_{i=0}^{q}(b(n;i)-\hat{b}(n;i))^2}{\sum_{n=0}^{1}\sum_{i=0}^{q}b^2(n;i)} \quad (30)$$

where $\hat{b}(n;i)$ is the estimated system parameter.

Example 1: The system impulse response coefficients were given by [2]

$$b(0;k) = [1, 0.1, -0.2], \quad (31)$$

$$b(1;k) = [1.5, 0.7, 0.4]. \quad (32)$$

The data length was chosen to be $N = 2 \times 4096$. Table I summarizes the mean $\pm$ standard deviation of the parameter estimates for $n = 0$ and $n = 1$ at SNR=0dB obtained by DG-2 approach based on (11) and this paper's two new algorithms ($l = 3$).

Example 2: The system impulse response coefficients were given by

$$b(0;k) = [1, 0.9, 0.385, -0.771], \quad (33)$$

$$b(1;k) = [1, -1.967, 0.5666, 0.2]. \quad (34)$$

The data length was chosen to be $N = 4 \times 4096$. Table II summarizes the mean $\pm$ standard deviation of the parameter estimates for $n = 0$ and $n = 1$ at SNR=10dB obtained by DG-2 approach and this paper's two new algorithms ($l = 3$).

From the parameter estimation results shown in Tables I and II, we have the following observations:

(1) From the MSE's, our two new algorithms perform better than DG-2 approach. This is because that DG-2 approach is sensitive to the estimation errors of TV-HOC's, however, our new algorithms are much robust to this kind of errors.

(2) The estimates obtained via *Algorithm* 2 are better (in terms of bias and deviation) than the estimates obtained via *Algorithm* 1. This is expected since the set of the output statistics exploited by *Algorithm* 2 is much larger than the data of the output statistics exploited by *Algorithm* 1; see *Remark* 3.

5. Conclusions

In this paper, we have established two normal equations for TVMA system identification and have developed two linear algebraic algorithms whose uniqueness is guaranteed (within some constant factor). These methods use only higher order cumulants and consequently yield consistent parameter estimation in the presence of additive colored noise with vanishing higher

order cumulants. Simulations have shown that the two new algorithms have better performance than the existing closed-form algorithms.

References

[1] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," Proc. IEEE, vol.79, pp.278-305, Mar. 1991.

[2] A. Dandawaté and G.B. Giannakis, "Modeling (almost) periodic moving average processes using cyclic statistics," IEEE Trans. Signal Processing, vol.44, pp.673-684, Mar. 1996.

[3] X.-D. Zhang and Y.-S. Zhang, "FIR system identification using higher order statistics alone," IEEE Trans. Signal Processing, vol.42, pp.2854-2858, Oct. 1994.

[4] X.-D. Zhang and Y.-S. Zhang, "Singular value decomposition-based MA order determination of non-Gaussian ARMA models," IEEE Trans. Signal Processing, vol.41, pp.2657-2664, Aug. 1993.

[5] A. Dandawaté and G.B. Giannakis, "Statistical tests for presence of cyclostationarity," IEEE Trans. Signal Processing, vol.42, pp.2355-2369, Sept. 1994.

[6] D.R. Brillinger, Time series: Data Analysis and Theory, Holden-day Inc., San Francisco, 1981.

TABLE I
Parameter Estimation Statistics for Example 1:
$N = 2 \times 4096$, SNR=0dB, 100 Monte Carlo Runs

Parameters	DG-2	Algorithm 1	Algorithm 2
$b(0;0) = 1.0$	0.4474 (± 0.9678)	1.1358 (± 0.7106)	1.0385 (± 0.5385)
$b(0;1) = 0.1$	0.1032 (± 0.1377)	0.1006 (±0.0608)	0.1103 (±0.0675)
$b(0;2) = -0.2$	−0.2453 (±0.1254)	−0.0884 (±0.4480)	−0.1672 (±0.1745)
$b(1;0) = 1.5$	1.2956 (±0.9438)	1.5039 (±0.4643)	1.4383 (±0.3876)
$b(1;1) = 0.7$	0.0675 (±0.7105)	0.7934 (±0.4621)	0.7515 ±(0.3933)
$b(1;2) = 0.4$	0.1745 (±0.1305)	0.3977 (±0.1366)	0.4057 (±0.1443)
MSE	−8.9167	−19.9852	−26.3467

TABLE II
Parameter Estimation Statistics for Example 2:
$N = 4 \times 4096$, SNR=10dB, 100 Monte Carlo Runs

Parameters	DG-2	Algorithm 1	Algorithm 2
$b(0;0) = 1.0$	0.0399 (±1.2045)	1.2573 (±0.7006)	1.0994 (±0.5694)
$b(0;1) = 0.9$	0.8884 ±(0.1058)	0.8482 (±0.3279)	0.8619 (±0.0881)
$b(0;2) = 0.385$	−0.1239 (±0.6297)	0.5268 (±0.3570)	0.4475 (±0.2465)
$b(0;3) = -0.771$	−0.7750 (±0.0580)	−0.7936 (±0.2497)	−0.7654 (±0.0670)
$b(1;0) = 1.0$	1.0052 (±0.0882)	1.0285 (±0.7053)	0.9445 (±0.2202)
$b(1;1) = -1.967$	−0.0603 (±2.3915)	−2.5787 (±1.4547)	−2.2176 (±1.1386)
$b(1;2) = 0.5666$	0.5800 (±0.0947)	0.6793 (±0.8961)	0.5978 (±0.1376)
$b(1;3) = 0.2$	0.2945 (±0.1860)	0.2034 (±0.3212)	0.2095 (±0.1274)
MSE	−2.0839	−12.1293	−19.7690

Fast Estimation of Wiener Kernels of Nonlinear Systems in the Frequency Domain

Michael A. Shcherbakov
IVS, State Technical University of Penza
Krasnaya 40, Penza 440017, Russia
mike@diamond.stup.ac.ru

Abstract

A method for identification of discrete nonlinear systems in terms of the Volterra-Wiener series is presented. It is shown that use of a special composite-frequency input signal as approximation to Gaussian noise provides a computational efficiency of this method, especially for high order kernels. Orthogonal functionals and consistent estimates for Wiener kernels in the frequency domain are derived for this class of noise input. A basis of the proposed computational procedure for practical identification is the fast Fourier transform (FFT) algorithm which is used both for generation of actions and for analysis of system reactions.

1. Introduction

Continued interest has been shown in the use of functional series in the modeling, identification and control of nonlinear systems [1-6] since the initial work by Wiener [1]. He considered the class of causal systems that produce an output with finite mean-square value when their input is a Gaussian white noise. The output $y(t)$ of an unknown nonlinear "black-box" system can be approximated by a series of functionals $G_m[h_m,x(t)]$ of the input $x(t)$ as

$$y(t) = \sum_{m=0}^{M} G_m[h_m, x(t)] \qquad (1)$$

where h_m is the Wiener kernel of the order m.

A main difficulty encountered when one wants to apply the Wiener approach to identification problems involves the measurement of the kernels which is computationally demanding. Several methods have been presented to find the Wiener kernels of nonlinear systems from given input and output pairs. Lee and Schetzen [7] showed that the kernels can be estimated by input-output crosscorrelation. French and Butz [8] used the fast Fourier and Walsh transform algorithms for calculation of the Wiener kernels. However, there are some difficulties involved in these methods:

- A white Gaussian process is unrealizable.
- Formula for kernels, $m \geq 2$, involves Dirac delta functions, when two or more kernel's arguments are equal.
- The required computation increases very rapidly with the order of the Wiener kernel being calculated.

This paper will resolve these difficulties by investigating discrete systems with special type of noise inputs generated using the FFT algorithm. We will construct the G-functionals for such inputs, and the formula for the Wiener kernels and the efficient identification algorithm in the frequency domain will be presented.

2. Forcing functions for nonlinear systems testing

It is necessary that the system stimulus must, on the one hand, be like a random noise to get maximum information about unknown system and, on the other hand, to simplify the G-functionals and the procedure of identification on the whole. Taking into consideration these circumstances, let us consider as a test input the following periodic noise approximation

$$x(n) = \sum_{k=-N_x}^{N_x} X(k) \exp j\frac{2\pi kn}{N}. \qquad (2)$$

Here $X(k)=A(k)\varphi(k)$ are complex Fourier coefficients, where the amplitudes $A(k)$ determine the power spectrum of the input, and the phases $\varphi(k)$ are independent random values with uniform distribution.

For a zero mean real signal, the complex valued Fourier coefficients have the following relationships: $X(0)=0$, $X(-k)=X^*(k)$. According to the Central Limit Theorem, the signal in the form of (2), being a sum of independent random quantities, has a nearly Gaussian distribution for sufficiently large N_x. For every set $\varphi_l(k)$ of the random phases, formula (2) determines the sequence $x_l(n)$ having N samples long which may be formed by the inverse FFT of the coefficients $X_l(k)$.

0-8186-8005-9/97 $10.00 © 1997 IEEE

3. G-functionals and Wiener kernels in the frequency domain

According to the proposed method of the test signal generation, the random input process $x(n)$ is determined by the set of input Fourier coefficients $X(k)$, k=0,...,N_x, and the corresponding response $y(n)$ can be characterized by the set of output Fourier coefficients $Y(k)$, k=0,...,N_y. Therefore, it is possible to rewrite input-output relationship (1) for nonlinear systems in the frequency domain as

$$Y_M(k) = \sum_{m=0}^{M} G_m[H_m, X(k)] \qquad (3)$$

where $H_m(k_1,...,k_m)$ is the multidimensional discrete Fourier transformation (DFT) of the Wiener kernel $h_m(n_1,...,n_m)$.

By using a Gram-Schmidt orthogonalization procedure, the G-functionals can be shown to be

$$G_m[H_m, X(k)] = \sum_{D_m} H_m(k_1,...,k_m)\delta^{k}_{k_1+...+k_m}\prod_{i=1}^{m} X(k_i) \qquad (4)$$

where the summation must extend over the m-D region D_m consisting of various combinations $(k_1,...,k_m)$ from integers $\{-N_x,...,-1,1,...,N_x\}$ such that $k_1 > k_2 > ... > k_m$, and $k_i \neq -k_j$, and δ_i^j is a Kronecker delta.

The Wiener kernels in the frequency domain for the model (3) can be determined by minimization of the mean square error between the DFT $Y(k)$ of the system and $Y_M(k)$ model responses

$$F = E\{\Delta^T \Delta^*\} \to \min$$

where $\Delta^T = [\delta_1,...,\delta_{Ny}]$ is the vector of the complex errors having the elements $\delta_k = Y(k) - Y_M(k)$, $E\{\bullet\}$ denotes the average operation, and $*$ represents the complex conjugate.

Minimizing this function, the optimal Wiener kernels become

$$H_m(k_1,...,k_m) = \frac{\mathrm{E}\{Y(k_1+...+k_m)X^*(k_1)\cdot...\cdot X^*(k_m)\}}{A^2(k_1)\cdot...\cdot A^2(k_m)}.$$

In order to construct an estimate of kernel $H_m(k_1,...,k_m)$ which would be suitable in practice let us introduce the periodogram

$$I^l_{yx...x}(k_1,...,k_m) = Y^l(k_1+...+k_m)\exp[-j\sum_{i=1}^{m}\varphi_l(k_i)]. \qquad (5)$$

Then we can get the estimate by averaging L periodograms for different blocks of data as follows

$$\hat{H}_m(k_1,...,k_m) = \frac{\sum_{l=1}^{L} I^l_{yx...x}(k_1,...,k_m)}{A^2(k_1)\cdot...\cdot A^2(k_m)}. \qquad (6)$$

4. Identification algorithm

If the random phases $\varphi(k)$ are formed by random sampling from set of numbers $2\pi r/R$, r=0,...,R-1, the equation (5) for the periodogram may be rewritten as

$$I^l_{yx...x}(k_1,...,k_m) = Y^l(k_1+...+k_m) \times \exp\left(-j\frac{2\pi}{R}\{s^l_{k_1}+...+s^l_{k_m}\} \bmod R\right) \qquad (7)$$

where s_k^l is l-th set of random integers, and $\{\bullet\} \bmod R$ denotes summation defined modulo R.

The calculation of the Wiener kernels for order $m \geq 2$ may be performed more effectively if it is noted that periodogram (7) may take limited number of values $Y_l(k)\exp\text{-}j2\pi i/N_x$, k=0,...,N_y, i=0,...,R-1. This allows us in advance to form the array of possible products for every DFT $Y_l(k)$. Thus, the algorithm of identification consists of the following steps:

1. Generation of the random integers $s^l_1,...,s^l_{N_x}$ and forming the complex Fourier coefficients $X_l(k)$.
2. Calculation the l-th block of the input signal by using of the inverse FFT algorithm
$$x_l(n) = \sum_{k=-N_x}^{N_x} X_l(k)\exp j\frac{2\pi kn}{N}, \quad n=0,...,N\text{-}1.$$
3. Stimulation of the system by the input signal $x_l(n)$ and registration of the response $y_l(n)$.
4. Calculation the complex Fourier coefficients $Y_l(k)$ by use the FFT algorithm
$$Y_l(k) = \frac{1}{N}\sum_{n=0}^{N-1} y_l(n)\exp -j\frac{2\pi kn}{N}, \quad k=0,...,N_y.$$
5. Definition of the array $Z_l(k,i)$ of all the possible values of the periodograms
$$Z_l(k,i) = Y_l(k)\exp\text{-}j2\pi i/R, \quad k=0,...,N_y,\ i=0,...,R\text{-}1.$$
6. Forming the periodograms from the array $Z_l(k,i)$
$$I^l_{yx...x}(k_1,...,k_m) = Z_l(k_1+...+k_m, \{s^l_{k_1}+...+s^l_{k_m}\} \bmod R).$$
7. Calculation of the kernel estimates using eqn.(6).

This algorithm of the kernels estimation bases also on idea of scanning the definition regions D_m, m=1,..,M, so that the partial sums $k_1+...+k_m$ and $\{s^l_{k_1}+...+s^l_{k_m}\} \bmod R$ obtained for m-order kernel estimate could be used for calculation of the $(m+1)$-order periodogram.

Next, we will describe a simple recursive method for generating of the domains D_m. The one-dimensional frequency domain consists of discrete frequencies $D_1 = \{0,1,...,N/2\}$. The two-dimensional frequency domain consists of the sum interaction domain

$$D_2^+ = \{(k_1,k_2): k_1 \leq k_2 \leq N/2 - k_1\}$$

and the difference interaction domain

$$D_2^- = \{(-k_1,k_2): k_1 \leq k_2 \leq N/2\}.$$

In the general case, m-dimensional frequency domain can be generated by (m-1)-dimensional one. The sum D_m^+ and the difference D_m^- interaction domains are expressed as follows:

$$D_m^+ = \{(k_1,\ldots,k_{m-1},k_m): \\ k_{m-1} \le k_m \le \min(N/2, N/2 - S_{m-1})\},$$

$$D_m^- = \{(-k_{m-1},\ldots,-k_1,k_m): \\ \max(S_{m-1}, -k_1) \le k_m \le N/2,\ S_{m-1} \ne 0\}$$

for all frequency points $(k_1,\ldots,k_{m-1}) \subseteq D_{m-1}$. Here S_m denotes the sum of indices $k_1+\ldots+k_m$. The whole domain D_m is formed as union of these two domains: $D_m = D_m^+ \cup D_m^-$. Note that for every new generated point $(k_1,\ldots,k_m)$ of the frequency domain the indices are ordered such that $k_1 \le k_2 \le \ldots \le k_m$.

Since the number C_m of the combinations $(k_1,\ldots,k_m)$ containing in the kernel definition region D_m increases very rapidly with order m of the kernel, and the number of multiplications, demanding for calculation of the array $Z_l(k,i)$ of all possible values of the periodograms, does not depend on m, the proposed algorithm, as compared with methods [7, 8], allows one to decrease the number of multiplications to a marked degree.

Actually, the most effective method [8] demands approximately $L(C_1+2C_2+\ldots+mC_m)$ complex multiplication, while as the proposed method requires only $LR(N_y+1)/2$ multiplications. For a comparison, in Table we give a reduction factor of multiplications for $N_x=N_y$, R=8 and various values m of the kernel order. Thus the computational efficiency rapidly increases with the kernel order.

Table. Reduction factor of multiplications

Kernel order m	The number N_Y of output frequencies			
	16	32	64	128
1	1	1	1	1
2	3	6	12	24
3	26	110	447	1804

5. Statistical properties of kernel estimates

Averaging the estimate (6) yields

$$\mathrm{E}\left\{\hat{H}_m(k_1,\ldots,k_m)\right\} = H_m(k_1,\ldots,k_m).$$

Therefore the estimate is unbiased.

A mean square error of the kernel estimate is given by the variance

$$\mathrm{Var}\left\{\hat{H}_m(k_1,\ldots,k_m)\right\} \qquad (8)$$

$$= \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} \mathrm{cov}\left\{I^{l_1}_{yx\ldots x}(k_1,\ldots,k_m) I^{l_2 *}_{yx\ldots x}(k_1,\ldots,k_m)\right\} \Big/ L^2 \prod_{i=1}^{m} A^2(k_i)$$

Taking into account the independence of the random phases $\varphi(k)$, for covariance we obtain

$$\mathrm{cov}\left\{I^{l_1}_{yx\ldots x}(k_1,\ldots,k_m) I^{l_2 *}_{yx\ldots x}(k_1,\ldots,k_m)\right\} \qquad (9)$$

$$= \begin{cases} 0 & l_1 \ne l_2 \\ \mathrm{E}\left\{\left|Y_l(k_1+\ldots+k_m)\right|^2\right\} - \left|H_m(k_1,\ldots,k_m)\right|^2 \prod_{i=1}^{m} A^2(k_i) & l_1 = l_2 = l. \end{cases}$$

Substituting (9) into (8), we have

$$\mathrm{Var}\left\{\hat{H}_m(k_1,\ldots,k_m)\right\} = \frac{1}{L}\Bigg(S_y(k_1+\ldots+k_m) \Big/ N\prod_{i=1}^{m} A^2(k_i) \\ - \left|H_m(k_1,\ldots,k_m)\right|^2 \Bigg) \qquad (10)$$

where $S_y(k)$ is the spectrum of the output process $y(n)$. Since the variance (10) goes to zero as $L\to\infty$, the estimate $\hat{H}_m(k_1,\ldots,k_m)$ is consistent.

In order to clarify sense of the derived expression (10), let us take advantage of using the following formula for output spectrum

$$S_y(k) = N\sum_{m=0}^{\infty}\sum_{D_m}\left|H_m(k_1,\ldots,k_m)\right|^2 \\ \times\delta^k_{k_1+\ldots+k_m}\prod_{i=1}^{m} A^2(k_i). \qquad (11)$$

Then output spectrum at the point $(k_1+\ldots+k_m)$ may be represented as sum

$$S_y(k_1+\ldots+k_m) = N\left|H_m(k_1,\ldots,k_m)\right|^2 \prod_{i=1}^{m} A^2(k_i) \\ + S^{\Delta}_y(k_1+\ldots+k_m). \qquad (12)$$

Here the first term is caused by kernel $H_m(k_1,\ldots,k_m)$ at the point $(k_1,\ldots,k_m)$, and the second - by values of the kernel in the rest points.

After substituting (12) into (10), we obtain

$$\mathrm{Var}\left\{\hat{H}_m(k_1,\ldots,k_m)\right\} = S^{\Delta}_y(k_1+\ldots+k_m) \Big/ NL\prod_{i=1}^{m} A^2(k_i).$$

Thus, the variance of the estimate $\hat{H}_m(k_1,\ldots,k_m)$ at the point $(k_1+\ldots+k_m)$ is directly proportional to the output spectrum except for the component caused by the value of the kernel being estimated. The less this value contributes to the total spectrum, the more error is involved. The estimates have tendency to be worse as nonlinearity becomes stronger since the output spectrum enriches by new frequency components. If the system is linear, the output spectrum $S_y(k)$ is entirely defined by the first order kernel $H_1(k)$. Therefore, in this case variance of estimate $\hat{H}_1(k)$ is equal to zero.

For modeling of the input-output power relationship by (11), kernel magnitudes squared is needed. But using $\left|\hat{H}_m(k_1,\ldots,k_m)\right|^2$ as the estimate of $\left|H_m(k_1,\ldots,k_m)\right|^2$ can leads to large errors because one becomes biased for finite L.

Possibly the best estimate is

$$|H_m(k_1,\ldots,k_m)|^2_{\text{est}} = \left|\hat{H}_m(k_1,\ldots,k_m)\right|^2 - \text{Var}\left\{\hat{H}_m(k_1,\ldots,k_m)\right\}. \quad (13)$$

Substituting (10) into (13), one can be defined as follows

$$|H_m(k_1,\ldots,k_m)|^2_{\text{est}} = \frac{L}{L-1}\left|\hat{H}_m(k_1,\ldots,k_m)\right|^2 - \hat{S}_y(k_1+\ldots+k_m)\Big/ N(L-1)\prod_{i=1}^{m} A^2(k_i) \quad (14)$$

where $\hat{S}_y(k) = \frac{N}{L}\sum_{l=1}^{L}|Y_l(k)|^2$ is an estimate of $S_y(k)$.

It can be shown that the estimate (14) is unbiased and, therefore, one is more suitable for modeling spectral relations in nonlinear systems.

6. Experimental results

In order to test the results described in the previous sections Matlab programs have been written on the basis of the proposed identification algorithm.

The method developed in this paper has been tested by applying two kinds of inputs to the known system in fig.1: one is a zero-mean Gaussian input, and another is the composite-frequency input signal in the form (2). The kernels calculated by Wiener method are compared to those obtained by the method developed in this paper. Both methods give approximately identical estimates of the kernels in the frequency domain. However, comparison these methods with respect to computation time on IBM PC shows a high speed of proposed algorithm and confirms the theoretical calculation given above in the table.

The estimates of the first three kernels of known system has been calculated by averaging L=700 blocks of data containing N=128 points each according (6). For stimulation the system, the test signal in the form (2) with R=16 random phases and N_x=20 spectral samples was used. Using obtained kernel estimates, relevant contributions to the total response were calculated. This is shown in Fig. 2. The effect of the addition of nonlinear contribution was analyzed by using a mean squared error test. We had the following averaged reduction for the sequence of the models: linear $y_1(n)$ - 39.6, quadratic $y_2(n)$ - 17.5 and cubic $y_3(n)$ - 15.4.

Figs. 3-4 show the first and second order kernel estimates, respectively. Note that nonlinearity leads to larger errors, especially, for those frequency components which bring in a negligible contribution to the output spectrum.

For modeling the spectral input-output relationship, the kernel magnitudes squared have been also calculated by using (14). These estimates were used to calculate the output spectrum $S_m(k)$ of the model and its components

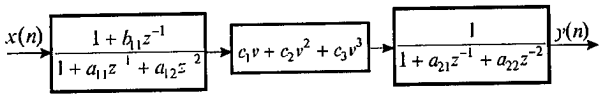


Figure 1. Known nonlinear system used for testing: a_{11}=-1.5681, a_{12}=0.64, b_{11}=-1, a_{21}=-1.2895, a_{22}=0.49, c_1=0.1, c_2=0.03, c_3=0.001

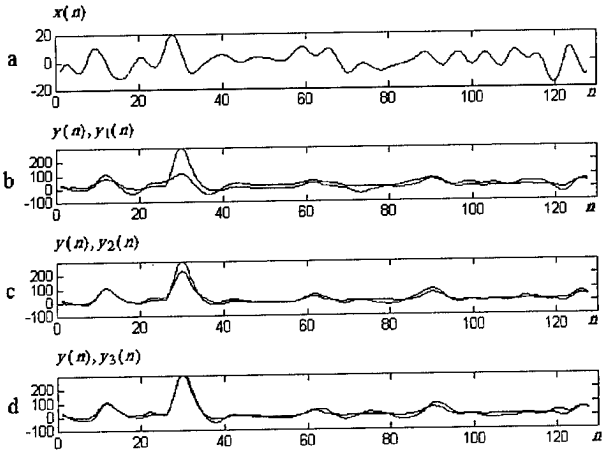


Figure 2. Simulation results: (a) input signal $x(n)$; (b-d) actual output signal $y(n)$ and model components $y_1(n)$, $y_2(n)$, $y_3(n)$.

$S_{m1}(k)$, $S_{m2}(k)$, $S_{m2}(k)$ contributed by the first three kernels. Fig. 5 shows these spectral estimates together with actual values $S_y(k)$, $S_{y1}(k)$, $S_{y2}(k)$, $S_{y3}(k)$ obtained by analytical way. Note that, obtained spectral estimates are unbiased and approximate the actual spectra fairly well. The model spectrum is also close to the spectral estimate $\hat{S}_y(k)$ obtained from output data directly.

The first order kernel gives a basic contribution to the output spectrum, and it can be evaluated most exactly. Since power of the quadratic term $S_{y2}(k)$ is mainly concentrated at the low frequencies, the errors of the estimate $\hat{H}_2(k_1,k_2)$ appear to be stronger as far as (k_1+k_2) goes towards high frequency region. The third term $S_{y3}(k)$ contributes the less power to the output spectrum and therefore one seems to be the most corrupted.

8. Conclusion

An algorithm for identification of discrete nonlinear systems in terms of the orthogonal series was presented. The process generated by inverse FFT algorithm was used as a test signal. For this input the, G-functionals and Wiener kernels were defined in the frequency domain. The proposed algorithm offers a significant reduction in computational complexity compared with the known methods since the number of multiplications does not depend on the kernel order.

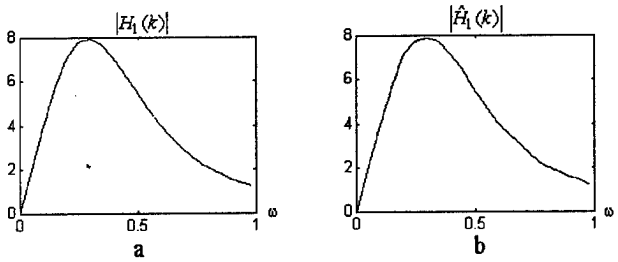


Figure 3. Magnitude of the first order kernel: (a) actual, (b) estimated.

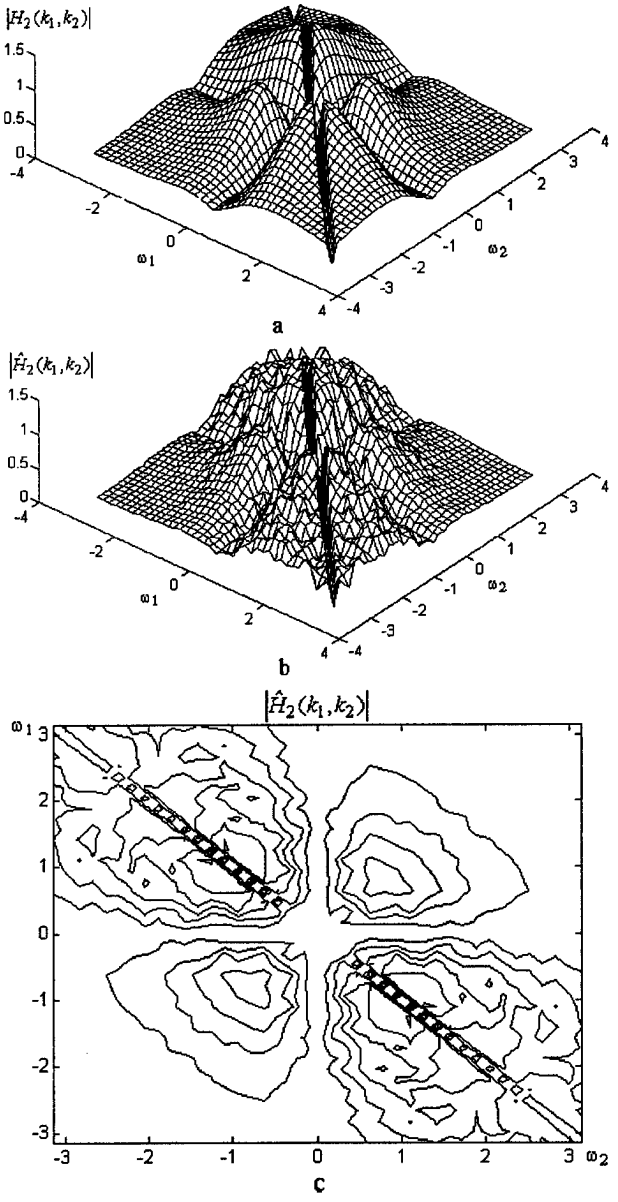


Figure 4. Magnitude of the second order kernel: (a) actual value; (b) kernel estimate; (c) contour plot of the kernel estimate.

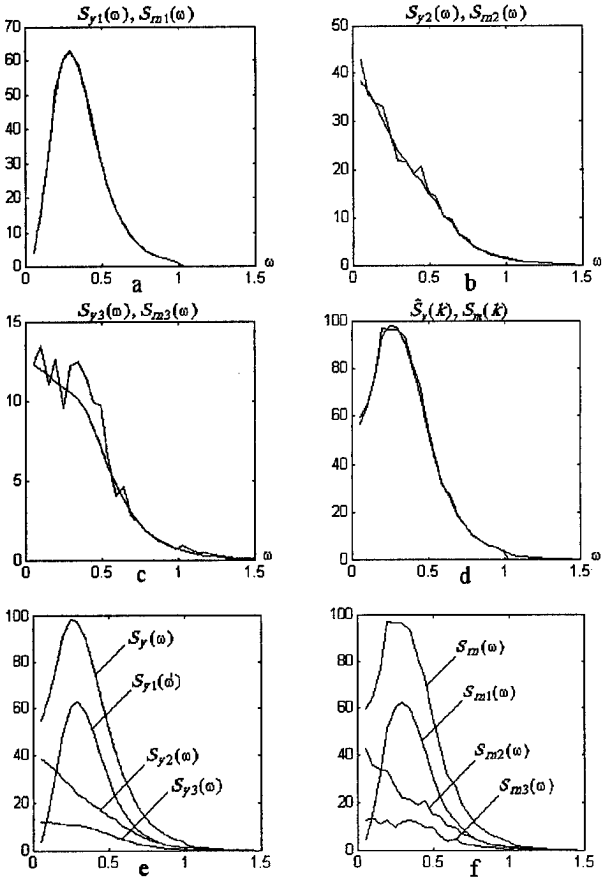


Figure 5. Modeling of the output spectral components.

References

[1] N. Wiener. *Nonlinear problems in random theory*. MIT Press, 1958.

[2] M. Schetzen. *The Volterra and Wiener theories of nonlinear systems*. John Wiley, New York, 1980.

[3] S.A. Billings. Identification of nonlinear systems-A survey. *Proc. IEE*, 127-D(6):272-285, 1980.

[4] K.I. Kim and E.I. Powers. A digital method of modeling quadratically nonlinear systems with a general random input. *IEEE Tras. on ASSP*, 36(11):1758-1769, 1988.

[5] B.G. Mertzios. Parallel modeling and structure of nonlinear Volterra discrete systems. *IEEE Trans. Circuits & Syst.* 41(5):359-371, 1994.

[6] P.Z. Marmarelis and V.Z. Marmarelis. *Analysis of physiological systems: the white-noise approach*. Plenum Press, New York, 1978.

[7] Y.W. Lee and M. Schetzen. Measurement of the Wiener kernels of nonlinear system by cross-correlation. *Int. J. Control* . (2):237-254, 1965.

[8] A.S. French and E.G. Butz. Measuring the Wiener kernels of nonlinear system using the fast Fourier algorithm. *Int. J. Control*, 17(3):529-539, 1973.

Second-Order Statistics versus HOS for the Estimation of Harmonics in Additive and Multiplicative Noise

Mounir GHOGHO
National Polytechnics Institute of Toulouse
ENSEEIHT-LEN7-SIC, 2, rue Camichel, 31071 Toulouse Cedex, France
e-mail: ghogho@len7.enseeiht.fr

Abstract

Second-Order Statistics (SOS) have been widely used for the detection and estimation of coherent sinusoids in additive wide-band noise. This paper addresses the detection and estimation of harmonics which have been corrupted by both multiplicative and additive noise. HOS are useful in estimating harmonics of zero mean amplitude where SOS generally fail. The paper analyses and compares the performance of SOS and HOS when the harmonic has both coherent and non coherent powers. We determine thresholds on the coherent-to-non coherent sine wave power ratio which delimitate the regions of optimality of SOS and HOS. Gaussian as well as non Gaussian noise sources are studied.

1 Introduction

Harmonic retrieval is one of the classical and important problems in statistical signal processing. The large existing literature mostly addresses harmonics in additive noise. Multiplicative noise models occur in several applications, e.g. underwater propagation signals [2], radar signals [1], fading communication channels [8] (see [7] for other applications).

Consider a discrete-time harmonic in multiplicative and additive noise

$$x(t) = (A + b(t))\, e^{j(\omega t+\phi)} + \nu(t), \quad t = 0, ..., N-1 \quad (1)$$

with the following assumptions: $AS1$) A is a positive deterministic constant; $AS2$) $b(t)$ is a real white multiplicative noise with variance σ_b^2 and kth-order moment m_{kb}, $k > 2$; $AS3$) ω and ϕ are the harmonic frequency and phase, assumed constant in the range $(0, \pi)$ and a uniformly random variable over $(-\pi, \pi]$ respectively; $AS4$) $\nu(t)$ is a complex white additive noise with variance σ_ν^2 and kth-order absolute moment $m_{k\nu}$, $k > 2$; $AS5$) the multiplicative and additive noise sources are mutually independent.

Several techniques based on Second-Order Statistics (SOS) have been proposed for harmonic retrieval in the additive noise case, i.e. $A \neq 0$ and $\sigma_b^2 = 0$. More recently, some authors (e.g. [6]) advocated the use of HOS when the additive noise is colored and Gaussian. This is motivated by the ability of HOS to suppress the effects of this kind of noise. On the other hand, there is no motivation to use HOS when the Gaussianity assumption is unsatisfied.

For the noise model (1), SOS are useless for harmonic retrieval if $A = 0$. HOS are then required provided $\sigma_b^2 \neq 0$. Fourth-order statistics have been shown to be an efficient tool for this particular harmonic retrieval problem [7][3].

Both SOS and HOS are useful for harmonic retrieval in the general case where $A \neq 0$ and $\sigma_b^2 \neq 0$. So, the following question must be posed: given a multiplicative and additive noisy environment, which statistics are optimal for the detection and estimation of hidden periodicities. The goal of this paper is to provide statistical tools that answer this question. We will show that HOS outperform SOS when the multiplicative noise power exceeds a threshold which depends upon A and the noise distributions.

2 Harmonic Retrieval using SOS and HOS

Under assumptions $AS2$-$AS5$ and $A = 0$, the moments and cumulants of $x(t)$ are identically zero except over a finite set of hyperplanes. In this paper, we focus on the fourth-order statistics. The fourth-order moment is non-zero only along the 1-D slice $m_{4x}(0, \tau, \tau)$. Since the additive noise is not necessarily Gaussian, cumulants are not considered here. Thus, the comparison of SOS and HOS performances will be based on $m_{2x}(\tau)$

0-8186-8005-9/97 $10.00 © 1997 IEEE

and $m_{4x}(0,\tau,\tau)$ which are obtained, in the general case (i.e. A, σ_b^2, $\sigma_\nu^2 \neq 0$), as

$$\begin{aligned} m_{2x}(\tau) &\triangleq E\{x^*(t)x(t+\tau)\} \quad (2)\\ &= A^2e^{j\omega\tau} + (\sigma_b^2+\sigma_\nu^2)\delta(\tau) \\ m_{4x}(0,\tau,\tau) &\triangleq E\{x^*(t)x^*(t)x(t+\tau)x(t+\tau)\} \quad (3)\\ &= \left(A^2+\sigma_b^2\right)^2 e^{j2\omega\tau} + \left[4A^2\sigma_b^2 + \right. \\ &\qquad 4Am_{3b} + m_{4b} - \sigma_b^4 + \\ &\qquad \left. 4\sigma_\nu^2(A^2+\sigma_b^2) + m_{4v}\right]\delta(\tau) \end{aligned}$$

where

$$\begin{aligned} m_{3b} &\triangleq E\left\{b^3(t)\right\} \\ m_{4b} &\triangleq E\left\{b^4(t)\right\} \\ m_{4\nu} &\triangleq E\left\{\left|\nu(t)\right|^4\right\} \end{aligned}$$

The spectrum $S_{2x}(\lambda)$ and trispectrum slice $S_{4x}(0,\lambda,\lambda)$ peak at ω and 2ω respectively:

$$\begin{aligned} S_{2x}(\lambda) &\triangleq \sum_{\tau=-\infty}^{\infty} m_{2x}(\tau)e^{-j\lambda\tau} \quad (4)\\ &= A^2\delta(\alpha-\omega) + \sigma_b^2 + \sigma_\nu^2 \\ S_{4x}(0,\lambda,\lambda) &\triangleq \sum_{\tau=-\infty}^{\infty} m_{4x}(0,\tau,\tau)e^{-j\lambda\tau} \quad (5)\\ &= \left(A^2+\sigma_b^2\right)^2\delta(\alpha-2\omega) + \\ &\qquad 4A^2\sigma_b^2 + 4Am_{3b} + m_{4b} - \sigma_b^4 \\ &\qquad + 4\sigma_\nu^2(A^2+\sigma_b^2) + m_{4v} \end{aligned}$$

Note that the $m_{4x}(0,\tau,\tau)$-based methods require ω to be in $(0,\pi/2)$ in order to avoid the aliasing phenomenon.

It is worth noting that $m_{4x}(0,\tau,\tau)$ is the covariance function of $x^2(t)$. The detection and estimation of a hidden periodicity in (1) can then be performed by applying to $x(t)$ or $x^2(t)$ one of the numerous techniques of detecting and estimating harmonics in additive noise. For example, the harmonic frequency can be estimated using the peak picking technique. That is, ω can be estimated consistently by

$$\widehat{\omega}_{SOS} = \arg\max_{\lambda>0} \widehat{S}_{2x}(\lambda) \quad (6)$$

provided $A \neq 0$; or

$$\widehat{\omega}_{HOS} = \frac{1}{2}\arg\max_{\lambda>0} \widehat{S}_{4x}(0,\lambda,\lambda) \quad (7)$$

provided $A \neq 0$ or/and $\sigma_b^2 \neq 0$, where

$$\begin{aligned} \widehat{S}_{2x}(\lambda) &= \frac{1}{N}\left|\sum x(t)e^{-j\lambda t}\right|^2 \quad (8)\\ \widehat{S}_{4x}(0,\lambda,\lambda) &= \frac{1}{N}\left|\sum x^2(t)e^{-j\lambda t}\right|^2 \quad (9) \end{aligned}$$

which are simply the periodograms of $x(t)$ and $x^2(t)$ respectively.

3 SOS versus HOS

The signal x_t can be written as

$$x_t = s_1(t) + \xi_1(t) \quad (10)$$

where

$$s_1(t) = Ae^{j(\omega t+\phi)}$$

and

$$\xi_1(t) = b(t)e^{j(\omega t+\phi)} + \nu(t).$$

The SOS-based methods regard $s_1(t)$ as the desired signal (coherent sinusoid) and $\xi_1(t)$ as a white additive noise. The signal and noise components are uncorrelated. Using this signal and noise decomposition, the SNR relative to SOS can be defined as

$$\begin{aligned} SNR_{SOS} &= \frac{E\left\{\left|s_1(t)\right|^2\right\}}{E\left\{\left|\xi_1(t)\right|^2\right\}} \\ &= \frac{A^2}{\sigma_b^2+\sigma_\nu^2} \quad (11) \end{aligned}$$

In the same manner, the signal and noise decomposition of x_t^2 is given by

$$x^2(t) = s_2(t) + \xi_2(t) \quad (12)$$

where

$$\begin{aligned} s_2(t) &= (A^2+\sigma_b^2)e^{j(2\omega t+2\phi)} \quad (13)\\ \xi_2(t) &= \left(b^2(t) - \sigma_b^2 + 2Ab(t)\right)e^{j(2\omega t+2\phi)} \quad (14)\\ &\quad + 2\nu(t)(A+b(t))e^{j(\omega t+\phi)} + \nu^2(t) \end{aligned}$$

$\xi_2(t)$ is a (wide-sense) white noise and decorrelated from $s_2(t)$. The SNR relative to HOS is then obtained as

$$SNR_{HOS} = \frac{\left(A^2+\sigma_b^2\right)^2}{4A^2\sigma_b^2 + 4Am_{3b} + m_{4b} - \sigma_b^4 + 2\sigma_\nu^2(A^2+\sigma_b^2) + m_{4\nu} - \sigma_\nu^4} \quad (15)$$

For both SOS and HOS, the greater the SNR, the larger the harmonic detection and estimation performance. We can therefore define the relative efficiency of HOS with respect to (w.r.t.) SOS as

$$\rho \triangleq \frac{SNR_{HOS}}{SNR_{SOS}} \quad (16)$$

Throughout this paper, HOS are said to be optimal if $\rho > 1$ and SOS are optimal otherwise.

Before treating the multiplicative noise effects, note that ρ is uniformly lower than 1 in the pure additive noise case (i.e. $\sigma_b^2 = 0$):

$$\rho = \left(1 + \kappa_{4\nu}\frac{\sigma_\nu^2}{A^2}\right)^{-1} < 1 \qquad (17)$$

where $\kappa_{4\nu}$ is the kurtosis:

$$\kappa_{4\nu} = \frac{m_{4\nu}}{\sigma_\nu^4}$$

Hence, SOS are the optimal statistics in this case. The following proves that this result can be wrong in the presence of multiplicative noise.

Below, we limit our study to multiplicative noise sources whose third-order moments vanish. Such condition is satisfied for symmetrically distributed multiplicative noise.

3.1 Pure Multiplicative Noise

In the absence of additive noise, SNR_{SOS} and SNR_{HOS} can be rewritten as

$$SNR_{SOS} = \frac{A^2}{\sigma_b^2} \qquad (18)$$

$$SNR_{HOS} = \frac{\left(\frac{A^2}{\sigma_b^2}+1\right)^2}{4\frac{A^2}{\sigma_b^2}+\kappa_{4b}-1} \qquad (19)$$

where κ_{4b} is the kurtosis of the multiplicative noise:

$$\kappa_{4b} = \frac{m_{4b}}{\sigma_b^4}$$

PROPOSITION 1. *HOS are optimal w.r.t. SOS if the coherent-to-non coherent sine wave power ratio satisfies the following condition*

$$\frac{A^2}{\sigma_b^2} < \theta_1 \qquad \textbf{(C1)}$$

where

$$\theta_1 = \frac{1}{6}\left(-(\kappa_{4b}-3) + \sqrt{(\kappa_{4b}-3)^2+12}\right) \qquad \square$$

θ_1 is a monotonic decreasing function of κ_{4b}. Thus, the lower κ_{4b}, the more the HOS becomes optimal. The SOS are certainly optimal if (using $\kappa_{4b} > 1$)

$$\frac{A^2}{\sigma_b^2} > \max(\theta_1) = 1 \qquad (20)$$

whatever the multiplicative noise pdf.

To study the relationship between the threshold θ_1 and the noise pdf, the generalized zero mean Gaussian pdf is used. It is a symmetric density function parameterized by two constants, the shape parameter $\alpha > 0$, and the scale or size parameter $\beta > 0$. It is defined by

$$f(b) = \frac{\alpha}{2\beta\Gamma(1/\alpha)}\exp -\left|\frac{b}{\beta}\right|^\alpha \qquad (21)$$

where $\Gamma(.)$ is the gamma function. The Gaussian case is obtained for $\alpha = 2$, and for lower values of α, the shape of f decays at a lower rate than in the Gaussian case. Thus, the generalized Gaussian pdf gives densities ranging from the Gaussian to those with much faster or much slower rates of exponential decay of their tails.

The corresponding absolute moments are given by

$$m_{kb} \triangleq E\left\{|b|^k\right\} = \frac{\Gamma(\frac{k+1}{\alpha})}{\Gamma(\frac{1}{\alpha})}\beta^k \qquad (22)$$

κ_{4b} is then found to be

$$\kappa_{4b} = \frac{\Gamma(\frac{1}{\alpha})\Gamma(\frac{5}{\alpha})}{\Gamma^2(\frac{3}{\alpha})} \qquad (23)$$

κ_{4b} is a decreasing function of α. Thus, the performance gain using the HOS increases with α. Figure 1 displays θ_1 as a function of α.

For Gaussian multiplicative noise ($\alpha = 2$), **(C1)** becomes

$$\frac{A^2}{\sigma_b^2} < \theta_1 = \frac{1}{\sqrt{3}} \simeq 0.57 \qquad (24)$$

SNR_{HOS} versus SNR_{SOS} is shown in figure 2. The SNR_{HOS} remains almost constant ($\cong 0.5$) for SNR_{SOS} ranging from 0 to 0.57. Figure 4 shows that the simulation results are in accordance with the theoretical predictions. The triperiodogram slice (9) provides superior enhancement of the line spectrum w.r.t the periodogram for small values of $\frac{A^2}{\sigma_b^2}$. The converse is true when $\frac{A^2}{\sigma_b^2}$ exceeds 0.57.

3.2 Multiplicative and Additive Noise Sources

In this section, we study the influence of the additive noise on the above optimality results. The following generalizes proposition 1.

PROPOSITION 2. *HOS are optimal w.r.t. SOS if the coherent-to-non coherent sine wave power ratio satisfies the following condition*

$$\frac{A^2}{\sigma_b^2} < \theta_2 = \frac{1}{6}\left(-\gamma + \sqrt{\gamma^2+12}\right) \qquad \textbf{(C2)}$$

where

$$\gamma = \left(1 + \frac{\sigma_\nu^2}{\sigma_b^2}\right)^{-1} \left(\kappa_{4\nu} \frac{\sigma_\nu^4}{\sigma_b^4} + 2\frac{\sigma_\nu^2}{\sigma_b^2} + \kappa_{4b} - 3\right) \qquad \square$$

For Gaussian multiplicative and additive noise sources, the optimality threshold in (**C2**) can be rewritten as

$$\theta_2 = -\frac{1}{3}\frac{\sigma_\nu^2}{\sigma_b^2} + \frac{1}{3}\sqrt{\frac{\sigma_\nu^4}{\sigma_b^4} + 3} \qquad (25)$$

Figure 3 displays the regions of optimality of SOS and HOS. It appears that for fixed values of A and σ_b^2, the additive noise does not significantly degrade the optimality performance of HOS provided $\frac{\sigma_\nu^2}{\sigma_b^2} \lessapprox 0.1$.

4 Conclusions

This paper has considered the problem of detecting and estimating a complex sine wave in both multiplicative and additive noise. HOS are shown to be preferable to SOS in certain multiplicative and additive noisy environments. This is true not only for zero mean harmonic amplitude, but also for values of the mean ranging from 0 to a threshold which depends on the probability density functions of the noise sources. The choice of the optimal statistics is crucial for short data record lengths. It is also shown that the performance gain using SOS increases for heavy-tailed noise pdfs. The optimality conditions established in this paper are also valid for the choice between the cyclic mean and cyclic variance methods proposed in [9] and [5]. More generally, this paper determines which input is best (in the sense of maximizing the SNR), $x(t)$ or $x^2(t)$, to present to linear processors (e.g. FFT-based detectors, LMS adaptive line enhancer [4], etc.). Finaly, our results can be easily extended to polynomial phase signals.

References

[1] O. Besson and F. Castanié, "On estimating the frequency of a sinusoid in autoregressive multiplicative noise," *Signal processing,* vol.30, no. 1, pp. 65-83, Jan. 1993

[2] R. F. Dwyer, "Fourth-order spectra of a Gaussian amplitude-modulated sinusoids," *J. Acoust. Soc. Am.* 90 (2), Pt. 1, Aug. 1991.

[3] M. Ghogho and B. Garel, "On AR modulated harmonics: CRB and parameter estimation," *Proc. ICASSP'96,* pp. 2507-2510, Atlanta, May 1996.

[4] M. Ghogho, M. Ibnkahla and Neil Bershad, "Analytical behavior of the LMS adaptive line enhancer for sinusoids corrupted by multiplicative and additive noise," submitted to *IEEE Trans. Signal Processing.*

[5] G. Giannakis and G. Zhou , "Harmonics in multiplicative and additive noise: parameter estimation using cyclic statistics," *IEEE Trans. Signal Processing,* vol. 43, pp. 2217-2221, Sep. 1995.

[6] A. Swami and J. M. Mendel, "Cumulant-based approach to harmonic retrieval and related problems," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 39, pp. 1099-1109, May 1991.

[7] A. Swami, "Multiplicative noise models: parameter estimation using cumulants," *Signal processing,* vol. 36, no. 3, Apr. 1994.

[8] H. L. Van Trees, *Detection, Estimation and Modulation Theory. Part* 1, Wiley, New York, 1968.

[9] G. Zhou and G. Giannakis, "Harmonics in multiplicative and additive noise: performance analysis of cyclic estimators," *IEEE Trans. Signal Processing,* vol. 43, pp. 1445-1460, June 1995

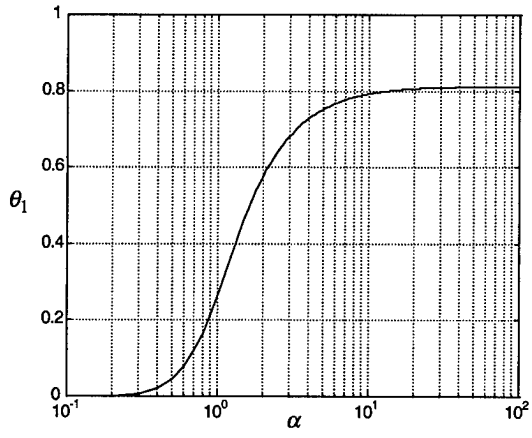


Fig. 1 : Plot of threshold θ_1 as a function of α.

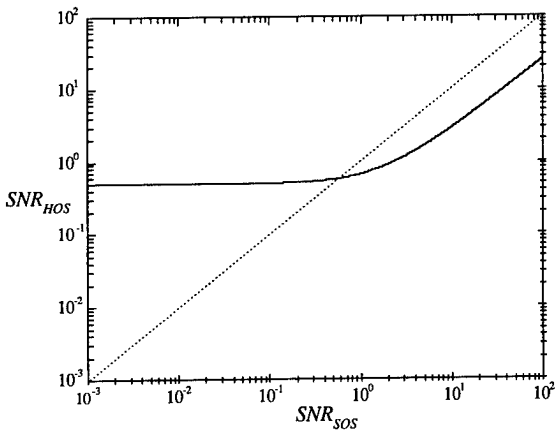


Fig. 2 : SNR_{HOS} versus SNR_{SOS}. Gaussian multiplicative noise

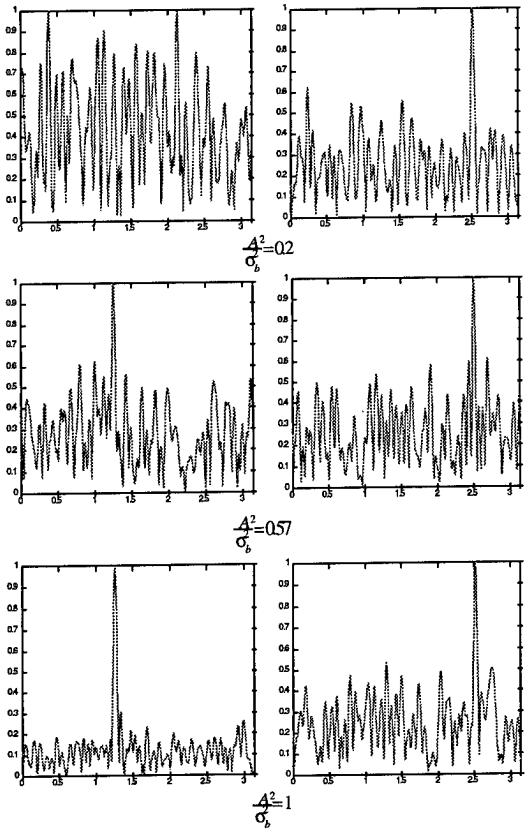


Fig. 4 : Normalized periodogram (left) and triperiodogram (right) for different values of $\frac{A^2}{\sigma_b^2}$. Gaussian multiplicative and additive noise.

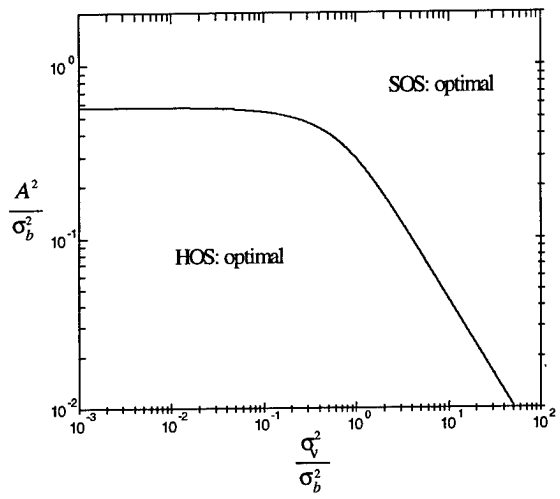


Fig. 3 : Regions of optimality of SOS and HOS . Gaussian multiplicative and additive noise

Design of a Volterra series-based Nonlinear Compensator

J.Y. Kim, K.Y. Cho, Y.N. Kim, J.H. Chung, and S.W. Nam
Dept. of Electrical Eng., Hanyang Univ., Seoul, 133-791, KOREA
Phone: (02) 290-0346; FAX: (02) 297-1569; Email: swnam@email.hanyang.ac.kr

Abstract

In this paper, a new Volterra series-based adaptive preprocessing technique is presented to linearize weakly nonlinear systems with their linear parts being invertible. In particular, a systematic but simple design procedure is proposed for the compensation of system nonlinearities up to a required order, yielding substantial reduction of computation burden. For the performance test of the proposed approach, some simulation results are also provided.

1. Introduction

There have been many signal processing applications where undesirable nonlinearities, inherent in a system, might degrade the overall system performance[1-5]. To compensate for such unwanted nonlinear effects, various linearization methods have been proposed[2-5], among which the Pth-order inverse is one of the commonly used techniques. Although the Pth-order inverse approach has a firm theoretical basis[1], its design complexity might limit its application in real situations: e.g., as the memory size of the system and the order of system nonlinearities increase, its design procedure becomes much more complex. To relieve such a complexity in the design procedure, an adaptive linearization approach[2] was proposed, where only the lowest-order nonlinear part, regarded as a dominant factor degrading the overall system performance, is scheduled to be canceled out in the total system output. However, even in case of such a lowest-order inverse approach, the overall system performance still depends on the amount of remaining nonlinear parts in the system output.

In this paper, a new adaptive nonlinear preprocessing technique, based on Volterra series, is proposed for the linearization of weakly nonlinear systems with their linear parts being invertible, where (i) its design procedure is still simple and (ii) system nonlinearities up to a certain order can be compensated for in a systematic way.

The approach of this paper is based on the extension of the lowest-order inverse[2] and cascade connection of J basic preprocessing (or predistortion) blocks, leading to substantial reduction of computational burden (compared with that in the conventional linearization methods: e.g., the Pth-order inverse).

In the following section, the problem of designing a nonlinear compensator, based on Volterra series, is considered, along with the contraction mapping theorem. Then, the computational complexity required in the proposed approach is discussed in section 3. Finally, to demonstrate the performance and applicability of the proposed approach, some simulation results using a baseband satellite communication channel model and a sample-and-hold circuit model are provided in section 4.

2. Design of a Volterra series-based preprocessor

The Nth-order Volterra series representation of a weakly nonlinear system can be given by

$$y[k]=\sum_{n=1}^{N}\sum_{i_1=0}^{M-1}\sum_{i_2=0}^{M-1}\cdots\sum_{i_n=0}^{M-1}h_n(i_1,i_2,\ldots,i_n)\bullet x[k-i_1]x[k-i_2]\cdots x[k-i_n]=\sum_{n=1}^{N}H_n(x) \quad (1)$$

where M is the memory size of the system, and $x[k]$ and $y[k]$ denote the input and the output, respectively. In addition, $h_n(i_1,\ldots,i_n)$ is an n_{th}-order Volterra kernel, and $H_n(\cdot)$ is the n_{th}-order Volterra operator[1].

The structure of the proposed Volterra series-based nonlinear compensator is presented in Fig.1, where $\hat{H}_1$ and $\sum_{n=2}^{N}\hat{H}_n$, respectively, denote the estimated linear and nonlinear kernels of a given system, obtained adaptively through the system identification procedure by utilizing second- and higher-order statistical information of the input and between the input and output. Also, $\hat{H}_1^{-1}$ is a linear inverse derived from $\hat{H}_1$. In Fig.2(a), the

0-8186-8005-9/97 $10.00 © 1997 IEEE

structure of the multi-block (here, J-block) compensator proposed in this paper is presented, where the J-block compensator can be implemented in a systematic way by adding a basic block G(see Fig.2(b)), composed of $\hat{H}_1^{-1}$ and $\sum_{n=2}^{N}\hat{H}_n$ of a given physical system, to the (J-1)-block compensator. Note that (i) application of the proposed J-block compensator leads to elimination of system nonlinearities up to (J+1)-order, when the second- and higher-order kernels of a given system are not equal to zero, and (ii) the structure of the lowest-order inverse[2] coincides with that of an one-block nonlinear compensator without compensating for linear distortion. Note that the lowest-order inverse compensator[2] is composed of a pure nonlinear filter($\sum_{n=2}^{N}\hat{H}_n$) followed by a linear inverse filter ($\hat{H}_1^{-1}$).

The new compensator proposed in this paper is shown in Fig. 1-2, where $v=\hat{H}_1^{-1}\sum_{n=2}^{N}\hat{H}_n\hat{H}_1^{-1}$ is in parallel with the linear inverse system ($\hat{H}_1^{-1}$). Then, the predistorted signal is $y=\hat{H}_1^{-1}(x_0)-V(x_0)$, and the output of the physical system to be linearized is given by

$$\hat{x}_0 = H_1\left(\hat{H}_1^{-1}(x_0)\right)-H_1\left(V(x_0)\right) + \sum_{n=2}^{N} H_n\left(\hat{H}_1^{-1}\left(x_0-\hat{H}_1\left(V(x_0)\right)\right)\right) \quad (2)$$

To meet the linearization condition $\hat{x}_0 \approx x_0$, the following should be satisfied:

$$H_1\left(V(x_0)\right)=\sum_{n=2}^{N} H_n\left(\hat{H}_1^{-1}\left(x_0-\hat{H}_1\left(V(x_0)\right)\right)\right) \quad (3)$$

For convenience, let $x=x_0-\hat{H}_1\left(V(x_0)\right)$ and define the following:

$$\hat{H}_1 H_1^{-1} = I + e_1 \quad (4)$$

$$\hat{H}_n = H_n + \Delta_n \quad (5)$$

where I is an identity operator, and e_1 and Δ_n are error terms due to imperfect kernel estimation.

Then, eq.(3) can be expressed as follows:

$$\begin{aligned} x &= x_0-\hat{H}_1\left(H_1^{-1}\left(\sum_{n=2}^{N} H_n\left(\hat{H}_1^{-1}(x)\right)\right)\right) \\ &= x_0-\sum_{n=2}^{N}\hat{H}_n\left(\hat{H}_1^{-1}(x)\right) \\ &\quad +\sum_{n=2}^{N}\Delta_n\left(\hat{H}_1^{-1}(x)\right)-e_1\left(\sum_{n=2}^{N} H_n\left(\hat{H}_1^{-1}(x)\right)\right) \end{aligned} \quad (6)$$

Note that if a weakly nonlinear system is identified exactly(i.e., $H_1=\hat{H}_1$ and $H_n=\hat{H}_n$), then, eq.(6) yields the following:

$$x = x_0-\sum_{n=2}^{N}\hat{H}_n\left(\hat{H}_1^{-1}(x)\right)=T(x) \quad (7)$$

In addition, eq.(7) (i.e., finding x) can be solved by applying two methods discussed below. Also, the output of the pre-compensator can be expressed by

$$y=\hat{H}_1^{-1}\left(x_0-\hat{H}_1V(x)\right)=\hat{H}_1^{-1}(x) \quad (8)$$

In particular, it can be shown from eq.(4)-(6) that the compensated output error (i.e. $\hat{x}_0-x_0$) due to the imperfect system identification can be expressed as follows:

$$E_{x_0} = -H_1\hat{H}_1^{-1}\left(\sum_{n=2}^{N}\Delta_n\left(\hat{H}_1^{-1}(x)\right)-e_1\left(\sum_{n=2}^{N} H_n\left(\hat{H}_1^{-1}(x)\right)\right)\right) + e_1^{-1}(x_0) \quad (9)$$

Now, let's consider two methods (Method I and Method II) to solve eq.(7). The difference between two approaches is that Method II uses the previously calculated solutions of eq.(7) for the current time solution to be calculated, but Method I dose not.

2-1. Method I: conventional solution

Let x be a fixed point of the mapping $T: X \rightarrow X$ in eq.(7). Then, it can be obtained by the following iterative procedure, based on the contraction mapping theorem providing the sufficient condition for the existence of the fixed point [3,5,6]:

$$x_j = x_0-\sum_{n=2}^{N}\hat{H}_n\left(\hat{H}_1^{-1}\left(x_{j-1}\right)\right)=T(x_{j-1}) \quad (10)$$

where x_0 is the initial point and the mapping $T(\cdot)$ is described in Fig. 2(b).

It can be shown that another linearization scheme[5], , where the linear inverse filter is located in front of the conpensator and the basic block T is composed of the pure nonlinear filter followed by the linear inverse filter, shows the same performance as the compensator of Fig.2.

2-2. Method II: more efficient solution

Now, consider the problem of designing the adaptive preprocessor (as shown in Fig.2) in an efficient way. It is assume that the linear part of a physical system being

linearized is invertible. As shown in Fig.3, previously calculated data $x_J[l]$, $l<k$ (k: current time), and the current output $x_j[k]$ can be utilized to obtain $x_J[k]$. However, to utilize the previously calculated data $x_J[l]$, each block system G need be expressed in a Volterra series form. To avoid it, the following notations are adopted with the structure of G being maintained:

$$X_j = \left[x_{j-1}[k], x_J[k-1], \ldots, x_J[k-M_L+1] \right] \quad (11)$$

$$Y_j = \left[f_j, y[k-1], y[k-2], \ldots, y[k-M+1] \right] \quad (12)$$

where (i) X_j is the input vector to the linear inverse filter in the j-th block, (ii) Y_j is the input vector to the pure nonlinear filter in the j-th block, (iii) M_L is the memory size of the linear inverse filter, and (iv) f_j is the intermediate value in the j-th block(see Fig. 3).

Now, we will derive the sufficient condition when the above approach yields a contraction mapping (also, see eq.(13)). To derive such condition, it is convenient to use Volterra series expression of the basic block G (here, let G_n be the n_{th}-order Volterra operator of G). Then, the input vector X_j' to the j-th block can be given by

$$x_j[k] = T\left(x_{j-1}[k]\right) = x_0[k] - \sum_{n=2}^{N} G_n\left(x_{j-1}[k]\right) \quad (13)$$

$$X_j' = [x_j[k], x_J[k-1], \ldots, x_J[k-M_L-M+2]] \quad (14)$$

For all $\|x\|, \|y\| \le r$,

$$\|T(x)-T(y)\| = \left\| \sum_{n=2}^{N} \left(G_n(x)-G_n(y)\right) \right\|$$

$$\le \sum_{n=2}^{N} \sum_{k=1}^{n} \sum_{i_1=0}^{M_G-1} \sum_{i_2=0}^{M_G-1} \cdots \sum_{i_n=0}^{M_G-1} \left\| g_n(\underbrace{i_1,\ldots,i_n}_{\#of\ 0 is\ k}) \right\| r^{n-k} \left\| x^k - y^k \right\| \quad (15)$$

$$\le \sum_{n=2}^{N} \sum_{k=1}^{n} \sum_{i_1=0}^{M_G-1} \sum_{i_2=0}^{M_G-1} \cdots \sum_{i_n=0}^{M_G-1} \left\| g_n(\underbrace{i_1,\ldots,i_n}_{\#of\ 0 is\ k}) \right\| k r^{n-1} \left\| x-y \right\|$$

and the sufficient condition for $T(\cdot)$ to be the contraction mapping[6] is given by

$$\alpha = \sum_{n=2}^{N} \sum_{k=1}^{n} \sum_{i_1=0}^{M_G-1} \sum_{i_2=0}^{M_G-1} \cdots \sum_{i_n=0}^{M_G-1} \left\| g_n(\underbrace{i_1,\ldots,i_n}_{\#of\ 0 is\ k}) \right\| k r^{n-1} < 1 \quad (16)$$

where $M_G = M_L + M - 1$ and r, satisfying the condition of eq.(16), is the upper bound for the applicable magnitude of the input to each basic block T.

3. Computational Complexity

To calculate the computational burden for the Volterra series-based preprocessor discussed in section 2.2, examine the number of multiplications required per system output. As shown in Fig. 2, each block G requires M_L multiplications for the linear inverse filter(LIF) and $\sum_{n=2}^{N} n\binom{n+M-1}{n}$ multiplications for the pure nonlinear part. Since each block is an independent Volterra system of same structure, the J-block compensator requires J times the number of multiplications per output required for one-block compensator(i.e., when Method I is applied).

As shown in Fig. 3(i.e., Method II), some operations, executed exactly in the same way, exist in every block G (i.e., in both the linear inverse and the pure nonlinear operations). Thus, it can be easily shown that the number of multiplications per output for J linear inverse operations is $(M_L-1)+J+M_L$ and the number of multiplications per output for J pure nonlinear operations is given by

$$\sum_{n=2}^{N} \left(\sum_{k=1}^{n-1} k\binom{k+M-2}{k} + n\binom{n+M-2}{n} \right) + J\sum_{n=1}^{N} n \quad (18)$$

For example, if J=3 $N=3$, $M=4, M_L=10$, then number of multiplications required in Method II is 85 and that needed in Method I is 220.

4. Simulation Results

To verify the performance of the proposed method, two simulation results, obtained by analyzing a baseband satellite communication channel model and a sample-and-hold circuit model, are given in the next two sections. In particular, note that a post-compensator, which can be also designed for the linearization of the sample-and-hold circuit has the same structure as the pre-compensator[1].

4-1. Simulation I

In this simulation, the performance of the proposed nonlinear pre-compensator is tested by utilizing a baseband model of a satellite communication channel[4], as shown in Fig.4: the transmit and receiver filters are given by TX=[0.8, 0.1] and RX=[0.9, 0.2, 0.1], respectively, and the nonlinearities of TWT(traveling wave tube) may be characterized by the following

AM/AM and AM/PM conversions:

$$A(r) = \frac{\alpha_a r}{1 + \beta_a r^2} \quad (19)$$

$$\Phi(r) = \frac{\alpha_\phi r^2}{1 + \beta_\phi r^2} \quad (20)$$

In eq.(19)-(20), $\alpha_a = 2, \beta_a = 1, \alpha_\phi = \pi/3, \beta_\phi = 1$, and r is the input amplitude. If θ is the phase of the TWT input, then the amplitude and phase of the TWT output can be expressed by $A(r)$ and $\Phi(r)+\theta$. Then, the satellite channel can be modeled by Volterra series with odd-order nonlinearities[3,4]. For this simulation, its third-order Volterra model, and 16-PSK signals with 0.68 magnitude (as the input to the predistortion compensator) are applied. The simulation results are given in Fig. 5(b) (i.e., the compensated output when Method II is applied) and in Table I (i.e., NMSE vs. # of basic blocks), where the efficiency of Method II is indicated, compared with that of Method I.

4-2. Simulation II

The nonlinearity in a sample-and-hold circuit may be caused by the input-independent and input-dependent timing jitters, and the relation between the input x[k] and output f(x[k]) can be approximated as follows[4]:

$$f(x[k]) \approx x[k] + \frac{c}{T}(x[k] - x[k-1])(1 - |x[k]|) \quad (21)$$

where c is a constant determined by the circuit parameters and T is a sampling interval. In this simulation, c=0.02 and the input is a random signal uniformly distributed between -3.0 and 3.0. The simulation results are shown in Fig. 6, with the convergent property of the NMSE in the compensated output when Method II is applied Note that the two-block compensator shows almost the same performance as the three-block one does, where little performance improvement for the three-block case is observed (this is closely related to modeling error, as mentioned in eq.(9), due to the third-order Volterra system identification of the sample-and-hold circuit).

For another test, a Volterra model of the sample-and-hold circuit (not using eq.(21)) with 10 MHz sampling frequency and 0.98 MHz mono-tone input signal [4] is utilized. In this case, the compensator can be designed by using the estimated kernels (i.e., by finding the inverse of the Volterra filter: thus, the error of eq.(9) does not exist in this case). The simulation results are plotted in Fig. 7, where it is shown that if more basic blocks are added to the system, higher-order nonlinear terms (e.g., 3-th , 5-th order, etc.) can be compensated for in a systematic way.

5. Conclusion

In this paper, a systematic, but simple, design procedure for a Volterra series-based adaptive preprocessor (or predistortor) is presented to linearize weakly nonlinear systems. In particular, it is shown that system nonlinearities up to a required order can be compensated for in a systematic way by adding appropriate basic compensation blocks to the corresponding system, and the computational burden for such linearization can be considerably reduced by utilizing the fact that each block in the compensator has the same structure. Finally, the performance of the proposed approach is verified by utilizing a baseband satellite communication channel model and a sample-and-hold circuit model.

References

[1] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear System*, John Wiley and Sons, Inc., New York, 1980.

[2] F.X.Y. Gao, "Adaptive Linear and Nonlinear Filters," Ph.D. dissertation, The Univ. of Toronto, Nov. 1991.

[3] I.S. Park and E.J. Powers, "A New Predistorter Design Technique for Nonlinear Digital Communication Channels," *Proc. Fourth Int. Sym. On Signal Processing and its Appl.*, Vol. 2, pp.618-621, Gold Coast, Australia, Aug., 1996.

[4] C.S. Eun, "Design and Comparison of Nonlinear Compensators," Ph.D. dissertation, The Univ. of Texas at Austin, Dec. 1995.

[5] R.D. Nowak and B.D. Van Veen, "Volterra Filter Equalization: A Fixed Point Approach," *IEEE Trans. on Signal Processing*, vol. 45, no. 2, pp. 377-388, Feb. 1997.

[6] E. Kreyszig, *Introductory Functional Analysis with Applications,* John Wiley and Sons, Inc., New York, 1978.

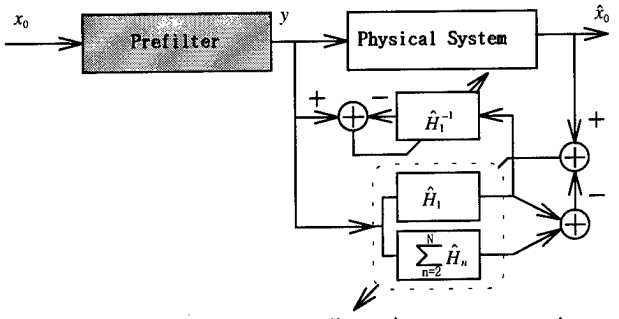


Fig. 1. The proposed predistortion compensator

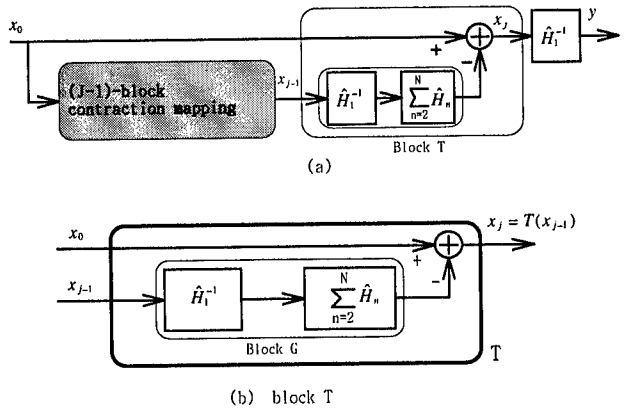


Fig. 2 The J-block compensator (Method I): (a) its structure; (b) contraction mapping block T used in (a) (The initial value of x_j is x_0).

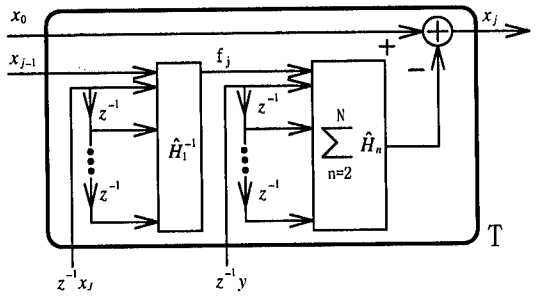


Fig. 3. The basic block T (Method II: here, the initial value of x_j is x_0).

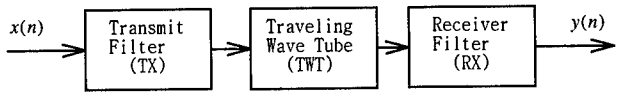


Fig. 4. The baseband model satellite communication channel

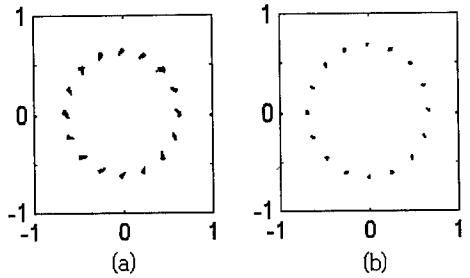


Fig. 5. The compensated output of a satellite communication channel when Method II is used: (a) 1-block, (b) 5-bolck

Table I. NMSE vs. # of blocks for the satellite communication channel.

# of blocks (J)	Method I	Method II
1-block	0.00659	0.00595
2-block	0.00183	0.00153
3-block	0.00057	0.00050
4-block	0.00030	0.00028
5-block	0.00015	0.00015

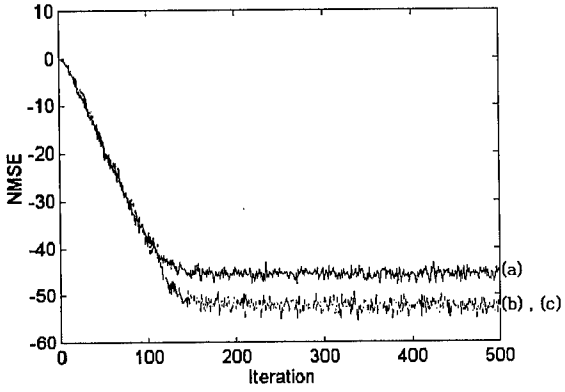


Fig. 6. The NMSE in the compensated output of the sample-and-hold model: (a) 1-block, (b) 2-block, and (c) 3-block.

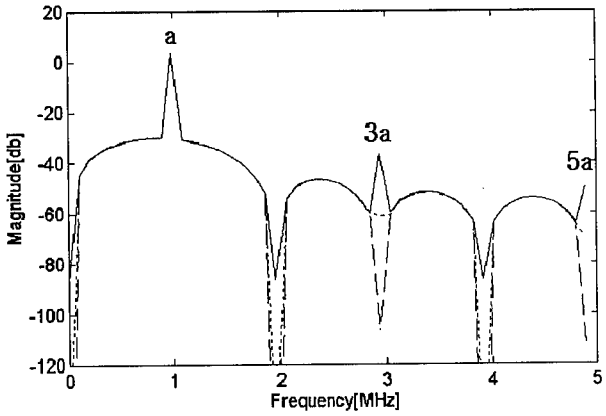


Fig. 7. Spectral distribution of the compensated output of the sample-and-hold circuit with a monotone(0.98MHz) input signal: dotted line: 1-block; dashed line: 3-block; solid line: uncompensated.

Nonlinearly Constrained Optimisation Using a Penalty-Transformation Method for Volterra Parameter Estimation

Tania Stathaki
Signal Processing and Digital Systems Section
Imperial College, UK

Abstract

This paper forms a part of a series of recent studies we have undertaken, where the problem of nonlinear signal modelling is examined. We assume that the observed "output" signal is derived from a Volterra filter that is driven by a Gaussian input. Both the filter parameters and the input signal are unknown and therefore the problem can be classified as blind or unsupervised in nature. In the statistical approach to the solution of the above problem we seek for equations that relate the unknown parameters of the Volterra model with the statistical parameters of the "output" signal to be modelled. These equations are highly nonlinear and their solution is achieved through a novel constrained optimisation formulation. The results of the entire modelling scheme are compared with recent contributions.

1. Introduction

Nonlinear filters constitute an important part of statistical signal processing for non gaussian processes since the use of linear filters to extract signals of interest yield suboptimal solutions in the presence of such processes.
Considerable attention has been focussed in published papers (for example see [1-5]), on the subject of nonlinear signal modelling by means of Volterra representations.
In the blind form of the problem, only the output is observable, and thus for any subsequent modelling we must use only the measured output data. A fundamental assumption underlying many of the approaches to the problem it to take the "input" to be a stationary random process with Gaussian statistics, an assumption which allows substantial simplification in the closed-form expressions for the Volterra kernels. It has been shown that, for a zero mean white Gaussian input, nonlinear expressions for the linear and quadratic transfer functions can be derived in terms of various spectral moments up to third order (i.e. the bispectrum) of the signal [4-5].

The outstanding difficulty in such modelling is related to the quest for finding solutions for these highly nonlinear equations that yield the Volterra kernels.
In this contribution we employ a quadratic Volterra filter form for the modelling of one dimensional signals.
In our recent studies [4-5] we used the nonlinear equations that yield the Volterra parameters to form an unconstrained optimisation problem which was solved using Lagrange Programming Neural Networks (LPNN) [6].
In [8-9] another approach is proposed for the determination of the Volterra kernels which is based on constrained optimisation again using LPNN. In [8-9] the contribution is that we pay particular attention to the reliability of the statistical measures used in the process. Indeed, the second order moments are known to be more reliable than higher moments. Use of this fact is made in the construction of the constraints and penalty functions for the optimisation problem.
In this paper we pay again particular attention to the second order statistical measures but in a different sense. More specifically we use the equations that relate the unknown parameters of the model with the autocorrelations of the signal to form a multivariate penalty function. This function is incorporated in the cost function and yields a so called transformation minimisation function. The method presented in this study belongs to a class of optimisation methods called penalty-transformation methods [10]. These seem to provide better results to the Volterra parameter estimation problem than the Lagrange methods, as far as robustness and rate of convergence are concerned.

2. Preliminaries

We represent the signal as the output of a non linear time invariant causal system driven by noise $x[n]$. The Volterra representation of the input/output relationship is given by

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$y[n] = h_o + \sum_i h_1[i]x[n-i] + \sum_i \sum_j h_2[i,j]x[n-i]x[n-j]$$

$$+ \sum_i \sum_j \sum_k h_3[i,j,k]x[n-i]x[n-j]x[n-k] + \ldots$$

where $h_1[i]$, $h_2[i,j]$, $h_3[i,j,k]$,... are the linear, quadratic, cubic etc. filter weights or kernels respectively, h_o is a constant term whose value depends on the input $x[n]$ and $0 \le i,j,k \le N-1$ where N denotes the filter length [4-5].

As a linear time-invariant (LTI) system is completely characterised by its unit impulse response, so a nonlinear system which can be represented by a Volterra series is completely characterised by its Volterra kernels.

Identification of systems which contain anything higher than second-order kernels is a very difficult task because of the excessive amount of computation. In this paper we consider second order, causal, Volterra filters with symmetric kernels. A symmetric kernel is a symmetric function of its arguments so that for n arguments $\tau_1, \tau_2, \ldots, \tau_n$ there are $n!$ possible interchanges that leave $h_n[\tau_1, \tau_2, \ldots, \tau_n]$ unchanged. Specifically, for $n=2$, $h_2[\tau_1, \tau_2]$ is symmetric if $h_2[\tau_1, \tau_2] = h_2[\tau_2, \tau_1]$.

The second order Volterra filter is given by the following relationship:

$$y[n] = \sum_i a[i]x[n-i] + \sum_i \sum_j b[i,j]x[n-i]x[n-j]$$

where we assumed that $h_o = 0$ without loss of generality [5]. By assuming that the input signal $x[n]$ is a discrete, stationary, zero mean, white Gaussian process, the output process is also discrete, stationary but non-Gaussian process and not necessarily zero mean. To simplify the expressions we assume that the output is zero mean in which case the following condition $\sum_i b[i,i] = 0$ must hold.

3. Second order statistical analysis

In this section the second order statistical analysis of the output will be considered. The autocorrelation function of the real process $y[n]$ is given by $R[k] = E\{y[n]y[n+k]\}$ and in view of the second order Volterra model can be written as:

$$R[k] = \sum_i a[i]E\{x[n-i]y[n+k]\}$$

$$+ \sum_i \sum_j b[i,j]E\{x[n-i]x[n-j]y[n+k]\} \quad (1)$$

The terms of equation (1) involve averaging over the product of one, two, three and four Gaussian random variables. It is known that the average of the product of an odd number of zero-mean jointly Gaussian random variables is identically zero irrespective of their mutual correlation. Moreover, the average of the product of an even number of zero-mean jointly Gaussian random variables is equal to the summation over all distinct ways of partitioning the random variables into products of averages of pairs. For example, if x_1, x_2, x_3, x_4 are zero-mean jointly Gaussian random variables, then:

$$E\{x_1x_2x_3\} = 0 \quad (2)$$

$$E\{x_1x_2x_3x_4\} =$$

$$E\{x_1x_2\}E\{x_3x_4\} + E\{x_1x_3\}E\{x_2x_4\} + E\{x_1x_4\}E\{x_2x_3\} \quad (3)$$

With (2) and (3), $R[k]$ reduces to the form:

$$R[k] = \beta \sum_i a[i]a[i+k] + 2\beta^2 \sum_i \sum_j b[i,j]b[i+k, j+k] \quad (4)$$

where β is the variance of the input driving noise.

The autocorrelation function given by (4) is not sufficient to solve the problem because the number of unknowns present in these equations is much greater than the number of useful samples of $R[k]$. For example for the one dimensional case if the kernels $a[i]$ and $b[i,j]$ have length N, then the number of equations provided is N, while the number of unknowns is $N(N+3)/2+1$. However, additional information can be provided by examining higher order statistics [5].

4. Third order statistical analysis

In this section additional information is by examining the higher order statistics of the output, and this information can be employed towards the solution of the nonlinear equations. If we define $M[k,l]$ to be the third order moment sequence of $y[n]$, then [5]:

$$M[k,l] = E\{y[n]y[n+k]y[n+l]\} \quad (5)$$

In the following, the third-order moment sequence of the second order Volterra filter is derived. First we use the following symbols:

$$G_1[k] = \sum_i a[i]x[n-i+k] \quad (6)$$

$$G_2[k] = \sum_i \sum_j b[i,j]x[n-i+k]x[n-j+k] \quad (7)$$

Based on (6), (7) one can easily expand (5) in the following compact form:

$$M[k,l] = E\{G_1[0]G_1[k]G_1[l] + G_1[0]G_1[k]G_2[l]$$

$$+G_1[0]G_2[k]G_1[l] + G_1[0]G_2[k]G_2[l] + G_2[0]G_1[k]G_1[l]$$

$$+G_2[0]G_1[k]G_2[l] + G_2[0]G_2[k]G_1[l] + G_2[0]G_2[k]G_2[l]\} \quad (8)$$

The first, fourth, sixth and seventh terms of (8) involve averaging over an odd number of zero-mean jointly Gaussian random variables. Therefore are identically zero. Equation (8) then becomes:

$$M[k,l] = E\{G_1[0]G_1[k]G_2[l] + G_1[0]G_2[k]G_1[l] + G_2[0]G_1[k]G_1[l] + G_2[0]G_2[k]G_2[l]\} \quad (9)$$

Each term of (9) involves averaging over an even number of zero-mean jointly Gaussian random variables. Keeping in mind the procedure we described in the previous paragraph, one can decompose the average of the product of an even number of jointly Gaussian random variables into a summation of products of averages of pairs. The first term of (9) (not using the fact that $b[i,j]$ is a symmetric kernel), can then be written as follows:

$$E\{G_1[0]G_1[k]G_2[l]\} = E\{(\sum_i a[i]x[n-i])(\sum_i a[i]x[n-i+k])(\sum_i\sum_j b[i,j]x[n-i+l]x[n-j+l])\}$$
$$= 2\beta^2 \sum_i \sum_j a[i+l]a[j+l-k]b[i,j]$$

Now we define $\phi_1[k,l]$ to be as follows:

$$\phi_1[k,l] = \sum_i \sum_j a[i]a[j]b[i+k, j+l]$$

It can be proven [8] that

$$E\{G_1[0]G_1[k]G_2[l]\} = 2\beta^2\phi_1[l, l-k] \quad (10)$$

Similarly, one can show that:

$$E\{G_1[0]G_2[k]G_1[l]\} = 2\beta^2\phi_1[k, k-l] \quad (11)$$

$$E\{G_2[0]G_1[k]G_1[l]\} = 2\beta^2\phi_1[-k, -l] \quad (12)$$

The fourth term of (9) is quite different from the first three terms. It involves averaging over the product of four Gaussian random variables as well as averaging over the product of six Gaussian random variables. The latter can be broken into the sum of fifteen terms, where each term involves a product of three averages of distinct pairs of random variables.

By doing so and defining:

$$\phi_2[x,y,z] = = \sum_i \sum_j \sum_k b[i,j]b[i+x, k+y]b[j+z, k]$$

we obtain:

$$E\{G_2[0]G_2[k]G_2[l]\} = 8\beta^3\phi_2[l, l-k, k] \quad (13)$$

which is valid only when $\sum_i b[i,i] = 0$ or $b[i,i] = 0$, $\forall i$. In other case equation (13) is more complicated.

We replace (10), (11), (12), (13) in (9) and we obtain $M[k,l]$. It is now possible to use (4) in conjunction with (9) to provide a sufficient number of nonlinear equations required to solve for the unknown system parameters a and b and the variance of the driving noise β.

5. Lagrange Programming Neural Networks (LPNN)

LPNN is a neural network primarily designed for general nonlinear programming. The methodology is based on the well-known Lagrange multiplier method for general constrained optimisation problems. The essential components of the approach are as follows [6-7].

Consider the following constrained nonlinear programming problem:

Minimise $X(f)$

Subject to $Y(f) = 0$

where $X: R^n \to R$ and $Y: R^n \to R^m$ are given functions and $m \le n$. The components of Y are denoted $Y_1, \ldots, Y_m$.

We can define the Lagrangian function $L: R^{n+m} \to R$ by

$$L(f, \lambda) = X(f) + \lambda^t Y(f)$$

where $\lambda \in R^m$ is referred to as the Lagrangian multiplier. The dynamic equations of the LPNN are then defined as:

$$\frac{df}{dt} = -\nabla_f L(f, \lambda) \quad (14)$$

$$\frac{d\lambda}{dt} = \nabla_\lambda L(f, \lambda) \quad (15)$$

where

$$\nabla_f L(f,\lambda) = \left[\frac{dL}{df_1}, \frac{dL}{df_2}, \ldots, \frac{dL}{df_n}\right]^t \text{ and}$$

$$\nabla_\lambda L(f,\lambda) = \left[\frac{dL}{d\lambda_1}, \frac{dL}{d\lambda_2}, \ldots, \frac{dL}{d\lambda_m}\right]^t, \text{ respectively.}$$

If the network is physically stable, then the equilibrium point (f^*, λ^*) defined by $\frac{df}{dt} = 0$ and $\frac{d\lambda}{dt} = 0$, satisfies the equations:

$$\nabla_f L(f^*, \lambda^*) = \nabla X(f^*) + \nabla Y(f^*)^t \lambda^* = 0$$

$$\nabla_\lambda L(f^*, \lambda^*) = Y(f^*) = 0$$

and thus provides a Lagrange solution to the optimisation problem.

Equations (4) furnish a set of relationships for the autocorrelations while (9) provide the third order moments and both of those quantities can be estimated by standard means from a given signal. We are seeking for the

parameters $\{a[i]\}$, $\{b[i,j]\}$ and β of a Volterra model that would produce these second and third order moments. The problem is therefore as follows: Given the autocorrelation estimates $\rho[k]$ and the third order moments estimates $\mu[k,l]$ obtained from the signal directly, to determine $\{a[i]\}$, $\{b[i,j]\}$ and β.

5.1 Unconstrained optimisation

In [4-5] we present a solution to the above problem in the LPNN sense, which starts initially with a suitable cost function such as the one given below

$$L(f) = \sum_i (\rho[i] - R[i,f])^2 + \sum_i \sum_j (\mu[i,j] - M[i,j,f])^2$$

where $f = (\underline{a}, \underline{b}, \beta)^t$ is a vector formed by the unknown parameters of the Volterra model and the unknown variance of the driving noise and $R[i,f]$, $M[i,j,f]$ are the theoretical expressions for the autocorrelations and third order moments respectively.

From this cost function and in accordance with the above formulations, the LPNN dynamic equations may be set up as in (14). Notice that no constraints have been incorporated. Thus, the Lagrange parameters are set to zero, or the corresponding Lagrange neurons are clamped to zero level.

The signal flow graph of these equations (14) describes the required dynamic neural network structure the steady state of which delivers the solution.

5.2 Constrained optimisation

In [8] the minimisation is carried out under certain constraints which we have chosen in a way to reflect the accuracy of our measurements and of the estimation procedures. For finite duration signals autocorrelations are more accurately estimated than higher order moments, so a constrained nonlinear programming problem is formulated as follows.

Minimise: $L(f) = \sum_i \sum_j (\mu[i,j] - M[i,j,f])^2$

subject to: $\rho[i] = R[i,f]$

In this form we have a constrained optimisation problem for which we form the following Lagrange function

$$L(f,\lambda) = L(f) + \sum_i \lambda_i (\rho[i] - R[i,f])$$

The dynamic equations of the LPNN (which are the updated equations for f and λ) are now defined as in (14) and (15).

The stability of the neural network and the optimality of the solution are guaranteed under some regularity and convexity conditions [6-7].

5.3 Penalty methods

In [9] we pay again particular attention to the second order statistical measures but in a different sense. More specifically we use the equations that relate the unknown parameters of the model with the autocorrelations of the signal to form a penalty function [7]. This function is incorporated to the cost function and yields a so called Augmented Lagrangian function. The problem now has as follows.

$$L(f,\lambda) = L(f) + \sum_i \lambda_i (\rho[i] - R[i,f]) + c_k \sum_i (\rho[i] - R[i,f])^2$$

where $\{c_k\}$ is a penalty parameter sequence satisfying

$$0 < c_k < c_{k+1} \ \forall k, \ c_k \to \infty$$

The development of the above method was motivated by the concept of maintaining implicit control over the violations of constraints by penalising the objective function at points which violate or perhaps tend to violate the constraints.

The methods presented in [8-9] provide solutions closer to the optimum than the unconstrained method in [4-5].

5.4 Penalty-transformation methods

In this study the problem is again formulated as in [8-9].

Minimise:

$$L(f) = \sum_i \sum_j (\mu[i,j] - M[i,j,f])^2$$

subject to:

$Y_i(f) = \rho[i] - R[i,f] = 0, \ i = 0,\ldots,N-1$

We use now a new penalty function [9] defined by

$P(Y(f),\theta,\sigma) = \sum_{i=0}^{N-1} \sigma_i (\theta_i + Y_i(f))^2$ where θ and σ are N – vectors of controlling parameters. N is the length of the Volterra model.

A *transformation function* now becomes

$$T(f,\theta,\sigma) = L(f) + \sum_{i=0}^{N-1} \sigma_i (\theta_i + Y_i(f))^2$$

This method is based on the observation that if $f^{(k)}$ minimizes $T(f,\theta,\sigma)$, then $f^{(k)}$ is also a solution of the problem: Minimise $L(f^{(k)})$ subject to $Y(f^{(k)})$.

The algorithm we use has as follows.

(i) We select $\theta^{(1)} = 0, \ \sigma^{(1)} = 1, \ k = 0$

(ii) $k = k+1$

(iii) Minimise $T(f,\theta^{(k)},\sigma^{(k)})$ to find $f^{(k)}$, starting the unconstrained minimisation from $f^{(k-1)}$

(iv) If $Y(f^{(k)})$ is converging sufficiently rapidly to zero, then

$$\theta^{(k+1)} = \theta^{(k)} + Y(f^{(k)})$$

$$\sigma^{(k+1)} = \sigma^{(k)} + c$$

where $0 < c < 1$ and return to (ii).

Otherwise

$$\theta^{(k+1)} = \frac{1}{\mu}\theta^{(k)}$$

$$\sigma^{(k+1)} = \sigma^{(k)}$$

where $\mu > 1$ and return to (ii).

In this method the penalty parameters that are incorporated into the cost function are changing according to the change of the cost function which is under explicit control. By doing this we try to avoid the possibility that the cost function will diverge because of an independent increase of the penalty parameters.

6. Simulations

In the simulations we use a synthetic one dimensional signal which is described by a quadratic of the following form:

$$y[n] = x[n] + 1.8x[n-1] + 0.5x[n]x[n-1]$$

The size of the signal is 1000 samples. We apply both the constrained optimisation approach introduced in [8] and the penalty-transformation approach presented in this paper for this signal. We are dealing with nonlinear and nonconvex functions so we might encounter difficulties concerning the convergence of the algorithm. In this work we repeat the same experiments starting from different initial points for the unknown parameters, chosen to be in the attraction basin of the global solution that can be approximately determined using *simulated annealing* algorithms. The mean values and the variances of the estimates are calculated. The first table below shows the solutions arising from the method which was presented in [8]. The second table contains the results for the same experiments but using the procedure described in this study.

It is observed from the tables below that the new approach yields improved results as expected also theoretically compared to the method in [8] and to the extent in [4-5]. These results arise from a large number of tests involving different signals.

Parameter	real value	mean estimated value	variance
$a[0]$	1	1.1652	0.0014
$a[1]$	1.8	2.05	0.0021
$b[0,1]$	0.5	0.3744	5.8518e-04
β	1.55	1.5834	0.0031

Parameter	real value	mean estimated value	variance
$a[0]$	1	0.92	1.459e-04
$a[1]$	1.8	2.056	0.001
$b[0,1]$	0.5	0.47	1.92e-04
β	1.55	1.5258	0.0021

REFERENCES

[1] T. Koh and E. J. Powers, "Second-order Volterra filtering and its application to nonlinear system identification", *IEEE, ASSP*, Vol. 33, No. 6. Dec. 1985

[2] K. I. Kim and E. J. Powers, "A digital method of modeling quadratically nonlinear systems with a general random input", *IEEE Trans. on ASSP*, Vol. 36, No. 11, November 1988.

[3] S. P. Parker and F. A. Perry, "A discrete ARMA model for nonlinear system identification", *IEEE Trans. on Circuits and Systems*, Vol. CAS-28, No. 3, March 1981.

[4] Tania Stathaki, "Two dimensional blind Volterra signal modelling", *Proc. of the IEEE Workshop on Higher Order Statistics,* 12-14 June 1995, SPAIN.

[5] Tania Stathaki, "Blind Volterra signal modelling", *Proc. of the ICASSP 96,* vol. 3, pages 1601-1604, May 7-10, 1996, Atlanta, Georgia, USA.

[6] S. Zhang and A. G. Constantinides "Lagrange programming neural networks", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol. 39, No. 7., July 1992.

[7] D. P. Bertsekas, *Constrained Optimisation and Lagrange Multipliers Method,* Academic Press, New York, 1982.

[8] Tania Stathaki, "Blind Volterra modelling using constrained optimisation", *Proc. of the Asilomar Conference on Signals, Systems and Computers*, November 3-6 1996.

[9] Tania Stathaki and Anne Scohyers, "A constrained optimisation approach to the blind estimation of Volterra kernels", *Proc. of the ICASSP 97,* vol. 3, pages 1601-1604, April 1997, Munich, GERMANY.

[10] P. E. Gill and W. Murray editors, *Numerical Methods for Constrained Optimisation*, Academic Press, London, 1974.

An iterative mixed norm image restoration algorithm

*Min-Cheol Hong **Tania Stathaki *Aggelos K. Katsaggelos

*Northwestern University
Dept. of Electrical and Computer Engineering
Evanston, IL 60201
Email: {hong,aggk}@ece.nwu.edu

**Imperial College
Signal Processing and Digital System Section
United Kingdom
Email: tania@ic.ac.uk

Abstract

In this paper, we propose an iterative mixed norm image restoration algorithm. A functional which combines the least mean squares (LMS) and the least mean fourth (LMF) functionals is proposed. A function of the kurtosis is used to determine the relative importance between the LMS and the LMF functionals. An iterative algorithm is utilized for obtaining a solution and its convergence is analyzed. Experimental results demonstrate the capability of the proposed approach.

1. Introduction

The image restoration problem has been extensively studied in the past years [1, 5]. In general, the mean squared error (MSE) norm has been used for formulating the restoration problem, resulting in the least mean squared (LMS) approach. The reason for this is that it leads to mathematically tractable solutions and yields optimal results when the contaminating noise has Gaussian distribution.

In a number of application, the contaminating noise may be non-Gaussian or a combination of several noise types [4]. In such cases, norms with order higher than two have been used in adaptive filtering [2, 6, 7]. It has been shown that the LMF approach outperforms the LMS under certain noise distributions, such as sub-Gaussian distributions [7].

In this paper, we propose a mixed norm image restoration algorithm to combine the advantages of the LMS and the LMF in dealing with several noise distributions. In order to control the relative contribution of the LMS and the LMF, a functional of the kurtosis is used.

This paper is organized as follows. In section 2 the mixed norm algorithm is proposed. In section 3 an iterative algorithm is described and its analysis is presented. In section 4 various forms of the functional of the kurtosis are discussed. Finally, experimental results are shown in section 5 and conclusions are discussed in section 6.

2. Mixed norm algorithm

An MxN dimensional image may be degraded due to uniform motion, defocusing, long term atmosphere turbulence, and a combination of them [1]. The typical degradation model has the form

$$y = Hx + n, \tag{1}$$

where vectors y, x, n are of size MNx1, and represent the lexicographically ordered observed degraded image, the original image, and the additive noise, respectively. H is the degradation matrix of size MNxMN which may represent a spatially invariant or spatially varying degradation.

For most applications the noise term n is assumed to be zero-mean Gaussian. There are application, however, for which the additive noise is characterized by other distribution, such as, uniform or Cauchy, or by a combination of them. In the Gaussian noise case an LMS approach is used. For other noise distribution, however, norms of higher order need to be used.

The work we present in this paper is motivated by the use of mixed norm in adaptive filtering [6, 7]. We therefore propose to minimize the following functional $M(x)$ with respect to x, for obtaining an estimate of the restored image

$$M(x) = (1 - \gamma(x))||y - Hx||_2^2 + \gamma(x)||y - Hx||_4^4, \tag{2}$$

where $\gamma(\cdot)$ $(0 \leq \gamma(x) \leq 1)$ is a parameter controlling the relative importance of two terms and $||\cdot||_p$ denotes the l_p norm.

As is the case in most image restoration application we assume that only one realization of the degraded image y is available. Then if we assume that n is ergodic, $||y - Hx||^2$ (divided by the number of pixels in the image) provides an estimate of $E[(y - Hx)_i^2]$, where $E[\cdot]$ denotes statistical expectation and $(z)_i$ is the ith element of a vector z.

The performance of the LMS and LMF algorithms have been investigated in the literature [2, 4, 7]. It was shown

0-8186-8005-9/97 $10.00 © 1997 IEEE

that under certain noise conditions (e.g., sub-Gaussian signals), LMF and other higher criteria exhibit improved performance compared to the LMS. The converse is true for Gaussian and super-Gaussian signals. By using Eq. (2) therefore, it is desired that in the extreme cases of only Gaussian or super-Gaussian noise the contribution of the fourth norm is negligible (i.e, $\gamma(x) \approx 0$), while for sub-Gaussian noise the relative contribution of the fourth norm is large (i.e, $\gamma \approx 1$). On the other hand, for cases of mixed noise, $\gamma(x)$ will be attain values which will control the relative contribution of the two norms. For all cases and for most practical situations, however, the noise distribution is not known. Therefore, it is desirable that the values of $\gamma(x)$ be determined using the available data.

A parameter which determines the Gaussianity of a signal is the kurtosis. For a random variable r it is defined by

$$\chi(r) = E[r^4] - 3E^2[r^2]. \tag{3}$$

The kurtosis is zero for Gaussian signal, it is positive for super-Gaussian or leptokurtic signals and negative for sub-Gaussian or platykurtic signals. In analogy to Eq. (3), we define the kurtosis of a signal when only one realization is available by

$$\chi(r) = \frac{1}{MN}||r||_4^4 - 3(\frac{1}{MN}||r||_2^2)^2. \tag{4}$$

Based on the above discussion we propose to determine $\gamma(x)$ as a function of $\chi(x)$. A number of forms of this function have been investigated, and are discussed in section 4.

3. Iterative solution and convergence analysis

3.1 Development

We propose to use a steepest-descent algorithm for obtaining a solution to the minimization problem of Eq. (2). The gradient of $M(x)$ with respect to x is equal to

$$\begin{aligned}&\nabla_x M(x) = -2(1-\gamma(x))H^T(y-Hx) \\ &-4\gamma(x)H^T(y-Hx)^3 + [-\nabla_x\gamma(x)||y-Hx||_2^2 \\ &+\nabla_x\gamma(x)||y-Hx||_4^4].\end{aligned} \tag{5}$$

A closer look at the last term inside the brackets reveals that it are very small, and therefore it is omitted in the following analysis. It is also confirmed experimentally that the restoration results are indistinguishable with and without the use of the last term. The resulting iteration then takes the form

$$\begin{aligned}&x_{k+1} = x_k + \beta[(1-\gamma(x_k))H^T(y-Hx_k) \\ &+2\gamma(x_k)H^T(y-Hx_k)^3] = \\ &x_k + \beta H^T[(1-\gamma(x_k))I + 2\gamma(x_k)P(x_k)](y-Hx_k),\end{aligned} \tag{6}$$

where β is the relaxation parameter which controls the convergence, as well as, the convergence rate and $P(x)$ is an MNxMN diagonal matrix with diagonal elements $P(x_k)_{i,i} = (y_i - (Hx_k)_i)^2$.

3.2 Convergence analysis

In this analysis we follow steps similar to the ones in [3]. Iteration (6) for two consecutive values can be rewritten as

$$\begin{aligned}&x_{k+1} - x_k = (x_k - x_{k-1}) \\ &+\beta[-H^TH(x_k - x_{k-1}) - H^T(F(x_k) - F(x_{k-1})) \\ &+H^TH(G(x_k) - G(x_{k-1})) \\ &+2H^T(K(x_k) - K(x_{k-1})) \\ &-2H^T(L(x_k) - L(x_{k-1})],\end{aligned} \tag{7}$$

where $F(x_k) = \gamma(x_k)y$, $G(x_k) = \gamma(x_k)x_k$, $K(x_k) = \gamma(x_k)P(x_k)y$, and $L(x_k) = \gamma(x_k)P(x_k)Hx_k$. The nonlinear factors $F(x_k), G(x_k), K(x_k)$, and $L(x_k)$ can be linearized, that is,

$$\begin{aligned}F(x_k) - F(x_{k-1}) &\approx J_F(x_k)(x_k - x_{k-1}), \\ G(x_k) - G(x_{k-1}) &\approx J_G(x_k)(x_k - x_{k-1}), \\ K(x_k) - K(x_{k-1}) &\approx J_K(x_k)(x_k - x_{k-1}),\end{aligned}$$

and

$$L(x_k) - L(x_{k-1}) \approx J_L(x_k)(x_k - x_{k-1}),$$

where $J_F(x_k)$, $J_G(x_k)$, $J_K(x_k)$ and $J_L(x_k)$ are the Jacobian matrices. Their mnth elements denoted by $J_F(x_k)_{m,n}$, $J_G(x_k)_{m,n}$, $J_K(x_k)_{m,n}$ and $J_L(x_k)_{m,n}$ and are equal to

$$\begin{aligned}J_F(x_k)_{m,n} &\approx 0, \\ J_G(x_k)_{m,n} &\approx \gamma(x_k)\delta_{m,n}, \\ J_K(x_k)_{m,n} &\approx \gamma(x_k)\frac{\partial P(x_k)_{m,m}}{\partial x_{k,n}}y_m,\end{aligned} \tag{8}$$

and

$$\begin{aligned}J_L(x_k)_{m,n} \approx\ & \gamma(x_k)\frac{\partial P(x_k)_{m,m}}{\partial x_{k,n}}\sum_i h_{m,i}x_i \\ &+\gamma(x_k)P(x_k)_{m,m}h_{m,n},\end{aligned} \tag{9}$$

where $\delta_{m,n} = 1$ when $m = n$. From Eqs. (8) and (9),

$$\begin{aligned}&J_K(x_k)_{m,n} - J_L(x_k)_{m,n} = \\ &\gamma(x_k)[-2(y_m - \sum_i h_{m,i}x_{k,i})h_{m,n}y_m \\ &+2(y_m - \sum_i h_{m,i}x_{k,i})h_{m,n}\sum_i h_{m,i}x_{k,i} \\ &-(y_m - \sum_i h_{m,i}x_{k,i})^2h_{m,n}] \\ &= -3\gamma(x_k)P(x_k)_{m,m}h_{m,n}.\end{aligned} \tag{10}$$

Therefore, Eq. (7) can be rewritten as

$$x_{k+1} - x_k = \quad (11)$$
$$[I - \beta((1-\gamma(x_k))H^TH + 6\gamma(x_k)H^TP(x_k)H)] \cdot$$
$$(x_k - x_{k-1}).$$

Considering the norm of both sides of Eq. (11), the sufficient condition for convergence becomes

$$||I - \beta((1-\gamma(x_k))H^TH + 6\gamma(x_k)H^TP(x_k)H)|| < 1. \quad (12)$$

Since $(1-\gamma(x_k))H^TH + 6\gamma(x_k)H^TP(x_k)H$ is a positive definite matrix, this condition becomes

$$|1-\beta(\lambda_{max}((1-\gamma(x_k))H^TH+6\gamma(x_k)H^TP(x_k)H))| < 1, \quad (13)$$

where $\lambda_{max}(A)$ denotes the maximum singular value of the matrix A. Condition (13) from a sufficient condition for convergence which needs to be satisfied at each iteration step. Using the triangular inequality this condition becomes

$$0 < \beta < \frac{2}{(1-\gamma(x_k)) + 6\gamma(x_k)\lambda_{max}(P(x_k))}, \quad (14)$$

because $\lambda_{max}(H^TH) = \lambda_{max}(H^T) = \lambda_{max}(H) = 1$. Since $0 \leq \gamma(x_k) \leq 1$ and $\lambda_{max}(P(x_k))$ is greater than one and even tighter upper bound for β is $\frac{2}{6\lambda_{max}(P(x_k))}$.

4. Choice of mixed norm parameter

The desirable properties of $\gamma(x)$ were mentioned in section 2. In addition, in section 3, we established upper and lower bounds for $\gamma(x)$ based on the convergence analysis of the resulting iteration.

A number of functions defining $\gamma(x)$ in terms of the kurtosis were tested. Two of them are the following, with $n_k = y - Hx_k$,

$$\gamma(x_k) = \gamma(x_{k-1}) - c_1\chi(n_k),\ c_1 > 0, \quad (15)$$

$$\gamma(x_k) = \frac{\exp(-c_4\chi(n_k))}{1+\exp(-c_4\chi(n_k))},\ c_4 > 0. \quad (16)$$

In both cases, $\gamma(x_k)$ is thresholded so that it belongs to the range of values $[0, 1]$.

The updating equations (15) and (16) ensure that for values $\chi(x_k) > 0$ (leptokurtic noise distribution) the importance of the fourth norm is reduced relatively to the second norm. For values of $\chi(x) < 0$ (platykurtic noise distribution) the importance of the fourth norm is increased relatively to the second norm. The important advantage of the proposed algorithm is that the relative importance of the two norms is adjusted automatically based on the partially restored image, without the need to specify in advance the noise distribution(s). This is particularly important when mixtures of noise distributions are used.

5. Experimental results

In our experiments, we used the 256x256 pixels image shown in Fig. 1. The original image is blurred by 7x7 uniform motion blur. Uniform (sub-Gaussian) and/or Cauchy (super-Gaussian) distributed noise signals were added to the blurred image. We tested the proposed algorithm for various signal to noise ratios (SNR). For evaluating the performance of the algorithm, the improvement in SNR (dB) was utilized. It is defined at the kth iteration by

$$\Delta_{SNR} = 10\log_{10}\frac{||y-x||_2^2}{||x_k-x||_2^2}. \quad (17)$$

The criterion

$$\frac{||x_{k+1}-x_k||_2^2}{||x_k||_2^2} \leq 10^{-6} \quad (18)$$

was used for terminating the iteration.

In Figs. 2 and 3, the noisy and blurred images with 20 dB Cauchy noise, and 20 dB uniform noise are shown, respectively. The corresponding restored images by the proposed algorithm, where $\gamma(x)$ in Eq. (16) was used with $c_4 = 10^{-5}$, are shown in Figs. 4 and 5. In Table 1, comparative results are shown. In all cases the blur used is the 7x7 uniform blur. The combination noise is the one resulting from equal contributions of uniform and Cauchy noise. The LMS and LMF approaches are compared with the results of the proposed algorithm. By LMMN1 and LMMN2 the proposed algorithm is denoted with the use respectively of Eqs. (15) and (16), with $c_1 = c_4 = 10^{-5}$.

By comparing the entries of Table 1 it is clear that the proposed algorithm performs similarly to the LMF algorithm for the uniform noise case and similarly to the LMS algorithm for the Cauchy case, as expected. For the combined noise case, the proposed algorithm (LMMN2) performs slightly better than both the LMS and LMF.

The important point to be made here is that with the proposed algorithm no knowledge of the noise distribution is required, and the relative contribution of the LMS and LMF approaches is adjusted based on the partially restored image.

The logarithm of the error between the original and restored images is compared for one experiment for the LMS, LMF and the proposed LMMN2 approaches in Fig. 6. The LMMN2 and LMF approach approaches converge to the same point, since the noise is uniformly distributed.

The value of the parameter $\gamma(x_k)$ as a function of the iteration number are shown in Figs. 7 and 8. In Fig. 7 for both noise cases, $\gamma(x_k)$ is close to 1, that is only one l_4 term is used, as expected. In Fig. 8, although the noise is uniform, $\gamma(x_k)$ is about 0.5 giving equal contributions to the l_2 and l_4 norms. This is explained by the fact that the

SNR is relatively high (20 dB) and therefore the absolute value of the kurtosis is not as large as for the 10 dB case (the kurtosis is negative in both case).

6. Conclusions

In this paper, we proposed an iterative mixed norm restoration algorithm. The algorithm perform similarly to the LMS approach for Gaussian and super-Gaussian noise distributions and similarly to the LMF approach for sub-Gaussian noise distributions, and for mixed noise distribution. The functional which determines the relative importance between the l_2 and l_4 in the mixed norm formulation is automatically adjusted per iteration based on a measure of the kurtosis of the noise using the partially restored image. Therefore, no prior knowledge of the noise distribution is required. The addition of a smoothing term weighted by the regularization functional in the functional to be minimized is currently under investigation.

References

[1] H. C. Andrews and B. R. Hunt. *Digital Image Restoration.* Prentice-Hall, New York, 1977.

[2] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides. Least mean mixed-norm adaptive filtering. *Electronics Letters*, 30(19):1574–1575, Sept. 1994.

[3] M. G. Kang and A. K. Katsaggelos. General choice of the regularization functional in regularized image restoration. *IEEE Trans. on Image Processing*, 4(5):594–602, May 1995.

[4] S. A. Kassam and H. V. Poor. Robust techniques for signal processing: A survey. *Proc. of IEEE*, 73(3):433–481, Mar. 1985.

[5] A. K. Katsaggelos, editor. *Digital Image Restoration*, volume 23 of *Springer Series in Information Sciences*. Heidelberg : Spring-Verlag, 1991.

[6] O. Tanrikulu and A. G. Constantinides. Least-mean kurtosis : A novel high-order statistics based adaptive filtering algorithm. *Electronics Letters*, 30:189–190, 1994.

[7] E. Walach and B. Widrow. The least mean fourth (lmf) adaptive algorithm and its family. *IEEE Trans. on Information Theory*, IT-30(2):275–283, Mar. 1984.

Figure 1. Original image

Figure 2. Noisy blurred image (7x7 uniform blur, 20 dB Cauchy noise)

Figure 3. Noisy blurred image (7x7 uniform blur, 20 dB uniform noise)

Figure 4. Restored image of Fig. 2 by proposed algorithm, $\Delta_{SNR} = 2.31$ dB

Figure 5. Restored image of Fig. 3 by proposed algorithm, $\Delta_{SNR} = 2.68$ dB

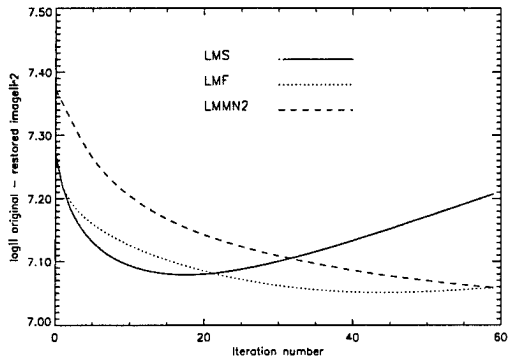


Figure 6. Mean squared error comparison ; 20 dB additive uniform noise

Noise type	SNR (dB)	LMS Δ_{SNR}	LMF Δ_{SNR}	LMMN1 Δ_{SNR}	LMMN2 Δ_{SNR}
Uniform	10	2.98	3.51	3.35	3.41
Uniform	20	2.30	2.71	2.58	2.68
Uniform	30	3.56	3.79	3.66	3.75
Cauchy	10	2.82	1.03	2.81	2.82
Cauchy	20	2.31	0.58	2.31	2.31
Cauchy	30	3.48	1.79	3.43	3.46
Gaussian	10	2.94	2.83	2.90	2.98
Gaussian	20	2.31	2.26	2.28	2.32
Gaussian	30	3.54	3.50	3.51	3.54
Combination	10	2.95	3.33	3.30	3.40
Combination	20	2.30	2.57	2.53	2.66
combination	30	3.61	3.78	3.75	3.81

Table 1. Performance comparison

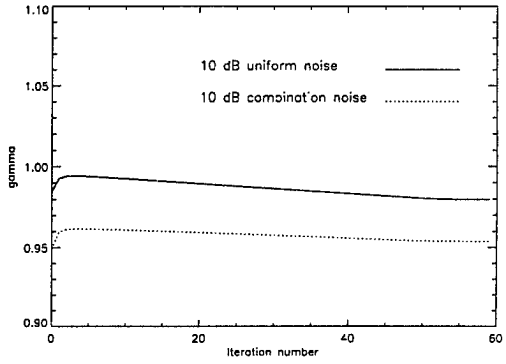


Figure 7. Values of $\gamma(x)$ as function of iteration number

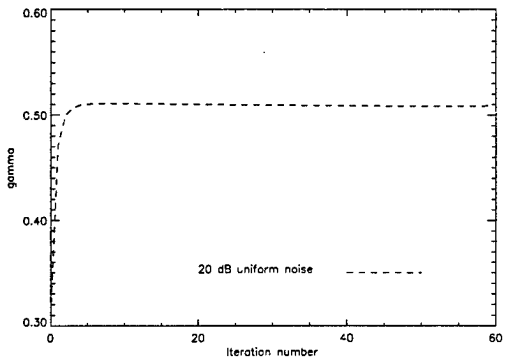


Figure 8. Values of $\gamma(x)$ as function of iteration number

On Identifying Volterra Transfer Functions of Cubically Nonlinear Systems Using Minimally Sampled Data

Ching-Hsiang Tseng
Department of Electrical Engineering
National Taiwan Ocean University
Keelung, Taiwan, R.O.C.
email: cht@hp720a.ee.ntou.edu.tw

Abstract

A practical method for identifying cubically nonlinear systems is presented in this paper. This method identifies the system by using the higher-order spectra of the system input and output. Compared to the conventional method, which requires the system output to be sampled at six times the bandwidth of the input, the proposed method only requires the system output to be sampled at twice the bandwidth of the system input. This greatly reduces the required computation and processing speed of the circuits. The advantage of the proposed method over the conventional one is demonstrated via computer simulation.

1. Introduction

The Volterra series [1, 2, 3, 4, 5] is a well studied model for describing nonlinear systems. In modeling nonlinear systems using Volterra series, the required coefficients in the model increase dramatically with the order of the Volterra series. Therefore, the practical application of the Volterra model is often limited to mildly nonlinear systems which can be modeled by a lower-order Volterra series. However, since the third-order Volterra series is the lowest-order Volterra model including both even and odd order nonlinear terms, the technique of using Volterra model for nonlinear system identification needs to be extended up to at least third order. In fact, one may find in the literature that many nonlinear effects in science and engineering can be appropriately described by a cubically nonlinear system and have been successfully modeled by a third-order Volterra series [6, 7, 8]. For such applications, a reliable and practical method for identifying the cubically nonlinear systems will be appreciated.

Building upon the many quadratically nonlinear system identification techniques [9, 10, 11, 12], the development of identification techniques for cubically nonlinear systems is conceptually straightforward. However, it is by no means trivial due to the great complexity associated with working in a 3-dimensional space. For determining the Volterra transfer functions of cubically nonlinear systems, a higher-order-spectrum [13, 14] based method was developed [15]. This method is applicable to a broader class of nonlinear problems in the sense that it doesn't assume particular statistics for the input and thus allows a larger variety of input characteristics (Gaussian and non-Gaussian).

Although the method developed in [15] (referred to as the conventional method in the remainder of this paper) is quite effective in determining Volterra transfer functions of a cubically nonlinear system, in digital implementation it requires the system output signal to be sampled at a frequency which is at least 6 times the bandwidth of the system input (in contrast to twice the bandwidth for the linear system case) [15]. Using such a high sampling frequency has two disadvantages. First, 3-times faster A/D converters and processing circuits are required. This suggests that expensive circuitry may be required to implement these methods. Second, for the Volterra transfer functions to achieve a required frequency resolution, a 3-times larger number of DFT (discrete Fourier transform) points is required. This increases the computational complexity of these methods since the required computation increases with the number of DFT points.

It is the objective of this paper to develop a practical technique for identification of cubically nonlinear systems with a random input using the input/output data sampled at the Nyquist sampling rate [16] of the input. Note that this sampling rate is the minimum one with-

0-8186-8005-9/97 $10.00 © 1997 IEEE

out leading to aliasing in the input data samples and thus is called the minimum sampling rate in this paper. Although the input data samples are immune from aliasing at the minimum sampling rate, the output data samples, however, may be aliased due to the fact that the bandwidth of the output of a cubically nonlinear system can be as large as 3 times the bandwidth of the input [15]. We show that, under the minimum-sampling-rate condition, the conventional method fails to achieve good estimates of the Volterra transfer functions. To deal with this problem, we propose a new model which properly accounts for the aliasing effects of the output. Based on this model, we develop a minimum-sampling-rate method. By properly accounting for the aliasing effects of the output, the proposed method is able to properly identify the Volterra transfer functions of a cubically nonlinear system by using the minimally sampled input/output data. The effectiveness of the proposed method is demonstrated by computer simulation.

2. The Conventional Method

In this section, we consider identification of cubically nonlinear systems in the discrete domain. This means the input/output signal is sampled for further processing. In taking samples of the signals, one needs to be aware of the fact that a nonlinear system may generate new output frequency components which are not in the input frequency band. For example, due to intermodulation and harmonic distortions, a cubically nonlinear system whose input has a bandwidth of W may generate an output with a bandwidth as large as $3W$ [15]. Therefore, according to the sampling theorem [16], the sampling frequency needs to be greater than or equal to $6W$ to avoid aliasing of the output data samples (in contrast to $2W$ for the linear system case). This is 3 times higher than the required sampling frequency for linear systems. Therefore, one needs to be careful in selecting the sampling frequency if the system under consideration is nonlinear.

Consider a cubically nonlinear system with an input $x(t)$ and an output $y(t)$, where $x(t)$ has a bandwidth of W. Assuming that $x(t)$ and $y(t)$ are sampled at the rate $f_s = 6W$ to prevent output data samples from aliasing, then the discrete frequency-domain third-order Volterra series representation of the system can be written as follows [15]:

$$Y(m) = Y_1(m) + Y_2(m) + Y_3(m) + \varepsilon(m), \quad -3M \le m \le 3M \tag{1}$$

$$Y_1(m) = \begin{cases} H_1(m)X(m) & |m| \le M \\ 0 & M+1 \le |m| \le 3M \end{cases} \tag{2}$$

$$Y_2(m) = \begin{cases} \sum_{\substack{-M \le i,j \le M \\ (i+j=m)}} H_2(i,j)X(i)X(j), \\ \quad |m| \le 2M \\ 0, \quad 2M+1 \le |m| \le 3M \end{cases} \tag{3}$$

$$Y_3(m) = \sum_{\substack{-M \le i,j,k \le M \\ (i+j+k=m)}} H_3(i,j,k)X(i)X(j)X(k), \quad |m| \le 3M \tag{4}$$

where $X(\cdot)$ and $Y(\cdot)$ are the discrete Fourier transforms (DFT's) of the input and output, and $Y_1(\cdot)$, $Y_2(\cdot)$, and $Y_3(\cdot)$ are the linear, quadratic, and cubic output terms of the system. The linear, quadratic, and cubic transfer functions of the system are denoted by $H_1(\cdot)$, $H_2(\cdot,\cdot)$, and $H_3(\cdot,\cdot,\cdot)$ respectively in (2)-(4). $\varepsilon(m)$ denotes the modeling error, and M is the maximum discrete frequency of $X(\cdot)$ with a non-zero value. If an N-point DFT is used, then for the sampling rate $f_s = 6W$ we have $M = N \times (W/f_s) = N/6$ (M is rounded to its nearest integer).

Since the system is excited by an input whose frequency range is from $-M$ to M, the estimation region of the quadratic transfer function (QTF) $H_2(i,j)$ in the 2-D (i,j) space is a square, as indicated by the square 'ABCD' in Fig. 1. Note that $Y_2(m)$ includes contributions from those $H_2(i,j)$'s which satisfy $i+j=m$. That is, all the points on the line segment $\overline{\mathrm{EF}}$ in Fig. 1 contribute to $Y_2(m)$. It can be shown that, by using the symmetry properties $H_2(i,j) = H_2(j,i)$ and $H_2^*(-i,-j) = H_2(i,j)$, the QTF in the square 'ABCD' can be uniquely determined as long as the QTF in the triangle 'ABO' is known [11]. Similarly, the estimation region for the cubic transfer function (CTF) $H_3(i,j,k)$ in the 3-D (i,j,k) space is a cube, as shown in Fig. 2. Since $Y_3(m)$ includes contributions from those $H_3(i,j,k)$'s which satisfy $i+j+k=m$, all the points on the hexagon 'ABCDEF' in Fig. 2 contribute to $Y_3(m)$. By using the symmetry properties $H_3(i,j,k) = H_3(j,i,k) = H_3(i,k,j)$ and $H_3^*(-i,-j,-k) = H_3(i,j,k)$, the CTF in the entire cubic region can be completely specified by the CTF in the pentahedron 'OIGHPN' shown in Fig. 3.

By taking into account the aforementioned symmetry properties of the Volterra kernels, we may rewrite (1) in a vector form as follows [15]:

$$Y(m) = \mathbf{X}^T(m)\mathbf{H}(m) + \varepsilon(m), \quad 0 \le m \le 3M \tag{5}$$

$$\mathbf{X}(m) = \begin{cases} [X(m)\ \mathbf{X}_2^T(m)\ \mathbf{X}_3^T(m)]^T, \\ \quad \text{if } 0 \le m \le M \\ [\mathbf{X}_2^T(m)\ \mathbf{X}_3^T(m)]^T, \\ \quad \text{if } M+1 \le m \le 2M \\ \mathbf{X}_3(m), \\ \quad \text{if } 2M+1 \le m \le 3M \end{cases} \tag{6}$$

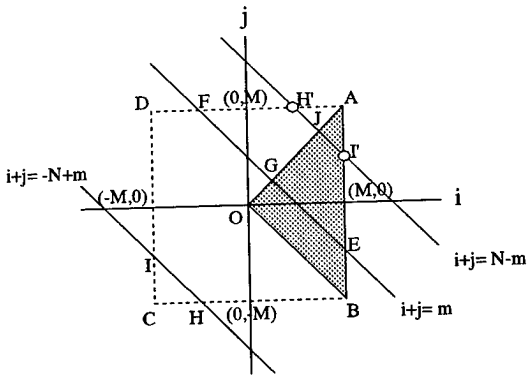


Figure 1. The estimation region for the quadratic transfer function.

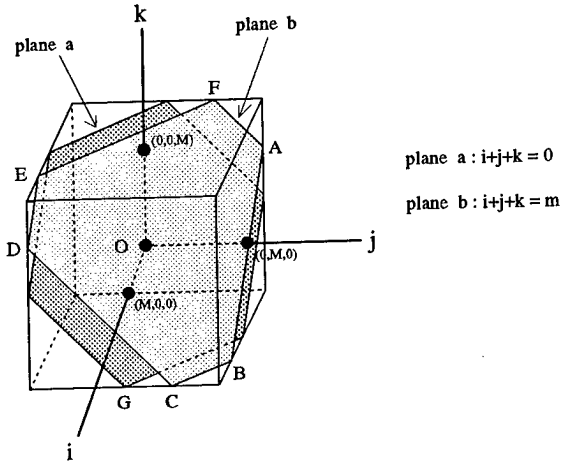


Figure 2. The estimation region for the cubic transfer function before considering symmetry properties.

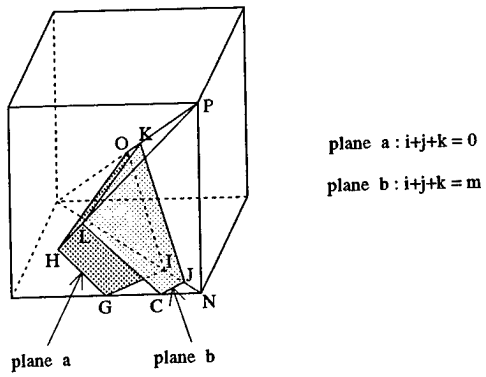


Figure 3. The estimation region for the cubic transfer function after considering symmetry properties.

$$\mathbf{H}(m) = \begin{cases} [H(m)\ \mathbf{H}_2^T(m)\ \mathbf{H}_3^T(m)]^T, \\ \quad \text{if } 0 \le m \le M \\ [\mathbf{H}_2^T(m)\ \mathbf{H}_3^T(m)]^T, \\ \quad \text{if } M+1 \le m \le 2M \\ \mathbf{H}_3(m), \\ \quad \text{if } 2M+1 \le m \le 3M \end{cases} \tag{7}$$

where the quadratic components $\mathbf{X}_2(m)$ and $\mathbf{H}_2(m)$ are defined by

$$\mathbf{X}_2(m) = \begin{bmatrix} S_2(M, m-M) \\ S_2(M-1, m-M+1) \\ \vdots \\ S_2(i_q(m), j_q(m)) \end{bmatrix} \tag{8}$$

$$\mathbf{H}_2(m) = \begin{bmatrix} H_2(M, m-M) \\ H_2(M-1, m-M+1) \\ \vdots \\ H_2(i_q(m), j_q(m)) \end{bmatrix} \tag{9}$$

$$S_2(i,j) = I(i,j)X(i)X(j) \tag{10}$$

$$I(i,j) = \begin{cases} 1 & \text{if } i = j \\ 2 & \text{if } i \ne j \end{cases} \tag{11}$$

$$i_q(m) = \lfloor \frac{m+1}{2} \rfloor \tag{12}$$

$$j_q(m) = m - i_q(m) \tag{13}$$

and the cubic components $\mathbf{X}_3(m)$ and $\mathbf{H}_3(m)$ are defined by

$$\mathbf{X}_3(m) = \begin{bmatrix} S_3(M, j_{cb}(m), k_{cb}(m)) \\ S_3(M, j_{cb}(m)-1, k_{cb}(m)+1) \\ \vdots \\ S_3(i_c(m), j_{ce}(m), k_{ce}(m)) \end{bmatrix} \tag{14}$$

$$\mathbf{H}_3(m) = \begin{bmatrix} H_3(M, j_{cb}(m), k_{cb}(m)) \\ H_3(M, j_{cb}(m)-1, k_{cb}(m)+1) \\ \vdots \\ H_3(i_c(m), j_{ce}(m), k_{ce}(m)) \end{bmatrix} \tag{15}$$

$$S_3(i,j,k) = I(i,j,k)X(i)X(j)X(k) \tag{16}$$

$$I(i,j,k) = \begin{cases} 6 & \text{if } i \ne j \ne k \\ 1 & \text{if } i = j = k \\ 3 & \text{otherwise} \end{cases} \tag{17}$$

$$i_c(m) = \lfloor \frac{m+2}{3} \rfloor \tag{18}$$

$$j_{cb}(m) = M - r[M-m] \tag{19}$$

$$k_{cb}(m) = m - M - j_{cb}(m) \tag{20}$$

$$j_{ce}(m) = \lfloor \frac{m - i_c(m) + 1}{2} \rfloor \tag{21}$$

$$k_{ce}(m) = m - i_c(m) - j_{ce}(m) \tag{22}$$

$$r[t] = \begin{cases} t & \text{if } t \ge 0 \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where $\lfloor v \rfloor$ denotes the largest integer being less than or equal to v. The reader is referred to [15] for the derivation of the above formulas.

By minimizing the mean square error $E[|\varepsilon(m)|^2]$ in (5), one obtains the optimal minimum mean square error estimate of $\mathbf{H}(m)$, say $\hat{\mathbf{H}}(m)$, as follows:

$$\hat{\mathbf{H}}(m) = E[\mathbf{X}^*(m)\mathbf{X}^T(m)]^{-1}E[\mathbf{X}^*(m)Y(m)] \quad (24)$$

Note that $E[\mathbf{X}^*(m)\mathbf{X}^T(m)]$ is composed of higher-order auto-moment spectra of $X(\cdot)$ up to 6th order, and $E[\mathbf{X}^*(m)Y(m)]$ is composed of higher-order cross-moment spectra between $X(\cdot)$ and $Y(\cdot)$ up to 4th order.

3. The Proposed Minimum-Sampling-Rate Method

In this section, we consider sampling the input/output signal of a cubically nonlinear system at the minimum sampling rate. That is, the sampling frequency f_s is set to $2W$, where W is the bandwidth of the system input. Since the output may have a bandwidth as wide as $3W$, the resulting output data samples are aliased. Under such a circumstances, the relation between $X(\cdot)$ and $Y(\cdot)$ described by (1) is no longer appropriate. In fact, the total quadratic output $Y_q(m)$ in $Y(m)$ can be written as follows [17]:

$$Y_q(m) = Y_2(m) + Y_2(-N+m), \quad 0 \le m \le M \quad (25)$$

Similarly, the total cubic output $Y_c(m)$ in $Y(m)$ can be written as follows:

$$Y_c(m) = Y_3(m) + Y_3(m+N) + Y_3(-N+m), \quad 0 \le m \le M \quad (26)$$

From the above analysis we see that an appropriate model for the minimum-sampling-rate case is as follows:

$$Y(m) = Y_l(m) + Y_q(m) + Y_c(m) + \varepsilon(m) \quad (27)$$

where $Y_l(m) = Y_1(m)$. The discrete model described by (27) can be rewritten in a vector form as follows:

$$Y(m) = \mathbf{H}^T(m)\mathbf{X}(m) + \varepsilon(m), \quad m = 0, \ldots, M \quad (28)$$

where

$$\mathbf{X}(m) = [X(m)\ \mathbf{X}_q^T(m)\ \mathbf{X}_c^T(m)]^T, \quad 0 \le m \le M \quad (29)$$

$$\mathbf{H}(m) = [H(m)\ \mathbf{H}_q^T(m)\ \mathbf{H}_c^T(m)]^T, \quad 0 \le m \le M \quad (30)$$

The model described by (28) is referred to as the minimum-sampling-rate (MSR) model in the remainder of this paper.

By comparing (28) to (5), we see similar vector forms. Therefore, the same MMSE technique used to solve (5) can be applied to (28). This method is referred to as the minimum-sampling-rate (MSR) method in the remainder of this paper.

4. Computer Simulation

To demonstrate the advantage of the proposed approach, we conducted a simulation using the following cubically nonlinear system:

$$y(n) = -0.64x(n)+x(n-2)+0.9x^2(n)+x^2(n-1) +0.6x(n)^3-0.3x(n-1)^3+e(n) \quad (31)$$

The input $x(n)$ and the additive noise $e(n)$ used in (31) are zero-mean white Gaussian random signals with a SNR of 40 dB.

We first collected 16,000 input/output data samples and divided them into 5,000 segments of 32 samples each. The 32-point DFT's of the data segments were calculated and then used by the conventional and the MSR methods to estimate the Volterra transfer functions. Note that since the input is white within the discrete frequency range (i.e., 0 to 16), the maximum discrete frequency of the input is equal to the discrete Nyquist frequency 16. This corresponds to the minimum-sampling-rate condition and thus the output data samples are aliased. For this case, the linear, quadratic, and cubic transfer function estimates obtained by the conventional method achieved a normalized mean square error (NMSE) of 0.0097, 0.0020, and 0.0356, respectively. On the other hand, the estimated linear, quadratic, and cubic transfer functions obtained by the MSR method achieved a NMSE of 0.0043, 0.0008, and 0.0065, respectively. These results indicate that the estimated Volterra transfer functions obtained by the MSR method are more accurate than the ones obtained by the conventional method under the minimum-sampling-rate condition.

5. Conclusion

In this paper, we have developed a practical method for cubically nonlinear system identification using input and output data sampled at the minimum sampling rate. The benefit of using the minimum sampling rate is that it significantly reduces the required processing speed of the circuits and the required computation. We showed that, under such a sampling rate, the proposed

method outperforms the conventional method. The advantage of the proposed method over the conventional one have been demonstrated by computer simulation.

Acknowledgment

This work was support in part by the National Science Council under Contract NSC 86-2213-E-019-007.

References

[1] D. R. Brillinger. Identification of polynomial systems by means of higher-order spectra. *J. Sound, Vibration.*

[2] S. Im and E. J. Powers. Extended principal domain for volterra models. In *Proc. IEEE Signal Processing ATHOS Workshop on Higher-Order Statistics*, pages 381–385, Begur, Girona, Spain, 1995.

[3] P. E. J. Kim, K.Ĩ. A digital method of modeling quadratically nonlinear systems with a general random input. *IEEE Trans. Acoust., Speech, Signal Processing*, Nov.

[4] T. Koh and E. J. Powers. Second-order volterra filtering and its application to nonlinear system identification. *IEEE Trans. Acoust., Speech, Signal Processing*, 33(6):1445–1455, Dec. 1985.

[5] S. W. Nam and E. J. Powers. Application of higher order spectral analysis to cubically nonlinear system identification. *IEEE Trans. Signal Processing*, 42(7), July 1994.

[6] C. L. Nikias and M. R. Raghuveer. Bispectrum estimation: a digital signal processing framework. *Proc. IEEE*, 75(7):869–891, July 1987.

[7] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1989.

[8] W. J. Rugh. *Nonlinear System Theory - The Volterra/Wiener Approach.* The Johns Hopkins Press Ltd., London, 1981.

[9] A. Samelis and D. Pavlidis. Mechanisms determining third order intermodulation distortion in algaas/gaas heterojunction bipolar transistors. *IEEE Trans. Microwave Theory and Techniques*, 40(12):2374–2380, Dec. 1992.

[10] I. W. Sandberg. The mathematical foundations of associated expansions for mildly nonlinear systems. *IEEE Trans. Circuits, Syst.*, 30(7):441–455, July 1983.

[11] I. W. Sandberg. On volterra expansions for time-varying nonlinear systems. *IEEE Trans. Circuits, Syst.*, 30(2), Feb 1983.

[12] I. W. Sandberg. Uniform approximation with doubly finite volterra series. *IEEE Trans. Signal Processing*, 40(6):1438–1442, June 1992.

[13] M. Schetzen. *The Volterra and Wiener Theories of Nonlinear System.* Wiley, New York, 1980.

[14] L. J. Tick. The estimation of transfer functions of quadratic systems. *Technometrics*, Nov.

[15] Y.-C. Tseng and D. A. Linebarger. Linear-quadratic system identification with completed frequency domain region of support. *IEEE Trans. Signal Process.*, 43(1):254–261, Jan. 1995.

[16] R. Tymerski. Volterra series modeling of power conversion systems. *IEEE Trans. Power Electronics*, 6(4):712–718, Oct. 1991.

[17] N. A. Zervos and R. L. Cupo. Adaptive nondisruptive measurement of harmonic distortion for voiceband data transmission. *IEEE Trans. Commun.*, 43(7):2184–2200, July 1995.

Higher-Order Statistics of Natural Images and their Exploitation by Operators Selective to Intrinsic Dimensionality

Gerhard Krieger, Christoph Zetzsche, Erhardt Barth
Ludwig-Maximilians-Universität München
Institut für Medizinische Psychologie
Goethestr. 31, D-80336 München, Germany
krieger@imp.med.uni-muenchen.de

Abstract

Natural images contain considerable statistical redundancies beyond the level of second-order correlations. To identify the nature of these higher-order dependencies, we analyze the bispectra and trispectra of natural images. Our investigations reveal substantial statistical dependencies between those frequency components which are aligned to each other with respect to orientation. We argue that operators which are selective to local intrinsic dimensionality can optimally exploit such redundancies. We also show that the polyspectral structure we find for natural images helps to understand the hitherto unexplained superiority of orientation-selective filter decompositions over isotropic schemes like the Laplacian pyramid. However, any essentially linear scheme can only partially exploit this higher-order redundancy. We therefore propose nonlinear i2D-selective operators which exhibit close resemblance to hypercomplex and end-stopped cells in the visual cortex. The function of these operators can be interpreted as a higher-order whitening of the input signal.

1. Introduction

The extraordinarily large amount of information contained in images of natural scenes calls for efficient data compression techniques. In order to exploit redundancies, most of today's image encoding schemes rely primarily on second-order statistics in conjunction with linear systems theory. Examples are predictive, sub-band (wavelet), and transform coding (e.g. KLT). However, even such common image structures like locally oriented lines and edges cannot be represented by second-order statistics (for a detailed argumentation see [10]). Vector quantization could theoretically overcome these limitations but suffers from combinatorial explosion.

An alternative approach can be derived from the concept of *intrinsic dimensionality* which relates the degrees of freedom provided by a signal domain to the degress of freedom actually used by a given signal [8]. This concept provides a hierarchy of local image signals in terms of different degrees of redundancy:

i0D-signals are constant, i. e. $u(x,y) = const.$ within a local window.

i1D-signals can locally be approximated by a function of only one variable, i. e. $u(x,y) = u(ax + by)$. Examples are straight lines and edges. Sinusoidal gratings, the eigenfunctions of linear systems, are also a member of this class.

i2D-signals are neither *i0D* nor *i1D*. Examples are corners, junctions, curved lines and edges, etc.

The hierarchy of intrinsic dimensionality seems to be reflected in biological vision systems, which are well adapted to the statistical properties of the environment. Evolution caused not only the development of linear isotropic and orientation-selective neurons, but also of nonlinear *i2D*-selective cells known as 'hypercomplex' or 'end-stopped'. For example, in primates more than half of the neurons in area V2 of the visual cortex are selective to *i2D*-stimuli [5]. It has also been shown that the least redundant *i2D*-information is already sufficient for a reconstruction of the original image signal [1].

2. Limitations of the second-order approach

Our statistical investigations of natural images revealed that the hierarchy in terms of intrinsic dimensionality is reflected in different class probabilties [6] [7]. In this hierarchy *i0D*-signals are the most common, and hence most redundant ones. The corresponding type of redundancy can

0-8186-8005-9/97 $10.00 © 1997 IEEE

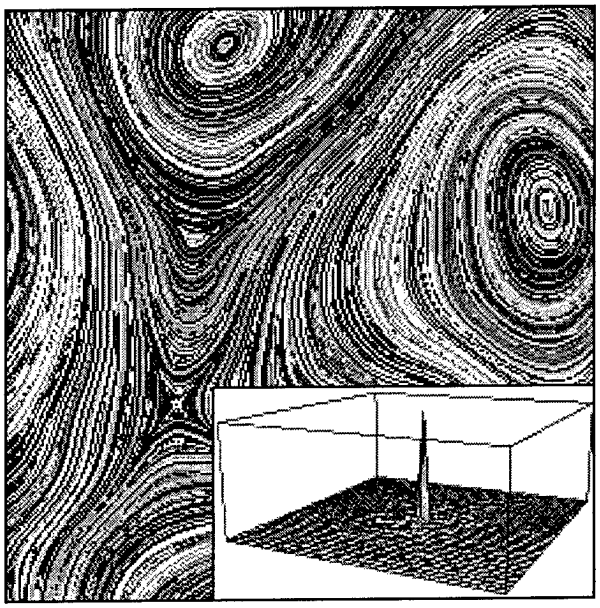


Figure 1. Sample image of an ergodic orientation-only process. This image completely lacks second-order correlations as can be seen from the autocorrelation function shown in the inlet (cf. [9]).

be related to second-order statistics, in form of the characteristic strong emphasis of low frequency components.

The next common class of signals are the *i1D*-signals, and any complete statistical description has to take them into account, too. Since the investigations of Oppenheim and Lim it is known, that the phase relations between the frequency components play an important role for the representation of the structural properties of images [4]. In this context the location of events such as lines or points has been identified as a property that cannot be represented by the Fourier amplitude, i. e. by second-order statistics, but requires the availability of Fourier phase. We have more recently extended this critique by suggesting that an further important deficit of the second-order approach has to be seen in its inability for the representation of *oriented structure* (*i1D*-signals) [10]. In order to illustrate this, we have constructed an orientation-only image that is full of oriented structure but lacks any second-order dependencies, i.e. it is 'white' (Fig. 1). Since second-order theory is insufficient to capture the oriented structure in this image, we have concluded that 'local orientedness' is a higher-order statistical property [10]. This point will be pursued further in the present paper.

3. Higher-order statistics of natural images

Surprisingly little is known about the higher-order statistics of natural images. In previous studies we noted substantial statistical dependencies remaining between the (second-order uncorrelated) outputs of orientation selective filter decompositions (wavelets) [6] [7] [10] [11]. Here we use higher-order spectra to investigate these dependencies in a more systematic way. The bispectrum $C_3^U(f_{x1}, f_{y1}, f_{x2}, f_{y2})$ is given by the Fourier transform of the third-order cumulant $c_3^u(x_1, y_1, x_2, y_2)$ of a stationary random process $\{u(x,y)\}$. Alternatively, the Fourier-Stieltjes representation of this process offers the possibility to express the bispectrum directly in terms of the components $dU(f_x, f_y)$ [3]:

$$E[dU(f_{x1}, f_{y1}) \cdot dU(f_{x2}, f_{y2}) \cdot dU^*(f_{x3}, f_{y3})] =$$

$$= \begin{cases} C_3^U(f_{x1}, f_{y1}, f_{x2}, f_{y2}) & \text{if } \binom{f_{x3}}{f_{y3}} = \binom{f_{x1}}{f_{y1}} + \binom{f_{x2}}{f_{y2}} \\ \quad \cdot df_{x1} df_{y1} df_{x2} df_{y2} & \\ 0 & \text{else} \end{cases}$$

From this equation it is apparent that the bispectrum is a measure for the statistical dependencies between three frequency components, the sum of which equals zero. A direct computation in the frequency domain can also be derived from this notation [3].

We have extensively investigated the bispectra of natural images and compared them to those of noise images with almost identical first and second-order statistics. Only for the natural images we found a concentration of bispectral 'energy' exactly in those regions where the frequency components are aligned to each other with respect to their orientation (an example is shown in Fig. 2). Systematic statistical dependencies of a similar type have also shown up in our preliminary investigations using trispectra (see Fig. 3). The trispectral 'energy' is concentrated around the aligned frequency triples for natural images as it is for samples from the orientation-only process, whereas it shows a substantially different distribution for random images which are equivalent to the natural ones in their 1st-order pdf and 2nd-order acf. This prompts us to conclude that the predominance of oriented *i1D*-structures in natural images is reflected in the concentration of polyspectral 'energy' of co-oriented frequency-tuples.

4. Exploitation of the higher-order statistics of natural images

4.1. Linear filter decompositions

If local 'orientedness' is a higher-order property not reflected in the standard second-order statistics, shouldn't nonlinear operators be required for its exploitation? This is only partially correct. While it is certainly true that an exhaustive exploitation of higher-order dependencies will, in

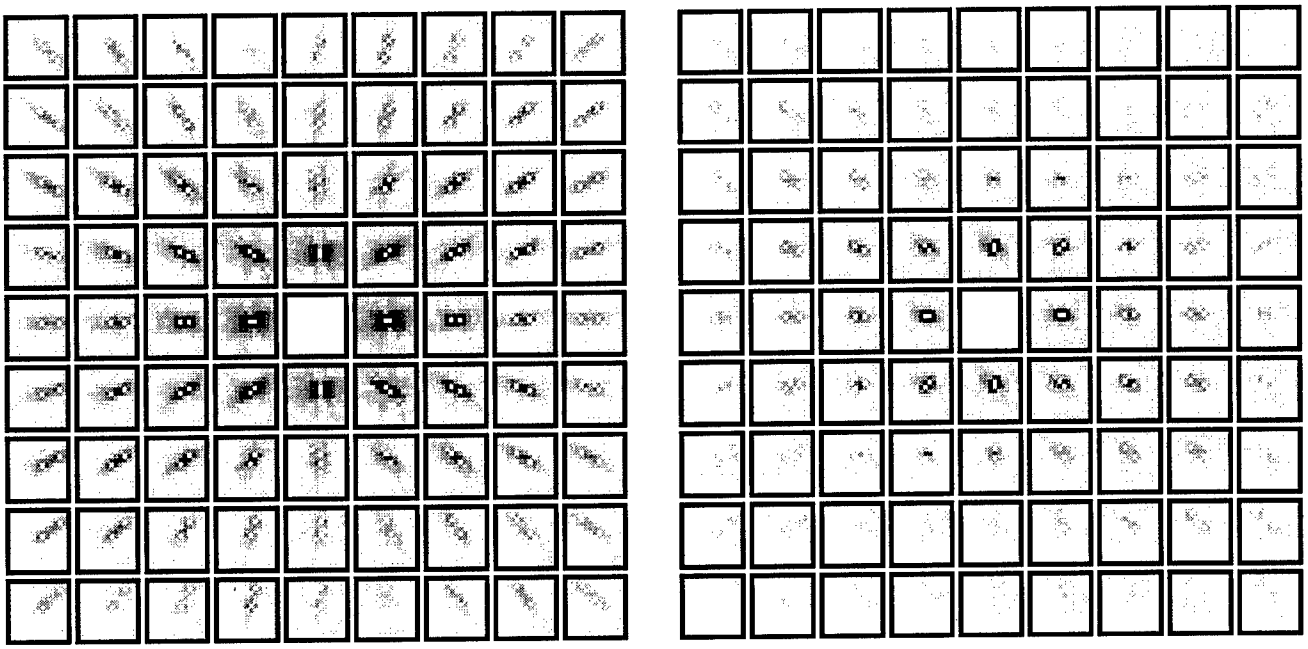


Figure 2. (a, left) Bispectral magnitude $|C_3(f_{x1}, f_{y1}, f_{x2}, f_{y2})|$ of a natural image. Shown are several sections with $f_{x1} = \text{const.}$ and $f_{y1} = \text{const.}$. Note the elongated concentration of bispectral 'energy' in regions where the frequency components are aligned to each other with respect to orientation. (b, right) Bispectrum of a noise image whose first-order pdf and second-order acf are approximately equivalent to that of the natural image. Here the 'energy'-concentration is more circular, around the void points $(f_{x1} + f_{x2} = 0, f_{y1} + f_{y2} = 0)$.

general, require nonlinear operations, a partial exploitation may well be achieved by suitable linear operators. In fact, the advantages of certain linear decompositions can only be properly conceived, once the higher-order dependencies are taken into account. We have suggested that this is the case with the encoding of images by orientation-selective decompositions like the Cosine transform or wavelet-like transforms [10].

In our view, the usual interpretation of the relationship between such transforms and the statistics of natural images requires a revision, although the current view appears to be consistent at a first glance: The KLT yields oriented basis functions, orientation obviously is an important structural property of natural images, hence usage of the optimally adapted, i.e. oriented basis functions for image compression will yield positive results. A closer look, however, reveals some problems with this interpretation.

In fact, strict application of the standard second-order reasoning would predict that the compression performance of an isotropic decomposition, like the Laplacian pyramid, should be close to the one achievable with a further orientation-selective splitting of the frequency bands [10]. This is due to the fact that no deviation from spectral flatness can be exploited along the orientation variable. However, practical experience and statistical investigations showed the contrary: substantial additional savings in data rate can be obtained if an image is not only decomposed by radial bandpass funtions but is also split up with respect to orientation , e.g. [12].

While this fact cannot be understood within the framework of second-order statistics, our results indicate that it might be explained by taking into account the structure of higher-order spectra since only orientation-selective filter decompositions can appropriately adapt to the concentration of higher-order spectral 'energy' along the '*i1D*-lines' as revealed by our measurements. An isotropic decomposition, like the Laplacian pyramid, will inevitably yield a mismatch to this specific type of structure.

4.2. Nonlinear operators

While an orientation-selective linear decomposition can partially exploit higher-order dependencies, nonlinear operators will be required in general. This raises the question of appropriate types of nonlinear operators. We will argue that *i2D*-selective operators are suitable candidates.

Starting from the Volterra-Wiener expansion of nonlinear functionals, a quite general class of *i2D*-selective operators can be derived. The Volterra series relates the input $u_1(x, y)$ of a nonlinear, shift invariant system to its output

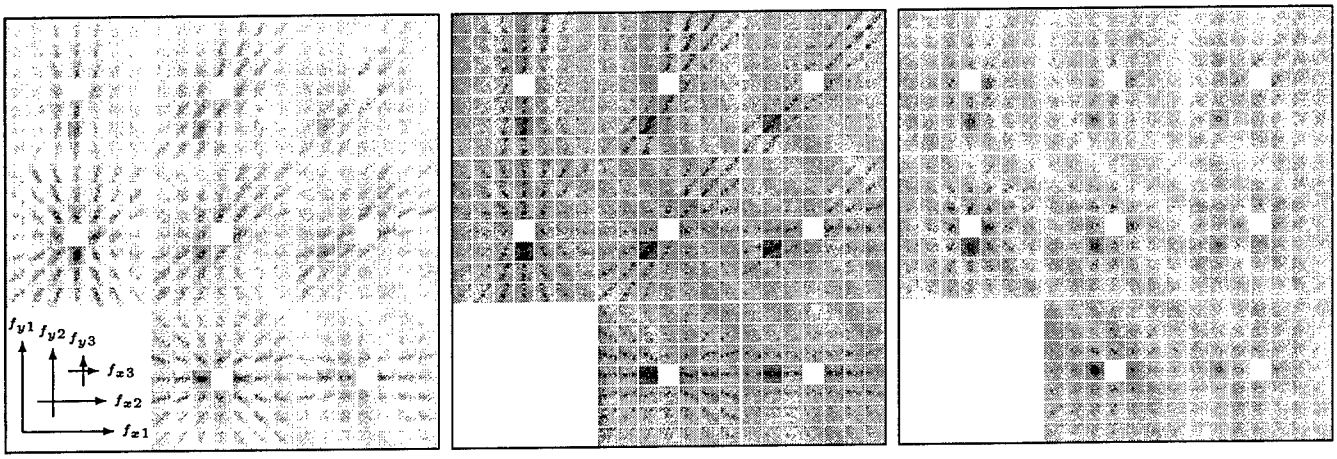


Figure 3. (a, left) Trispectral magnitude $|C_4(f_{x1}, f_{y1}, f_{x2}, f_{y2}, f_{x3}, f_{y3})|$ of a natural image. Note the elongated concentration of trispectral 'energy' in regions where the frequency components are aligned to each other. (b, middle) Trispectrum of the contour noise image given in Fig. 1. (c, right) Trispectrum of a noise image whose first-order pdf and second-order acf are approximately equivalent to that of the natural image. Here the 'energy'-concentration is more circular.

$u_2(x, y)$ in the following way:

$$u_2(x,y) = h_0 + \tag{1}$$
$$+ \iint h_1(x_1, y_1) \cdot u_1(x - x_1, y - y_1) \cdot dx_1 dy_1 +$$
$$+ \iiiint h_2(x_1, y_1, x_2, y_2) \cdot u_1(x - x_1, y - y_1) \cdot$$
$$\cdot u_1(x - x_2, y - y_2) \cdot dx_1 dy_1 dx_2 dy_2 + \ldots$$

The quadratic part of Eq. (2) may be expressed in the frequency domain as

$$U_2(f_x, f_y) = \iint \tilde{U}_2(f_{x1}, f_{y1}, f_x - f_{x1}, f_y - f_{y1})\, df_{x1} df_{y1}$$

where

$$\tilde{U}_2(f_{x1}, f_{y1}, f_{x2}, f_{y2}) = \tag{2}$$
$$= H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2}) \cdot U_1(f_{x1}, f_{y1}) \cdot U_1(f_{x2}, f_{y2})$$

is the expanded output spectrum and $H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2})$ is the Fourier transform of the second-order Volterra kernel $h_2(x_1, y_1, x_2, y_2)$. Note that Eq. (2) may be regarded as the weighting of an AND-like conjunction between frequency components. A necessary and sufficient condition for a quadratic Volterra operator to be insensitive to *i0D*- and *i1D*- signals is given by [2]:

$$H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2}) = 0 \quad \forall \quad f_{x1} \cdot f_{y2} = f_{y1} \cdot f_{x2}$$

The operation of systems adhering to this condition can be regarded as a blocking of aligned frequency components (see Fig. 4, an example is provided in Fig. 5). Since the redundancies of natural images are reflected in the concentration of polyspectral magnitude at aligned frequency components, *i2D*-selective systems have the potential to perform a kind of 'higher-order whitening', just as differentiating linear operators have the potential to whiten the $1/f^2$-decay of the power spectrum of natural images.

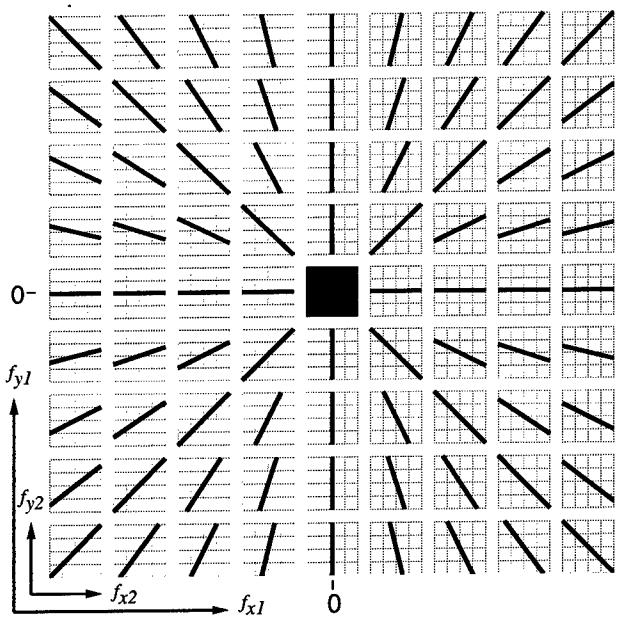


Figure 4. Illustration of the forbidden zones (indicated as black lines) for an *i2D*-selective quadratic Volterra kernel $H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2})$ in the frequency-domain. A system whose symmetric kernel vanishes at the forbidden zones may be regarded as blocking frequency components of equal orientation [2].

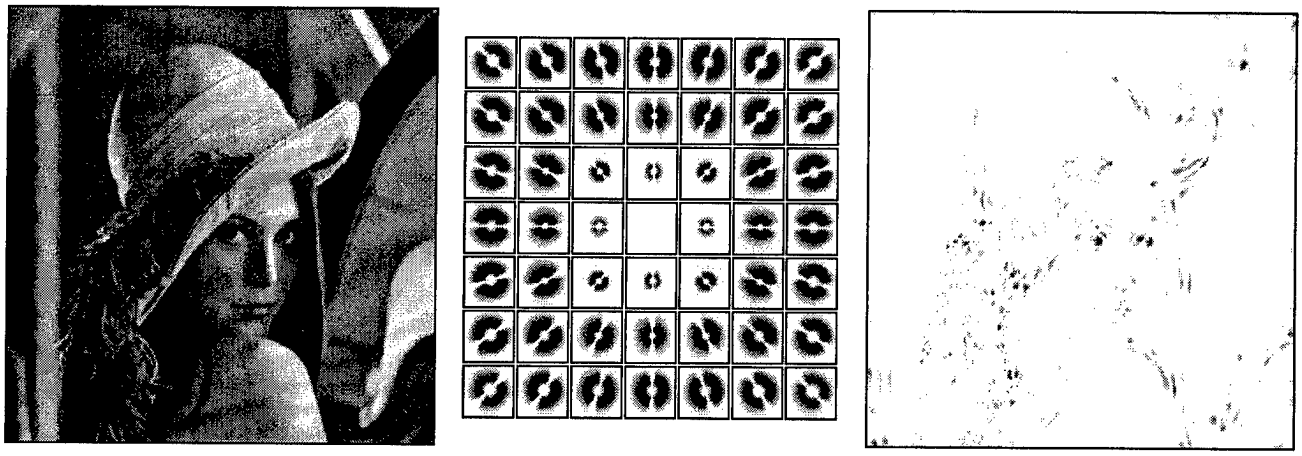


Figure 5. Application of *i2D*-selective operators to natural images: (a, left) Example of a natural image. (b, middle) Frequency-domain Volterra kernel of an *i2D*-selective operator. Note that the *i2D*-selectivity is guaranteed since the 'forbidden zones' of Fig. 4 are taken into account. (c, right) Image resulting from the application of the *i2D*-selective Volterra operator (only positive responses are shown). The original image can be reconstructed from such *i2D*-only representations if multiple scales are taken into account [1].

5. Conclusion

The frequent occurence of local oriented features is a basic structural property of natural images that cannot be captured by second-order statistics. Investigations of the polyspectra revealed that this property is reflected in the concentration of polyspectral magnitude in those regions where the frequency components are aligned in orientation. This fact can partially be exploited by orientation-selective linear filter decompositions. However, a full exploitation requires nonlinear schemes, for which *i2D*-selective operators are suitabe candidates.

Just like the distinction between a constant and a varying signal can be regarded as an elementary operation in linear signal processing, the distinction between *i1D* and *i2D*-signals can be regarded as a basic operation in the nonlinear processing of images. In this sense, *i2D*-operators can be seen as performing a 'nonlinear differentiation' beeing capable of exploiting the elementary redundancies of natural images.

References

[1] E. Barth, T. Caelli, and C. Zetzsche. Image encoding, labelling and reconstruction from differential geometry. *CVGIP:GRAPHICAL MODELS AND IMAGE PROCESSING*, 55(6):428–446, 1993.

[2] G. Krieger and C. Zetzsche. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *[Special Issue on Nonlinear Image Processing] - IEEE Trans. Image Processing*, 5(6):1026–1042, 1996.

[3] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework.* Prentice-Hall, 1993.

[4] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proc. IEEE*, 69:529–541, 1981.

[5] G. A. Orban. *Neuronal operations in the visual cortex.* Springer, Heidelberg, 1984.

[6] B. Wegmann and C. Zetzsche. Statistical dependence between orientation filter outputs used in a human vision based image code. In M. Kunt, editor, *Visual communication and image processing*, volume 1360 of *Proc. SPIE*, pages 909–923, Bellingham, WA, 1990.

[7] B. Wegmann and C. Zetzsche. Feature-specific vector quantization of images. *[Special Issue: Vector Quantization] - IEEE Trans. Image Processing*, 5:274–288, 1996.

[8] C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–1117, 1990.

[9] C. Zetzsche, E. Barth, G. Krieger, and B. Wegmann. Neural network models and the missing link between cortical orientation selectivity and the natural environment. *submitted*, 1997.

[10] C. Zetzsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. Watson, editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.

[11] C. Zetzsche, E. Barth, and B. Wegmann. Nonlinear aspects of primary vision: entropy reduction beyond decorrelation (Invited Address). In J. Morreale, editor, *SID International Symposium - Digest of Technical Papers*, volume XXIV, pages 933–936. Soc. Inform. Display, Playa del Ray, CA, 1993.

[12] C. Zetzsche and W. Schönecker. Orientation selective filters lead to entropy reduction in the processing of natural images. *Perception*, 16:229, 1987.

A Normalized Block LMS Algorithm for Frequency-Domain Volterra Filters *

Sungbin Im
Department of Information and Telecommunication Engineering
Soongsil University, Seoul, Korea 156-743
TEL: +82-2-820-0906, FAX: +82-2-814-3627, Email: sbi@interpia.net

Abstract

The objective of the paper is to introduce a new adaptive filtering algorithm for estimating frequency-domain second-order Volterra filter coefficients. The approach rests upon the normalized LMS (NLMS) algorithm and the frequency-domain block LMS algorithm. The utilization of the normalized LMS algorithm facilitates choice of a proper step size, with which the adaptive frequency-domain Volterra filter is guaranteed to be convergent in the mean-squared sense, and improves convergence rate. The frequency-domain block LMS algorithm estimates frequency-domain second-order Volterra filter coefficients which correspond to the DFT of the time-domain Volterra filter coefficients.

1. Introduction

In the last two decades, nonlinear digital filters, and in particular, digital Volterra filters [1], have become the subject of increasing interest. In many cases, the core of the problem is the identification of an unknown nonlinear system. In relation to the identification methodology, the computational simplicity of the well-know LMS algorithm has considerable appeal. However, it is well-known that an LMS-type Volterra filter suffers from slow convergence even for independent identically distributed Gaussian inputs[2]. This phenomenon is rooted from the fact that the quadratic component of the second-order Volterra filters takes the products of the linear inputs as its input sequence. Such correlated inputs result in a large eigenvalue spread of the corresponding correlation matrix, which leads to slow convergence. This holds for the time- and frequency-domain adaptive Volterra filters.

*This work was supported by GYONAE Research Project 97.

Several authors have demonstrated the utility of time- and frequency-domain Volterra filters updated using an LMS-type procedure for a range of applications [3, 4, 5], where different step sizes were applied to linear and quadratic filters in order to overcome the slow convergence problem. However, it is still difficult to choose the optimal step sizes for the linear and quadratic filters. In addition to the problems mentioned previously, the frequency-domain adaptive Volterra filters suffer from the wrap-around error because of the periodicity of the discrete frequency domain. This can be avoided by the block LMS algorithm [6] based on the overlap-save method.

The objective of the paper is to introduce a new adaptive filtering algorithm for estimating frequency-domain second-order Volterra filter coefficients, which correspond to the DFT of the time-domain Volterra filter coefficients. In order to avoid the wrap-around error in the frequency-domain processing, the frequency-domain Volterra filter utilizes the frequency domain BLMS algorithm [6]. According to [7], the conventional linear adaptive filters based on the NLMS algorithm obtain rapid convergence for highly correlated signals, by improving a large eigenvalue spread. With similar reasoning, our approach rests upon the NLMS algorithm in order to solve the slow convergence problem.

The remainder of this paper is organized as follows. The following section describes the second-order time- and frequency-domain Volterra filters, and presents the BLMS algorithm for the frequency-domain second-order Volterra filter. In Section 3, we develop the normalized block LMS algorithm for the frequency-domain Volterra filters. Section 4 contains a numerical example to demonstrate the validity of the proposed algorithm by comparing the time-domain LMS algorithm with multiple step sizes. Finally, the paper is concluded in Section 5.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Second-Order Volterra Filters

If we assume that the nonlinear system to be identified is stable and of finite memory, and has nonlinearities up to second order, the second-order Volterra filter can approximate the output of the nonlinear system by its sampled data form. The output can be represented as follows;

$$\begin{aligned} y_t(n) &= \sum_{i=0}^{N-1} h_1(i)x(n-i) \\ &+ \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} h_2(i,j)x(n-i)x(n-j) \\ &+ e(n) \end{aligned} \tag{1}$$

where $y_t(\cdot)$ and $x(\cdot)$ are observed responses and excitations, respectively. $h_1(\cdot)$ and $h_2(\cdot,\cdot)$ are linear and quadratic Volterra filter coefficients. When $y(n)$ is the output of the second-order Volterra filter,

$$e(n) = y_t(n) - y(n). \tag{2}$$

In the discrete frequency domain based on an M-point DFT, the second-order Volterra filter can be represented as follows[8];

$$\begin{aligned} Y_t(m) &= H_1(m)X(m) \\ &+ \sum_{p=0}^{M-1}\sum_{q=0}^{M-1} H_2(p,q)X(p)X(q)\delta_M(m-p-q) \\ &+ E(m) \end{aligned} \tag{3}$$

where

$$\delta_M(m) = \begin{cases} 1, & (m \text{ modulo } M) = 0 \\ 0, & (m \text{ modulo } M) \neq 0 \end{cases} \tag{4}$$

In (3), $Y_t(\cdot)$, $X(\cdot)$, $H_1(\cdot)$, $H_2(\cdot,\cdot)$ and $E(\cdot)$ are DFT's of $y_t(\cdot)$, $x(\cdot)$, $h_1(\cdot)$, $h_2(\cdot,\cdot)$, and $e(\cdot)$, respectively. Note that the frequency-domain Volterra filter (3) selects input frequency component pairs (p,q) using the delta function defined by modulo M.

Let us denote the input of the frequency domain Volterra filter (3) by the complex vector $\underline{X}(m) = [X_0, X_1, \ldots, X_M]^T$ where the superscript T denotes transpose, and the number of elements is $M+1$. Each X_k is determined by

$$X_k = \begin{cases} X(m), & k = 0 \\ X(p_k)X(q_k), & k = 1,\ldots,M \end{cases} \tag{5}$$

where (p_k, q_k) is selected by $\delta_M(m - p_k - q_k) = 1$. Denote the frequency-domain Volterra filter coefficient vector by $\underline{H}(m) = [H_0, H_1, \ldots, H_M]$, where

$$H_k = \begin{cases} H_1(m), & k = 0 \\ H_2(p_k, q_k), & k = 1,\ldots,M \end{cases} \tag{6}$$

Then, the output of the frequency-domain Volterra filter is given by

$$Y(m) = \underline{H}(m)\underline{X}(m) \tag{7}$$

The block update equations [6] are given by

$$\underline{H}^{\kappa+1}(m) = \underline{H}^{\kappa}(m) + \mu\underline{\nabla}^{\kappa}(m) \tag{8}$$

where κ represents the κ-th update time, μ is a step size. $\underline{\nabla}^{\kappa}(m)$ denotes an estimate of the gradient at time κ, and is given by

$$\begin{aligned} \underline{\nabla}^{\kappa}(m) &= [\hat{\nabla}_1^{\kappa}(m), \hat{\nabla}_2^{\kappa}(p_1,q_1), \ldots, \\ &\quad \hat{\nabla}_2^{\kappa}(p_k,q_k), \ldots, \hat{\nabla}_2^{\kappa}(p_M,q_M)], \end{aligned} \tag{9}$$

where (p_k, q_k) satisfies $\delta_M(m - p_k - q_k) = 1$, and the components $\hat{\nabla}_1^{\kappa}(m)$ and $\hat{\nabla}_2^{\kappa}(p_k,q_k)$ are computed by the procedure given in Appendix.

Recall that for a linear adaptive filter, its convergence is guaranteed only if the step size μ is less than $2/\lambda_{max}$ [9], where λ_{max} is the largest eigenvalue of the correlation matrix of the inputs. In addition, when the eigenvalues are widely spread, (that is, ratio of $\lambda_{max}/\lambda_{min}$ is large) the rate of convergence is limited by the smallest eigenvalue. Similar conditions hold for the Volterra filter [2]. As shown in (5), the quadratic input components consist of the products of the linear components, which results in high correlation among the input components.

3. A Normalized Block LMS Algorithm

In order to improve the convergence characteristics of the frequency-domain Volterra filter, we apply the normalized LMS algorithm[9] to (8), because for correlated input signals, the NLMS algorithm improves a eigenvalue spread, which influences the rate of convergence [7].

The normalized block update equation is given by

$$\underline{H}^{\kappa+1}(m) = \underline{H}^{\kappa}(m) + \mathbf{G}\{\alpha^{\kappa}(m)\underline{X}^{\kappa H}(m) * \underline{E}^{\kappa}(m)\} \tag{10}$$

where the superscript H denotes complex conjugate transpose and $*$ represents array multiplication. Note that the step size μ in (8) is replaced with

$$\alpha^{\kappa}(m) = \frac{\beta}{[\epsilon + \underline{X}^{\kappa H}(m)\underline{X}^{\kappa}(m)]}, \tag{11}$$

where β is a fixed scalar with $0 < \beta < 2$ and ϵ is a small positive constant that bounds $\alpha^{\kappa}(m)$ when $\underline{X}^{\kappa}(m)$ is momentarily small. $\underline{X}^{\kappa H}(m)\underline{X}^{\kappa}(m)$ is an estimate of the input signal power to the Volterra filter in the m-th frequency

bin at time κ. In (10), $\mathbf{G}$ stands for the operation of imposing a constraint on $\underline{X}^{\kappa H}(m) * \underline{E}^{\kappa}(m)$ in order to achieve a linear correlation. The error vector $\underline{E}^{\kappa}(m)$ is defined by

$$\underline{E}^{\kappa}(m) = [E_1^{\kappa}(m), E_2^{\kappa}(p_1, q_1), \ldots, E_2^{\kappa}(p_M, q_M)] \quad (12)$$

where each component is obtained through the procedure given in Appendix.

Table 1. Linear and Quadratic Filter Coefficients Used in the Volterra System

$h_1(i)$		$h_2(i,j)$				
i		(i,j)	0	1	2	3
0	0.0582	0	0.5181	-0.0624	-0.1609	-0.5075
1	-0.0175	1	-0.0624	0.2914	0.2408	-0.0601
2	-0.0785	2	-0.1609	0.2408	-0.6678	0.8759
3	0.2165	3	-0.5075	-0.0601	0.8759	0.2876

4. An Example

The objective of this example is to demonstrate the validity of the proposed algorithm and to compare its performance with that of the time-domain LMS algorithm [2]. In this example, the adaptive Volterra filter is used in the system modeling mode. The proposed adaptive Volterra filter has the equivalent order and number of coefficients in the frequency domain consistent with the time-domain Volterra system to be identified. The time-domain adaptive filter based on the LMS algorithm has the same order and number of coefficients.

In this example, we compare the performance of the proposed algorithm and the time-domain LMS algorithm in terms of coefficient error NMSEC and filter output error NMSEO. The coefficient error NMSEC is defined by

$$NMSEC(\kappa) = \frac{||\underline{h} - \underline{h}^{\kappa}||^2}{||\underline{h}||^2} \quad (13)$$

where $\underline{h}$ is the time-domain Volterra coefficient vector, whose elements are given in Table 1. For the proposed algorithm, $\underline{h}^{\kappa}$ is the inverse DFT of the estimate of the frequency-domain Volterra coefficients at time κ while $\underline{h}^{\kappa}$ is the estimate of the time-domain Volterra coefficients for the time-domain LMS algorithm. The filter output error is defined by

$$NMSEO(\kappa) = \frac{|||[e(\kappa N), \ldots, e(\kappa N + N - 1)]|||^2}{|||[y_t(\kappa N), \ldots, y_t(\kappa N + N - 1)]|||^2} \quad (14)$$

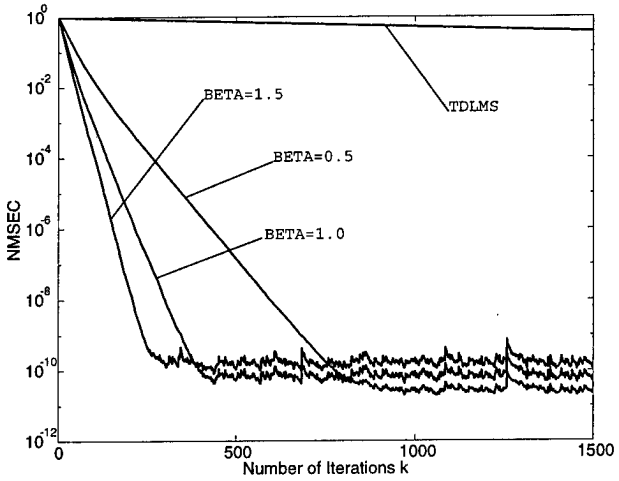


Figure 1. Coefficient error versus number of iterations.

where $[e(\kappa N), \ldots, e(\kappa N + N - 1)]$ and $[y_t(\kappa N), \ldots, y_t(\kappa N + N - 1)]$ are an output error vector at time κ and an observed output vector at time κ, respectively.

In this example, we consider the second-order Volterra system with the filter coefficients shown in Table 1, while the input sequence is chosen to be Gaussian-distributed. A zero-mean white Gaussian noise of SNR= 100 dB is added to the system output. For the proposed algorithm, β's are set to 0.5, 1, and 1.5. For the time-domain LMS algorithm, as proposed in [2], the step size of the linear filter is set to $1/\lambda_{max}^{L}$, where λ_{max}^{L} is the largest eigenvalue of the autocorrelation matrix of the linear input vectors and that of the quadratic filter is set to $1/\lambda_{max}^{Q}$, where λ_{max}^{Q} is the largest eigenvalue of the autocorrelation matrix of the quadratic input vectors. The results shown in the figures are ensemble averaged over 100 independent experiments. For the time-domain LMS algorithm, the mean value of the linear step size is about 0.6613×10^{-4} over 100 experiments, while that of the quadratic step size is about 0.0926×10^{-4}.

The coefficient error curves and the output error curves of the two algorithms are shown in Figures 1 and 2, respectively. In the figures, the curve labeled "TDLMS" is obtained by the time-domain LMS algorithm, while the remaining curves, labeled "BETA=0.5", "BETA=1.0", and "BETA=1.5", are obtained by the proposed algorithm with different β's of 0.5, 1.0, and 1.5. Note in Figure 1 that the coefficient estimates by the proposed algorithm converge to the true value with error of about 10^{-10} after about 900 iterations ($\beta = 0.5$) while the time-domain LMS algorithm seems to need more iterations. As β is reduced, the rate

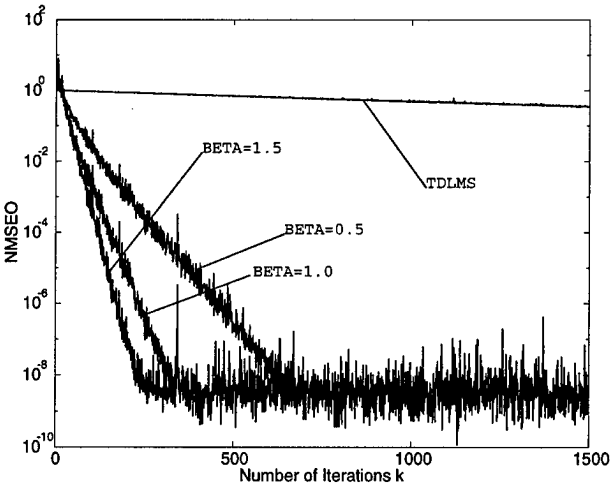


Figure 2. Output error versus number of iterations.

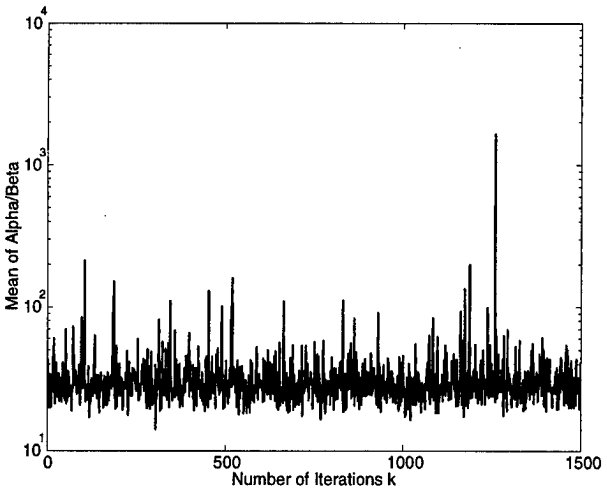


Figure 3. Mean of $\alpha^{\kappa}(m)/\beta$ over m's versus number of iterations.

of convergence of the proposed algorithm is correspondingly reduced. From Figure 2, we observe that a reduction in the value of β has the effect of reducing the fluctuation in the curve. Figure 3 shows the mean values of the $\frac{1}{[\epsilon+\underline{X}^{\kappa H}(m)\underline{X}^{\kappa}(m)]}$ with $\epsilon = 10^{-6}$ over 8 frequency bins.

The proposed algorithm limits the range of β. Thus, it is somewhat easier to choose a proper step size with which the proposed algorithm is convergent in the mean-squared sense, while the time-domain LMS algorithm requires the information on the largest eigenvalue of the correlation matrix of the input sequence for determining the optimal step size.

5. Summary

We presented a normalized block LMS algorithm for estimating frequency-domain second-order Volterra filter coefficients. The validity of the proposed algorithm was demonstrated though computer simulation. The use of the normalized LMS algorithm facilitates choice of a proper step size because the algorithm limits the range of β, and improves the convergence rate. With use of the block algorithm, the estimated frequency-domain coefficients converge to the DFT of the time-domain Volterra filter coefficients. Even though we did not demonstrate in this paper, the block LMS algorithm provides large savings in the computational complexity by replacing the one- and two-dimensional convolutions and correlations with the frequency-domain multiplications using the efficient FFT algorithm.

6. Appendix

In the followings, the bar denotes complex conjugate, and 1D DFT and 2D DFT represent M-point and (M, M)-point DFT's, respectively. It is also assumed that $M = 2N$, where N is a time-domain memory length and M is a block size for DFT.

6.1. Output Error Sequence and DFT

- 1D Error Sequence

$$\mathbf{e}_1^{\kappa} = [\underbrace{0,\ldots,0}_{N}, e(\kappa N),\ldots,e(\kappa N+N-1)]^T \quad (15)$$

where $e(n)$ is given by (2).

$$\mathbf{E}_1^{\kappa} = [E_1^{\kappa}(0),\ldots,E_1^{\kappa}(M-1)]^T = \text{1D DFT}\{\mathbf{e}_1^{\kappa}\}. \quad (16)$$

- 2D Error Sequence

$$E_2^{\kappa}(p,q) = \frac{1}{M}E_1^{\kappa}((p+q)_M), \quad (17)$$

where $(\cdot)_M$ represents the modulo-M operation.

6.2. Gradient Constraints

- For the linear component

$$\hat{\nabla}_1^{\kappa} = \text{1D DFT}\{[g_1^{\kappa}(0),\ldots,g_1^{\kappa}(N-1),0,\ldots,0]^T\}. \quad (18)$$

where

$$g_1^{\kappa}(i) = \text{1D IDFT}\{\bar{X}^{\kappa}(m)E^{\kappa}(m)\}, \quad i=0,\ldots,2N-1 \quad (19)$$

- For the quadratic component

$$\hat{\nabla}_2^{\kappa} = \text{2D DFT}\{[g_2^{\kappa}(i,j)\mathcal{W}_2(i,j)]\} \qquad (20)$$

where

$$\mathcal{W}_2(i,j) = \begin{cases} 1, & 0 \le i,\ j \le N-1 \\ 0, & \text{otherwise} \end{cases} \qquad (21)$$

and

$$[g_2^{\kappa}(i,j)] = \text{2D IDFT}\{[\bar{X}^{\kappa}(p)\bar{X}^{\kappa}(q)E_2^{\kappa}(p,q)]\} \qquad (22)$$

where i, j, p, and $q = 0, \ldots, M-1$.

References

[1] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley and Sons, New York, 1980.

[2] P.M. Clarkson and M.V. Dokic, "Stability and Convergence Behaviour of Second-Order LMS Volterra Filter," *Electronics Letters*, vol. 27, no. 5, pp. 441-443, February 1991.

[3] T. Koh and E.J. Powers,"Second-Order Volterra Filtering and its Application to Nonlinear System Identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1445-1455, 1985.

[4] D. Mansour and A. H. Gray, "Frequency Domain Non-Linear Adaptive Filter," *Proc. ICASSP 81*, pp. 550-553, 1981.

[5] A. Gutierrez and W. E. Ryan, "Performance of Adaptive Volterra Equalizers on Nonlinear Satellite Channels," *Proc. of ICC '95*, pp. 488-492, 1995.

[6] S. Im and E.J. Powers, "A Block LMS Algorithm for Third-Order Frequency-Domain Volterra Filters," *IEEE Signal Processing Letters*, vol. 4, no. 3, pp. 75-78, March 1997.

[7] M. Rupp, "The Behavior of LMS and NLMS Algorithms in the Presence of Spherically Invariant Processes," *IEEE Trans. Signal Processing*, vol. 41, no. 3, pp. 1149-1160, March 1993.

[8] S. Im and E.J. Powers, "A Fast Method of Discrete Third-Order Volterra Filtering," *IEEE Trans. on Signal Processing*, vol. 44, no. 9, pp. 2195-2208, September 1996.

[9] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, N.J., 1991.

Windows and Volterra Transfer Function Estimation

Hyungsuk Yoo and Edward J. Powers
Department of Electrical and Computer Engineering
and Electronics Research Center
The University of Texas at Austin
Austin, Texas 78712-1084, USA

Abstract

The effects of conventional data windows on Volterra transfer function estimation are investigated. The input/output data for two known second-order systems are utilized to estimate the transfer functions, and the results are compared with true values. In addition, the use of window correction factors to offset the bias introduced into the higher-order moment spectra, by the fact that the data is attenuated at the beginning and end of a record, is investigated. In all cases, it is found that the rectangular window results in the smallest NMSE (normalized mean square error) for the estimated quadratic transfer functions.

1. Introduction

Volterra transfer function estimation in the frequency domain using higher-order spectra has been studied for many years and applied to analysis of nonlinear phenomena. For nonGaussian inputs, second-order Volterra transfer function estimation in the frequency-domain utilizes estimates of the auto-spectra of the input up to fourth-order and cross-spectra of input and output up to third-order. So the quality of the transfer function estimation is directly related to that of polyspectral estimation.

It is well known in classical power spectral estimation[1] that if the signal is non-periodic, or periodic with the data length not equal to an integer multiple of periods, then spectral leakage exists. Thus we usually employ conventional data windows to reduce the spectral leakage. The research focus of bispectral estimation has been on the variance of estimation. In that sense many researchers have suggested using a data window in bispectral estimation in order to reduce the variance of estimation[2, 3]. K. Sasaki, et al.[4] derived the bias and variance bound for the indirect class of bispectral estimation. V. Chandran and S.L. Elgar[5] analyzed the bias and variance of the direct class of bispectral estimation including leakage effects.

When a data window is used, the beginning and end of the data record to be Fourier transformed are attenuated. Thus the mean square value, mean cube value, mean quartic value, etc. of the windowed data are reduced, and, hence, the corresponding higher-order moment spectra (power spectrum, bispectrum, trispectrum, etc.) will be underestimated, i.e., they will be biased. One way to correct (approximately) for this bias is to multiply the relevant polyspectrum by an appropriate window correction factor (w.c.f.).

Empirical evidence has suggested that the use of a rectangular window (which is equivalent to no data window) yields the "best" Volterra transfer function estimators. This "evidence" is not quantitative in that it is often based on physical insight into the nonlinear system/phenomena being modeled. For this reason, the results of a systematic study of the effects of windows and their corresponding w.c.f.'s on frequency domain Volterra transfer function estimation is presented in this paper.

In Sec. 2 we overview the frequency-domain second-order Volterra model, and the estimation of the Volterra transfer functions using an appropriate hierarchy of auto- and cross-spectra. In Sec. 3 we consider two different second-order systems characterized by known linear and quadratic transfer functions. The Volterra transfer functions are then estimated using a variety of data windows, with and without w.c.f.'s. Both Gaussian and exponentially distributed random inputs are utilized. A normalized mean square error (between the true and estimated quadratic Volterra transfer function coefficients) is used as a "goodness" criterion. In all cases examined, we find that use of a rectangular window (i.e., no window at all) yields the smallest normalized mean square error. The paper is summarized and concluded in Sec. 4.

0-8186-8005-9/97 $10.00 © 1997 IEEE

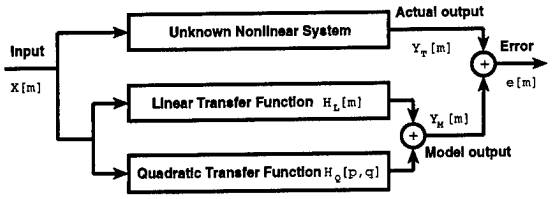


Figure 1. The conceptual block-diagram of second-order Volterra model in the discrete frequency domain

2 Frequency-domain Volterra model

The frequency-domain second-order Volterra model at discrete frequency m ($0 \leq m \leq N/2$) is as follows.

$$Y_M[m] = H_L[m]X[m] + \sum_{p,q} H_Q[p,q]X[p]X[q]\delta_N[p+q-m] \quad (1)$$

where $X[m]$ is the DFT of the input, $Y_M[m]$ is the DFT of the model output, $H_L[m]$ is the linear transfer function, $H_Q[p,q]$ is the quadratic transfer function.

A block diagram of second-order Volterra model in the frequency-domain is shown in Fig. 1. Since this model of (1) is linear in terms of the kernels H_L and H_Q, it can be represented in vector multiplication form

$$Y_M[m] = [H_L[m]\ \mathbf{H}_Q^T] \begin{bmatrix} X[m] \\ \mathbf{X}_Q \end{bmatrix} \quad (2)$$

where $\mathbf{H}_Q, \mathbf{X}_Q$ are vectors made up of $H_Q[p,q], X[p]X[q]$, and $p+q=m$. Minimization of the mean square error $E[e^2[m]]$ in Fig. 1 leads to the following normal equation,

$$\begin{bmatrix} E[Y_T[m]X^*[m]] & E[Y_T[m]\mathbf{X}_Q^*] \end{bmatrix} = [H_L[m]\ \mathbf{H}_Q^T] \times \begin{bmatrix} E[X[m]X^*[m]] & E[X[m]\mathbf{X}_Q^*] \\ E[\mathbf{X}_Q X^*[m]] & E[\mathbf{X}_Q\mathbf{X}_Q^*] \end{bmatrix} \quad (3)$$

where $E[\cdot]$ denotes expectation. So we need to estimate a hierarchy of higher-order auto-spectra of the input signal up to 4th order, and cross-spectra of the input and output signals up to 3rd order . Clearly, the goodness of the estimation of the transfer function is directly related to the goodness of the polyspectral estimators.

The use of data window attenuates the original signal and that results in bias of the polyspectral estimators. The factor $W_p = [1/N\sum_n w^2[n]]^{-1}$ is well known as the window correction factor (w.c.f.) for power spectral estimation. The generalized window correction factors[2] for higher-order spectral estimations are as follows, where W_b is the w.c.f. for bispectral estimation and W_t is the w.c.f. for 4th-order moment spectral estimation.

$$W_b = [\frac{1}{N}\sum_n w^3[n]]^{-1} \quad (4)$$

$$W_t = [\frac{1}{N}\sum_n w^4[n]]^{-1} \quad (5)$$

Multiplying the windowed polyspectral estimations by these factors will give bias-compensated polyspectral estimations, in the sense that the corresponding moments (mean square, mean cube, mean quartic values) will be conserved (approximately) between time and frequency domains.

3 Simulation results

The simulations are carried out for two known second-order nonlinear systems using the direct class of higher-order spectral estimation. The transfer function estimation method is valid for general input statistics, so we used exponentially distributed and normally distributed signals. And for comparison, we used various conventional windows[1] versus no window (or effectively a rectangular window). For the goodness criteria of the transfer function estimators, we use the normalized mean square error (NMSE) of the quadratic transfer function coefficients defined as follows.

$$NMSE(H_Q) = \frac{1}{K}\sum_{p,q}\frac{|H_{Qt}(p,q) - H_{Qe}(p,q)|^2}{|H_{Qt}(p,q)|^2} \quad (6)$$

Here $H_{Qt}(\cdot)$ and $H_{Qe}(\cdot)$ are the true value and estimated value of $H_Q(\cdot)$, respectively. The index pair (p,q) ranges over all relevant values of discrete frequency pairs in the two-dimensional frequency plane where H_Q is defined, and K is the total number of relevant pairs (p,q). Because of space limitations we focus in this paper on the relationship between data windows and quadratic transfer function estimation.

3.1. System 1

The first second-order nonlinear system model[6] in the time domain has the following representation.

$$y[n] = -0.64x[n] + x[n-2] + 0.9x^2[n] + x^2[n-1] \quad (7)$$

The frequency domain transfer functions are as follows.

$$H_L[m] = -0.64 + e^{-j4\pi m/N} \quad (8)$$

$$H_Q[p,q] = 0.9 + e^{-j2\pi(p+q)/N} \quad (9)$$

	Exponential distribution		Gaussian distribution	
	with w.c.f.	w/o w.c.f.	with w.c.f.	w/o w.c.f.
Rectangular	0.0244	0.0244	0.0144	0.0144
Hamming	0.0673	0.0976	0.0167	0.0418
Triangular	0.0651	0.1354	0.0185	0.0849
Hanning	0.1102	0.1518	0.0222	0.0474

Table 1. Normalized mean square error of the quadratic transfer function estimates of the system 1

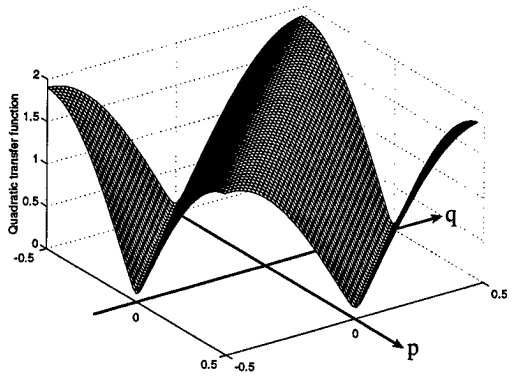


Figure 2. True quadratic transfer function of the system 1.

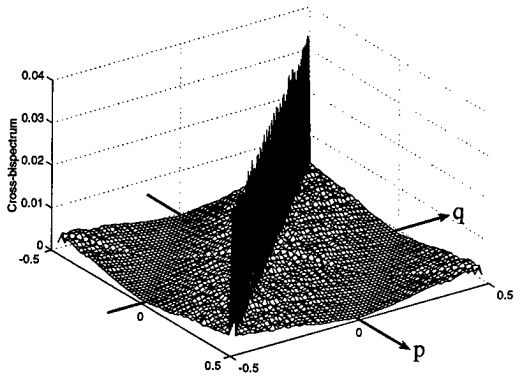


Figure 3. Cross bispectrum S_{yxx} of system 1 for exponentially distributed input.

Here N represents the FFT length, and the frequency indices m,p,q range from $-N/2$ to $N/2-1$. The frequency axes in the figures are normalized to the FFT length N.

The quadratic transfer function shown in Fig. 2 is strong in low frequency region and weak in high frequency region. Note the values of the quadratic transfer function do not exceed 2 (actually the maximum is 1.9 from Eq. 9). We are going to compare the estimated transfer functions with this figure.

Fig. 3 shows the cross-bispectral estimate $\hat{S}_{yxx}$ of input and output of the system 1 with a rectangular window and exponential input. Note the strong peak line along $-45°$ which has the following representation:

$$\begin{aligned} S_{yxx}[p,-p] &= E[Y[0]\cdot X^*[p]\cdot X^*[-p]] \\ &= E[Y[0]\cdot X^*[p]\cdot X[p]] \\ &= E[Y[0]\cdot |X[p]|^2] \end{aligned} \quad (10)$$

Due to the strong low frequency quadratic transfer function, the output dc value is large. Furthermore, the inputs used are white (i.e., not band limited). Thus the resulting $S_{yxx}[p,-p]$, which is roughly the input signal power times the output dc value, becomes large. If we employ a conventional data window, there corresponds a widening of the main-lobe. Since the peak line is relatively sharp compared to the neighboring strip, the widening of main-lobe can contaminate the transfer function estimation.

Fig. 4 and Fig. 5 show the quadratic transfer function estimate for the exponential input case with a rectangular window and a Hamming window with w.c.f., respectively. Fig. 6 and Fig. 7 show the quadratic transfer function estimate for the Gaussian input case with rectangular window and Hamming window with w.c.f., respectively. Note that the estimates with rectangular window are smoother than those with Hamming window. Also note that the estimates with Hamming window have a strong discontinuity along a strip centered along the $-45°$ line, which is mainly due to the widening of main-lobe.

The normalized mean square error (NMSE) of the quadratic transfer function estimates for both exponential and normal (Gaussian) distribution inputs over various data windows are presented in Table 1. The table also shows how the NMSE varies with whether we employ a window correction factor (w.c.f.) or not. Note that the w.c.f. for the rectangular window is 1, so the NMSE for a rectangular window with w.c.f. and without w.c.f. are the same. We can observe that in all cases the window correction factor reduces the NMSE, which indicates that the w.c.f. can compensate the bias of polyspectral estimation. Furthermore, note that the rectangular window gives the smallest normalized mean square error among the conventional data windows.

Let us compare the NMSE of a rectangular window and Hamming window with w.c.f. The NMSE of Hamming

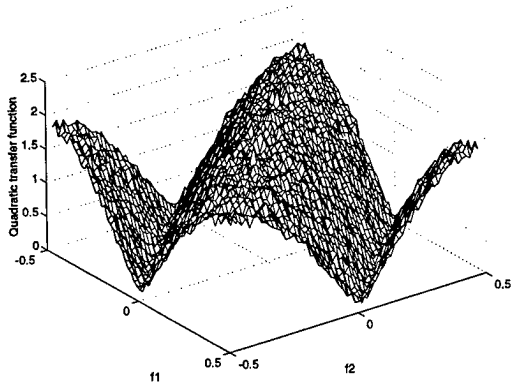


Figure 4. Quadratic transfer function estimate of the first system using Rectangular window. The input signal is exponentially distributed.

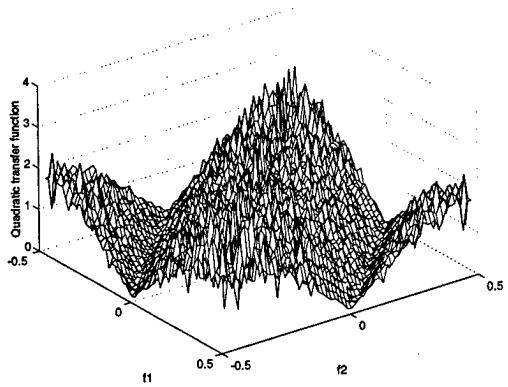


Figure 5. Quadratic transfer function estimate of the first system using Hamming window and appropriate w.c.f. The input signal is exponentially distributed.

window with w.c.f. for exponentially distributed input is slightly less than 3 times that of rectangular window for the same input. The difference is visible when comparing Fig. 4 and Fig. 5. Note the scale difference in z-axis of two figures. The estimated quadratic transfer function values using rectangular window in Fig. 4 do not exceed 2.5, whereas those values using Hamming window with w.c.f. have a maximum between 3 and 4 which is quite large compared to the true maximum value 1.9 (Eq. 9) in Fig. 2. Next consider the Gaussian input case. Although the NMSE of Hamming window with w.c.f. is almost the same as that of the rectangular window, we observe, when comparing Fig. 6 and Fig. 7, that their appearances are quite different. Most discrepancies are concentrated around the -45 degree line. So in this case the NMSE difference is not adequate to show the difference in estimation quality.

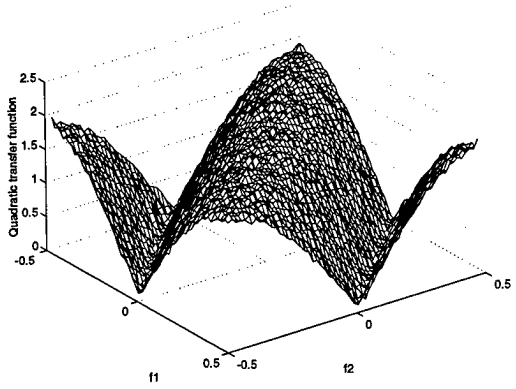


Figure 6. Quadratic transfer function estimate of the first system using Rectangular window. The input signal is normally distributed.

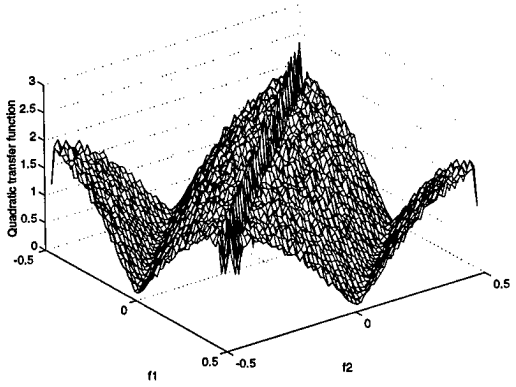


Figure 7. Quadratic transfer function estimate of the first system using Hamming window and appropriate w.c.f. The input signal is normally distributed.

3.2. System 2

The next second-order nonlinear system model in the time domain has the following representation.

$$y[n] = -0.7x[n] - 0.9x[n-2] + 0.8x^2[n] - 0.5x^2[n-2] \quad (11)$$

The frequency domain transfer functions are as follows.

$$H_L[m] = -0.7 - 0.9e^{-j4\pi m/N} \quad (12)$$

$$H_Q[p,q] = 0.8 - 0.5e^{-j4\pi(p+q)/N} \quad (13)$$

The quadratic transfer function is shown in Fig. 8. We observe that this quadratic transfer function is strong in mid-frequency range and weak in the low and high frequency range.

The normalized mean square error (NMSE) of the quadratic transfer function estimates for both exponential and normal (Gaussian) distribution inputs over various data

	Exponential distribution		Gaussian distribution	
	with w.c.f.	w/o w.c.f.	with w.c.f.	w/o w.c.f.
Rectangular	0.0138	0.0138	0.0043	0.0043
Hamming	0.0646	0.0919	0.0150	0.0391
Triangular	0.0636	0.1296	0.0172	0.0822
Hanning	0.1001	0.1363	0.0190	0.0482

Table 2. Normalized mean square error of the quadratic transfer function estimates of the system 2.

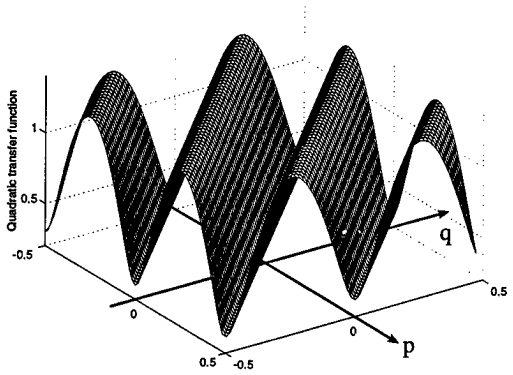


Figure 8. True quadratic transfer function of the system 2.

windows are presented in Table 2. The table also shows how the NMSE depends on whether we employ window correction factor (w.c.f.) or not. We can observe that the window correction factor reduces the NMSE, and that the rectangular window gives the smallest normalized mean square error among the conventional data windows, with and without w.c.f.'s.

4. Conclusions

In this paper, it was shown that the use of a rectangular window (v.s. tapered windows) yields the best estimates of Volterra quadratic transfer functions. Since this result is based on two simulation experiments, it obviously is not general. Nevertheless, it is consistent with empirical observations of those modeling nonlinear physical systems. Space has not permitted a discussion of linear transfer function results. Suffice it to say that both rectangular and conventional windows give comparable results. Also the w.c.f.'s cancel out in linear transfer function estimation. However, this is not the case for quadratic transfer function estimation. Space limitations have also precluded a discussion of the role of the extended principal domain[7] in the quadratic transfer function estimation. These issues will be discussed in a subsequent full length paper.

Acknowledgment

This work is supported in part by the Joint Service Electronics Program AFOSR Contract F-49620-95-C-0045, Department of Navy Grant N00014-92-J-1046, and NSF Grant CDR-8721512 through the Offshore Technology Research Center.

References

[1] F.J. Harris. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. *Proc. IEEE*, 66(1):51–83, January 1978.

[2] P.J. Huber, et. al. Statistical Methods for Investigating Phase Relations in Stationary Stochastic Processes. *IEEE Transactions on Audio and Electroacoustics*, AU-19(1):78–86, 1971.

[3] D.R. Brillinger and M. Rosenblatt. Asymptotic Theory of estimates of kth Order Spectra. *Spectral Analysis of Time Series*, B. Harris, ed., 135–188, Wiley, New York.

[4] K. Sasaki, T. Sato, and Y. Yamashita. Minimum Bias Windows for Bispectral Estimation. *J. Sound & Vibr.*, 40(1):139–148, 1975.

[5] V. Chandran and S.L. Elgar. Mean and Variance Estimates of the Bispectrum of a Harmonic Random Process - An Analysis Including Leakage Effects. *IEEE Trans. Signal Processing*, 39(12):2640–2651, December 1991.

[6] K.I. Kim and E.J. Powers. A Digital Method of Modeling Quadratically Nonlinear Systems with a General Random Input. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 36(11):1758–1769, November 1988.

[7] S. Im and E.J. Powers. Extended Principal Domain for Volterra Models. *Proceedings of the IEEE Signal Processing ATHOS Workshop on Higher-Order Statistics*, 381–385, Begur, Spain, June 12-14, 1995.

PERFORMANCE ANALYSIS OF VOLTERRA KERNEL ESTIMATORS WITH GAUSSIAN INPUTS

Arthur J. Redfern *G. Tong Zhou*

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250, USA.

ABSTRACT

The focus of this paper is on Volterra nonlinear system identification from input-output data. When the system is linear-quadratic and the input is Gaussian, closed-form expressions for the kernels were derived by Tick based on input-output cross-cumulants. However, there have been no known variance expressions for the kernel estimates. In this paper, we analyze the performance of the first- and second-order kernel estimates when the input is zero-mean white Gaussian, and the additive noise has unknown color and distribution. Closed-form variance expressions are presented and verified by simulations.

1. INTRODUCTION

For physical systems exhibiting mild nonlinearities, Volterra series modeling has been demonstrated to be a viable approach (see e.g., [1], [3], [6]). A discrete-time Volterra series relates the input $w(n)$ and the output $x(n)$ via

$$x(n) = \sum_{k=1}^{\infty} \sum_{u_1} \cdots \sum_{u_k} h_k(u_1, \ldots, u_k) \times \prod_{m=1}^{k} w(n-u_m), \quad n = 0, \ldots, N-1, \quad (1)$$

where $h_k(u_1, \ldots, u_k)$ is called the k^{th}-order kernel. In a practical setting, this series is often truncated in kernel order and kernel length. Sandburg has shown that the "doubly finite" Volterra series can uniformly approximate a large class of nonlinear systems [10]. Without loss of generality, we assume that $h_k(u_1, \ldots, u_k)$ is symmetric in its arguments, because we can symmetrize the kernel if otherwise [11].

A common problem encountered in Volterra modeling is to estimate the kernels from input-output data and several algorithms have been developed [4], [5], [7], [8], [12]. Higher-order statistics (cross-correlations, cumulants, or polyspectra) are natural with nonlinearities and are essential tools in deriving performance analysis results. To the best of the authors' knowledge, there have been no known closed-form variance expressions for the kernel estimates – the primary reason of which seems to be the complexity of analysis.

In this paper, we focus on linear-quadratic systems and white Gaussian inputs. We first review Tick's method [12] of estimating the kernels from input-output data and then show their variance expressions. The latter not only allow us to quantify the relative contribution from various system, input, and noise parameters but also establish benchmark performance measures to be compared with alternative algorithms such as in [8], [13].

2. IDENTIFICATION OF LINEAR AND QUADRATIC KERNELS

A linear-quadratic system (i.e., a second-order Volterra system) is described by

$$x(n) = \sum_{u_1} h_1(u_1) w(n-u_1) + \sum_{u_1} \sum_{u_2} h_2(u_1, u_2) w(n-u_1) w(n-u_2) + v(n), \quad (2)$$

where $w(n)$ is the input, $x(n)$ is the output, and $v(n)$ is the additive noise. In this paper, we assume that $w(n)$ is zero-mean, i.i.d. Gaussian with variance σ_w^2, and $v(n)$ is zero-mean, with unknown color and distribution, and is independent of $w(n)$.

The cumulant is the main tool used in this paper to derive kernel and variance expressions. We refer the reader to Section 2.3 of [2] for the definition and properties of cumulants.

2.1. Identification of the First-Order Kernel

The first-order kernel coefficients are found by computing the cross-cumulant between $x(n)$ and $w(n)$ [12]. Using the multilinearity property of cumulants (see p.19 in [2]), we find,

$$\begin{aligned} c_{xw}(\tau_1) &\triangleq \text{cum}\{x(n), w(n-\tau_1)\} && (3)\\ &= \sum_{u_1} h_1(u_1) \text{cum}\{w(n-u_1), w(n-\tau_1)\} && (4)\\ &\quad + \sum_{u_1} \sum_{u_2} h_2(u_1, u_2) \\ &\quad \times \text{cum}\{w(n-u_1) w(n-u_2), w(n-\tau_1)\} && (5)\\ &\quad + \text{cum}\{v(n), w(n-\tau_1)\} && (6)\\ &= \sum_{u_1} h_1(u_1)\ \text{cum}\{w(n-u_1), w(n-\tau_1)\} && (7)\\ &= \sum_{u_1} h_1(u_1)\ \sigma_w^2\ \delta(u_1-\tau_1) && (8)\\ &= \sigma_w^2\ h_1(\tau_1). && (9) \end{aligned}$$

The cumulant in (5) is zero because $w(n)$ is zero-mean Gaussian, whereas the cumulant in (6) is zero because $w(n)$ and $v(n)$ are mutually independent and zero-mean. Therefore, $h_1(\tau_1)$ is expressed in terms of the input-output cross-cumulant as

$$h_1(\tau_1) = \frac{c_{xw}(\tau_1)}{\sigma_w^2}. \quad (10)$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

2.2. Identification of the Second-Order Kernel

The second-order kernel coefficients are found from the cross-cumulant between $x(n)$ and two copies of $w(n)$ [12]:

$$c_{xww}(\tau_1,\tau_2) \triangleq \text{cum}\{x(n), w(n-\tau_1), w(n-\tau_2)\}$$

$$= \sum_{u_1} h_1(u_1)\text{cum}\{w(n-u_1), w(n-\tau_1), w(n-\tau_2)\} \quad (11)$$

$$+ \sum_{u_1}\sum_{u_2} h_2(u_1,u_2)\ \text{cum}\{w(n-u_1)w(n-u_2), w(n-\tau_1), w(n-\tau_2)\} \quad (12)$$

$$+\text{cum}\{v(n), w(n-\tau_1), w(n-\tau_2)\} \quad (13)$$

$$= \sum_{u_1}\sum_{u_2} h_2(u_1,u_2)\text{cum}\{w(n-u_1)w(n-u_2), w(n-\tau_1), w(n-\tau_2)\} \quad (14)$$

$$= \sum_{u_1}\sum_{u_2} h_2(u_1,u_2)\left[\sigma_w^4\delta(u_1-\tau_1)\delta(u_2-\tau_2) + \sigma_w^4\delta(u_1-\tau_2)\delta(u_2-\tau_1)\right] \quad (15)$$

$$= \sigma_w^4\ h_2(\tau_1,\tau_2) + \sigma_w^4\ h_2(\tau_2,\tau_1) \quad (16)$$

$$= 2\sigma_w^4\ h_2(\tau_1,\tau_2). \quad (17)$$

The Leonov-Shiryaev formula [2, p. 10] is used to obtain (15) from (14). We have also used the kernel symmetry assumption; i.e., $h_2(\tau_1,\tau_2) = h_2(\tau_2,\tau_1)$, to simplify (16) to (17). Therefore, $h_2(\tau_1,\tau_2)$ is expressed in terms of the input-output cross-cumulant as follows:

$$h_2(\tau_1,\tau_2) = \frac{c_{xww}(\tau_1,\tau_2)}{2\sigma_w^4}. \quad (18)$$

3. PERFORMANCE ANALYSIS OF KERNEL ESTIMATES

In practice, we estimate the kernels from N samples of input-output data. It is important to know the performance of these estimates under various input and system conditions. Closed-form variance expressions will also enable us to infer the data length necessary to meet certain (low) kernel variance requirements. Due to space limitations, we present here only the final variance expressions of the kernel estimates; detailed derivations can be found in [9]. For tractability of analysis, we assume that σ_w^2 is known – this is the case when we have control of the input; otherwise, we estimate σ_w^2 from $N^{-1}\sum_{n=0}^{N-1} w^2(n)$.

3.1. The First-Order Kernel

The cross-cumulant used in calculating the first-order kernel is estimated from a finite amount of input-output data as

$$\hat{c}_{xw}(\tau_1) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)w(n-\tau_1). \quad (19)$$

Inserting (19) into (10), we find

$$\hat{h}_1(\tau_1) = \frac{\hat{c}_{xw}(\tau_1)}{\sigma_w^2} = \frac{1}{N\sigma_w^2}\sum_{n=0}^{N-1} x(n)w(n-\tau_1). \quad (20)$$

Now substitute (2) into (20) to obtain

$$\hat{h}_1(\tau_1) = \frac{1}{N\sigma_w^2}\sum_{n=0}^{N-1}\sum_{u_1} h_1(u_1)w(n-u_1)w(n-\tau_1) \quad (21)$$

$$+\frac{1}{N\sigma_w^2}\sum_{n=0}^{N-1}\sum_{u_1}\sum_{u_2} h_2(u_1,u_2) \times w(n-u_1)w(n-u_2)w(n-\tau_1) \quad (22)$$

$$+\frac{1}{N\sigma_w^2}\sum_{n=0}^{N-1} v(n)w(n-\tau_1). \quad (23)$$

The variance expression of the above linear kernel estimate can be shown to contain eight terms [9]:

$$\text{var}\{\hat{h}_1(\tau_1)\} = \frac{1}{N}\sum_{u_1} h_1^2(u_1) + \frac{1}{N}\sum_{u_1} h_1(u_1)h_1(2\tau_1-u_1) + \frac{2\sigma_w^2}{N}\sum_{u_1}\sum_{u_2} h_2^2(u_1,u_2) + \frac{\sigma_w^2}{N}\left[\sum_{u_1} h_2(u_1,u_1)\right]^2 + \frac{4\sigma_w^2}{N}\left[\sum_{u_1} h_2(u_1,\tau_1)\right]^2 + \frac{4\sigma_w^2}{N}\sum_{u_1}\sum_{u_2} h_2(u_1,u_1)h_2(u_2,\tau_1) + \frac{4\sigma_w^2}{N}\sum_{u_1}\sum_{u_2} h_2(u_1,\tau_1+u_1-u_2) \times h_2(u_2,\tau_1+u_2-u_1) + \frac{\sigma_v^2}{N\sigma_w^2}. \quad (24)$$

3.2. The Second-Order Kernel

The second-order cross-cumulant is estimated from

$$\hat{c}_{xww}(\tau_1,\tau_2) = \frac{1}{N}\sum_{n=0}^{N-1}(x(n)-\hat{m}_x)w(n-\tau_1)w(n-\tau_2), \quad (25)$$

where $\hat{m}_x = N^{-1}\sum_{n=0}^{N-1} x(n)$. To simplify analysis, we substitute $\hat{m}_x$ by

$$m_x = E[x(n)] = \sum_u h_2(u,u)\sigma_w^2. \quad (26)$$

The resulting variance will be slightly larger than the actual variance with $\hat{m}_x$ but our simulations show that the difference is small.

Substituting (25) into (18), we find

$$\hat{h}_2(\tau_1,\tau_2) = \frac{1}{2\sigma_w^4}\ \hat{c}_{xww}(\tau_1,\tau_2)$$

$$= \frac{1}{2N\sigma_w^4}\left[\sum_{n=0}^{N-1}\sum_u h_1(u) \times w(n-u)w(n-\tau_1)w(n-\tau_2)\right. \quad (27)$$

$$+\sum_{n=0}^{N-1}\sum_{u_1}\sum_{u_2} h_2(u_1,u_2) \times w(n-u_1)w(n-u_2)w(n-\tau_1)w(n-\tau_2) \quad (28)$$

$$+\sum_{n=0}^{N-1} v(n)w(n-\tau_1)w(n-\tau_2) \quad (29)$$

$$\left.-\sum_{n=0}^{N-1}\sum_u h_2(u,u)\sigma_w^2 w(n-\tau_1)w(n-\tau_2)\right] \quad (30)$$

The variance expression of the above second-order kernel estimate can be shown to contain twenty-eight terms [9]:

$$
\begin{aligned}
&\mathrm{var}\{\hat{h}_2(\tau_1,\tau_2)\} = \\
&\quad \frac{1}{4N\sigma_w^2}\sum_u h_1^2(u) + \frac{1}{2N\sigma_w^2}h_1^2(\tau_1) + \frac{1}{2N\sigma_w^2}h_1^2(\tau_2) \\
&\quad + \frac{1}{2N\sigma_w^2}h_1(\tau_1)h_1(\tau_2) \\
&\quad + \frac{1}{2N\sigma_w^2}h_1(2\tau_1-\tau_2)h_1(2\tau_2-\tau_1) \\
&\quad + \frac{1}{4N\sigma_w^2}\sum_u h_1^2(u)\delta(\tau_1-\tau_2) \\
&\quad + \frac{1}{4N\sigma_w^2}\left[\sum_u h_1(u)\right]^2\delta(\tau_1-\tau_2) \\
&\quad + \frac{1}{N\sigma_w^2}\sum_u h_1(u)h_1(\tau_1)\delta(\tau_1-\tau_2) \\
&\quad + \frac{1}{2N}\sum_{u_1}\sum_{u_2} h_2^2(u_1,u_2) + \frac{2}{N}\sum_{u_1} h_2^2(u_1,\tau_1) \\
&\quad + \frac{2}{N}\sum_{u_1} h_2^2(u_1,\tau_2) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_1)h_2(\tau_1,2\tau_2-u_1) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_2)h_2(2\tau_1-u_1,\tau_2) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_1)h_2(u_1+\tau_1-\tau_2,\tau_2) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_2)h_2(u_1+\tau_2-\tau_1,\tau_1) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_1)h_2(\tau_1+\tau_2-u_1,\tau_2) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,\tau_2)h_2(\tau_1+\tau_2-u_1,\tau_1) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,2\tau_1-\tau_2)h_2(u_1+\tau_2-\tau_1,2\tau_2-\tau_1) \\
&\quad + \frac{1}{N}\sum_{u_1} h_2(u_1,2\tau_2-\tau_1)h_2(u_1+\tau_1-\tau_2,2\tau_1-\tau_2) \\
&\quad + \frac{1}{2N}\sum_{u_1} h_2(u_1,u_1+\tau_2-\tau_1) \\
&\qquad \times h_2(2\tau_1-u_1,\tau_1+\tau_2-u_1) \\
&\quad + \frac{1}{2N}\sum_{u_1} h_2(u_1,u_1+\tau_1-\tau_2) \\
&\qquad \times h_2(2\tau_2-u_1,\tau_1+\tau_2-u_1) \\
&\quad + \frac{1}{2N}\sum_{u_1}\sum_{u_2} h_2(u_1,u_2)h_2(u_1+\tau_2-\tau_1,u_2+\tau_2-\tau_1) \\
&\quad + \frac{4}{N}\sum_{u_1}\sum_{u_2} h_2(\tau_1,u_2)h_2(u_1,u_1-u_2+\tau_1)\delta(\tau_1-\tau_2) \\
&\quad + \frac{1}{2N}\sum_{u_1}\sum_{u_2}\sum_{v_1} h_2(u_1,u_2)h_2(u_1-u_2+v_1,v_1) \\
&\qquad \times\delta(\tau_1-\tau_2) \\
&\quad + \frac{\sigma_v^2}{4N\sigma_w^4} + \frac{\sigma_v^2}{2N\sigma_w^4}\delta(\tau_1-\tau_2) + \frac{1}{4N}\left[\sum_u h_2(u,u)\right]^2 \\
&\quad - \frac{1}{4N}\left[\sum_{u_1} h_2(u_1,u_1)\right]^2\delta(\tau_1-\tau_2). \qquad (31)
\end{aligned}
$$

3.3. Observations

We make the following observations regarding the variance expressions (24) and (31).

1. Since $\mathrm{var}\{\hat{h}_1(\tau_1)\}$ and $\mathrm{var}\{\hat{h}_2(\tau_1,\tau_2)\}$ are $O(N^{-1})$, the kernel estimates are consistent.
2. We make no assumption on the color or distribution of the additive noise.
3. Increasing the additive noise variance σ_v^2 increases the variance of both the first and second-order kernels.
4. The variance of the input sequence σ_w^2 has significant contributions to the variance expressions. Interestingly, it affects the first and second-order kernels in an opposite fashion: the variance of $\hat{h}_1(\tau_1)$ tends to increase while the variance of $\hat{h}_2(\tau_1,\tau_2)$ tends to decrease with increasing σ_w^2.
5. The variance of $\hat{h}_2(\tau_1,\tau_1)$ contains additional terms as compared to that with $\tau_1 \neq \tau_2$. Therefore, diagonal kernel estimates tend to have larger variance.
6. Even for τ_1 such that $h_1(\tau_1) = 0$, we find $\mathrm{var}\{\hat{h}_1(\tau_1)\} \neq 0$ due to τ_1-independent terms such as $N^{-1}\sum_{u_1} h_1^2(u_1)$ in (24). Similarly, even for (τ_1,τ_2) with $h_2(\tau_1,\tau_2) = 0$, we will have $\mathrm{var}\{\hat{h}_2(\tau_1,\tau_2)\} \neq 0$ due to constant terms such as $(4N\sigma_w^2)^{-1}\sum_u h_1^2(u)$ in (31).

4. SIMULATIONS

The linear-quadratic system is described as in (2) with the following parameters: $h_1(0) = 0.8$, $h_1(1) = 1.2$, $h_1(2) = 0.4$, $h_2(0,0) = 1$, $h_2(0,1) = h_2(1,0) = 0.5$, and $h_2(1,1) = 0.3$. The input $w(n)$ is zero-mean, white Gaussian with variance $\sigma_w^2 = 1$. Additive noise $v(n)$ is zero-mean, white Gaussian noise with variance $\sigma_v^2 = 0$ or $\sigma_v^2 = 0.1$. We first formed $\hat{h}_1(\tau_1)$ and $\hat{h}_2(\tau_1,\tau_2)$ from N samples of input-output data and then obtained the empirical variances of $\hat{h}_1(\tau_1)$ and $\hat{h}_2(\tau_1,\tau_2)$ based on 500 independent realizations. In Figs. 1-4, we display the empirical (solid line) and theoretical (dotted line) $N\mathrm{var}\{\hat{h}_1(\tau_1)\}$ (Figs. 1 and 2) and $N\mathrm{var}\{\hat{h}_2(\tau_1,\tau_2)\}$ (Figs. 3 and 4) – they are seen to agree fairly well.

5. CONCLUSIONS

Analytical variance expressions of the kernel estimates are obtained for linear-quadratic systems with white Gaussian input. The estimates are shown to be consistent. The variance expressions describe explicit dependence on the data length, input variance, additive noise variance, and true kernel values. Our simulation results agree well with those predicted by the theory.

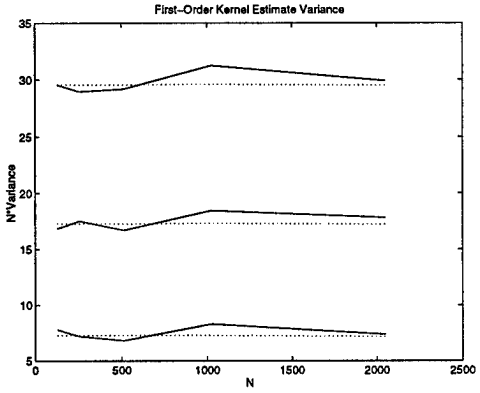


Figure 1. First order kernel estimates, $\sigma_v^2 = 0$.

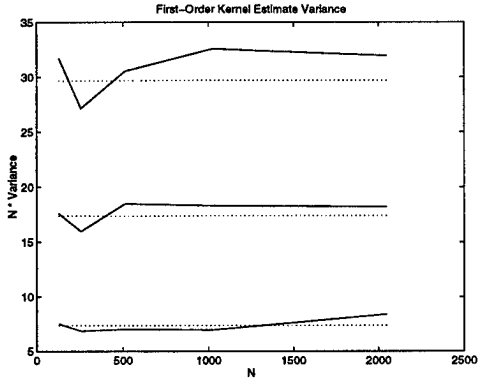


Figure 2. First order kernel estimates, $\sigma_v^2 = 0.1$.

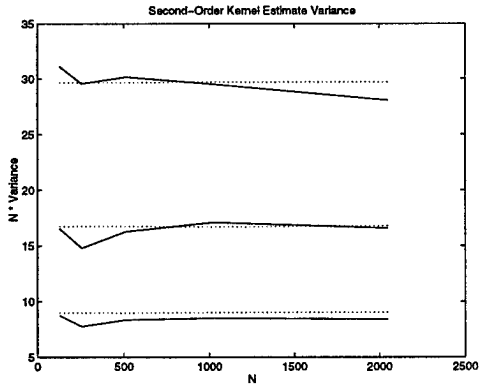


Figure 3. Second order kernel estimates, $\sigma_v^2 = 0$.

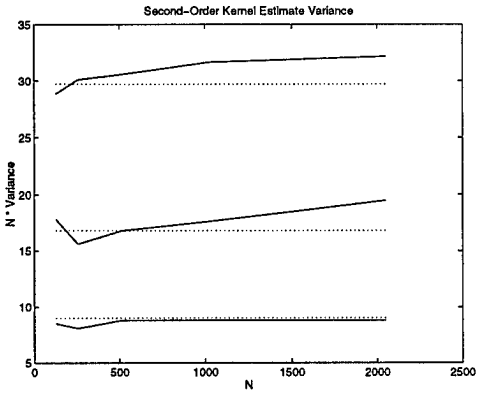


Figure 4. Second order kernel estimates, $\sigma_v^2 = 0.1$.

REFERENCES

[1] S. Benedetto, E. Biglieri and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links – a Volterra series approach," *IEEE Trans. on Aerospace and Electronic Systems,* vol. 15, no. 4, pp. 494-507, 1979.

[2] D.R. Brillinger, *Time Series: Data Analysis & Theory.* Holden-Day, 1981.

[3] R. Hermann, "Volterra modeling of digital magnetic saturation recording channels," *IEEE Trans. on Magnetics,* vol. 26, no. 5, pp. 2125-2127, 1990.

[4] P. Koukoulas and N. Kalouptsidis, "Nonlinear system identification using Gaussian inputs," *IEEE Trans. on Signal Processing,* vol. 43, no, 8, pp. 1831-1841, 1995.

[5] Y.W. Lee and M. Schetzen, "Measurement of the Wiener kernels of a non-linear system by cross-correlation," *International Journal of Control,* vol. 2, pp. 237-254, 1965.

[6] V.Z. Marmarelis, "Coherence and apparent transfer function measurements for nonlinear psychological systems," *Annals of Biomedical Engineering,* vol. 16, pp. 143-157, 1988.

[7] R.D. Nowak and B.D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification," *IEEE Trans. on Signal Processing,* vol. 42, no, 8, pp. 2124-2135, 1994.

[8] Y.S. Cho and E.J. Powers, "Quadratic system identification using higher order spectra of I.I.D. signals," *IEEE Trans. on Signal Processing,* vol. 42, no. 5, pp. 1268-1271, May 1994.

[9] A.J. Redfern and G.T. Zhou, "Performance analysis of linear and quadratic Volterra kernel estimates with Gaussian inputs," *Technical Report,* School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, 1997 (available upon request).

[10] I.W. Sandberg, "Uniform approximation with doubly finite Volterra series," *IEEE Trans. on Signal Processing,* vol. 40, no. 6, pp. 1438-1442, 1992.

[11] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems.* New York: Wiley, 1980.

[12] L.J. Tick, "The estimation of transfer functions of quadratic systems," *Technometrics,* vol. 3, no. 4, pp. 562-567, 1961.

[13] G.T. Zhou and G.B. Giannakis, "Nonlinear channel identification and performance analysis with PSK inputs," *Proc. 1st IEEE Signal Processing Workshop on Advances in Wireless Communications,* pp. 337-340, Paris, France, April 1997.

Stochastic Resonance in a Discrete Time Nonlinear SETAR(1,2,0,0) Model

Steeve Zozor and Pierre-Olivier Amblard
CEPHAG-ENSIEG UPRESA CNRS 5086
B.P. 46, 38402 Saint Martin d'Hères Cedex
Steeve.Zozor@cephag.inpg.fr Bidou.Amblard@cephag.inpg.fr

Abstract

We present in this paper the stochastic resonance phenomenon in a discrete time context. Indeed, stochastic resonance has been commonly investigated in continuous-time. Analytical results given by a simple bistable nonlinear SETAR$(1,2,0,0)$ *are studied. Then, the ability of such a system to be used in signal processing is discussed.*

1. Introduction

Stochastic Resonance (SR) is a physical phenomenon occurring in bistable dynamic systems excited by random noise and a pure frequency. Suppose that a particle is moving in the bistable potential $U(x)$ pictured in Figure 1. Under the assumption of strong friction, the equation of motion reduces to $\dot{x}(t) = -\nabla_x U(x)$. In that case, the particle will fall and stay in one well of the potential. If a random noise is added, $\dot{x}(t) = -\nabla_x U(x) + n(t)$, the particle will have a non-zero probability to jump from one well to the other. When thresholded, the output signal $x(t)$ will be roughly a telegraph signal. Now, if we add a sinusoid in the input, the potential will be modulated, and provided a fine tune of the parameters (sine and noise amplitudes), a cooperative effect between the sinusoid and the noise takes place: the particle will jump from one well to the other at the frequency of the sine. The interesting fact is that the output Signal-to-Noise Ratio (SNR) at the frequency of the sinusoid presents a maximum when plotted against the variance of the input noise! (see Figures 3 and 5).

This fact has lead several researchers to examine the ability of SR to detect small amplitude periodical signal [3, 4, 5]. This motivated this study. The theory of SR in continuous time systems is difficult. However, under some hypothesis, several approximate theories exist that explain that phenomenon [2]. In this paper, we examine SR in a discrete time context, and this is the first attempt to do so. The model we use is a nonlinear SETAR(1,2,0,0) model [7]. We consider two systems, which read

$$\begin{cases} \text{SETAR(1,2,0,0)} & : & x_n & = & \Phi(x_{n-1}) + e_n \\ \text{output of system 1} & : & y_n & = & x_n \\ \text{output of system 2} & : & y_n & = & \Psi(x_n) \end{cases}$$

where $\Phi(x) = c\,\text{sign}(x)$ and $\Psi(x) = V\,\text{sign}(x)$. The input is $e_n = b_n + \varepsilon \sin(2\pi n\lambda_0 + \varphi_0)$ where b_n is a Gaussian zero-mean white noise of variance σ^2. ε is supposed to be small compared to σ, and $\varepsilon \leq c$ (with only the modulation, system 1 and system 2 are not able to jump from $\pm c$ to $\mp c$). System 2 can be viewed as a discrete Schmitt Trigger, known in continuous time to provide SR [1]. In the following section, we give the theoretical study of those systems, which are difficult because of the presence of the sinusoid. Section 3 presents a discussion on the results and on the use of SR in Signal Processing.

2. Theoretical study of the systems

2.1. System 1

The output of the SETAR(1,2,0,0) is Markovian, thus, we can use the Chapmann-Kolmogorov equation to evaluate the probability density function (pdf) of x_n, $f_n(x)$,

$$f_n(x) = \int_{\mathbb{R}} f_b(x - \Phi(y) - \varepsilon_n)\, f_{n-1}(y)\, \mathrm{d}y \qquad (1)$$

where f_b denotes the pdf of the noise b_n and where $\varepsilon_n = \varepsilon \sin(2\pi n\lambda_0 + \varphi_0)$ represents the sine. Denoting α_n as

$$\alpha_n = \int_0^{+\infty} f_n(x)\, \mathrm{d}x \qquad (2)$$

the probability function can be expressed as

$$\begin{aligned} f_n(x) &= (1 - \alpha_{n-1})\, f_b(x + c - \varepsilon_n) \\ &\quad + \alpha_{n-1}\, f_b(x - c - \varepsilon_n) \end{aligned} \qquad (3)$$

Now integrating (3) over $[0\,;\,+\infty[$ leads to

$$\alpha_n = K_n\, \alpha_{n-1} + \text{erf}\left(\frac{-c + \varepsilon_n}{\sigma}\right) \qquad (4)$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

where $K_n = \text{erf}(\frac{c+\varepsilon_n}{\sigma}) - \text{erf}(\frac{-c+\varepsilon_n}{\sigma})$ and $\text{erf}(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})\,dx$. Under the assumption $\varepsilon \ll \sigma$, $\text{erf}(\frac{\pm c+\varepsilon_n}{\sigma}) \approx \text{erf}(\frac{\pm c}{\sigma}) + f_b(c)\varepsilon_n$, then α_n verifies the following equation

$$\alpha_n = K\alpha_{n-1} + \text{erf}(-\frac{c}{\sigma}) + L\varepsilon_n \tag{5}$$

where $K = \text{erf}(\frac{c}{\sigma}) - \text{erf}(\frac{-c}{\sigma})$ and $L = f_b(c)$. (5) can easily be studied, and it can be shown that α_n is asymptotically expressed as

$$\alpha_n = \frac{1}{2} + \frac{L}{|\alpha|}\xi_n \tag{6}$$

where $\alpha = 1 - K\mathrm{e}^{-2i\pi\lambda_0}$, and where $\xi_n = \varepsilon \sin(2\pi n\lambda_0 + \varphi_0 - \text{Arg}(\alpha))$.

Equation (6) shows that signal x_n is almost cyclostationary, and that system 1 improved a beating phenomenon: the weights α_n and $1-\alpha_n$ of $f_b(x \pm c - \varepsilon_n)$ pulse. This phenomenon is coupled to a moving mean phenomenon: the locations of $f_b(x \pm c - \varepsilon_n)$ periodically move.

The autocorrelation function of x_n, $\Gamma(n,q) = E[x_n x_{n+q}]$, is then expressed as

$$\Gamma(n,q) = \int_{\mathbb{R}^2} x_n x_{n+q} f(x_n, x_{n+q})\,dx_n\,dx_{n+q} \tag{7}$$

where $f(x_n, x_{n+q})$ denotes the joint probability density function of (x_n, x_{n+q}).

Without additive modulation ($\varepsilon = 0$), signal x_n is stationary and the correlation function is denoted $\Gamma_0(q)$. Considering that $q \geq 1$, $\Gamma_0(q+1)$ is given by

$$\begin{aligned}\Gamma_0(q+1) &= \int_{\mathbb{R}^2} x_n x_{n+q+1} f(x_n, x_{n+q+1})\,dx_{n+q+1}\,dx_n \\ &= \int_{\mathbb{R}^2} x_n \Phi(x_{n+q}) f(x_n, x_{n+q})\,dx_{n+q}\,dx_n \end{aligned} \tag{8}$$

where we have used the Chapmann-Kolmogorov equation within the integral. Continuing the recursion we get

$$\begin{aligned}\Gamma_0(q+1) = &\int_{\mathbb{R}^2} x_n f(x_n, x_{n+q-1}) \int_{\mathbb{R}} \Phi(x_{n+q}) \\ &f_b(x_{n+q} - \Phi(x_{n+q-1}))dx_{n+q}\,dx_{n+q-1}\,dx_n \\ = &\int_{\mathbb{R}^2} x_n c[\text{erf}(\frac{\Phi(x_{n+q-1})}{\sigma}) - \text{erf}(-\frac{\Phi(x_{n+q-1})}{\sigma})] \\ &f(x_n, x_{n+q-1})\,dx_{n+q-1}\,dx_n\end{aligned}$$

Using the following equality

$$c\,[\text{erf}(\frac{\Phi(x_{n+q-1})}{\sigma}) - \text{erf}(-\frac{\Phi(x_{n+q-1})}{\sigma})] = c\,\text{sign}(x_{n+q-1})K$$

we get

$$\Gamma_0(q+1) = K \int_{\mathbb{R}^2} x_n \Phi(x_{n+q-1}) f(x_n, x_{n+q-1})\,dx_{n+q-1}\,dx_n \tag{9}$$

Using (8) and (9) it can be concluded that for $q \geq 1$ the correlation function is recursively expressed as

$$\Gamma_0(q+1) = K\,\Gamma_0(q) \tag{10}$$

We evaluate $\Gamma_0(1) = Kc^2 + 2cL\sigma^2$ and $\Gamma_0(0) = c^2 + \sigma^2$. The autocorrelation function of x_n in the absence of modulation is then given by

$$\begin{cases} \Gamma_0(0) = c^2 + \sigma^2 \\ \Gamma_0(q) = (K.c^2 + 2cL\sigma^2).K^{q-1} \quad \text{for } q \geq 1 \end{cases} \tag{11}$$

For $\varepsilon \neq 0$, x_n is almost cyclostationary, and using the same way (which is a quite more complicate in this case) we show that the autocorrelation function of x_n is expressed as

$$\Gamma(n,q) = \Gamma_0(q) + \mu_{n+q}\mu_n \tag{12}$$

where

$$\begin{cases} \mu_n &= M\varepsilon \sin(2\pi n\lambda_0 + \varphi_0 - \varphi_1) \\ M^2 &= |a + i\,b|^2 \\ \varphi_1 &= \text{Arg}(a + i\,b) \\ a &= -\frac{2cL}{|\alpha|^2}\sin(2\pi\lambda_0) \\ b &= 1 + \frac{2cL}{|\alpha|^2}(\cos(2\pi\lambda_0) - K) \end{cases} \tag{13}$$

It can be noticed that $E[x_n] = \mu_n$. To get a spectrum at the zero-cyclic frequency, $\Gamma(n,q)$ is averaged over n. This leads to

$$\Gamma(q) = \Gamma_0(q) + \frac{\varepsilon^2 M^2}{2}\cos(2\pi q\lambda_0) \tag{14}$$

Then, the SNR of the output is evaluated as the ratio between the local power of the output sine around λ_0 and the local power spectral density of the output noise at λ_0. The SNR is then expressed as $SNR_o = \frac{\varepsilon^2 M^2}{4S_0(\lambda_0)}$, where $S_0(\lambda) = \sum_{q=-\infty}^{+\infty} \Gamma_0(q)\mathrm{e}^{-2i\pi\lambda}$ is the Fourier Transform of $\Gamma_0(q)$, the correlation of the output noise.

SNR_o depends on the frequency λ_0, and on the variance of the input noise σ^2. But SNR_o is not inversely proportional to σ^2 because of the nonlinearity of the system. The interacting effect between the sine and the noise can be seen through the SNR, and also through the amplification of the cosine M^2 on the autocorrelation function. These results will be more discussed in section 3

2.2. System 2

The output y_n of the second system is a two-states signal, and is also Markovian. Then, y_n is characterized by its

probability vector $\underline{p}(n)$, and its transition matrix $\underline{\underline{P}}(n)$ from step $n-1$ to n. In our context, we assume that ε is small compared to σ. Then, using (6) we get

$$\underline{p}(n) = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{L}{|\alpha|}\xi_n \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{15}$$

y_n is also almost cyclostationary, and is also submitted to a beating phenomenon. Nevertheless the moving mean phenomenon seen in x_n has been cancelled by the thresholding. The components of the transition matrix $\underline{\underline{P}}(n)$ are expressed by $P_{i,j}(n) = \Pr[y_n = (2i-3)V|_{y_n=(2j-3)V}]$ where $i,j \in \{1,2\}$ (we insist that with our notation $\underline{p}(n) = \underline{\underline{P}}(n)\,\underline{p}(n-1)$). Hence,

$$P_{1,1}(n) = \frac{\Pr[x_n < 0 \text{ and } x_{n-1} < 0]}{\Pr[x_{n-1} < 0]} \tag{16}$$

This leads to

$$P_{1,1}(n) = \frac{\int_{-\infty}^{0}\int_{-\infty}^{0} f(x_{n-1}, x_n)\,\mathrm{d}x_n\,\mathrm{d}x_{n-1}}{1-\alpha_{n-1}} \tag{17}$$

The joint density probability function of (x_{n-1}, x_n) is expressed as $f(x_{n-1}, x_n) = f(x_n|_{x_{n-1}})f(x_{n-1})$, i.e. $f(x_{n-1}, x_n) = f_b(x_n - \Phi(x_{n-1}) - \varepsilon_n)f(x_{n-1})$. Hence $P_{1,1}(n)$ is expressed as

$$P_{1,1}(n) = \mathrm{erf}(\frac{c-\varepsilon_n}{\sigma}) \tag{18}$$

i.e.

$$P_{1,1}(n) \approx \mathrm{erf}(\frac{c}{\sigma}) - L\varepsilon_n \tag{19}$$

Similarly

$$P_{2,2}(n) \approx \mathrm{erf}(\frac{c}{\sigma}) + L\varepsilon_n \tag{20}$$

It can be seen that $\mathrm{erf}(\pm\frac{c}{\sigma}) = \frac{1}{2}(1 \pm K)$. $\underline{\underline{P}}(n)$ is a stochastic matrix, hence $\underline{\underline{P}}(n)$ is expressed as

$$\underline{\underline{P}}(n) = \underline{\underline{P_o}} + L.\underline{\underline{D}}.\varepsilon_n \tag{21}$$

where

$$\underline{\underline{P_o}} = \frac{1+K}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1-K}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{22}$$

and where

$$\underline{\underline{D}} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \tag{23}$$

Notice that $\underline{\underline{P_o}}$ is the transition matrix without modulation ($\varepsilon = 0$). In the presence of the modulation, this transition matrix is modulated.

The transition matrix from step n to step $n+q$ is then expressed as $\underline{\underline{\Pi}}(n,q) = \underline{\underline{P}}(n+q)...\underline{\underline{P}}(n+1)$, that is approximatively $\underline{\underline{\Pi}}(n,q) \approx \underline{\underline{P_o}}^q + \sum_{k=0}^{q-1} \underline{\underline{P_o}}^{q-1-k}\,\underline{\underline{D}}\,\underline{\underline{P_o}}^k \varepsilon_{n+1+k}$

Using the following equations (which are evident)

$$\begin{cases} \underline{\underline{P_o}}\,\underline{\underline{D}} &= K\,\underline{\underline{D}} \\ \underline{\underline{D}}\,\underline{\underline{P_o}} &= \underline{\underline{D}} \end{cases} \tag{24}$$

it can be seen that

$$\underline{\underline{\Pi}}(n,q) = \underline{\underline{P_o}}^q + (\sum_{k=0}^{q-1} K^{q-1-k}\,\varepsilon_{n+1+k})\underline{\underline{D}} \tag{25}$$

Thus, we get

$$\underline{\underline{\Pi}}(n,q) = \underline{\underline{P_o}}^q + \frac{1}{|\alpha|}(\xi_{n+q} - K^q\xi_n)\underline{\underline{D}} \tag{26}$$

Using (22) it can be seen that

$$\underline{\underline{P_o}}^q = \frac{1+K^q}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1-K^q}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{27}$$

The autocorrelation function of y_n is defined by $\Gamma(n,q) = E[y_n y_{n+q}]$. This function is then expressed as $\Gamma(n,q) = V^2.[\Pi_{1,1}(n,q).p_1(n) + \Pi_{2,2}(n,q).p_2(n) - \Pi_{1,2}(n,q).p_2(n) - \Pi_{2,1}(n,q).p_1(n)]$. Using the previous results we get

$$\Gamma(n,q) = V^2[K^q(1 - \frac{4L^2}{|\alpha|}) + \frac{4L^2}{|\alpha|^2}\xi_{n+q}\xi_n] \tag{28}$$

Notice that $E[y_n] = \frac{2VL}{|\alpha|}\xi_n$. Again, averaging over n leads to

$$\Gamma(q) = V^2.K^q + \frac{2V^2L^2\varepsilon^2}{|\alpha|^2}\cos(2\pi q\lambda_0) \tag{29}$$

The correlation function of the noise part of the output is then $\Gamma_{\text{noise}}(q) = V^2(1 - \frac{4L^2}{|\alpha|})K^q$, hence $S_{\text{noise}}(\lambda) = \frac{V^2(1-\frac{4L^2}{|\alpha|})(1-K^2)}{1+K^2-2K\cos(2\pi\lambda)}$. The SNR of the output is then approximatively expressed as

$$SNR_o = \frac{\varepsilon^2 L^2}{1-K^2} \tag{30}$$

SNR_o does not depend of V, but the most important result is that SNR_o does not depend on the frequency λ_0.

SNR_o tends to 0 when σ^2 tends to 0 or $+\infty$. It yields that it exists an optimal noise variance σ^2_{opt} for which the SNR presents a maximum. This fact confirms that system 2 produces Stochastic Resonance. Moreover $S_0(\lambda)$ tends to V^2 when σ tends to the infinity: The output noise spectrum is then constant. This fact is a direct consequence of the thresholding, and not a consequence of the SR phenomenon (unlike the study of Fauve in [1] should suggest). Indeed this fact also appears without modulation. For the Schmitt Trigger [1] the thresholding is natural and the SR probably

appears when the power spectral density of the output noise is roughly constant. Hence Fauve thought that it was as if the sine have "pumped" energy to the noise. Nevertheless the sine effectively "pump" energy to the noise, and this at all the frequencies (cf. $S_{\text{noise}}(\lambda)$). Furthermore the sum of the power spectral density of the noise at λ_0 and the local power of the sine is constant for all ε, and in particular for $\varepsilon = 0$. That confirms the effect of "energy transfer" from the noise to the sine. This cooperative effect can be seen either on the SNR plotted against the variance of the input noise or on the local power of the sine plotted against σ^2. The previous results will be more discussed in the following section.

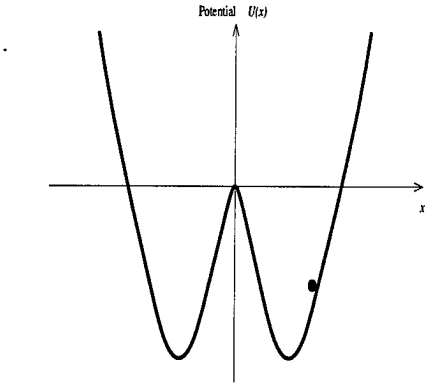


Figure 1. Example of bimodal potential

3. Discussion

Notice first that the SNR of the output of system 1 tend to $+\infty$ when $\sigma \to 0^+$, unlike the SNR of system 2 (figures 3 and 5). Indeed, with the modulation ($\varepsilon \neq 0$) and without noise ($\sigma = 0$), a particle in a bistable potential does not jump from one well to the other, but periodically moves in one well : the SNR is infinite. Thresholding the output (as in the second system) makes it constant ($\pm V$), since the particle stays in one well. It can also be concluded that SNR_o of the second system is less than SNR_o of the first one.

The maximum of SNR is then explained by a cooperative effect between the sine and the noise. A multi-spectral analysis of those systems may stress the interaction between the modulation and the noise.

Nevertheless, these two systems seem to be not interesting to detect a small amplitude noisy sine by adapting the parameters of the systems to the variance of the noise. Indeed, the SNR of the output presents a maximum for an optimal variance, but it can be observed in Figures 3 and 5 that the gains of the two systems $G = SNR_o/SNR_i$, where SNR_i is the SNR of the input, are always less or equal to 1. The study of G had never been done [1, 2, 3, 4, 5, 6], but we think that the same result may be found in continuous time systems. At the present time we study systems similar to system 2 using SETAR$(1, N, 0, \ldots, 0)$ models and thresholding the output, to generalize some of these results. The approach is to find systems for which the maximum of the local power of the sine (plotted against σ^2) appears while the spectrum of the noise $S_0(\lambda)$ is roughly constant (cf. [1]). It is expected that the maximum of the SNR is important and that the gain is greater than 1 for such systems.

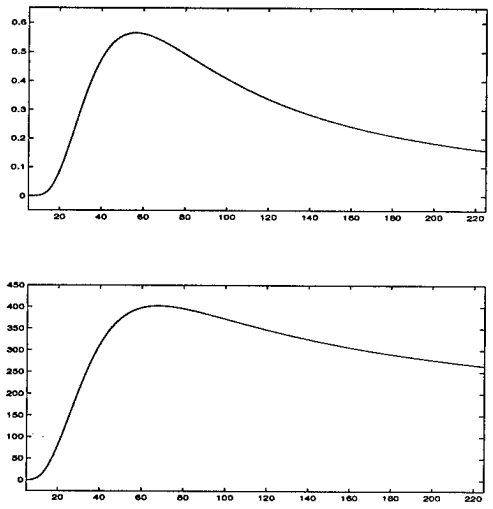


Figure 2. Theoretical local power of the output sine (top figure) and theoretical power spectral density of the output noise at λ_0 (bottom figure) of system 1 plotted against σ^2. $c = 10$, $\lambda_0 = .04$ and $\varepsilon = 1$).

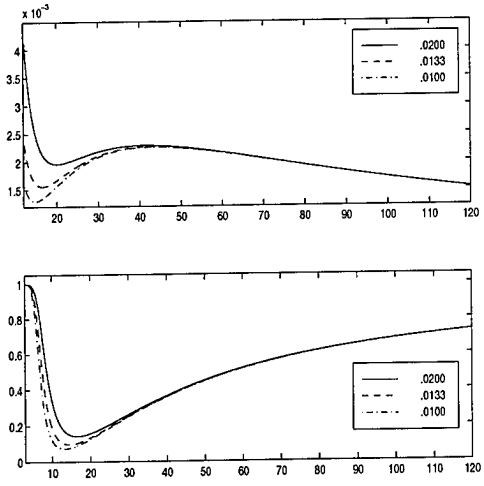


Figure 3. Theoretical output SNR (top) and theoretical gains (bottom) of system 1 plotted against σ^2 for several λ_0. $c = 10$ and $\varepsilon = 1$).

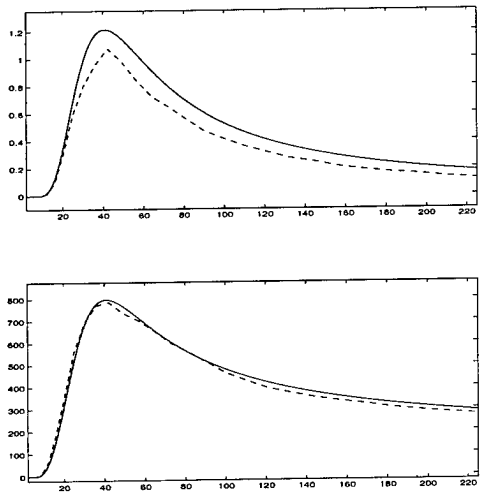


Figure 4. Top figure: Theoretical (line) and experimental (dotted line) local power of the output sine of system 2 plotted against σ^2; bottom figure: theoretical (line) and experimental (dotted line) power spectral density at λ_0 of the output noise of system 2 plotted against σ^2. $V = c = 10$ and $\varepsilon = 1$).

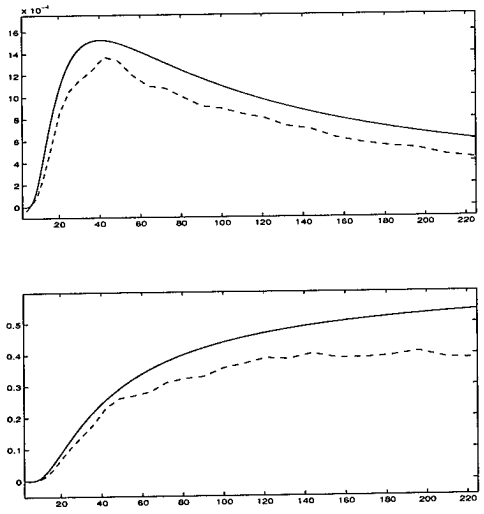


Figure 5. Top figure: Theoretical (line curve) and experimental (dotted line) output SNR of system 2 plotted against σ^2; bottom figure: theoretical (line curve) and experimental (dotted line) gain of system 2 plotted against σ^2. $V = c = 10$ and $\varepsilon = 1$.

References

[1] S. Fauve and F. Heslot *Stochastic resonance in a bistable system*, Physical Letters, Vol. 97A, p. 5, 1983

[2] McNamara, B.; Wiesenfeld, K. *Theory of stochastic resonance*, Physical Review A, Vol. 39, no. 9, pp. 4854-4869, 1989

[3] Jung, P.; Hanggi, P. *Amplification of small signals via stochastic resonance*, Physical Review A, Vol. 44, no. 12, pp. 8032-8042, 1991

[4] A.R. Bulsara, L. Gammaitoni *Tuning in to Noise*, Physics today, pp. 39-45, March 1996

[5] A. S. Asdi and A. H. Tewfik *Detection of weak signals using adaptive stochastic resonance*,IEEE International conference on Speech, Acoustic and, Signal Processing, vol. 2, pp. 1332-1335, 1995

[6] F. Moss *Stochastic resonance: from the ice ages to the monkey's ear*, in Some Contemporary Problems in Statistical Physics, edited by G. Weiss (SIAM, Philadelphia), pp. 205-53, 1994

[7] H. Tong *Non-linear Times Series: A Dynamical System Approach*, Oxford Science Publication, New York, 1990

Texture classification using third order correlation tools

Christophe COROYER
LACOS
Université du Havre
Place Robert Schumann, 76610 Le Havre, France
coroyer@iut.univ-lehavre.fr

David DECLERCQ, Patrick DUVAUT
E.T.I.S. - E.N.S.E.A.
6, Avenue du Ponceau
95014 CERGY Cedex, France
duvaut@ensea.fr

Abstract

This study presents a new method for textures classification based on higher order statistics (HOS). We propose the use of third order correlation tools for texture analysis. We compare the performance of three different tools : the bicorrelation in the spatial domain, the bispectrum in the frequency domain and the bicorspectrum which is a spatial/frequency representation in that case. We test classification on representative textures of Brodatz album.

1. Introduction

We can find a lot of textures analysis methods in litterature, based on very different principles. We can cite for the most used :

- the coocurrence or SGBD matrices [6]
- wavelets based method [7]
- Gabor filters [1]
- Markov based methods [5]

All these methods describe only the second order statistics.

We propose to test higher order statistics for texture analysis. We know that moments are not sufficient to describe efficiently a texture. We use dynamic tools, that is to say tool depending on spatial of frequency parameters as features for classification.

2. Gaussianity of textures

2.1. Gaussianity tests

Before using higher order statistics (HOS), it is crucial to check if the process is Normal or not. Within this framework, skewness and kurtosis tests (third and fourth order cumulants) form a set of relevant parameters when measuring normality ; they also exhibit a good compromise between numerical complexity and accuracy. We chose several representative Brodatz textures. Sub-images for estimating cumulants have 64x64 pixels. The precision of the estimators only depends on their mean and variance. Thus, 4096 samples are enough to get a good precision for both the skewness and the kurtosis. The measures of the skewness and kurtosis for normalized estimators, presented on table 1, show that most of the textures have non symetric p.d.f's (non zero skewness) [3].

2.2. Image of textures

We present 1 the texture basis. We choose 9 textures of the Brodatz album [2], with different properties qualifying these textures. There are fine or coarse textures, isotropic or not, ...

2.3. Results

We present Table 1 results of Skewness and Kurtosis tests for the textures. We process the tests on 4096 samples.

Texture/Ordre	$\hat{S}$	$\hat{K}$
D9 grass lawn	0.20	−0.82
D17 herringbone weave	−0.48	−0.64
D19 woolen cloth	0.41	−0.42
D24 pressed calf leather	0.65	−0.81
D29 beach sand	−0.54	−0.098
D38 water	−0.18	−0.23
D68 wood grain	−1.17	0.67
D84 raffia	0.030	−1.07
D112 plastic bubbles	0.91	0.55

Table 1.

0-8186-8005-9/97 $10.00 © 1997 IEEE

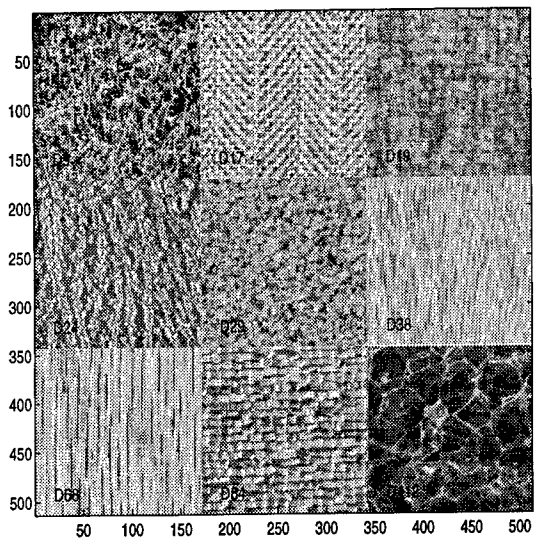


Figure 1. Base of Brodatz textures used

These values of the two cumulants exhibit that most of the textures are non Gaussian. Only the water can be considered as Gaussian. Moreover, the textures are well represented in the third order statistics. We can thus focus on the third order statistics, by using third order dynamic tools.

Remark 1 *The use of fourth order tools such as tricorrelation have two problem : on the one hand, the variance of estimators are larger and images have very limited numbers of samples. On the other hand, results of tricorelation is less readable, because it is in a three dimension domain for a 1D signal.*

3. Choice of a third order tool

We test three kind of third order correlation :

- the bicorrelation in the spatial domain
- the bispectrum in the frequency domain : we choose a non-parametric estimator, the 2D Fourier transform of the bicorrelation
- the bicorspectrum, a new spatial/frequency representation containing third order correlation

The three tools presented above give a 2D result for a 1D process. We calculate a sampled tool and give the result of an averaged estimation.

3.1. Denition of tools

The bicorrelation of a process is the dynamic version of the third order cumulant [8]. The bicorrelation of a real stationary process with zero mean is defined by :

$$\gamma_3[m,n] = E[x_c[t] \cdot x_c[t+m] \cdot x_c[t+n]] \quad (1)$$

We present on figures (2, 3) results of bicorrelations on two textures of Brodatz album : the wool image (D19 woolen cloth) and the wood image (D68 wood grain). We can see that the bicorrelation seems to be discriminant for analysing textures.

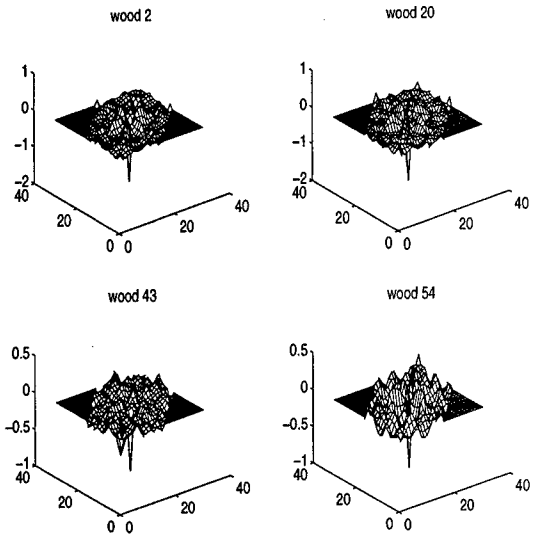


Figure 2. Sampling bicorrelation on rows of wood image

We present on figures (4, 5) results of bicorrelations on two textures of Brodatz album : the wool image (D19 woolen cloth) and the bubbles image (D68 plastic bubles).

We can see that the bispectrum seems also to be discriminant for analysing textures.

The Bicorspectrum or BET is defined by [4] :

$$\beta_{xw}(t,f) = \frac{1}{2T} \int_{\Re^2} x_w(u-t) \cdot x_w(u+\frac{\tau}{2}) \cdot x_w^*(u-\frac{\tau}{2}) \cdot e^{-j2\pi f\tau} d\tau du \quad (2)$$

where x_w is the windowed signal x of length $2T$.

We present on figures (6, 7) results of bicorrelations on two textures of Brodatz album : the weave image (D17 herringbone weave) and the grass image (D9 grass lawn). We can see that the bicorspectrum seems also to be discriminant for analysing textures.

3.2. Choice of the bicorrelation

By comparing the performance of the different third order tools, we have shown that the best feature is the bicorrelation. We have compared each tools itself, by

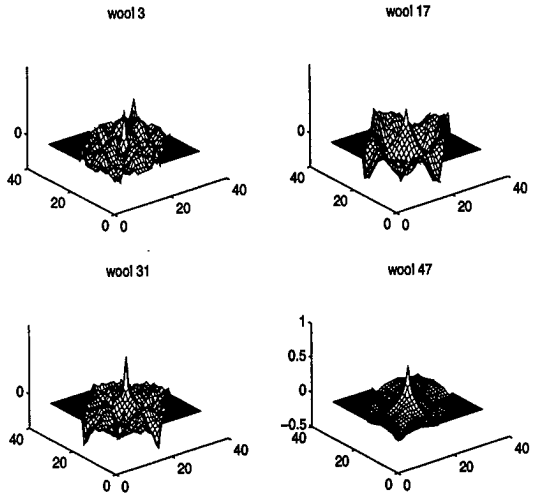


Figure 3. Sampling bicorrelation on rows of woolen cloth image

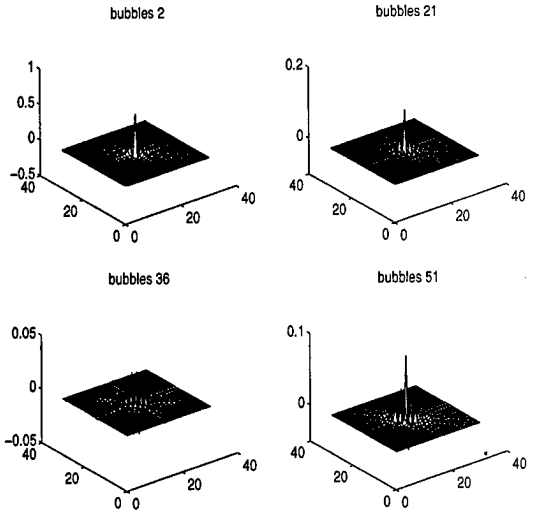


Figure 4. Sampled bispectrum on rows of plastic bubles image

testing their proper rate of good classification, using a simple distance to discriminate textures each other.

We use a biased estimator of bicorrelation (for $N-1 \geq m \geq n \geq 0$):

$$\chi_{3,b}[m,n] = \frac{1}{N} \sum_{t=0}^{N-m-1} x_c[t] \cdot x_c[t+m] \cdot x_c[t+n] \quad (3)$$

4. Texture Classification using Bicorrelation

4.1. The use of Bicorrelation

For a one dimension process, the sampled bicorrelation gives us a two dimensional result. To characterise textures, we want to define a simple vector of one dimensional parameters.Thus, we have defined a monodimensional tool by only keeping the values for $m = n$, that is to say on the principal axe.

4.2. Characterisation of textures

We consider texture as two stationary 1-D random processes rows and columns, as in image compression by wavelets. This type of image description conserves all the information contained in the initial 2D image.

The mean and the variance are very important parameters to describe textures, but they are not sufficient to obtain good performance for classification. The bicorrelation is estimated both for the rows and

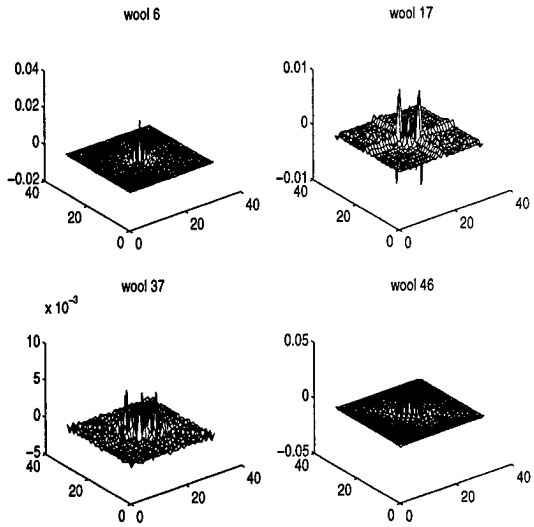


Figure 5. Sampled bispectrum on rows of woolen cloth image

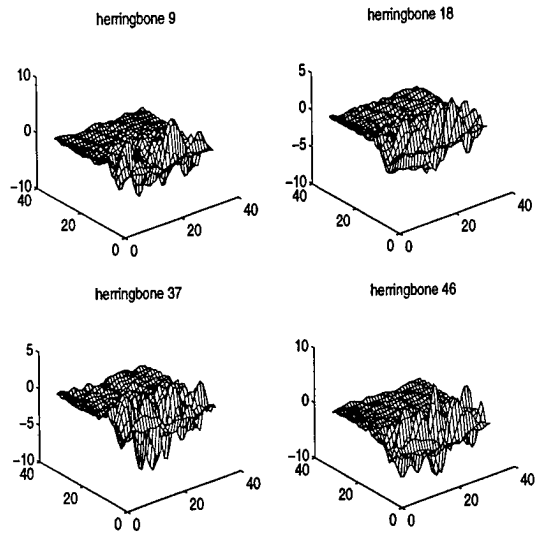


Figure 6. Sampling bicorspectrum on rows of herringbone weave image

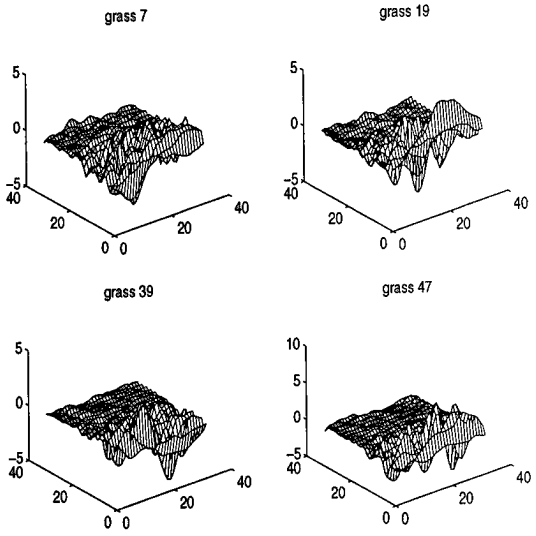


Figure 7. Sampling bicorspectrum on rows of grass image

the columns and the result, that is to say two vectors of bicorrelation, are given by evaluating the mean of the estimated bicorrelations along the rows and the columns. The sampling bicorrelation is estimated only on the first delays along the principal axe. To complete the statistical study, we add mean, variance, skewness, kurtosis and three delays of autocorrelation to the vector of parameters. We measure the difference of performance between second order and when we add third order tools.

We compare the results with a basic method, the SGBD or coocurrences matrices. We have used the four parmeters proposed by Haralick [6], for the two unit vectors, $\vec{i}$ and $\vec{j}$. These parameters are : contrast, entropy, correlation and two order angular moment.

We choose a neural network for discriminating textures. The learning rules are based on a backpropagation algorithm. The functions of associated decisions can then be written in form of production rules.

5. Performance on Brodatz Textures

We present the results of classification and compare to the results with parameters of coocurrence matrices in same directions. We use a leave one out validation method for testing the performance of the different tools. Surprisingly for large sub-images of test, we obtain better results by using correlations than coocurrence matrix.

1. a set 56 images 64x64 for each texture (leave one out rules).

Tool	Order 2	Order 3	Coocurrences
Good classification	96,8%	98,8%	95,4%
Non-classification	0,2%	0,2%	0,2%
Bad classification	3.0%	1.0%	4.4%

2. a set 224 images 32x32 for each texture (10% for learning set).

Tools	Order 2	Order 3	Coocurrences
Good classification	90,3%	92,5%	90,5%
Non-classification	6,2%	4,1%	3,5%
Bad classification	3.5%	3.4%	6%

With 10% for the test set and 90% for the learning set, the results are quite good with the decreasing number of samples to estimate bicorrelation. We obtain more than 92% of good classification over the test set.

6. Conclusion

To conclude, we propose to give a set of parameters, "adapted" for classification, depending on the size of analysis sub-image :

- for 64x64 or 32x32 sub-images, the set is composed with the mean, the variance, the skewness, the kurtosis, the first three lags of autocorrelation and firt two lags of bicorrelation, both for the rows and the columns.

- for 16x16 or less sub-images, SGBD matrices reach better performance

References

[1] A. C. Bovic, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell.*, 12:55–73, 1990.

[2] P. Brodatz. *Texture- A photographic album for artists and designers.* New York. Reinhold, 1968.

[3] C. Coroyer, C. Jorand, and P. Duvaut. Roc curves of skweness and kurtosis statistical tests: Application to textures. In *Proc. of Eur. Signal Proc. Conf. (EUSIPCO)*, pages 450–454, 1994.

[4] P. Duvaut, T. Gouraud, and C. Coroyer. Bet, the bicorspectral transform. a new tool for non-gaussian processes. In *IEEE ATHOS Workshop*, pages –, 1995.

[5] S. Geman and D. Geman. Stochastic relaxation and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6:–, 1984.

[6] R. Haralick, K. Shammugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. Syst., Man, Cyb.*, 3:610–621, 1973.

[7] A. Laine and J. Fan. Texture classification by wavelet packet signatures. *IEEE Trans. Pattern Anal. Machine Intell.*, 15:1186–1190, 1993.

[8] C. Nikias and J. M. Mendel. Signal processing with higher order spectra. *IEEE Signal Processing Magazine*, pages 10–37, 1993.

ON THE IDENTIFIABILITY OF BILINEAR STOCHASTIC SYSTEMS

Christophe Bourin and Pascal Bondon

Laboratoire des Signaux et Systèmes - CNRS
Plateau de Moulon, 91192 Gif-sur-Yvette, France
bourin@lss.supelec.fr, bondon@lss.supelec.fr

ABSTRACT

Bilinear systems are useful to model nonlinear time series. They can be described by a nonlinear recursive equation involving a finite number of parameters. Their analysis and particularly the estimation of the parameters is of central interest. In this paper we establish difference equations between lagged moments and cumulants up to third-order of a simple bilinear model, and show how to use these relations to estimate the parameters.

1. INTRODUCTION

In many signal processing problems the time series are represented by linear models. But in several situations the linear representation fails to describe the involved series satisfactorily. Thus, some nonlinear models have been proposed to handle such time series (see [7] and [3] for a recent review). The most well known of these models is the Volterra series expansion, which was first studied by Wiener. The input-output relation of a Volterra system is written by means of kernels. In theory, a nonlinear series can be expanded with an infinite number of kernels, but in practice the Volterra expansion must be truncated. Unfortunately, there does not exist any approximation criterion which allows us to state *a priori* at which degree of nonlinearity we may stop the expansion. Furthermore, the numerical complexity increases exponentially with this degree. That is the reason why only quadratic or cubic systems have been studied in the literature.

Recently, a new class of nonlinear models have been introduced in control theory [8, 5]. This is the class of bilinear systems. The input-output relation of such a system is described by a nonlinear recursive equation. The main advantage of this representation is that it involves only a finite number of parameters while the Volterra series expansion of a bilinear system requires an infinite number of kernels [3]. In other words bilinear models have the same kind of relationship with Volterra systems as ARMA models with linear systems. The counterpart of the reduced number of parameters is that the analysis is more complicated than that of a truncated Volterra system and this is due to the nonlinear recursive equation. Despite this aspect, bilinear models seems to be an appropriate tool when dealing with nonlinear time series. In section 2 we briefly present generalities on bilinear time series. Then we restrict our study to a simple model. In section 3 we give the expressions of lagged moments up to third-order and in section 4 we apply these formulae to the estimation of the model parameters. The cases where the input $\{e_n\}$ is available, and where $\{e_n\}$ is unobservable are considered.

2. BILINEAR MODELS

The general discrete-time bilinear series, denoted BL(p, q, r, k), is written

$$x_n = e_n + \sum_{i=1}^{p} a_i x_{n-i} + \sum_{i=1}^{q} c_i e_{n-i} + \sum_{i=1}^{r}\sum_{j=1}^{k} b_{ij} e_{n-i} x_{n-j} \quad (1)$$

where $\{e_n\}$ is a sequence of independent and identically distributed zero-mean random variables

0-8186-8005-9/97 $10.00 © 1997 IEEE

with variance $\sigma^2 < \infty$. The coefficients $\{a_i\}$ and $\{c_i\}$ correspond respectively to the AR and the MA part of (1) while the coefficients $\{b_{ij}\}$ correspond to the bilinear part. If $b_{ij} = 0$ for all $i > j$, the model is called superdiagonal. In a similar way, if $b_{ij} = 0$ for $i < j$, the model is subdiagonal. The statistical analysis of a subdiagonal model is much more difficult. This is due to the fact that in the last term on the right-hand side of (1), the contribution of $e_{n-i}x_{n-j}$ is nonzero when x_{n-j} is posterior to e_{n-i}. When $\{x_n\}$ is a causal function of $\{e_n\}$, x_{n-j} is then correlated with e_{n-i}, and this is why the analysis is more complicated. Finally, if $b_{ij} = 0$ for $i \neq j$, the model is said to be diagonal.

The study of the statistical properties of a bilinear model began with the works of Granger, Andersen [2] and Subba Rao [10]. The conditions to ensure the existence and uniqueness of a stationary and ergodic solution of (1) are more difficult to obtain than in the linear case. They have been studied by several authors in some special cases [6, 1]. Contrary to the linear case, the stationarity condition depends on the variance of the input $\{e_n\}$. For example, a BL(1,0,1,1) model is strictly stationary if and only if $|a_1^2 + b_{11}^2\sigma^2| < 1$.[1] When using a model (linear or nonlinear) for prediction, we require that this model be invertible, i.e., that its input can be expressed as a function of its output. When the model is linear, this condition is easy to obtain and does not depend on the distribution of the output. When the model is bilinear, this is not the case and the condition is much more difficult to obtain.

The parameters of a bilinear model can be estimated either by a maximum likelihood method, [11], or by a method of moments. Although the maximum likelihood estimator is usually more efficient than the moment estimator, the latter is simpler to implement. The method of moments consists in using recursive equations involving moments or cumulants and replacing the theoretical values by their estimates. The second-order properties of a bilinear system are similar to those of a linear system. More precisely, a BL(p, q, p, q) model has the same covariance structure as an ARMA(p, q) model. Consequently, we cannot distinguish a bilinear series from a linear one with a classical covariance analysis. It is then necessarily to also study some high-order cumulants. A method of moments has been used to identify very simple models in [2, 9, 4]. In particular, the MA part is absent in these models ($q = 0$ in (1)). To our knowledge, the case of the general bilinear model (1) is not solved. The reason is that the complexity of the calculations increases considerably even when only one term $e_{n-i}x_{n-j}$ is added.

[1] Note that this condition does not necessarily imply the existence of the moments.

3. EXPRESSIONS OF THE MOMENTS UP TO THIRD-ORDER

In the rest of the paper, we consider the bilinear model BL(1,1,1,1)

$$x_n = e_n + ax_{n-1} + ce_{n-1} + be_{n-1}x_{n-1} \quad (2)$$

where $\{e_n\}$ is a sequence of independent and identically distributed Gaussian and zero-mean random variables with variance σ^2. Furthermore, we assume that

$$a^2 + b^2\sigma^2 < 1 \quad (3)$$

which is a sufficient condition for the existence and uniqueness of a stable, causal and strictly stationary solution for the model (2).

The moments can be calculated directly using (2), but the calculations are tedious because of the diagonal term $e_{n-1}x_{n-1}$. Nevertheless, under condition (3) the model can be written with the Markovian representation

$$x_n = z_{n-1} + e_n \quad (4)$$

$$z_n = \alpha_n z_{n-1} + \beta_n e_n \quad (5)$$

$$\alpha_n = a + be_n \quad (6)$$

$$\beta_n = a + c + be_n. \quad (7)$$

As we are interested in a causal solution, we deduce that z_n depends on e_{n-k} only when $k \geq 0$. Consequently, z_n and e_{n+k} are independent when $k > 0$, and the use of equations (4)-(7) instead of (2) reduces considerably the complexity of the calculations to express the moments.

In the following, we derive the difference equations between lagged moments up to third-order

defined by $m_1 = \mathsf{E}(x_n)$, $m_2(k) = \mathsf{E}(x_n x_{n-k})$ and $m_3(k_1, k_2) = \mathsf{E}(x_n x_{n-k_1} x_{n-k_1-k_2})$. Note that the definition of $m_3(k_1, k_2)$ is different from the usual definition $\mathsf{E}(x_n x_{n-k_1} x_{n-k_2})$. This definition is well suited to the structure of the third-order moments of (2), but its counterpart is that $m_3(k_1, k_2) \neq m_3(k_2, k_1)$. Because of symmetry relations, we can reduce our study to the case where k, k_1, k_2 are nonnegative. The details of the calculations are omitted.

• The first-order moment of $\{x_n\}$ can be written

$$m_1 = \frac{b\sigma^2}{1-a}.$$

Note that according to (3), $a \neq 1$, and then m_1 is well-defined.

• The second-order moments of $\{x_n\}$ can be written

$$\begin{aligned} m_2(0) &= \sigma^2[2b^2\sigma^2 + c^2 + 2ac + 1 \\ &\quad + 2b(2a+c)m_1]/(1 - a^2 - b^2\sigma^2) \\ m_2(1) &= am_2(0) + 2b\sigma^2 m_1 + c\sigma^2 \\ m_2(k) &= am_2(k-1) + b\sigma^2 m_1 \end{aligned}$$

for all $k > 1$. The stationary condition (3) is sufficient to ensure the existence of the function m_2.

• The expression of the function m_3 is as follows. We have

$$m_3(0,0) = \frac{B_1 m_2(0) + C_1 m_1 + D_1}{1 - a^3 - 3ab^2\sigma^2}$$

where

$$\begin{aligned} B_1 &= 3b\sigma^2(3b^2\sigma^2 + 3a^2 + 2ac) \\ C_1 &= 3\sigma^2[6b^2\sigma^2(a+c) + 2a^2c + ac^2 + 1] \\ D_1 &= 3\sigma^4 b(2b^2\sigma^2 + 4ac + 3c^2). \end{aligned}$$

On the other hand

$$m_3(0,1) = A_2 m_3(0,0) + B_2 m_2(0) + C_2 m_1 + D_2$$

where

$$\begin{aligned} A_2 &= a^2 + b^2\sigma^2 \\ B_2 &= 2b\sigma^2(3a+c) \\ C_2 &= \sigma^2(b^2\sigma^2 + c^2 + 4ac + 1) \\ D_2 &= 4bc\sigma^4 \end{aligned}$$

and

$$m_3(0,k) = A_3 m_3(0,k-1) + B_3 m_2(k-1) + C_3 m_1$$

for all $k > 1$, with

$$\begin{aligned} A_3 &= A_2 \\ B_3 &= 2b\sigma^2(2a+c) \\ C_3 &= \sigma^2(2b^2\sigma^2 + c^2 + 2ac + 1). \end{aligned}$$

Lastly

$$\begin{aligned} m_3(1,0) &= am_3(0,0) + 3b\sigma^2 m_2(0) + 2c\sigma^2 m_1 \\ m_3(1,k) &= am_3(0,k) + 2b\sigma^2 m_2(k-1) + c\sigma^2 m_1 \end{aligned}$$

for all $k > 0$, and

$$m_3(k_1, k_2) = am_3(k_1 - 1, k_2) + b\sigma^2 m_2(k_2)$$

for all $k_1 > 1$ and $k_2 \geq 0$. The previous expressions are obtained under the assumption

$$|a|(a^2 + 3b^2\sigma^2) < 1. \tag{8}$$

This is a sufficient condition to ensure the existence of the function m_3. Note that the stationarity condition (3) does not necessary imply (8) and, as a consequence, the process $\{x_k\}$ can be strictly stationary but not third-order stationary since the function m_3 may not exist.

4. IDENTIFIABILITY

In this section we use the explicit relations between moments and cumulants and the expressions of the moments given in the previous section to derive the lagged cumulants of the bilinear process up to third-order. Then we address the problem of the determination of the parameters of (2) using these cumulants. We assume that the conditions (3) and (8) hold. Let $c_1 = \mathsf{Cum}(x_n)$, $c_2(k) = \mathsf{Cum}(x_n, x_{n-k})$ and $c_3(k_1, k_2) = \mathsf{Cum}(x_n, x_{n-k_1}, x_{n-k_1-k_2})$.

• The second-order cumulant function c_2 is

given by

$$c_2(0) = \sigma^2 \frac{b^4\sigma^4 + b^2\sigma^2(a+2c+1)(1-a)}{(1-a^2-b^2\sigma^2)(1-a)^2} + \frac{\sigma^2(c^2+2ac+1)}{1-a^2-b^2\sigma^2} \quad (9)$$

$$c_2(1) = ac_2(0) + c\sigma^2 + \frac{b^2\sigma^4}{1-a} \quad (10)$$

$$c_2(k) = ac_2(k-1) \quad (11)$$

for all $k > 1$.

• The third-order cumulant function c_3 satisfies

$$c_3(0,k) = (a^2+b^2\sigma^2)c_3(0,k-1) + 2b\sigma^2[a+c + b^2\sigma^2/(1-a)]c_2(k-1) \quad (12)$$

for all $k > 1$,

$$c_3(1,0) = ac_3(0,0) + 2b\sigma^2 c_2(0) \quad (13)$$

$$c_3(1,k) = ac_3(0,k) + b\sigma^2 c_2(k) \quad (14)$$

for all $k > 0$, and

$$c_3(k_1,k_2) = ac_3(k_1-1,k_2) \quad (15)$$

for all $k_1 > 1$ and $k_2 \geq 0$. If $c_2(k) = 0$ for all $k \neq 0$ and $c_3(k_1,k_2) = 0$ for all $k_1 > 0$ and $k_2 \geq 0$, it can be shown that we have $a = b = c = 0$ and $\sigma^2 = c_2(0)$, which means that $\{x_n\}$ is a Gaussian independent process. Now suppose that there exist one index $k' > 0$ such that $c_2(k') \neq 0$ or two indices $k'_1 > 0$ and $k'_2 \geq 0$ such that $c_3(k'_1,k'_2) \neq 0$. Then the AR coefficient a can be estimated either from (11) where $k = k'+1$, and we have

$$a = \frac{c_2(k'+1)}{c_2(k')} \quad (16)$$

or from (15) where $k_1 = k'_1 + 1$ and $k_2 = k'_2$ and we have

$$a = \frac{c_3(k'_1+1,k'_2)}{c_3(k'_1,k'_2)}. \quad (17)$$

The parameters b and c cannot be directly estimated. By introducing $\lambda = b\sigma^2$ and $\mu = c\sigma^2$, we deduce from (13) that

$$\lambda = \frac{c_3(1,0) - ac_3(0,0)}{2c_2(0)} \quad (18)$$

and then it results from (10) that

$$\mu = c_2(1) - ac_2(0) - \frac{\lambda^2}{1-a}. \quad (19)$$

Note that the identification of the parameters a, λ and μ is not possible using only the second-order cumulants. On the one hand, if σ^2 is known, the model is identifiable and we have $b = \lambda/\sigma^2$ and $c = \mu/\sigma^2$. On the other hand, if σ^2 is not known, it must be estimated from the relations between the cumulants. Equation (12) is equivalent to

$$H_1(k)\sigma^2 = H_2(k) \quad (20)$$

for all $k > 1$, where

$$H_1(k) = c_3(0,k) - a^2c_3(0,k-1) - 2\lambda a c_2(k-1)$$

$$H_2(k) = \lambda^2 c_3(0,k-1) + 2\lambda[\mu + \lambda^2/(1-a)] \times c_2(k-1).$$

It is not always possible to obtain σ^2 from (20) since the terms $H_1(k)$ and $H_2(k)$ may be zero for all $k > 1$. When $H_2(k_0) \neq 0$ for some $k_0 > 1$, the model is identifiable and theoretically σ^2 can be estimated from (20) when $k = k_0$. However, from a practical point of view this latter equation is not numerically stable and it would be better to use equation (9) which can be written as a second-order polynomial in σ^2. This polynomial admits two solutions but it can be shown after some tedious calculations that only the greater solution verifies equation (20). When $H_2(k_0) = 0$ for all $k_0 > 1$, the model is not identifiable and there exists two sets of parameters that are solutions of the problem.

The method of moments consists in replacing the theoretical cumulants by their estimates obtained from a finite number of output samples and solving for the parameters. When the input is available, σ^2 is directly estimated from the data and we can then identify the four parameters of the model. On the other hand, when the input is not observable, σ^2 is not known and has to be estimated using the output cumulants as described above. To illustrate these results we have generated 100 realizations, each of 1000 samples, for two sets of parameters $a = -0.5, b = 0.3, c = 0.1, \sigma^2 =$

1 and $a = 0.2, b = -0.2, c = 0.2, \sigma^2 = 1$. Assuming that only the output is available, we have estimated the parameters using equations (16-17), (18), (19) and (9). The means and variance of the estimates are given in Tables 1 and 2. We can see that the results are satisfactory and are in accordance with the results obtained in [9] for the identification of a BL(1,0,1,1) model which does not contain an MA part.

Table 1

	a	b	c	σ^2
true value	-0.5	0.3	0.1	1
mean	-0.5043	0.2985	0.0901	1.0064
variance	0.0028	0.0011	0.0056	0.0023

Table 2

	a	b	c	σ^2
true value	0.2	-0.2	0.2	1
mean	0.1902	-0.1912	0.2088	0.9795
variance	0.0053	0.0016	0.0048	0.0066

5. REFERENCES

[1] M. Bhaskara Rao, T. Subba Rao and A. M. Walker, "On the Existence of Some Bilinear Time Series Models", *J. Time Ser. Anal.*, vol. 4, pp. 95-110, 1983.

[2] C. W. J. Granger and A. P. Andersen, *An Introduction to Bilinear Time Series Analysis,* Vandenhoeck and Ruprecht: Göttingen, 1978.

[3] D. Guégan, *Séries Chronologiques Non Linéaires à Temps Discret,* Economica: Paris, 1994.

[4] W. K. Kim, L. Billard and I. V. Basawa, "Estimation for the First-Order Bilinear Time Series Model", *J. Time Ser. Anal.*, vol. 11, pp. 215-229, 1990.

[5] R. R. Mohler, *Bilinear Control Processes,* Academic Press: New York, 1973.

[6] T. D. Pham and L. T. Tran, "On the First-Order Bilinear Time Series Model", *J. Appl. Prob.*, vol. 18, pp. 617-627, 1981.

[7] M. B. Priestley, *Non-Linear and Non-Stationary Time Series Analysis,* Academic Press: London, 1988.

[8] A. Ruberti, A. Isidori and P. d'Alessandro, *Theory of Bilinear Dynamical Systems,* Springer Verlag: Berlin, 1972.

[9] S. A. O. Sesay and T. Subba Rao, "Yule-Walker Type Difference Equations for Higher-Order Moments and Cumulants for Bilinear Time Series Models," *J. Time Ser. Anal.*, vol. 9, no. 4, pp. 385-401, 1988.

[10] T. Subba Rao, "On the Theory of Bilinear Time Series Models," *J. R. Statist. Soc.*, ser. B 43, pp. 224-255, 1981.

[11] T. Subba Rao and M. M. Gabr, *An Introduction to Bispectral Analysis and Bilinear Time Series Models,* Springer-Verlag: New-York, 1984.

PLENARY

Signal Processing with Alpha-Stable Distributions: Current and Future Trends

Dr. Chrysostomos L. Nikias, *Univ. of Southern California (USA)*

The importance of extending the statistical signal processing methodology to the alpha-stable framework is apparent. First, scientists and engineers have started to appreciate lower-order moments and the elegant scaling and self-similarity properties of stable distributions. Additionally, real life applications exist in which impulsive channels tend to produce large-amplitude, short-duration interferences more frequently than Gaussian channels do. The stable law has been shown to successfully model noise over certain impulsive channels. In this talk, we present an overview of alpha-stable processes and lower-order statistics. In addition, we review recent advanced techniques for detection, parameter estimation, and system identification in the presence of signals/noise modeled as stable processes. Finally, we address future trends on signal processing within the alpha-stable framework.

TAP: System Identification & Deconvolution

Chair: Ed Powers
University of Texas-Austin, USA

Blind Channel Identification Based on Higher-Order Cumulant Fitting Using Genetic Algorithms

S. Chen †, Y. Wu † and S. McLaughlin ‡

† Department of Electrical and Electronic Engineering, University of Portsmouth
Anglesea Building, Portsmouth PO1 3DJ, U.K.
Email: schen@ee.port.ac.uk

‡ Department of Electrical Engineering, University of Edinburgh
King's Buildings, Edinburgh EH9 3JL, U.K.

Abstract

A family of blind equalisation algorithms identifies a channel model based on a higher-order cumulant (HOC) fitting approach. Since HOC cost functions are multimodal, gradient search techniques require a good initial estimate to avoid converging to local minima. We present a blind identification scheme which uses genetic algorithms (GAs) to optimise a HOC cost function. Because GAs are efficient global optimal search strategies, the proposed method guarantees to find a global optimal channel estimate. A micro-GA implementation is adopted to further enhance computational efficiency. As is demonstrated in computer simulation, this GA based scheme is robust and accurate, and has a fast convergence performance.

1. Introduction

An important class of blind equalisation algorithms uses techniques based on HOCs [1]–[4]. A two-stage strategy is usually adopted, which first identifies a channel model using HOC fitting algorithms and then employs the estimated channel model to design an equaliser. The key step of this approach is its ability to obtain an accurate channel model. Once a channel model is available, a variety of existing equaliser design methods can be employed, ranging from simple linear inverse filter to sophisticated maximum likelihood sequence estimator, depending on a trade-off between performance and complexity. Therefore, we concentrate on blind channel identification using the HOC fitting approach in this paper.

HOC fitting cost functions are however multimodal, and conventional gradient techniques [3],[4] may converge to "wrong" solutions unless a good initial value for the channel parameters is provided, which is not always possible. To overcome the problem of local minima, simulated annealing has been implemented to optimize a HOC cost function [5]. We propose to use GAs [6]–[9] for blind channel identification based on HOC fitting. This GA based scheme is very robust and achieves a global optimal solution regardless of initial value of channel estimate. Furthermore, the number of parameters to be optimised in the problem of blind channel estimation is usually small, and GAs are particularly effective to solve this kind of optimisation problems. The micro-GA implementation [8] is adopted to further improve convergence rate.

The channel is modelled as a finite impulse response filter with an additive Gaussian white noise:

$$r(k) = \sum_{i=0}^{n_a} a_i s(k-i) + e(k) \qquad (1)$$

Blind channel identification refers to the determination of the channel model $\mathbf{a} = [a_0\ a_1 \cdots a_{n_a}]^T$ using only the noisy received signal $\{r(k)\}$ and some knowledge of statistic properties of $s(k)$. We will assume a real-valued channel and a PAM symbol constellation. Extension to complex-valued channels and modulation schemes is straightforward.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Higher-order cumulant fitting

Since $3rd$-order cumulants of $r(k)$ are zero, $4th$-order cumulants have to be used [10]. The $4th$-order cumulant fitting cost function adopted in our application is defined by:

$$J_4(\hat{\mathbf{a}}) =$$

$$\sum_{\tau=-\hat{n}_a}^{\hat{n}_a} \left(\hat{C}_{4,r}(\tau,\tau,\tau) - \gamma_{4,s} \sum_{i=\max\{0,-\tau\}}^{\min\{\hat{n}_a,\hat{n}_a-\tau\}} \hat{a}_i \hat{a}_{i+\tau}^3 \right)^2 \quad (2)$$

where $\hat{C}_{4,r}(\tau,\tau,\tau)$ is the diagonal slice of the time estimate $\hat{C}_{4,r}(\tau_1,\tau_2,\tau_3)$ of the $4th$-order cumulant, $\gamma_{4,s}$ the kurtosis of $s(k)$, $\hat{n}_a$ an estimated channel length and $\hat{\mathbf{a}}$ the channel estimate.

Most of the existing algorithms for HOC fitting employ gradient search techniques. HOC fitting cost functions are well-known to be multimodal, and gradient based optimisation methods may fail to work. Even with measures of providing good initial channel estimate, it has been observed that gradient algorithms sometimes converge to local minima [3]. Using GAs to optimize the cost function (2) has the advantage of guaranteeing to find a global optimal channel estimate.

A channel model (1) typically contains a few taps. Thus the number of parameters to be optimized in (2) is small, and GAs are efficient in solving this kind of "small-dimensional" optimisation problems. In reality, the channel order n_a is unknown. A simple method is to overfit with $\hat{n}_a \geq n_a$. Although this will complicate the cost function and may cause problems to gradient based methods, the GA based method is capable of identifying those nonexisting taps with (near) zero values. An inspection of the obtained channel estimate allows us to delete those insignificant taps.

3. Genetic algorithms

GAs belong to a problem solving approach based loosely on the evolution of species in nature. They differ from gradient optimisation techniques mainly in four aspects [7]. Firstly, GAs work with an encoding of the parameter set to be searched, not the parameters themselves. Secondly, unlike gradient techniques, which concentrate their efforts on a single potential solution in the search space, GAs search with a population of potential solutions. Thirdly, GAs use the value of the objective function (termed fitness), not derivatives, to evaluate potential solutions. Lastly, GAs use probabilistic transition rules. The seemingly undirected search is guided by the fitness value of each individual and how it compares with others.

A popular encoding scheme is the bit-string encoding [6], which is adopted in our application. A simple GA usually consists of three operations, namely selection, crossover and mutation [7], at each cycle or generation. An "elitist" strategy, which automatically copies a few of the best solutions in the population into the next genneration, is often incorporated. In crossover operation, we adopt multiple crossover points [7], and the number of crossover points in our application is 4.

Since our goal is to find a global optimum solution quickly, the micro-GA [8] offers certain advantages. The population size in a micro-GA is much smaller than those in "standard" GAs. This feature of the micro-GA not only makes it particularly useful for nonstationary optimisation problems but also improves convergence rate in general [8]. Simply adopting a very small population size and letting the search converge just once is not very useful apart from quickly allocating some local optimum. In a micro-GA, after the search has converged, the population is reinitialised with random values while the best individual found so far is copied to the newly generated population. The reinitialisation is repeated until no further improvement can be achieved.

In general, the more complex the search space is, the larger the population size should be. The population size in our micro-GA is 2 times of the number of parameters. The crossover rate is set to 1.0 to facilitate a high rate of information exchange while the mutation rate is set to 0.0 as the reinitialisation of the population will keep the diversity of potential solutions fairly well. Most of GAs adopt the proportional selection [7]. Due to small population size of the micro-GA, the tournament selection is used in choosing parents [8].

4. GA based cumulant fitting scheme

Our GA based scheme is summarized as follows:

Step 1. Given a set of $\{r(k)\}_{k=1}^{N}$, assume an overlength $\hat{n}_a$ and compute the time estimate of the required cumulant to form the cost function (2).

Step 2. With a set of initial channel parameter vectors $\{\hat{\mathbf{a}}_i\}_{i=1}^{n_p}$, where n_p is the population size, use the micro-GA to optimise the cost function (2).

The fitness function value f_i for $\hat{\mathbf{a}}_i$ is defined as

$$f_i = \frac{1}{J_4(\hat{\mathbf{a}}_i)} \tag{3}$$

A measure is used to ensure that each $\hat{\mathbf{a}}_i$ satisfies

$$\sum_{j=0}^{\hat{n}_a} \hat{a}_{i,j}^4 = \frac{\hat{C}_{4,r}(0,0,0)}{\gamma_{4,s}} \tag{4}$$

In the population initialisation, the taps of each $\hat{\mathbf{a}}_i$ first take values randomly from $(-1,\ 1)$, and each chosen $\hat{\mathbf{a}}_i$ is normalised using (4). When a new population is produced, each member is also normalised. This ensures that the population is inside the feasible set of channel models and has an effect of improving convergence rate.

It is well-known that sign and time-shift ambiguities exist in blind identification. Sign ambiguity is due to the fact that both $\mathbf{a}$ and $-\mathbf{a}$ are global optimal solutions of (2). Time-shift ambiguity can be illustrated as follows. Let the true channel be $[a_0 \cdots a_{n_a}]^T$, $a_0 \neq 0$ and $a_{n_a} \neq 0$. Let $\hat{n}_a = n_a + 2$. Then $[a_0 \cdots a_{n_a}\ 0\ 0]^T$, $[0\ a_0 \cdots a_{n_a}\ 0]^T$ and $[0\ 0\ a_0 \cdots a_{n_a}]^T$ are all global optimal solutions. A solution to time-shift ambiguity is to fix the first tap. We do not fix a tap but use the following measure. If the first tap of a member is zero (absolute value smaller than a threshold), a shifting is performed to ensure that the first tap is always nonzero. Complexity of the GA based scheme is determined by the number of cost-function evaluations. The micro-GA employed is specifically designed to minimize this complexity.

5. Simulation results

The two channels used in the simulation were:

Channel 1 $\mathbf{a} = [-0.21\ -0.50\ 0.72\ 0.36\ 0.21]^T$
Channel 2 $\mathbf{a} = [0.227\ 0.460\ 0.688\ 0.460\ 0.227]^T$

The performance of the algorithm was assessed through the best cost function value $J_4(\tilde{\mathbf{a}})$, where $\tilde{\mathbf{a}}$ was the best channel model in the population, and the mean tap error (MTE) defined by:

$$\text{MTE} = \| \pm \tilde{\mathbf{a}} - \mathbf{a}\|^2 = \sum_{i=0}^{\hat{n}_a} (\pm\tilde{a}_i - a_i)^2 \tag{5}$$

In the expression (5), $-\tilde{\mathbf{a}}$ is used if $\tilde{\mathbf{a}}$ converges to $-\mathbf{a}$. Otherwise $\tilde{\mathbf{a}}$ is used.

8-PAM data symbols were transmitted and 50000 noisy received data samples were used to compute the time estimate of the $4th$-order cumulant. All the results were averaged over 100 different runs. Figs. 1 and 2 depict evolutions of the cost function and the MTE for channel 1 with different signal-to-noise ratio (SNR) conditions and assumed channel lengths $\hat{n}_a$, respectively. Table 1 summarises the results (mean$\pm$variance) for channel 1 with a SNR of 20 dB. Results for channel 2 are similarly given in Figs. 3 and 4 and table 2.

Some observations can be drawn. The GA based scheme always finds a global optimal channel estimate and the optimisation process converges fast. Compared with other existing methods of HOC fitting, our method is more accurate and robust, as is demonstrated by very small variances of estimated channel taps. Our method is capable of identifying nonexisting channel taps with (near) zero values (at least an order smaller than values for existing taps). Thus, model order selection can simply be carried out by first assuming an overlength $\hat{n}_a$ and then inspecting the obtained channel estimate to delete those insignificant taps. We also performed a range of simulation with a data length of 25000 samples and 16-PAM symbols. The results, not shown here, also confirm the above observations.

6. Conclusions

A GA based scheme has been developed for blind channel identification based on HOC fitting. Apart from ensuring a global optimal channel estimate regardless of initial conditions, the proposed method is highly accurate and very robust. Small variances of channel estimates and insensitivity to noise achieved in our simulation was not reported previously in other existing methods. Our application also demonstrates advantages of using the

micro-GA for fast global optimisation of multimodal cost functions.

References

[1] K.S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, Vol.10, pp.1195–1208, 1982.

[2] G.B. Giannakis and J.M. Mendel, "Identification of nonminimum phase system using higher order statistics," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-37, pp.360–377, 1989.

[3] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Blind equalization of nonminimum phase channels: higher order cumulant based algorithm," *IEEE Trans. Signal Processing*, Vol.41, pp.681–691, 1993.

[4] J.K. Tugnait, "Blind equalization and estimation of digital communication FIR channels using cumulant matching," *IEEE Trans. Communications*, Vol.43, pp.1240–1245, 1995.

[5] J. Ilow, D. Hatzinakos and A.N. Venetsanopoulos, "Blind equalizers with simulated annealing optimization for digital communication systems," *Int. J. Adaptive Control and Signal Processing*, Vol.8, pp.501–522, 1994.

[6] J.H. Holland, *Adaptation in Natural and Artificial Systems.* Ann Arbor, MI.: University of Michigan Press, 1975.

[7] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning.* Reading, MA.: Addison-Wesley, 1989.

[8] K. Krishnakumar, "Micro-genetic algorithms for stationary and non-stationary function optimization," in *Proc. SPIE Intell. Cont. Adapt. Syst.*, 1989, Vol.1196, pp.289–296.

[9] L. Yao and W.A. Sethares, "Nonlinear parameter estimation via the genetic algorithm," *IEEE Trans. Signal Processing*, Vol.42, pp.927–935, 1994.

[10] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, Vol.79, pp.278–305, 1991.

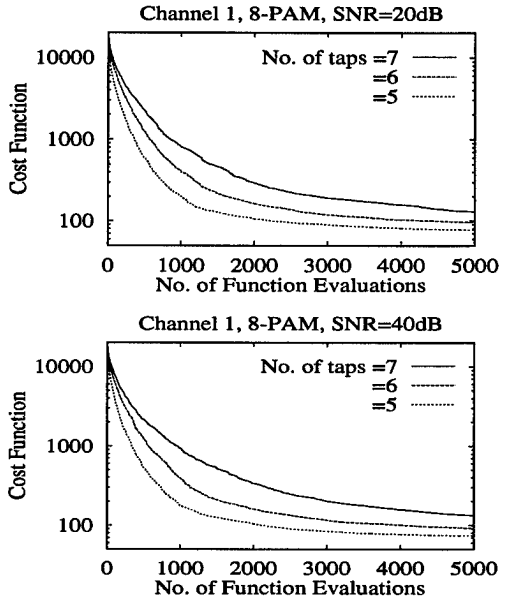


Figure 1. Cost function versus number of function evaluations for channel 1.

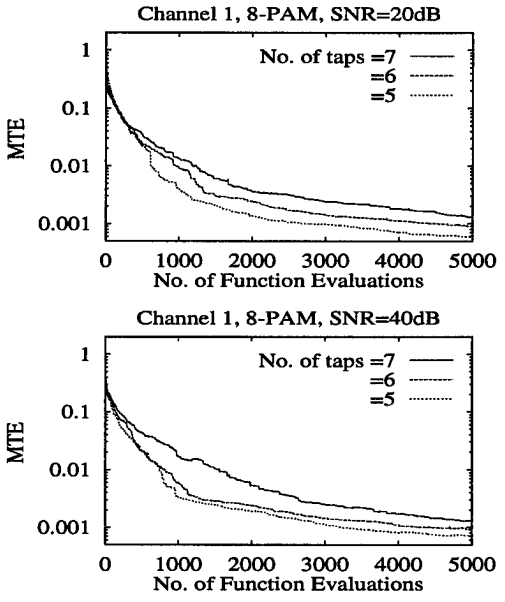


Figure 2. Mean tap error versus number of function evaluations for channel 1.

	true	estimate (mean±variance)		
	$n_a = 4$	$\hat{n}_a = 4$	$\hat{n}_a = 5$	$\hat{n}_a = 6$
a_0	-0.21	-0.20975±0.00014	-0.21124±0.00022	-0.21098±0.00020
a_1	-0.50	-0.49957±0.00003	-0.49975±0.00004	-0.49835±0.00005
a_2	0.72	0.72092±0.00002	0.71965±0.00002	0.72087±0.00003
a_3	0.36	0.35788±0.00006	0.35859±0.00010	0.35823±0.00011
a_4	0.21	0.21073±0.00016	0.21058±0.00013	0.20870±0.00017
a_5	—	—	-0.00041±0.00050	-0.00184±0.00050
a_6	—	—	—	-0.00173±0.00055

Table 1. Identification results for channel 1 with 8-PAM and SNR=20dB.

	true	estimate (mean±variance)		
	$n_a = 4$	$\hat{n}_a = 4$	$\hat{n}_a = 5$	$\hat{n}_a = 6$
a_0	0.227	0.22731±0.00080	0.22221±0.00099	0.21743±0.00119
a_1	0.460	0.45727±0.00097	0.45679±0.00127	0.45481±0.00242
a_2	0.688	0.68913±0.00070	0.68095±0.00066	0.67408±0.00181
a_3	0.460	0.45744±0.00052	0.46426±0.00077	0.46870±0.00262
a_4	0.227	0.22564±0.00079	0.22322±0.00074	0.22265±0.00217
a_5	—	—	0.01577±0.00781	0.01514±0.00653
a_6	—	—	—	0.00359±0.00485

Table 2. Identification results for channel 2 with 8-PAM and SNR=20dB.

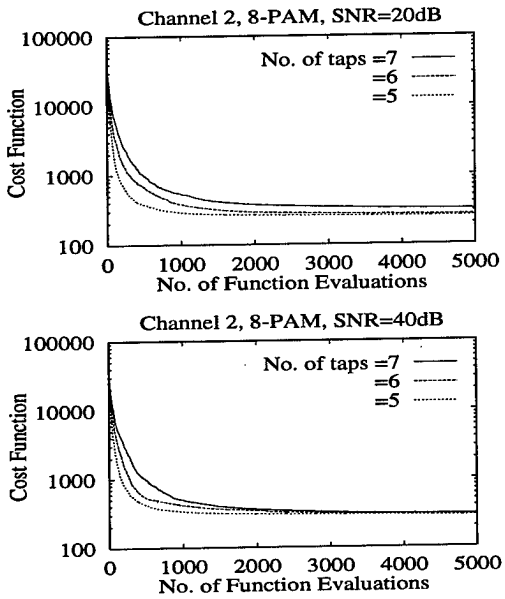


Figure 3. Cost function versus number of function evaluations for channel 2.

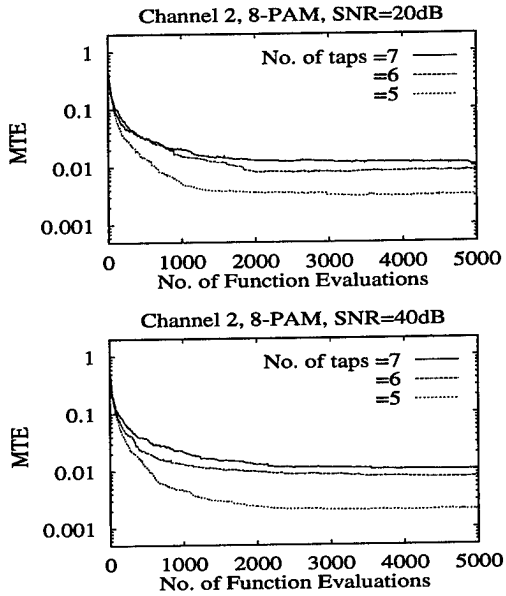


Figure 4. Mean tap error versus number of function evaluations for channel 2.

Blind System Identification Using Fourth Order Spectral Analysis of Complex Signals

Cécile HUET and Joël LE ROUX
I3S, University of Nice Sophia Antipolis - CNRS
250 rue Albert Einstein Sophia Antipolis 06560 Valbonne FRANCE
Tel: +33 04.92.94.26.82; fax: +33 04.92.94.28.96
e-mail: huet@alto.unice.fr - leroux@alto.unice.fr

Abstract

In this paper we give an analytic optimal solution to the identification problem of non minimum phase systems using the fourth order spectra. We show that this solution is in first approximation equivalent to the solution given by the well-known kurtosis maximization method. The proposed solution gives the phase of the system transfer function, the modulus can be obtained from the second order statistics. However, this solution requires a trispectrum phase unwrapping as the trispectrum phase is known in the interval $[-\pi, \pi[$ *but need to be unwrapped in the interval* $[-4\pi, 4\pi[$ *in order to obtain the optimal solution. Therefore, we present different phase unwrapping solutions. Next, we propose a method to improve the trispectrum phase estimation using a factorizability condition. Simulation results are given and the algorithm shows good behavior even with few data.*
Keywords : HOS - Fourth Order Statistics - Blind Identification.

1. INTRODUCTION

Fourth order statistics of complex signals are among the possible tools for treating the problem of channel identification and equalization in digital communications [1]. On one hand, time domain methods have been studied for several years [2],[1]. These are parametric methods and requires a correct determination of the system order. On the other hand, some frequency domain methods have been developed [1],[3],[4],[5]. Pierce [6],[7] and Shalvi & Weinstein [8] have treated the case of complex signals with applications in the field of radar signals and equalization respectively.

In this paper, we present an optimal solution to the phase reconstruction problem that use all the trispectrum information. However, this solution requires a previous phase unwrapping step to determine the trispectrum phase in the interval $[4\pi; 4\pi[$. In order to ensure this correct determination, unwrapping methods [9] require that the maximum error on the estimated trispectrum phase is less than $\pi/4$. In this paper, we develop different phase unwrapping solutions. Next, we present a method to improve the trispectrum phase estimation. This method is based on a factorizability condition [10]. Indeed, the optimal trispectrum obtained from the measurements is the factorizable trispectrum which minimizes the distance from the measured values. Finally simulation results show the results enhancement provided by such an improvement method.

2. DEFINITIONS

2.1. Fourth order spectrum of complex signals

Our developments are based on the fourth order cumulant which never vanishes even under circularity hypothesis [11][12]. The fourth order cumulant of a complex random sequence $x(t)$ is given by :

$$\begin{aligned} C_4^x(t_1,t_2,t_3) &= E\{x^*(\tau)x(\tau+t_1)x(\tau+t_2)x^*(\tau+t_3)\} \\ &\quad -E\{x^*(\tau)x(\tau+t_1)\}E\{x(\tau+t_2)x^*(\tau+t_3)\} \\ &\quad -E\{x^*(\tau)x(\tau+t_2)\}E\{x(\tau+t_1)x^*(\tau+t_3)\} \\ &\quad -E\{x(\tau+t_1)x(\tau+t_2)\}E\{x^*(\tau)x^*(\tau+t_3)\}, \quad (1) \end{aligned}$$

and the corresponding fourth order spectrum is [6][7] :

$$\begin{aligned} T_4^X(\omega_1,\omega_2,\omega_3) &= \quad (2) \\ &E\{X(\omega_1)X(\omega_2)X^*(-\omega_3)X^*(\omega_1+\omega_2+\omega_3)\} \\ &-E\{X(\omega_1)X^*(\omega_1)\}E\{X(\omega_2)X^*(\omega_2)\}\delta(\omega_2+\omega_3) \\ &-E\{X(\omega_2)X^*(\omega_2)\}E\{X(\omega_1)X^*(\omega_1)\}\delta(\omega_1+\omega_3) \\ &-E\{X^*(\omega_3)X^*(-\omega_3)\}E\{X(\omega_1)X(-\omega_1)\}\delta(\omega_1+\omega_2). \end{aligned}$$

where $X(\omega)$ is the Fourier transform of the sequence $x(t)$ and $\delta(\omega) = 1$ if $\omega = 0$ and $\delta(\omega) = 0$ elsewhere.

- Remark: the three planes appearing in (2) hold no phase information.
- Note that this definition is equivalent to the one given in [13] if we note $\omega_3 = -\omega_3$.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2.2. Non-redundant trispectrum domain for complex signals

The trispectrum definition implies the following symmetries :

$$\begin{aligned}\psi(\omega_1,\omega_2,\omega_3) &= \psi(\omega_2,\omega_1,\omega_3)\\ &= -\psi(\omega_1+\omega_2+\omega_3,-\omega_3,-\omega_1)\\ &= -\psi(\omega_1+\omega_2+\omega_3,-\omega_3,-\omega_2)\\ &= -\psi(-\omega_3,\omega_1+\omega_2+\omega_3,-\omega_1)\\ &= -\psi(-\omega_3,\omega_1+\omega_2+\omega_3,-\omega_2)\\ &= \psi(\omega_1,\omega_2,-(\omega_1+\omega_2+\omega_3))\\ &= \psi(\omega_2,\omega_1,-(\omega_1+\omega_2+\omega_3))\end{aligned}$$

Using these trispectrum symmetries, the non-redundant trispectrum support is defined as follows :

- Inside the stationary domain :

$$-\tfrac{1}{2} \le \omega_1,\omega_2,\omega_3 < \tfrac{1}{2} \qquad -\tfrac{1}{2} \le \omega_1+\omega_2+\omega_3 < \tfrac{1}{2};$$
$$\omega_2 \le -\omega_3 \qquad \omega_1 \le -\omega_3 \qquad \omega_1 \ge \omega_2 .$$

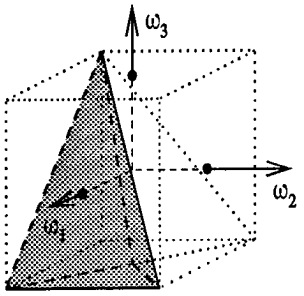


Figure 1. Non redundant domain inside the stationary support ($\frac{1}{8}$ of the entire domain inside the stationary support).

- Outside the stationary domain :

$$\omega_1,\omega_2,\omega_3 < \tfrac{1}{2} \qquad \omega_1+\omega_2+\omega_3 > \tfrac{1}{2};$$
$$\omega_1 \ge \omega_2 \qquad \omega_1+\omega_2+2\omega_3 > 1.$$

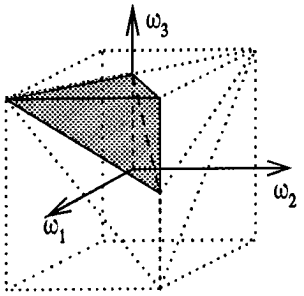


Figure 2. Non redundant domain outside the stationary support ($\frac{1}{8}$ of the entire domain outside the stationary support).

2.3. Identification using the trispectrum

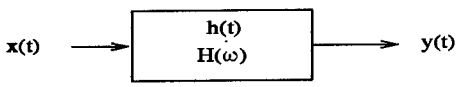


Figure 3. Identification scheme.

If the analyzed signal is the output of a LTI system driven by a non-gaussian zero-mean IID complex sequence $x(t)$ (Fig. 3), the fourth order spectrum phase satisfies :

$$\begin{aligned}\psi_4^Y(\omega_1,\omega_2,\omega_3) &= \varphi^H(\omega_1)+\varphi^H(\omega_2) \\ &\quad -\varphi^H(-\omega_3)-\varphi^H(\omega_1+\omega_2+\omega_3)+k\pi,\end{aligned} \tag{3}$$

where $\psi_4^Y(\omega_1,\omega_2,\omega_3)$ is the output trispectrum phase, $\varphi^H(\omega)$ is the system Fourier transform phase and $k = 0$ or 1 depending on the input kurtosis sign. However the input kurtosis sign being equal to the output kurtosis sign, the value of k is known from the output measurements, and (3) can be written :

$$\begin{aligned}\psi_4(\omega_1,\omega_2,\omega_3) &= \psi_4^Y(\omega_1,\omega_2,\omega_3)-k\pi = \\ \varphi^H(\omega_1)+\varphi^H(\omega_2)-\varphi^H(-\omega_3)-\varphi^H(\omega_1+\omega_2+\omega_3).\end{aligned} \tag{4}$$

In this paper, we use the fourth order statistics to reconstruct the transfer function phase only, as its magnitude can be obtained from the second order statistics, when there is no measurement noise.

3. LEAST SQUARES PHASE RECONSTRUCTION FROM THE TRISPECTRUM PHASE

The optimal phase reconstruction $\hat{\varphi}^H$ obtained from the measured trispectrum phase $\psi_4(\omega_1,\omega_2,\omega_3)$ uses all redundant trispectrum measures. However this reconstruction requires a prior phase unwrapping (cf. [9]). The criterion and the general formula for real signals are given in [4]. In the complex case, the optimal fourth order solution is given by the phase which minimizes:

$$\sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1}\left[\psi_4(\omega_1,\omega_2,\omega_3)-\hat{\psi}_4^H(\omega_1,\omega_2,\omega_3)\right]^2, \tag{5}$$

where

$$\begin{aligned}\hat{\psi}_4^H(\omega_1,\omega_2,\omega_3) &= \hat{\varphi}^H(\omega_1)+\hat{\varphi}^H(\omega_2)-\hat{\varphi}^H(-\omega_3)\\ &\quad -\hat{\varphi}^H(\omega_1+\omega_2+\omega_3).\end{aligned} \tag{6}$$

The minimum is obtained when

$$\hat{\varphi}^H(\omega) = \frac{1}{N^2}\sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\psi_4(\omega,\omega_1,\omega_2)+K, \tag{7}$$

where K is an arbitrary constant. However, due to the divisions in (7), the value of $\psi_4(\omega_1,\omega_2,\omega_3)$ must be known in the interval $[-4\pi;4\pi[$, but from the measurements, this value is only known in $[-\pi;\pi[$. In section 5, we will give different algorithms to obtain the unwrapped phase from its wrapped value.

4. IDENTIFICATION AND KURTOSIS MAXIMIZATION

This section shows that the previous least-squares identification is in first approximation equivalent to the solution given by the kurtosis maximization criterion.

4.1. Kurtosis maximization criterion

The scheme used in the equalization context is given in Figure 4. The criterion was proposed by D. Donoho and

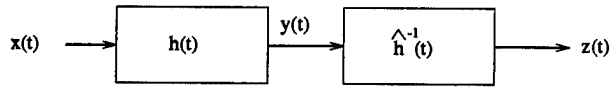


Figure 4. Equalization scheme.

later by O. Shalvi & E. Weinstein [14] [15] [16] in order to recover the input sequence $x(t)$. They estimate $\hat{h}^{-1}(t)$ through the maximization of $|K(z)|$ under the constraint $E\left\{|z|^2\right\} = E\left\{|x|^2\right\}$, where $K(z)$ is the kurtosis of z. In a first step O. Shalvi & E. Weinstein propose to whiten the output signal ; by this way, they reconstruct the channel Fourier transform phase as in the reconstruction algorithms in the frequency domain.

4.2. Expression of the Kurtosis maximization in the frequency domain

The kurtosis of the output sequence $z(t)$ is given by :

$$K(z) = C_4^z(0,0,0) = \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1} T^Z(\omega_1,\omega_2,\omega_3)$$
$$= \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1} \left|T^Z(\omega_1,\omega_2,\omega_3)\right| e^{j\psi_4^Z(\omega_1,\omega_2,\omega_3)} \quad (8)$$

Since $K(z)$ is a real number, its modulus is :

$$|K(z)| = \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1} \left|T^Z(\omega_1,\omega_2,\omega_3)\right| e^{j\left[\psi_4^Z(\omega_1,\omega_2,\omega_3)-k\pi\right]}, \quad (9)$$

where $k = 1$ if $K(z) < 0$ and $k = 0$ if $K(z) \geq 0$.

- If the output is whitened, $|T^Z|$ is constant.
- $|K(z)|$ being a positive real number, the complex exponentials in (9) are replaced by their real part.

Then the Fourier phase maximizing the kurtosis of the output is the phase which maximizes :

$$J = \frac{|K(z)|}{|T^Z|} = \quad (10)$$
$$\sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1}\cos\left[\psi_4^Y(\omega_1,\omega_2,\omega_3) - \hat{\psi}_4^H(\omega_1,\omega_2,\omega_3) - k\pi\right].$$

- Since the kurtosis sign of z is the kurtosis sign of x, the value of k is the same as in (3); then we can replace $\psi_4^Y(\omega_1,\omega_2,\omega_3) - k\pi$ by $\psi_4(\omega_1,\omega_2,\omega_3)$ (cf. 4).

Finally, (10) becomes :

$$J = \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1}\cos\left[\psi_4(\omega_1,\omega_2,\omega_3) - \hat{\psi}_4^H(\omega_1,\omega_2,\omega_3)\right]. \quad (11)$$

4.3. Taylor expansion of the kurtosis

If the trispectrum is factorizable and its phase is accurately estimated, the difference between $\psi_4^Y(\omega_1,\omega_2,\omega_3)$ and $\hat{\psi}_4^H(\omega_1,\omega_2,\omega_3)$ will be small and we can expand eq. (11) :

$$J = \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1}\left[1 - \frac{1}{2}\left[\psi_4(\omega_1,\omega_2,\omega_3) - \hat{\psi}_4^H(\omega_1,\omega_2,\omega_3)\right]^2 + \frac{1}{24}\left[\psi_4(\omega_1,\omega_2,\omega_3) - \hat{\psi}_4^H(\omega_1,\omega_2,\omega_3)\right]^4 + \ldots\right]. \quad (12)$$

If we limit the development to the second term, the maximization of this criterion is reduced to the minimization of the LS criterion obtained in the frequency domain (5). Under this hypothesis, kurtosis maximization is reduced to the minimization of a quadratic criterion for which the analytic solution is known (7). However it needs a phase unwrapping which has a solution if the trispectrum phase estimate is accurate enough.

5. PHASE UNWRAPPING ALGORITHMS

Some phase unwrapping techniques are described in this section:

- the first solution extend Marron's bispectrum phase unwrapping to the trispectrum case;
- the second method uses in a first step a recursive algorithm [17] to obtain a first approximation of the Fourier phases $\hat{\varphi}^H(\omega)$. The modulo 8π trispectrum phases, obtained from these phases using (4) give the interval where must lie the unwrapped value of $\psi_4(\omega_1,\omega_2,\omega_3)$;
- finally a method combining a recursive algorithm and the optimal method of section 3 is presented.

5.1. Extension of Marron's bispectrum phase unwrapping algorithm

Marron *et al.* [9] have shown that it is possible to deduce the N^2 unwrapped bispectrum phases $\psi_3(\omega_1,\omega_2)$ from the $N-1$ modulo 2π bispectrum phases :

$$\psi_3(1,\omega), \omega = 1,2,\ldots,N-1.$$

To obtain the unwrapped phases, they use the following equation:

$$\psi_3(\omega_1,\omega_2) = \psi_3(\omega_1-1,\omega_2+1) + \psi_3(1,\omega_2) - \psi_3(1,\omega_1-1). \quad (13)$$

The Marron's algorithm is extended to the fourth order spectrum of complex signals as follows [18]:

$$\psi_4(z+v+w,x,y)+\psi_4(w,v,z)= \psi_4(w,x,y)+\psi_4(w+x+y,v,z), \quad (14)$$

for all v,w,x,y,z.
This formula can be generalized to any order. It is always a four terms identity.
It is possible to deduce all the N^3 unwrapped trispectrum phases from the $(N-1)$ modulo 2π trispectrum phases used in a multiresolution recursive algorithm using (14)[18].
The trispectrum phases deduced by this unwrapping process are compared to the measured modulo 2π trispectrum phases and a phase of $2q\pi$ is added to these measured values so that their difference will be less than π.
There are other approaches for performing phase unwrapping. The next two methods are also efficient in practice.

5.2. Multiresolution used to unwrap the trispectrum phase

The multiresolution method [17] gives a first approximation of the channel phases $\hat{\varphi}^H(\omega)$. This recursive algorithm does not need any phase unwrapping. Those values are then used to compute all the modulo 8π trispectrum phases in the interval $[-4\pi, 4\pi[$ using equation (4).

5.3. Multiresolution combined with the optimal method

Such a combination is possible thanks to the recursive structure of the multiresolution method [17]. The multiresolution will be combined with the optimal method as follows :
- at step n of the algorithm, the multiresolution method gives a first estimate of $\hat{\varphi}^H\left(\frac{2m+1}{2^n}\right)$ for $m = 0,1,...,2^{n-1}-1$.
- these values, and those calculated in the previous steps (at lower resolutions), are used to unwrap the trispectrum phases : $\psi_4\left(\frac{p}{2^n},\frac{q}{2^n},\frac{r}{2^n}\right)$ for $p,q,r = 0,1,...,2^n-1$ using (4).
- Next, the LS estimation (7) uses these trispectrum phases to give an improved estimation of $\hat{\varphi}^H\left(\frac{2m+1}{2^n}\right)$ for $m = 0,1,...,2^{n-1}-1$.
- Finally, these last estimates will be used to initialize the next step $(n+1)$ of the algorithm.

6. Trispectrum estimation enhancement

The main drawback of the projection method of section 3 is due to the difficulties in performing correctly phase unwrapping when the trispectrum estimation is very noisy. We propose an attempt to treat this problem in improving this trispectrum estimation. We look for the trispectrum phase $\psi_{FACTO}(\omega_1,\omega_2,\omega_3)$ which

- satisfies the factorizability condition [10]:

$$\begin{aligned}\psi_{FACTO}(\omega_1,\omega_2,\omega_3) &= \psi_{FACTO}(\omega_1,x,y) \\ &+\psi_{FACTO}(\omega_1+x+y,\omega_2,\omega_3) \\ &-\psi_{FACTO}(\omega_1+\omega_2+\omega_3,x,y) \quad (15)\end{aligned}$$

$\forall\omega_1,\omega_2,\omega_3$ for a given(x,y),

- and minimizes:

$$\sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1}\sum_{\omega_3=0}^{N-1}\left[\psi(\omega_1,\omega_2,\omega_3)-\psi_{FACTO}(\omega_1,\omega_2,\omega_3)\right]^2. \quad (16)$$

Note that these two conditions (eq. 15 and eq. 16) need only be satisfied for the modulo 2π values of $\psi_{FACTO}(\omega_1,\omega_2,\omega_3)$ and $\psi(\omega_1,\omega_2,\omega_3)$.
The following recursion gives good results in practice : starting from the measured trispectrum phases, we compute recursively :

$$e^{j\,\psi_{new}(\omega_1,\omega_2,\omega_3)} = \quad (17)$$
$$\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left[e^{j\,\psi_{old}(\omega_1,x,y)}e^{j\,\psi_{old}(\omega_1+x+y,\omega_2,\omega_3)}\; e^{-j\,\psi_{old}(\omega_1+\omega_2+\omega_3,x,y)}\right]$$

Moreover, we impose that $\psi_{new}(\omega_1,\omega_2,\omega_3)$ satisfies the trispectrum symmetries (cf. section 2.2). In our simulations, the improved trispectrum was obtained after about 5 iterations.

7. SIMULATION RESULTS

The channel used in the simulations is a 5th order filter. The input was a 4-QAM IID signal. This system is defined by:

$$\begin{aligned}y[k] &= x[k]+0.1\,x[k-1]-1.87\,x[k-2] \\ &+3.02\,x[k-3]-1.435\,x[k-4]+0.49\,x[k-5]\end{aligned}$$

Note this channel has two zeros close the unit circle ($\rho = 0.99$). The trispectrum is estimated from 450 samples of the observed signal $y[k]$, using the direct method. Figure 5 shows the results obtained from the optimal phase reconstruction algorithm of section 3 using the phase unwrapping method of section 5.3. This poor result is due to the errors in the phase unwrapping step. Indeed, the trispectrum estimation computed from so few data is very noisy. In figure 6, we first apply the trispectrum enhancement method before the phase reconstruction step. This result illustrates the advantage of using such an enhancement method. Finally figure 7, shows the mean results and the standard deviations obtained from 50 Monte Carlo runs obtained in the same conditions than the results of figure 6. As the simulation results show, this method yields good results, even when few data are used and errors on the trispectrum phase exceeds $\pi/4$.

8. CONCLUSIONS

The reconstruction combining phase unwrapping and projection formula (7) gives acceptable results, at least when phase unwrapping works well. However, if the phase is not correctly unwrapped, the quality of the reconstruction reduces. It seems that in this case the research of an optimal solution under factorizability constraint as proposed in section 6 improves the results significantly even when the amount of data is less than one thousand signal samples. This seems to confirm that the principal difficulty in blind identification based on higher order spectral analysis lies essentially in accurate phase unwrapping.

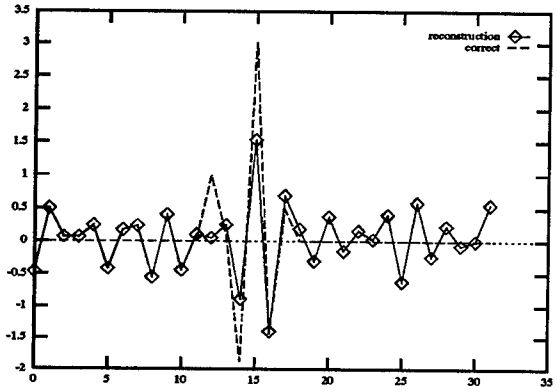


Figure 5. Impulse response reconstruction using the combination method.

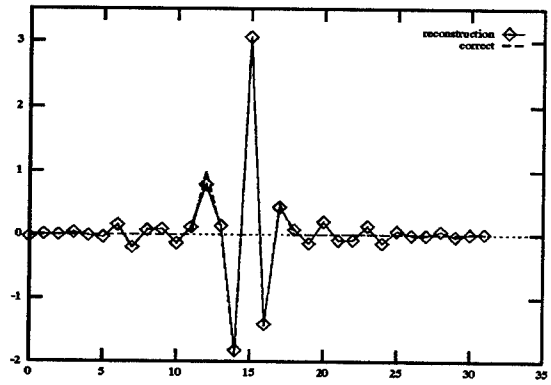


Figure 6. Impulse response reconstruction using first the trispectrum enhancement method, and then the combination method.

References

[1] J.M. Mendel. Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory : Theoretical Results and Some Applications. *Proceedings of the IEEE*, 79(3): 278-305, March 1991.

[2] G. Giannakis and J.M. Mendel. Identification of non-minimum phase systems using Higher order statistics. *IEEE Trans. On ASSP*, 37(3): 360-377, March 1989.

[3] R. Pan and C.L. Nikias. Phase Reconstruction in the Trispectrum Domain. *IEEE Trans. On Signal Processing*, 35(6): 895-897, June 1987.

[4] J. Le Roux, P. Sole, A.M. Tekalp and A.T. Erdem. Tekalp-Erdem estimator gives the LS estimate for Fourier phase and log-Fourier modulus.

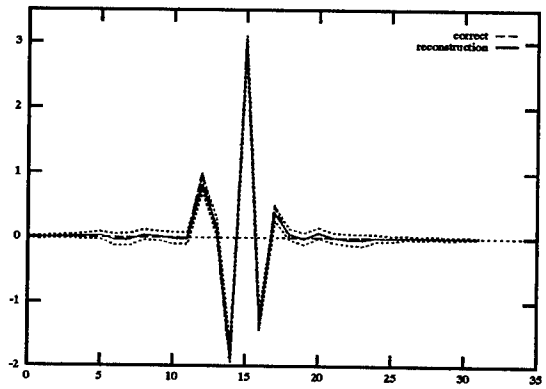


Figure 7. Mean results and standard deviation obtained from 50 Monte Carlo runs

[5] J. Le Roux and P. Sole. Least-squared error reconstruction of a sampled signal Fourier transform from its n-th order polyspectrum. *Signal Processing*, 35: 75-81, 1994.

[6] R. Pierce. Application of Higher Order Spectra to High Resolution Radar Measurement. *IEEE Signal Processing ATHOS Workshop on HOS*, pages 207-210, Chamrousse, France, July 1991.

[7] R. Pierce. Velocity Measurement of Radar Target using HOS. *IEEE Signal Processing ATHOS Workshop on HOS*, pages 164-167, Lake Tahoe, June 1993.

[8] O. Shalvi and E. Weinstein. Super-Exponential methods for blind deconvolution. *In Proc. SPIE 1991 Adaptative Signal Processing*, 1565: 143-152, January 1991.

[9] J.C. Marron, P.P. Sanchez and R.C. Sullivan. 'Unwrapping algorithm for least-squares phase recovery from the modulo 2π bispectrum phase. *JOSA A*,7:14-20, January 1990.

[10] J. Le Roux and C.Huet. Factorizability of Complex Signals Higher Order Spectra : A Necessary and Sufficient Condition. *IEEE Signal Processing ATHOS Workshop on HOS*, Banff, Canada, July 1997.

[11] P. Comon. Circularité et signaux aléatoires à temps discrets. *Traitement du signal*, 11(5):417-420, 1994.

[12] B. Picinbono. On circularity. *IEEE Trans. On Signal Processing*, 42(12): 3473-3482, December 1994.

[13] P.O. Amblard, M.Geata and J.L.Lacoume. Statistics for complex variables and signals - PartII : Signals. *Signal Processing*,53: 15-25,1996.

[14] D. Donoho. On Minimum Entropy Deconvolution. *Applied Time Series Analysis II,D.F.Fiendley ed. Academic Press, New York*, pages 565-608, 1981.

[15] O. Shalvi and E. Weinstein. New criteria for blind deconvolution of nonminimum phase systems (channels). *IEEE Trans. On Information theory*, 36(2):312-321, March 1990.

[16] J. Tugnait. Comments on "New criteria for blind deconvolution of nonminimum phase systems (channels)". *IEEE Trans. On Information theory*, 38(1): 210-213, January 1992.

[17] J. Le Roux, D. Rossille and C.Huet. A multiresolution extension of Lohman Weigelt and Wirnitzer recursion for computing a Fourier transform phase from the third order spectrum phase. *IEEE Signal Processing ATHOS Workshop on HOS*, pages 315-319, Begur, Spain, June 1995.

[18] C. Huet and J. Le Roux. Some properties and algorithms for fourth order spectral analysis of complex signals. *Proceedings of EUSIPCO 1996*, pages 1441-1444, Trieste, Italy, Sept., 1996.

Fractionally-Spaced Signal Reconstruction Based on Maximum Likelihood*

Boaz Porat
Department of Electrical Engineering,
Technion—Israel Institute of Technology, Haifa 32000, Israel.

Benjamin Friedlander
Department of Electrical and Computer Engineering,
University of California, Davis, CA 95616, USA.

Abstract

This paper proposes a scheme for maximum-likelihood-based signal reconstruction. The scheme extends a previous work by Yellin and Friedlander to the case of fractionally-spaced data. The use of fractionally-spaced data obviates the need for timing-phase recovery. Batch and adaptive algorithms are derived and illustrated by examples. The algorithms are useful for equalization of digital communication channels.

1. Introduction

Yellin and Friedlander, in [1], [2], and [3] presented a family of algorithms for multi-channel blind signal reconstruction based on maximum likelihood. The underlying assumption in these works is that the source signals are i.i.d., mutually independent, and the received signals are sampled at the symbol rate. In this paper we extend the work [1] to the case of sampling at higher than the symbol rate.

The main advantage of sampling at higher than the symbol rate is reduced sensitivity to timing phase errors. When sampling at the symbol rate, the sampling instants must coincide with the most open point of the eye. Usually, a communication channel has bandwidth larger than half the symbol rate (so-called excess bandwidth), so a sampling rate equal to the symbol rate voilates Nyquist's rule. Elimination (or minimization) of intersymbol interference then depends critically on accurate knowledge of the channel's impulse (or frequency) response and on accurate timing. On the other hand, when the sampling rate is increased to meet Nyquist's condition, no information is lost in the sampling, therefore we expect to be able to reconstruct the source signal more reliably. In most practical cases, the bandwidth excess is less than 100 percent, so doubling the sampling rate (i.e., sampling twice per symbol) is sufficient.

In this paper we consider only the case of a single source and a single receiver. Although the probability distribution of the source is not restricted in principle, in practice we are interested only in discrete distributions, typical of communication signals. We propose a fractionally-spaced equalizer based on maximum likelihood. The equalizer's coefficients are estimated directly, without explicit estimation of the channel response. We derive both batch and adaptive versions of the algorithm. Study of the multi-source multi-receiver case (the so-called *signal separation* problem) is left to a future paper.

2. The Signal Model

We use the variable t to denote continuous time, the integer n to denote the time index of signals sampled at the symbol rate, and the integer n' to denote signals sampled at L times the symbol rate, where L is the oversampling factor. Parentheses are used for continuous time and square brackets for discrete time. Let $s[n]$ denote the transmitted symbol sequence, T the symbol interval, and $c(t)$ the continuous-time response of the channel to a discrete-time unit sample. Both $s[n]$ and $c(t)$ are complex valued. The sequence $s[n]$ is assumed to be zero-mean i.i.d. The continuous-time received

*This work was supported by the Office of Naval Research under Contract No. N00014-95-1-0912 and by the University of California MICRO program and Applied Signal Technology, Inc.

0-8186-8005-9/97 $10.00 © 1997 IEEE

signal is given by

$$y(t) = \sum_{n=-\infty}^{\infty} s[n]c(t - t_0 - nT), \tag{1}$$

where t_0 is an unknown delay. Here and throughout the derivation we assume that there is no noise.

The discrete-time (oversampled) received signal is given by

$$y[n'] = y(n'T/L) = \sum_{n=-\infty}^{\infty} s[n]c((n' - nL)(T/L) - t_0)$$

$$= \sum_{n=-\infty}^{\infty} s[n]c[n' - nL], \tag{2}$$

where

$$c[n'] \triangleq c(n'(T/L) - t_0) \tag{3}$$

is the equivalent discrete-time channel. Note that the discrete-time channel depends on the delay t_0.

3. The Fractionally-Spaced Equalizer

Let us split the received signal $y[n']$ into L parallel subchannels:

$$y^{(l)}[n] = y[nL + l], \quad 0 \le l \le L - 1. \tag{4}$$

Let the equalizer consist of L "subequalizers", where the lth subequalizer has coefficients $\{h^{(l)}[k],\ 0 \le k \le 2K\}$. We denote by $\boldsymbol{\theta}$ the set of real parts and imaginary parts of all $h^{(l)}[k]$ (this set therefore contains $2(2K + 1)L$ elements). The outputs of the subequalizers are defined as

$$u^{(l)}[n] = \sum_{k=0}^{2K} h^{(l)}[k]y^{(l)}[n - k]. \tag{5}$$

The reconstructed symbol sequence is, by definition,

$$u[n] = \sum_{l=0}^{L-1} u^{(l)}[n]. \tag{6}$$

We regard the signal $u[n]$ as an estimate of $s[n-K]$; in other words, ideally we would like to get $u[n] = s[n-K]$ after convergence.

If the equalizer were ideal, $u[n]$ would be equal to $s[n]$, therefore it would be independent of past and future values of arbitrary functions of $u[n]$, that is, we would have for all $\tau \neq 0$

$$E\{f(u[n])\bar{u}[n - \tau]\} = E\{f(u[n])\}E\{\bar{u}[n - \tau]\} = 0, \tag{7}$$

for any function $f(\cdot)$ (where $\bar{u}$ denotes conjugation of u). In [1], the main idea is to force the condition (7) to hold, at least approximately, for a user-chosen nonlinear function $f(\cdot)$. Different functions give rise to different equalizers. Choosing $f(u[n]) = u[n]$ is equivalent to forcing the sequence $u[n]$ to be uncorrelated, therefore the resulting equalizer belongs to the family of equalizers based on second-order statistics. As is well known, such equalizers cannot, in general, contend with nonminimum-phase channels. This explains why $f(\cdot)$ must be a nonlinear function. In particular, if $s[n]$ has a continuous and differentiable probability density function $p(s)$ and we choose

$$f(s) = \frac{1}{p(s)}\left(\frac{\partial p(s)}{\partial \Re\{s\}} + j\frac{\partial p(s)}{\partial \Im\{s\}}\right), \tag{8}$$

then the solution of (7) can be regarded as an approximate maximum likelihood estimation of the equalizer coefficients [1]. The function $f(s)$ in (8) is the *score function* of $p(s)$. For discrete distributions, a particularly attractive choice is

$$f(s) = s - \langle s \rangle, \tag{9}$$

where $\langle s \rangle$ is the symbol nearest to s in the source alphabet. The function $f(s)$ in (9) is an approximation of the score function of a Gaussian mixture density, where the individual Gaussians are centered around the discrete symbol values and their widths approach zero; see [4]. In practice, it reaches the best asymptotic relative efficiency under noise-free conditions; see [2].

The main idea in this work is to replace (7) by the stronger condition

$$E\{f(u[n])\bar{u}^{(l)}[n - \tau]\} = 0,$$
$$\text{for all } \tau \neq 0 \text{ and } 0 \le l \le L - 1. \tag{10}$$

Our premise is that, if we can force (10) to hold, at least approximately, then (7) will hold a fortiori. In practice, we will consider (10) only for $|\tau| \le Q$, where Q is a user-chosen parameter satisfying $Q \ge K$. We assume that enough data $y[n']$ are available to generate $f(u[n])$ and $u^{(l)}[n - \tau]$ in the range $1 \le n \le N$ for all $-Q \le \tau \le Q$. We define

$$R^{(l)}[\tau] = \frac{1}{N}\sum_{n=1}^{N} f(u[n])\bar{u}^{(l)}[n - \tau]$$
$$- E\{f(u[n])\bar{u}^{(l)}[n]\}\delta[\tau],$$
$$0 \le l \le L - 1, \quad -Q \le \tau \le Q. \tag{11}$$

Our aim is to cause the $R^{(l)}[\tau]$ to be as close to zero as

possible. To this end, we will introduce a cost function

$$V = \frac{1}{2}\sum_{l=0}^{L-1}\sum_{\tau=-Q}^{Q} |R^{(l)}[\tau]|^2, \tag{12}$$

and seek to minimize V as a function of the subequalizers coefficients $h^{(l)}[k]$.

Batch-type algorithms for minimization of the cost function V can be derived using standard optimization techniques: steepest descent, Newton-Gauss, quasi-Newton, and so on. Such optimization methods rely on analytic expressions for the partial derivatives of the cost function with respect to the parameters. We now derive expressions for these derivatives. Differentiation of the cost function with respect to an arbitrary parameter θ gives

$$\frac{\partial V}{\partial \theta} = \sum_{l=0}^{L-1}\sum_{\tau=-Q}^{Q} \Re\left\{\bar{R}^{(l)}[\tau]\frac{\partial R^{(l)}[\tau]}{\partial \theta}\right\}, \tag{13}$$

where

$$\frac{\partial R^{(l)}[\tau]}{\partial \theta} = \frac{1}{N}\sum_{n=1}^{N}\frac{\partial f(u[n])}{\partial \theta}\bar{u}^{(l)}[n-\tau] + f(u[n])\frac{\partial \bar{u}^{(l)}[n-\tau]}{\partial \theta}. \tag{14}$$

In the complex case, $f(u)$ is not always a differentiable (analytic) function of u; however, when it is not, it is usually differentiable with respect to u and $\bar{u}$ seperately. For example, the function $f(u) = |u|^2 = u\bar{u}$ is not analytic in u, but we have

$$\frac{\partial f(u)}{\partial u} = \bar{u}, \quad \frac{\partial f(u)}{\partial \bar{u}} = u.$$

Denote by $g_1(u)$ the partial derivative of $f(u)$ with respect to u and by $g_2(u)$ its partial derivative with respect to $\bar{u}$. Then

$$\frac{\partial R^{(l)}[\tau]}{\partial \theta} = \frac{1}{N}\sum_{n=1}^{N}\left\{g_1(u[n])\frac{\partial u[n]}{\partial \theta} g_2(u[n])\frac{\partial \bar{u}[n]}{\partial \theta}\right\}\bar{u}^{(l)}[n-\tau] + f(u[n])\frac{\partial \bar{u}^{(l)}[n-\tau]}{\partial \theta}. \tag{15}$$

Note that, for $f(\cdot)$ given in (9), $g_1(\cdot) \equiv 1$ and $g_2(\cdot) \equiv 0$.

Let

$$h^{(m)}[k] = h_{\mathrm{r}}^{(m)}[k] + j h_{\mathrm{i}}^{(m)}[k].$$

Then

$$\frac{\partial u^{(l)}[n-\tau]}{\partial h_{\mathrm{r}}^{(m)}[k]} = y^{(m)}[n-\tau-k]\delta[l-m], \tag{16a}$$

$$\frac{\partial u^{(l)}[n-\tau]}{\partial h_{\mathrm{i}}^{(m)}[k]} = jy^{(m)}[n-\tau-k]\delta[l-m], \tag{16b}$$

and

$$\frac{\partial u[n]}{\partial h_{\mathrm{r}}^{(m)}[k]} = y^{(m)}[n-k], \tag{17a}$$

$$\frac{\partial u[n]}{\partial h_{\mathrm{i}}^{(m)}[k]} = jy^{(m)}[n-k]. \tag{17b}$$

Substitute (16) and (17) in (15) to get

$$\frac{\partial R^{(l)}[\tau]}{\partial h_{\mathrm{r}}^{(m)}[k]} = \frac{1}{N}\sum_{n=1}^{N}\{g_1(u[n])y^{(m)}[n-k] + g_2(u[n])\bar{y}^{(m)}[n-k]\}\bar{u}^{(l)}[n-\tau] + f(u[n])\bar{y}^{(m)}[n-\tau-k]\delta[l-m], \tag{18a}$$

$$\frac{\partial R^{(l)}[\tau]}{\partial h_{\mathrm{i}}^{(m)}[k]} = \frac{j}{N}\sum_{n=1}^{N}\{g_1(u[n])y^{(m)}[n-k] - g_2(u[n])\bar{y}^{(m)}[n-k]\bar{u}^{(l)}[n-\tau] - f(u[n])\bar{y}^{(m)}[n-\tau-k]\delta[l-m]. \tag{18b}$$

4. An Adaptive Fractionally-Spaced Equalizer

An adaptive equalizer enables tracking of the parameters of a time-varying channel. In our case, since the equivalent discrete-time channel depends on the delay t_0, an adaptive equalizer also enables tracking of a variable delay, for example due to source or receiver motion. The adaptive equalizer we derive in this section is based on the stochastic gradient approach. More complex algorithms, based on stochastic quasi-Newton, can be derived but are not discussed here.

We assign the subscript "old" to quantities based on data up to time $n-1$, and the subscript "new" to quantities based on data up to time n. In particular, we define

$$R_{\mathrm{new}}^{(l)}[\tau] = \lambda R_{\mathrm{old}}^{(l)}[\tau] + f(u[n])\bar{u}^{(l)}[n-\tau] - E\{f(u[n])\bar{u}^{(l)}[n]\}\delta[\tau], \quad 0 \le l \le L-1, \quad -Q \le \tau \le Q, \tag{19}$$

where $\lambda < 1$ for exponentially decaying memory, and $\lambda = 1$ for growing memory. Correspondingly,

$$V_{\text{new}} = \frac{1}{2} \sum_{l=0}^{L-1} \sum_{\tau=-Q}^{Q} |R_{\text{new}}^{(l)}[\tau]|^2, \tag{20}$$

The partial derivatives of V_{new} are given by

$$\frac{\partial V_{\text{new}}}{\partial \theta} = \sum_{l=0}^{L-1} \sum_{\tau=-Q}^{Q} \Re \left\{ \bar{R}_{\text{new}}^{(l)}[\tau] \frac{\partial R_{\text{new}}^{(l)}[\tau]}{\partial \theta} \right\}. \tag{21}$$

An assumption commonly made in deriving adaptive algorithms is that the algorithm has already converged to a stationary point at time $n-1$, meaning that

$$\frac{\partial V_{\text{old}}}{\partial \theta} = \sum_{l=0}^{L-1} \sum_{\tau=-Q}^{Q} \Re \left\{ \bar{R}_{\text{old}}^{(l)}[\tau] \frac{\partial R_{\text{old}}^{(l)}[\tau]}{\partial \theta} \right\} = 0. \tag{22}$$

We also assume that $R_{\text{old}}^{(l)}$ and $R_{\text{new}}^{(l)}$ are sufficiently close to each other, so that (22) implies

$$\frac{\partial V_{\text{old}}}{\partial \theta} \approx \sum_{l=0}^{L-1} \sum_{\tau=-Q}^{Q} \Re \left\{ \bar{R}_{\text{new}}^{(l)}[\tau] \frac{\partial R_{\text{old}}^{(l)}[\tau]}{\partial \theta} \right\} \approx 0. \tag{23}$$

When we substitute (23) and (19) in (21) we get

$$\frac{\partial V_{\text{new}}}{\partial \theta} = \sum_{l=0}^{L-1} \sum_{\tau=-Q}^{Q} \Re \left\{ \bar{R}_{\text{new}}^{(l)}[\tau] \left[\left(g_1(u[n]) \frac{\partial u[n]}{\partial \theta} + g_2(u[n]) \frac{\partial \bar{u}[n]}{\partial \theta} \right) \bar{u}^{(l)}[n-\tau] + f(u[n]) \frac{\partial \bar{u}^{(l)}[n-\tau]}{\partial \theta} \right] \right\}. \tag{24}$$

The partial derivatives of $u[n]$ and $u^{(l)}[n-\tau]$ are as given in (17) and (16), respectively.

The parameter vector $\boldsymbol{\theta}$ is updated using

$$\boldsymbol{\theta}_{\text{new}} = \boldsymbol{\theta}_{\text{old}} - \mu[n] \frac{\partial V_{\text{new}}}{\partial \boldsymbol{\theta}}, \tag{25}$$

where $\mu[n] = \mu_0/n$ in the growing memory case and $\mu[n] = \mu_0$ in the decaying memory case (in either case, μ_0 is a user-chosen parameter).

In summary, the adaptive algorithm performs the following computations at each time point:

1. Computation of $\{u^{(l)}[n],\ 0 \le l \le L-1\}$ and $u[n]$, using (5) and (6), respectively, with the current values of $h^{(l)}[k]$.
2. Updating the $R^{(l)}[\tau]$, using (19).
3. Computation of the new instantaneous gradient, using (24), (17) and (16).
4. Updating the equalizer coefficients, using (25).

5. Examples

We illustrate the adaptive equalizer by two examples. The purpose of these examples is to show the main advantage of using a fractionally-spaced equalize, namely: obviating the need for timing recovery. The source signal in both examples is 16-QAM. The channel in has a raised-cosine impulse response with 40 percent bandwidth excess. When the output of such a channel is sampled at the symbol rate, it is ISI-free if the sampling instants are properly timed. In the first example, the oversampling ratio is $L = 2$ and the sampling instants are delayed by $0.125T$ and $0.625T$ with respect to the ideal timing instant. The corresponding coefficients of the equivalent discrete-time channel (truncated to length 14) are:

$$\begin{aligned} c[n'] = &-0.0063, 0.0088, 0.0469, -0.0368, -0.1593, \\ &0.1239, 0.7678, 0.9722, 0.4436, -0.0891, -0.1191, \\ &0.0270, 0.0324, -0.0053. \end{aligned}$$

The parameters of the equalizer were chosen as

$$K = 6, \quad Q = 8, \quad \mu = 10^{-6}, \quad \lambda = 0.999.$$

The initial equalizer coefficients where chosen as $h^{(0)}[K] = h^{(1)}[K] = 0.5$, and $h^{(l)}[k] = 0$ otherwise.

Figure 1 shows the equalized symbol constellation during four stages of the equalizer's operation. Part a shows the first 500 symbols; part b shows the symbols at time points 1501 through 2000; part c shows the symbols at time points 4501 through 5000; finally, part d shows the symbols at time points 9501 through 10000. As we see, the equalizer indeed converges from an almost closed-eye initial state to an almost open-eye final state.

The second example compares the fractionally-spaced equalizer with a symbol-rate equalizer based on the same optimization criterion. In this example we added white noise with standard deviation 0.2 (this represents an SNR of about 20 dB). When the sampling instants coincide with the most open point of the eye, both equalizers perform equally well. However, when there is a delay of $0.25T$, the symbol-rate equalizer collapses. This is shown in parts a and b of Figure 2. Part a shows the first 500 symbols at the output of the symbol-rate equalizer and part b shows the symbols at time points 9501 through 10000. As we see, the symbol-rate equalizer does not converge in this case. Parts c and d show the corresponding symbols at the output of the fractionally-spaced equalizer, when the sampling instants are $0.25T$ and $0.75T$. As we see, the fractionally-spaced equalizer properly converges to an open-eye state.

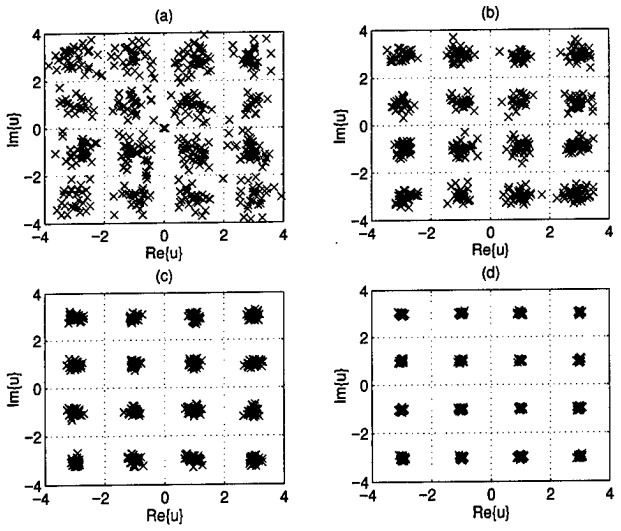


Figure 1. The equalized symbol constellations in the first Example: (a) time points 1–500; (b) time points 1501-2000; (c) time points 4501-5000; (d) time points 9501-10000.

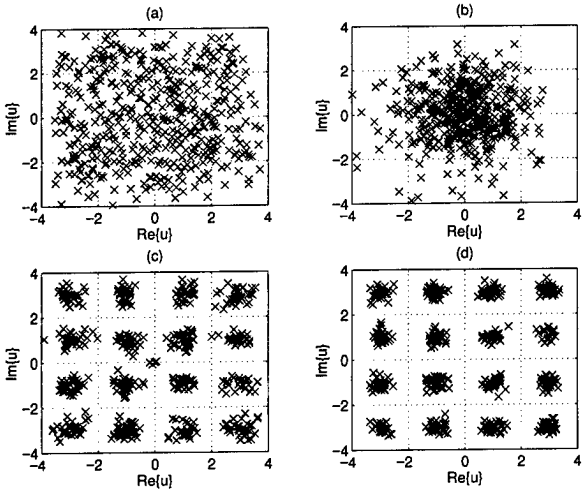


Figure 2. The equalized symbol constellations in the second Example: (a) symbol-rate equalizer, time points 1–500; (b) symbol-rate equalizer, time points 9501-10000; (c) fractionally-spaced equalizer, time points 1-500; (d) fractionally-spaced equalizer, time points 9501-10000.

References

[1] D. Yellin and B. Friedlander, "A Unified Maximum Likelihood Approach to Multi-Channel Signal Reconstruction", submitted for publication.

[2] D. Yellin and B. Friedlander, "Multi-Channel System Identification and Deconvolution: Performance Bounds", submitted for publication.

[3] D. Yellin and B. Friedlander, "A Maximum Likelihood Approach to Blind Separation of Narrowband Digital Communication Signals", *Proc. 30th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, November 1996.

[4] S. Bellini and F. Rocca, "Asymptotically Efficient Blind Deconvolution", *Signal Processing*, Vol. 20, pp. 193–209, 1990.

A Class of Blind Deconvolution and Equalization Algorithms for Nonminimum Phase Multi-Input Multi-Output Systems

Chih-Chun Feng and Chong-Yung Chi *
Department of Electrical Engineering
National Tsing Hua University,
Hsinchu, Taiwan, Republic of China
cychi@ee.nthu.edu.tw

Abstract

A unified class of inverse filter criteria using two cumulants, which includes Wiggins' criterion, Donoho's criteria and Tugnait's criteria as special cases, has been proposed by Chi and Wu for blind deconvolution and equalization of real single-input single-output (SISO) linear time-invariant (LTI) systems. In this paper, we extend this class of (single channel) inverse filter criteria to a family of multistage and a family of single stage criteria for deconvolution and equalization of real (or complex) multi-input multi-output (MIMO) LTI systems with only non-Gaussian measurements. It can be shown that the two families of inverse filter criteria lead to perfect equalization for MIMO systems under some conditions. Some simulation results for the optimum inverse filter using gradient type iterative optimization algorithms were provided to support the proposed criteria. Finally, we draw some conclusions.

1. Introduction

Multichannel blind deconvolution and equalization is a problem to estimate a desired signal $\mathbf{u}(n) = [u_1(n), \ldots, u_p(n)]^{\mathrm{T}}$ with only a set of measurements $\mathbf{x}(n) = [x_1(n), \ldots, x_q(n)]^{\mathrm{T}}$ given by the following convolutional model:

$$\mathbf{x}(n) = \mathbf{H}(n) * \mathbf{u}(n) = \sum_{k=-\infty}^{\infty} \mathbf{H}(k)\mathbf{u}(n-k) \quad (1)$$

where $\mathbf{H}(n)$ is the $q \times p$ impulse response matrix sequence of a p-input q-output linear time-invariant (LTI) system. The problem has recently drawn extensive attention in wireless communications, such as mobile communications with asynchronous direct-sequence code-division multiple access (DS-CDMA), data communications over dually polarized multipath channel, and array signal processing for base-station with spatial division multiple access (SDMA).

Higher-order statistics (HOS) and inverse filters have been used for multichannel blind deconvolution and equalization [1-5]. Inouye and Sato [2] extended Shalvi and Weinstein's (single channel) inverse filter criterion [6] to both multistage (MS) and single stage (SS) multichannel inverse filter criteria. Tugnait [3] also extended the single channel inverse filter criteria reported in [7] to MS multichannel inverse filter criteria. These criteria use only second- and third- or fourth-order cumulants of signals. On the other hand, Chi and Wu [8] proposed a class of single channel inverse filter criteria $J_{r,m}$ using an rth-order (even) and an mth-order ($> r$) cumulants of real data. This class includes Wiggins' criterion (associated with $J_{2,4}$) [9], Donoho's criteria (associated with $J_{2,m}$) [10], and Tugnait's (single channel) criteria $J_{2,3}$, $J_{2,4}$ and $J_{4,6}$ [7] as special cases. Note that Shalvi and Weinstein's criteria [6] are actually the same as Donoho's criteria for real data. In this paper, we extend this class of inverse filter criteria $J_{r,m}$ to a family of MS and a family of SS inverse filter criteria for real (or complex) multi-input multi-output (MIMO) LTI systems.

2. Model assumptions and problem formulation

Assume that $\mathbf{x}(n)$, $n = 0, \ldots, N-1$ are a given set of real (or complex) non-Gaussian measurements generated from (1) with the following assumptions:

(A1) The components $u_i(n)$, $i = 1, \ldots, p$, of the input vector $\mathbf{u}(n)$ are real (or complex), zero-mean,

*This work is supported by the National Science Council under Grant NSC 86-2213-E-007-037.

0-8186-8005-9/97 $10.00 © 1997 IEEE

i.i.d. non-Gaussian, and all of them are statistically independent of each other.

(A2) The unknown p-input q-output LTI system

$$\mathcal{H}(z) = \sum_{n=-\infty}^{\infty} \mathbf{H}(n) z^{-n} \tag{2}$$

is real (or complex) stable with possibly nonminimum phase.

(A3) $q \geq p$.

(A4) The MIMO LTI system $\mathcal{H}(z)$ is of full rank on the unit circle, i.e., $rank\{\mathcal{H}(z)\} = p$ for $|z| = 1$.

As mentioned in [3], the assumptions (A3) and (A4) are a set of sufficient conditions for the existence of the stable inverse filter of $\mathcal{H}(z)$.

Assume that $\mathcal{V}(z)$ is a stable $p \times q$ MIMO LTI system and $\mathbf{V}(n)$ is the impulse response matrix sequence of $\mathcal{V}(z)$. Let

$$\begin{aligned} \mathbf{e}(n) &= [e_1(n), \ldots, e_p(n)]^{\mathrm{T}} \\ &= \mathbf{V}(n) * \mathbf{x}(n) = \mathbf{G}(n) * \mathbf{u}(n) \end{aligned} \tag{3}$$

where

$$\mathbf{G}(n) = \mathbf{V}(n) * \mathbf{H}(n) \quad \text{(see (1))} \tag{4}$$

is the $p \times p$ impulse response matrix sequence of the combined overall system, denoted $\mathcal{G}(z)$.

With $\mathcal{V}(z)$ as an estimate for $\mathcal{H}(z)$'s inverse system, the goal of multichannel deconvolution and equalization is to find an optimum $\mathcal{V}(z)$ such that

$$\mathcal{G}(z) = \mathcal{V}(z) \cdot \mathcal{H}(z) = \mathbf{P} \cdot \mathcal{D}(z) \tag{5}$$

(perfect equalization) where $\mathbf{P}$ is a (nonsingular) permutation matrix and $\mathcal{D}(z)$ is a $p \times p$ diagonal matrix given by

$$\mathcal{D}(z) = \mathrm{diag}(\alpha_1 z^{-\tau_1}, \cdots, \alpha_p z^{-\tau_p}) \tag{6}$$

in which $\alpha_i, i = 1, \ldots, p$ are unknown real (or complex) scale factors and τ_i, $i = 1, \ldots, p$ are unknown time delays. Consequently, it can be easily observed from (3) and (5) that the optimum equalized signal

$$\mathbf{e}(n) = [\alpha_{i_1} u_{i_1}(n - \tau_{i_1}), \ldots, \alpha_{i_p} u_{i_p}(n - \tau_{i_p})]^{\mathrm{T}} \tag{7}$$

where $\{i_1, \ldots, i_p\}$ is a permuted sequence of the sequence $\{1, \ldots, p\}$ associated with $\mathbf{P}$.

3. Inverse filter criteria for MIMO LTI systems

For ease of later use, let $h_{ij}(n)$, $v_{ij}(n)$ and $g_{ij}(n)$ denote the (i,j)th elements of the matrices $\mathbf{H}(n)$, $\mathbf{V}(n)$ and $\mathbf{G}(n)$, respectively. Moreover, let

$$C^{e_i}_{l_1,l_2} = \mathrm{CUM}(\underbrace{e_i(n), \ldots, e_i(n)}_{l_1 \text{ terms}}, \underbrace{e_i^*(n), \ldots, e_i^*(n)}_{l_2 \text{ terms}}) \tag{8}$$

denote the $(l_1 + l_2)$th-order cumulant of the ith equalized signal ($i \in \{1, \ldots, p\}$)

$$e_i(n) = g_{i1}(n) * u_1(n) + \cdots + g_{ip}(n) * u_p(n) \tag{9}$$

where the superscript '*' denotes complex conjugation.

The new multichannel inverse filter criteria to be presented below are based on the following theorem:

Theorem 1. Let the $(l_1 + l_2)$th-order cumulant of $u_i(n)$ be $\gamma^{u_i}_{l_1,l_2}$ where $i = 1, \ldots, p$. Assume that all the $(2s)$th-order cumulants of $u_i(n)$, $i = 1, \ldots, p$, have the same sign, i.e.,

$$\mathrm{sign}\{\gamma^{u_1}_{s,s}\} = \mathrm{sign}\{\gamma^{u_2}_{s,s}\} = \cdots = \mathrm{sign}\{\gamma^{u_p}_{s,s}\} \tag{10}$$

Then under (A1) through (A4) and $l_1 + l_2 > 2s \geq 2$,

$$\widetilde{J}_i \triangleq \frac{|C^{e_i}_{l_1,l_2}|^{2s}}{|C^{e_i}_{s,s}|^{l_1+l_2}} \tag{11}$$

$$\leq \kappa_{\max} \triangleq \max\{\kappa_j,\ j = 1, \ldots, p\} \tag{12}$$

where

$$\kappa_j = \frac{|\gamma^{u_j}_{l_1,l_2}|^{2s}}{|\gamma^{u_j}_{s,s}|^{l_1+l_2}} \tag{13}$$

The equality of (12) holds if and only if

$$g_{ij}(n) = \alpha_i \delta(n - \tau_i)\delta(j - j_0), \quad j_0 \in \{1, \ldots, p\} \tag{14}$$

where α_i is an unknown real (or complex) scale factor, τ_i is an unknown time delay, and j_0 is the index associated with the maximum value of κ_j, $j = 1, \ldots, p$ (see (12)). □

Note that the results presented in Theorem 1 for $(s, l_1, l_2) = (1, 2, 1)$ and $(1, 2, 2)$ have been proposed by Tugnait [3], while those for other choices of (s, l_1, l_2) such as $(1, 3, 1)$, $(1, 3, 2)$, ... are new. Next, based on Theorem 1, two families of multichannel inverse filter criteria, which follows, in part, the ideas proposed by Inouye and Sato [2], are presented for finding the optimum inverse filter $\mathbf{V}(n)$.

A. Family of MS Criteria

Let $\mathbf{v}_i(n)^{\mathrm{T}}$ denote the ith row vector of $\mathbf{V}(n)$. The family of MS criteria, which, at ith stage, try to find the optimum $\mathbf{v}_i(n)$ using $\widetilde{J}_i$ with some uncorrelatedness constraints on the obtained $(i-1)$ inverse filter output processes $e_1(n)$, ..., $e_{i-1}(n)$, is as follows:

Stage 1: Estimation of $\mathbf{v}_1(n)$.

Maximize

$$J_1^{(\mathrm{MS})} = \widetilde{J}_1 \tag{15}$$

where $\widetilde{J}_1$ is given by (11) and $l_1 + l_2 > 2s \geq 2$.

Stage i: Estimation of $\mathbf{v}_i(n)$ for $i \geq 2$.

Maximize

$$J_i^{(\mathrm{MS})} = \widetilde{J}_i - \lambda_i \sum_{k=1}^{i-1} \frac{\sum_{\tau \in Z} |C_{s,s}^{e_i e_k}(\tau)|^2}{|C_{s,s}^{e_i}||C_{s,s}^{e_k}|} \tag{16}$$

where $l_1 + l_2 > 2s \geq 2$, λ_i is a positive real constant, Z is the set of all integers, and

$$C_{s,s}^{e_i e_k}(\tau) = \mathrm{CUM}(\overbrace{e_i(n), \ldots, e_i(n)}^{s \text{ terms}}, \underbrace{e_k^*(n-\tau), \ldots, e_k^*(n-\tau)}_{s \text{ terms}}) \tag{17}$$

The deconvolution and equalization capabilities of the proposed MS criteria are supported by the following theorem:

Theorem 2. The MS criteria given by (15) and (16) lead to a solution for $\mathcal{G}(z)$ that satisfies (5) (perfect equalization), provided that λ_i given by (16) are chosen such that $\lambda_i \geq \kappa_{\max}$ (see (12)). □

B. Family of SS Criteria

Again, with some uncorrelatedness constraints on the inverse filter output processes, the family of SS criteria simultaneously estimates $\mathbf{v}_1(n)$, $\mathbf{v}_2(n)$, ..., $\mathbf{v}_p(n)$ by maximizing

$$J^{(\mathrm{SS})} = \sum_{i=1}^{p} \widetilde{J}_i - \lambda \sum_{i=2}^{p} \sum_{k=1}^{i-1} \frac{\sum_{\tau \in Z} |C_{s,s}^{e_i e_k}(\tau)|^2}{|C_{s,s}^{e_i}||C_{s,s}^{e_k}|} \tag{18}$$

where $l_1 + l_2 > 2s \geq 2$ and λ is a positive real constant, and meanwhile their deconvolution and equalization capabilities are supported by the following theorem.

Theorem 3. Under the assumption that all κ_i, $i = 1, 2, , ..., p$ (see (13)), are the same, the SS criteria $J^{(\mathrm{SS})}$ lead to a solution for $\mathcal{G}(z)$ that satisfies (5) (perfect equalization). □

Two worthy remarks regarding the proposed MS and SS criteria are given as follows:

(R1) Theorems 2 and 3 are counterparts to the ones proposed by Inouye and Sato [2] for their MS and SS criteria, respectively, while the latter further requires a quite restrictive condition that $\gamma_{1,1}^{u_i} = 1$ for all i.

(R2) With $(s, l_1, l_2) = (1, 2, 1)$ or $(1, 2, 2)$, Tugnait's MS approach [3] obtains the equalized signal $e_i(n)$ as the estimate of an input signal $u_j(n)$ by unconstrained maximization of $\widetilde{J}_i$ (i.e., without any uncorrelatedness constraints on the inverse filter output processes) at the ith stage. However, this approach has to estimate the channel impulse responses $h_{lj}(n)$, $l = 1, \ldots, q$, for each stage i followed by removing the contribution of $u_j(n)$ from the measurements, leading to an MIMO system with q outputs and $(p - i)$ inputs for the next stage. On the other hand, besides many new choices for (s, l_1, l_2), the optimum inverse filter matrix $\mathbf{V}(n)$ is directly estimated by both of the proposed MS and SS criteria without need of estimation of the channel impulse responses for system dimension reduction.

4. Optimization algorithms for MS and SS criteria

To find the optimum inverse filter $\mathbf{V}(n)$ using the proposed MS and SS criteria given by (15), (16) and (18) with a given set of data, the cumulants used in these criteria can be simply replaced by the associated sample cumulants, and a causal FIR filter of order L can be used as an approximation to $\mathbf{V}(n)$. Since the objective function $J_i^{(\mathrm{MS})}$ given by (15) and (16) for stage i is a highly nonlinear function of

$$\mathbf{v}_i \triangleq [v_{i1}(0), \ldots, v_{i1}(L), \ldots, v_{iq}(0), \ldots, v_{iq}(L)]^{\mathrm{T}} \tag{19}$$

and the objective function $J^{(\mathrm{SS})}$ given by (18) is also a highly nonlinear function of

$$\mathbf{v} \triangleq [\mathbf{v}_1^{\mathrm{T}}, \mathbf{v}_2^{\mathrm{T}}, \ldots, \mathbf{v}_p^{\mathrm{T}}]^{\mathrm{T}} \tag{20}$$

We use gradient type iterative optimization algorithms to find the optimum $\mathbf{v}_i$. Specifically, at the kth iteration, $\mathbf{v}_i$ (associated with MS criteria) is updated by

$$\begin{cases} \mathbf{v}_i[k] = \mathbf{v}_i[k-1] + \rho \cdot \left. \dfrac{\partial J_i^{(\mathrm{MS})}}{\partial \mathbf{v}_i^*} \right|_{\mathbf{v}_i = \mathbf{v}_i[k-1]} \\ \mathbf{v}_i[k] = \mathbf{v}_i[k] / ||\mathbf{v}_i[k]|| \end{cases} \tag{21}$$

where ρ is a positive real constant. Note that the second line of (21) (i.e., normalization for $\mathbf{v}_i[k]$) is due to the fact that $J_i^{(\mathrm{MS})}$ is invariant to any scaled version of $\mathbf{v}_i$ (see (14)). For SS criteria, $\mathbf{v}$ given by (20) is also updated in a similar way except that $J_i^{(\mathrm{MS})}$ is replaced by $J^{(\mathrm{SS})}$ and normalization operation is performed for each $\mathbf{v}_i$ in $\mathbf{v}$ (see (20)).

5. Simulation results

In this section, let us present two simulation examples to demonstrate the proposed multichannel inverse filter criteria. In the two examples, synthetic noisy data $\mathbf{x}(n)$ for SNR = 20 dB were generated from (1), to which q uncorrelated white Gaussian noises were added.

Example 1. (Real signals)

Let us consider a real 2-input 2-output (*i.e.*, $p = q = 2$) MA(6) system $\mathcal{H}(z)$ (taken from [3]) given by

$$\begin{aligned}
\mathcal{H}_{11}(z) &= 0.6455 - 0.3227z^{-1} + 0.6445z^{-2} \\
&\quad -0.3227z^{-3} \\
\mathcal{H}_{12}(z) &= 0.6140 + 0.3684z^{-1} \\
\mathcal{H}_{21}(z) &= 0.3873z^{-1} + 0.8391z^{-2} + 0.3227z^{-3} \\
\mathcal{H}_{22}(z) &= -0.2579z^{-1} - 0.6140z^{-2} + 0.8442z^{-3} \\
&\quad +0.4421z^{-4} + 0.2579z^{-6}
\end{aligned}$$

The driving inputs $u_i(n)$, $i = 1, 2$, were assumed to be real, zero-mean, Exponentially i.i.d. with $\gamma_{1,1}^{u_i} = 1$, $\gamma_{2,1}^{u_i} = 2$, $\gamma_{2,2}^{u_i} = 6$, and $\gamma_{3,2}^{u_i} = 24$, $i = 1, 2$. The MS criteria for $(s, l_1, l_2) = (1, 2, 1)$, $(1, 2, 2)$ and $(1, 3, 2)$ were used to obtain the inverse filters $\mathbf{v}_i$, $i = 1, 2$, with filter order $L = 14$ and $\lambda_2 = \kappa_{\max}$ (i.e., $\lambda_2 = 4$, 36 and 576 for $(s, l_1, l_2) = (1, 2, 1)$, (1, 2, 2) and (1, 3, 2), respectively) as required by Theorem 2. Thirty independent runs for $N = 2048$ and 4096 were performed.

A performance index MISI defined as (taken from [1])

$$\begin{aligned}
\text{MISI} &= \sum_{i=1}^{p} \frac{\sum_{j=1}^{p}[\sum_n |g_{ij}(n)|^2] - \max\{|g_{ij}(n)|^2, \forall j, n\}}{\max\{|g_{ij}(n)|^2, \forall j, n\}} \\
&\quad + \sum_{j=1}^{p} \frac{\sum_{i=1}^{p}[\sum_n |g_{ij}(n)|^2] - \max\{|g_{ij}(n)|^2, \forall i, n\}}{\max\{|g_{ij}(n)|^2, \forall i, n\}}
\end{aligned} \tag{22}$$

was used as a measure of the multichannel intersymbol interference after equalization. Note that MISI = 0 when $\mathcal{G}(z)$ satisfies (5). Table 1 shows average values of MISI's, denoted < MISI >, which were calculated with the obtained thirty independent estimates of $g_{ij}(n)$. One can see that all the MISI's after equalization are smaller than the MISI before equalization (the top row of Table 1) with the best MISI improvement (around 14 dB) for $(s, l_1, l_2) = (1, 2, 1)$.

Example 2. (Complex signals)

A real 2-input 2-output nonminimum-phase MA(2) system

$$\mathcal{H}(z) = \begin{bmatrix} 1 - 0.3z^{-1} + 0.8z^{-2} & -0.92z^{-1} \\ z^{-2} & 1 - 0.5z^{-1} + 0.2z^{-2} \end{bmatrix}$$

whose zeros are $0.5946 \pm j1.0738$ and $-0.1946 \pm j0.2614$, was used. The input $u_1(n)$ was assumed to be a 8-PSK signal of unity variance and the other input $u_2(n)$ was a 16-QAM signal of unity variance. The MS criteria for $(s, l_1, l_2) = (1, 2, 2)$ were used to obtain the inverse filters $\mathbf{v}_i$, $i = 1, 2$, with filter order $L = 20$ and $\lambda_2 = \kappa_{\max} = 1$ as required by Theorem 2. A single realization was performed for data length $N = 4096$.

The MISI before equalization is 7.8460 dB and the MISI after equalization is -13.6245 dB for this example, i.e., the use of the proposed MS criteria led to around 21 dB improvement in MISI. Moreover, Figures 1(a) and 1(b) show the unequalized signal constellations (i.e., eye patterns) associated with $x_1(n)$ and $x_2(n)$, respectively, for $n = 0 \sim 4095$. Figures 1(c) and 1(d) show the equalized signal constellations associated with $e_1(n)$ and $e_2(n)$, respectively, for $n = 0 \sim 4095$. One can see from these figures that the eye patterns after equalization are open to a sufficient degree.

6. Conclusions

Based on Theorem 1 and Inouye and Sato's ideas, we have extended Chi and Wu's single channel inverse filter criteria to a family of MS and a family of SS criteria for blind deconvolution and equalization of real (or complex) MIMO LTI systems. Furthermore, we proved (Theorems 2 and 3) that under some conditions, both of the proposed MS and SS criteria lead to perfect equalization, and their efficacy was justified through computer simulations.

References

[1] Y. Inouye and T. Habe, "Multichannel blind equalization using second- and fourth-order cumulants," *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, Girona, Spain, June 1995, pp. 96–100.

[2] Y. Inouye and T. Sato, "Unconstrained optimization criteria for blind equalization of multichannel linear systems," *Proc. IEEE 1996 SP Workshop on Statistical Signal and Array Processing*, Corfu, Greece, June 1996, pp. 320–323.

[3] J.K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher-order statistics and inverse filter criteria," to appear in *IEEE Trans. Signal Processing*.

[4] M.K. Tsatsanis, "Inverse filter criteria for CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 102–112, Jan. 1997.

[5] J.K. Tugnait, "Blind spatio-temporal equalization of impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 268–271, Jan. 1997.

[6] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.
[7] J.K. Tugnait, "Estimation of linear parametric models using inverse filter criteria and higher order statistics," *IEEE Trans. Signal Processing*, vol. 41, pp. 3196–3199, Nov. 1993.
[8] C.-Y. Chi and M.-C. Wu, "Inverse filter criteria for blind deconvolution and equalization using two cumulants," *Signal Processing*, vol. 43, no. 1, pp. 55–63, Apr. 1995.
[9] R.A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, pp. 21–35, 1978.
[10] D.L. Donoho, "On minimum entropy deconvolution," *Applied Time Series Analysis II*, D.F. Findly, ed., Academic Press, New York, 1981.

Table 1. Average values of MISI's over thirty independent runs for SNR = 20 dB.

Initial MISI = 9.8293 dB (before equalization)		
(s, l_1, l_2)	< MISI > (in dB)	
	$N = 2048$	$N = 4096$
(1,2,1)	−4.1371	−4.7832
(1,2,2)	0.4316	−2.4964
(1,3,2)	5.8262	2.3185

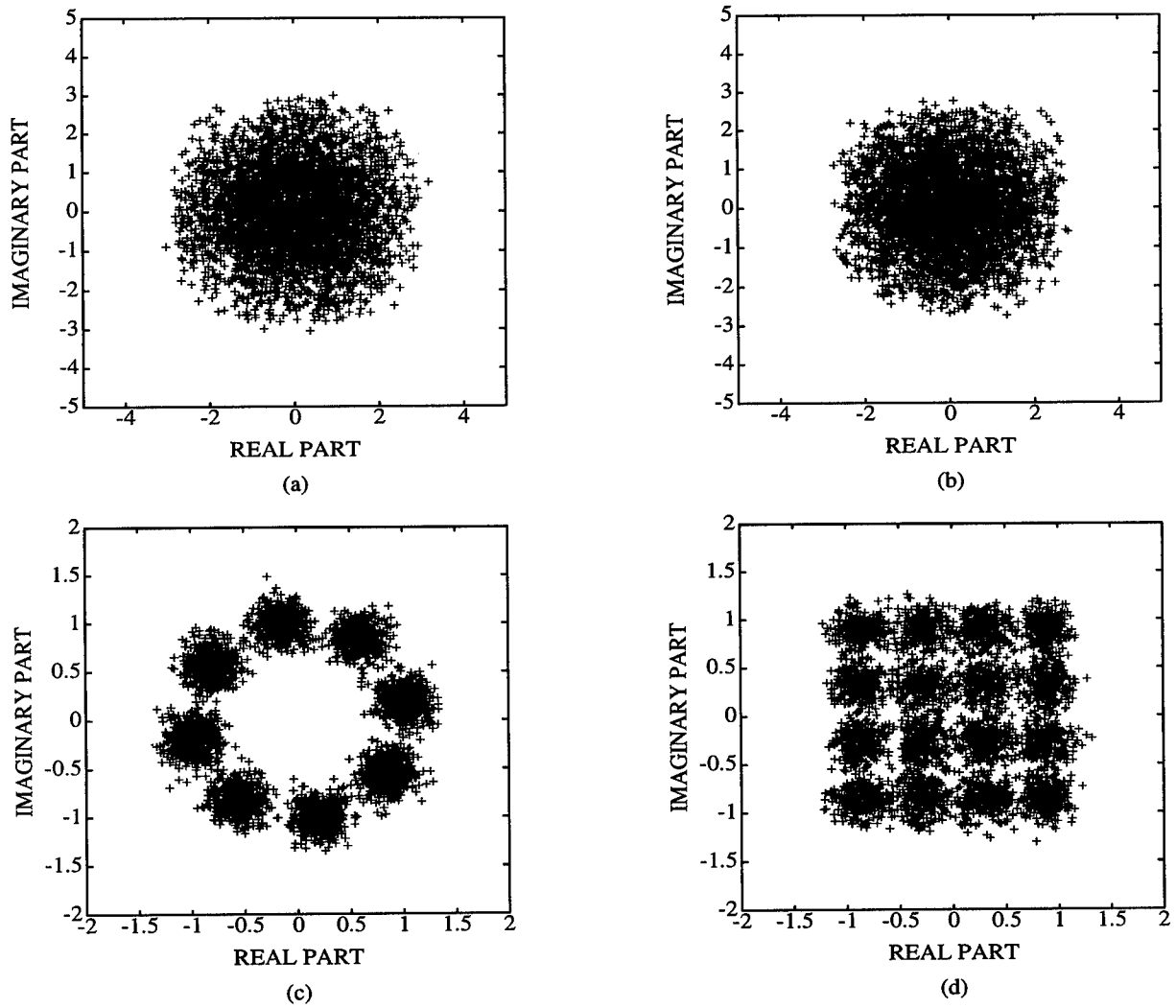


Figure 1. (a) and (b) Unequalized signal constellations associated with $x_1(n)$ and $x_2(n)$, respectively, for $n = 0 \sim 4095$ and SNR = 20 dB; (c) and (d) equalized signal constellations associated with $e_1(n)$ and $e_2(n)$, respectively, for $n = 0 \sim 4095$ using the MS criteria with $(s, l_1, l_2) = (1, 2, 2)$.

On-line Algorithms for Blind Deconvolution of Multichannel Linear Time-Invariant Systems

Yujiro Inouye and Takehito Sato

Department of Systems Engineering,
Faculty of Engineering Science,
Osaka University,
1-3 Machikaneyama, Toyonaka, Osaka 560, Japan
E-mail: inouye@sys.es.osaka-u.ac.jp

Abstract

Blind deconvolution and blind equalization have been important interesting topics in diverse fields including data communication, image processing and geophysical data processing. Recently, Inouye and Habe proposed a constrained multistage criterion for attaining blind deconvolution of multichannel linear time-invariant (LTI) systems [2]. In this paper, based on their constrained criterion, we present an iterative algorithm for solving the blind deconvolution problem of multichannel LTI systems. Inouye and Sato proposed new unconstrained criteria for accomplishing the blind deconvolution of multichannel LTI systems [3]. Based on their unconstrained criteria, we show iterative algorithms for solving the blind deconvolution of multichannel LTI systems. Simulation examples are included to examine the proposed algorithms.

1 Introduction

Blind deconvolution and blind equalization have been important interesting topics in diverse fields including data communication, image processing and geophysical data processing. Recently, Inouye and Habe proposed a constrained multistage criterion for attaining blind deconvolution of multichannel linear time-invariant (LTI) systems [2]. In this paper, based on their constrained criterion, we present an iterative algorithm for solving the blind deconvolution problem of multichannel LTI systems. Inouye and Sato proposed a new unconstrained multistage criterion for accomplishing the blind deconvolution of multichannel LTI systems [3]. Under the assumption that all the magnitudes of the fourth-order auto-cumulants of components of the input vector process are identical, Inouye and Sato also presented a new unconstrained single-stage maximization criterion for attaining the blind deconvolution of multichannel LTI systems [3]. Based on their unconstrained criteria, we show iterative algorithms for solving the blind deconvolution of multichannel LTI systems. Simulation examples are included to examine the proposed algorithms.

In this paper, we use the same notation as in [2], [3].

2 Problem Formulation

Let us consider the system shown Fig. 1. It is a cascade connection of an unknown multichannel system $\boldsymbol{H}$ and a multichannel equalizer $\boldsymbol{W}$. We make the following assumptions on the system and the signals involved.

(A1) The unknown system $\boldsymbol{H}(z)$ is described by

$$\boldsymbol{y}(t) = \sum_{k=-\infty}^{\infty} \boldsymbol{H}(k)\boldsymbol{u}(t-k) \tag{1}$$

where $\boldsymbol{y}(t)$ is a real/complex n-column output vector, $\boldsymbol{u}(t)$ is a real/complex n-column input vector, and $\{\boldsymbol{H}(k)\}$ is a real/complex $n \times n$ matrix sequence called the **impulse response**. The system is stable, that is, the impulse response satisfies the absolute summability condition

$$\sum_{k=-\infty}^{\infty} ||\boldsymbol{H}(k)|| < \infty. \tag{2}$$

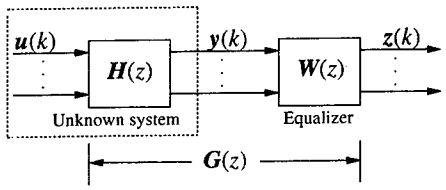


Figure 1. Unknown system and equalizer

0-8186-8005-9/97 $10.00 © 1997 IEEE

(A2) The transfer function defined by

$$\boldsymbol{H}(z) := \sum_{k=-\infty}^{\infty} \boldsymbol{H}(k) z^{-k} \tag{3}$$

is of full rank on the unit circle $|z| = 1$ (this implies it has no zero on the unit circle).

(A3) The input process $\{\boldsymbol{u}(t)\}$ is a zero-mean, non-Gaussian vector process, whose component processes $\{u_i(t)\}$, $i = 1, \cdots, n$, are mutually independent. Moreover, each component process $\{u_i(t)\}$ is an independently and identically distributed (i.i.d.) process with non-zero variance $\sigma_{u_i}^2 \neq 0$ and non-zero fourth-order auto-cumulant $\kappa_{4,u_i} \neq 0$.

(A4) The equalizer $\boldsymbol{W}(z)$ is described by

$$\boldsymbol{z}(t) = \sum_{k=-\infty}^{\infty} \boldsymbol{W}(k) \boldsymbol{y}(t-k) \tag{4}$$

where $\boldsymbol{z}(t)$ is a real/complex n-column vector, called the **equalizer output**, and $\{\boldsymbol{W}(k)\}$ is a real/complex $n \times n$ matrix sequence. It is assumed that the equalizer $\boldsymbol{W}$ is also stable.

There are inherent ambiguities in the solution to the multichannel deconvolution problem as follows: In general, we cannot identify the order of the arrangement of the components $u_1(t), \cdots, u_n(t)$ of input vector $\boldsymbol{u}(t)$, the time origin of each component $u_i(t)$, and the magnitude of each component $u_i(t)$.

Taking these ambiguities into account, the multichannel blind deconvolution problem is formulated such that it is to find an equalizer $\boldsymbol{W}$ so that the transfer function $\boldsymbol{G}(z)$ of the combined system takes the form of

$$\boldsymbol{G}(z) = \boldsymbol{P}\boldsymbol{\Lambda}(z)\boldsymbol{D} \tag{5}$$

where $\boldsymbol{P}$ is a permutation matrix, $\boldsymbol{\Lambda}(z)$ is a diagonal matrix with diagonal entries $\lambda_{ii}(z) = z^{l_i}$, $i = 1, \cdots, n$ (where l_i is an integer), and $\boldsymbol{D}$ is a constant diagonal matrix. Moreover, if we know all the magnitudes of the variances of the components of the input process ahead, we can constrain to make the diagonal matrix $\boldsymbol{D}$ in (5) be equal to a diagonal matrix with the diagonal entries all being unit magnitude.

We require the following notions for stationary random vector processes in this paper. It is said that a stationary random process $\{\boldsymbol{u}(t)\}$ satisfies the **normalized condition** if the variance of each component of the vector process $\{\boldsymbol{u}(t)\}$ is equal to unit. It is said to satisfy the **normalized whitening condition** if all the component processes $\{u_i(t)\}, i = 1, \cdots, n$, of $\{\boldsymbol{u}(t)\}$ are white random processes with unit variance and they are mutually uncorrelated.

3 Blind Deconvolution

To begin with, let us assume that the input process $\{\boldsymbol{u}(t)\}$ satisfies the normalized condition by dividing each component $\{u_i(t)\}$ by the square root of variance $\sigma_{u_i}^2$ to eliminate the magnitude ambiguity.

First we consider the following multistage maximization criterion (A):
(Stage 1) Maximize $|\kappa_{4,z_1}|$ subject to $\sigma_{z_1}^2 = 1$.
(Stage k) Maximize $|\kappa_{4,z_k}|$ subject to $\sigma_{z_k}^2 = 1$ and $r_{z_k,z_i^*}(\tau) = 0$ for all $\tau \in Z$, and all $i = 1, 2, \cdots, k-1$. Here k moves successively from 2 to n.

Then we obtain the following theorem.

Theorem 1[2]: Under the normalized condition of the input process, the multistage maximization criterion (A) gives a solution to the multichannel blind deconvolution problem.

4 Iterative Algorithms

Based on the multistage maximization criterion (A), we develop an iterative algorithm for multichannel blind deconvolution by the gradient or steepest-descent method. At Stage k in the criterion (A), the maximization of each criterion function should be subjected to the following first constraint or the first and second constraints:

(C1) $\sigma_{z_k}^2 = 1$,

(C2) $r_{z_k,z_i^*}(\tau) = 0, \quad \tau \in Z, \quad i = 1, \cdots, k-1$.

The first constraint is the same as the one treated in the single-channel case [1], but the second one appears to be new in the multichannel case. In order to interpret these constraints into constraints for the coefficients (tap-coefficients) of the equalizer, we require a whitening operation for the output process $\{\boldsymbol{y}(t)\}$ ahead. This is also the case for single-channel blind deconvolution [1]. Let $\{\tilde{\boldsymbol{y}}(t)\}$ be a whitened process of $\{\boldsymbol{y}(t)\}$. Such a filter exists, but it is not unique. Let us denote the transfer function of an equalizer for the cascade system of the unknown system and the whitening filter by $\boldsymbol{C}(z)$. We make the following assumption on the equalizer $\boldsymbol{C}$.

(A5) The equalizer $\boldsymbol{C}$ is assumed to be a causal FIR system of sufficient length, so that the truncation effects can be ignored.

Then we can represent the transfer function $\boldsymbol{C}(z) = [c_{kj}(z)]_{k,j=1}^n$ by

$$c_{kj}(z) = \sum_{l=0}^{l_k-1} c_{kj}(l) z^{-l}; \qquad k, j = 1, \cdots, n \tag{6}$$

where $\{c_{kj}(l) : l = 1, \cdots, l_k\}_{k,j=1}^n$ denotes the sequence of the coefficients (or tap-coefficients) of equalizer $\boldsymbol{C}$.

Now we can represent the constraints of **(C1)** and **(C2)** by the tap-coefficients of equalizer $\boldsymbol{C}$. Since $\{\tilde{\boldsymbol{y}}(t)\}$ satisfies the normalized whitening condition, it follows from **(A5)** along with (6) that

$$\sigma_{z_k}^2 = c_{2,z_k,z_k^*}(0) = \|\boldsymbol{c}_k\|^2 \tag{7}$$

where

$$\boldsymbol{c}_k = \left[\boldsymbol{c}_{k1}^T, \cdots, \boldsymbol{c}_{kn}^T\right]^T \in C^{nl_k} \tag{8}$$

$$\boldsymbol{c}_{kj} = \left[c_{kj}(0), \cdots, c_{kj}(l_k - 1)\right]^T \in C^{l_k}. \tag{9}$$

This mean the first condition **(C1)** is equivalent to

$$\|\boldsymbol{c}_k\|^2 = 1. \tag{10}$$

Similarly, it follows from **(A5)** along with (6) that

$$r_{z_k, z_i^*}(-\tau) = \sum_{t=0}^{l_k-1} \{q^\tau c_{i1}^*(t) c_{k1}(t) + \cdots + q^\tau c_{in}^*(t) c_{kn}(t)\};$$
$$\tau \in Z,\ i = 1, \cdots, k-1, \tag{11}$$

where q is the forward shift operator defining by $qc_{ij}^*(t) = c_{ij}^*(t+1)$, and q^{-1} is the backward shift operator defined by $q^{-1}c_{ij}^*(t) = c_{ij}^*(t-1)$.

Let $\boldsymbol{c}_i$ denote the column vector consisting of the coefficients $\{c_{ij}(0), \cdots, c_{ij}(l_i - 1)\}_{j=1}^n$ defined in the same way as the column vector given by (8) and (9). In order to treat the above relations in (11) on a unitary space, we embed the vectors $\boldsymbol{c}_1, \cdots, \boldsymbol{c}_{k-1}$ into the unitary space C^{nl_k} containing $\boldsymbol{c}_k$. This means that $\{l_i\}_{i=1}^n$ is an increasing sequence and that we reset $\boldsymbol{c}_i$ as

$$\boldsymbol{c}_i = \left[\boldsymbol{c}_{i1}^T, \cdots, \boldsymbol{c}_{in}^T\right]^T \in C^{nl_k} \tag{12}$$

$$\boldsymbol{c}_{ij} = \left[c_{ij}(0), \cdots, c_{ij}(l_k - 1)\right]^T \in C^{l_k} \tag{13}$$

where $c_{ij}(\tau) = 0$ for $\tau \geq l_i$. Then it follows from (11)

$$r_{z_k, z_i^*}(-\tau) = (q^\tau \boldsymbol{c}_i)^H \boldsymbol{c}_k = < \boldsymbol{c}_k, q^\tau \boldsymbol{c}_i >, \tag{14}$$

where $<,>$ denotes the inner product on space C^{nl_k}, and $q^\tau \boldsymbol{c}_i$ is defined by

$$q^\tau \boldsymbol{c}_i = \left[q^\tau \boldsymbol{c}_{i1}^T, \cdots, q^\tau \boldsymbol{c}_{in}^T\right]^T \in C^{nl_k} \tag{15}$$

$$q^\tau \boldsymbol{c}_{ij} = \left[q^\tau c_{ij}(0), \cdots, q^\tau c_{ij}(l_k - 1)\right]^T \in C^{l_k}. \tag{16}$$

Since $c_{ij}(\tau) = 0$ for $\tau < 0$ or $\tau \geq l_i$, we have $q^\tau \boldsymbol{c}_i = 0$ for $\tau \leq -l_k$ or $\tau \geq l_i$. Therefore, (14) means that the second condition **(C2)** is equivalent to

$$< \boldsymbol{c}_k, q^\tau \boldsymbol{c}_i >= 0;$$
$$-(l_k - 1) \leq \tau \leq l_i - 1;\ i = 1, \cdots, k-1. \tag{17}$$

Let m_k be the number of the constraints in (17) defined by

$$m_k = \sum_{i=1}^{k-1} (l_k + l_i - 1) \tag{18}$$

and $\boldsymbol{B}_k$ be an $nl_k \times m_k$ dimensional matrix defined by

$$\boldsymbol{B}_k = [\boldsymbol{C}_1, \cdots, \boldsymbol{C}_{k-1}] \in C^{nl_k \times m_k} \tag{19}$$

$$\boldsymbol{C}_i = \left[q^{l_i - 1}\boldsymbol{c}_i, \cdots, q\boldsymbol{c}_i, \boldsymbol{c}_i, q^{-1}\boldsymbol{c}_i, \cdots, q^{-(l_k-1)}\boldsymbol{c}_i\right] \in C^{nl_k \times (l_k + l_i - 1)} \tag{20}$$

We decompose the unitary space C^{nl_k} into two orthogonal subspaces as

$$C^{nl_k} = \mathrm{Im}\boldsymbol{B}_k \oplus (\mathrm{Im}\boldsymbol{B}_k)^\perp, \tag{21}$$

where $\mathrm{Im}\boldsymbol{B}_k$ denotes the image space of matrix $\boldsymbol{B}_k$ and $(\mathrm{Im}\boldsymbol{B}_k)^\perp$ denotes the complementary orthogonal subspace of $\mathrm{Im}\boldsymbol{B}_k$ in C^{nl_k}. Then it follows from (14) and (17) that the second condition **(C2)** is equivalent to

$$\boldsymbol{c}_k \in (\mathrm{Im}\boldsymbol{B}_k)^\perp. \tag{22}$$

In order that there is non-trivial solution $\boldsymbol{c}_k \neq 0$ of (22), it holds $nl_k > m_k$, provided that all the column vectors of $\boldsymbol{B}_k$ are linearly independent. This means along with (18) that the sequence $\{l_i\}_{i=1}^n$ is an increasing sequence satisfying

$$l_k > \sum_{i=1}^{k-1} (l_i - 1), \quad k = 1, \cdots, n. \tag{23}$$

We now describe an iterative procedure for solving the constrained maximization problems in the multistage criterion (A) by the gradient method. The procedure for finding a solution which attains the maximum of the criterion function of Stage 1 subject to **(C1)** with $k = 1$ is just the same as Shalvi and Weinstein [1], and thus each iteration of the procedure at Stage 1 consists of the following two steps:

$$\boldsymbol{c}_1' = \boldsymbol{c}_1 + \mu_1 \nabla_{\boldsymbol{c}_1} F_1, \tag{24}$$

$$\boldsymbol{c}_1'' = \frac{1}{\|\boldsymbol{c}_1'\|}\, \boldsymbol{c}_1' \tag{25}$$

where $\boldsymbol{c}_1$ is the vector of the first-row tap-coefficients (defined by (8) and (9) with $k = 1$) before the iteration, $\boldsymbol{c}_1''$ is vector of the tap-coefficients after iteration, μ_1 is a positive constant that regulates the step-size, and $\nabla_{\boldsymbol{c}_1} F_1$ is the gradient of the criterion function F_1 of Stage 1 with respect to $\boldsymbol{c}_1$.

At Stage k (where $2 \leq k \leq n$) in addition to **(C1)** we should take account of second condition **(C2)** which is equivalent to $\boldsymbol{c}_k \in (\mathrm{Im}\boldsymbol{B}_k)^\perp$. Let $\boldsymbol{P}_k$ denote the matrix of the projection onto the subspace $\mathrm{Im}\boldsymbol{B}_k$ along $(\mathrm{Im}\boldsymbol{B}_k)^\perp$, and $\boldsymbol{P}_k^\perp$ denote the matrix of the projection onto the subspace $(\mathrm{Im}\boldsymbol{B}_k)^\perp$ along $\mathrm{Im}\boldsymbol{B}_k$. Then any vector $\boldsymbol{c}_k$ in C^{nl_k} has an orthogonal decomposition given by

$$\boldsymbol{c}_k = \boldsymbol{c}_k^{(1)} \oplus \boldsymbol{c}_k^{(2)} \tag{26}$$

where

$$\boldsymbol{c}_k^{(1)} = \boldsymbol{P}_k \boldsymbol{c}_k \quad \text{and} \quad \boldsymbol{c}_k^{(2)} = \boldsymbol{P}_k^\perp \boldsymbol{c}_k. \tag{27}$$

It is proved that

$$\nabla_{\boldsymbol{c}_k^{(2)}} F_k = \boldsymbol{P}_k^\perp \nabla_{\boldsymbol{c}_k} F_k. \tag{28}$$

It follows from (27) and (28)

$$\boldsymbol{c}_k^{(2)} + \mu_k \nabla_{\boldsymbol{c}_k^{(2)}} F_k = \boldsymbol{P}_k^\perp \boldsymbol{c}_k + \mu_k \boldsymbol{P}_k^\perp \nabla_{\boldsymbol{c}_k} F_k$$
$$= \boldsymbol{P}_k^\perp \{\boldsymbol{c}_k + \mu_k \nabla_{\boldsymbol{c}_k} F_k\}, \tag{29}$$

where $\mu_k > 0$.

Based on the preceding discussions, each iteration of the procedure for finding a point which attains the maximum of the criterion function F_k of Stage k subject to **(C1)** and **(C2)**

(where $2 \le k \le n$) is made of the following two steps:

$$\boldsymbol{c}_k' = \boldsymbol{P}_k^{\perp}\{\boldsymbol{c}_k + \mu_k \nabla_{\boldsymbol{c}_k} F_k\}, \tag{30}$$

$$\boldsymbol{c}_k'' = \frac{1}{\|\boldsymbol{c}_k'\|}\,\boldsymbol{c}_k' \tag{31}$$

where $\boldsymbol{c}_k$ is the vector before the iteration, and $\boldsymbol{c}_k''$ is the vector after the iteration.

The Stochastic Gradient Algorithm

The algorithm goes on recursively in time t. Assume the coefficients of the equalizer have been computed at time $t-1$. Then they are updated at time t from the data $\{z_j(t), \tilde{y}_j(t-\tau); j = 1, \cdots, n; 0 \le \tau \le \max(l_i)\}$ by the following steps:

Step 1. Compute $\{c_{1j}''(\tau); \tau = 0, \cdots, l_1 - 1\}_{j=1}^n$ by

$$c_{1j}'(\tau) = c_{1j}(\tau) + \mu_1 \operatorname{sgn}\kappa_{4,z_1} \times \left[|z_1(t)|^2 z_1(t)\tilde{y}_j^*(t-\tau) - \langle z_1(t)^2\rangle z_1^*(t)\tilde{y}_j^*(t-\tau)\right]$$

$$c_{1j}''(\tau) = \frac{1}{\sqrt{\sum_{j=1}^n \sum_{\tau=0}^{l_1} |c_{1j}'(\tau)|^2}}\, c_{1j}'(\tau)$$

$$\langle z_1(t)^2\rangle = (1-\mu_{\varepsilon_1})\langle z_1(t-1)^2\rangle + \mu_{\varepsilon_1} z_1(t)^2$$

where μ_1 is a positive constant that regulates the step-size, and μ_{ε_1} is a positive constant that regulates the step-size used for estimating $E\{z_1(t)^2\}$. Then set

$$c_{1j}(\tau) = c_{1j}''(\tau) \quad \text{for} \quad j = 1, \cdots, n;\ \tau = 0, \cdots, l_1 - 1.$$

Step 2. Set $k = 2$.

Step 3. Define $\{\boldsymbol{c}_i\}_{i=1}^{k-1}$ by

$$\boldsymbol{c}_i = \left[\boldsymbol{c}_{i1}^T, \cdots, \boldsymbol{c}_{in}^T\right]^T \in C^{nl_k}$$

$$\boldsymbol{c}_{ij} = \left[c_{ij}(0), \cdots, c_{ij}(l_k - 1)\right]^T \in C^{l_k}$$

where $c_{ij}(\tau) = 0$ for $\tau \ge l_i$.

Step 4. Define $\{\boldsymbol{C}_i\}_{i=1}^{k-1}$ by

$$\boldsymbol{C}_i = \left[q^{l_i-1}\boldsymbol{c}_i, \cdots, q\boldsymbol{c}_i, \boldsymbol{c}_i, q^{-1}\boldsymbol{c}_i, \cdots, q^{-(l_k-1)}\boldsymbol{c}_i\right].$$

Step 5. Compute the projection matrix $P_k^{\perp}$ presenting the projection onto the complementary orthogonal subspace of the image space of matrix $[\boldsymbol{C}_1, \cdots, \boldsymbol{C}_{k-1}]$ (This computation may be carried by finding a SVD (singular value decomposition) of $[\boldsymbol{C}_1, \cdots, \boldsymbol{C}_{k-1}]$).

Step 6. Compute $\{c_{kj}'(\tau); \tau = 0, \cdots, l_k - 1\}_{j=1}^n$ by

$$c_{kj}'(\tau) = c_{kj}(\tau) + \mu_k \operatorname{sgn}\kappa_{4,z_k} \times \left[|z_k(t)|^2 z_k(t)\tilde{y}_j^*(t-\tau) - \langle z_k(t)^2\rangle z_k^*(t)\tilde{y}_j^*(t-\tau)\right]$$

$$\langle z_k(t)^2\rangle = (1-\mu_{\varepsilon_k})\langle z_k(t-1)^2\rangle + \mu_{\varepsilon_k} z_k(t)^2$$

where μ_k is a positive constant that regulates the step-size, and μ_{ε_k} is a positive constant that regulates the step-size for estimating $E\{z_k(t)^2\}$.

Step 7. Define $\boldsymbol{c}_k'$ by

$$\boldsymbol{c}_k' = \left[\boldsymbol{c}_{k1}'^T, \cdots, \boldsymbol{c}_{kn}'^T\right]^T \in C^{nl_k}$$

$$\boldsymbol{c}_{kj}' = \left[c_{kj}'(0), \cdots, c_{kj}'(l_k - 1)\right]^T \in C^{l_k}.$$

Step 8. Compute $\boldsymbol{c}_k''$ by

$$\boldsymbol{c}_k'' = \boldsymbol{P}_k^{\perp}\boldsymbol{c}_k'.$$

Step 9. Define $\{c_{kj}''(\tau); \tau = 0, \cdots, l_k - 1\}_{j=1}^n$ by

$$\left[c_{kj}''(0), \cdots, c_{kj}''(l_k - 1)\right]^T = \boldsymbol{c}_{kj}''$$

$$\left[\boldsymbol{c}_{k1}''^T, \cdots, \boldsymbol{c}_{kn}''^T\right]^T = \boldsymbol{c}_k''.$$

Step 10. Compute $\{c_{kj}(\tau); \tau = 0, \cdots, l_k - 1\}_{j=1}^n$ by

$$c_{kj}(\tau) = \frac{1}{\sqrt{\sum_{j=1}^n \sum_{\tau=0}^{l_k-1} |c_{kj}''(\tau)|^2}}\, c_{kj}''(\tau).$$

Step 11. If $k < n$, set $k = k+1$ and go to Step 3. If $k = n$, the go to next time $t+1$. To go to the next time, we set $t = t+1$ and collect new data $\{z_j(t+1), \tilde{y}_j(t+1); j = 1, \cdots, n\}$ from each output, where $z_j(t+1)$'s are the equalizer outputs (at $t+1$) computed using the new coefficients of the equalizer updated at time t. Then go back to Step 1.

5 The Unconstrained Criterion

It is generally more difficult to solve a maximization problem with constraints than to solve a constraint-free maximization problem equivalent to the original one. In the sequel, we develop constraint-free criteria for solving the multichannel blind deconvolution.

Let us assume that we know all the magnitudes of the fourth-order auto-cumulants of the components of the vector process ahead and that they satisfies the following decreasing sequence condition

$$|\gamma_1| \ge |\gamma_2| \ge \cdots \ge |\gamma_n| \tag{32}$$

where $\gamma_i := \kappa_{4,u_i}$ for $i = 1, \cdots, n$. Consider the following potential function [1] defined by

$$\phi_i(z_i) := |\kappa_{4,z_i}| + |\gamma_i| f(\sigma_{z_i}^2) \tag{33}$$

where $f(\cdot)$ is a continuous real-valued function over $[0, \infty)$ such that

$$p(x) := x^2 + f(x) \tag{34}$$

monotonically increases in $0 \le x < 1$, monotonically decreases for $x > 1$, and has a unique maximum at $x = 1$. Such a function, for example, is given by $p(x) = 2\alpha x - \alpha x^2, \alpha > 0$.

Corresponding to the multistage maximization criterion (A), we consider the following unconstrained multistage maximization criterion (B):

(Stage 1): Maximize

$$J_1 := |\kappa_{4,z_1}| + |\gamma_1| f(\sigma_{z_1}^2) \tag{35}$$

(Stage k): Maximize

$$J_k := |\kappa_{4,z_k}| + |\gamma_k| f(\sigma_{z_k}^2) - \lambda_0 (\sum_{i=1}^{k-1} \sum_{\tau \in Z} |r_{z_i,z_k^*}(\tau)|^2)^2 \tag{36}$$

where λ_0 is a positive constant greater than $|\gamma_1|$, *i.e.*, $\lambda_0 \geq |\gamma_1|$.

Based on Theorem 1, we have the following theorem.

*Theorem 2[3] :*Under the normalized whitening condition of the input process $\{\boldsymbol{u}(t)\}$, the unconstrained multistage maximization criterion (B) gives a solution to the multichannel deconvolution problem.

Proof: The proof is omitted.

Under the assumption that all the magnitudes of the fourth-order auto-cumulants of the components of the input vector process are identical, Inouye and Sato also presented a new unconstrained single-stage maximization criterion for attaining the blind deconvolution of multichannel LTI systems[3]. Based on their unconstrained criteria, we can show also on-line (or stochastic gradient) algorithms for solving the blind deconvolution of multichannel LTI systems. However, the algorithms are omitted for a page limit.

6 Simulation Results

In order to see the performance of the proposed algorithms, we examined the on-line (or stochastic gradient) algorithm (A) based on the constrained multistage maximization criterion, the on-line algorithm (B) based on the unconstrained multistage maximization criterion, and the on-line algorithm (C) based on the unconstrained single-stage maximization criterion.

We considered the following scenario: Since the algorithm (A) require (multichannel) spectral prewhitening of the output process of the unknown system as in [1], we chose an all-pass LTI system for the unknown system, which is a 2-input and 2-output system described by

$$\boldsymbol{H}(z) = \begin{pmatrix} \frac{0.5+z^{-1}}{1+0.5z^{-1}} & 0 \\ 0 & \frac{0.2+z^{-1}}{1+0.2z^{-1}} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}. \tag{37}$$

Hence we need not to perform a whitening operation for the output process ahead in applying the algorithm (A), and thus we used the same symbol $\boldsymbol{W}$ as an equalizer in applying the three algorithms. The channel input signals $u_1(t)$ and $u_2(t)$ were data sequences from mutually independent 4-PSK (phase-shift keying) sources with unit variance. We used a 2-input and 2-output equalizer $\boldsymbol{W}$. The length l_1 and the length l_2 of equalizer $\boldsymbol{W}$ were chosen to be 12 and 24, respectively. As a measure of performance we used

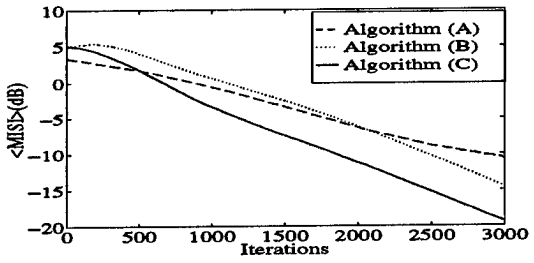


Figure 2. Performances of the three on-line algorithms.

the multichannel intersymbol interference denoted by M_{ISI}, which was introduced in [2] and [3]. The initial M_{ISI} in logarithmic (dB) scale was 5.0803 dB for the three algorithms. The three algorithms were also examined in 10 independent Monte Carlo runs using 3,000 data sample points at the channel outputs for each Monte Carlo run.

We note that both the algorithm (B) and the algorithm (C) need not performing prewhitening of the channel outputs even if this is not the case, but the gradients of the criterion functions in the algorithm (B) and the algorithm (C) involve more terms than those of the criterion functions in the algorithm (A). In Fig. 2, we plotted the average M_{ISI}, denoted by $\langle M_{ISI} \rangle$, over 10 independent Monte Carlo runs. It can be seen from Fig. 2 that the algorithm (C) exhibits the best performance of all the algorithms.

7 Conclusions

We proposed the on-line algorithm (A) based on the constrained multistage criterion. Then we presented the on-line algorithm (B) based on the unconstrained multistage criterion and the on-line algorithm (C) based on the unconstrained single-stage criterion. Simulation results have shown illustrate the performance of the three algorithms, (A), (B), and (C).

References

[1] O. Shalvi, and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems(channels)", *IEEE Trans. Inform. Theory*, vol. 36, pp. 312-321, Mar. 1990.

[2] Y. Inouye and T. Habe, "Multichannel blind equalization using second- and fourth-order cumulants", *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, pp. 96-100, 1995.

[3] Y. Inouye and T. Sato, "Unconstrained optimization criteria for blind equalization of multichannel linear systems," *Proc. IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 320-323, 1996.

Fourth-Order HOS Deconvolution in Wideband Communication Applications with Lattice Geometric Motion

James C. Solinsky, Ph.D.
Director of Advanced Technology
Cubic Defense Systems, Inc. (CDS)
9333 Balboa Avenue, MS 8A3
San Diego, CA 92123
e-mail: j.solinsky@cubic.com

ABSTRACT

A number of fourth-order HOS approaches are reviewed for applicability to RF propagation in wideband application in urban environments. A lattice geometry is used to model constraints in the propagation which constrain the mean direction of arrival. Errors in estimation of this angle, and methods of increasing the signal level can potentially be mitigated using blind deconvolution with a distributed set of sources. With a moving receiver, a time average can also potentially improve these estimates. Data results show some of these modeled propagation effects, which have dramatic angular changes with slight frequency shifts.

1. Introduction

In current digital personal cellular mobile communications systems (PCS) and wireless digital local-area networks (WLAN) used in urban environments, the effects of multiple propagation paths (multipaths) received at the same point can cause significant signal-loss effects and phase distortion of the transmitted data. These effects for the higher-frequency bands (e.g., 902-928 MHz, 2.5 GHz) can have large link margin losses over short ranges (e.g., 100 dB). Wideband operation for digital transmission in excess of 1 Mbps requires a few MHz of bandwidth and signal-to-noise ratios (SNR) in excess of 10 dB to achieve tolerable bit-error rates (i.e., BER < 10^{-3}). The effects of multipath fading and local phase distortion can cause SNR losses of over 20 dB and long time delay spreads, which combine to severely limit PCS/WLAN applications in urban environments. An approach to mitigating multipath effects using higher-order statistics (HOS) is presented and related to past theoretical work and recent data measurements.

Simplistic Representation

One can view an urban propagation environment as a collection of free-pathways and building-blocked pathways on a rectangular grid. This lattice structure can be analogous to a solid-state crystal, with a unit cell having a "diatomic" cubic structure consisting of an "L" shaped free-path, and a square shaped solid path as shown in Figure 1. A transmitter source (TX) propagates at a wavelength λ between these unit cells, of length 2a, and cell location (k, l), with Snell's Law reflection from the solid squares. The received propagation at a singular point (RX) consists of direct, line-of-sight (LOS) propagation, and N, multiple scattered paths, from 1 to m(e.g., 7) reflections. A few example pathways are shown in Figure 1, with a mean received pathway direction of arrival (DOA), $\bar{\succ}$, being different from the direct path.

If the datalink transmitted signal is represented as a complex modulated function $\tilde{s}(t)$, at frequency f_o, then the received signal $\tilde{r}(t)$, at range R, is just a phase delayed LOS direct path signal, with amplitude reduced by α as:

$$\tilde{r}(t) = \alpha\tilde{s}(t - \tau)\, e^{-i2\prec f_o-} \qquad (1)$$

For a multipath environment, each m^{th} path contributes to the received multipath signal, $\tilde{r}_m(t)$, which can be represented by a sum of signals as:

$$\tilde{r}_m(t) = \sum_{m\ 1}^{N} \hbar \mapsto_m(t)\tilde{s}(t - _ _m(t))\, e^{-i2\prec f_{o-m}} \qquad (2)$$

with amplitude distortion $\alpha_m(t)$ and time delay $\tau_m(t)$ being time-dependent functions from changes in the RX/TX location, or from changes in the propagation path from other motion. Generally in open fields, N is only 2 from a single ground reflection, but in urban areas, N can be large.

0-8186-8005-9/97 $10.00 © 1997 IEEE

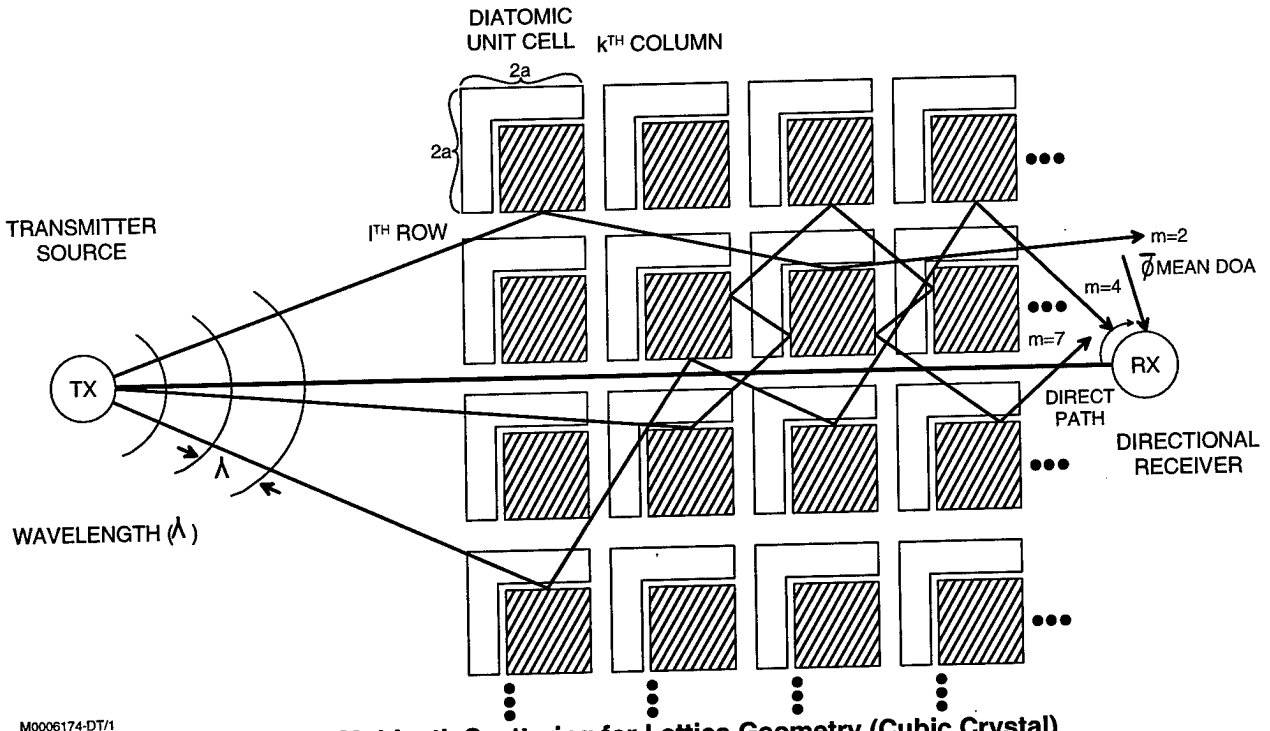


Figure 1. Multipath Scattering for Lattice Geometry (Cubic Crystal)

The analogy of the lattice propagation in cubic solid-state crystals[1] is analyzed by the atomic oscillation modes being either longitudinal or transverse to the direction of travel. Here the diatomic cell has two atomic masses (i.e., M_1, M_2) with force constant C, and each propagation mode has two vibrations for a given wavelength at frequency f ($f = c/\lambda$, where c is the speed of propagation). Nearest neighbor interaction (such as in a single reflection point per unit cell of Figure 1) simplifies the equation of motion to two traveling wave solutions, with ω frequency dependence ($\omega = 2\pi f$) for wave number $k=2\pi/\lambda$, and the first Brillouin zone is at $k_{max} = \pm \pi/2a$. At k_{max}, the two propagation modes are limited to $\omega^2 = 2C/M_2$ (optic) and $\omega^2 = 2C/M_1$ (acoustic) for $M_1 > M_2$. These two wave number solutions are analogous to the reflection of Figure 1 residing with one corridor (i.e., example m = 2) or between two corridors (i.e., example m = 4), and have two distinct dispersion relationships (ω vs. k in Figure 17 of Ref. 1).

A ray-tracing propagation example was performed in simulation of an urban section of Rosslyn, Virginia[2], modeled with a direct path and multipaths down to 100 dB in reduced SNR. The average path delay spread was 16.2 nsec with an average of N = 6.3 path components. This lattice model of restricted DOA was supported by a histogram of 319 location measurements along one street of the simulation having dominant peaks at ±90° (Figure 2 of Ref. 2). Thus, referring back to Equation 2, it is possible to generalize the summation for large N to a convolution, as:

$$\tilde{r}_m(t) \; \hbar \; \hbar\text{-}(t)\tilde{s}(t - \text{-}(t))\, e^{-i2\text{-}f_{o}\text{-}(t)} \frac{dt}{T} \qquad (3)$$

for time periods $T >> \tau$.

If one considers the motion of a receiver within the urban lattice geometry model as a collection of omni-receivers superimposed as an array receiver during the averaging time T, then one can make an analogy to other models of DOA. Merge and Wong[3] examine the DOA for a linear array of four elements at $\lambda/2$ spacing (e.g., $\lambda = 4a$), and solve for the eigenvalues of a fourth-order cumulant matrix (Ck). These values are limited to a signal subspace of dimension 2, with an identical separation with signal spread (DOA) to the dispersion of the lattice model in Ref. 1 (see Figure 1 of Ref. 3). This supports the analogy of using a convolution model for urban propagation, which has a dispersion limitation on the DOA propagation in averaging times performed over a long time relative to the propagation delays.

The key concept in using HOS is to deconvolve the effects of multipath over short motion times as follows. The inverse Fourier transform (-FFT) of the product of the transformed functions (FFT) is:

$$\tilde{r}_m(t) = [\tilde{R}_m(f)]^{-FFT}, \text{ with } \tilde{R}_m(f) \; \hbar \; \tilde{S}(f)\tilde{A}(f) \quad (4), (5)$$

where $\tilde{S}(f) = [\tilde{s}(t)]^{FFT} e^{i2\pi f_o \tau}$, and (6)

$\tilde{A}(f) = [h(t) e^{i2\pi f_o \tau}]^{FFT}$ (7)

The deconvolution or equalization involves a complex division of the estimate of $\tilde{A}(f)$ as $1/\langle \tilde{A}(f) \rangle$, with estimation $\langle\ \rangle$, and the inversion of Equation (3), by solving for $\tilde{s}(t - \tau)$ as:

$$\tilde{s}(t - \tau) = [\tilde{R}_m(f) / \hbar \tilde{A}(f) \hbar]^{-FFT} \quad (8)$$

Since the estimate of $\langle \tilde{A}(f) \rangle$ can be noise contaminated, this deconvolution using second-order estimates can have singularities. However, since $\tilde{s}(t)$ in general is non-Gaussian, a HOS approach to estimating the deconvolution can be free of Gaussian noise contamination, if done in fourth-order.

Paper Outline

The paper extends this propagation analogy in the next section into the various publications on deconvolution using HOS, with particular focus on the fourth order formulations. In most multipath applications the received pathway is unknown or blind. A summary of these results is used to gain insight on DOA data measurements made in urban environments, which show dispersive effects with motion and change in frequency. These results are summarized in a conclusion section to argue for further theoretical work in fourth-order *blind* deconvolution for removing the multipath effects in wideband PCS and WLAN propagation for enhanced SNR and higher digital transmission rates.

2. Background

Typical HOS blind deconvolution techniques have used higher-order spectra and cepstra for two receivers and a Gaussian model for the convolution noise[4]. The estimation of the kernel involves a nonlinear transformation and mean-square error (MSE) technique. A number of fourth-order techniques are reviewed.

Fourth-Order HOS

Fourth-order HOS cumulant (HOC) techniques[5] have reduced simulated MSE signal error by up to 30 dB and preserved the phase for 10^3 estimation iterations. Such iterative techniques may not be possible in long coherence-length, direct-sequence spread-spectrum (DSS) PCS/WLAN modulation techniques for large numbers of code division multiple access (CDMA) users. However, recent fourth-order techniques in data-link applications have predicted success in larger sample sizes[6]. The use of multiple receiver data sets with the Kolmogorov theoreom[7] suggests the use of a neural network (i.e., an extended Kalman filter) with a moving, single receiver can synthesize the single source separation in the DOA problem for a two source simulation, which in turn reduces the phase error of the deconvolution.

Since the digital datalinks under consideration using DSS modulation are non-Gaussian, the recent fourth-order HOS approaches[8,9] were shown to be superior in DOA applications with correlated Gaussian noises, and have been extended to spatially distributed signals[3]. Other work[10] has shown superiority of fourth-order estimation in digital communication applications for SNR < 12 dB in blind deconvolution under noisy conditions. Also, multipath application of fourth-order HOS has shown very long filter deconvolutions for SNR = 10 dB[11], and many independent signal points.

Multipath Urban Propagation

Theoretical multipath propagation in urban environments has shown agreement with data collection using a free-space/conducting-slab model in 3D[12], with SNR range roll-off to powers of R^n at n = -2.6 to -4.4. Similar data were collected in urban environments by CDS at a number of locations and frequencies. These losses had a n = -3 roll-off in the urban "lattice", with an additional loss of 8-50 dB inside buildings. Narrow band, indoor sweeps of frequencies in the 226 MHz and 910 MHz bands showed multipath fades of up to 30 dB over bandwidths of 30 MHz. This means that improved, indoor PCS/WLAN applications require the estimation of a deconvolution kernel for moving receivers in urban environments, over short distance and in short times, and under low SNR. The next section highlights the changes in the DOA of these received signals.

3. Data Results

Figure 2 shows an approximated lattice geometry for the actual buildings used in the CDS data measurements. Calibration and sweep location sites are shown for the different multi-storied buildings. Figure 3 shows a 1.2 MHz narrow band sweep in power (3a) and DOA (3b), with each point having a mean ($\bar{\phi}$ as solid) and standard deviation value (open) averaged over many samples in a fixed geometry. The true LOS DOA is over 45° in error from these mean measurements, with a few degrees of fluctuation error. This dispersion in DOA with frequency and reduced SNR (dBm level) clearly shows correlation. Figure 4 is the same data measurement, but with motion near the receiver of human bodies in the room. Even these

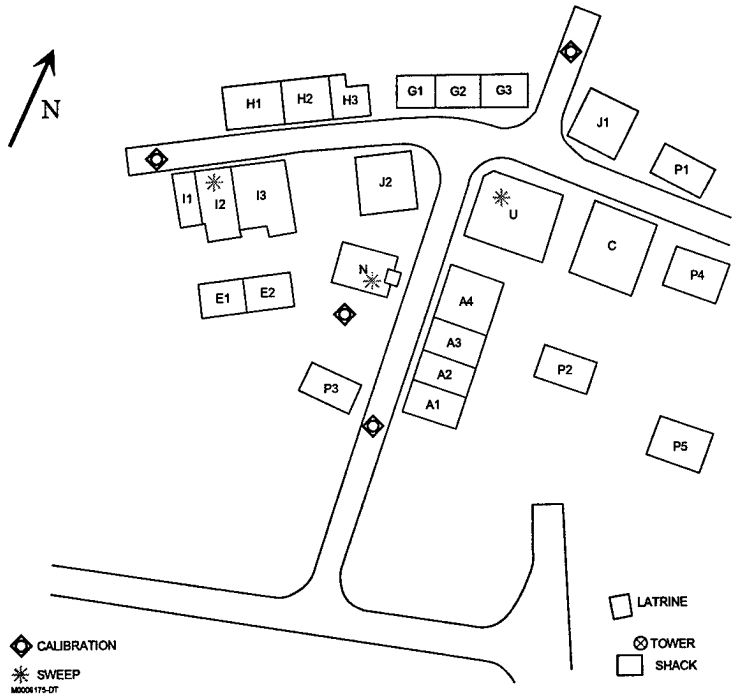


Figure 2. Urban Measurement Environment for DOA Sweeps

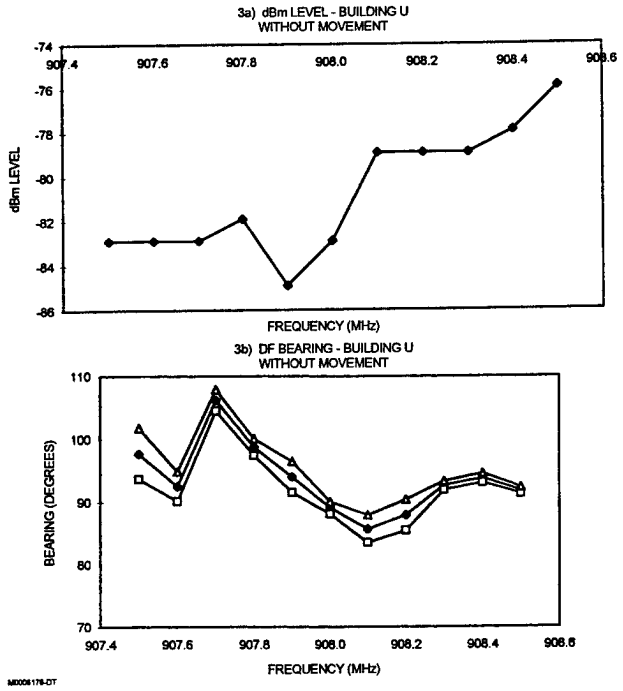


Figure 3. RF Sweep Building U - No Movement - 2nd Story A (Frequency 907.4 - 908.6 MHz F.S.; LOS DOA = 145°)

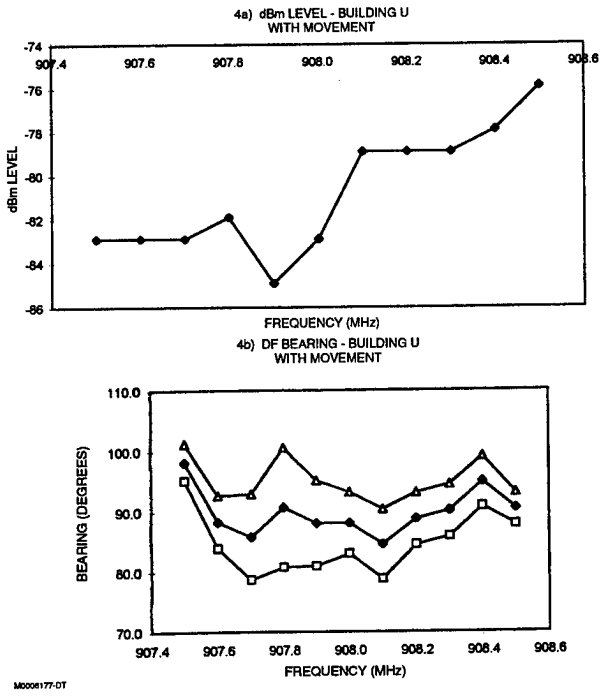


Figure 4. RF Sweep Building U - Movement - 2nd Story A (Frequency 907.4 - 908.6 MHz F.S.; LOS DOA = 145°)

slight reflections disturb the measurement by a factor of 5 in greater estimation noise. This multipath reflection noise cannot be considered to be stochastic over short times, but places a new constraint on DOA estimation using HOS —stochastic motion noise. Similar results were measured in other buildings, with nonlinear and sometimes abrupt changes in DOA with slight frequency changes. This supports a lattice model effect in constraining urban RF propagation to switch between constrained pathways, with slight changes in wavelength.

4. Conclusion

An analysis of urban geometries and lattice wave propagation models has shown constrained propagation dependence with frequency, particularly in the DOA measurements. Recent fourth-order theoretical and simulated-data tests with HOS algorithms, show promise in blind deconvolution approaches, which could potentially improve SNR and DOA estimation. However, multipath measurements show that slight changes in the mean DOA ($\bar{\phi}$) estimation occur from motion of reflecting/blocking humans, and would be more severe from solid object motion. This example can be viewed as a distributed set of receiver measurements, if time averaging estimation is used in the blind deconvolution approach. These results point to a need for new theoretical algorithm development for improving PCS/WLAN technology.

5. Acknowledgments

The data measurements were made under a CDS IRAD program with the assistance of Mike Bredimus, John Roes, Allen Ripingill, and Ed Kutney and the data reduction efforts of Aaron Lindsdau and Jerome Cruz.

6. References

1) C. Kittel, *Introduction to Solid State Physics*, 3rd Edition, John Wiley, N.Y., NY, p. 147 (1967).

2) Glomo Project Web Page, "Simulation of CDMA Using Spatial Filtering in Multipath Environments," Virginia Tech @ address:
http://www.mprg.ee.vt.edu/research/glomo/antenna/CDMA_sim.html

3) Y. Meng, K.M. Wong, "A Direction Finding Algorithm Based on Higher-Order Statistics for Distributed Signals," IEEE Sig. Process. ATHOS Workshop in HOS, Begur. Girona. SPAIN, p. 91, (1995).

4) C.L. Nikios, A.P. Petropule, *Higher-Order Spectra Analysis*, Prentice Hall, Englewood Cliffs, NJ, p. 253 ff.(1993).

5) F. Zheng, S. McLaughlin, B. Mulgrew, "A 2nd and 4th-Order Cumulant Based Blind Equalization Algorithm for Nonminimum Phase Channels", HOS — J.L. Lacoume, Editor, Elsevier Sci. Pub., p. 129 (1992).

6) S. Bellini, "Blind Estimation of a Complex Channel Impulse Response," IEEE Sig. Process. Workshop on HOS, Tahoe, CA, p. 131 (1993).

7) M. Nájar, M.A. Lagumas, "Simultanious DOA Estimation Based on Kolmogorov's Theorem", Ibid, p. 280 (1993).

8) G. Scarano, A.G. Matioli, and G. Jacovitti, "On DOA Estimation Based on HOS," Ibid, p. 285 (1993).

9) N. Yuen, B. Friedlander, "DOA Estimation in Multipath Based on Fourth-Order Cumulants," IEEE Sig. Process ATHOS Workshop on HOS, Begur. Girona. SPAIN, p. 71 (1995).

10) J.M. Brossler, A. Essebbar, "Synchronization and Equalization Using HOS," Ibid, p. 325 (1995).

11) M. Gaeta, F. Mondoloni, S. Jespers, "Application of a Blind Deconvolution Technique to Multipath Situations", Ibid., p. 488 (1995).

12) G.M. Whitman, K.S. Kim, E. Niver, "A Theoretical Model for Radio Attenuation Inside Buildings", IEEE Trans. on Veh. Tech., 44 (1) p. 621 (1995).

Higher-Order Statistics-based Deconvolution of Ultrasonic Nondestructive Testing Signals

Ahmed Yamani
K.F.U.P.M. Box 1811
Dhahran 31261, Saudi Arabia
email:myamani@dpc.kfupm.edu.sa

Maamar Bettayeb
K.F.U.P.M. Box 1730
Dhahran 31261, Saudi Arabia
e-mail:maamarab @dpc.kfupm.edu.sa

Lahouari Ghouti
K.F.U.P.M. Box 1128
Dhahran 31261, Saudi Arabia
e-mail:g940143@dpc.kfupm.edu.sa

Abstract

Pulse-echo reflection techniques are used for ultrasonic flaw detection in most commercial instruments. As the measured pulse echo signal is assumed to be the result of linearly convolving the defect impulse response (IR) with the measurement system response, the objective is thus, to remove the effect of the measurement system through a deconvolution operation and extract the defect impulse response. The major drawbacks of conventional second-order statistics (SOS)-based deconvolution techniques are their inability to identify non-minimum phase systems, and their sensitivity to additive Gaussian noise. Our contribution is to show that higher-order statistics (HOS)-based deconvolution techniques are more suitable to unravel the effects of the measurement systems and the additive Gaussian noise. Synthetic as well as real ultrasonic signals are used to support this claim.

1. Introduction

Pulse-echo reflection techniques are used for ultrasonic flaw detection in most commercial instruments [1]. The ultrasonic wave, generated by a piezoelectric transducer coupled to the test specimen, propagates through the material and part of its energy is reflected if the wave encounters an inhomogeneity or discontinuity in its path, while the remainder is reflected by the back surface of the test specimen. A typical oscilloscope display is shown in Fig. 1. The first wavelet represents the initial voltage applied to the transducer in order to generate the wave, while the successive echoes represent the voltage generated by the reflected wave (from the flaw and the back echo respectively) impinging on the transducer. The flaw echo in Fig. 1 contains information regarding the material discontinuity that the ultrasonic wave has encountered in its path. For this, signal processing is used on the flaw echo only, and the other echoes are discarded from subsequent signal display.

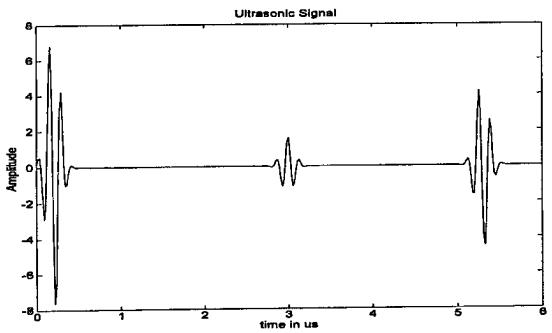


Fig. 1: Typical oscilloscope display of an ultrasonic examination.

0-8186-8005-9/97 $10.00 © 1997 IEEE

Flaw echo signals are masked by the characteristics of the measuring instruments, the propagation paths taken by the ultrasonic wave, and are corrupted by additive noise. It is assumed that the measured flaw echo is obtained by linearly convolving the flaw or the defect impulse response with the measurement system response. Deconvolution operation therefore, seeks to undo the effect of the convolution and extract the defect impulse response which is essential for defect identification.

Conventional deconvolution techniques (CDT) such as least square, Wiener filter, and minimum variance deconvolution [2] are based on a priori knowledge of second-order statistics (SOS) of the noise and the input signal. In practice however, the acoustic noise due to scattering from the grains inside the propagation medium does not have a readily known statistic [3]. Moreover, ultrasonic pulse echoes are found to be non-minimum phase systems. SOS-based deconvolution techniques, being phase-blind cannot therefore, accurately estimate the defect impulse response.

The objective of this paper is to formulate the defect ultrasonic model in the polyspectrum domain where the processing is more suitable to unravel the effect of the measurement system and the additive Gaussian noise. Thereafter, the defect impulse response is recovered from its noise-free polyspectrum. Synthesized, as well as real ultrasonic signals are used to show that the proposed technique excels conventional SOS-based deconvolution techniques commonly used in NDT.

2. Theory

A measured ultrasonic flaw signal, *y(t),* can be modeled as the convolution of the measurement system response function, *x(t),* with the flaw's impulse response function, *h(t),* plus noise, *N(t)*. This model can be written as

$$y(t) = x(t) \otimes h(t) + N(t) \qquad (1)$$

where $\otimes$ denotes the convolution operation. With this model, defect of a particular geometry would be completely characterized by its impulse response. Estimation of *h(t)* in (1), from the knowledge of *y(t)*, and *x(t)* is variously known as system identification, filtering, or simply as deconvolution. Many deconvolution techniques have been developed in different engineering areas such as seismic exploration, military applications, and medical imaging. Chen [2] has studied the feasible applications of these deconvolution techniques to ultrasonic NDT, and has concluded that Wiener filter is a good candidate for such application. The main drawbacks of CDT are their inability to identify non-minimum phase systems, and their optimal implementation requires a priori knowledge of the noise statistics. These drawbacks can be completely alleviated when using HOS-based deconvolution techniques as is shown in this paper.

Equation (1) can be written in the polyspectrum domain as [4]

$$C_n^{\,y}(w_1, w_2, \ldots, w_{n-1}) = C_n^{\,x}(w_1, w_2, \ldots, w_{n-1}) H(w_1) H(w_2) \cdots H(w_{n-1}) H^*(w + w_2 + \ldots + w_{n-1}) + C_n^{\,N}(w_1, w_2, \cdots, w_{n-1}) \qquad (2)$$

where $C_n^{\,s}(w_1, w_2, \ldots, w_{n-1})$ is the nth-order spectrum of the signal *s(t)* (which could be *y(t), x(t),* or *N(t)*), *H(w)* is the Fourier transform of the defect impulse response *h(t),* and *w* is the angular frequency. Without loss of generality, (2) can be rewritten as

$$C_n^{\,y}(w_1, w_2, \ldots, w_{n-1}) = C_n^{\,x}(w_1, w_2, \ldots, w_{n-1}) C_n^{\,h}(w_1, w_2, \ldots, w_{n-1}) + C_n^{\,N}(w_1, w_2, \ldots w_{n-1}) \qquad (3)$$

For Gaussian noise, the polyspectrum *(n>2),* of *N(t)* is zero and thus, the noise-free polyspectrum of the defect impulse response can be calculated from (3), and used to recover *h(t)*. Alternatively, if the bispectrum is used , i.e., *n=3* above, then the noise does not have to be Gaussian to be filtered out from (2) and (3). It can have any symmetric probability density function (PDF). With one of these noise assumptions in mind, equation (2) and (3) represent the basis for the HOS-based deconvolution technique used in this paper.

3. Results

In this section, the HOS-based deconvolution technique is tested on synthesized as well as real ultrasonic signals obtained from artificial defects [2]. For computational efficiency, bispectra of the input-output signals are used only. In addition, as the recovery of a signal from its bispectrum is not a one-to-one transformation, we calculate the bicepstrum using the relationship between biceptrum and bispectrum defined by Pan and Nikias [5], thereafter, the defect impulse response is recovered using the bicepstral parameters [6].

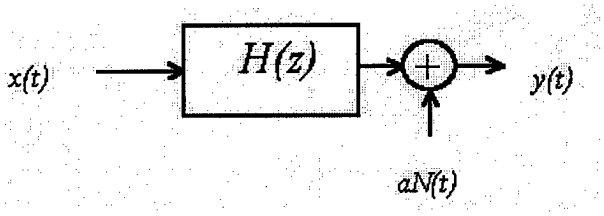


Fig. 2: The ultrasonic defect model

3.1 Synthesized Data

With reference to Fig. 2, the input signal, *x(t)*, is taken as a Gaussian pulse that is amplitude modulating a single tone carrier whose frequency lies in the ultrasonic range. The noise, *N(t)*, having a normally distributed PDF, is scaled by a constant, α, to account for different signal - to-noise ratios (SNR). Three different linear time-invariant systems are considered in this paper, namely:

- A non minimum phase moving average (MA) system whose transfer function is given by

$$H_{MA}(z) = 0.2197z^2 - 0.747z + 0.6085 + 0.1533z^{-1} \quad (4)$$

- A minimum phase autoregressive (AR) system whose transfer function is given by

$$H_{AR}(z) = \frac{1}{1 - 0.7z^{-1} + 0.6z^{-2} - 0.3z^{-3}} \quad (5)$$

- A non minimum phase autoregressive moving average (ARMA) system whose transfer function is given by

$$H_{ARMA}(z) = \frac{1 - 3.25z^{-1} + 3.5399z^{-2} - 1.2487z^{-3}}{1 - 1.86z^{-1} + 1.47z^{-2} - 0.5246z^{-3}} \quad (6)$$

For a given SNR, the output signal, *y(t)*, is computed using the model of Fig. 2. The bispectrum (*n=3)* of the system impulse response (SIR) is obtained from (3), and used to recover *h(t)* using the bicepstral parameters as stated above. To test the performance of the proposed technique, the variance of the error signal (between the true and estimated SIR signals of the MA system above) is computed for each SNR. Fig. 3 shows this result for a SNR as low as -5 dB. For comparison, a similar error variance is computed when the MA SIR is estimated using Wiener filter and is shown in the same figure. It can be clearly seen that the HOS-based deconvolution technique excels its counterpart CDT represented here by Wiener filter.

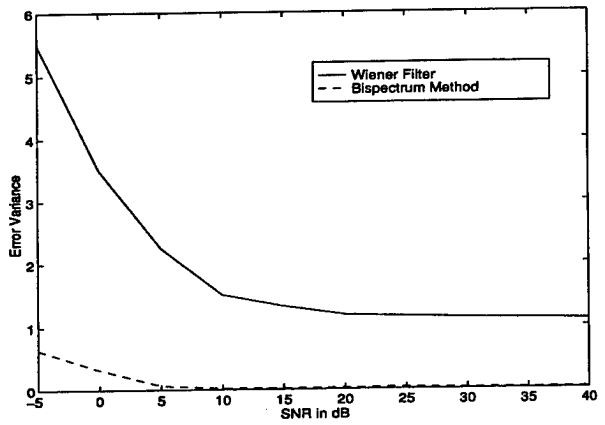


Fig.3: Error variance of the MA system impulse response.

To quantify the effect of these error variances on the estimated MA SIR, a plot of the later at a SNR=5dB is shown in Fig. 4. It can be seen that while the estimated impulse response obtained from the HOS technique is faithfully reproduced, the Wiener filter, at an error variance of about 2.2, fails completely.

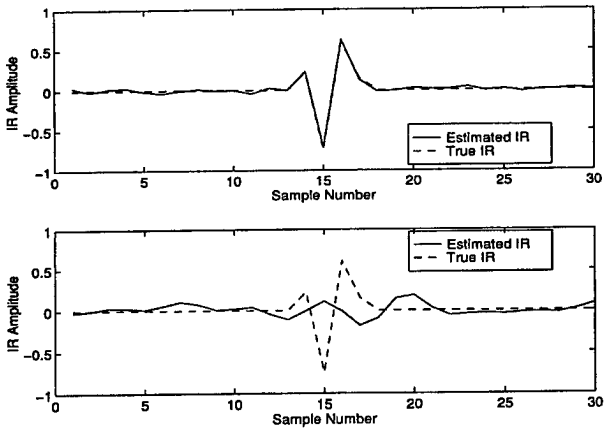


Fig. 4: MA system impulse responses obtained from HOS (top), and Wiener filter (bottom) deconvolution techniques for 5 dB SNR.

To complete this section, the AR and ARMA systems as defined by (5) and (6), are tested and their corresponding SIR estimated using the proposed technique for a SNR= 5 dB, are shown in Figs. 5 and 6 respectively.

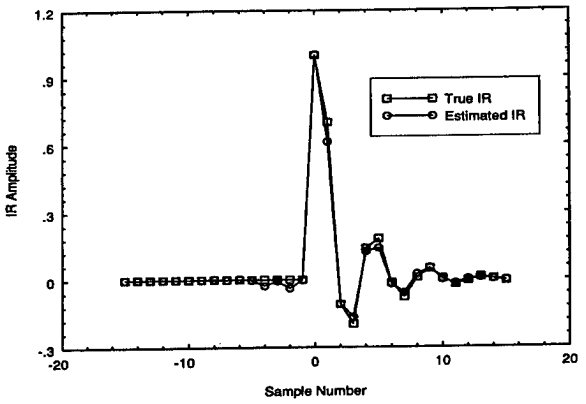


Fig. 5: AR system impulse response : true and estimated SIR for a SNR = 5 dB.

Again, the HOS-based deconvolution technique, with its potential of preserving the phase information, faithfully reproduces the SIR of both minimum (Fig. 5), and non minimum (Fig. 6) phase systems even at extremely low SNR. The small variations shown in Figs. 5 and 6 may be attributed to errors made in computing the cepstral parameters from the bicepstrum [5].

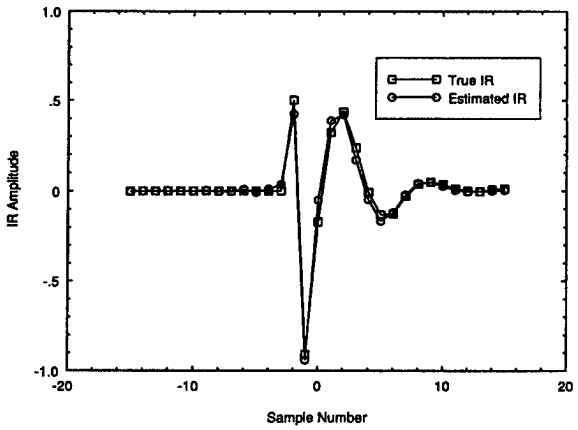


Fig.6 : ARMA system impulse response : true and estimated SIR for a SNR = 5 dB.

3.2 Real Ultrasonic Data

The proposed deconvolution technique is tested using real ultrasonic data [2], which is part of a larger data set obtained from the Army's Material Technology Laboratory, (Watertown, MA). The input signal, *x(t),* is measured in practice, from a flawless sample (A0), having the same characteristics as the specimen under test. It is referred to; sometimes; as the reference signal. Two artificial defects are considered; namely a flat-cut (A1), and an angular-cut hole (A2), in aluminum blocks, (see [2] for an illustration of these defect geometries). The center frequency of the transducer used is 15 MHz, and the A-scan signals contain 512 data points digitized at a rate of 100 MHz. The pulse-echo signals corresponding to A0, A1, and A2 samples are represented by T15A0, T15A1, and T15A2 respectively. For clarity, the signal T15A1 is shown in Fig. 7.

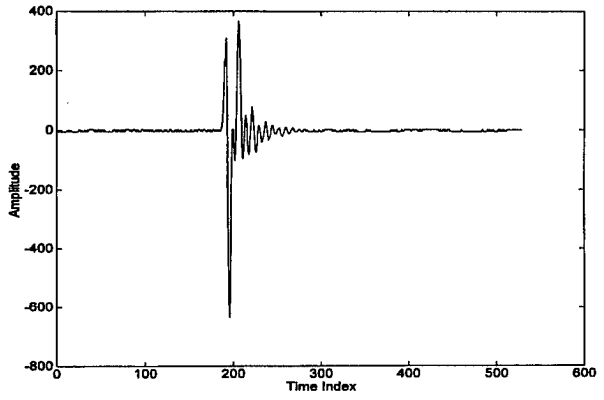


Fig. 7:Ultrasonic Pulse echo measured from sample A1.

When the bispectrum-based deconvolution technique is applied to real ultrasonic signals, namely; T15A1 and T15A2, with T15A0 taken as the reference signal, smooth, oscillation-free impulse responses are obtained as shown in Figs. 8 and 9. For comparison with CDT, the reader is referred to [7] where the same signals have been used.

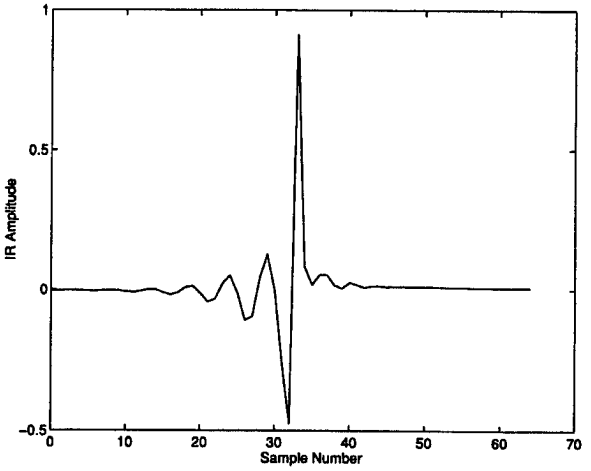


Fig. 8: Impulse response of the flat-cut hole (A1).

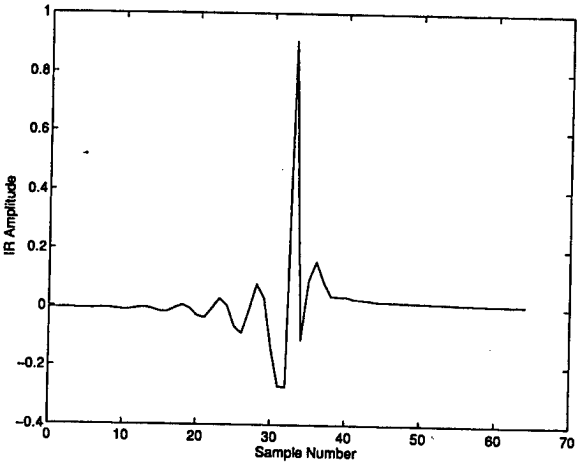


Fig. 9: Impulse response of the angular-cut hole (A2).

4. Conclusion

In this paper, we have shown that the drawbacks of the SOS-based CDT are completely removed when HOS-based deconvolution techniques are used. Synthesized, as well as real ultrasonic signals have been used to demonstrate this claim. Although we have focused on a non parametric deconvolution technique, and the bispectrum case of the polyspectra, higher-order based, parametric blind deconvolution techniques can also be used to remove the effect of the measured reference signal. Future work will be directed towards the influence of different reference signal models on the deconvolved

defect impulse response using both polyspectra and polycepstra of real ultrasonic signals.

Acknowledgments

This work is supported in part through research grant provided by SABIC under the grant EE/ SABIC/ 96-8. The Authors would like to express their gratitude to Prof. C. H. Chen of the University of Massachusetts for the experimental data he supplied and his valuable help and comments.

5. References

[1] J. Krautkramer and H. Krautkramer: Ultrasonic Testing of materials, 4th ed. Springer-Verlag, 1990, pp. 160-220.

[2] C. H. Chen, "High resolution techniques and their application to NDE", Inter. Journ. Imag. Tech., vol. 1, 1989, pp. 223-242.

[3] S. P. Neal, and K. D. Donohue, " Testing for non Gaussian fluctuations in grain noise", QNDE, vol. 15, 1996, pp. 1-9.

[4] C. L. Nikias, and A. P. Petropulu ' High-Order Specra Analysis: A nonlinear Signal Processing Framework', Prentice Hall, Englewood Cliffs, NJ, 1993.

[5] R. Pan, and C. L. Nikias, " The complex cepstrum of higher-order cumulants and non minimum phase system identification", IEEE Trans. On ASSP, vol. 36, no. 2, pp. 185-205, Feb. 1988.

[6] A. V. Oppenheim, and R. W. Schafer, ' Discrete-time signal processing', Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[7] S. K. Sin, and C. H. Chen, " A comparison of deconvolution techniques for the ultrasonic nondestructive evaluation of materials", IEEE Trans. On Image Proc., vol. 1, no. 1, pp. 3-10,Jan. 1992.

Combination of HOS Based Blind Equalization Algorithms for Use in Mobile Communications

Asoke K. Nandi and Christian Schmidt
University of Strathclyde, Department of Electronic and Electrical Engineering
204 George Street, Glasgow, G1 1XW, UK
asoke@eee.strath.ac.uk

Abstract

Mobile communication links require adaptive equalization with a fast rate of convergence while keeping computational effort at reasonable levels. In this paper we propose to combine known algorithms for blind equalization in order to exploit their desirable properties to reach this goal. A switching criterion is proposed which is based on the change in the equalizer impulse response between iterations of the adaption algorithm and may be used to detect changes of the channel impulse response. Algorithms under consideration include Godard's algorithm, Stop-and-Go algorithm, and tricepstrum equalization algorithm (TEA).

1. Introduction

Blind higher order statistics (HOS) based equalization algorithms are in use in a wide area of signal processing. Some of these applications, like biomedical and seismic, permit offline processing of data while others, like reduction of intersymbol interference (ISI) in mobile communications, require continuing processing of received data.

ISI introduced by channel fading is a major limitation for mobile link performance. Due to movements of users and shading and reflecting obstacles in such environments, channel characteristics are changing continually. These rapid changes require adaptive equalization algorithms with fast convergence. Such algorithms are often computationally expensive while others that are less computationally expensive have a slower rate of convergence.

In this paper, we propose to use fast but complex algorithms only for a startup phase and switch to slower but more simple algorithms for tracking after a fixed number of iterations. A metric is presented to detect significant changes of the channel impulse response, allowing to switch back to the faster algorithm.

The organisation of the paper is as follows. In section 2 the existing algorithms used in this paper are described and compared. In section 3 the concept of combining algorithms and the switching criterion are explained. Section 4 gives the simulation results for the proposed algorithms.

2. Existing algorithms

HOS based algorithms for blind equalization are either Bussgang type which use HOS implicitly or are based explicitly on HOS. Bussgang type algorithms based on LMS adaption use a substitute for the desired data, obtained by a memoryless nonlinear transform of the received data. Various algorithms using different nonlinearities have been proposed in the literature [1], [2], [3], [4], [5].

An algorithm based explicitly on HOS is the tricepstrum equalization algorithm (TEA) [6]. This algorithm identifies the (possibly nonminimum phase) impulse response of the channel and uses its inverse for the equalizer.

2.1. Bussgang type algorithms

The basic principle of a Bussgang type equalizer is shown in Fig. 1. Two Bussgang type algorithms used in this paper are Godard algorithm [2] and Stop-and-Go algorithm [4].

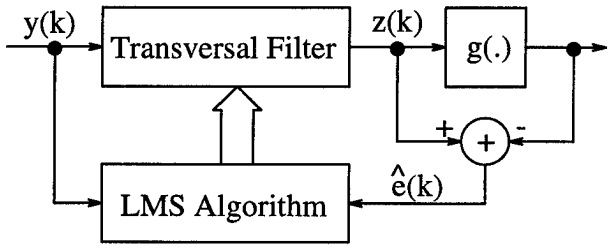


Figure 1. Bussgang algorithm equalizer

Godard's algorithm uses a nonlinearity of the form

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$g[z(k)] = \frac{z(k)}{|z(k)|} \cdot \left[|z(k)| + R_p|z(k)|^{p-1} - |z(k)|^{2p-1}\right] \quad (1)$$

with

$$R_p = \frac{E[|x(k)|^{2p}]}{E[|x(k)|^p]} \quad (2)$$

where $z(k)$ is the signal at the equalizer output and $x(k)$ are the transmitted symbols. In our simulations we chose $p = 2$. The transmitted symbols are assumed to be independent and identically distributed (i.i.d.) with non-Gaussian distribution. The perturbations introduced by the channel are ISI and additive white Gaussian noise.

As the cost function of the Godard algorithm is not convex, the convergence to the global minimum of ISI is not guaranteed. However, proper choice of step size parameter and initial equalizer gain can ensure convergence [2].

The Stop-and-Go algorithm [4] is a modification of the decision directed (DD) algorithm [1]. The DD algorithm is known to require a good initial estimate to be able to track the signal (open eye condition). The Stop-and-Go algorithm extends the DD algorithm to have the capability of blind equalization by stopping equalization for the current iteration whenever the probability that the decided sign is the correct one is too small. The corresponding memoryless nonlinearity has the form

$$g[z(k)] = z(k) - \frac{1}{2}\left(f(k)_R\hat{e}(k)_R - f(k)_I\hat{e}(k)_I\right) \quad (3)$$

where $\hat{e}(k)_R = Real[z(k) - \hat{x}(k)]$ and $\hat{e}(k)_I = Imag[z(k) - \hat{x}(k)]$ are the real and imaginary parts of the difference between the received symbol and the decided symbol $\hat{x}(k)$. The flags $f(k)_R, f(k)_I \in \{0, 1\}$ are chosen to be one if and only if the signs of the real and imaginary part, respectively, of $\hat{e}(k)$ and the Sato-like error $\tilde{e}(k) = z(k) - \beta \cdot \mathrm{csgn}[z(k)]$ agree. The constant β is real. The complex sign function is denoted by csgn[·]. The Stop-and-Go algorithm typically converges in two phases: a startup phase with slow convergence and a rapidly converging second stage.

Common to both Bussgang type algorithms is that the phase of the input symbols cannot be detected correctly.

2.2. Explicit HOS based algorithms

An explixitly HOS based algorithm is the tricepstrum equalization algorithm (TEA) [6]. This algorithm proceeds in 3 steps: 1) (adaptive) estimation of a number of second order moments and fourth order cumulants, 2) forming an overdetermined set of linear equations for the coefficients of the tricepstrum and solving this system, and 3) calculating the equalizer impulse response from the tricepstrum coefficients. Detailed descriptions of the TEA algorithm can be found in [6], [7]. The structure of an equalizer using TEA is shown in Fig. 2.

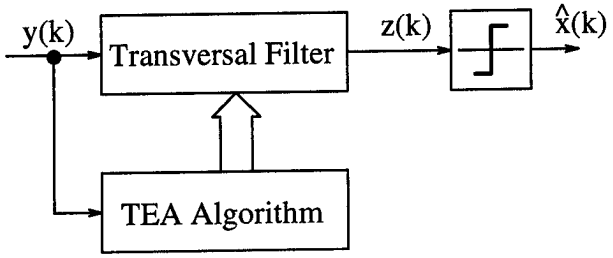


Figure 2. TEA equalizer

The advantages of this technique over Bussgang algorithms are the speed of convergence, the correct estimation of the channel phase, and the absence of convergence problems. However, the most severe problem is the significant computational cost in the estimation of the cumulants and solving the equation system. Another drawback is that the steady-state level of intersymbol interference is significantly higher than for Bussgang algorithms. This condition can be improved by proper choice of the number of cumulants used for the algorithm.

3. Algorithm combinations

3.1. Motivation

Mobile communication channels undergo rapid changes. At high data rates the channel is assumed to be quasi static over a certain number of symbols. Data is then transmitted in packets with a length according to the variation rate of the channel. At the beginning of transmission of each data packet a training sequence is sent to allow the estimation of the equalizer coefficients. During data transmission, the equalizer impulse response is kept constant or a simple algorithm (like DD) is used for tracking. When the channel undergoes a significant change tracking will be lost.

An improvement will be to use blind equalization algorithms which do not require a training sequence. Examination of existing algorithms shows that there is a trade-off between rate of convergence and computational complexity. It is possible to use a complex but fast converging algorithm for an initial startup and switch later to a simple algorithm for tracking. This scheme can be further improved by using a switching criterion on the available information at the receiver to detect significant channel changes.

3.2. Single channel

Simulations have been performed using a single static channel. As a first stage equalization has been carried out using the TEA algorithm. After a fixed number of iterations the algorithm has been changed to a Bussgang type. It turned out that as little as 100 samples for the TEA are enough to make equalization twice as fast as using Bussgang type algorithms alone (see Fig. 3, 4). The performance criterion intersymbol interference (ISI) is explained in section 4. Using more samples for TEA equalization does not lead to faster convergence. After switching from TEA to Bussgang an increase in ISI can be observed before it converges to its final value. An important feature of the combined algorithm is that the initial phase estimate of the TEA algorithm is kept by the following Bussgang equalization. Bussgang algorithms themselves are unable to estimate the correct phase of the data.

3.3. Multiple channels

When the channel changes significantly it is desirable to detect these changes and switch back to the faster algorithm. A decision criterion has to be found which is only based on information available at the receiver, i.e. on the detected data or on the equalizer impulse response. The criterion proposed in the following is derived from the equalizer impulse response. The first step is to find a measure of the state of equalization. The change of equalizer impulse response

$$q(k) = 10 \cdot \log_{10} \left(\sum_n |w_{0n}(k) - w_{0n}(k-1)|^2 \right) \quad [dB] \tag{4}$$

with the w_{0n} are the tap weights of the normalized equalizer impulse response ($\sum_n w_{0n} = 1$), seems to be a natural choice. When the equalizer is not initialized the first changes in the equalizer impulse response will be large until the tracking mode is reached. Then smaller changes result from noise and random input data. When the channel changes considerably, the equalizer impulse response will show larger variations again. To recognize a trend averaging is performed on this criterion in a way that the mean value of recent iterations is compared to the mean value and the standard deviation of a much larger number of previous samples:

$$m_1(k) = \frac{1}{N_1} \sum_{n=k-N_1+1}^{k} q(n) \tag{5}$$

$$m_2(k) = \frac{1}{N_2} \sum_{n=k-N_1-N_2+1}^{k-N_1} q(n) \tag{6}$$

$$\sigma_2(k) = \sqrt{\frac{1}{N_2} \sum_{n=k-N_1-N_2+1}^{k-N_1} \left(q(n) - m_2(k)\right)^2} \tag{7}$$

Switching between the algorithms occurs, when

$$m_1(k) > m_2(k) + \alpha \cdot \sigma_2(k) \tag{8}$$

where α is a real constant. A channel change is assumed as soon as the number of samples from $q(k)$ rises above a certain level over the mean of the past samples. The fraction of the standard deviation is used to avoid changes caused by noise and to provide an adaptive measure for switching. The choice of N_1, N_2, and α is essential for a reliable detection of channel changes.

4. Simulation results

A number of simulations have been carried out to demonstrate viability of the above concepts. Three different channel impulse responses with four taps each have been considered.

channel1	channel2	channel3
-0.0810+0.0612i	-0.1411-0.0040i	0.1820 + 0.0009i
0.9249-0.9003i	0.9496-0.9194i	-0.4942+0.4470i
0.2825+0.3650i	0.2764+0.3704i	0.6461-0.794i
-0.0779-0.0887i	0.0779-0.887i	0.2

Channels 1 and 2 have similar rms delay spread (0.3633 and 0.3643), while channel 3 has a rms delay spread of 0.5654.

For faster operation the tricepstrum equalization algorithm has been implemented such that the equation system is updated only once every 100 iterations while the cumulants are updated at each iteration. This means that for the combined algorithm the computationally expensive part of the TEA has to be performed only once.

Another crucial point of implementation is the proper choice of the parameters N_1, N_2, and α of the switching criterion. The following dependencies have been observed: A value of $N_1 = 100$ is sufficient. N_2 should be at least three times larger than N_1. The larger N_2 the more reliable is the criterion. The parameter α has an optimal value. Is α too small, normal fluctuations of $q(k)$ without changes of the channel cause false alarms. If α is too large, no channel changes will be detected. The parameters depend on the algorithm used. The ISI of the cascaded channel and equalizer impulse response s with the largest absolute tap value $|s|_{max}$ is [8]

$$ISI(s) = 10 \cdot log_{10} \left(\frac{\sum |s_n|^2 - |s|^2_{max}}{|s|^2_{max}} \right) \quad [dB] \tag{9}$$

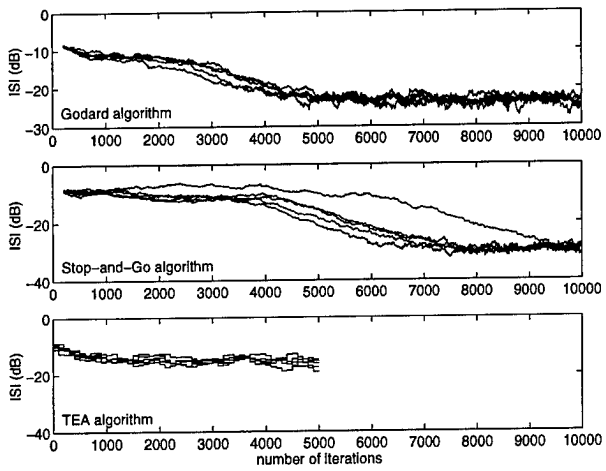


Figure 3. ISI performance for channel 1 and 5 Monte Carlo simulations

Figure 3 shows the ISI performance of the basic algorithms for channel 1. For each algorithm 5 Monte Carlo simulations have been performed. It can be seen that the Godard algorithm converges faster than Stop-and-Go, but settles at a higher level of ISI. The performance of LMS adaptive algorithms strongly depends on the step size parameter chosen. The implemented TEA algorithm converges much faster than the Bussgang type algorithms but settles at an even higher level of ISI.

Figure 4 shows how the speed of convergence is improved by combining algorithms. Again, the channel used for the simulations is channel 1. After a fixed number of iterations the TEA algorithm is switched off and Bussgang algorithms are used. It can be seen that the number of TEA iterations performed does not have a significant influence on the speed of convergence of the following algorithm.

A disturbing fact is that after switching the ISI goes up instead of down. It is desirable to reduce this effect and to obtain a continuous decrease of ISI.

Figures 5 and 6 demonstrate the effect of changes in the channel on ISI and the error criterion. To allow recognizing a trend in the error plots averaging over 100 samples has been performed. The first channel simulated in Fig. 5 and Fig. 6 is channel1. After 8000 samples the channel has been switched to channel 2 or 3, respectively.

5. Conclusions

It is indeed possible to speed up the convergence of Bussgang algorithms by combining them with an initial stage of TEA initialization. Since TEA algorithm uses upto fourth order statistics and the initial stage takes only 100 samples it may be expected that the contribution of the TEAalgorithm

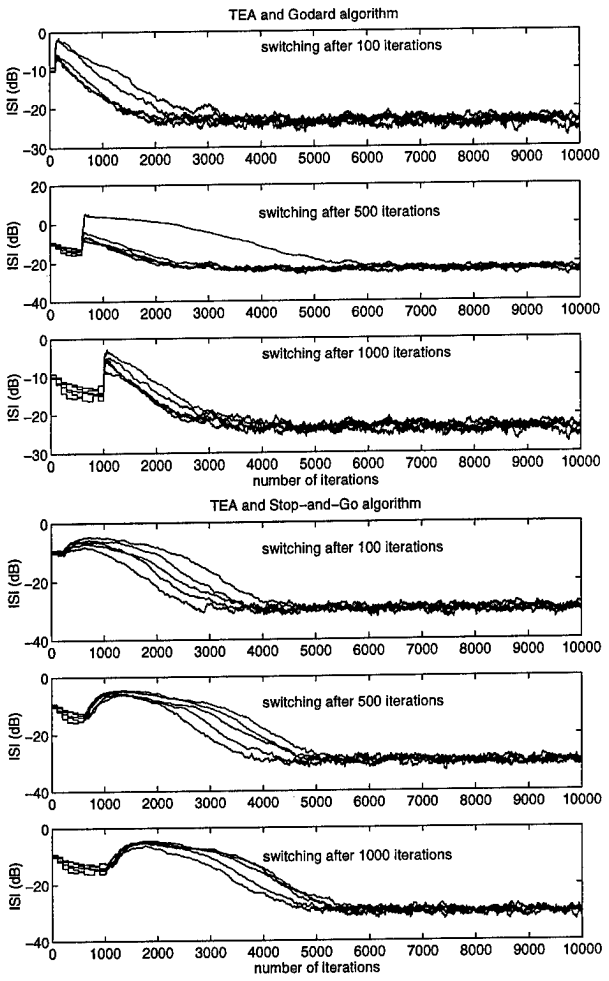


Figure 4. ISI performance for combined algorithm, channel 1 and 5 Monte Carlo simulations

to equalization is rather small. More important is that a correct estimate of the channel's phase relations is given which is retained by the Bussgang algorithm. This is of particular interest since Bussgang algorithms are not able to estimate the correct phase relations.

The second point of this paper is the derivation of a reliable criterion to detect channel changes. It turns out that the change of equalizer impulse response is able to give this measure. The parameters of the criterion have been found experimentally and depend on the algorithm used. Higher reliability is obtained using the Stop-and-Go algorithm.

Further research in this subject has to be done. This includes consideration of more realistic channels. A number of statistical models for both indoor and outdoor mobile communication exists [9]. The increase of ISI after switching from TEA to Bussgang algorithms has to be investigated

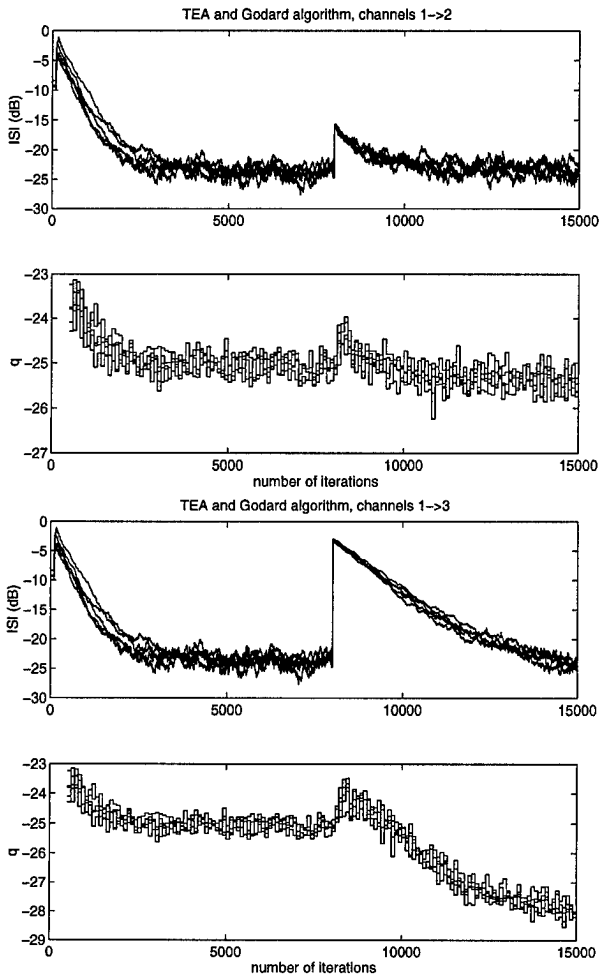


Figure 5. Multiple channels, TEA and Godard algorithm

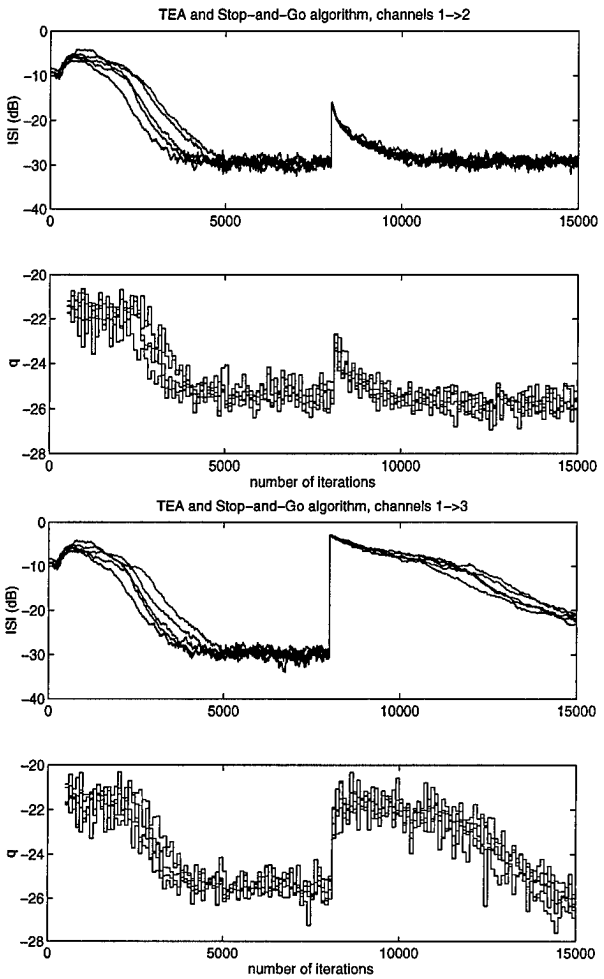


Figure 6. Multiple channels, TEA and Stop-and-Go algorithm

to improve the rate of convergence even further.

References

[1] Y. Sato. A method for self-recovering equalization for multilevel amplitude-modulation systems. *IEEE Trans. Commun.*, COM-23(6):679-682, June 1975.

[2] D. N. Godard. Self-recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems. *IEEE Trans. Commun.*, COM-28(11):1867-1875, November 1980.

[3] A. Benveniste and M. Goursat. Blind Equalizers. *IEEE Trans. Commun.*, COM-32(8):871-883, August 1984.

[4] G. Picchi and G. Prati. Blind Equalization and Carrier Recovery Using a Stop-and-Go Decision-Directed Algorithm. *IEEE Trans. Commun.*, COM-35(9):877-887, September 1986.

[5] J. K. Tugnait and U. Gummadavelli. Blind Channel Estimation. *J. Franklin Institute*, 333(B):311-337, 1996.

[6] D. Hatzinakos and C. L. Nikias. Blind Equalization Using a Tricepstrum-Based Algorithm. *IEEE Trans. Commun.*, 39(5):669-681.

[7] C. L. Nikias and A. P. Petropolu. *Higher Order Spectra Analysis*, Prentice Hall, Englewood Cliffs, 1993.

[8] J. G. Proakis. *Digital Communications*, McGrawHill, New York, 1983.

[9] P. F. M. Smulders. *Broadband wireless LANs: A feasibility Study*. PhD Thesis, Eindhoven University of Technology, 1995.

SYSTEM RECONSTRUCTION FROM SELECTED HOS REGIONS *

Haralambos Pozidis and Athina P. Petropulu
Electrical and Computer Engineering Department
Drexel University, Philadelphia, PA 19104
Tel. (215) 895-2358 Fax. (215) 895-1695
pozidis@cbis.ece.drexel.edu, athina@artemis.ece.drexel.edu

Abstract

System reconstruction from arbitrarily selected slices of the n-th order output spectrum is considered. We establish that unique identification of the impulse response of a system can be performed, up to a scalar and a circular shift, based on any two horizontal slices *of the discretized n-th order output spectrum, ($n \geq 3$), as long as the distance between the slices and the grid size satisfy a simple condition. For the special case of real systems, one slice suffices for reconstruction. The ability to select the slices to be used for reconstruction enables one to avoid regions of the n-th order spectrum where the estimation variance is high, or where the ideal bispectrum is expected to be zero, as in the case of bandlimited systems. We propose a mechanism for selecting slices that result in improved system estimates. We also demonstrate via simulations the superiority, in terms of estimation bias and variance, of the proposed method over existing approaches in the case of bandlimited systems.*

1. Introduction

System reconstruction has been a very active field of research in the recent years. Higher-order spectra (HOS) have been applied successfully to the problem of system reconstruction, mainly because of their ability to preserve the true system phase, and their robustness to additive Gaussian noise of unknown covariance. System reconstruction methods can be divided into two main categories: parametric and non-parametric. Parametric methods fit a specific model to the output observations and use the output statistics to identify the model parameters. Sensitivity to model order mismatch is their main disadvantage. Non-parametric methods reconstruct the system by recovering its Fourier phase and magnitude. In this paper we consider non-parametric system reconstruction methods.

Non-parametric methods can be divided into two main sub-categories, those that utilize the whole bispectrum information ([4], [1], [6], [8], [5]), and those that require bispectrum slices ([3], [2]). Methods that use fixed bispectrum slices cannot be applied for the reconstruction of bandlimited systems, since, depending on the system, the ideal bispectrum along these slices can be zero. Moreover, in the presence of noise and finite data records, bispectrum estimates along fixed slices can exhibit high estimation variance, and since single slices are used, there is no averaging mechanism to reduce estimation errors.

Several bispectrum slices within the bispectrum principal domain have been used in a method proposed in [1], and averaging was performed over the frequency response sample estimates. However, as the sample number decreases, the number of independent realizations of the corresponding frequency response sample over which averaging can be performed decreases too. As a result, low frequency samples of the frequency response exhibit considerable estimation variance in the presence of noise or data of finite lengths. Errors in low frequency samples can propagate to the remaining samples, since the method is iterative in nature. However, the method performs well in the case of wideband systems. The methods in [4] and [8] can also be viewed as methods that use several bispectrum slices. According to these methods, the non-redundant bispectrum is used to form a linear system of equations, which can be solved for the unknown frequency response samples. Although the system of equations is overdetermined and a solution could be obtained even if some slices were discarded, there is no mechanism for selecting the "best" slices to use, nor is there any guarantee that certain bispectrum regions can be avoided.

*This work was supported by NSF under grant MIP-9553227.

0-8186-8005-9/97 $10.00 © 1997 IEEE

In this paper, we consider the possibility of system reconstruction from *arbitrarily selected* higher-order spectra slices of the system output. We first establish that unique identification of the impulse response of an *arbitrary* system can be performed, up to a scalar and a circular shift, based on *any two horizontal slices* of the output discretized n-th order spectrum, of any order $n \geq 3$, of the system, as long as the distance between the slices and the grid size satisfy a simple condition. When the system is real, one slice suffices for system reconstruction. We then propose a method for system reconstruction based on a pair of selected HOS slices of the system output. In [7] it was shown that the obtained system estimates are asymptotically unbiased and consistent. We propose a mechanism to select the slices that will result in improved system estimates. We demonstrate via simulation examples the superiority of the proposed method over existing approaches, for the case of bandlimited systems.

2. System reconstruction from any pair of horizontal slices of the HOS of the system output

We define a *horizontal slice* of the n-th order spectrum $C_n^x(\omega_1, \omega_2, \ldots, \omega_n)$ of a signal $x(n)$ as the one-dimensional sequence that arises when we fix all indices $\omega_2, \ldots, \omega_n$ to certain real numbers, and allow ω_1 to take all possible values in $(-\pi, \pi]$. Throughout the paper, the term "slice" will be used instead of "horizontal slice". The distance between two slices $C_n^x(\omega, \alpha_1, \ldots, \alpha_{n-1})$ and $C_n^x(\omega, \beta_1, \ldots, \beta_{n-1})$ is defined as the l^1-norm of the vector $\boldsymbol{\alpha} - \boldsymbol{\beta}$, i.e., $\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_1 = \sum_{i=1}^{n-1} \mid \alpha_i - \beta_i \mid$, where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_{n-1}]^T$ and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_{n-1}]^T$.

In this paper we consider reconstruction from third-order spectra. A generalization of the results to the n-th order spectra case can be found in [7]. Consider a stationary process $x(n)$ given by:

$$x(n) = e(n) * h(n) + w(n), \qquad (1)$$

where $e(n)$ is an i.i.d. non-Gaussian process with zero mean and finite n-th order cumulant $\gamma_n^e \neq 0$, for $n \geq 2$; $w(n)$ is a stationary zero-mean Gaussian process of unknown covariance which is assumed independent of $e(n)$; $h(n)$ is the impulse response of an exponentially stable, generally mixed-phase, *complex* LTI system which has to be estimated from the output $x(n)$. It is initially assumed that $h(n)$ does not have zeros on the unit circle, however this assumption is relaxed later.

The frequency-domain bispectrum of $x(n)$ is given by

$$C_3^x(\omega_1, \omega_2) = \gamma_3^e H(\omega_1) H(\omega_2) H(-\omega_1 - \omega_2), \qquad (2)$$

with $H(\omega)$ denoting the frequency response of the system. The following proposition holds:

Proposition 1 *For the process $x(n)$ described by eq. (1), $h(n)$ is always identifiable, within a (complex) constant and a circular shift, from any two slices of the discretized output bispectrum, i.e, $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l_1)$ and $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l_2)$, $k = 0, \ldots, N-1$, if and only if N and $r = |l_1 - l_2|$ are relatively prime integers. If $h(n)$ is real, then it is identifiable, within a constant and a circular shift, based on a single slice of the discretized output bispectrum, i.e., $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l)$, if and only if N and $r = 2l$ are coprime integers.*

The proof of this proposition can be found in [7].

3. The reconstruction method

By evaluating (2) at discrete frequencies $\omega = \frac{2\pi}{N}k$, $k \in [0, \ldots, N-1]$, we obtain the discrete bispectrum of $x(n)$, i.e.,

$$C_3^x(k, l) = \gamma_3^e H(k) H(l) H(-k - l). \qquad (3)$$

Let us consider two slices of the discrete bispectrum at distance r, i.e. slices $(:, l)$ and $(:, l + r)$, with l arbitrarily chosen, i.e.,

$$\begin{aligned} C_3^x(k, l) &= \gamma_3^e H(k) H(l) H(-k - l) \\ C_3^x(k, l + r) &= \gamma_3^e H(k) H(l + r) H(-k - l - r) \end{aligned} \qquad (4)$$

Taking natural logarithms of both sides in (4) and subtracting we get:

$$\log C_3^x(k, l) - \log C_3^x(k, l + r) = \log \frac{H(l)}{H(l + r)} + \log \frac{H(-k - l)}{H(-k - l - r)} \qquad (5)$$

Substituting $m = -k - l$ in (5) we get:

$$\log H(m) = \log H(m - r) + \log H(l + r) - \log H(l) + \log C_3^x(-m - l, l) - \log C_3^x(-m - l, l + r), \qquad (6)$$

where it can be shown that:

$$\log H(l + r) - \log H(l) = \frac{1}{N} \sum_{k=0}^{N-1} \left[\log C_3^x(k, l + r) - \log C_3^x(k, l) \right]. \qquad (7)$$

From (6) and (7) we would be able to calculate the frequency response of the system recursively, provided that the initial conditions $\{\log H(k),\ k = 0, 1, \ldots, r-1\}$ were known a priori. However, a solution can still be obtained, without the need of any a priori information.

Let $\mathbf{h}_l = [\log H(1), \ldots, \log H(N-1)]^T$ be the $(N-1) \times 1$ vector of the unknown samples of the logarithm of the frequency response of the system. By substituting $s = m - r$ in (6), and letting s take the values $0, 1, \ldots, N-2$, we form the linear system of equations:

$$\mathbf{A}\mathbf{h}_l = \mathbf{c}, \tag{8}$$

where $\mathbf{c}$ is a $(N-1) \times 1$ vector of bispectrum values along the slices $(:, l)$ and $(:, l+r)$, with

$$\mathbf{c}_i = \log C_3^x(-i-r-l, l) - \log C_3^x(-i-r-l, l+r) + c_{l,r} \tag{9}$$

where $i = 0, 1, \ldots, N-2$, and

$$c_{l,r} = \log H(l+r) - \log H(l). \tag{10}$$

Matrix $\mathbf{A}$ is a sparse matrix with special structure; it is bidiagonal if $r = 1$ and tridiagonal otherwise, and contains 1's and (-1)'s. It can be proved, [7], that *matrix* $\mathbf{A}$ *is nonsingular if and only if* N *and* r *are relatively prime integers* and that *if* $\mathbf{A}$ *is nonsingular, then* $\det \mathbf{A} = 1$.

In this method, the logarithm of a pair of bispectrum slices is used to recover the logarithm of the frequency response of a system. Although the phase of the bispectrum appears implicitly in the expressions, only the principal argument of the bispectrum is actually needed. To see that, let $\tilde{\psi}(k, l)$ denote the principal argument of the phase sample $\psi(k, l)$. Then it holds:

$$\psi(k, l) = \tilde{\psi}(k, l) + 2\pi I(k, l), \tag{11}$$

where $I(k, l)$ is an integer function of k, l. The solution of (8), when $\mathbf{c}$ is computed based on $\tilde{\psi}(k, l)$, becomes:

$$\begin{aligned}\tilde{\mathbf{h}}_l &= \mathbf{A}^{-1}\left(\mathbf{c} + j2\pi\mathbf{I}_1(l) + j\frac{2\pi}{N}\mathbf{I}_2(l)\right) \\ &= \mathbf{h}_l + \frac{\mathrm{adj}(\mathbf{A})}{\det(\mathbf{A})}\left(j2\pi\mathbf{I}_1(l) + j\frac{2\pi}{N}\mathbf{I}_2(l)\right)\end{aligned} \tag{12}$$

where $\mathbf{I}_1(l), \mathbf{I}_2(l)$ are vectors of integer values. However, since $\det(\mathbf{A}) = 1$, the reconstructed frequency response of the system will only differ from the true one by a complex exponential factor of the form $e^{jm\frac{2\pi}{N}I(l)}$, where $I(l)$ is an integer function of l, which corresponds to a circular shift of the impulse response of the system. Therefore, $h(n)$ is reconstructed within a multiplicative scalar and a circular shift, as stated previously.

Throughout the derivation of the method it has been assumed that the system $h(n)$ does not have zeros on the unit circle, since then $H(k) = 0$ for some k. However, if $C_3^x(k, l) = 0$ for some (k, l), then we can change the spacing between samples, or equivalently re-estimate the bispectrum in a different grid of frequency points to surpass that problem, as suggested in [1] and [8].

The procedure outlined in this Section is valid for any pair of bispectrum slices, subject to the condition given in Proposition 1. Therefore, by using different pairs of slices, we can average the reconstructed systems in the time-domain (after scaling and shifting them appropriately), thus reducing the effects of noise and finite data lengths in the estimation of cumulants.

4. Simulations

In this Section we demonstrate the performance of the proposed method and compare it to the methods of [1] (BLW) and [8] (RG) for the reconstruction of bandlimited systems. The BLW and RG methods use several bispectrum slices within the bispectrum principal domain and are applicable to the reconstruction of bandlimited systems. Since the RG method was developed for the reconstruction of real systems, comparisons are performed for a real system.

We implemented all methods using the bispectrum of the system output for simplicity. We generated the input process $e(n)$ as an i.i.d. sequence with zero mean and nonzero skewness, and added zero-mean, Gaussian noise to the output of the system, at various signal-to-noise ratios (SNRs). The bispectrum of the output signal was estimated using the *indirect method*. Data of length L were segmented into non-overlapping records of length 256 symbols. The third-order cumulants of each record were estimated in a square grid of $(2M-1) \times (2M-1)$ lags, and then averaged over all $\frac{L}{256}$ records. Finally, a $N \times N$ two-dimensional FFT was applied on the averaged cumulants to obtain the discrete bispectrum of the output process.

Although the system considered was real-valued, we used two slices for the reconstruction procedure instead of only one slice (see Proposition 1), since an FFT size of $N = 64$, a power of 2, was used, in order to speed up computations. The reconstruction procedure was repeated using several different pairs of slices, and the estimated systems were averaged in the time-domain, in order to reduce noise effects.

The comparison method BLW determines the frequency response of the unknown system except for a linear-phase term, $e^{jk\phi(1)}$, where $\phi(\omega)$ is the phase response. Although this is not a drawback, it creates representation problems for the time-domain recovered signal $h(n)$ if $\phi(1)$ is not an integer, since it then cor-

responds to a non-integer time-delay. To circumvent that problem, we supplied the correct value of $\phi(1)$ to the BLW algorithm.

We generated a bandpass system as:

$$h(t) = 0.77^{|t|}\cos(2\pi 0.49\,t)+0.8(0.65)^{|t|}\sin(2\pi 0.38\,t+\frac{\pi}{5})$$

where $-4.5 \leq t \leq 3$ seconds. A discrete-time signal $h(n)$ of length 16 symbols was generated by sampling $h(t)$ every 0.5 seconds. Since the true length of the signal $h(n)$ is assumed unknown, we used $M = 20$ in the computation of third-order cumulants.

Although any pair of slices, as soon as it satisfies the identifiability conditions, should provide identical results, it was found in practice that some slices provide better results than others. Let us define the term "frequency content" as

$$\int_{-\pi}^{\pi} \mid C_3^x(\omega, l) \mid d\omega$$

Based on our experience with simulations, we found that the use of bispectrum slices with low *frequency content* resulted in poor performance; the opposite was also found to be true. In order to select the slices for better reconstruction, we run 100 different simulations and computed the frequency content of each slice at each run. The average frequency content over all runs, is shown in Fig. 1 where the shaded area indicates standard deviation. It can be seen that slices 11-18 exhibit a consistently higher frequency content than all others.

Then we run 100 simulations of the proposed method, using slices 13-14 (corresponding to "good" slices) and 23-24 (corresponding to "bad" slices), and the results are shown in Fig. 2 for $L = 2048$ output samples and SNR equal to ∞ and 10 dB. Clearly, the use of slices 13-14 produces superior results, both in bias and variance, compared to those of slices 23-24.

Next, we run 100 simulations of the proposed method, using averaging over slices 10-17 (in pairs of consecutive slices), and of methods BLW and RG. The results are shown in Figs. 3 and 4 for SNR of ∞ and 10 dB respectively. In both figures, graphs (a), (b) and (c) correspond to $L = 1024$, while (d), (e) and (f) to $L = 2048$ output samples used. The graphs on the left, center and right correspond to the proposed, BLW and RG methods respectively. It can be seen that method RG performs better than BLW both in terms of bias and variance, but both methods are outperformed by the proposed one. This can be attributed to the fact that the actual system $h(n)$ is bandpass, therefore its output bispectrum contains regions of low magnitude, where the useful signal information is significantly corrupted by noise. The inclusion of such regions in the reconstruction procedure is responsible for poor performance. On the other hand, selection of regions with higher signal information only, as in the proposed method, leads to better results.

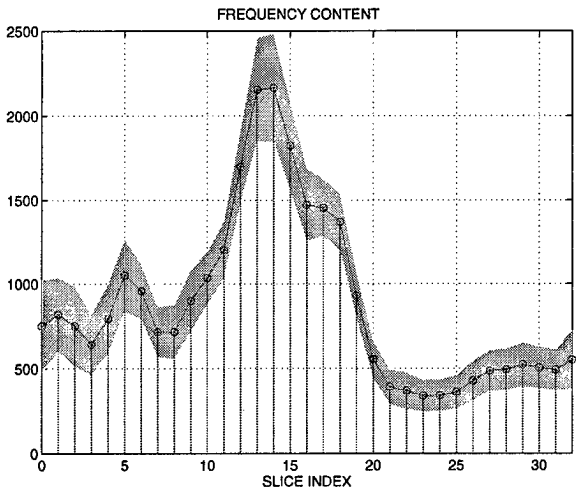


Figure 1. Frequency content for slices 1-33 of the output bispectrum of the system. Circles and solid line represent the average over 100 simulations, while shaded area indicates sample standard deviation.

5. Conclusions

A non-parametric method for system reconstruction based on HOS slices was presented. It was shown that an arbitrary system can be uniquely identified based on any two slices of the output discretized n-th order spectrum, with $n \geq 3$, as long as the distance between the slices and the grid size satisfy a simple condition. By using the logarithm of two n-th order spectrum slices, the logarithm of the frequency response of the system was obtained as the solution of a linear system of equations. A mechanism to select slices that result in improved system estimates was proposed. Simulation examples confirmed the superiority of the proposed method as compared to existing schemes, when tested on bandlimited systems. The flexibility of the proposed method in selecting the slices to be used for reconstruction could allow one to avoid regions of the n-th order spectrum where the useful system information is limited or distorted by noise.

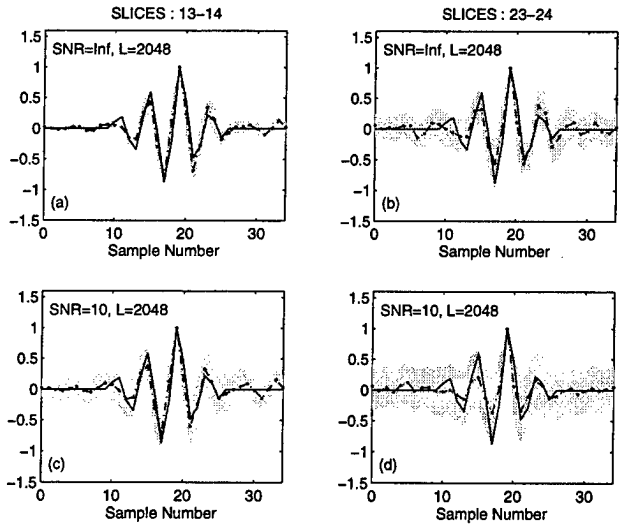


Figure 2. Results of the proposed method using slices 13-14 and 23-24 of the output bispectrum with SNR ∞ dB ((a) and (b)), and SNR 10 dB ((c) and (d)), and $L = 2048$ samples. Actual system is in solid lines, the average over 100 estimates in dash-dotted lines, and shaded area indicates standard deviation.

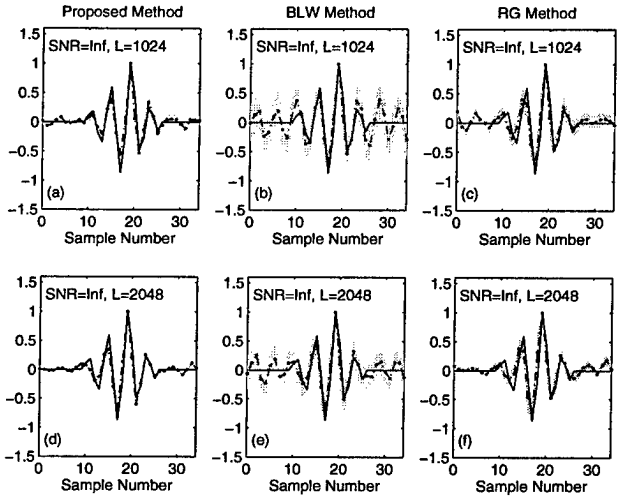


Figure 3. Comparison of the proposed, BLW and RG methods for $L = 1024$ ((a), (b) and (c)), and $L = 2048$ output samples ((d), (e) and (f)), and SNR=∞ dB. Actual system is in solid lines, the average over 100 estimates in dash-dotted lines, and shaded area indicates standard deviation.

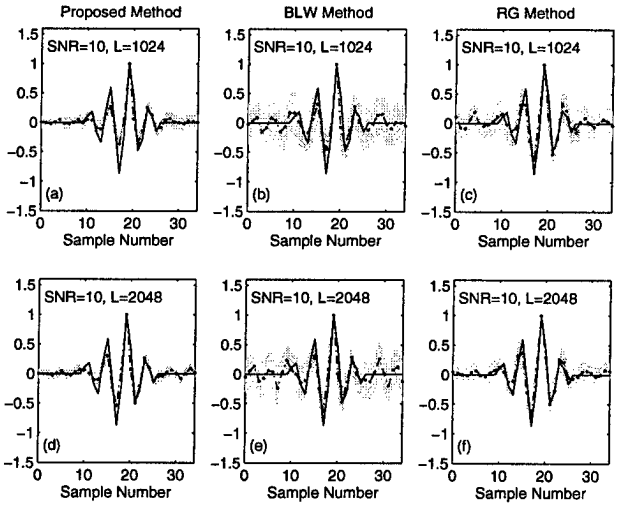


Figure 4. Comparison of the proposed, BLW and RG methods for $L = 1024$ ((a), (b) and (c)), and $L = 2048$ output samples ((d), (e) and (f)), and SNR=10 dB. Actual system is in solid lines, the average over 100 estimates in dash-dotted lines, and shaded area indicates standard deviation.

References

[1] H. Bartelt, A.W. Lohmann and B. Wirnitzer, "Phase and amplitude recovery from bispectra", *Applied Optics*, vol. 23, no. 18, pp. 3121-3129, Sept. 1984.

[2] S.A. Dianat and M.R. Raghuveer, "Fast Algorithms for Phase and Magnitude Reconstruction from Bispectra", *Opt. Eng.*, vol. 29, pp. 504-512, May 1990.

[3] K.S. Lii and M. Rosenblatt, "Deconvolution and Estimation of Transfer Function Phase and Coefficients for Nongaussian Linear Processes", *The Annals of Statistics*, vol. 10, pp. 1195-1208, 1982.

[4] T. Matsuoka and T.J. Ulrych, "Phase Estimation Using the Bispectrum", *Proc. IEEE*, vol. 72, pp. 1403-1411, Oct. 1984.

[5] R. Pan and C.L. Nikias, "The Complex Cepstrum of Higher Order Cumulants and Nonminimum Phase System Identification", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 186-205, Feb. 1988.

[6] A.P. Petropulu and C.L. Nikias, "Signal Reconstruction from the Phase of the Bispectrum", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 40, pp. 601-610, March 1992.

[7] H. Pozidis, and A.P. Petropulu, "System Reconstruction Based on Selected Regions of the Discretized HOS", *IEEE Trans. Sig. Proc.*, submitted in 1997.

[8] M. Rangoussi and G.B. Giannakis, "FIR Modeling Using Log-Bispectra: Weighted Least-Squares Algorithms and Performance Analysis", *IEEE Trans. Circ. and Sys.*, vol. 38, no. 3, pp. 281-296, Mar. 1991.

Blind System Identification in an Impulsive Environment

Jijun Yin and Athina P. Petropulu
Electrical and Computer Engineering Department
Drexel University, Philadelphia, PA 19104
Tel. (215) 895-2358 Fax. (215) 895-1695
yin@cbis.ece.drexel.edu, athina@artemis.ece.drexel.edu

Abstract

[1] *A method is presented for blind system identification in an impulsive environment, where the system output is described by a symmetric α-stable($S\alpha S$) law. The method employs either the phase or the magnitude of the recently proposed α-Spectrum [1] of the system output. It is much simpler than the method proposed in [1] that also relies on the phase or magnitude of the α-Spectrum, and provides the system cepstrum via closed form expressions.*

1. Introduction

The Gaussian distribution is frequently used in signal processing to model data. In most cases it is a reasonable assumption and can also be justified by central limit theorem. Many signals and noises encountered in the real world, however, are non-Gaussian. This paper considers non-Gaussian phenomena that can be characterized as impulsive, such as lightening in the atmosphere, switching transients in power lines, car ignitions, accidental hits in telephone lines, underwater acoustic signals [5]-[7]. Signals in this class exhibit sharp spikes or bursts corresponding to values that significantly deviate from the process mean. This class of signals can be described by an α−stable ($0< \alpha \leq 2$) model [4]. The main difference between the stable and the Gaussian densities is that the tails of first are heavier than those of the second. Frequently used members of the stable family are the Gaussian (α=2) and Cauchy (α=1) distributions.

The problem of modeling $S\alpha S$ processes has recently received attention [9],[1]. The $S\alpha S$ process is modeled as the response of a LTI system excited by white symmetric α-stable noise. In [1] the notion of α-Spectrum has been introduced and used to identify the system frequency response based on the system output only. In this paper we propose new methods for system identification based on α-Spectrum of the system output, that rely on a simple relationship between the cepstrum of the system impulse response and the inverse Fourier transform of the phase, or the log-magnitude of the α-Spectrum. The main advantage of the proposed approach over the ones in [1] is simplicity of expressions and computations, which renders them less susceptible to computation errors.

[1]This work was supported by NSF under grant MIP-9553227

2. Mathematical Background

Stable distributions are the only class of distributions that can be the limit distributions of sums of i.i.d random variables (Generalized central limit theorem). Their density function does not have a closed form, thus they are usually described by their characteristic function:

$$\Phi(\omega) = E\{exp(j\omega x)\}$$

$$= \begin{cases} exp(ja\omega - \gamma|\omega|^{\alpha}(1 - j\beta\frac{\omega}{|\omega|}tan\frac{\pi\alpha}{2})) & \alpha \neq 1 \\ exp(ja\omega - \gamma|\omega|^{\alpha}(1 - j\beta\frac{2}{\pi}\frac{\omega}{|\omega|}ln|\omega|)) & \alpha = 1. \end{cases} \quad (1)$$

where $\alpha \in (0,2]$ is the characteristic exponent, $\beta \in [-1,1]$ is the symmetry index, $\gamma > 0$ is the spread parameter, and a is the location parameter. A symmetric α-stable $S\alpha S$ process is described by the characteristic function given in (1), where $\beta = 0$.

A *standard* $S\alpha S$ processes is a $S\alpha S$ process with $\gamma = 1$. A complex random variable is $S\alpha S$ if its real and imaginary part are jointly $S\alpha S$. An important complex $S\alpha S$ random variables is the isotropic complex $S\alpha S$ random variable with characteristic function $\Phi(\omega_1, \omega_2) = exp(-\gamma(\omega_1^2 + \omega_2^2)^{\frac{\alpha}{2}})$.

0-8186-8005-9/97 $10.00 © 1997 IEEE

α-stable processes have finite moments of order p, where $-1 < p < \alpha$. For processes with $\alpha < 2$ second- or higher-order moments are not defined. Fractional-order moments, however, exist and their definitions can be found in [4]. Similarly, for such process the covariance is not defined. Its fractional-order equivalent, the covariation is defined (only for stable processes with $\alpha > 1$) and its definition can be found in [4]. The covariation estimator of any two jointly $S\alpha S$ random variables X and Y is [1]:

$$[X,Y]_\alpha = \frac{E\{XY^{\langle p-1\rangle}\}}{E\{|Y|^p\}}[Y,Y]_\alpha, \ p<\alpha, \qquad (2)$$

where

$$Y^{\langle p\rangle} = \begin{cases} |Y|^{p-1}Y^*, \ Y \ is \ complex; \\ |Y|^p sign(Y), \ Y \ is \ real, \end{cases} \qquad (3)$$

and $[Y,Y]_\alpha$ equals the dispersion of Y, when Y is real or isotropic complex, while it equals some unknown real number in all other cases for Y [11].

The following are some properties of the covariation: 1) linearity exists for the first term: if X_1, X_2, Y are jointly SαS, then

$$[aX_1 + bX_2, Y]_\alpha = a[X_1,Y]_\alpha + b[X_2,Y]_\alpha. \qquad (4)$$

2) pseudo-linearity exists for the second term: if Y_1, Y_2 are independent and X, Y_1, Y_2 are jointly SαS, then

$$[X, aY_1+bY_2]_\alpha = a^{\langle\alpha-1\rangle}[X,Y_1]_\alpha + b^{\langle\alpha-1\rangle}[X,Y_2]_\alpha. \qquad (5)$$

3) if X, Y are independent, then $[X,Y]_\alpha = 0$.

3. Blind System identification

Let $y(n)$ be the output of a LTI system excited by an i.i.d. real or isotropic complex standard $S\alpha S$ random input $x(n)$. Let $h(n)$, $H(z)$ be the system impulse response and transfer function, respectively.

The $\alpha-$Spectrum of the system output was defined in [1] to be the covariation between $y(n)$ and a function of $y(n)$, i.e.,

$$S_\alpha(z) = [y(n), w_z(n)]_\alpha, \qquad (6)$$

where

$$w_z(n) = \sum_{i=-q}^{i=q} y(n-i)z^i, \qquad (7)$$

with q being the length of the FIR approximation of $h(n)$.

In [1] it was also shown that the α-Spectrum is related to the system transfer function as:

$$\begin{aligned} S_\alpha(z) &= H((\frac{1}{z})^{\langle\alpha-1\rangle})(H(z))^{\langle\alpha-1\rangle} \\ &= H(\frac{z}{|z|^\alpha})\frac{|H(z)|^\alpha}{H(z)}. \end{aligned} \qquad (8)$$

Evaluating the α-Spectrum $S_\alpha(z)$ on the unit circle ($|z|=1$) yields

$$S_\alpha(e^{j\omega}) = H(e^{j\omega})(H(e^{j\omega}))^{\langle\alpha-1\rangle} = |H(e^{j\omega})|^\alpha, \qquad (9)$$

from where we get the system magnitude response only. In order to recover the channel, we need to evaluate the $\alpha-$Spectrum on circles other than the unit circle, i.e., on a circle with radius r ($r \neq 1$). Taking the logarithm of both sides of (8) yields:

$$\begin{aligned} \log S_\alpha(re^{j\omega}) &= \log|S_\alpha(re^{j\omega})| + jarg\{S_\alpha(re^{j\omega})\} \\ &= log|H(r^{1-\alpha}e^{j\omega})| + (\alpha-1)log|H(re^{j\omega})| \\ &\quad +j(\phi(r^{1-\alpha}e^{j\omega}) - \phi(re^{j\omega})), \end{aligned} \qquad (10)$$

where $\phi(\omega)$ is the system phase response. In order to ensure the existence of complex cepstrum of the α-Spectrum, r and $r^{1-\alpha}$ must be taken in the range (R_{min}, R_{max}), where R_{min} and R_{max} denote respectively the location of the zeros with the minimum and the maximum magnitude .

Let $f_1(n), f_2(n)$ denote respectively the inverse Fourier transform of the phase of $\alpha-$Spectrum evaluated on circles with radii r_1 and r_2. Let us define:

$$\begin{aligned} f_i(n) &= F^{-1}\{arg\{S_\alpha(re^{j\omega})\}\} \\ &= F^{-1}\{\phi(r_i^{1-\alpha}e^{j\omega}) - \phi(r_i e^{j\omega})\} \\ &= F^{-1}\{Im\{logH(r_i^{1-\alpha}e^{j\omega})H^*(r_i e^{j\omega})\}\} \\ &= \frac{a_i(n)\widehat{h}(n) + b_i(n)\widehat{h}(-n)}{2j}, \ i=1,2 \end{aligned} \qquad (11)$$

where

$$a_i(n) = r_i^{(\alpha-1)n} - r_i^{-n} \qquad (12)$$

$$b_i(n) = r_i^n - r_i^{-(\alpha-1)n}, \qquad (13)$$

and $\widehat{h}(n)$ denotes the complex cepstrum of the system impulse response. In (11) it was used the fact that the complex cepstrum corresponding to $H(re^{j\omega})$ equals $r^{-n}\widehat{h}(n)$.

Similarly, we define:

$$\begin{aligned} g_i(n) &= F^{-1}\{log|S_\alpha(re^{j\omega})|\} \\ &= F^{-1}\{log|H(r^{1-\alpha}e^{j\omega})| + (\alpha-1)log|H(re^{j\omega})|\} \\ &= \frac{a_i'(n)\widehat{h}(n) + b_i'(n)\widehat{h}(-n)}{2}, \ i=1,2 \end{aligned} \qquad (14)$$

where

$$a_i'(n) = r_i^{(\alpha-1)n} + (\alpha-1)r_i^{-n} \quad (15)$$

$$b_i'(n) = r_i^{(1-\alpha)n} + (\alpha-1)r_i^{n}, \quad (16)$$

Thus, the cepstrum of the system can be reconstructed using the α-Spectrum phase by:

$$\widehat{h}(n) = \frac{2j[b_2(n)f_1(n) - b_1(n)f_2(n)]}{a_1(n)b_2(n) - a_2(n)b_1(n)}, \quad for \quad n \neq 0, \quad (17)$$

or using the α-Spectrum magnitude by:

$$\widehat{h}(n) = \frac{2[b_2'(n)g_1(n) - b_1'(n)g_2(n)]}{a_1'(n)b_2'(n) - a_2'(n)b_1'(n)}, \quad for \quad n \neq 0, \quad (18)$$

If $\alpha = 2$ the denominator of both (17) and (18) becomes zero. Thus, for $\alpha = 2$ system identification is not possible, which comes at no surprise since in this case the system input is a Gaussian process. It can be verified that when $\alpha \neq 2$, the denominator of (17) is different than zero for any value of n, and for any positive r1 and r2 not equal to 1, as long as $r_1 \neq r_2$. Thus, reconstruction from the α-Spectrum phase can be carried out based on any two circles with radii r_1 and r_2 ($r_1 \neq r_2$). For $\alpha > 1$ the denominator of (18) does become zero depending on the values of r_1, r_2 and n. It can, however, be verified that it is nonzero for every n , as long as $r_1 \neq r_2$,

Based on its complex cepstrum, $h(n)$ can be recovered within a scalar and a delay based on inverse cepstrum operations.

Implementation Issues

- The covariation needed in the estimation of the α−Spectrum is obtained as:

$$[y(n), w_z(n)]_\alpha = \frac{E\{y(n)w_z(n)^{\langle p-1\rangle}\}}{E\{|w_z(n)|^p\}}[w_z(n), w_z(n)]_\alpha. \quad (19)$$

  In the case of a real $S\alpha S$ input, $w_z(n)$, in general, is not real nor isotropic complex, thus $[w_z(n), w_z(n)]_\alpha$ cannot be estimated [11]. In such a case, although reconstruction from the α-Spectrum magnitude would not be possible, reconstruction from the α-Spectrum phase would be possible, since this unknown term is real, thus does not contribute to the phase of the α-Spectrum required in (17).

- The fractional moment order, p, needed in the covariation estimate, can be taken to be anywhere in the range $(0, \alpha)$. However, it was demonstrated in [10] that when p is in the range $(1/2, \alpha/2)$, (19) leads to better covariation estimates.

- When the length q, required in (7), is not known, a length overestimate can be used.

- For the computation of the phase of the α-Spectrum phase unwrapping is required. The time domain equivalent of the α-Spectrum is real, hence, phase unwrapping can be carried out as usual.

- In practice, we should choose r1 and r2 that are close to 1 unless we have some priori knowledge about the positions of poles and zeros of the system.

4. Simulation results

In this section we demonstrate the validity of the algorithm proposed in this paper. The channel was taken to be:

$$H(z) = 1.5924z^{-1}(1 - (0.5462 + 0.3758j)z^{-1}) \\ (1 - (0.5462 - 0.3758j)z^{-1})(1 + 0.628z), \quad (20)$$

which corresponds to a nonminimum phase FIR system of order 3.

Two cases of input were considered, complex isotropic and real. In both cases the input was taken to be white i.i.d., 5,000 sample long, generated based on $\alpha = 1.5$ and $\gamma_x = 1$. The system reconstruction was performed based on 20 independent input realizations. The covariation estimator was obtained based on $p = 0.6$ for the first input and $p = 0.7$ for the second input, α was taken to be equal to its true value, and the length of the FIR approximation of $h(n)$ needed in (7) was taken to be $q = 10$. The radii of the two circles are $r_1 = 1.2$ and $r_2 = \frac{1}{r_1}$. The estimation results of the 20 Monte Carlo simulations for the complex isotropic input are shown in Fig. 1, while those corresponding to the real input are shown in Fig. 2.

Another set of simulations was performed based on the estimated α. In each Monte Carlo run, α was estimated via the method of [1]. The estimated phase and magnitude response based on 20 Monte Carlo runs are shown in Fig. 3. The mean of the estimated α was 1.3946.

References

[1] Xinyu Ma, C. L. Nikias ,"Parameter Estimation and Blind Channel Identification in Impulsive Signal Environments," *IEEE Trans. on Signal*

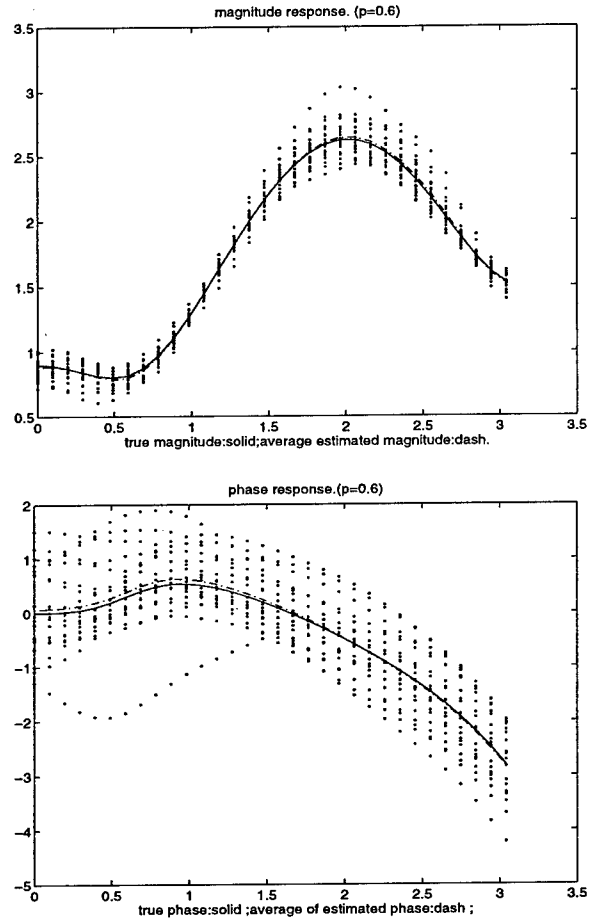


Figure 1. Estimation of the mixed-phase FIR channel with white i.i.d. isotropic complex SαS input (5000 samples, the radii of the two circles are 1.2 and $\frac{1}{1.2}$, using true alpha).

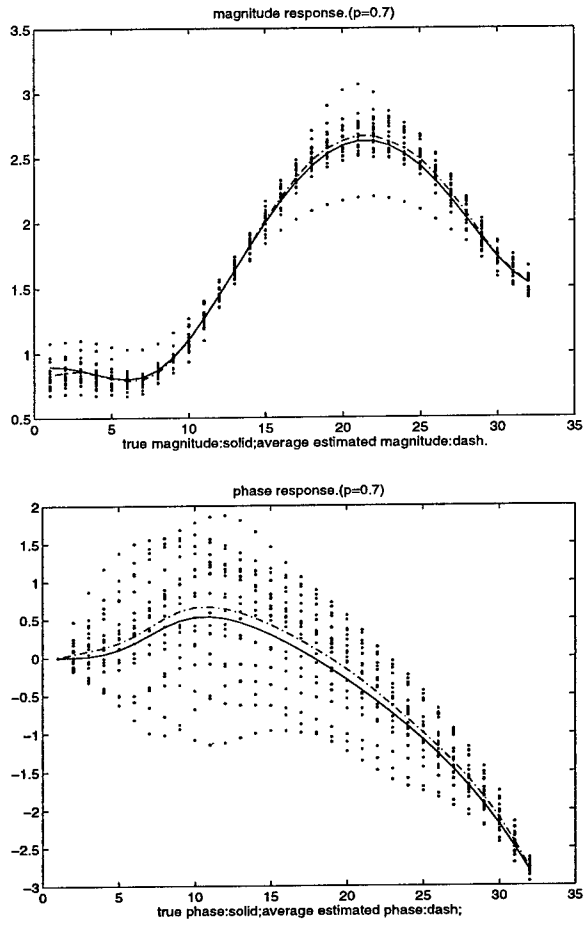


Figure 2. Estimation of the mixed-phase FIR channel with white i.i.d. real SαS input (5000 samples, the radii of the two circles are 1.2 and $\frac{1}{1.2}$, using estimated alpha).

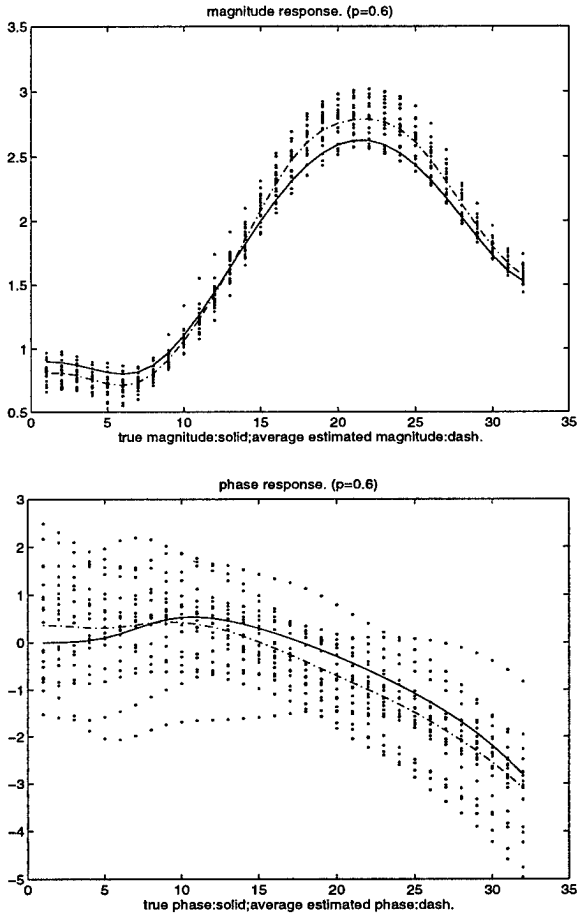


Figure 3. Estimation of the mixed-phase FIR channel with white i.i.d. isotropic SαS input (5000 samples, the radii of the two circles are 1.2 and $\frac{1}{1.2}$, using estimated alpha).

Processing, vol. 43, no. 12, pp. 2884-2897, December 1995.

[2] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random processes: Stochastic Models with Infinite Variance*, New York: Chapman and Hall, 1994.

[3] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*, Prentice Hall Inc., 1993.

[4] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, New York: Wiley, 1995.

[5] P. Mertz ,"model of impulsive noise for data transmission," *IRE Trans. Comm. Syst.*, Vol. CS-9, pp.130-137, 1961.

[6] B. W. Stuck and B. Kleiner, "A statistical analysis of telephone noise," *Bell Systems Technical Journal*, vol. 53, pp. 1262-1320, 1974.

[7] E. J. Wegman, S. G. Schwartz, and J. B. Thomas, Eds., *Topics in Non-Gaussian Signal Processing*, New York: Academic, 1989.

[8] S. Cambanis and G. Miller, "Linear Problems in P-th Order stable Processes," *SIAM J.Appl. Math.*, Vol. 41, pp.43-69, Aug, 1981.

[9] M. I. Gurelli and C. L. Nikias, "Blind Identification and Array processing Application of Generalized Higher-order statistics," Proc. of Milcom, 1996.

[10] P. Tsakalides and C. L. Nikias, "The Robust Covariation-Based MUSIC (ROC-MUSIC) Algorithm for Bearing Estimation in Impulsive Noise Environments," *IEEE Trans. on Signal Processing*, vol. 44, no. 7, pp. 1623-1633, July 1996.

[11] M. Hernandez and C. Houdre, "Disjointness results for some classes of inequalities," *Studia Mathematica*, vol. 105, no. 3, pp. 235-252, 1993.

Non-Causal ARMA Model Identification by Maximizing the Kurtosis

J.-L. VUATTOUX and E. LE CARPENTIER
Institut de Recherche en Cybernétique de Nantes, UMR C.N.R.S. n° 6597,
Ecole Centrale de Nantes / Université de Nantes,
1 rue de la Noë, B.P. 92101, 44321 Nantes Cedex 03, France,
E-mail : Jean-Luc.Vuattoux@lan.ec-nantes.fr

Abstract

The problem of estimating the parameters of a non-causal ARMA system, driven by an unobservable input noise is addressed. We propose a method based on a generalized version of the prediction error minimum variance approach and on the maximum kurtosis properties. Firstly, a spectrally equivalent (SE) model is identified with the generalized minimum variance approach. Secondly, the kurtosis allows us to identify the phase of the true model by localizing its zeros and poles from the SE model. Finally, we propose a new method which is a closed-loop form of the preceding method allowing to improve the accuracy of the parameter estimation and to obtain a better reconstruction of the estimated model phase. Simulations results seems to confirm the good behavior of the proposed methods compared to methods using higher order statistics.

1. Introduction

This paper deals with parameter identification of a non-causal stable linear time invariant (LTI) system driven by an unknown zero-mean, stationary, non-gaussian white noise (NGWN). Many methods, using principally higher-order cumulants, were proposed in the literature [8], [3], [7], [5]. Here, we are interested in a method proposed by Boumahdi [1], which uses the second-order statistics (SOS) information and the maximum kurtosis property to identify non-causal MA or AR, or causal ARMA models. The interest of this approach is that it provides parameter estimates with lower variance than if higher-order statistics (HOS) are used.

We propose to extend this method to non-causal ARMA models and to improve the estimation accuracy of this method in the sense of the error covariance matrix of the parameter estimate. We identify firstly the spectrally equivalent (SE) minimum phase model by using SOS, then we use the maximum kurtosis property to recover the phase of the true model (non-causal MA, AR or ARMA model). But our approach differs from the one of Boumahdi [1] by the fact that we use a prediction error method (PEM) to identify the SE minimum phase model. Effectively, we propose a generalization of the minimum variance approach to the case of non-causal ARMA models by introducing a model separating the causal and anti-causal parts. And finally, to improve the estimation accuracy, we introduce a recurrence on the two steps of variance minimization and kurtosis maximization, this being based on a simple idea.

We present in section 2 the ARMA models and the hypotheses used. In section 3, the minimum variance approach is detailed. In section 4, the maximum kurtosis property is presented. We proposed a new identification scheme in section 5. Finally, simulation results are summarized in section 6. In section 7, a conclusion is given.

2. ARMA models and hypotheses

The observed process $\mathbf{y} = \{y[n]\}_{n\in\mathbf{Z}}$ is modeled as the output of a discrete non-causal stable LTI system of impulse

0-8186-8005-9/97 $10.00 © 1997 IEEE

response (IR) h_{θ_0} with unobservable, zero-mean, stationary, non-gaussian white input $\mathbf{x} = \{x[n]\}_{n\in\mathbb{Z}}$:

$$\mathbf{y} = h_{\theta_0} * \mathbf{x} \tag{1}$$

where $*$ is the convolution sum and θ_0 is the actual parameter in the parameterized form of the IR h_θ with z transform $\mathcal{Z}h_\theta(z)$ defined by :

$$\mathcal{Z}h_\theta(z) = \frac{\mathcal{Z}\beta(z)}{\mathcal{Z}\alpha(z)} = \frac{1+\sum_{i=1}^{q}\beta[i]\,z^{-i}}{1+\sum_{i=1}^{p}\alpha[i]\,z^{-i}} \tag{2}$$

with $\theta = [\beta[1]\ldots\beta[q]\ \alpha[1]\ldots\alpha[p]]^T$.
The IR h_θ is supposed to be stable with stable inverse. So, we suppose that the ARMA(p,q) model, of orders p and q known, is non-causal, free of zero-pole cancelations, with no zeros or poles on the unit circle (UC).

Let $(P_i)_{1\le i\le n_a}$ $\left((\tilde{P}_i)_{1\le i\le n_{\tilde{a}}}\right)$ be the system poles inside (outside) the UC with $n_a + n_{\tilde{a}} = p$, and let $(Z_i)_{1\le i\le n_b}$ $\left((\tilde{Z}_i)_{1\le i\le n_{\tilde{b}}}\right)$ be the system zeros inside (outside) the UC with $n_b + n_{\tilde{b}} = q$. We can express $\mathcal{Z}h_\theta(z)$ (2) by :

$$\begin{aligned}\mathcal{Z}h_\theta(z) &= \frac{\prod_{i=1}^{n_b}(1-Z_i\,z^{-1})\cdot\prod_{i=1}^{n_{\tilde{b}}}(1-\tilde{Z}_i\,z^{-1})}{\prod_{i=1}^{n_a}(1-P_i\,z^{-1})\cdot\prod_{i=1}^{n_{\tilde{a}}}(1-\tilde{P}_i\,z^{-1})}\\ &= G_\theta\cdot z^{-s_\theta}\cdot\mathcal{Z}h'_\theta(z)\end{aligned} \tag{3}$$

with $G_\theta = \prod_{i=1}^{n_{\tilde{b}}}(-\tilde{Z}_i)/\prod_{i=1}^{n_{\tilde{a}}}(-\tilde{P}_i)$, $s_\theta = n_{\tilde{b}} - n_{\tilde{a}}$ and

$$\begin{aligned}\mathcal{Z}h'_\theta(z) &= \frac{\prod_{i=1}^{n_b}(1-Z_i\,z^{-1})}{\prod_{i=1}^{n_a}(1-P_i\,z^{-1})}\cdot\frac{\prod_{i=1}^{n_{\tilde{b}}}(1-\tilde{Z}_i^{-1}\,z)}{\prod_{i=1}^{n_{\tilde{a}}}(1-\tilde{P}_i^{-1}\,z)}\\ &= \mathcal{Z}h'_{ca}(z)\cdot\mathcal{Z}h'_{ac}(z^{-1})\end{aligned} \tag{4}$$

where $\mathcal{Z}h'_{ca}(z)$ ($\mathcal{Z}h'_{ac}(z^{-1})$) is the causal (anti-causal) part of the ARMA model. n_a and $n_{\tilde{a}}$ (n_b and $n_{\tilde{b}}$) denote the number of poles (zeros) of $\mathcal{Z}h'_{ca}(z)$ and $\mathcal{Z}h'_{ac}(z^{-1})$, respectively.

Then the output process $\mathbf{y}$ can be modeled of equivalent manner to (1), as showed on the figure 1, by :

$$\mathbf{y} = h'_{\theta_0} * \mathbf{x}' \tag{5}$$

where $\mathcal{Z}x'(z) = G_{\theta_0}\cdot z^{-s_{\theta_0}}\cdot\mathcal{Z}x(z)$. $\mathbf{x}'$ is then a scaled shifted version of $\mathbf{x}$.

Note that this alternate representation (4)-(5) of a non-causal non-minimum phase ARMA system is used in [6] to develop maximum likelihood estimates of ARMA models. Moreover, we showed that this modeling verifies some interesting properties.

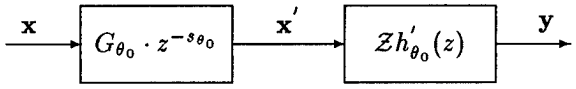


Figure 1. Model of the system.

Indeed, the power spectrum of $\mathbf{y}$ is defined by :

$$\begin{aligned}S_y(\omega) &= S_{x'}(\omega)\,|\mathcal{Z}h'(e^{\jmath\omega})|^2\\ &= \nu'\ \mathcal{Z}h'_{ca}(e^{\jmath\omega})\ \mathcal{Z}h'_{ac}(e^{-\jmath\omega})\ \mathcal{Z}h'_{ca}(e^{-\jmath\omega})\ \mathcal{Z}h'_{ac}(e^{\jmath\omega})\end{aligned}$$

where $S_{x'}(\omega)$ is the power spectrum of the white input $\mathbf{x}'$ and ν' its variance.
We can show easily the property which follows :

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_y(\omega)d\omega = \log\nu' \tag{6}$$

by using the fact that $\frac{1}{2\pi}\int_{-\pi}^{\pi}\log(1+a\,e^{\pm\jmath\omega})\,d\omega = 0$ for all $a\in\mathbb{C}$ such that $|a|<1$, where $\mathbb{C}$ denotes the complex numbers set. This formula is simply demonstrated by using power series expansion of the logarithm. It permits us to prove that $\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\mathcal{Z}h'_{ca}(e^{\pm\jmath\omega})d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\mathcal{Z}h'_{ac}(e^{\pm\jmath\omega})d\omega = 0$, and to obtain the relation (6). It is important to note that the integral of the logarithm of the power spectral density of $\mathbf{y}$ is generally not equal to the logarithm of the variance of $\mathbf{x}$, except in the causal minimum phase case.

Now, by using the property (6), we demonstrate that :

$$\mathrm{E}\{y^2[n]\} \ge \nu' \tag{7}$$

where $\mathrm{E}\{y^2[n]\}$ is the variance of $\mathbf{y}$.
Because we have :

$$\begin{cases}\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_y(\omega)d\omega &= \log\nu'\\ \frac{1}{2\pi}\int_{-\pi}^{\pi}S_y(\omega)d\omega &= \mathrm{E}\{y^2[n]\}\end{cases}$$

So, we obtain :

$$\frac{\mathrm{E}\{y^2[n]\}}{\nu'} - 1 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{S_y(\omega)}{\nu'} - \log\frac{S_y(\omega)}{\nu'} - 1\right)d\omega$$

But $u-\log u-1 > 0$ if $u\neq 1$ and $u-\log u-1 = 0$ if $u = 1$. Like this, $\frac{\mathrm{E}\{y^2[n]\}}{\nu'} - 1 > 0$, hence $\mathrm{E}\{y^2[n]\} > \nu'$ and we have $\mathrm{E}\{y^2[n]\} = \nu'$ if and only if for all ω, $S_y(\omega) = \nu'$, i.e. if $\mathcal{Z}h'_{\theta_0}(z)$ is the transfer function of an all-pass filter. Note that we cannot say anything about the variance of the input $\mathbf{x}$.

3. Minimum variance criterion

For all parameters vector θ and all signal $\mathbf{y}$, we defined the sequences $\varepsilon(\theta) = \{\varepsilon(\theta)[n]\}_{n\in\mathbf{Z}}$ and $\varepsilon'(\theta) = \{\varepsilon'(\theta)[n]\}_{n\in\mathbf{Z}}$ on the figure 2, where $h_\theta'^{-1}$ is the inverse of h'_θ for the convolution sum. The deconvolution is performed by forward and backward filtering (see [7]).

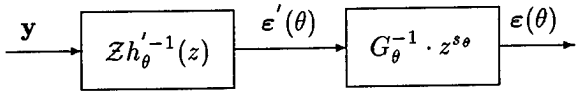


Figure 2. Deconvolution scheme.

We can prove, similarly to the property (7), that

$$\mathrm{E}\left\{\varepsilon'^2(\theta)[n]\right\} \geq \nu'$$

where the equality is obtained if and only if $h'_{\theta_0} * h_\theta'^{-1}$ is the impulse response of an all-pass filter. But we cannot say anything about the variance of $\varepsilon(\theta)$, because G_θ depends on the zeros and poles of h_θ which are outside the unit circle. (Obviously, second order methods are not sufficient to identify non-causal models).

Then, the minimization of the following criterion :

$$J(\theta) = \mathrm{E}\left\{\varepsilon'^2(\theta)[n]\right\} \tag{8}$$

gives at least $2^{(p_r+q_r+\frac{p_c+q_c}{2})}$ solutions, p_r (p_c) and q_r (q_c) being the numbers of real (complex) poles and zeros, respectively, of the true model parameterized by θ_0. These solutions correspond to the SE models (obtained by replacing poles and zeros of the actual model by their conjugated inverses). But this is true only if the actual ARMA model does not contain an all-pass factor.

4. Maximum kurtosis property

Assuming that $\mathbf{x}'$ is fourth order white noise, the normalized kurtosis, noted $\overline{\gamma}_{4,\varepsilon'}$, of the prediction error $\varepsilon'(\theta)$ is given by :

$$\overline{\gamma}_{4,\varepsilon'} = \frac{\gamma_{4,\varepsilon'}}{\gamma^2_{2,\varepsilon'}} = \overline{\gamma}_{4,x'} \frac{\sum_{i=-\infty}^{+\infty}(h'_{\theta_0} * h_\theta'^{-1})^4[i]}{(\sum_{i=-\infty}^{+\infty}(h'_{\theta_0} * h_\theta'^{-1})^2[i])^2} \tag{9}$$

where $\gamma_{2,\varepsilon'}$ is the variance of $\varepsilon'(\theta)$, $\gamma_{4,\varepsilon'}$ its kurtosis (defined as the value at the origin of its fourth order cumulant) and $\overline{\gamma}_{4,x'}$ is the normalized kurtosis of $\mathbf{x}'$. Its value is "a distance" between the statistics of $\mathbf{x}'$ and the gaussianity.

So, the relation (9) leads to :

$$|\overline{\gamma}_{4,\varepsilon'}| \leq |\overline{\gamma}_{4,x'}| \tag{10}$$

if we supposed that h'_{θ_0} and h'_θ are stable with stable inverses. The equality is obtained if and only if $h'_\theta = h'_{\theta_0}$.

The relation (10) implies that all filtered version of a NGWN is "more gaussian" than the noise itself. So, as proposed by Donoho [2], a mean for a good non-minimum phase blind identification is to maximize the absolute value of the normalized kurtosis of the estimated input.

The method proposed by Boumahdi [1] consists in two steps. Firstly, identification of the SE minimum phase model by using classical methods assuming causality and minimum phase. Then, the estimated model is the one which maximizes $|\overline{\gamma}_{4,\varepsilon'}|$ among the $2^{(p_r+q_r+\frac{p_c+q_c}{2})}$ SE models. The interest of this method, that we noted the "maximum kurtosis" (MK) method, is that we estimate the numerical values of the non-causal ARMA parameters by using only SOS, then with lowest estimation variance, the HOS (kurtosis) being used as criterion of choice among all the SE models to recover the phase of the true model.

5. Improvement of the method

In practice, given N consecutive samples of the system output $\{y[n]\}_{n\in[1,N]}$, we want to estimate the true parameters θ_0. So, to estimate a SE model, we minimize the following criterion :

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left[\varepsilon'(\theta)[i]\right]^2 \tag{11}$$

Theoretically, the knowledge of the minimum of (8) corresponding to a SE model, allows us to determine the minima of (8) corresponding to the others SE models. But it is false in practice with the criterion (11), due to the finite signal length. The principle of the improvement of the MK method is to minimize again the criterion (11), in the direction of the minimum for which the phase correspond to the true model (after maximization of the kurtosis). It leads to the following algorithm :

- *Initialization :* Minimization of $J(\theta)$ (convergence to a local minimum).
- *Recurrence* on the two following steps :

a) determination of the model phase by maximization of $|\overline{\gamma}_{4,\varepsilon'}|$ among all the SE models,

b) minimization of the criterion (11) with initialization by the model maximizing $|\overline{\gamma}_{4,\varepsilon'}|$.

We stop the recurrence when we obtain the same result on two successive iterations. This recurrence allows to obtain a better reconstruction of the phase of the estimated model and to improve parameter estimation accuracy, compared to the method MK and those given in [1]. Note that this method is nothing but the maximum likelihood estimation method under gaussian hypothesis, which gives local minima around all the models SE to the actual one. The maximization of the kurtosis permits to determine the actual phase model. Hence, the covariance matrix of parameters estimates can be calculated using formula found in [6]. The main disadvantage of the "improved maximum kurtosis" (IMK) method is the impossibility to identify ARMA models having all-pass factor.

6. Simulation results

In this section, we present three simulation examples to illustrate the proposed method IMK and to compare it with others methods proposed in the literature. To reduce realization dependency of the simulations, we averaged over 100 independent Monte Carlo runs (MCR). The input white noise $\mathbf{x}$ used is exponentially distributed ($\gamma_{2,x} = 1$ and $\gamma_{3,x} = 2$). Tables 1, 2 and 3 display the estimated parameters (mean and one standard deviation (Std)) for each method.

Example 1 : We consider the following true MA(5) model [8] :

$$\mathcal{Z}h_{\theta_0}(z) = 1 + 0.1\,z^{-1} - 1.87\,z^{-2} + 3.02\,z^{-3} - 1.435\,z^{-4} + 0.49\,z^{-5}$$

with zeros located at $0.25 \pm \jmath 0.433$, $0.7 \pm \jmath 0.7$ and -2.

The signal length used is $N = 1024$ samples. The method IMK is compared to the methods presented in [8], [3] and [1]. The methods of [8] and [3] are "linear algebra solutions" and use different slices of second and third order cumulants. Table 1 shows the simulation results. We can see that the IMK method gives estimates having lower bias and variance than the other methods.

Table 1. Parameter estimates : Example 1, 100 MCR, N=1024, mean, Std.

	$\beta[i]$	$\beta[1]$	$\beta[2]$	$\beta[3]$	$\beta[4]$	$\beta[5]$
	True value	0.100	-1.870	3.020	-1.4350	0.490
[8]	Mean	-0.2684	-0.7191	1.1103	-0.3138	0.0246
	Std	0.5212	0.7051	0.8711	0.8518	0.4240
[3]	Mean	-0.2897	-1.0359	1.9145	-0.9173	0.3283
	Std	0.6067	0.9641	1.0694	0.5525	0.3172
[1]	Mean	0.1405	-1.8038	2.8593	-1.2613	0.4531
	Std	0.5171	0.9751	1.6299	1.6843	0.5440
IMK	Mean	0.1404	-1.9367	3.0718	-1.4375	0.4860
	Std	0.2288	0.3593	0.2865	0.1038	0.0650

Example 2 : Let the following non-causal AR(3) model [7] :

$$\mathcal{Z}h_{\theta_0}(z) = \frac{1}{1 - 2.05\,z^{-1} + 1.65\,z^{-2} - 0.8125\,z^{-3}}$$

with poles located at $0.4 \pm \jmath 0.7$ and 1.25.

The signal length used is $N = 256$ samples. The method IMK is compared to the method presented in [1] and [4, Theorem 2], this last using third order cumulants. Table 2 displays the simulation results. We can see that the IMK method has lower variance than the other methods and the method of Boumahdi [1] is better than those of [4], which uses only HOS.

Example 3 : The true non-causal ARMA(2,1) model [5] used is given by :

$$\mathcal{Z}h_{\theta_0}(z) = \frac{1 - 0.5\,z^{-1}}{1 - 0.5\,z^{-1} - 0.9375\,z^{-2}}$$

with two poles located at -0.75 and 1.25, and a zero at 0.5.

The signal length used is $N = 512$ samples. The method IMK is compared to the method MK (version of [1] allowing identification of non-causal ARMA models). The simulation results are presented in Table 3. We can see that the IMK method provides estimates with lower variance and bias than the method MK.

Table 2. Parameter estimates : Example 2, 100 MCR, N=256, mean, Std.

	$\alpha[i]$	$\alpha[1]$	$\alpha[2]$	$\alpha[3]$
	True value	-2.050	1.650	-0.8125
[4]	Mean	-1.5618	0.9837	-0.3444
	Std	0.1353	0.1778	0.1158
[1]	Mean	-2.0782	1.6673	-0.8192
	Std	0.1547	0.1501	0.0959
IMK	Mean	-2.0743	1.6741	-0.8229
	Std	0.1500	0.1467	0.0947

Table 3. Parameter estimates : Example 3, 100 MCR, N=512, mean, Std.

	$\alpha[i]\|\beta[i]$	$\alpha[1]$	$\alpha[2]$	$\beta[1]$
	True value	-0.500	-0.9375	-0.500
MK	Mean	-0.5315	-0.9625	-0.4829
	Std	0.1078	0.1028	0.0837
IMK	Mean	-0.5309	-0.9594	-0.4908
	Std	0.0998	0.1000	0.0844

The simulation results show that the IMK method (MK method improved) ameliorates the accuracy of the parameter estimation facing to the MK method. Because the improvement (see section 5) permits to obtain the good minimum of the criterion (11) and to decrease the rate of false reconstruction of the phase.

7. Conclusion

We proposed in this paper a new method (IMK) to identify non-causal ARMA models, which is an extension of the Boumahdi method [1], using the maximum kurtosis property. This extension is based on a generalization of the minimum variance approach to the non-causal case. We have showed that it is necessary to introduce an ARMA modeling with separation of causal and anti-causal parts. Then, the maximum likelihood estimator (under gaussian hypothesis) has several local minima corresponding to all the models SE to the actual one. Finally, we used the maximum kurtosis property to identify the phase of the true model from one of these local minima (MK method).

Then, we have presented the method IMK which consists in a closed-loop form of the MK method. The recurrence introduced in the method IMK permits to improve the accuracy of the parameter estimation and to obtain a better reconstruction of the phase of the estimated model, in comparison with the MK method. Moreover, for the IMK method, the covariance matrix of parameters estimates is known and can be calculated with the formula found in [6].

The problem is that this method cannot identify ARMA models with all-pass factors. But the simulation results showed the good performances of the method IMK, compared to methods using HOS, for the identification of non-causal MA, AR and ARMA models without all-pass factors. Its interest is that it uses all the SOS information for identification, HOS being used only as criterion of choice among all the SE models to determine the model phase.

References

[1] M. Boumahdi. Blind identification using the kurtosis with applications to field data. *Signal Processing*, 48(3):205–216, 1996.

[2] D. Donoho. *On minimum entropy deconvolution. In Applied Time Series Analysis, II.* D.F. Findley, Academic Press, New York, 1981.

[3] J. Fonollosa and J. Vidal. System identification using a linear combination of cumulant slices. *IEEE Trans. on S.P.*, 41(7):2405–2412, July 1993.

[4] G. Giannakis and A. Swami. Identifiability of general ARMA processes using linear cumulant based estimators. *Automatica*, 28(4):771–779, 1992.

[5] A. Petropulu. Noncausal nonminimum phase ARMA modeling of non-gaussian processes. *IEEE Trans. on S.P.*, 43(8):1946–1954, August 1995.

[6] O. Shalvi and E. Weinstein. Maximum likelihood and lower bounds in system identification with non-gaussian inputs. *IEEE Trans. on Inform. Theory*, 40(2):328–339, March 1994.

[7] J. Tugnait. Fitting non-causal autoregressive signal plus noise models to noisy non-gaussian linear processes. *IEEE A.C.*, AC-32(6):547–551, June 1987.

[8] J. Tugnait. New results on FIR system identification using higher order statistics. *IEEE Trans. on S.P.*, 39(10):2216–2221, October 1991.

IDENTIFICATION OF MULTIVARIABLE STOCHASTIC LINEAR SYSTEMS USING INTEGRATED POLYSPECTRUM GIVEN NOISY INPUT-OUTPUT DATA

Jitendra K. Tugnait

Department of Electrical Engineering
Auburn University, Auburn, Alabama 36849, USA

ABSTRACT

The problem considered is that of identification of unknown parameters of multivariable, linear "errors-in-variables" models, i.e., linear systems where measurements of both input and output of the system are noise contaminated. Attention is focused on frequency-domain approaches where the integrated polyspectrum (bispectrum or trispectrum) of the input and the integrated cross-polyspectrum, respectively, of the given time-domain input-output data are exploited. We first develop (computable) expressions for the covariance matrix of the system transfer function estimate and show that the system transfer function matrix estimate is asymptotically complex Gaussian. Then we propose and analyze a pseudo-maximum likelihood (PML) estimator of system parameters using the developed statistics of the system transfer function estimate. Finally two simulation examples are presented.

1. INTRODUCTION

It is a common practice in system identification literature to assume that measurements of the system output are noisy but the measurements of the input to the system are *perfect*. This assumption is not necessarily true in system and control applications where the input is not under the analyst's (complete) control; rather, it can only be measured. Clearly, it may not always be possible to neglect the noise introduced by the sensor itself, or by the ambient environment. In this paper we propose to investigate multivariable system identification under "symmetric modeling" [1],[2] of stochastic systems by allowing the input measurements also to be noise contaminated, in addition to having noisy output measurements. Such models are called *errors-in-variables* models in the econometrics literature. Past work on dynamic system identification with noisy input has concentrated overwhelmingly on exploitation of second order statistics [1]-[4]. It is known that, in general, there does not exist a unique solution when only second order statistics are used. Use of higher-order statistics can alleviate this problem at the cost of higher-variance estimates [5]-[9]. Our recent work [9] dealing with single-input/single-output models has utilized frequency-domain approaches coupled with a novel entity, integrated polyspectrum, which leads to computationally simpler and statistically more accurate parameter estimates than heretofore possible with higher-order statistics, including the approaches of [5]-[8]. In [10] identifiability and consistency issues for MIMO systems were considered. No simulation examples were provided in [10].

The objective of this paper is threefold:

- To develop (computable) expressions for the covariance matrix of the system transfer function estimate (which is based upon certain integrated polyspectrum estimates). To show that the system transfer function estimate is asymptotically complex Gaussian.
- To propose and analyze a pseudo-maximum likelihood estimator of system parameters using the developed statistics of the system transfer function estimate.
- To provide computer simulation examples.

2. MODEL ASSUMPTIONS

The true system (i.e. the one generating the data) is given by

$$\mathbf{s}(t) = \mathcal{H}(q)\mathbf{u}(t) \qquad (2-1)$$

where $p-$column vector $\mathbf{s}(t)$ is the system output, t is discrete-time, and $m-$column vector $\mathbf{u}(t)$ is the system input. With q^{-1} denoting the backward-shift operator (i.e. $q^{-1}\mathbf{u}(t) = \mathbf{u}(t-1)$), the linear system $\mathcal{H}(q)$ represents an IIR (infinite impulse response) system:

$$\mathcal{H}(q) = \sum_{i=0}^{\infty} H(i)q^{-i}. \qquad (2-2)$$

Noisy measurements of the system input and output are available as

$$\mathbf{x}(t) = \mathbf{u}(t) + \mathbf{v}_i(t), \quad \mathbf{y}(t) = \mathbf{s}(t) + \mathbf{v}_o(t). \qquad (2-3)$$

Given an input-output (IO) record $\{\mathbf{x}(t), \mathbf{y}(t), \quad t = 1, 2, \cdots\}$, but the underlying true system model $\mathcal{H}(q)$ ***unknown***, it is of much interest in control, communications and signal processing applications to fit a rational function model

$$\mathcal{G}(q) := A^{-1}(q)B(q) = [I + \sum_{i=1}^{n_a} A_i q^{-i}]^{-1}[\sum_{i=1}^{n_b} B_i q^{-i}] \qquad (2-4)$$

to given input-output record. Define the unknown parameter vector

$$\theta = \{\text{vec}\{A_i\},\ 1 \le i \le n_a, \quad \text{vec}\{B_j\},\ 1 \le j \le n_b\} \qquad (2-5)$$

where vec denotes the column stacking operator [18]. We wish to estimate θ given certain statistics of the input-output data record $\{\mathbf{x}(t), \mathbf{y}(t),\ t = 1, 2, \cdots\}$. When it exists true value of θ will be denoted by θ_0.

This work was supported by the National Science Foundation under Grant ECS-9504878.

0-8186-8005-9/97 $10.00 © 1997 IEEE

The following conditions are assumed to be true.

(AS1) The true system transfer function $\mathcal{H}(q)$ is causal and stable. Therefore, $\sum_{i=0}^{\infty} \|H(i)\|^2 < \infty$.

(AS2) All the processes involved (i.e., $\mathbf{x}(t)$, $\mathbf{y}(t)$, $\mathbf{v}_i(t)$, and $\mathbf{v}_o(t)$) are zero-mean and jointly stationary. Furthermore, the noise sequences $\{\mathbf{v}_i(t)\}$ and $\{\mathbf{v}_o(t)\}$ are independent of $\{\mathbf{u}(t)\}$, hence of $\{\mathbf{s}(t)\}$. The integrated polyspectrum (defined in Sec. 3) $S_{r_{lu}u}(\omega)$ (l=2 or 3) of $\{\mathbf{u}(t)\}$ is invertible for all $\omega \in [0, \pi]$ if the proposed approaches utilize the entire frequency range $[0, \pi]$. If a finite number of frequencies are used then $S_{r_{lu}u}(\omega)$ need be invertible only for this frequency set. Finally, $\mathbf{v}_o(t) \neq \mathcal{H}(q)\mathbf{v}_i(t)$ and $S_{vv}(\omega)$ is invertible whenever $S_{r_{lu}u}(\omega)$ is, where $\mathbf{v}(t) := \mathbf{v}_o(t) - \mathcal{H}(q)\mathbf{v}_i(t)$.

(AS3) The cumulant/ cross-cumulant sequences of the various processes involved satisfy the following summability conditions:

$$\sum_{\tau_1,\cdots,\tau_{k-1}=-\infty}^{\infty} [1 + |\tau_j|]|C_{z_1 z_2 \cdots z_k}(\tau_1, \cdots, \tau_{k-1})| < \infty,$$

for each $j = 1, 2, \cdots, k-1$ and each $k = 2, 3, \cdots$ where $z_i(t)$ denotes a component of any of the involved processes such as $\mathbf{u}(t)$, $\mathbf{s}(t)$, $\mathbf{v}_i(t)$, etc. and $C_{z_1 z_2 \cdots z_k}(\tau_1, \cdots, \tau_{k-1})$ denotes the k-th joint cumulant of the random variables $\{z_1(t+\tau_1), \cdots, z_{k-1}(t+\tau_{k-1}), z_k(t)\}$.

(AS4) The noise processes are jointly Gaussian if we exploit the integrated trispectrum (defined in Sec. 3) of the data. The noise processes are assumed to have vanishing bispectrum (defined in Sec. 3) if we exploit the integrated bispectrum of the data.

3. TRANSFER FUNCTION ESTIMATOR

Define $m-$column vectors $\mathbf{r}_{3u}(t)$ and $\mathbf{r}_{2u}(t)$ whose $i-$th components $r_{3ui}(t)$ and $r_{2ui}(t)$, respectively, are given by

$$r_{3ui}(t) = u_i^3(t) - 3u_i(t)E\{u_i^2(t)\} - E\{u_i^3(t)\}, \quad (3-1)$$

$$r_{2ui}(t) = u_i^2(t) - E\{u_i^2(t)\}, \quad (3-2)$$

where $u_i(t)$ is the $i-$th component of $\mathbf{u}(t)$. We will use the notation $\mathbf{r}_{3x}(t)$ and $\mathbf{r}_{2x}(t)$ to denote the above vectors when $\mathbf{u}(t)$ is replaced with $\mathbf{x}(t)$. Mimicking the univariate formulation of [9], we now define a "component-by-component" integrated bispectrum of $\mathbf{u}(t)$ as

$$S_{r_{2u}u}(\omega) = \sum_{k=-\infty}^{\infty} E\{\mathbf{r}_{2u}(t)\mathbf{u}^T(t+k)\}e^{-j\omega k}. \quad (3-3)$$

Similarly we have integrated trispectrum

$$S_{r_{3u}u}(\omega) = \sum_{k=-\infty}^{\infty} E\{\mathbf{r}_{3u}(t)\mathbf{u}^T(t+k)\}e^{-j\omega k}. \quad (3-4)$$

Consider the cross-spectrum ($l = 2, 3$)

$$S_{r_{lx}y}(\omega) = \sum_{k=-\infty}^{\infty} E\{\mathbf{r}_{lx}(t+k)\mathbf{y}^T(t)\}e^{-j\omega k}. \quad (3-5)$$

Therefore, if $S^{-1}_{r_{lu}u}(\omega)$ exists, then we have the transfer function matrix expression

$$\mathcal{H}(e^{j\omega}) = [S^{-1}_{r_{lx}x}(\omega)S_{r_{lx}y}(\omega)]^{\mathcal{H}} \quad (l = 2, 3). \quad (3-6)$$

Given a record of length T, let $\mathbf{X}(\omega)$ denote the DFT of $\{\mathbf{x}(t),\ 1 \leq t \leq T\}$ given by

$$\mathbf{X}(\omega_k) = \sum_{t=0}^{T-1} \mathbf{x}(t+1)\exp(-j\omega_k t) \quad (3-7)$$

where

$$\omega_k = \frac{2\pi}{T}k, \quad k = 0, 1, \cdots, T-1. \quad (3-8)$$

Similarly define $\mathbf{Y}(\omega_k)$. Also let $\mathbf{R}_{lx}(\omega)$ denote the DFT of $\{\mathbf{r}_{lx}(t)\}$ obtained by using the relations (recall (3-1) and (3-2))

$$r_{3xj}(t) = x_j^3(t) - 3x_j(t)\widehat{\mu}_{2xj}, \quad (3-9)$$

$$\widehat{\mu}_{2xj} = \frac{1}{N}\sum_{t=1}^{N} x_j^2(t), \quad (3-10)$$

$$r_{2xj}(t) = x_j^2(t), \quad (3-11)$$

($j = 1, 2, \cdots, m$). Given the above DFT's, following [11, Sec. 7.4] we define the cross-spectrum estimators as

$$\widehat{S}_{r_{lx}x}(k) = \frac{1}{T(2m_T+1)} \sum_{i=-m_T}^{m_T} \mathbf{R}_{lx}(\omega_{k-i})[\mathbf{X}(\omega_{k-i})]^{\mathcal{H}}, \quad (3-12)$$

$$\widehat{S}_{r_{lx}y}(k) = \frac{1}{T(2m_T+1)} \sum_{i=-m_T}^{m_T} \mathbf{R}_{lx}(\omega_{k-i})[\mathbf{Y}(\omega_{k-i})]^{\mathcal{H}}, \quad (3-13)$$

In light of (3-12) define a coarser frequency grid:

$$\omega_l = \frac{2\pi l}{L_T} = \frac{2\pi l(2m_T+1)}{T} \quad (3-14)$$

with $l = 0, 1, \cdots, L_T - 1$ where $L_T = \lfloor \frac{T}{2m_T+1} \rfloor$.

Using the estimated spectra we have an estimator of the system transfer function at frequency ω_k (as in [11, Chapter 8])

$$\widehat{\mathcal{H}}(e^{j\omega_k}) = \widehat{S}_{yu}(k)\widehat{S}^{-1}_{uu}(k) \quad (3-15)$$

provided that $\widehat{S}^{-1}_{uu}(k)$ exists. Choose m_T to be such that as $T \to \infty$, we have $m_T T^{-1} \to 0$ and $m_T \to \infty$. If $S^{-1}_{uu}(\omega_k)$ exists, then it follows from [11, Thm. 8.11.1] that

$$\lim_{T\to\infty} \widehat{\mathcal{H}}(e^{j2\pi f}) = \lim_{T\to\infty} \widehat{S}_{yu}(k(T))\widehat{S}^{-1}_{uu}(k(T))$$

$$= \mathcal{H}(e^{j2\pi f}) \quad \text{w.p.1} \quad (3-16)$$

where $\lim_{T\to\infty} k(T)/T = f$. Convergence in (3-16) is uniform in f. In addition, $\widehat{\mathcal{H}}(e^{j\omega_k})$ for $k = l(2m_T+1)$, $l = 0, 1, \cdots, (L_T/2) - 1$, are (asymptotically) independent.

Let $k_l(T)$ with $T = 1, 2, \cdots$ be a sequence of integers such that $\lim_{T\to\infty} k_l(T)/T = f_l$, a fixed frequency (in Hz). We may take these integers to belong to the coarser grid $\{k \mid k = l(2m_T+1),\ \ l = 0, 1, \cdots, (L_T/2) - 1\}$. Under

the cumulant summability conditions stated in assumption (AS3), it follows from [11, Thm. 7.4.1] that

$$\lim_{T\to\infty} E\{\widehat{S}_{r_{1x}x}(k_l(T))\} = S_{r_{1x}x}(2\pi f_l), \qquad (3-17)$$

$$\lim_{T\to\infty} E\{\widehat{S}_{r_{1x}y}(k_l(T))\} = S_{r_{1x}y}(2\pi f_l). \qquad (3-18)$$

Moreover, it follows from [11, Thm. 7.4.3] that for $0 \le f_l, f_m \le 0.5$

$$\lim_{T\to\infty} \Delta_T \operatorname{cov}\{[\widehat{S}_{r_{1x}x}(k_l(T))]_{i_1,j_1}, [\widehat{S}_{r_{1x}y}(k_m(T))]_{i_2,j_2}\}$$
$$= \delta(l-m)[S_{r_{1x}r_{1x}}(2\pi f_l)]_{i_1,i_2}[S_{yx}(2\pi f_l)]_{j_2,j_1}, \qquad (3-19)$$

$$\lim_{T\to\infty} \Delta_T \operatorname{cov}\{[\widehat{S}_{r_{1x}x}(k_l(T))]_{i_1,j_1}, [\widehat{S}^*_{r_{1x}y}(k_m(T))]_{i_2,j_2}\}$$
$$= \delta(l-m)[S_{r_{1x}y}(2\pi f_l)]_{i_1,j_2}[S_{r_{1x}x}(2\pi f_l)]_{i_2,j_1}, \qquad (3-20)$$

where

$$\Delta_T = 2m_T + 1 \qquad (3-21)$$

$\operatorname{cov}\{X,Y\} = E\{XY^{\mathcal{H}}\} - E\{X\}E\{Y^{\mathcal{H}}\}$, $[A]_{ij}$ denotes the ij-th element of matrix A, and other covariances may be deduced in a similar manner. Convergence in (3-17)-(3-20) is uniform in f.

Consider a fixed set of M frequencies λ_l for $l = 1, 2, \cdots, M$ such that $0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_M < \pi$, where $\lambda_l = 2\pi f_l$. Using (3-17)-(3-21) and a perturbation expansion of $\widehat{\mathcal{H}}^{\mathcal{H}}(e^{j\lambda_k})$ yields (see also proofs of Thm. 8.7.1 and Lemma P8.1 in [11, pp. 449-450] pertaining to power spectrum and cross-power spectrum)

$$\widehat{\mathcal{H}}^{\mathcal{H}}(e^{j\lambda_l}) = (E\{\widehat{S}_{r_{1x}x}(k_l(T))\})^{-1} E\{\widehat{S}_{r_{1x}y}(k_l(T))\}$$
$$+(E\{\widehat{S}_{r_{1x}x}(k_l(T))\})^{-1}[\widehat{S}_{r_{1x}y}(k_l(T)) - E\{\widehat{S}_{r_{1x}y}(k_l(T))\}]$$
$$-(E\{\widehat{S}_{r_{1x}x}(k_l(T))\})^{-1}[\widehat{S}_{r_{1x}x}(k_l(T)) - E\{\widehat{S}_{r_{1x}x}(k_l(T))\}]$$
$$\times(E\{\widehat{S}_{r_{1x}x}(k_l(T))\})^{-1} E\{\widehat{S}_{r_{1x}y}(k_l(T))\}$$
$$+ o_p(\Delta_T^{-0.5}), \qquad (3-22)$$

where $o_p(a_n)$ implies a random sequence $\{z_n\}$ such that $\lim_{n\to\infty} z_n a_n^{-1} = 0$ in probability (i.p.). Using the asymptotic distribution of the estimators (3-12) and (3-13) on the coarse grid (3-14), [11, Thm. P5.2] and (3-6), it follows (as in [11, Thm. 8.8.1]) that $\widehat{\mathcal{H}}(e^{j\lambda_l})$ for $l = 1, 2, \cdots, M$ are (asymptotically) jointly complex (circularly symmetric) Gaussian such that

$$\lim_{T\to\infty} \Delta_T \operatorname{cov}\left(\operatorname{vec}(\widehat{\mathcal{H}}(e^{j\lambda_k})), \operatorname{vec}(\widehat{\mathcal{H}}(e^{j\lambda_l}))\right)$$
$$= \Sigma(\lambda_k)\,\delta(k-l) \qquad (3-23)$$

where

$$\Sigma(\lambda_k) := S_{uu}^{-1}(\lambda_k) \otimes \left[S_{yy}(\lambda_k) - S_{yu}(\lambda_k) S_{uu}^{-1}(\lambda_k) S_{uy}(\lambda_k)\right] \qquad (3-24)$$

and $\otimes$ denotes the Kronecker product. Thus, $\widehat{\mathcal{H}}(e^{j\omega_k})$ on the coarse grid (3-14) is asymptotically a complex Gaussian (in the sense of [11, Sec. 4.2]) random variable, independent at distinct frequencies on the coarse grid over $(0, \pi)$, with the covariance structure (3-17).

Remark 1. In the rest of the paper we will use ω_k to denote a frequency on the coarse grid (3-14) with $k = 0, 1, \cdots, L_T - 1$ but we will use λ_k to denote a fixed frequency independent of the record length T. •

4. A PSEUDO-MAXIMUM LIKELIHOOD SOLUTION

As in [10] consider a quadratic transfer function matching approach. Let $\mathcal{G}(e^{j\omega};\theta)$ denote the transfer function of (1-3) with the system parameters θ as defined in (2-1). We choose θ to minimize the cost

$$J'_{2T}(\theta; M, \Omega_L, \Omega_U) = \sum_{l=1}^{M} \left\| [\widehat{\mathbf{h}}(e^{j\lambda_l}) - \mathbf{g}(e^{j\lambda_l};\theta)] \right\|^2_{\mathbf{W}(\lambda_l)} \qquad (4-1)$$

where $0 \le \Omega_L \le \lambda_1 < \lambda_2 < \cdots < \lambda_M \le \Omega_U \le \pi$,

$$\widehat{\mathbf{h}}(e^{j\lambda_l}) := \operatorname{vec}\{\widehat{\mathcal{H}}(e^{j\lambda_l})\}, \quad \mathbf{g}(e^{j\lambda_l};\theta) := \operatorname{vec}\{\mathcal{G}(e^{j\lambda_l};\theta)\}, \qquad (4-2)$$

$\|\mathbf{e}\|^2_{\mathbf{W}(\omega_l)} = \mathbf{e}^{\mathcal{H}}\mathbf{W}(\omega_l)\mathbf{e}$ and $\mathbf{W}(\omega_l)$ is a positive-definite, Hermitian weighting matrix. This is a nonlinear iterative optimization problem which is initialized by the consistent, closed-form equation-error solution of Sec. 3.2 of [10].

In light of the results of Sec. 3, let us pick

$$\mathbf{W}(\lambda_l) = \Delta_T^{-1}\Sigma(\lambda_l) = \operatorname{cov}\{\widehat{\mathbf{h}}(e^{j\lambda_l}), \widehat{\mathbf{h}}(e^{j\lambda_l})\}. \qquad (4-3)$$

Then (4-1) leads to a (pseudo) maximum likelihood (ML) parameter estimator (asymptotically). Since, in practice, $\Sigma(\lambda_l)$ is unknown, we replace it with its consistent estimator $\widehat{\Sigma}(\lambda_l)$ which is obtained from (3-18) by replacing all the quantities therein with their consistent estimators such as (3-12) and (3-13). This leads to the cost

$$J_{2T}(\theta; M, \Omega_L, \Omega_U)$$
$$= \sum_{l=1}^{M} \left[\widehat{\mathbf{h}}(e^{j\lambda_l}) - \mathbf{g}(e^{j\lambda_l};\theta)\right]^{\mathcal{H}} \widehat{\Sigma}^{-1}(\lambda_l) \left[\widehat{\mathbf{h}}(e^{j\lambda_l}) - \mathbf{g}(e^{j\lambda_l};\theta)\right]. \qquad (4-4)$$

It follows [11] that under (AS1)-(AS4), we have $\lim_{T\to\infty} \widehat{\Sigma}(\lambda_l) = \Sigma(\lambda_l)$ w.p.1 uniformly in λ_l.

Proof of the following result is omitted.

Lemma 1. (A) Under (AS1)-(AS4) such that $0 \le \Omega_L \le \lambda_1 < \lambda_2 < \cdots < \lambda_M \le \Omega_U \le \pi$, $\lim_{T\to\infty} J_{2T}(\theta; M, \Omega_L, \Omega_U) \overset{w.p.1}{=} J_{2\infty}(\theta; M, \Omega_L, \Omega_U)$ uniformly in θ for $\theta \in \Theta_C$, any compact set, where

$$J_{2\infty}(\theta; M, \Omega_L, \Omega_U)$$
$$= \sum_{l=1}^{M} \left[\mathbf{h}(e^{j\lambda_l}) - \mathbf{g}(e^{j\lambda_l};\theta)\right]^{\mathcal{H}} \Sigma^{-1}(\lambda_l) \left[\mathbf{h}(e^{j\lambda_l}) - \mathbf{g}(e^{j\lambda_l};\theta)\right] \qquad (4-5)$$

and $\mathbf{h}(e^{j\lambda_l}) := \operatorname{vec}\{\mathcal{H}(e^{j\lambda_l})\}$.

(B) In addition, let the set of frequencies above become dense in the interval $[\Omega_L, \Omega_U]$ as $M \to \infty$ where λ_i's are spaced uniformly in this interval. Then $\lim_{M,T\to\infty} M^{-1} J_{2T}(\theta; M, \Omega_L, \Omega_U) \overset{w.p.1}{=} J_{2\infty}(\theta; \infty, \Omega_L, \Omega_U)$ uniformly in θ for $\theta \in \Theta_C$ where

$$J_{2\infty}(\theta; \infty, \Omega_L, \Omega_U) = \frac{1}{\pi}\int_{\Omega_L}^{\Omega_U} \left[\mathbf{h}(e^{j\omega}) - \mathbf{g}(e^{j\omega};\theta)\right]^{\mathcal{H}}$$
$$\times \Sigma^{-1}(\omega) \left[\mathbf{h}(e^{j\omega}) - \mathbf{g}(e^{j\omega};\theta)\right] d\omega \bullet \qquad (4-6)$$

5. CONVERGENCE ANALYSIS

Define

$$\widehat{\theta}^{(2)}_{TM} = \arg\{\min_\theta J_{2T}(\theta; M, \Omega_L, \Omega_U)\}, \quad (5-1)$$

$$\overline{\theta}^{(2)}_{M} = \arg\{\min_\theta J_{2\infty}(\theta; M, \Omega_L, \Omega_U)\} \quad (5-2)$$

and

$$\overline{\theta}^{(2)} = \arg\{\min_\theta J_{2\infty}(\theta; \infty, \Omega_L, \Omega_U)\}. \quad (5-3)$$

Using Lemma 1 and some standard arguments we can establish Theorem 1.

Theorem 1. Under the hypotheses of Lemma 1, it follows that

$$\lim_{T\to\infty} \widehat{\theta}^{(2)}_{TM} \overset{\text{w.p.1}}{\in} \mathcal{D}^{(2)}_M(\Omega_L, \Omega_U) := \left\{\theta \,\middle|\, J_{2\infty}(\theta; M, \Omega_L, \Omega_U) = J_{2\infty}(\overline{\theta}^{(2)}_M; M, \Omega_L, \Omega_U)\right\} \quad (5-4)$$

and

$$\lim_{M\to\infty}\lim_{T\to\infty} \widehat{\theta}^{(2)}_{TM} \overset{\text{w.p.1}}{\in} \mathcal{D}^{(2)}_\infty(\Omega_L, \Omega_U) := \left\{\theta \,\middle|\, J_{2\infty}(\theta; \infty, \Omega_L, \Omega_U) = J_{2\infty}(\overline{\theta}^{(2)}; \infty, \Omega_L, \Omega_U)\right\} \bullet \quad (5-5)$$

Proof: Mimic the proof of Theorems 1 and 3 in [5] using Lemma 1. Note that the convergence to the set $\mathcal{D}^{(2)}$ is to be interpreted in the sense of Ljung [12, p. 215]. □

Next we have the consistency result. It follows from [10, Thm. 1], Theorem 1, and the conditions required for uniqueness of the full polynomial form of the matrix fraction description of MIMO transfer functions [3, Sec. 6.3]. First we need a definition.

Def. **Sufficient Order Model Set:** The true model $\mathcal{H}(q)$ is of the type (1-3) such that the true model orders n_{a0} and n_{b0} satisfy $\min(n_a - n_{a0}, n_b - n_{b0}) \geq 0$. •

Theorem 2. Under (AS1)-(AS4), $\Omega_L > 0$, $\Omega_U < \pi$ and sufficient order modeling such that $(2p-1)n_a + n_b \leq 2M$, it follows that

$$\lim_{T\to\infty} \widehat{\theta}^{(2)}_{TM} \overset{\text{w.p.1}}{\in} \mathcal{D}^{(so)}$$

where

$$\mathcal{D}^{(so)} := \left\{\theta \,\middle|\, A^{-1}(q;\theta)B(q;\theta) = \mathcal{H}(q)\right\}.$$

If $\min(n_a - n_{a0}, n_b - n_{b0}) = 0$, $A(q;\theta_0)$ and $B(q;\theta_0)$ are coprime and $\text{rank}[A_{n_{a0}} \vdots B_{n_{b0}}] = p$, then $\mathcal{D}^{(so)}$ is a singleton equaling $\{\theta_0\}$. •

6. SIMULATION EXAMPLES

6.1. Example 1.

Referring to (1-1)-(1-3), the 2 × 2 true system model is taken to be

$$\mathcal{H}(q) = A^{-1}(q)B(q) \quad (6-1)$$

with ($I_{2\times 2}$ denotes the 2 × 2 identity matrix)

$$A(q) = I_{2\times 2} + \begin{bmatrix} -1.5 & 0 \\ 0 & -1.8 \end{bmatrix} q^{-1} + \begin{bmatrix} 0.7 & 0 \\ 0 & 0.81 \end{bmatrix} q^{-2} \quad (6-2)$$

and

$$B(q) = \begin{bmatrix} 1.0 & 0 \\ 1.0 & 1.0 \end{bmatrix} q^{-1} + \begin{bmatrix} 0.5 & 0.8 \\ -0.9 & 0. \end{bmatrix} q^{-2}. \quad (6-3)$$

The (noise-free) inputs are generated as

$$u_i(t) = -0.8u_i(t-1) + w_i(t) \quad (i = 1, 2) \quad (6-4)$$

where $\{w_i(t)\}$ is i.i.d., binary (±1 with prob. 0.5 each) with $\{w_1(t)\}$ independent of $\{w_2(t)\}$. The output measurement noise is colored and its components are given by

$$v_{oi}(t) = 1.3725v_{oi}(t-1) - 0.9608v_{oi}(t-2) + \epsilon_{oi}(t) \quad (i = 1, 2) \quad (6-5)$$

where $\{\epsilon_{oi}(t)\}$ is i.i.d. zero-mean Gaussian with $\{\epsilon_{o1}(t)\}$ independent of $\{\epsilon_{o2}(t)\}$. The input measurement noise is colored and its components are given by

$$v_{ij}(t) = 1.3725v_{ij}(t-1) - 0.9608v_{ij}(t-2) + \epsilon_{ij}(t) \quad (j = 1, 2) \quad (6-6)$$

where $\{\epsilon_{ij}(t)\}$ is i.i.d. zero-mean Gaussian with $\{\epsilon_{i1}(t)\}$ independent of $\{\epsilon_{i2}(t)\}$. The input noise $\mathbf{v}_i(t)$ is independent of $\mathbf{v}_o(t)$. The two noise sequences are scaled to achieve the desired SNR at the output (= $E\{\|\mathbf{s}(t)\|^2\}/E\{\|\mathbf{v}_o(t)\|^2\}$ where $\mathbf{s}(t) := \mathcal{H}(q)\mathbf{u}(t)$) and at the input (= $E\{\|\mathbf{u}(t)\|^2\}/E\{\|\mathbf{v}_i(t)\|^2\}$).

Given the AR coefficient estimates $\widehat{A}^{(l)}_i$ in the l–th Monte Carlo run, define normalized mean-square error (NMSE) in estimating the AR coefficients averaged over M_c runs as

$$\text{NMSE}_A = \frac{M_c^{-1}\sum_{l=1}^{M_c}\sum_{i=1}^{n_a}\|A_i - \widehat{A}^{(l)}_i\|^2}{\sum_{i=1}^{n_a}\|A_i\|^2}.$$

Similarly NMSE in estimating the MA coefficients is defined as

$$\text{NMSE}_B = \frac{M_c^{-1}\sum_{l=1}^{M_c}\sum_{i=1}^{n_b}\|B_i - \widehat{B}^{(l)}_i\|^2}{\sum_{i=1}^{n_b}\|B_i\|^2}.$$

The simulation results based on 100 Monte Carlo runs for two different record lengths and SNR = 10 dB (at both input as well as output) are shown in Table I for the approaches proposed in [10, Sec 3.2] and Sec. 4 under the sufficient-order case with $n_a = n_{a0} = 2$ and $n_b = n_{b0} = 2$ (recall the definitions of Sec. 4). In applying the proposed approaches, we selected $2m_T + 1 = 11$ and 45 for record lengths of $T = 1024$ and 4096, respectively. The number of frequency points M was taken to be all the points on the coarse grid (3-14) that lie in $(0, \pi)$. For comparison we also show the results obtained using the classical least-squares algorithm ([3, Sec. 7.1]) and the output error method [3],[12]. For the output error method (which requires nonlinear optimization) the initial guess was selected as

$$\widehat{A}(q) = I_{2\times 2} + \begin{bmatrix} -0.6 & 0 \\ 0 & -1.4 \end{bmatrix} q^{-1} + \begin{bmatrix} 0.1 & 0 \\ 0 & 0.5 \end{bmatrix} q^{-2} \quad (6-7)$$

and

$$\widehat{B}(q) = I_{2\times 2} + \begin{bmatrix} 0. & 0. \\ 0. & 0. \end{bmatrix} q^{-2}. \quad (6-8)$$

Comparing (6-7)-(6-8) with (6-2)-(6-3) notice that the initial guess is not too far from the true values.

It is seen from Table I that the proposed approaches perform quite well. Use of the statistics of the transfer function estimates does improve the accuracy of the parameter estimates. Both of the proposed approaches perform better than the output error method.

6.2. Example 2.

This example is the same as Example 1 except for the input and output noise sequences which are now dependent. The output measurement noise $\mathbf{v}_o(t)$ is given by

$$\mathbf{v}_o(t) = -0.9\mathbf{v}_o(t-1) + \begin{bmatrix} \epsilon_{o1}(t) \\ \epsilon_{o2}(t) \end{bmatrix} \qquad (6-9)$$

where $\{\epsilon_{oi}(t)\}$ is i.i.d. zero-mean Gaussian with $\{\epsilon_{o1}(t)\}$ independent of $\{\epsilon_{o2}(t)\}$. The input measurement noise $\mathbf{v}_i(t)$ is given by

$$\mathbf{v}_i(t) = 0.8\mathbf{v}_i(t-1) + \mathbf{v}_o(t). \qquad (6-10)$$

The simulation results based on 100 Monte Carlo runs for a record length of 4096 and SNR = 10 dB (at both input as well as output) is shown in Table II under the sufficient-order case. In applying the proposed approaches, we selected $2m_T+1 = 45$. The number of frequency points M was taken to be all the points on the coarse grid (3-14) that lie in $(0, \pi/2)$. It is seen from Table II that the proposed approaches are vastly superior to the output error method.

7. CONCLUSIONS

The parametric frequency domain approaches presented in [9] for estimation of the parameters of single-input single-output, linear errors-in-variables models given time-domain noisy input-output data, were extended to multivariable models in [10]. Attention was focused on frequency-domain approaches where the integrated polyspectrum (bispectrum or trispectrum) of the input and the integrated cross-polyspectrum, respectively, of the given time-domain input-output data are exploited. In [10] identifiability and consistency issues for MIMO systems were considered. No simulation examples were provided in [10]. In this paper we first developed expressions for the covariance matrix of the system transfer function estimate (which is based upon certain integrated polyspectrum estimates) and showed that the system transfer function matrix estimate is asymptotically complex Gaussian. Then we proposed and analyzed a pseudo-maximum likelihood (PML) estimator of system parameters using the developed statistics of the system transfer function estimate. Finally two simulation examples were provided to illustrate two approaches: the proposed PML and the equation-error approach of [10].

8. REFERENCES

[1] M. Deistler, "Symmetric modeling in system identification," in H. Nijmeijer and J.M. Schumacher (Eds.), *Three Decades of Mathematical System Theory*, Lecture Notes in Control & Information Sciences, Springer, 1989.

[2] R.E. Kalman, "Identifiability and modeling in econometrics," in Development in Statistics, IV, P.R. Krishnaiah (Ed.), Academic Press, 1983.

[3] T. Söderström and P. Stoica, *System Identification.* Prentice Hall Intern.: London, 1989.

[4] P. Stoica and A. Nehorai, "On the uniqueness of prediction error models for systems with noisy input-output data," *Automatica*, vol. 23, pp. 541-543, 1987.

[5] J.K. Tugnait, "Stochastic system identification with noisy input using cumulant statistics," *IEEE Transactions on Automatic Control*, vol. AC-37, pp. 476-485, April 1992.

[6] Y. Inouye and H. Tsuchiya, "Identification of linear systems using input-output cumulants," *Intern. J. Control*, vol. 53, pp. 1431-1448, 1991.

[7] A. Delopoulos and G.B. Giannakis, "Consistent identification of stochastic linear systems with noisy input-output data," *Automatica*, vol. 30 , Aug. 1994.

[8] Y. Inouye and Y. Suga, "Identification of linear systems with noisy input using input-output cumulants," *Intern. J. Control*, vol. 59, pp. 1231-1253, May 1994.

[9] J.K. Tugnait and Y. Ye, "Stochastic system identification with noisy input-output measurements using polyspectra," *IEEE Trans. Automatic Control*, vol. AC-40, pp. 670-683, April 1995.

[10] J.K. Tugnait, "Identifiability of multivariable errors-in-variables models using integrated polyspectrum," in *Proc. IEEE Signal Proc./ATHOS Workshop on Higher-Order Statistics*, pp. 39-43, June 12-14, 1995, Begur, Spain.

[11] D.R. Brillinger, *Time Series Data Analysis and Theory.* New York: Holt, Rinehart and Winston, 1975.

[12] L. Ljung, *System Identification: Theory for the User.* Prentice-Hall: Englewood Cliffs, N.J., 1987.

Table I: Example 1: Normalized mean-square errors (NMSE) (in dB) in estimating the AR and MA coefficients: SNR = 10 dB, averages over 100 Monte Carlo runs.
OEM – output error method (time-domain),
EE – Eqn.-error [10, Sec. 3.2]

Approach	NMSE_A (dB)		NMSE_B (dB)	
	T=1024	T=4096	T=1024	T=4096
EE	−13.45	−21.39	−10.66	−18.35
Sec. 4	−19.48	−36.52	−11.78	−25.35
OEM	−17.54	−18.02	−10.72	−12.73

Table II: Example 2: Normalized mean-square errors (NMSE) (in dB) in estimating the AR and MA coefficients: SNR = 10 dB, averages over 100 Monte Carlo runs.
OEM – output error method (time-domain),
EE – Eqn.-error [10, Sec. 3.2]

Approach	NMSE_A (dB)	NMSE_B (dB)
	T=4096	T=4096
EE	−29.78	−16.44
Sec. 4	−39.73	−22.81
OEM	−3.91	4.70

ON THE OPTIMALITY OF BUSSGANG AND SUPER EXPONENTIAL BLIND DECONVOLUTION METHODS

Gaetano Scarano, Giovanni Jacovitti

Dip. INFOCOM, Università di Roma "La Sapienza" via Eudossiana 18, I-00184 Roma, Italy

E-mail: gaetano@infocom.ing.uniroma1.it

Abstract

In this contribution, a generalization of Super Exponential blind deconvolution method is discussed. The generalization consists in the definition of a "spikyness" criterion involving nonlinearities rather than only powers. This allows to rephrase Bussgang deconvolution in the framework of Super Exponential deconvolution using a spikyness criterion which takes into account to the pdf of the input series to be recovered. Improved performance are expected when generalized Super Exponential deconvolution is tuned to suitable optimality criteria.

I Introduction

Among blind deconvolution techniques, *Bussgang* deconvolution [1, 2] and *Super Exponential* methods [3] have received great interest due to their accuracy and their simplicity of formulation and implementation.

While a certain degree of sub–optimality can be advoked in Bussgang deconvolution, Super Exponential methods do not rely on any statistical optimality criterion; these latter methods are rather formulated considering an iterative nonlinear algorithm able to drive the overall channel+equalizer impulse response toward a (possibly scaled and time–shifted) unit sample.

Interestingly enough, both the methods, by exploiting the non–Gaussianity of the unknown i.i.d. signal to be recovered, result in iterative algorithms which require the solution of a linear set of equations involving higher–order moments and/or cumulants of the signals at hand.

In this contribution, we establish a link between the two blind deconvolution methods by showing that Bussgang deconvolution can be obtained as a generalization of Super Exponential methods explicitely taking into account the probabilistic description of the signal to be recovered.

Recalling the blind deconvolution method described in [4, 5, 6] where it has been also showed that Bussgang deconvolution is an approximate implementation of optimal (in a MMSE sense) blind deconvolution, we also conclude that Super Exponential methods result in optimal MMSE solutions when the generalized form is considered. In this case, all the higher–order cumulants are implicitely used through expectation of a suitable nonlinearity depending on the pdf of the unknown signal.

Finally, we note that these estimation methods, based on non–Gaussianity measured through cumulants or nonlinear moments, are implemented in simple iterative schemes constituting the nonlinear minimization of suitable cost functions. The simplicity of the iterative scheme allows to control the solution at the generic iteration step, giving a concrete possibility to avoid solution corresponding to local minima/maxima (due to numerical bad–conditioning induced by sample statistics) which is a difficult task to perform when the nonlinear optimization is implemented using (closed) general purposes optimization software packages. For the same reason, a statistical analysis can be reasonably performed [3, 10] despite the analytical complications arising when (sample) higher–order cumulants are considered.

II Generalized Super Exponential Deconvolution

The key point of this contribution is constituted by the generalization of Super Exponential methods for blind deconvolution in order to obtain the solving equations in a form comparable to those obtained in Bussgang deconvolution. To this end, let us recall the basics of Super Exponential methods. With reference to Fig.1, let us denote by

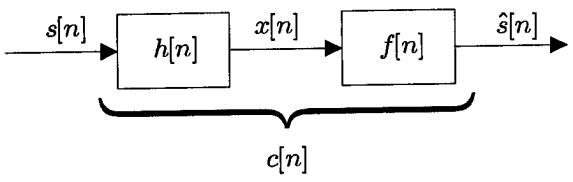


Figure 1: *Blind deconvolution.*

$f[n]$ the impulse response of the equalizer of the linear and shift–invariant (LSI) transformation $h[n]$ which has linearly distorted the i.i.d. series $s[n]$; this latter has to be recovered from the noise–free observed series $x[n]$. Moreover, let us denote by $c[n] = h[n] * f[n]$ the impulse response of

0-8186-8005-9/97 $10.00 © 1997 IEEE

the overall LSI transformation acting on $s[n]$.

Any deconvolution method attempts to synthetize $f[n]$ in order to drive $c[n]$ as close as possible to $\alpha \cdot \delta[n-d]$, *i.e.* to realize an approximate ideal channel apart a gain α and a delay d. Super Exponential methods iteratively determine $g_i[n]$ at iteration i in such a way that the overall response $c_i[n] = h[n] * f_i[n]$ is more "spiky" than the previous iteration. The spikyness is obtained through elevation to the integer power p and normalization[1]

$$c_i[n] = (c_{i-1}[n])^p \cdot \left(\mathcal{E}\left\{(c_{i-1}[n])^p\right\}\right)^{-\frac{1}{2}} \qquad (1)$$

where $\mathcal{E}\{\cdot\}$ extracts the energy of the sequence in its argument. The normalization in (1) induces the spiking behaviour of the overall LSI $c_i[n]$ transformation which naturally tends to $\delta[n-d]$ as iterations proceed.[2] The resulting LMS equalizer $f_i[n]$ at iteration i is obtained from the following set of linear equation (see [3])

$$\mathbf{R}_x \cdot \mathbf{f}_i' = \frac{\sigma_s^2}{\kappa_s^{p+1}} \cdot \boldsymbol{\kappa}_{s_{i-1}x}^{(p,1)} \qquad (2)$$

$$\mathbf{f}_i = \mathbf{f}_i' \cdot \left(\mathbf{f}_i'^T \cdot \mathbf{R}_x \cdot \mathbf{f}_i'\right)^{-\frac{1}{2}} \qquad (3)$$

where σ_s^2 and κ_s^{p+1} are the second–order cumulant (variance) and the p-order cumulant of the input signal $s[n]$, respectively, $||\mathbf{f}_i||_m = f_i[m]$ is a vector which collects the equalizer coefficients, $||\mathbf{R}_x||_{km} = R_x[k-m] \stackrel{\text{def}}{=} \mathrm{E}\{x[k] \cdot x[m]\}$ is the covariance matrix of the observed series $x[n]$, and the vector $||\boldsymbol{\kappa}_{s_{i-1}x}^{(p,1)}||_m = \kappa_{s_{i-1}x}^{(p,1)}[m]$ collects the diagonal slice of cross-cumulants of order $(p,1)$

$$\kappa_{s_{i-1}x}^{(p,1)}[m] \stackrel{\text{def}}{=} \mathrm{cum}\left(\overbrace{s_{i-1}[n], \cdots, s_{i-1}[n]}^{p}, x[n-m]\right)$$

being $s_{i-1}[n] = f_{i-1}[n] * x[n]$ the estimation of the unknown series $s[n]$ at iteration $(i-1)$.

Note that, the right–hand side of (2) involves the cumulants of order $(p+1)$ due to the choice of the power of order p in the spiking criterion (1).

The generic cross–cumulant of order $(p,1)$ can also be obtained as the following nonlinear "hybrid" moment [7]

$$\kappa_{s_{i-1}x}^{(p,1)}[m] = \mathrm{E}\{x[n-m] \cdot \gamma_p(s_{i-1}[n])\} \qquad (4)$$

where $\gamma_p(s)$ is a suitable polynomial of order p, with coefficients depending on the moments (up to the order p) of the random variable s. For instance, using $\gamma_2(s) = s^2$ yields the third–order slice cross–cumulants for zero–mean series while $\gamma_3(s) = s^3 - 3 \cdot \sigma_s^2 \cdot s$ obtains the fourth–order cross–cumulants for zero–mean series with variance σ_s^2.

Question: *What happens if we consider general nonlinearities $g(\cdot)$ in (4) and use these nonlinear moments in the solving equations (2) rather than the $(p+1)$–order cumulants?*

The answer is: *In the spiking equation (1) another nonlinearity $\eta(\cdot)$ (depending on $g(\cdot)$ but not the same!) is considered in lieu of the power of order p!*

This result generalizes the Super Exponential methods including all those nonlinearities $\eta(\cdot)$ resulting in a spiking action on the overall transformation $c[n]$. The proof of the previous claim relies on the representation of the nonlinear moments $\mathrm{E}\{x[n-m] \cdot g(x[n])\}$ by the cumulant series decomposition introduced in [8, 9]. In fact, let us consider a nonlinearity $g(s_{i-1}[n])$ instead of $\gamma_p(s_{i-1}[n])$ in (4)

$$r_{gx}[m] \stackrel{\text{def}}{=} \mathrm{E}\{x[n-m] \cdot g(s_{i-1}[n])\} \qquad (5)$$

and express it through a cumulant series expansion [8, 9]

$$\mathrm{E}\{x[n-m] \cdot g(s_{i-1}[n])\} = \sum_{p=1}^{\infty} \frac{1}{p!} \mathrm{E}\left\{g^{(p)}(s_{i-1}[n])\right\} \kappa_{s_{i-1}x}^{(p,1)}[m] \qquad (6)$$

having denoted by $g^{(p)}(\cdot)$ the p-order derivative of $g(\cdot)$.

Now, let us consider the generalized problem of determining an equalizer $f_i[n]$ such that the overall response $c_i[n]$ is "spiked" through a nonlinearity $\eta(\cdot)$ rather than a power; expanding $\eta(\cdot)$ in Taylor series, we have:

$$c_i[n] = \eta(c_{i-1}[n]) = \sum_{p=1}^{\infty} \alpha_p (c_{i-1}[n])^p \qquad (7)$$

It is straightforwardly shown that the LMS equalizer is obtained from the following set of "generalized" normal equations

$$\mathbf{R}_x \cdot \mathbf{f}_i = \sum_{p=1}^{\infty} \alpha_p \cdot \frac{\sigma_s^2}{\kappa_s^{p+1}} \cdot \boldsymbol{\kappa}_{s_{i-1}x}^{(p,1)} \qquad (8)$$

Note that (8) has the same structure of (2); the only difference is that the right–hand side of (8) involves all the higher–order cross–cumulants $\kappa_{s_{i-1}x}^{(p,1)}$, weighed by the coefficients of the Taylor expansion of the spiking nonlinearity $\eta(\cdot)$ in (7).

Finally, we see that the right–hand side of (8) coincides with the nonlinear cross-correlations of (6) when it results

$$\alpha_p = \frac{1}{p!} \cdot \frac{\kappa_s^{p+1}}{\sigma_s^2} \cdot \mathrm{E}\left\{g^{(p)}(s_{i-1}[n])\right\}$$

In other words, the nonlinear "spiking" criterion

$$c_i[n] = \sum_{p=1}^{\infty} \frac{1}{p!} \cdot \frac{\kappa_s^{p+1}}{\sigma_s^2} \mathrm{E}\left\{g^{(p)}(s_{i-1}[n])\right\} \cdot (c_{i-1}[n])^p$$

[1] For notation simplicity and for readibility purposes, in this paper we refer to the case of real sequences and LSI transformations. The complex case does not involve any special consideration and it can be obtained by properly defining higher–order cumulants.

[2] Apart pathological situations of multimodal $c[n]$ having extremal values numerically equal.

is implicitely adopted when the nonlinear cross–correlations (6) are considered in the right–hand side of the normal equations (8) which obtain the equalizer. Note also that, advoking again the cumulant series expansion, the spiking criterion (7) can also be written in the following form

$$c_i[n] = \eta\left(c_{i-1}[n]\right) = \mathrm{E}\left\{s[n] \cdot g\left(s_{i-1}[n]\right)\right\}$$

where the dependence of $c_{i-1}[n]$ is hidden in the nonlinearity $g\left(s_{i-1}[n]\right)$.

In summary, the nonlinearity $\eta(\cdot)$ acting on the overall LSI transformation $c_i[n]$ takes into account all the higher–order cumulants of the non–Gaussian input $s[n]$, weighing them by expectations of derivatives of the nonlinearity $g(\cdot)$. Moreover, it should be expected a perfomance improvement if the weights are suitable chosen, *i.e.* the nonlinearity $g(\cdot)$ is obtained according to some optimality criterion. For instance, in [3] performance analysis of the Super Exponential deconvolution shows that sample fourth order cumulants suffice to obtain zero ISI at convergence when constant modulus constellations are considered as input. In [10], performance analysis of the generalized Super Exponential deconvolution herewith presented shows that, for (odd) polynomial nonlinearities of seventh order $g(s)=a\cdot s+b\cdot s^3+c\cdot s^5+d\cdot s^7$, there exist optimal values of coefficients (a, b, c, d) that result in zero ISI also when non constant modulus constellations are considered.

III Bussgang Deconvolution

As above outlined, among the nonlinearities $g(\cdot)$ optimal choices could be made according to suitable optimality criteria. To investigate in this direction, let us briefly recall Bussgang deconvolution algorithm [1, 2]. The equalizer $\mathbf{f}_i$ is obtained (at iteration i) solving the classical linear set of *normal* equations (Wiener filter) and normalizing:

$$\mathbf{R}_{\mathrm{x}} \cdot \mathbf{f}_i' = \mathbf{r}_{\hat{s}_i x} \tag{9}$$

$$\mathbf{f}_i = \mathbf{f}_i' \cdot \left(\mathbf{f}_i'^{\mathrm{T}} \cdot \mathbf{R}_{\mathrm{x}} \cdot \mathbf{f}_i'\right)^{-\frac{1}{2}} \tag{10}$$

where the vector

$$||\mathbf{r}_{\hat{s}_i x}||_m = R_{\hat{s}_i x}[m] \stackrel{\mathrm{def}}{=} \mathrm{E}\left\{x[n-m] \cdot \hat{s}_i[n]\right\}$$

collects the cross-correlation lags between the observations $x[n]$ and an estimate of the unknown series $\hat{s}_i[n]$.

The key idea in Bussgang deconvolution is to obtain a sample estimate of the cross-correlations $\mathrm{E}\left\{\hat{s}_i[n] \cdot x[n-m]\right\}$ using a MMSE estimate of $s[n]$ drawn from the following signal–in–noise model for the Wiener estimate at iteration $(i-1)$

$$z[n] \stackrel{\mathrm{def}}{=} s_{i-1}[n] = f_i[n] * x[n] = s[n] + w_i[n] \tag{11}$$

In (11), the "deconvolutional" noise

$$w_i[n] \stackrel{\mathrm{def}}{=} \left(f_i[n] - f_i[0]\delta[n]\right) * x[n]$$

is assumed Gaussian, white and independent of $s[n]$. In [1, 2], it is claimed that this assumption can be reasonably substained at convergence, *i.e.* when $w_i[n]$ becomes "long and oscillatory".[3] In these hyphoteses, the MMSE estimation of $s[n]$, given $z[n]$ in (11), is the conditional *a posteriori* mean and does not depend on the time index n. Dropping it for simplicity, we have

$$\hat{s}_i(z) = \int z \cdot p_{S_i/Z}(z)\, dz \tag{12}$$

where $p_{S_i/Z}(\cdot)$ is the conditional *a posteriori* pdf of $s_i[n]$ given $z[n]$. Note that if the deconvolutional noise $w_i[n]$ is white and independent of $s[n]$, the conditional mean is simply a nonlinearity acting sample–by–sample on $z[n] \stackrel{\mathrm{def}}{=} s_{i-1}[n]$, *i.e.* the zero–memory MMSE estimator takes the form $\hat{s}[n]=g\left(s_{i-1}[n]\right)$ where the nonlinearity $g(\cdot)$ depends on the pdf of $s[n]$ and $w_i[n]$. Note that the shape of the conditional estimator (12) changes at each iteration since it actually depends on the variance of the Gaussian "deconvolutional" noise $w_i[n]$, which, in turn, diminishes as iterations proceed. This estimate is used to form the right–hand side of (9) through the nonlinear cross-correlations[4]

$$R_{\hat{s}_i x}[m] = \mathrm{E}\left\{g\left(s_{i-1}[n]\right) \cdot x[n-m]\right\} \tag{13}$$

Now, look at the solving equations of the two blind deconvolution methods (2) and (9): they are identical in structure, the only difference being constituted by the right–hand sides. The generic elements of the right–hand sides are expressed by (4) and (13), respectively: they are nonlinear expectations, the only difference being the nonlinearity, chosen to obtain the cross–cumulants (corresponding to spiking through powers) in (4) and according to the conditional expectation (extraction of signal in white deconvolutional noise) in (13).

In the framework of the above described generalization of Super Exponential methods, we conclude that Bussgang deconvolution obtain the equalizer by "spiking" the overall impulse response $c[n]$ using all the available probabilistic description in the conditional pdf $p_{S_i/Z}(\cdot)$.
For instance, in Fig.2 the implicit "spikyness" criterion is plotted in the case of a binary signaling communications scheme, *i.e.* considering an i.i.d. input series $s[n]$ assuming the values ± 0.5 with identical probabilities. The curves in Fig.2 have been obtained by approx-

[3] The assumption of Gaussianity of the deconvolutional noise $w_i[n]$ has been discussed also in [11] where it has been shown that Gaussianity cannot be substained even at converence. Since the estimator (12) does not take properly into account the pdf of the convolutional noise, it has to be considered as a suboptimal estimate; it is adopted mainly for its simplicity.

[4] The *Bussgang* denomination is due to the fact that the equalizer satisfies $\mathbf{f}_i \simeq \mathbf{f}_{i-1}$ at convergence and this in turn implies the invariance of the cross-correlation under a nonlinear transformation $\mathrm{E}\left\{x[n-m] \cdot s_i[n]\right\} \propto \mathrm{E}\left\{g\left(s_i[n] \cdot x[n-m]\right)\right\}$.

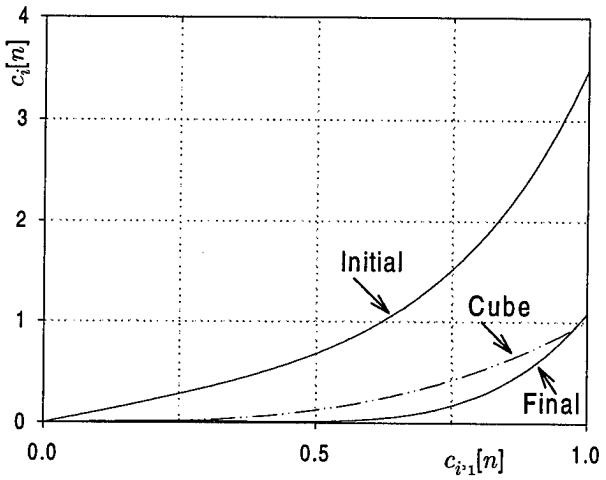


Figure 2: *The implicit "spikyness" criterion in Bussgang deconvolution of a binary signaling communications scheme.*

imating (in an uniformly MMSE sense) the nonlinear estimator (12) as a fifth order polynomial. The deconvolutional noise $w[n]$ has been assumed white, zero–mean, Gaussian with variance $\sigma)w^2=0.1$, *i.e.* $\simeq 4$ dB below the binary signal power. The curve labelled "Initial" in Fig.2 refers to the case of uniformly distributed observations $z[n]$ over $(-1,1)$, *i.e.* to the initial iterations of the algorithm; instead, the curve labelled "Final" refers to the case of binary (perfectly equalized) observations $z[n]$, *i.e.* near the convergence of the algorithm. For comparison purposes, also the "Cube" curve is reported, corresponding to the spiking through fourth–order cumulants. We see that the spiking in Bussgang deconvolution is "adapted" to the state of the equalized data: initially, high values are enhanced without suppressing small values (this reflects some initial uncertainity about the true position of the spike), while, near the covergence, small values are shrinked without enhancing the already adjusted spike (which is assuming value $\simeq 1$. On the other hand, the cubic law acts on high and small values always in the same fashion during the iterations.

IV Optimality of Bussgang Deconvolution

Despite its simple formulation, Bussgang Deconvolution has been shown to be an approximate algorithmic implementation of an optimal Bayesian solution derived in [4] and reported also in [5, 6]. Loosely speaking, the normal equations (9) and the conditional mean (12) results from the differentiation of the quadratic cost function w.r.t. the unknown equalizer paremeters and w.r.t. the unobserved input series $s[n]$, respectively.

This confirms the fact that nonlinearities can be optimally chosen in the generalized Super Exponential blind deconvolution. Bussgang deconvolution offers the possibility of controlling the "spikyness" as iterations proceed by adapting the estimation to the amount of "deconvolutional" noise not yet removed; on the other hand, optimal "fixed" parameters polynomial nonlinearities can be optimally designed to drive ISI to zero, as performance analysis conducted in [3, 10] have been shown.

V Conclusion

A generalization of Super Exponential blind deconvolution method has been presented. The inherent "spikyness" criterion is allowed to include general nonlinearities rather than only powers as in the original formulation.

This generalization allows to interpretate Bussgang deconvolution as a particular instance of generalized Super Exponential deconvolution obtained by selecting the spikyness criterion according to the known pdf of the input series to be recovered.

The usefulness of such an approach is confirmed by performance analysis which shows that zero ISI can be obtained also when non constant modulus constellations are considered.

References

[1] R. Godfrey, F. Rocca, "Zero–memory nonlinear deconvolution", *Geophysical Prospecting*, vol.29, 1981.

[2] S. Bellini, "Bussgang techniques for blind deconvolution and equalization", in *Blind Deconvolution*, S. Haykin ed., Prentice–Hall, 1994.

[3] O. Shalvi, E. Weinstein, "Universal methods for blind deconvolution", in *Blind Deconvolution*, S. Haykin ed., Prentice–Hall, 1994.

[4] G. Jacovitti, A. Neri, G. Scarano, "Complex reflectivity based non-minimum phase deconvolution", in *Deconvolution and Inversion*, M.H. Worthington ed., Blackwell Scientific Publications, London, U.K., 1987.

[5] G. Jacovitti, A. Neri, G. Scarano, "An iterative approach to non-minimum phase AR identification by non-linear moments", presented at *Workshop on Higher-Order Spectral Analysis*, Vail (Co), U.S.A., June 1989.

[6] A. Neri, G. Scarano, G. Jacovitti, "A Bayesian iterative method for blind deconvolution", *SPIE Annual Symposium*, San Diego (CA), USA, July 1991.

[7] G. Scarano, G. Jacovitti, "Applications of generalized cumulants to array processing', *Sig. Proc.*, vol. 53, no. 2, 1996.

[8] G. Scarano, "Cumulant series expansion of hybrid non-linear moment of complex random variables", *IEEE Transactions on Signal Processing*, vol.39, no.4, April 1991.

[9] G. Scarano, D. Caggiati, G. Jacovitti, "Cumulant series expansion of hybrid non-linear moment of *n* variates", *IEEE Trans. Sig. Proc.*, vol.41, no.1, January 1993.

[10] E. De Angelis, "Prestazioni di algoritmi per la deconvoluzione alla cieca", *Tesi di Laurea, Dip. INFOCOM, Università di Roma "La Sapienza"*, July 1997.

[11] Z. Ding, C.R. Johnson, R.A. Kennedy, "Global convergence issues with linear blinad adaptive equalizers", in *Blind Deconvolution*, S. Haykin ed., Prentice–Hall, 1994.

Recursive Estimation Algorithm for FIR Systems Using the 3rd and 4th Order Cumulants

Hyoungill Kim, Bumki Jeon, Taewon Yang, and Koeng-Mo Sung
the School of Electrical Engineering, College of Eng.,
Seoul National University, Seoul, 151-742, KOREA

Abstract

A recursive estimation algorithm for FIR systems is proposed using the 3rd and 4th order cumulants. From the 3rd and 4th order cumulants relationship, we construct a certain matrix form whose entry is consists of the system output sequence. Using this matrix form, the proposed recursive algorithm is developed by Overdetermined Recursive Instrumental Variable(ORIV) method. The proposed algorithm provides improved estimation accuracy when additive Gaussian noise is present and can be applied to a time varying system as well. Simulation results are presented to compare the performance with other HOS-based algorithms.

I. Introduction

Recently, new identification algorithms using higher order cumulants have appeared in the literature: GM-method [1] and Adaptive IIR algorithm based on GM-equation [2], *etc.* The GM-method and its modified algorithms show erratic behaviors when the output data are corrupted by additive Gaussian noise since they include the correlation terms. The new estimation algorithm using the 3rd and 4th order cumulants [3] has been proposed to avoid the above-mentioned erratic behaviors, alleviating the effect of additive white/colored Gaussian noise. This algorithm uses the 3rd and 4th order cumulants but not the correlation terms. As a result, it is blind to additive white/colored Gaussian noise even if the SNR is low.

In general, higher than second order cumulants are commonly of high variance and therefore a large number of record is required to obtain reliable estimates of higher order cumulants. Compared with the batch form algorithms, recursive algorithms have some merits, i.e., tracking capability of a time varying system and less strict requirement on the number of data. Hence, we propose a new recursive algorithm for an FIR system as an extension of the algorithm in [3]. At first, we transform the relationship between 3rd and 4th cumulant into the matrix form as a least square solution. This matrix is consists of 3rd and 4th order cumulants. To derive the recursive estimation algorithm using this matrix form, we reformulate the component of matrix whose entries are consist of output sequence by introducing the instrumental variable. Using reformulated matrix, we get the time recursive estimation algorithm by following the ORIV derivation procedure.

The outline of the paper is as follows: In Section II, we introduce the GM equation and the relationship between 3rd and 4th order cumulants. And, the proposed recursive algorithm based on 3rd and 4th order cumulants is presented. In Section III, we present a set of numerical examples illustrating the behavior of the proposed algorithm with comparison to the other algorithm, and show the performance of tracking capability of a time varying system..

II. Proposed Algorithm

Let an observed process be

$$y(n) = x(n) + w(n) = \sum_{i=0}^{q} b(i)v(n-i) + w(n) \qquad (1)$$

where $x(n)$ is the system output, $\{b(n)\}$ is the impulse response of an unknown system, and $\{v(n)\}$ is i.i.d. non-Gaussian and has an asymmetric p.d.f. with $E\{v(n)\} = 0$, $\gamma_{3,v} = E\{v^3(n)\} \neq 0$, $\gamma_{4,v} = E\{v^4(n)\} - 3[E\{v^2(n)\}]^2 \neq 0$. The additive noise $\{w(n)\}$ is Gaussian and independent of $\{v(n)\}$.

The GM equation[1] is show the relationship between correlation and cumulant as follows.

$$c_{2,y}(m) + \sum_{i}^{q} b^2(i)c_{2,y}(m-i) =$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$\varepsilon_{2,3}\left[c_{3,y}(m)+\sum_{i=1}^{q}b(i)c_{3,y}(m-i)\right], \text{ for } -q\le m\le 2q \qquad (2)$$

where $c_{3,y}(m)=c_{3,y}(m,m)=E\{y(n)y(n+m)^2\}$, $c_{2,y}(m)$ is correlation term of output sequence with m lags, and $\varepsilon_{2,3}=\gamma_{2,y}/\gamma_{3,y}$. The GM equation based algorithms are performed well for system identification even if the system is nonminimum phase. However, they fail to estimates when the output sequence is corrupted by additive noise due to bias of correlation term in sprite that the cumulants is blinded to additive noise. In this motivation, the relationship between 3rd and 4th order cumulants is proposed as was shown in [3],

$$c_{3,y}(m)+\sum_{i=1}^{q}b^3(i)c_{3,y}(m-i)= \varepsilon\left[c_{4,y}(m)+\sum_{i=1}^{q}b^2(i)c_{4,y}(m-i)\right], \text{ for } -q\le m\le 2q \qquad (3)$$

where, $\varepsilon=\gamma_{3,y}/\gamma_{4,y}$ and $c_{4,y}(m)=c_{4,y}(m,m,m)=E\{y(n)y(n+m)^3\}-3E\{y(n)y(n+m)\}\,c_{2,y}(0)$. Eq. (3) can be recast in a following matrix form as,

$$A\cdot\theta=b, \qquad (4)$$

and

$$A=\begin{bmatrix} c_{4,y}(-q) & & 0 & 0 & \cdots & 0 \\ c_{4,y}(-q+1) & \ddots & \vdots & c_{3,y}(-q) & \ddots & \\ \vdots & \ddots & c_{4,y}(-q) & \vdots & \ddots & 0 \\ c_{4,y}(q) & & c_{4,y}(-q+1) & c_{3,y}(q) & & c_{3,y}(-q) \\ \vdots & \ddots & \vdots & & \ddots & \vdots \\ 0 & & c_{4,y}(q) & 0 & & c_{3,y}(q) \end{bmatrix},$$

$$\theta=\begin{bmatrix}\varepsilon_{3,4} & \varepsilon_{3,4}b^2(1) & \cdots & \varepsilon_{3,4}b^2(q) & -b^3(1) & \cdots & -b^3(q)\end{bmatrix}^T,$$

$$b=\begin{bmatrix}c_{3,y}(-q) & \cdots & \cdots & c_{3,y}(q) & 0 & \cdots & 0\end{bmatrix}^T,$$

where, A is the $(3q+1)\times(2q+1)$ matrix, and b is the $(3q+1)$ vector. In general, we can obtain the least squares solution of (4) as $\hat{\theta}=(A^TA)^{-1}A^Tb$. To derive the recursive estimation algorithm from (4), we propose, instead, to replace A and b by the following estimates. Define

$$\hat{A}=Z_n^T[C_{4,n}|C_{3,n}], \quad \hat{b}=Z_n^Tc_{3,n} \qquad (5)$$

where,

$$Z_n=\begin{bmatrix} y(0) & & 0 \\ \vdots & \ddots & \\ \vdots & & y(0) \\ \vdots & & \vdots \\ y(n) & \cdots & y(n-3q)\end{bmatrix}, \quad c_{3,n}=\begin{bmatrix}0\\ \vdots\\ 0\\ y^2(0)\\ \vdots\\ y^2(n-q)\end{bmatrix},$$

$$C_{3,n}=\begin{bmatrix} 0 & & 0 \\ \vdots & \ddots & \\ 0 & & 0 \\ y^2(0) & \ddots & \vdots \\ \vdots & \ddots & 0 \\ \vdots & & y^2(0) \\ \vdots & & \vdots \\ y^2(n-q) & \cdots & y^2(n-2q)\end{bmatrix},$$

$$C_{4,n}=\begin{bmatrix} 0 & & 0 \\ \vdots & \ddots & \\ 0 & & 0 \\ y^3(0)-\alpha y(0) & \ddots & \vdots \\ \vdots & & 0 \\ \vdots & & y^3(0)-\alpha y(0) \\ \vdots & & \vdots \\ y^3(n-q)-\alpha y(n-q) & \cdots & y^3(n-2q)-\alpha y(n-2q)\end{bmatrix},$$

where, $\alpha=3c_{2,y}(0)$, and $\hat{A}$ and $\hat{b}$ is estimate of A and b, respectively. Z_n acts as instrumental variable with $(n+1)\times(3q+1)$ matrix, $c_{3,n}$ is the vector with $(n+1)\times 1$, and $C_{4,n}$ and $C_{3,n}$ is the matrix with $(n+1)\times(q+1)$ and $(n+1)\times q$, respectively. Note that the elements of $Z_n^TC_{4,n}$ and $Z_n^TC_{3,n}$ are the consistent estimates of $c_{4,y}(m)$ and $c_{3,y}(m)$, for $-q\le m\le q$, respectively. This is in agreement with the matrix appearing in the left-hand side of (4). Similarly, $Z_n^Tc_{3,n}$ is also a consistent estimate of the right-hand side of (4).

The estimates of the parameter vector is given by the least squares solution to (5);

$$\hat{\theta}_n=\left(Y_n^TZ_nZ_n^TY_n\right)^{-1}Y_n^TZ_nZ_n^Tc_{3,n} \qquad (6)$$

where $Y_n=\left[C_{4,n}|C_{3,n}\right]$. Following the derivation of ORIV[4], the time update of parameter vector can be separately involved as follows:

$$P_n=\left(Y_n^TZ_nZ_n^TY_n\right)^{-1}, S_n=Y_n^TZ_n, L_n=Z_n^Tc_{3,n}. \qquad (7)$$

The P_n and S_n matrix can be updated as was described in ORIV method. The most significant difference between the proposed and the algorithm of [2] lies in L_n which correspond to the $\hat{b}$ matrix. It can be recursively updated by

$$L_{n+1} = \lambda L_n + z_{n+1} y^2(n+1), \tag{8}$$

where, the $0 \le \lambda \le 1$ is an exponential forgetting factor which enables the algorithm to track time varying parameter, and z_{n+1} is $(n+1)$th column vector of Z_{n+1}. From (7), we note that

$$\begin{aligned}\hat{\theta}_{n+1} &= P_{n+1} S_{n+1} L_{n+1} \\ &= P_{n+1}\left(\lambda S_n + y_{n+1} z_{n+1}^T\right)\left(\lambda L_n + z_{n+1} y^2(n+1)\right)\end{aligned} \tag{9}$$

From time update of P_{n+1} and S_{n+1}, we obtain the following time update equation of parameter vector $\hat{\theta}_{n+1}$

$$\hat{\theta}_{n+1} = \hat{\theta}_n + P_{n+1}\phi_{n+1}\Lambda_{n+1}^T\left(V_{n+1} - \phi_{n+1}^T\hat{\theta}_n\right) \tag{10}$$

where, ϕ_{n+1} and Λ_{n+1} are same as was discussed in ORIV method, and $V_{n+1} = [z_{n+1}^T L_n \;\; y^2(n+1)]^T$. Finally, the proposed recursive algorithm using the 3rd and 4th order cumulants is summarized in Table 1. We refer the reader to [4] for details.

In estimating the 4th order cumulant, the estimate of correlation $c_{2,y}(0)$ is needed. The corresponding time update formula would be

$$c_{2,y,n+1}(0) = \left(n \cdot c_{2,y,n}(0) + y^2(n+1)\right)/n+1 \tag{11.1}$$

or

$$c_{2,y,n+1}(0) = \beta \cdot c_{2,y,n}(0) + (1-\beta) y^2(n+1) \tag{11.2}$$

Eq. (11.1) and (11.2) can be used for time invariant system case and time variant system case, respectively.

III. Simulation Results

The proposed algorithm is compared with the recursive algorithm using the 2nd and 3rd order cumulants [2], batch algorithm in [3], and the least square solution of (6). The input sequence $\{v(n)\}$, which is an independent exponentially distributed zero-mean random sequence with $\sigma_v^2 = 1$, $\gamma_{3,v} = 2$, and $\gamma_{4,v} = 6$, was generated, and additive Gaussian noise was added to the observed output sequence. We performed 100 Monte Calro runs for each of the algorithms with 3000 data samples. The simulated system was

$$x(n) = v(n) - 1.25v(n-1) \text{ (nonminimum system).} \tag{12}$$

The colored Gaussian noise is generated after passing a white, zero-mean Gaussian noise through an MA filter with coefficients [1, -2.33, 0.75, 0.5, -1.3, -1.4].

Table I.
SUMMARY OF THE PROPOSED ALGORITHM

Initialization

$S_0 = \mu[I_{2q+1}|0]$ $(2q+1)\times(3q+1)$, $P_0 = \mu^{-1} I_{2q+1}$

$L_0 = 0$ $(3q+1)\times 1$, $\hat{\theta}_0 = 0$ $(2q+1)\times 1$

For $n = 0,1,2,\cdots$

$z_{n+1} = [y(n+1)\cdots y(n-3q+1)]^T$

$$\begin{cases} c_{2,y,n+1}(0) = (n \cdot c_{2,y,n}(0) + y^2(n+1))/(n+1) \\ c_{2,y,n+1}(0) = \alpha \cdot c_{2,y,n}(0) + (1+\alpha) y^2(n+1) \end{cases}$$

$x_{n+1} = [y^3(n-q+1)\cdots y^3(n-2q+1)|y(n-q)\cdots y(n-2q+1)]^T$

$x_{n+1} = x_{n+1} - 3c_{2,y,n+1}(0)[y(n-q+1)\cdots y(n-2q+1)|0\cdots 0]^T$

$w_{n+1} = S_n z_{n+1}$, $S_{n+1} = \lambda S_n + x_{n+1} z_{n+1}^T$

$Q_{n+1} = [w_{n+1} x_{n+1}]$, $A_{n+1} = \begin{bmatrix} -z_{n+1}^T z_{n+1} & \lambda \\ \lambda & 0 \end{bmatrix}$

$K_{n+1} = P_n Q_{n+1}\left(A_{n+1} + Q_{n+1}^T P_n Q_{n+1}\right)^{-1}$

$P_{n+1} = \lambda^{-2}[P_n - K_{n+1} Q_{n+1} P_n]$

$v_{n+1} = \left[z_{n+1}^T L_n \;\; y^2(n-q+1)\right]^T$, $L_{n+1} = \lambda L_n + z_{n+1} y^2(n-q+1)$

$\hat{\theta}_{n+1} = \hat{\theta}_n + K_{n+1}(v_{n+1} - Q_{n+1}^T \hat{\theta}_{n+1})$

Table II.
Simulation Results
(true parameter $b(1)$=-1.25, N = 3000)

	SNR	0 dB	5dB	10dB	∞dB
proposed algorithm	mean	-1.1473	-1.2711	-1.2763	-1.2767
	std.	0.4944	0.1049	0.0843	0.0747
[2] (2'nd and 3'rd)	mean	-1.3721	-1.4368	-1.4442	-1.3444
	std.	0.4629	0.1769	0.1432	0.1274
[3] batch algorithm	mean	-1.0753	-1.1305	-1.1308	-1.1568
	std.	0.0451	0.0339	0.0261	0.0286
batch alg. Of Eq.(4)	mean	-1.0753	-1.1257	-1.1524	-1.1564
	std.	0.0514	0.0315	0.0253	0.0276

The solution of (6) does directly not tell us the sign of $b(k)$. The two choices are exist but we will take following choice since the division by the estimate of $\varepsilon_{3,4}$ as a process of obtaining the estimates of $b^2(k)$ may give rise to a very large magnitude of estimates by small $\varepsilon_{3,4}$.

$$\hat{b}(k) = \sqrt[3]{estimates\ of\ b^3(k)} \qquad (13)$$

The obtained curve is depicted in Fig. 1, and the statistics (mean ± variance) of estimated coefficients are listed in Table II.

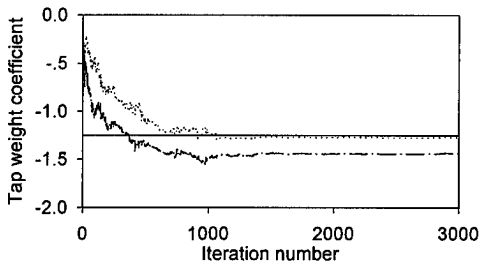


Fig. 1. The mean trajectories of estimated parameters, in 5 dB SNR case.
Solid line: true, Dotted line: proposed,
Dashed-dotted line: alg. in [2] (2nd and 3rd)

As we see, the batch algorithm based on [3] and the recursive algorithm in [2] show consistent performance for noiseless case. However, when the output sequence is corrupted by additive Gaussian noise, these estimate the parameter with bias value. On the contrary, the proposed algorithm shows the consistent performance even if the additive Gaussian noise is not unknown whether it is white or colored. For the comparison of batch algorithm, the estimates of recursive algorithm are more exact than those of batch algorithm. However, we can see that the standard deviations of estimates using recursive algorithm are larger than those of estimates using batch algorithm. From fig. 1, we can see the convergence of the proposed algorithm is obtained at 1000 points which is relatively smaller than data record length requirements for batch algorithm.

For the identification of a time varying system, we initially generated data using $x(n) = v(n) - 0.8v(n-1)$, and switched after 2500 data samples to another MA system described by the above nonminimum system. Fig. 2 shows the time histories of $\hat{b}(1)$. Note that the propose algorithm indeed tracks the variation of MA parameter, with some delay.

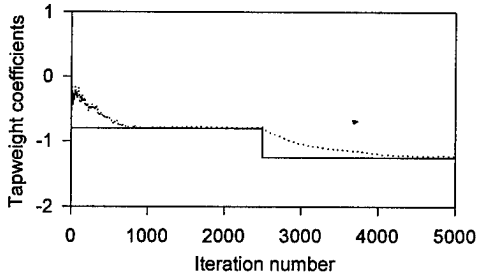


Fig. 2. The mean trajectory of estimated parameter in a time varying case.
Solid line: true, Dotted line: proposed

IV. Conclusions

A cumulant based recursive parameter estimation algorithm for MA system identification was considered. We have developed ORIV type recursive algorithm, and applied it to a particular set of equations involving the 3rd and 4th order cumulants. Since the proposed algorithm did not use the correlation terms, it had a blindness property against additive white/color Gaussian noise. Hence, the proposed algorithm was useful when the SNR was low, or the a priori information of noise was not known. In addition, the proposed algorithm showed satisfactory performance when we estimated the time varying MA system.

Refèrences

[1] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360-377, Mar, 1989.

[2] B. Friedlander and B. Porat, "Adaptive IIR Algorithms Based on Higher-Order Statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 485-495, April 1989.

[3] Y. J. Na, K. S. Kim, I. Song, and T. Kim, "Identification of Nonminimum Phase FIR Systems Using the Third- and Fourth-Order Cumulants," *IEEE Trans. Signal Processing*, vol. 43, pp. 2018-2022, Aug. 1995.

[4] B. Friedlander, "The Overdetermined Recursive Instrumental Variable Method," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 353-356, April 1984.

TAL: Alpha-Stable Distributions

Chair: Brian Sadler
Army Research Laboratories, USA

Zero-Order Statistics: A Signal Processing Framework for Very Impulsive Processes *

Juan G. Gonzalez, David W. Griffith, and Gonzalo R. Arce
Department of Electrical and Computer Engineering
University of Delaware, Newark, DE 19716, USA
gonzalez@ee.udel.edu

Abstract

Techniques based on conventional higher-order statistics fail when the underlying processes become impulsive. Although methods based on fractional lower-order statistics (FLOS) have proven successful in dealing with heavy-tailed processes, they fail in general when the noise distribution has very heavy *algebraic tails, i.e., when the algebraic tail constant is close to zero. In this paper we introduce a signal processing framework that we call* Zero-Order Statistics *(ZOS). ZOS are well defined for any process with algebraic or lighter tails, including the full class of α-stable distributions. We introduce zero-order scale and location statistics and study several of their properties. The intimate link between ZOS and FLOS is presented. We also show that ZOS are the optimal framework when the underlying processes are* very impulsive.

All figures, simulations and source code utilized in this paper are reproducible and freely accessible in the Internet at http://www.ee.udel.edu/~gonzalez/PUBS/HOS97a

1. Introduction

Second-order processes have been historically the main subject of study in statistical signal processing. Second-order-based estimation techniques are commonly recognized as the natural tools to be used in the presence of Gaussian noise. Research efforts on higher-order statistics (HOS) have led to the development of improved estimation algorithms for non Gaussian environments, but this work has been based on the assumptions that second-order and higher-order statistics of the processes exist and are finite [7]. Important non Gaussian impulsive processes, found in radar and mobile communications for example, can be efficiently modeled by infinite variance processes for which the theory of HOS is not useful [3, 4, 5, 6, 8].

It has been shown repeatedly in the literature that heavy-tailed processes that appear in practice are well modeled by probability distributions with algebraic tails, i.e., random variables for which $\Pr(|X| > x) = O(x^{-\alpha})$, for some fixed $\alpha > 0$. Examples of such noise processes include those modeled by α-stable distributions [8], Hall's generalized t-model [3], the generalized Cauchy model [6], and the Pareto distribution. The tail-heaviness of these distributions is mainly determined by the tail constant α, with increased impulsiveness corresponding to smaller values of α.

Algebraic-tailed random variables exhibit finite absolute moments for orders less than α; i.e., $E|X|^p < \infty$ if $p < \alpha$. Conversely, if $p \geq \alpha$, the absolute moments become infinite, making them unsuitable for statistical analysis. When $\alpha < 2$, the processes present infinite variance, and the standard second or higher-order statistics cannot be successfully applied. Attempts to characterize the behavior of impulsive signals in this scenario have relied on *fractional lower-order statistics* (FLOS) in the context of non Gaussian α-stable distributions ($\alpha < 2$). Here, given a fixed α, appropriate choices of p in the interval $(0; \alpha)$ can lead to useful characterizations of the process structure [8].

While FLOS have been proven useful when dealing with impulsive signals, they present several shortcomings. First, FLOS do not provide a universal framework for dealing with algebraic-tailed processes. Since p is usually restricted to the interval $(0; \alpha)$, constructing a valid FLOS requires the previous knowledge (or estimation) of α in order to pick an appropriate value of p. On the other hand, for any given $p > 0$, there will always be a "remaining" class of very impulsive processes (those with $\alpha \leq p$) for which the associated FLOS are not appropriate.

In this paper we introduce the preliminaries of a new theory of statistics which is well defined over all distributions with algebraic or less-heavy tails. Unlike lower or higher-order statistics, these *zero-order statistics* or *ZOS*, as we

*This research was funded in part by the National Science Foundation under the Grant MIP-9530923, and through collaborative participation in the Advanced Telecommunications/Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under Cooperative Agreement No. DAAL01-96-2-0002.

0-8186-8005-9/97 $10.00 © 1997 IEEE

will call them, provide a common ground for the analysis of basically *any* known distribution of practical use. In the same way as p^{th}-order moments constitute the basis of FLOS and HOS techniques, the theory of zero-order statistics is based on logarithmic "moments" of the form $E\log|X|$. We study the fundamental properties of ZOS and prove their optimality when the tail constant α tends to zero. This important property insinuates the strong potential of ZOS in the robust estimation of very impulsive processes.

2. Logarithmic-Order Processes

The following theorem provides a natural motivation for characterizing the class of processes of interest in this paper:

Theorem 1 *Let X be a random variable with algebraic or less heavier tails. Then, $E\log|X| < \infty$.*

Proof [1]: See [1].

Since our main goal is to develop a signal processing framework under which all random processes with algebraic tails can be characterized, Theorem 1 allows us to restrict our attention to processes with finite logarithmic moments. We will refer to such processes as being of "logarithmic order", in analogy with the term "second order", used to denote processes with finite variance.

Of particular interest in this paper is the class of logarithmic-order processes with infinite variance, which includes all the algebraic-tailed processes with tail constant $\alpha < 2$. Classical statistical methods, usually derived from the theory of second-order processes, present serious limitations when dealing with this kind of processes. For example, design techniques based on the minimization of the mean square error (MSE) become practically useless since the MSE turns out to be infinite all the time. Least squares regression (LS) and general linear techniques are also widely acknowledged as inappropriate in infinite variance environments [8]. Spectral characterizations, of paramount importance in engineering applications, are no longer meaningful since the (second-order) power of the processes become infinite. Even the mean parameter $\mu = EX$, basic ingredient of classical probability theory, does not exist for algebraic-tailed distributions with $\alpha < 1$. Also, fundamental estimators of location and scale such as the sample average and the sample standard deviation present serious shortcomings. The sample average rapidly loses efficiency when the tail constant departs from $\alpha = 2$ to smaller (more impulsive) values. For tail constants $\alpha \leq 1$ the sample average reaches inconsistency, becoming more harmful than useful when $\alpha < 1$. The sample standard deviation, on the other hand, is not consistent for any $\alpha < 2$.

[1] Due to lack of space, we have omitted the proofs in this paper, limiting the discussion to the most significant insights. The reader is referred to the paper website [1] for detailed proofs.

The evident limitations of classical methods in the context of logarithmic-order processes make it necessary to develop new basic statistics under which an efficient theory of estimation can be built.

3. The Geometric Power

In this section we introduce a new scale indicator, namely the *geometric power*[2], which overcomes many of the limitations of second-order theory in the framework of logarithmic-order processes. Next, in Section 4, a location indicator intimately linked with the geometric power is developed. These two parameters constitute the underpinnings of the theory of ZOS. Much as the (second-order) power and the expected value play a central role in the theory of second-order processes, we believe that the geometric power and its related location parameter are of fundamental importance in the development of a theory for logarithmic-order processes.

Definition 1 *Let X be a logarithmic-order random variable. The* geometric dispersion *of X is defined as*

$$S_0 = S_0(X) = e^{E\log|X|} \tag{1}$$

3.1. Properties

It can be easily shown that the geometric power is a scale parameter, and as such, it can be effectively used as an indicator of process strength or "power" in situations where second-order methods are inadequate. In the following, we enumerate some of the most important properties of this new parameter. The reader is referred to the paper website [1] for the proofs.

1. *S_0 is a scale parameter.* For any given constant c, $S_0(cX) = |c|S_0(X)$.
2. *S_0 is an indicator of "power" or process strength.* For all X, $S_0(X) \geq 0$. For any given constant c, $S_0(c) = |c|$. In addition, $S_0(X) = 0$ if and only if $\Pr(X = 0) > 0$, which implies that zero power is only attained when there is a discrete probability mass located in zero.
3. *Triangular Inequality.* For any pair of random variables X and Y, $S_0(X+Y) \leq S_0(X) + S_0(Y)$
4. *Chebyshev Inequality.* Let X be a symmetric random variable with symmetry center μ, such that $S_0(X-\mu) > 1$. Then,

$$\Pr(|X-\mu|^{1/c} \geq S_0(X-\mu)) \leq \frac{1}{c}. \tag{2}$$

[2] We coined the name *geometric* because of the intimate link of this parameter with the geometric mean. Unfortunately there is not enough space in this paper to illustrate this link.

5. *Multiplicativity.* For any pair of random variables X and Y, $S_0(XY) = S_0(X)S_0(Y)$

3.2. Geometric power of α-stable processes

Alpha-stable processes constitute one of the most important infinite-variance families in the logarithmic-order class. Since they are the only processes that satisfy a form of Generalized Central Limit Theorem, they can appear in practice as a result of natural stochastic phenomena. For an introduction to α-stable distributions in a signal processing context see [8].

The geometric power of a zero-centered symmetric α-stable ($S\alpha S$) distribution can be calculated as

$$S_0 = \frac{(C\gamma)^{1/\alpha}}{C}, \tag{3}$$

where $C = \exp(C_e) \approx 1.78$, is the exponential of the Euler constant, γ is the dispersion parameter, and α is the characteristic exponent of the distribution. A simple derivation of this expression is available on the website of this paper [1].

Figure 1 illustrates the soundness of the geometric power as an indicator of signal strength. The scatter on the left side was generated from a stable distribution with $\alpha = 1.99$ and $S_0 = 1$. On the right-hand side, the scatter comes from a Gaussian distribution ($\alpha = 2$) also with unitary geometric power. After an intuitive inspection of Fig. 1, it is reasonable to conclude that both of the generating processes possess the same strength, in accordance with the values of the geometric power. Contrarily, the values of the second-order power lead to the misleading conclusion that the process on the left is much stronger than the one on the right.

It is easy to find examples like the above that also disqualify FLOS-based indicators of strength in the logarithmic-order framework. In fact, fractional moments of order p present the same type of discontinuities like the one illustrated in Fig. 1 for processes with tail constants close to $\alpha = p$. The geometric power, on the other side, is consistently continuous along all the range of values of α, giving a more intuitive appealing in the context of logarithmic-order processes. This "universality" of the geometric power provides a general framework for comparing the strengths of any pair of logarithmic-order signals, in the same way as the (second-order) power is used in the classical framework.

3.3. Relation with FLOS

The geometric power is intimately linked to FLOS parameters as indicated in the following:

Theorem 2 *Let $S_p = (E|X|^p)^{1/p}$ denote the scale parameter derived from the p^{th}-order moment of X. If S_p exists for sufficiently small values of p, then*

$$S_0 = \lim_{p \to 0} S_p. \tag{4}$$

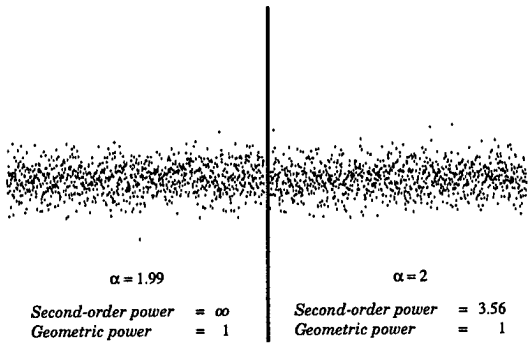


Figure 1. Comparison of second-order power Vs. geometric power for i.i.d. α-stable processes. Left: $\alpha = 1.99$. Right: $\alpha = 2$.

Furthermore, $S_0 \leq S_p$, for any $p > 0$.

Proof: See [1].

The above result indicates that techniques derived from the geometric power are the limiting *zero-order* relatives of FLOS. We coined the name "Zero-Order Statistics" as a consequence of this property.

4. Zero-Order Location Parameter

In the framework of second-order processes the mean of a random variable can be conveniently described as the parameter μ_2 that minimizes the power of the shifted variable $X - \mu$. This is,

$$\mu_2 = \arg\min_{\mu} S_2(X - \mu) \tag{5}$$

In the same way, the geometric power motivates the definition of a location parameter of fundamental importance in the theory of logarithmic-order processes:

Definition 2 *Let X be a logarithmic-order variable. We define the zero-order indicator of location as the parameter μ_0 that minimizes the geometric power of the shifted variable $X - \mu$. This is,*

$$\mu_0 = \arg\min_{\mu} S_0(X - \mu) \tag{6}$$

4.1. Zero-Order Estimation of Location

The definition of the ZOS location parameter in (6) motivates the development of a new estimator with strong potential for locating very impulsive processes. Substituting (1) in (6), we can reformulate the definition of μ_0 as

$$\mu_0 = \arg\min_{\mu} E \log|X - \mu|. \tag{7}$$

In order to propose an estimator of μ_0, an intuitive consideration of (7) would advise us to replace the expected value operator by the sample average so that we obtain

$$\hat{\mu}_0 = \arg\min_{\mu} \sum_{i=1}^{N} \log|x_i - \mu|. \tag{8}$$

This expression has the problem that its argument equals $-\infty$ at every sample value (i.e. whenever $\mu = x_i$), which results in the minimization giving a multiple tie among all the samples. Although this tie is a serious problem that has to be resolved, it gives us very important information: $\hat{\mu}_0$ is a "selection-type" estimator, in the sense that it is always equal to one of the sample values[3]. In order to resolve the tie, we can constrain the minimization problem in (8) to a compact set that does not include indeterminacies of the cost function. This can be done by defining

$$\hat{\mu}_\delta = \arg\min_{\beta \in \Delta} \sum_{i=1}^{N} \log|x_i - \beta|, \tag{9}$$

where

$$\Delta = \Re - \bigcup_{i=1}^{N} (x_i - \delta; x_i + \delta), \tag{10}$$

$\Re$ is the real numbers set and δ is a small positive constant. The tie is easily solved by letting

$$\hat{\mu}_0 = \lim_{\delta \to 0} \hat{\mu}_\delta. \tag{11}$$

Mathematical manipulation of (11) leads to a surprisingly simple expression for $\hat{\mu}_0$:

$$\hat{\mu}_0 = \arg\min_{x_j \in \mathcal{M}} \prod_{i=1, x_i \neq x_j}^{N} |x_i - x_j|, \tag{12}$$

where $\mathcal{M}$ is the set of most repeated values in the sample[4]. According to (12), $\hat{\mu}_0$ will always be one of the most repeated values in the sample, resembling the behavior of a sample mode. This *mode property*, as we call it, intuitively insinuates a high effectivity of the estimator in the presence of heavy impulsive noise. In the following sections we report both theoretical and experimental evidence of this intuition.

4.2. Properties of the ZOS location estimator

In addition to the "selection" and mode properties reported above, $\hat{\mu}_0$ presents a rich set of interesting properties. In the following we enumerate some of the most important ones. The interested reader is referred to the paper website [1] for discussion on the proofs.

[3]This "selection" property, shared also by the median, would make $\hat{\mu}_0$ very appropriate for image processing applications [9, 2].

[4]See the paper website [1] for a rigurous derivation of this formula.

1. *Shift and scale invariance.* Let $z_i = ax_i + b$, for $i = 1, \cdots, N$. Then, $\hat{\mu}_0(z_1, \cdots, z_N) = a\hat{\mu}_0(x_1, \cdots, x_N) + b$.

2. *No overshoot/undershoot.* The following bounds always hold for $\hat{\mu}_0$:

$$x_{(2)} \leq \hat{\mu}_0 \leq x_{(N-1)}, \tag{13}$$

where $x_{(i)}$ denotes the i^{th} order statistic of the sample. A direct consequence of this property is stated next.

3. *If $N = 3$, $\hat{\mu}_0$ is equivalent to the sample median.*

4. *Consistency.* Under very general regularity conditions on the distribution of the underlying process, $\hat{\mu}_0$ is a consistent estimator of μ_0. When the underlying distribution is symmetric, $\hat{\mu}_0$ converges asymptotically to the center of symmetry.

5. *Optimality in very impulsive environments.* Let $\{f_\alpha\}_{\alpha>0}$ denote a family of symmetric and algebraic-tailed density functions parameterized by the tail constant α. Let $T_\alpha(x_1, \cdots, x_N)$ denote the maximum likelihood estimator associated with f_α. Then,

$$\lim_{\alpha \to 0} T_\alpha(x_1, \cdots, x_N) = \hat{\mu}_0(x_1, \cdots, x_N), \tag{14}$$

independently of the density functions f_α.

The last two properties constitute a solid theoretical argument compelling the use of ZOS in estimation problems that involve very heavy tails. As it should be expected, the proofs are very elaborated and require advanced mathematical tools. The interested reader is referred again to the paper website [1] for details.

A word of caution shall be stated here. Being mode-type, $\hat{\mu}_0$ presents a large breakdown point in the sense that two outliers with exactly the same value may have catastrophic effects in the performance of the estimator. However, for random variables with continuous distributions this does not represent a serious danger. As it is stated by Properties 4 and 5, the possibility of such an event does not preclude the estimator from being consistent and efficient over a wide spectrum of logarithmic-order distributions.

5. Performance of ZOS in α-stable noise

The performance of the ZOS location estimator was comparatively evaluated using Montecarlo simulation. Figure 2 shows the estimated mean absolute errors (MAE) of the sample mean, the sample median and the ZOS estimator when used to locate a size 5 i.i.d. symmetric α-stable sample. The values of α ranged from 2 (Gaussian case) down to 0.3 (very impulsive). Values close to 2 indicate a distribution close to the Gaussian, in which case the sample mean

outperforms both the median and the ZOS estimator. As α is decreased, the noise becomes more impulsive and the sample mean rapidly loses efficiency, being outperformed by the sample median for values of α less than 1.7. More interestingly, as α approaches 1, the estimated MAE of the sample mean explodes. In fact, it can be shown that, for $\alpha < 1$ it is more efficient to use any of the sample values than the sample mean itself, making the estimator totally useless. As α continues to decrease, the sample median loses progressively more efficiency with respect to the ZOS estimator, and at $\alpha \approx 0.87$, the ZOS begins to outperform the sample median. This is an expected result given the optimality of the ZOS estimator for small values of α. For the last value in the plot, $\alpha = 0.3$, the ZOS estimator has an estimated efficiency ten times larger than the median. This increase in relative efficiency is expected to grow without bounds as α approaches 0 (the optimality point of the ZOS estimator). Unfortunately, the variance of the MAE also increases for decreasing values of α, making the evaluations via simulation difficult. More theory must be developed that allows the evaluation and comparison of estimators for small values of α.

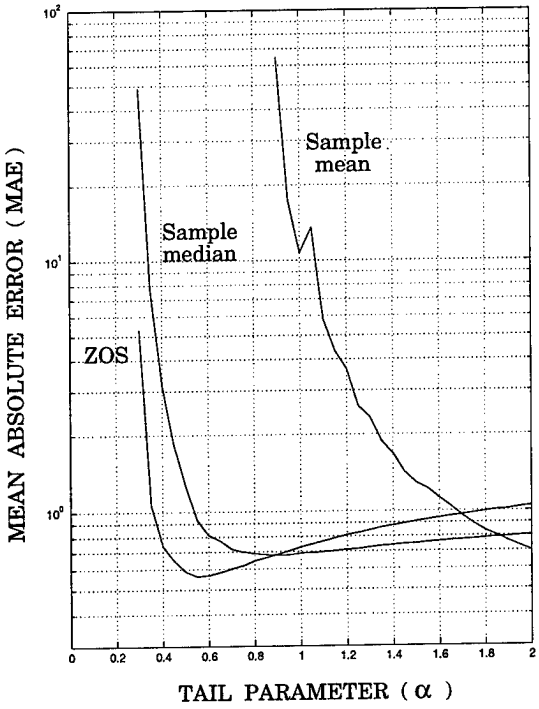


Figure 2. Estimated Mean Absolute Error of the sample mean, sample median and ZOS location estimator in α-stable noise ($N = 5$). The ZOS estimator outperforms both mean and median for small values of the tail constant (approx. $\alpha < 0.87$).

6. Conclusions

We have introduced the preliminaries of a theory of *zero-order statistics* (ZOS) which is sound and consistent for all processes with finite logarithmic moments. This "logarithmic-order" class, includes impulsive-type processes modeled by algebraic tails, and basically embraces *all* the probability models with known practical use. Two new parameters, the geometric power and the zero-order location, constitute the underpinnings of ZOS theory, playing a role similar to the power and the expected value in the context of second-order processes. Properties of the new parameters were reported, and a novel mode-type estimator with optimality properties in very impulsive processes was directly derived from the zero-order location. Given the limitations of second-order and fractional-order theory, ZOS are an attractive alternative in signal processing applications with infinite variance processes.

Further work is to be done in developing a theory of zero-order correlation that successfully attacks estimation problems in very impulsive non i.i.d. (bursty) environments. More work is also necessary in developing estimators of ZOS parameters, establishing a spectral theory, and defining more convenient evaluators of estimator performance in very impulsive noise.

References

[1] J. G. Gonzalez. *Zero-order Statistics: A Signal Processing Framework for Very Impulsive Processes.* Article website, http://www.ee.udel.edu/~gonzalez/PUBS/HOS97a, 1997.

[2] J. G. Gonzalez, D. L. Lau, and G. R. Arce. Towards a general theory of robust nonlinear filtering: Selection filters. *Proc. IEEE ICASSP'97*, Apr. 1996.

[3] H. M. Hall. *A New Model for Impulsive Phenomena: Application to Atmospheric-Noise Communication Channels.* Technical Report. Stanford Electronics Laboratories, Stanford University, Stanford, California, 1966.

[4] J. Ilow. *Signal Processing in Alpha-Stable Noise Environments: Noise Modeling, Detection and Estimation.* Ph.D. Diss. Dep. Electr. Eng. Univ. of Toronto, 1996.

[5] B. Mandelbrot. Long-run linearity, locally gaussian processes, h-spectra and infinite variances. *Interant. Econ. Rev.*, 10:82–111, 1969.

[6] J. H. Miller and J. B. Thomas. Detectors for discrete time signals in nongaussian noise. *IEEE Trans. on Info. Theory*, pages 241–250, Mar. 1972.

[7] C. L. Nikias and A. T. Petropulu. *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework.* Prentice-Hall, New Jersey, 1993.

[8] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications.* John Wiley and Sons, New York, 1995.

[9] P. Zurbach, J. G. Gonzalez, and G. R. Arce. Weighted myriad filters for image processing. *Proc. IEEE ISCAS'96*, May 1996.

Comparison between Asymmetric Generalized Gaussian (AGG) and Symmetric-α-Stable (SαS) Noise Models for Signal Estimation in Non Gaussian Environments

A. Tesei, R. Bozzano, and C.S.Regazzoni
Dept. of Biophysical and Electronic Engineering (DIBE) - University of Genoa - Genoa, ITALY
E-Mail: tale@dibe.unige.it

Abstract

The paper is focused on the problem of multilevel digital signal estimation in the presence of generic noise in a communication system. Noise is assumed unimodal, independent identically distributed, generically non Gaussian, that is eventually asymmetric, impulsive or not. The proposed solution is based on a previously developed estimator which requires the analytical probability density function model of the noise.

The selected estimator was originally applied under the assumption of SαS noise distribution. In this paper, the asymmetric generalized Gaussian (agg) model is selected as a suitable model to describe the noise processes: hence, it is discussed and compared with the SαS distributions in terms of decoding performances.

Tests were performed on simulated binary sequences corrupted by interference generated as SαS processes. Test results outlines comparable performances of the two families of parametric noise models.

1. Introduction

The present work is addressed to *multilevel digital signal estimation* in the presence of *generic noise*.

This problem is relevant in a variety of fields, in particular those involving telecommunication problems (e.g., radar, sonar, etc.), in which many processes cannot be realistically described as Gaussian.

For example, if a communication receiver was built for working under conditions of Gaussian noise and is subjected to non Gaussian (impulsive or not) interference, its performances strongly degrade and may get unacceptable.

Most of previous works proposed solutions under the assumption of *Gaussian or Gaussian-mixture noise*. Other recent works ([1] and related papers) were focused on the estimation of signals corrupted by *impulsive* interference, by first computing the a-posteriori density, and then using Symmetric-α-Stable (SαS) distributions in the estimator.

Actually, the stable model was validated with a variety of real impulsive data, which are very common as generated by natural and man-made sources, and were shown to fit well all of them [2]. Many works investigated on the analytical properties, and good data-fitting capabilities of SαS distributions mainly in detection and estimation problems [2][3]. Other papers were published about the generation of processes with SαS distribution [2][3] and the estimation of parameters of alpha-stable interference [4].

Although a generic eventually asymmetric version of alpha-stable model was formulated [2], signal processing applications were commonly based on the particular case of the symmetric SαS probability density function (pdf).

For taking into account generalized cases, the present approach, starting from the estimator proposed in [1], extends its use to *arbitrary, eventually non-impulsive and/or asymmetric noise*, modeled by a parametric pdf having a very simple analytical closed form.

This model depends on few parameters of the Second and Higher Order Statistics (SOS and HOS) [5], to estimate from data. It is the so-called *asymmetric generalized Gaussian* (*agg*) pdf [6], and derives from the asymmetric Gaussian and the generalized Gaussian [7] pdfs.

Its main characteristics consist in: dependence on statistical parameters up to the fourth orders, being simple to estimate and allowing the model to describe densities with variable skewness and shape (in terms of sharpness and tail heaviness); mathematical similarities with the Gaussian function, hence inheriting many properties of the Gaussian pdf; robustness to accurately fit many real distributions.

In this paper, the selected model is briefly outlined; in particular its analytical properties are discussed and compared with the SαS model capabilities from a theoretical point of view.

Results obtained from numerical simulations by applying the selected estimator under the two noise model assumptions are presented and evaluated in terms of accuracy.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. The problem of signal estimation

The problem of multilevel digital signal estimation is mathematically formalized as follows.

The scalar discrete state-space system [5]:

$$\begin{cases} X_{k+1} = A_k X_k + W_k & k = 1,2,\ldots,N \\ Z_k = H_k X_k + V_k & A_k, H_k \in \Re \end{cases} \quad (1)$$

is considered, where $\{W_k\}$ is the independent identically distributed (iid) noise having pdf $p_W(.)$ and added to the model of interference $\{X_k\}$, and $\{V_k\}$ is the L-level signal ($Pr\{V_k=d_i\}=q_i$, i=1,2,...L, $\sum_{i=1}^{L} q_i = 1 \ \forall k$). The sequences $\{W_k\}$ and $\{V_k\}$ are assumed to be mutually independent. Let $Z^k=\{Z_1, Z_2,\ldots,Z_k\}$ be the set of possible observations up to time k; for any fixed k, the past measurement $z^k=\{z_l: Z_l=z_l, l=1,2,\ldots k\}$ is given. $\hat{X}_k$ is the filtered estimator of the true X_k, $\hat{X}_{k+1}$ its prediction. The objective of the algorithm is finding $\hat{X}_k$ so that the error $|X_k - \hat{X}_k|$ is minimum, thus removing interference from observations and recovering the signal.

2.1. The recursive estimator

Different types of estimators can be designed, once computed the a-posteriori pdf $p_{X_k/Z^k}(x_k / z^k)$. The method proposed by Shen and Nikias in [1] is selected here: it depends explicitly on the analytical model of the noise pdf, hence taking into account eventual non Gaussianity and allowing one to select the best model that fits data.

In this section a brief description of the method is included in order to clarify the employment of the selected parametric models in expressing analytically the pdf $p_W(.)$ of the generic noise $\{W_k\}$.

By applying the Bayesian law, a useful theorem was proved from (1): given the initial a-priori pdf $p_{X_1}(x_1)$, the a-posteriori pdfs are uniquely determined by:

$$p_{X_k/Z^{k-1}}(x_k / z^{k-1}) = \frac{\sum_{i=1}^{L} q_i p_{X_{k-1}/Z^{k-2}}(t_{k-1,i} / z^{k-2}) p_W(x_k - A_{k-1} t_{k-1,i})}{\sum_{i=1}^{L} q_i p_{X_{k-1}/Z^{k-2}}(t_{k-1,i} / z^{k-2})} \quad k=3,4,\ldots \quad (2)$$

$$p_{X_k/Z^k}(x_k / z^k) = \frac{\sum_{i=1}^{L} q_i p_{X_k/Z^{k-1}}(t_{k,i} / z^{k-1}) \delta(x_k - t_{k,i})}{\sum_{i=1}^{L} q_i p_{X_k/Z^{k-1}}(t_{k,i} / z^{k-1})} \quad k=2,3,\ldots \quad (3)$$

given

$$t_{k,i} = \frac{z_k - d_i}{H_k} \quad i=1,\ldots,L \quad k=1,2,\ldots,N \quad (4)$$

and the initial densities:

$$p_{X_1/Z_1}(x_1 / z_1) = \frac{p_{X_1}(x_1) p_{Z_1/X_1}(z_1 / x_1)}{p_{Z_1}(z_1)} \quad (5)$$

$$p_{X_2/Z_1}(x_2 / z_1) = \frac{\sum_{i=1}^{L} q_i p_{X_1}(t_{1,i}) p_W(x_2 - A_1 t_{1,i})}{\sum_{i=1}^{L} q_i p_{X_1}(t_{1,i})} \quad (6)$$

An optimal criterion based on the L^p-distance was selected, for which $\hat{X}_k(Z^k)$ must be chosen so that $\left\| \hat{X}_k(Z^k) - X_k \right\|_p = \min$, $p \in (0,+\infty)$.

In the present work attention is focused on the only $p=1$ case (i.e., the so-called AVC method), as considered less restrictive about error distribution and practically equivalent about performances with respect to the $p=2$ case [1]. The AVC procedure follows.

After densities initializations, for $k=1,2\ldots,N$ repeat:

1. From the data set Z^k and the signal levels $\{d_i\}$ the measurement set $T_k=\{t_{k,i}, i=1,2,\ldots,L\}$ is computed.
2. The numerical weights $p_{X_k/Z^{k-1}}(t_{k,i} / z^{k-1})$ $i=1,2,\ldots,L$ of the L^1-distance can be evaluated from the mentioned theorem.
3. The objective function
$$J(x) = \sum_{i=1}^{L} q_i p_{X_k/Z^{k-1}}(t_{k,i} / z^{k-1}) \left| t_{k,i} - x \right| \quad (7)$$
can be computed for any data point in T_k, where x is the current interference value to estimate. Its optimum value, $\hat{x}_k^{opt}(z^k)$, minimizing $J(x)$, allows one to recover the signal value d_{i_o}.
4. Finally, $p_{X_{k+1}/Z^k}(./z^k)$ can be predicted from (2).

It is clear that the algorithm needs a realistic analytical model of the $\{W_k\}$ distribution.

3. Noise modeling for estimation optimization

A parametric pdf is proposed for modeling the $\{W_k\}$ distribution. It is compared in terms of statistical and mathematical properties, and estimation performances with the SαS distribution [2], whose application in the selected estimator was presented and discussed in [1].

3.1. The asymmetric generalized Gaussian (agg) pdf

A parametric pdf is selected, able to fit in analytically simple and realistic way iid unimodal generically non Gaussian noise, being impulsive or not, symmetric or not.

For describing deviation from symmetry, it was found [7] that the information provided by the 3rd-order parameter skewness is equivalent to that yielded by the combination of the empirical left and right variances:

$$\sigma_l^2 = \frac{1}{N_l - 1}\sum_{k=1, x_k < m_k}^{N_l} (x_k - m_k)^2 \quad (8)$$

$$\sigma_r^2 = \frac{1}{N_r - 1}\sum_{k=1, x_k > m_k}^{N_r} (x_k - m_k)^2 \quad (9)$$

where m_x is the pdf mode, and N_l (N_r) is the number of x_k samples $<m_x$ ($>m_x$). If they are inserted in the generalized Gaussian model, having variable shape in terms of sharpness, the asymmetric generalized Gaussian is derived:

$$p_{agg}(x) = \begin{cases} \frac{c\gamma_a}{\Gamma(1/c)} e^{-\gamma_l^c [-(x-m_x)]^c} & x < m_x \\ \frac{c\gamma_a}{\Gamma(1/c)} e^{-\gamma_r^c [(x-m_x)]^c} & x \geq m_x \end{cases} \quad (10)$$

where:

$$\gamma_l = \frac{1}{\sigma_l}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}, \gamma_r = \frac{1}{\sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}, \quad (11)$$

$$\gamma_a = \frac{1}{\sigma_l + \sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}. \quad (12)$$

An approximated analytical model matching the theoretical parameter c with the kurtosis β_2 was achieved [7]; it was developed in the limit case of $\sigma_l^2 = \sigma_r^2$, but it is proved to be suitable also for any other variance combinations since the MSE value between the c-based and the β_2- based pdfs is always of the order of 10^{-7}.

The kurtosis-based *agg* pdf is defined for $\beta_2 > 1.865$. It is continuous, and derivable everywhere. It was proved [8] to be approximately stable in the case of $\beta_2 \geq 3$, i.e. when it is heavy-tailed and refers to impulsive variables.

3.2. Comparison with the SαS pdf model

A comparison with the α-stable distributions [4] can be useful, as they were selected in [1] for validating the estimator employed here under impulsive-noise conditions.

The SαS pdf $p_a(\gamma,\delta,x)$ of the stochastic variable x can be defined by the inverse Fourier transform integral:

$$p_a(\gamma,\delta,x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{j\delta\omega - \gamma|\omega|^\alpha} e^{-j\omega x} d\omega \quad (13)$$

and is completely determined by fixing three parameters, the characteristic exponent α ($0<\alpha\leq 2$), the dispersion γ ($\gamma>0$), and the location parameter δ ($\delta\in\Re$). The corresponding characteristic function is given by the exponential:

$$F_a(\gamma,\delta,x) = e^{j\delta\omega - \gamma|\omega|^\alpha} \quad (14)$$

The parameter α is responsible of the pdf capability of varying the heaviness of the pdf tails; the smaller its value the heavier the tails. In this sense it is, at least intuitively, related to kurtosis (and to c parameter of the generalized Gaussian pdf). For particular values of α the family reduces to some limit cases: for $\alpha=2$ a Gaussian function, for $\alpha=1$ a Cauchy model are obtained. For other α values the model does not present closed-form pdf, which can be a significant limitation in its analytical employment, although it can be numerically computed as suggested in [4].

The dispersion γ is related to the spread of the SαS model which is the same information described by variance in the Gaussian and generalized Gaussian pdfs.

The location parameter δ is the pdf point of symmetry, hence is analogous to the mean value.

Also the SαS pdfs present similarities with the Gaussian model, first of all, the *stability* property, which is maintained for any non Gaussian SαS pdf [4].

However, only the moment of the related variable having order $p<\alpha$ are finite (conversely, all the Gaussian moments are finite). This property is rigorously true only in theory: in most of practical applications, in which a finite, whatever high, number of data samples with finite, whatever big, values, is available, moments can be estimated also in the case of impulsive stochastic variables.

The *agg* pdf class extends SαS capabilities by taking into account also non impulsive (sub-Gaussian) pdfs, although a lower kurtosis bound is imposed.

Both the SαS distribution and the *agg* pdf can be easily extended to multivariate cases. Their common main limitations consist in unimodality and needing of estimating their parameters on the basis of long series of data.

4. Simulation results and discussion

In the simulation tests, let's assume the noise density as *agg* and compare performances with the SαS assumption proposed in [1]. Comparison is carried out on the data corrupted by simulated impulsive noise.

A single random double-level signal $\{V_k\}$ ($P_{-1}=P_{+1}=1/2$) of length N (N=1000) was generated by using standard uniform distribution routines. The experimental results were obtained by taking this sequence as a common reference for all the runs.

Several noise sequences $\{W_k\}$ of the same length N were generated as SαS processes [3]: the characteristic exponent α varies within the range [0.4-2.0], the dispersion γ goes from 0.2 up to 5.0, while the location parameter is fixed to 0.

Figure 1 plots one of the generated SαS noise

sequences given α=1.5, γ=2.0, δ=0. The estimated *agg* pdf (μ_x=0.02, m_x=-0.13, σ_l^2=25.2, σ_r^2=37.4, β_2=68.1), superimposed on the related histogram, shows a good curve fitting (see Fig. 2).

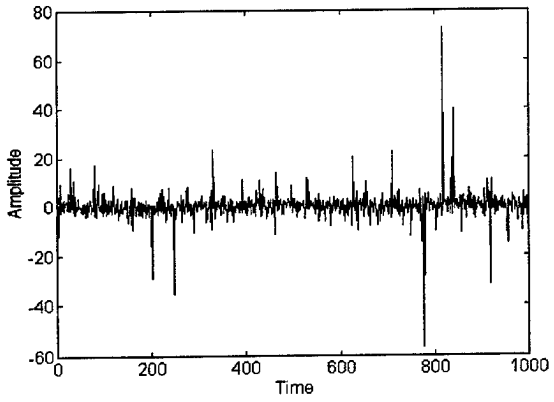


Figure 1. One realization of SαS noise (α=1.5,γ=2.0,δ=0).

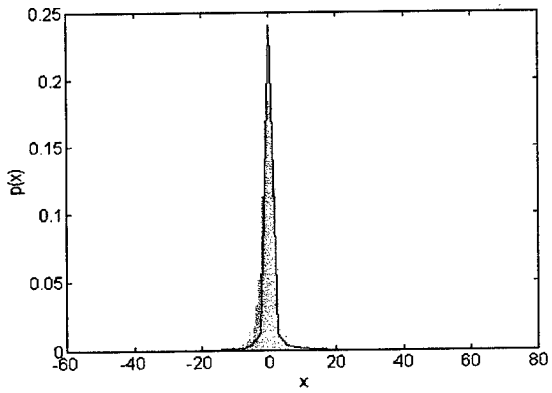


Figure 2. Fitting of the *agg* pdf with the data histogram.

In order to apply the *agg* pdf model, it is necessary to estimated its characteristic parameters from each noise sequence. This operation is performed by simply using the parameter operative definitions.

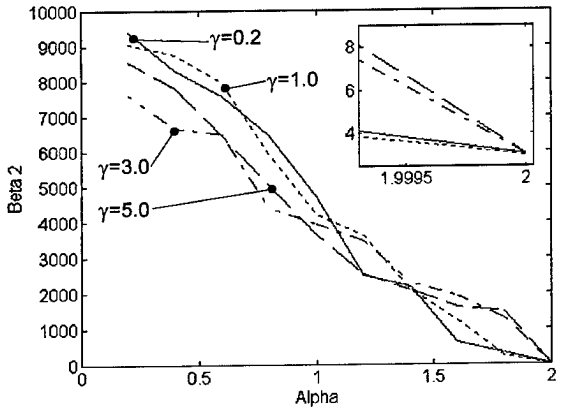


Figure 3. β_2 estimated in terms of the α parameter from SαS sequences for some γ fixed values.

Figure 3 presents the statistical β_2 trends as α varies for some γ values, obtained by averaging the estimates on 50 sequences each composed by 10000 samples.

The system (1) was applied for obtaining the interference and the measure data, with A_k=0.7 and H_k=1 fixed for each run and for each discrete time instant k. The same number (M=*10*) of independent runs for each combination of the parameters α and γ were performed and the obtained results averaged.

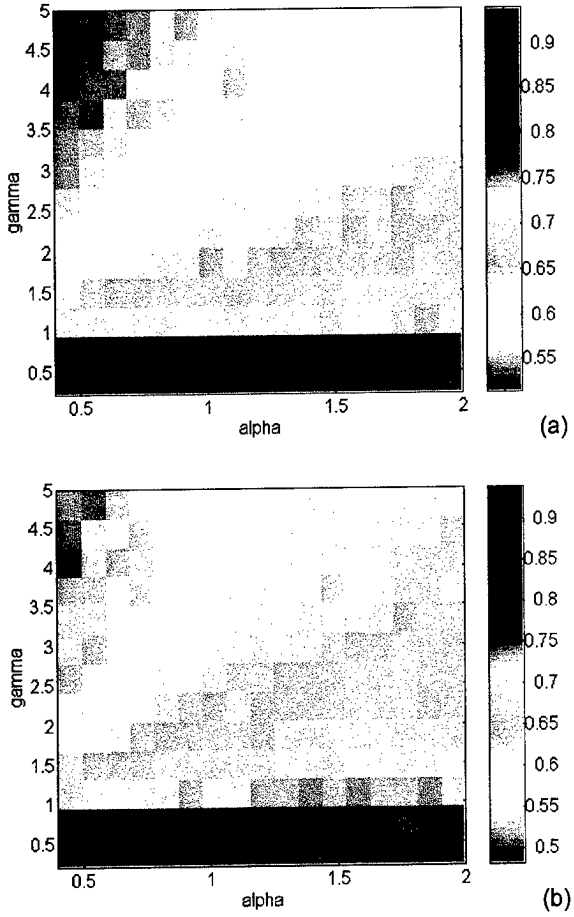


Figure 4. (a) SαS and (b) *agg* decoding performances for combination of α and γ.

Figure 4 shows the correct-decoding probability [0-1] as obtained by modeling noise as an SαS (Fig. 4a) and an *agg* (Fig. 4b) variable. Results are shown for combinations of the parameters α and β_2.

Some examples are extracted from Fig. 4 when γ is fixed (Fig. 5a) and when α is fixed (Fig. 5b).

Performances are comparable and prove that both approaches are robust and accurate as α varies even in the case of very impulsive noise.

Hence, the *agg* model is shown to be adequate for representing the distribution of a variety of stochastic processes.

Notice that the comparison is performed under the ideal condition of exact knowledge of the SαS parameters, while in real applications they should be estimated as the *agg* parameters are

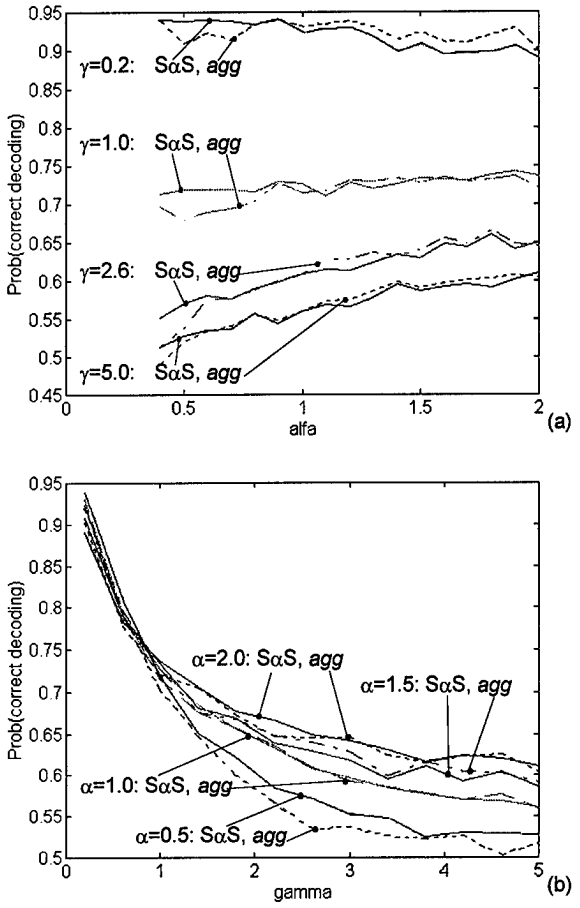


Figure 5. Comparison of the correct decoding probability for different fixed values of (a) γ and (b) α.

5. Conclusions and future perspectives

The paper has addressed the problem of multilevel digital signal estimation in the presence of generic noise. Noise has been assumed unimodal, iid, generically non Gaussian, that is eventually asymmetric, impulsive or not. The proposed solution consists in using a known estimator requiring the analytical pdf model of interference as a parameter. At this aim the asymmetric generalized Gaussian pdf has been selected, discussed and compared with the SαS distributions, which present appreciable analytical characteristics and fit well symmetric impulsive noise. The selected estimator was originally applied under the assumption of SαS noise distribution, hence this model provides a valuable reference for validating *agg* pdf adequacy in solving signal estimation problems.

Tests have outlined comparable performances of the two families of parametric models; anyway the main *agg* advantage yields in its capability of representing and well-fitting a wide set of possible real distributions by taking into account not only variable tail heaviness (including sub-Gaussianity, disregarded by the SαS family), but also eventual deviation from symmetry.

Future activities concern the extension of the selected model to multivariate cases, which is feasible thanks to its analytical similarities with the Gaussian function, and consequent application to multidimensional signals, e.g. images.

References

[1] J. Shen, and C. L. Nikias, "Estimation of multilevel digital signals in the presence of arbitrary impulsive interference," IEEE Trans. on Signal Processing, Vol. 43 (1), pp. 196-203, January 1995.

[2] C.L. Nikias, M. Shao, *Signal Processing with Alpha-stable distributions and applications*, J. Wiley and Sons, 1995.

[3] G.A. Tsihrintzis, and C.L. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an Alpha-Stable process," IEEE Trans. on Communications, Vol. 42 (3), pp. 904-914, February/March/April 1996.

[4] G.A. Tsihrintzis, and C.L. Nikias, "Fast estimation of the parameters of Alpha-Stable impulsive interference," IEEE Trans. on Signal Processing, Vol. 44 (6), pp. 1492-1503, June 1996.

[5] C.L. Nikias, and J. Mendel, "Signal processing with Higher-Order Spectra," IEEE SP Magazine, pp. 10-37, July 1993.

[6] A. Tesei, and C.S. Regazzoni, "Application to locally optimum detection of a new noise model," Proc. IEEE ICASSP'96, Vol. 5, pp. 2467-2470, 1996.

[7] A. Tesei, and C.S. Regazzoni, "HOS-based noise models for signal-detection optimization in non-Gaussian environments," Proc. IEEE ISIT'95, p. 296, 1995.

[8] A. Tesei, C.S. Regazzoni, and G. Tacconi, "HOS-based modeling of low frequency underwater acoustic noise for signal detection in shallow waters," Proc. IEEE OCEANS '96, pp. 1421-1426, 1996.

No Evidence of Stable Distributions in Radar Clutter *

Jacek Ilow
Technical University of Nova Scotia,
Department of Electrical Engineering
Halifax, NS, B3J 2X4, Canada
e-mail: ilow@tuns.ca

Henry Leung
Defence Research Establishment Ottawa
Surface Radar Section
Ottawa, Ont., K1A 0Z4, Canada
Henry.Leung@dreo.dnd.ca

Abstract

The alpha-stable distribution is a theoretical model for impulsive noise that currently enjoys wide success. In this paper, we test its applicability to high resolution radars that are capable of resolving fine structure of the sea surface. The received sea clutter signal by such systems is not well modeled by a Gaussian process, and we expected that stable distributions may provide better description of noise statistics than the conventional non-Gaussian models such as the K-distribution. However, in the important for radar low probability of false alarm region, we found that the K-distribution fits better the sea-clutter amplitude statistics than the alpha-stable distribution. In the application considered, we explain this failure of stable model based on the analytical stable noise modeling.

1 Problem Formulation

Although stable distribution models are recently often used in statistical signal processing, their histogram-fitting and modeling flexibility in most cases are only analyzed theoretically [8], [10]. In this paper, we fill this gap in an attempt to fit stable distributions to radar clutter data. We compare the fit of the stable model to the real world data with that of the K-distribution model. To achieve this, first, we present a computationally efficient method to evaluate the cumulative distribution function (cdf) of stable random variables (RVs) based on the Fourier-Bessel expansion. Then, we employ different sets of scanning radar data for model validation. The data sets used are from OHGR'93 IPIX database of DREO. The detailed description of the IPIX radar and the database can be found in [9]. In general, we had available the sequences of in-phase (I) and quadrature (Q) receiver outputs, which formed the stationary realization of the bivariate, circularly symmetric (CS), random variable (RV). The objective was to compare the fit of the K-distribution and the fit of the envelope of CS alpha stable RV to the amplitude statistics of the observed clutter.

Our conclusion from the data analysis conducted is that, in the low probability of false alarm region, the K-distribution fits sea-clutter amplitude statistics better than the stable distribution. The theoretical explanation for that is in the physics of the clutter generating process. Specifically, the spatio-temporal distribution of scattering centers contributing to noise corresponds better to the analytical K-distributed noise modeling than than to the stable noise modeling introduced in [4].

2 Preliminaries of Alpha-stable Distributions

The univariate symmetric α-stable ($S\alpha S$) RV with "zero-mean" is defined based on its characteristic function (cf) [8]:

$$\phi_\alpha(\omega) = \exp(-\gamma|\omega|^\alpha), \tag{1}$$

where α is the characteristic exponent, and dispersion γ is a quantity analogous to the variance. The characteristic exponent controls the heaviness of the probability density function (pdf) tails: a small positive value of α indicates severe impulsiveness, while a value close to 2 indicates a more Gaussian type of behavior. In this paper, we are primarily concerned with the bivariate circularly symmetric (CS) $\alpha-$stable random vectors for which the cf is given as [8]

$$\psi_\alpha(\omega_1,\omega_2) = \exp(-\gamma|\omega_1^2+\omega_2^2|^{\frac{1}{2}\alpha}) = \phi_\alpha(-\gamma||\omega||^\alpha), \tag{2}$$

where $||\omega||$ is the L_2 norm of ω.

*This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

0-8186-8005-9/97 $10.00 © 1997 IEEE

The estimation of parameters for CS α-stable distributions is described in [7].

A circularly symmetric distribution is the two dimensional (2-D) case of a wider class of spherically symmetric distributions.

Definition 1 *An $n \times 1$ random vector $\mathbf{X}$ in $\mathbb{R}^n$ is said to have a spherically symmetric (SS) distribution if for every $\mathbf{\Gamma} \in \mathcal{O}(n)$, $\mathbf{\Gamma X}$ has the same distribution as $\mathbf{X}$, where $\mathcal{O}(n)$ denotes the set of $n \times n$ orthogonal matrices:*

$$\forall_{\mathbf{\Gamma} \in \mathcal{O}(n)} \qquad \mathbf{\Gamma X} \stackrel{d}{=} \mathbf{X}. \tag{3}$$

Here $\stackrel{d}{=}$ means that the two sides have the same distribution.

From **Definition 1**, follows an important property needed to discuss further SS RVs [5]:

Proposition 1 *An $n \times 1$ RV $\mathbf{X}$ has a spherical distribution if and only if its characteristic function $\phi(\mathbf{t})$ satisfies one of the following equivalent conditions:*

1. *$\phi(\mathbf{\Gamma t}) = \phi(\mathbf{t})$ for any $\mathbf{\Gamma} \in \mathcal{O}(n)$;*
2. *There exist a function $\psi(\cdot)$ of a scalar variable such that $\phi(\mathbf{t}) = \psi(\sqrt{\mathbf{t}^T\mathbf{t}}) = \psi(||\mathbf{t}||)$, where $||\mathbf{t}||$ is the L_2 norm of $\mathbf{t}$.*

We write $\mathbf{X} \sim S_n(\psi)$ to mean that $\mathbf{X}$ has a cf of the form $\psi(||\mathbf{t}||)$, where $\psi(\cdot)$ is a function of a scalar variable, called the characteristic generator of the spherical distribution. For SS stable distributions, the characteristic generator is of the form: $\psi(x) = \exp(-\gamma|x|^\alpha)$.

The main problem, when we start to work with α-stable distributions, is that no closed-form expressions exist for pdfs and cdfs of α-stable RVs, except for the Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1$) and Lévy ($\alpha = \frac{1}{2}$) distributions. There are convergent (valid for $x << 1$) or asymptotic (valid for $x >> 1$) power series representations for the cdf of stable distributions [10], however they are difficult to handle in numerical calculations because of the oscillating character of coefficients in the series. In the following section, we demonstrate how to use the Fourier-Bessel series expansion for the efficient calculation of the cdf for the envelope of bivariate CS α-stable RV.

3 The Fourier-Bessel Series in Calculating PDF

In general, the pdf of $n \times 1$ a random vector $\mathbf{X}$ is related to the characteristic function $\phi(\mathbf{t})$ by the n-D inverse Fourier transform

$$p_{\mathbf{X}}(\mathbf{x}) = p(||\mathbf{x}||) = (2\pi)^{-n} \int \exp(j\mathbf{x}^T\mathbf{t})\phi(\mathbf{t})d\mathbf{t}, \tag{4}$$

where the integration is over the whole $\mathbf{t}$ space. The change of co-ordinates in (4) from Cartesian to polar for SS distributions results in the following relation [6]:

$$p_{\mathbf{X}}(\mathbf{x}) = p(r) = \frac{r^{1-n/2}}{(2\pi)^{n/2}} \int_0^\infty \lambda^{n/2}\psi(\lambda)\mathcal{J}_{(n-2)/2}(\lambda r)d\lambda, \tag{5}$$

where $r = ||\mathbf{x}|| = \sqrt{\sum_{i=1}^n x_i^2}$, and $\mathcal{J}_k(\cdot)$ represents the ordinary Bessel function of the first kind of order k. The relation (5) is usually referred to as the Hankel transform. In the same fashion, it can be shown that [6]

$$\phi(\mathbf{t}) = \psi(r) = \frac{(2\pi)^{n/2}}{r^{n/2-1}} \int_0^\infty \lambda^{n/2}p(\lambda)\mathcal{J}_{(n-2)/2}(\lambda r)d\lambda. \tag{6}$$

In applications considered in this paper, we are more interested in the pdf $P(r)$ and cdf $F(r)$ of $||\mathbf{X}||$. One of the consequences of **Proposition 1** is that the pdf P(r) of $||\mathbf{X}||$ is related to the pdf of $\mathbf{X}$ as

$$P(r) = \frac{2\pi^{n/2}r^{n-1}}{\Gamma(n/2)}p(r). \tag{7}$$

Therefore, from (5) and (6), we get

$$P(r) = \frac{2^{-n/2+1}r^{n/2}}{\Gamma(n/2)} \int_0^\infty \lambda^{n/2}\psi(\lambda)\mathcal{J}_{(n-2)/2}(\lambda r)d\lambda \tag{8}$$

$$\psi(r) = \frac{\Gamma(n/2)}{(\frac{r}{2})^{n/2-1}} \int_0^\infty \lambda^{1-n/2}P(\lambda)\mathcal{J}_{(n-2)/2}(\lambda r)d\lambda. \tag{9}$$

If $\psi(\cdot)$ is known, we will calculate first $P(r)$ as in (8) using a Fourier-Bessel series approximation. Assuming that $P(r) = 0$ for $r > \mathsf{Y}$, (8) can be written as [2]:

$$P(r) = \frac{2^{-n/2+2}r^{n/2}}{\mathsf{Y}^{\frac{n}{2}+1}\Gamma(n/2)} \sum_{k=1}^\infty \frac{\zeta_k^{n/2-1}F_k}{\mathcal{J}_{\frac{n}{2}}^2(\zeta_k)}\mathcal{J}_{\frac{n}{2}-1}(\zeta_k\frac{r}{\mathsf{Y}}), \tag{10}$$

where ζ_k is the kth real, positive, zero of $\mathcal{J}_{\frac{n}{2}-1}(\cdot)$, and F_k is the Fourier-Bessel coefficient

$$F_k \triangleq \frac{\Gamma(n/2)}{(\frac{\zeta_k}{2\mathsf{Y}})^{n/2-1}} \int_0^{\mathsf{Y}} \lambda^{1-n/2}P(\lambda)\mathcal{J}_{(n-2)/2}(\frac{\zeta_k}{\mathsf{Y}}\lambda)d\lambda. \tag{11}$$

The Fourier-Bessel coefficient is almost the same as the characteristic function in (9) evaluated at $\frac{\zeta_k}{\mathsf{Y}}$, so we may try to replace F_k with $\psi(\frac{\zeta_k}{\mathsf{Y}})$. Now, integrating in (10) term-by-term, the cdf of $||\mathbf{X}||$ is

$$F(r) = \frac{2^{-n/2+2}r^{n/2}}{\mathsf{Y}^{\frac{n}{2}}\Gamma(n/2)} \sum_{k=1}^\infty \frac{\zeta_k^{n/2-2}\psi(\frac{\zeta_k}{\mathsf{Y}})}{\mathcal{J}_{\frac{n}{2}}^2(\zeta_k)}\mathcal{J}_{\frac{n}{2}}(\zeta_k\frac{r}{\mathsf{Y}}). \tag{12}$$

In numerical calculations, the series in (12) is truncated at some term L. Another source of error is the replacement of F_k with $\psi(\zeta_k/\Upsilon)$ which corresponds to aliasing in calculating FFT. The detailed analysis of truncation errors and coefficient approximations can be found in [2]. The dimension we are interested in this paper is: $n = 2$. In this case, ζ_k are not available in closed-form, and we use approximations from [1]. For $k = 1, \cdots, 20$, ζ_k are calculated in [1], and the rest is well approximated by the series $\zeta_k \sim (k\pi - \frac{1}{4}\pi) + \frac{1}{8(n\pi - \pi/4)} + \cdots$.

4 Data Analysis

The cdf of the K-distribution is $F(\rho) = 1 - \frac{2}{\Gamma(\nu)}(\frac{b\rho}{2})^{\nu} K_{\nu}(b\rho)$ where $K_{\nu}(\cdot)$ is the modified Bessel function of the third kind of order ν. The model parameters for K-distributions are computed from $< x^2 >= (\frac{2}{b})^2 \nu$ and $< x^4 >= 2(1 + \frac{1}{\nu}) < x^2 >^2$, the second and forth moment, respectively [3].

The data used in this study are from the dual polarized X-band instrumentation-quality radar operating at 9.4 GHz. Quadrature demodulators in receiver channels provide in-phase (I) and quadrature (Q) outputs for further signal processing. Each component is represent using 8 bits. The radar pulse-width is 200ns, which results in the cell resolution of 15m. The sampling period was $T_s = 4ms$. The data sets length for analysis was 10,000 samples from one range-bin. The data were collected at Cape Bonavista, Canada, and the radar was located on a cliff edge overlooking the Atlantic ocean at the height of 22 m above the sea level. The environmental conditions were chosen as to simulate shipboard radars where the grazing angle is low.

In Fig.1(a), we present the CDF of a vertically (VV) polarized noise (or rather its amplitude) based on the (i) empirical CDF, (ii) K-distributed model and (iii) alpha-stable model. For the particular data set presented it is hard to determine which model provides a better fit based on the visual evaluation. Note however, that from radar perspective, standard statistical tests such as the chi-square or Kolgomorov-Smirnov goodness-of-fit tests are of limited use for clutter data [3]. This is because in radar, goodness of fit is mainly important in the low region of probability of false alarm (P_{fa}) i.e., $1 - cdf(\cdot)$. When we go to the tails of the amplitude distribution as in Fig.1(b), where we show $\log(1 - cdf(\cdot))$ for the same data set, it is still hard to resolve the question of a better fit. However, this data set is not characteristic to the database available. Figure 2 compares the $\log(1 - \text{cdf})$ for the representative data sets from the OHGR database for HH and VV polarized clutter using the empirical cdf and those calculated for two models considered. From this figure, it is evident that in the low regions of P_{fa}, the K-distribution provides a better fit to the data than the stable-distribution which has tails much to heavy.

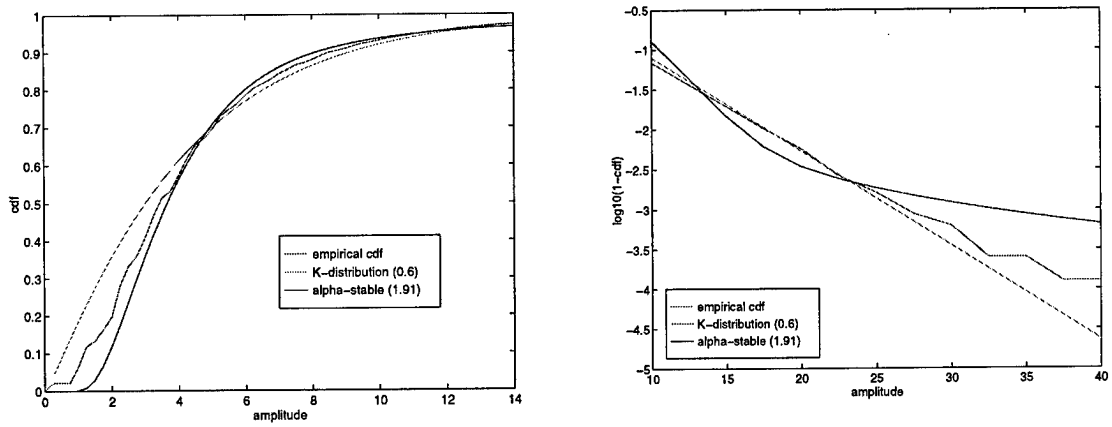
The possible source of erroneous analysis could include: (i) incorrect parameter estimation and (ii) problems with calculation of the cdf for stable distributions. However, before applying our tests on real data sets, we tested our procedures on synthetic data, and the results obtained corroborated our observations for radar data.

Discussion

When a radar illuminates a large patch of the sea, it is usually found that the pdf of the envelope of the backscattered echo is well approximated by a Rayleigh distribution. This is a consequence of the Central Limit Theorem (CLT) because the signal can be thought as the summation of a "infinite" number of random components from independent scattering centers. However, for a narrow beam and short pulse-length radar, the conditions (such as the large number of scatterers) for the CLT to hold are not met, and the Gaussian statistics are no longer appropriate. This was the reason for introducing the K-distribution into clutter modeling. This model can account for the fluctuating but finite number of scatterers in the small area of sea surface, with the average number of contributing scatterers proportional to the area of the sea surface illuminated by the radar. In analytical stable modeling, one can account for the finite number of interferers, as well as propagation conditions, however one cannot ignore scattering centers close to the receiver. In high resolution radar, close to the receiver scattering centers do not contribute to the interference, and the backscattering effect does not comply with analytical stable modeling [4] (the size of the patch illuminated is much smaller than the distance of this patch from the radar).

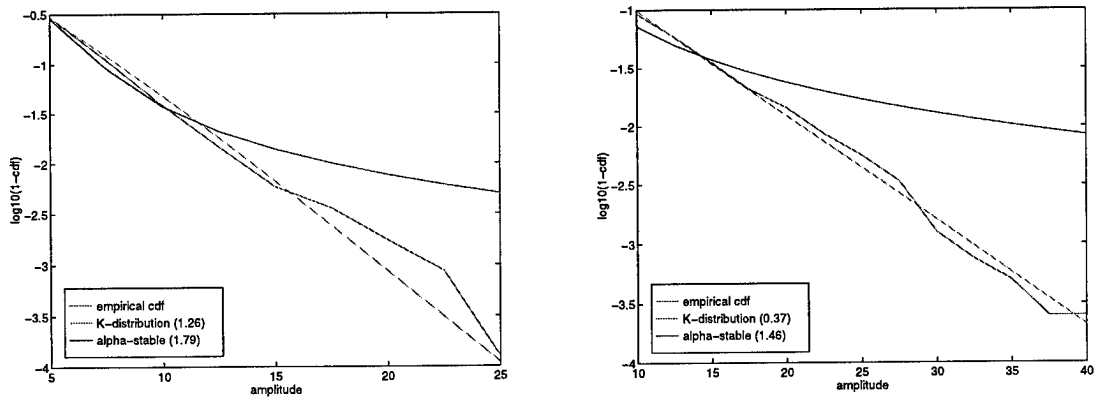
5 Conclusion

In this paper, we compared the fit of the K-distribution and the alpha-stable distribution to the real-world radar data by performing delicate tests on the distributions tail characteristics. Our general observation is that for most of the data sets available, the K-distribution provides better fit to the experimental data than the alpha-stable distribution. We provide possible theoretical explanation for this behavior. As an additional result, we presented a computationally efficient method to calculated the cdf of stable distributions.



a) at the origin b) at the tail

Figure 1 CDF of one data set from the OHGR database for the VV polarized clutter.



a) Tail of CDF for VV polarized clutter b) Tail of CDF for HH polarized clutter

Figure 2 Representative tails of CDF for the OHGR database clutter.

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions.* National Bureau of Standards, Washington, 1972.

[2] J. Bird. Error performance of binary NCFSK in the presence of multiple tone interference and system noise. *IEEE Trans. Commun.*, COM 33, No.3:203–209, March 1985.

[3] H. Chan. Radar sea-clutter at low grazing angles. *IEE Proceedings, Pt.F*, Vol.137, No.2:102–112, April 1990.

[4] J. Ilow. *Signal Processing in Alpha-Stable Noise Environments: Noise Modeling, Detection and Estimation.* Ph.D. thesis, University of Toronto, URL: http://www.comm.toronto.edu/ ~ilow/ilow.html, Toronto, 1996.

[5] K. Fang, S. Kotz and Ng Kai. *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, 1990.

[6] R. Lord. The use of the Hankel transform in statistics. *Biometrica*, 41:44–55, 1954.

[7] X. Ma and C. Nikias. Parameter estimation and blind channel identification in impulsive signal environments. *IEEE Trans. Signal Processing*, 43, No.12:2884–2897, Dec. 1995.

[8] C. Nikias and M.Shao. *Signal Processing with Alpha-Stable Distributions and Applications.* John Wiley & Sons, Inc., New York, 1995.

[9] T. Nohara and S. Haykin. Growler detection in sea clutter using Gaussin spectrum models. *IEE Proceedings, Pt.F*, Vol.141, No.5:285–291, Oct. 1994.

[10] G. Tsihrintzis and C. Nikias. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. *IEEE Trans. Commun.*, 43, No.2/3/4:904–914, March 1995.

On the Modeling of Network Traffic and Fast Simulation of Rare Events using α-Stable Self-Similar Processes

Anestis Karasaridis and Dimitrios Hatzinakos
University of Toronto
Dept. of Electrical and Computer Engineering
Toronto, Ont. M5S 3G4, Canada
email: {anestis,dimitris}@comm.toronto.edu

Abstract

We present a new model for aggregated Network traffic based on α-Stable Self-Similar processes which captures the burstiness and the Long Range Dependence of the data. We show how the Fractional Gaussian noise assumption fails and why our proposed model fits well by comparing real and synthesized network traffic. In addition, we show that we can speed up the simulation times for estimation of rare event probabilities, such as cell losses in ATM switches, by up to three orders of magnitude using α-Stable modeling and Importance Sampling.

1 Introduction

In the area of Telecommunication networks, recent high-quality, high-resolution extensive measurements of Local Area Network traffic [7], Variable-Bit-Rate compressed video streams[2], Wide Area Network traffic [12] and Web client-server traffic [3], have shown that the aggregated traffic in networks does not exhibit Poisson characteristics but rather it follows similar statistics for a wide range of time-scales. The assumption that in general the traffic is a Self-Similar process (SSP) can explain the existence of Long Range Dependence (LRD) (i.e. slowly decaying correlation) and the persistent burstiness regardless of the time-scale.

One of the main implications of the above observations in Network engineering is that the aggregation of traffic in statistical multiplexers and ATM switches does not smooth out the overall traffic if the individual sources are bursty. Similarly, deterministic traffic smoothers have objectionable value. Also, according to new results in buffer occupancy distributions for Self-Similar inputs, the buffer cell loss probability decreases with the buffer size only algebraically in contrast to Markovian models where the decrease is exponential. The average cell delay always increases with the buffer size whereas in Markovian models it does not exceed a certain limit regardless of the buffer size [8],[10].

Since the main properties of the high speed network traffic are high variability and self-similarity, two independent modeling approaches emerged among others. The first is the use of heavy-tailed distributions (e.g. Pareto) to account for the high-variability [7], and the other is the use of Self-Similar processes (e.g. Fractional Gaussian noise) to account for the statistical self-similarity of the data [11]. While these approaches give better results than the simple Poisson or Compound Poisson models, they fail to unify the desired model properties.

On the other hand, α-Stable SSP's can capture both the high variability (burstiness), since the underlying distribution is heavy-tailed, and the self-similarity. Furthermore, they provide a physical interpretation on how the observed data appear as the superposition of independent effects according to the Generalized Central Limit Theorem.

This paper addresses two problems: **a)** Network Traffic modeling using α-Stable Self-Similar models and **b)** Fast simulation of Rare Events based on α-Stable modeling. A new simple and parsimonious model is proposed that provides better connection between the model parameters and the physical generation of Self-Similar processes. The validity of the proposed model is tested by fitting real network traffic data obtained from the ftp site of Bellcore. It is also shown that the proposed model is a powerful tool in deriving new fast simulation methods of rare events (such as cell losses in ATM switches) in network traffic.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2 Network traffic modeling

As a measure of network traffic we use the packet counts (number of packets arrived/time unit) or the interarrival times of the packets. Real data were obtained from Bellcore's ftp site (ftp.bellcore.com) and the modeling was applied to external traffic data (file OctExt.TL) which give the time-stamps and length of Ethernet packets through the main gateway. The external data show more variability than the internal and to the authors' knowledge have not yet been modeled successfully.

Our proposed model is based on the well balanced *Linear Fractional Stable Noise* (LFSN) which is defined as follows [13]:

$$L_{\mathbf{S},H}(i) = \int_{-\infty}^{\infty} (|i+1-x|^{H-1/\alpha} - |i-x|^{H-1/\alpha}) M_{\mathbf{S}}(dx), \quad (1)$$

where $M_{\mathbf{S}}(dx)$ is an α-Stable random measure with Lebesgue control measure, H is the *Hurst* or *self-similarity parameter*, α is the *characteristic exponent* of the α-Stable distribution and $\mathbf{S}$ is the vector of the α-Stable distribution parameters $(\alpha, \sigma, \beta, \mu)$. It is a stationary process with LRD if $H > 1/\alpha$ and it is a generalization of the Fractional Gaussian noise which is obtained by replacing α with 2. The above integral (1) in discrete form can be approximated by the convolution of the linear filter $h_d(i) = |i+1|^d - |i|^d, d = H - 1/\alpha$, $i \in \mathcal{Z}$ with iid α-Stable random variables denoted by $S_{\alpha,\sigma,\beta,\mu}(i)$:

$$L_{\alpha,\sigma,\beta,\mu,H}(i) = h_d(i) * S_{\alpha,\sigma,\beta,\mu}(i). \quad (2)$$

The proposed α-Stable Self-Similar model is linear and has the following expression:

$$\begin{aligned} M(i) &= c_1 \cdot L_{\alpha,H}(i) + c_2 \quad (3) \\ &= c_1 \cdot h_d(i) * S_{\alpha,1,1,0}(i) + c_2, \quad c_1, c_2 \in \mathcal{R}^+, \end{aligned}$$

where c_1 and c_2 are positive real constants and $L_{\alpha,H}(i)$ is Linear Fractional Stable Noise (LFSN) with $\beta = 1, \sigma = 1$ and $\mu = 0$ and $H > 1/\alpha$ to ensure long range dependence.

Since $\beta = 1$ the LFSN process is *totally skewed*. This does not imply that the density function has support only on the positive X axis for all α's. It is strictly positive only for $\alpha < 1$ but this condition is very restrictive for our modeling since we impose the inequality $\alpha > 1/H$, where $0 < H \leq 1$. Also the condition that α is greater than 1, ensures that the mean of the LFSN exists, according to the properties of α-Stable distributions.

The expected value of the model is

$$\mathbf{E}[M(i)] = c_1 \cdot \sum_k h(k) \cdot \mathbf{E}[S_{\alpha,1,1,0}(i-k)] + c_2 = c_2, \quad (4)$$

since $\mathbf{E}[S_{\alpha,1,1,0}(i)] = \mu = 0$. The positive constant c_1 is equivalent to the scaling parameter σ of the α-Stable distribution S. In other words the model in (4) is equivalent to the following: $M(i) = h_d(i) * S_{\alpha,c_1,1,0}(i) + c_2$.

The Characteristic Function (CF) of $S_{\alpha,1,1,0}$, $\alpha \neq 1$ is:

$$\Phi_S(\omega) = \exp\left[-|\omega|^\alpha (1 - j\mathrm{sgn}(\omega)\tan(\pi\alpha/2))\right], \quad (5)$$

and from (4), LFSN's CF becomes

$$\ln \Phi_L(\omega) = -|\omega|^\alpha \cdot \sum_i |h_d(i)|^\alpha (1 - j\mathrm{sgn}(\omega)\tan(\pi\alpha/2)). \quad (6)$$

As a result the CF of the model process M is:

$$\begin{aligned} \ln \Phi_M(\omega) \quad = \quad & -|c_1 \cdot \omega|^\alpha \cdot \sum_i |h_d(i)|^\alpha \cdot \\ & (1 - j\mathrm{sgn}(\omega)\tan(\pi\alpha/2)) + jc_2\omega. \quad (7) \end{aligned}$$

2.1 Model parameter estimation

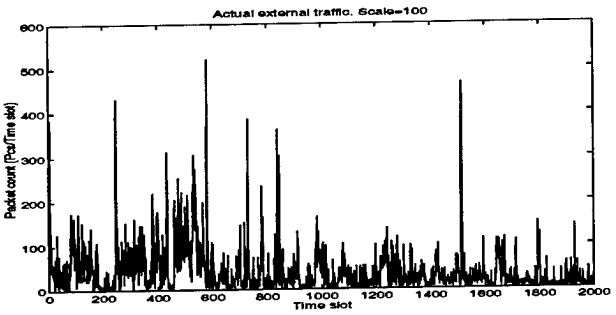
The suggested parametric model is quite parsimonious since it depends on a set of only four parameters: (H, α, c_1, c_2). The self-similarity parameter H is estimated first, by using the R/S statistic [9],[1]. Estimates of the R/S statistic were calculated for both the packet-count and the interarrival-time processes and show that they are long range dependent since $H > 0.5$ in both cases ($H = 0.85$ for the packet counts and $H = 0.78$ for the interarrival times).

Parameter c_2 is calculated as the mean of the modeled process as we saw in (4). In practice c_2 has to be adjusted slightly since the generator of LFSN according to (4) might give a small non zero average.

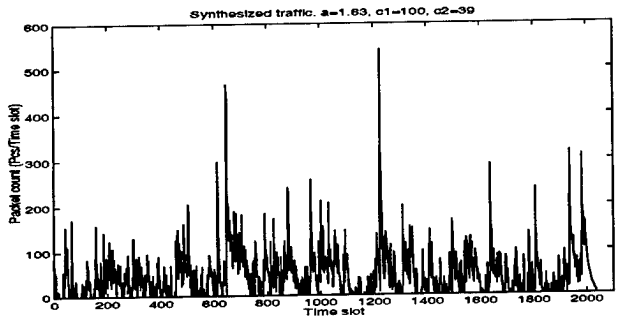
The other parameters α and c_1 can be calculated by minimizing the mean absolute error between the model and the real data:

$$\min_{1/H < \alpha \leq 2, c_1 > 0} \mathbf{E}|\mathbf{X} - c_1 \cdot \mathbf{L}_{a,H} - c_2|, \quad (8)$$

where $\mathbf{X}$ is the vector of the real data corresponding to either the packet-count or the interarrival-time process. The existence of the mean absolute error is guaranteed since $\alpha > 1$ for the permissible range of H, as mentioned above. Results from simulations using the above procedure for parameter estimation are shown in figure 1 where we see that the synthesized traffic approximates very well the real one. We also note that



(a)



(b)

Figure 1. (a) Actual traffic (packet counts) with scaling 10sec, and (b) Synthesized traffic using the new model proposed (see eq. 4) for $\alpha = 1.63, c_1 = 100$ and $c_2 = 39$. The model parameters were estimated using the minimum absolute error approach (see eq. 8).

the estimated value α is far from 2 which means that the model is far from being Gaussian. It was suggested also in [11] that this type of traffic cannot be modeled using Fractional Gaussian Noise.

Another method to estimate the parameters α and c_1 is by using the expression of the Characteristic Function of the suggested model. An unbiased and consistent estimate of the CF can be obtained by employing the Sample Characteristic Function (SCF) [6]. From (7) it follows that:

$$\ln(-\ln|\Phi_M(\omega)|) = \alpha \cdot \ln|\omega| + \ln c_1^{\alpha} + \ln \sum_i |h_d(i)|^{\alpha} \quad (9)$$

which can be solved for α and c_1 using linear regression over uniformly spaced frequencies [6].

The last method which uses the SCF is faster over the first one that uses the average absolute error but it is very sensitive to estimation errors of the CF, especially when the scale c_1 of the real data is large. On the other hand, the first method might converge to different sets of parameters depending on the initial estimates.

In the following section we provide a link between the use of α-Stable distributions for network traffic modeling and the feasibility of fast simulations of network traffic for the estimation of the probability of rare events such as cell losses or excessive delays.

3 Fast simulation of networks using α-Stable modeling and Importance Sampling

In Network performance evaluation we are interested in estimating the probabilities of unwanted rare events that compromise the quality of service for users. Some of those unwanted events are packet losses, excessive waiting times, or component failures in systems. We need to know how the probabilities of such events change from node to node and also how they affect the end-to-end performance of the system. One disadvantage of moving away from traditional Poisson modeling is that analytical results for delays and overflow probabilities become very hard to compute, especially for non-trivial topologies with networks of queues and complex admission and routing procedures. In such cases, simulation is the only viable solution to characterize the network traffic locally, as well as end-to-end.

Suppose now that we need to simulate an ATM switch and verify that under normal conditions the probability of cell loss remains close to a typical value of 10^{-9}. Therefore, we need at least 10^9, or typically two orders of magnitude more packets (to reduce the variance of the estimation), to go through the switch during the simulation. Thus, there will be a need for generation of very long data records and many repetitions of the experiment which will require very long simulation times.

Importance Sampling (IS) is a method to reduce simulation times by infusing an increased number of rare events into the system and unbiasing the output by multiplying it by a likelihood function. In the following, we will give a short introduction to IS and show how α-Stable modeling can be used in this framework to build fast simulators.

Suppose that the probability of a rare event P_{RE} is equal to the probability that the random variable $\mathbf{X}$, with density function $p(x)$, is in the set A, $P(\mathbf{X} \in A)$:

$$P_{RE} \triangleq P(\mathbf{X} \in A) = \int 1_{(\mathbf{X} \in A)} p(x)\, dx = E_p[1_{(\mathbf{X} \in A)}] \quad (10)$$

The idea behind IS is to replace $p(x)$ by $\frac{p(x)}{g(x)} g(x) \triangleq L(x)g(x)$ in order to infuse more events $(\mathbf{X} \in A)$ into

the system by using a different density function $g(x)$ instead of $p(x)$. From (10) we get:

$$P_{RE} = \int 1_{(\mathbf{X}\in A)} \frac{p(x)}{g(x)} g(x)\, dx = E_g[1_{(\mathbf{X}\in A)} L(x)] \quad (11)$$

An unbiased estimate of the probability of the rare event is given by taking N samples generated by the PDF $g(x)$, $(\mathbf{X}_1, \ldots, \mathbf{X}_N)$ and then using the likelihood function $L(x)$ to unbiase the estimate:

$$\hat{P}_{RE} = \frac{1}{N} \sum_{n=1}^{N} 1_{(\mathbf{X}_n \in A)} L(\mathbf{X}_n) \quad (12)$$

If $p(x)$ is α-Stable with characteristic exponent α_1 and the rare event is in $A = \{\mathbf{X} : \mathbf{X} > x_o\}$, where $\mathbf{X}$ is the queue length and x_o is the buffer size, then $g(x)$ is chosen as α-Stable with characteristic exponent α_2, where $\alpha_2 < \alpha_1$. Then the probability of a rare event $P(\mathbf{X} > x_o)$ increases, since for a general α-Stable distribution $S_{\alpha,\sigma,\beta,\mu}(x)$ the tails behave as follows:

$$\lim_{x_o \to \infty} P(\mathbf{X} > x_o) = 0.5 C_\alpha (1+\beta) \sigma^\alpha x_o^{-\alpha}, \quad (13)$$

where C_α is a constant.

In the following section we provide results of simulations that show how accurate the estimator is over different simulation-time gains, and how its variance is affected by the choice of the twisting parameter α_2.

3.1 Simulations

We assume that the input process follows a symmetric α-Stable distribution (SaS) with $\alpha_1 = 1.8$ and that the rare event is $A = \{\mathbf{X} : \mathbf{X} > x_o\}$, where x_o is selected from fractile tables so that $P_{RE} = P(\mathbf{X} > x_o) = 10^{-4}$. For the above parameters P_{RE} and α_1, we find that $x_0 = 44.28$. The reason that we chose to simulate a regular α-Stable process rather than a Self-Similar α-Stable is that there exist very detailed fractile tables and also analytical expressions that provide very accurate calculation of the P_{RE} of our choice and a basis for comparison with the simulation results.

Since there is no closed form PDF for α-Stable distributions except for a few special cases, we used numerical computation by Fourier inversion of the general form of the Characteristic Function.

We carried out simulations for different numbers of samples $N = 10^2, 10^3, 5 \cdot 10^3$ and $N = 10^4$ for a range of α_2, where we calculated the average estimated P_{RE} and it's variance for 10 iterations of the same experiment. Note that if P_{RE} is given by (10) and is estimated by simulation, $\hat{P}_{RE} = 1/N \sum_{n=1}^{N} I_n$, where I_n is the indicator function of the set A. Then for a 10% error in the estimation with 99% confidence, we would need $N = 6.635 \times 10^6$ samples, by the Central Limit Theorem. Figures 2 and 3 show the results of the simulations for $N = 1000$ and $N = 5000$ respectively. We see that the estimates approach the theoretical value of 10^{-4} as the twisting parameter α_2 becomes smaller and also that the variance of the estimates decreases. Even though the sample size is three orders of magnitude smaller than the one required by the Central Limit Theorem without Importance Sampling, the error of the estimate is less than 10% for all $\alpha_2 < 1.1$ with $N = 1000$ and less than 5% with N=5000.

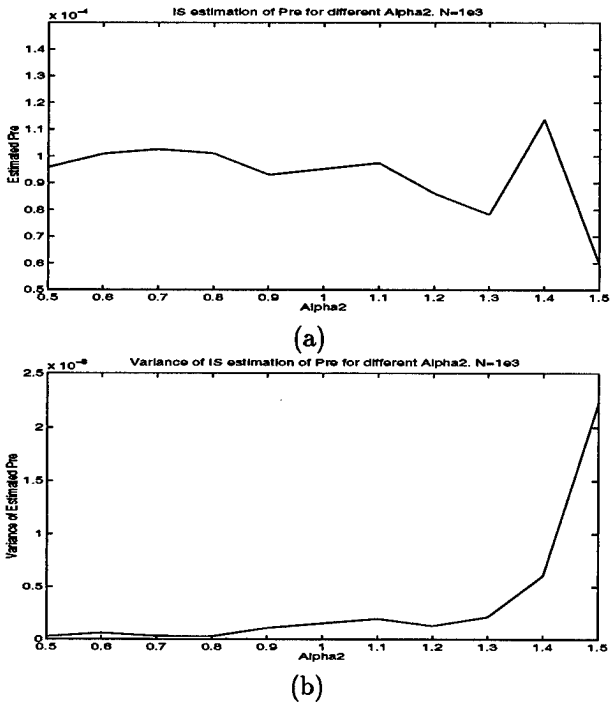


Figure 2. Importance Sampling fast simulation for sample size N=1000. (a) Estimated P_{RE} and, (b) Variance of estimated P_{RE}. We see that as the twisting parameter α_2 is decreased the estimates become more accurate and the variance decreases.

4 Conclusions

We showed that α-Stable Self-Similar models can be a very effective tool in modeling future high-speed network traffic. In addition, they can provide the foundation for fast network simulators for effective network evaluation and management.

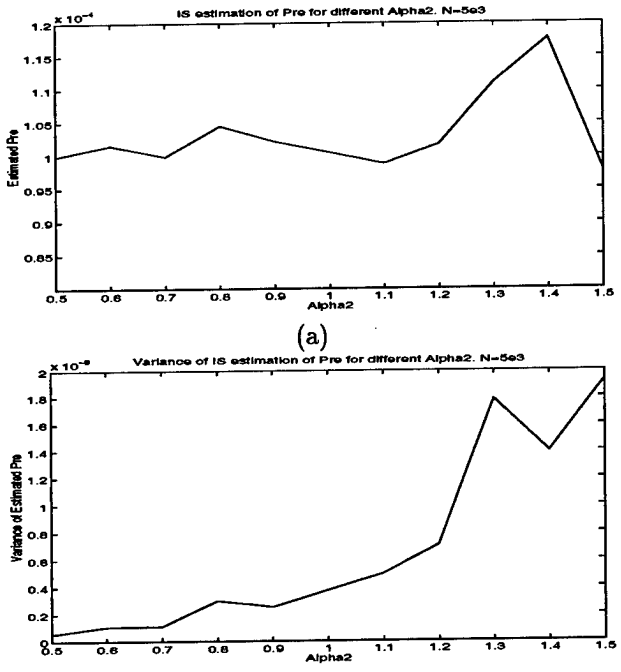


Figure 3. Importance Sampling fast simulation for sample size N=5000. (a) Estimated P_{RE} and, (b) Variance of estimated P_{RE}.

References

[1] Beran J., *Statistics for Long-Memory Processes*, Chapman & Hall 1994.

[2] Beran J., Sherman R., Taqqu M.S., Willinger W., "Long-Range Dependence in Variable-Bit-Rate Video Traffic", IEEE Trans. Comm., Vol. 43, No. 2/3/4, Feb/Mar/Apr 1995.

[3] Crovella M.E., Bestavros A., "Explaining World Wide Web Traffic Self-Similarity", Technical Report TR-95-015, Boston University, 1995.

[4] Heidelberger P., "Fast Simulation of Rare Events in Queuing and Reliability Models", ACM Trans. on Model. and Comp. Simula., Vol. 5, No. 1, Jan. 1995.

[5] Kogon S.M., Manolakis D.G., "Signal Modeling with Self-Similar α-Stable Processes: The Fractional Levy Stable Motion Model", IEEE Trans. Sig. Proc., Vol. 44, No. 4, April 1996.

[6] Koutrouvelis I.A., "Regression-Type Estimation of the Parameters of Stable Laws", J. Amer. Stat. Assoc., Vol. 75, No. 372, Dec. 1980.

[7] Leland W.E., Taqqu M.S., Willinger W., Wilson D.V., "On the Self-Similar Nature of Ethernet Traffic (Extended Version)", IEEE Trans. Networking, Vol. 2, No. 1, Febr. 1994.

[8] Likhanov N., Tsybakov B., Georganas N.D., "Analysis of an ATM Buffer with Self-Similar (Fractal) Input Traffic", Proc. INFOCOM 95, Boston, MA.

[9] Mandelbrot B.B., Van Ness H. W., "Fractional Brownian Motions, Fractional Noises and Applications", SIAM Rev., Vol. 10, pp. 422-436, Oct. 1968.

[10] Norros I, "A storage model with self-similar input", Queueing Systems, 16(1994), pp. 387-396.

[11] Norros I., "On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks", IEEE Sel. Areas in Comm., Vol. 13, No. 6, Aug. 1995.

[12] Paxson V., Floyd S., "Wide-Area Traffic: The Failure of the Poisson Modeling", IEEE/ACM Trans. Networking, 3(3), pp. 226-244, June 1995.

[13] Samorodnitsky G., Taqqu M.S., *Stable Non-Gaussian Random Processes*, Chapman & Hall, 1994.

TDE, DOA AND RELATED PARAMETER ESTIMATION PROBLEMS IN IMPULSIVE NOISE

Ananthram Swami *Brian Sadler*

Army Research Lab
AMSRL-IS-TA, 2800 Powder Mill Road, Adelphi, MD 20783-1197, USA
e-mail: (aswami,bsadler)@emh3.arl.mil

ABSTRACT

We address the problem of time-delay estimation (TDE) and direction-of-arrival (DOA) estimation in the presence of symmetric alpha-stable noise. We show that these problems can be handled by conventional correlation or cumulant based techniques, provided that the noisy signals are first passed through a generic zero-memory non-linearity. This pre-processing is also useful in the detection context. We also address the problem of blind linear system identification, where the input is an iid alpha-stable process; we show that consistent estimates of the possibly non-minimum phase ARMA parameters can be obtained by using self-normalized correlations and cumulants. Theoretical arguments are supported by simulations.

1 INTRODUCTION

In typical signal processing applications, additive noise can often be well modeled as the sum of a nominal stationary Gaussian component (thermal noise, etc) and high amplitude non-stationary, non-Gaussian (NG) components [16]. The NG component may be modeled as the output of a linear filter excited by an inhomogeneous Poisson process or random high-amplitude transients [16, pp 184-196]. In the stationary setting, the NG component can be modeled by a filtered Bernoulli-Gaussian product process [9], or as the realization of a Gaussian mixture process, which are special cases of the Middleton model [16, pp 137-215].

More recently, a lot of attention has been given to the stationary symmetric alpha stable (SαS) process. The characteristic function of the SαS random variable (rv) is given by [8, 10], $E\{\exp(jvx)\} = \exp(-\gamma|v|^\alpha)$, $0 < \alpha \leq 2$, $\gamma > 0$. Parameter α is the index or characteristic exponent, γ is the dispersion, and $\zeta := \gamma^{1/\alpha}$ is called the scale parameter. Apart from the lack of a closed-form pdf ($\alpha \neq 1, \alpha \neq 2$), the S$\alpha$S rv possesses various interesting properties such as, tail probabilities of order $|x|^{-\alpha}$ and $E|x|^\rho = \infty$, $\rho > \alpha$, $\alpha < 2$, etc [10, 8]. Note that sample estimates of $|x|^\rho$ are consistent only for $-1/2 < \rho < \alpha/2$.

Because of the 'infinite variance' of the SαS rv, the *covariation*,

$$C_\alpha(x,y) := \gamma_y E\{X|Y|^{p-1}Y^*\}/E|Y|^p, p < \alpha \leq 2 \qquad (1)$$

and the covariation coefficient, $\lambda_{x,y} := C_\alpha(x,y)/\gamma_y$, have been used instead of the correlation in various applications [5]. The covariation is defined *only* if both X and Y are SαS with the *same* α. In [13], we have shown that estimates of the normalized correlation show much less variability than those of the covariation. Motivated by [1], we used normalized correlations to estimate the 'spectrally' equivalent minimum-phase (SEMP) parameters of a linear SαS process. Here, we will define and use normalized cumulants to estimate parameters of mixed-phase models.

A classical approach to detection in impulsive noise is to use a locally optimum detector (LOD); but this requires knowledge of the noise pdf. An alternative is to use a zero-memory non-linearity (ZMNL) which captures the LOD behavior for a class of pdfs. Pre-processing the data by using a signed fractional power, such as in the covariation, is equivalent to using a particular ZMNL, but is not optimal in any sense. We use ZMNL pre-processing followed by correlation-based processors for both detection and estimation. We demonstrate its use in the context of harmonic retrieval, direction-of-arrival (DOA) estimation, and time-delay estimation (TDE).

2 LINEAR SαS PROCESSES

We address the problem of estimating the ARMA parameters of the process,

$$y(n) = -\sum_{k=1}^{p} a(k)y(n-k) + \sum_{k=0}^{q} b(k)u(n-k), \qquad (2)$$

where $u(n)$ is iid SαS. Impulse response (IR), $h(n)$, is assumed to satisfy the α-summability condition: $\exists \delta \in (0,\alpha) \cap (0,1)$ st $\sum_m |h(m)|^\delta < \infty$.

Hannan [2] showed that the correlation-based Yule-Walker equations lead to consistent estimates for the pure AR case. Davis and Resnick [1] showed that the sample estimates of the normalized correlation of a linear SαS process are consistent,

$$\widehat{R}_{xx}(m) := \sum_{n=1}^{N} x(n)x^*(n+m) \;/\; \sum_{n=1}^{N} |x(n)|^2 \qquad (3)$$

$$\xrightarrow{p} R_{hh}(m) := \sum_{j=-\infty}^{\infty} h(j)h(j+m) \;/\; \sum_{j=-\infty}^{\infty} h^2(j) \qquad (4)$$

where the convergence is in probability. Mikosch *et al* [7], extending the results of [1], showed that the sample estimate of the self-normalized periodogram is consistent, and that Whittle-type estimators based on it lead to consistent estimates of the ARMA parameters. All of these papers also established rates of convergence, of the form $N^{1/\alpha}$. The ML estimator for a MA(1) Cauchy process was derived in [15]. The α-spectrum was introduced by Ma-Nikias [4] who showed that it retains both amplitude and phase information, thus establishing that non-minimum phase linear systems can, in principle, be blindly estimated. However, consistency of the sample estimate of the α-spectrum, or of the resulting (truncated) cepstral coefficients was not established.

Motivated by the results of [1, 7], we proposed to use correlation-based algorithms to estimate the SEMP model

0-8186-8005-9/97 $10.00 © 1997 IEEE

parameters in [13]; of course, inherent all-pass factors cannot be identified via this approach. We define normalized fourth-order moments via

$$\widehat{M}_{4x}(\tau_1,\tau_2,\tau_3) := \frac{\sum_n x(n)x(n+\tau_1)x(n+\tau_2)x(n+\tau_3)}{(\sum_n |x(n)|^2)^2} . \quad (5)$$

Normalized moments of other orders (and conjugations in the complex case) can be defined similarly. Fourth-order cumulants can now be defined in the usual way,

$$\widehat{C}_{4x}(\tau_1,\tau_2,\tau_3) = \widehat{M}_{4x}(\tau_1,\tau_2,\tau_3) - \widehat{R}_{xx}(\tau_1)\widehat{R}_{xx}(\tau_2-\tau_3) - \widehat{R}_{xx}(\tau_2)\widehat{R}_{xx}(\tau_3-\tau_1) - \widehat{R}_{xx}(\tau_3)\widehat{R}_{xx}(\tau_1-\tau_2) \quad (6)$$

where $\widehat{R}_{xx}$ was defined in (3). We claim that

$$\widehat{C}_{4x}(\tau_1,\tau_2,\tau_3) \xrightarrow{p} \frac{\sum_n h(n)h(n+\tau_1)h(n+\tau_2)h(n+\tau_3)}{\kappa\,(\sum_n |h(n)|^2)^2}$$

where κ is a constant. Self-normalized polyspectra can then be defined as the Fourier transforms of the cumulant sequences. Rates of convergence and limiting distributions can presumably be established along the lines of [1].

Having estimated normalized correlations and fourth-order cumulants, we can use various algorithms to resolve the true zeros of the linear model via cumulants [6]. We note that our approach holds for $0 < \alpha < 2$, not just for $1 < \alpha < 2$.

Example 1. We simulated a MA(3) process for two different values of α; additive white Gaussian noise (AWGN) with variance σ_g^2 was added to the linear process. We estimated the normalized autocorrelation from $N = 2000$ samples; we fitted a long AR model (order 27) and then a MA model to it. Since SOS are used, the algorithm yields estimates of the SEMP model. Mean and standard deviations of the estimates, averaged over 100 trials, are given in Table 1, Notice that when α is small, the additive noise has very little effect. In the noisefree case, the estimates are unbiased and have very low variance. As α becomes larger, the signal becomes more Gaussian, and it is harder to suppress the effects of the AWGN with shorter data lengths. The smaller the value of α, the more spiky are the data. The normalization suppresses all but the strongest impulses; thus, the 'normalized data', consists of relatively well-separated impulse responses, which facilitates the estimation of the model parameters. The normalization also helps to suppress the effects of additive finite variance noise.

In order to resolve the true zeros of the model, we estimated the normalized fourth-order cumulants in (6), and then used them in the 'GM RC' least-squares algorithm [6]. Estimates of the mixed phase model are shown in Table 2. Note the good performance of the estimator. The variance of the estimate increases as α increases; this is because the rates of convergence of the moment estimates is of order $N^{1/\alpha}$, i.e., for a fixed N, we expect better performance when α is smaller (more non-Gaussian). We note that the non-parametric α-spectrum approach of [4] required over 5×10^5 samples to yield good estimates.

Example 2. We simulated an AR(4) process, with two different values of α, and $N = 2000$. The AR parameters were estimated from the (normalized) correlation (20 lags) via the Yule-Walker equations. Means and standard deviations of the parameter estimates, averaged over 100 trials, are reported in Table 3. Notice the good performance of the estimates, as expected from the results of [2].

Example 3. The observed signal was $y(t) = u(t)/A_I(z) + g(t)/A_n(z)$, where signal $u(t)$ was SαS with $\alpha = 1, \gamma = 1$; noise $g(t)$ was white Gaussian with variance σ_g^2. The AR parameters were $A_I = [1, 0, 0.75]$ and $A_n = [1, -0.4, 0.6]$. We estimated the A_I parameters using normalized second- and fourth-order statistics. Mean and standard deviations estimated from 100 trials, with $N = 1024$ (first two rows), and $N = 4096$ (last two rows) are shown in Table 4. Notice the good performance of the estimates even in the noisy case. The cumulant-based estimates appear to show an advantage for small N and smaller α.

Example 4. We simulated an ARMA(3,3) process with parameters AR= $[1, -1.6, 1.21, -0.36]$, MA= $[1.5, -6.21, 9.1514, -4.1147]$, $\gamma = 1$, $\alpha = 0.5, 1.5$. The data length was $N = 4000$. We used SOS to estimate the SEMP parameters both in the noisefree case, as well as in the noisy case (AWGN, with $\sigma_g^2 = 100$). Mean and σ-bounds of the estimates, averaged over 100 trials, are shown in Fig. 1. The circles show the true SEMP IR. Notice the excellent performance of the estimator. We note that the α-spectrum approach of [4] required about 10^5 samples in the noisefree case.

We conclude this section by noting that normalized correlations and cumulants can be used to obtain consistent estimates of the ARMA parameters, even in the presence of finite variance additive noise (arbitrary color and pdf). Low-variance estimates are obtained even with moderate data lengths ($N = 2000$).

3 ESTIMATION IN SαS NOISE

The observed signal is described by

$$x(n) = c(n\,|\theta) + w(n), \quad n = 1, ..., N, \quad (7)$$

Here $c(n\,|\theta)$ may be either deterministic or a finite-variance signal, $w(n)$ is iid SαS, and we want to estimate the non-random parameter vector θ. The use of MLE's is hampered by the lack of closed-form expressions for the pdfs. Since $E(|u|^2) = \infty$, we are clearly in the weak signal regime. The obvious approach is to clip the impulsive noise by passing it through some ZMNL, which is linear near the origin. We use a ZMNL which was suggested by Ljung [9],

$$\begin{aligned} \bar{\sigma} &= \text{median}\{|x(n) - \text{median}\{x(n)\}\} \\ \widehat{\sigma}_N &= \bar{\sigma}/\sqrt{0.7} \quad \delta = 3\widehat{\sigma}_N \\ \phi(u) &= \begin{cases} -\delta \exp(-(u+\delta)^2/2\sigma^2)) & u < -\delta \\ u & |u| < \delta \\ \delta \exp(-(u-\delta)^2/2\sigma^2)) & u > \delta \end{cases} \end{aligned} \quad (8)$$

σ is a tuning parameter which we arbitrarily set to $\widehat{\sigma}_N$. If the signal is not very weak, it is better to set δ to $3\widehat{\sigma}_N + \text{median}\{x(n)\}$. A good ZMNL will essentially transform the SαS noise to a more Gaussian looking one, leaving the signal intact. Hence, algorithms which assume that the additive noise is nominally Gaussian can be used with the ZMNL output. An attractive feature of the above ZMNL is that it is data-adaptive; second, it is analytically tractable, and the computational load is not high. The ZMNL involves computation of the median; in the context of SαS processes, we know that the median is a consistent estimator of location.

We applied this idea to three estimation problems: harmonic retrieval, DOA and TDE. Analytic derivation of performance bounds and comparisons with CRB's will be reported elsewhere (see also Section 5).

Since $w(n)$ has infinite variance, a generalized SNR defined by $E\{|c(t)|^2\}/\gamma_w$ has been proposed; however, the empirical SNR for a fixed γ varies tremendously with α. We suggest that a better definition of SNR is $E\{|c(t)|^2\}/E\{\phi^2(w(t))\}$, where the denominator is the variance of the ZMNL output.

We applied ZMNL pre-processing to standard parameter estimation problems: harmonic retrieval, DOA and TDE.

3.1 Harmonic Retrieval

The noisy signal is given by $y(n) = \sum_{k=1}^{p} A_k \exp(j2\pi f_k n) + w(n)$, $n = 1, ..., N$, where $w(n)$ is complex isotropic SαS, A_k's are complex, perhaps random, amplitudes, and f_k are the unknown frequencies to be estimated. We compare our correlation-based method to covariation-based methods.

Example 5. We let $p = 2$, $f_1 = 0.1$, $f_2 = 0.2$, $|A_1| = |A_2| = 1$, $\gamma_w = 1$. We passed the noisy complex signal through a circularly symmetric version of the ZMNL in (8); the correlation of the ZMNL output was estimated, and ESPRIT was then applied to estimate the frequencies. Each realization had 640 samples (16 realizations with random phases, each with 40 samples). We also applied ESPRIT to the raw data, and to the covariation matrix estimated with $p = \alpha/4$. The mean estimates (100 trials) are shown in Fig. 2 as a function of α, for the two frequencies; the corresponding standard deviations are shown in Fig. 3. In the figures, $*$, $+$, o and x denote estimates based on ZMNL-ESPRIT, raw ESPRIT, covariations and the normalized covariations. Notice that the ZMNL-based method yields low variance estimates for $\alpha > 0.5$; the raw correlation-based method performs poorly, as expected. Rather surprisingly, covariation-based ESPRIT did not work well; it has been reported in the literature that the Bartlett estimator based on covariations outperforms that based on correlations.

3.2 TDE

The observed signals are $x(n) = s(n) + u(n)$, $y(n) = As(n - D) + v(n), n = 1, ..., N$, where signal $s(n)$ is deterministic or random with finite variance, $u(t)$ and $v(t)$ are the sensor noises, assumed to be SαS, A is the unknown attenuation factor, and D is the unknown delay to be estimated. The sensor signals are passed through the ZMNL in (8) to clip the noise; the delay is estimated by locating the peak of the 'ML-windowed' autocorrelation.

Example 6. We let $D = 10$, $A = 0.9$, $N = 4000$. As in [5], $s(n)$ was white Gaussian, with $\sigma_s^2 = 1$; $u(n)$ and $v(n)$ were iid SαS processes with $\gamma = 1$. Figure 4 shows the performance (100 trials) of the proposed algorithm for $\alpha = 0.7, 1, 1.8, 2.0$.

As noted in [5], covariation-based techniques can be used only if $\alpha > 1$, and do not appear to work well for $\alpha < 1.5$. Two new techniques were reported in [5]: In one method the sensor signals are raised to a (possibly different) fractional power, $0 \leq p < \alpha/2$, and then cross-correlated. A second method minimizes the p-norm. Both these methods required $N = 20,000$ samples to obtain results comparable to ours. With ZMNL pre-processing, the correlation-based TDE worked extremely well.

3.3 DOA

The approach used in the harmonic retrieval problem extends to the DOA problem, even for $\alpha < 1$; in contrast, the ROC-MUSIC approach of [14] is applicable only for $\alpha \geq 1$, (and, indeed, assumes that all the source signals and the noise are jointly SαS with the same α).

Example 7. The complex SαS noise had unit dispersion; the two source signals were white Gaussian with unit variance. An eight sensor ULA was assumed, and 500 snapshots were used. In the first case, the source bearings were $\pm 5^o$, and in the second case $\pm 10^o$. The bearings were estimated via correlation-based ESPRIT, both with (o) and without ($+$) the ZMNL pre-processing. Because of the poor performance of covariation-based ESPRIT in the harmonic retrieval problem, we did not consider it here. Means and standard-deviations of the estimates, averaged over 100 runs, are shown in Figure 5, as a function of α. The proposed estimator is seen to work very well.

The data-adaptive ZMNL of (8) is very effective in suppressing the effects of SαS noise. Its performance is better than that of covariation-based techniques; more importantly, it works well for all kinds of noises: Gaussian, non-Gaussian, stable, finite variance. The colored noise case is handled along the lines in [9], and results comparable to those in the iid case are obtained.

4 DETECTION

One of the classical approaches to detection in the presence of impulsive noise (weak signal) is to use a locally optimum detector (LOD). Designing a LOD requires knowledge of the noise pdf; an alternative is to use a ZMNL which captures the LOD behavior for a class of pdfs; this appears to work well even in SαS noise [3]. The LOD non-linearity for SαS rv's is linear near $x = 0$, with slope $\Gamma(3/\alpha)/\Gamma(1/\alpha)$; the linear approximation is valid for $|x|^2 \ll \Gamma(1/\alpha)/\Gamma(3/\alpha)$; its tails are given by $(\alpha + 1)/|x|$. This partly explains the 'robustness' of Cauchy detectors. The ZMNL in (8) has an exponential tail, but works well. As we saw earlier, this ZMNL is very useful in the estimation context.

As we argued before, the ZMNL essentially transforms the impulsive noise to a more Gaussian-looking one; hence, standard detectors, which nominally assume additive Gaussian noise, can be used at the output of the ZMNL. We demonstrate this idea by considering a classical detection problem.

Example 8. The observed signal is $y(n) = A + u(n)$, $n = 1, ..., N$, where $u(n)$ is SαS. We want to test whether or not $A = 0$ (i.e., detect the signal). The standard approach is to estimate the mean (or median) and see if it is statistically different from zero. When $u(n)$ is Gaussian, we have the Student t-test, which detects the signal if $\sqrt{N}\ |\hat{\mu}| \ / \ \sqrt{\widehat{\sigma^2}} > \tau$, where $\hat{\mu}$ and $\widehat{\sigma^2}$ are sample estimates of mean and variance, and τ is a threshold (1.96 for P_{fa} of 0.05). Obviously this test will perform poorly if $u(n)$ is SαS. We passed the noisy data through the ZMNL in (8), and then applied the Student t-test, with $\tau = 1.96$. ROC curves for $\alpha = 1$ are shown in Fig. 6 (solid line: w/o ZMNL, '+' with ZMNL). In these curves $P_{fa} = 0.05$ corresponds to $P_d = 0.9257$ (.2728) with (without) the ZMNL when N=50, and $P_d = 0.9965$ (.2763) when $N = 100$. Note that for $N \geq 100$, the ROC with the ZMNL is essentially flat at $P_d = 1$. These curves were computed from a set of 10^6 trials.

We repeated the experiment, again using default values for the $\phi(\cdot)$ in (8), and with fixed threshold $\tau = 1.96$, but varying α. P_d and P_{fa} estimated from 10^6 runs are shown as a function of α in Fig 7. The solid line is P_d using the ZMNL, and the $+$'s denote P_d w/o the ZMNL. We remark that the performance can be further improved by tuning the parameters of the ZMNL.

5 CRAMER-RAO BOUNDS

Consider the parameter estimation problem for deterministic signals in iid SαS noise. The observed signal is $y(n) = \mu c(n\ |\theta) + w(n)$, $n = 1, ..., N$, where μ and θ are unknown non-random constants, and $w(n)$ is iid SαS. The waveform $c(.)$ is completely specified by the vector θ. The developments in [11, 12] (which consider the finite variance case) are directly applicable here. Let $I_f(\alpha)$ denote the Fisher Information Matrix (FIM) for location for a SαS process with parameter α and unit dispersion, $\gamma = 1$. Then, the FIM for θ is given by

$$J_{ij} = I_f(\alpha) \frac{\mu^2}{\gamma_w^{2/\alpha}} \sum_{n=1}^{N} \frac{\partial c(n)}{\partial \theta_i} \frac{\partial c(n)}{\partial \theta_j} .$$

Consider the multiplicative noise models, $y(n) = w(n)c(n\ |\theta)$, where $w(n)$ is iid SαS. Assume that $1 < \alpha$ and $E\{w(n)\} = \mu$. Let $I_\gamma(\alpha)$ denote the FIM for the squared

scale parameter $\gamma^{2/\alpha}$ of a standard SαS rv (i.e., with $\gamma = 1$). Following [12], the FIM is given by,

$$J_{ij} = \left[4I_\gamma(\alpha) + \frac{\mu^2}{\gamma_w^{2/\alpha}} I_f(\alpha)\right] \left[\sum_{n=1}^{N} \frac{\partial c(n)}{\partial \theta_i} \frac{\partial c(n)}{\partial \theta_j} \frac{1}{c^2(n)}\right]$$

where we assume that $c(n) \neq 0\ \forall n$. The FIM for location and scale are finite, and are shown in Fig. 8 as a function of α ($\gamma = 1$). Note that $I_f(\alpha)$ is essentially linear in α, and $I_\gamma(\alpha)$ is weakly quadratic; notice the change in convexity at $\alpha = 1$. The FIM's were numerically evaluated via the STABLE program written by Prof. John Nolan at American University.

Now consider the parameter estimation problem for AR SαS processes, $y(n) = -\sum_{k=1}^{p} a(k)y(n-k)$. Following [11, 12], the FIM for the AR parameters is given by

$$J(\mathbf{a}) = (N-p) I_f(\alpha) E\{u^2\} R_{hh},$$

where R_{hh}, the deterministic auto-correlation of the impulse response $h(n)$, and $I_f(\alpha)$ are both finite. But $E\{u^2\} = \infty$; hence, the FIM is singular. The argument extends to the general linear model. Hence, in the context of linear SαS processes, it is important to establish rates of convergence of the estimator. Letting $C = \sigma_u^2 R_{hh}$ as above, the FIM's in (10) and (16) of [11] hold for $\alpha > 1$.

6 CONCLUSIONS

Normalized correlations and cumulants can be used to obtain consistent low-variance estimates of the ARMA parameters of linear SαS processes. Simple ZMNL pre-processing helps in detection/estimation problems, when the noise is heavy-tailed, and perhaps α-stable. Cramer-Rao bounds were derived for the parameter estimation problem. The performance of our algorithms was illustrated via various examples.

REFERENCES

[1] R.A. Davis and S.I. Resnick, *Ann Stat*, 14(2), 533-558, 1986.

[2] E.J. Hannan and M. Kanter, *J. Appl. Prob.* 411-415, 1977.

[3] J. Ilow, D. Hatzinakos and A.N. Venetsanopoulos, *Proc. Milcom-94*, 287-290, 1994.

[4] X. Ma and C.L. Nikias, *IEEE Trans Sig Proc*, 43(12), 2884-2897, Dec 1995.

[5] X. Ma and C.L. Nikias, *IEEE Trans Sig Proc*, 2669-2687, Nov 1996.

[6] J.M. Mendel, *Proc. IEEE*, 79, 278-305, 1991.

[7] T. Mikosch *et al*, *Annals of Stat*, 305-326, 1995.

[8] C.L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*, John Wiley, 1995.

[9] B. Sadler, *IEEE Trans Sig Proc*, 2793-2800, Nov 1996.

[10] G. Samorodnitsky and M.S. Taqqu, *Stable non-Gaussian random processes*, Chapman & Hall, 1994

[11] D. Sengupta and S.M. Kay, *IEEE Trans ASSP*, 38(10), 1661-76, Oct 1990.

[12] A. Swami, *Signal Processing*, vol 53, pp 231-244, Sept 1996.

[13] A. Swami, *Proc. ICASSP-97*, Munich, Germany, Vol 5, pp 3541-44, April 1997.

[14] P. Tsakalides and C.L. Nikias, *IEEE Trans Sig Proc*, 1623-33, July 1996.

[15] G.A. Tsihrintzis and C.L. Nikias, *IEEE Sig Proc Lett*, 2(12), 224-226, Dec 1995.

[16] E.J. Wegman, S.C. Schwartz and J.B. Thomas, *Topics in non-Gaussian signal processing*, Springer-Verlag, 1989.

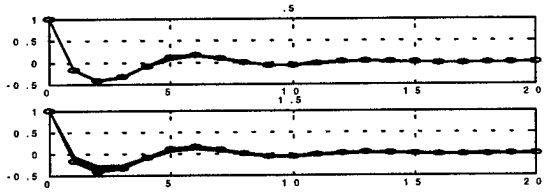


Fig 1. (Ex 4) Estimated IR's, $\mu \pm \sigma$ bounds; top: $\alpha = 0.5$, bottom: $\alpha = 1.5$

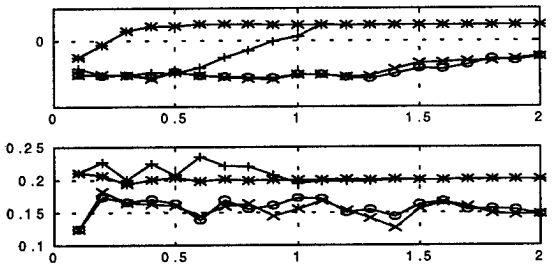


Fig 2. (Ex 5) Mean estimates of f_1 and f_2 vs α

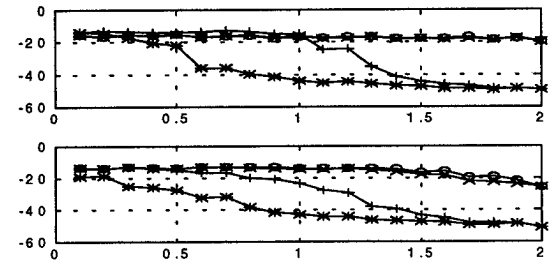


Fig 3. (Ex 5) Std-dev of estimates of f_1 and f_2 vs α

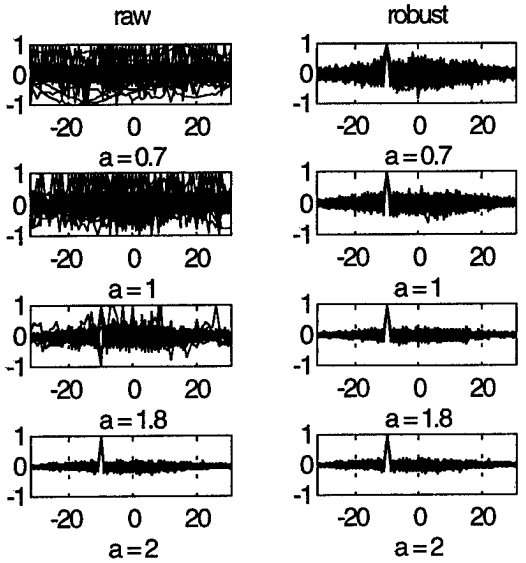


Fig 4. (Ex 6) TDE w and w/o ZMNL

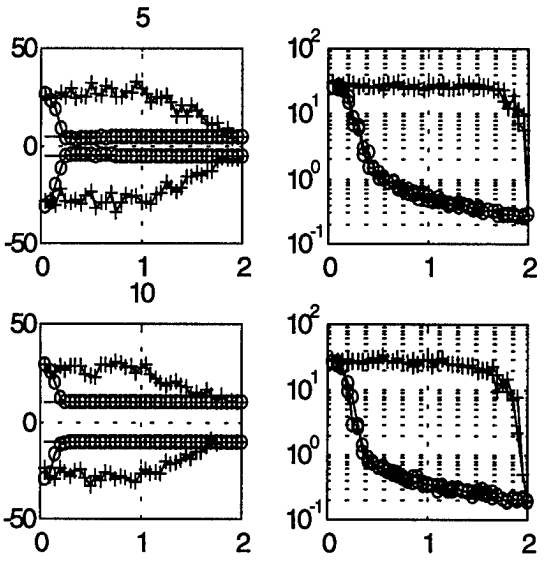


Fig 5. (Ex 7) DOA estimates: means (left col) and std-deviations (right col), for bearings of $\pm 5^o$ (top row) and $\pm 10^o$ (bottom row).

	b(1)	b(2)	b(3)
true	0.5000	-1.3000	0.7000
MP	-0.4644	-0.2464	0.2760
	$\alpha = 0.5$		
$\sigma_g = 0$	-0.4642 (0.0017)	-0.2462 (0.0023)	0.2753 (0.0043)
$\sigma_g = 10$	-0.4642 (0.0017)	-0.2462 (0.0023)	0.2753 (0.0043)
	$\alpha = 1.5$		
$\sigma_g = 0$	-0.4639 (0.0200)	-0.2394 (0.0344)	0.2706 (0.0238)
$\sigma_g = 10$	-0.1709 (0.0994)	-0.1200 (0.0566)	0.1026 (0.0625)

Table 1: (Example 1) SEMP MA estimates, based on normalized R_{xx}.

	b(1)	b(2)	b(3)
true	0.5000	-1.3000	0.7000
$\alpha = 0.5$	0.5112 (0.0511)	-1.2932 (0.0618)	0.6952 (0.0359)
$\alpha = 1.0$	0.5089 (0.1067)	-1.2861 (0.1375)	0.6913 (0.0842)
$\alpha = 1.2$	0.4944 (0.1989)	-1.3272 (0.1610)	0.7120 (0.0794)
$\alpha = 1.5$	0.4958 (0.2420)	-1.3316 (0.2482)	0.7074 (0.1163)

Table 2: (Example 1) Mixed-phase MA estimates, based on normalized R_{xx} and C_{4x}.

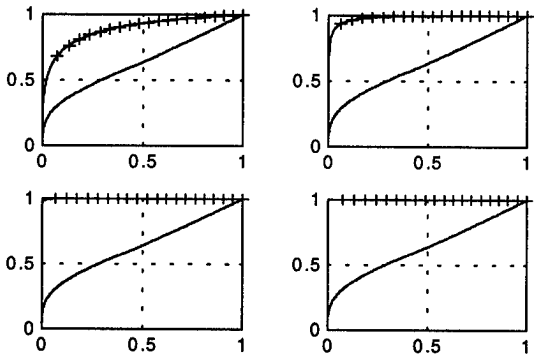


Fig 6. (Ex 8) ROC curve, $\alpha = 1$, $N = 20, 50, 100, 200$

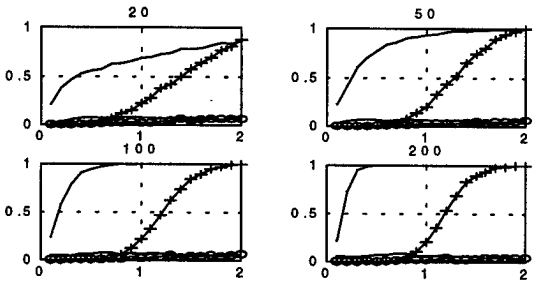


Fig 7. (Ex 8) P_d vs α, $\tau = 1.96$, for $N = 20, 50, 100, 200$

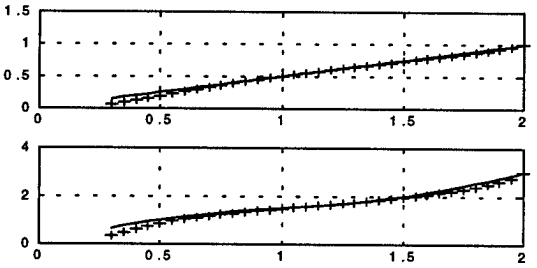


Fig 8. FIM for location (top) and scale (bottom) vs α, with $\gamma = 1$

	a(1)	a(2)	a(3)	a(4)
true	-1.6000	1.9000	-1.0000	0.4000
$\alpha = 0.5$	-1.5947 (0.0419)	1.8930 (0.0665)	-0.9947 (0.0573)	0.3991 (.0253)
$\alpha = 1.5$	-1.5890 (0.0415)	1.8822 (0.0714)	-0.9836 (0.0684)	0.3939 (0.0297)

Table 3: (Example 2) AR parameter estimates based on R_{xx}.

	$\sigma_g = 1$		$\sigma_g = 10$	
true	0.0000	0.7500	0.0000	0.7500
c2	-0.0020 (0.0182)	0.7496 (0.0180)	-0.0223 (0.0344)	0.7297 (0.0316)
c4	-0.0022 (0.0256)	0.7442 (0.0159)	-0.0042 (0.0287)	0.7418 (0.0198)
c2	0.0001 (0.0113)	0.7479 (0.0111)	-0.0053 (0.0123)	0.7423 (0.0133)
c4	-0.0002 (0.0144)	0.7467 (0.0136)	0.0000 (0.0149)	0.7466 (0.0141)

Table 4: (Example 3) AR estimates from noisy data

TPL: Array Processing

Chair: Ben Friedlander
University of California-Davis, USA

Which cumulants should be selected for steering vector estimation ?

Thomas Kaiser*
Gerhard-Mercator University
Department of Communication Engineering
FB 9, FG NT, Bismarckstrasse 81
47048 Duisburg, Germany
th.kaiser@uni-duisburg.de

Jerry M. Mendel
University of Southern California
Signal and Image Processing Institute
3740 McClintock Ave
Los Angeles, CA 90089-2564, USA
mendel@sipi.usc.edu

Abstract

Cumulants have been successfully applied in the area of narrowband array signal processing. This motivates a performance analysis to find out the strengths and the weaknessess of each new algorithm. Hitherto, most of the known performance analyses are based on the asymptotic covariance of sample cumulants and are therefore called asymptotic performance analyses. Recently, explicit formulas for the finite-sample covariances of second-, third-, and fourth-order sample cumulants for any kind of signal, any kind of noise, any array shape and arbitrary sensors are derived [8],[9]. These formulas enable a finite-sample performance analysis.

In the single source case the steering vector is proportional to a vector built up by a product of second-order cumulants or by fourth-order cumulants. This means that the finite-sample (co)variance of the steering vector can be investigated by using the formulas for the finite-sample covariance of the second- and fourth-order sample cumulant. Hence, the open question "Which cumulants should be selected for steering vector estimation ?" - is addressed in this paper.

1 Introduction

Since existing second-order statistics based methods cannot solve numerous problems in narrowband array signal processing, many higher-order cumulant based algorithms have been recently proposed. For example, the virtual-ESPRIT-Algorithm (VESPA) for direction-finding and recovery of independent sources can also calibrate an array of unknown configuration [5]. Also VESPA has been extended for direction-finding of highly correlated or coherent sources [7], which is often the case in practice due to multipath propagation or "smart" jamming. Since higher-order cumulants are insensitive to additive colored or white Gaussian noise, the suppression of Gaussian noise is always accomplished by VESPA and extended VESPA. These new methods and algorithms require a performance analysis to find out their strengths and weaknessess.

Some performance analyses for higher-order cumulant based methods have already been done. For example, *Cardoso and Moulines* [2] derived closed-form expressions for the asymptotic covariance of MUSIC-like direction-of-arrival (DOA) estimates based on two different fourth-order cumulant matrices. *Yuen and Friedlander* [11] have compared the performance of ESPRIT [10], higher-order ESPRIT [4], and VESPA-algorithm. Performance analyses are usually based on *asymptotic* covariances of second- or fourth-order sample cumulants; therefore, to provide a basis for investigation of the finite-sample case, we have recently derived the *finite-sample* covariance of second-, third-, and fourth-order sample cumulants [8], [9].

This paper emphasizes the finite-sample case to understand the excellent performance for modest numbers of snapshots observed by *Doğan and Mendel* and *Gőnen et al.* [7]. Furthermore, investigating the finite-sample case also allows the analyses of adaptive algorithms. Another noticeable advantage of an analytic formula, compared to often used Monte-Carlo simulations, is the easier and more reliable recognition of properties of the *finite-sample* covariance. This can be seen later by some examples.

The purpose of this paper is to study the accuracy of steering vector estimates based on second- or fourth-order sample cumulants as a function of the array shape, the DOA's, the kind and number of sensor's, the kind of source and noise signals, the number of samples and the signal-to-noise ratio (SNR). Indeed, by using the finite-sample covariance formula, the *necessary number of samples* for a required estimation accuracy can be predicted. Alternatively, the *optimal source signals* in the sense of highest estimation accuracy can be obtained by optimization. The structure of the paper is as follows. In Section 2, we formulate the problem and introduce the neccessary notation. An analysis of the single-source case is given in Section 3. In the last Section some interesting examples are shown.

*This work was supported by a scholarship of the NATO Scientific Board and co-ordinated by the Deutscher Akademischer Austauschdienst.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2 Problem Formulation

Let P narrowband plane waves centered at a known frequency ω_0 impinge on an arbitrary array composed of M sensors. The received signal vector $\mathbf{r}(t) = (r_1(t), ..., r_M(t))^T$ can be modeled as

$$\mathbf{r}(t) = \mathbf{A}(\phi, \vartheta)\, \mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{s}(t) = (s_1(t), ..., s_P(t))^T$ is a $P \times 1$ vector which contains the P zero-mean source signal complex envelopes at time t, and $\mathbf{A}(\phi, \vartheta)$ is an $M \times P$ steering matrix. $\phi \in [-\pi, \pi]$ is the azimuth angle, $\vartheta \in [0, \pi]$ is the elevation angle and $\mathbf{n}(t)$ is an $M \times 1$ vector composed of M zero-mean noise signals. Any noise signal $n_m(t)$, $m = 1(1)M$ is[1] independent of any source signal $s_p(t)$, $p = 1(1)P$. Furthermore, all signals $s_p(t)$, $n_m(t)$ are modeled as sequences of independent and identically, but *arbitrarily* distributed complex random variables with finite moments up to the eighth-order. This model is commonly used in the far-field case, but it can be extended to coherent signals [7] and to the near-field case [3].

In this article, we focus our attention on *second-* and *fourth-order cumulants* (see [8] for third-order cumulants), which are respectively defined as

$$c_{k,l} = \mathrm{E}\{r_k(t) r_l^*(t)\} \tag{2}$$

and

$$\begin{aligned} c_{k,l,m,n} \;=\; & \mathrm{E}\{r_k(t) r_l^*(t) r_m^*(t) r_n(t)\} \\ & - \mathrm{E}\{r_k(t) r_l^*(t)\}\, \mathrm{E}\{r_m^*(t) r_n(t)\} \\ & - \mathrm{E}\{r_k(t) r_m^*(t)\}\, \mathrm{E}\{r_l^*(t) r_n(t)\} \\ & - \mathrm{E}\{r_k(t) r_n(t)\}\, \mathrm{E}\{r_l^*(t) r_m^*(t)\}, \end{aligned} \tag{3}$$

since $\mathrm{E}\{r_m(t)\} = 0 \,\forall\, m = 1(1)M$. The second-order, $\hat{c}_{k,l}$, and fourth-order *sample* cumulant, $\hat{c}_{k,l,m,n}$, are obtained by replacing the expected values $\mathrm{E}\{\}$ by time averages $\frac{1}{N}\sum_{t=1}^{N}$, where N is the number of samples. Hence, given a realization of the random vector $\mathbf{r}(t)$, the sample cumulants can be calculated. Note that although the estimator $\hat{c}_{k,l}$ is unbiased, the proposed estimator $\hat{c}_{k,l,m,n}$ is *biased* for the assumed model. In order to obtain an unbiased estimator, each sum of $\hat{c}_{k,l,m,n}$ must be multiplied with a proper constant (see [8]).

We omit the tedious calculations and the formulas of the finite-sample covariances. For example, the finite-sample covariance of the fourth-order sample cumulants consist of 1 eighth-, 12 sixth-, 76 fourth-, and 132 second-order moments of $r_k(t)$, resulting in several thousand terms of moments of $s_p(t)$, $n_m(t)$ [8],[9]. All these formulas are published in different computer languages on the World-Wide-Web under `http://fb9nt-1n.uni-duisburg.de/mitarbeiter/kaiser/hoswshop.97/hoswshop97.html`.

[1] $m = 1(1)M$ is a more compact form of $m = 1, 2, ..., M$. The number in parentheses means the increment of the sequence.

3 Analysis of the Single-Source Case

In the single source case, the model (1) can be written as

$$\mathbf{r}(t) = \mathbf{a}(\phi, \vartheta)\, s(t) + \mathbf{n}(t), \tag{4}$$

where $\mathbf{a}(\phi, \vartheta) = (a_1(\phi, \vartheta), ..., a_M(\phi, \vartheta))^T$,

$$a_m(\phi, \vartheta) = g_m(\phi, \vartheta)\, \mathrm{e}^{\mathrm{j}\frac{2\pi}{\lambda}(x_m \sin\vartheta\cos\phi + y_m \sin\vartheta\sin\phi + z_m\cos\vartheta)}, \tag{5}$$

$g_m(\phi, \vartheta)$ is the response, (x_m, y_m, z_m) are the coordinates of the m-th sensor, $\lambda = 2\pi c/\omega_0$ is the source wavelength, and c is the propagation speed of the plane wave.

Using some well-known properties of cumulants [6], it can be shown that

$$c_{k,l} = \gamma_{2,s}\, a_k(\phi, \vartheta)\, a_l^*(\phi, \vartheta) \quad \forall k \neq l \tag{6}$$

and

$$c_{k,l,m,n} = \gamma_{4,s}\, a_k(\phi, \vartheta)\, a_l^*(\phi, \vartheta)\, a_m^*(\phi, \vartheta)\, a_n(\phi, \vartheta) \tag{7}$$

if k, l, m, n are not equal at all, or, if they are equal, $n_k(t)$ must be Gaussian distributed. This latter condition is due to the fact that higher-order cumulants of Gaussian random variables are zero [6]. $\gamma_{2,s}$ is the second- and $\gamma_{4,s}$ is the fourth-order cumulant of the random variable $s(t)$.

Eq. (6) suggests estimating the steering vector, up to the scale factor $\gamma_{2,s}^2\, a_M^*(\phi, \vartheta)\, a_1(\phi, \vartheta) a_{M-1}^*(\phi, \vartheta)$, by

$$\hat{\mathbf{a}}^{(2)}(\phi, \vartheta) = \begin{pmatrix} \hat{c}_{1,M}\, \hat{c}_{1,M-1} \\ \hat{c}_{2,M}\, \hat{c}_{1,M-1} \\ \vdots \\ \hat{c}_{M-1,M}\, \hat{c}_{1,M-1} \\ \hat{c}_{1,M}\, \hat{c}_{M,M-1} \end{pmatrix}. \tag{8}$$

This can be easily proven by substituting $\hat{c}_{k,l}$ for $c_{k,l}$ and inserting (6). Similiarily, an estimator based on fourth-order cumulants is obtained from eq. (7), up to the scale factor $\gamma_{4,s}\, a_k(\phi, \vartheta)\, a_l^*(\phi, \vartheta) a_m^*(\phi, \vartheta)$, as

$$\hat{\mathbf{a}}^{(4)}_{k,l,m}(\phi, \vartheta) = \begin{pmatrix} \hat{c}_{k,l,m,1} \\ \vdots \\ \hat{c}_{k,l,m,M} \end{pmatrix}. \tag{9}$$

The only restriction on k, l, m in eq. (9) is that one of them must be unequal to another. This suggests collecting steering vector estimates $\hat{\mathbf{a}}^{(4)}_{k,l,m}(\phi, \vartheta)$ for as many different indices as possible, and using the mean of them to reduce the variance. However, inserting a realization of $\mathbf{r}(t)$ into eq. (9), and rewriting it as a function of $\mathbf{a}(\phi, \vartheta)$, yields

$$\hat{\mathbf{a}}^{(4)}_{k,l,m}(\phi, \vartheta) = f_{k,l,m}\, \mathbf{a}(\phi, \vartheta) + g_{k,l,m} \tag{10}$$

where $f_{k,l,m}$, $g_{k,l,m}$ are independent of $\mathbf{a}(\phi, \vartheta)$. Consequently, all these vectors are linearly dependent; thus,

in the following (k,l,m) is always fixed to $(3,2,1)$ and the indices of $\hat{\mathbf{a}}^{(4)}_{k,l,m}(\phi,\vartheta)$ will be discarded.

We mentioned earlier an *unbiased* estimator for $c_{k,l,m,n}$; therefore, $\hat{\mathbf{a}}^{(4)}(\phi,\vartheta)$ can also be estimated unbiased. In contrast, it is not possible to obtain an unbiased estimator for $\hat{\mathbf{a}}^{(2)}(\phi,\vartheta)$, because the bias

$$\mathrm{B}\{\hat{c}_{k,l}\ \hat{c}_{m,n}\} = \frac{1}{N}\left(\mathrm{E}\left\{r_k(t)r_l^*(t)r_m(t)r_n^*(t)\right\} - \mathrm{E}\left\{r_k(t)r_l^*(t)\right\}\mathrm{E}\left\{r_m(t)r_n^*(t)\right\}\right).$$

always remains a function of the fourth-order moment. This is the price paid for using the second-order based estimator.

4 Simulations and Analytic Evaluations

In this Section, we study the performance of the proposed steering vector estimators by Monte-Carlo simulation and analytic evaluation. In all the cases, $n_m(t)$ is Gaussian distributed for all $m = 1(1)M$, the number of realizations to obtain the simulated variance is fixed to 1000, the fourth-order cumulant indices are $(k,l,m) = (3,2,1)$, and one source with $\mathrm{E}\left\{|s(t)|^2\right\} = 1$ is always assumed $(P = 1)$. If nothing else is mentioned in the following, then $N = 20$, $M = 20$, $SNR = 0\mathrm{dB}$, a linear array of omnidirectional $(g_m(\phi,\vartheta) = 1)$ sensors with spacing $\lambda/2$ is assumed, $s(t)$ is a BPSK-signal, and the azimuth and elevation direction of arrival are $\phi = 0^o$, $\vartheta = 0^o$. Due to the inherent scale factor, all figures show the *normalized* variance and bias of $\hat{\mathbf{a}}^{(2)}$ and $\hat{\mathbf{a}}^{(4)}$. Note also that always the *mean*

$$\bar{V}^{(2)} = \frac{1}{M}\sum_{m=1}^{M}\mathrm{Var}\left\{\frac{\hat{a}_m^{(2)}}{a_M^* a_{M-1}^* a_1 \gamma_{2,s}^2}\right\}$$

$$\bar{B}^{(2)} = \frac{1}{M}\sum_{m=1}^{M}\left|\mathrm{B}\left\{\frac{\hat{a}_m^{(2)}}{a_M^* a_{M-1}^* a_1 \gamma_{2,s}^2}\right\}\right|^2$$

of the normalized variance and squared bias, respectively, will be plotted for easier comparison. $\bar{V}^{(4)}$ and $\bar{B}^{(4)}$ are defined equivalently. The estimated variance obtained from the Monte-Carlo simulations will be denoted as $\hat{\bar{V}}^{(2,4)}$. $\bar{B}^{(2)}$ and $\bar{V}^{(2)}$ will always be plotted as dashed lines, $\bar{B}^{(4)}$ and $\bar{V}^{(4)}$ as solid lines, $\hat{\bar{V}}^{(2)}$ as plus-signs, and $\hat{\bar{V}}^{(4)}$ as circles.

4.1 Example 1: Dependence on the number of samples N

One often mentioned deficiency of higher-order statistics is the increased number of data to achieve the same estimation accuracy obtained from second-order statistics; hence, an investigation of the number of samples N, which is chosen in this example as $N = 10(1)100$, is very interesting.

In Fig. 1, observe that $\hat{a}^{(4)}$ exhibits not only a lower mean-squared bias $(\bar{B}^{(4)} = 0 \forall N)$, but also a lower mean variance than $\hat{a}^{(2)}$. This is a surprising result, especially since it is also valid for very small N. The good agreement between the simulation and analytic evaluation confirms this result.

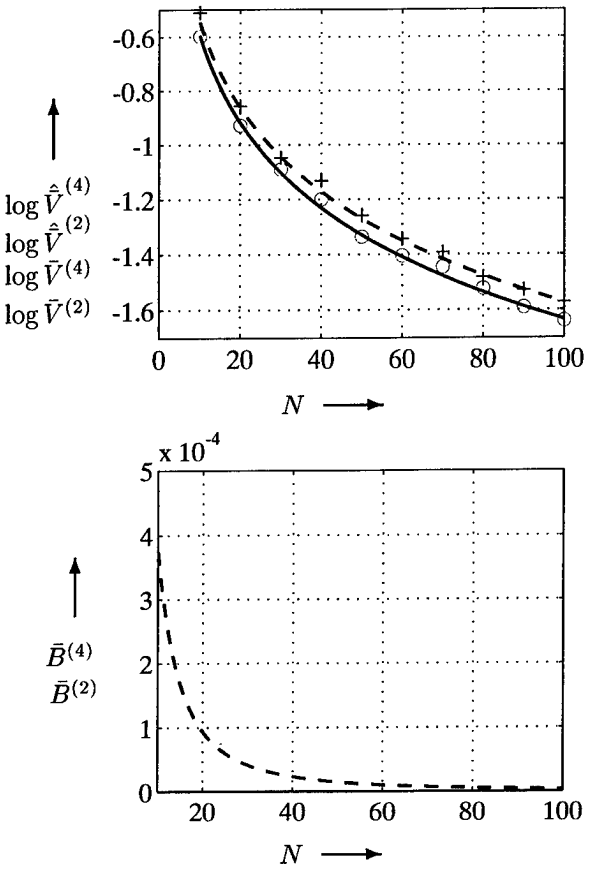


Figure 1.

4.2 Example 2: Dependence on the array shape

In this example we have varied the location $x_1 = 0$, $y_1/\lambda = 0(0.1)2$ of the first sensor of a (non)-linear array composed of $M = 5$ sensors. Since $\bar{B}^{(2)}$ is always very small compared to $\bar{V}^{(2)}$, it will be omitted in the following.

Observe from Fig. 2, that $\bar{V}^{(2)}$ and $\bar{V}^{(4)}$ are *independent* of the sensor location. Many additional examples with non-linear arrays, which are not shown here, confirm this interesting result. We will take up this property again in the next example. Note that this independence cannot be reliably recognized by $\hat{\bar{V}}^{(2)}$ and $\hat{\bar{V}}^{(4)}$. For this set of parameters, fourth-order sample cumulants are always more accurate estimators than second-order sample cumulants for *any* array shape.

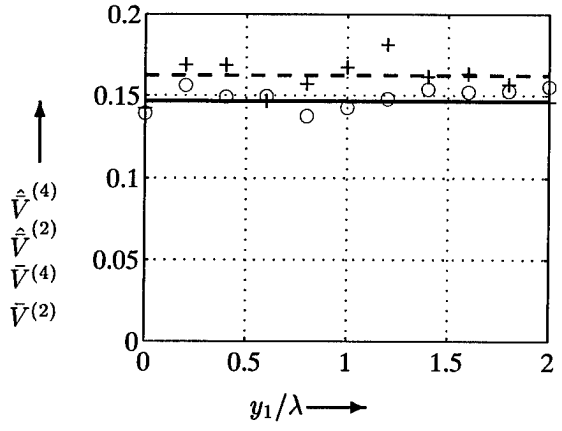


Figure 2.

4.3 Example 3: Dependence on the direction-of-arrival

Now the azimuth direction-of-arrival $\phi = -90^o(1^o)90^o$ is varied.

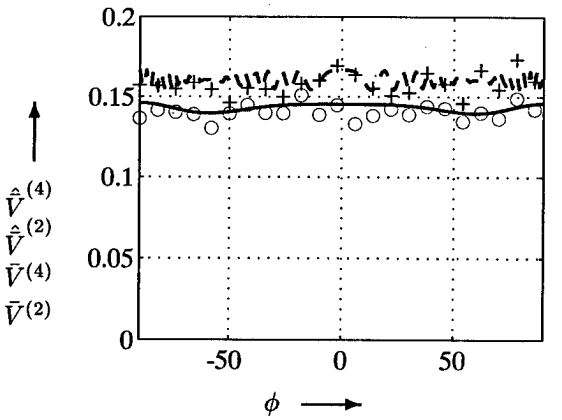


Figure 3.

It can be seen from Fig. 3, that $\bar{V}^{(2)}$ and $\bar{V}^{(4)}$ are only slightly dependent on the azimuth direction-of-arrival. Since the direction of arrivals and the sensor location only affects the angle of the steering vector elements for omnidirectional sensors (see eq. (5)), we investigate now the dependence of $\bar{V}^{(2)}$ and $\bar{V}^{(4)}$ on $|a_m(\phi, \vartheta)|$. For this purpose we consider a linear array composed of $M = 5$ dipoles with $g_m(\phi, \vartheta) = \sin(\phi - \alpha_m)$, where $(\alpha_1, ..., \alpha_M)^T = (0^o, 0^o, 5^o, -3^o, 10^o)^T$ is the orientation vector of the dipoles [5].

Observe, from Fig. 4, that here the fourth-order sample cumulants are not always more accurate estimators than second-order sample cumulants. Especially in the region around $\phi = 0^o$ where the dipoles are very insensitive, the $\bar{V}^{(2)}$ is considerably lower than $\bar{V}^{(4)}$. The dramatic increase in this region is due to the normalization, since the scaling factor is zero at $\phi = \alpha_1, ..., \alpha_M$.

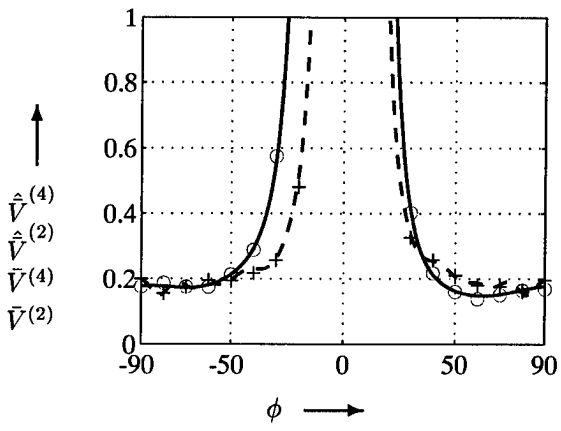


Figure 4.

A further example of practical interest is the behaviour of a linear array composed of $M = 20$ radar antennas with response [1] $g_m(\phi, \vartheta) = 1/\sin(\vartheta) \quad \forall \vartheta \in (20^o, 70^o)$.

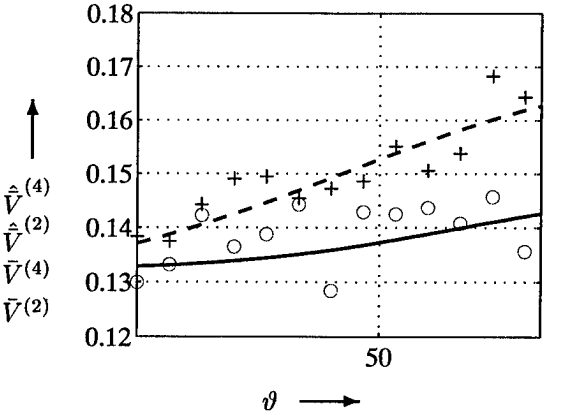


Figure 5.

The results in Fig. 5 are in contrast to the dipole array. Here, $\bar{V}^{(4)}$ is considerably lower for *any* elevation angle ϑ compared $\bar{V}^{(2)}$. The latter examples suggest that for direction-sensitive sensors fourth-order sample cumulants do not always lead to a higher estimation accuracy.

4.4 Example 4: Dependence on the number of sensors

In this example the number of sensors $M = 3(1)20$ is varied. Observe from Fig. 6 that $\bar{V}^{(4)}$ is always lower than $\bar{V}^{(2)}$ The difference between them is slightly increasing with increasing M.

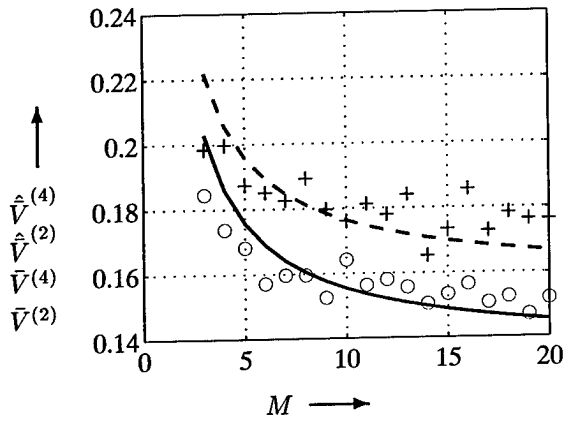


Figure 6.

4.5 Example 5: Dependence on SNR and kind of signals

Up to now we fixed the SNR to 0dB and assumed a BPSK-signal. *Yuen and Friedlander* [11] have noticed that the use of 4-QAM signals instead of BPSK signals decreases the performance of fourth-order cumulant based algorithms compared to second-order cumulant based algorithms; therefore, at this point we will not only vary the SNR but also the kind of signal. Fig. 7 shows the logarithm of $\bar{V}^{(2)}$, $\bar{V}^{(4)}$ for $SNR = -20$db(5db)40db and for BPSK-, 4-QAM-, and 64-QAM signals.

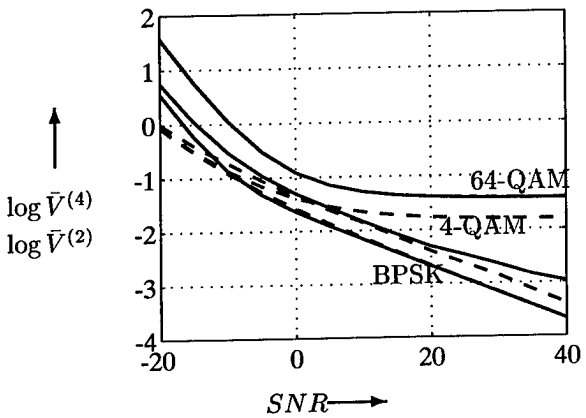


Figure 7.

For BPSK-signals (lower dashed and solid curves) fourth-order sample cumulants give somewhat better results than second-order sample cumulants for moderate and high SNR; however, the opposite is true for 4-QAM (middle dashed and solid curves) and especially for 64-QAM (upper dashed and solid curves) signals, which confirms the statement of [11] also in the small sample case.

5 Conclusions

In this paper we studied the finite-sample variance of steering vector estimates based on second- and fourth-order sample cumulants in the single source case. For BPSK-signals and moderate or high SNR's, fourth-order sample cumulants usually exhibit a lower finite-sample variance than second-order sample cumulants. This result is very general and also holds for other parameters, like the array shape, number of sensors, and especially the number of samples. On the other hand, for very low SNR-environments, or for 4-QAM and 64-QAM signals, second-order sample cumulants should be preferred. Of course, all these results are only valid for the conditions of the model defined in eq. (1), like, for example, the independence of the noise signals. Future work will investigate the finite-sample behaviour of VESPA for one source where noise-free second-order cumulants are calculated by using fourth-order cumulants. In addition, a combination of $\hat{a}^{(2)}$ and $\hat{a}^{(4)}$, e.g. the mean, will be analyzed and the multiple source case will be considered. As mentioned in the problem formulation, the model in eq. (1) can be extended to the near-field and coherent sources cases, which can also be investigated using the tools of this paper and [8], [9].

References

[1] A. Brunner, "Possibilities of Dimensioning Doubly Curved Reflectors for Azimuth-Search Radar Antennas," *IEEE Trans. on Ant. and Prop.*, vol. AP-19, no. 1, pp. 52-57, Jan. 1971.

[2] J. F. Cardoso and E. Moulines, "Asymptotic Performance Analysis of Direction-Finding Algorithms Based on Fourth-Order Cumulants," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 214-224, January 1995.

[3] R. N. Challa and S. Shamsunder, "Higher-Order subspace Algorithms for Passive Localization of Near-Field Sources," *Twenty-Ninth Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, pp. 377-381, October 30 - November 1, 1995.

[4] H. H. Chiang and C. L. Nikias, "The ESPRIT algorithm with higher-order statistics," *Proc. of Vail Workshop on Higher-Order Spectral Analysis*, pp. 163-168, Vail, CO, June 1989.

[5] M. C. Doğan and J. M. Mendel, "Applications of Cumulants to Array Processing," Part I: Aperture extension and Array Calibration", *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1200-1216, May 1995.

[6] M. C. Doğan and J. M. Mendel, "Cumulant-Based Blind Optimum Beamforming," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 722-741, July 1994.

[7] E. Gönen, M. C. Doğan and J. M. Mendel, "Applications of Cumulants to Array Processing: Direction-Finding in Coherent Signal Environment," *Twenty-Eighth Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, pp. 633-637, October 31-November 2, 1994.

[8] T. Kaiser and J. M. Mendel, "Finite-Sample Covariances of Second-, Third-, and Fourth-Order Sample Cumulants in Narrowband Array Processing," *Technical Report, Signal and Image Processing Institute, University of Southern California*, Los Angeles, September 1996.

[9] T. Kaiser and J. M. Mendel, "Covariance of Finite-Sample Cumulants in Array Processing," *IEEE Signal Processing Workshop on Higher-Order Statistics*, Banff, Canada, July 1997.

[10] A. Paulraj, R.Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proceedings of the IEEE*, vol. 74, no. 7, pp. 1044-1045, July 1986.

[11] N. Yuen and B. Friedlander, "Asymptotic Performance Analysis of ESPRIT, Higher-Order ESPRIT, and Virtual ESPRIT Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, October 1996.

IMPROVING THE THRESHOLD PERFORMANCE OF HIGHER-ORDER DIRECTION FINDING METHODS VIA PSEUDORANDOMLY GENERATED ESTIMATOR BANKS

Alex B. Gershman *Johann F. Böhme*

Signal Theory Group, Dept. of Electr. Eng.
Ruhr University, D-44780 Bochum, Germany
gsh@sth.ruhr-uni-bochum.de

ABSTRACT

Recently reported estimator bank approach [1] is extended below to the fourth-order direction finding algorithms. The essence of our approach is to exploit "parallel" underlying eigenstructure based estimators for removing the outliers and improving the direction finding performance in the threshold domain. The pseudorandomly generated weighted fourth-order MUSIC estimators are exploited as underlying techniques for estimator bank. Motivated by the superior performance and reduced computational complexity of beamspace and root modifications of the second-order eigenstructure techniques, beamspace root implementations of fourth-order MUSIC and fourth-order estimator bank are developed. Simulations show dramatical improvements of the threshold performance.

1. INTRODUCTION

The performances of direction finding algorithms are known to degrade as the Signal to Noise Ratio (SNR) goes down below a certain threshold or as a number of snapshots becomes small. This phenomenon referred to as *threshold effect*, is especially strong for higher-order methods [2]-[4]. Recently, a promising *estimator bank* approach [1], [5] allowing to lower the SNR threshold has been developed for second-order eigenstructure based algorithms. The algorithm proposed in [1], [5] has been referred to as Pseudo-Random Joint Estimation Strategy (PR-JES). The essence of this approach is to exploit the additional information arising when several *underlying* Direction Of Arrival (DOA) estimators are calculated simultaneously for a same batch of data (for single data record) [6]. To improve the performance, the evolutionary principles are used, i.e. PR-JES chooses and exploits most "successful" estimators (having no failure in the preliminary estimated source localization sectors) from the full set of underlying techniques.

In this paper, we extend this approach to the fourth-order direction finding algorithms based on *contracted quadricovariance* [2], [4]. Weighted fourth-order MUSIC estimators with pseudorandomly generated weighting matrix are introduced and exploited as underlying techniques for the fourth-order estimator bank. Motivated by the superior performance and reduced computational complexity of the second-order beamspace root eigenstructure techniques [7], beamspace root implementations of fourth-order MUSIC and fourth-order estimator bank are then developed. Simulations with QAM sources and unknown colored Gaussian noise demonstrate very significant improvements of the threshold performance relative to fourth-order MUSIC.

2. FOURTH-ORDER MUSIC AND WEIGHTED MUSIC

Consider a linear array of n sensors. Assume that there are $q < n$ narrowband stationary zero-mean mutually uncorrelated far-field sources. The ith array vector snapshot can be modeled as [3], [4]

$$\boldsymbol{x}(i) = \boldsymbol{A}\boldsymbol{s}(i) + \boldsymbol{n}(i), \quad i = 1, 2, \ldots, M \qquad (1)$$

where $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_q)]$ is the $n \times q$ direction matrix, $\theta_1, \theta_2, \ldots, \theta_q$ are the signal DOA's, $\boldsymbol{a}(\theta)$ is the $n \times 1$ steering vector, $\boldsymbol{s}(i)$ is the $q \times 1$ vector of random non-Gaussian circular source waveforms, and $\boldsymbol{n}(i)$ is the $n \times 1$ vector of unknown colored Gaussian sensor noise.

The DOA estimation problem considered below is that given the measurements $\boldsymbol{x}(i)$, $i = 1, 2, \ldots, M$, the signal DOA's $\theta_1, \theta_2, \ldots, \theta_q$ should be estimated. Define the second- and fourth-order moments [3], [4]

$$\mu_2(l, k) = [\boldsymbol{R}]_{lk} = \mathrm{E}\,[x_l(i)x_k^*(i)], \quad 1 \le l, k \le n \qquad (2)$$

This research was supported in parts by the SASPARC Project of INTAS, by the grant Bo 568/22-1 of DFG, and by Alexander von Humboldt Foundation.

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$\mu_4(l,k,m,p) = \mathrm{E}\left[x_l(i)x_k^*(i)x_m(i)x_p^*(i)\right], \quad 1 \le l,k,m,p \le n \tag{3}$$

where $\boldsymbol{R} = \mathrm{E}\{\boldsymbol{x}(i)\boldsymbol{x}^H(i)\}$ is the $n \times n$ covariance matrix, $(\cdot)^H$ and $(\cdot)^*$ denote the Hermitian transpose and complex conjugate, respectively. The consistent sample estimates of (2) and (3) are given by

$$\hat{\mu}_2(l,k) = [\hat{\boldsymbol{R}}]_{lk} = \frac{1}{M}\sum_{i=1}^{M} x_l(i)x_k^*(i) \tag{4}$$

$$\hat{\mu}_4(l,k,m,p) = \frac{1}{M}\sum_{i=1}^{M} x_l(i)x_k^*(i)x_m(i)x_p^*(i) \tag{5}$$

The quadricovariance is defined as a set of fourth-order cumulants [4] and can be written under circularity assumption as

$$\kappa_4(l,k,m,p) = \mu_4(l,k,m,p) - \mu_2(l,k)\mu_2(m,p) - \mu_2(l,p)\mu_2(m,k), \tag{6}$$

The consistent estimate of (6) is given by

$$\hat{\kappa}_4(l,k,m,p) = \hat{\mu}_4(l,k,m,p) - \hat{\mu}_2(l,k)\hat{\mu}_2(m,p) - \hat{\mu}_2(l,p)\hat{\mu}_2(m,k) \tag{7}$$

The contracted quadricovariance is defined as [2], [4]

$$c_{lk} = \sum_{m=1}^{n} \kappa_4(l,k,m,m), \tag{8}$$

Using (6), rewrite (8) in matrix notation as [4]

$$\boldsymbol{C} = \mathrm{E}\left[\boldsymbol{x}^H(i)\boldsymbol{x}(i)\,\boldsymbol{x}(i)\boldsymbol{x}^H(i)\right] - \mathrm{Tr}\,(\boldsymbol{R})\,\boldsymbol{R} - \boldsymbol{R}^2 \tag{9}$$

The sample estimate of (9) is given by

$$\hat{\boldsymbol{C}} = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{x}^H(i)\boldsymbol{x}(i)\,\boldsymbol{x}(i)\boldsymbol{x}^H(i) - \mathrm{Tr}\,(\hat{\boldsymbol{R}})\,\hat{\boldsymbol{R}} - \hat{\boldsymbol{R}}^2 \tag{10}$$

It can be shown [4] that (9) can be represented as

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{Z}\boldsymbol{A}^H \tag{11}$$

where $\boldsymbol{Z}$ is a $q \times q$ Hermitian matrix. According to (11), the structure of $\boldsymbol{C}$ is similar to that of the noise-free covariance matrix in the second-order case [4]. Hence, the conventional MUSIC algorithm can be exploited for DOA estimation. Express the eigendecomposition of (10) as

$$\hat{\boldsymbol{C}} = \sum_{i=1}^{n} \hat{\lambda}_i \hat{\boldsymbol{e}}_i \hat{\boldsymbol{e}}_i^H = \hat{\boldsymbol{E}}_S \hat{\boldsymbol{\Lambda}}_S \hat{\boldsymbol{E}}_S^H \tag{12}$$

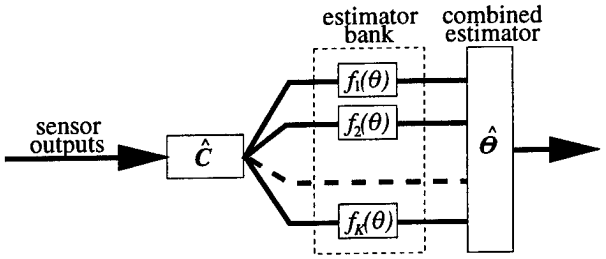


Figure 1: Concept of estimator bank.

where the $n \times \hat{q}$ and $n \times (n-\hat{q})$ matrices $\hat{\boldsymbol{E}}_S$ and $\hat{\boldsymbol{E}}_N$ contain the sample signal and noise subspace eigenvectors, respectively, and $\hat{q}$ is an estimate of q. In turn, the $\hat{q} \times \hat{q}$ and $(n-\hat{q}) \times (n-\hat{q})$ matrices $\hat{\boldsymbol{\Lambda}}_S$ and $\hat{\boldsymbol{\Lambda}}_N$ contain the signal and noise subspace sample eigenvalues, respectively. The contracted quadricovariance based MUSIC [2], [4] estimates the source DOA's as locations of $\hat{q}$ highest peaks of spectral function

$$f_{\text{CQ-MUSIC}}(\theta) = [\boldsymbol{a}^H(\theta)\hat{\boldsymbol{E}}_N\hat{\boldsymbol{E}}_N^H\boldsymbol{a}(\theta)]^{-1} \tag{13}$$

Weighted contracted quadricovariance MUSIC estimator can be defined as [5]

$$f_{\boldsymbol{W}}^{\text{CQ}}(\theta) = [\boldsymbol{a}^H(\theta)\hat{\boldsymbol{E}}_N\boldsymbol{W}\hat{\boldsymbol{E}}_N^H\boldsymbol{a}(\theta)]^{-1} \tag{14}$$

where $\boldsymbol{W}$ is the $(n-\hat{q}) \times (n-\hat{q})$ nonnegative definite weighting matrix.

3. ESTIMATOR BANKS AND PR-JES

The essence of the estimator bank approach [1], [5] is to generate multiple "independent" estimators for the same batch of data snapshots. To improve the performance, the evolutionary principles are used, i.e. the underlying estimators are involved into a concurrence and only most "successful" estimators (having no failure in the preliminary estimated source localization sectors) are used while all failed estimators are attributed to outliers and removed from the further consideration. The results of "successful" estimators are then combined in the final estimate as shown in Fig. 1. Assume that arbitrary underlying estimators $f_l(\theta)$, $l = 1,2,\ldots,K$ are calculated in a parallel manner, using the single batch of array data $\boldsymbol{x}(i)$, $i = 1,2,\ldots,M$, or, equivalently, using the single estimate $\hat{\boldsymbol{C}}$ of the contracted quadricovariance matrix. Similarly to the second-order case, define the fourth-order estimator bank as [5]

$$\mathcal{F} = \{f_l(\theta),\ l = 1,2,\ldots,K\}_{\hat{\boldsymbol{C}}} \tag{15}$$

It is suitable to generate the underlying estimators in a pseudorandom manner, using the set of weighted MUSIC estimators (14) with rank-one weighting matrices

$\boldsymbol{W} = \boldsymbol{w}\boldsymbol{w}^H$ withdrawn from the Gaussian random generator

$$\boldsymbol{w}_l \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}), \quad l = 1, 2, \ldots, K \tag{16}$$

where $\mathrm{E}[\boldsymbol{w}_l \boldsymbol{w}_k^H] = \delta_{lk}\boldsymbol{I}$, $\mathrm{E}[\boldsymbol{w}_l \boldsymbol{w}_k^T] = 0$, $\boldsymbol{I}$ is the identity matrix, and $(\cdot)^T$ denotes the transpose. Using (16), as many "independent" estimators as necessary can be generated to satisfy the required compromise between computational cost and threshold performance. A choice of rank-one weighting matrices for (14) is dictated by computational reasons [5].

Define the following hypothesis which is formulated for an arbitrary DOA estimator:

$\mathcal{H}$: *The estimator spectral function has more than* $\hat{q} - 1$ *separate spectral peaks localized in* Θ_S.

Here, Θ_S are the preliminary estimated angular sectors of source localization. Assume that the source localization sectors are specified as L non-overlapping intervals

$$\Theta_S = [\theta_{a,1}, \theta_{b,1}] \cup [\theta_{a,2}, \theta_{b,2}] \cup \cdots \cup [\theta_{a,L}, \theta_{b,L}] \tag{17}$$

and that the estimate of the number of sources $\hat{q}$ is known. The PR-JES technique [1], [5] can be written for given sample contracted quadricovariance matrix $\hat{\boldsymbol{C}}$ and K pseudo-randomly generated underlying estimators as the following sequence of steps:

Step 1: Calculate the MUSIC estimator (13) and test the hypothesis $\mathcal{H}$ for this estimator alone. If $\mathcal{H}$ is accepted for this estimator then estimate the source DOA's as the locations of $\hat{q}$ highest peaks of its spectral function (13), and terminate the algorithm (i.e., go to step 4). If $\mathcal{H}$ is not accepted for the MUSIC estimator then go to the next step.

Step 2: Generate K different random vectors $\boldsymbol{w}_l$, $l = 1, 2, \ldots, K$ using the random generator (16) and calculate the underlying DOA estimators (14). Denote their spectral functions $f_l(\theta)$, $l = 1, 2, \ldots, K$. As a result of this step, the estimator bank (15) is completed.

Step 3: Test the hypothesis $\mathcal{H}$ for each DOA estimator from the estimator bank.

If $\mathcal{H}$ is accepted for any J $(0 < J \leq K)$ estimators $\tilde{f}_l(\theta)$, $l = 1, 2, \ldots, J$ from the total number of K estimators $f_l(\theta)$, $l = 1, 2, \ldots, K$ then estimate the k-th DOA θ_k as

$$\hat{\theta}_k = \mathrm{med}\,\{\tilde{\theta}_k^{(1)}, \tilde{\theta}_k^{(2)}, \ldots, \tilde{\theta}_k^{(J)}\}, \quad k = 1, 2, \ldots, \hat{q} \tag{18}$$

where $\tilde{\theta}_1^{(l)} < \tilde{\theta}_2^{(l)} < \cdots < \tilde{\theta}_{\hat{q}}^{(l)}$ is the ordered set of angles, corresponding to the $\hat{q}$ highest peaks of function $\tilde{f}_l(\theta)$ that are localized in Θ_S, and

$$\mathrm{med}\,\{b_1, \ldots, b_m\} = \begin{cases} (c_{\frac{m}{2}} + c_{\frac{m}{2}+1})/2, & \text{even } m \\ c_{\frac{m+1}{2}}, & \text{odd } m \end{cases} \tag{19}$$

Here $\{c_1, \ldots, c_m\} = \mathrm{sort}\,\{b_1, \ldots, b_m\}$, and $\mathrm{sort}\,\{\cdots\}$ denotes the operator of sorting in ascending (descending) order.

If $\mathcal{H}$ is not accepted for all DOA estimators from the estimator bank then estimate the k-th DOA θ_k as

$$\hat{\theta}_k = \mathrm{med}\,\{\theta_k^{(1)}, \theta_k^{(2)}, \ldots, \theta_k^{(K)}\}, \quad k = 1, 2, \ldots, \hat{q} \tag{20}$$

where $\theta_1^{(l)} < \theta_2^{(l)} < \cdots < \theta_{\hat{q}}^{(l)}$ is the ordered set of angles, corresponding to the $\hat{q}$ highest peaks of function $f_l(\theta)$ that are localized in the whole array field of view $[-90°, 90°]$.

Step 4: Stop. □

Testing of $\mathcal{H}$ is used for detecting failures in the underlying estimates. It helps to remove the outliers and groups "successful" estimators. Step 1 guarantees the asymptotic behaviour of PR-JES to be not worse than that of MUSIC. I.e., if MUSIC is decided to have no failure within Θ_S then no more processing is necessary because MUSIC is known to provide excellent asymptotic performance. Otherwise, the estimator bank is used for lowering the SNR threshold.

The role of sector information (17) is very important. Each separate sector may include several unresolved sources without affecting the performance of the following processing steps. However, the performance of PR-JES may degrade if some of the true DOA's are not included in Θ_S or if these sectors are inadequately wide compared to the real source clusters. Note that the similar sector information is used in beamspace processing and the similar problems arise if this information is inadequate.

4. BEAMSPACE ROOT IMPLEMENTATION

Motivated by the success of the previous research on beamspace and root high-resolution methods and on combining of beamspace and root algorithms into one scheme [7], let us develop a beamspace root implementation of fourth-order PR-JES for a Uniform Linear Array (ULA). The importance of beamspace root implementation can be motivated by the following: i) polynomial rooting will provide significant computational savings compared with spectral search, ii) signal concentration in beamspace and polynomial rooting are expected to further improve the threshold and asymptotic performances and robustness against mismodeling [7], iii) the preliminary knowledge of source localization sectors can be simultaneously exploited for the hypothesis testing procedure in PR-JES, as well as for the design of beamspace transformation matrix.

Taking into account the analogy between contracted quadricovariance and covariance matrices, it is logical to apply the beamspace transformation directly to the sample contracted quadricovariance matrix (10). Hence, the $p \times p$ sample beamspace contracted quadricovariance matrix can be introduced as

$$\hat{\boldsymbol{C}}_{\mathrm{B}} = \boldsymbol{T}^H \hat{\boldsymbol{C}} \boldsymbol{T} \tag{21}$$

where $\boldsymbol{T}$ is the $n \times p$ $(q < p \leq n)$ beamspace transformation matrix satisfying $\boldsymbol{T}^H\boldsymbol{T} = \boldsymbol{I}$. Write the eigendecomposition of (21) as

$$\hat{\boldsymbol{C}}_{\mathrm{B}} = \hat{\boldsymbol{U}}_S\hat{\boldsymbol{\Gamma}}_S\hat{\boldsymbol{U}}_S^H + \hat{\boldsymbol{U}}_N\hat{\boldsymbol{\Gamma}}_N\hat{\boldsymbol{U}}_N^H \qquad (22)$$

where the $\hat{q} \times \hat{q}$ and $(p-\hat{q}) \times (p-\hat{q})$ diagonal matrices $\hat{\boldsymbol{\Gamma}}_S$ and $\hat{\boldsymbol{\Gamma}}_N$ contain the signal and noise eigenvalues, and the columns of the $p \times \hat{q}$ and $p \times (p-\hat{q})$ matrices $\hat{\boldsymbol{U}}_S$ and $\hat{\boldsymbol{U}}_N$ contain the signal and noise eigenvectors, respectively. Using the analogy between spectral and root estimators and between element and beam spaces, the beamspace root version of the contracted quadricovariance based MUSIC estimator (13) can be expressed as

$$f_{\mathrm{CQ-MUSIC}}(z) = \boldsymbol{a}^T(z^{-1})\boldsymbol{T}\hat{\boldsymbol{U}}_N\hat{\boldsymbol{U}}_N^H\boldsymbol{T}^H\boldsymbol{a}(z) \qquad (23)$$

where $\boldsymbol{a}(z) = (1, z, \ldots, z^{n-1})^T$. The signal DOA's can be estimated from the closest to the unit circle roots of the polynomial (23). In turn, we introduce the beamspace root version of the weighted MUSIC estimator (14) as

$$f_{\boldsymbol{W}}^{\mathrm{CQ}}(z) = \boldsymbol{a}^T(z^{-1})\boldsymbol{T}\hat{\boldsymbol{U}}_N\boldsymbol{W}\hat{\boldsymbol{U}}_N^H\boldsymbol{T}^H\boldsymbol{a}(z) \qquad (24)$$

where $\boldsymbol{W}$ is the $(p-\hat{q}) \times (p-\hat{q})$ weighting matrix.

Consider a general case of multiple sectors (multiple source clusters). In this case, different beamspace transformation matrices are required for each sector, and PR-JES must be applied to the underlying polynomial sets rather than to underlying estimators as in Section 3. I.e., each underlying estimator in the estimator bank is represented now as a polynomial set of the dimension L, where L is the total number of angular sectors in (17), and where polynomials from the same set differ by the beamspace transformation matrix $\boldsymbol{T}$.

Consider first the polynomial set $\{f_i(z)\}_{i=1}^L$ associated with arbitrary root estimator $f(z)$. Let the ith polynomial $f_i(z)$ from this set has N_i roots (for exactness, hereafter we consider only the roots lying inside the unit circle) associated with the angles localized within arbitrary interval $[\theta_{a,i}, \theta_{b,i}]$. The following hypothesis will be used in the sequel for sorting out the "successful" estimators:

$$\mathcal{H}: \quad \sum_{i=1}^{L} N_i \geq \hat{q} \qquad (25)$$

Now, we are able to formulate the beamspace root implementation of PR-JES which can be represented as the following sequence of steps.

Step 1: For each interval $[\theta_{a,i}, \theta_{b,i}]$ compute the matrix $\boldsymbol{T}$ and the non-weighted beamspace MUSIC polynomial (23). Denote this polynomial $f_i(z)$. As a result of this step, L different polynomials $f_i(z)$, $i = 1, 2, \ldots, L$ are available for different intervals $[\theta_{a,i}, \theta_{b,i}]$, $i = 1, 2, \ldots, L$.

Step 2: For each interval $[\theta_{a,i}, \theta_{b,i}]$, $i = 1, 2, \ldots, L$, find the roots $\{z_{i,1}, z_{i,2}, \ldots, z_{i,N_i}\}$ of $f_i(z)$ associated with the angles localized within $[\theta_{a,i}, \theta_{b,i}]$. Test the hypothesis (25). If (25) is accepted then estimate source DOA's from the $\hat{q}$ closest to the unit circle roots selected from the roots $\{z_{1,1}, z_{1,2}, \ldots, z_{1,N_1}, z_{2,1}, z_{2,2}, \ldots, z_{2,N_2}, \ldots, z_{L,1}, z_{L,2}, \ldots, z_{L,N_L}\}$ and stop (go to step 5). If (25) is not accepted then go to the next step.

Step 3: Generate $\boldsymbol{w}_l$, $l = 1, 2, \ldots, K$ using (16) and compute K underlying polynomials (24) for each interval $[\theta_{a,i}, \theta_{b,i}]$ using the previously computed beamspace transformation matrices for these intervals. Denote these polynomials $f_i^{(l)}(z)$, $i = 1, 2, \ldots, L$, $l = 1, 2, \ldots, K$. As a result of this step, K polynomial sets $\{f_i^{(l)}(z)\}_{i=1}^L$, $l = 1, 2, \ldots, K$ are available.

Step 4: For each interval $[\theta_{a,i}, \theta_{b,i}]$, $i = 1, 2, \ldots, L$, and each polynomial $f_i^{(l)}(z)$ corresponding to this interval, find the roots $\{z_{i,1}^{(l)}, z_{i,2}^{(l)}, \ldots, z_{i,N_i^{(l)}}^{(l)}\}$ associated with the angles localized within $[\theta_{a,i}, \theta_{b,i}]$. Here $N_i^{(l)}$ is the number of such roots of the polynomial $f_i^{(l)}(z)$. For each polynomial set $\{f_i^{(l)}(z)\}_{i=1}^L$, set $N_i = N_i^{(l)}$ and test (25). If (25) is accepted for any J $(0 < J \leq K)$ polynomial sets $\{\tilde{f}_i^{(l)}(z)\}_{i=1}^L$, $l = 1, 2, \ldots, J$ from the total number of K polynomial sets $\{f_i^{(l)}(z)\}_{i=1}^L$, $l = 1, 2, \ldots, K$ then for each polynomial $\tilde{f}_i^{(l)}(z)$ find the roots $\{\tilde{z}_{i,1}^{(l)}, \tilde{z}_{i,2}^{(l)}, \ldots, \tilde{z}_{i,\tilde{N}_i^{(l)}}^{(l)}\}$ associated with the angles localized within $[\theta_{a,i}, \theta_{b,i}]$. Here $\tilde{N}_i^{(l)}$ is the number of such roots of the polynomial $\tilde{f}_i^{(l)}(z)$. Estimate the kth DOA using (18) where $\tilde{\theta}_1^{(l)} < \tilde{\theta}_2^{(l)} < \cdots < \tilde{\theta}_{\hat{q}}^{(l)}$ is the ordered set of angles associated with the $\hat{q}$ closest to the unit circle roots selected from the roots $\{\tilde{z}_{1,1}^{(l)}, \tilde{z}_{1,2}^{(l)}, \ldots, \tilde{z}_{1,\tilde{N}_1^{(l)}}^{(l)}, \tilde{z}_{2,1}^{(l)}, \tilde{z}_{2,2}^{(l)}, \ldots, \tilde{z}_{2,\tilde{N}_2^{(l)}}^{(l)}, \ldots, \tilde{z}_{L,1}^{(l)}, \tilde{z}_{L,2}^{(l)}, \ldots, \tilde{z}_{L,\tilde{N}_L^{(l)}}^{(l)}\}$. If (25) is not accepted for all K polynomial sets $\{f_i^{(l)}(z)\}_{i=1}^L$, $l = 1, 2, \ldots, K$ then estimate the kth DOA using (20) where $\theta_1^{(l)} < \theta_2^{(l)} < \cdots < \theta_{\hat{q}}^{(l)}$ is the ordered set of angles localized in the whole array field of view $[-90°, 90°]$, and associated with the $\hat{q}$ closest to the unit circle roots selected from the overall number of roots of the polynomial set $\{f_i^{(l)}(z)\}_{i=1}^L$.

Step 5: Stop. □

5. SIMULATION RESULTS

In our simulations, a ULA of ten omnidirectional sensors with half-wavelength spacing and two 4-QAM equal-power far-field sources impinging from $\theta_1 = 10°$, $\theta_2 = 12.5°$ were assumed. The number of independent

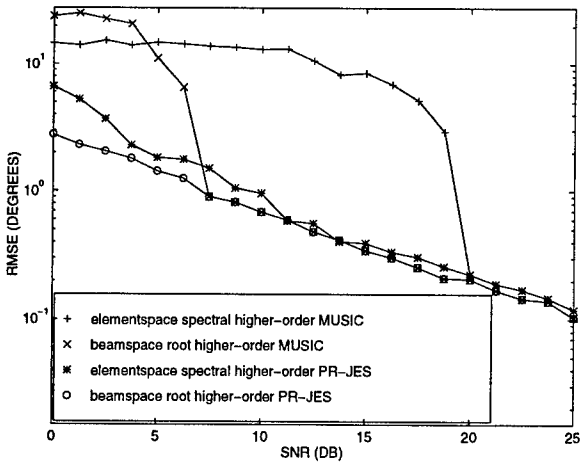


Figure 2: Experimental RMSE's of DOA estimation versus SNR. $K = 20$.

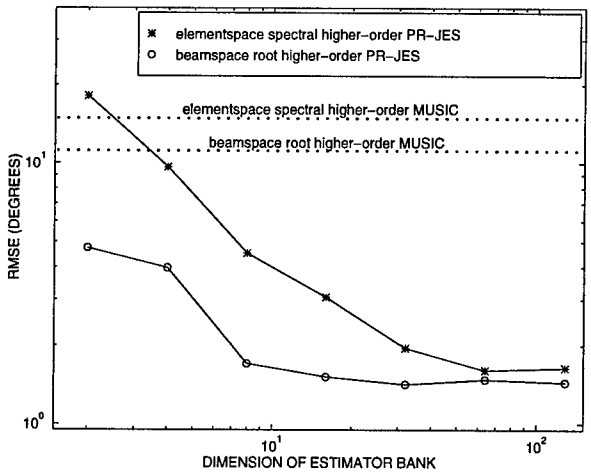


Figure 3: Experimental RMSE's of DOA estimation versus K. SNR = 5 dB.

snapshots in each simulation run was $M = 100$. A total of 100 independent runs were performed to obtain each experimental Root Mean Square Error (RMSE) value for simulated DOA estimators. The results were additionally averaged over the number of sources. When modeling PR-JES, the randomly generated vectors (16) were renewed in each simulation run (in order to examine the performance averaged over the random choice of weighting vector). The SNR was defined as the ratio of signal-to-noise powers in a single sensor. In our simulations, we always assumed that $\hat{q} = q$ and $\Theta_S = [3.5, 19.0]$. Unknown spatially colored Gaussian noise was modeled as a "spatially autoregressive" process [3] with the covariance matrix $[\boldsymbol{Q}]_{lk} = \rho^{|l-k|}$, where $\rho = 0.8$. The beamspace dimension $p = 7$ was assumed. The matrix $\boldsymbol{T}$ was designed within the sector Θ_S using the so-called "spheroidal sequences" approach [8].

Figs. 2 shows the DOA estimation RMSE's versus SNR. Elementspace spectral fourth-order MUSIC and PR-JES are compared with their beamspace root implementations. In this figure, the dimension of estimator bank was fixed ($K = 20$). Fig. 3 compares the RMSE's of these fourth-order methods versus the dimension of estimator bank K. In this figure, the SNR was fixed and equal to 5 dB. Both figures demonstrate that PR-JES performs dramatically better than MUSIC in the threshold domain. Interestingly, for known q our technique allows to avoid the abrupt threshold phenomenon at all. Fig. 2 verifies that the asymptotic (high SNR) performances of PR-JES and MUSIC are similar. From Fig. 3 one can observe that the performance of PR-JES can be improved via increasing the dimension of estimator bank. As expected, the beamspace root implementation of fourth-order MUSIC performs better in the threshold and asymptotic domains than elementspace spectral fourth-order MUSIC. Similarly, beamspace root PR-JES has better performance than elementspace spectral PR-JES.

REFERENCES

[1] A.B. Gershman, J.F. Böhme, "Improved DOA estimation via pseudorandom resampling of spatial spectrum," *IEEE SP Lett.*, vol. 4, pp. 54-57, Feb. 1997.

[2] R. Pan and C. Nikias, "Harmonic decomposition in cumulant domains," in *Proc. ICASSP'88*, New York, pp. 2356-2359, April 1988.

[3] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics," *IEEE Trans. SP*, vol. 39, pp. 2016-2023, Sept. 1991.

[4] J.-F. Cardoso and E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants," *IEEE Trans. SP*, vol. 43, pp. 214-224, Jan. 1995.

[5] A.B. Gershman, "Pseudo-randomly generated estimator banks: a new tool for improving the threshold performance of direction finding," *IEEE Trans. SP*, in review.

[6] A.B. Gershman, "High-resolution direction finding using eigendecomposition based estimators and joint estimation strategy," *Electronics Lett.*, vol. 27, 1991, pp. 2308–2309.

[7] M.D. Zoltowski, G.M. Kautz, and S.D. Silverstein, "Beamspace root-MUSIC," *IEEE Trans. SP*, vol. 41, pp. 344-364, Jan. 1993.

[8] P. Forster and G. Vezzosi, "Application of spheroidal sequences to array processing," in *Proc. ICASSP'87*, Dallas, TX, pp. 2268-2271, May 1987.

MUSIC's and Cramér-Rao Bound in Fourth-Order Cumulant Domain

Huan Wu, Zheng Bao and Kehu Yang
Institute of Electronics Engineering, Xidian University
Xi'an, Shaanxi 710071, P. R. China
wyz01@xidian.edu.cn

Abstract

A unifying asymptotic performance analysis of a class of MUSIC algorithms for direction-of-arrival (DOA) estimation in fourth-order cumulant domain (FOCD-MUSIC) is presented in this paper. A simple and explicit formula for the asymptotic variances of DOA estimation by FOCD-MUSIC's is given. The Cramér-Rao bound for DOA estimation in fourth-order cumulant domain (FOCD-CRB) is also derived. The performances of three typical FOCD-MUSIC's and the conventional covariance-based MUSIC are compared. It is shown that the FOCD-MUSIC's are inefficient and they are not superior to the conventional MUSIC algorithm in any case. Nevertheless, the FOCD-MUSIC's outperform the conventional MUSIC with reduced variances and improved robustness when the spatial sources are closely spaced and the signal-to-noise ratios (SNR's) are relatively low. Simulations are included to validate the analytical results.

1. Introduction

In recent years, DOA estimation algorithms based on higher-order cumulants have drawn a lot of attention [3-9] due to their ability to improve the robustness of the covariance-based techniques. Pan and Nikias [3] were the first to report on the fourth-order cumulant-based MUSIC [2]. They used the diagonal slice to form a spatial cumulant matrix. Porat and Friedlander [4] used all spatial fourth-order cumulants to develop a MUSIC-like algorithm. Moulines and Cardoso [5] proposed the contracted quadricovariance by smoothing part of the spatial fourth-order cumulants. Chen and Lin [6] presented a class of fourth-order cumulant matrices to applied to MUSIC. In [7] we developed a direction finding algorithm based on the maximal set of nonredundant fourth-order cumulants (MSNC algorithm). In [8] we proposed a Toeplitz approximation method in fourth-order cumulant domain (CTAM algorithm). The common point of the above algorithms is that the true (infinite snapshots) cumulant matrices are similar to the true noise-free covariance matrix. The differences of them, which result in different performance, rely on the way to compose the estimated cumulant matrices. We notice that these algorithms can be put into one category, that is, MUSIC algorithms in fourth-order cumulant domain (FOCD-MUSIC's).

Few FOCD-MUSIC's have been evaluated analytically. Moulines and Cardoso [5] made an asymptotic analysis of the algorithms in [3] and [5]. But the process and result of their analysis, compared to the those presented in this paper, are very complicated. Fan and Younan [9] also made a statistical analysis of the algorithm in [3]. Their analysis is based on an assumption that the sample cumulant errors due to finite data length are uncorrelated random variables with identical variance. We point out in this paper that this assumption is generally *not* true, especially when the spatial emitters are spaced closely (e.g., less than half of the beamwidth).

In this paper, we present a unifying and explicit analysis of the class of FOCD-MUSIC's. The Cramér-Rao lower bound for DOA estimation in fourth-order cumulant domain (FOCD-CRB) is also derived. Both the analysis and the derivation are based on the statistical characterization of the sample fourth-order cumulants, which is also presented in this paper.

2. Problem Formulation

Assume P narrowband plane waves are incident on an arbitrary array of M ($M>P$) sensors. The output of the mth sensor of the array can be written by

$$y_m(t)=\sum_{p=1}^{P} s_p(t)v_m(\theta_p)+w_m(t)\,,\ t=1,\ldots,N\,, \quad (1)$$

where N is the number of snapshots, $s_p(t)$ is the complex envelope, $p = 1, ..., P$; $w_m(t)$ is the additive noise and $v_m(\theta)=\exp\{j\vec{k}_\theta\cdot\vec{r}_m\}$, $\vec{k}_\theta$ and $\vec{r}_m$ are wave number and sensor location vectors, $m = 1, ..., M$. We make the follow-

This work has been supported by the Natural Science Foundation of China under Grant 69602008.

0-8186-8005-9/97 $10.00 © 1997 IEEE

ing assumptions on the array outputs. (AS1): $s_p(t)$'s are zero mean stationary non-Gaussian process with non-zero kurtosis; (AS2): $w_m(t)$'s are zero mean circular Gaussian process; (AS3): $s_p(t)$'s and $w_m(t)$'s are statistically independent themselves and of each other. Based on these assumptions and the properties of cumulants, the true second- to eighth-order cumulants can be easily derived as follows:

$$R_y(k,m) = \sum_{p=1}^{P} \gamma_{2,p} v_k(\theta_p) v_m^*(\theta_p) + R_w(k,m) \quad (2)$$

$$C_{4y}(k_1,k_2,m_1,m_2) = \sum_{p=1}^{P} \gamma_{4,p} v_{k_1}(\theta_p) v_{k_2}(\theta_p) v_{m_1}^*(\theta_p) v_{m_2}^*(\theta_p) \quad (3)$$

$$C_{6y}(k_1 \cdots k_3, m_1 \cdots m_3) = \sum_{p=1}^{P} \gamma_{6,p} v_{k_1}(\theta_p) \cdots v_{k_3}(\theta_p) v_{m_1}^*(\theta_p) \cdots v_{m_3}^*(\theta_p) \quad (4)$$

$$C_{8y}(k_1 \cdots k_4, m_1 \cdots m_4) = \sum_{p=1}^{P} \gamma_{8,p} v_{k_1}(\theta_p) \cdots v_{k_4}(\theta_p) v_{m_1}^*(\theta_p) \cdots v_{m_4}^*(\theta_p) \quad (5)$$

The sample fourth-order cumulant can be estimated by

$$\hat{C}_{4y}^N(k_1,k_2,m_1,m_2) = \frac{1}{N}\sum_{t=1}^{N} y_{k_1}(t) y_{k_2}(t) y_{m_1}^*(t) y_{m_2}^*(t) + \frac{1}{N^2}\sum_{t,u=1}^{N} [\, y_{k_1}(t) y_{m_1}^*(t) y_{k_2}(u) y_{m_2}^*(u) + y_{k_1}(t) y_{m_2}^*(t) y_{k_2}(u) y_{m_1}^*(u) \,] \quad (6)$$

3. Asymptotic Characterization of Sample Fourth-Order Cumulants

The sample cumulant error due to finite snapshots is defined by

$$\Delta C_{4y}^N(k_1,k_2,m_1,m_2) = \hat{C}_{4y}^N(k_1,k_2,m_1,m_2) - C_{4y}(k_1,k_2,m_1,m_2) \quad (7)$$

Theorem 1. Under (AS1)-(AS3), the estimate of fourth-order cumulant (6) is asymptotically unbiased. The asymptotic covariance of the sample cumulants is given by

$$\lim_{N\to\infty} N \cdot \mathrm{E}\{\Delta C_{4y}^N(k_1,k_2,m_1,m_2)\Delta C_{4y}^N(k_3,k_4,m_3,m_4)\} = G_8 + G_{62} + G_{44} + G_{422} + G_{2222} \quad (8a)$$

that is

$$\mathrm{AE}\{\Delta C_{4y}^N(k_1,k_2,m_1,m_2)\Delta C_{4y}^N(k_3,k_4,m_3,m_4)\} = (G_8 + G_{62} + G_{44} + G_{422} + G_{2222})/N \quad (8b)$$

where "AE" denotes asymptotic mean; G_8, G_{62}, G_{44}, G_{422}, and G_{2222} are functions of cumulants:

$G_8 = C_{8y}(k_1,k_2,k_3,k_4,m_1,m_2,m_3,m_4)$;

$G_{62} =$

$C_{6y}(k_1,k_2,k_3,m_1,m_3,m_4)R_y(k_4,m_2) + C_{6y}(k_1,k_2,k_3,m_2,m_3,m_4)R_y(k_4,m_1)$
$+C_{6y}(k_1,k_2,k_4,m_1,m_3,m_4)R_y(k_3,m_2) + C_{6y}(k_1,k_2,k_4,m_2,m_3,m_4)R_y(k_3,m_1)$
$+C_{6y}(k_1,k_3,k_4,m_1,m_2,m_3)R_y(k_2,m_4) + C_{6y}(k_1,k_3,k_4,m_1,m_2,m_4)R_y(k_2,m_3)$
$+C_{6y}(k_2,k_3,k_4,m_1,m_2,m_3)R_y(k_1,m_4) + C_{6y}(k_2,k_3,k_4,m_1,m_2,m_4)R_y(k_1,m_3)$;

$G_{44} = C_{4y}(k_1,k_2,m_1,m_3)C_{4y}(k_3,k_4,m_2,m_4)$
$+C_{4y}(k_1,k_2,m_1,m_4)C_{4y}(k_3,k_4,m_2,m_3) + C_{4y}(k_1,k_2,m_2,m_3)C_{4y}(k_3,k_4,m_1,m_4)$
$+C_{4y}(k_1,k_2,m_2,m_4)C_{4y}(k_3,k_4,m_1,m_3) + C_{4y}(k_1,k_2,m_3,m_4)C_{4y}(k_3,k_4,m_1,m_2)$
$+C_{4y}(k_1,k_3,m_1,m_2)C_{4y}(k_2,k_4,m_3,m_4) + C_{4y}(k_1,k_3,m_1,m_3)C_{4y}(k_2,k_4,m_2,m_4)$
$+C_{4y}(k_1,k_3,m_1,m_4)C_{4y}(k_2,k_4,m_2,m_3) + C_{4y}(k_1,k_3,m_2,m_3)C_{4y}(k_2,k_4,m_1,m_4)$
$+C_{4y}(k_1,k_3,m_2,m_4)C_{4y}(k_2,k_4,m_1,m_3) + C_{4y}(k_1,k_3,m_3,m_4)C_{4y}(k_2,k_4,m_1,m_2)$
$+C_{4y}(k_1,k_4,m_1,m_2)C_{4y}(k_2,k_3,m_3,m_4) + C_{4y}(k_1,k_4,m_1,m_3)C_{4y}(k_2,k_3,m_2,m_4)$
$+C_{4y}(k_1,k_4,m_1,m_4)C_{4y}(k_2,k_3,m_2,m_3) + C_{4y}(k_1,k_4,m_2,m_3)C_{4y}(k_2,k_3,m_1,m_4)$
$+C_{4y}(k_1,k_4,m_2,m_4)C_{4y}(k_2,k_3,m_1,m_3) + C_{4y}(k_1,k_4,m_3,m_4)C_{4y}(k_2,k_3,m_1,m_2)$;

$G_{422} =$

$C_{4y}(k_3,k_4,m_1,m_2)R_y(k_1,m_3)R_y(k_2,m_4) + C_{4y}(k_3,k_4,m_1,m_2)R_y(k_1,m_4)R_y(k_2,m_3)$
$+C_{4y}(k_1,k_2,m_3,m_4)R_y(k_3,m_1)R_y(k_4,m_2) + C_{4y}(k_1,k_2,m_3,m_4)R_y(k_3,m_2)R_y(k_4,m_1)$
$+C_{4y}(k_1,k_3,m_1,m_3)R_y(k_2,m_4)R_y(k_4,m_2) + C_{4y}(k_1,k_3,m_2,m_4)R_y(k_2,m_3)R_y(k_4,m_1)$
$+C_{4y}(k_1,k_3,m_1,m_4)R_y(k_2,m_3)R_y(k_4,m_2) + C_{4y}(k_1,k_3,m_2,m_3)R_y(k_2,m_4)R_y(k_4,m_1)$
$+C_{4y}(k_2,k_4,m_1,m_3)R_y(k_1,m_4)R_y(k_3,m_2) + C_{4y}(k_2,k_4,m_2,m_4)R_y(k_1,m_3)R_y(k_3,m_1)$
$+C_{4y}(k_2,k_4,m_1,m_4)R_y(k_1,m_3)R_y(k_3,m_2) + C_{4y}(k_2,k_4,m_2,m_3)R_y(k_1,m_4)R_y(k_3,m_1)$
$+C_{4y}(k_1,k_4,m_1,m_3)R_y(k_2,m_4)R_y(k_3,m_2) + C_{4y}(k_1,k_4,m_2,m_4)R_y(k_2,m_3)R_y(k_3,m_1)$
$+C_{4y}(k_1,k_4,m_1,m_4)R_y(k_2,m_3)R_y(k_3,m_2) + C_{4y}(k_1,k_4,m_2,m_3)R_y(k_2,m_4)R_y(k_3,m_1)$
$+C_{4y}(k_2,k_3,m_1,m_3)R_y(k_1,m_4)R_y(k_4,m_2) + C_{4y}(k_2,k_3,m_2,m_4)R_y(k_1,m_3)R_y(k_4,m_1)$
$+C_{4y}(k_2,k_3,m_1,m_4)R_y(k_1,m_3)R_y(k_4,m_2) + C_{4y}(k_2,k_3,m_2,m_3)R_y(k_1,m_4)R_y(k_4,m_1)$;

$G_{2222} =$

$R_y(k_3,m_1)R_y(k_4,m_2)R_y(k_1,m_3)R_y(k_2,m_4) + R_y(k_3,m_1)R_y(k_4,m_2)R_y(k_2,m_3)R_y(k_1,m_4)$
$+R_y(k_4,m_1)R_y(k_3,m_2)R_y(k_1,m_3)R_y(k_2,m_4) + R_y(k_4,m_1)R_y(k_3,m_2)R_y(k_2,m_3)R_y(k_1,m_4)$

where $R_y(...)$, $C_{4y}(...)$, $C_{6y}(...)$ and $C_{8y}(...)$ are second- to eighth-order cumulants of $y(t)$, as shown in (2)-(5).

Proof: The derivation of (8) is somewhat lengthy but it is straightforward by using (6) and the moment-to-cumulant formulas [1]. ∎

Remark: The statistical characterization of sample fourth-order cumulants of a *single realization* of random process was discussed in [10] and [11]. In [4] an approximate expression for the covariance of the sample fourth-order cumulants of *multiple realizations (snapshots)* of the process used in array processing was derived. But the expression is a function of higher-order moments. The result presented in this paper is a function of higher-order cumulants and is convenient for computing in most applications such as array processing.

It is seen form the theorem that the sample cumulants can not generally be regarded as uncorrelated and equivariace random variables. However, the joint distribution of sample cumulants is asymptotically Gaussian (not circular in general)[11, 12].

4. Asymptotic Performance of FOCD-MUSIC

Let $\hat{C}$ be the estimated cumulant matrix used in FOCD-MUSIC. The true cumulant matrix has a general form of

$$\boldsymbol{C} = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^H \tag{9}$$

where

$$\boldsymbol{\Sigma} = \mathrm{diag}[\gamma_{4,1}, \cdots, \gamma_{4,P}], \tag{10}$$

$$\gamma_{4,p} = \mathrm{Cum}\{ s_p(t), s_p(t), s_p^*(t), s_p^*(t) \} \tag{11}$$

$$\boldsymbol{B} = [\boldsymbol{b}(\theta_1), \cdots, \boldsymbol{b}(\theta_P)] \tag{12}$$

$$\boldsymbol{b}(\theta) = [b_1(\theta), \cdots, b_{M_C}(\theta)]^T. \tag{13}$$

$\boldsymbol{b}(\theta)$ is the steering vector in cumulant domain and M_c is the dimension of the cumulant matrix, which depend on the cumulant matrix specified by the algorithm.

Theorem 2. The FOCD-MUSIC DOA estimator is asymptotically unbiased. The asymptotic variance of $\hat{\theta}_p$ is given by

$$\mathrm{AE}\{ (\Delta\theta_p)^2 \} =$$

$$2\,\mathrm{Re}\Bigg[\sum_{q_1,q_2=1}^{P} \sum_{\substack{i_1,i_2=1 \\ i_1 \neq q_1, i_2 \neq q_2}}^{M_c} \frac{(\boldsymbol{u}_{i_1} \otimes \boldsymbol{u}_{i_2})^H \mathrm{AE}\{ \Delta\boldsymbol{C} \otimes \Delta\boldsymbol{C} \} (\boldsymbol{u}_{q_1} \otimes \boldsymbol{u}_{q_2})}{(\lambda_{i_1} - \lambda_{q_1})(\lambda_{i_2} - \lambda_{q_2})}$$

$$\cdot \boldsymbol{G}_{p,q_1}^T \boldsymbol{u}_{i_1} \boldsymbol{u}_{i_2}^T \boldsymbol{G}_{p,q_2} + \frac{(\boldsymbol{u}_{i_1} \otimes \boldsymbol{u}_{i_2}^*)^H \mathrm{AE}\{ \Delta\boldsymbol{C} \otimes \Delta\boldsymbol{C}^* \} (\boldsymbol{u}_{q_1} \otimes \boldsymbol{u}_{q_2}^*)}{(\lambda_{i_1} - \lambda_{q_1})(\lambda_{i_2} - \lambda_{q_2})}$$

$$\cdot \boldsymbol{G}_{p,q_1}^T \boldsymbol{u}_{i_1} \boldsymbol{u}_{i_2}^H \boldsymbol{G}_{p,q_2}^* \Bigg] \Big/ \ddot{S}^2(\theta_p, \boldsymbol{U}_s) \tag{14}$$

where $\Delta\theta_p = \hat{\theta}_p - \theta_p$, $\otimes$ denotes Kronecker product, λ_i's and $\boldsymbol{u}_i$'s are eigenvalues and the corresponding normalized eigenvectors of $\boldsymbol{C}$, with $|\lambda_i|$ in decreasing order ($\lambda_i = 0$ for $i = P+1, \ldots, M$), $\boldsymbol{U}_s = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_P]$, and

$$\Delta\boldsymbol{C} = \hat{\boldsymbol{C}} - \boldsymbol{C}, \quad \boldsymbol{d}(\theta_p) = \dot{\boldsymbol{b}}(\theta_p) \tag{15}$$

$$\boldsymbol{G}_{p,q} = -\boldsymbol{d}^*(\theta_p)\boldsymbol{u}_q^H \boldsymbol{b}(\theta_p) - \boldsymbol{b}^*(\theta_p)\boldsymbol{u}_q^H \boldsymbol{d}(\theta_p) \tag{16}$$

$$\ddot{S}(\theta_p, \boldsymbol{E}_s) = 2\boldsymbol{d}^H(\theta_p)(\boldsymbol{I}_{M_c} - \boldsymbol{U}_s\boldsymbol{U}_s^H)\boldsymbol{d}(\theta_p) \tag{17}$$

Proof: The derivation of (14) require some approximations on the perturbed signal eigenvectors [13]. The details of the derivation will be given in another paper[8] and are omitted here due to space limitation. ∎

By using theorem 1 the computation of (14) is straightforward when an algorithm of FOCD-MUSIC is specified.

5. Cramér-Rao Bound in Fourth-Order Cumulant Domain

The asymptotic Cramér-Rao lower bound for cumulant-based DOA estimation may be determined in cumulant domain since the sample cumulants are asymptotically Gaussian distributed [12]. We choose all of the nonredundant sample cumulants as the "observations" in cumulant domain without loss of information.

Theorem 3. The maximal set of nonredundant sample cumulants is given by $\mathcal{C}$:

$$\mathcal{C} = \{ \hat{C}_{4y}^N(k_1, k_2, m_1, m_2) | (k_1, k_2, m_1, m_2) \in \mathcal{T} \} \tag{18}$$

$$\mathcal{T} = \{ (k_1, k_2, m_1, m_2) | k_1, k_2, m_1, m_2 \text{ satisfy CONDS} \} \tag{19}$$

where the CONDS is given by:

$$\begin{cases} k_1,\ m_1 \in [1,\ M];\ k_2 \in [1,\ k_1];\ m_2 \in [1,\ m_1] \\ (m_1 - 1)m_1 + 2m_2 \leq (k_1 - 1)k_1 + 2k_2 \end{cases} \tag{20}$$

the number of elements in $\mathcal{C}$ is Q,

$$Q = M(M+1)(M^2 + M + 2)/8. \tag{21}$$

Proof: The result can be obtained by examining the symmetry of the fourth-order cumulants. (20) is a revised version of that in [7]. See [7] for more details. ∎

Let $\tau_q = (k_1^{(q)}, k_2^{(q)}, m_1^{(q)}, m_2^{(q)}) \in \mathcal{T}$. We stack the nonredundant sample cumulants in a vector $\hat{\xi}$

$$\hat{\xi} = [\hat{C}_{4y}^N(\tau_1), \cdots, \hat{C}_{4y}^N(\tau_Q)]^T = \hat{\xi}_R + j\hat{\xi}_I \tag{22}$$

Since the sample cumulants are not circular Gaussian in general, we create a $2Q$-dimensional real Gaussian vector $\hat{\zeta}$, $\hat{\zeta} = [\hat{\xi}_R^T, \hat{\xi}_I^T]^T$. In the case of white additive Gaussian noise, parameters to be estimated in cumulant domain are: DOAs $\theta = \{\theta_p\}$, signal cumulant $\gamma = \{\gamma_{2,p}, \gamma_{4,p}, \gamma_{6,p}, \gamma_{8,p}\}$, $p = 1, \ldots, P$, and noise power σ_w^2. Then the Fisher information matrix (FIM) [14] is given by

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}^{\theta,\theta} & \boldsymbol{J}^{\theta,\gamma} & \boldsymbol{J}^{\theta,\sigma_w^2} \\ (\boldsymbol{J}^{\theta,\gamma})^T & \boldsymbol{J}^{\gamma,\gamma} & \boldsymbol{J}^{\gamma,\sigma_w^2} \\ (\boldsymbol{J}^{\theta,\sigma_w^2})^T & (\boldsymbol{J}^{\gamma,\sigma_w^2})^T & J^{\sigma_w^2,\sigma_w^2} \end{bmatrix} \tag{23}$$

where

$$\boldsymbol{J}_{i,k}^{\theta,\theta} = \frac{\partial\boldsymbol{\zeta}^T}{\partial\theta_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{\zeta}}{\partial\theta_k} + \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\theta_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\theta_k} \} \tag{24}$$

$$\boldsymbol{J}_{i,k}^{\gamma,\gamma} = \frac{\partial\boldsymbol{\zeta}^T}{\partial\gamma_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{\zeta}}{\partial\gamma_k} + \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\gamma_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\gamma_k} \} \tag{25}$$

$$J^{\sigma_w^2,\sigma_w^2} = \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\sigma_w^2} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\sigma_w^2} \} \tag{26}$$

$$\boldsymbol{J}_{i,k}^{\theta,\gamma} = \frac{\partial\boldsymbol{\zeta}^T}{\partial\theta_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{\zeta}}{\partial\gamma_k} + \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\theta_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\gamma_k} \} \tag{27}$$

$$\boldsymbol{J}_i^{\theta,\sigma_w^2} = \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\theta_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\sigma_w^2} \} \tag{28}$$

$$\boldsymbol{J}_i^{\gamma,\sigma_w^2} = \frac{1}{2}\mathrm{tr}\{ \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\gamma_i} \boldsymbol{R}_\zeta^{-1} \frac{\partial\boldsymbol{R}_\zeta}{\partial\sigma_w^2} \}, \tag{29}$$

where $J_{i,k}^{\theta,\theta}$ is the element at ith row and kth column of $J^{\theta,\theta}$, and so on. It is easy to find that

$$R_\zeta = \frac{1}{2}\begin{bmatrix} \mathrm{Re}\{R_{\xi_1} + R_{\xi_2}\} & -\mathrm{Im}\{R_{\xi_1} - R_{\xi_2}\} \\ \mathrm{Im}\{R_{\xi_1} + R_{\xi_2}\} & \mathrm{Re}\{R_{\xi_1} - R_{\xi_2}\} \end{bmatrix}, \quad (30)$$

where

$$R_{\xi_1} = \mathrm{AE}\{\Delta\xi\Delta\xi^H\} = [\mathrm{AE}\{\Delta C_{4y}^N(\tau_{q_1})(\Delta C_{4y}^N(\tau_{q_2}))^*\}]_{Q\times Q} \quad (31)$$

$$R_{\xi_2} = \mathrm{AE}\{\Delta\xi\Delta\xi^T\} = [\mathrm{AE}\{\Delta C_{4y}^N(\tau_{q_1})\Delta C_{4y}^N(\tau_{q_2})\}]_{Q\times Q} \quad (32)$$

Finally, we get the asymptotic Cramér-Rao lower bound for DOA estimation in fourth-order cumulant domain (FOCD-CRB):

$$\mathrm{AE}\{\Delta\theta_p^2\}_{\mathrm{FOCD-CRB}} = J_{p,p}^{-1}, \ p = 1,\ldots, P. \quad (33)$$

where $J_{p,p}^{-1}$ represents the element at pth row and column of J^{-1}. Using the result of theorem 1 the FOCD-CRB can be evaluated numerically.

6. Simulations

Simulation experiments were performed on a uniform linear array (ULA) separated by half a wavelength of the narrowband signals. Two emitters (P=2) broadcast BPSK waveforms from θ_1 =10° and θ_2 =10° with respect to the broadside of the array. The additive noise is white Gaussian. Three algorithms within the class of FOCD-MUSIC, named MUSIC-like[4], MSNC[7] and CTAM[8], were evaluated and compared to the conventional MUSIC[2] and FOCD-CRB by numerical theory results and Monte Carlo simulations. Each of these simulations is based on 100 independent runs. Shown in Fig. 1 are the estimation mean-square-errors (MSE's) of θ_1 versus signal-to-noise ratios (SNR's). The experiment condition is: number of sensors M = 4; number of snapshots N = 500. The close agreement between theoretical predictions and simulation results is clearly evident when the SNR is above 5 dB, and good for CTAM algorithm [8] even when SNR is as low as -5 dB. The MSNC algorithm [7] is almost of no difference with the MUSIC-like algorithm [4] while the former is less computationally expensive than the latter. It can be seen that the FOCD-MUSIC's outperform the conventional MUSIC [2] only when the SNR is somewhat low (less that 10 dB approximately). Fig. 2 and Fig. 3 plot the MSE's versus the number of snapshots (N) while the number of sensors (M) varies from 4 in Fig. 2 to 6 in Fig. 3 (which means the relative spatial separation between the two emitters is increased). Again, the analytical predictions compare favorably with the simulation results when the number of snapshots is not quite low (greater than 100 approximately). It is shown that the FOCD-MUSIC's outperform the conventional MUSIC when spatial emitters are rather closely spaced (less than half of the beamwidth approximately). This is also shown by Fig. 4, where the MSE's are versus the variation of the difference of spatial angles of the two emitters.

7. Conclusions

A unifying asymptotic analysis of FOCD-MUSIC was presented and the FOCD-CRB was also derived. Simulation experiments validated the analytical results. Three typical FOCD-MUSIC's were evaluated and compared to covariance-based MUSIC and FOCD-CRB. The MUSIC-like and MSNC have an almost identical performance as predicted in [7]. CTAM behaves rather robustnessly in all scenarios. The FOCD-MUSIC's are not superior to the conventional MUSIC algorithm in all cases. Nevertheless, they outperform the conventional MUSIC with reduced variances and improved robustness when the spatial sources are closely spaced and the signal-to-noise ratios (SNRs) are relatively low. The FOCD-MUSIC's are inefficient. It can be seen from Fig. 2 and Fig. 3 that the estimation covariances decrease parallel with the FOCD-CRB while the number of snapshots goes large, which means that the FOCD-MUSIC's are not asymptotically efficient.

References

[1] J. M. Mendel, "Tutorial on higher-order statistics(spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, 79(3), 278-305, Mar. 1991.

[2] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. AP*, 34(3), 276-280, Mar. 1986.

[3] R. Pan and C. L. Nikias, "Harmonic decomposition methods in cumulant domains," *Proc. IEEE ICASSP-88*, 2356-2359.

[4] B. Porat and B. Friedlander, "Direction finding algorithms based on higher-order statistics," *IEEE Trans. SP*, 39(9), 2016-2024, Sept. 1991.

[5] E. Moulines and J. F. Cardoso, "Direction finding algorithms using fourth order statistics. Asymptotic performance analysis," *Proc. IEEE ICASSP-92*, II-437-440.

[6] Y.-H. Chen and Y.-S. Lin, "Fourth-order cumulant matrices for DOA estimation," *IEE Proc.-Radar, Sonar and Navig.*, 141(3), 144-148, June 1994.

[7] H. Wu, Y.-K. Jia and Z. Bao, "Direction finding and array calibration based on maximal set of nonredundant cumulants," *Proc. IEEE ICASSP-96*, 2626-2629.

[8] H. Wu and Z. Bao, "Robustness of DOA estimation in cumulant domain," submitted to *IEEE Trans. SP*, under revising, 1996.

[9] X. Fan and N. H. Younan, "Asymptotic analysis of the cumulant-based MUSIC method in the presence of sample cumulant errors," *IEEE Trans. SP*, 43(3), 799-802, 1995.

[10] Fonollosa, J. A. R., "Sample cumulants of stationary process: Asymptotic results," *IEEE Trans. SP,* 43(4), 967-977, 1995.

[11] P. Tichavsky and A. Swami, "Statistical characterization of sample fourth-order cumulants of a noisy complex sinusoidal process," *IEEE Trans. SP,* 43(7), 1620-1630, July 1995.

[12] G. B. Giannakis and M. K. Tsatsanis, "A unifying maximum-likelihood view of cumulant and polyspectral measures for non-Gaussian signal classification and estimation," *IEEE Trans. IT,* 38(2), 386-406, Mar. 1992

[13] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the Minimum-Norm algorithms in resolving plane waves in noise," *IEEE Trans. ASSP* 34(2), 331-341, Apr. 1986.

[14] J. M. Francos and B. Friedlander, "Bounds for estimation of complex exponentials in unknown colored noise," *IEEE Trans. SP,* 43(9), 2176-2185, Sept. 1995.

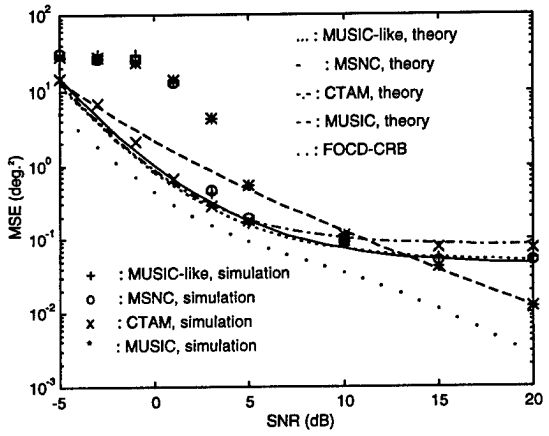


Fig. 1 Mean-square-errors (MSE's) versus signal-to-noise ratios (SNR's) for $\theta_1 = 10°$. $M = 4$, N=500.

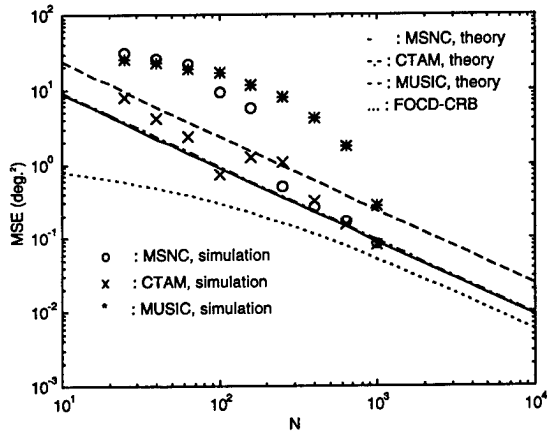


Fig. 2 Mean-square-errors (MSE's) versus number of snapshots (N) for $\theta_1 = 10°$. $M = 4$, SNR= 5 dB.

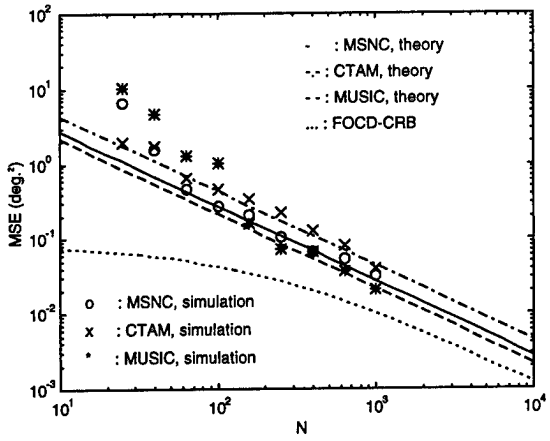


Fig. 3 Mean-square-errors (MSE's) versus number of snapshots (N) for $\theta_1 = 10°$. $M = 6$, SNR = 5 dB.

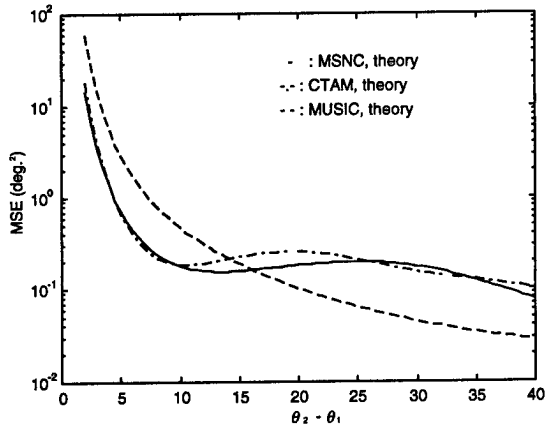


Fig. 4 Mean-square-errors (MSE's) versus the angle differences of the two emitters for $\theta_1 = 10°$. $M = 4$, N =500, SNR =5 dB.

Adaptive Separation of an Unknown number of Sources

Z. Malouche and O. Macchi
Laboratoire des Signaux et Systèmes
CNRS/Univ Paris Sud - GDR-PRC ISIS
Supélec, Plateau du Moulon, 91192 Gif-sur-Yvette Cedex, France
E-mail: malouche@lss.supelec.fr

Abstract

The problem of separation of mixed sources is adressed in this paper. To solve this problem, at least as many observations as sources are needed. In particular, the number of sources can be unknown. The separation system is a linear network updated with a stochastic descent algorithm to minimize some separation criterion. A first algorithm separates sources with positive kurtosises while a second one separates sources with negative kurtosises. For both, the performances are independent of the mixture. Besides, in the noisy case, when there are more sensors than sources, the additional outputs merely generate noise.

1. Introduction and problem statement

The unsupervised separation of several (p) independent zero-mean sources s_i contained in several (m) observed linear mixtures x_j has become a classical problem in signal processing. It has applications in mobile-radio communications, array processing, biological signal processing, feature classification, etc... In most previous contributions, it is assumed that the number p of sources has already been estimated [1],[3],[5]-[7], [13]. In [4], the case of an unknown p is considered but only in the noiseless context.

In [12] we have proposed an adaptive algorithm which can separate an unknown number of sources, thanks to some optimality criterion. But the results are restricted to sources with negative kurtosis. In this contribution, starting with the same criterion, we derive two other algorithms respectively efficient with sources having positive or negative kurtosis. They both have satisfactory achievement in the noisy case.

The mixture model under investigation is

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{w} \tag{1}$$

where $\boldsymbol{s} \triangleq (s_1, \ldots, s_p)^T$, $\boldsymbol{x} \triangleq (x_1, \ldots, x_m)^T$, $\boldsymbol{A}$ is the $m \times p$ mixture matrix. The $m \times 1$ vector $\boldsymbol{w}$ accounts for the presence of a zero-mean, Gaussian noise vector that is assumed independent of the source vector. The problem is unsupervised when $\boldsymbol{s}$, $\boldsymbol{A}$ and $\boldsymbol{w}$ are all unknown and only $\boldsymbol{x}$ is observed. It is an inherent feature of this problem that the order of sources s_i cannot be recovered. We assume that the rank of $\boldsymbol{A}$ is p, which implies that $m \geq p$. Consider a separation system with q outputs, where q is some a priori overestimate of p ($q \geq p$). The output vector is

$$\boldsymbol{y} \triangleq (y_1, \ldots, y_q)^T = \boldsymbol{H}^T \boldsymbol{x}. \tag{2}$$

where the $m \times q$ matrix $\boldsymbol{H}$ is characterizes the separating system. Denoting

$$\boldsymbol{C} \triangleq \boldsymbol{A}^T \boldsymbol{H} \tag{3}$$

the $p \times q$ matrix of the over-all system (mixture followed by separation system), the output is

$$\boldsymbol{y} = \boldsymbol{C}^T \boldsymbol{s}. \tag{4}$$

Clearly separation is achieved iff

$$\boldsymbol{C} = \boldsymbol{\Lambda}[\boldsymbol{I}\ \boldsymbol{Q}]\boldsymbol{P} \tag{5}$$

where $\boldsymbol{P}$ is an arbitrary $q \times q$ permutation matrix. The bracket means concatenation of the $p \times p$ invertible diagonal matrix $\boldsymbol{\Lambda}$ and the $p \times (q-p)$ matrix $\boldsymbol{Q}$ whose columns contain at most one non zero entry. Equation (5) means that all of the p sources s_i are recovered on at least one output y_j with arbitrary non zero gain and arbitrary order, the other outputs being additional copies of certain sources.

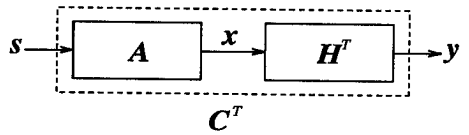


Figure 1. Overall system: mixture matrix followed by separating system

0-8186-8005-9/97 $10.00 © 1997 IEEE

2 Separation criterion and algorithms

The separation criterion is to minimize

$$J(\boldsymbol{H}) = \sum_{j=1}^{q} E\{y_j^4 - 2\sigma y_j^2\} - 2\alpha \mathrm{Ln}(\det \boldsymbol{H}^T\boldsymbol{H}) \quad (6)$$

where σ and α are two scaling positive quantities, as already proposed in [12]. The first part of this criterion is a sum of simple extracting criteria [11] associated to one output. The second term forces the rank of the matrix $\boldsymbol{H}$ to be equal to q by avoiding the case where $\boldsymbol{H}^T\boldsymbol{H}$ is singular. In the noiseless case, it can be shown that under some mild condition, the separating points in $\boldsymbol{H}$ are forced onto the minima of the criterion (6).

The stochastic gradient of the criterion (6) reads

$$\nabla J = \boldsymbol{x}\boldsymbol{y}^T \mathrm{diag}(\boldsymbol{y}\boldsymbol{y}^T) - \sigma \boldsymbol{x}\boldsymbol{y}^T - \alpha \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1} \quad (7)$$

where $\mathrm{diag}(\boldsymbol{y}\boldsymbol{y}^T) \stackrel{\Delta}{=} \mathrm{diag}(y_1^2, \ldots, y_q^2)$. The increment of the associated adaptive algorithm is

$$\Delta_0 \boldsymbol{H} = -\mu \nabla J \ , \ \mu > 0 \quad (8)$$

where μ is a (small) step-size. The computation of (7) is awkward since it requires to invert the matrix $\boldsymbol{H}^T\boldsymbol{H}$. In [2] and [1],[5], two different stochastic descent algorithms have been proposed with respective increments

$$\Delta_1 \boldsymbol{H} = -\mu \boldsymbol{H} \nabla J^T \boldsymbol{H} \quad (9)$$

$$\Delta_2 \boldsymbol{H} = -\mu \boldsymbol{H}\boldsymbol{H}^T \nabla J \quad (10)$$

It follows from (7) that

$$\Delta_r \boldsymbol{H} = -\mu \boldsymbol{H} \delta_r(\boldsymbol{y}) \ , \ r = 1, 2 \quad (11)$$

where

$$\delta_1(\boldsymbol{y}) = \delta_2(\boldsymbol{y})^T = \mathrm{diag}(\boldsymbol{y}\boldsymbol{y}^T)\, \boldsymbol{y}\boldsymbol{y}^T - \sigma \boldsymbol{y}\boldsymbol{y}^T - \alpha \boldsymbol{I} \quad (12)$$

Clearly Δ_1 and Δ_2 require less computation than Δ_0 since the inversion of $\boldsymbol{H}^T\boldsymbol{H}$ has disappeared. Moreover, the increments (11) show an "equivariant" property according to the definition in [5]: in the noiseless case, the separation performances do not depend on the mixture $\boldsymbol{A}$. This can be seen on the so-called "transformed" algorithm that is obtained for the overall system $\boldsymbol{C}$ by premultiplying eq. (11) with $\boldsymbol{A}^T$. The corresponding increment reads

$$\Delta_r \boldsymbol{C} = -\mu \boldsymbol{C} \delta_r(\boldsymbol{y}) \ , \ r = 1, 2 \quad (13)$$

which only involves the overall system $\boldsymbol{C}$ and its output $\boldsymbol{y}$, but neither $\boldsymbol{A}$ nor $\boldsymbol{x}$. Therefore the separation performances are not affected by a possible ill- conditionning of $\boldsymbol{A}$, contrary to the gradient increment (8).

3. Stationary points

The stationary points $\overline{\boldsymbol{H}}$ of the two algorithms (11) are those of the associated ordinary differential equations (ODE)

$$\frac{d\boldsymbol{H}}{dt} = -E\{\boldsymbol{H}\delta_r(\boldsymbol{y})\} \ , \ r = 1, 2. \quad (14)$$

It is interesting to investigate the corresponding transformed ODE

$$\frac{d\boldsymbol{C}}{dt} = -E\{\boldsymbol{C}\delta_r(\boldsymbol{y})\} \ , \ r = 1, 2. \quad (15)$$

Separating stationary points When $\sigma = 1$, it can be shown that the separating stationary points of (15) are

$$\overline{\boldsymbol{C}} = \overline{\boldsymbol{\Lambda}} \, [\boldsymbol{I} \ \overline{\boldsymbol{Q}}] \, \overline{\boldsymbol{P}} \quad (16)$$

where

- $\overline{\boldsymbol{\Lambda}} = \mathrm{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_p)$, and $\overline{\lambda}_i$ is a root of

$$E\{a_i^4\}\lambda_i^4 - E\{a_i^2\}\lambda_i^2 - \frac{\alpha}{1 + r_i} = 0, \quad (17)$$

- r_i denotes the number of supplementary times where the source s_i is extracted. The r_i only depend on initialization and satisfy $\sum_{i=1}^{p} r_i = q - p$,
- the $p \times (q - p)$ matrix $\overline{\boldsymbol{Q}}$ has the particular form

$$\overline{\boldsymbol{Q}} = \begin{pmatrix} \overbrace{1 \cdots 1}^{r_1} & 0 \cdots 0 & & 0 \\ 0 \cdots 0 & \overbrace{1 \cdots 1}^{r_2} & & \\ & & \ddots & \\ 0 & & & \overbrace{1 \cdots 1}^{r_p} \end{pmatrix} \quad (18)$$

when $q > p$. It is empty when $q = p$.

Clearly a source may be extracted more than once. However, all the sources are extracted at least once. For both algorithms $(15)_{r=1,2}$, the study of stability for a separating stationary point provides a condition on the source statistics, but its expression is somewhat intricate. Fortunately, some sufficient (unnecessary) stability conditions can be stated easily:

Result 1: If all the sources have positive (resp. negative) kurtosis the algorithms $\Delta_1 \boldsymbol{C}$ (resp. $\Delta_2 \boldsymbol{C}$) is stable at the separating stationary point (16)
where the kurtosis of a source s_i is $\kappa_i = E\{a_i^4\} - 3E^2\{a_i^2\}$. Moreover, separation of sources whose kurtosises have different signs may be possible in some cases.

Non-separating stationary points Both ODE (15) exhibit non-separating stationary points whose expressions are not explicited in this paper. However, they are unstable when the kurtosis condition of *result 1* holds.

4 Behaviour of the algorithm in H

In fact, the behaviour of the transformed algorithms (13) in C does not always describe correctly the behaviour of the direct algorithms (11) written with H. This depends on whether $m = q = p$ or $m \geq q > p$, and whether the system is affected or not by some noise.

The case $m = q = p$ In this case A is invertible. Hence a one to one correspondance between algorithms (11) and algorithms (13). Therefore, the convergence of C to $\overline{C}$ implies the convergence of H to $\overline{H} = A^{-T}\overline{C}$.

The case $m \geq q > p$

• *The noise-free case.* We have shown that

$$\Delta H = M \Delta C + N. \tag{19}$$

where M is an $m \times p$ constant matrix and N is an $m \times q$ non zeromean stochastic matrix. Consequently, $E\{\Delta C\} = 0$ does not imply $E\{\Delta H\} = 0$. In other words, although the separating stationary state is reached by C, the matrix H still wanders and never reaches any stationary point.

There exists a linear subspace $\overline{\mathcal{H}}$ of dimension $(m - p)$ in which the matrices $\overline{H}$ are such that $\overline{C} = A^T \overline{H}$. Since the algorithms (11) minimize -Ln(det$H^T H$), this logarithm will reach $-\infty$. So H wanders in $\overline{\mathcal{H}}$ until det$H^T H$ gets infinite. Consequently, the entries of H are divergent.

• *The noisy case.* The m noise components assumed independent can be viewed as other sources. Therefore, the new number of independent sources is $p' = p + m$ and the condition $m \geq p'$ no longer holds. Then perfect separation of the p' sources is impossible. However, by considering a low noise power, it can be shown that after convergence, the separating system outputs read (up to a permutation of indices)

$$\overline{y}_i \equiv \begin{cases} s_i + \epsilon_1^T s + \epsilon_2^T w & , \quad i = 1, \ldots, p \\ \gamma^T w + \epsilon_3^T s & , \quad i = p+1, \ldots, q \end{cases} \tag{20}$$

where γ denotes a vector and $\epsilon_1, \epsilon_2, \epsilon_3$ three vectors with small entries. Hence a quasi-separating stationnary state. Note that when $m > p$ increases, the norm of ϵ decreases.

5. Simulations

In order to confirm the theoretical results, two simulations have been run with $p = 2$ sources, $m = q = 4$ sensors, and for a mixture matrix

$$A = \begin{pmatrix} 1.0 & 0.5 \\ 0.4 & 1.0 \\ 0.9 & 0.4 \\ 0.3 & 1.1 \end{pmatrix}. \tag{21}$$

The noise components w_i are zero-mean, independent Gaussian variables with power $P_w = 0.01$. Simulation 1 (resp. 2) corresponds to the increment $\Delta_1 H$ (resp. $\Delta_2 H$) and uses 2 sources with positive (resp. negative) kurtosis whose distribution is depicted in Fig.2(a) (resp. in Fig.2(b)). The histograms of the mixtures x_j are respectively depicted in Fig.3 (resp. in Fig.6). The 4 output signals y_j are plotted in 5 and 8 respectively, which show that separation is reached after approximatively 2000 iterations for the two algorithms. The histograms of the output signals are depicted in Fig.4 and Fig.7 respectively. In simulation 1 the sources are recovered by outputs (y_1,y_4) while in simulation 2 the sources are recovered by outputs (y_2,y_3)). The residual noise components are noticeble in the histogram of the recovered sources. Besides, the histograms of the remaining outputs exhibit the characteristic Gaussian shape corresponding to a mixture of noise components. As a result, a rough a priori knowledge about the sources statistics makes them easy to detect in number and value.

6. Conclusion

Two new simple adaptive separation algorithms are presented in this paper. Both are derived from the same optimization criterion. They both work with an unknown number of sources. The first one separates sources with positive kurtosises while the second one works with sources having negative kurtosis. For a low additive Gaussian noise, each source is recovered once. If there are more sensors than sources, the other outputs recover some combinations of the noise components. Therefore, the number of sources is easily estimated.

References

[1] S. Amari, A. Cichocki and H. Yang, A new learning algorithm for blind signal separation *Proc. NIPS*, Denver, 1995.

[2] J. J. Atick and A. N. Redlich, Convergent algorithm for sensory receptive field development *Neural Computation*, Vol. 5, pp. 45-60, 1993.

[3] A. J. Bell and T. J. Sejnowski, An information-maximization approach to blind source separation and blind deconvolution. *Neural Computation*, N.7, Vol. 7, pp. 1129-1159, 1995.

[4] A. Cichocki, W. Kasprzak and S.I. Amari, Neural network approach to blind separation and enhancement of images. *Proc. EUSIPCO*, pp. 579-582, September 1996.

[5] J. F. Cardoso and B. Laheld, Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, N.44, Vol. 12, pp. 3017-3030, December 1996.

[6] P. Comon. Independent component analysis, a new concept? *Signal Processing*, N.36, Vol. 3, pp. 287-314, April 1994.

[7] N. Delfosse and P. Loubaton, Adaptative blind separation of independent sources: a deflation approach, *Signal Processing*, Vol. 45, pp. 59-83, 1995.

[8] J. Hérault and C. Jutten, Blind separation of sources, part I : An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, Vol. 24, pp. 1-10, 1991.

[9] A. Hyvärinen, Simple one-unit neural algorithms for blind source separation and blind deconvolution *Proc. ICONIP*, Honk Kong, September, 25-27 1996.

[10] Z. Malouche and O. Macchi, Extended anti-hebbian adaptation for unsupervised source extraction. *Proc. ICASSP*, Vol. 3, pp. 1665-1668, May 1996.

[11] Z. Malouche and O. Macchi, Adaptive unsupervised extraction of one component of a linear mixture with a single neuron. *Submitted to IEEE transactions on Neural Networks*, June 1996.

[12] Z. Malouche and O. Macchi, A linear adaptive neural network for extraction of independent components *Proc. ESANN*, pp. 261-266 , April 16-18, 1997.

[13] E. Moreau and O. Macchi, High order contrasts for self-adaptive source separation, *International Journal of Adaptive Control and Signal Processing*, Vol. 10, No. 1, pp 19-46, January 1995.

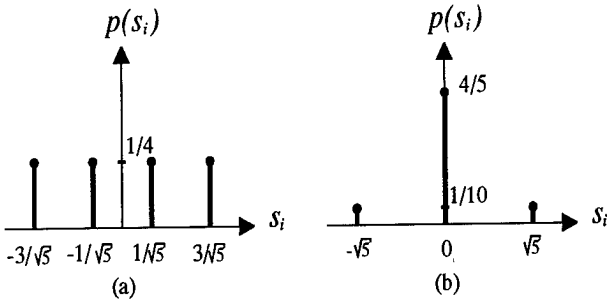


Figure 2. Probability distribution functions of the two sources: (a) with positive kurtosis $\kappa_i = 2$ (b) with negative kurtosis $\kappa_i = -1.36$

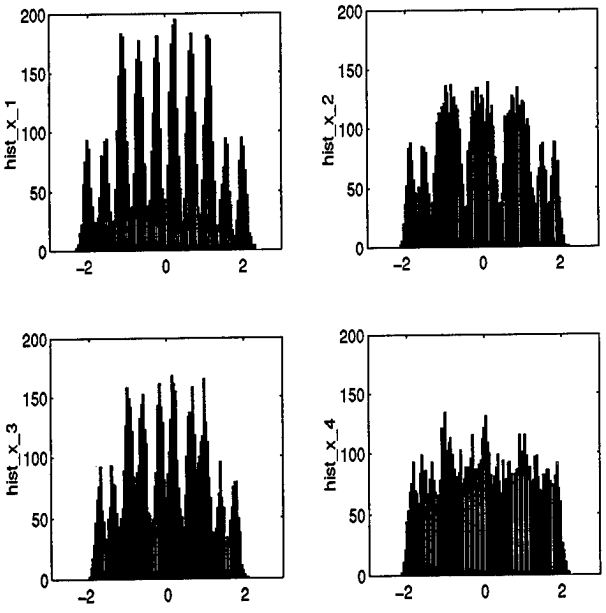


Figure 3. Simulation 1: histograms of the mixtures x_j

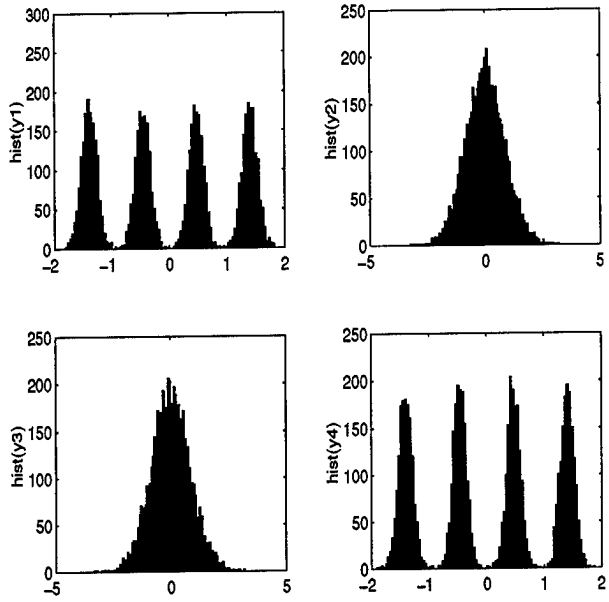


Figure 4. Simulation 1: histograms of the outputs y_j

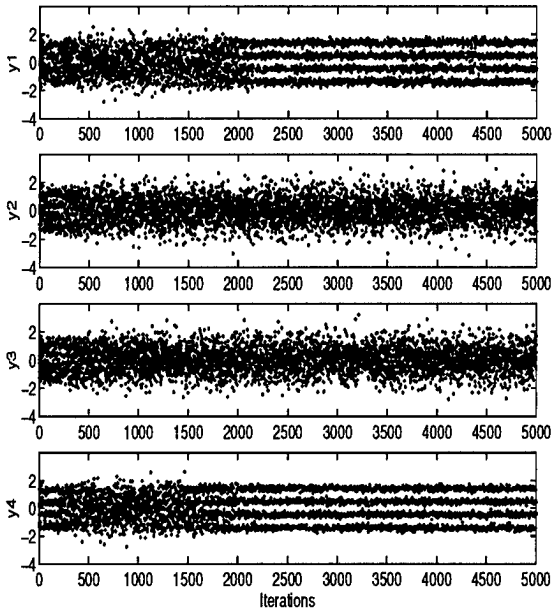


Figure 5. Simulation 1: output signals y_j

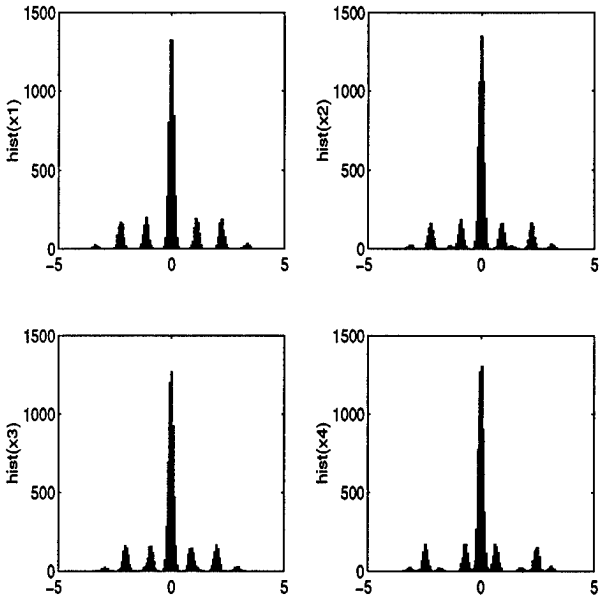


Figure 6. Simulation 2: histograms of the mixtures x_j

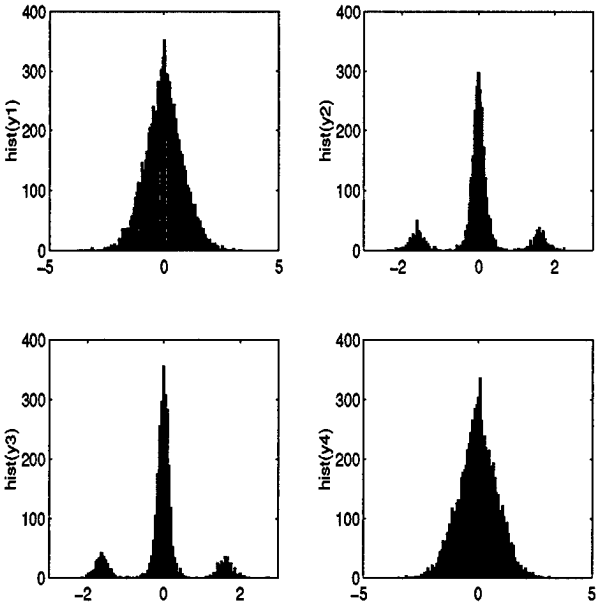


Figure 7. Simulation 2: histograms of the outputs y_j

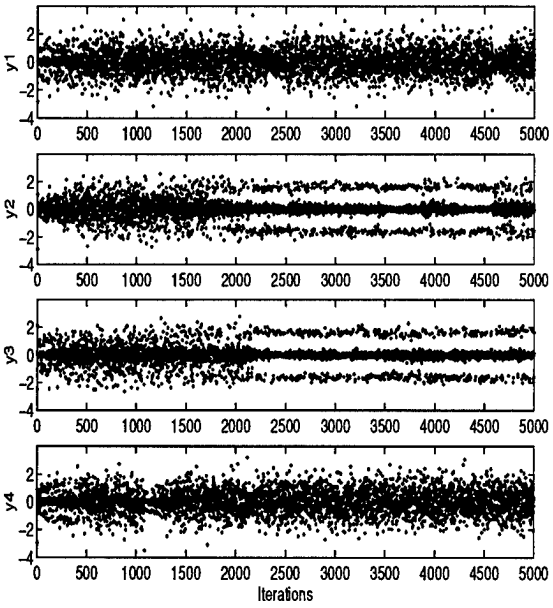


Figure 8. Simulation 2: output signals y_j

NEAR-FIELD LOCALIZATION USING INVERSE FILTER CRITERIA-BASED BLIND SEPARATION AND CUMULANT MATCHING

Anil Govindaraju and *Jitendra K. Tugnait*

Department of Electrical Engineering
Auburn University, Auburn, AL 36849, USA
Tel: (334)844-1846 FAX: (334)844-1809
E-mail: govinan@eng.auburn.edu tugnait@eng.auburn.edu

ABSTRACT

This paper is concerned with the problem of near-field source localization. The problem is tackled using the method of blind separation of independent signals (sources) from their linear instantaneous (memoryless) mixtures. The various signals are assumed to be zero-mean non-Gaussian but not necessarily linear or i.i.d. Approaches using higher-order cumulants are developed using the fourth-order normalized cumulants of the "beamformed" data. The instantaneous mixture matrix is obtained by cross-correlating the extracted inputs with the observed outputs. The first approach is source-extractive, i.e., the sources are extracted and cancelled one-by-one. The other approach is based upon cumulant matching of the estimated and model-based cumulants parametrized by the source parameters (range, bearing and cumulant). Illustrative simulation examples are provided.

1. INTRODUCTION

This paper is concerned with the problem of near-field source localization using blind separation of independent signals (sources) from their linear instantaneous (memoryless) mixtures The problem consists of extracting the sources without any distortion from the observation of unknown linear mixtures of the unknown sources. Such a model arises in a wide variety of array processing applications. In this paper we study the problem of near-field localization via blind source separation. Prior work on this problem includes [1] and [2]. General source separation problem has been considered in [5]-[7],[11].

It has been pointed out in [1] and [2] that the second-order statistics based methods (such as that in [9], and others (see [1],[2])) are computationally complex requiring nonlinear optimization whose convergence to the desired solution is not guaranteed. Moreover, they yield biased estimates in spatially correlated Gaussian measurement noise. In [1] and [2], two methods have been proposed for near-field source localization. One of them is a search-free solution (TLS-ESPIRIT) and the other is a cumulant-matching based solution which requires nonlinear optimization and which is initialized using the TLS-ESPIRIT solution for reliable convergence. In this paper we investigate alternatives to [1]-[2]. Use of blind separation for near-field localization has been suggested in [4].

2. SYSTEM MODEL

Given measurements $y_i(k), (i = 1, 2, ..., N)$, at time k at N sensors, let these measurements be a linear instantaneous mixture of M source signals $s_m(k), (m = 1, 2, ..., M)$:

$$y_i(k) = \sum_{m=1}^{M} f_{im} s_m(k) + n_i(k), \quad i = 1, 2, ..., N, \tag{1}$$

$$\Rightarrow \mathbf{y}(k) = \mathbf{F}\mathbf{s}(k) + \mathbf{n}(k), \tag{2}$$

where $\mathbf{y}(k) = [y_1(k) \vdots y_2(k) \vdots \cdots \vdots y_N(k)]^T$, similarly for $\mathbf{s}(k)$ and $\mathbf{n}(k)$, $s_m(k)$ is the m-th input at sampling time k, $y_i(k)$ is the i-th output, $n_i(k)$ is the additive Gaussian measurement noise, and f_{im} is the scalar memoryless transfer function with $s_m(k)$ as the input and $y_i(k)$ as the output. We will call the matrix $\mathbf{F}$ (with im-th component f_{im}) the mixing matrix. We allow all of the above values to be complex-valued. The vector sequence $s(k)$ is assumed to be zero-mean and spatially independent (i.e. various source signals are independent). Also assume that the fourth-cumulant or the third-cumulant of $s_m(k)$ is nonzero $\forall m$. In blind source separation problems the mixing matrix $\mathbf{F}$ is unknown and the problem is to recover $\mathbf{s}(k)$ (upto scaling) given data.

Consider a uniform linear array with inter-element distance d and number of sensors N. The sensors are assumed to be omnidirectional and identical, indexed as $1, 2, ..., N$. The impinging signals arrive from M narrowband, co-channel sources with center frequency ω_0 rad/sec and wavelength λ. Let the m-th emitter's signal arrive at an angle θ_m w.r.t. the array broadside and let this emitter be at a range r_m, both these parameters being measured w.r.t. the sensor '1'. Under these assumptions f_{im} is given by [1],[2],[9]

$$f_{im} = \exp\{jP_{im}\} \tag{3}$$

where

$$P_{im} = \omega_m(i-1) + \phi_m(i-1)^2, \tag{4}$$

$$\omega_m := -2\pi \frac{d}{\lambda} \sin(\theta_m), \quad \phi_m := \pi \frac{d^2}{\lambda r_m} \cos^2(\theta_m). \tag{5}$$

The near-field localization problem consists of estimation of the m-th source azimuth angle θ_m and range r_m given the observation vector $\mathbf{y}(k), k = 1, 2, ..., T$.

The following conditions have been imposed in [1]-[2] to arrive at their solution:

- **(SC-1)** The N-element array is uniform and linear with spacing $d \leq \lambda/4$ and moreover, $N - 2 > M$.
- **(SC-2)** The range parameters of the sources are not equal, i.e., $\phi_k \neq \phi_l$ for $k \neq l$ in (5).
- **(SC-3)** The source signals $s_m(k)$ are zero-mean, mutually independent, non-Gaussian, narrowband, stationary sources with non-zero kurtosis. The sensor noise $n_i(k)$ is independent of the source signals and is Gaussian.

We provide a solution based upon the following conditions:

This work was supported by the National Science Foundation under Grant MIP-9312559.

0-8186-8005-9/97 $10.00 © 1997 IEEE

(GT-1) The N-element array is uniform and linear with spacing $d \leq \lambda/2$ and moreover, $N \geq M$.

(GT-2) The source signals $s_m(k)$ are zero-mean, mutually independent, non-Gaussian, narrowband, stationary sources with non-zero kurtosis such that $(\theta_m, r_m) \neq (\theta_n, r_n)$ for $n \neq m$. The sensor noise $n_i(k)$ is independent of the source signals and is Gaussian.

Note in particular, that for the same number of array elements we allow for larger aperture (hence better accuracy): compare **(SC-1)** with **(GT-1)** – we allow $d = \lambda/2$. As a matter of fact we consider two solutions:

- **Iterative Source Separation Using Inverse Filter Criteria.** Here we follow [3] and [8] to estimate a scaled and permuted version of $\mathbf{F}$ (see also [11]). Note that unlike [3] and [8] we only use spatial filtering; there is no temporal filtering as the mixture in (1)-(2) is instantaneous. This method yields columns of $\mathbf{F}$. Next each extracted column of $\mathbf{F}$ is used to estimate θ_m and r_m pair using (3)-(5). Here we first use a 'crude' closed-form solution followed by nonlinear optimization using the closed-form solution as an initial guess.
- **Cumulant matching.** Here we estimate $\mathbf{F}$ (or three parameters, range, bearing and signal cumulant, per source) via cumulant matching using the iterative source separation solution as an initial guess. Also we use all cumulants combinations, not a just a few as in [2]. The advantage of cumulant matching is that the resulting solution is consistent in noise.

3. ESTIMATION CRITERIA

The iterative separation step is based on estimation using higher-order statistics. In the rest of this section the noise $\mathbf{n}(k)$ in (1) is assumed to be negligible. Let $\mathrm{CUM}_4(w)$ denote the fourth-order cumulant of the complex-valued zero-mean random variable w, defined as

$$\begin{aligned}\mathrm{CUM}_4(w) &:= \mathrm{cum}_4\{w, w^*, w, w^*\} \\ &= E\{|w|^4\} - 2[E\{|w|^2\}]^2 - |E\{w^2\}|^2. \end{aligned} \quad (6)$$

The notation $\gamma_{4sm} = \mathrm{CUM}_4(s_m(k))$ and $\sigma^2_{sm} = E\{|s_m(k)|^2\}$ is used. Consider an 1 x N row-vector $\mathbf{C^T}$, with its i-th entry denoted by c_i, operating on the data vector $\mathbf{y}(k)$. Let the product be denoted by $e(k)$. Then it follows that

$$e(k) = \sum_{i=1}^{N} c_i y_i(k) = \sum_{m=1}^{M} h_m s_m(k), \quad (7)$$

where

$$h_m := \sum_{i=1}^{N} c_i f_{im} \quad \text{and} \quad \mathbf{H} := [h_1 h_2 ... h_M]^T = \mathbf{C^T F}. \quad (8)$$

From [10], the fourth-order cumulant can be expressed as

$$\mathrm{CUM}_4(e(k)) = \sum_{m=1}^{M} \gamma_{4sm} |h_m|^4 = \sum_{m=1}^{M} \overline{\gamma}_{4sm} |\overline{h}_m|^4, \quad (9)$$

$$E\{|e(k)|^2\} = \sum_{m=1}^{M} \sigma^2_{sm} |h_m|^2 = \sum_{m=1}^{M} |\overline{h}_m|^2, \quad (10)$$

where

$$\overline{h}_m := \sigma_{sm} h_m \quad \text{and} \quad \overline{\gamma}_{4sm} := \frac{\gamma_{4sm}}{\sigma^4_{sm}}. \quad (11)$$

Consider the cost function

$$J_{42}(\mathbf{C}) := |\mathrm{CUM}_4(e(k))| / |E\{|e(k)|^2\}|^2. \quad (12)$$

Define $|\overline{\gamma}_{4max}| = \max_{1 \leq m \leq M} |\overline{\gamma}_{4sm}|$. We get from (9)

$$\begin{aligned} |\mathrm{CUM}_4(e(k))| &\leq |\overline{\gamma}_{4max}| \sum_{m=1}^{M} \frac{|\overline{\gamma}_{4sm}|}{|\overline{\gamma}_{4max}|} |\overline{h}_m|^4 \\ &\leq |\overline{\gamma}_{4max}| \left[\sum_{m=1}^{M} |\overline{h}_m|^2 \right]^2 \end{aligned} \quad (13)$$

with equality if and only if

$$\overline{h}_m = d\delta(m - m_0), \quad m_0 \in \{1, 2, ..., M\}, \quad (14)$$

where d is some complex constant and m_0 is such that $|\overline{\gamma}_{4sm_0}| = |\overline{\gamma}_{4max}|$. Note that (14) follows from the fact that $\sum_{m=1}^{M} |\overline{h}_m|^4 = [\sum_{m=1}^{M} |\overline{h}_m|^2]^2$ iff (14) is true for some m_0. Thus we have

$$J_{42}(\mathbf{C}) \leq |\overline{\gamma}_{4max}|, \quad (15)$$

with equality iff (14) holds (except for $|\overline{\gamma}_{4max}| = |\overline{\gamma}_{4sm}|$). Thus, when (14) holds true, (7) reduces to

$$e(k) = d s_{m_0}(k), \quad (16)$$

i.e., the processed output is a possibly scaled version of one of the sources. This can be utilized to separate the sources.

3.1. Does such a solution exist?

It follows from (11) and (14) that

$$h_j = \sum_{i=1}^{N} c_i f_{ij} = \begin{cases} d & \text{if} \quad j = j_0 \\ 0 & \text{if} \quad j \neq j_0, \; j = 1, 2, \cdots, M \end{cases}$$

In the matrix notation we have

$$\mathbf{F}^T \mathbf{C} = \mathbf{E} \quad (17)$$

where

$$\mathbf{E} = [0 \; 0 \; \cdots \; 0 \; d \; 0 \; \cdots \; 0]^T.$$

Note that the $M-$column vector $\mathbf{E}$ has the nonzero entry in its j_0-th row with zeros every place else. A set of sufficient conditions for the existence of the desired solution is: $N \geq M$ and $\mathbf{F}$ has full column-rank. These conditions are satisfied if **(GT-1)** and **(GT-2)** hold true.

3.2. Stationary Points

Suppose that we choose to optimize (12) via a gradient-based optimization method. Then we will converge to a stationary point of (12). The set of stationary points of the cost $J_{42}(\mathbf{C}) = G_{42}(\mathbf{h})$ w.r.t. the composite response $\overline{h}_j$ is the set of solutions to

$$\frac{\partial G_{42}(\mathbf{h})}{\partial \overline{h}^*_m} = 0. \quad (18)$$

From (9),(10),(12) and (18) we have either $\overline{h}_m = 0$ or

$$\overline{\gamma}_{4sm} |\overline{h}_m|^2 = \mathrm{CUM}_4(e(k)) \,/\, E\{|e(k)|^2\} \quad (19)$$

where the right-side of (19) is evaluated at the solution to (18) under consideration. Since the cost $G_{42}(\mathbf{h})$ $(=J_{42}(\mathbf{C}))$

is invariant to any scaling of $\mathbf{h}$ or of $\mathbf{C}$, in practice this ambiguity is resolved by requiring unit norm for the equalizer tap gains (i.e. $\mathbf{C}$); see Sec. 4. This therefore rules out the solution $\overline{h}_j = 0\ \forall j$ to (18). Therefore, we will confine our attention to solutions that have at least one nonzero element $\overline{h}_j$.

Let P be the number of nonzero components of

$$\mathbf{h} := [\overline{h}_1, \cdots, \overline{h}_M]^T$$

for which (18) holds true. For a prespecified P ($P = 1, 2, \cdots M$), the solutions to (18) are all vectors

$$\mathbf{h}^{(P)} := [\overline{h}_1^{(P)}, \cdots, \overline{h}_M^{(P)}]^T$$

such that

$$|\overline{h}_j^{(P)}|^2 = \begin{cases} \beta_P / \overline{\gamma}_{4sj} & \text{if } j \in I_P \\ 0 & \text{otherwise} \end{cases} \qquad (20)$$

where $\beta_P = \mathrm{CUM}_4(e(k))/E\{|e(k)|^2\}$ is evaluated at $\mathbf{h}^{(P)}$ and I_P is any P-element subset of $\{1, 2, \cdots, M\}$.

By perturbation analysis (see also [3]) we can show that all the vectors $\mathbf{h}^{(P)}$, $P = 2, 3, \cdots, M$, are unstable equilibria (saddle points) of $G_{42}(\mathbf{h})$. Next we examine the second-order derivatives of G_{42} at $\mathbf{h}^{(1)}$ which implies a solution of the type (14). By expanding $G_{42}(\mathbf{h})$ in a Taylor series around $\mathbf{h}^{(1)}$, we can establish that the class of solutions such as (14) (for any choice of j_0) is a "2-D plane" of locally stable equilibria (local maxima): any perturbation in (14) except for the one in d, only decreases $G_{r2}(\mathbf{h})$ below $G_{r2}(\mathbf{h}^{(1)})$; changes in d alone leave it unchanged.

The above discussion pertains to the stationary points of J_{42} w.r.t. the composite response $\overline{h}_j$. However, in practice, we can only adjust the beamformer weights. So what about the stationary points of J_{42} w.r.t. the beamformer coefficients? This aspect is discussed next. Using (8) and conditions **(GT-1)**-**(GT-2)**, it is easy to show that $\partial J_{42}(\mathbf{C})/\partial c_m = 0$ for $m = 1, 2, \cdots, N$ implies that

$$\frac{\partial G_{42}(\mathbf{h})}{\partial \overline{h}_j} = 0 \quad \text{for} \quad j = 1, 2, \cdots, M. \qquad (21)$$

Thus, all stationary points of J_{42} w.r.t. the beamformer coefficients are described by the stationary points of G_{42} w.r.t. the composite response $\overline{h}_j$.

Finally, the above discussion carries over to $J_{32}(\mathbf{C})$.

4. COST OPTIMIZATION

An iterative, batch, steepest descent method is used to optimize (12). The cost $J_{42}(\mathbf{C})$ is invariant to any scaling of $\mathbf{C}$. To rectify this, at every iteration the equalizer tap gain vector $\mathbf{C}$ are normalized to unit norm. Let $\widehat{J}_{42}(\mathbf{C})$ denote the data-based (12) with its explicit dependence upon $\mathbf{C}$. Let $\mathbf{C}$ denote the initial guess which sets all taps to be equal and unit norm. For calculation purposes, (6) is substituted in (12) to get

$$\widehat{J}_{42}(\mathbf{C}) = \frac{E\{|e(k)|^4\} - |E\{e^2(k)\}|^2}{|E\{|e(k)|^2\}|^2} - 2. \qquad (22)$$

The steps taken to optimize the cost $\widehat{J}_{42}(\mathbf{C})$ are

(i) Set the step size $\rho = 1$.

(ii) Calculate $\mathbf{C}' = \mathbf{C} + \rho \frac{\partial \widehat{J}_{42}(\mathbf{C})}{\partial \mathbf{C}^*}$ and the resulting cost $\widehat{J}_{42}(\mathbf{C}')$.

(iii) If $\widehat{J}_{42}(\mathbf{C}') > \widehat{J}_{42}(\mathbf{C})$, then accept $\mathbf{C}'' = \mathbf{C}'/\|\mathbf{C}'\|$ as the new equalizer vector, set $\mathbf{C} \leftarrow \mathbf{C}''$, and then go to (i). Else set $\rho = \rho/2$ and go to (ii).

5. ITERATIVE SOURCE SEPARATION

Maximization of (12) w.r.t. the vector $\mathbf{C}$ leads to the solution (16) . Given (16) we can estimate and remove the contribution of $s_{m_0}(k)$ from (1). Then we have a linear mixture with N measurements but $M-1$ sources (instead of M sources as in (1)-(2)). Repeat the process, i.e., maximize (12) w.r.t a new set of weights $\mathbf{C}$ to get a solution $e(k) = d' s_{m_0'}(k)$ where $m_0' \in (\{1, 2, ..., M\} - \{m_0\})$. Simple cross-correlation technique may be used for extracted source cancellation; a by-product of this is estimates of the gains $f_{im_0}, (i = 1, 2, ..., N)$, with the extracted $s_{m_0}(k)$ as input and the N measurements as outputs. This leads to the following procedure:

1. Maximize (12) w.r.t $\mathbf{C}$ to obtain (16).
2. Cross-correlate $\{e(k)\}$ (of (16)) with the given data (1) and define the estimates

$$\widehat{f}_{im_0} := \frac{E\{y_i(k)e^*(k)\}}{E\{|e(k)|^2\}} \quad \text{and} \quad \widehat{\gamma}_{m_0} := \mathrm{CUM}_4(e(k)), \qquad (23)$$

where $\widehat{\gamma}_{m_0}$ is the fourth-order estimated cumulant for the source m_0. The reconstructed contribution $\widehat{y}_{i,m_0}(k)$ of $e(k)$ to the data $y_i(k) (i = 1, 2, ..., M)$, is

$$\widehat{y}_{i,m_0}(k) := \widehat{f}_{im_0} e(k). \qquad (24)$$

3. Remove the above contribution from the data to define a linear mixture with N outputs and $M-1$ sources. These are given by

$$y_i'(k) := y_i(k) - \widehat{y}_{i,m_0}(k). \qquad (25)$$

4. If $M > 1$, set $M \leftarrow M-1, y_i(k) \leftarrow y_i'(k)$, and go back to step 1, else quit.

Analyzing the above algorithm we have

$$E\{y_i(k)e^*(k)\} = \sum_{m=1}^{M} f_{im} E\{s_m(k)e^*(k)\} = f_{im_0} d^* \sigma^2_{sm_0}. \qquad (26)$$

Using (26) in (23) we have

$$\widehat{f}_{im_0} = \frac{f_{im_0} d^* \sigma^2_{sm_0}}{|d|^2 \sigma^2_{sm_0}} = f_{im_0}/d. \qquad (27)$$

It follows from (24) and (27) that

$$\widehat{y}_{i,m_0}(k) = f_{im_0} s_{m_0}(k). \qquad (28)$$

Now use (25) and (28) to deduce that

$$y_i'(k) = \sum_{m=1, m \neq m_0}^{M} f_{im} s_m(k), \quad i = 1, 2, ..., N. \qquad (29)$$

Using (3)-(5), we get

$$\widehat{f}_{im_0} = e^{j[\omega_{m_0}(i-1) + \phi_{m_0}(i-1)^2]}, \quad i = 1, 2, ..., N, \qquad (30)$$

where

$$\omega_{m_0} = -2\pi \frac{d}{\lambda} \sin(\theta_{m_0}) \quad \text{and} \quad \phi_{m_0} = \pi \frac{d^2}{\lambda r_{m_0}} \cos^2(\theta_{m_0}). \qquad (31)$$

It follows from the preceding developments that under the conditions **(GT-1)**-**(GT-2)** and noise $\mathbf{n}(k) \equiv 0$, the

proposed iterative approach is capable of blind source separation and blind identification of the mixing matrix $\mathbf{F}$ up to a scaling and a permutation matrix. That is, given $\mathbf{F}$, we end up with a $\mathbf{G}$ where the two are related via

$$\mathbf{G} = \mathbf{FAP} \tag{32}$$

where $\mathbf{A}$ is an $M \times M$ diagonal scaling matrix (recall d in (14)), and $\mathbf{P}$ is an $M \times M$ permutation matrix (recall m_0 in (14), we don't "know" which source it refers to). When noise $\mathbf{n}(k) \not\equiv 0$, the resultant solution will be biased; how much is problem dependent. This bias motivates the cumulant-matching solution considered in Sec. 7. Since the noise is assumed to be Gaussian, cumulant matching will yield consistent estimates within the class (32).

6. SOURCE COORDINATE EXTRACTION

To estimate ω_{m_0} and ϕ_{m_0} given $\mathbf{F}$, consider

$$\widehat{x}_i \triangleq \frac{f_{(i+1)m_0}}{f_{im_0}} = e^{j[\omega_{m_0}+(2i-1)\phi_{m_0}]}, \quad i = 1, 2, ..., N-1. \tag{33}$$

Next define

$$\widehat{z}_i \triangleq \frac{\widehat{x}_{(i+1)m_0}}{\widehat{x}_{im_0}} = e^{j[2\phi_{m_0}]}, \quad i = 1, ..., N-2. \tag{34}$$

Solving (34) for $\widehat{\phi}_{m_0}$, we have

$$\widehat{\phi}_{m_0} = \frac{1}{N-2}\mathrm{Re}\left[\frac{1}{2j}\sum_{i=1}^{N-2}\ln \widehat{z}_i\right], \tag{35}$$

which yields the correct solution if no phase-wrapping problems occur. The only condition necessary for no phase wrapping to occur is

$$|2\widehat{\phi}_{m_0}| \leq \pi, \tag{36}$$

which follows from (34). From (36) and (31), we get

$$r_{m_0} \geq 2\frac{d^2}{\lambda}, \tag{37}$$

since $\cos^2(\theta_{m_0}) \leq 1$. Since $d \leq \lambda/2$, we get $r_{m_0} \geq \lambda/2$ as the condition for no phase wrapping to occur. Using (33) and (35), define

$$\widetilde{x}_i \triangleq \widehat{x}_i e^{-j\widehat{\phi}_{m_0}(2i-1)} \approx e^{j\omega_{m_0}}, \quad i = 1, ..., N-1. \tag{38}$$

We can then solve for $\widehat{\omega}_{m_0}$ as

$$\widehat{\omega}_{m_0} = \frac{1}{N-1}\mathrm{Re}\left[\frac{1}{j}\sum_{i=1}^{N-1}\ln \widetilde{x}_i\right]. \tag{39}$$

The source coordinates of the source m_0 can then be estimated as

$$\widehat{\theta}_{m_0} = \arcsin\left(-\frac{\widehat{\omega}_{m_0}}{2\pi d/\lambda}\right), \quad \widehat{r}_{m_0} = \pi\frac{d^2}{\widehat{\phi}_{m_0}\lambda}\cos^2(\widehat{\theta}_{m_0}). \tag{40}$$

To further refine the estimates of the source coordinates, a two-dimensional grid search is used in the vicinity. This was the method used in the simulations to get the initial guess for cumulant matching (direct approach).

7. CUMULANT MATCHING

7.1. Indirect Approach

Cumulant matching strives to reduce the error between the data-based cumulants and the model-based cumulants by varying the parameters. Let $\widehat{C}_4(i_1, i_2, i_3, i_4)$ denote the data-based consistent estimate of $C_4(i_1, i_2, i_3, i_4)$ where

$$C_4(i_1, i_2, i_3, i_4) = \mathrm{cum}_4(y_{i_1}(k), y^*_{i_2}(k), y_{i_3}(k), y^*_{i_4}(k)). \tag{41}$$

For model-based cumulant calculations we fix the fourth-order cumulant γ_m of the source m to be

$$\gamma_m = \mathrm{sgn}(\widehat{\gamma}_m), \tag{42}$$

where $\widehat{\gamma}_m$ is given by (23) and $\mathrm{sgn}(x)$ is the sign function. Any non-identity $|\gamma_m|$ is "absorbed" in f_{im}'s. The parameter-based cumulant is then given by

$$C_4(i_1, i_2, i_3, i_4|\mathbf{F}) = \sum_{m=1}^{M}\gamma_m f_{i_1 m} f_{i_2 m} f_{i_3 m} f_{i_4 m}, \tag{43}$$

where γ_m is fixed to the value given by (42) and $f_{i_1 m}, f_{i_2 m}, f_{i_3 m}$ and $f_{i_4 m}$ are initialized using

$$f_{im} = \widehat{f}_{im}(|\widehat{\gamma}_m|)^{1/4}, \tag{44}$$

where $\widehat{\gamma}_m$ and $\widehat{f}_{im}$ are given by (23). Cumulant matching is then used to solve

$$\min_{\mathbf{F}} \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{i_3=1}^{N}\sum_{i_4=1}^{N} |\widehat{C}_4(i_1, i_2, i_3, i_4) - C_4(i_1, i_2, i_3, i_4|\mathbf{F})|^2, \tag{45}$$

with the solution being consistent within the class specified by (32).

7.2. Direct Approach

This is as in Sec. 7.1 except that now each source is parametrized by three parameters: γ_m, θ_m and r_m. Therefore, the model-based cumulants are now parametrized by these three parameters using (3)-(5) and (43). We now have a much more parsimonious parametrization particularly when $N \gg M$.

8. SIMULATION EXAMPLES

8.1. Example 1.

We consider an example from [2]. There are two ($M = 2$) near-field non-Gaussian (BPSK) sources of equal power at locations ($\theta_1 = 20°, r_1 = 1.8\lambda$) and ($\theta_2 = -5°, r_2 = 5\lambda$). The number of sensors in the uniform linear array is 11 ($N = 11$). We considered two inter-element spacings: $d = \lambda/4$ as in [2], and $d = \lambda/2$ unlike [2]. The measurement noise is circular-white Gaussian and the number of data samples is $T = 1000$. Table I shows the results based upon 100 Monte Carlo runs for SNR = 10dB; the results quoted from [2] are for SNR = 15dB. It is seen that the proposed approaches outperform [2] (inspite of having lower SNR).

8.2. Example 2.

We modify the example from [2] to include the case of mixed sources. The source at location ($\theta_2 = -5°, r_2 = 5\lambda$) is now taken to be an exponential source; the rest is as in Example 1. Table II shows the results based upon 100 simulations for SNR = 10dB.

8.3. Example 3.

We consider the example from [9] and modify it to generate correlated noise. There are now two BPSK sources of equal power at locations $(\theta_1 = 5°, r_1 = 20\lambda)$ and $(\theta_2 = 17°, r_2 = 15\lambda)$. Gaussian sources could not be used as in [9] since the cumulant of a Gaussian random variable is zero. To generate correlated noise, a Gaussian source at location $(\theta_n = 11°, r_n = 25\lambda)$ was included. The number of sensors in the uniform linear array is 5 $(N = 5)$ and the inter-element spacing is taken to be $d = \lambda/2$. The measurement noise is circular-white Gaussian and the number of data samples is $T = 4000$. Table III shows the results based upon 100 simulations for SNR = 3dB.

Remark. For Examples 1 and 2, the SNR is computed as $\text{SNR} = \frac{S}{2\sigma_n^2}$ where S is the signal power given by

$$S = \frac{1}{N} \sum_{m=1}^{M} \sum_{i=1}^{N} |f_{im}|^2 E\{|s_m(k)|^2\}, \qquad (46)$$

and σ_n^2 is the noise variance given by $\sigma_n^2 = (2N)^{-1} \sum_{i=1}^{N} E\{|n_i(k)|^2\}$. Notice that (46) yields signal power summed over all sources; it is not average signal power per source. For Example 3, the SNR is computed as

$$\text{SNR} = \frac{S}{2(\sigma_n^2 + \frac{1}{N} \sum_{i=1}^{N} |g_i|^2 \sigma_n^2)}, \qquad (47)$$

where g_i is the gain at the i-th sensor for the Gaussian source.

9. REFERENCES

[1] S. Shamsunder and R.N. Challa, "High-order subspace based algorithms for passive localization of near-filed sources", in *Proc. 29th Asilomar Conf. Signals Systems Computers*, pp. 777-781, Pacific Grove, CA, Oct. 29 - Nov. 1, 1995.

[2] S. Shamsunder and R.N. Challa, "Subspace methods based on cumulants for near-field localization", *IEEE Trans. Signal Proc.*, in review.

[3] J.K. Tugnait, "Identification of multichannel linear non-Gaussian processes using higher-order statistics", in *Proc. 29th Asilomar Conf. Signals Systems Computers*, pp. 792-797, Pacific Grove, CA, Oct. 29 - Nov. 1, 1995. [Also *IEEE Trans. Signal Processing*, vol. SP-45, pp. 658-672, March 1997.]

[4] S. Haykin (Ed.), *Advances in Spectrum Analysis and Array Processing, Vol. III.* Prentice-Hall, 1995, pp.446-448.

[5] N. Yuen and B. Friedlander, "Asymptotic performance analysis of blind signal copy using fourth-order cumulants", *Intern. J. Adaptive Control & Signal Proc.*, vol. 10, pp.239-265, 1996.

[6] J.F. Cardoso and A. Soulomiac, "Blind beamforming for non-Gaussian signals", *IEE Proc.-F, Radar and Signal Processing*, vol. 140, pp.362-270, Dec. 1993.

[7] P. Comon, "Independent component analysis, a new concept?", *Signal Processing*, vol. 36, No. 3, pp.287-314, 1994.

[8] J.K. Tugnait, "On blind separation of convolutive mixtures of independent linear signals", in *Proc. Eighth IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp.312-315, Corfu, Greece, June 24-26, 1996.

[9] D. Starer and A. Nehorai, "Passive localization of near-field sources by path following", *IEEE Transactions on Signal Processing*, vol. SP-42, pp. 677-680, March 1994.

[10] M. Rosenblatt, *Stationary Sequences and Random Fields.* Birkhäuser: Boston, 1985.

[11] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing*, vol. 45, No. 1, pp. 59-83, July 31, 1995.

The approaches used in tables below are:
(i): Proposed (Inverse Filter: Secs. 5 & 6)
(ii): Proposed (Cumulant Matching: Sec. 7.2)
(iii): [2] (Subspace)
(iv): [2] (Cumulant Matching)

TABLE I: Results for Example 1 – $N = 11, M = 2$

2 BPSK sources
T = 1000, SNR = 10dB

Parameters		θ_1	r_1	θ_2	r_2
True Values		20°	1.8λ	−5°	5λ
$d = \lambda/4$					
Approach (i)	mean	20.555°	1.765λ	−5.022°	5.008λ
	σ	±0.156°	±0.010λ	±0.342°	±0.091λ
Approach (ii)	mean	20.010°	1.800λ	−5.001°	5.001λ
	σ	±0.452°	±0.027λ	±0.417°	±0.109λ
Approach (iii)	mean	19.95°	1.78λ	−4.94°	5.22λ
	σ	±0.67°	±0.05λ	±0.64°	±0.47λ
Approach (iv)	mean	19.95°	1.80λ	−4.94°	5.22λ
	σ	±0.32°	±0.12λ	±0.30°	±1.10λ
$d = \lambda/2$					
Approach (i)	mean	20.007°	1.800λ	−4.994°	4.999λ
	σ	±0.081°	±0.004λ	±0.080°	±0.019λ
Approach (ii)	mean	19.999°	1.800λ	−4.993°	4.999λ
	σ	±0.135°	±0.008λ	±0.129°	±0.032λ

TABLE II: Results for Example 2 – $N = 11, M = 2$

One BPSK and one exponential source
T = 1000, SNR = 10dB, $d = \lambda/4$

Parameters		θ_1	r_1	θ_2	r_2
True Values		20°	1.8λ	−5°	5λ
Approach (i)	mean	20.344°	1.779λ	−5.136°	5.052λ
	σ	±0.792°	±0.049λ	±0.858°	±0.256λ
Approach (ii)	mean	19.874°	1.810λ	−5.072°	5.037λ
	σ	±1.131°	±0.079λ	±1.132°	±0.318λ

TABLE III: Results for Example 3 - $N = 5, M = 2$

2 BPSK sources
T = 4000, SNR = 3dB, $d = \lambda/2$

Parameters		θ_1	r_1	θ_2	r_2
True Values		5°	20λ	17°	15λ
Approach (i)	mean	4.947°	20.507λ	16.862°	15.301λ
	σ	±0.314°	±1.947λ	±0.159°	±0.647λ
Approach (ii)	mean	4.995°	20.334λ	16.957°	15.337λ
	σ	±0.404°	±2.611λ	±0.382°	±1.471λ

TPP: Source Separation & Array Processing

Chair: Joel LeRoux
University of Nice, France

Covariance of Finite-Sample Cumulants in Array-Processing

Thomas Kaiser*
Gerhard-Mercator-University
Department of Communication Engineering
FB 9, FG NT, Bismarckstrasse 81
47048 Duisburg, Germany
th.kaiser@uni-duisburg.de

Jerry M. Mendel
University of Southern California
Signal and Image Processing Institute
3740 McClintock Ave
Los Angeles, CA 90089-2564, USA
mendel@sipi.usc.edu

Abstract

In this paper we provide explicit formulas for the covariances of second-, third-, and fourth-order sample cumulants as used in narrowband array processing. These covariances provide a basis for analysing the performance of cumulant based algorithms for finite-sample length, which is in contrast to usual asymptotic performance analyses. The use of these formulas, which consist of several thousand terms, will be demonstrated, and a rough idea of their applicability to a performance analysis for finite numbers of samples will be given.

1 Introduction

Since existing second-order statistics based methods cannot solve numerous problems in narrowband array signal processing, many higher-order cumulant based algorithms have been recently proposed. For example, the virtual-ESPRIT-Algorithm (VESPA) for direction-finding and recovery of independent sources can also calibrate an array of unknown configuration (see *Doğan and Mendel* [5]). Also VESPA can be extended to direction-finding of highly correlated or coherent sources (see *Gőnen et al.* [8]), which is often the case in practice due to multipath propagation or "smart" jamming. Furthermore, since higher-order cumulants are theoretically insensitive to additive white or colored Gaussian noise, they always suppress such noise. These new higher-order cumulant based algorithms require a performance analysis not only among themselves, but also with second-order statistics based algorithms.

In the past decade, array signal processing methods based on eigen-decomposition of the covariance matrix have received considerable attention, since they provide high resolution with low computational complexity. *Xu and Buckley* [12] presented a bias analysis of the MUSIC location estimator using a second-order Taylor expansion of the derivative of the null spectrum. They show in their simulations that the bias expression can be accurately applied even for a very limited number N of independent snapshots ($N = 20$). *Reed et al.* [11] derived an exact output signal-to-interference-plus-noise ratio (SINR) formula as a function of N for an optimum beamformer. *Feldman and Griffiths* [7] also investigated the SINR formula of Reed and showed by simulation, that this formula can be closely approximated by a simple expression if N is larger than approximately 50. These results show that, perhaps, a small number of samples is sufficient for practical use of many asymptotic performance analyses.

Some performance analyses for higher-order cumulant based methods have already been done. *Cardoso and Moulines* [2] have derived closed-form expressions of the asymptotic covariance of MUSIC-like direction-of-arrival (DOA) estimates based on two different fourth-order cumulant matrices. *Yuen and Friedlander* [14] presented an asymptotic performance analysis of ESPRIT [13], higher-order ESPRIT [4], and VESPA.

All these performance analyses are based on the *asymptotic* covariances of second-, or fourth-order sample cumulants. To provide a basis for investigation of the finite-sample case, we will derive covariances of the second-, third- and fourth-order sample cumulants, the *finite-sample* covariances. Given these results, we can address the question of whether higher-order cumulant-based algorithms need more samples than second-order cumulant based algorithms, and we can also study adaptive algorithms.

2 Problem Formulation

Let P narrowband plane waves centered at a known frequency ω_0 impinge on an arbitrary array composed of M sensors. The received signal vector $\mathbf{r}(t) =$

*This work was supported by a scholarship of the NATO Scientific Board and co-ordinated by the Deutscher Akademischer Austauschdienst.

0-8186-8005-9/97 $10.00 © 1997 IEEE

$(r_1(t),...,r_M(t))^T$ can be modeled as

$$\mathbf{r}(t) = \mathbf{A}(\phi,\vartheta)\,\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{s}(t) = (s_1(t),...,s_P(t))^T$ is a $P \times 1$ vector which contains the P zero-mean non-Gaussian source signal complex envelopes at time t, and $\mathbf{A}(\phi,\vartheta)$ is an $M \times P$ steering matrix. $\phi \in [-\pi,\pi]$ is the azimuth angle, $\vartheta \in [0,\pi]$ is the elevation angle and $\mathbf{n}(t)$ is an $M \times 1$ vector composed of M zero-mean noise signals. Any noise signal $n_m(t)$, $m = 1(1)M$ is[1] independent of any source signal $s_p(t)$, $p = 1(1)P$. Furthermore, all signals $s_p(t)$, $n_m(t)$ are modeled as sequences of independent and identically, but *arbitrarily* distributed, complex random variables with finite moments up to the eighth-order. This model is commonly used in the far-field case, but it can be extended to coherent signals [8] and to the near-field case [3].

In this paper, we focus our attention on *second-*, *third-*, and *fourth-order cumulants*, which are respectively defined as

$$c_{k,l} = \mathrm{E}\{r_k(t)r_l^*(t)\} \tag{2}$$

$$c_{k,l,m} = \mathrm{E}\{r_k(t)r_l^*(t)r_m(t)\} \tag{3}$$

$$\begin{aligned} c_{k,l,m,n} = {} & \mathrm{E}\{r_k(t)r_l^*(t)r_m^*(t)r_n(t)\} \\ & -\mathrm{E}\{r_k(t)r_l^*(t)\}\,\mathrm{E}\{r_m^*(t)r_n(t)\} \\ & -\mathrm{E}\{r_k(t)r_m^*(t)\}\,\mathrm{E}\{r_l^*(t)r_n(t)\} \\ & -\mathrm{E}\{r_k(t)r_n(t)\}\,\mathrm{E}\{r_l^*(t)r_m^*(t)\}, \end{aligned} \tag{4}$$

since $\mathrm{E}\{r_m(t)\} = 0\,\forall\, m = 1(1)M$. The second-order, $\hat{c}_{k,l}$, and third-order *sample* cumulants, $\hat{c}_{k,l,m}$, are obtained by replacing the expected values, $\mathrm{E}\{\}$, by time averages $\frac{1}{N}\sum_{t=1}^{N}$, where $N \in \mathbf{N}$ is the number of samples . In contrast, the fourth-order *sample* cumulant $\hat{c}_{k,l,m,n}$ is estimated more generally by

$$\begin{aligned} \hat{c}_{k,l,m,n} = {} & \frac{1}{\alpha}\sum_{t=1}^{N} r_k(t)r_l^*(t)r_m^*(t)r_n(t) \\ & -\frac{1}{\beta}\sum_{t=1}^{N} r_k(t)r_l^*(t)\sum_{p=1}^{N} r_m^*(p)r_n(p) \\ & -\frac{1}{\beta}\sum_{t=1}^{N} r_k(t)r_m^*(t)\sum_{p=1}^{N} r_l^*(p)r_n(p) \\ & -\frac{1}{\beta}\sum_{t=1}^{N} r_k(t)r_n(t)\sum_{p=1}^{N} r_l^*(p)r_m^*(p), \end{aligned} \tag{5}$$

where α and β are functions of N. If the fourth-order sample cumulant is an *unbiased* estimate,

$$\alpha = \frac{N^2-N}{N+2}, \quad \beta = N^2 - N, \quad N > 1, \tag{6}$$

[1] $m = 1(1)M$ is a more compact form of $m = 1,2,...,M$. The number in parentheses means the increment of the sequence.

otherwise, the commonly used eqs.

$$\alpha = N, \quad \beta = N^2, \quad N > 0, \tag{7}$$

lead to a *biased estimator*. Note that the second- and the third-order sample cumulants are always unbiased.

3 The Finite-Sample Covariances

To demonstrate the tedious calculations of the finite-sample covariances, first the derivation of the second-order finite-sample covariance, and then the basic equation of the fourth-order finite-sample covariance will be given. All these formulas are published in different computer languages on the World-Wide-Web under `http://fb9nt-ln.uni-duisburg.de/mitarbeiter/kaiser/hoswshop.97/hoswshop97.html` and also in [9].

The second-order finite-sample covariance can be written as

$$\begin{aligned} \mathrm{Cov}\left(\hat{c}_{k_1,l_1},\hat{c}^*_{k_2,l_2}\right) &= \mathrm{E}\left\{\hat{c}_{k_1,l_1}\hat{c}^*_{k_2,l_2}\right\} - \mathrm{E}\left\{\hat{c}_{k_1,l_1}\right\}\mathrm{E}\left\{\hat{c}^*_{k_2,l_2}\right\} \\ &= \frac{1}{N}\left(\mathrm{E}\left\{r_{k_1}(t)r^*_{l_1}(t)r^*_{k_2}(t)r_{l_2}(t)\right\}\right. \\ &\quad \left. - \mathrm{E}\left\{r_{k_1}(t)r^*_{l_1}(t)\right\}\mathrm{E}\left\{r^*_{k_2}(t)r_{l_2}(t)\right\}\right). \end{aligned} \tag{8}$$

and is therefore equal to the asymptotic covariance

$$\mathrm{Cov}^{\infty}\left(\hat{c}_{k_1,l_1},\hat{c}^*_{k_2,l_2}\right) = \lim_{N\to\infty} N\mathrm{Cov}\left(\hat{c}_{k_1,l_1},\hat{c}^*_{k_2,l_2}\right) \tag{9}$$

divided by N. Equivalently, the third-order finite-sample covariance is given by

$$\begin{aligned} &\mathrm{Cov}\left(\hat{c}_{k_1,l_1,m_1},\hat{c}^*_{k_2,l_2,m_2}\right) = \\ &\quad \frac{1}{N}\left(\mathrm{E}\left\{r_{k_1}(t)r^*_{l_1}(t)r_{m_1}(t)r^*_{k_2}(t)r_{l_2}(t)r^*_{m_2}(t)\right\} - \right. \\ &\qquad \left. \mathrm{E}\left\{r_{k_1}(t)r^*_{l_1}(t)r_{m_1}(t)\right\}\mathrm{E}\left\{r^*_{k_2}(t)r_{l_2}(t)r^*_{m_2}(t)\right\}\right), \end{aligned}$$

and is also equal to the asymptotic covariance $\mathrm{Cov}^{\infty}\left(\hat{c}_{k_1,l_1,m_1},\hat{c}^*_{k_2,l_2,m_2}\right)$ divided by N. This means that any asymptotic performance analyses based on second-, and third-order cumulants are also valid for the finite-sample case. Unfortunately, as we will see later, this property does not hold for fourth-order sample cumulants.

For further evaluation of (8) the fourth- and the second-order moments of $r_k(t)$ must be calculated using (1). We will omit the variables t, ϕ, ϑ in the following, to simplify the notation. After some algebra, we obtain

$$\begin{aligned} &\mathrm{E}\left\{r_{k_1}r^*_{l_1}r^*_{k_2}r_{l_2}\right\} = \\ &\mathrm{E}\{\mathbf{a}^T_{k_1}\mathbf{s}\mathbf{s}^H\mathbf{a}^*_{l_1}\mathbf{s}^H\mathbf{a}^*_{k_2}\mathbf{a}^T_{l_2}\mathbf{s}\} + \mathrm{E}\{\mathbf{a}^T_{k_1}\mathbf{s}\mathbf{s}^H\mathbf{a}^*_{l_1}\}\mathrm{E}\{n^*_{k_2}n_{l_2}\} + \\ &\mathrm{E}\{\mathbf{a}^T_{k_1}\mathbf{s}\mathbf{s}^H\mathbf{a}^*_{k_2}\}\mathrm{E}\{n^*_{l_1}n_{l_2}\} + \mathrm{E}\{\mathbf{a}^T_{k_1}\mathbf{s}\mathbf{a}^T_{l_2}\mathbf{s}\}\mathrm{E}\{n^*_{l_1}n^*_{k_2}\} + \\ &\mathrm{E}\{\mathbf{s}^H\mathbf{a}^*_{l_1}\mathbf{s}^H\mathbf{a}^*_{k_2}\}\mathrm{E}\{n_{k_1}n_{l_2}\} + \mathrm{E}\{\mathbf{s}^H\mathbf{a}^*_{l_1}\mathbf{a}^T_{l_2}\mathbf{s}\}\mathrm{E}\{n_{k_1}n^*_{k_2}\} \\ &+\mathrm{E}\{\mathbf{s}^H\mathbf{a}^*_{k_2}\mathbf{a}^T_{l_2}\mathbf{s}\}\mathrm{E}\{n_{k_1}n^*_{l_1}\} + \mathrm{E}\{n_{k_1}n^*_{l_1}n^*_{k_2}n_{l_2}\} \end{aligned} \tag{10}$$

where $\mathbf{a}_{k_1} = (a_{1,k_1}, ..., a_{M,k_1})^T$ is the k_1-th column vector of the steering matrix $\mathbf{A}$ and the elements of $\mathbf{a}_{k_1}$ can be written as

$$a_{m,k_1} = g_m(\phi_{k_1}, \vartheta_{k_1}) \, e^{j\frac{2\pi}{\lambda}(x_m \sin\vartheta_{k_1} \cos\phi_{k_1} + y_m \sin\vartheta_{k_1} \sin\phi_{k_1} + z_m \cos\vartheta_{k_1})} \quad (11)$$

in which $g_m(\phi_{k_1}, \vartheta_{k_1})$ is the response of the m-th sensor on a source signal impinging on the array with angles ϕ_{k_1}, ϑ_{k_1}, (x_m, y_m, z_m) are the coordinates of the m-th sensor, $\lambda = 2\pi c/\omega_0$ is the source wavelength, and c is the propagation speed of the plane wave. To further calculate (10), the random variables ($\mathbf{s}$) must be separated from the deterministic variables ($\mathbf{a}_k$). Since we will also be faced with this problem for third-, and fourth-order finite-sample covariances, we look for a general notation. Between the scalars $\mathbf{s}^H\mathbf{a}_k^*$, $\mathbf{a}_k^T\mathbf{s}$, we can simply introduce a Kronecker product ($\otimes$) (instead of transposing). Using the rule $(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A}\otimes\mathbf{C})\,(\mathbf{B}\otimes\mathbf{D})$ we obtain, after some algebra,

$$\begin{aligned}
\mathrm{Cov}\left(\hat{c}_{k_1,l_1}, \hat{c}^*_{k_2,l_2}\right) &= \frac{1}{N}\left(\left(\mathbf{a}_{k_1}^T \otimes \mathbf{a}_{l_1}^H \otimes \mathbf{a}_{k_2}^H \otimes \mathbf{a}_{l_2}^T\right)\right. \\
&\left(\mathrm{E}\left\{\mathbf{s}\otimes\mathbf{s}^*\otimes\mathbf{s}^*\otimes\mathbf{s}\right\} - \mathrm{E}\left\{\mathbf{s}\otimes\mathbf{s}^*\right\} \otimes \mathrm{E}\left\{\mathbf{s}^*\otimes\mathbf{s}\right\}\right) + \\
&\left(\mathbf{a}_{k_1}^T \otimes \mathbf{a}_{k_2}^H\right) \mathrm{E}\left\{\mathbf{s}\otimes\mathbf{s}^*\right\} \mathrm{E}\left\{n^*_{l_1} n_{l_2}\right\} + \\
&\left(\mathbf{a}_{k_1}^T \otimes \mathbf{a}_{l_2}^T\right) \mathrm{E}\left\{\mathbf{s}\otimes\mathbf{s}\right\} \mathrm{E}\left\{n^*_{l_1} n^*_{k_2}\right\} + \\
&\left(\mathbf{a}_{l_1}^H \otimes \mathbf{a}_{k_2}^H\right) \mathrm{E}\left\{\mathbf{s}^*\otimes\mathbf{s}^*\right\} \mathrm{E}\left\{n_{k_1} n_{l_2}\right\} + \\
&\left(\mathbf{a}_{l_1}^H \otimes \mathbf{a}_{l_2}^T\right) \mathrm{E}\left\{\mathbf{s}^*\otimes\mathbf{s}\right\} \mathrm{E}\left\{n_{k_1} n^*_{k_2}\right\} + \\
&\left.\mathrm{E}\left\{n_{k_1} n^*_{l_1} n^*_{k_2} n_{l_2}\right\} - \mathrm{E}\left\{n_{k_1} n^*_{l_1}\right\} \mathrm{E}\left\{n^*_{k_2} n_{l_2}\right\}\right). \quad (12)
\end{aligned}$$

Given the statistics of the source- and noise signals up to order 4, the second-order finite-sample covariance can be calculated.

Due to lack of space, we omit the formula for the third-order finite-sample covariance. The fourth-order finite-sample covariance can be written, after tedious calculations, as

$$\begin{aligned}
&\mathrm{Cov}\left(\hat{c}_{k_1,l_1,m_1,n_1}, \hat{c}^*_{k_2,l_2,m_2,n_2}\right) = \\
&\frac{(3\alpha-\beta)^2 N}{\alpha^2\beta^2}\left(\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1} r^*_{k_2} r_{n_2} r_{m_2} r^*_{n_2}\} - \right. \\
&\qquad \left.\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} r_{n_2} r_{m_2} r^*_{n_2}\}\right) + \\
&\frac{(3\alpha-\beta)N(N-1)}{\alpha\beta^2}\left(\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1} r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{k_2} r_{n_2}\} + \right. \\
&\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1} r_{n_2} r^*_{n_2}\}\mathrm{E}\{r^*_{k_2} r_{m_2}\} + \mathrm{E}\{r_{k_1} r^*_{m_1} r_{n_1} r_{m_2} r_{n_2}\}\mathrm{E}\{r^*_{k_2} r^*_{n_2}\} + \\
&\mathrm{E}\{r_{k_1} r^*_{k_2} r^*_{n_2} r_{n_1} r_{m_2} r_{n_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1}\} - 2\mathrm{E}\{r_{k_1} r_{n_1} r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{k_2} n_2\}\mathrm{E}\{r^*_{l_1} r^*_{m_1}\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} n_2 r^*_{n_2} r_{m_2} r^*_{l_1} r^*_{m_1}\} + \mathrm{E}\{r^*_{l_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r_{k_1} r^*_{k_2} n_2 r_{m_2} r^*_{n_2}\} - \\
&2\mathrm{E}\{r_{k_1} r^*_{m_1}\}\mathrm{E}\{r^*_{k_2} r_{m_2} n_2 r^*_{n_2}\}\mathrm{E}\{r_{n_1} r^*_{l_1}\} + \mathrm{E}\{r_{k_1} r^*_{m_1}\}\mathrm{E}\{r^*_{k_2} n_2 r^*_{l_1} r_{n_1} r_{m_2} r^*_{n_2}\} + \\
&\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1} r^*_{k_2} r^*_{n_2}\}\mathrm{E}\{r_{m_2} n_2\} - 2\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} r^*_{n_2}\}\mathrm{E}\{r_{m_2} n_2\} + \\
&\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1} r^*_{k_2} r_{m_2}\}\mathrm{E}\{r^*_{n_2} n_2\} - 2\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{m_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} r_{m_2}\}\mathrm{E}\{r^*_{n_2} n_2\} + \\
&\mathrm{E}\{r_{k_1} r^*_{l_1} r^*_{k_2} n_2 r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{m_1} r_{n_1}\} + \mathrm{E}\{r_{k_1} r^*_{l_1}\}\mathrm{E}\{r^*_{m_1} r_{n_1} r^*_{n_2} r_{m_2} r^*_{k_2} n_2\} + \\
&\mathrm{E}\{r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{l_1} r_{k_1} r^*_{m_1} r_{n_1} r_{k_2} n_2\} - 2\mathrm{E}\{r^*_{l_1} r_{k_1} r^*_{m_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} n_2\}\mathrm{E}\{r^*_{n_2} r_{m_2}\}) + \\
&\frac{(N^2-3N+2)N}{\beta^2}\left(\mathrm{E}\{r_{k_1} r_{n_1} r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1}\}\mathrm{E}\{r^*_{k_2} n_2\} + \right. \\
&\qquad \left.\mathrm{E}\{r_{k_1} r_{n_1} r^*_{n_2} n_2\}\mathrm{E}\{r^*_{k_2} r_{m_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1}\} + \right. \\
&\mathrm{E}\{r_{k_1} r_{n_1} r_{m_2} r_{n_2}\}\mathrm{E}\{r^*_{k_2} r^*_{n_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1}\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} r^*_{n_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1} r_{m_2} n_2\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{k_2} r_{m_2}\}\mathrm{E}\{r^*_{l_1} r^*_{m_1} r^*_{n_2} n_2\} + \mathrm{E}\{r_{k_1} r^*_{m_1}\}\mathrm{E}\{r^*_{k_2} n_2\}\mathrm{E}\{r^*_{l_1} r_{n_1}\} + \\
&\mathrm{E}\{r_{k_1} r^*_{m_1}\}\mathrm{E}\{r_{m_2} r^*_{k_2}\}\mathrm{E}\{r^*_{l_1} r_{n_1} n_2 r^*_{n_2}\} + \mathrm{E}\{r_{k_1} r^*_{m_1}\}\mathrm{E}\{r_{m_2} r^*_{n_2}\}\mathrm{E}\{r^*_{l_1} r_{n_1} n_2\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{k_2} n_2 r^*_{n_2}\}\mathrm{E}\{r_{m_2}\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r_{m_2} r^*_{k_2} n_2\}\mathrm{E}\{r^*_{n_2}\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{k_2} r_{m_2} r^*_{n_2}\}\mathrm{E}\{n_2\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{k_2} n_2\}\mathrm{E}\{r_{m_2} r^*_{n_2}\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r_{m_2} r^*_{k_2} n_2\}\mathrm{E}\{r^*_{n_2}\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{k_2} n_2\}\mathrm{E}\{r_{m_2} r^*_{n_2}\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r_{m_2} r^*_{n_2}\}\mathrm{E}\{n_2\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{n_2} n_2\}\mathrm{E}\{r_{m_2}\} + \\
&\mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r_{m_2} n_2\}\mathrm{E}\{r^*_{n_2}\} + \mathrm{E}\{r_{k_1} r_{n_1}\}\mathrm{E}\{r^*_{m_1} r^*_{n_2} n_2\}\mathrm{E}\{r_{m

4 Simulations and Analytic Evaluations

The purposes of this Section are twofold, namely: 1) to "verify" the formulas, especially the fourth-order finite-sample covariance, by Monte-Carlo simulation and analytic evaluation; and, 2) to give a rough idea of applicability of an asymptotic performance analysis for finite N. For the second purpose we consider the *relative error* between the *finite-sample* and the *asymptotic* variance of the fourth-order sample cumulant

$$e_{k,l,m,n} = \frac{N\mathrm{Var}\left(\hat{c}_{k,l,m,n}\right) - \mathrm{Var}^{\infty}\left(\hat{c}_{k,l,m,n}\right)}{N\mathrm{Var}\left(\hat{c}_{k,l,m,n}\right)}. \qquad (16)$$

where Var^{∞} is defined equivalently to (9). In the case of the biased estimator (7), $e_{k,l,m,n}$ can be written as

$$e_{k,l,m,n} = \frac{Nb_2 + b_3}{N^2 b_1 + Nb_2 + b_3} \overset{N\gg 1}{\approx} \frac{b_2}{b_1 N}, \qquad (17)$$

and, for the unbiased estimator (6),

$$e_{k,l,m,n} = \frac{b_5 + b_6}{(N-1)b_4 + Nb_5 + b_6} \overset{N\gg 1}{\approx} \frac{b_5 + b_6}{(b_4 + b_5)N}. \qquad (18)$$

4.1 Example 1: Comparison of Monte-Carlo Simulations and Analytic Evaluation

To demonstrate the correctness of our formulas, we avoid symmetries in the parameters (e. g., we choose a non-linear array) and important parameters are not zero (e. g., we choose skewed distributed noise with odd moments not to zero). We consider a non-linear array [5] of $M = 5$ radar antennas with response $g_m(\phi, \vartheta) = 1/\sin\vartheta$ [1]. Two 64-QAM signals consist of $N = 20$ samples impinging on this array from directions $\phi_1 = 0^o$, $\phi_1 = 20^o$, $\vartheta_1 = \vartheta_2 = 40^o$ with $\mathrm{E}\left\{|s_i(t)|^2\right\} = 1$ for $i = 1, 2$. The signal-to-noise ratio varies from $SNR = -60\mathrm{dB}(10\mathrm{dB})60\mathrm{dB}$ and the additive noise is logarithmic Gaussian distributed with skewness

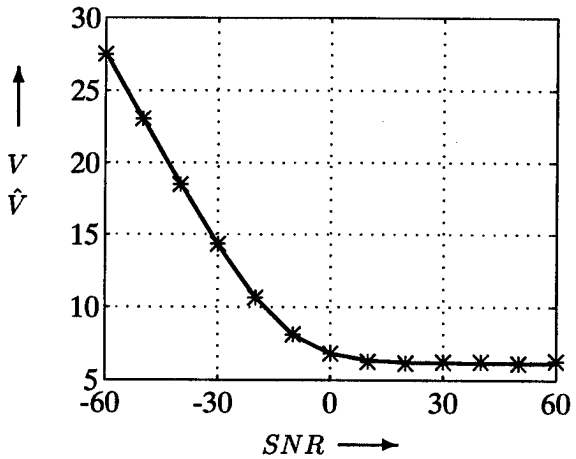


Figure 1. $V = \log_{10}\mathrm{Var}\left(\hat{c}_{4,3,2,1}\right)$ **(solid) and** $\hat{V} = \log_{10}\widehat{\mathrm{Var}}\left(\hat{c}_{4,3,2,1}\right)$ **(stars) for** $N = 20$ **samples.**

Fig. 1 shows a very good agreement between the analytically calculated variance V and the estimated variance $\hat{V}$ (by Monte-Carlo simulations). The clear increase of the variance for low SNR comes from the large higher-order *noise* moments. Equivalently, the slightly decrease for high SNR comes from the small higher-order *signal* moments. Figure 2 shows V and $\hat{V}$ for SNR fixed to 0dB and the second azimuth angle $\phi_2 = -20^o(1^o)20^o$. All other parameters remain unchanged. It can be seen that the agreement between V and $\hat{V}$ is not as good as in Fig. 1, especially for $\phi_2 \approx \phi_1 = 0$. This means that, for different parameters (SNR, ϕ), a different number of Monte-Carlo simulations is necessary to get reliable results about the (co)variance of finite-sample cumulants. This can be avoided by use of our analytic formulas. The example's results demonstrate the validity of our formulas; however, numerous other experiments, which are not shown here, also confirm this.

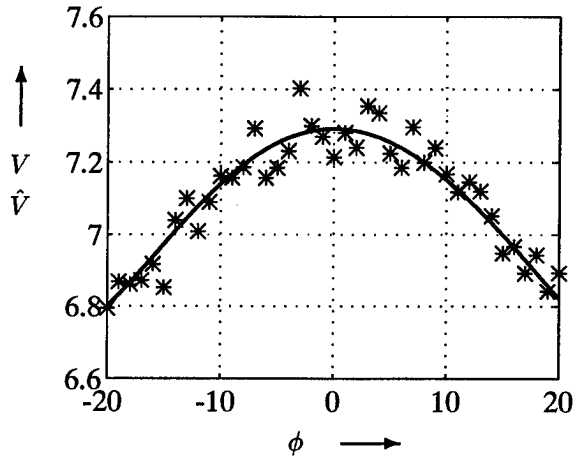


Figure 2. $V = \log_{10}\mathrm{Var}\left(\hat{c}_{4,3,2,1}\right)$ **(solid) and** $\hat{V} = \log_{10}\widehat{\mathrm{Var}}\left(\hat{c}_{4,3,2,1}\right)$ **(stars) for** $N = 20$ **samples.**

4.2 Example 2: Dependence of $e_{k,l,m,n}$ on SNR, M, P, ϕ, θ

In this example, we assume a linear array of 10 omnidirectional ($g_m(\phi_k, \vartheta_k) = 1\ \forall k = 1(1)P,\ m = 1(1)M$) sensors with spacing $\lambda/2$. The azimuth and elevation angles are $\phi = 0^o$, $\vartheta = 0^o$ and (k, l, m, n) is fixed to $(1, 1, 1, 1)$. $s(t)$ is a BPSK-signal , $SNR = -20(10)20$, and $P = 1$. For $N \gg 1$ the behaviour of both relative errors in the Fig. 3 plots is linear as expected from (17), (18). On the other hand, the curves are non-linear for small values of $N \leq 50$, which is caused by b_3 or b_6. Somewhat surprisingly, the relative error for $SNR = -20\mathrm{dB}$ exhibits a sharp break near $N = 8$ in the case of the biased estimator, because $Nb_2 + b_3$ is near zero for $N = 8$. This means that the finite-sample and asymptotic variance are approximately equal for

$N = 8$. Therefore, it can happen that an asymptotic performance analysis could be in nearly perfect agreement to a Monte-Carlo simulation even for very small N. In addition, the biased estimator leads to a smaller relative error for low SNR, but, for higher SNR, the relative error of the unbiased estimator is smaller. Other experiments have shown that not only the arrival angles ϕ, ϑ, but also the number of sensors M have no impact on the upper curves. This also remains true for two equal-powered sources.

Observe from Fig. 3, that only $N \approx 300$ sample are necessary for a relative error of 0.01; hence, asymptotic performance analysis can often be applied with sufficient accuracy in narrowband array processing for a small number of samples and for low SNR.

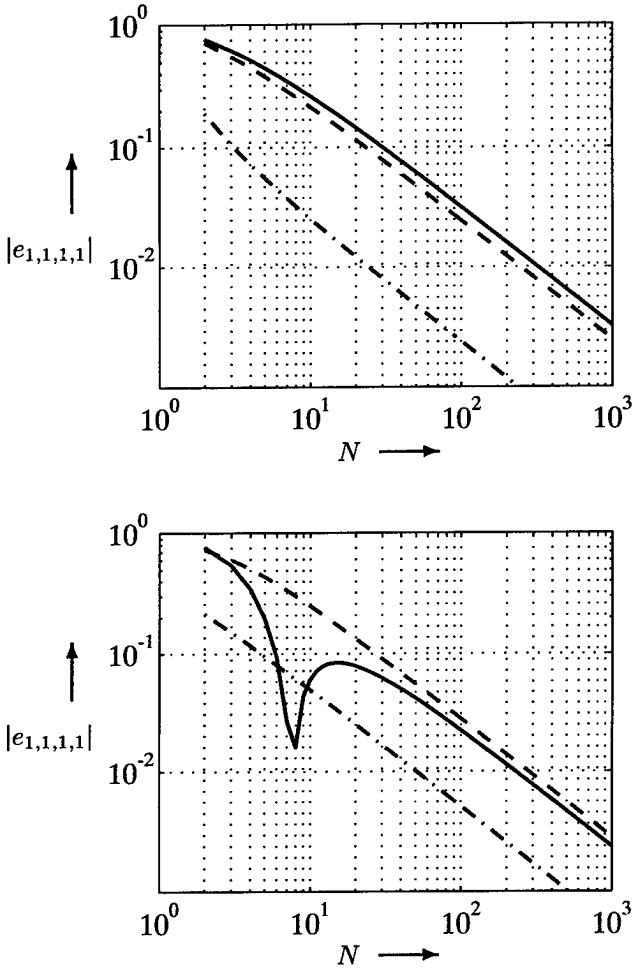


Figure 3. Relative error as a function of N for the unbiased- (upper) and biased (lower) fourth-order sample cumulant at different SNR ($SNR = -20$dB solid, $SNR = 0$dB dashed, $SNR = 20$dB dash-dotted).

5 Conclusions

In this paper we have provided explicit formulas for the covariances of second-, third-, and fourth-order sample cumulants for narrowband array processing. These equations can be used not only for finite-sample performance analyses of cumulant-based algorithms but also for optimization of some parameters to obtain highest estimation accuracy, e. g., the kind of source signals. We demonstrated the correct working of these formulas by examples and showed that asymptotic performance analysis can often be applied with sufficient accuracy in narrowband array processing for small number of samples and for low SNR.

References

[1] A. Brunner, "Possibilities of Dimensioning Doubly Curved Reflectors for Azimuth-Search Radar Antennas," *IEEE Trans. on Ant. and Prop.*, vol. AP-19, no. 1, pp. 52-57, Jan. 1971.

[2] J. F. Cardoso and E. Moulines, "Asymptotic Performance Analysis of Direction-Finding Algorithms Based on Fourth-Order Cumulants", *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp.214-224, January 1995.

[3] R. N. Challa and S. Shamsunder, "Higher-Order subspace Algorithms for Passive Localizationof Near-Field Sources", *Twenty-Ninth Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, October 30 - November 1, 1995.

[4] H. H. Chiang and C. L. Nikias, "The ESPRIT algorithm with higher-order statistics", *Proc. of the Vail Workshop on Higher-Order Spectral Analysis*, pp. 163-168, Vail, CO, June 1989.

[5] M. C. Doğan and J. M. Mendel, "Applications of Cumulants to Array Processing", Part I: Aperture extension and Array Calibration", *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1200-1216, May 1995.

[6] M. C. Doğan and J. M. Mendel, "Cumulant-Based Blind Optimum Beamforming", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 722-741, July 1994.

[7] D. D. Feldman and L. J. Griffiths, "A Constraint Projection Approach For Robust Adaptive BEamforming", Proceedings of ICASSP, vol. 2, pp. 1381-1384, May 1991.

[8] E. Gönen, M. C. Doğan and J. M. Mendel, "Applications of Cumulants to Array Processing: Direction-Finding in Coherent Signal Environment", *Twenty-Eight Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, October 31-November 2, 1994.

[9] T. Kaiser and J. M. Mendel, "Finite-Sample Covariances of Second-, Third-, and Fourth-Order Sample Cumulants in Narrowband Array Processing," *Technical Report, Signal and Image Processing Institute, University of Southern California*, Los Angeles, September 1996.

[10] T. Kaiser and J. M. Mendel, "Which Cumulants should be used for Steering Vector Estimation ?," *IEEE Sig. Proc. Workshop on Higher-Order Statistics*, Banff, Canada, July 1997.

[11] I.S. Reed, J.D. Mallet and L.E. Brennan, "Rapid Convergence Rate in Adaptive Arrays", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 10, pp. 853-863, November 1974.

[12] X-L. Xu and K. M. Buckley, "Bias Analysis of the MUSIC Location Estimator", *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2559-2569, October 1992.

[13] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proceedings of the IEEE*, vol. 74, no. 7, pp. 1044-1045, July 1986.

[14] N. Yuen and B. Friedlander, "Asymptotic Performance Analysis of ESPRIT, Higher Order ESPRIT, and Virtual ESPRIT Algorithms," *IEEE Trans. on Sig. Proc.*, vol. 44, no. 10, Oct. 1996.

DOA Estimation Using Blind Separation of Sources

Mehrez Hirari
Earth Observation Research Center
National Space Agency of Japan.
E-mail: mehrez@aurora.ee.uec.ac.jp.

Masashi Hayakawa
Dept. of Electronic Engineering,
The Univ. of Electro-Communications,
Chofu city, Tokyo.

Abstract

In the present paper we propose a new approach for the estimation of DOA for polarized EM waves using blind separation of sources. In this approach we use a vector-sensor, a sensor whose output is a complete set of the EM field components of the irradiating wave and we reconstruct the waveforms of all the original signals; that is, all the EM components of the source's field. The blind separation of sources is made iteratively using a recurrent Hopfield-like single layer neural network. The simulation results for two sources have been investigated. We have considered coherent and incoherent sources, and also the case of varying DOA's vis-à-vis to the sensor and a varying polarization.

1 Introduction

The problem of estimating the DOA has so far been tackled using different approaches. Most of them employed sensor arrays to carry out space-time processing by means of spatial filtering or beamforming. The performance of this approach is, however, directly dependent upon the physical size of the array, regardless of the size and quality (in the sense of signal-to-noise ratio) of the available data [1, 2]. The limitation of this approach due to its poor resolution (ability to distinguish between two sources) and the need of extension to more than one source required in many applications have motivated the research in statistical signal processing. This finally resulted in the emergence of the signal parameter estimation approach as an active area. Important results include, to cite only a few, AR/MA, maximum entropy (ME)[3], and MUltiple SIgnal Classification (MUSIC)[4], [5].

This paper considers multiple electromagnetic source localization using sensors whose outputs correspond to the complete electric and magnetic fields at the sensor. We assume that the wave is traveling in a nonconductive, homogeneous, and isotropic medium. Additionally we assume a plane wave, which is equivalent to a farfield assumption (or a maximum wavelength that is much smaller than the source-to-sensor distance), a point-source assumption (i.e., the source size is much smaller than the source-to-sensor distance), and a point like sensor (i.e., the dimension of the sensors is small compared with the minimum wavelength). Such a wave can be expressed as

$$s(t) = \alpha_i \exp j(w_i t). \tag{1}$$

Thus, the vector of signals incident upon a linear equispaced array of n sensors is described by

$$y(t) = As + \text{noise} = \sum_{i=1}^{m} \alpha_i \exp j(w_i t) d_i + \eta(t), \tag{2}$$

where m is the number of sources irradiating the array, α_i, and $w_i = 2\pi f_i$ are the amplitude and frequency of the ith source respectively. The n-element vector d_i is the direction vector (called also steering vector) corresponding to each source

$$d_i = (1, e^{-j\tau_i}, e^{-2j\tau_i}, ..., e^{-(n-1)j\tau_i}), \qquad 1 \le i \le n \tag{3}$$

$$\tau_i = \frac{2\pi D}{\lambda_i} \sin\theta_i. \tag{4}$$

The vector-sensor array processing was addressed in [6] where a direct solution was given using the Cramer-Rao bound. Using these vector sensors should outperform the scalar sensors used by the earlier described methods because they allow us to make full use of the EM field inner characteristics and the information therein. Exploiting directly the waveforms enables us to use the phase information and hence, estimate the DOA for coherent sources. Noticeable is that determining the DOA for coherent sources is a very challenging task for spectral methods.

0-8186-8005-9/97 $10.00 © 1997 IEEE

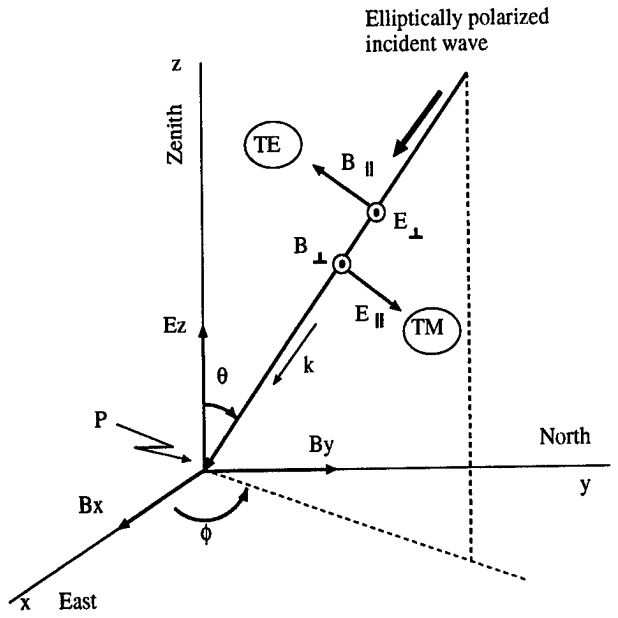


Figure 1. Field observation

2 Measurement Model

In this section we study the measurement model used in this approach and consequently derive the DOA and polarization of the incident source. Figure 1 shows the configuration of the measurement model in which a downgoing wave makes an incident angle θ with the vertical direction and an azimuthal angle Φ with the East. In the figure, $H_{\|}$ and $E_{\|}$ indicate the magnetic and electric field components in the plane of incidence, while $H_{\perp}$ and $E_{\perp}$ those perpendicular to it. The EM field components at the point P whose height is small compared with a wavelength are given by the following equations.

$$\begin{aligned} H_x &= 2A_{\perp}e^{jwt}\{(\cos\Phi - x\cos\theta\sin\Phi) + jy\cos\theta\sin\Phi\} \\ &= A_x\cos(wt+\psi_x) \\ H_y &= 2A_{\perp}e^{jwt}\{-(\sin\Phi + x\cos\theta\cos\Phi) + jy\cos\theta\cos\Phi\} \\ &= A_y\cos(wt+\psi_y) \\ E_z &= -2A_{\perp}e^{jwt}\sin\theta \\ &= A_z\cos(wt). \end{aligned} \tag{5}$$

In the above notations $H_{\perp} = A_{\perp}\cos wt$ and $H_{\|} = A_{\|}\cos(wt+\alpha)$, $(j=\sqrt{-1})$. The wave polarization is defined by

$$\begin{aligned} p &= \frac{H_{\|}}{H_{\perp}} = \frac{A_{\|}}{A_{\perp}}\exp(j\alpha) \\ &= x + jy. \end{aligned} \tag{6}$$

As we can see from the above equations, it is easy to calculate the DOA and polarization of the incident wave if we can determine the waveforms of the three field components H_x, H_y, and E_z. Let denote by M_{iz} and ψ_{iz} the amplitude ratios and phase differences (with E_z as a reference) defined by

$$\begin{aligned} M_{iz} &= \frac{H_i}{E_z} = \frac{A_i}{A_z} \\ \psi_{iz} &= \psi_z - \psi_i \qquad i = x, y. \end{aligned} \tag{7}$$

By simple manipulations on Eqs. (5) and (7) we get the DOA and the polarization [10],

$$\begin{aligned} \tan\Phi &= -\frac{M_{xz}\sin\psi_{xz}}{M_{yz}\sin\psi_{yz}} \\ &= -\frac{A_x\sin\psi_{xz}}{A_y\sin\psi_{yz}} \\ \sin\theta &= (M_{yz}\cos\psi_{yz}\sin\Phi - M_{xz}\cos\psi_{xz}\cos\Phi)^{-1} \\ &= \left\{\frac{(A_xA_z\sin\psi_{xz})^2+(A_yA_z\sin\psi_{yz})^2}{A_x^2A_y^2\sin^2(\psi_{xz}-\psi_{yz})}\right\}^{1/2}, \end{aligned} \tag{8}$$

$$\begin{aligned} x &= \frac{\tan\theta}{\sin\Phi}\left(M_{xz}\cos\psi_{xz} + \frac{\cos\Phi}{\sin\theta}\right) \\ y &= \frac{\tan\theta}{\sin\Phi}M_{xz}\sin\psi_{xz}. \end{aligned} \tag{9}$$

In the following section we show how to estimate $H_i(t)$ and $E_z(t)$.

3 Blind Separation of the EM Field Components

3.1 Blind separation of sources

We consider the output of an array of sensors and suppose that the output signals are the sum of many independent but completely unknown original source signals as expressed by Eq.(2). The blind separation of the original signals consists in their recovery from the observation of a linear mixture observed at the array output. Note that the independence of sources is the only a priori information available.

Let denote by $x(t)$ the solution to the problem; that is, the estimated sources. Since this solution is made upon a test of the statistical independence (all sources are mutually independent), any scaled permutation of these sources will satisfy this test. This leads to an indeterminacy in the solution in both its scale and its arrangement and can be expressed by

$$x(t) = \Gamma P s(t). \tag{10}$$

Γ is a diagonal scaling matrix and P is any permutation matrix. $x(t)$ and $s(t)$ are the estimated and original signals respectively. We will show later that even though it is serious in blind separation theory, this indeterminacy is not harmful in the present application.

The problem of separation was first addressed by J. Herault and C. Jutten who proposed a solution based on neural computing [7, 11] as shown in Fig. 2. The

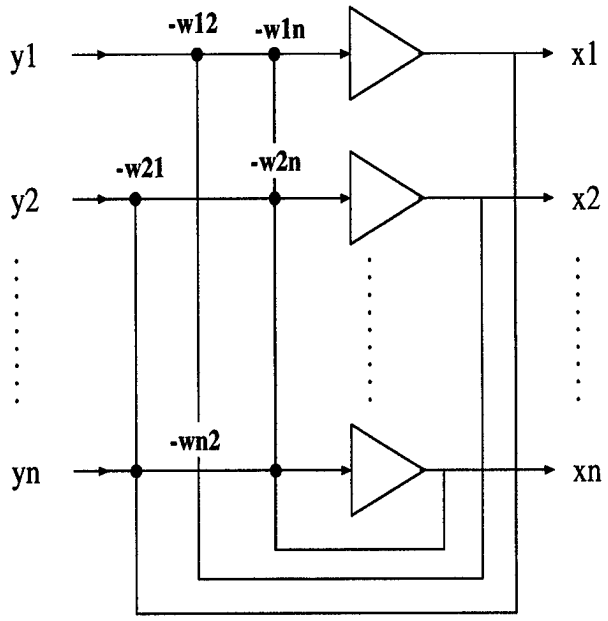


Figure 2. Architecture of the neural solution.

estimated solution x in the figure is a weighted sum of the sensors' output y as given by Eq. (2). It has been shown theoretically that the architecture of such a network allows us to separate the unknown sources. Practically, this requires some additional information in order to bring the network to the optimal solution. In the present case, the only available a priori information is that the original sources s_i are statistically independent. Therefore after the separation, the output signals x_i must be independent as well. For simplicity of our explanation let us assume at the beginning that the average values of the input signals are zero, i.e.

$$E_M(s_i) = 0 \quad \text{and} \quad E_M(x_i) = 0, \quad i = 1.....m, \tag{11}$$

where $E_M(.)$ is the statistical mean. Two signals are statistically independent are at least decorrelated (zero covariance), i.e. the expected values of their products are null; that is,

$$E_M(s_i s_j) = 0 \quad \text{and} \quad E_M(x_i x_j) = 0. \tag{12}$$

However, the zero covariance of Eq.(12) is not a sufficient condition for the statistical independence. A much stronger requirement for the statistical independence is,

$$E_M(s_i^k s_j^p) = 0 \quad \text{and} \quad E_M(x_i^k x_j^p) = 0, \tag{13}$$

for any odd integer value of k and p. On the basis of this requirement Herault and Jutten proposed a learning rule of the following form

$$\frac{dw_{ij}}{dt} = \mu f(x_i(t)) g(x_j(t)), \tag{14}$$

where $\mu > 0$ is a positive learning rate and $f()$ and $g()$ are two odd nonlinear functions. Thus, the weights are updated iteratively by the HJ algorithm starting from randomly chosen small values.

3.2 Separation of the EM field components

We apply the previous algorithm to separate each of the field components. If we consider the case of three components and the case of two sources, we have a set of three pairs of signals: (E_z^1, E_z^2), (H_x^1, H_x^1), and (H_y^1, H_y^2). $(.)^1$ are the components of the first source and $(.)^2$ are those of the second source. Figure 3 shows the separation diagram for three components field and two sources. In order to separate two sources ($m = 2$),

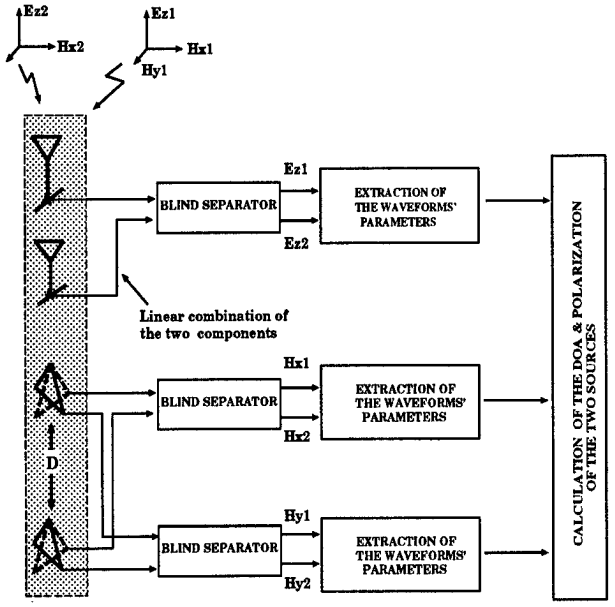


Figure 3. Diagram of DOA finder based on blind separation.

we need to set at least two antennae ($n = 2$) at a distance D of each other and measure the three components of the EM field. The output of the sensors is given by Eq.(2) for each component. If we denote by s any of these components (H_x,H_y,E_z), the output y of its corresponding sensors are given by

$$\begin{aligned} y_1 &= a_{11}s_1 + a_{12}s_2 e^{-jw\tau_2} + \eta_1 \\ y_2 &= a_{21}s_1 e^{-jw\tau_1} + a_{22}s_2 + \eta_2, \end{aligned} \tag{15}$$

where τ_i is given by Eq.(4) and η is an additive noise. By analogy to Eq.(2), the mixing matrix is of the form

$$A = \begin{pmatrix} a_{11} & a_{12}e^{-jw\tau_2} \\ a_{21}e^{-jw\tau_1} & a_{22} \end{pmatrix}. \qquad (16)$$

To solve the system of Eq.(15) we need to have the coefficients a_{ij} at least to be constant over the estimation time. However, if the incident angle is constant over the separation time and for a given frequency w, the term $e^{-jw\tau}$ is a constant complex. In other words, the incident angle should be constant over the separation time. If this requirement is met and in the presence of an additive noise we get the system of Eq.(2) and we can adopt the HJ algorithm that requires a constant matrix A.

4 simulation

In this section we show some results of computer simulation. We consider the estimation of the DOA and polarization for two sources, and the sources are assumed to be elliptically polarized and to have different incident and azimuthal angles, θ and ϕ. Figure 4

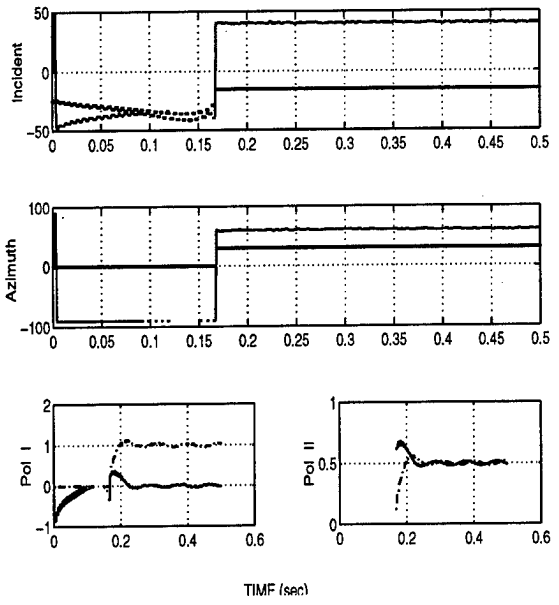


Figure 4. Blind estimation of DOA and polarization for two coherent sources.

shows the estimated results of DOA and polarization for two sources using the proposed algorithm. The upper and middle boxes show the incident and azimuth angles, θ and Φ respectively, while the lower boxes show the polarization for the two sources. In

		Without noise		With noise	
	Assum. (Deg.)	Estim. (Deg.)	σ (10^{-2})	Estim. (Mean)	σ (Deg.)
θ_1	40	39.90	1.0	40.23	1.66
θ_2	-15	-15.0	1.0	-14.7	1.96
Φ_1	60	59.90	1.0	59.38	2.62
Φ_2	30	29.90	1.0	34.0	12.75
p_1	0,1	0,.99	1.0	.014,1.0	.08,.07
p_2	.5, .5	.5,.5	1.2	.49,.51	.14,.09

Table 1. Summary of estimation results.

the two upper boxes, the solid and dashed lines stand for source **I** and **II**, respectively. The source **I** is assumed to be circularly right-handed polarized ($p_1 = j$) and its assumed DOA is ($\theta_1 = 40°, \Phi_1 = 60°$), while the source **II** is elliptically polarized with $p_2 = \frac{1}{2}(1+j)$ and its DOA is ($\theta_2 = -15°, \Phi_2 = 30°$). We first consider the estimation case for coherent sources, which is a rather difficult and challenging task especially for spectral methods where the the frequency of the source is the only available information. This approach overcomes this difficulty by benefiting from the phase information calculated directly from the reconstructed waveforms. The DOA's in a noiseless conditions are estimated to be ($\hat{\theta}_1 = 39.9°, \hat{\Phi}_1 = 59.9°$) and ($\hat{\theta}_2 = -15.0°, \hat{\Phi}_2 = 29.9°$) for the sources **I** and **II** respectively with an error of almost 1% as it is shown in Table 4. The mean polarization for each source is $p_1 = (0.99 \pm 0.024)j$ and $p_2 = 0.5 + 0.5j \pm 0.012$. The two cells network is seen to converge after $158ms$ as shown in Fig. 4. Table 4 summarizes the estimation results for two coherent sources. We also considered the case where the DOA is slightly varying with time; that is, the incident and/or azimuthal angle is not fixed vis-à-vis to the sensor. This happens in some applications when the transmitter is moving. The upper box in Fig. 5 indicates the estimated incident angles θ_1 and θ_2 where the original or assumed angle θ_1 is varying sinusoidally with time, and is plotted in dashed lines. The comparison of the curves in the top of Fig. 5 indicates the estimation is delayed with the original θ_1 by 40 ms, which is partly due to the presence of filters and other processing units that induce small delays. We also investigated the case of varying polarization and the case where the two source have vey close polarizations. We have mentioned in our theoretical study that the solution given by the HJ algorithm is the subject to an indeterminacy on the scale and the permutation order. For this application, it is easy to distinguish the two sources by using the inner statistical properties of the

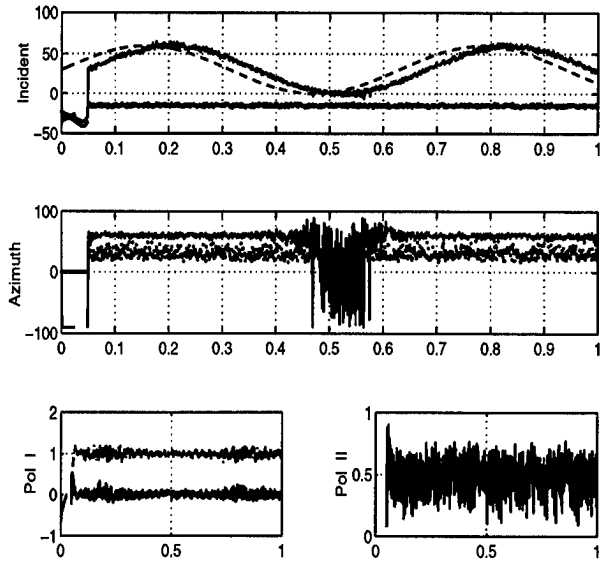


Figure 5. The estimation of incident and azimuthal angles and the polarization of the two sources when the incident angle θ_1 is varying sinusoidally with time.

field components. As a matter of fact, it suffices to run a correlation test on two components to separate the components of one source. While, for the scale indeterminacy, the simulation results have shown that this scale factor is almost equal for all components of one field, and because we use only the amplitude ratios, this scale factor does not have any effect on the estimation results. The other plausible result is concerned with the polarization. So far, most of the approaches concerning the estimation of the DOA of polarized electromagnetic waves have been imposing some specific assumptions on the source polarization in order to obtain an explicit theoretical form for the DOA [12, 13]. The present approach is found to be advantageous over the previous studies, since we can get all the information from the waveform, and we did not need any additional assumption on the source polarization.

References

[1] S.P. Applebaum, *Adaptive Arrays.* IEEE Trans. Antennas and Propagation, 24(9): 585-598, Sept. 1976

[2] Don H. Johnson and Dan E. Dudgeon, *Array Signal Processing- Concepts and Techniques.* Prentice-Hall, Englewood Cliffs, NJ, 1993.

[3] S.M. Kay and S.L. Marple, *Spectrum analysis - A modern perspective.* Proc. IEEE, vol. 69, Nov. 1981, pp. 1380-1419.

[4] Don H. Johnson and S.R. DeGraaf, *Improving resolution of bearing in passive sonar arrays be eigenvalue analysis,* IEEE Trans. Acous. Speech Signal Processing vol. ASSP-30, Aug. 1982.

[5] H. Krim and M. Viberg, *Two decades of array signal processing research,* IEEE Signal Proc. Magazine July 1996.

[6] A. Nehorai and E. Paldi, *Vector-sensor array processing for electromagnetic source localization* , IEEE Trans. Signal Processing, Vol. 42, No. 2, Feb. 1994.

[7] C. Jutten and J. Herault and B. Ans, *Détection de grandeurs primitives dans un message composite par une architecture de calcul neurométrique en apprentissage non supervisé,* Proc. Xème colloque GRETSI, Nice, France, 20-24 May 1985, pp. 1017-1022.

[8] Storey, L.R.O. and F. Lefeuvre, *The analysis of 6-component measurements of a random electromagnetic wave field in a magnetoplasma -I. The direct problem,* Geophys. J. R. Astr. Soc., Vol. 56, pp. 255-269, 1979.

[9] Storey, L.R.O. and F. Lefeuvre, *The analysis of 6-component measurements of a random electromagnetic wave field in a magnetoplasma -II. The integration kernels,* Geophys. J. R. Astr. Soc., Vol. 62, pp. 173-194, 1980.

[10] M. Hayakawa, *Whistlers,* Handbook of Atmospheric Electrodynamics, Vol II, Edit by Hans Volland, CRC Press, 1995.

[11] C. Jutten and J. Herault, *Blind separation of sources, PartI: an adaptive algorithm based on neurometric architecture,* Signal Processing, Vol. 24, pp. 1-10, 1991

[12] M. Hirari and M. Hayakawa, *A neural network for the DOA of VLF/ELF radio waves,* IEICE Trans. Commun, Vol. E79-B, No. 10 October 1996.

[13] M. Hirari, *Statistical methods and artificial neural networks in the direction finding of VLF/ELF radio waves,* Ph.D thesis, The Univ. of Electro-Communications, Tokyo, 1996.

Dimensionality Reduction in Higher-Order-Only ICA

Lieven De Lathauwer Bart De Moor Joos Vandewalle
K.U.Leuven - E.E. Dept.- ESAT - SISTA/COSIC
Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
{ Lieven.DeLathauwer Bart.DeMoor Joos.Vandewalle } @esat.kuleuven.ac.be

Abstract

Most algebraic methods for Independent Component Analysis (ICA) consist of a second-order and a higher-order stage. The former can be considered as a classical Principal Component Analysis (PCA), with a three-fold goal: (a) reduction of the parameter set of unknowns to the manifold of orthogonal matrices, (b) standardization of the unknown source signals to mutually uncorrelated unit-variance signals, and (c) determination of the number of sources. In the higher-order stage the remaining unknown orthogonal factor is determined by imposing statistical independence on the source estimates. Like all correlation-based techniques, this set-up has the disadvantage that it is affected by additive Gaussian noise. However it is possible to solve the problem, in a way that is conceptually blind to additive Gaussian noise, by resorting only to higher-order cumulants. The purpose of this paper is to explain how the dimensionality of the ICA-model can algebraically be reduced to the true number of sources in higher-order-only schemes.

1. Introduction

In this paper, the basic ICA-model will be denoted as follows:

$$Y = \mathbf{M}X + N \tag{1}$$

in which the observed vector Y, the source vector X and the noise vector N are zero-mean random vectors with values in $\mathbb{R}$ or $\mathbb{C}$. The components of X are assumed to be mutually statistically independent, as well as statistically independent from the noise components; the mixing matrix $\mathbf{M}$ has linearly independent columns. As is well-known, the goal of ICA consists of the estimation of the mixing matrix and the corresponding realizations of X, given only realizations of Y.

ICA-algorithms often have the form of a two-stage procedure. First, in a prewhitening step, the second-order statistics of the observation vector Y are used to "standardize" the problem: Y is transformed in a vector $\tilde{Y}$ with unit covariance. In the more-sensors-than-sources case this also implies a projection of Y onto the dominant subspace of the data. The standardization step allows to restrict the unknown mixing matrix to the manifold of orthogonal matrices. In the second step, the orthogonal factor is then obtained from a higher-order cumulant tensor of $\tilde{Y}$. Examples of algebraic prewhitening-based ICA-methods are [4, 3, 7, 10].

On the other hand, it is also possible to identify the mixing matrix by using *only* the higher-order cumulant tensor. For algebraic techniques, we refer to [2, 1, 5, 11]. Such higher-order-only methods have the interesting feature that they allow to boost signal-to-noise ratios when the noise is Gaussian. Although there is no prewhitening stage, one may still want to reduce the dimensionality of the higher-order cumulant in the more-sensors-than-sources case, as this may significantly reduce the computational efforts in the actual algorithm. In this paper, we will investigate how the dimensionality reduction can be achieved, from a multilinear algebraic point of view.

In Section 2, we will introduce some necessary background material on multilinear algebra. In Section 3 the actual core of our contribution is addressed: where the problem of dimensionality reduction in PCA can essentially be formulated as the best rank-R approximation of a given matrix, the higher-order-only problem leads to a clear multilinear equivalent, denoted as the best rank-$(R, R, \ldots, R)$ approximation (in least-squares sense) of the given higher-order cumulant tensor. In Section 4 we discuss if and how the approximation problem can be solved by a tensorial generalization of truncating the Eigenvalue Decomposition (EVD) or Singular Value Decomposition (SVD). Section 5 deals with the local minimisation of the quadratic error criterion by an Alternating Least Squares (ALS) algorithm that can be interpreted as a higher-order generalization of the technique of orthogonal iterations [12]. Quite some research on related topics has been done in the field of Psychometrics; the main ideas behind Sections 4 and 5 can

0-8186-8005-9/97 $10.00 © 1997 IEEE

also be found in [13]. In Section 6 we give an explicit expression for the gradient of the objective function, for use in gradient-based optimization routines. Section 7 summarizes the conclusions.

Section 5 forms a generalisation of a previous paper [9], on the best rank-1 approximation of a given tensor.

For simplicity of notation, our discussion is restricted to the case of real-valued data. The results are readily generalized for complex tensors.

2. Basic definitions and notations

Multilinear algebra is the algebra of higher-order tensors. For the purpose of this paper, higher-order tensors can intuitively be imagined as multi-indexed arrays of numerical values. The tensors of interest in HOS (higher-order moments, cumulants, ...) are super-symmetric: in the real-valued case, this means that they remain unchanged under arbitrary index permutations. We subsequently touch on (a) some rank-related definitions, (b) the multiplication of a tensor by a matrix, (c) a standardized representation in matrix format, and (d) some geometry-related concepts, such as scalar product, orthogonality and norm.

Definition 1 (Rank-1 tensor) *An Nth-order tensor $\mathcal{A}$ has* rank 1 *when it equals the outer product of N vectors U, V, ... Z:*

$$\mathcal{A} = U \circ V \circ \ldots \circ Z \tag{2}$$

i.e. $\mathcal{A}_{i_1 i_2 \ldots i_N} = U_{i_1} V_{i_2} \ldots Z_{i_N}$ for all index values.

Definition 2 (Rank) *The rank of an arbitrary Nth-order tensor $\mathcal{A}$, denoted by $r = \mathtt{rank}(\mathcal{A})$, is the minimal number of rank-1 tensors that yield $\mathcal{A}$ in a linear combination.*

Definition 3 (n-rank) *The n-rank of $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_N}$, denoted by r_n, is defined as the dimension of the vector space generated by the $I_1 I_2 \ldots I_{n-1} I_{n+1} \ldots I_N$ vectors, obtained by varying the index i_n of $\mathcal{A}_{i_1 i_2 \ldots i_N}$ while keeping the other indices fixed.*

The definition of n-rank generalizes the definitions of column and row rank of matrices. Def. 2 gives an other way to extend the concept of rank to higher-order tensors. In contrast to the matrix case, the different n-ranks of a higher-order tensor are not necessarily the same; and when they are equal (e.g. as a result of super-symmetry), they can still be different from the rank of the tensor. An Nth order tensor with n-ranks $r_1, r_2, \ldots, r_N$, will be denoted as a rank-$(r_1, r_2, \ldots, r_N)$ tensor. An interesting discussion of rank-related issues for super-symmetric tensors, with bibliographic pointers to the early literature, can be found in [5].

The product of two matrices can be generalized to the product of a tensor and a matrix as follows:

Definition 4 *The n-mode product of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{U}$, is an $(I_1 \times I_2 \times \ldots J_n \ldots \times I_N)$-tensor given by*

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \ldots j_n \ldots i_N} = \sum_{i_n} a_{i_1 i_2 \ldots i_n \ldots i_N} u_{j_n i_n}$$

for all index values.

The n-mode product allows us to express the effect of basis transformations on a given tensor. As an example, for the matrices $\mathbf{F} \in \mathbb{R}^{I_1 \times I_2}$, $\mathbf{U} \in \mathbb{R}^{J_1 \times I_1}$, $\mathbf{V} \in \mathbb{R}^{J_2 \times I_2}$ the matrix product $\mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^T$ can be written as the "symmetric" expression $\mathbf{F} \times_1 \mathbf{U} \times_2 \mathbf{V}$.

It is often more familiar to formulate tensor expressions in a matrix language. We define the following standard matrix representation of a higher-order tensor:

Definition 5 *The* mode-n matrix unfolding *of $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_N}$, denoted by $\mathbf{A}_{(n)}$, is the $(I_n \times I_1 I_2 \ldots I_{n-1} I_{n+1} \ldots I_N)$-matrix that contains the element $\mathcal{A}_{i_1 i_2 \ldots i_N}$ on the position given by row number i_n and column number $(i_1 - 1) I_2 I_3 \ldots I_{n-1} I_{n+1} \ldots I_N + (i_2 - 1) I_3 \ldots I_{n-1} I_{n+1} \ldots I_N + \ldots + i_N$.*

Remark that the n-mode product $\mathcal{B} = \mathcal{A} \times_n \mathbf{U}$ can be translated as $\mathbf{B}_{(n)} = \mathbf{U} \cdot \mathbf{A}_{(n)}$, and that the rank of $\mathbf{A}_{(n)}$ corresponds to the n-rank of $\mathcal{A}$, even in a numerical sense. For super-symmetric tensors, the various matrix unfoldings are equal, and will shortly be denoted as $\mathbf{A}$.

The definitions of scalar product, orthogonality and Frobenius-norm can be generalized in a trivial way:

Definition 6 *The scalar product $\langle \mathcal{A}, \mathcal{B} \rangle$ of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_N}$ is defined as*

$$\langle \mathcal{A}, \mathcal{B} \rangle \stackrel{\text{def}}{=} \sum_{i_1} \sum_{i_2} \cdots \sum_{i_N} \mathcal{A}_{i_1 i_2 \ldots i_N} \mathcal{B}_{i_1 i_2 \ldots i_N} \tag{3}$$

Definition 7 *Tensors of which the scalar product equals 0, are mutually orthogonal.*

Definition 8 *The Frobenius-norm of a tensor $\mathcal{A}$ is given by*

$$\|\mathcal{A}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle} \tag{4}$$

The Frobenius-norm can be interpreted as the "size" of the tensor. The square of this norm can be seen as the "energy" in the tensor.

3. Dimensionality reduction in ICA with and without prewhitening

Let us first remind some more formal aspects of the procedure of prewhitening. The ICA-model implies that the

covariance $\mathbf{C}_2^Y$ of Y is given by

$$\mathbf{C}_2^Y = \mathbf{C}_2^X \times_1 \mathbf{M} \times_2 \mathbf{M} + \mathbf{C}_2^N = \mathbf{M}\mathbf{C}_2^X\mathbf{M}^T + \mathbf{C}_2^N \quad (5)$$

in which $\mathbf{C}_2^X$ and $\mathbf{C}_2^N$ symbolize the covariance of X resp. N. $\mathbf{C}_2^X$ is diagonal, since the source signals are uncorrelated. We can additionally assume that the source signals have unit variance, which at most involves an irrelevant rescaling of the columns of $\mathbf{M}$. As a consequence, a first estimate of $\mathbf{M}$, up to an orthogonal factor $\mathbf{V}$, can be obtained in the form of a square root $\mathbf{F}$ of $\mathbf{C}_2^Y$:

$$\mathbf{C}_2^Y = \mathbf{F}\mathbf{F}^T = (\mathbf{F}\mathbf{V}^T)(\mathbf{F}\mathbf{V}^T)^T \simeq \mathbf{M}\mathbf{M}^T \quad (6)$$

Substitution of the Singular Value Decomposition (SVD) $\mathbf{M} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T$ of the mixing matrix, leads to a classical PCA

$$\mathbf{C}_2^Y = \mathbf{U}\mathbf{S}^2\mathbf{U}^T + \mathbf{C}_2^N = (\mathbf{U}\mathbf{S})(\mathbf{U}\mathbf{S})^T + \mathbf{C}_2^N \quad (7)$$

in which the left singular matrix $\mathbf{U}$ and the matrix of singular values $\mathbf{S}$ are estimated with an EVD of the observed covariance (or, numerically preferable, the direct SVD of the dataset), while the right singular matrix $\mathbf{V}$ remains unknown. Let us denote the dimension of Y by I and the number of sources by R. If $I > R$, then the dominant R-dimensional eigenspace of $\mathbf{C}_2^Y$ is retained. If the noise is known to be spatially white, then its variance σ_N^2 can be estimated as the mean of the smallest $I - R$ eigenvalues of $\mathbf{C}_2^Y$.

On the other hand, the fourth-order cumulant $\mathcal{C}_4^Y$ of Y can be expressed in a similar way as the covariance in Eq. (5):

$$\mathcal{C}_4^Y = \mathcal{C}_4^X \times_1 \mathbf{M} \times_2 \mathbf{M} \times_3 \mathbf{M} \times_4 \mathbf{M} + \mathcal{C}_4^N \quad (8)$$

In this equation, $\mathcal{C}_4^X$ and $\mathcal{C}_4^N$ symbolize the fourth-order cumulant tensors of X resp. N; the latter vanishes if the noise is Gaussian. Like $\mathbf{C}_2^X$, $\mathcal{C}_4^X$ is diagonal, as the source signals are statistically independent. Due to the diagonality property, Eq. (8) can be rewritten as

$$\mathcal{C}_4^Y \simeq \sum_{r=1}^{R} \kappa_r^X M_r \circ M_r \circ M_r \circ M_r \quad (9)$$

in which κ_r^X represents the marginal cumulant of the rth source (we assume that the sources are kurtic), and in which M_r denotes the rth column of $\mathbf{M}$. Eq. (9) is a decomposition of $\mathcal{C}_4^Y$ in a minimal number of rank-1 terms, as the columns of $\mathbf{M}$ are assumed to be linearly independent; it can even be proved that the decomposition is essentially unique, up to some trivial indeterminacies [14]. As a consequence, the aim of higher-order-only ICA can be formulated as the computation of a rank-revealing decomposition of $\mathcal{C}_4^Y$, taking into account that the sample cumulant equivalent of Eq. (9) may be perturbated by non-Gaussian noise components, finite datalength effects, model misfit, etc.

In addition, according to the definition of n-rank, every tensor satisfying the assumptions of Eq. (9) is a rank-(R, R, R, R) tensor. To deal with the situation in which $I > R$, we propose to perform an orthogonal projection of the sample cumulant $\hat{\mathcal{C}}_4^Y$ on the manifold of super-symmetric rank-(R, R, R, R) tensors, before addressing the harder problem of the further projection on the submanifold of rank-R tensors and the actual computation of decomposition (9). Notice that in the second-order case both projections coincide, as n-rank and rank are necessarily the same, which corresponds to the single best rank-R approximation problem, associated with Eq. (7).

Formally, the problem we want to solve, can be formulated as follows:
Given a sample cumulant $\hat{\mathcal{A}} \in \mathbb{R}^{I\times I\times I\times I}$, find a super-symmetric tensor $\tilde{\mathcal{A}} \in \mathbb{R}^{I\times I\times I\times I}$, with $r_1 = r_2 = r_3 = r_4 = R$, that minimizes the least-squares criterion

$$f(\tilde{\mathcal{A}}) = \|\hat{\mathcal{A}} - \tilde{\mathcal{A}}\|^2 \quad (10)$$

The conditions imply that $\tilde{\mathcal{A}}$ can be decomposed as:

$$\tilde{\mathcal{A}} = \mathcal{B} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{U} \times_4 \mathbf{U} \quad (11)$$

in which $\mathbf{U} \in \mathbb{R}^{I\times R}$ has orthonormal columns and $\mathcal{B} \in \mathbb{R}^{R\times R\times R\times R}$ is super-symmetric. After computation of the best rank-(R, R, R, R) approximation, the actual ICA-algorithm can be applied to $\mathcal{B}$ instead of $\hat{\mathcal{A}}$; $\mathcal{B}$ can be interpreted as the cumulant tensor of the R-dimensional stochastic vector $Z = \mathbf{U}^T Y$.

4. Best rank-(R, R, R, R) approximation and Higher-Order Eigenvalue Decomposition

In this section we discuss the tensorial equivalent of the computation of the best rank-R approximation of a symmetric matrix by truncation of its symmetric EVD. First, the decomposition itself can be generalized as follows:

Theorem 1 *Every super-symmetric Nth order tensor $\mathcal{A} \in \mathbb{R}^{I\times I\times\ldots\times I}$ can be written as*

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \ldots \times_N \mathbf{U} \quad (12)$$

in which:

- $\mathbf{U} = [U_1 U_2 \ldots U_I]$ *is an orthogonal $(I \times I)$-matrix*
- *the core tensor $\mathcal{S}$ is a super-symmetric $(I \times I \times \ldots I)$-tensor of which the subtensors $\mathcal{S}_{i_n=\alpha}$, obtained by fixing one arbitrary index to α, have the properties of: - all-orthogonality: two subtensors $\mathcal{S}_{i_n=\alpha}$ and $\mathcal{S}_{i_n=\beta}$*

are orthogonal for all possible values of n, α and β subject to $\alpha \neq \beta$:

$$\langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \rangle = 0 \quad when \quad \alpha \neq \beta \tag{13}$$

- ordering:

$$\|\mathcal{S}_{i_n=1}\| \geq \|\mathcal{S}_{i_n=2}\| \geq \ldots \geq \|\mathcal{S}_{i_n=r_n}\| > 0 \tag{14}$$

and

$$\|\mathcal{S}_{i_n=r_n+1}\| = \ldots = \|\mathcal{S}_{i_n=I}\| = 0 \tag{15}$$

The Frobenius-norms $\|\mathcal{S}_{i_n=i}\|$, symbolized by s_i, are the higher-order eigenvalues of $\mathcal{A}$ and the vector U_i is the ith higher-order eigenvector.

This Higher-Order Eigenvalue Decomposition (HOEVD) is the super-symmetric case of the Higher-Order Singular Value Decomposition (HOSVD), discussed in [8]. Clearly, the HOEVD is a formal generalization of the EVD of symmetric matrices. Moreover, it can be proved that the HOEVD of a second-order symmetric tensor essentially boils down to its matrix EVD.

The matrix $\mathbf{U}$ can be computed as the left singular matrix of the matrix unfolding $\mathbf{A}$. The core tensor $\mathcal{S}$ can then be found from the following form of Eq. (12):

$$\mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^T \times_2 \mathbf{U}^T \ldots \times_N \mathbf{U}^T \tag{16}$$

The higher-order eigenvalues correspond to the singular values of $\mathbf{A}$; as such, they give numerical information on the n-rank of $\mathcal{A}$. If Eq. (9) is exactly satisfied, then the number of non-zero higher-order eigenvalues corresponds to the number of kurtic sources. In practice, the number of sources will be estimated by determining the hypothesized gap between "signal" and "noise" higher-order eigenvalues.

The properties of the higher-order eigenvectors and eigenvalues, together with the ordering convention, suggest that setting $s_{R+1} = s_{R+2} = s_N = 0$, will result in a good rank-(R, R, R, R) approximation of $\mathcal{A}$. Remarkably enough, it turns out that this approximation is in general not the globally optimal one. A second important difference with matrices, is that the criterion function $f(\hat{\mathcal{A}})$ can exhibit spurious local optima. On the other hand, we have observed that the HOEVD-approximation generally belongs to the valley of f that contains the global optimum. There is not an absolute guarantee: we have been able to generate ill-conditioned cases in which gradient-descent, starting from the HOEVD-truncate, eventually leads to a local optimum with a close to optimal fit, but nevertheless different from the global optimum [6]. In practice however, no problems of this kind have been observed.

5. Higher-order orthogonal iterations

In this section we briefly discuss an ALS algorithm for the local optimisation of f, that can be interpreted as a higher-order generalization of the method of orthogonal iterations [12].

First, observe that the minimisation of f is equivalent to the maximisation of the following function $g(\mathbf{U})$ over the matrices with mutually orthonormal columns:

$$g(\mathbf{U}) = \|\mathcal{A} \times_1 \mathbf{U}^T \times_2 \mathbf{U}^T \times_3 \mathbf{U}^T \times_4 \mathbf{U}^T\|^2 \tag{17}$$

Proving this relationship falls outside the scope of this article, but as a motivation we can remark that, similarly, the computation of the best rank-R approximation of a symmetric matrix $\mathbf{H}$ is equivalent to the determination of a columnwise orthonormal matrix $\mathbf{U}$ that maximizes the Frobenius-norm of $\mathbf{U} \cdot \mathbf{H} \cdot \mathbf{U}^T$.

Next, let us embed the maximisation of g in the maximisation, over the matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}, \mathbf{U}^{(4)}$ with orthonormal columns, of the function $\tilde{g}$, defined as:

$$\tilde{g}(\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(4)}) = \|\mathcal{A} \times_1 \mathbf{U}^{(1)^T} \ldots \times_4 \mathbf{U}^{(4)^T}\|^2 \tag{18}$$

Because of the super-symmetry of $\mathcal{A}$, the global optimum will satisfy $\mathbf{U}^{(1)} = \ldots = \mathbf{U}^{(4)}$.

Now imagine that e.g. the matrices $\mathbf{U}^{(2)}$, $\mathbf{U}^{(3)}$ and $\mathbf{U}^{(4)}$ are fixed and that $\tilde{g}$ in Eq. (18) is merely a quadratic expression in the components of the unknown matrix $\mathbf{U}^{(1)}$. In matrix notation, we have:

$$\tilde{g} = \|\mathbf{U}^{(1)^T} \cdot [\mathbf{A}_{(1)} \cdot (\mathbf{U}^{(2)} \otimes \mathbf{U}^{(3)} \otimes \mathbf{U}^{(4)})]\|^2 \tag{19}$$

in which $\otimes$ denotes the Kronecker product. Hence the columns of $\mathbf{U}^{(1)}$ can be found as an orthonormal basis for the dominant subspace of the column space of the matrix between square brackets. Repeating this procedure for different mode numbers leads to an ALS algorithm for the (local) maximization of $f(\tilde{\mathcal{A}})$: in each step the estimate of one of the matrices $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(4)}$ is optimized, while the other matrix estimates are kept constant.

Clearly this technique is a higher-order extension of the orthogonal iteration method for matrices. A major difference is that each iteration step does not involve a linear but a multilinear transformation. A second important difference is that an iteration step also implies the estimation of a dominant subspace, instead of a mere orthonormalisation. It is inherent to the ALS-mechanism that higher-order orthogonal iterations yield unsymmetric intermediate results.

6. Gradient-based optimisation

As an alternative for ALS-iterations, one could consider the use of more general-purpose optimisation routines, in

which the unknown matrix $\mathbf{U}$ is situated on the Stiefel manifold of column-wise orthonormal $(I \times R)$-matrices. The symmetry can be retained during iteration by simply treating all the modes in the same way. For the optimisation procedure it is possible to exploit knowledge of the gradient. More specifically, the gradient of g over the Stiefel manifold takes the following form:

$$\nabla_{\mathbf{U}} g = 8\,(\mathbf{I} - \mathbf{U}\mathbf{U}^T) \cdot \tilde{\mathbf{U}} \cdot \tilde{\mathbf{U}}^T \cdot \mathbf{U} \qquad (20)$$

in which

$$\tilde{\mathbf{U}} \stackrel{\text{def}}{=} \mathbf{A} \cdot (\mathbf{U}^{(2)} \otimes \mathbf{U}^{(3)} \otimes \mathbf{U}^{(4)}) \qquad (21)$$

Expression (20) consists of a sum of 4 equal terms, as Eq. (17) involves the matrix $\mathbf{U}$ 4 times, in a super-symmetric way. For each of the terms, the symmetric version of Eq. (19) can be interpreted as a classical subspace tracking problem, with fixed matrix $\tilde{\mathbf{U}}$.

7. Conclusions

In an ICA set-up with more observation channels than source components, the number of sources can conceptually be estimated without resorting to second-order statistics. The corresponding reduction of the dimensionality of the problem takes the form of a multilinear generalization of the best rank-R approximation of matrices. A first approximative solution can be obtained by truncation of a higher-order EVD/SVD-type decomposition, in much the same way as for matrices. The determination of the optimal result generally requires some additional fine-tuning. With this respect, we have paid some attention to an ALS-type descent technique and to gradient-based optimisation over the Stiefel manifold.

Acknowledgements

This work has partly been supported by the Flemish Government (Concerted Research Action GOA-MIPS, the F.W.O. (Fund for Scientific Research - Flanders) project G.0292.95, the F.W.O. project G.0256.97, the F.W.O. Research Communities ICCoS and Advanced Numerical Methods for Mathematical Modelling), the Belgian State (Interuniversity Poles of Attraction IUAP P4-02 and IUAP P4-24) and the European Commission (Human Capital and Mobility Network SIMONET and SCIENCE-ERNSI SC1-CT92-0779). Bart De Moor is a senior Research Associate with the F.W.O. Lieven De Lathauwer is a Research Assistant with the I.W.T. (Flemish Institute for Science and Technology in Industry). The scientific responsibility is assumed by the authors.

References

[1] J.-F. Cardoso. Super-symmetric decomposition of the fourth-order cumulant tensor; blind identification of more sources than sensors. In *Proc. ICASSP-91*, pages 3109–3112, 1991.

[2] J.-F. Cardoso. Iterative techniques for blind source separation using only fourth-order cumulants. In *Proc. EUSIPCO-92*, pages 739–742, Brussels, Belgium, August 1992.

[3] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *IEE Proceedings-F*, 140(6):362–370, December 1994.

[4] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, April 1994.

[5] P. Comon and B. Mourrain. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53(2-3):93–108, September 1996.

[6] L. De Lathauwer. *Signal Processing by Multilinear Algebra.* PhD thesis, K.U.Leuven, E.E. Dept., October 1997.

[7] L. De Lathauwer, B. De Moor, and J. Vandewalle. Blind source separation by higher-order singular value decomposition. In *Proc. EUSIPCO-94*, pages 175–178, Edinburgh, Scotland, U.K., September 1994.

[8] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. ESAT/SISTA report 94-31, K.U.Leuven, December 1994. Submitted to: *SIAM J. Mat. Anal. Appl.*

[9] L. De Lathauwer, B. De Moor, and J. Vandewalle. Higher-order power method - application in independent component analysis. In *Proc. NOLTA'95*, pages 91–96, Vegas, USA, December 1995.

[10] L. De Lathauwer, B. De Moor, and J. Vandewalle. Blind source separation by simultaneous third-order tensor diagonalization. In *Proc. EUSIPCO-96*, pages 2089–2092, Trieste, Italy, September 1996.

[11] L. De Lathauwer, B. De Moor, and J. Vandewalle. Independent component analysis based on higher-order statistics only. In *Proc. SSAP-96*, pages 356–359, Corfu, Greece, June 1996.

[12] G. Golub and C. Van Loan. *Matrix computations.* Johns Hopkins University Press, second edition, 1991.

[13] P. Kroonenberg. *Three-Mode Principal Component Analysis.* DSWO Press, Leiden, 1983.

[14] J. Kruskal. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.*, (18):95–138, 1977.

Blind Source Separation with Noisy Sources

C. Servière

CEPHAG-ENSIEG BP 46 38402 Saint-Martin d'Hères cedex FRANCE

Abstract

A new method of source separation with noisy observations is proposed in the case of two sensors. Each observation contains a mixture of two signals with noise. The objective is to estimate the frequency spectra of the linear filters that combine the two signals in the data stream. The main characteristic of the method is to take into account additive noises. No hypotheses on their probability densities are made. We derive for that an original objective function, based on nonlinear functions of the observations. Specific properties of these functions, chosen as exponential functions, and the hypothesis of independent sources lead to a direct solution for the estimation of the filters. An analytic solution may be computed from it, using only the data. The convergence speed of the method and its robustness against non gaussian noise are illustrated in the paper with simulation results.

1 Introduction

In this paper, source separation based on a two-sensor scenario is considered. A general model for this scenario may be described as follows. Consider two random signals s(k) and r(k), called 'sources' hereafter. Suppose they are propagating through a deterministic linear medium such that we receive two linear combinations x(k) and y(k) on two sensors. Assuming the sources are statistically independent, the problem consists of recovering them from observed signals only. This problem is generally called source separation or blind identification.

It is also supposed that sources are stationary, non Gaussian and with zero-mean. A general model for the observations is given in the frequency-domain by :

(1) $Xi(n) = Si(n) + f(n)\, Ri(n) + Vi(n)$

(2) $Yi(n) = g(n)\, Si(n) + Ri(n) + Wi(n)$

where Si(n), Ri(n), Xi(n), Yi(n), Vi(n) and Wi(n) are the N-point discrete Fourier transforms (DFT) of the i-th data blocks of s(k), r(k), x(k), y(k), v(k) and w(k) at frequency bin n. Vi(n) and Wi(n) are assumed to be independent additive noises, and no assumptions are made regarding the statistical distribution. On the contrary, the source separation methods assume that the complex sources Si(n) and Ri(n) are not strictly complex normal. The model (1) (2) represents either instantaneous mixtures of narrow-band signals, or convolutive mixtures of wide-band signals. In the second case, f(n) and g(n) are the complex weights of two linear filters F and G, at frequency bin n. The goal of source separation is to estimate Si(n) and Ri(n). Because of the presence of the noises Vi(n) and Wi(n), it is in general impossible to recover exactly the signal. We may only give estimates $\hat{f}(n)$ and $\hat{g}(n)$ of f(n) and g(n).

This problem of source separation has been widely discussed in recent years and has been used for instance in radar and sonar processing, speech enhancement, separation of rotating machine noises, independent component analysis or localization in array processing.

2 Source separation methods

Recently proposed methods for source separation are based on the hypothesis of source independence and test different measurements of the statistical independence. As it has been shown that the second order statistics are not sufficient to give a solution, information contained in the higher order statistics has been exploited. Various solutions relying on the use of the fourth-order moments or cumulants have already been reported (formulation of contrast functions, maximization of normalized cumulants) [1] [2] [3] [4]. Other approaches make use of nonlinear functions and also make higher order statistics appear (including for example the information maximization principle) [5] [6].

The authors generally consider a model with no additive noise (case of adaptive methods or neural networks) or with gaussian noises.

We propose here a new source separation method which derives from an original objective function, based on nonlinear functions of the observations. Specific properties of these functions, chosen as exponential functions, and the hypothesis of independent sources lead to a direct solution for the estimation of the coefficients f(n), g(n). Proposed expressions are deduced from the objective function without any hypotheses on the probability density of the noise signals.

3 A new method of source separation with noisy observations

Let us define the estimation of the sources as two linear combinations of Xi(n) and Yi(n) :

(3) $X'i(n) = Xi(n) - \alpha(n)\, Yi(n)$

(4) $Y'i(n) = Yi(n) - \beta(n)\, Xi(n)$

0-8186-8005-9/97 $10.00 © 1997 IEEE

The two sources are separated if the pair $(\alpha(n), \beta(n))$ is equal to $(f(n),g(n))$ or $(1/g(n),1/f(n))$. The approach is to formulate a two-dimensional complex function $F(\alpha(n), \beta(n))$ such that its mimimum value is obtained for $(f(n),g(n))$ and $(1/g(n),1/f(n))$, and this value is unchanged, whatever the power or the statistical distribution of the noises. In order to avoid the problems generally inherent to adaptive algorithms (local minima, influence of the initialization, low convergence speed), we search a direct computation for the estimation of f(n) and g(n).
We define first nonlinear functions of the combinations X'i(n) and Y'i(n). The influence of the noises can be eliminated only if the non linear functions can be separated into a product of functions of Si(n), Ri(n), Vi(n) or Wi(n). The nonlinearity is necessarily chosen as an exponential complex function, for example $e^{jXi(n)}$ and $e^{jYi(n)}$. This point will be discuss in §3.4.

3.1 Determination of a family of non linear functions

Let us develop the following non linear functions :
$E\{X'i(n).Y'i(n)^*.e^{jXi(n)}\}$, $E\{X'i(n).Y'i(n)^*.e^{jYi(n)}\}$, where * represents the conjugate of a complex value.
This particular function can be separated into a product of functions of Si(n), Ri(n) and Vi(n). It will allow us to eliminate firstly the terms containing the sources Si(n), Ri(n), and secondly the terms containing the noise signals Vi(n), Wi(n).
Due to the hypothesis of independence between the source and noise signals Si(n), Ri(n), Vi(n) and Wi(n), and the exponential non linearity, we obtain complex polynomials of second degrees in $\alpha(n)$ and $\beta(n)$. The complex coefficients contain: $E\{|Si(n)|^2.e^{jSi(n)}\}$, $E\{|Ri(n)|^2.e^{jRi(n)}\}$, $E\{|Vi(n)|^2.e^{jVi(n)}\}$ and $E\{|Wi(n)|^2.e^{jWi(n)}\}$.
These terms can be computed with only non linear functions of the observations such as :
$E\{Xi(n)e^{jXi(n)}\}$, $E\{Xi(n).Yi(n)^*e^{jXi(n)}\}$, $E\{e^{jXi(n)}\}$, $E\{|Xi(n)|^2.e^{jXi(n)}\}$, $E\{Yi(n)e^{jXi(n)}\}$, ...
It allows us to write two two-dimensional complex functions $F(\alpha(n),\beta(n))$, such that $F(f(n),g(n))$ and $F((1/g(n),1/f(n))$ depend only on f(n), g(n) and the observations. More details are explained in appendix.

(5) $F1(\alpha(n),\beta(n)) = J0(n) + J1(n)\,\alpha(n) + J2(n)\,\alpha(n)\beta(n)^* + J3(n)\,\beta(n)^*$

where :

$$J0(n)=E\{Xi(n)Yi(n)^*e^{jXi(n)}\}-E\{e^{jXi(n)}\}.E\{Xi(n)Yi(n)^*\} - \frac{E\{Xi(n)e^{jXi(n)}\}}{E\{e^{jXi(n)}\}}E\{Yi(n)^*e^{jXi(n)}\}$$

$$J1(n)=-E\{|Yi(n)|^2e^{jXi(n)}\}+E\{e^{jXi(n)}\}.E\{|Yi(n)|^2\} + \frac{E\{Yi(n)e^{jXi(n)}\}}{E\{e^{jXi(n)}\}}E\{Yi(n)^*e^{jXi(n)}\}$$

$$J2(n) = J0(n)^*$$

$$J3(n)=-E\{|Xi(n)|^2e^{jXi(n)}\}+E\{e^{jXi(n)}\}.E\{|Xi(n)|^2\} + \frac{E\{Xi(n)e^{jXi(n)}\}}{E\{e^{jXi(n)}\}}E\{Xi(n)^*e^{jXi(n)}\}$$

The second function $F2(\alpha(n),\beta(n))$ is similar to $F1(\alpha(n),\beta(n))$, replacing $e^{jXi(n)}$ with $e^{jYi(n)}$.
Each function $F(\alpha(n),\beta(n))$ verifies :

(6) $$\frac{F(f(n),\,g(n))}{F(1/g(n),1/f(n))} = g(n)^*.\,f(n)^*$$

It also can be written that the pairs $(f(n),g(n))$ and $(1/g(n),1/f(n))$ cancel the functions $G(\alpha(n),\beta(n))$ defined as :

(7) $G(\alpha(n),\beta(n))=F(\alpha(n),\beta(n))- \alpha(n)^*\beta(n)^*F(1/\beta(n),1/\alpha(n))$

Previous equation (7) has an infinite number of solutions but we see from it that for the expected solutions $(f(n),g(n))$ and $(1/g(n),1/f(n))$, the ratio (8)

$$\frac{1+\frac{J0(n)}{J1(n)}\beta(n)}{1+\frac{J0(n)^*}{J1(n)}\beta(n)^*}$$ is equal to:

$$\frac{1-f(n)g(n)}{1-f(n)^*g(n)^*}$$ and $$\frac{1-f(n)g(n)}{1-f(n)^*g(n)^*}\cdot\frac{f(n)^*g(n)^*}{f(n)g(n)}$$. We remark that it depends only on the values of f(n), g(n) and not on the signals.
Suppose now the new observations :

(9) $NXi(n) = \lambda Xi(n) = \lambda Si(n) + f(n)\,\lambda Ri(n) + \lambda Vi(n)$
$NYi(n) = \lambda Yi(n) = g(n)\,\lambda Si(n) + \lambda Ri(n) + \lambda Wi(n)$

where λ is a real number different from 1.
The values NJ0(n) and NJ1(n), computed from equation (9) with NXi(n) and NYi(n), are different from J0(n) and J1(n), computed with Xi(n) and Yi(n). As previously, the ratio (8) depends only on f(n) and g(n) and not on the signals. As a result, we obtain :

(10) $$\frac{1+\frac{J0(n)}{J1(n)}\beta(n)}{1+\frac{J0(n)^*}{J1(n)}\beta(n)^*} = \frac{1+\frac{NJ0(n)}{NJ1(n)}\beta(n)}{1+\frac{NJ0(n)^*}{NJ1(n)}\beta(n)^*}$$

From (10), we deduce that the complex solutions $\beta(n)$ (including g(n) and 1/f(n)) belong to a circle C0, using a geometrical interpretation of the complex space.

(11) $$[\frac{J0(n)\,NJ0(n)^*}{J1(n)\;NJ1(n)}-\frac{J0(n)^*NJ0(n)}{J1(n)\;NJ1(n)}|\beta(n)|^2 + [\frac{J0(n)}{J1(n)}-\frac{NJ0(n)}{NJ1(n)}]\beta(n)+[\frac{NJ0(n)^*}{NJ1(n)}-\frac{J0(n)^*}{J1(n)}]\beta(n)^*=0$$

Due to the nonlinearity of the exponential function, NJ0(n) and NJ1(n) depend on λ and are not proportional to J0(n) and J1(n). It defines a family of non linear functions $F1(\alpha(n),\beta(n),\lambda)$.
C0 is identified using (11). J0(n)/J1(n) and NJ0(n)/NJ1(n) are computed for two values of λ (for example λ=1 and λ=0.5). The statistical moments are estimated by averages on data.

3.2 Proof of an analytical solution

Analyse now the second family of functions $F2(\alpha(n),\beta(n),\lambda)$, by replacing Xi(n) with Yi(n) in the exponential term (or NXi(n) with NYi(n)).

(12)

$$F2(\alpha(n),\beta(n)) = K0(n) + K1(n)\,\alpha(n) + K2(n)\,\alpha(n)\beta(n)^* + K3(n)\,\beta(n)^*$$

$$K0(n)=E\{Xi(n)Yi(n)^*e^{j(Yi(n))}\}-E\{e^{j(Yi(n))}\}.E\{Xi(n)Yi(n)^*\}$$
$$-E\{Xi(n)e^{j(Yi(n))}\}E\{e^{j(Yi(n))}\}^{-1}\,E\{Yi(n)^*e^{j(Yi(n))}\}$$

$$K1(n)=-E\{|Yi(n)|^2e^{j(Yi(n))}\}+E\{e^{j(Yi(n))}\}.E\{|Yi(n)|^2\} + \frac{E\{Yi(n)e^{j(Yi(n))}\}}{E\{e^{j(Yi(n))}\}}E\{Yi(n)^*e^{R(Yi(n))+I(Yi(n))}\}$$

$$K2(n) = K0(n)^*$$

$$K3(n)=-E\{|Xi(n)|^2e^{j(Yi(n))}\}+E\{e^{j(Yi(n))}\}.E\{|Xi(n)|^2\} + \frac{E\{Xi(n)e^{j(Yi(n))}\}}{E\{e^{j(Yi(n))}\}}E\{Xi(n)^*e^{j(Yi(n))}\}$$

As we did before, we compute the values of $F2(\alpha(n),\beta(n))$ for the pairs (f(n),g(n)) and (1/g(n),1/f(n)) :

(13) $F2(f(n), g(n)) = f(n).\,M(Wi(n),Yi(n))$

where M depends on noise and the observations:

$$M(Wi(n),Yi(n))=-E\{Wi(n)^*.e^{j(Yi(n))}\}+E\{|Wi(n)|^2\}.E\{e^{j(Yi(n))}\}$$

(14) $F4(1/g(n),1/f(n))=1/g(n).M(Wi(n),Yi(n))$

(13) and (14) provide a second equation (15), which depends only on the observations Xi(n) and Yi(n), exactly as in (6).

(15) $F2(\alpha(n),\beta(n)) = \alpha(n)\beta(n)\,F2(1/\beta(n), 1/\alpha(n))$

(f(n),g(n)) and (1/g(n),1/f(n)) are necessarily solutions of (15).

(15) may be developed as :

$$\left[1 + \frac{K0(n)^*}{K3(n)}\alpha(n)\right]\left[1 - \alpha(n)^*\,\beta(n)^*\right] = \left[1 + \frac{K0(n)}{K3(n)}\alpha(n)^*\right]\left[1 - \alpha(n)\,\beta(n)\right]$$

As in section 3.1, using two values of λ, we deduce from (15) that the complex solutions $\alpha(n)$ (including f(n) and 1/g(n)) belong to a circle C1 of equation (16).

$$\left[\frac{K0(n)^*}{K3(n)}\frac{NK0(n)}{NK3(n)} - \frac{K0(n)}{K3(n)}\frac{NK0(n)^*}{NK3(n)}\right]|\alpha(n)|^2 + \left[\frac{K0(n)^*}{K3(n)} - \frac{NK0(n)^*}{NK3(n)}\right]\alpha(n) + \left[\frac{NK0(n)}{NK3(n)} - \frac{K0(n)}{K3(n)}\right]\alpha(n)^* = 0$$

Consequently, the values of $1/\alpha(n)$ (including 1/f(n) and g(n)) belong to a straight line L1, which may be easily computed from C1. As we know that 1/f(n) and g(n) also belong to the circle C0, the expected values 1/f(n) and g(n) are necessarily the two intersections between L1 and C0.

The cancellation of the four non linear functions $F(\alpha(n),\beta(n))$ provides only two pairs of solutions (f(n),g(n)) and (1/g(n),1/f(n)). Indeed, the four equations $(G(\alpha(n),\beta(n))=0)$ can be combined into two polynomials in $\alpha(n)$ and two polynomials in $\beta(n)$ such as :

(17) $K0(n)\,|\alpha(n)|^2+K1(n)\alpha(n)-K1(n)^*\alpha(n)^*=0$

$L0(n)^*\,\alpha(n) - L0(n)\,\alpha(n)^* + L1(n) = 0$

We deduce from (17) that the complex solutions $\alpha(n)$ (respectively $\beta(n)$) belong to a circle C0 and a straight line L0 (respectively C1 and L1). We verify that the two intersections between C0 and L0 (respectively C1 and L1) are f(n) and 1/g(n) (respectively g(n), 1/f(n)).

The complex coefficients K0(n), K1(n), L0(n) and L1(n) depend only on the statistics of non linear functions of the observations Xi(n) and Yi(n) and may be estimated from data.

3.3 Influence of the noise powers

Note however, that equations (17) are verified only if noise powers are not negligible. If not :

(18) $F1(f(n),g(n)) = F1(1/g(n),1/f(n)) = 0$

$F2(f(n),g(n)) = F2(1/g(n),1/f(n)) = 0$

An easier solution may be proposed in this case :

(19) $E\{(\overline{Xi}(n) - \alpha(n)\,\overline{Yi}(n))\,.\,(Yi(n) - \beta(n)Xi(n))^*\,.\,e^{R(Xi(n))+I(Xi(n))}\} = 0$

The hypothesis of uncorrelation between sources leads to :

(20) $E\{(Xi(n) - \alpha(n)Yi(n))\,.\,(Yi(n) - \beta(n)\,Xi(n))^*\} = 0$

(19) and (20) are sufficient to solve the problem. It yields to a second order equation in $\beta(n)^*$ whose solutions are $g(n)^*$ and $(1/f(n))^*$. The influence of the noise powers is illustrated with simulation results in §4.

3.4 Determination of the exponential functions

The non linear functions must be separated into a product of functions of Si(n), Ri(n) and Vi(n). It will allow us to eliminate firstly the terms containing the sources Si(n), Ri(n), and secondly the terms containing the noise signals Vi(n), Wi(n). The nonlinearity is necessarily chosen as an exponential complex function, $e^{H(Xi(n))}$ and $e^{H(Yi(n))}$. We see from the previous computations that $e^{H(Xi(n))}$ must be separated into product of $e^{H(Si(n))}$, $e^{H(Ri(n))}$, ... in order to obtain a direct solution of the complex gains f(n) and g(n). It is the only constraint of linearity of the function H. We propose in the paper a particular function H, another one is proposed in [7] : R[Xi(n)]+I[Xi(n)] where R and I are respectively the real and imaginary part of the complex Xi(n). The first one is more robust, regarding the choice of the coefficients λ.

4 Simulation results

We verify the robustness of the algorithm against non gaussian noise. The observations are instantaneous mixtures of complex time sequences. One source admits the four equiprobable values (1, -1, j, -j). The second one is unsymmetrically distributed, with zero mean and unit variance: it is a QAR1 process, constructed from a Gaussian noise. The noise signals are uniformly

distributed with zero mean and unit variance. The complex coefficients f and g are equal to (2,1) and (0.5,1).
In the proposed algorithm, some statistical moments must be estimated, in order to compute the quantities J0(n), J1(n), NJ0(n), NJ1(n), K0(n), K3(n), NK0(n), and NK3(n). For each sample i, we estimated these moments, using time-averages in a recursive way. We observe (figure 1), the error between f of value (2,1) and its estimate $\hat{f}(i)$, $|f - \hat{f}(i)|$. The error has been drawn for two different noise powers. These powers have been computed in order to obtain the two signal to noise ratios, 0dB and 5dB, between noise powers and source powers. We notice that the convergence speed depends on the signal to noise ratio, which seem to be usual. It is about 2000 samples for 5dB and 3500 for 0dB.

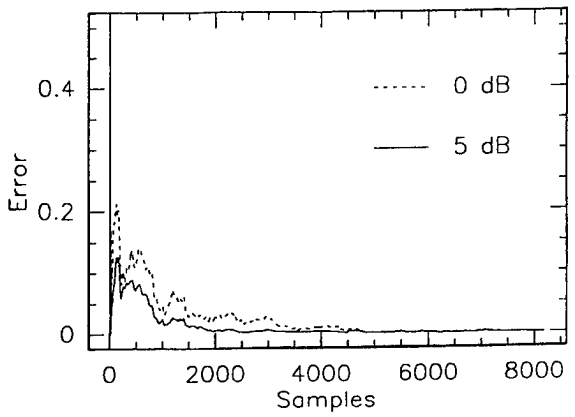


Figure 1 : Estimation error of f in function of time samples

Secondly, we verify the robustness of the algorithm against noise. For that, "asymptotic" performances are analyzed. We show residual errors, computed with 8192 samples : $|f - \hat{f}(8192)|$ for several signal to noise ratios from -8dB to 4dB.

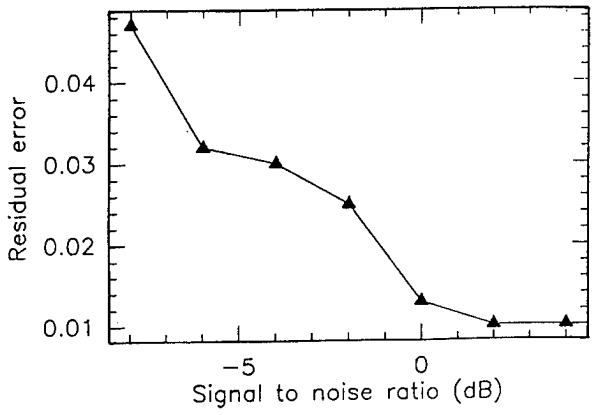


Figure 2 : Residual error in function of signal to noise ratio

We notice on figure 2 that the algorithm is quite robust against noise (low residual error for any signal to noise ratio). However, performances are specially good for a signal to noise ratio higher than 0 dB.
According to what has been shown previously §3.3, the method is available if the noise powers are not negligible. For the proposed example in figure 3, the method remains available up to a 10dB signal to noise ratio. For a signal to noise ratio higher than 10dB, the residual error rapidly increases, and it is better to solve equations (19) (20).

5 Conclusion

We propose a method of source separation, taking into account additive noises. No hypotheses on their probability densities are made. The main characteristic of the method is the use of an objective function based on nonlinear functions (exponential). Specific properties of these functions and the hypothesis of independent sources make it possible to write an analytic solution for the estimation of the two linear filters that combine sources. The convergence speed of the proposed method and its robustness against noise are then illustrated with several simulation results.

Bibliography

[1] P.Comon "Blind identification in presence of noise", in *Proc. of EUSIPCO-92*, Brussels, 24-27 August 1992, Elsevier, Amsterdam, 1992, pp. 835-838
[2] P. Comon "Independent component analysis, a new concept?", *Signal Processing*, vol 36, n°3, April 1994, pp287-314
[3] J.K.Tugnait "On Blind Separation of Convolutive Mixtures of Independent Linear Signals", *IEEE Sig. Proc. Workshop on SSAP*, June 24-26 1996, Corfu, pp 312-315
[4] JF.Cardoso, S.Bose, B.Friedlander "An optimal source separation based on second and fourth order cumulants", *IEEE Sig. Proc. Workshop on SSAP*, June 24-26 1996, Corfu, pp 198-201
[5] AJ.Bell, T.J.Sejnowski "An information-maximization approach to blind separation and blind deconvolution", Neural Computation 7(6) 1995, pp1129-1159
[6] C.Jutten, J.Herault "Blind separation of sources, Part I *Signal Processing*, vol 24, 1991, pp. 1-10
[7] C.Servière, D.Baudois "Source separation with noisy observations: a noise cancelling application", *Signal Processing*, vol 42, 1995, pp 45-57

Appendix

Let us define :
$(1)G1(\alpha(n),\beta(n)) = E\{X'i(n) \,.\, (Y'i(n))^* \,.\, e^{j(Xi(n))}\}$
$=E\{ (Xi(n) - \alpha(n)\, Yi(n)) \,.\, (Yi(n) - \beta(n)\, Xi(n))^* \,.\, e^{j(Xi(n))}\}$
where * represents the conjugate of a complex value.
In the computation of G1(f(n),g(n)) and G1((1/g(n),1/f(n)), we notice the existence of some terms which depend only

on sources such as $E\{Si(n).e^{j(Si(n))}\}$ and $E\{Ri(n)^* e^{j(f(n).Ri(n))}\}$

Now replace the observations Xi(n), Yi(n) with $\overline{X}i(n)$, $\overline{Y}i(n)$, where:

$$(2)\ \overline{X}i(n) = Xi(n) - \frac{E\{Xi(n)e^{j(Xi(n))}\}}{E\{e^{j(Xi(n))}\}}$$

$$(3)\ \overline{Y}i(n) = Yi(n) - \frac{E\{Yi(n)e^{j(Xi(n))}\}}{E\{e^{j(Xi(n))}\}}$$

Consequently, a new model for source separation is given by :

(4) $\overline{X}i(n) = \overline{S}i(n) + f(n)\ \overline{R}i(n) + \overline{V}i(n)$

(5) $\overline{Y}i(n) = g(n)\ \overline{S}i(n) + \overline{R}i(n) + Wi(n)$

Signals $\overline{S}i(n)$, $\overline{X}i(n)$ and $\overline{V}i(n)$ are deduced from Si(n), Ri(n) and Vi(n) by subtraction from a nonzero value such that $\overline{S}i(n).e^{j(Xi(n))}$, $\overline{R}i(n).e^{j(Xi(n))}$ and $\overline{V}i(n)e^{j(Xi(n))}$ have zero means. The noise signal Wi(n) is not modified while Wi(n) is independent of Xi(n). Thus $Wi(n)e^{j(Xi(n))}$ already has a zero mean.

In order to eliminate the terms containing Si(n), Ri(n), in G1(f(n),g(n)) and G1((1/g(n),1/f(n)), we propose to modify the function F1, by replacing $(Xi(n)-\alpha(n)Yi(n))$ with $(\overline{X}i(n)-\alpha(n)\overline{Y}i(n))$.

Consider now:

(6)

$G2(\alpha(n),\beta(n))=E\{(\overline{X}i(n) - \alpha(n)\ \overline{Y}i(n))\ .\ (Yi(n) - \beta(n)\ Xi(n))^*\ .\ e^{j(Xi(n))}\}$

In the computation of G2(f(n),g(n)) and G2((1/g(n),1/f(n)), Si(n) and Ri(n) appear in the terms $E\{\overline{S}i(n).e^{j(Si(n))}\}$ and $E\{\overline{R}i(n)^* e^{j(f(n).Ri(n))}\}$, which are zero, owing to the definition of $\overline{S}i(n)$ and $\overline{R}i(n)$ and the hypothesis of independence of Si(n), Ri(n), Vi(n) and Wi(n), leaving :

(7) $G2(f(n), g(n)) = - g(n)^*.\ E\{\overline{V}i(n)Vi(n)^*\ .\ e^{j(Xi(n))}\}$

$- f(n)\ .\ E\{|Wi(n)|^2\}\ .\ E\{e^{j(Xi(n))}\}$

(8) $G2(1/g(n),1/f(n))=-1/f(n)^*.\ E\{\overline{V}i(n)Vi(n)^*.e^{j(Xi(n))}\}$

$-1/g(n)\ .\ E\{|Wi(n)|^2\}\ .\ E\{e^{j(Xi(n))}\}$

Subsequently, we eliminate the terms containing Wi(n) in (7), (8). For that, we use the information of uncorrelation between Si(n), Ri(n), Vi(n) and Wi(n), which provides :

(9) $-E\{(Xi(n)-f(n)Yi(n)).(Yi(n)-g(n)Xi(n))^*\} = f(n)\ E\{|Wi(n)|^2\} + g(n)^*\ E\{|Vi(n)|^2\}$

Now substitute $E\{|Wi(n)|^2\}$ from (9) to (7) and (8) which yields :

(10) $G2(f(n),g(n))-E\{e^{j(Xi(n))}\}.E\{[Xi(n)-f(n)Yi(n)].[Yi(n)-g(n)Xi(n)]^*\} = g(n)^*.\ L(Vi(n),Xi(n))$

where :

$L(Vi(n),Xi(n))=-E\{\overline{V}i(n)Vi(n)^*.e^{j(Xi(n))}\}+ E\{|Vi(n)|^2\}.E\{e^{j(Xi(n))}\}$

(11) $G2(1/g(n),1/f(n))-E\{e^{j(Xi(n))}\}.E\{[Xi(n)-1/g(n)(n)Yi(n)].[Yi(n)-1/f(n)Xi(n)]^*\}$

$=1/f(n)^*.\ L(Vi(n),Xi(n))$

As a result, let the new two-dimensional function G3, be defined by :

(12)

$G3(\alpha(n),\beta(n))=F2(\alpha(n),\beta(n))-E\{e^{j(Xi(n))}\}.E\{[Xi(n)-\alpha(n)Yi(n)].[Yi(n)-\beta(n)Xi(n)]^*\}$

$=E\{\ [\overline{X}i(n) - \alpha(n)\ \overline{Y}i(n)]\ .\ [Yi(n) - \beta(n)\ Xi(n)]^*\ .\ e^{j(Xi(n))}\} - E\{e^{j(Xi(n))}\}\ .\ E\{[Xi(n) - \alpha(n)Yi(n)].\ [Yi(n) - \beta(n)\ Xi(n)]^*\}$

G3 may also be written as :

(13) $G3(\alpha(n),\beta(n)) = J0(n) + J1(n)\ \alpha(n) + J2(n)\ \alpha(n)\beta(n)^* + J3(n)\ \beta(n)^*$

with :

$J0(n)=E\{Xi(n)Yi(n)^*e^{j(Xi(n))}\}- E\{e^{j(Xi(n))}\}.E\{Xi(n)Yi(n)^*\}$

$$- \frac{E\{Xi(n)e^{j(Xi(n))}\}}{E\{e^{j(Xi(n))}\}} E\{Yi(n)^*e^{j(Xi(n))}\}$$

$J1(n)= -E\{|Yi(n)|^2e^{j(Xi(n))}\} + E\{e^{j(Xi(n))}\}.E\{|Yi(n)|^2\}$

$$+ \frac{E\{Yi(n)e^{j(Xi(n))}\}}{E\{e^{Rj(Xi(n))}\}} E\{Yi(n)^*e^{R(Xi(n))}\}$$

$J2(n) = J0(n)^*$

$J3(n)=-E\{|Xi(n)|^2e^{j(Xi(n))}\} + E\{e^{j(Xi(n))}\}.E\{|Xi(n)|^2\}$

$$+\frac{E\{Xi(n)e^{j(Xi(n))}\}}{E\{e^{j(Xi(n))}\}} E\{Xi(n)^*e^{j(Xi(n))}\}$$

We observe from (9) and (10) that :

$$(14)\ \frac{F3(f(n),\ g(n))}{F3(1/g(n),1/f(n))} = g(n)^*.\ f(n)^*$$

Consequently, the nonlinear function $G3(\alpha(n),\beta(n))$ verifies:

(15) $G3(f(n),g(n))=g(n)^*.f(n)^*.[J0(n)+J1(n).1/g(n)+ J2(n).1/g(n).1/f(n)^*+J3(n).1/f(n)^*]$

where J0(n), J1(n), J2(n) and J3(n) depend only on the observations Xi(n) and Yi(n).

A similar expression may be written for G3(1/g(n),1/f(n)) :

(16)

$G3(1/g(n),1/f(n))=1/g(n)^*.1/f(n)^*.[J0(n)+J1(n).f(n)+J2(n).f(n).g(n)^*+J3(n).g(n)^*]$

(13), (14) and (15) naturally provide a nonlinear equation whose pairs (f(n),g(n)) and (1/g(n),1/f(n)) are necessarily solutions :

(17) $G3(\alpha(n),\beta(n)) = \alpha(n)^*\beta(n)^*G3(1/\beta(n),1/\alpha(n))$

It may be developed as :

$$\left[1+\frac{J0(n)}{J1(n)}\beta(n)\right]\left[1-\alpha(n)^*\beta(n)^*\right] =$$

$$\left[1+\frac{J0(n)^*}{J1(n)}\beta(n)^*\right]\left[1-\alpha(n)\beta(n)\right]$$

We obtain then the equation (15).

Blind Separation of Convolutive Mixtures: A Gauss-Newton Algorithm

Sergio Cruces
TSC Group, ESI Telecomunicación
Universidad de Sevilla
Av. Reina Mercedes, 41012-Sevilla, Spain
sergio@viento.us.es

Luis Castedo
Dpto. Electrónica y Sistemas
Universidad de La Coruña
Campus de Elviña, 15071-La Coruña, Spain
luis@sol.des.fi.udc.es

Abstract

*This paper addresses the blind separation of convolutive mixtures of independent and non-Gaussian sources. We present a block-based Gauss-Newton algorithm which is able to obtain a separation solution using only a specific set of output cross-cumulants and the hypothesis of soft mixtures. The order of the cross-cumulants is chosen to obtain a particular form of the Jacobian matrix that ensures convergence and reduces computational burden. The method can be seen as an extension and improvement of the Van-Gerven's symmetric adaptive decorrelation (**SAD**) method. Moreover, the convergence analysis presented in the paper provides a theoretical background to derive an improved version of the Nguyen-Jutten algorithm.*

1. Introduction

Blind source separation is receiving a growing interest due to its applications in diverse fields such as array processing, multiuser communications, etc ([10]-[8]). While most of the work has been developed in the context of instantaneous mixtures, the more difficult problem of convolutive mixtures has received less attention. Recently, Van Gerven *et al.* [15] have proposed and analyzed the Symmetric Adaptive Decorrelation (**SAD**) algorithm that only uses second order statistics and exhibits interesting signal separation properties. However, it is well known that methods based on second order statistics are not sufficient to solve the blind source separation problem and higher order statistics are necessary. For this reason, Nguyen and Jutten [12] have proposed an algorithm (**NJ**) that cancels fourth order cross-cumulants of the output signals to achieve separation. Nevertheless, the convergence of this algorithm has not been analyzed yet.

*This work has been supported by the National Research Plan of Spain, CICYT (Grant No. TIC96-0500-C10-08) and Xunta de Galicia (Grant No. XUGA 10502A96)

In this paper we present a method that generalizes and improves the **SAD** algorithm using higher order cross-cumulants. From the given method, it also can be derived an algorithm which is similar the **NJ** algorithm for two sources except for a term that ensures asymptotic stability.

2. Signal model

Let us assume the signal model presented in Figure 1. There are N independent sources $\mathbf{s}[n] = [s_1[n], \ldots, s_N[n]]^T$ where at most one is Gaussian. The Darmois-Skitovich theorem [3] ensures that the sources can be separated by imposing pairwise independence between them. The sources are convolutively mixed through a multichannel linear time-invariant system, with memory, to obtain the observations vector $\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^T$, where $M \geq N$. The impulse response of the mixing system is characterized by the sequence of $M \times N$ matrices $\mathbf{A}[n] = [a_{ij}[n]]$. We will assume that the mixing system is **FIR** and (possibly) non-causal. As a consequence, $\mathbf{A}[n] \neq \mathbf{0}$ for $-L_{a1} \leq n \leq L_{a2}$, and the relationship between sources and observations can be written as

$$\mathbf{x}[n] = \mathbf{A}[n] * \mathbf{s}[n] = \sum_{k=-L_{a1}}^{L_{a2}} \mathbf{A}[k]\mathbf{s}[n-k]$$

In order separate the sources, the observations are processed by another multichannel **LTI** system, with memory, characterized by its $M \times N$ impulse response sequence matrix $\mathbf{B}[n]$. Again, the separating system will be **FIR**, (possibly) non causal and non-zero in the interval $-L_{b1} \leq n \leq L_{b2}$. Therefore, the output vector $\mathbf{y}[n] = [y_1[n], \ldots, y_N[n]]^T$ is given by

$$\mathbf{y}[n] = \mathbf{B}[n] * \mathbf{x}[n] = \sum_{k=-L_{b1}}^{L_{b2}} \mathbf{B}[k]\mathbf{x}[n-k]$$

There are several indeterminacies in the source separation problem [14]. Our approach can handle the sources

0-8186-8005-9/97 $10.00 © 1997 IEEE

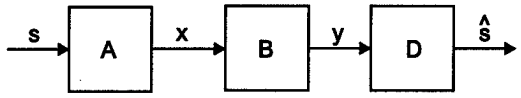


Figure 1. Feed-forward system model.

ordering indeterminacy, but we must avoid the scaling and delay indeterminacy. Towards this aim, we will suppose that the diagonal terms of the separation sequence matrices $\mathbf{B}[k]$ and that of $\mathbf{A}[k]\mathbf{P}$ are equal to the unit impulse $\delta[k]$ ($\mathbf{P}$ is an unknown permutation matrix).

Once the signals $\mathbf{y}[n]$ are separated, it is necessary to introduce a post-processing system $\mathbf{D}(z)$ in order to remove the undesired correlations originated by the separation process. For this task, it can be shown [5] that the $\mathbf{D}(z)$ transfer function should be diagonal with elements given by

$$D_{ii}(z) = \frac{Adj_{ii}(\mathbf{B}(z))}{Det[\mathbf{B}(z)]} \qquad i = 1, \ldots, N. \tag{1}$$

where $Adj_{ii}(\cdot)$ is the adjoint operator on the (i,i) matrix element. At the output of the post-processing system we obtain the vector $\hat{\mathbf{s}}[n]$ of estimated sources.

The following hypothesis will be needed for the algorithm to perform adequately

1. The mixture is soft in the sense that each observed signal receives a dominant contribution from one of the sources. This typically occurs when each sensor is closer to a different source and these have similar power.

2. There is a subset of cumulants of the sources in which the cumulants of the same order do not have great differences in their magnitude. This is obviously true when the sources are identically distributed.

The soft mixture assumption is appropiate for the separation case of convolutive mixtures, since a sufficient condition for the stability of the post-processing filters also depends on this hypothesis. It can also be shown that the soft mixtures assumption will lead to a $\mathbf{D}(z)$ that is not far from the identity matrix, fact that will be important in later demonstrations.

The separation filter length $L_b = L_{b1} + L_{b2} + 1$ is chosen to match the length of the filters in the numerator of the inverse multichannel transfer function. Therefore $L_{b1} = L_{a1}(N-1)$ and $L_{b2} = L_{a2}(N-1)$ if $M = N$, and $L_{b1} = 3L_{a1}(N-1)$ and $L_{b2} = 3L_{a2}(N-1)$ when $M > N$. The inverse of a **FIR** multichannel transfer function has some **IIR** components, these components will be provided by $\mathbf{D}(z)$ in the two sources case. Nevertheless, in the multiple sources case, our approach will be only approximate. We can't reach exactly the solution but, under the soft mixture hypothesis, the **IIR** part will be negligible or approximable by setting a bit longer the separation filters length.

3. Statistical dependence measure

We define the cumulant tensor $\mathbf{C}_{\alpha_1,\ldots,\alpha_l}(\mathbf{y}[n],\ldots,\mathbf{y}[n])$ as the tensor where the $(i_1,\ldots,i_l)$ element is $C_{\alpha_1,\ldots,\alpha_l}(y_{i_1}[n],\ldots,y_{i_l}[n]) \triangleq Cum(y_{i_1}[n]\times\alpha_1,\ldots,y_{i_l}[n]\times\alpha_l)$ the $(\alpha_1,\ldots,\alpha_l)$-order cross-cumulant between the signals $y_{i_1}[n],\ldots,y_{i_l}[n]$.

We propose that the adjustable elements of the separating matrix $\mathbf{B}(z)$ be selected to jointly diagonalize the set of cumulant matrices $\mathbf{C}_{\alpha,\beta}(\mathbf{y}[n],\mathbf{y}[n-k])$ for the different lags $k = -L_{b1},\ldots,L_{b2}$ and for the set Ω of n_Ω pairs (α,β) of natural numbers. Towards this aim, we propose to minimize the following cost function

$$\mathbf{\Phi_\Omega} = \sum_{(\alpha,\beta)\in\Omega} w_{(\alpha,\beta)}\phi_{(\alpha,\beta)} \tag{2}$$

$$\phi_{(\alpha,\beta)} = \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{k=-L_{b1}}^{L_{b2}} \left(C_{\alpha,\beta}(y_i[n],y_j[n-k])\right)^2 \tag{3}$$

The $w_{\alpha,\beta}$ are a set of weighting coefficients. Note that $\mathbf{\Phi}_\Omega$ can be interpreted as a statistical dependence measure of the output vector $\mathbf{y}[n]$. Initially, we will consider the set Ω containing all the possible pairs (α,β). However, it will be shown that the minimization of the subset $(1,\beta)$ is sufficient to ensure separation under the soft mixture assumption. Trivial non-separation solutions, where any of the outputs is equal to zero, occur when the mixing or separating matrices are not full row rank. The convergence to these solutions is avoided by the soft mixture hypothesis since, when this hypothesis holds, we found to be on the basin of attraction of the correct separation solution.

4. Minimization algorithm

For simplicity reasons, we will rewrite the dependence measure Φ_Ω in a vector form. Let us define the vector $\mathbf{z}_{(\alpha,\beta)} = \{C_{\alpha,\beta}(y_i[n],y_j[n-k]);\quad i,j|_{j\neq i} = 1\cdots N,\ k = -L_{b1},\cdots,L_{b2}\}$. If we now define the vector $\mathbf{z} = [\sqrt{w_{(\alpha,\beta)_1}}\mathbf{z}^T_{(\alpha,\beta)_1},\ldots,\sqrt{w_{(\alpha,\beta)_{n_\Omega}}}\mathbf{z}^T_{(\alpha,\beta)_{n_\Omega}}]^T$ with $(\alpha,\beta)_i \in \Omega$, the dependence measure can be written as the inner product $\mathbf{\Phi_\Omega} = \mathbf{z}^T\mathbf{z}$. It is also convenient to rearrange the separation variables in a vector $\mathbf{b} = \{b_{ij}[k];\quad i = 1,\ldots,N;\ j|_{j\neq i} = 1,\ldots,M;\ k = -L_{b1},\ldots,L_{b2}\}$.

Different algorithms can be used to minimize $\mathbf{\Phi}_\Omega$. Due to the soft mixture hypothesis we are initially close to the separation solution and the vector $\mathbf{z}$ of output cross-cumulants is small (with norm close to zero). In this situation, we can use the Gauss-Newton (**GN**) method [6] to

find the desired minimum, because the Hessian matrix of the adaptation can be well approximated from the Jacobian matrix. This reason and the quadratic convergence of the **GN** method reduces the computational burden.

The Gauss-Newton method uses the gradient and the Hessian of $\mathbf{\Phi}_\Omega$ with respect to the separation coefficients. The gradient is given by: $\nabla_{\mathbf{b}}\mathbf{\Phi}_\Omega = 2\mathbf{J}\mathbf{z}$ where $\mathbf{J} = [\frac{\partial \mathbf{z}}{\partial b_1}, \cdots, \frac{\partial \mathbf{z}}{\partial b_{L_b M(N-1)}}]^T$ is the Jacobian matrix. The global Jacobian matrix can also be defined in terms of the sub-Jacobian matrices $\mathbf{J}_{(\alpha,\beta)} = \nabla_{\mathbf{b}}\mathbf{z}_{(\alpha,\beta)}$, since $\mathbf{J} = [\sqrt{w_{(\alpha,\beta)_1}}\mathbf{J}_{(\alpha,\beta)_1}, \cdots, \sqrt{w_{(\alpha,\beta)_{n_\Omega}}}\mathbf{J}_{(\alpha,\beta)_{n_\Omega}}]$. Near the solution the Hessian can be approximated by $\mathbf{H} \approx 2\mathbf{J}\mathbf{J}^T$. Then, the Gauss-Newton method consists in the following iteration

$$\mathbf{b}|_{n+1} = \mathbf{b}|_n + \mu\Delta\,, \qquad \mathbf{H}\cdot\Delta = -2\mathbf{J}\mathbf{z} \tag{4}$$

where $\mu \in (0,2)$ is the step-size. Denoting by $\mathbf{J}^{T\#}$ the pseudo-inverse of $\mathbf{J}^T$, the increment Δ is calculated as $\Delta = -\mathbf{J}^{T\#}\mathbf{z}$ or, in an equivalent way, by solving the following system of linear equations

$$\left(\sum_{i=1}^{n_\Omega} w_{(\alpha,\beta)_i}\mathbf{J}_{(\alpha,\beta)_i}\mathbf{J}^T_{(\alpha,\beta)_i}\right)\Delta = -\sum_{i=1}^{n_\Omega} w_{(\alpha,\beta)_i}\mathbf{J}_{(\alpha,\beta)_i}\mathbf{z}_{(\alpha,\beta)_i}$$

4.1. Jacobian structure

In this section, we will choose an adequate set Ω to reduce the computational burden of the algorithm (4). We will exploit the Jacobian structure for this selection, showing that some few elements of the Jacobian matrix has a strong dominance over the rest. This enables to approximate the Jacobian matrix by an strong sparse matrix. As a consequence, the computational burden of the algorithm will be reduced for two reasons: there are less cumulants to estimate and there exist efficient methods to solve linear sparse systems [7].

Each sub-Jacobian $\mathbf{J}_{(\alpha,\beta)}$ is a $L_b M(N-1)$-by-$L_b N(N-1)$ rectangular matrix. Its elements are proportional to

$$\begin{aligned}\frac{\partial C_{\alpha,\beta}(y_i[n], y_j[n-k])}{\partial b_{rs}[m]} &= \\ = \;& \delta_{ir}\,\alpha\, C_{1,\alpha-1,\beta}(x_s[n-m], y_i[n], y_j[n-k]) + \\ & \delta_{jr}\,\beta\, C_{1,\alpha,\beta-1}(x_s[n-k-m], y_i[n], y_j[n-k])\end{aligned} \tag{5}$$

where $\delta_{ir} = \{1 \text{ if } i = r,\ 0 \text{ else}\}$, and where the index (i,j,k,r,s,m) belong to the range $R = \{i,j|_{j\neq i}, r = 1,\ldots,N;\ s|_{s\neq r} = 1,\ldots,M;\ m,k = -L_{b1},\ldots,L_{b2}\}$. The resulting expression shows that, for (α,β) both greater than 1 and under the assumption of soft mixtures, all the Jacobian elements are close to zero, since the signals $y_i[n]$ and $y_j[n]$ will have an small and decreasing dependence between each other. Therefore, $\Phi_{(\alpha,\beta)}$ has a slow variation or sensitivity with respect to the separation coefficients, and this makes this kind of pairs undesirable to use in a gradient based algorithm. When α or β are equal to 1, the Jacobian is no longer close to zero, therefore avoiding the former problem. We will set $\alpha = 1$ since this choice leads to a well behaved structure for the Jacobian matrix while $\beta = 1$ does not. Substituting $\alpha = 1$ in expression (5) and neglecting the non-dominant terms around the separation solution, we arrive at this first approximation for the Jacobian elements

$$\frac{\partial C_{1,\beta}(y_i[n], y_j[n-k])}{\partial b_{rs}[m]} \approx \delta_{ir} C_{1,\beta}(x_s[n-m], y_j[n-k])$$

The above expression is not true for the case $\beta = 1$, but even in that case constitutes a valid approximation for the Jacobian.

When the sources can be considered locally stationary with respect to the delay L_b (e.g. voice signals) or when we work with a block based method of n_d samples where $n_d >> L_b$, the cross-cumulant estimates can be considered stationary. This simplifies the last approximation to

$$\frac{\partial C_{1,\beta}(y_i[n], y_j[n-k])}{\partial b_{rs}[m]} \approx \tag{6}$$
$$\begin{cases} \delta_{ir} C_{1,\beta}(x_s[n-m+k], y_j[n]) & \text{if } k \leq m \\ \delta_{ir} C_{1,\beta}(x_s[n], y_j[n-k+m]) & \text{if } k \geq m \end{cases}$$

The Jacobian can now be computed from $n_\Omega 2L_b M(N-1)$ cumulant estimates. We see that the structure of the sub-Jacobians $\mathbf{J}_{(\alpha,\beta)}$ is sparse, they are block toeplitz with blocks of dimension $M(N-1)$, and each block is block diagonal formed by sub-blocks of dimension $(N-1)$.

When the sources are white with respect to its cumulants of order $1+\beta$ or when the mixture is instantaneous, the approximation (6) can be further simplified in the following way

$$\frac{\partial C_{1,\beta}(y_i[n], y_j[n-k])}{\partial b_{rs}[m]} \approx \delta_{ir}\delta_{km} C_{1,\beta}(x_s[n], y_j[n]) \tag{7}$$

In this case, the Jacobian precises only $n_\Omega M(N-1)$ cumulant computations. The sub-Jacobian matrices $\mathbf{J}_{(\alpha,\beta)}$ are block diagonal with blocks of dimension $N-1$. The proposed algorithm for the two white sources case reduces to the coefficients adaptation:

$$\begin{aligned} b_{ij}[k]|_{n+1} &= b_{ij}[k]|_n - \mu\cdot \\ &\frac{\sum_{(1,\beta)\in\Omega} w_\beta\, C_{1,\beta}(x_j[n], y_j[n])\, C_{1,\beta}(y_i[n], y_j[n-k])}{\sum_{(1,\beta)\in\Omega} w_\beta\, C^2_{1,\beta}(x_j[n], y_j[n])} \\ & i,j|_{j\neq i} = 1,\ldots,2; k = -L_{b1},\ldots,L_{b2} \end{aligned} \tag{8}$$

Recall that this iteration only requires $2n_\Omega(L_b+1)$ cumulants!

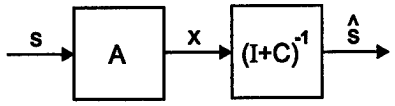


Figure 2. Feed-backward system model.

5. Link with another separation methods

For the moment, we have used the feed-forward (**FF**) separation structure (Figure 1). Another possibility is the feed-backward (**FB**) filtering structure that can be seen in Figure 2. This structure has the advantage of avoiding post-filtering by a post-processing matrix, but its analysis is not simple due to the feedback. In this structure, the transfer function of the separation filter is $(\mathbf{I}+\mathbf{C}(z))^{-1}$ where $\mathbf{C}(z)$ is a matrix with the same dimensions as $\mathbf{A}(z)$ and zero diagonal elements. We can establish a connection between our feed-forward algorithm and a similar feed-backward one under the assumption of soft mixtures.

The relation between the feed-backward and the feed-forward structure is given by $\mathbf{D}(z)\mathbf{B}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1}$. For the two sources case, the solution of this equation reduces to $\mathbf{C}(z) = \mathbf{I} - \mathbf{B}(z)$. This establishes the connection between both filtering approaches. For more than two sources and soft mixtures, we may approximate $(\mathbf{I} + \mathbf{C}(z))^{-1}$ by the linear part of its Taylor series expansion around $\mathbf{C}(z) = \mathbf{0}$ which is $(\mathbf{I} + \mathbf{C}(z))^{-1} \approx \mathbf{I} - \mathbf{C}(z)$. Since for the soft mixtures hypothesis we can approximate the post-processing matrix by the identity $\mathbf{D}(z) \approx \mathbf{I}$, the connection between both filtering approaches still takes the form $\mathbf{C}(z) \approx \mathbf{I} - \mathbf{B}(z)$. Having this in mind, we can propose a feed-backward separation algorithm simply updating the coefficients $c_{ij}[k]$ in the same way the feed-forward algorithm does, but changing the adaptation sign and replacing $\mathbf{y}[n]$ by $\hat{\mathbf{s}}[n]$. The algorithm equivalent to (8) in the feed-backward case is

$$c_{ij}[k]|_{n+1} = c_{ij}[k]|_n + \mu \cdot \frac{\sum_{(1,\beta)\in\Omega} w_\beta \, C_{1,\beta}(x_j[n], \hat{s}_j[n]) \, C_{1,\beta}(\hat{s}_i[n], \hat{s}_j[n-k])}{\sum_{(1,\beta)\in\Omega} w_\beta \, C^2_{1,\beta}(x_j[n], \hat{s}_j[n])}$$
$$i, j|_{j\neq i} = 1, \ldots, 2; k = -L_{a1}, \ldots, L_{a2} \qquad (9)$$

This algorithm will also be asymptotically stable for $\beta > 1$ since it converges to the Gauss-Newton algorithm when approaching to the separation solution. When $n_\Omega = 1$ it simplifies to

$$c_{ij}[k]|_{n+1} = c_{ij}[k]|_n + \mu \frac{C_{1,\beta}(\hat{s}_i[n], \hat{s}_j[n-k])}{C_{1,\beta}(x_j[n], \hat{s}_j[n])}$$
$$i, j|_{j\neq i} = 1, \ldots, 2; k = -L_{a1}, \ldots, L_{a2}$$

It can seen that when $\beta = 1$ this algorithm reduces to the **SAD** algorithm [15]. For $\beta = 3$ it takes a similar, but improved, form of the **NJ** algorithm [12]. The improvement over this last algorithm is three-fold: the proposed algorithm has a greater speed of convergence, it works with non causal mixtures, and it is asymptotically stable under the given hypothesis while **NJ** is not stable.

6 Separation example

Before beginning with the separation example we will introduce some additional definitions. The combined multichannel impulse response sequence of mixing and separating systems will be denoted by $\mathbf{H}[n] = \mathbf{B}[n] * \mathbf{A}[n]$. Let us define the matrix $\mathbf{H_E} = \left[\sum_{k=-\infty}^{\infty} H^2_{ij}[k]\right]_{ij}$ whose elements are the energy of $\mathbf{H}[n]$. We normalize the energy matrix in the following way: $\bar{\mathbf{H}}_\mathbf{E} = \frac{1}{2}(\mathbf{D}_r^{-1}\mathbf{H_E} + \mathbf{H_E}\mathbf{D}_c^{-1})$ where $\mathbf{D}_r$ and $\mathbf{D}_c$ are diagonal matrices whose diagonal elements are the maxima of the rows and columns of $\mathbf{H_E}$, respectively. This performance index is similar to the one used in [11]. The normalized matrix $\bar{\mathbf{H}}_\mathbf{E}$ is invariant to scalation by diagonal matrices and time shifts of the transfer functions. Its maxima coefficients by columns or rows are equal to 1, and will be associated with the normalized energy of the direct transfer functions. The rest will be associated to the normalized energy of the interference transfer functions.

Let us suppose we have two independent white source signals $\mathbf{s}[n]$ (of 5000 samples) which are mixed through the channel $\mathbf{A}(z)$ to obtain a convolutive mixture $\mathbf{y}[n]$. The mixing matrix is formed by five taps FIR filters, and verify the soft mixture condition. The initial filters energy matrix is $\bar{\mathbf{H}}_\mathbf{E} = [\,1\,,\ 0.21\ ;\ 0.98\,,\ 1\,]$. We run a few iterations of the feed-forward separation algorithm presented in (8) to separate sources. We will use the set of cumulants $\Omega = \{(1,3)\}$ for the dependence measure. At each iteration we show in Figure 3 and Figure 4 the evolution and performance of the algorithm. In Figure 3 we show how the interference energy is reduced whereas the direct path signals energy is preserved. The final energy matrix is $\bar{\mathbf{H}}_\mathbf{E} = [\,1\,,\ 0.0009\ ;\ 0.0009\,,1\,]$ which is close to the identity matrix. The algorithm does not converge exactly to the identity matrix when using cross-cumulant estimates due to the finite length of the data. If we have the exact cross-cumulants at each iteration the convergence to the separation solution is achieved with an arbitrary precision. Figure 4 presents the minimization of the dependence measure. It can be seen that the theoretical convergence of the algorithm is quadratic, and therefore, reaches the separation solution in a few iterations.

7 Conclusions

A new signal separation algorithm is presented for the convolutive mixture of independent and non-Gaussian

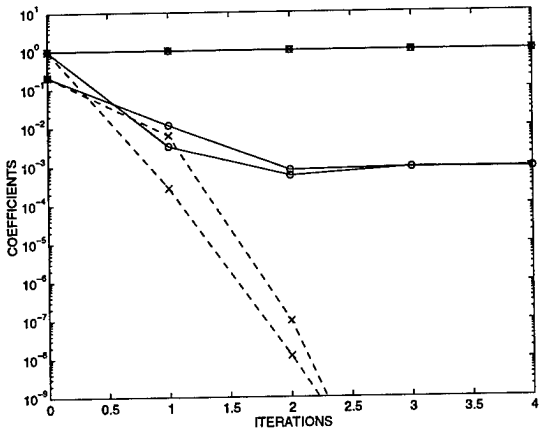


Figure 3. Coefficients of the filter energy matrix $\bar{H}_E$ vs. iterations when using: '–o– ' cross-cumulants estimates, '- -x- -' exact cross-cumulants.

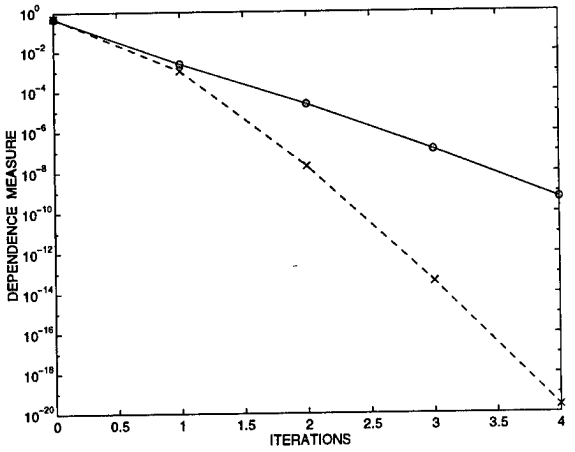


Figure 4. Dependence measure Φ_Ω vs. iterations: '–o–' using cross-cumulants estimates, '- -x- -' using exact cross-cumulants.

sources. It was shown that it generalizes and improves a class of convolutive separation algorithms that have been previously proposed, giving then a theoretical justification for their convergence in the soft mixtures case. The algorithm was obtained for feed-forward and feed-backward structures, and has the following advantages: it does not depend on the signals cumulants sign to converge, it has a fast convergence since it is a Gauss-Newton method, and it can work with source signals of arbitrary probability density functions which may have a set of non-zero cumulants, while the moment based algorithms are limited to signals of even pdfs. Finally, it is robust in the sense that it still allows source separation when the signals cumulants of a certain order are zero provided that Ω contains other cumulants orders which do not vanish, e.g. this allows to achieve separation when there is one Gaussian signal if $(1,1) \in \Omega$.

Acknowledgment

The authors wish to thank J.I. Acha from the University of Seville (Spain) for his cooperation and fruitful discussions on the signal separation topic.

References

[1] X.-R. Cao and R. wen Liu. General approach to blind source separation. *IEEE Transactions on Signal Processing*, 44(3):562–571, Mar. 1996.

[2] J. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, 44(12):3017–3030, Dec. 1996.

[3] P. Comon. Independent component analysis, a new concept? *Signal Processing*, (36):287–314, 1994.

[4] P. Comon. Contrasts for multichannel blind deconvolution. *IEEE Transactions on Signal Processing*, July 1996.

[5] S. Cruces, R. Martin, and J. Acha. Blind separation of sources: Stability results and comparisons. In *Fourth COST 229 Workshop on Adaptive Methods and Emergent Techniques for SP and Com.*, June 1996.

[6] R. Fletcher. *Practical Methods of Optimization*. John Willey and Sons, second edition edition, 1987.

[7] G. Golub and C. Van-Loan. *Matrix Computations*. Johns Hopkins, third edition, 1996.

[8] C. Jutten and J.-F. Cardoso. Separation of sources: Really blind? In *International Symposium on Nonlinear Theory and Applications (NOLTA)*, Las Vegas (USA), Dec. 1995.

[9] C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, (24):1–10, 1991.

[10] J. Karhunen. Neural approaches to independent component analysis and source separation. *Proc 4th European Symposium on Artificial Neural Networks (ESANN '96)*, 1996.

[11] E. Moreau and O. Macchi. High-order contrasts for self-adaptive source separationcriteria for complex source separation. *International Journal of Adaptive Control and Signal Proc.*, 10:19–46, 1996.

[12] H.-L. Nguyen-Thi and C. Jutten. Blind source separation for convolutive mixtures. *Signal Processing*, (45):209–229, 1995.

[13] E. Sorouchyari. Blind separation of sources, part iii: Stability analysis. *Signal Processing*, (24):21–29, 1991.

[14] L. Tong, R.-W. Liu, V.-C. Soon, and Y.-F. Huang. Indeterminacy and identifiability of blind identification. *IEEE Trans. on Circuits and Systems*, 38(5):499–509, May 1991.

[15] S. Van-Gerven and D. Van-Compernolle. Signal separation by symmetric adaptive decorrelation: Stability, convergence, and uniqueness. *IEEE Transactions on Signal Processing*, 43(7):1602–1612, July 1995. 1995.

Identifiability of the Superimposed Signals Using Third-Order Cumulants

Danilo Korže, Damjan Zazula
Faculty of Electrical Engineering and Computer Science
University of Maribor
Smetanova 17, 2000 Maribor, Slovenia
korze@uni-mb.si, zazula@uni-mb.si

Abstract

The paper deals with the variance analysis estimating the third-order cumulant of the MA system output signal which was excited with the Bernoulli distributed white noise on the input. These scheme was used to simulate the superimposed systems, or signals, which often arise in the practical approaches (EMG signal decomposition). The equations were derived which show dependencies of the variances of the third-order output cumulant values on the lenght of the output signal and on the values of the samples in the impulse response of the unknown system as well. The structure of the variance inside the third-order cumulant array was observed. The speed and the convergence of the identification procedure could be predicted using these results for the case of superimposed MA systems. The knowledge of the structure of the variance is helpfull in designing new identification methods. Some examples are given to show the correctness of the derived equations.

1. Introduction

Despite the variaty of different methods based on cumulants which are theoretically correct [1, 2], it could be noticed in computer simulations that the results are strongly dependent on the values in the impulse response of the uknown system being identified. That effect could be noted with all kind of system, i.e. minimum, maximum or mixed phase. Besides, two very similar MA systems could give much different quality of the identification results, if the variance of the result is the measure of the quality.

On the other hand, higher-order statistical methods have proved successful in decomposition of superimposed signals modelled as outputs of multichannel systems with Bernoulli input excitation [3, 4]. To connect the dependencies of the results on the type of the impulse response used and on the type of the superimposed MA systems observed, the paper develops an analytical variance model of the third-order cumulants, whose outcomes are further used in evaluation of a possible signal decomposition.

In Section 2, the variance for the third-order cumulant is derived, with the Bernoulli distributed input white noise. In Section 3, some simulation results are described and the structure of variances in the third-order cumulant array is observed. The conclusions are gathered in Section 4.

2. Variance analysis for the third-order cumulant

An MA(q) system with the Bernoulli distributed input is considered:

$$W \sim \begin{array}{c|c} 1-p & -p \\ \hline p & 1-p \end{array}, \tag{1}$$

where p is the probability of the positive pulse in the input signal.

The probability function for the output signal $y(n)$ of the MA system could be written as:

$$Y \sim \begin{array}{c|c|c|c|c} Y_0 & \cdots & Y_i & \cdots & Y_{2^{q+1}} \\ \hline (1-p)^{q+1} & \cdots & p^{k_{1,i}}(1-p)^{k_{0,i}} & \cdots & p^{q+1} \end{array} \tag{2}$$

where the expression $k_{1,i}$ in (2) denotes the number of ones and the expression $k_{0,i}$ the number of zeros in the binary presentation of i. Exactly 2^{q+1} different values could be obtained in the output signal, because each sample in the system impulse response is multiplied either by the value of $-p$ or by the value of $1-p$ in the Bernoulli distributed input noise. The values in the output signal are computed as:

$$Y_i = \sum_{j=1}^{k_{1,i}} h(i_j) - p \sum_{j=0}^{q} h(j), \tag{3}$$

where values $h(i_j)$ represent those samples in the impulse response which were multiplied by the positive pulse $1-p$ in the input noise. That are the samples in the impulse response whose corresponding bit in the binary representation

0-8186-8005-9/97 $10.00 © 1997 IEEE

of i is set to 1. Equally to the input noise the output signal is zero mean.

The third-order cumulant $C_{3,y}(\tau_1, \tau_2)$ of the output signal $y(n)$ for an MA system of order q, which was excited with non-Gaussian distributed white noise, was derived in [5] as

$$C_{3,y}(\tau_1, \tau_2) = \gamma_{3,w} \sum_{n=0}^{q} h(n)h(n+\tau_1)h(n+\tau_2). \quad (4)$$

In simulations, the third-order cumulant value $C_{3,y}(\tau_1, \tau_2)$ is estimated from N samples of the output signal $y(n)$ as follows:

$$\hat{C}_{3,y}(\tau_1, \tau_2) = \frac{1}{K} \sum_{n=0}^{K} y(n)y(n+\tau_1)y(n+\tau_2), \quad (5)$$

where $K = N - \max(\tau_1, \tau_2)$. Denote partial products in (5) with random variable $C_{3,y}^{\tau_1,\tau_2}$. For the case of the superimposed signals, each of them depends on the system response $h(n)$ and $q + \max(\tau_1, \tau_2) + 1$ subsequent samples of the Bernoulli distributed input noise. Together, $2^{q+\max(\tau_1,\tau_2)+1}$ different values could be obtained for one partial product.

Based on the structure of the superimposed signal $y(n)$ Eq. (2) and using the calculation scheme for the third-order cumulant, the following probability distribution function results for partial products $C_{3,y}^{\tau_1,\tau_2}$ from Eq. (5):

$$C_{3,y}^{\tau_1,\tau_2} \sim \begin{array}{c|c|c} X_0 & \cdot & X_i \\ \hline (1-p)^{q+\max(\tau_1,\tau_2)+1} & \cdot & p^{k_{1,i}}(1-p)^{k_{2,i}} \\ \hline \cdots & & \\ \hline \cdots & & X_{2^{q+\max(\tau_1,\tau_2)+1}} \\ \hline & & p^{q+\max(\tau_1,\tau_2)+1} \end{array} \quad (6)$$

The values of the random variable $C_{3,y}^{\tau_1,\tau_2}$ are indexed from X_0 to $X_{2^{q+\max(\tau_1,\tau_2)+1}}$. The relation between the values X_i of random variable $C_{3,y}^{\tau_1,\tau_2}$ and the values of random variable Y is the following:

$$\begin{aligned} X_i &= Y_{(i \bmod 2^{q+1})} Y_{((i \bmod 2^{q+1+\tau_1}) \div 2^{\tau_1})} \\ &\cdot Y_{((i \bmod 2^{q+1+\tau_2}) \div 2^{\tau_2})}, \end{aligned} \quad (7)$$

where the operator mod means the remainder of the integer division and the operator $\div$ stands for integer division.

For an arbitrary lag (τ_1, τ_2) in the third-order cumulant, N such partial products are summed up and averaged giving the following variance of $\hat{C}_{3,y}(\tau_1, \tau_2)$:

$$\begin{aligned} Var_N &= \Big(\frac{1}{N}\sum_{i=1}^{N} X_i - m\Big)^2 p(\sum_i x_i) \\ &= \Big(\frac{1}{N}\sum_{i=1}^{N}(X_i - m)\Big)^2 p(\sum_i x_i) \\ &= \Big(\frac{1}{N^2}\sum_{i=1}^{N}(X_i - m)^2 \\ &+ \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(X_i - m)(X_j - m)\Big)p(\sum_i x_i) \\ &= \frac{1}{N}\Big(\frac{1}{N}\sum_{i=1}^{N}(X_i - m)^2 p(\sum_i x_i)\Big) \\ &+ \frac{2}{N}\Big(\frac{1}{N}\Big(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} X_i X_j \\ &- m\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(X_i + X_j) \\ &+ \frac{N(N-1)}{2}m^2\Big)\Big)p(\sum_i x_i) \\ &= \frac{1}{N}Var_1 + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} X_i X_j p(\sum_i x_i) \\ &- \frac{(N-1)}{N}m^2, \end{aligned} \quad (8)$$

where Var_1 means:

$$Var_1 = \frac{1}{N}\sum_{i=1}^{N}(X_i - m)^2 p(\sum_i x_i). \quad (9)$$

The second part of the Eq. (8) is marked as Cov_1 and written:

$$\begin{aligned} Cov_1 &= \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} X_i X_j p(\sum_i x_i) - \frac{(N-1)}{N}m^2 \\ &= \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} X_i p(X_i) X_j p(X_{j-i}) \\ &- \frac{N-1}{N}m^2 \\ &= \frac{2}{N^2}\sum_{i=1}^{N-1} X_i p(X_i) \sum_{j=i+1}^{N} X_j p(X_{j-i}) \\ &- \frac{N-1}{N}m^2. \end{aligned} \quad (10)$$

The cumulant value $C_{3,y}(\tau_1, \tau_2)$ is obtained as a sum of arbitrary number of random variables $C_{3,y}^{\tau_1,\tau_2}$. But the values X_i in this random variable are dependent: when one partial product has the value of X_i with the probability $p(X_i)$, its successor has either the value of $X_{(i \div 2)+2^{q+\max(\tau_1,\tau_2)}}$ with the probability $p(X_i)p$ or the value of $X_{(i \div 2)}$ with the probability $p(X_i)(1-p)$:

$$\begin{aligned} X_i^{(n)} &\longrightarrow X_{(i \div 2)+2^r}^{(n+1)} \quad \text{or} \\ X_i^{(n)} &\longrightarrow X_{(i \div 2)}^{(n+1)}. \end{aligned} \quad (11)$$

system	order q	phase	impulse response
1	4	mixed	h1=[1 -2 3 1.5 -0.8]
2	4	minimum	h2=[1 1.3312 0.3272 0.2352 0.2568]
3	1	maximum	h3=[1 4]
4	1	maximum	h4=[1 -4]

Table 1. The impulse responses of the tested MA systems

In Eq. (11) the value $r = q + \max(\tau_1, \tau_2)$ was used. Let's mark $E(C_{3,y}^{\tau_1,\tau_2}) = m$ and compute the average of partial sums, when any two partial products are summed up:

$$\begin{aligned} m_2 &= \sum_{i=0}^{2^{r+1}-1} \Big\{ \Big[\frac{X_i + X_{(i\div 2)+2^r}}{2} \Big] p(X_i) p \\ &+ \Big[\frac{X_i + X_{(i\div 2)}}{2} \Big] p(X_i)(1-p) \Big\} \\ &= \frac{1}{2} \Big[\underbrace{\Big\{ \sum_{i=0}^{2^{r+1}-1} X_i p(X_i) \Big\}}_{m} \\ &+ \sum_{i=0}^{2^{r+1}-1} X_{(i\div 2)+2^r} p(X_i) p \\ &+ \sum_{i=0}^{2^{r+1}-1} X_{(i\div 2)} p(X_i)(1-p) \Big] \\ &= \frac{1}{2} \Big[m + \sum_{i=0}^{2^r-1} X_i \underbrace{[p(X_{2i}) + p(X_{2i+1})](1-p)}_{p(X_i)} \\ &+ \sum_{i=2^r}^{2^{r+1}-1} X_i \underbrace{[p(X_{2i-2^{r+1}}) + p(X_{2i-2^{r+1}+1})]p}_{p(X_i)} \Big] = \\ &= \frac{1}{2}(m+m) = m. \end{aligned} \tag{12}$$

Let's prove that the underbraced expressions in Eq. (12) equal $p(X_i)$. But first we have to prove the following:

$$p(X_{2i})p = p(X_{2i+1})(1-p). \tag{13}$$

Following the definition in (6), $p(X_{2i}) = (1-p)^{k_1} p^{r+1-k_1}$ and $p(X_{2i+1}) = (1-p)^{k_1-1} p^{r+2-k_1}$. The constant k_1 represents the number of zeros in $(r+1)$-digit long binary represention of the number $2i$. If the probability for $p(X_{2i})$ is multiplied by p and probability for $p(X_{2i+1})$ by $1-p$, the same expression results on the both sides of Eq. (13). We can write as well the relation:

$$p(X_i) = p(X_{2i}), \tag{14}$$

because the number of ones in the binary digit will remain the same if the number is multiplied by 2 (shift for one digit left). Considering Eqs. (13) and (14), we proved that the first underbrace in Eq. (12) equals to $p(X_i)$. Let's prove the same for the second underbrace. Because in the binary representation of the number $2i+1$ we have one 1 more as in the binary representation of the number $2i - 2^{r+1} + 1$, if $i \geq 2^r$, the following auxiliary result could be written:

$$p(X_{2i-2^{r+1}+1}) = \frac{1-p}{p} p(X_{2i+1}). \tag{15}$$

If $p(X_{2i+1})$ is expressed from Eq. (13) and put into Eq. (15), the value of the second underbrace in Eq. (12) is proven as well.

With derivation (12) it was shown that the average of partial sums, when summing up two partial products, equals the average of random variable $C_{3,y}^{\tau_1,\tau_2}$, which gives the possible values of the partial products calculating the cumulant value $C_{3,y}(\tau_1, \tau_2)$. Following the same procedure, it could be proven that this average remains always the same, summing up N partial products, and equals $m_N = m$.

But the main question is how the variance is changing? Our starting point is the variance of the random variable, $Var_1 = D(C_{3,y}^{\tau_1,\tau_2})$, which represents the values of partial products. How the variance Var_N is changing when, calculating the cumulant value $C_{3,y}(\tau_1, \tau_2)$, N partial products are summed up? First, the variance Var_1 could be written as follows:

$$Var_1 = \sum_{i=0}^{2^{r+1}-1} X_i^2 p(X_i) - m^2. \tag{16}$$

When $j - i > q + \max(\tau_1, \tau_2)$ in Eq. (10), the expression becomes constant and equal to m^2. When we deal with length N and the MA system of order q, we have $\frac{(N-q-\max(\tau_1,\tau_2))(N-q-\max(\tau_1,\tau_2)-1)}{2}$ sumations, which gives m^2. The part with the value m^2 is equal to:

$$\begin{aligned} &\Big[\frac{(N-q-\max(\tau_1,\tau_2))(N-q-\max(\tau_1,\tau_2)-1)}{N^2} \\ &\qquad - \frac{N(N-1)}{N^2} \Big] m^2 \\ &= \frac{[-2N+q+\max(\tau_1,\tau_2)+1][q+\max(\tau_1,\tau_2)]}{N^2} m^2. \end{aligned} \tag{17}$$

Other parts, when $j - i \leq q + \max(\tau_1, \tau_2)$, could be

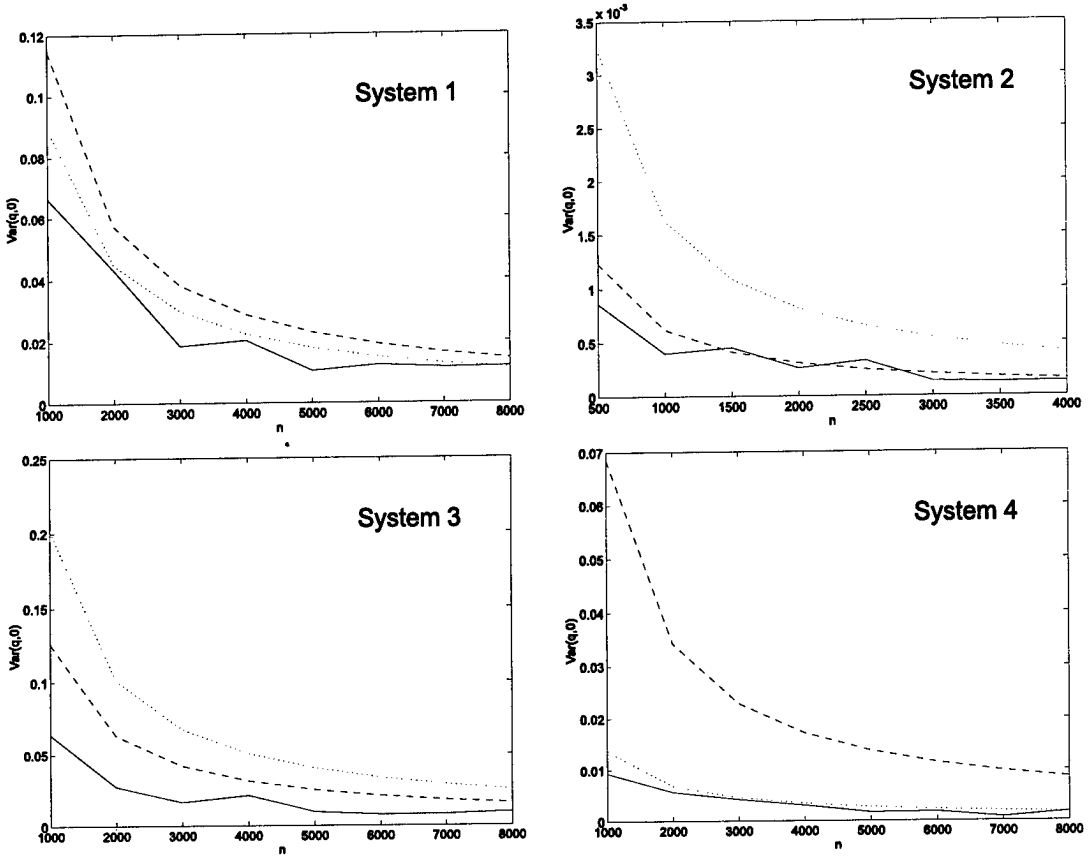


Figure 1. Variance of the cumulant value $C_{3,y}(q,0)$ depening on the length of the output signal $y(n)$ and the system used

calculated as:

$$R_j = \sum_{i=0}^{2^{r+1}-1} X_i p(X_i) \sum_{k=0}^{2^j-1} p^{k_{0,k}} (1-p)^{k_{1,k}} \cdot X_{(i \div 2^j)+rev(compl(k))\cdot 2^{2q-j}}, \quad (18)$$

where $rev(compl(k))$ stands for the reversed-bit-order value of the 1's complement of k, and $\div$ is the integer division.

At length N, we have $N-1$ terms at the distance $j-i=1$, $N-2$ terms at the distance $j-i=2$, and so on. At the distance $j-i=q$, we have $N-q$ terms, and at the distance $j-i=q+\max(\tau_1,\tau_2)$ only $N-q-\max(\tau_1,\tau_2)$ terms. The final expression for the variance of the cumulant value $C_{3,y}(\tau_1,\tau_2)$ using N samples long output signal y is:

$$Var_N = \frac{Var_1}{N} + \frac{2}{N^2}\sum_{i=1}^{r}(N-i)R_i + \frac{[-2N+q+\max(\tau_1,\tau_2)+1][q+\max(\tau_1,\tau_2)]}{N^2} m^2. \quad (19)$$

The variance of the sum of random variables denoting partial products was split into the variance Var_1 and covariance Cov_1 of the individual partial products. From Eq. (19) is evident that both parts decrease when longer output signals are used for the cumulant calculation. The contribution of Var_1 is significant, because it gives the answer how fast the error of the estimation will decrease. The covariance, Cov_1, which is the second part in Eq. (19), could lower the total variance, but less than for one class of magnitude. With the help of Eq. (19), it could be predicted how fast, or how accurate, the cumulants of the system will be computed with the given length of the system output signal.

3. Variance analysis and simulation results

In Fig. 1, four examples of the variance calculation are presented, first with a mixed-phase system, the second one with a minimum-phase system, and last two with a maximum-phase system. The dotted line depicts the result of the analytical model in Eq. (19), the dashed line the

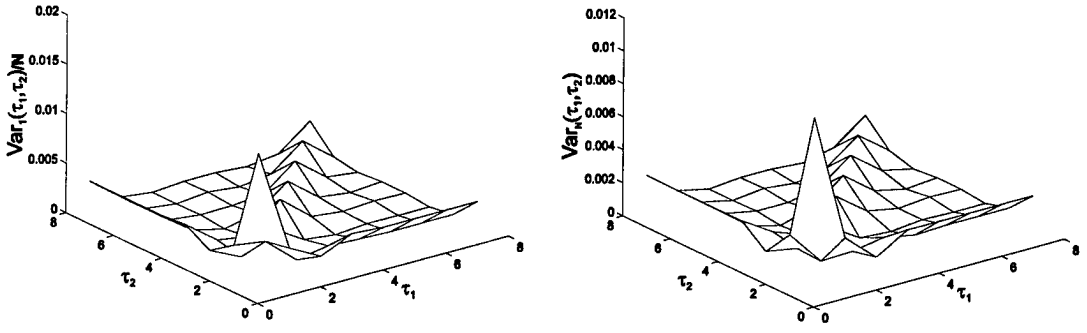


Figure 2. Analytical variances of the cumulant values $C_{3,y}(\tau_1, \tau_2)$ depending on lags τ_1 and τ_2 (system h=[1 0.8 -1.6 1.2]); left $Var_1(\tau_1, \tau_2)/N$, right $Var_N(\tau_1, \tau_2)$; $N = 1000$

approximation from Eq. (5), and the solid line the variance obtained in computer simulations with signals of the lengths indicated on x-axis.

In Table 1, the actual impulse responses with their orders q are gathered. In all four simulation tests 30 independent Monte Carlo runs were performed. It was noticed that the simulation variance was well estimated with expression (19). On the other hand, the simple estimation Var_1/n is not so exact, but in some examples where the sum of the samples in the impulse response is strongly positive, it happens that the simpler estimation Var_1/n is even more accurate as the complete expression (19).

Additionally, the variance structure inside the third-order cumulant was studied (Fig. 2). The difference between the approximation $Var_1(\tau_1, \tau_2)/N$ and the analytical model $Var_N(\tau_1, \tau_2)$ is shown. The system used has mixed phase. It is evident that the variances are close to zero when the lags τ_1 and τ_2 are bigger ($\tau_1 > q$, $\tau_2 > q$ and $|\tau_1 - \tau_2| > q$). It could be noticed as well that the variance is much larger for the cases $\tau_1 = \tau_2$, $\tau_1 = 0$ and $\tau_2 = 0$. This should be considered when new identification methods are designed.

4. Conclusions

Starting with the analytical variance model of the third-order cumulant $C(\tau_1, \tau_2)$, the variance of the identified system response can be obtained analytically as well, respecting the identification method applied. Such a variance model leads, on the one hand, to a confidence measure of the system identified at a certain N, on the other hand, it indicates the systems that generate too high variance to be identified in finite conditions.

Observing Eqs. (19) and (18), it is evident that computational complexity of the variance model developed is $\mathcal{O}(2^{q+max(\tau_1,\tau_2)+1})$ and it is not dependent on the signal length N used in the cumulant estimation. Therefore, this SISO variance model may also be implemented as a building block in the multichannel MISO and MIMO variance models with a reasonable complexity.

References

[1] J. M. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proc. IEEE*, 79:278-305, 1991.

[2] C. L. Nikias, M. R. Raghuveer. Bispectrum estimation: A digital signal processing framework. *Proc. IEEE*, 75:869-891, 1987.

[3] D. Zazula, D. Korže, A. Šoštarič, D. Korošec. Study of Methods for Decomposition of Superimposed Signals with Application to Electromyograms. *Neuroprosthetics from Basic Research to Clinical Applications (Eds. Pedotti et al.)*. Springer-Verlag Berlin, Heidelberg, 377-389, 1996.

[4] D. Zazula, D. Korže. Higher-Order Cumulants in System Identification and Signal Decomposition. *Visual Modules: Proceedings of the 19th ÖAGM and 1st SDRV Workshop*, Oldenbourg Wien, München, 275-284, 1995.

[5] G. B. Giannakis. *Signal processing using higher-order statistics*. PhD Thesis, Los Angeles: Dep. Elec. Eng., Univ. Southern California, 1987.

Minimum Entropy Approach for Multisensor Data Fusion

Yifeng Zhou
Telexis Corporation Canada
2427 Holly Lane, Ottawa, Ontario, Canada K1V 7P2

Henry Leung
Surface Radar Section
Defence Research Establishment Ottawa, Ottawa, Ontario, Canada K1A OK2

Abstract

In this paper, we present the minimum entropy fusion approach for multisensor data fusion in non-Gaussian environments. We represent the fused data in the form of the weighted sum of the multisensor outputs and use the varimax norm as the information measure. The optimum weights are obtained by maximizing the varimax norm of the fused data. The minimum entropy fusion solution only depends on the empirical distribution of the sensor data and makes no specific distribution assumptions about the sensor data. Numerical simulation results are provided to show the effectiveness of the proposed fusion approach.

1. Introduction

In recent years, the problem of multisensor fusion has received considerable attention [1][2]. They play an important role in the operation of intelligent machines and systems. The primary aim of sensor fusion is to gain more accurate information with reduced uncertainty by utilizing redundancy information from multiple sensors. The potential advantages of integrating multiple sensory information correspond to the notions of redundancy, complementarity, timeliness and cost of the information, respectively [1]. In the past decades, there have been developed many fusion techniques [3][4][5]. The minimum variance approach by Durrant-Whyte [3] is based on the uncertainty modeling of multisensor systems and uses the Bayesian inference theory. The weighted least squares method by Matthies and Shafer [4] is a variation of the minimum variance approach which takes the covariance information into consideration. The geometric approach proposed by Nakamura and Xu [5] is to associate geometric uncertainty with the covariance ellipsoid of the fused sensory information. The optimum fusion is obtained by minimizing the geometric volume of the ellipsoid. These techniques are developed under the Gaussian assumption and their performances usually deteriorate when the underlying signal and noise distribution deviate from the assumed Gaussian. It is known that the Gaussian assumption is for the mathematical convenience and is usually only justified in theory by the central limit theorem. In practical applications, however, it does not generally hold and is often violated.

In this paper, we present the minimum entropy fusion approach for multisensor data fusion in non-Gaussian environments. We consider the fusion in the scope of linear combination, that is, the fused information is represented as the weighted sum of the sensor outputs. We use the varimax norm as the information measure. The varimax norm has been defined by data analysts [6][7] in trying to find a simple representation of set of orthogonal vectors. It measures the simplicity of a signal and maximizing the varimax norm has the effects of simplifying the appearance or the entropy of a signal. The minimum entropy deconvolution (MED) [8] by Wiggins is one successful application of the varimax norm to the blind deconvolution problem. We propose the minimum entropy fusion solution which minimizes the entropy, or maximizing the varimax norm of the fused sensor data. The minimum entropy fusion solution depends only on the empirical distribution of the sensor data and makes no distribution assumptions about the sensor data. Numerical simulation results are provided to show the effectiveness of the proposed fusion approach.

0-8186-8005-9/97 $10.00 © 1997 IEEE

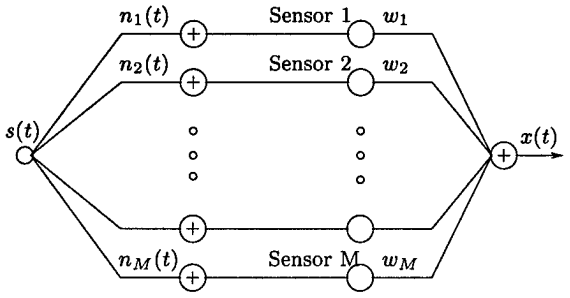


Figure 1. Multisensor data fusion model.

2. Mathematical Model for Multisensor Data Fusion

Consider a multisensor fusion system with M ($M \geq 2$) sensors, as shown in Figure 1. We assume that each sensor outputs are corrupted with additive noise with arbitrary distribution. The mth sensor measurement can be expressed as

$$x_m(t) = s(t) + n_m(t), \quad m = 1, 2, \ldots, M, \qquad (1)$$

where $s(t)$ denotes the signal and $n_m(t)$ represents the additive noise of the mth sensor. We assume that $n_m(t)$ is a identically independently distributed (i.i.d) process of zero mean. The objective of fusion is to combine the sensory information in a systematic manner to get a good consensus $s(t)$. We consider fusion in the scope of linear combination, that is, the fused information is given as a linear combination of the M sensor measurements

$$x(t) = \sum_{m=1}^{M} w_m x_m(t), \qquad (2)$$

where $\{w_m(t), m = 1, 2, \ldots, M\}$ denote the weighting coefficients. The objective of fusion is to find the optimum weights that minimize the uncertainty of the fused information.

3. Minimum Variance Fusion Approach

The most widely used uncertainty measure in fusion is the variance. The minimum variance fusion can been derived from different perspectives; they basically minimize the variance of the fused sensory information. Consider the mean of the fused information $x(t)$

$$E[x(t)] = \sum_{m=1}^{M} w_m E[x_m(t)] = \{\sum_{t=1}^{M} w_m\} E[s(t)], \qquad (3)$$

where E denotes the expectation. To make $x(t)$ an unbiased estimate of $s(t)$, we impose the constraint $\underline{a}^T \underline{w} = 1$ where $\underline{a} = [1, 1, \ldots, 1]^T$. The variance of $x(t)$ is given by

$$J_{MV} = E\{\underline{w}^T \underline{n}(t) \underline{n}^T(t) \underline{w}\} = \underline{w}^T R_n \underline{w}, \qquad (4)$$

where

$$\begin{aligned} \underline{w} &= [w_1, w_2, \ldots, w_M]^T \\ \underline{n}(t) &= [n_1(t), n_2, \ldots, n_M(t)]^T \end{aligned}$$

and R_n denotes the covariance of $\underline{n}(t)$. The minimum variance fusion solution can be solved by the following linearly constrained optimization problem

$$\underline{w}_{MV} = \arg\min_{\underline{w}} J_{MV}, \qquad (5)$$

subject to the constraint $\underline{a}^T \underline{w} = 1$. Using the method of Lagrange multipliers, the optimum solution $\underline{w}_{MV}$ can be obtained as

$$\underline{w}_{MV} = R_n^{-1} \underline{a} [\underline{a}^T R_n^{-1} \underline{a}]^{-1}. \qquad (6)$$

4. Varimax Norm and Minimum Entropy Fusion

The minimum variance fusion solution is based on the second-order statistics of the sensory information and is optimum only under the Gaussian assumption. Its performance usually deteriorates when the underlying Gaussian model is violated. It is known that Gaussian processes are characterized by its second-order statistics while non-Gaussian processes contain valuable information in their high order moments which can be exploited for improved performance. In this paper, we use the varimax norm as an information measure to study the problem of multisensor data fusion in non-Gaussian environments.

The varimax norm has been defined by data analysts [6][7] in trying to find a simple representation of set of orthogonal vectors. It measures the simplicity of a signal and maximizing the varimax norm has the effects of simplifying the appearance or the entropy of a signal. We usually regard the varimax norm as another form of representing the entropy. The varimax norm of the fused information $x(t)$ is defined as

$$V[x(t)] = \frac{C[x(t)]}{L^2[x(t)]} = \frac{\frac{1}{N}\sum_{t=1}^{N} \mid x(t) \mid^4}{\{\frac{1}{N}\sum_{t=1}^{N} \mid x(t) \mid^2\}^2}, \qquad (7)$$

where N is the number of sensor data. Note that $V(y)$ is related to the sampled Kurtosis of $\{y(t)\}$ which is defined as the all zero lag of the fourth-order cumulant normalized by the variance. The varimax norm only depends on the empirical distribution of sensor data $\{y(1), y(2), \ldots, y(t)\}$ viewed as a simple random sample. The varimax norm is also scale invariant. When using the varimax norm as the uncertainty measure for fusion, some constraints are needed to be imposed on the weighting coefficients to make the optimization solution unique. We propose the minimum entropy fusion solution as

$$\underline{w}_{ME} = \arg\max_{\underline{w}} V[x(t)], \tag{8}$$

subject to the constraint $\underline{a}^T\underline{w} = 1$. The optimization (8) is nonlinear and numerical techniques are needed.

5. Performance Analysis

In this section, we show some properties of the minimum entropy fusion solution. The theoretical results in [9] do provide some useful intuitive justifications for it.

Let $\mathcal{X}$ be a set of random variables with finite variances which is closed under linear combinations, i.e., $\sum_i a_i\mathcal{X}_i + c_i$ is in $\mathcal{X}$ when $\mathcal{X}_i$ are where a_i and c_i are constants. We give the definitions of equivalency and inequality for random variables as follows.

Definition For two random variables x and y, $x \doteq y$, if for some constants a and $c \neq 0$, $ax + c$ has the same probability distribution as y; and $x\cdot \geq y$, if for a set of constants a_i with $\sum_i a_i^2 < \infty$, $y \doteq \sum a_i x_i$, where x_i are independent samples of x. Notation "$\cdot >$" means "$\cdot \geq$ but not $\doteq$".

The variables x and y are regarded as equivalent when $x \doteq y$. Notation $\cdot \geq$ is a partial order on $\mathcal{X}$ and has the properties of transitivity and asymmetry, that is, if $x\cdot \geq y$ and $y\cdot \geq z$ then $x\cdot \geq z$; if $x\cdot \geq y$ and $y\cdot \geq x$, then $x \doteq y$. The equivalence $\doteq$ and the partial order $\cdot \geq$ also characterize the Gaussianity of random variables. The following lemma provides the order relations of random variables [9].

Lemma For x in $\mathcal{X}$ and z Gaussian,

$$x\cdot \geq \sum_i a_i\mathcal{X}_i\cdot \geq z, \tag{9}$$

and (9) is strict unless either (a) $\mathcal{X}$ is Gaussian or (b) $\mathcal{X}$ is not Gaussian, but the linear combination is trivial (no two a_i's are nonzero).

Relation (9) implies that the linear combinations of independent random variables are more Gaussian than any individual component of the combination. Thus, we usually interpret the partial order $x\cdot \geq y$ as y being more Gaussian than x.

The Kurtosis has been traditionally used in in statistics for measuring the Gaussianity of a random variable (the Kurtosis of a Gaussian process is zero). In [9], it has been shown that the varimax norm agrees with $\cdot \geq$ on $\mathcal{X}$, that is,

$$x\cdot \geq y \text{ implies } V(x) \geq V(y), \tag{10}$$

for every x and y which are filtered white samples from $\mathcal{X}$.

Assume that $s(t)$ and $\{n_m(t),\ m = 1, 2, \ldots, M\}$ are independent samples from a non-Gaussian $\mathcal{X}$. With the constraint $\underline{a}^T\underline{w} = 1$, the fused information can be written as

$$x(t) = s(t) + n_f(t), \tag{11}$$

where $n_f(t) = \sum_{m=1}^{M} w_m n_m(t)$ denotes the fused sensor noise information. Then, we have $x(t)\cdot \leq n_f(t)$ and

$$V[y(t)] \leq V[n_f(t)]. \tag{12}$$

As indicated by (12), when no optimization is performed over $\{w_m\}$, the fused information would have more uncertainty than the fused noise information. However, maximizing the varimax norm of the fused information has the effects of maximizing the varimax norm of the fused noise information, or in other words, minimizing the uncertainty of the fused noise information.

6. Numerical Simulations

We provide a numerical example to show the effectiveness of the proposed minimum entropy solution. More complete numerical studies are still under investigation. In the simulation, we use $M = 4$ and $N = 200$. Assume that the signal is Gaussian distributed and the sensor noise follows the one side exponential distribution. Figure 2 shows the mean square error of the fused sensory information versus noise variance. The comparison is made with the minimum variance fusion approach. It is evident that the minimum entropy fusion shows improved performance over the minimum variance fusion for all the noise variances used in the test.

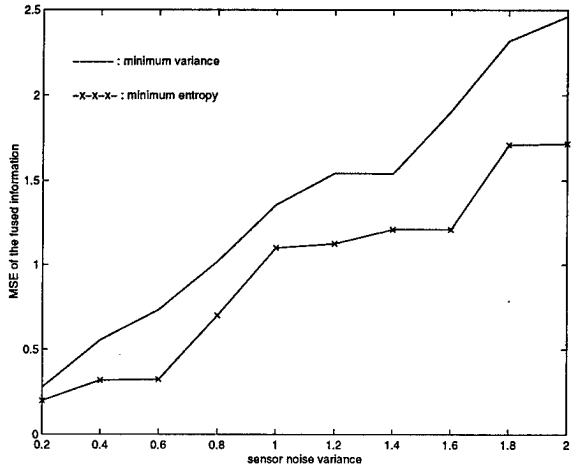


Figure 2. Fusion by the minimum entropy and the minimum variance approaches.

References

[1] R. C. Luo and M. G. Kay, "Data fusion and sensor integration : State-of-the-art 1990s", In M. A. Abidi and R. C. Gonzalez (ed.), *Data Fusion in Robotics and Machine Intelligence*, Academic Press, Inc., 1992

[2] J. K. Aggarwal, *ed.*, *Multisensor Fusion for Computer vision*, Springer-Verlag, Berlin, 1993

[3] H. F. Durrant-Whyte, "Integrating distributed sensor information, an application to a robot system coordinate", Proc. IEEE Int. Conf. Syst., Man, and Cyber., pp. 415-419, 1985

[4] L. Matthies and S. A. Shafer, "Error modeling in stereo navigation", IEEE J. Robotics Automation, vol. 3, pp. 239-248, 1987

[5] Y. Nakamura, Y. Xu, "Geometric fusion method for multi-sensor robotic systems", Proc. IEEE International Conference on Robotics and Automation, pp. 668-673, Scottsdale, Arizona, 1989

[6] H. F. Kaiser, "The varimax criterion for analytic rotation in factor analysis", Psychometrika, Vol. 23, pp. 187-200, 1958

[7] W. W. Cooley and P. R. Richards, *Multivariate analysis*, Wiley : New York, 1971

[8] R. A. Wiggins, "Minimum entropy deconvolution", Geoexploration, Vol. 16, pp. 21-35, 1978

[9] D. Donoho, "On minimum entropy deconvolution", in *Applied time series analysis II*, Academic Press, pp. 565-608, 1981

Blind separation of sources applied to convolutive mixtures in shallow water

M. Gaeta [1], F. Briolle [2], Ph. Esparcieux [1]

1: Vibria, 185 Espace Athena, 83190 Ollioules, France. Email: gaeta@vibria.francenet.fr
2: DCE / CTSN / DLSM, le Brusc, 83140, France.

Abstract :

In underwater acoustics, the signal received by sensors is a mixture of different elementary sources, filtered by the environment. In blind separation of sources, we can isolate each source from different mixtures of sources without any a priori information, except for assuming statistical independence of the different sources. Two French researchers, J. Herault and C. Jutten had earlier proposed a neuromimetic solution to the problem.

In our work, we use this solution to separate convolutive mixtures of simulated complex underwater signals in shallow water environment. To allowed the multipath identification a whitening step have to be introduced. We propose a local whitening procedure that does not impact the separated signal output and preserve the signal characteristics

This promising technique can be improved using non causal whitening filters more adapted to the target environment.

Keywords : shallow water propagation, source separation, multipath identification, whitening.

1. Introduction

The Blind Separation of Sources problem arises from different fields such as astronomy, astrophysics, underwater acoustics, medical applications. According to the domain, one can focus on an object (star, ship, electrocardiogram wave, etc.) that produces some signals (optical, electromagnetic, acoustics signals, etc..)

In underwater acoustic, the signal received by sensors is a mixture of different elementary sources, filtered by the environment. For example, these different sources could be the signatures of vessels, the noise made by the environment, self-noise, etc..

The optimal use of the sensors will be to separate all these sources. This issue is called blind separation of sources because we do not have any a priori knowledge about these sources : the only hypothesis made is that they are statistically independent.

The sources path trough a transmission field are received on a set of sensors.

The transmission field is supposed to be isotropic deterministic, to present a stationary property and furthermore to be linear. So the received signals can be considered as linear mixtures of the initial sources.

Using only the knowledge of the received signals, one can try to characterise the initial sources.

The sources are supposed to be in a shallow water environment The signals are simulated and the mixtures received by the sensors are convolutive.

The separated signals will be used for several purposes such as detection, noise reduction, classification etc.. For this applications it is important to preserve as much as possible the initial sources characteristics. The proposed algorithm is design in this way.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Propagation characteristics

In the shallow water environment, the signal received by the hydrophone is the result of the direct sound added with the echoes of the source.

Depending on the position of the source and the receiver, the propagation effect can be modelled as a linear time-invariant filter H_{ij}. The impulse response $h_{ij}(t)$ of the filter can be estimated with GAMARAY, a ray-based model developed by E. K. Westwood [5].

Most of the taps of the impulse response are equal to zero ; the non null taps are representative of the direct sound and the echoes. Typical impulse responses of the filters, for two different positions source-receiver, are shown on the figure 1 :

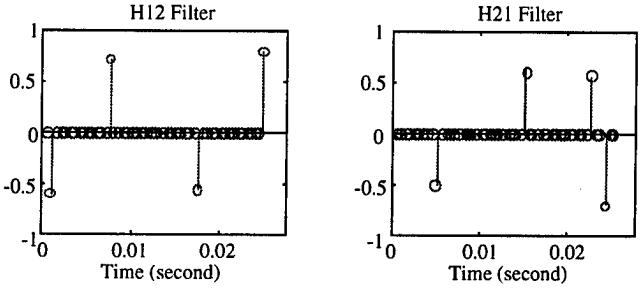


Figure 1 : Propagation effect in shallow water environment. Impulse response of two different filters H_{12} and H_{21} estimated by GAMARAY.

The received signal $r_j(t)$ on the hydrophone j is the contribution of each source signal , filtered by the environment :

$$r_j(t) = \Sigma\, h_{ij}(t) * s_i(t) \quad i = \{1, ..., M\} \quad (1)$$

For two hydrophones R_1 and R_2 and two sources N_1 and N_2, the received signal can be modelled as convolutive mixtures as shown on the figure 2.

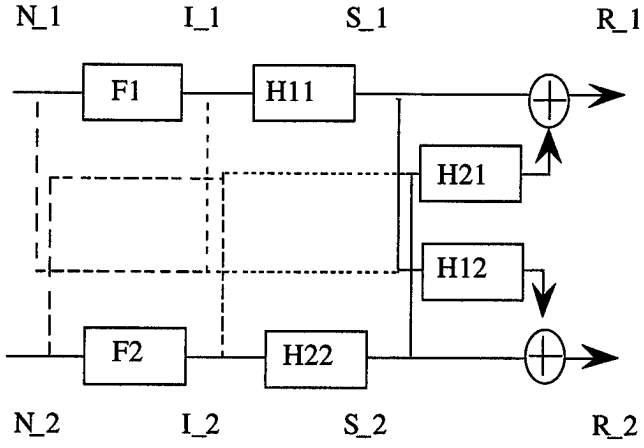


Figure 2 : Received signals generation

Starting from white noise processes N_1 and N_2, the initial sources I_1 and I_2 are obtained by linear filtering respectively F1 and F2.

Then they propagate trough the media. During the propagation to each sensor R_1 and R_2 a transformation is performed respectively by the filters H_{11} and H_{22} and the coloured sources S_1 and S_2 are obtained.

It is clear that the filters H_{11} and H_{22} could not be estimated using independence properties of the sources because all the signals of a defined line (N_i, I_i, S_i) are independent to the signals taken on an other line (N_i, I_i, S_i). Further hypothesis have to be stated to define the point to be reached.

The signals received on R_1 and R_2 are convolutive mixtures of S_1 and S_2. The filters H_{12} and H_{21} are a model of the propagation in shallow water.

The filters H_{12} and H_{21} could be estimated by restoring the independence properties of the output of a system build as a mirror image of the propagation.

3. Source separation principle

The algorithm proposed in [3] can be seen as an invert propagation procedure.

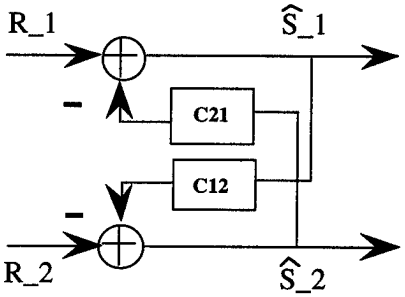


Figure 3 : Estimation of the sources S1 and S2 by invert propagation procedure.

The problem is now an identification one using a neural net.

The estimation of the separating filters C_{12} and C_{21} is based on the minimisation of the cross information of the network output. The cross information can be quantify using various statistics. Quadratic, higher order or non linear solutions exist and the choice depends on the application type.

From our experience, the results are similar in terms of performance and tracking capability. The different memory parameters involved in the various algorithms have to be optimised depending on the local or statistical properties of the sources.

An important point to underline is that the separating filters C_{12} and C_{21} would be equal to the mixture filters H_{12} and H_{21} if the sources S_1 and S_2 are equivalent to a white noise [4].

As an example let us consider a mixture of two white noises. The first one is a realisation of a uniform distribution and the second is obtained by taking the Arctang of a gaussian process.

The filters H_{11} and H_{22} are equal to 1 and the impulse responses of filters H_{12} and H_{21} present a sparse corresponding to a shallow water propagation distribution (Figure 1).

The independence criteria is the fourth order cross moment of the estimated signal. 10 000 points are used for this simulations.

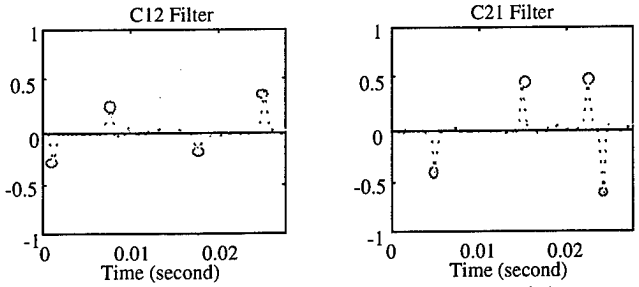


Figure 4 : Estimation of the impulse response of the two separating filters C_{12} and C_{21}

The rough results are presented on figure 4, using a dash line. One can see that a post processing is possible by identifying the main propagation path : the impulse responses of the filters C_{12} and C_{21} are close to the typical impulse responses used for this simulation (cf. figure 1).

Because of the particular shape of the filters H_{12} and H_{21}, only maxima values have to be taken into account to identify the impulse responses of the separating filters C_{12} and C_{21}.

This reconditioning will remove the variance estimation of the null taps of the filters. In some cases, those variations can represent the major part of the separating filters and lead to poor separation results.

4. Local whitening Algorithm

The underwater sources are not white processes so a whitening step have to be done to come back to the previous case and identify the main path components.

If a pre-whitening is performed the separated sources have to be coloured to be used for identification purposes. To avoid doing this back and forth procedure, we propose to do the whitening step just where it is needed. This is to say for the separation filters C_{12} and C_{21} updates.

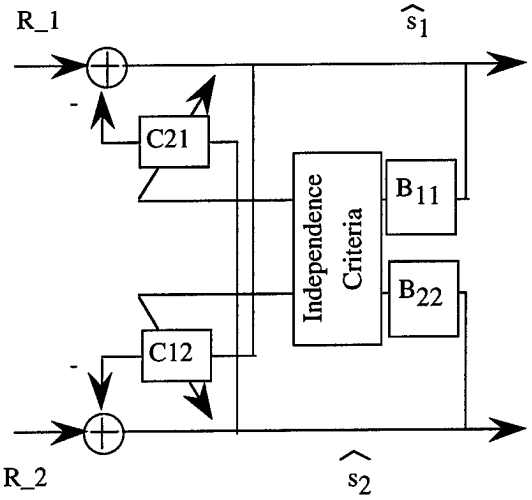


Figure 5 : Estimation of the sources S1 and S2 by invert propagation procedure and local whitening.

The output equation is still the same :

(2)

But the updating independence criteria are not estimated directly on the output signal but on the whitened versions of them.

The algorithm can be summarised with the following lines of pseudocode.

```
For all samples
  Begin
  Estimated the whitening filter taps B1
     and B2 using the output signal.
  Estimate the independence criteria
     (Indep) using the whitened output signal.
  Update all the coefficients using :
     Cij (t+1) = Cij (t) + μ Indep
  Estimated the output signals
  End
```

The following example presents the results of underwater signal mixed using the same filters than previous example.

The first source is a submarine signal and the second a ship noise.

The sampling frequency is 2000 Hz, and the signal duration is equal to 10 seconds.

Using the local whitening algorithm, the estimated impulse responses of the filters C_{12} and C_{21} are presented on figure 6.

The whitening filters B_{11} and B_{22} are 50 taps FIR filters.

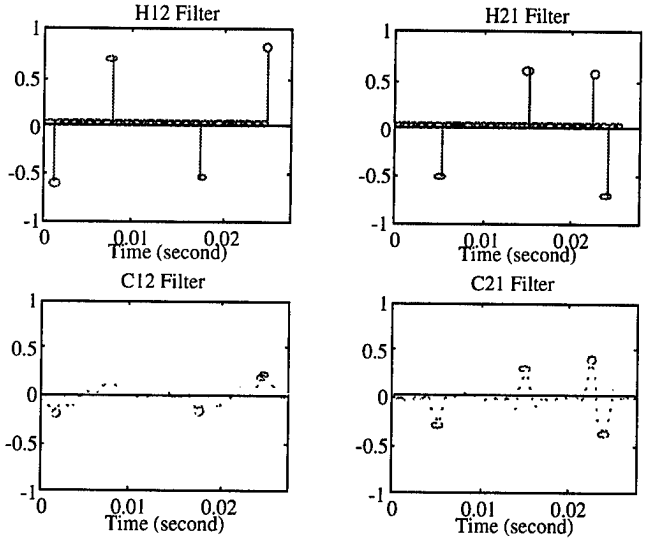


Figure 6: Estimation of the two separating filters C_{12} and C_{21} using local whitening algorithm.

The results are promising - the estimated filters C_{12} and C_{21} are close to H_{12} and H_{21} - but can be improved. So the identification of the main propagation path could be done for coloured sources.

The source separation filters are only related to the propagation and not to the colour of the sources.

By doing a local whitening, the estimation of the filters are only related to the independence properties of the sources and not to their respective colours.

5. Conclusion

We present here the first results of sources separation technique applied to underwater acoustic in shallow water.

We propose a local whitening version of the neuromimetic approach proposed in [4]. The purpose is to avoid unnecessary prewhitening / unwhitening operations. Moreover by doing a local whitening procedure for the separating filters estimation, the impulse response of the propagation channel can be identified.

This offer the opportunity to constraint the estimation using the knowledge of the media, and improve the results.

Some improvement of the whitening procedure have to be done and come closer to the results obtained using white noise sources.

Two ways for improvement can be investigate : the use of IIR filter and the estimation of non minimum phase filters. This non minimum phase filters will be mandatory as long as the filters will no longer be equal to 1. Various Blind deconvolution algorithm using non linear criteria can then be used [1], [2].

The first results confirm the possible use of sources separation technique for underwater applications.

References :

[1] M. Gaeta, M. Nikias, "A new method for multiple source detection and identification from array data using cumulants and its application to shock waves propagation", IEEE Signal processing workshop on higher order statistics, June 7-9, 1993 Lake Tahoe, California USA.

[2] W. Gray, "Variable Norm Deconvolution", PhD dissertation, Department of geophysics, Stanford University, 1979.

[3] C. Jutten, J. Herrault, "Blind sources separation : an adaptive algorithm based on neuromimetic architecture", Signal Processing, 24:1-10, 1991.

[4] H.L. Nguyen Thi and C. Jutten, "Blind sources separation for convolutive mixtures", Signal Processing, 45:209-229, 1995.

[5] E.K. Westwood, C.T. Tindel, " Shallow water time series simulation using ray theory ", J. Acoust Soc. Am., 81: 1752-1761, 1987.

Separation of Sinusoidal Sources

C. Servière, V. Capdevielle, J-L. Lacoume

CEPHAG-ENSIEG BP 46 38402 Saint-Martin d'Hères cedex FRANCE

ABSTRACT

A particular source separation problem is addressed in this paper. We mainly focus on the separation of convolutive mixtures of rotating machine noises when the rotation speeds are close. Three specific points are developed. In the first point, we study the feasibility of the separation of periodic signals, regarding the hypothesis of random and non gaussian sources. We also discuss about the hypothesis of independence between the sources, in function of the rotation speeds. In the second point, we analyze the performances of the source separation for close rotation speeds. They are linked to a partial correlation between the machine noises. Then, we propose a new method for very close rotation speeds, which takes into account this partial dependence between the sources.

1. Introduction

A particular source separation problem is addressed in this paper. We mainly focus on the separation of convolutive mixtures of rotating machine noises. The influence of the rotation speeds on the performances of source separation methods are detailed. These methods consist in recovering the signals emitted by p sources $\underline{s}(t)$ from M observed linear mixtures $\underline{r}(t)$ of these sources. For that purpose, well-known methods classically use the mere uncorrelation of the emitted source signals but need extra information to achieve their objective. In a general context of source separation, this a priori knowledge is replaced with higher order statistic information. The only three assumptions are the non-gaussianity of the source signals, their mutual independence and the linearity and stationarity of the propagation.

Since ten years, many solutions have been proposed which test different measurements of the statistical independence. They are based on the use of fourth-order moments or cumulants [1], [2], [3], [4], [5], [6] nonlinear functions which also make higher order statistics appear [7] [8] [9], or contrast functions [10]. Recently, a deflation procedure has been developed in [11]. The linear filters (which characterize the propagation from the sources to the sensors) can be estimated using adaptive or nonadaptive algorithms minimizing or looking for zeros of different independence criteria. Some procedures are also based on the maximum likelihood principle [12] [13]. An original geometry-based procedure has also been proposed in the case of n-valued signals [14].

Most of these works reconstruct the source signals from instantaneous mixtures : the observations at any time t only depend on the sources at any time t. In the case of a convolutive mixture of narrow-band signals, the model can be reduced to an similar instantaneous mixture with complex coefficients of the mixture. Consequently, similar methods may be used to separate these sources .

In §3, we study the feasibility of the separation of periodic signals, regarding the hypothesis of random and non gaussian sources when the rotation speeds are different enough. In §4, we study the feasibility of the source separation methods, regarding the hypothesis of independent sources. More precisely, these methods test the cancellation of second and fourth-order cross-cumulants. Then we analyze the performances of the source separation versus the difference between the rotation speeds. Indeed, this quantity measures the dependence between the sources. In §5, we propose a new method for very close (or identical) rotation speeds, which takes into account this partial dependence between the sources.

2. Modelization of the problem

In a general blind source separation problem, the observed M-dimensional data vector $\underline{r}(t)$ can be represented in frequency-domain by an instantaneous complex mixture for each frequency bin n, which leads to the following model :

(1) $\quad \underline{Ri}(n) = \underline{\underline{A}}(n)\ \underline{Si}(n) + \underline{Vi}(n) \quad n=0, ..., N-1$

where $\underline{Ri}(n)$ is the N-point Discrete Fourier Transform (DFT) of the ith data block of the observation $\underline{r}(t)$. $\underline{Si}(n)$ represents the DFT of the ith data block of the p-dimensional data vector of the sources $\underline{s}(t)$. $\underline{\underline{A}}(n)$ is a matrix (M.p) which characterizes the linear propagation from sources to sensors and $\underline{Vi}(n)$ represents an additive M-dimensional gaussian noise. The problem consists first in identifying the matrix $\underline{\underline{A}}(n)$. After a singular value decomposition, the mixing matrix $\underline{\underline{A}}(n)$ is expressed as the product of three matrices.

(2) $\quad \underline{\underline{A}}(n) = \underline{\underline{F}}(n)\ \underline{\underline{D}}(n)\ \underline{\underline{\Pi}}(n)$

0-8186-8005-9/97 $10.00 © 1997 IEEE

where $\underline{\underline{F}}(n)$ and $\underline{\underline{\Pi}}(n)$ are two (M.M) and (p.p) unitary matrices. $\underline{\underline{D}}(n)$ is an (M.p) diagonal matrix.

The two matrices $\underline{\underline{F}}(n)$ and $\underline{\underline{D}}(n)$ are identified thanks to second-order statistic criteria. They respectively contain the eigenvectors and the eigenvalues of the spectral matrix of the observation $\underline{R}i(n)$.

After projection of the observation vector $\underline{R}i(n)$ in the signal subspace (which is spanned by the eigenvectors associated to the dominant eigenvalues) and normalization, the components of the p-dimensional vector, noted $\underline{X}i(n)$, are uncorrelated and normalized. They are relied to the components of the normalized source vector, noted $\underline{s}'(t)$, by:

(3) $\underline{X}i(n) = \underline{\underline{\Pi}}(n)\ \underline{S}'i(n)$

where $\underline{S}'i(n)$ is the DFT of $\underline{s}'(t)$.

$\underline{\underline{\Pi}}(n)$ can be expressed as a product of Givens rotations and estimated thanks to fourth-order criteria, by testing different measurements of the statistical independence of the sources [1] [2] [3], ..., [10] , as presented in §1.

3. Separation of periodic signals

This part of the paper is devoted to the feasibility of the separation of periodic signals, after Discrete Fourier Transform. As it has been shown before in §2 in the case of random signals, the source separation methods lay on the additional information provided by fourth-order statistics. This information only exists under the hypothesis of independent and non gaussian sources. Whatever the tested criterion, the variance of the estimate of matrix $\underline{\underline{\Pi}}(n)$ is inversely proportional to the kurtosis of the sources. In a similar way, in frequency domain, we study the distance of the DFT to gaussianity thanks to the spectral kurtosis. It is defined as a section of the general trispectrum of the normalized sources.

Let K(Si(n)) be the kurtosis of Si(n), defined by :

$$(4) \qquad K(Si(n)) = \frac{cum(Si(n), Si(n)^*, Si(n), Si(n)^*)}{cum(Si(n), Si(n)^*)^2}$$

where cum represents the second and fourth-order cumulants and * the complex conjugate.

In practice, the source separation methods are applied on a record of the random process $\underline{X}i(n)$, under the hypothesis of ergodicity. Consequently, the algorithms lay on the non cancellation of the estimate K of the spectral kurtosis : the statistical quantities are estimated by time averaging over L data blocks.

In the case of periodic (thus deterministic) signals, the statistical kurtosis has non sense. However, as previously, the estimate K can always be computed by averaging over L data blocks on each component of $\underline{S}i(n)$.

Let s(t) be a periodic signal of period T. For example, let s(t) be a sinusoid of deterministic frequency and phase, noted ω and ϕ :

$$(5) \qquad s(t) = A \sin(2\pi\omega t + \phi)$$

Let us compute the DFT of s(t) on the ith data block of Si(n), for n close to ω : (6)

$$Si(n) = \sum_{k=0}^{N-1} s(k+i)\exp(-2\pi jkn/N)$$

$$Si(n) = \frac{A}{2j} e^{j(\phi + 2\pi\omega i + \pi(N-1)(\omega - \frac{n}{N})} . F(\omega)$$

$$\text{with } F(\omega) = \frac{\sin(\pi(\omega - \frac{n}{N})N)}{\sin(\pi(\omega - \frac{n}{N}))}$$

After some computations, K is equal to [15] [16]:

$$(7) \quad K = -1 - \left| \frac{1}{L} \sum_{i=1}^{L} \exp(j4\pi\omega i) \right|^2 = -1 - \frac{G(2\omega)}{L^2}$$

$$\text{with } G(2\omega) = \left| \frac{\sin(2\pi\omega L)}{\sin(2\pi\omega)} \right|^2$$

If ω is superior to (3/4L), then $G(2\omega)/L^2$ is inferior to the limit value $(2/3\pi)$ for large L. The limit of the ratio $G(2\omega)/L^2$ (for ω=3/4L) is almost constant with L. The maxima of this function are decreasing with $(L \sin(2\pi\omega))^{-2}$. Consequently the estimate of the kurtosis K(Si(n)) tends towards (-1) for a fixed value of ω, if L is large enough and n close to ω. Periodic signals can be decomposed in Fourier series and K tends towards -1 for the harmonic frequency bins. As a result, the non cancellation of K provides additional equations that allow the estimate of matrix $\underline{\underline{\Pi}}(n)$ for the harmonic frequencies n. This point theoretically proves the feasibility of the source separation of rotating machine noises.

4. Separation of sinusoids with close frequencies

4.1 Discussion on the hypothesis of independent sources

In this part, we analyze the performances of the source separation. They depend on the difference between rotation speeds. For that, suppose s1(t) and s2(t), two periodic signals. They are decomposed in Fourier series as combinations of sinusoids of deterministic frequencies ω1 and (ω1+$\Delta\omega$), amplitudes A1 and A2 and phases ϕ1 and ϕ2. The term $\Delta\omega$ is proportional to the difference between the machine rotation speeds. The source separation rests on the non cancellation of the estimate K of the spectral kurtosis of each source. This point is always verified since K tends towards -1 when L tends towards infinity, whatever the frequencies (ω1 or (ω1+$\Delta\omega$)). It is also based on the independence of the sources, for random signals. In the case of deterministic signals, it means that the estimates of the second and fourth-order normalized cross-cumulants tend towards 0

when L tends towards infinity. We suppose that N is large enough, so that there is no interaction due to the DFT between several harmonics.
They are respectively equal to : (8)

$$\hat{C}(S1,S2^*) = \frac{1}{L}\sum_{i=1}^{L} S1i(n)S2i(n)^* = \frac{A1A2}{4L}.$$

$$e^{j(\phi 1 - \phi 2 - \pi(N-1)\Delta\omega)} \sum_{i=1}^{L} \exp(j2\pi\Delta\omega i) F(\omega 1) F(\omega 1 + \Delta\omega)$$

(9) The symmetrical fourth-order cross-cumulant is equal to

$$\hat{C}(S1,S1^*,S2,S2^*) = \frac{1}{L}\sum_{i=1}^{L} S1i(n)S1i(n)^* S2i(n)S2i(n)^*$$

$$-|\hat{C}(S1,S2^*)|^2 - |\frac{1}{L}\sum_{i=1}^{L} S1i(n)S2i(n)|^2 - \hat{C}(S1,S1^*)\hat{C}(S2,S2^*)$$

$$= (\frac{A1A2}{4})^2 F(\omega 1)^2 F(\omega 1 + \Delta\omega)^2 [- \left|\frac{1}{L}\sum_{i=1}^{L} \exp(j2\pi\Delta\omega i)\right|^2$$

$$- \left|\frac{1}{L}\sum_{i=1}^{L} \exp(j2\pi(2\omega 1 + \Delta\omega)i)\right|^2]$$

From their expressions, we conclude that these quantities decrease with L, after normalization, according to $|\frac{1}{L}\sum_{i=1}^{L} \exp(j2\pi\Delta\omega i)|$ and $-\left|\frac{1}{L}\sum_{i=1}^{L} \exp(j2\pi\Delta\omega i)\right|^2$. As detailed in §3, if $\Delta\omega$ remains superior to (3/2L), these quantities are inferior to $|L \sin(2\pi\Delta\omega)|^{-1}$ and $|L \sin(2\pi\Delta\omega)|^{-2}$ (or $|L\, 2\pi\Delta\omega|^{-1}$ and $|L\, 2\pi\Delta\omega|^{-2}$ for weak values of $\Delta\omega$). These quantities tend towards 0 if $\Delta\omega$ is not strictly equal to zero. If $\Delta\omega$ is equal to zero, then the second and fourth-order normalized cross-cumulants are respectively equal to 1 and (-1).
As a conclusion, $|G(\omega)| \in [0,1]$ in function of L and ω.

4.2 Computation of the estimated sources, using source separation methods

Suppose two mixtures of the two sources s1(t) and s2(t). The estimation of these sources $\hat{S}1(n)$ and $\hat{S}2(n)$ can be modelized as :

(10) $$\hat{S}1(n) = (\hat{C}(\hat{S}1,\hat{S}1^*))^{-1/2}[S1(n) + \alpha(n)S2(n)]$$

$$\hat{S}2(n) = (\hat{C}(\hat{S}2,\hat{S}2^*))^{-1/2}[\beta(n)S1(n) + S2(n)]$$

The cancellation of the second-order and (for example) the symmetrical fourth-order normalized cross-cumulants leads to the two following equations : (11)

$$\beta(n)^* A1^2 F(\omega 1)^2 + \alpha(n)\beta(n)^* A1A2F(\omega 1)F(\omega 1+\Delta\omega)G(-\Delta\omega)$$

$$+A1A2F(\omega 1)F(\omega 1+\Delta\omega)G(\Delta\omega) + \alpha(n)(A2F(\omega 1+\Delta\omega))^2 = 0$$

(12)

$$(A1A2F(\omega 1)F(\omega 2))^2 (\alpha(n)^*\beta(n)^* + \alpha(n)\beta(n))(1 - |G(\Delta\omega)|^2) +$$

$$\alpha(n)^*\beta(n)(G(2\Delta\omega) - G(\Delta\omega)^2) - |\hat{C}(\hat{S}1,\hat{S}2^*)|^2$$

$$+\alpha(n)\beta(n)(G(-2\Delta\omega) - G(-\Delta\omega)^2))) = 0$$

If $|G(\Delta\omega)|=0$, then (11) and (12) can be simplified into :

(13) $$\beta(n)^* A1^2 F(\omega 1)^2 + \alpha(n)(A2F(\omega 1+\Delta\omega))^2 = 0$$

(14)

$$|\beta(n)|^2 C(S1,S1^*,S1,S1^*) + |\alpha(n)|^2 C(S2,S2^*,S2,S2^*) = 0$$

As the fourth-order cumulants of sources S1(n) and S2(n) are negative, (14) assures the separation of the sources. The unique solution is $\alpha(n)=\beta(n)=0$. Equation (13) is also verified. It is the classical case of independent sources, in the sense that the estimations of the second-order and fourth-order cross-cumulants of the sources are zero.
If $0 < |G(\Delta\omega)| < 1$, then (12) assures that $\alpha(n)\beta(n)=0$ (if $C(\hat{S}1,\hat{S}2^*)=0$). However, equation (11) is not verified for $\alpha(n)=0$ and $\beta(n)=0$, since the sources are correlated. This solution does not lead to the separation of the sources. In that case, the estimated sources will be equal to : (for example if $\beta(n)=0$)

(15) $$\hat{S}1(n) = (\hat{C}(\hat{S}1,\hat{S}1^*))^{-1/2}[S1(n) + \alpha(n)S2(n)]$$

$$\hat{S}2(n) = (\hat{C}(\hat{S}2,\hat{S}2^*))^{-1/2}[S2(n)]$$

with : $$\alpha = \frac{-A1.F(\omega 1)G(\Delta\omega)}{A2.F(\omega 1+\Delta\omega)}$$

We verify that $\alpha(n)$ tends toward zero with $G(\Delta\omega)$.

4.3 Computation of the rejection level

We characterize the performances of blind source separation methods in terms of rejection levels. After separation, it remains in each estimated sources a residual contribution of the others sources. The rejection level is defined as the power spectrum of the contribution of the residual source at frequency $\omega 1$ in the estimate of the source at frequency $\omega 1+\Delta\omega$. We see that it is equal in that case to :

$$|\alpha(n)|^2 (\hat{C}(\hat{S}1,\hat{S}1^*))^{-1}$$

It is decreasing with $\alpha(n)$, and therefore with $G(\Delta\omega)$.
If $|G(\Delta\omega)|$ tends towards 1 (or $\Delta\omega$ towards 0), then α tends towards (-A1/A2). It proves that the rejection level is bounded for totally correlated sources.
We see in figure 1 the rejection level in function of $(L.\Delta\omega)$ on simulation results. We remark that it is correctly decreasing with $(L.\Delta\omega)$, since the estimation of the source cross-cumulants are more and more close to zero.

5. Case of sinusoids with identical frequencies

5.1 Number of sensors equal to the source number

In that case, if $\Delta\omega$ (or L) is not large enough, the sources appear as correlated ones and consequently not independent. In that case, the source signals can be modelized more precisely as the addition of several sinusoids of different frequencies and non gaussian random noises. In frequency domain, the kth source $Si^k(n)$ contains one or several sinusoidal components (for example one) and can be written as : (16)

$$Si^k(n) = \frac{Ak}{2j} e^{j(\phi k + 2\pi\omega i + \pi(N-1)(\omega - \frac{n}{N})} .F(\omega)$$

$$+Ni^k(n)$$

Random processes $N^k(n)$ are assumed to verify the previous hypotheses of independent signals. We propose a new method of source separation which takes into account the partial dependence between the sources. This partial dependence is due to the phase term :

$$(17) \quad e^{j(2\pi\omega i + \pi(N-1)(\omega - \frac{n}{N})} .F(\omega).$$

The first step of the source separation method at second-order is suppressed : the projection and the normalization of the observation vector $\underline{Ri}(n)$ in the signal subspace cannot provide here orthogonal mixtures of the sources. The mixture is then not modelized with Givens rotations, unless the sinusoids of same frequencies are considered as a supplementary source in the bin n. It means constraints on the source number and on the sensor number.

Suppose the first model:

(18) $\underline{Ri}(n) = \underline{\underline{A}}(n)\ \underline{Si}(n) + \underline{Vi}(n) \quad n=0, ..., N-1$

with :

$$Si^k(n) = \frac{Ak}{2j} e^{j(\phi k + 2\pi\omega i + \pi(N-1)(\omega - \frac{n}{N})} .F(\omega)$$

$$+Ni^k(n)$$

$\underline{\underline{A}}(n)$ is correctly identified when the estimates of each source $\hat{Si}^k(n)$ contain only one random independent process $\hat{Ni}^k(n)$ and a sinusoidal component. In that case, the second-order cross-cumulants provide no information, as their value depends on the signal to noise ratio in each source. The fourth-order cumulants $\hat{C}(Sk, Sk^*, Sj, Sj^*)$ are not equal to zero (for $k \neq j$) as usual in source separation hypotheses. They are equal to the fourth-order cumulants of the sinusoidal components (9), since the estimates $\hat{Ni}^k(n)$ and $\hat{Ni}^j(n)$ must be independent (for $k \neq j$). As the sinusoidal components cannot have been normalized at second-order, $\hat{C}(Sk, Sk^*, Sj, Sj^*)$ depends on their amplitudes $(Ak)^2(Aj)^2$.

$\hat{C}(Sk, Sk^*, Sj, Sj^*)$ can be normalized, in order to obtain the kurtosis of a sinusoid, which is equal to -1. For that, we have to eliminate the influence of the noises $\hat{Ni}^k(n)$ and $\hat{Ni}^j(n)$. The estimate of the fourth-order cross-cumulants is divided with the squared modulus of the second-order cross-cumulant (and not the cumulant):

$$(19) \quad C = \frac{\hat{C}(Sk(n), Sk(n)^*, Sj(n), Sj(n)^*)}{|\hat{C}(Sk(n), Sj(n)^*)|^2}$$

If the estimates of each source $\hat{Si}^k(n)$ contain only one random independent process $\hat{Ni}^k(n)$ and a sinusoidal component, then C is equal to -1. It is the only constant quantity that characterizes the sources. Inversely, we prove that the estimate of $\underline{\underline{A}}(n)$, such that the estimated sources verify C=-1, achieves the separation source. Consequently, there is no spurious solutions.

Indeed: (20)

$$\hat{C}(Sk(n), Sk(n)^*, Sj(n), Sj(n)^*) = -|E(Sk(n), Sj(n)^*)|^2 +$$

$$E(Sk(n), Sk(n)^*, Sj(n), Sj(n)^*) - E(|Sk(n)|^2)E(|Sj(n)|^2)$$

Using equation (20), (C=-1) is equal to (21) for each couple of sources: (21)

$$E(Sk(n), Sk(n)^*, Sj(n), Sj(n)^*) = E(|Sk(n)|^2)E(|Sj(n)|^2)$$

Equation (21) is realized if and only if the sources are independent or with constant modulus, or are the addition of independent noises and constant modulus sources. The Discrete Fourier Transform of periodic signals verifies this condition of constant modulus source at the harmonic frequency bins.

$$(22) \quad |Si^k(n)| = \frac{Ak}{2}.|F(\omega)|$$

$Si^k(n)$ does not depend on the time index i. Consequently, it is a constant modulus source.

5.2 Sensor number superior to the source number

In that case, the first step of the source separation method is unchanged : the projection of the observation vector $\underline{Ri}(n)$ in the signal subspace provides uncorrelated signals if the sinusoids of same frequencies are considered as a supplementary source in the bin n. As previously, the signals, after projection and normalization, are linked to the sources by the matrix $\underline{\underline{\Pi}}(n)$. This matrix can be expressed as a product of Givens rotations. When the sensor number is superior to the source number, the matrix $\underline{\underline{\Pi}}(n)$ is correctly identified when the estimates of each source are independent. It means that they contain

only one random independent process $\hat{N}i^k(n)$ or one sinusoidal component. In that case, the fourth-order cumulants $\hat{C}(Sk,Sk^*,Sj,Sj^*)$ for $k \neq j$, are equal to zero (as the usual hypotheses of source separation methods).

However the complete estimation of the source $\hat{S}i^k(n)$ must contain one independent noise and one sinusoidal component, in order to separate the vibrations issued from two different machines. Besides, the estimates of the sources are considered as normalized sources, which is not realistic. The power of each component of the sources is then estimated with the help of second-order cross-cumulants between the sensors and the normalized sources [15].

6. Conclusion

A particular source separation problem is addressed in this paper. We mainly focus on the separation of convolutive mixtures of rotating machine noises when the rotation speeds are close. Three specific points are developed. In the first point, we study the feasibility of the separation of periodic signals, regarding the hypothesis of random and non gaussian sources. We also discuss about the hypothesis of independence between the sources, in function of the rotation speeds. In the second point, we analyze the performances of the source separation for close rotation speeds. They are linked to a partial correlation between the machine noises. Then, we propose a new method for very close rotation speeds, which takes into account this partial dependence between the sources.

BIBLIOGRAPHY

[1] P.Comon, "Independent Component Analysis, a new concept?, *Signal Processing*, vol36, pp287-314, 1994

[2] D.Yellin, E.Weinstein, "Criteria for multichannel signal separation", *IEEE Tr. SP*, vol 42, n°8, 1994, pp2158-2168

[3] J.K.Tugnait "On Blind Separation of Convolutive Mixtures of Independent Linear Signals", *IEEE Sig. Proc. Workshop on SSAP*, June 24-26 1996, Corfu, pp 312-315

[4] J-F. Cardoso "Iterative techniques for blind source separation using only fourth-order cumulants", in J. Vandewalle, R. Boite, M. Moonen, A. Oosterlinck Eds, *Signal Processing VI, Theory and Applications, Proceeding of EUSIPCO-92*, Brussels, August 1992, Elsevier, Amsterdam, 1992, pp. 739-742

[5] J.L. Lacoume, P.Ruiz, "Separation of independent sources from correlated inputs", *IEEE Trans. on Signal Processing*, vol 40, n°12, December 1992, pp 3074-3078

[6] J-F. Cardoso, S. Bose, B. Friedlander "Output cumulant matching for source separation", *Proc. Int. Sig. Proc. Workshop on Higher Order Statistics*, Girona, June 1995, pp. 44-48

[7] H.L. Nguyen Thi, Ch. Jutten "Blind source separation for convolutive mixtures", *Signal Processing*, vol 45, 1995, pp 209-229

[8] Ch. Jutten, J. Herault "Blind separation of sources, Part I : An adaptive algorithm based on neuromimetic architecture", *Signal Processing*, vol 24, n°1, July 1991, pp. 1-10

[9] P. Comon, Ch. Jutten, J. Herault "Blind separation of sources, Part II : Problems statement, *Signal Processing*, vol 24, n°1, July 1991, pp. 11-20

[10] E. Moreau, O. Macchi "High Order Contrasts for Self-Adaptive Source Separation", Int. Journal of Adaptive Control and Signal Processing, vol 10, n°1, pp 19-46, January 1996

[11] N. Delfosse, Ph. Loubaton "Adaptive blind separation of independent sources: A deflation approach", *Signal Processing*, vol 45, n°1, July 1995, pp 59-83

[12] D.T. Pham, Ph. Garat, C. Jutten "Separation of a mixture of independent sources through a maximum likelihood approach", in J. Vandewalle, R. Boite, M. Moonen, A. Oosterlinck Eds, *Signal Processing VI, Theory and Applications, Proceeding of EUSIPCO-92*, Brussels, August 1992, Elsevier, Amsterdam, 1992, pp. 771-774

[13] M. Gaeta, J-L. Lacoume "Source separation versus hypothesis", *Proc. Int. Sig. Proc. Workshop on Higher Order Statistics*, Chamrousse, France, July 1991, pp. 269-272

[14] C.G. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, J. Ortega "Separation of sources : a geometry-based procedure for reconstruction of n-valued signals", *Signal Processing*, vol 46, 1995

[15] V.Capdevielle, "Séparation de sources large bande à l'aide de moments d'ordre supérieur", *Ph.D. Thesis*, INPGrenoble, 1995

[16] V. Capdevielle, C. Servière, J-L. Lacoume, "Blind separation of wide-band sources in the frequency domain, *Proc. ICASSP95*, pp 2080-2083, 1995

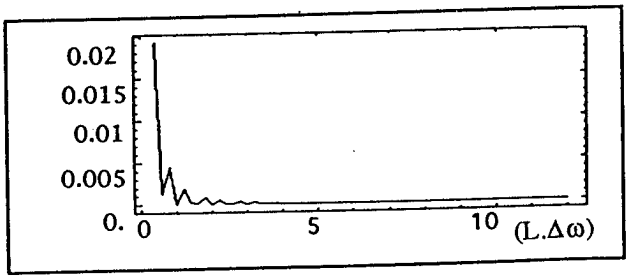


Figure 1 : Rejection level in function of (L.Δω)

A Linear Feedforward Neural Network with Lateral Feedback Connections for Blind Source Separation

Seungjin CHOI
Lab for Artificial Brain Systems
Frontier Research Program, RIKEN
2-1 Hirosawa, Wako-shi
Saitama 351-01, Japan
schoi@lion.riken.go.jp

Andrzej CICHOCKI
Lab for Artificial Brain Systems
Frontier Research Program, RIKEN
2-1 Hirosawa, Wako-shi
Saitama 351-01, Japan
cia@hare.riken.go.jp

Abstract

We presents a new necessary and sufficient condition for the blind separation of sources having non-zero kurtosis, from their linear mixtures. It is shown here that a new blind separation criterion based on both odd ($f(y) = y^3$) and even ($f(y) = y^2$) functions, presents desirable solutions, provided that all source signals have negative kurtosis (sub-Gaussian) or have positive kurtosis (super-Gaussian). Based on this new separation criterion, a linear feedforward network with lateral feedback connections is constructed. Both theoretical and computer simulation results are presented.

1. Introduction

Blind source separation is a fundamental problem encountered in many applications such as array signal processing, sonar, digital communications, some biomedical applications. The goal of blind source separation is to recover the independent source signals from their linear mixtures without the knowledge of mixing matrix. Specially when the propagation media is slowly changing (the mixing matrix is slowly changing), an adaptive system for blind source separation is necessary.

An adaptive blind source separation was first introduced by Jutten and Herault [11]. Later it was further developed by others [12, 7, 4, 1, 2, 8, 5]. Most of existing methods mentioned above are based on nonlinear odd functions (for example, $f(y) = y^3$ or $f(y) = tanh(y)$) because they assume that the probability distributions of all source signals are symmetric. Recently Choi *et al* [6] have shown that a new learning algorithm based on even nonlinearity (e.g. a quadratic function, $f(y) = y^2$) enables the separation of source signals having non-zero skewness (the 3rd-order cumulant) without spurious equilibria. We extend this result to the separation of source signals having the same sign of non-zero kurtosis (the 4th-order cumulant). A new necessary and sufficient condition for blind source separation is presented. It is shown here that the source signals can be separated by a linear transformation, if and only if all the 2nd- and 4th-order *cross*-cumulants of the output of the network are zero, provided that the source signals are statistically independent and each of them has the same sign of non-zero kurtosis. Based on this criterion, a linear feedforward network with lateral feedback connections is constructed with associated adaptive learning algorithms. The proposed learning algorithms are based on both cubic and quadratic functions to force all 2nd- and 4th-order cross-cumulants of the output of the network to vanish to zero.

2 A New Blind Source Separation Criterion

Consider the case where the observation vector $\mathbf{x}(t) \in \mathbb{R}^n$ and the source vector $\mathbf{s}(t) \in \mathbb{R}^n$ are related by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the mixing matrix. The problem of blind sources separation is to recover the source signals $\mathbf{s}(t)$ from the observation vector $\mathbf{x}(t)$ without the knowledge of the mixing matrix $\mathbf{A}$. Let $\mathbf{y}(t)$ be the output of the network, i.e., $\mathbf{y}(t) = \mathbf{W}(t)\mathbf{x}(t) = \mathbf{W}(t)\mathbf{A}\mathbf{s}(t)$. It is desired to update $\mathbf{W}(t)$ such that the global system $\mathbf{G}(t) = \mathbf{W}(t)\mathbf{A}$ converges to a matrix,

$$\mathbf{G} = \mathbf{P}\mathbf{\Lambda} \tag{2}$$

as $t \to \infty$, for some permutation matrix $\mathbf{P}$ and nonsingular diagonal matrix $\mathbf{\Lambda}$.

Throughout this paper, the following assumptions are made:

0-8186-8005-9/97 $10.00 © 1997 IEEE

A1: $\mathbf{A} \in \mathbb{R}^{n \times n}$ is nonsingular.
A2: At each time t, the components of $\mathbf{s}(t)$ are statistically independent.
A3: Each component of $\mathbf{s}(t)$ is a zero mean ergodic stationary process with a non-zero variance.
A4: Each component of $\mathbf{s}(t)$ has a non-zero kurtosis, i.e.,

$$\begin{aligned} cum_4\{s_i(t)\} &= E\{s_i^4(t)\} - 3E^2\{s_i^2(t)\}, \\ &\neq 0, \quad \text{for } i = 1, \cdots, n, \end{aligned} \tag{3}$$

where E denotes the expectation operator.

Without loss of generality, we can assume that the variance of each source signal is unity, i.e., $\mathbf{R}_{ss} = E\{\mathbf{s}(t)\mathbf{s}^T(t)\} = \mathbf{I}$, where $\mathbf{I}$ denotes identity matrix. Let the vector sequence $\mathbf{y}(t)$ be a linear mixture of the source vector sequence $\mathbf{s}(t)$,

$$\mathbf{y}(t) = \mathbf{G}\mathbf{s}(t). \tag{4}$$

Let us define two different 4th-order cumulant matrices of $\mathbf{y}(t)$ as follows:

$$\begin{aligned} [\mathbf{C}_{13y}]_{ij} &= cum_{13}\{y_i(t), y_j(t)\} \\ &= cum\{y_i(t), y_j(t), y_j(t), y_j(t)\}, \end{aligned} \tag{5}$$

$$\begin{aligned} [\mathbf{C}_{22y}]_{ij} &= cum_{22}\{y_i(t), y_j(t)\} \\ &= cum\{y_i(t), y_i(t), y_j(t), y_j(t)\}, \end{aligned} \tag{6}$$

where $[\cdot]_{ij}$ denotes the (i,j)th component of the matrix. The covariance matrix $\mathbf{R}_y = E\{\mathbf{y}(t)\mathbf{y}^T(t)\}$ is decomposed as

$$\mathbf{R}_y = \mathbf{G}\mathbf{G}^T. \tag{7}$$

For the following decomposition, we introduce Hadamard product (element-wise product). Let $\circ$ denotes the Hadamard product. For example, the (i,j)th element of $\mathbf{G} \circ \mathbf{G}$ is g_{ij}^2.

Property 1 *The 4th-order cumulant matrices* $\mathbf{C}_{13y}$ *and* $\mathbf{C}_{22y}$ *have the following decompositions:*

$$\mathbf{C}_{13y} = \mathbf{G}\mathbf{K}_s[\mathbf{G} \circ \mathbf{G} \circ \mathbf{G}]^T, \tag{8}$$

$$\mathbf{C}_{22y} = [\mathbf{G} \circ \mathbf{G}]\mathbf{K}_s[\mathbf{G} \circ \mathbf{G}]^T, \tag{9}$$

where $\mathbf{K}_s$ *is the nonsingular diagonal matrix whose ith diagonal element is* $cum_4\{s_i(t)\}$.

Proof: It can be proved by using multi-linearity of cumulant. Consider the (i,j)th element of $\mathbf{C}_{13y}$. Time index t is omitted throughout the proof. ($y_i = y_i(t)$, $s_i = s_i(t)$)

$$\begin{aligned} & cum_{13}\{y_i, y_j\} \\ &= cum\{\sum_{l_1} g_{il_1}s_{l_1}, \sum_{l_2} g_{jl_2}s_{l_2}, \sum_{l_3} g_{jl_3}s_{l_3}, \sum_{l_4} g_{jl_4}s_{l_4}\} \\ &= \sum_{l_1}g_{il_1}\sum_{l_2}g_{jl_2}\sum_{l_3}g_{jl_3}\sum_{l_4}g_{jl_4}cum\{s_{l_1}, s_{l_2}, s_{l_3}, s_{l_4}\} \\ &= \sum_k g_{ik}g_{jk}^3 cum_4\{s_k\} \end{aligned}$$

Similarly, the decomposition (9) can be proved.

Theorem 1 *Let the source signals* $\mathbf{s}(t)$ *satisfy the assumptions given in A1 through A4. Suppose that the kurtosis of all source signals* $\mathbf{s}(t)$ *are either positive or negative. Then* $\mathbf{G}$ *has decomposition (2), i.e.,*

$$\mathbf{G} = \mathbf{P}\mathbf{\Lambda}, \tag{10}$$

if and only if the following conditions are satisfied:

$$\mathbf{G}\mathbf{G}^T = \mathbf{\Lambda_1}, \tag{11}$$

$$\mathbf{G}\mathbf{K}_s[\mathbf{G} \circ \mathbf{G} \circ \mathbf{G}]^T = \mathbf{\Lambda_2}, \tag{12}$$

$$[\mathbf{G} \circ \mathbf{G}]\mathbf{K}_s[\mathbf{G} \circ \mathbf{G}]^T = \mathbf{\Lambda_3}, \tag{13}$$

where $\mathbf{\Lambda_1}$, $\mathbf{\Lambda_2}$, *and* $\mathbf{\Lambda_3}$ *are nonsingular diagonal matrices.*

Proof: For the sake of simplicity, let us assume that $\mathbf{\Lambda_1} = \mathbf{I}$. Then, $\mathbf{G}$ is an orthogonal matrix. Let $\mathbf{\Lambda_2} = diag\{\alpha_1, \cdots, \alpha_n\}$ and $\mathbf{K}_s = diag\{\kappa_1, \cdots, \kappa_n\}$. From (12), the (i,j)th element of the matrix $\mathbf{G}$, g_{ij} should satisfy $g_{ij} = 0$ or $g_{ij}^2 = \frac{\alpha_i}{\kappa_j}$. Using this fact, the (i,j)th element $(i \neq j)$ of LHS in (13) is given by

$$\alpha_i\alpha_j \sum_{l \in \mathcal{K}} \frac{1}{\kappa_l}, \tag{14}$$

where $\mathcal{K}$ is a set for the collection of all possible cases where $\mathbf{G} \neq \mathbf{P}\,\mathring{\mathbf{I}}$ for some permutation matrix $\mathbf{P}$ and the diagonal matrix $\mathring{\mathbf{I}}$ whose diagonal elements are either +1 or -1. Note that $\alpha_i \neq 0$ for $\forall i$ and $\sum_{l \in \mathcal{K}} \frac{1}{\kappa_l} \neq 0$. Thus, $\mathcal{K}$ is empty set. Therefore, $\mathbf{G} = \mathbf{P}\,\mathring{\mathbf{I}}$ □

Note that most of the existing algorithms satisfy the first two conditions (11), (12) so that they could converge to spurious equilibria. For instance, it can be easily shown that the following $\mathbf{G}$ satisfies (11) and (12), but it is not a solution to blind source separation:

$$\mathbf{G} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}. \tag{15}$$

3. Implementation

We present two different learning algorithms whose equilibria satisfy the conditions given in (11), (12), and (13). The network as shown in Figure 1, consists of a feedforward network $\mathbf{W}(t)$ and a feedback network $\mathbf{U}(t)$. The feedback network $\mathbf{U}(t)$ can be a lower triangular matrix with zero diagonal elements (this feedback connections are referred to be lateral feedback connections) or can be a full matrix (a fully recurrent network).

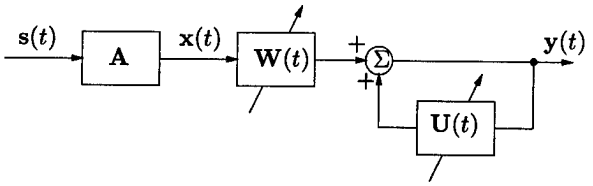


Figure 1. Our approach to blind source separation

3.1. The network with lateral feedback connections

We consider the case where $\mathbf{U}(t)$ is a lower triangular matrix with zero diagonal elements as shown in Figure 2. The output of the network $y_i(t)$, $i = 1, \cdots, n$ is described by

$$y_i(t) = \sum_{j=1}^{n} w_{ij}(t)x_j(t) + \sum_{j<i} u_{ij}(t)y_j(t), \qquad (16)$$

where $w_{ij}(t)$ is the synaptic weight between the ith output $y_i(t)$ and the jth input $x_j(t)$ and $u_{ij}(t)$ is the lateral connection between $y_i(t)$ and $y_j(t)$. Or in matrix form,

$$\mathbf{y}(t) = \mathbf{W}(t)\mathbf{x}(t) + \mathbf{U}\mathbf{y}(t), \qquad (17)$$

where $\mathbf{W}(t) = [w_{ij}(t)]_{n\times n}$ is a full matrix and $\mathbf{U} = [u_{ij}(t)]_{n\times n}$ is a lower triangular matrix with $u_{ij}(t) = 0$ for $i \leq j$.

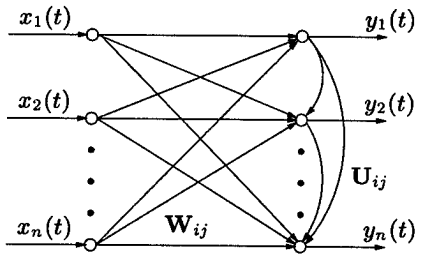


Figure 2. The structure of the feedforward network with lateral feedback connections

For such a neural network, we have developed the following adaptive learning algorithm:

$$\frac{d\mathbf{W}(t)}{dt} = \eta_w(t)\{\alpha\mathbf{I} - \alpha\mathbf{y}(t)\mathbf{y}^T(t) - sgn(\kappa_s)\Lambda(t) + sgn(\kappa_s)[\mathbf{y}(t)\circ\mathbf{y}(t)\circ\mathbf{y}(t)]\mathbf{y}^T(t)\}\mathbf{W}(t), \qquad (18)$$

$$\frac{du_{ij}(t)}{dt} = \eta_u(t)\{1 - y_i^2(t)y_j^2(t)\}, \quad \text{for } i > j, \qquad (19)$$

where $\Lambda(t)$ is a diagonal matrix whose ith diagonal element is $y_i^4(t)$. For sub-Gaussian source signals, $sgn(\kappa_s) = -1$, and for super-Gaussian source signals, $sgn(\kappa_s) = +1$. When the convergence of learning algorithm (18) and (19) is achieved, we have

$$E\{y_i(t)y_j(t)\} = 0, \qquad (20)$$

$$E\{y_i(t)y_j^3(t)\} = 0, \qquad (21)$$

$$E\{y_i^2(t)y_j^2(t)\} = 1, \qquad (22)$$

for $i \neq j$, and

$$E\{y_i^2(t)\} = 1, \quad \text{for } i = 1, \cdots, n. \qquad (23)$$

Note that

$$cum_{22}\{y_i(t), y_j(t)\} = E\{y_i^2(t)y_j^2(t)\} - E\{y_i(t)y_j(t)\} - 2E\{y_i^2(t)\}E\{y_j^2(t)\}. \qquad (24)$$

It can be easily seen that when these conditions (20), (21), (22), and (23) are satisfied, all 2nd- and 4th-order cross-cumulants of $\mathbf{y}(t)$ become zero. By Theorem 1, these equilibria are desirable solutions.

3.2. The network with full feedback connections

We consider the case where the feedback connection matrix $\mathbf{U}(t)$ is a full matrix (a fully recurrent network including the self-inhibitions connections from each output node back to itself). The output of the network $\mathbf{y}(t)$ is still given by

$$\mathbf{y}(t) = \mathbf{W}(t)\mathbf{x}(t) + \mathbf{U}(t)\mathbf{y}(t). \qquad (25)$$

The feedforward connections $\mathbf{W}(t)$ is trained to force all 4th-order cross-cumulants of the output $\mathbf{y}(t)$ to vanish to zero. The decorrelation learning algorithm with the feedback connections $\mathbf{U}(t)$ was motivated from the natural gradient-based algorithm [1]. We have developed the following learning algorithm

$$\frac{d\mathbf{W}(t)}{dt} = \eta_w(t)\{\mathbf{\Gamma}(t) - [\mathbf{y}(t)\circ\mathbf{y}(t)][\mathbf{y}(t)\circ\mathbf{y}(t)]^T - sgn(\kappa_s)\mathbf{y}(t)[\mathbf{y}(t)\circ\mathbf{y}(t)\circ\mathbf{y}(t)]^T + sgn(\kappa_s)[\mathbf{y}(t)\circ\mathbf{y}(t)\circ\mathbf{y}(t)]\mathbf{y}^T(t)\}\mathbf{W}(t), \qquad (26)$$

$$\frac{d\mathbf{U}(t)}{dt} = \eta_u(t)\{\mathbf{I} - \mathbf{U}(t)\}\{\mathbf{I} - \mathbf{y}(t)\mathbf{y}^T(t)\}, \qquad (27)$$

where the ith diagonal elements of $\mathbf{\Gamma}(t)$ is $\mathbf{y}_i^4(t)$ and all off-diagonal elements of $\mathbf{\Gamma}(t)$ are unity. It can be shown that stable equilibria of (26) and (27) satisfy the conditions (11), (12), and (13).

4. Discussion

We will discuss the extension of Theorem 1 to the case where sub-Gaussian (negative kurtosis) and super-Gaussian (positive kurtosis) source signals are mixed. Even in this case, Theorem 1 still holds if any possible sum of the inverse of kurtosis of more than two source signals, is not zero, i.e., $\sum_{l \in \mathcal{F}} \frac{1}{\kappa_l} \neq 0$, where $\mathcal{F}$ is a set of all combination of indices greater than or equal to two, out of $(1, 2, \cdots, n)$. This condition guarantees that all diagonal elements of $\mathbf{\Lambda_3}$ are non-zero. However, in situation where $\mathbf{x}(t)$ contains the mixtures of both sub-Gaussian and super-Gaussian source signals, there is a possibility that the sum of the inverse of the kurtosis might be zero. Moreover, the learning algorithms become unstable. Thus additional techniques are required [3].

The learning algorithm (18) and (19) associated with the first proposed network is working well for sub-Gaussian sources and super-Gaussian sources (not for mixtures of sub- and super-Gaussian). For the stability of learning algorithm, we introduce the function $sgn(\kappa_s)$. At this moment, we do not have theoretical result for stability analysis.

The neural network architecture shown in Figure 1 and 2 can be used for extraction of source signals one by one. Maximization or minimization of the normalized kurtosis with deflation approach is found in [9, 10]. We are developing the learning algorithm for sequential extraction of source in the hierarchical network as shown in 2.

5. Computer Simulations

The computer simulations are conducted to evaluate the performance of the proposed algorithm (18), (19) and (26), (27). The global system $\mathbf{G} = (\mathbf{I} + \mathbf{U})^{-1}\mathbf{WA}$ is evaluated to check the efficiency of the algorithm. The global system $\mathbf{G}$ approach the generalized permutation when the learning algorithm converges to desirable solutions.

5.1. Computer Simulation 1

Three i.i.d. binary sources drawn from uniform distribution (negative kurtosis) are used. The mixing matrix $\mathbf{A}$ is chosen randomly as

$$\mathbf{A} = \begin{bmatrix} 0.1549 & 0.1405 & 0.3916 \\ 0.5258 & 0.2041 & 0.9370 \\ 0.2047 & 0.5108 & 0.4310 \end{bmatrix} \qquad (28)$$

All elements of $\mathbf{G}(t) = (\mathbf{I} + \mathbf{U}(t))^{-1}\mathbf{W}(t)\mathbf{A}$ are plotted as shown in Figure 3 and 4.

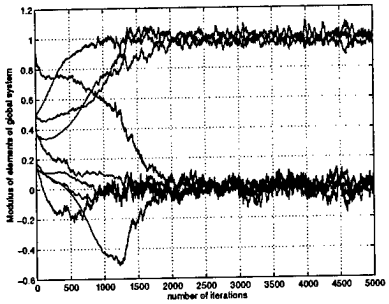


Figure 3. The modulus of each element of the global system $\mathbf{G}(t)$ when the first proposed network with the learning algorithm (18), (19). The learning rates were set $\eta_w(t) = .003$, $\eta_u(t) = .0007$.

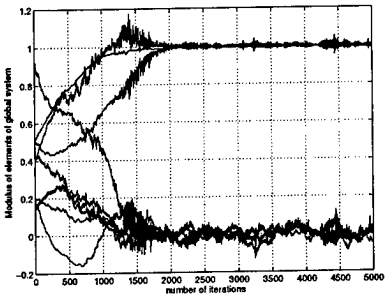


Figure 4. The modulus of each element of the global system $\mathbf{G}(t)$ when the second proposed network with the learning algorithm (26), (27). The learning rates were set $\eta_w(t) = .003$, $\eta_u(t) = .009$.

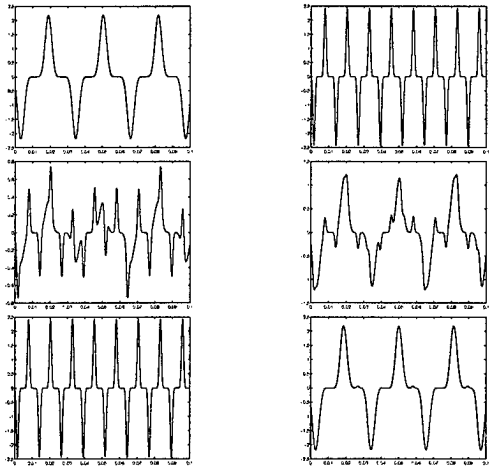


Figure 5. Sampling rate was 10kHz. Each plot shows the signal for the duration of .1 second. The original source signals are shown in the first row. Two sensor outputs are shown in the second row. The recovered signals are shown in the third row by using the learning algorithm (18) and (19). The similar result can be obtained by using (26) and (27).

5.2. Computer Simulation 2

Two independent source signals having positive kurtosis were used to generate the observation data $\mathbf{x}(t)$ by using randomly generated mixing matrix $\mathbf{A}$. Two different source signals are given by

$$\begin{aligned} s_1(t) &= sin^7(\omega_1 t), \\ s_2(t) &= sin^9(\omega_2 t). \end{aligned} \tag{29}$$

Figure 5 shows original source signals $\mathbf{s}(t)$, the sensor outputs $\mathbf{x}(t)$, and the recovered signals $\mathbf{s}(t)$.

6. Conclusion

In this paper, we presented a new necessary and sufficient condition for blind source separation. It was shown that if all 2nd- and 4th-order cross-cumulants of the output $\mathbf{y}(t)$ become zero, then the source signals can be recovered. It was proved by algebraic properties of 2nd- and 4th-order cumulants for the case of n sources. Based on this criterion, we constructed a linear feedforward network followed by a feedback network. Two different learning algorithms were presented and confirmed by the computer simulations.

References

[1] S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation," *Advances in Neural Information Processing Systems*, Vol. 8, 1995.

[2] S. Amari, A. Cichocki, and H. H. Yang, "Recurrent neural networks for blind separation of sources," in *Proc. NOLTA*, Vol. 1, pp. 37-42, 1995.

[3] S. Amari, T.-P. Chen, and A. Cichocki, "Stability analysis of learning algorithms for blind source separation," *Neural Networks*, 1997. (in print)

[4] A. Bell and T. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," *Neural Computation* 7, pp. 1129-1159, 1995.

[5] J. -F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *IEEE Trans. on Signal Processing*, pp. 3017-3030, December 1996.

[6] S. Choi and R. Liu, "A learning algorithm for the blind separation of non-zero skewness source signals with no spurious equilibria," in *Proc. ESANN* April 1997. (in print)

[7] A. Cichocki, R. Unbehauen, L. Moszczynski, and E. Rummert, "A new on-line adaptive learning algorithm for blind separation of source signals," in *Int. Symp. on Artificial Neural Networks*, pp. 406-411 1994.

[8] A. Cichocki, S. Amari, M. Adachi, and W. Kasprzak, "Self-adaptive neural networks for blind separation of sources'," in *Proc. ISCAS*, Vol.2, pp. 157-160, 1996.

[9] A. Cichocki, S. Amari and R. Thawonmas, "Blind signal extraction using self adaptive non-linear Hebbian learning rule," in *Proc. NOLTA*, pp. 377-380, October 1996.

[10] A. Cichocki, R. Thawonmas and S. Amari "Sequential blind signal extraction in order specified by stochastic properties", *Electronics Letters*, Vol.33, No.1, pp.64-65, 1997.

[11] C. Jutten and J. Herault, "Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture," *Signal Processing*, 24, pp. 1-10, 1991.

[12] E. Moreau and O. Macchi, "Two novel architectures for the self adaptive separation of signals," in *IEEE Conf. Comm.*, pp. 1154-1159, 1993.

[13] E. Oja, J. Karhunen, and A.Hyvarinen, "From Neural PCA to Neural ICA," *Neural Information Processing Systems Workshop on Blind Signal Processing and its Applications*, December 1996.

Fast Wideband Source Separation Based on Higher-Order Statistics

S. Bourennane, M. Frikel and A. Bendjama
C.M.C.S URA CNRS 2053
B.P. 52, Quartier Grossetti, 20250 Corte France
e-mail : bourenna@univ-corse.fr

Abstract

In this paper we develop an algorithm to improve the accuracy of the estimation of the direction of arrival of the wide-band sources. It is well known that when the noise cross-spectral matrix is unknown, these estimates may be grossly inaccurate. Using both the fourth order cumulant for suppression of the Gaussian noise, the transformation matrices for estimating the coherent signal subspace and a noneigenvector algorithm a robust method for the source characterisation problem in the presence of noise with an unknown cross-spectral matrix is developed. We shall show that the performance of bearing estimation algorithms improves substantially when our robust algorithm is used. Simulation results are presented for the unknown noise spectral matrix.

1. Introduction

The estimation of directions of arrival of multiple narrow-band or wideband sources in the noise is a classic problem in array signal processing[1-6]. Many bearing estimation procedures have been reported in the literature, among which the various eigenanalysis-based methods have been the focus of many studies[1-5]. The eigenstructure procedure is computationally costly when the number of sensors is large. These methods are also dependent on the structure of the noise cross-spectral matrix. A fundamental assumption for most direction finding algorithms, developed in the last decade, is that the noise is spatially and temporally white or the spatial correlation structure of the background noise is known to within a multiplicative scalar. Then, the localisation algorithm can be usually modified in a straightforward manner to include it in the treatment. In practice, this assumption is rarely fulfilled. This is due to the fact that the noise has several origins, such as traffic noise, ambient sea noise, or flow noise, and sometimes the source signals with low power or undetected are assimilated to the noise, which are often spatially correlated. In recent years, there has been a growing interest in the problem of improving high resolution eigenstructure techniques with objective of lowering the signal to noise ratio resolution threshold or the spatially colored noise[7-13]. The ambient noise is unknown in practice, therefore its modelisation or its estimation is necessary. The methods developed for this problem are very few and there is not a definitive solution to this problem. There are some practical methods : in[11] two methods are obtained by optimisation of criterion and by using AR or ARMA modelling of noise. In[12-13] the spatial correlation matrix of noise is modelled by the known Bessel functions. As in[8] the ambient noise spectral matrix is modelled by a sum of hermitian matrices known up to multiplicative scalar. In[14] this estimate is obtained by measuring the array cross-spectral matrix when no signals are present. This procedure assume that the noise is not varying with time, which is not fulfilled in several domain applications. Another possibility[15] arises when the correlation structure is known to be invariant under a translation or rotation. The so-called covariance differencing technique can be then applied to reduce the noise influence. In this method, two identical translated and/or rotated measurements of the array cross-spectral matrix are required and hypothesises the invariance of the

0-8186-8005-9/97 $10.00 © 1997 IEEE

noise cross-spectral matrix, while the source signals change between the two measurements. The estimate of the noise cross-spectral matrix is eliminated by simple subtraction. Furthermore, this method cannot be applied when the source cross-spectral matrix satisfies the same invariance property or when only one measurement is available. In[8-9] a particular modelling structure noise spectral matrix, which takes into account the characteristic noise relative to its origins, is given. In general, even if the individual treatments is different in these articles, the obtained noise structure matrix is the same.

In this paper, the direction-of-arrivals of the sources are estimated in the presence of Gaussian noise with unknown spectral matrix. By means of the property of the fourth-order cumulant, the noise contribution is suppressed. Thus we use the coherent propagator method with spatial fourth order cumulant matrices. Instead of spectral matrices which are used by the coherent signal subspace algorithm, the focusing matrices are used in order to transform the propagator of the fourth order cumulant matrices at each analysis frequency bin into the propagator at the selected frequency.

2. Problem formulation

Consider an uniform linear array composed of N identical sensors separated from each other by a distance d. Let P, (P<N), sources impinge on the array from the directions $\{\theta_1, \theta_2, \theta_3, \ldots, \theta_p\}$. The signal received at the ith sensor can be expressed as :

$$r_i(t) = \sum_{p=1}^{P} s_p(t - \tau_{ip}) + n_i(t), \quad i = 1, \ldots, N$$

$-T/2 < t < T/2$ (1)

where $n_i(t)$ is the additive noise at the ith sensor, $s_p(t)$ is the signal emitted by the pth source and τ_{ip} is the propagation delay associated with the pth source and the ith sensor. Rewriting (1) in matrix notation, in the frequency domain, we obtain :

$$\mathbf{r}(f_j) = \mathbf{A}(f_j)\mathbf{s}(f_j) + \mathbf{n}(f_j), \quad j = 1, \ldots, M \tag{2}$$

where $\mathbf{r}(f_j)$, $\mathbf{s}(f_j)$ and $\mathbf{n}(f_j)$ are the Fourier transforms of the observation, signal and noise vectors, and $\mathbf{A}(f_j)$ is the NxP transfer matrix of the source-sensor array systems with respect to some chosen reference point, $\mathbf{A}(f_j) = [\mathbf{a}(f_j, \theta_1), \mathbf{a}(f_j, \theta_2), \ldots, \mathbf{a}(f_j, \theta_p)]$, $\mathbf{a}(f_j, \theta_i)$ is the steering vector of the array toward the direction, θ_i at the frequency f_j. It is assumed that $\mathbf{A}(f_j)$ is full rank. In other words, for each fj the steering vectors $\mathbf{a}(f_j, \theta_i)$ j=1,....,M and i=1,...,P are linearly independent. For a large T samples of the observation vector at different frequency bins are uncorrelated.

Assume that the signals and the additive noises are uncorrelated, and the noises are assumed to be Gaussians. It follows from these assumptions that the spatial cross-spectral matrix of the observation vector at frequency fj is given by :

$$\Gamma(f_j) = E[\mathbf{r}(f_j)\mathbf{r}^+(f_j)]$$

using the above assumptions, we obtain :

$$\Gamma(f_j) = \mathbf{A}(f_j)\Gamma_s(f_j)\mathbf{A}^+(f_j) + \Gamma_n(f_j) \tag{3}$$

where the superscript + represents the hermitian transpose.

The universal spectral matrix in the coherent subspace method can be shown as :

$$\Gamma = \frac{1}{M}\sum_{j=1}^{M} \mathbf{T}(f_j)\Gamma(f_j)\mathbf{T}^+(f_j) \tag{4}$$

where $\mathbf{T}(f_j)$ is the j-th focusing matrix, such that $\mathbf{T}(f_j)\mathbf{A}(f_j) = \mathbf{A}(f_c)$, where f_c is the selected focusing frequency. In[16] an unitary version of the coherent subspace method is introduced which is based on choosing $\mathbf{T}(f_j)$ to

$$\min \left\| \mathbf{A}(f_c) - \mathbf{T}(f_j)\mathbf{A}(f_j) \right\|^2 \quad j = 1, \ldots, M \tag{5}$$

s.t. $\mathbf{T}^+(f_j)\,\mathbf{T}(f_j) = \mathbf{I}$

where $\mathbf{A}(f_c)$ is the focusing location matrix of the array and $\|.\|$ is the Frobenius norm matrix. The matrix $\mathbf{T}(f_j)$ solution of (5) is the focusing matrix of the unitary coherent subspace method.

The focusing matrices transform the signal subspace at the j-th frequency bin into the focusing signal subspace and it has been shown that the unitary coherent subspace method does not create focusing loss[16]. The focusing loss is defined as the ratio of the signal-to-noise ratio after focusing to the signal to noise ratio before focusing. To determine the focusing matrix $\mathbf{T}(f_j)$ from (5), it is assumed that the matrices $\mathbf{A}(f_j)$ and $\mathbf{A}(f_c)$ are known. In practice, an ordinary beamformer is used to estimate the initial azimuths of the sources. These angles are, then, used as the focusing angles. The spatial fourth order cumulant matrix at j-th temporal frequency bin is defined as

$$\mathbf{H}(f_j)=Cum\left\{\begin{bmatrix} r_1(f_j) & r_1{}^*(f_j) & r_1(f_j) \\ . & . & . \\ r_N(f_j) & r_N{}^*(f_j) & r_N(f_j) \end{bmatrix} \cdot [r_1{}^*(f_j),.....,r_N{}^*(f_j)]\right\}$$

where "*Cum*" is the abbreviation for cumulant and * denotes complex conjugate. Estimate the following fourth-order cumulant matrix:

$$\mathbf{H}_{11}(f_j)=cum(r_1(f_j),r^*{}_1(f_j),\mathbf{r}(f_j),\mathbf{r}^+(f_j))$$

Substituting (1) into the above expression and taken account the noise is Gaussian, after some calculations[17-18] we obtain

$$\mathbf{H}_{11}(f_j)=\sum_{i=1}^{P} h_{si}(f_j)\left|\mathbf{A}_j(1,i)\right|^2 \mathbf{a}_i(f_j)\mathbf{a}_i{}^+(f_j)$$

$$\mathbf{H}_{11}(f_j)=\mathbf{A}(f_j)\mathbf{H}_S(f_j)\mathbf{A}^+(f_j) \qquad (6)$$

where $\mathbf{H}_s(f_j)$ is the diagonal matrix

$$\mathbf{H}_S(f_j)=diag[\left|\mathbf{Aj}(1,1)\right|^2 h_{s1}(f_j),$$

$$\left|\mathbf{Aj}(1,2)\right|^2 h_{s2}(f_j),....,\left|\mathbf{Aj}(1,P)\right|^2 h_{sp}(f_j)]$$

with

$$h_{si}(f_j)=Cum[s_i(f_j),s_i{}^*(f_j),s_i(f_j),s_i{}^*(f_j)]$$

of the sources signals. In equation (6) the suppression of the noise spectral matrix is due to the fact the noise is assumed to be Gaussian then their fourth order cumulants are zero. In practice the fourth-order cumulant matrix $\mathbf{H}_{11}(f_j)$ is estimated from the set of received vectors $\mathbf{r}_k(f_j)$, j =1,.....,M and k=1,.., K.

3. Propagator method

Using the assumption, that the matrix $\mathbf{A}(f_j)$ is of rank P, there exists a (N-P) x P matrix such that :

$$\mathbf{A}_2(f_j)=\Pi^+(f_j)\mathbf{A}_1(f_j), \text{ for j=1,...,M} \qquad (7)$$

where $\mathbf{A}_1(f_j)$ and $\mathbf{A}_2(f_j)$ are two block matrices of the transfert matrix $\mathbf{A}(f_j)$, of dimensions (PxP) and (N-P)xP respectively. The matrix $\Pi(f_j)$ is the propagator operator. For estimating $\Pi(f_j)$, the fourth-order cumulant matrix(6) is used. Using (7), equation (6) can be written[19-20] :

$$\mathbf{H}_{11}(f_j)=\begin{bmatrix} \mathbf{H}_{11}^{11}(f_j) & \mathbf{H}_{11}^{12}(f_j) \\ \mathbf{H}_{11}^{21}(f_j) & \mathbf{H}_{11}^{22}(f_j) \end{bmatrix}, \text{ where}$$

$$\mathbf{H}_{11}^{11}(f_j)=\mathbf{A}_1(f_j)\mathbf{H}_s(f_j)\mathbf{A}_1{}^+(f_j) \qquad (8)$$

$$\mathbf{H}_{11}^{12}(f_j)=\mathbf{H}_{11}^{11}(f_j)\Pi(f_j) \qquad (9)$$

$$\mathbf{H}_{11}^{21}(f_j)=\Pi^+(f_j)\mathbf{H}_{11}^{11}(f_j) \qquad (10)$$

$$\mathbf{H}_{11}^{22}(f_j)=\Pi^+(f_j)\mathbf{H}_{11}^{11}(f_j)\Pi(f_j). \qquad (11)$$

For example, one can estimate the propagator by :

$$\Pi(f_j)=\left(\mathbf{H}_{11}^{11}(f_j)\right)^{-1}\mathbf{H}_{11}^{12}(f_j). \qquad (12)$$

The obtained propagator is used to calculate the localisation function, given by :

$$L(\theta)=\frac{1}{\left|\mathbf{Q}^+(f_j)a(f_j,\theta)\right|^2}, \text{ for } \theta\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right],$$

where $\mathbf{Q}^+(f_j)=\left[\Pi^+(f_j)\ \middle|-\mathbf{I}\right]$. (13)

In order to exploit the advantage provided by the fourth-order cumulant domain and the coherent treatment, the transformation matrices are applied to the propagator operator at each frequency bin $\mathbf{T}(f_j)\Pi(f_j)=\hat{\Pi}(f_c)$ for j =1,..,M, where $\hat{\Pi}(f_c)=\left[\hat{\mathbf{A}}_1(f_c)\hat{\mathbf{H}}_s(f_c)\hat{\mathbf{A}}_1^{+}(f_c)\right]^{-1}\mathbf{H}_{11}^{12}(f_c)$ and $\hat{\mathbf{H}}_s(f_c)=\frac{1}{M}\sum_{j=1}^{M}\hat{\mathbf{A}}_1^{-1}(f_j)\mathbf{H}_{11}^{11}(f_j)\left(\hat{\mathbf{A}}_1^{+}(f_j)\right)^{-1}$,

$\hat{\mathbf{A}}_1(f_j)$ is obtained by using an initial estimates of the direction of arrival of the sources. The transformation matrix is given by :

$$\mathbf{T}(f_j)=\hat{\Pi}(f_c)\left(\mathbf{H}_{11}^{12}(f_j)\right)^{+}\left[\mathbf{H}_{11}^{12}(f_j)\mathbf{H}_{11}^{12+}(f_j)\right]^{-1}\left(\mathbf{H}_{11}^{11}(f_j)\right)^{+}$$

The coherent propagator is then :

$$\tilde{\Pi}(f_c)=\frac{1}{M}\sum_{j=1}^{M}\mathbf{T}(f_j)\Pi(f_j).$$

Finally, $\tilde{\Pi}(f_c)$ may be used in (13) for the estimation of the direction-of-arrivals of the wide-band sources.

4. Numerical example

We consider a linear array of N=14 sensors uniformly spaced at half wavelength of a central temporal frequency. The source signals are temporally stationary zero-mean, non-Gaussian wide-band. Four sources impinge on the array θ_1 = 8°, θ_2 = 10°, θ_3 = 30° and θ_4 = 32° respectively, in the presence of the colored and Gaussian noise. The array noise is stationary zero-mean, independent of the signals. The signal to noise ratio is 10 dB. From the array outputs, 256 snapshots of 64 samples each were selected and the frequency components were obtained via FFT. The directions 9° and 31° given by the beamformer method are used to estimate the focusing matrices. The MUSIC spectra of spatial spectral matrix and cumulant matrix based on the coherent propagator method are plotted in Figures 1 and 2 respectively. From Figure 2 it is obvious that the four sources are localized by using the cumulant matrices while they are not resolved when the spectral matrices are used.
One can remark the Music method can not separate the four sources even the number of the sources is taken equal to 4, but the proposed method gives the exact azimuth of the sources.

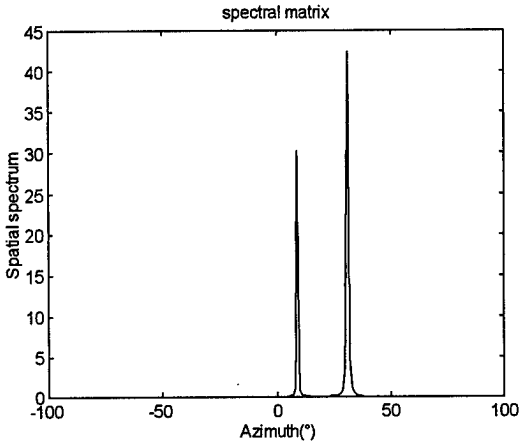


Figure1 : Spectral Matrix and Coherent Signal Subspace Method.

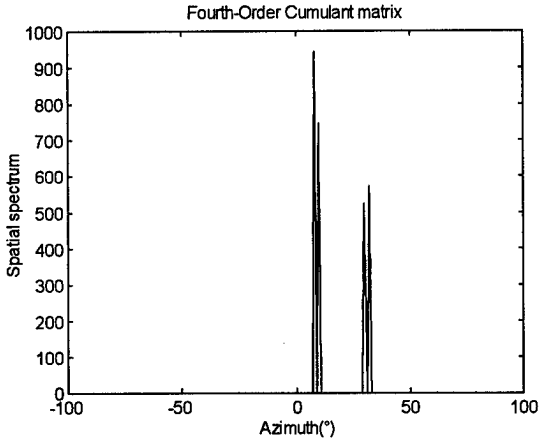


Figure 2 : Fourth-Order Cumulant and Coherent propagator Method.

5. Conclusion

In this study, we have developed a coherent propagator method based on the fourth-order cumulant for locating the wide-band sources from the received data in the presence of unknown noise

fields. The simulation results show that the proposed algorithm has asymptotically a performance similar to (i.e., exact and unbiased) the standard eigenstructure algorithm applicable in known noise fields.

References

[1] G. Bienvenu, L. Kopp, Optimality of high resolution array processing using the eigensystem approach, IEEE Trans. ASSP,n° 5, pp. 1235-1247, oct.(1983).

[2] H. Wang, M. Kaveh, Coherent signal subspace processing for the detection and estimation of angles of arrival of multiple wide band sources, IEEE Trans. ASSP, n°4, Aug.(1985).

[3] S. Bourennane, M. Frikel, Localisation of the wide band sources with estimation of the antenna shape, IEEE-Workshop on Statistical signal and array processing, Corfu, Greece, June (1996).

[4] S. Bourennane, M. Frikel, New method for locating the wide band sources, IEEE-DSP, Loen, Norway, Sept.(1996).

[5] S. Bourennane, M. Frikel, Optimal coherent subspace for bearing estimation, IEEE-TENCON DSP Applications, Perth Australia, Nov.(1996).

[6] J. Munier, Identification de fronts d'ondes corrélés et distordus, Revue traitement du signal, n°4, (1987).

[7] C. Latombe, I. Tas, Improvement of passive array treatment by estimation of the spectral matrix of the noises, Revue Traitement du signal, n° 2, (1987).

[8] B. Göranson, Direction finding in the presence of spatially correlated noise fields, EUSIPCO'94, (1994).

[9] B. Friedlander, A. J. Weiss, Direction finding using noise covariance modelling, IEEE Trans. SP,n° 2, Jul.(1995)

[10] Q. T. Zhang,K.M. Wong, Information theoretic criteria for the determination of the number of signals in spatially correlated noise, IEEE Trans. SP, n° 4, April (1993).

[11] J.P. Le cadre, Parametric methods for spatial processing in the presence of unknown colored noise fields, IEEE Trans. ASSP, n°7, (1989).

[12] G. Bienvenu, L. Kopp, Multiple detection using eigenvalues when the noise spatial coherence is partially unknown, NATO ASI, n° 7, Jul. (1989).

[13] G. Bienvenu, L. Kopp, Méthode haute résolution pour la localisation de sources rayonnantes, Onde électrique, n° 4, (1984).

[14] A. Paulraj, T. Kailath, Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase, ICASSP'85, (1985).

[15] A. Paulraj, T. Kailath, Eigenstructure methods for direction of arrival estimation in the presence of unknown noise fields, IEEE Trans. ASSP, n° 1, (1986).

[16] H. Hung and M. Kaveh, Focusing matrices for coherent signal subspace processing, IEEE Trans. ASSP, pp.1272-1281 (1988).

[17] C.L. Nikias and A.P. Petropulu, Higher order spectra analysis, A non linear signal processing Framework, prentice Hall, (1993).

[18] J.M. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory : theoretical results and some applications, Proc. IEEE, vol.79, n°3, pp.278-305, March (1991).

[19] S. Valaee and P. Kabal,Wide-band array processing using a two-sided correlation transformation, IEEE Trans. On S.P, jan.(1995).

[20] M. Frikel and S. Bourennane, Fast Algorithm for the wide-band array processing using a two-sided correlation transformation, EUSIPCO'96, (1996).

[21] V. Capdevielle, Ch. Serviere, J.L. Lacoume, Blind separation of wideband sources : Application to rotating machine signals, EUSIPCO'96, pp. 2085-2088.

COMPUTATIONAL SAVINGS USING NONLINEAR STATISTICS IN DOA ESTIMATION

Patrizio Campisi, Gaetano Scarano***

*Dip. Ingegneria Elettronica, Universitá di Roma III
via della Vasca Navale 84, I-00146 Roma, Italy
E-mail: campisi@ele.uniroma3.it

**Dip. INFOCOM, Universitá di Roma "La Sapienza"
via Eudossiana 18, I-00184 Roma, Italy
E-mail: gaetano@infocom.ing.uniroma1.it

Abstract

The computational aspects related to sample estimation of moments involving certain "piecewise" nonlinearities are addressed with application to DOA estimation. In particular, the accuracy vs. computational saving tradeoff associated to "soft–limiting" nonlinearities can be exploited to simplify the computation of sample covariances without resulting in significative accuracy loss. It is also shown how, in sample cumulants evaluation, this approach can be employed to reduce the overall number of aritmethic operations using nonlinearities which act separately on the real and the imaginary parts of complex numbers.

I Introduction

High resolution Direction of Arrival (DOA) estimation such as MUSIC, ESPRIT, *etc.*, exploit the geometric structure of the covariance matrix of the signal measured through an array of sensors. While some attention has been devoted in literature regarding the computational aspects of the underlying eigen–decomposition of the covariance matrix (*e.g.* [1]) or concerning the proposition of alternative methods avoiding such eigen–decomposition (*e.g.* [2]), the reduction of the complexity associated to the computation of the covariance has not received equal attention.

In [3, 4], nonlinear statistics have been proposed for DOA estimation in the framework of the so–called *Hybrid Nonlinear Moments* (HNL); particular *fast* statistics based on the extraction of the *signum* of real and imaginary parts of the complex valued signal samples have been described, extending techniques for estimation of covariance of Gaussian processes based on Bussgang theorem [5]. An extension of Bussgang theorem is found in [6] for complex processes while [7] address the multivariate real case. In [8], a similar technique has been adopted for estimation of time–frequency distribution of polynomial phase signal and its accuracy has been thoretically evaluated. Recently, a different dithering/hard-clipping technique has been extendend to computation of higher–order moments in [9].

The main drawback of signum based nonlinearities is related to the loss of accuracy introduced by the hard–limiters.

In this contribution, we introduce *piecewise* nonlinearities aimed at saving computations in the numerical evaluation, while preserving the desired degree of accuracy without increasing the size of the sample. The nonlinearities herewith addressed have the following form

$$p(x) = \begin{cases} f_1(x) & \text{for } |x| < \alpha \\ f_2(x) & \text{for } |x| > \alpha \end{cases} \tag{1}$$

where the parameter α controls the choice among the nonlinearities $f_1(\cdot)$ and $f_2(\cdot)$ which act on medium–low signal levels and medium–high signal levels, respectively.

The computational savings obtainable from nonlinearities (1) are related to particular forms assumed by $f_1(\cdot)$ or $f_2(\cdot)$. For instance, the parameter α controls the tradeoff between covariance (accuracy) and signum (fast computation) when $f_1(x) = x$ and $f_2(x) = \text{sign}(x)$.

II DOA Estimation using Nonlinear Statistics

We are concerned with the following linear model of observations drawn from a M sensors uniform linear array (ULA)

$$\mathbf{x} = \mathbf{A}(\boldsymbol{\omega}) \cdot \mathbf{s} + \mathbf{w} \tag{2}$$

where $\mathbf{s}$ is a L vector of non Gaussian, independent sources, $\mathbf{w}$ is a M vector of (possibly coloured) Gaussian noises, independent of the sources, the DOA's are collected in the vector $\boldsymbol{\omega} \stackrel{\text{def}}{=} [\omega_1 \cdots \omega_L]^{\text{T}}$ and

$$\mathbf{A}(\boldsymbol{\omega}) \stackrel{\text{def}}{=} [\mathbf{a}(\omega_1) \ \cdots \ \mathbf{a}(\omega_L)]$$

is the matrix of DOA's related steering vectors

$$\mathbf{a}(\omega_l) = \left[1 \ e^{j\omega_l} \cdots e^{j\omega_l(M-1)}\right]^{\text{T}}$$

Modern high–resolution methods of DOA estimation rely on the subspace decomposition of the covariance matrix

$$\mathbf{R}_{\text{x}} \stackrel{\text{def}}{=} \text{E}\{\mathbf{x} \cdot \mathbf{x}^{\text{H}}\} = \mathbf{A} \cdot \mathbf{R}_{\text{s}} \cdot \mathbf{A}^{\text{H}} + \mathbf{R}_{\text{w}} \tag{3}$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

where $\mathbf{R}_s$ and $\mathbf{R}_w$ are the covariance matrices of the sources and the noises, respectively, and the dependence on the DOA vector ω has been dropped out for readibility purposes. If the number of sources L is less than the number of sensors M, suitable subspace decomposition of $\mathbf{R}_x$ constitutes the basis for DOA estimation. In [3, 4], an extension of this approach to nonlinear statistics is described and it is summarized below.

Let $g(x,\overline{x})$ a complex nonlinearity, having denoted complex conjugation by an overbar ; taking the cross–correlations between $\mathbf{x} \stackrel{\text{def}}{=} [x_1 \cdots x_M]^{\mathrm{T}}$ and its nonlinear version $g(\mathbf{x}) \stackrel{\text{def}}{=} [g(x_1,\overline{x}_1) \cdots g(x_M,\overline{x}_M)]^{\mathrm{T}}$, we obtain a class of matrices having the same structure of the covariance matrix (3). Thus, subspace decomposition based methods can be still applied. In fact, we have [3, 4]

$$\mathbf{G}_{\mathbf{x}} \stackrel{\text{def}}{=} \mathrm{E}\{\mathbf{x} \cdot g(\mathbf{x})^{\mathrm{H}}\} = \mathbf{A} \cdot \mathbf{H}_{\mathbf{x}} \cdot \mathbf{A}^{\mathrm{H}} + k_g \cdot \mathbf{R}_{\mathrm{w}} \qquad (4)$$

where $\mathbf{H}_{\mathbf{x}}$ is a matrix depending on the pdf of the sources and of the noises and

$$k_g \stackrel{\text{def}}{=} \mathrm{E}\left\{\frac{\partial}{\partial x} g(x,\overline{x})\right\} \qquad (5)$$

is the (Bussgang) proportionality factor relating nonlinear cross-correlations of Gaussian processes.[1]

In particular, we investigate on the use of a complex *limiter*, [2] *i.e.*

$$g(x_r, x_i) = \mathrm{lmt}(x_r) + j\mathrm{lmt}(x_i) \qquad (6)$$

where the real nonlinearity $\mathrm{lmt}(v)$ is defined as follows:

$$\mathrm{lmt}(v) = \begin{cases} v & \text{for } |v| < \alpha \\ \alpha \cdot \mathrm{sign}(v) & \text{for } |v| > \alpha \end{cases} \qquad (7)$$

where the parameter α trades off between accuracy and computational savings. The saving is obtained since, for $|x_r| > \alpha$ or $|x_i| > \alpha$ the multiplication needed in the computation of the sample statistics reduce to a sign bit adjustement. An analytical perfomance evaluation is currently under development for the Gaussian case, following the indication of [13]. Here, we will present some simulation results concerning the accuracy/saving tradeoff.

Other computational savings are possible when the sources are non–Gaussian distributed. In this case, DOA estimation based on higher–order cumulants offers some advantages concerning the blind rejection of coloured Gaussian unwanted components or the detection of phase coupling between the harmonics [10, 11, 12].

Nonlinear statistics possessing similar properties to cumulants can be obtained by subtracting an adequate quantity of covariance (second–order cumulant) from a general nonlinear moment. In fact, the expectation

$$\mathrm{E}\{x \cdot \overline{\gamma}(y,\overline{y})\} \stackrel{\text{def}}{=} \mathrm{E}\{x \cdot \overline{g}(y,\overline{y})\} - k_g \cdot \mathrm{E}\{x \cdot \overline{y}\} \qquad (8)$$

does not contain any second–order (covariance) contribution when the constant k_g is chosen as indicated by Bussgang theorem [3, 4], namely

$$k_g = \mathrm{E}\left\{\frac{\partial}{\partial y} g(y,\overline{y})\right\} \qquad (9)$$

As an utilization of (8), we will show how power nonlinearities acting separately on the real and the imaginary parts may offer some computational savings. For instance, let us consider the following fourth–order slice cross–cumulant between the circularly complex variates x and y with variances σ_x^2 and σ_y^2, respectively:

$$\mathrm{cum}\,(x,\overline{y},y,\overline{y}) \stackrel{\text{def}}{=} \mathrm{E}\{x \cdot \overline{y} \cdot y \cdot \overline{y}\} - 2 \cdot \sigma_y^2 \cdot \mathrm{E}\{x \cdot \overline{y}\} \qquad (10)$$

and consider also the following nonlinearity acting separately on the real and the imaginary part of $y \stackrel{\text{def}}{=} y_r + jy_i$

$$g_3(y_r, y_i) \stackrel{\text{def}}{=} y_r^3 + jy_i^3 \qquad (11)$$

The following nonlinear moment

$$\mathrm{E}\{x \cdot \overline{\gamma}_3(y,\overline{y})\} = \mathrm{E}\{x \cdot \overline{g}_3(y,\overline{y})\} - k_{g_3} \cdot \mathrm{E}\{x \cdot \overline{y}\} \qquad (12)$$

still posseses the cumulants property of rejecting Gaussian additive random variables since

$$\begin{aligned} k_{g_3} &= \frac{1}{2}\mathrm{E}\left\{\left(\frac{\partial}{\partial y_r} - j\frac{\partial}{\partial y_r}\right)\left(y_r^3 + jy_i^3\right)\right\} \\ &= \frac{3}{2}\left(\sigma_{y_r}^2 + \sigma_{y_i}^2\right) = \frac{3}{2} \cdot \sigma_y^2 \end{aligned} \qquad (13)$$

Moreover, mimicking the limiter behaviour, a computational parsimonious nonlinearity, acting separately on the real and imaginary parts, is

$$f_3(y,\overline{y}) = f(y_r) + jf(y_i) \qquad (14)$$

[1] The partial differentiation in (5) is defined in terms of differentiations w.r.t. the real and imaginary parts of $x \stackrel{\text{def}}{=} x_r + jx_i$ as follows

$$\frac{\partial}{\partial x} \stackrel{\text{def}}{=} \frac{1}{2}\left(\frac{\partial}{\partial x_r} - j\frac{\partial}{\partial x_i}\right)$$

Analogously,

$$\frac{\partial}{\partial \overline{x}} \stackrel{\text{def}}{=} \frac{1}{2}\left(\frac{\partial}{\partial x_r} + j\frac{\partial}{\partial x_i}\right)$$

For analytic functions, the former defines the complex differention rule due to Cauchy–Riemann conditions, while the latter vanishes for the same reason. The nonlinearies $g(x,\overline{x})$ herewith considered are non analytic functions w.r.t. the complex variable x; rather they should be considered as complex valued functions of the two real variables (x_r, x_i). Regarding x and $\overline{x}$ as independent variables, they can be represented by biargumental analytic functions (see [6]).

[2] With a little abuse of notation, we use the same notation, *e.g.* $g_3(y,\overline{y})$ and $g_3(y_r, y_i)$, to represent a complex nonlinearity either when the independent variables are $(x,\overline{x})$ or (x_r, x_i)

where

$$f(v) = \begin{cases} v & \text{for } |v| < \alpha \\ v \cdot \dfrac{v^2}{\alpha^2} & \text{for } |v| > \alpha \end{cases} \qquad (15)$$

The Bussgang constant

$$k_{f_3} \stackrel{\text{def}}{=} \mathrm{E}\left\{ \frac{\partial}{\partial y} f(y, \overline{y}) \right\}$$

associated to the nonlinearity (15) needs the first derivative

$$\frac{\partial}{\partial y} f_3(y, \overline{y}) = \frac{1}{2} \left(\dot{f}(y_r) + \dot{f}(y_i) \right)$$

where

$$\dot{f}(v) = \begin{cases} 1 & \text{for } |v| < \alpha \\ 3 \cdot \dfrac{v^2}{\alpha^2} & \text{for } |v| > \alpha \end{cases}$$

The savings offered by the nonlinear expectation $\mathrm{E}\{x \cdot \overline{f}_3(y, \overline{y})\} - k_{f_3} \cdot \mathrm{E}\{x \cdot \overline{y}\}$ consist in avoiding the square and two real multiplications involved in (12) whenever it results $|y_r| < \alpha$ or $|y_i| < \alpha$ in sample estimation. This constitutes a 50% saving w.r.t. extra computations needed to cope with the nonlinearity $g_3(\cdot, \cdot)$ in (12).

III Simulation Results and Conclusion

Computer simulations have been performed to support the feasibility of the above described approaches.

The considered scenario consists of two equiamplitude waves, having independent and uniformly distributed phases, impinging on an array of 8 half–wavelenght spaced sensors. First, the tradeoff between accuracy and computation of nonlinear moments involving the complex limiter (LMT) of (6) has been addressed in the case of white additive Gausian noises for different SNR values. To test the accuracy of the limiter based estimation, the angular difference between the DOA's is diminished from 10° in Fig.1 (DOA's 15° and 25°) to 5° in Fig.2 (DOA's 15° and 20°). In the same figures, the MSE relative to the DOA 15° is reported also for the covariance based estimation (curves labelled "COV") and the so–called "complex hybrid signum" based estimation (curves labelled "CHS"); this latter is obtained for $\alpha = 0$ in (6). All the estimations are drawn from Root-MUSIC applied to sample statistics of size $N = 100$ independent snapshots. Averages over 100 MonteCarlo runs have been performed. The value of the parameter α has been chosen equal to half the standard deviation of the signal received at the generic sensor: we have observed that this roughly corresponds to a 25% savings of real multiplications and additions. We see that, for DOA spacing of 5°, the complex signum (CHS) based estimation presents a sensible accuracy loss while the limiter based estimation (LMT) shows the expected intermediate behavior between covariance and signum based estimates.

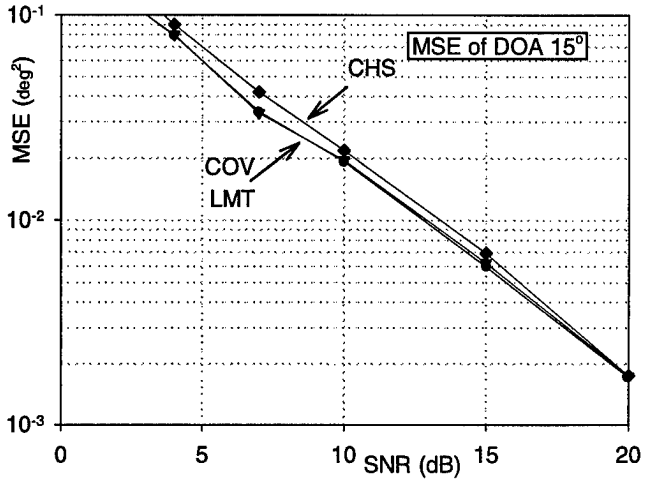


Figure 1: *Two sources from 15° and 25° in white noises: MSE of Root-MUSIC estimates of DOA 15° from 100 snapshots and 100 MonteCarlo runs.*

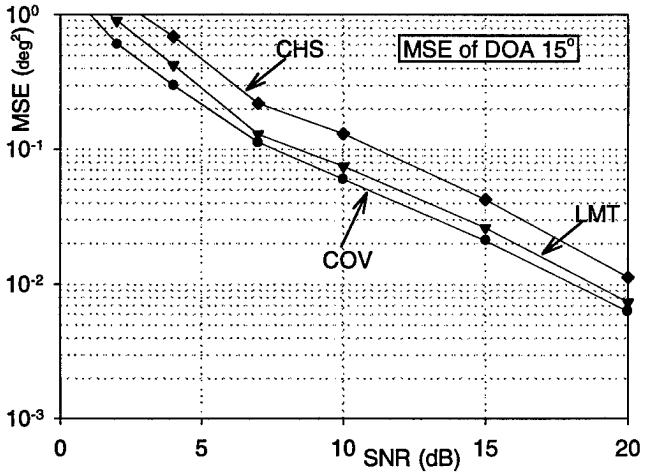


Figure 2: *Two sources from 15° and 20° in white noises: MSE of Root-MUSIC estimates of DOA 15° from 100 snapshots and 100 MonteCarlo runs.*

In Figs.3 and 4, the case of highly correlated noises, with correlation coefficient very close to one, has been considered. Here, estimation drawn from the nonlinearity $f_3(\cdot, \cdot)$ in (14) is considered along with that relative to fourth–order slice cumulants defined in (10), and the mean squared estimation errors of DOA 15° and of DOA 25° are plotted in Fig.3 and Fig.4, respectively, for a sample size of $N = 256$ independent snapshots. The value of α is chosen as already

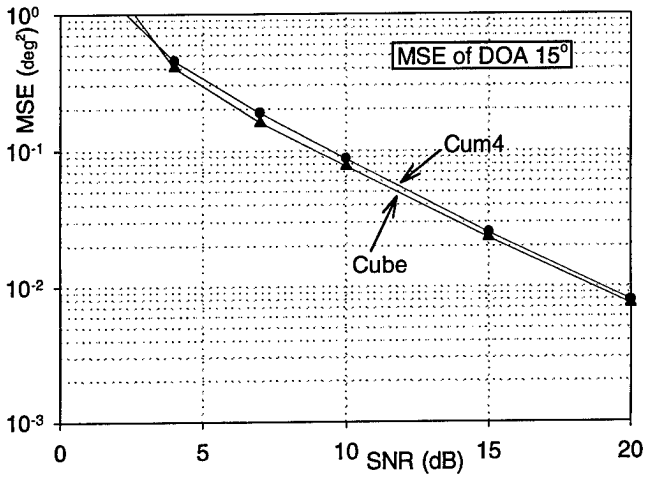


Figure 3: *Two sources from 15° and 25° in highly correlated noises: MSE of Root-MUSIC estimates of DOA 15° from 256 snapshots and 100 MonteCarlo runs.*

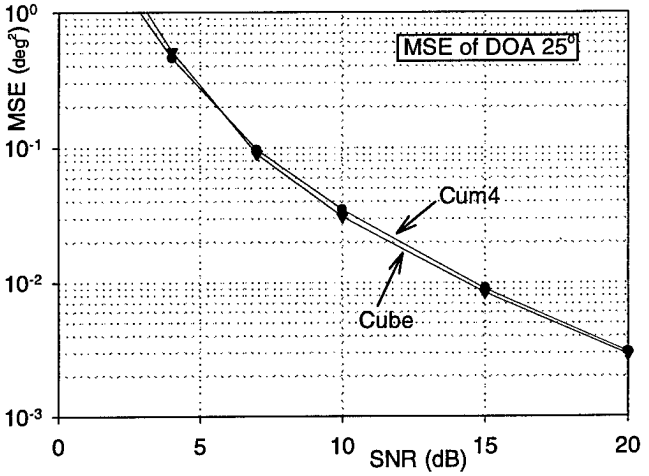


Figure 4: *Two sources from 15° and 25° in highly correlated noises: MSE of Root-MUSIC estimates of DOA 25° from 256 snapshots and 100 MonteCarlo runs.*

done in Figs.1 and 2 and, also in this case, it roughly corresponds to a 25% savings of real multiplications and sums. We see that the accuracy of the estimation based on (14), (curves labelled by "Cube") is even slightly better than that obtained from fourth–order cumulants (curves labelled by "Cum4'), which is far from being an "optimal" statistics. In conclusion, the use of nonlinear statistics in DOA estimation can save computations without affecting the overall accuracy. A performance analysis of the complex limiter (6) based estimation is currently under investigation in the reference case of Gaussian sources in Gaussian noise.

References

[1] P. Comon,, G.H. Golub, "Tracking a few extremal singular values and vectors in signal processing", *Proc. IEEE*, vol.78, no.8, August 1990.

[2] S. Marcos, A. Masal, M. Benedir, "The propagator method for source bearing estimation", *Signal Processing*, vol.42, 1995.

[3] G. Jacovitti, G. Scarano, "Hybrid non-linear moments in array processing and spectrum analysis", *IEEE Tran. on Sig. Proc.*, vol.42, no.7, July 1994.

[4] G. Scarano, G. Jacovitti, "Applications of generalized cumulants to array processing", *Sig. Proc.*, vol. 53, no. 2, 1996.

[5] G. Jacovitti, A. Neri, R. Cusani, "Methods for estimating the autocorrelation function of complex Gaussian processes", *IEEE Trans. on Acoust., Speech and Sig. Proc.*, vol.35, no.8, August 1987.

[6] G. Scarano, "Cumulant series expansion of hybrid non-linear moment of complex random variables", *IEEE Trans. on Sig. Proc.*, vol.39, no.4, April 1991.

[7] G. Scarano, D. Caggiati, G. Jacovitti, "Cumulant series expansion of hybrid non-linear moment of *n* variates", *IEEE Trans. on Sig. Proc.*, vol.41, no.1, January 1993.

[8] A. Scaglione, S.Barbarossa, A. Porchia, G. Scarano, "Nonlinear time frequency distributions with multiplication–free kernels ", *8–th IEEE Workshop on Statist. Sig. and Array Proc.*, Corfu, Greece, June 1996.

[9] A. Swami, B. Sadler, "Fast estimation of higher–order moments using sign bit and reference signals", *Proc. VII IEEE Dig. Sig. Proc. Workshop*, Loen, Norway, 1-4 September, 1996.

[10] A. Swami, J.M. Mendel, "Cumulant-based approach to the harmonic retrieval and related problem," *IEEE Trans. on Sig. Proc.*, vol.39, no.5, May 1991.

[11] M.C. Dogan, J.M. Mendel, "Cumulant–based blind optimum beamforming", *IEEE Trans. on Aerosp. and Electr. Sys.*, vol.30, no.3, July 1994.

[12] J.M.M. Anderson, G.B. Giannakis, A. Swami, "Harmonic retrieval using higher order statistics: a deterministic formulation", *IEEE Transactions on Signal Processing*, vol.43, no.8, August 1995.

[13] P. Stoica, A. Nehorai, "MUSIC, maximum likelihood, and Cramer–Rao bound", *IEEE Trans. Acous. Sp. Sig. Proc.*, vol. 37, no. 5, May 1989.

BEHAVIOUR OF HIGHER ORDER BLIND SOURCE SEPARATION METHODS IN THE PRESENCE OF CYCLOSTATIONARY CORRELATED MULTIPATHS

Pascal Chevalier[1], *Véronique Capdevielle*[1] *and Pierre Comon*[2]

[1]Thomson-CSF Communications, UTTC/TSI, 66 rue du Fossé Blanc, 92231 Gennevilliers, France
E-Mail : pascal.chevalier@thomcom.thomson.fr
[2]Eurecom Institute, 2229 route des crêtes, BP 193, 06904 Sophia Antipolis Cedex, France
E-Mail : Comon@Eurecom.fr

ABSTRACT

Over the last decade, higher order (HO) methods have been strongly developed in particular to blindly separate instantaneous mixtures of statistically independent stationary sources. However, in many situations of practical interest, the received sources are (quasi)-cyclostationary (digital radiocommunications) and are not always statistically independent but may be correlated to each other, which occurs in particular for HF links or in mobile radiocommunications contexts where propagation multipaths are omnipresent. In such situations, the behaviour of the classical HO blind source separation methods is not known, which may be a limitation to the use of these methods in operational contexts. The purpose of this paper is precisely to fill the gap previously mentionned by analysing the behaviour, in radiocommunications contexts, of three classical HO blind source separation methods when several potentially correlated paths of each source, assumed (quasi)-cyclostationary, are received by the array.

1. INTRODUCTION

Over the last decade, higher order (HO) methods have been strongly developed in particular to blindly separate instantaneous mixtures of statistically independent and stationary sources [1-5]. In [6-7], the performance of two of these methods, corresponding to the so-called JADE method [2] and to the one which optimizes a constrast function squaring the samples fourth-order autocumulants [3], have been analysed for arbitrary statistically independent and stationary sources scenari. In a same way, the performance of a third method [5], optimizing a constrast function which is not squaring the samples fourth-order autocumulants, have been presented recently in [8] still for statistically independent stationary sources.

However, in many situations of practical interest, the received sources are (quasi)-cyclostationary (digital radiocommunications) and are not always statistically independent but may be correlated to each other, which occurs in particular for HF links or in mobile radiocommunications contexts where propagation multipaths are omnipresent. In such situations, the behaviour of the previous HO blind source separation methods is not known, which may be a limitation to the use of these methods in operational contexts.

The purpose of this paper is precisely to fill the gap previously mentionned by analysing the behaviour, in radiocommunications contexts, of the three HO blind source separation methods introduced in [2], [3] and [5] respectively, when several potentially correlated paths of each source, assumed (quasi)-cyclostationary, are received by the array.

2. HYPOTHESIS AND PROBLEM FORMULATION

Consider an array of N Narrow-Band (NB) sensors and let us call $\boldsymbol{x}(t)$ the vector of the complex amplitudes of the signals at the output of these sensors. Each sensor is assumed to receive a noisy mixture of P statistically independent NB (quasi)-cyclostationary sources, with their associated propagation multipaths. Under these assumptions, the observation vector $\boldsymbol{x}(t)$ can be written as follows

$$\boldsymbol{x}(t) = \sum_{i=1}^{P} \sum_{k=1}^{M_i} \alpha_{ik}\, m_i(t-\tau_{ik})\, e^{-j\omega_0\tau_{ik}}\, \boldsymbol{a}_{ik} + \boldsymbol{b}(t)$$

$$\triangleq \sum_{i=1}^{P} A_i\, \boldsymbol{m}_i(t) + \boldsymbol{b}(t) \triangleq A\, \boldsymbol{m}(t) + \boldsymbol{b}(t) \quad (2.1)$$

where $\boldsymbol{b}(t)$ is the noise vector, assumed zero-mean, stationary, spatially white and Gaussian, ω_0 is the carrier pulsation, M_i is the number of paths associated to the source i, $m_i(t)$ is the complex envelope of the source i, α_{ik}, τ_{ik} and $\boldsymbol{a}_{ik}$ are the complex attenuation, the delay and the steering vector of the path k of the source i, $\boldsymbol{m}_i(t)$ is the $(M_i \times 1)$ vector which components are the complex

*this work has been supported by D.R.E.T (French Administration)

0-8186-8005-9/97 $10.00 © 1997 IEEE

signals $m_{ik}(t) \triangleq \alpha_{ik}\, m_i(t - \tau_{ik})\, e^{-j\omega_0\tau_{ik}}$ $(1 \le k \le M_i)$, A_i is the $(N \times M_i)$ matrix of the sources steering vectors $\boldsymbol{a}_{ik}$ $(1 \le k \le M_i)$, $\boldsymbol{m}(t)$ is the $(M \times 1)$ vector obtained by concatenation of the vectors $\boldsymbol{m}_i(t)$ and A is the $(N \times M)$ matrix of all the vectors $\boldsymbol{a}_{ik}$, where M is the sum of the M_i, $(1 \le i \le P)$.

In these conditions, the correlation matrix of the observation vector, $R_x(t) \triangleq E[\boldsymbol{x}(t)\boldsymbol{x}(t)^\dagger]$, can be written as

$$R_x(t) = A\, R_m(t)\, A^\dagger + \eta_2\, I \qquad (2.2)$$

where $\dagger$ means transposition and conjugation, η_2 is the mean power of the noise per sensor, I is the Identity matrix and $R_m(t) \triangleq E[\boldsymbol{m}(t)\, \boldsymbol{m}(t)^\dagger]$ is the correlation matrix of the vector $\boldsymbol{m}(t)$.

In a same way, the quadricovariance $Q_x(t)$ of the observation vector $\boldsymbol{x}(t)$, which components, defined by $Q_x(i, j, k, l)(t) \triangleq \mathrm{Cum}(x_i(t), x_j(t)^*, x_k(t)^*, x_l(t))$, are the fourth order cumulants of $\boldsymbol{x}(t)$, can be written as

$$Q_x(t) = (A \otimes A^*) Q_m(t) (A \otimes A^*)^\dagger \qquad (2.3)$$

where $Q_m(t)$ is the quadricovariance of the vector $\boldsymbol{m}(t)$, $*$ means complex conjugation and $\otimes$ corresponds to the Kronecker product.

In fact, the expression (2.1) describes N particular convolutive mixtures of P statistically independent sources at the output of the sensors. These mixtures could be processed by every blind separators of convolutive mixtures developed these last years [9]. However, in practical situations, for some reasons such as, for example, that of the numerical complexity, it may be chosen to process the vectorial mixture (2.1) as an instantaneous one, considering a propagation path as a particular source. This is the philosophy we adopt in this paper. In these conditions, although in (quasi)-cyclostationary contexts it may be advantageous to implement a Poly-Periodic (PP) and Widely Linear structure of array filtering [10], the problem of sources separation we address in this paper is to find the Linear and Time Invariant (TI) $(N \times M)$ NB separator W, outputing the vector $\boldsymbol{y}(t) \triangleq W^\dagger \boldsymbol{x}(t)$ and giving, to within a diagonal and a permutation matrix, the best estimate of the vector $\boldsymbol{m}(t)$. In the following sections, we study the behaviour of the tree HO blind source separators W introduced in [2], [3] and [5] respectively, for different scenari of sources and paths, for several digital modulations and relative time delays between the paths.

3. HO BLIND SOURCE SEPARATION OF (QUASI)-CYCLOSTATIONARY SOURCES

3.1 Possible HO blind source separators

In (quasi)-cyclostationary contexts, the matrices (2.2) and (2.3) become Time-Dependent and more precisely PP. As a consequence, the matrices $R_x(t)$ and $Q_x(t)$ have Fourier serial expansions which show off in particular the cyclic frequencies of the observations. It may be very useful to exploit the information contained in all the cyclic frequencies of the observations to improve the performance of the HO blind source separators, as it has been shown recently in [11]. However, for particular reasons such as the numerical complexity, we may prefer to still use, in (quasi)-cyclostationary contexts, the classical methods of HO blind source separation introduced in [2], [3] or [5], which, in this case, exploit only the information contained in the cyclic frequency zero of $R_x(t)$ and $Q_x(t)$, i.e. in the temporal mean $R_x \triangleq \langle R_x(t) \rangle$ and $Q_x \triangleq \langle Q_x(t) \rangle$ of $R_x(t)$ and $Q_x(t)$ respectively, which is the choice we adopt in this paper.

Note that for stationary sources, the temporal mean R_x and Q_x of the 2nd and 4th order statistics correspond to the statistics themselves, which is not the case for (quasi)-cyclostationary sources. In this latter case, R_x and Q_x can still be written as (2.2) and (2.3) but where $R_m(t)$ and $Q_m(t)$ are replaced by their temporal mean noted R_m and Q_m respectively. Thus, the temporal mean operation obviously preserves the algebraic structure of $R_x(t)$ and $Q_x(t)$ and also the potential second and fourth order statistical independence of the paths (R_m is still diagonal and the non zero elements of Q_m are still the 4-th order autocumulants of the paths when the latter are independent).

3.2 Statistics estimation

It is well known that for zero-mean, stationary and ergodic sources, the classical estimators of the 2nd and 4th order cumulants provide asymptotically unbiaised estimates of the 2nd and 4th order cumulant of the data, which variance tends to zero when the number of independent samples increases. However, in the presence of (quasi)-cyclostationary and cyclo-ergodic sources, we must wonder whether these classical estimators still generate asymptotically unbiaised estimates of the data statistics temporal mean. The answer to this question has been given in [12] and is negative in the general case for the 4th order cumulant. More precisely, noting $R_x^\alpha(i, j)$, $C_x^\beta(i, j)$ and $M_x^\gamma(i, j, k, l)$ the coefficients associated to the cyclic frequencies α, β and γ in the Fourier serial expansion of $R_x(i, j)(t) \triangleq E[x_i(t)x_j(t)^*]$, $C_x(i, j)(t) \triangleq E[x_i(t)x_j(t)]$ and $M_x(i, j, k, l)(t) \triangleq E[x_i(t)x_j(t)^* x_k(t)^* x_l(t)]$ respectively, it has been shown in [12] that

$$Q_x(i, j, k, l) = M_x^0(i, j, k, l) - \sum_\alpha R_x^\alpha(i, j)\, R_x^{-\alpha}(l, k) - \sum_\alpha R_x^\alpha(i, k)\, R_x^{-\alpha}(l, j) - \sum_\beta C_x^\beta(i, l)\, C_x^\beta(j, k)^* \qquad (3.1)$$

while the classical 4th order cumulant estimator generates asymptotically an apparent 4th order cumulant given by

$$Q_{xa}(i, j, k, l) \triangleq M_x^0(i, j, k, l) - R_x^0(i, j)\, R_x^0(l, k) - R_x^0(i, k)\, R_x^0(l, j) - C_x^0(i, l)\, C_x^0(j, k)^* \quad (3.2)$$

Comparing (3.1) and (3.2), we deduce that for (quasi)-cyclostationary sources, the classical estimators of the 4th-order cumulant do not generate, in the general case, the true value of the latter, which must be taken into account in the behaviour analysis of the classical HO blind separators in (quasi)-cyclostationary contexts and which may even prevent (in very particular situations) the separation of statistically independent non Gaussian sources [13].

3.3 Classical HO blind separators description

The indirect HO methods presented in [2], [3] and [5] aim at blindly identifying the sources steering vectors before the effective sources separation, the latter being done by implementing a spatial filtering operation from the steering vectors estimates [6]. The blind identification of the latter requires a prewhitening of the data, by the pseudo-inverse, $R_s^{-1/2}$, of a square root of $R_s \triangleq A\, R_m\, A^\dagger$, which aims at orthonormalizing (for statistically independent paths) these steering vectors so as to search for the latter through a unitary matrix simpler to handle. For each of the methods presented in [2], [3] and [5], this unitary matrix must maximize a particular blind criterion which is theoretically a function of the Q_z elements, where Q_z is the temporal mean of the quadricovariance of the whitened data $z(t) \triangleq R_s^{-1/2} x(t)$. However practically, the blind criterion optimized by the classical HO blind separators is a function of the Q_{za} elements, where Q_{za} is the apparent quadricovariance of the whitened data.

Using (2.1), the vector $z(t)$ can be written as

$$z(t) = \sum_{i=1}^{P} \sum_{k=1}^{M_i} m'_{ik}(t)\, \boldsymbol{a}'_{ik} + \boldsymbol{b}'(t) \triangleq A'\boldsymbol{m}'(t) + \boldsymbol{b}'(t) \quad (3.3)$$

where $m'_{ik}(t)$ is the normalized complex envelope of the path ik, A' is the $(M \times M)$ matrix of the whitened paths steering vectors $\boldsymbol{a}'_{ik}$ and $\boldsymbol{b}'(t)$ is the whitened noise vector. Consequently, the Q_z and Q_{za} matrices can be written as

$$Q_{z(a)} = (A' \otimes A'^*) Q_{m'(a)} (A' \otimes A'^*)^\dagger \quad (3.4)$$

where $Q_{m'}$ and $Q_{m'a}$ are the true and the apparent temporal mean of the quadricovariance of $\boldsymbol{m}'(t)$ respectively.

For M stationary and statistically independent total paths, the HO blind separators introduced in [2], [3] and [5] have high performance [6-8], which is directly related to the fact that the M vectors $\boldsymbol{a}'_{ik} \otimes \boldsymbol{a}'_{ik}$ are orthonormal eigenvectors of $Q_z = Q_{za}$ associated to the non zero eigenvalues. We must then wonder whether this result still holds for (quasi)-cyclostationary and potentially correlated paths and if not, what is the behaviour of these separators in such situations, which is the purpose of the following.

4. FOURTH ORDER CORRELATION PROPERTIES OF DIGITAL MODULATIONS

The analysis of the eigenstructure of Q_z and Q_{za} in the presence of several potentially correlated (quasi)-cyclostationary paths requires the analysis of $Q_{m'}$ and $Q_{m'a}$ in the same context and in particular the analysis of the 4th-order correlation properties of digital modulations. For this purpose, we consider in this section only one source ($P = 1$) with two paths ($M_1 = 2$), we assume that $\tau_{11} = 0$, $\tau_{12} = \tau$, we note $m_1(t)$ simply $m(t)$ and we analyse the evolution of the 16 $Q_{m'}$ and $Q_{m'a}$ elements as a function of τ. Note that due to the particular symetries of these matrices, the 16 elements of each matrix can be deduced from only 5 elements corresponding to the element (1, 1, 1, 1) (temporal mean of the 4th order true or apparent autocumulant) and to the 4 elements (1, 1, 1, 2), (1, 1, 2, 2), (1, 2, 2, 1) and (1, 2, 2, 2) (temporal mean of the 4th order true or apparent crosscumulants).

Recalling that the 2nd order correlation coefficient between the two considered paths is defined by $\rho_2(\tau) \triangleq \langle E[m'(t)\, m'(t-\tau)^*] \rangle e^{j\omega_0 \tau}$, we define, for each of the two matrices $Q_{m'}$ and $Q_{m'a}$, four 4th-order correlation coefficients (associated to the indices $ijkl$ = 1112, 1122, 1221 and 1222) defined by

$$\rho_{4(a)}[ijkl](\tau) \triangleq Q_{m'(a)}[ijkl] \,/\, Q_{m'(a)}[1111] \quad (4.1)$$

The four coefficients $\rho_4[ijkl](\tau)$ and the four others $\rho_{4a}[ijkl](\tau)$ characterize the true and the apparent 4th-order correlation of all the modulations respectively. From these coefficients, it is also possible to define, for each matrix $Q_{m'}$ and $Q_{m'a}$, an average 4th order correlation coefficient which modulus can be defined by the following expression

$$|\rho_{4(a),av}(\tau)| \triangleq (1/14)[4|\rho_{4(a)}[1112](\tau)| + 4|\rho_{4(a)}[1122](\tau)| + 2|\rho_{4(a)}[1221](\tau)| + 4|\rho_{4(a)}[1222](\tau)|] \quad (4.2)$$

In order to quantify the 4th-order correlation of some modulations, let us consider the linear modulations, characterized by a complex envelope $m(t)$ given by

$$m(t) = \sum_n a_n\, v(t - nT) \quad (4.3)$$

where the complex symbols a_n are i.i.d. random variables, T is the symbol duration and $v(t)$ is a real-valued pulse function. Under these assumptions, it is possible to show that :

- $|\rho_2(\tau)|$ and the modulus of the four true 4th-order correlation coefficient only depend on τ and $v(t)$ but do not depend on the symbol statistics.

- the modulus of the four apparent 4th-order correlation coefficient depend on τ, $v(t)$ and also on the 2th and 4th-order symbol statistics, which confirms the fact that the classical 4th-order cumulant estimator changes the 4th-order correlation of the linear modulations.

For example, if we choose the square pulse such that $v(t) = 1$ if $0 \le t < T$ and $v(t) = 0$ elsewhere, we find that $\rho_2(\tau)$ and the four true 4th-order correlation coefficients have the same modulus equal to $1 - |\tau|/T$. This implicates that $|\rho_{4,av}(\tau)| = |\rho_2(\tau)|$ and shows that generally a 2nd order correlation between paths generates also a 4th order correlation.

The previous results are illustrated at figure 1 which shows the variations of $|\rho_{4,av}(\tau)|$ and $|\rho_{4a,av}(\tau)|$ as a function of $|\rho_2(\tau)|$ for a BPSK and a QPSK modulation and for two pulse functions corresponding to the square and the half-Nyquist function with a roll-off of 0.25.

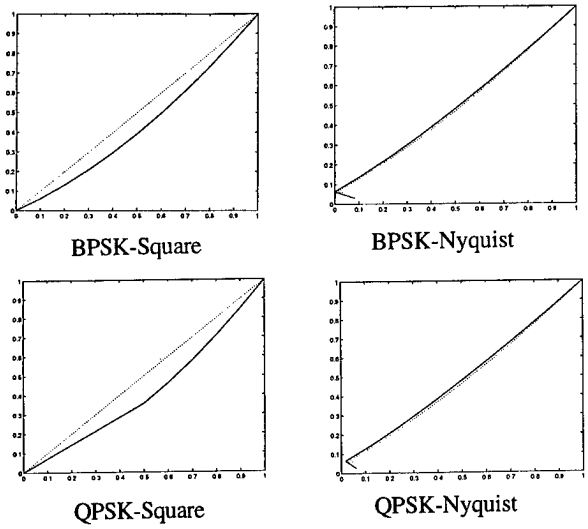


Fig. 1 - $|\rho_{4,av}(\tau)|$ *and* $|\rho_{4a,av}(\tau)|$ *as a function of* $|\rho_2(\tau)|$
--- : $|\rho_{4a}|$, ···· : $|\rho_4|$

5. EIGENSTRUCTURE OF Q_z AND Q_{za}

We analyse, in this section, the eigenstructure of the matrices Q_z and Q_{za} in the presence of correlated paths of (quasi)-cyclostationary independent sources.

5.1 Prewhitening of the data

The prewhitening stage of the data by the matrix $R_s^{-1/2}$ transforms (2.1) into (3.3). It is then possible to show that in the presence of correlated paths, the whitened steering vectors $\boldsymbol{a}'_{ik}$ are no longer orthonormalized. More precisely, it can be shown that :

- $\boldsymbol{a}'_{ik}$ is orthogonal to $\boldsymbol{a}'_{jl}$ if and only if $m'_{ik}(t)$ and $m'_{jl}(t)$ are not correlated
- $\boldsymbol{a}'_{ik}$ is normalized if and only if $m'_{ik}(t)$ is uncorrelated with all the $m'_{jl}(t)$ $(jl \neq ik)$

For example, in the case where $P = 1$ and $M_1 = 2$, noting ρ_2 the 2nd order correlation coefficient between $m'_{11}(t)$ and $m'_{12}(t)$, we find that

$$\boldsymbol{a}'^{\dagger}_{11}\boldsymbol{a}'_{11} = \boldsymbol{a}'^{\dagger}_{12}\boldsymbol{a}'_{12} = 1 / (1 - |\rho_2|^2) \qquad (5.1)$$

$$\boldsymbol{a}'^{\dagger}_{11}\boldsymbol{a}'_{12} = -\rho_2 / (1 - |\rho_2|^2) \qquad (5.2)$$

5.2 Eigenstructure of Q_z

The statistical independence of the P considered sources implicates that the Q_z matrix, defined by (3.4), can be written as

$$Q_z = \sum_{i=1}^{P} (A'_i \otimes A'^{*}_i)\, Q'_{mi}\, (A'_i \otimes A'^{*}_i)^{\dagger} \triangleq \sum_{i=1}^{P} Q_{zi} \qquad (5.3)$$

where A'_i is the $(M \times M_i)$ matrix of the sources steering vectors $\boldsymbol{a}'_{ik}$ $(1 \le k \le M_i)$ and Q'_{mi} is the temporal mean of the quadricovariance of the vector $\boldsymbol{m}'_i(t)$ which components are the $m'_{ik}(t)$ $(1 \le k \le M_i)$. The orthogonality of the vectors $\boldsymbol{a}'_{ik}$ and $\boldsymbol{a}'_{jl}$ for $i \neq j$ (section 5.1) implicates that for $1 \le i \le P$, the eigenvalues and eigenvectors of Q_{zi} are also eigenvalues and eigenvectors of Q_z. Consequently, the eigenvalues and eigenvectors of Q_z correspond to the reunion of the eigenvalues and eigenvectors of the matrices Q_{zi}. In other words, statistical independent sources contribute to the eigenstructure of Q_z without any interaction between themselves. The rank r of Q_z is then equal to the sum of the rank, r_i, of the matrices Q_{zi}.

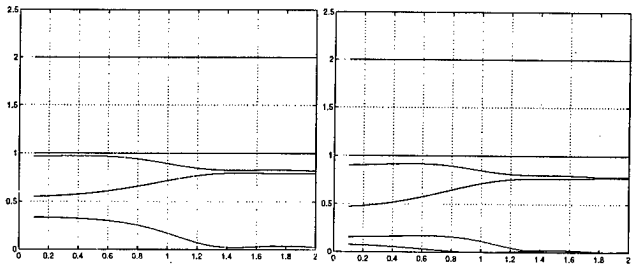
On the other hand, the rank of Q_{zi}, r_i, may vary between M_i (independent paths of the source i) and M_i^2 (all the paths of the source i are correlated to each other). However, it can be shown that for linear modulations, even when all the paths of the source i are correlated to each other, $r_i < M_i^2$. Besides, whatever the kind of modulation, it can be shown that the eigenvalues of Q_{zi} and thus those of Q_z do not depend on the mixture matrix A.

5.3 Eigenstructure of Q_{za}

The modification of the 4th-order correlation of the sources by the classical 4th-order cumulant estimators implicates that it may exists situations for which the apparent 4th-order cross-cumulant temporal mean of two statistically independent sources is not zero [13]. Consequently, Q_{za} may have a structure not similar to that described by (5.3) and the results of section 5.2 may not be applied for Q_{za}. However, in most practical cases the structure (5.3) still holds exactly or approximately for Q_{za}, with Q'_{mi} and Q_{zi} replaced by Q'_{mia} and Q_{zia}, and the results of section 5.2 can still be applied, despite of the fact that the apparent 4th-order autocumulants are no longer equal to the true ones. However note that for linear modulations, the rank of Q_{zia} is often equal to M_i^2 when all the paths of the source i are correlated to each other.

5.4 Illustrations

The figure 2 illustrates the previous results by showing the values of the non zero eigenvalues modulus of Q_z and Q_{za} in the presence of $P = 3$ independent sources (QPSK-Nyquist, QPSK-Square, BPSK-Square) with $M_1 = 2$, $M_2 = 1$ and $M_3 = 1$, for several values of τ/T, where τ is the relative time delay between the two paths of the source 1. The varations of τ does not modify the 2 highest values.



Q_z Q_{za}

Fig. 2 - *Eigenvalues modulus of Q_z and Q_{za}, for $P = 3$ with $M_1 = 2$, $M_2 = 1$ and $M_3 = 1$, as a function of τ/T*

6. BLIND IDENTIFICATION AND SOURCE SEPARATION

In the presence of one source with several correlated paths, the blind identification of the paths steering vectors cannot be done exactly since the whitened steering vectors are not orthogonal to each other (section 5.1). In this case, the blind estimates of the paths steering vectors are linear combination of the true ones with coefficients directly related to the 4th-order correlation between the paths. Consequently, the separation of correlated paths cannot be optimal but still occurs up to a relatively high level of 2nd-order correlation, depending on the Q_z matrix which is exploited (true or apparent), the modulation and the pulse function for linear modulations.

In the presence of several independent sources with their own paths, although it exists situations for which the separation of the different sources (not paths) fails, even from the exploitation of the true Q_z, in most practical cases, this separation occurs even from the use of Q_{za}.

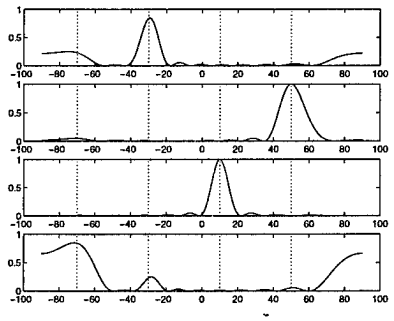


Fig. 3 - *Spatial correlation coefficient as a function of θ*

The Figure 3 illustrates the previous results by showing the spatial correlation coefficient between the 4 blind steering vector estimates and the array manifold of a ULA of 10 sensors as a function of θ for one QPSK-Nyquist source (ro 0.25) with two paths which DOA are $-70°$ and $-30°$ and such that $\tau/T = 0.4$, a QPSK-Square (30°) and a BPSK-Square (50°) with one path each. Note that Q_z and Q_{za} give, in that case, the same good estimation of the paths DOA by this DF method called Blind-Maxcor [14].

8. CONCLUSION

The behaviour of the classical indirect HO blind source separation methods has been analysed in the presence of correlated paths of several (quasi)-cyclostationary sources, through the analyses of the 4th-order correlation of digital modulation and the eigenstructure of the whitened quadricovariance. The choice of the 4th-order cumulant temporal mean estimator has been shown to be crutial in some cases. In most practical situations, the classical methods do not mix paths of independent sources and separate correlated paths up to a high 2nd order correlation.

REFERENCES

[1] C. JUTTEN, J. HERAULT, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture", *Signal Processing*, Vol 24, pp 1-10, 1991.

[2] J.F. CARDOSO, A. SOULOUMIAC, "Blind Beamforming for Non Gaussian Signals", *IEE Proc-F*, Vol 140, N°6, pp 362-370, Dec 1993.

[3] P. COMON, "Independent Component Analysis", *Signal Processing*, Vol 36, N°3, Special Issue On Higher Order Statistics, pp 287-314, Apr. 1994.

[4] L. DeLATHAUWER, B. DeMOOR, J. VANDEWALLE, "Independent Component Analysis based on Higher Order Statistics only ", *Proc. IEEE SP Workshop on SSAP*, Corfu (Greece), pp 356-359, June 1996.

[5] P. COMON, E. MOREAU, "Improved Contrast dedicated to Blind Separation in Communications", *Proc. ICASSP*, Munich (Germany), pp 3453-3456, Apr. 1997.

[6] P. CHEVALIER, "On the Performance of Higher Order Blind Source Separation Methods", *Proc. IEEE ATHOS Workshop on Higher Order Stat.*, Begur (Spain), pp 30-34, June 1995.

[7] P. CHEVALIER, "Méthodes Aveugles de Filtrage d'Antenne", *Revue d'Electronique et d'Electricité*, SEE, N°3, pp 48-58, Sept. 1995.

[8] P. COMON, P. CHEVALIER, V. CAPDEVIELLE, "Performance of Contrast-Based Blind Source Separation", *Proc. IEEE SP Workshop on SP Advances in Wireless Communications*, SPAWC, Paris (France), pp 345-349, April 1997.

[9] J.K. TUGNAIT, "On Blind Separation of Convolutive Mixture of Independent Linear Systems", *IEEE SP Workshop on SSAP*, Corfu , pp 312-315, June 1996.

[10] P. CHEVALIER, A. MAURICE, "Constrained Beamforming for Cyclostationary Signals", *Proc. ICASSP*, Munich (Germany), pp 3789-3792, Apr. 1997.

[11] A. FERREOL, P. CHEVALIER, "Higher Order Blind Source Separation using the Cyclostationarity Property of the Signals", *Proc. ICASSP*, Munich, pp 4061-4064, Apr. 1997.

[12] P. MARCHAND, D. BOITEAU, "Higher Order Statistics for QAM Signals : A Comparison between cyclic and stationary representations", *Proc. EUSIPCO*, Triestre (Italy), pp 1531-1534, Sept. 1996.

[13] P. CHEVALIER, A. FERREOL, "Limites des Estimateurs Classiques de Cumulant d'Ordre Quatre pour la Séparation Aveugle de Sources Cyclostationnaires", *Proc. GRETSI*, Grenoble (France), Sept. 1997.

[14] P. CHEVALIER, G. BENOIT, A. FERREOL, "DF after Blind Identification : Blind-Maxcor and Blind-MUSIC", *EUSIPCO*, Triestre (Italy), pp 2097-2100, Sept. 1996.

Towards non Symmetrical Optimal Source Separation Contrasts

Eric Moreau

MS-GESSY, ISITV, Université de Toulon et du Var
avenue Georges Pompidou, BP56
F-83162 La Valette du Var, Cedex, FRANCE
e-mail : moreau@isitv.univ-tln.fr

Abstract

We consider the problem of sources separation and particularly criteria based approaches. A generalized definition of contrast function is given in order to consider non symmetrical and/or non scale invariant functions. Two generalized contrasts involving high-order cumulants are proposed. In the case of two sources, we derive the optimal non symmetrical coefficient by minimizing a performance index. Finally computer simulations are presented in order to illustrate the results and to show the interest in considering a non symmetrical contrast.

1. Introduction

The source separation problem can be simply formulated as follows: several unknown linear mixtures of certain independent signals called sources are observed. The goal is to recover the original sources without knowing the mixing system. Hence this must be realized from the only observations and this is the reason why this problem is often qualified as "blind".

Among the great number of approaches that have been proposed in the recent literature, we are primary concerned with high-order statistics criteria based approaches, e.g. [1]-[9]. In this field, contrast functions constitute separation criteria in the sense that their maximization solve the source separation problem. As defined in [2], contrasts are imposed to be symmetrical and scale invariant functions. The global maximization of such a contrast is a necessary and sufficient condition for source separation. Even if this is a good thing, one can be interested in finding "only" sufficient conditions. This is the main purpose of this communication to show that we can take advantage of considering non symmetrical contrast according to a generalized definition. Notice that, even if it is not clearly said, non symmetrical criteria have been derived in [1] but with another approach and non scale invariant functions have been given in [7].

2. Notations and assumptions

The classical linear memoryless mixture model is considered. It reads

$$\boldsymbol{x}(k) = \boldsymbol{G}\boldsymbol{a}(k) \tag{1}$$

where $\boldsymbol{x}(k)$ is the $(N,1)$ vector of observations, $\boldsymbol{a}(t)$ the $(N,1)$ vector of statistically independent sources, k the discrete time and $\boldsymbol{G}$ the (N,N) *invertible* mixture matrix. The source separation problem consists in estimating a matrix $\boldsymbol{H}$ such that the $(N,1)$ vector

$$\boldsymbol{y}(k) = \boldsymbol{H}\boldsymbol{x}(k) \tag{2}$$

restores the N input sources $a_i(k), i \in \{1,\ldots,N\}$. We define the matrix $\boldsymbol{S}$ of the global system as

$$\boldsymbol{S} \triangleq \boldsymbol{H}\boldsymbol{G}\,, \tag{3}$$

hence according to (1) and (2)

$$\boldsymbol{y}(k) = \boldsymbol{S}\boldsymbol{a}(k)\,. \tag{4}$$

Because sources are assumed inobservable, there are some inherent indeterminations in their restitution. That is, in general, we cannot identify the power and the order of each sources. Hence they are said separated if and only if the global matrix reads

$$\boldsymbol{S} = \boldsymbol{D}\boldsymbol{P} \tag{5}$$

where $\boldsymbol{D}$ is an invertible diagonal matrix and $\boldsymbol{P}$ a permutation matrix.

The following assumptions are made

`A1a.` The sources $a_i(k), i \in \{1,\ldots,N\}$, are zero-mean, unit power and statistically mutually independent;

0-8186-8005-9/97 $10.00 © 1997 IEEE

A1b. $\boldsymbol{a}(k)$ is a random vector stationary up to order under consideration, i.e. $\forall i \in \{1,\ldots,N\}$, the cumulant Cum $(\underbrace{a_i(k),\ldots,a_i(k)}_{p\times},\underbrace{a_i^*(k),\ldots,a_i^*(k)}_{q\times})$ is an independent function of k, denoted by $\mathsf{C}_{p,q}a_i$;
A1c. Given p and q, the cumulants are assumed to satisfy one of the two following conditions:
c1. $|\mathsf{C}_{p,q}a_1| \geq \cdots \geq |\mathsf{C}_{p,q}a_N| > 0$;
c2. $|\mathsf{C}_{p,q}a_1| \geq \cdots \geq |\mathsf{C}_{p,q}a_{N-1}| > |\mathsf{C}_{p,q}a_N| = 0$.
In particular this means that at most one of the cumulants $\mathsf{C}_{p,q}a_i$, $i \in \{1,\ldots,N\}$ is zero.
A2. Global matrix $\boldsymbol{S}$ is such that $\boldsymbol{S}\boldsymbol{S}^H = \boldsymbol{S}^H\boldsymbol{S} = \boldsymbol{I}$.

The set of random vectors satisfying assumptions A1a to A1c is denoted by $\mathcal{A}$. The set of matrix satisfying assumption A2 is denoted by $\mathcal{U}$. The subset of $\mathcal{U}$ of matrices satisfying (5) is denoted by $\mathcal{P}$. Finally the set of random vectors $\boldsymbol{y}(k)$ satisfying (4) where $\boldsymbol{a}(k) \in \mathcal{A}$ and $\boldsymbol{S} \in \mathcal{U}$ is denoted by $\mathcal{Y}$.

3. Source separation criteria

We first recall the initial definition of a contrast [2]:

Definition 1 A contrast on $\mathcal{Y}$ is a multivariate mapping $\mathcal{I}$ from the set $\mathcal{Y}$ to $\mathbb{R}$ which satisfies the following three requirements:
*1. $\forall \boldsymbol{y} \in \mathcal{Y}, \forall \boldsymbol{S} \in \mathcal{P}, \mathcal{I}(\boldsymbol{S}\boldsymbol{y}) = \mathcal{I}(\boldsymbol{y})$;
*2. $\forall \boldsymbol{a} \in \mathcal{A}, \forall \boldsymbol{S} \in \mathcal{U}, \mathcal{I}(\boldsymbol{S}\boldsymbol{a}) \leq \mathcal{I}(\boldsymbol{a})$;
*3. $\forall \boldsymbol{a} \in \mathcal{A}, \forall \boldsymbol{S} \in \mathcal{U}, \mathcal{I}(\boldsymbol{S}\boldsymbol{a}) = \mathcal{I}(\boldsymbol{a}) \Leftrightarrow \boldsymbol{S} \in \mathcal{P}$.

Such contrasts are symmetrical and scale invariant functions (*1) which have to be maximized (*2) to get separation (*3). According to this definition, numerous contrasts have been proposed, see e.g. [6]. Now, in order to consider non symmetrical (and also non scale invariant) functions, we propose the following "generalized" definition of a contrast:

Definition 2 A contrast on $(\mathcal{Y}, \mathcal{P}_d)$ is a multivariate mapping $\mathcal{I}$ from the set $\mathcal{Y}$ to $\mathbb{R}$ which satisfies, $\forall \boldsymbol{a} \in \mathcal{A}$ and $\forall \boldsymbol{S} \in \mathcal{S}$, the following two requirements:
R1. $\mathcal{I}(\boldsymbol{S}\boldsymbol{a}) \leq \mathcal{I}(\boldsymbol{a})$;
R2. $\exists \mathcal{P}_d \subset \mathcal{P}, \mathcal{P}_d \neq \emptyset \;/\; \mathcal{I}(\boldsymbol{S}\boldsymbol{a}) = \mathcal{I}(\boldsymbol{a}) \Leftrightarrow \boldsymbol{S} \in \mathcal{P}_d$.

Such contrasts are not imposed to be, a priori, symmetrical or scale invariant. Moreover all the global maxima constitute a non empty subset of the set of separating matrices. Hence the maximization of the contrast as defined in definition 2 is no longer a necessary and sufficient condition for source separation but a sufficient one. It can also be noticed that contrasts in the sense of definition 1 are (fortunately) contrasts in the sense of definition 2. The converse is not, in general, true as exemplified in the following.

Define

$$\mathcal{I}^{\gamma}_{p,q}(\boldsymbol{y}) = \sum_{i=1}^{N} \gamma_i |\mathsf{C}_{p,q} y_i| \tag{6}$$

where the real numbers γ_i, $i \in \{1,\ldots,N\}$, are such that

$$\gamma_1 \geq \cdots \geq \gamma_N > 0\,.$$

We can now propose the following result

Proposition 1 If $p+q \geq 3$, the function $\mathcal{I}^{\gamma}_{p,q}(\boldsymbol{y})$ is a contrast in the sense of definition 2.

Because of lack of place, the proof is reported in a full-length paper [8]. Now let us consider the specific case of sources with identical sign ε_p of their (p,p) order cumulants, i.e. $\forall i$,

$$\mathrm{sgn}(\mathsf{C}_{p,p}a_i) = \varepsilon_p\,,$$

then we have the following result:

Proposition 2 If $p \geq 2$, the function

$$\mathcal{J}^{\gamma}_{p,p}(\boldsymbol{y}) = \varepsilon_p \sum_{i=1}^{N} \gamma_i \mathsf{C}_{p,p} y_i$$

is a contrast in the sense of definition 2.

4. Gradient based algorithm

4.1. Algorithm

In order to have A2, we consider that a first stage realizes a whitening of the observations. This "classical" stage will not be discussed here. The whiteness of $\boldsymbol{y}$ is then ensured if $\boldsymbol{H}$ is orthonormal. In order to find such an orthonormal matrix $\boldsymbol{H}$ which separates the sources, a gradient-based algorithm is proposed in order to maximize one of the contrasts. For this task we use a parametrization of $\boldsymbol{H}$ thanks to planar (Givens) rotations. For the sake of simplicity, we only consider the case of $N = 2$ *real sources* and contrast $\mathcal{J}^{\gamma}_{2,2}$ which will be denoted $\mathcal{J}$ in the following. Hence $\boldsymbol{H}$ is parametrized thanks to one angle:

$$\boldsymbol{H} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

and the gradient based updating rule reads

$$\theta(n) = \theta(n-1) + \mu' \frac{\partial \mathcal{J}}{\partial \theta} \tag{7}$$

where we have, after some algebra,

$$\frac{\partial \mathcal{J}}{\partial \theta} = 4\varepsilon_2(\gamma_1 \mathsf{E}[y_1^3 y_2] - \gamma_2 \mathsf{E}[y_1 y_2^3])\,. \tag{8}$$

In practice we use a stochastic version of (7) by dropping the expectation operator in (8), hence

$$\theta(n) = \theta(n-1) + \mu\varepsilon_2 y_1(n) y_2(n)(y_1^2(n) - \delta y_2^2(n)) \quad (9)$$

where

$$\mu = 4\gamma_1\mu' \quad \text{and} \quad \delta = \gamma_2/\gamma_1 \,.$$

4.2. Convergence analysis

Let us consider the deviation $v = \theta - \tilde{\theta}$ where $\tilde{\theta}$ denotes the true value of parameter in order to have $y_1 = a_1$ and $y_2 = a_2$. Using $\theta(n) = v(n) + \tilde{\theta}$, adaptation (9), writen in terms of deviation, can be approximated (at order one) by the following linear recursive equation

$$v(n) = (1 - \mu A(n))v(n-1) + \mu z(n) \quad (10)$$

where

$$z(n) = \varepsilon_2(a_1^3(n)a_2(n) - \delta a_1(n)a_2^3(n)) \,; \quad (11)$$

$$A(n) = \varepsilon_2(a_1^4(n) + \delta a_2^4(n) - 3(1+\delta)a_1^2(n)a_2^2(n)) \,. \quad (12)$$

Assuming that each sources are independent and identically distributed random sequences, then it is easily seen that the mean of $v(n)$ converge to zero if $0 < \mu < 2/\mathsf{E}A$ where

$$\mathsf{E}A = |\mathsf{C}_{2,2}a_1| + \delta|\mathsf{C}_{2,2}a_2| \,. \quad (13)$$

Now in the mean square, denoting $V(n) = \mathsf{E}v^2(n)$, we have

$$V(n) = (1 - \mu\lambda)V(n-1) + \mu^2\mathsf{E}z^2 \quad (14)$$

where

$$\lambda = 2\mathsf{E}A - \mu\mathsf{E}A^2$$

and

$$\mathsf{E}z^2 = \mathsf{E}a_1^6 + \delta^2\mathsf{E}a_2^6 - 2\delta\mathsf{E}a_1^4\mathsf{E}a_2^4 \quad (15)$$

It is easily found that $V(n)$ converge to

$$\mathsf{E}v^2 = \mu\frac{\mathsf{E}z^2}{\lambda} \quad (16)$$

if $0 < \mu < 2\mathsf{E}A/\mathsf{E}A^2$.

4.3. Performance of the algorithm

In order to take into account both the convergence speed and the mean square error, we define a performance index Q as

$$Q = \frac{\mathsf{E}v^2}{\mu\lambda} \quad (17)$$

where $\mu\lambda$ charaterizes the dB wins per iterations. Now for small enough μ, we have $\lambda \approx 2\mathsf{E}A$ and thanks to (16)

$$Q = \frac{\mathsf{E}z^2}{4(\mathsf{E}A)^2} \,. \quad (18)$$

This index depends on the statistics of the sources and on the free parameter δ. One can ask if there exists a minimum for Q w.r.t. δ. One can easily obtain the value of δ denoted δ_o such that $\partial Q/\partial\delta = 0$:

$$\delta_o = \frac{\mathsf{E}a_1^6|\mathsf{C}_{2,2}a_2| + \mathsf{E}a_1^4\mathsf{E}a_2^4|\mathsf{C}_{2,2}a_1|}{\mathsf{E}a_2^6|\mathsf{C}_{2,2}a_1| + \mathsf{E}a_1^4\mathsf{E}a_2^4|\mathsf{C}_{2,2}a_2|} \,. \quad (19)$$

It can also be shown that the second derivative of Q w.r.t. δ for $\delta = \delta_o$ is positive and thus this is a minimum for Q.

Given sources, if $\delta_o \leq 1$ then it can be used in the adaptation (9) in order to get the best performances in the sense that Q is minimum.

For example, if the two sources have the same statistics then $\delta_o = 1$. In that case, there is no interest in considering a non symmetrical contrast. Now if the second source (a_2) is Gaussian, i.e. $\mathsf{E}a_2^4 = 3$ and $\mathsf{E}a_2^6 = 15$, then $\delta_o = \mathsf{E}a_1^4/5$ and if the first source a_1 is binary then $\mathsf{E}a_1^4 = 1$ and $\delta_o = 1/5 = 0.2$. It can be noticed that there exists some cases where $\delta_o > 1$ (when $\mathsf{E}a_1^4 > 5$ in the previous case) and thus, sometimes, it occurs that no optimal contrast can be built from Q.

5. Computer simulations

In order to illustrate the hereabove results, some computer simulations are presented in the case of two sources ($N = 2$). Convergence of the algorithm based on adaptation (9) are illustrated thanks to an index defined on the global matrix $\boldsymbol{S}$ according to

$$\mathrm{ind}_\alpha(\boldsymbol{S}) \triangleq \frac{1}{2}\left[\sum_i\left(\sum_j \frac{|s_{ij}|^\alpha}{\max_\ell |s_{i\ell}|^\alpha} - 1\right) + \sum_j\left(\sum_i \frac{|s_{ij}|^\alpha}{\max_\ell |s_{\ell j}|^\alpha} - 1\right)\right] \quad (20)$$

where $\alpha \geq 1$. This positive index is indeed zero if $\boldsymbol{S}$ satisfies (5) and a small value indicates the proximity to the desired solutions. The mixing matrix is taken orthonormal such that the prewhitening stage is not necessary

$$\boldsymbol{G} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix}, \qquad \theta_1 = 20\frac{\pi}{180} \,. \quad (21)$$

In all the simulations the index ind_2 is considered.

Three cases are presented: i) one binary source (i.e. source taking the two values ± 1 with equal probability) and one Gaussian source; ii) one 4-PAM source (i.e. source taking the four values $\pm 3/\sqrt{5}$, $\pm 1/\sqrt{5}$ with equal probability) and one Gaussian source and finally iii) one binary source and one 4-PAM source. We plot the evolution of $\theta(n)$ and of the index of convergence w.r.t. iterations. They are compared to those of the classical symmetrical case (i.e. $\delta = 1$).

In all cases, Fig.1 and 2 for case i), Fig.3 and 4 for case ii) and Fig.5 and 6 for case iii), one can see an advantage in considering a non symmetrical contrast.

Acknowledgment: The author would like to thank Dr. N. Thirion for helpful comments and insights.

References

[1] J.F. Cardoso, S. Bose and B. Friedlander, "On Optimal Source Separation Based on Second and Fourth Order Cumulants", in Proc. *IEEE SP Workshop on SSAP*, Corfu, Greece, pp 198-201, June 1996.

[2] P. Comon, "Independent Component Analysis, a New Concept?", *Signal Processing*, Vol. 36, pp 287-314, 1994.

[3] N. Delfosse and P. Loubaton, "Adaptive Blind Separation of Independent Sources: a Deflation Approach", *Signal Processing*, Vol. 45, pp 59-83, 1995.

[4] Y. Inouye and T. Sato, "Unconstrained Optimization Criteria for Blind Equalization of Multichannel Linear Systems", in Proc. *IEEE SP Workshop on SSAP*, Corfu, Greece, pp 320-323, June 1996.

[5] A. Mansour and C. Jutten, "Fourth-Order Criteria for Blind Source Separation", *IEEE Trans. on Signal Processing*, Vol. 43, pp 2022-2025, August 1995.

[6] E. Moreau and J.-C. Pesquet, "Independence/Decorrelation Measures with Applications to Optimized Orthonormal Representations", to appear in Proc. *ICASSP'97*, Munich, Germany, April 1997.

[7] E. Moreau and N. Thirion, "Multichannel Blind Signal Deconvolution Using High Order Statistics", in Proc. *IEEE SP Workshop on SSAP*, Corfu, Greece, pp 336-339, June 1996.

[8] E. Moreau and N. Thirion, "Non Symmetrical Contrasts for Sources Separation", in Preparation.

[9] J.K. Tugnait, "On Blind Separation of Convolutive Mixtures of Independent Linear Systems", in Proc. *IEEE SP Workshop on SSAP*, Corfu, Greece, pp 312-315, June 1996.

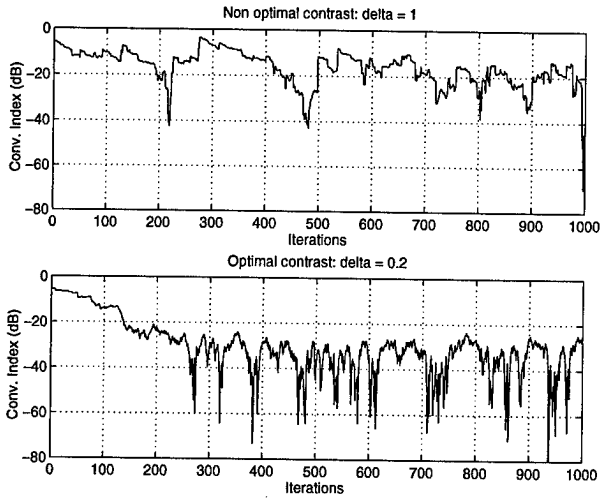


Figure 1. Evolution of the convergence index with the classical contrast and the non symmetrical optimal one in case i).

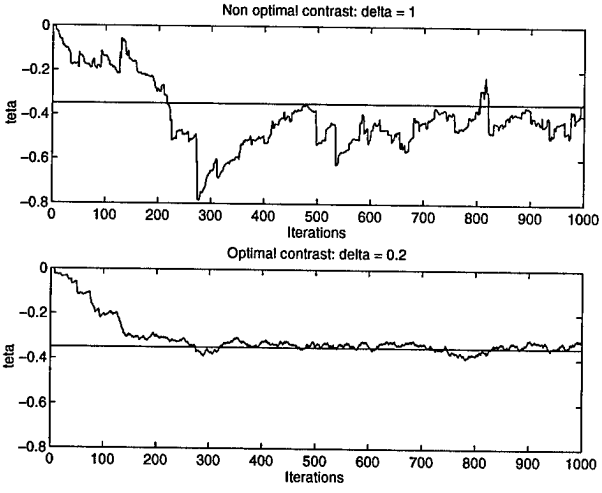


Figure 2. Evolution of the angle of the estimated rotation with the classical contrast and the non symmetrical optimal one in case i).

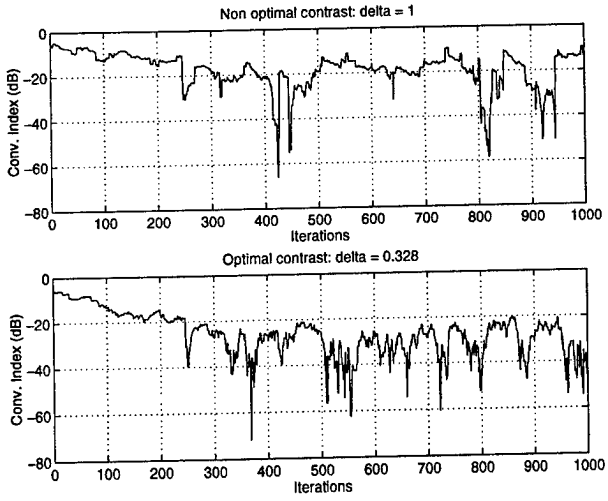


Figure 3. Evolution of the convergence index with the classical contrast and the non symmetrical optimal one in case ii).

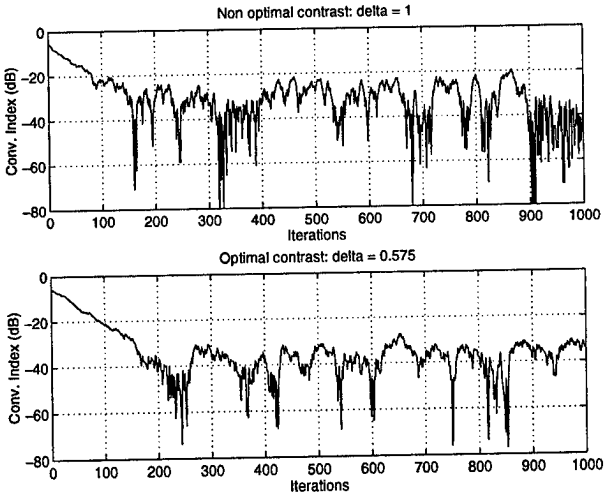


Figure 5. Evolution of the convergence index with the classical contrast and the non symmetrical optimal one in case iii).

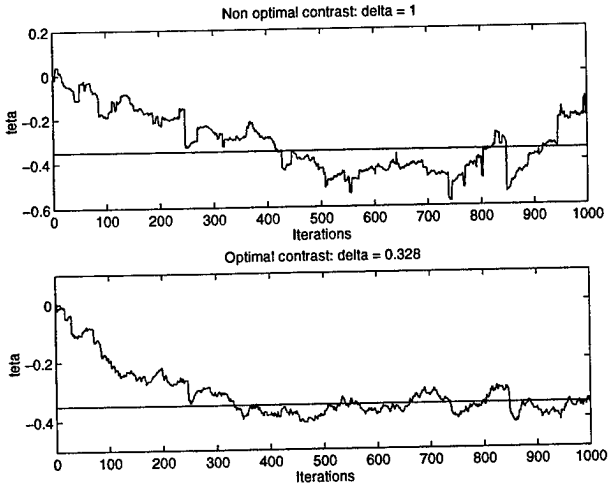


Figure 4. Evolution of the angle of the estimated rotation with the classical contrast and the non symmetrical optimal one in case ii).

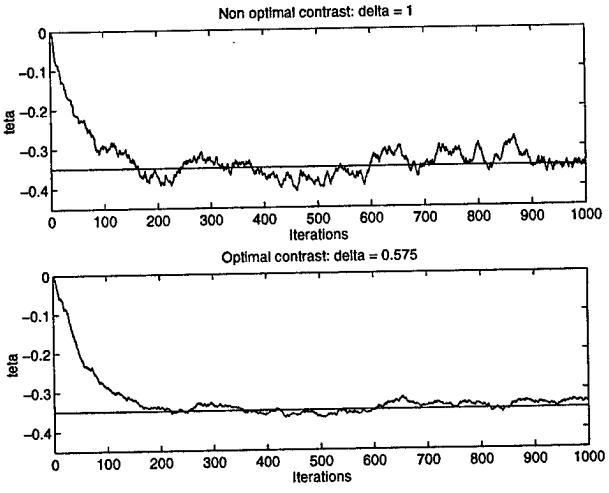


Figure 6. Evolution of the angle of the estimated rotation with the classical contrast and the non symmetrical optimal one in case iii).

Narrow Band Source Separation in Wide Band Context Applications to Array Signal Processing

Jérôme GALY[1], Claude ADNET[2] and Eric CHAUMETTE[2]

[1] E.N.S.I.C.A, 1 place Emile Blouin
31056 Toulouse Cedex (France)
galy@ensica.fr

[2] THOMSON-CSF AIRSYS, 7 rue des mathurins
92223 Bagneux Cedex (France)
Tel : 01-40-84-21-86 adnet@airsys.thomson.fr

ABSTRACT

Blind source separation is now a well known problem. Various methods have been proposed for instantaneous and convolutive mixtures of sources. Conventional antenna array processing techniques are based on the use of second order statistics but rest on restrictive assumptions. Thus, when a priori informations about the propagation or the geometry of the array are not available, the model can be generalized to a blind sources separation model. It supposes the statistical independence of the sources and their non-gaussianity.
In this paper, we focus on the narrow band source separation problem embedded in wide band jammers. We show that JADE algorithm [1] made for instantaneous mixture is still valid in a wide band context where only the signals of interest are narrow-band. We also prove that a wide band signal tends to occupy all the degrees of freedom of the covariance matrix and modifies the signal subspace dimension.

1 PROBLEM STATEMENT

In this paper, we address the narrow band source separation problem embedded in wide band jammers. The classical approach consists in employing a noise reduction algorithm based on second-order statistics of the received signals and using a priori information on signals, wave propagation or antenna.
Nevertheless, in the case of independent signals and jammers, one can resort to a more original approach which relies on higher order statistics to achieve a 'blind' source separation.

This study is supported by STSIE (Service Technique des Systèmes d'Information et de l'Electronique).

We consider an array of m sensors and n sources (denoted s(t)), the array output denoted x(t) is corrupted by independent additive gaussian noise (n(t)) :

$$x(t) = As(t) + n(t). \quad (1)$$

where A is an m x n complex matrix.

Blind source separation assumes no a priori information on A and consists in finding the initial signals $s_i(t)$ from the observations $x_i(t)$ of the mixture (1).
In the case of instantaneous mixtures, among various blind source separation methods, available in literature [3][4][5], Cardoso and Souloumiac give us an algorithm, called JADE [1], asymptotically optimal [2][6] and of less implementation complexity.
This technique for blind source separation can be developed in two steps : normalization by whitening the signal part (using 2-order statistics) and joint diagonalization of a set of eigenmatrices (using higher order statistics).
We know that if (A_0, s_0, n_0) is a solution of (1) then

$$(A_0 \Lambda P, P^{+} \Lambda^{-1} s_0, n_0)$$

is also solution of (1) where Λ is a diagonal matrix and P a permutation matrix.
This source separation technique has been used for interference cancellation (synthesis in [7]).

2 WIDE BAND CONTEXT

We first deal with the bandwidth influence on covariance matrix eigenvalues. Thus, the realistic simulation of the wave propagation through an array of sensors allows to take into account the wide band effects on the spatial-temporal structure of the signals.

0-8186-8005-9/97 $10.00 © 1997 IEEE

The narrow band model permits to liken the signals propagation delays between sensors to simple phase shift independent of the frequency.
In the case of wide band signals, we must take into account the bandwidth influence on the signal subspace dimension. We can find in literature [8],[9], [10] some algorithms specific to wide band signals or convolutive mixtures.
We want to show in this paper the JADE algorithm behavior in a wide band context.
Many simulations have been proposed using different signals (BPSK, QPSK, gaussian jammers...) with variable frequencies : we give some interesting results.

3 BANDWIDTH INFLUENCE

The spatial covariance matrix of x(t) can be denoted :

$$R_x = R_s + \sigma^2 I$$

where

$$R_s = \int_{\nu_0 - B/2}^{\nu_0 + B/2} \Phi(\nu)\Phi^H(\nu)d\nu$$

where B is the signal bandwidth, ν_0 the carrier frequency and $\Phi(\nu)$ the steering vector of the signal.
We can express $\Phi(\nu)$ near ν_0 with a Taylor series expansion and after, a simple variable change ($\nu = \nu - \nu_0$) we have :

$$\Phi(\nu+\nu_0) = \Phi(\nu_0) + \nu\dot{\Phi}(\nu_0) + \frac{\nu^2}{2}\ddot{\Phi}(\nu_0) + \frac{\nu^3}{3!}\dddot{\Phi}(\nu_0)$$

$$= \sum_{n=0}^{N-1} \frac{\nu^n}{n!}\overset{(n)}{\Phi}(\nu_0)$$

where $\dot{\Phi}(\nu)$ is the derivative of $\Phi(\nu)$ with respect to ν. So we can say that :

$$\Phi(\nu+\nu_0)\Phi^H(\nu+\nu_0) = \sum_{\substack{n=0\\m=0}}^{N-1} \frac{\nu^{n+m}}{n!m!}\overset{(n)}{\Phi}(\nu_0)\overset{(m)H}{\Phi}(\nu_0)$$

$$\text{So : } R_s = \int_{-B/2}^{B/2} \sum_{\substack{n=0\\m=0}}^{N-1} \frac{\nu^{n+m}}{n!m!}\overset{(n)}{\Phi}(\nu_0)\overset{(m)H}{\Phi}(\nu_0)d\nu$$

$$R_s = \sum_{\substack{n=0\\m=0}}^{N-1} \frac{1}{n!m!}\overset{(n)}{\Phi}(\nu_0)\overset{(m)H}{\Phi}(\nu_0)\int_{-B/2}^{B/2}\nu^{n+m}d\nu$$

We can see that only the odd terms (n+m+1) are non zero. The spatial covariance matrix of x(t) can be written as :

$$R_s = B\Phi(\nu_0)\Phi^H(\nu_0) + \frac{2B^3}{24}\dot{\Phi}(\nu_0)\dot{\Phi}^H(\nu_0)$$

$$+\frac{B^3}{24}\left[\Phi(\nu_0)\ddot{\Phi}^H(\nu_0) + \ddot{\Phi}(\nu_0)\Phi^H(\nu_0)\right]$$

if we stop the expansion at the third order.
In the case of an uniform rectilinear array with equidistant sensors, we can give the expression of the steering vectors :

$$\Phi(\nu) = (1, \Phi_1, \Phi_2, \ldots, \Phi_{N-1})^T$$

with $\Phi_k = e^{j2\pi(k-1)\frac{d\sin\theta}{c}\upsilon}$, c is the light propagation velocity, d the distance between two adjacent sensors and θ the wave-front direction.

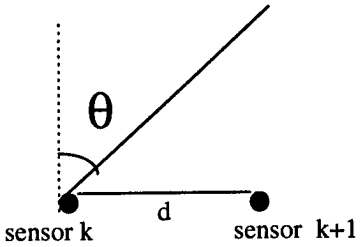


Figure 1 : Array of sensors

In this context, we can see that :

$$\dot{\Phi}_k = j2\pi(k-1)\frac{d\sin\theta}{c}\Phi_k$$

so :

$$\dot{\Phi} = \tilde{D}\Phi$$

where $\tilde{D}$ is a diagonal matrix. We can also write :

$$\ddot{\Phi} = \tilde{D}^2\Phi$$

We can now diagonalize the covariance matrix and find the eigenvalues

$$\lambda_1 = N - \frac{1}{36}B^2\frac{N(N+1)}{N-1}\left(\frac{\pi d\sin\theta}{c}\right)^2$$

and

$$\lambda_2 = \frac{1}{36}B^2\frac{N(N+1)}{N-1}\left(\frac{\pi d\sin\theta}{c}\right)^2$$

We can find another eigenvalue but if we stop the computation at the second order, two eigenvalues only appear.

If we consider the apparent width of the antenna considering the source called D_a, ($D_a = (N-1)d\sin\theta$ for an uniform rectilinear array) the eigenvalues become

$$\lambda_1 = N - \frac{1}{36}\frac{N(N+1)}{(N-1)^3}\left(\frac{\pi BD_a}{c}\right)^2$$

and

$$\lambda_2 = \frac{1}{36}\frac{N(N+1)}{(N-1)^3}\left(\frac{\pi BD_a}{c}\right)^2$$

In the case of an uniform circular array with equidistant sensors, we have for the eigenvalues of the covariance matrix :

$$\lambda_1 = N$$

$$\lambda_2 = \frac{N}{6}\left(\frac{\pi BD_a}{c}\right)^2$$

$$\lambda_3 = \frac{N}{320}\left(\frac{\pi BD_a}{c}\right)^4$$

To illustrate the bandwidth influence, we choose an array with eight sensors and we are interested in eigenvalues of the covariance matrix.
A wide-band signal occupies three eigenvalues in the covariance matrix (figure 2).

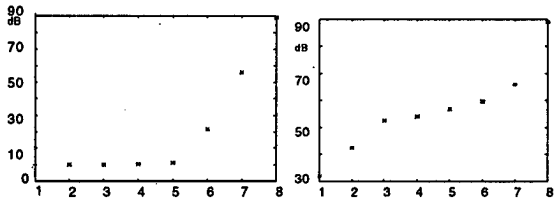


Figure 2

Figure 3

Eigenvalues

If we increase the bandwidth of the signal, it tends to occupy all the degree of freedom of R_s (figure 3)
In this context, we want to know if it's possible to separate a narrow band signal from a mixture composed by wide band jammers.

4 JADE IN WIDE-BAND CONTEXT

The first step of the JADE algorithm is the normalization by whitening the signal part using 2-order statistics.
When a signal occupies more than one freedom degree, it is very difficult to estimate the signal subspace dimension.
For example, on figure 2 we do not know if we have only one signal witch occupies three freedom degrees or three narrow band signals.
The equation (1) is : $x(t) = As(t) + b(t)$
where A is the mixing matrix of signals of interest and nuisance signals.

If we consider that A can be decomposed in :

$$A = U\Sigma_s^{-1/2}\Pi$$

with U and Π unitary matrices.

If $E[s(t)s(t)^H] = I$ then the covariance matrix of the output array is given by :

$$R_x = E[x(t)x(t)^H] = U\Sigma_s U + \sigma^2 I$$

whitening : we can decompose R_x in eigen-vectors and eigenvalues :

$$R_x = [U \quad V]\begin{bmatrix}\Sigma_s + \sigma^2 I & 0\\ 0 & \sigma^2 I\end{bmatrix}[U \quad V]^H$$

We can see that :

$$\Sigma_s^{-1/2}U^H x(t) = \Pi s(t) + \Sigma_s^{-1/2}U^H b(t)$$

The whitening step is directly liken to the estimation of the signal subspace dimension.
An error on this estimation will prevent JADE from separating the signals.
Without thermic noise , we can find $\hat{s}(t) = \Pi s(t)$.
Thus we can determine s(t) up to unitary matrix.

estimation of $\hat{s}(t)$: with the normalized problem, the identification of mixing matrix consists only in estimating a unitary matrix with fourth order statistics of whiten signals.
For JADE, the unitary matrix maximize the criterion :

$$\sum_{i,k,l=1}^{M} |Cum(x_i, x_i^*, x_k, x_l^*)|^2$$

where Cum represents the fourth order cumulant in our special case.
Two hypothesis are required for the estimation of this matrix, the signals $s_i(t)$ must be independent and non-gaussians (we can separate at most one gaussian signal).
So Cardoso and Souloumiac proposed [3] a unitary transform of whiten signals with joint diagonalization of a set of eigen-matrices.

5 EXPERIMENTAL RESULTS

We saw that a wide band signal tends to occupy all the degrees of freedom of the covariance matrix and to modify the signal subspace dimension.
Moreover, the realistic simulation of the wave propagation through an array of sensors allows to take into account the wide band effects on the spatial-temporal structure of the signals.

We consider now various examples for simulation.

The first case is the mixture of a narrow-band signal (BPSK (figure 4a1)) and two wide band jammers (a BPSK jammer (figure 4a2) and a gaussian jammer (figure 4a3)).

We can imagine that these signals occupy three degrees of freedom of the covariance matrix. But in fact, after the mixture on a array antenna, these signals occupy 7 degrees of freedom of the covariance matrix (figure 4c).

If we only take three for the signal subspace dimension in the whitening, the second stage of the separation (joint diagonalization) is not possible.
In fact, the first stage of the JADE algorithm must take into account this signal subspace dimension.
In our case, we normalize on a dimension of order six.
The figure 4b shows the 6 beams formed by the JADE algorithm and the separation of the narrow band signal from the other components of the initial mixture.

6 CONCLUSION

This paper shows that it is possible to extend the separation abilities of blind source separation algorithms initially designed for instantaneous mixture to a more general class of problems where only the signals of interest are narrow-band.
In this context, we adapted a specific algorithm called JADE to some cases where the jammers are wide-band.
We saw on simulations that this algorithm is perfectly fitted to array antenna.

REFERENCES

[1] J.F. CARDOSO, A. SOULOUMIAC, "An efficient technique for blind separation of complex sources." in Proc IEEE SP Workshop on Higher-Order Stat, Lake Tahoe, pp 275-279, 1993.

[2] A. SOULOUMIAC, J.F. CARDOSO, "Performances en séparation de sources." in Actes du quatorzième colloque GRETSI, pp 321-324, Juan-Les-Pins, France, 13-16 Septembre 1993.

[3] A. SOULOUMIAC, "Utilisation des statistiques d'ordre supèrieur pour le filtrage et la séparation de sources en traitement d'antenne." Thèse de doctorat de l'ENST, Février 1993.

[4] P. COMON, "Independent component analysis, a new concept ?" Signal Processing, vol. 36, n° 3, pp 287-314, 1994.

[5] M. GAETA, J.L. LACOUME, "Source separation versus hypothesis." in Proc Int. Workshop on Higher-Order Stat, Chamrousse, France pp 271-274, 1991.

[6] E. CHAUMETTE, P. COMON, D. MULLER, "Application of ICA to airport surveillance." in Proc IEEE SP Workshop on Higher-Order Stat, Lake Tahoe, pp 210-214, 1993.

[7] J. GALY, C. ADNET, E. CHAUMETTE "Blind Methods for Interference Cancellation in Array Processing." in Proc DSP 1997, Santorini, Greece.

[8] V. CAPDEVIELLE, C. SERVIERE, J. L. LACOUME, "Separation of wide band sources." in Proc IEEE SP/ATHOS Workshop on Higher-Order Stat, Girona, Spain, June 1995, pp 66-70.

[9] B. EMILE, "Estimation de temps de retard." in Ecole des techniques avancées Signal, Image, Parole "De l'ordre 2 aux ordres supèrieurs en traitement du signal." Grenoble 2-6 Septembre 1996.

[10] N. DELFOSSE, "Séparation aveugle adaptative de mélanges de sources indépendantes." Thèse de doctorat de l'ENST, Décembre 1995.

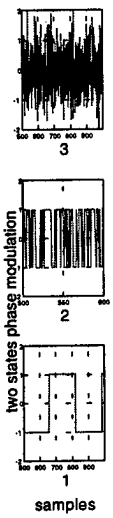


Figure 4a
Source and jammer signals

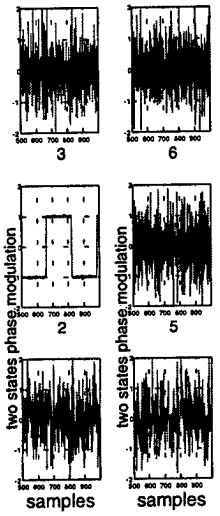


Figure 4b
6 beams with JADE

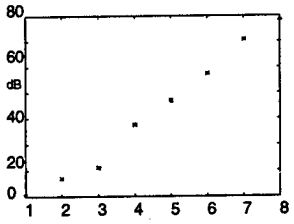


Figure 4c : Eigenvalues

PLENARY

Cumulant Tensors

Dr. Pierre Comon, *EURECOM (FRANCE)*

Cumulants of multidimensional random variables satisfy the properties that allow them to be considered as tensors. Most cumulant-based signal processing algorithms actually use slices of those tensors, mainly because numerical algorithms available today are only able to manipulate matrices. In addition, very few works in the literature have addressed the problem of decomposing or factorizing tensors. This would be a discrepancy if the problem was not so difficult, as emphasized in the talk. However, it is still possible to establish some links with the Eigenvalue decomposition, the congruent diagonalization, or the Cholesky factorization of matrices. But striking differences can be pointed out. It is hoped that these first basic statements will motivate further developments of tensor-based algorithms.

WAP: Detection, Estimation, Modeling & Analysis of Non-Gaussian Signals

Chair: Karl-Dirk Kammeyer
University of Bremen, Germany

ON REPRESENTATION FOR NONGAUSSIAN ARMA PROCESSES

K.Chandrasekhar and Shiv Dutt Joshi
Department of Electrical Engineering
Indian Institute of Technology
New Delhi-110016
INDIA.
E-Mail:ksekhar,sdjoshi@ee.iitd.ernet.in
sekhar@recw.ernet.in

ABSTRACT

A Generalised Predictor Space representation (of nonlinearity order two and memory M) for nonGaussian and nonminimum phase ARMA processes is proposed here, by expanding the underlying Hilbert space of finite L_2 norm random variables, which is now composed of linear combinations of linear as well as second order nonlinear terms of the process samples. Here the higher order statistical information enters into the picture in a natural way through the nonlinear terms. It is expected that the geometrical structure provided by the proposed predictor space would simplify the modeling of these processes. A set of new innovation vectors is defined on this space. Some of the properties of the new space are presented. The finite dimensionality of the proposed predictor space, when the underlying process admits a nonGaussian and non minimum phase ARMA representation is proved. The application of the proposed theory to estimate nonGaussian and nonminimum phase ARMA process parameters is also discussed.

1. Introduction:

In general Higher Order Statistics[HOS], in conjunction with the second order statistics provide additional information regarding statistical properties of random processes and parameters of the systems, to what can be obtained only from the latter. Thus they provide us with more accurate characterization of random signals and systems. An important motivation behind the use of HOS is based on the fact that the second order statistics yield only the spectrally equivalent minimum phase factors of a nonminimum phase process. On the other hand, HOS yield the true phase of the process upto a linear phase term [4]. In addition the use of HOS can potentially improve the estimation accuracy, i.e., reduce the variance of the estimated parameters. Also it is known that,when the underlying nonGaussian ARMA process is nonminimum phase, the optimal predictor in the mean square sense is not linear, but a nonlinear form of the past data. Here we present a generalised predictor space representation [of nonlinearity order q=2 and memory N], which is constructed by expanding the Hilbert space considered previously for the Gaussian processes by including the nonlinear terms also to take care of the nonGaussianity of the process.

The organisation of the paper is as follows. Section 1 presents a brief introduction to the general area of HOS and also the motivation behind the proposed representation. Section 2 presents the basic definition of the generalised predictor space of nonlinearity order two and memory N. A new innovation vector is defined on the proposed space. Some of the properties of the innovation vectors are also presented. This also contains a discussion on p-step ahead predictor in terms of the past innovations and also on the similarity of the proposed representation with the well known Wold's decomposition. In section 3, we prove the finite dimensionality of the proposed space, when the process admits nonGaussian ARMA representation. The implications of the proposed predictor space approach and its application to nonGaussian ARMA parameter estimation are discussed in section 4. Section 5 contains the conclusions drawn from the results obtained.

2. Preliminaries:

The basic definition of the generalised predictor space of nonlinearity order two and memory N is as follows:

.0-8186-8005-9/97 $10.00 © 1997 IEEE

Let $\{ y(t), t \in Z \}$ be a zero mean stationary random sequence with finite statistics upto fourth order and y(t) is not necessarily Gaussian, i.e.,

$E[|y^i(t)|] < \infty, \quad i = 2,3,4.$ (1)

Let us define some of the Hilbert spaces associated with this nonGaussian process and are of interest to us as follows:

$Y_t^+ = Closure\ Span\ \{ 1, y(t+i), \quad i \geq 0$
$y(t+i)y(t+j), i,j \geq 0, i-N \leq j \leq i \}$ (2)

$Y_t^- = Closure\ Span\ \{ 1, y(t+i), i < 0$
$y(t+i)y(t+j), i,j < 0, i-N \leq j \leq i \}$ (3)

Notice that Y_t^+ and Y_t^- are the generalizations of Hilbert spaces corresponding to the closure span of linear future and linear past terms respectively, associated with the Gaussian case [1],[7]. The generalised predictor space [of memory N and nonlinearity order q=2] is defined as the Hilbert space generated by projecting the space Y_t^+ onto Y_t^- where $(\cdot|\cdot)$ defines projection.

$P_t = Y_t^+ / Y_t^- = Span\ \{ z / Y_t^- ; z \in Y_t^+ \}$ (4)

Assume that the inner product defined on the Hilbert space for random variables X and Y is E[XY]. Let

$y(t) \triangleq \{ y(t), y^2(t), y(t)y(t-1), \ldots\ldots, y(t)y(t-N) \}$

$y_t \triangleq$ *The space spanned by the set of process vectors at time t,* $y(t)$.

$z_t \triangleq$ *The space spanned by the set of innovation vectors* $z(t)$.

$z(t) \triangleq y(t) - y(t) / Y_t^-$ (5)

(Because of '1' $z(t)$ *becomes a zero mean vector.)*

$$
\begin{aligned}
z(t-i) \triangleq\ & z(t-i) && y(t-i) && - y(t-i) / Y_{t-i}^- \\
& z(t-i,t-i) && y(t-i)y(t-i) && - y(t-i)\, y(t-i) / Y_{t-i}^- \\
& z(t-i,t-i-1) && y(t-i)y(t-i-1)) && - y(t-i)\, y(t-i-1) / Y_{t-i}^- \\
& \ldots\ldots && \ldots && \ldots\ldots \\
& z(t-i,t-i-N) && y(t-i)y(t-i-N) && - y(t-i)\, y(t-i-N) / Y_{t-i}^-
\end{aligned}
\quad (6)
$$

2.1 Decomposition of the observation space:

Theorem 1: $Y_t^- = z_{t-1} \oplus z_{t-2} \oplus \ldots\ldots \oplus z_{t-n} \oplus Y_{t-n}^-$ (7)

Proof: By induction on n, it is easy to prove this by using the orthogonal decomposition theorem.

Corollary 1.1: The innovation vectors are mutually orthogonal, i.e.,

$Span\{z(t-i)\} \perp Span\{z(t-j)\}, \forall\ i \neq j$ (8)

Proof: Follows from theorem (1).

Corollary 1.2: $z(t)$ is orthogonal to the projection of $y(t-i)$ onto Y_{t-j}^- for all $j > 0$

$i.e., Span\{ z(t) \} \perp y(t-i) / Y_{t-j}^-, \forall\ j > 0.$ (9)

Proof: The proof follows from the fact that $z(t)$ is orthogonal to the space Y_t^- : and Y_{t-j}^- is a subspace of Y_t^- for all $j > 0$ ■

2.2. Discussion:

Theorem 1 and its corollaries have a number of interesting implications. These can be summarised as follows:

(a) The linear space Y_t^- of all past observations and their nonlinear combinations can be orthogonally decomposed into two subspaces, namely, a subspace spanned by n immediately past innovations vectors and a subspace spanned by all observations upto n steps (i.e. Y_{t-n}^-).

(b) Since $y(t-i) / Y_{t-j}^-$ can be regarded as the minimum variance estimate of $y(t-i)$ based on the observations spanning the space Y_{t-j}^-, corollary 1.2 implies that the innovation process at time t is orthogonal to all i-step ahead predictors at time $(t-j)$, with $j > 0$; and i being an arbitrary integer. This is a significant and interesting result.

We next consider the structure of the process $y(t)$ in terms of innovations process.

2.3 The representation of $y(t)$ and its predictors in terms of innovations:

(Generalised Wold's Decomposition)

Let $y(t)$ be a discrete time stationary process and let the innovation be defined as above. Define $E[z(t)\, z(t)'] = G$ and define $Trace(G) < \infty$. Then it can be easily shown that $y(t)$ can be written as a weighted linear combination of the present and the past values of innovations vectors, i.e, we have in the mean square sense

$$y(t) = \sum_{i=0}^{\infty} H_i\ z(t-i) \quad (10)$$

where H_i are appropriate row vectors (1 x N+2) and

$$\sum_{i=0}^{\infty} H_i H_i' < \infty. \quad (11)$$

Proof: Using theorem 1, we can write the direct sum decomposition of Y_{t+1}^- as:

$$Y_{t+1}^- = z_t \oplus z_{t-1} \oplus \ldots\ldots \oplus z_{t-p} \oplus Y_{t-p}^- \quad (12)$$

An orthogonal projection of y(t) onto the space Y_{t+1}^- yields:

$$y(t) / Y_{t+1}^- = y(t) / z_t + y(t) / z_{t-1} + \ldots$$

$$\ldots + y(t) / z_{t-p} + y(t) / Y_{t-p}^- \quad (13)$$

$$X / Z = \langle X,Z \rangle \langle Z,Z \rangle^{-1} Z \quad (14a)$$

where $\langle X,Z \rangle = E[XZ]$ (14b)

$y(t) / z_{t-i} = H_i z(t-i)$

where $H_i = E[y(t) z(t-i)'] E[z(t-i) z(t-i)']$

From (13) we get

$y(t) / Y_{t+1}^- = y(t) = H_0 z_t + H_1 z_{t-1} + ... + H_p z_{t-p}$

$+ y(t) / Y_{t-p}^-$ (15a)

$(\because y(t) \in Y_{t+1}^-)$

Since by definition

$y(t) / Y_t^- = H_1 z_{t-1} + ... + H_p z_{t-p} + y(t) / Y_{t-p}^-$

and $z(t) \triangleq y(t) - y(t) / Y_t^-$

we get $H_0 = [1\ 0\ 0 \ ...\ 0]$

$\therefore y(t) / Y_{t-p}^- = y(t) - H_0 z_t - H_1 z_{t-1} - ... - H_p z_{t-p}$ (15b)

In order to prove the lemma it is sufficient to prove that:

$Lim \quad \| y(t) / Y_{t-p}^- \|^2 \Rightarrow 0$ (16)

$As\ p \to \infty$

Using the result of theorem (1) again, this time for Y_{t-p}^-, we get

$$Y_{t-p}^- = z_{t-p-1} \oplus Y_{t-p-1}^-$$

Orthogonal projection of $y(t)$ on Y_{t-p}^- yields:

$y(t) / Y_{t-p}^- = y(t) / z_{t-p-1} + y(t) / Y_{t-p-1}^-$

$= H_{p+1} z_{t-p-1} + y(t) / Y_{t-p-1}^-$ (17)

$Since \ \| y(t) / Y_{t-p}^- \|^2$

$= \| H_{p+1} z_{t-p-1} + y(t) / Y_{t-p-1}^- \|^2$ (18)

Using corollary (1.2) we can write

$$\langle z(t-i) , y(t)/ Y_{t-k}^- \rangle = 0, for\ k \geq i$$

from which it follows that (18) can be reduced to

$\| y(t) / Y_{t-p}^- \|^2 = H_{p+1} G H_{p+1}' +$

$\| y(t) / Y_{t-p-1}^- \|^2$ (19)

where the first term on R.H.S is positive, being square of a norm, it follows that

$\| y(t) / Y_{t-p}^- \|^2 > \| y(t) / Y_{t-p-1}^- \|^2$ or

$$\frac{\| y(t) / Y_{t-p-1}^- \|}{\| y(t) / Y_{t-p}^- \|} < 1$$

Thus $\| y(t) / Y_{t-p}^- \|$ is a strictly decreasing sequence. Further noting that $Y_{-\infty}^-$ is a null space, we have $y(t) / Y_{-\infty}^- = 0$ and we can write

$$Lim_{p \to \infty} \| y(t) / Y_{t-p}^- \| \to 0$$

Thus proving the generalised Wold's decomposition theorem.

Thus we have

$$Lim_{p \to \infty} \| y(t) - \sum_{i=0}^{p} H_i z(t-i) \|^2 \to 0$$

(or)

$$\| Lim_{p \to \infty} [y(t) - \sum_{i=0}^{p} H_i z(t-i)] \|^2 \to 0$$

(since norm is continuous)

(or)

we have in the mean square sense

$$y(t) = \sum_{i=0}^{\infty} H_i z(t-i)$$ ■

2.3.1 Discussion:

This is similar to the Wold's decomposition

$$y(t) = \sum_{i=0}^{\infty} h_i z(t-i)$$

where $z(t) \triangleq y(t) - y(t) / Y_t^-$ and the result is obtained for Gaussian process in a linear framework and this can provide only a minimum phase solution of a nonminimum phase process. But in our work we got a result which is in a multiinput and single output scenario; where $y(t)$ is the output of a multi input system whose inputs are the elements of the innovation vector $z(t-i)$ and the resulting structure can provide a complete solution for a nonminimum phase process[6].

If $H_i = h_i [1\ 0 \ ...0]$, (10) reduces to the well known Wold's decomposition in the linear Gaussian case.

2.3.2 Corollary 1.3 :

The i-step ahead predictor of $y(t)$ can be written as :

$$y(t+i) / Y_{t+1}^- = y(t+i/t) = y(t+i) - \sum_{m=0}^{i-1} H_m z(t+i-m)$$

where $y(t+i/t)$ denotes the i-step ahead predictor.

Proof: The direct sum orthogonal decomposition of Y_{t+i+1}^- can be written using theorem 1 as:

$Y_{t+i+1}^- = z_{t+i} \oplus z_{t+i-1} \oplus .... \oplus z_{t+1} \oplus Y_{t+1}^-$

Projecting $y(t+i)$ orthogonally onto this space and noting that $y(t+i) \in Y_{t+i+1}^-$

we get

$$y(t+i) \mid Y_{t+i+1}^{-} = y(t+i) = y(t+i) \mid z_{t+i} +$$

$$y(t+i) \mid z_{t+i-1} + \ldots + y(t+i) \mid z_{t+1} + y(t+i) \mid Y_{t+1}^{-}$$

and hence the result

$$y(t+i) \mid Y_{t+1}^{-} = y(t+i/t) = y(t+i) - \sum_{m=0}^{i-1} H_m \; z(t+i-m)$$

follows as desired. ■

Corollary 1.4 :

The *i*-step ahead predictor of $y(t)$ can also be expressed as:

$$y(t+i) \mid Y_{t+1}^{-} = y(t+i/t) = \sum_{m=0}^{\infty} H_{m+i} \; z(t-m) \qquad (20)$$

Proof: We know that

$$y(t+i) = \sum_{j=0}^{\infty} H_j \; z(t+i-j) \qquad (21)$$

From (20) we have

$$y(t+i) \mid Y_{t+1}^{-} = y(t+i/t) = y(t+i) - \sum_{m=0}^{i-1} H_m \; z(t+i-m)$$

$$= \sum_{j=0}^{\infty} H_j \; z(t+i-j) - \sum_{m=0}^{i-1} H_m \; z(t+i-m)$$

$$= \sum_{j=i}^{\infty} H_j \; z(t+i-j)$$

$$= \sum_{n=0}^{\infty} H_{n+i} \; z(t-n)$$

Hence the result (20). ■

2.3.2.1 Discussion:

The results of corollaries 1.3 and 1.4 relate the *i*-step ahead predictors and prediction errors to the past innovations. Equation (20) expresses the *i*-step ahead prediction error $[y(t+i) - y(y+i/t)]$ as a weighted sum of the '*i*' immediately past innovations with $[H_j, j=0,1,..,i-1]$ as the weighting coefficient vectors.

Corollary 1.4 gives a representation of the *i*-step ahead predictor in terms of past innovation vectors. Note the interesting similarity of the representation of $y(t)$ in (10) and of $y(t+i/t)$ in (21).

The most important and useful feature of corollary 1.3 is that unlike in corollary 1.4, this result enables us to express the *i*-step ahead predictor in terms of a finite dimensional basis. More specifically, the *i*-step ahead predictor is obtained here subtracting from $y(t+i)$, the information not contained in the infinite dimensional space Y_{t+1}^{-} .

3. Finite dimensionality of the generalised predictor space:

With the results of the earlier section at our disposal, in this section we will prove the finite dimensionality of the generalised predictor space, if the process admits a linear nonGaussian ARMA representation.

Theorem 2: Let $y(t)$ be an ARMA process of order (p,q), then the predictor space $P_{t-k,2,N}$ where N=max(p,q) is finite dimensional and is spanned by the predictors $[\, y(t-1) \mid Y_{t-N}^{-},\; y(t-2)/Y_{t-N}^{-},\; y(t-3) \mid Y_{t-N}^{-},\; '\ldots,\; y(t-N) \mid Y_{t-N}^{-} \,]$, *Where* $y_0 = 1$.

Proof:

$$y(t) + a_1 \, y(t-1) + \ldots\ldots + a_p \, y(t-p) = w(t) + b_1 \, w(t-1) + \ldots\ldots + b_q \, w(t-q) \qquad (22)$$

where $w(t)$ are i.i.d nonGaussian white process.

Assuming p>q, $N = max(p,q) = p$

Projecting both sides of (22) onto Y_{t-N}^{-} we get

$$y(t) \mid Y_{t-N}^{-} + a_1 \, y(t-1) \mid Y_{t-N}^{-} + \ldots + a_p \, y(t-p) \mid Y_{t-N}^{-} = w(t) \mid Y_{t-N}^{-} + b_1 \, w(t-1) \mid Y_{t-N}^{-} + \ldots + b_p \, w(t-p) \mid Y_{t-N}^{-} \qquad (23)$$

Clearly R.H.S of (23) = 0 .

By using induction principle we can show that

$$y(t+i) \mid Y_{t-N}^{-} = \text{Linear combination of } [\, y(t-1) \mid Y_{t-N}^{-}, \ldots\ldots, y(t-p) \mid Y_{t-N}^{-} \,] \text{ for any } i>0.$$

Multiplying by y(t-i) on both sides ($i \geq N+1$) , we get

$$y(t)y(y-i) + a_1 \, y(t-1)y(t-i) + \ldots + a_p \, y(t-p)y(t-i) = w(t)y(t-i) + b_1 \, w(t-1)y(t-i) + \ldots + b_q \, w(t-q)y(t-i) \qquad (24)$$

Projecting both sides onto Y_{t-N}^{-} we get

$$\sum_{j=0}^{p} a_j \, [y(t-j)y(t-i)] \mid Y_{t-N}^{-} = 0, \text{ for } i \geq N+1, a_0 = 0. \qquad (25)$$

Considering (24) for $i \leq N$ and by using the relation

$$y(t) = \sum_{k=0}^{\infty} h_k \; w(t-k)$$

(where h_k are impulse response coefficients)

we get

$$\sum_{j=0}^{p} a_j \, [y(t-j)y(t-i)] \mid Y_{t-N}^{-} = \sum_{j=0}^{q} b_j \, [w(t-j)y(t-i)] \mid Y_{t-N}^{-}$$

$$= \sum_{j \geq i}^{q} b_j \, [w(t-j)y(t-i)] \mid Y_{t-N}^{-}$$

$$= \sum_{j \geq i}^{q} b_j \, h_{j-i} \quad = k_i \qquad (26)$$

(where k_i are constants)

Combining (25) and (26) we get

$$y(t)y(t-i) \mid Y_{t-N}^{-} + a_1\, y(t-1)y(t-i) \mid Y_{t-N}^{-} + \ldots\ldots$$

$$+a_p\, y(t-p)y(t-i) \mid Y_{t-N}^{-} - k_i \mid Y_{t-N}^{-} = 0, \forall\, i, \qquad (27)$$

where k_i are some arbitrary constants and k_i are equal to zero for all $i \geq N+1$.

Using (27) we can show that all the predictors formed by the nonlinear terms *i.e.*, $y(t+i)y(t+i)) \mid Y_{t-N}^{-}$, $y(t+i)\, y(t+i-1) \mid Y_{t-N}^{-}$, ..., $y(t+i)y(t+i-p) \mid Y_{t-N}^{-}$ *for all* $i \geq 0$, can be expressed in terms of the assumed basis of P_{t-kN}^{2}, *i.e.*, $\boldsymbol{y(t-1) \mid Y_{t-N}^{-}}$, $\boldsymbol{y(t-2)/Y_{t-N}^{-}}$, $\boldsymbol{y(t-3) \mid Y_{t-N}^{-}}$, '..., $\boldsymbol{y(t-N) \mid Y_{t-N}^{-}}$.

This completes the proof. ■

4. Application to nonGaussian ARMA processes:

Here we present a result for representation of nonGaussian ARMA processes, which will be useful in the estimation of parameters of the process; and the relevant results will be published elsewhere. It is expected to provide consistent estimates of the AR parameters, as it makes use of more than p+1 slices of higher order statistical terms due to the expanded space[2].

Theorem 3:

If $y(t)$ admits the nonGaussian ARMA(N,N) representation,

$$y(t) + a_1\, y(t-1) + \ldots + a_N\, y(t-N) = w(t) + b_1\, w(t-1) + \ldots + b_N\, w(t-N)$$

Then y(t) admits the following representation:

$$y(t) + a_1\, y(t-1) + \ldots\ldots + a_N\, y(t-N) =$$

$$B_0\, z(t) + B_1\, z(t-1) + \ldots + B_N\, z(t-N), \qquad (28)$$

where

$\boldsymbol{z(t-i)}$ is as defined in(6) and N is the memory of nonlinearity in the predictor space representation.

Proof:

$$y(t) + a_1\, y(t-1) + \ldots\ldots + a_N\, y(t-N) = w(t) + b_1\, w(t-1) + \ldots\ldots + b_N\, w(t-N)$$

Projecting both sides onto Y_{t-N}^{-} we get

$$y(t) \mid Y_{t-N}^{-} + a_1\, y(t-1) \mid Y_{t-N}^{-} + \ldots + a_p\, y(t-p) \mid Y_{t-N}^{-} = 0$$

By writing the predictors in terms of the innovations $\boldsymbol{z(t-i)}$ we get the desired result. Where

$$H_0 = B_0$$

$$H_1 + a_1 H_0 = B_1$$

....

$$H_N + a_1 H_{N-1} + a_2 H_{N-2} + \ldots\ldots + a_N H_0 = B_N$$

This results in the desired expression (28). ■

5. Conclusions:

An approach to representing nonGaussian/ nonminimum phase random processes has been proposed. The results are proposed for second order nonlinearity for ease of presentation; though the generalisation of the theory to pth order nonlinearity is straight forward. Application of this theory to the parameter estimation of nonGaussian ARMA process is being investigated. We can also expect better estimates using this extended space as we are able to utilise more statistical information of the process. A result to this effect has been given by Bondon et.al[3]. A more generalised version of the representation for which the present case forms a special case is also obtained, the results of which are to be published elsewhere.

REFERENCES

[1] Akaike," *Markovian representation of stochastic processes and its application to the analysis of Auto Regressive Moving Average processes*," Annals of Institute of Statistical Mathematics, Vol.26, No.3, 1974.

[2] G.B.Giannakis, "*On the identifiability of nonGaussian ARMA models using cumulants*", IEEE trans. on Automatic Control, Vol.35, pp.18-26,1990.

[3] P.Bondon, P.L.Combettes and Picinibono," *Volterra filtering and higher order whiteness*", IEEE trans. on Signal Processing, Vol.43, No.9, September 1995.

[4] K.S.Lii and M.Rosenblatt, "*Deconvolution and estimation of transfer function phase and coefficients for nonGaussian linear processes*," Annals of Statistics, Vol.10, pp.1195-1208, 1982.

[5] J.M.Mendel, "*Tutorial on Higher order Statistics (Spectra) in signal processing and system theory: Theoretical results and some applications*", Proc. IEEE, Vol.79, pp.279-305, 1991.

[6] M.Rosenblatt, "*Comments on structure of and estimation for nonGaussian linear processes*" in *Topics in nonGaussian signal processing*, edited by E.J.Wagman, S.C.Schwartz and J.B.Thomas, Springer Verlog, New York Inc., 1989.

[7] Shivdutt Joshi, "*A novel representation for ARMA processes: Theory and applications*," Ph.D. dissertation, Department of Electrical Engineering, Indian Institute of Technology, Delhi, 1987.

Factorizability of Complex Signals Higher (Even) Order Spectra : a Necessary and Sufficient Condition

Joël Le Roux and Cécile Huet
I3S, University of Nice CNRS Bat. 4, tel 33 4 92 94 26 82
250 rue Albert Einstein Sophia Antipolis fax 33 4 92 94 28 98
06560 Valbonne France email : leroux@essi.fr

Abstract

This communication presents a necessary and sufficient condition for the factorizability of higher order spectra of complex signals. This condition is based on the symmetries of higher order spectra and on an extension of a formula proposed by Marron, Sanchez and Sullivan for unwrapping phases of third order spectra [1]. It is an identity between products of higher order spectra. Our factorisability test requires no phase unwrapping.
[1] J. C. Marron , P. P. Sanchez and R. C. Sullivan (1990): "Unwrapping algorithm for least-squares phase recovery from the modulo 2π bispectrum phase," J. Opt. Soc. Am. A/Vol.7, pp 14-20, January 1990.)

1. Introduction

The problem of higher order factorizability has been studied by several researchers. Tekalp and Erdem [13] and Pan and Nikias [10], base their development on the higher order "cepstrum" and check whether the support of this cepstrum reduces to a precise subspace or not. The difficulty in their approach is related to the multiple determination of a complex number logarithm. This implies the use of a phase unwrapping procedure. Dianat and Raghuveer [6] propose an approach that is valid only when MA modelling at a given order is possible. Alshebeili [1] has proposed a condition in the form of a LDU decomposion of the cumulant matrix (at order 3). The three kinds of methods show the validity of the factorisability essentially in computing functions closely related to the higher order spectral factor *i.e.* the values of the system impulse response.

For a different purpose, Marron, Sanchez and Sullivan [9] have given a recursive formula performing phase unwrapping of third order spectra. Their method is based on a relation between different values of a linear combination of the third order spectrum phases. It does not perform an implicit computation of the higher order spectral factor.

In this communication we intend to show that Marron, Sanchez and Sullivan (MSS) formula can actually be considered as a basis for developping a necessary and sufficient condition for factorizability : (1) we generalize MSS third order formula to the case of higher order spectra of complex signals ; (2) we give the expression of Marron's formula generalization directly in function of the higher order spectrum and not of the higher order spectrum logarithm (modulus logarithm and phase) ; the formula is an identity between products of higher order spectra in the frequency domain ; (3) we show that this generalization of Marron's formula is equivalent to the expression of factorizability in the cepstral domain [13] provided that the higher order spectral symmetries are satisfied ; this gives the necessary and sufficient conditions for factorizability ; (4) we show that this formulation in the frequency domain does not raise a phase unwrapping problem.

2. A generalization of MSS formula

We consider the third order spectrum $S_3(u,v)$ of a non gaussian real random process,

$$S_3(u,v) = E\{X(u)X(v)X(-u-v)\}. \qquad (1)$$

When $S_3(u,v)$, is factorizable and if its phase is correctly unwrapped, the following identity is satisfied [9]:

$$\psi_3(u+w,v)+\psi_3(u,w)=\psi_3(u+v,w)+\psi_3(u,v) \qquad (2)$$

where $\psi_3(u,v)$ is the phase of $S_3(u,v)$. This formula can be generalized as shown below (a detailed development is given in [7].)

We consider complex signals N-th order spectra, Fourier transforms of cumulants,

$$S_N(w_1,\ldots,w_{N-1}) = \rho_N(w_1,\ldots,w_{N-1}) \exp j\psi_N(w_1,\ldots,w_{N-1}). \qquad (3)$$

If factorizability is satisfied, then

$$S_N(w_1,\ldots,w_{N-1}) = \left[\prod_{p=1}^{N/2} H(w_p)\right] \qquad (4)$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$\times \left[\prod_{q=N/2+1}^{N-1} H^*(-w_q) \right] H^*\left(\sum_{r=1}^{N-1} w_r\right),$$

where $H(w)$ is the N-th order spectral factor. When (4) is satisfied the spectral factor $H(w)$ is unique up to a linear phase term of the form $e^{j(a+b.w)}$. Eq. (4) generalizes expressions found in [11][12], as done in [2] with a different choice of the variables. The expression of (4) given in terms of phases is

$$\psi_N(w_1,\ldots,w_{N-1}) = \sum_{p=1}^{N/2} \varphi(w_p) - \sum_{q=N/2+1}^{N-1} \varphi(-w_q) - \varphi\left(\sum_{r=1}^{N-1} w_r\right), \quad (5)$$

where $\varphi(w)$ is the phase of $H(w)$.

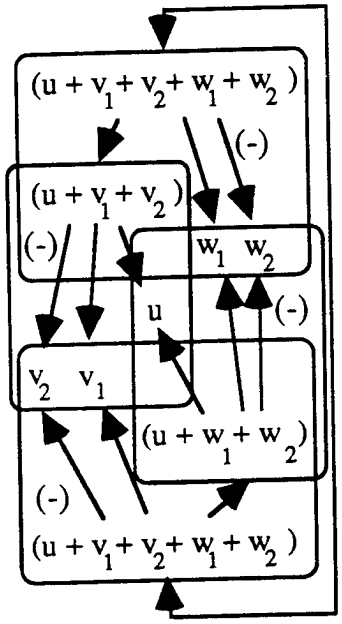


Fig. 1. MSS formula at order 4

When (5) holds, the *unwrapped* higher order spectrum phases satisfy :

$$\psi_N(u+\sum_{p=1}^{N-2} w_p, v_1,..,v_{N-2}) + \psi_N(u,w_1,..,w_{N-2}) = \psi_N(u+\sum_{p=1}^{N-2} v_p, w_1,..,w_{N-2}) + \psi_N(u,v_1,..,v_{N-2}). \quad (6)$$

(See the illustration in fig. 1). The result is verified by direct development. A similar equation holds in the case of the modulus logarithm (note the terms in $\log\rho_2$):

$$\log\rho_N(u+\sum_{p=1}^{N-2} w_p, v_1,\ldots,v_{N-2}) + \log\rho_N(u,w_1,\ldots,w_{N-2}) - \log\rho_2(u+\sum_{p=1}^{N-2} w_p) = \log\rho_N(u+\sum_{p=1}^{N-2} v_p, w_1,..,w_{N-2}) + \log\rho_N(u,v_1,\ldots,v_{N-2}) - \log\rho_2(u+\sum_{p=1}^{N-2} v_p). \quad (7)$$

Consequently, we have the following identity which holds even when the phase is not unwrapped (this identity is verified by direct development) :

$$\frac{S_N(u+\sum_{p=1}^{N-2} w_p, v_1,\ldots,v_{N-2}) S_N(u,w_1,\ldots,w_{N-2})}{S_2(u+\sum_{p=1}^{N-2} w_p)} = \frac{S_N(u+\sum_{p=1}^{N-2} v_p, w_1,\ldots,w_{N-2}) S_N(u,v_1,\ldots,v_{N-2})}{S_2(u+\sum_{p=1}^{N-2} v_p)}. \quad (8)$$

3. Symmetries of complex signals higher order spectra

The factorizability condition is based on MSS formula, but also the symmetries of complex signals higher order spectra. Studies on these symmetries can be found in [3][4][8]. Here we give the matricial operators associated with the transformations of the space that keep the higher order spectrum invariant. We consider only even orders. We have the two types of operations ($M = N/2-1$):

(a) The permutations of the $(M+1)$ first variables or of the M last variables do not modify $S_N(u,v_1,\ldots,v_M,w_1,\ldots,w_M)$. Any of the $(M+1)!M!$ matrices of the form

$$P = \begin{bmatrix} Q_{M+1} & 0 \\ 0 & Q_M \end{bmatrix}, \quad (9)$$

where Q_{M+1} and Q_M are permutation matrices of dimension $(M+1)$ and M respectively keep the higher order spectrum invariant.

(b) The following change of variables :

$$\begin{aligned} u' &= u + v_1 + \ldots + v_M + w_1 + \ldots + w_M, \\ v_1' &= -w_1, \ldots, v_M' = -w_M, \\ w_1' &= -v_1, \ldots, w_M' = -v_M, \end{aligned} \quad (10)$$

transforms the higher order spectrum in its complex conjugate. As a consequence, when applied to the variables, the operator T ,

$$T=\left[\begin{array}{c|ccccc|ccccc} 1 & 1 & \cdots & \cdots & \cdots & 1 & 1 & \cdots & \cdots & \cdots & 1 \\ \hline 0 & 0 & \cdots & \cdots & \cdots & 0 & -1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & \vdots & 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & & \ddots & \vdots & 0 & \cdots & 0 & -1 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & 0 & -1 \\ \hline 0 & -1 & 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & 0 & -1 & 0 & \cdots & 0 & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \ddots & & \vdots \\ \vdots & 0 & \cdots & 0 & -1 & 0 & \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & -1 & 0 & \cdots & \cdots & \cdots & 0 \end{array}\right], \quad (11)$$

keeps the higher order spectrum invariant (if we except the complex conjugation). One can verify that products of matrices of type P and T form a finite goup that keep invariant the subspace support of higher order spectra of band limited stationary signals.

4. A necessary and sufficient condition for factorizability

Theorem : $S_N(u_1,...,u_{N-1})$ is factorizable in the form (4) if and only if (8) is satisfied for a nonzero value of $(v_1,...,v_{N-2})$ and $S_N(u_1,...,u_{N-1})$ satisfies the higher order spectral symmetries of section 3.

The choice of the vector $(v_1,...,v_{N-2})$ requires some care: if $S_N(u_1,...,u_{N-1})$ is periodic, its period must be different from a multiple of $(v_1,...,v_{N-2})$. If $S_N(u_1,...,u_{N-1})$ is periodic and discrete, which is often the case in applications based on the DFT, it is always possible to choose $(v_1,...,v_{N-2})=(\Delta\nu,0,...,0)$ where $\Delta\nu$ is the frequency resolution.

Proof : The fact that the condition is necessary was shown in section 2.

4.1. Proof of the condition on the phases when they are correctly unwrapped

Symmetry (10) implies that $\psi_N(0,...,0)=0$.

If we know a correct (for example the continuous) determination of the higher order spectrum phase, and not only its determination in the range [-π, π[, we have

$$\psi_N(u+\sum_{p=1}^{N-2}w_p,v_1,..,v_{N-2})+\psi_N(u,w_1,..,w_{N-2}) \quad (12)$$
$$=\psi_N(u+\sum_{p=1}^{N-2}v_p,w_1,..,w_{N-2})+\psi_N(u,v_1,..,v_{N-2}).$$

(a) expression of the higher order spectrum phase as a sum

When $(v_1,...,v_{N-2})$ is fixed, the difference between the two function of $(N-1)$ variables in (12), $\psi_N(u+\sum_{p=1}^{N-2}v_p,w_1,...,w_{N-2})$ and $\psi_N(u,w_1,...,w_{N-2})$, is the difference between two functions of one variable :

$$\psi_N(u,w_1,..,w_{N-2})-\psi_N(u+\sum_{p=1}^{N-2}v_p,w_1,..,w_{N-2})= \quad (13)$$
$$\psi_N(u,v_1,..,v_{N-2})-\psi_N(u+\sum_{p=1}^{N-2}w_p,v_1,..,v_{N-2}).$$

If $\psi_N(u,w_1,...,w_{N-2})$ is a periodic function of u, we suppose that we have chosen a vector $(v_1,...,v_{N-2})$ such that $\ell=\sum_{p=1}^{N-2}v_p$ and its multiples are different from its period ; otherwise, MSS formula could not be used to check the factorizability. In order to take advantage of this property , we decompose $\psi_N(u,w_1,...,w_{N-2})$ in a sum :

$$\Psi_N(u,w_1,...w_{N-2})=f(u)$$
$$+g(u+\sum_{p=1}^{N-2}w_p)+h(u,w_1,...w_{N-2}), \quad (14)$$

where $h(u,w_1,...,w_{N-2})$ holds no additive terms depending only on u or $(u+\sum_{p=1}^{N-2}w_p)$.

(b) $h(u,w_1,...,w_{N-2})$ does not depend on u

Using (14) in computing the difference (13) yields :

$$\psi_N(u,w_1,...,w_{N-2})-\psi_N(u+\ell,w_1,...,w_{N-2}) \quad (15)$$
$$=f(u)+g(u+\sum_{p=1}^{N-2}w_p)+h(u,w_1,..,w_{N-2})$$
$$-f(u+\ell)+g(u+\ell+\sum_{p=1}^{N-2}w_p)$$
$$+h(u+\ell,w_1,..,w_{N-2})$$
$$=\psi_N(u,v_1,..,v_{N-2})-\psi_N(u+\sum_{p=1}^{N-2}w_p,v_1,..,v_{N-2}),$$

with $\ell=\sum_{p=1}^{N-2}v_p$. The last line of (15) holds only two terms functions respectively of u and $(u+\sum_{p=1}^{N-2}w_p)$. Consequently,

$$h(u,w_1,..,w_{N-2})=h(u+\ell,w_1,..,w_{N-2}). \quad (16)$$

If $(v_1,...,v_{N-2})$ is chosen appropriately, $\psi_N(u,w_1,...,w_{N-2})$ is not periodic of period ℓ in u, and neither is $h(u,w_1,...,w_{N-2})$. So, this last function

does not depend on u:

$$h(u,w_1,\ldots,w_{N-2}) = h'(w_1,\ldots,w_{N-2}), \quad (17)$$

and (14) becomes

$$\psi_N(u,w_1,\ldots,w_{N-2}) = f(u) + g(u+\sum_{p=1}^{N-2} w_p) + h'(w_1,\ldots,w_{N-2}). \quad (18)$$

(c) Application of the symmetries

Now we can apply the changes of variables of section 3, that keep the higher order spectrum invariant. First the permutation of u and any of the first $(N/2-1)$ variables $w_1,\ldots,w_{N/2-1}$ yields:

$$h'(w_1,..,w_p,..,w_{N/2-1},w_{N/2},..,w_{N-2}) = f(w_p) + h'(w_1,..,u,..,w_{N/2-1},w_{N/2},..,w_{N-2}) - f(u), \quad (19)$$

so that

$$h'(w_1,\ldots,w_{N/2-1},w_{N/2},\ldots,w_{N-2}) = \sum_{p=1}^{N/2-1} f(w_p) + h''(w_{N/2},\ldots,w_{N-2}). \quad (20)$$

If we use the transformation T,

$$h''(w_{N/2},\ldots,w_{N-2}) = -\sum_{p=N/2}^{N-2} f(-w_p), \quad (21)$$

and

$$g(u) = -f(u). \quad (22)$$

Finally

$$\psi_N(u,w_1,..,w_{N-2}) = f(u) + \sum_{p=1}^{N/2-1} f(w_p) - \sum_{p=N/2}^{N-2} f(-w_p) - f(u+\sum_{p=1}^{N-2} w_p), \quad (23)$$

which expresses the factorizability.

4.2. On the computation of the correct phase determination (phase unwrapping)

If (12) is satisfied, we have the expression (23) of $\psi_N(u,w_1,\ldots,w_{N-2})$ provided that $\psi_N(u,w_1,\ldots,w_{N-2})$ is a correct determination of the HOS phase in the range $[-4\pi,4\pi[$. We intend to show that it is always possible to compute the correct determination of $\psi_N(u,w_1,\ldots,w_{N-2})$ from the one that is known $\psi_N^0(u,w_1,\ldots,w_{N-2})$(given in the range $[-\pi,\pi[$):

$$\psi_N(u,w_1,..,w_{N-2}) = \psi_N^0(u,w_1,..,w_{N-2}) + 2\pi m(u,w_1,\ldots,w_{N-2}), \quad (24)$$

where m is an integer. We consider the values of $\psi_N^0(u,v,0,\ldots,0)$ (the third order spectum phase).A subset of these data can be used to reconstruct the spectral factor phase $f(u)$ using a recursive multiresolution algorithm without raising phase unwrapping difficulties [8] . The recursive structure of the algorithm allows, in theory, the reconstruction of $f(u)$ for all u at any frequency resolution. When $f(u)$ is known, in using (5), it is possible to compute the correct determination (24) of $\psi_N(u,w_1,\ldots,w_{N-2})$ satisfying (6). (Besides, this was the initial object of MSS recursion) : Starting from the HOS data used in this multiresolution method, there is only one manner of reconstructing recursively $\psi_N(u,w_1,\ldots,w_{N-2})$ satisfying (6). So, as the phase reconstructed through the computation of $f(u)$ satisfies (6), it gives the determination satisfying (24).

Consequently when (8) is satisfied, it is always possible to reconstruct the determination of the phase allowing the development of section 4.1 since the spectral factor is unique. The knowledge of $\psi_N^0(u,w_1,\ldots,w_{N-2})$ in the range $[-\pi,\pi[$ is sufficient.

4.3. Factorizability of the higher order spectrum modulus

A similar development applies to the modulus logarithm. When the HOS symmetries and (7) are satisfied, the modulus logarithm can be written in the form

$$\log\rho_N(u,w_1,\ldots,w_{N-2}) = \lambda(u) + \lambda(u+w_1+\ldots+w_{N-2}) + \lambda(w_1)+\ldots+\lambda(w_{N/2-1}) + \lambda(-w_{N/2})+\ldots+\lambda(-w_{N-2}). \quad (25)$$

When (7) is satisfied,

$$2\lambda(u) = \log|S_2(u)|, \quad (26)$$

which gives the modulus of the spectral factor. This implies the factorizability of the modulus.

Consequently, if (8) and the HOS symmetries hold, $S_N(u,w_1,\ldots,w_{N-2})$ is factorizable.

References

[1] S. A. Alshebeili, "New results on bispectrum factorization and nonminimum phase FIR system identification," *IEEE Signal Processing ATHOS Workshop on HOS*, Begur, Spain, June 1995, pp 439-443.

[2] P. O. Amblard, M. Gaeta and J. L. Lacoume, "Statistics for complex variables and signals, part II: signals", *Signal Processing*, n° 53, (1996), pp 15-25.

[3] D. R. Brillinger and M. Rosenblatt, "Computation and interpretation of the k-th order spectra", *Spectral analysis of time series*, B. Harris, ed. J. Wiley, 1966, pp 190-232.

[4] V. Chandran and S. Elgar, "A general procedure for the derivation of principal domains of higher order spectra", IEEE trans. on ASSP, vol. 42, n° 1, 1994, pp. 229-233.

[5] J. W. Dalle Molle, "A procedure to generate the nonvanishing support of the n-th order polyspectrum", *4th international IEEE signal processing workshop on higher order statistics*, Begur, Spain, June 1995, pp. 386-390.

[6] S. A. Dianat and M. R. Raghuveer : "Polyspectral factorization : necessary and sufficient condition for finite extent cumulant sequences," *Proc. Int Conf. on ASSP*, Glasgow, Scotland, U.K.,, May 1989 pp. 2322-2324.

[7] C. Huet, "Factorisation de spectres d'ordre quatre des signaux complexes", *thèse de doctorat*, Univ. Nice, 1997.

[8] J. Le Roux, D. Rossille and C. Huet, "A multiresolution extension of Lohman, Weigelt and Wirnitzer recursion for the reconstruction of the Fourier transform phase from the bispectrum phase", *4th international IEEE signal processing workshop on higher order stats*, Begur, Spain, June 1995, pp. 315-319.

[9] J. C. Marron , P. P. Sanchez and R. C. Sullivan : "Unwrapping algorithm for least-squares phase recovery from the modulo 2π bispectrum phase," *J. Opt. Soc. Am.* A/Vol.7, January 1990, pp. 14-20.

[10] R. Pan and C. L. Nikias, "The complex cepstrum of higher order cumulants and nonminimum phase system identification," *IEEE Trans. on Signal Processing*, vol. ASSP-36(2), Feb 1988, pp. 186-205.

[11] R. Pierce, "Velocity Measurement of Radar Target using HOS," *IEEE Signal Processing ATHOS Workshop on HOS , Lake Tahoe* , June 1993, pp. 164-167.

[12] O. Shalvi and E. Weinstein, "Super-Exponential methods for blind deconvolution," *In Proc. SPIE 1991 Adaptative Signal Processing* , Vol 1565, Jan. 1991, pp. 143-152.

[13] A. M. Tekalp and A.T. Erdem : "Higher-order spectrum factorization with applications in signal modeling and nonminimum phase system identification," *IEEE Trans. Acoust Speech Signal Processing*, Vol 37(10), 1989, pp. 1537-1549.

The estimation of stable distribution parameters

Stephen Bates and Steve McLaughlin
Department of Electrical Engineering,
University of Edinburgh, Mayfield Rd, Edinburgh, EH9 3JL

Abstract

This paper concerns the estimation of the parameters that describe a stable distribution. Stable distributions are characterised by four parameters which can be estimated using a number of methods and although approximate maximum likelihood estimation (MLE) techniques do exist, they are computationally intensive. There are a number of techniques that are much faster than MLE and these are the focus of this paper. These techniques are compared and contrasted both for stable random variables and for teletraffic data.

1 Introduction

A wide range of modelling areas that encompass economics, telecommunications, channel noise modelling and signal interpolation have received treatment with stable distributions (see section 1.4 in [10]). The reasons for this are that the environment to be modelled has tended to be more impulsive than the Gaussian case and the development of more powerful computers has made working with stable processes a more realistic proposition.

The lack of closed form expression for the pdf of almost all stable distributions makes MLE techniques impossible and it was only in 1971 that DuMouchel developed an approximate MLE technique [2]. However this technique is computationally intensive and does not lend itself to the real time estimation of stable parameters. We wish to concern ourselves with the quicker, more approximate techniques that have been developed.

This paper begins by introducing the salient properties of the stable distribution. It then goes on to discuss the methods we consider in this paper. The estimation of random variables drawn from a truly stable distribution is the focus of section 4 and this allows us to compare how different techniques perform when the actual parameter values are known. In section 5 we go on to estimate stable parameters for compressed video data. In this case the correct values for the stable distribution are not known. However we use an inverse Fourier transfer technique to determine the accuracy of each techniques estimates. Finally we finish with a number of conclusions and in appendix A we give details of how the software used in this paper may be obtained.

2 Stable Distribution Parameters

All stable distributions can be uniquely expressed by their characteristic function which is given by $\varphi(t) = \exp\{jat - |\gamma t|^{\alpha}[1 + j\beta \text{sign}(t)\omega(t,\alpha)]\}$, where, $\omega(t,\alpha) = \tan(\alpha\pi/2)$ for $\alpha \neq 1$ and $(2/\pi)\log|t|$ for $\alpha = 1$ and $\text{sign}(t) = 1$ for $t > 0$, 0 for $t = 0$ and -1 for $t < 0$.

The parameters a, α, β and γ describe completely a stable distribution. a ($-\infty \leq a \leq \infty$) is the location parameter, α ($0 \leq \alpha \leq 2$) is termed the characteristic exponent, β ($-1 \leq \beta \leq 1$) is the index of skew and γ ($0 \leq \gamma \leq \infty$) is termed the dispersion parameter. Some estimation techniques find a value for c which is equal to $\gamma^{\frac{1}{\alpha}}$.

In almost every case stable distributions do not have a closed form probability density function (pdf). The only exceptions to this are the case when $\alpha = 2$ (Gaussian), $\alpha = 1$, $\beta = 0$ (Cauchy) and $\alpha = 0.5$, $\beta = -1$ (Pearson). For more details about the properties of stable distribution see [10] and [11].

3 Estimators

The techniques under consideration in this paper fall into three categories. Firstly there are those that perform their estimates based on the quantiles of the distribution [9]. Then there is the technique developed by Kogon which employs an empirical estimate of the characteristic function [4]. Finally there are the fractional lower order moment (FLOM) based methods developed by Nikias, Ma and Tsihrintzis [8], [12]. In this section we briefly introduce these techniques and finish with a summary of their valid estimation ranges.

3.1 The quantile based method

Estimation techniques based on the quantiles of an empirical sample were first suggested by Fama and Roll [3].

0-8186-8005-9/97 $10.00 © 1997 IEEE

However their technique was limited to symmetric distributions and suffered from a small asymptotic bias. McCulloch developed a technique that uses five quantiles from a sample and estimates α, β, c and a over a wider range of the parameter space without asymptotic bias [9].

3.2 The characteristic function based method

Stable distributions can be uniquely defined by their characteristic function (the Fourier transform of their pdf). This implies that an empirical approximation of the function may be used to estimate the stable parameters for that sample. It was Koutrouvelis who first developed a method that took advantage of this approach [5], [6]. However in [4] Kogon develops a modified approach that yields better results in almost all cases and reduces the amount of computation required.

3.3 The FLOM based methods

A stable random variable with characteristic exponent $1 \leq \alpha < 2$ has infinite variance but finite mean. If the random variable has $\alpha < 1$ then both the mean and variance are infinite. More generally we can say $\mathbf{E}(|X|^p) = \infty$ if $p \geq \alpha, < \infty$ if $p < \alpha$. We can therefore use the behaviour of these FLOMs to determine an estimate for α. In [8] Ma and Nikias present the log FLOM method and in [12] Tsihrintzis and Nikias present the extreme order method.

3.4 Summary

Table 1 summarises the permissible range of parameter values for each of the estimation methods. Parameters that can be estimated by a given technique are marked with a *.

Name	α	β	γ	a
McCulloch	$(0.6, 2)^*$	$(-1, 1)^*$	$(0, \infty)^*$	$(-\infty, \infty)^*$
Kogon	$(0, 2)^*$	$(-1, 1)^*$	$(1)^*$	(0)
Ma	$(0, 2)^*$	(0)	$(0, \infty)^*$	(0)
Tsihrintzis	$(0, 2)^*$	(0)	$(0, \infty)^*$	$(-\infty, \infty)^*$

Table 1. *Permissible ranges of estimation.*

4 Estimating parameters for stable random variables

In this section we test the accuracy of the estimation techniques when the data is drawn from a stable distribution with known parameter values. In all the cases in this section we performed the estimation for 1000 independent sets of data, each of length 1000, and recorded the mean and the 95% confidence interval of the estimate.

4.1 Generating stable random variables

We generated the stable random variables using a method based on the work of Chambers, Mellow and Stuck (CMS) [1] and reproduced in [11].

4.2 Results

The first results consider the case of estimating α when $\beta = 0$, $c = 1$ and $a = 0$. Figure 1 plots the mean and 95% confidence interval for α for the four estimation techniques.

The next series of results is concerned with the estimation of α and β for series of stable random variables drawn from a skewed distribution. We only used the results from those estimators capable of estimating β i.e The estimators of Kogon and McCulloch. We have plotted the results for β =-1, -0.5, 0.5 and 1 in figures 2-5.

Next we wished to consider the effects of mis-applying a symmetric estimator to an asymmetric series. In this case we used the FLOM based techniques to estimate α from data drawn from a skewed distribution (β=-0.5 and 0.5, figure 6).

To investigate the estimation of γ we generated data and then scaled it using the mapping $X \rightarrow X * k$. Since the CMS algorithm generates random variables with $\gamma = 1$ the expected estimate of the scale parameter, c ($c = \gamma^{\frac{1}{\alpha}}$) is k. We considered the symmetric case for k=0.1 and k=10 (figures 7 and 8).

4.3 Discussion

There are many points that can be made about the results in section 4.2, the main ones we wish to make are listed below.

1. In figure 1 McCulloch's method performs best when $\alpha > 0.5$ (as expected) with a slight drop in the accuracy of $\hat{\alpha}$ as $\alpha \rightarrow 2$.

2. Tsihrintzis's method tends to overestimate α although it gives very similar results for $\beta = 0$ and $\beta = \pm 0.5$. This is unexpected as the method is based on the assumption of symmetry.

3. Kogon's estimation technique tends to perform better then McCulloch's over the range $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $c = 1$ and $a = 0$ (figures 2 - 5).

4. The method of Ma performs just as well at estimating α when $\beta = \pm 0.5$ as when $\beta = 0$. This suggests that the FLOM techniques may be more robust to the value of β than previously thought.

5. Kogon's $\hat{c}$ are very poor when $k = 10$ but improve when $k = 0.1$. This is due to the flattening effect larger γ have on the characteristic function.

6. Tsihrintzis's' $\hat{c}$ are overbiased when $k = 10$ and are underbiased when $k = 0.1$.

7. McCulloch and Ma's estimates for $\hat{c}$ are superior than the other two methods. Ma slightly outperforms McCulloch over the range $\alpha \in (0, 2]$.

5 Estimating parameters for real data

In this section we turn our attention to attempting to determine stable parameters for real data, to which we are attempting to fit a stable distribution. One problem is determining whether the estimated parameters produce a distribution which matches the data well. To test this we perform a goodness of fit test between the estimated pdf of the data and the pdf of the estimated stable distribution. We have already mentioned that no closed form expression for the pdf exists in most cases of stable distribution but we can determine it by calculating the inverse Fourier transform of the characteristic equation (section 2.3 in [10]).

5.1 The video data

The video data consists of 171000 values representing the numbers of bits/frame for a JPEG encoded version of the movie Star Wars. The salient statistics of the data are

N = 171000	σ = 6254.2	min = 8622
μ = 27791.2	max = 78459	

5.2 The Ethernet data

The packet data was generated from a well studied Ethernet trace collected at Bellcore Labs [7]. From this trace we constructed the "activity" per second data set (measured in bytes/s). We then differenced this set to obtain a distribution that was less one sided and more suitable for estimation. The statistics of the data are

N = 3143	σ = 79018	min = -377870
μ = 24.3	max = 395970	

5.3 Estimation procedure and results

In section 4.2 we discovered several important points about the performance of the estimation techniques. One point is that Kogon's technique gives good α and β estimates but only when c is close to 1. Therefore McCulloch's technique was applied first to the real data and it was scaled using $\hat{c}$ and $\hat{a}$. Since we found evidence to suggest the FLOM methods work well even when $\beta \neq 0$ we applied them to the data. Finally Kogon's technique was applied and all the estimated values were recorded in tables 2 and 3.

Technique	$\hat{\alpha}$	$\hat{\beta}$	$\hat{c}$	$\hat{a}$
McCulloch	1.898	1.000	4130	27856
Ma	1.985	-	1.052	-
Tsihrintzis	NaN	-	NaN	-0.008
Kogon	1.944	4.338	1.018	-

Table 2. *Results for JPEG video data.*

Technique	$\hat{\alpha}$	$\hat{\beta}$	$\hat{c}$	$\hat{a}$
McCulloch	1.511	0.046	41924	1577
Ma	1.426	-	0.878	-
Tsihrintzis	4.349	-	0.000	-0.033
Kogon	1.788	0.696	1.123	-

Table 3. *Results for Ethernet activity data.*

JPEG results

The $\hat{\alpha}$ was close to 2 for every technique (except Tsihrintzis) which suggests that the $\hat{\beta}$ are very unstable. This explains the large (and impossible) value for $\hat{\beta}$ from Kogon's estimator.

The NaN (Not a Number) results for the Tsihrintizis method result from the fact that the method must take the square-root of a value which in this case was negative. Since no similar result occurred when testing with stable random variables this may suggest the data in question does not approximate to a stable distribution.

In order to determine the performance of the estimation techniques we calculated the approximate pdf of the stable distribution with the estimates found using McCulloch's, Ma's and Kogon's techniques. The results are compared with the pdf of the original data in figure 9.

Ethernet results

In this case, even though the distribution was again symmetric in appearance (see figure 10) the $\hat{\alpha}$ were significantly less than 2 in all cases. This suggests that the distribution is more impulsive than for the JPEG case. It also means that $\hat{\beta}$ are more stable and have more influence over the shape of the distribution. An interesting point is that it is the Ma technique's estimate of α that achieves the best match, even though Kogon's technique estimates $|\beta| > 0$.

6 Conclusions

In this paper we have directly compared four quick stable parameter estimation techniques, with both stable random variables and two forms of real data. The results contain a wealth of information but we have only space to make a few salient observations.

These experiments were carried out to learn more about how we should apply these techniques to real data estimation. In that respect we suggest that:-

1. Scaling is critical for Kogon's technique.

2. Tsihrintzis and Ma's techniques can be applied, even if it appears that $|\beta| > 0$.

3. More than one technique should be applied until more about the performance of these techniques is understood.

Tables 2 and 3 record our estimates for two real data sets and the difference between the results for the different techniques are marked in some cases. Until we learn more about why this is the case we should be careful when applying these techniques to real data.

Appendix A

C code for the generation and estimation techniques discussed in this paper is available from the authors on request.

References

[1] J. M. Chambers, C. L. Mallows, and B. W. Stuck. A method for simulating stable random variables. *Journal of the American Statistical Association*, 71(354):340–344, June 1976.

[2] W. H. DuMouchel. *Stable distributions in statistical inference*. PhD thesis, Yale University, 1971.

[3] E. F. Fama and R. Roll. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66(334):331–338, June 1971.

[4] S. Kogon and D. Williams. On the characterization of impulsive noise with α-stable distributions using fourier techniques. In *Proceedings of the 29th Asilomar Conference of Signals ,Systems and Computing*, 1995.

[5] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, December 1980.

[6] I. A. Koutrouvelis. An iterative procedure for the estimation of the parameters of stable laws. *Communications in Statistics - Simulation and Computation*, 10(1):17–28, 1981.

[7] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–14, February 1994.

[8] X. Ma and C. L. Nikias. Parameter estimation and blind channel identification in impulsive signal enviroments. *IEEE Transactions on Signal Processing*, 43(12):2884–2897, December 1995.

[9] J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications on Statistics - Simulation*, 15(4):1109–1136, 1986.

[10] C. L. Nikias and M. Shao. *Signal processing with alpha-stable distributions and applications*. John Wiley & sons, 1996.

[11] G. Samorodnitsky and M. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models and Infinite Variance*. Chapman & Hall, 1994.

[12] G. A. Tsihrintzis and C. L. Nikias. Fast estimation of the parameters of alpha-stable impulsive interference. *IEEE Transactions on Signal Processing*, 44(6):1492–1503, June 1996.

Figures

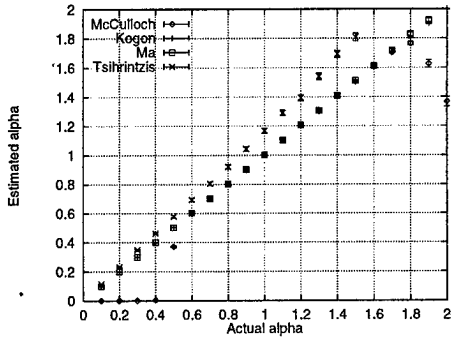


Figure 1. *Results for $\hat{\alpha}$ when $\beta = 0$.*

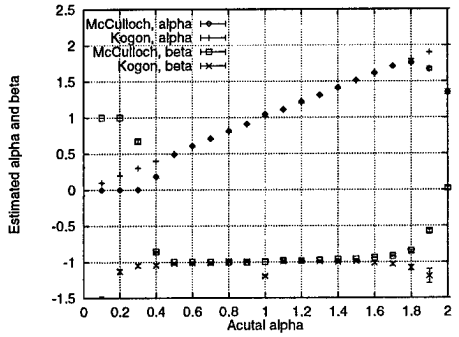


Figure 2. *Results for $\hat{\alpha}$ and $\hat{\beta}$ when $\beta = -1$.*

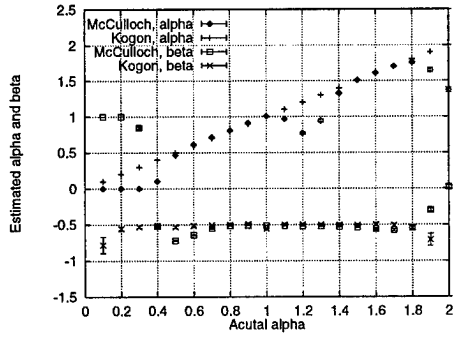


Figure 3. *Results for $\hat{\alpha}$ and $\hat{\beta}$ when $\beta = -0.5$.*

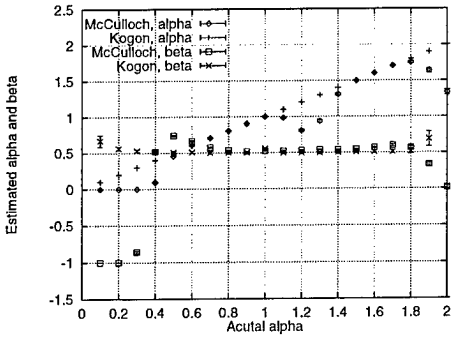


Figure 4. *Results for $\hat{\alpha}$ and $\hat{\beta}$ when $\beta = 0.5$.*

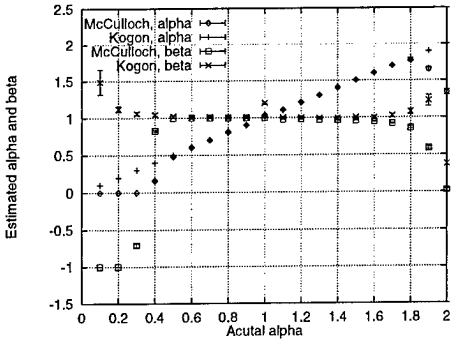


Figure 5. *Results for $\hat{\alpha}$ and $\hat{\beta}$ when $\beta = 1$.*

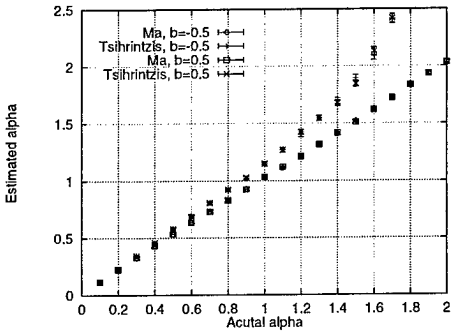


Figure 6. *Results for $\hat{\alpha}$ when the FLOM based techniques are applied to a skewed distribution.*

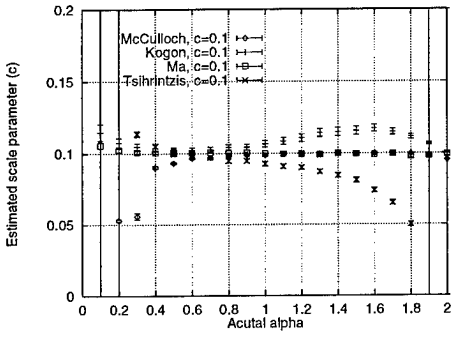


Figure 7. *Results for $\hat{c}$ when k=0.1.*

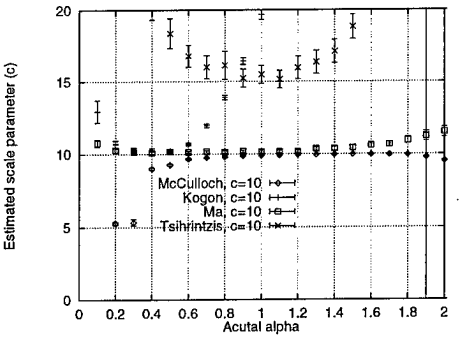


Figure 8. *Results for $\hat{c}$ when k=10.*

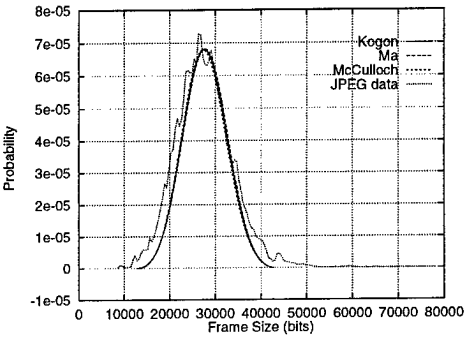


Figure 9. *The real and estimated stable pdfs for the JPEG data.*

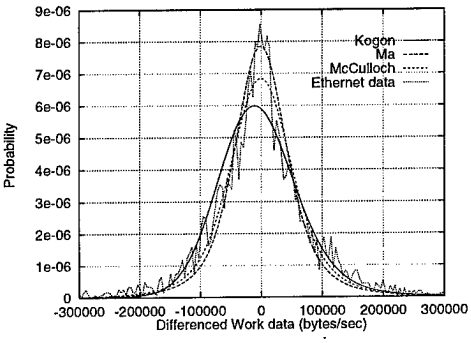


Figure 10. *The real and estimated stable pdfs for the Ethernet data.*

Fourier Series Based Nonminimum Phase Model for Second- and Higher-Order Statistical Signal Processing

Chong-Yung Chi *
Department of Electrical Engineering
National Tsing Hua University,
Hsinchu, Taiwan, Republic of China
cychi@ee.nthu.edu.tw

Abstract

In the paper, a parametric Fourier series based model (FSBM) for or as an approximation to an arbitrary nonminimum-phase linear time-invariant (LTI) system is proposed for statistical signal processing applications where a model for LTI systems is needed. Based on the FSBM, a (minimum-phase) linear prediction error (LPE) filter for amplitude estimation of the unknown LTI system together with the Cramer Rao (CR) bounds is presented. Then an iterative algorithm for obtaining the optimum mean-square LPE filter with finite data is presented which is also an approximate maximum likelihood algorithm when data are Gaussian. Then three iterative algorithms using higher-order statistics with finite non-Gaussian data are presented for estimating parameters of the FSBM followed by some simulation results to support the efficacy of the proposed algorithms. Finally, we draw some conclusions.

1. Introduction

In many statistical signal processing areas such as signal modeling, power spectral and polyspectral estimation, system identification, deconvolution and equalization, a widely known problem is the identification and estimation of an unknown linear time-invariant (LTI) system $h(n)$ driven by an unknown random signal $u(n)$ with only a given set of output measurements $x(n)$

$$x(n) = u(n) * h(n) = \sum_{k=-\infty}^{\infty} u(k)h(n-k) \qquad (1)$$

*This work is supported by the National Science Council under Grant NSC 86-2213-E-007-037.

The system function $H(z)$ is often modeled as a parametric rational function such as autoregressive (AR) model, moving average (MA) model and autoregressive moving average (ARMA) model, and therefore finding a rational model approximation to the system from data becomes a parameter estimation problem.

Recently, Chien, Yang and Chi [1] proposed a parametric cumulant based method for estimating the phase of the unknown system $h(n)$ through allpass filtering of measurements $x(n)$ when $x(n)$ is non-Gaussian. Their method is applicable for both 1-D and 2-D systems. They used a Fourier series based model (FSBM) for an allpass filter which leads to a consistent estimate for the system phase by maximizing a single absolute higher-order cumulant of the allpass filter output.

In this paper, an FSBM for or as an approximation to an arbitrary nonminimum-phase LTI system is proposed for applications in the aforementioned statistical signal processing areas. A linear prediction error (LPE) filter based on the FSBM for amplitude estimation of the system is proposed, and then estimation of FSBM (amplitude and phase) parameters is presented followed by some simulations results and conclusions.

2. Nonminimum-phase FSBM

Assume that $h(n)$ is a real nonminimum-phase LTI system with the frequency response $H(\omega) = H(z = \exp\{j\omega\}) = H^*(-\omega)$ defined as

$$H(\omega) = \exp\left\{\sum_{i=1}^{p} \alpha_i \cos(i\omega) + j\sum_{i=1}^{p} \beta_i \sin(i\omega)\right\} \qquad (2)$$

Two advantages of the pth-order nonminimum-phase FSBM defined by (2) over the rational model (i.e., AR, MA and ARMA models) for stable LTI systems are discussed as follows.

0-8186-8005-9/97 $10.00 © 1997 IEEE

The FSBM given by (2) is always a stable IIR system no matter whether it is causal or noncausal. Therefore, in practical applications where the system to be designed is a noncausal stable system such as noncausal inverse filter $1/H(\omega)$ (when $h(n)$ is not minimum-phase) in blind deconvolution and channel equalization, the FSBM given by (2) is more suitable than the ARMA model because the stability issue is never existent for the former, thus leading to more efficient design and processing procedure.

The complex cepstrum $\widetilde{h}(n)$ (inverse Fourier transform of $\ln[H(\omega)]$), which has been used in speech deconvolution, source separation of speech signals and seismic deconvolution, associated with the FSBM given by (2) can be easily shown to be

$$\begin{aligned}\widetilde{h}(n) &= \tfrac{1}{2}\textstyle\sum_{i=1}^{p}\alpha_i\{\delta(n+i)+\delta(n-i)\} \\ &+\tfrac{1}{2}\textstyle\sum_{i=1}^{p}\beta_i\{\delta(n+i)-\delta(n-i)\}\end{aligned} \tag{3}$$

without need of finding poles and zeros of the system [2] as the ARMA model requires.

Next, let us present minimum-phase FSBM and allpass FSBM. The FSBM for $H(\omega)$ given by (2) can also be expressed as

$$H(\omega) = H_{\rm MP}(\omega)\cdot H_{\rm AP}(\omega) \tag{4}$$

where the FSBM

$$H_{\rm MP}(\omega) = \exp\left\{\sum_{i=1}^{p}\alpha_i\cos(i\omega) - j\sum_{i=1}^{p}\alpha_i\sin(i\omega)\right\} \tag{5}$$

can be shown to be a causal stable minimum-phase system with $|H_{\rm MP}(\omega)| = |H(\omega)|$ and $h_{\rm MP}(0) = 1$, and $H_{\rm AP}(\omega)$ is also an allpass FSBM given by

$$H_{\rm AP}(\omega) = \exp\left\{j\sum_{i=1}^{p}\gamma_i\sin(i\omega)\right\} \tag{6}$$

where

$$\gamma_i = \alpha_i + \beta_i \tag{7}$$

3. FSBM for LPE filters

Let us briefly review the conventional LPE filter for ease of later need for the presentation of the FSBM for LPE filters.

A. Conventional LPE filters

Assume that $x(n)$ is a real stationary random process modeled by (1) where $h(n)$ is a stable LTI system driven by a white noise $u(n)$ with zero mean and variance σ^2. The conventional pth-order LPE filter [3]

$$A_p(z) = 1 + \sum_{i=1}^{p} a_i z^{-i} \tag{8}$$

(a causal FIR filter) processes $x(n)$ such that the prediction error

$$e(n) = x(n) * a_n = x(n) + \sum_{k=1}^{p} a_k x(n-k) \tag{9}$$

has minimum variance or average power $E[e^2(n)]$. The optimum LPE filter $\widehat{A}_p(z)$ is minimum-phase and can be obtained by the following orthogonality principle:

$$E[e(n)x(n-k)] = 0, \quad k = 1,2,...,p \tag{10}$$

leading to a set of symmetric Toeplitz linear equations of $\boldsymbol{a}_p = (a_1,...,a_p)^T$. When $x(n)$ is an AR(p) Gaussian process, for any unbiased estimates $\widehat{\boldsymbol{a}}_p$ and $\widehat{\sigma^2}$ with finite data, their covariance matrix is lower bounded by the following Cramer Rao (CR) bounds [3]:

$$\mathbf{C}_{\widehat{\boldsymbol{a}}_p,\widehat{\sigma^2}} \geq \frac{\sigma^2}{N}\begin{bmatrix}\mathbf{R}_{xx}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 2\sigma^2\end{bmatrix} \tag{11}$$

where N is the total number of data and $\mathbf{R}_{xx} = E[\mathbf{x}\mathbf{x}^T]$ in which $\mathbf{x} = (x(n),...,x(n-p+1))^T$.

B. LPE filters with FSBM

Let the pth-order LPE filter $v_p(n)$ be a causal stable minimum-phase IIR filter with $v_p(0) = 1$ and

$$V_p(\omega) = \exp\left\{\sum_{i=1}^{p}a_i\cos(i\omega) - j\sum_{i=1}^{p}a_i\sin(i\omega)\right\} \tag{12}$$

and thus the prediction error is given by

$$e(n) = x(n) * v_p(n) = x(n) + \sum_{k=1}^{\infty} v_p(k)x(n-k) \tag{13}$$

Note that we have used the same notations a_i for parameters of both the proposed FSBM LPE filter $V_p(\omega)$ and the conventional LPE filter $A_p(z)$. The optimum LPE filter $\widehat{V}_p(\omega)$ is described in the following theorem.

Theorem 1. Assume that $H(\omega)$ is an FSBM given by (2) with order equal to q instead of p, and $e(n)$ is the prediction error given by (13) with the LPE filter order $p \geq q$. Then the optimum LPE filter $\widehat{V}_p(\omega) = 1/H_{\rm MP}(\omega)$ with $\min\{E[e^2(n)]\} = E[u^2(n)] = \sigma^2$.

The optimum prediction error $e(n)$ must satisfy $\partial E[e^2(n)]/\partial a_k = 0$ from which one can prove the following orthogonality principle:

$$E[e(n)e(n-k)] = r_{ee}(k) = 0, \quad k = 1,2,...,p \tag{14}$$

However, (14) forms a set of nonlinear equations $\boldsymbol{a}_p$. Nevertheless, $e(n)$ will be a white process as p is sufficiently large which implies that $\widehat{V}_p(\omega) = \widehat{A}_p(\omega)$ (identical whitening filter) for $p = \infty$.

Based on Theorem 1, an iterative algorithm is used to estimate or find an approximation to $H_{\rm MP}(\omega)$ with finite data $x(0)$, $x(1)$, ..., $x(N-1)$ as follows:

Algorithm 1. Amplitude estimation

(S1) Set $p_{\max}$ (maximum of p), parameter L, increment parameter $s \geq 1$, convergence parameter ξ and $t = 0$ (iteration number).

(S2) Set $t = t + 1$, $p = s \times t$ and $\boldsymbol{a}_p = (a_1, ..., a_p)^T$. Search for the minimum of the objective function

$$J(\boldsymbol{a}_p) = \frac{1}{N-L} \sum_{n=L}^{N-1} e^2(n) \qquad (15)$$

and the associated optimum $\widehat{\boldsymbol{a}}_p$ by a gradient type iterative optimization algorithm (such as the well-known Fletcher-Powell algorithm) with the initial condition $\boldsymbol{a}_p(0) = (\widehat{\boldsymbol{a}}_{p-s}^T, \boldsymbol{0}^T)^T$.

(S3) If $p \leq p_{\max}$ and $|J(\widehat{\boldsymbol{a}}_p) - J(\widehat{\boldsymbol{a}}_{p-s})|/J(\widehat{\boldsymbol{a}}_{p-s}) \geq \xi$, go to (S2), otherwise

$$\widehat{H}_{\rm MP}(\omega) = 1/\widehat{V}_p(\omega) \ \ (\text{or } \widehat{\alpha}_i = -\widehat{a}_i) \qquad (16)$$

$$\widehat{\sigma^2} = J(\widehat{\boldsymbol{a}}_p) \qquad (17)$$

The optimum prediction error $e(n) = x(n) * \widehat{v}_p(n)$ corresponds to amplitude equalized data, and the gradient of $J(\boldsymbol{a}_p)$ with respect to a_k needed in (S2) can be shown to be

$$\frac{\partial J(\boldsymbol{a}_p)}{\partial a_k} = 2\widehat{r}_{ee}(k) = \frac{2}{N-L} \sum_{n=L}^{N-1} e(n)e(n-k) \qquad (18)$$

When $H(\omega)$ is a pth-order FSBM and $x(n)$ is Gaussian, it can be shown that both $\widehat{\boldsymbol{\alpha}}_p = -\widehat{\boldsymbol{a}}_p$ and $\widehat{\sigma^2}$ are approximate maximum-likelihood estimates. Moreover, the CR bounds associated with any unbiased estimates $\widehat{\boldsymbol{\alpha}}_p$ and $\widehat{\sigma^2}$ are given by

$$\mathbf{C}_{\widehat{\boldsymbol{\alpha}}_p, \widehat{\sigma^2}} \geq \frac{1}{N} \begin{bmatrix} \mathbf{I} & \boldsymbol{0} \\ \boldsymbol{0}^T & 2\sigma^4 \end{bmatrix} \qquad (19)$$

where $\mathbf{I}$ is a $p \times p$ identity matrix. Note that the CR bounds associated with AR parameters (see (11)) depend on correlations of $x(n)$, while those associated with α_i are uniform and independent of correlations of $x(n)$. The CR bound associated with σ^2 is the same for both FSBM and AR model.

4. Estimation of FSBM parameters

In this section, further with the assumption that $u(n)$ is non-Gaussian with nonzero Mth-order (≥ 3) cumulant (and thus $x(n)$ is also non-Gaussian), three iterative algorithms are to be presented for the estimation of parameters of the FSBM $H(\omega)$ given by (2).

The first two algorithms estimate the system amplitude using Algorithm 1 and system phase using Chien, Yang and Chi's phase estimation algorithm [1] which maximizes a single absolute Mth-order (≥ 3) sample cumulant, denoted $|\widehat{C}_{M,y}|$, of the phase equalized (allpass filtered) data

$$y(n) = x(n) * g_{\rm AP}(n) \qquad (20)$$

where $g_{\rm AP}(n)$ is a pth-order allpass FSBM

$$G_{\rm AP}(\omega) = \exp\left\{ j \sum_{i=1}^{p} b_i \sin(i\omega) \right\} \qquad (21)$$

It has been shown in [1] that the optimum $\widehat{G}_{\rm AP}(\omega)$ turns out to be a phase equalizer, i.e.,

$$\arg\{\widehat{G}_{\rm AP}(\omega)\} = -\arg\{H(\omega)\} + \omega\tau \qquad (22)$$

Because $|\widehat{C}_{M,y}|$ is a highly nonlinear function of b_i, one can use gradient type iterative algorithms for finding the optimum b_i. It has also been proven in [1] that

$$\frac{\partial y(n)}{\partial b_i} = \frac{1}{2}\{y(n+i) - y(n-i)\} \qquad (23)$$

which is needed for computing the gradient of $|\widehat{C}_{M,y}|$ with respect to b_i. Next, let us present Algorithms 2 and 3, respectively,

Algorithm 2.

(S1) Estimate $H_{\rm MP}(\omega)$ and σ^2 using Algorithm 1.

(S2) Find the optimum allpass FSBM $G_{\rm AP}(\omega)$ given by (21) using a gradient type iterative algorithm such that $|\widehat{C}_{M,y}|$ is maximum where $y(n) = x(n) * g_{\rm AP}(n)$. Then obtain the estimate $\widehat{\beta}_i = -\widehat{b}_i$.

Algorithm 3.

(S1) Estimate $H_{\rm MP}(\omega)$ and σ^2, and obtain the optimum prediction error $e(n) \simeq u(n) * h_{\rm AP}(n)$ using Algorithm 1.

(S2) Find the optimum allpass FSBM $G_{\rm AP}(\omega)$ given by (21) using a gradient type iterative algorithm such that $|\widehat{C}_{M,y}|$ is maximum where $y(n) = e(n) * g_{\rm AP}(n)$. Then obtain the estimate $\widehat{\gamma}_i = -\widehat{b}_i$.

The last algorithm (Algorithm 4) estimates the system amplitude and phase simultaneously using inverse filter criteria [4-7]. Chi and Wu [4] proposed a family of inverse filter criteria which includes Tugnait's criteria [5], Wiggins' criterion [6] and Shalvi and Weinstein's

criterion [7] as special cases. The inverse filter $h_{\rm INV}(n)$ is estimated by maximizing

$$J_{r,m} = \frac{|\widehat{C}_{m,\widehat{u}}|^r}{|\widehat{C}_{r,\widehat{u}}|^m} \tag{24}$$

where r is even and $m > r$, and

$$\widehat{u}(n) = x(n) * h_{\rm INV}(n) \tag{25}$$

It has been shown in [4] that $\widehat{u}(n) = bu(n-\tau)$ when $h(n)$ is an arbitrary stable LTI system where b is a scale factor and τ is an unknown time delay. Next, let us present Algorithm 4.

Algorithm 4.

(S1) Set integer $r \geq 2$ (even) and integer $m > r$. Let $H_{\rm INV}(\omega) = 1/H(\omega)$ as given by (2) with α_i and β_i replaced by $-\alpha_i$ and $-\beta_i$, respectively.

(S2) Find the optimum $H_{\rm INV}(\omega)$ (i.e., α_i and β_i) using a gradient type iterative algorithm such that $J_{r,m}$ is maximum. Then σ^2 is estimated as the sample variance of the obtained optimum inverse filter output $\widehat{u}(n)$.

It can be shown that

$$\frac{\partial \widehat{u}(n)}{\partial \alpha_i} = -\frac{1}{2}\{\widehat{u}(n+i) + \widehat{u}(n-i)\} \tag{26}$$

$$\frac{\partial \widehat{u}(n)}{\partial \beta_i} = -\frac{1}{2}\{\widehat{u}(n+i) - \widehat{u}(n-i)\} \tag{27}$$

which are needed for computing the gradient of $J_{r,m}$ with respect to α_i and β_i in (S2), respectively.

Notice that when $H(\omega)$ is not minimum-phase, the FSBM (noncausal stable IIR system) is well suited for the noncausal inverse filter $H_{\rm INV}(\omega) = 1/H(\omega)$ as mentioned in Section 2. When $H(\omega)$ is an FSBM, the optimum $\widehat{u}(n) \simeq u(n)$ (without scale factor and time delay between $\widehat{u}(n)$ and $u(n)$).

5. Simulation results

In this section, let us show two sets of simulation results. For the first simulation, the driving input $u(n)$ was assumed to be a zero-mean i.i.d. Exponentially distributed random sequence. An FSBM of order $p = 5$ was used for the system $H(\omega)$, whose amplitude and phase (solid lines) are shown in Figures 1a and 1b, respectively. Thirty independent runs were performed using Algorithm 3 with $p_{\max} = s = 5$, $L = 0$ and $M = 3$. Mean (dashed line) and mean$\pm$standard deviation (dotted lines) of the obtained thirty amplitude and phase estimates for $N = 1024$ and $SNR = 20$ dB (white Gaussian noise) are also shown in Figures 1a and 1b, respectively. One can see that both the amplitude and phase estimates are unbiased with small variance. Moreover, the results obtained by Algorithms 2 and 4 are similar to those shown in Figure 1.

The second set of simulation results for seismic deconvolution was obtained with $u(n)$ assumed to be a Bernoulli-Gaussian sequence and the system (source wavelet) $h(n)$ to be a third-order nonminimum-phase causal ARMA system (taken from [4]) instead of an FSBM. Figure 2a shows the synthetic data $x(n)$ for $N = 512$ and $SNR = 20$ dB (white Gaussian noise). Figures 2b and 2c show the (noncausal) estimate $\widehat{h}(n)$ (dotted line) and the deconvolved signal $\widehat{u}(n)$ (dotted line), respectively, obtained using Algorithm 4 with $p = 12$, $r = 2$ and $m = 4$, where the scale factor as well as the time delay between $\widehat{h}(n)$ and $h(n)$ (solid line) and the time delay between $\widehat{u}(n)$ and $u(n)$ (solid line) were artificially removed. One can see that $\widehat{h}(n)$ and $\widehat{u}(n)$ are good approximations to $h(n)$ and $u(n)$, respectively. The above simulation results support the efficacy of the proposed algorithms.

6. Conclusions

We have presented an FSBM for or as an approximation to an arbitrary nonminimum-phase LTI system for applications in the statistical signal processing areas mentioned in the introduction section. Based on the FSBM, an LPE filter for amplitude estimation together with the CR bounds, an algorithm for obtaining the optimum LPE filter, and three algorithms for estimating FSBM parameters were presented. All the gradient type iterative optimization algorithms used in the proposed algorithms have a computationally efficient parallel structure (FIR filter banks with only two nonzero coefficients $1/2$ or $-1/2$) (see (23), (26) and (27)). However, the gradient computation associated with Algorithm 1 (see (18)) does not need any further processing to the prediction error $e(n)$. Finally, two sets of simulation results were presented to support the efficacy of the proposed algorithms.

References

[1] H.-M. Chien, H.-L. Yang and C.-Y. Chi, "Parametric cumulant based phase estimation of 1-D and 2-D nonminimum phase systems by allpass filtering," to appear in IEEE Trans. Signal Processing.

[2] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, 1989.

[3] S.M. Kay, *Modern Spectral Estimation*, Prentice-Hall, 1988.

[4] C.-Y. Chi and M.-C. Wu, "Inverse filter criteria for blind deconvolution and equalization using two cumulants," Signal Processing, vol. 43, no. 1, pp. 55-63, April 1995.
[5] J.K. Tugnait, "Estimation of linear parametric models using inverse filter criteria and higher order statistics," IEEE Trans. Signal Processing, vol. 41, no. 11, pp. 3196-3199, Nov. 1993.
[6] R.A. Wiggins, "Minimum entropy deconvolution," Geoexploration, vol. 16, pp. 21-35, 1978.
[7] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," IEEE Trans. Information Theory, vol. 36, pp. 312-321, March 1990.

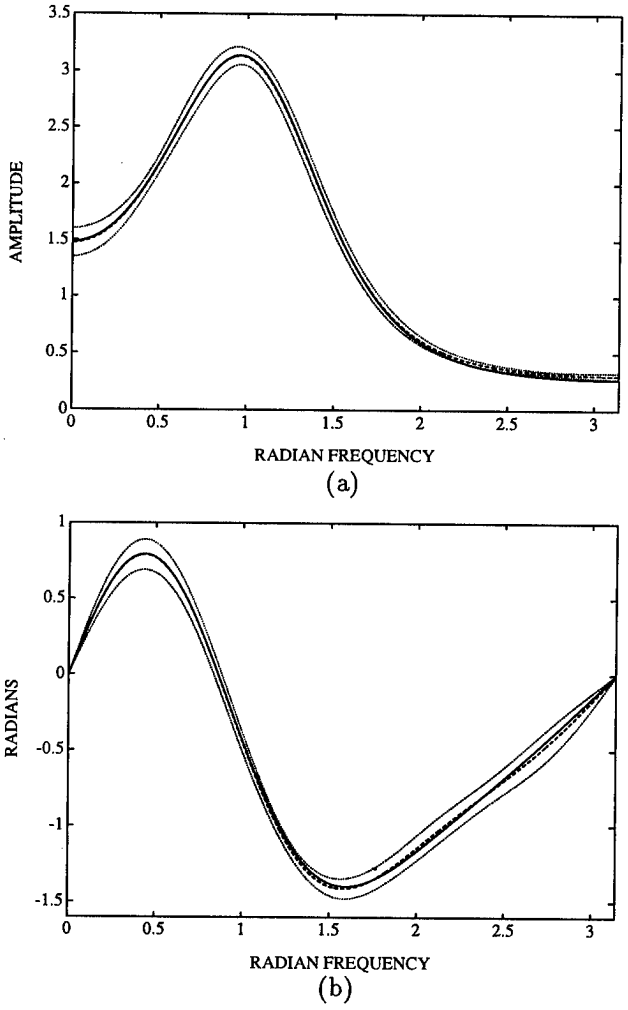


Figure 1. Simulation results for $N = 1024$ and $SNR = 20$ dB using Algorithm 3. Mean (dashed lines) and mean±standard deviation (dotted lines) of thirty (a) amplitude and (b) phase estimates together with the amplitude and phase (solid lines) of the system.

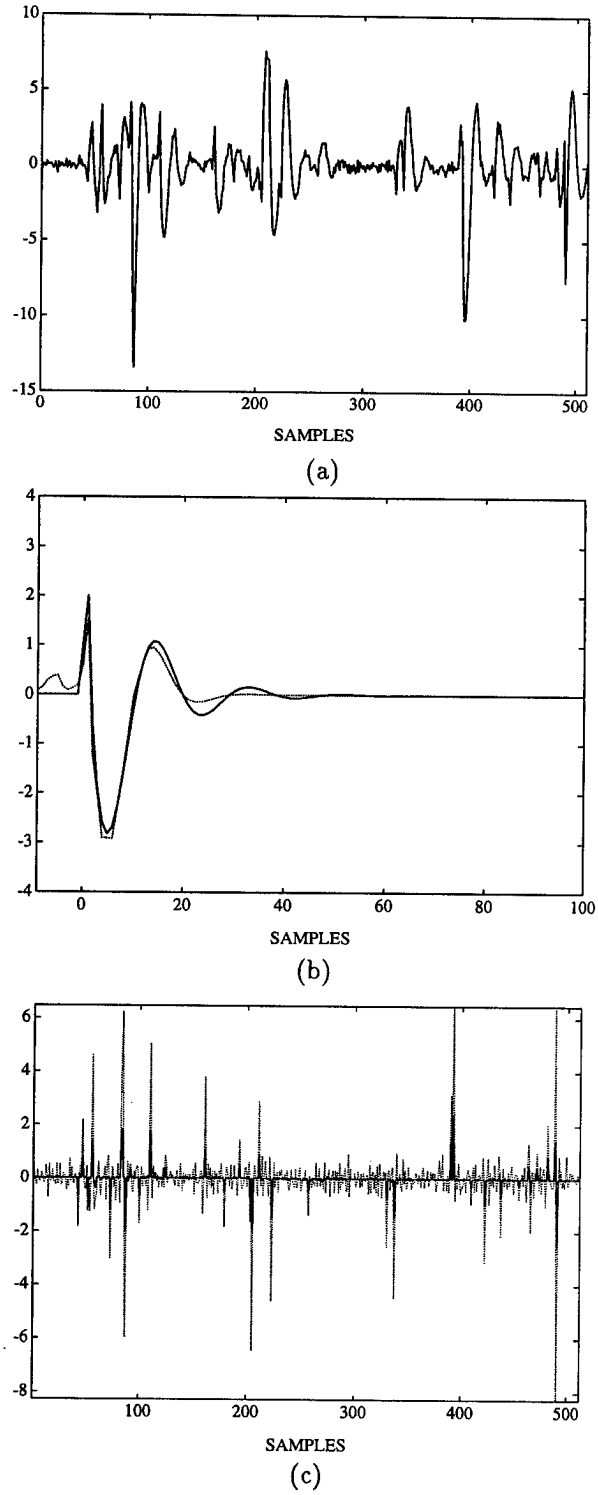


Figure 2. Simulation results for seismic deconvolution ($N = 512$ and $SNR = 20$ dB) using Algorithm 4. (a) Synthetic data $x(n)$; (b) source wavelet $h(n)$ (solid line) and estimate $\widehat{h}(n)$ (dotted line); and (c) input $u(n)$ (solid line) and deconvolved signal $\widehat{u}(n)$ (dotted line).

Variability in Higher Order Statistics of Measured Shallow-Water Shipping Noise

Lisa A. Pflug
Naval Research Laboratory, Code 7173, Stennis Space Center, MS 39529
Tel: 601 688-5574 FAX: 601 688-5341 , E-mail: pflug@pfilbert.nrlssc.navy.mil

George E. Ioup, Juliette W. Ioup
Department of Physics, University of New Orleans, New Orleans, LA 70148

Pamela Jackson
Naval Research Laboratory, Code 7173, Stennis Space Center, MS 39529

Abstract

Many underwater acoustic signal processing algorithms are designed for use in stationary and/or Gaussian noise. While these assumptions are often valid for applications in deep water ocean areas, they may not be appropriate for shallow water areas, especially in the presence of local shipping activity. Local shipping also produces spatial correlation in the noise and introduces additional complexity for multichannel processing. In this paper, two 30-minute sets of ambient ocean noise, recorded near the San Diego, California coast, are analyzed for stationarity and Gaussianity using the Kolmogorov-Smirnov test. Since processing algorithms based on higher order statistics often assume Gaussianity, time-dependent fluctuations in the third and fourth order cumulants are also analyzed. The analysis reveals significant variability in the time lengths of stationary periods, and episodic periods of nonGaussianity that last for up to five minutes. Statistical fluctuations appear predominantly in the second and fourth order cumulants rather than the third order cumulant. The shipping noise is also shown to be correlated between pairs of hydrophones with the level of correlation varying over time and the correlation ranging from positive to negative with increasing channel separation.

1. Introduction

The assumption that ambient noise is Gaussian has led to attempts to exploit the higher order statistics of underwater acoustic signals for detection. Algorithms using higher order statistics for both stationary and transient signal detection have shown good performance in simulations with Gaussian noise. However, if the noise is nonGaussian, then the performance of these detectors can be degraded. Real noise dominated by shipping is likely to have time-varying cumulants, leading to problems with algorithms that assume stationarity. The individual requirements of a signal processing algorithm must be taken into consideration when testing for stationarity and Gaussianity. For example, power signal detectors, which use averaging to achieve noise suppression, perform best in the presence of long periods of stationarity, while transient signal detectors generally require short periods of stationarity or none at all.

In underwater acoustics, it is well accepted that shipping noise is a dominant noise source at frequencies between 20 and 500 Hz and that there are time-dependent fluctuations in the noise [1]. The time scale of fluctuations depends on the number, density, and operating characteristics of ships, the proximity of shipping traffic to the receiver, the receiver location in the environment, and the propagation characteristics of the area. The complexity of shallow-water environments in conjunction with local shipping requires careful consideration before making assumptions about ambient noise.

Several studies of the higher order statistics of ocean noise have been published. In Ref. [2], Brockett et al. use a third order Gaussianity test to show that ambient ocean noise is Gaussian for periods of over a minute, but nonGaussian for shorter periods, while local shipping noise (due to one nearby ship) is nonGaussian for both the longer and shorter time periods. A normalized third order statistic is used in Ref. [3] to determine that a 45s

0-8186-8005-9/97 $10.00 © 1997 IEEE

deep-water noise sequence is nonGaussian. In Ref. [4], the noise generated by the towing platform in an experiment is shown to have strong bispectral components and is thus nonGaussian. The trispectrum is used in Ref. [5] to conclude that a time series generated by two ships passing about 460m and 635m from a sonobuoy does not differ significantly from a Gaussian distribution. These studies indicate that even in deep water, ship proximity is an important factor in the noise field.

In this work, two 30-minute sets of measured shallow water ambient noise dominated primarily by local shipping are investigated with the goal of determining whether stationarity and Gaussianity assumptions are appropriate, and if so, for what time periods. Third and fourth order cumulant calculations of the noise are used to give further insight into the potential effects of the noise on algorithms that are based on higher order statistics. Spatial correlation in the data is also investigated.

The noise was recorded at 1500 samples/second on a vertical array placed in 200m water near the port of San Diego, California [6]. The stationarity and Gaussianity analyses are restricted to noise recorded on two of the hydrophones, or channels: channel 2 at a depth of approximately 192 m, and channel 43 at a depth of approximately 116 m. The correlation analysis includes data from channels 16 (170m) and 31 (140m) in addition to channels 2 and 43. Ship locations during the experiment were tracked using radar and updated approximately every two minutes [7]. In both noise sets, several ships are moving close to the array [8].

2. Time-dependent cumulant fluctuations

A 1-second sliding window is used to calculate the fluctuating first through fourth order cumulants of the noise. A 15-point overlap of the sliding window corresponding to a cumulant sampling rate of 100 samples/second was found sufficient to avoid aliasing distortions in the time-variation of the second, third, and fourth order cumulants. The 30-minute noise sets are divided into segments of approximately 1-minute duration for processing. The average cumulants for the set 1 noise, channels 2 and 43, are shown in Figs. 1 and 2. In these figures, each circle represents the average cumulant over a 1-minute segment, and the dashed curves represent the sliding window minimum and maximum cumulant values within the 1-minute segment. The corresponding curves for the set 2 noise are shown in Figs. 3 and 4.

Visual inspection of the cumulants over the two noise sets reveals local periods of nonstationarity in the noise, primarily evident in the second and fourth order cumulants. Increases in the fourth order cumulant also indicate periods of nonGaussianity. The level of shipping activity near the array is moderate to high during times that the noise in set 1 and set 2 was recorded. The increased variance and nonzero kurtosis during the 0-5 minute and 20-30 minute segments of the set 1 noise, and during the 7-10 and 17-24 minute segments of the set 2 noise can be correlated with ship activity in a 2 km radius of the receiver site. The skewness shows minor fluctuations during these periods. The two periods of increased variances in the set 2 noise are due to two transient noise events, the first having a duration of approximately 2 minutes and the second having a duration of 3-6 minutes, While these noise

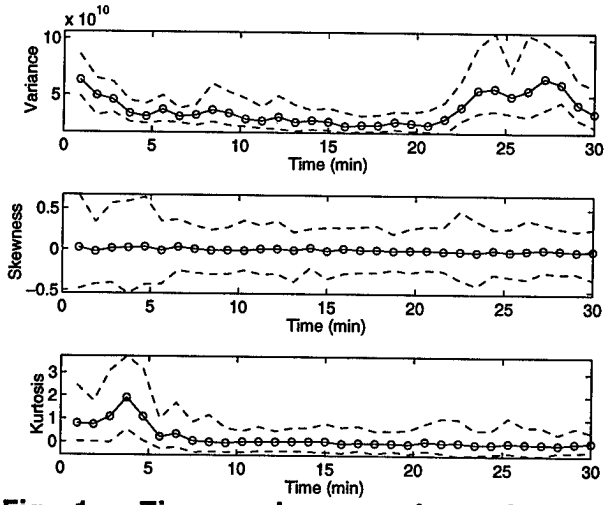


Fig. 1. Time-varying cumulants for the channel 2, set 1 ambient noise.

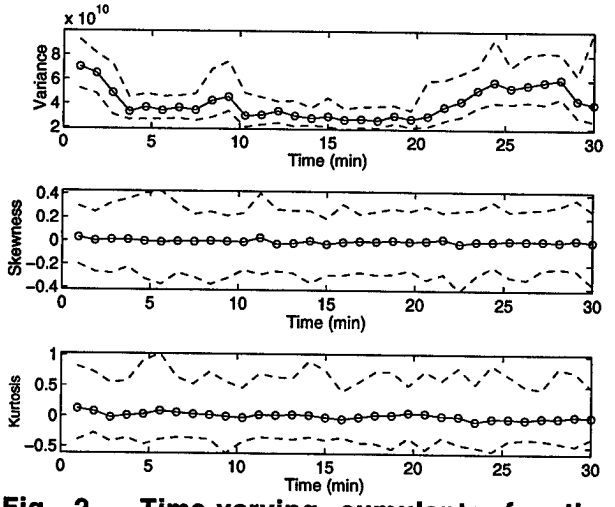


Fig. 2. Time-varying cumulants for the channel 43, set 1 ambient noise.

sources are of unknown origin, they are likely due to one of two ships that are traveling within 0.8 km of the array site during these times. When shipping activity remained outside of a 2 km radius of the receiver site for the two sets of noise, little variability in the noise variance and kurtosis was found, even though there were sometimes ships traveling within 5 km of the receiver site.

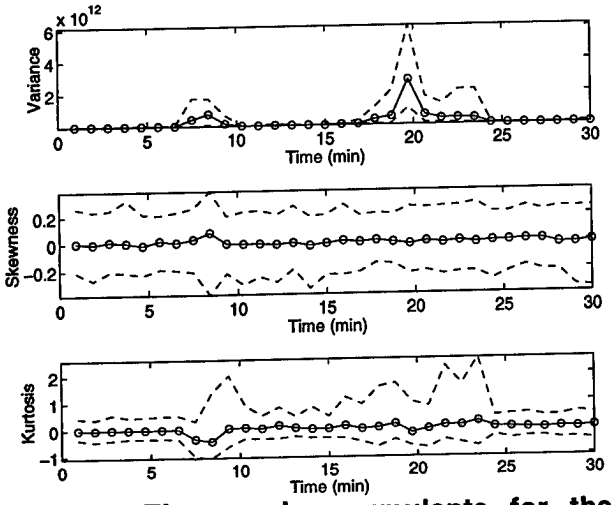


Fig. 3. Time-varying cumulants for the channel 2, set 2 ambient noise.

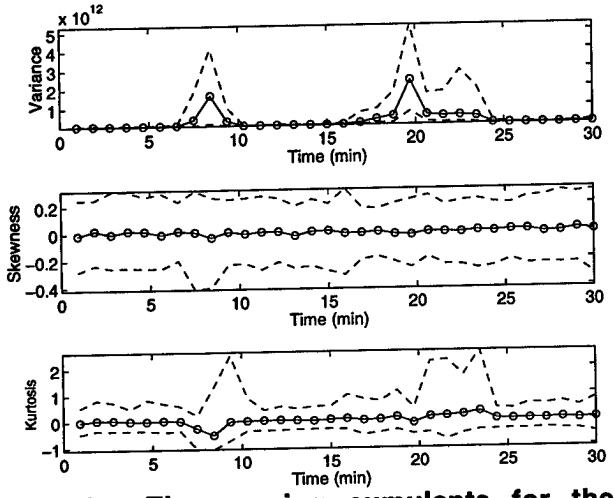


Fig. 4. Time-varying cumulants for the channel 43, set 2 ambient noise.

3. Stationarity and Gaussianity

The cumulant curves in Figs. 1-4 indicate that the shipping noise contains significant periods of nonstationarity. A two-sample Kolmogorov-Smirnov (K-S) test for goodness of fit is used to quantify the periods of stationarity/nonstationarity in the noise. To reduce the computational burden, the first 27.3 minutes of noise from the 30-minute sets are separated into three 9.1-minute segments for analysis. The test is applied successively over the data, using pairs of 1s windows with no overlap. Breaks in stationarity are indicated by significant changes in the cumulative probability distribution of the noise (at the 95% confidence level), as determined by the K-S test. The test reveals that stationary periods in the noise are episodic and vary in time duration from the minimum test resolution of 1.0s to 9.1 minutes, as shown in Fig. 5 for channels 2 and 43 from noise sets 1 and 2. Each circle in Fig. 5 represents a time point at which two successive probability distributions are determined to be different by the K-S test. Conversely, time periods where circles are absent represent periods of stationarity. Trends that appear in the stationarity test results generally correspond to the fluctuating cumulants in Figs. 1-4.

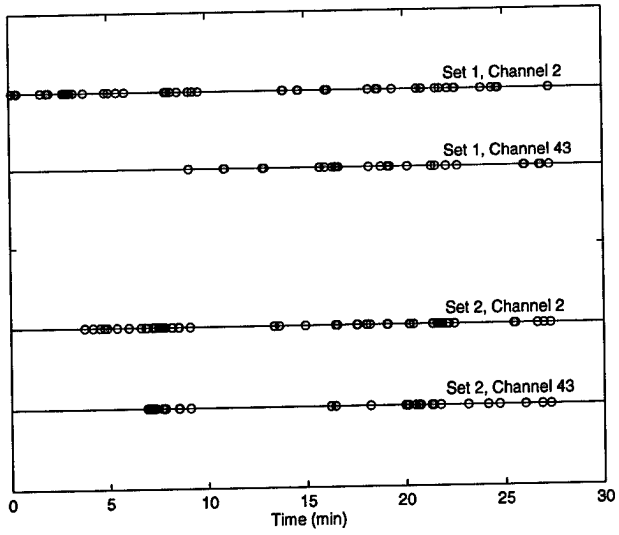


Fig. 5. Stationarity test results for a 1-second processing window.

A discrepancy between the cumulant calculations and the stationarity test appears during the set 1, channel 43 noise during the first 3 minutes of the noise in that the noise appears stationarity by the K-S test but the noise variance is declining. This can be accounted for by noting that the stationarity test results are based on the rate of change in the cumulants within each successive 1-second processing window and is thus most sensitive to abrupt changes, while the cumulant curves in Figs. 1-4 show the average cumulant level over 1 minute. Note

that there are periods of nonstationarity even when the cumulants indicate that the noise is Gaussian, due to changes in the noise variance.

To quantify the periods of nonGaussianity, a one-sample K-S test is applied to the two 30-minute noise sets. Each processing window is determined to be either Gaussian or nonGaussian. The percent of 1-second processing windows that are nonGaussian in each 1-minute noise segment is calculated and shown in Fig. 6. Test results indicate the existence of periods of nonGaussianity in the noise that correlate with the results of the cumulant calculations and some of the periods of nonstationarity. The periods of nonGaussianity generally last for several minutes. However, except for the transient noise events in the set 2 noise, and the 0-5 minute noise segment in the set 1 noise, the noise in most of processing windows is Gaussian. Note that this figure does not reveal the levels of nonGaussianity, as these can be observed in Figs. 1-4, and are discussed in detail in Ref. [8].

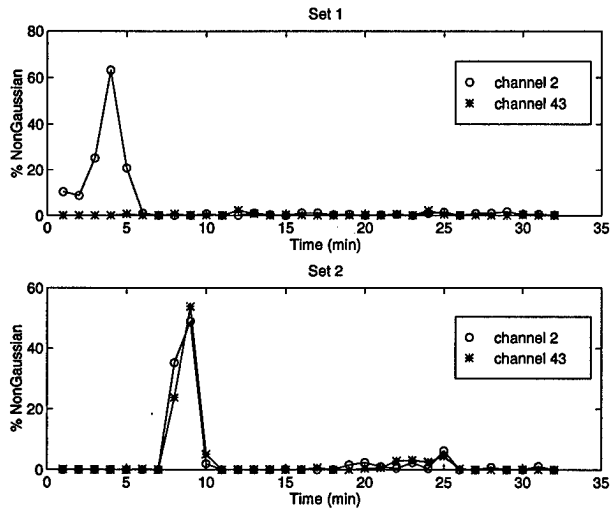


Fig. 6. Percent of nonGaussian processing windows in each one minute segment for channels 2 and 43, noise sets 1 and 2.

4. Spatial correlation

The ambient shipping noise is spatially correlated, i.e., correlated between the different hydrophones. This can affect signal processing algorithms, whose performance is often predicted and/or benchmarked using independent noise. Since the energy detector and some higher order statistics detectors make use of only the zero lag correlation value, the correlation coefficient is a useful measure of the spatial correlation of the measured noise. The zero lag correlation coefficient between two channels of data is given by

$$c_2 = \frac{\sum_{k=0}^{N-1} x_1(t)x_2(t)\Delta t}{\left(\sum_{k=0}^{N-1} x_1^2(t)\Delta t\right)^{1/2}\left(\sum_{k=0}^{N-1} x_2^2(t)\Delta t\right)^{1/2}}$$

with $t = k\Delta t$, and is 1.0 when $x_1(t)=x_2(t)$. In Fig. 7, the time-dependent correlation coefficients of the first three minutes of the set 1 measured noise are shown, for channel 2 correlated with each of channels 16, 31, and 43. These values are calculated using a 1-second processing window and 50% overlap. For independent Gaussian noise and the same processing parameters, the average absolute value of c_2 is 0.02 and the maximum is approximately 0.09. By comparison, the shipping noise is significantly correlated between the pairs of channels for this 3-minute segment. Channels 2 and 16 are positively correlated over the 3-minute time duration, channels 2 and 31 are positively correlated over about half the duration and negatively correlated, or anticorrelated, for the remaining duration. Channels 2 and 43 are negatively correlated for nearly the entire 3-minute segment. These results show that the degree of interchannel correlation changes from positively correlated to negatively correlated as the depth spacing between channels increases. Independent noise is theoretically zero correlated. A positively correlated

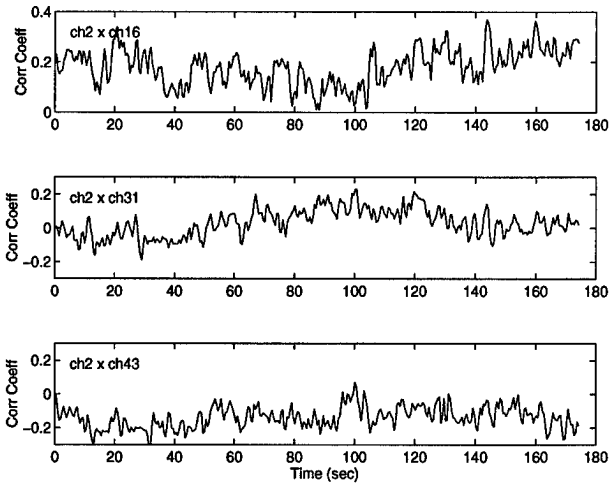


Fig. 7. Time-dependent correlation coefficients from the set 1 noise between channel 2 and (a) channel 16, (b) channel 31, and (c) channel 43 .

signal can be detected better in negatively correlated noise than in either independent or positively correlated noise. Negative correlation could result from an odd number of surface bounces in the path to one hydrophone and an even number to the other for dominant noise sources. This suggests that one should detect on positive and negative signal peaks when signal and noise can have opposite correlation.

To examine the interchannel correlation of the noise over the entire 30-minute segment, the channel 2 and channel 43 noise is divided into approximately 1-minute durations for processing (as done for the cumulant calculations shown above). The mean, minimum, and maximum c_2 values in each 1-minute duration are shown in Fig. 8 for noise sets 1 and 2. These two channels are negatively correlated for most of the 30-minute duration in both sets 1 and 2. The maximum and minimum correlation coefficients over the 30-minute duration are 0.20 and -0.32 for set 1, and 0.86 and -0.75 for set 2.

For the set 1 noise, the degree of negative correlation is larger during the 0-7 minute period when the channel 2 noise is nonGaussian and the channel 43 noise is Gaussian. For the set 2 noise, strong positive and negative c_2 values appear during the two transient noise events during the 7-10 and 17-24 minute segments.

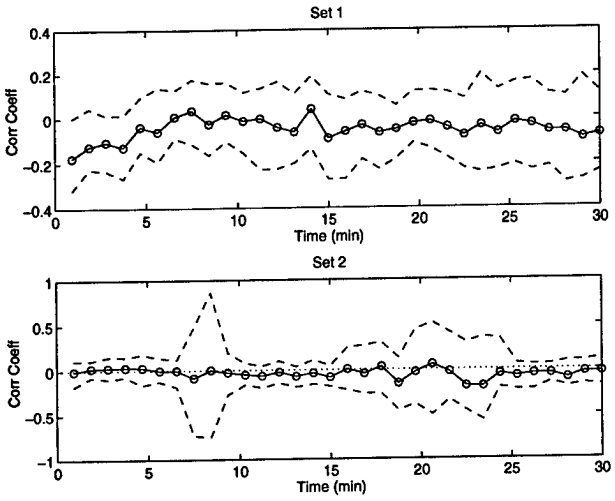


Fig. 8. Time-varying correlation coefficients between channels 2 and channel 43 from (a) the set 1 noise, and (b) the set 2 noise.

5. Conclusions

Two 30-minute sets of ambient shipping noise from a shallow-water harbor area are analyzed for cumulant variability, stationarity, nonGaussianity, and spatial correlation. Calculations of second through fourth order cumulants for two channels of noise recorded on a vertical array indicate that there are significant changes in both variance and kurtosis, in some instances due to loud transient noise events lasting 2-6 minutes, which are probably generated by ships moving within 1 km of the array. A two-sample K-S test is used quantify the periods of stationarity, which last from the minimum test resolution of 1-second to over 9.1 minutes. A one-sample K-S test reveals periods of nonGaussianity that are consistent with the time-varying levels of kurtosis and are episodic, lasting for periods of up to several minutes. Except for these periods, the noise is Gaussian, even though it may be nonstationary. Second and fourth order correlation coefficients of the noise reveal that the noise is spatially correlated between pairs of hydrophones with the correlation ranging from positive to negative as the sensor spacing increases.

Acknowledgments

This work was supported by the Office of Naval Research and the Naval Research Laboratory.

References

[1] Wenz, G. W., "Acoustic Ambient Noise in the Ocean: Spectra and Sources," J. Acoust. Soc. Am., 34, 1936-1956 (1962).

[2] Brockett, P. L., Hinich, M., and Wilson, G. R., "Nonlinear and Non-Gaussian Ocean Noise," J. Acoust. Soc. Am., 82, 1386-1394 (1987).

[3] Richardson, A. M. and Hodgkiss, W. S., "Bispectral Analysis of Underwater Acoustic Data," J. Acoust. Soc. Am., 96, 828-837 (1994).

[4] Hinich, M. J., Marandino, D., and Sullivan, E. J., "Bispectrum of Ship-Radiated Noise," J. Acoust. Soc. Am., 85, 1512-1517 (1989).

[5] Dalle Molle, J. and Hinich, M. J., "Trispectral Analysis of Stationary Random Time Series," J. Acoust. Soc. Am., 97, 2963-2978 (1995).

[6] Bachman, Richard T., Schey, Philip W., Booth, Newell O., and Ryan, Frank J., "Geoacoustic Databases for Matched-Field Processing: Preliminary Results in Shallow Water Off San Diego, California," J. Acoust. Soc. Am., 4, 2077-2085 (1996).

[7] Holliday, D. V., "Observations of Surface Vessel Activity off San Diego Harbor During the Summer of 1994," Tracor Report No. T-95-56-0001-U (1995).

[8] Pflug, L. A., Ioup, G. E., Jackson, P., and Ioup, J. W., "Stationarity and Gaussianity of Shallow-Water Ambient Noise Dominated by Local Shipping," submitted to J. Acoust. Soc. Am. (1997).

Higher Order Moment Detection of Transients in Measured Shallow-Water Noise

Lisa A. Pflug
Naval Research Laboratory, Code 7173, Stennis Space Center, MS 39529
Tel: 601 688-5574 FAX: 601 688-5341, E-mail: pflug@pfilbert.nrlssc.navy.mil

George E. Ioup, Juliette W. Ioup
Department of Physics, University of New Orleans, LA 70148

Abstract

A comparison between second and fourth order moment detectors is made for three low frequency transient signals embedded in measured ambient shipping noise from a shallow-water area. Detector evaluations must be made over short periods of time because of the nonstationary nature of the noise. Results indicate that the fourth order moment detects better than the second order moment for the two nonGaussian transient signals, but not as well for the more Gaussian transient signal. This suggests that the fourth order moment detector be used in addition to the second order moment detector, rather than as a replacement. The results of detection simulations with independent Gaussian noise are given for comparison to the results with the measured shipping noise. The comparison indicates that gains for the fourth order moment detector over the second order moment detector are higher in the shipping noise than in the Gaussian noise for the two nonGaussian signals.

1. Introduction

Since sonar targets have recently become quieter and their signatures more difficult to detect, especially in coastal and shallow-water areas, attempts are being made to exploit target-generated transient signals to improve detection. One approach is to capitalize on the nonGaussian nature that often characterizes transient signals through the use of higher order statistics. The specific detection problem addressed in this paper is that of passively detecting a single occurrence of a deterministic transient, or energy, signal in nonstationary noise[1]-[6]. The detectors do not assume noise or signal stationarity. The signal is assumed to be received on either one or two hydrophones, or channels, with channels repeated in the higher order moments as needed. While the ambient shipping noise is spatially correlated, it is assumed that the signal to be detected is more correlated than the noise.

2. Moments and detection methodology

The p-th order moment detector uses the zero lag of the p-th order correlation as a test statistic and compares it to a predetermined threshold. A decision is then made as to whether the test data contains a signal embedded in noise or only noise. The p-th order moment of a single channel of input data, x(t), is

$$m_p^x = \sum_{k=0}^{N-1} x^p(t)\Delta t$$

with $t = k\Delta t$. For two channels of input data, the second order moment is

$$m_2^x = \sum_{k=0}^{N-1} x_1(t)x_2(t)\Delta t \ ,$$

and the fourth order moment is defined by

$$m_4^x = \sum_{k=0}^{N-1} x_1^3(t)x_2(t)\Delta t.$$

The latter could also be defined by squaring $x_1(t)$ and $x_2(t)$ to form m_4^x, but this generally leads to poorer detection results when the noise is spatially correlated [7]. The p-th order moment is equal to the area, volume, or hypervolume beneath the corresponding moment spectrum, i.e.,

$$m_2^x = \sum_{j=-N/2}^{N/2} ES = \sum_{j=-N/2}^{N/2} X(f)X^*(f)\Delta f = \sum_{j=-N/2}^{N/2} |X(f)|^2 \Delta f,$$

and

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$m_4^x = \sum_{j_1=-N/2}^{N/2} \sum_{j_2=-N/2}^{N/2} \sum_{j_3=-N/2}^{N/2} TS$$

$$= \sum_{j_1=-N/2}^{N/2} \sum_{j_2=-N/2}^{N/2} \sum_{j_3=-N/2}^{N/2} \left[\begin{array}{l} X(f_1)X(f_2)X(f_3) \\ \bullet X^*(f_1+f_2+f_3)\Delta f_1 \Delta f_2 \Delta f_3 \end{array} \right]$$

for a single channel of input data, where ES is the energy spectrum, TS is the trispectrum, and X(f) is the Fourier transform of x(t) with $f = j\Delta f$. In this sense, the trispectrum is a higher-order extension of the energy.

Since transient signals have finite energy and moments that vary over time, averaging to achieve noise suppression must be used cautiously, and is generally possible only if the transient signal is of long duration and/or the sampling rate of the sonar system is high. To avoid these assumptions and the assumption of noise stationarity, the model used here assumes no averaging.

An additional processing step may be added to the detection algorithm. Previous simulations of moment detector performance in Gaussian noise have shown that the performance of the fourth order moment detector can improve relative to the second order moment detector through the use of simple passband filtering [2], [3]. That is, one-dimensional passband filtering of the received channel data prior to applying the detector has a nonlinear and advantageous effect in higher order moments. This process is referred to as prefiltering. In passive detection, it is unlikely that a good estimate of the signal passband will be available unless one is searching for a specific class of signals. However, prefiltering can be used by dividing the data passband into smaller frequency bands and applying the moment detectors successively in the smaller bands. Thus, the use of prefiltering in its simplest form can determine whether or not it has merit for passive sonar applications and whether it should be pursued in more practical applications.

3. Test data

The ambient shipping noise used for evaluation of the moment detectors was recorded during the Shallow Water Evaluation Cells Exercise-3 (SWellEX-3) which occurred during July-August 1994 near the port of San Diego, California [8]. Ambient noise measurements were recorded on a vertical array of hydrophones, located in water approximately 200 m deep. The area is characterized by heavy shipping traffic, including military, commercial, and recreational vessels. The noise is sampled at 1500 samples/second. Over 29 minutes of ambient shipping noise, received at two channels, is used for the detection analysis. The deeper channel is located at depth 192 m and the shallower channel is located at depth 116 m.

Three low frequency synthetic test signals are used in simulations with the measured noise. The first signal is a simulated *whale* cry and is almost Gaussian with a kurtosis of -0.73. This signal is not expected to be easily detectable with the higher order moments. The second is an exponentially damped *sinusoid* with a moderately large kurtosis of 5.89. The third signal is a narrow-time *pulse*. This signal is broadband and significantly nonGaussian (kurtosis = 262.65). The three test signals and their energy spectra are shown in Figs. 1 and 2. Simulations are performed with processing windows equal to the signal durations of one second.

Passbands for prefiltering are chosen to be 10-25 Hz for the whale, 0-40 Hz for the sinusoid, and 0-256 Hz for the pulse. A priori knowledge of the signal passband is generally not available. However, these detectors can be applied successively in sequential frequency bands. Use of such an algorithm will likely give somewhat different results than the broadband application tested here since the moments of the partial signal in each frequency band determine detection, rather than the moments of the entire signal. The effects of this kind of processing are currently being investigated.

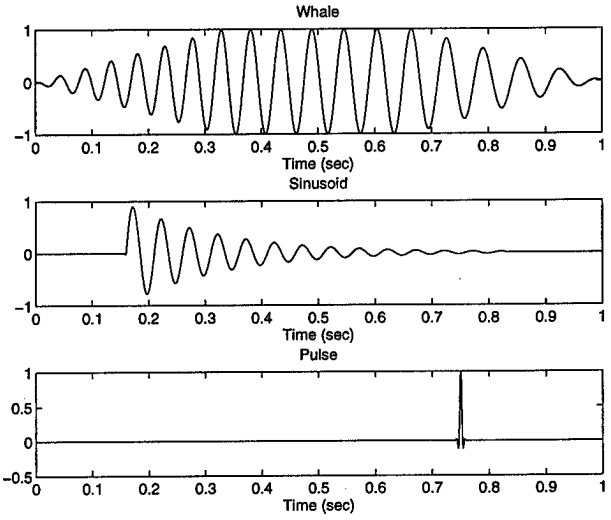


Fig. 1. Three test signals. The whale is almost Gaussian, the sinusoid is moderately nonGaussian, and the pulse is strongly nonGaussian.

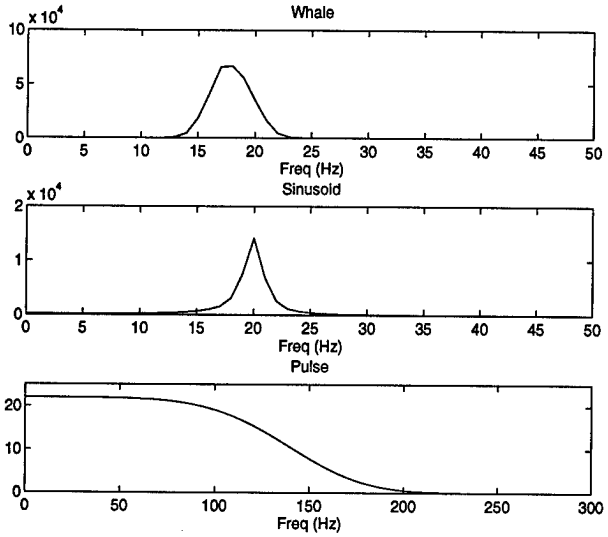


Fig. 2. Energy spectra of test signals.

4. Simulations and results

Since the measured noise is nonstationary, statistical measures of performance are not derived from simulations. Instead, an average measure of performance over short periods of time is given. The 29 minutes of ambient noise are separated into segments lasting 1.82 minutes, and each segment is divided into 109 processing windows. For a range of signal-to-noise ratios (SNRs), the signal is added to the noise in each processing window, one of the detectors is applied, and a resulting curve of probability of detection (P_d) versus SNR is obtained for probability of false alarm (P_{fa}) equal to 0.009. The P_d versus SNR curve is interpolated at P_d=0.5 to give a measure of performance that consists of a single number. The SNR gain is defined as the difference between these two numbers for the second and fourth order moment detectors, with a positive SNR gain indicating that the fourth order moment detector performs better than the second order moment detector.

The results of these simulations using one and two channels of input data and the three test signals are shown in Figs. 3-6. The single channel case uses noise received at the shallower hydrophone. Each point in these curves represents the SNR gain over a 1.82 minute time period. The average SNR gains over the entire 29 minute time periods are given in Tables 1 and 2.

For comparison, detection results using simulated independent Gaussian noise are included in Tables 3 and 4. Since this noise is stationary, 1000 realizations of noise per channel are used to calculate the SNR gain in a 1.82 minute segment.

Comparison of the second and fourth order moment detectors indicates that the fourth order moment can detect transient signals better than the second order moment, or energy, detector for the two nonGaussian signals when either one or two channels of input are used in the moment definitions, and no prefiltering is used (see Figs. 3 and 4). In some cases, about 1 dB of gain is achieved by using two channels of input for detection over using a single channel. Variations in the SNR gains over time are generally as high as 3 dB.

The fourth order moment detector performs better than the second order moment detector for the sinusoid signal for the majority of the 1.82 minute time periods. The average SNR gains over the 29 minutes are 0.73 dB and 0.66 dB for the one and two channel cases. For the highly nonGaussian pulse, the fourth order moment always detects best, with SNR gains of over 4 dB and 29-minute average SNR gains of 5.94 dB and 6.02 dB for the one and two channel cases. The fourth order moment detector does not usually show improvement over the second order moment for the whale signal. The average SNR gains here are -0.24 dB and -0.27 dB for the one and two channel cases.

When prefiltering is used, the SNR gains increase for the sinusoid signal (compare Figs. 3 and 5). The 29 minute average SNR gains for the one and two channel cases approximately double. Average SNR gains change less than 5% when prefiltering is added to the detection processing for the pulse due to the broadband and low frequency nature of the signal. The higher noise frequencies that are filtered (257-750 Hz) contain much less energy that the lower frequencies for the shipping noise. Prefiltering causes the fourth order moment detector to perform poorer for the whale signal. This is not the case with Gaussian noise, where prefiltering increases the SNR gain of the fourth order detector.

Additional differences occur when comparing Gaussian noise to the measured ambient shipping noise. Without prefiltering, the SNR gains for the pulse are 5.94 dB and 6.02 dB for the shipping noise with one and two channels of input, but are only 3.47 dB and 3.53 dB for the Gaussian noise. Higher gains also occur for the shipping noise and the pulse when prefiltering is used, and for the sinusoid with and without prefiltering.

Decreases in the average SNR gains with prefiltering occur only in the measured noise cases. Since the average Gaussian noise is uniformly distributed across frequencies, prefiltering advantages in the fourth order moment over the second order moment are always greater [2], [3]. However, in the nonstationary measured shipping noise, which is not uniformly distributed across

frequencies, the fourth order advantage compared to second order is somewhat unpredictable.

	No Prefiltering	Prefiltering
Whale	-0.24	-0.57
Sinusoid	0.73	1.48
Pulse	5.94	6.11

Table 1. Average SNR gains (in dB) for the tricorrelation detector over the cross correlation detector in the shipping noise with one channel of input.

	No Prefiltering	Prefiltering
Whale	-0.27	-0.98
Sinusoid	0.66	1.16
Pulse	6.02	5.73

Table 2. Average SNR gains (in dB) for the tricorrelation detector over the cross correlation detector in the shipping noise with two channels of input.

	No Prefiltering	Prefiltering
Whale	-0.44	-0.24
Sinusoid	-0.23	0.45
Pulse	3.47	4.40

Table 3. SNR gains (in dB) for the tricorrelation detector over the cross correlation detector in independent Gaussian noise with one channel of input.

	No Prefiltering	Prefiltering
Whale	-1.10	-0.45
Sinusoid	-1.29	0.84
Pulse	3.53	5.14

Table 4. SNR gains (in dB) for the tricorrelation detector over the cross correlation detector in independent Gaussian noise with two channels of input.

5. Conclusions

Simulations with three low frequency test signals embedded in both independent Gaussian noise and measured ambient shipping noise from a shallow-water area indicate that the fourth order moment can provide improved detection capability over the second order moment, or energy, detector if the transient signal is sufficiently nonGaussian. Since the measured noise is nonstationary, measures of relative detector performance are calculated over short periods of time. Results indicate that the SNR gains for the fourth order moment detector over the second order are higher on average in the shipping noise than they are in Gaussian noise for the nonGaussian transient signals. Since the fourth order moment detector does not perform as well as the second order for the transient signal that is nearly Gaussian, it may be best used in addition to the second order moment detector, rather than a replacement.

Acknowledgments

This work was supported by the Office of Naval Research and the Naval Research Laboratory.

References

[1] Pflug, L. A., Ioup, G. E., Ioup, J. W., Barnes, K. H., Field, R. L., and Rayborn, G. H., "Detection of Oscillatory and Impulsive Transients Using Higher Order Correlations and Spectra," J. Acoust. Soc. Am., 91, 2763-2776 (1992).

[2] Ioup, G. E., Pflug, L. A., Ioup, J. W., and Field, R. L., "Prefiltering for Higher Order Advantage," Signal Processing Workshop on Higher-Order Statistics, Lake Tahoe, 309-313 (1993).

[3] Pflug, L. A., Ioup, G. E., Ioup, J. W., and Field, R. L., "Prefiltering for Improved Correlation Detection of Bandlimited Transient Signals," J. Acoust. Soc. Am., 95, 1459-1473 (1994).

[4] Pflug, L. A., Ioup, G. E., and Ioup, J. W., "Performance Prediction Formulas for Higher Order Correlation Detection of Energy Signals," IEEE Signal Processing/ATHOS Workshop on Higher-Order Statistics, Begur, Spain, 152-156 (1995).

[5] Pflug, L. A., Ioup, G. E., Ioup, J. W., and Field, R. L., "Prediction of SNR Gain for Passive Higher Order Correlation Detection of Energy Transients," J. Acoust. Soc. Am., 98, 248-260 (1995).

[6] Walsh, D. O. and Delaney, P. A., "Detection of Transient Signals in Multipath Environments," IEEE J. of Oceanic Engineering, 20, 131-138 (1995).

[7] Pflug, L. A., Ioup, G. E., and Ioup, J. W., "The Effects of Ambient Shipping Noise on the Performance of Single and Multiple Channel Moment Detectors for Unknown Transient Signals," Naval Research Laboratory Report, NRL/FR/7173--97-0003 (1997).

[8] Bachman, Richard T., Schey, Philip W., Booth, Newell O., and Ryan, Frank J., "Geoacoustic Databases for Matched-Field Processing: Preliminary Results in Shallow Water Off San Diego, California," J. Acoust. Soc. Am., 100, 2077-2085 (1996).

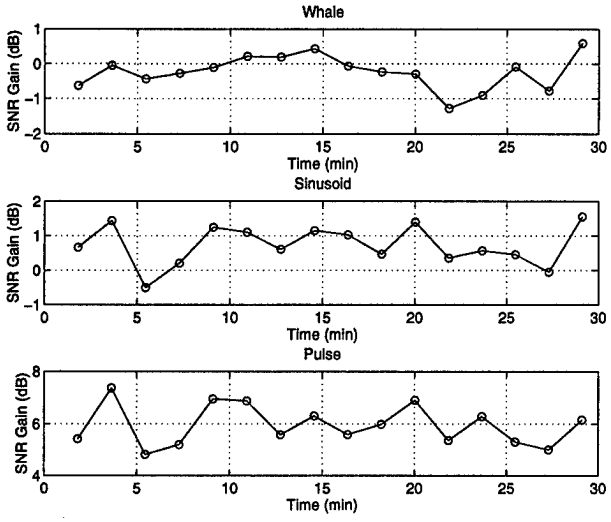


Fig. 3. SNR gains for the tricorrelation detector over the cross correlation detector using one channel of input data and no prefiltering.

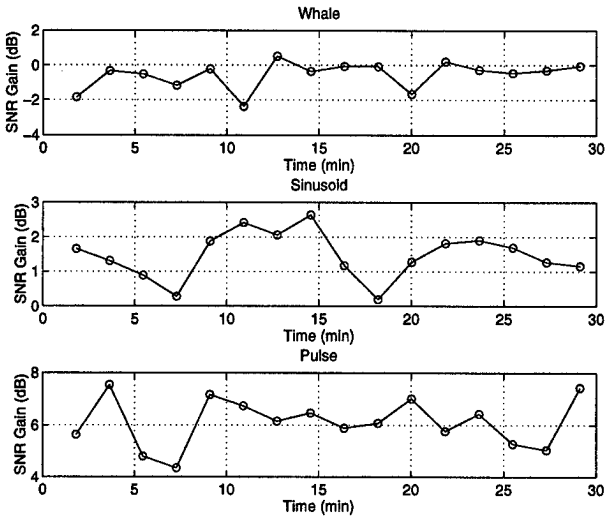


Fig. 5. SNR gains for the tricorrelation detector over the cross correlation detector using one channel of input data and prefiltering.

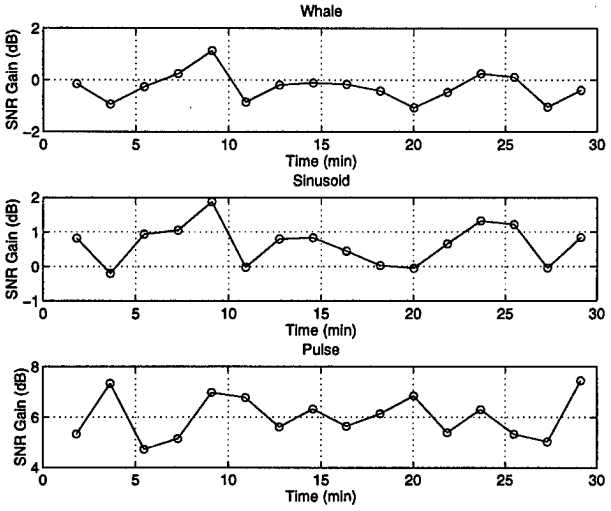


Fig. 4. SNR gains for the tricorrelation detector over the cross correlation detector using two channels of input data and no prefiltering.

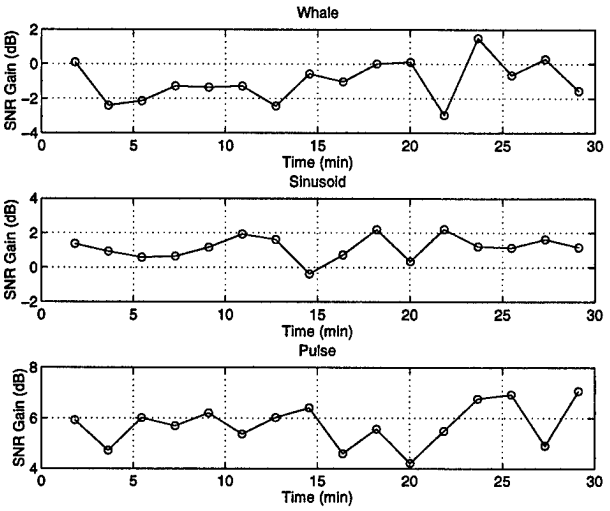


Fig. 6. SNR gains for the tricorrelation detector over the cross correlation detector using two channels of input data and prefiltering.

Detection and Classification of Spectrally Equivalent Processes: A Parametric Approach

Martial COULON, Jean-Yves TOURNERET and Mounir GHOGHO
National Polytechnics Institute of Toulouse, LEN7-GAPSE
2 rue Camichel, BP 7122, 31071 Toulouse Cedex 7, France
Tel: 05 61 58 83 67
e-mail : coulon@len7.enseeiht.fr

Abstract

The detection of two Spectrally Equivalent (SE) processes is addressed. The two SE processes are modeled using two SE parametric models: the noisy AR model and the ARMA model. Higher-order statistics are shown to be an efficient tool for the SE process detection problem. A new detector based on the Higher-Order Yule-Walker matrix singularity is studied. The detector performance is compared in supervised and unsupervised learning. The model order mismatch is then studied.

1 Introduction

The detection of Spectrally Equivalent (SE) processes is an interesting problem in many signal processing applications. SE processes have been observed in spread spectrum communications, including cellular mobil systems, military applications, geodesic survey, position location for vehicles [5]. The SE detection problem has also received much attention in SAR image processing. For instance, the detection of non-linearities contained in SAR images using Volterra models can reduce to compare signals having the same spectrum [4]. The detection of signals, which have very close spectra over a large bandwidth, is difficult to achieve using techniques based on second-order statistics.

This paper proposes to model SE processes by two SE parametric models: the $ARMA$ model and the noisy AR model. This modeling was studied successfully in [8], where it was proved that SE $ARMA$ and noisy AR processes cannot have the same Higher-Order Statistics (HOS). This property induced a binary hypothesis test based on higher-order cumulants. Unfortunately, this test cannot be easily generalized to composite hypotheses.

This paper studies new SE signal detectors based on the HOS Yule-Walker equations (HOYWE). Section 3 shows that the Higher-order Yule-Walker matrix singularity is an efficient tool for the SE process detection. A Likelihood Ratio Detector (LRD) based on the determinant of this matrix is developed, when the signal parameters are known. However, the signal parameters are unknown in most practical applications, and have to be estimated. Section 4 generalizes the proposed detector to composite hypotheses. This generalization is very simple, since the test statistic does not depend on the model parameters. Section 5 studies the model order mismatch problem. This section shows that an overestimated model order does not deteriorate the detector performance. Simulation results and conclusions are reported in sections 6 and 7.

2 Problem Formulation

The study is restricted to stationary and ergodic processes, with symmetric continuous spectra, which can be modeled by linear processes. The modeling of two SE processes by parametric $ARMA$ and noisy AR processes can be justified by the following property:

If S is a symmetric continuous spectrum, there exists an $AR(p)$ process, an $MA(q)$ process and an $ARMA(p,q)$ process whose spectra approximate S as close as desired [1].

Consider the two following signal models :

• An $AR(p)$ process $x(n)$ driven by an input $e(n)$ and corrupted by an additive random **Gaussian** noise $b(n)$:

$$y_0(n) = x(n) + b(n) \tag{1}$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

with:

$$x(n) = -\sum_{j=1}^{p} a_j x(n-j) + e(n) \qquad (2)$$

$e(n)$ and $b(n)$ are assumed to be i.i.d. and mutually independent.

• An $ARMA(p,p)$ process $y_1(n)$, driven by a **non-Gaussian** i.i.d. input $g(n)$, with the same mean and power spectral density as $y_0(n)$, defined by

$$y_1(n) = -\sum_{j=1}^{p} \widetilde{a}_j y_2(n-j) + \sum_{j=0}^{p} b_j g(n-j) \qquad (3)$$

The spectral density of $y_0(n)$ is a special $ARMA(p,p)$ spectrum [3]. The spectral equivalence for stable processes $y_0(n)$ and $y_1(n)$ implies that the $ARMA$ and noisy AR model parameters are linked by a one-to-one transformation.

3 Cumulant-Based Detector with Known Model Parameters ("KP Detector")

The SE process detection problem can be expressed as a binary hypothesis testing problem:

$$\begin{array}{ll} H_0 \left(\begin{array}{c}\text{Noisy AR}\\ \text{process}\end{array}\right): & y(n) = y_0(n)\ ,\ n = 0,\ldots,N-1 \\ H_1 \left(\begin{array}{c}\text{ARMA}\\ \text{process}\end{array}\right): & y(n) = y_1(n)\ ,\ n = 0,\ldots,N-1 \end{array} \qquad (4)$$

The Likelihood Ratio Test (LRT) for hypotheses H_0 and H_1 cannot be easily derived, for non-Gaussian AR and $ARMA$ processes. Consequently, suboptimal detectors have to be considered. This paper proposes to use the Higher-Order Yule-Walker Equations for the SE process detection. The HOYWE for an $AR(p)$ process are defined by [6]:

$$\sum_{j=0}^{p} a_j C_k^x(m-j, 0, \ldots, 0) = 0\ ,\ \forall m > 0 \qquad (5)$$

where $C_k^x(\tau_1, \tau_2, \ldots, \tau_{k-1})$ denotes the AR process kth-order cumulant at lag $\underline{\tau} = (\tau_1, \tau_2, \ldots, \tau_{k-1})$. The property $C_k^{y_0}(\underline{\tau}) = C_k^x(\underline{\tau})$ holds $\forall k > 2, \forall \underline{\tau} \in \mathbb{Z}^{k-1}$, since the additive noise is Gaussian and independent of $x(n)$. Denote $\mathbf{\Delta}_p(\boldsymbol{\xi}) = \det(\mathbf{R}_p(\boldsymbol{\xi}))$ the determinant of the following matrix

$$\mathbf{R}_p(\boldsymbol{\xi}) = \begin{bmatrix} \xi_{p+1} & \xi_p & \cdots & \xi_1 \\ \xi_{p+2} & \ddots & \ddots & \cdots \\ \cdots & \ddots & \ddots & \xi_p \\ \xi_{2p+1} & \cdots & \cdots & \xi_{p+1} \end{bmatrix} \qquad (6)$$

with $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{2p+1}) \in \mathbb{R}^{2p+1}$. Denote

$$\mathbf{C}_k^0 = (C_k^{y_0}(1-p, 0, \ldots, 0), \ldots, C_k^{y_0}(1+p, 0, \ldots, 0))^T \qquad (7)$$

The concatenation of eq.'s (5) for $m \in \{1, \ldots, p+1\}$ leads to

$$\mathbf{R}_p(\mathbf{C}_k^0).(1, a_1, \ldots, a_p)^T = 0 \qquad (8)$$

hence

$$\Delta_k^0 \triangleq \mathbf{\Delta}_p(\mathbf{C}_k^0) = 0 \qquad (9)$$

On the other hand, for an $ARMA(p,p)$ process, eq.'s (5) hold for $m > p$, but not for $m \in \{1, \ldots, p\}$. It is then legitimate to assume that

$$\Delta_k^1 \triangleq \mathbf{\Delta}_p(\mathbf{C}_k^1) \neq 0\ , \qquad (10)$$

where $\mathbf{C}_k^1$ is defined as in (7) with $ARMA$ process cumulants. The SE noisy AR and $ARMA$ process detection can then be expressed as a binary simple hypothesis testing problem:

$$\begin{array}{lll} H_0: & \text{(Noisy AR process)} & \Delta_k = \Delta_k^0 = 0 \\ H_1: & \text{(ARMA process)} & \Delta_k = \Delta_k^1 \neq 0 \end{array} \qquad (11)$$

Define $\widehat{\mathbf{C}}_k$ as the sample cumulant vector obtained by replacing the true cumulants in (7) by their usual estimates, and denote $\widehat{\Delta}_k \triangleq \mathbf{\Delta}_p(\widehat{\mathbf{C}}_k)$. The noisy AR and $ARMA$ process cumulant vector estimate is asymptotically an unbiased Gaussian vector such that: $\lim_{N\to+\infty} NE[(\widehat{\mathbf{C}}_k - \mathbf{C}_k^i)(\widehat{\mathbf{C}}_k - \mathbf{C}_k^i)^T \mid H_i] = \mathbf{\Sigma}_k^i$ [6]. According to ([1], p. 211), this property implies that the determinant estimate $\widehat{\Delta}_k$ is asymptotically an unbiased Gaussian variable with:

$$\lim_{N\to+\infty} NE[(\widehat{\mathbf{\Delta}}_k - \mathbf{\Delta}_k^i)^2 / H_i] = D_k^{iT} \mathbf{\Sigma}_k^i D_k^i \triangleq \sigma_i^2 \qquad (12)$$

In (12), D_k^i is a vector whose j^{th} element is $D_k^i(j) = (\partial \mathbf{\Delta}(\boldsymbol{\xi}) / \partial \xi_j)(\mathbf{C}_k^i)$. It is then straightforward to prove that:

$$D_k^i(j) = \sum_{m=1}^{|\xi_j|} (Cof(\mathbf{R}_p(\mathbf{C}_k^i))_m \qquad (13)$$

where $|\xi_j|$ denotes the number of ξ_j in $\mathbf{R}_p(\boldsymbol{\xi})$ and $(Cof(\mathbf{R}_p(.)))_m$ the $\mathbf{R}_p(.)$ matrix cofactor computed at the m^{th} occurence of ξ_j. The statistical properties of $\widehat{\Delta}_k$ can then be asymptotically derived under both hypotheses:

$$\begin{array}{lll} H_0: & \text{Noisy AR process} & \widehat{\Delta}_k \sim N(0, \sigma_0^2(N)) \\ H_1: & \text{ARMA process} & \widehat{\Delta}_k \sim N(\Delta_k^1, \sigma_1^2(N)) \end{array} \qquad (14)$$

with $\sigma_i^2(N) = \frac{1}{N}\sigma_i^2$. The LRT for these two hypotheses can then be expressed as:

$$H_0 \text{ rejected if } T_1 \triangleq \frac{\widehat{\Delta}_k^2}{\sigma_0^2(N)} - \frac{(\widehat{\Delta}_k - \Delta_k^1)^2}{\sigma_1^2(N)} > k_1 \quad (15)$$

In (15), k_1 is a threshold depending on the False Alarm Probability (FAP). Under both hypothesis H_i ($i \in \{0,1\}$), T_1 can be normalized by $T_{1,i}=\lambda_i \widetilde{T}_i + \mu_i$ where $\widetilde{T}_i$ is distributed according to a non-central χ^2 distribution with 1 degree of freedom and non-central parameter $\nu_i = \sigma_i^2(N)\Delta_1^2/(\sigma_0^2(N) - \sigma_1^2(N))^2$. For a fixed FAP, the Probability of Detection (PD) is then given by:

- First case: $\sigma_1^2 > \sigma_0^2$,

$$PD = 1-\chi_1^2[\nu_1]\left(\left\{\lambda_0\chi_1^2[\nu_0]^{-1}(1-FAP) + \mu_0 - \mu_1\right\}/\lambda_1\right) \quad (16)$$

- Second case: $\sigma_0^2 > \sigma_1^2$,

$$PD = \chi_1^2[\nu_1]\left(\left\{\lambda_0\chi_1^2[\nu_0]^{-1}(FAP) + \mu_0 - \mu_1\right\}/\lambda_1\right) \quad (17)$$

where $\chi_d^2[\nu](.)$ and $\chi_d^2[\nu]^{-1}(.)$ denote the cumulative distribution function (cdf) and its inverse for a non-central χ^2 distribution with d degrees of freedom and non-central parameter ν.

4 Cumulant-based Detector with Unknown Parameters ("UP Detector")

This section addresses the SE process detection problem in unsupervised learning (ARMA and noisy AR parameters unknown). Assume that M independent realizations of $\widehat{\Delta}_k$, denoted $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$, are available. These M measurements can be obtained from one single signal by segmentation. This procedure consists in considering an N-sample signal as M segments of K samples. The estimation $\widehat{\Delta}_{k,j}$ is computed from the j^{th} segment samples. Two remarks have to be pointed up:

- the segment size K has to be large enough to obtain approximately normally distributed determinant $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$.
- the slices cannot be adjacent: two consecutive slices have to be sufficiently separated to be independent. The separation between two segments can be computed using the procedure proposed in [2].

Note that the previous conditions regarding the segment length and the slice separation cannot be satisfied for short signals. Consider the sequence $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$ of M independent one-dimensional normal variables. Define S^2 as the sample variance of the sequence $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$:

$$S^2 \triangleq \frac{1}{M}\sum_{j=1}^{M}\left(\widehat{\Delta}_{k,j} - \overline{\Delta}_k\right)^2, \text{ with } \overline{\Delta}_k = \frac{1}{M}\sum_{j=1}^{M}\widehat{\Delta}_{k,j} \quad (18)$$

It is well known that $T_2 \triangleq \frac{\overline{\Delta}_k - E\left[\widehat{\Delta}_{k,j}\right]}{S}\sqrt{M-1}$ has a Student distribution with $M-1$ degrees of freedom [7]. Consequently, the distribution of T_2 under H_0 does not depend on the model parameters (since $E\left[\widehat{\Delta}_{k,j}\right] = 0$ under H_0). The detection strategy is then defined by:

$$H_0 \text{ rejected if } T_2^2 > \tau_2 \quad (19)$$

In (19), τ_2 is a threshold independent of the model parameters. The distribution of T_2 under hypothesis H_1 is unknown (since $E\left[\widehat{\Delta}_{k,j}\right] \neq 0$). Consequently, the theoretical PD corresponding to (19) cannot be obtained. It has to be numerically computed using Monte-Carlo simulations.

Note that the Generalized Likelihood Ratio detector based on the observations $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$ leads to the same test statistic T_2 (see appendix).

5 Spectral Equivalence Model Order Mismatch

Assume that the SE model order has been overestimated. Eq. (5) can then be written:

$$\sum_{j=0}^{p'} a'_j C_k^x(m-j,0,\ldots,0) = 0\ ,\ \forall m > 0 \quad (20)$$

with

$$\begin{cases} a'_j = a_j & \forall j \in \{1,\ldots,p\} \\ a'_j = 0 & \forall j \in \{p+1,\ldots,p'\} \end{cases} \quad (21)$$

p' is the overestimated order ($p' > p$). Denote $\mathbf{C}_k^{'0}$ the vector obtained by replacing p by p' in (7). Eq. (20) then yields:

$$\mathbf{\Delta}_{p'}(\mathbf{C}_k^{'0}) = \det\left(\mathbf{R}_{p'}(\mathbf{C}_k^{'0})\right) = 0 \quad (22)$$

Moreover, it is still legitimate to assume that $\mathbf{\Delta}_{p'}(\mathbf{C}_k^{'1}) \neq 0$. Consequently, the hypothesis testing problem is exactly the same as (11). The only difference is that larger cumulant slices are involved. As a conclusion, the model order p can be estimated using a conventional technique such as Akaike's [3]. A corrected estimate greater than p' will be prefered in case of uncertainty.

6 Simulations Results

Many simulations have been performed to validate the previous theoretical results. In this paper, the simulations are presented for an $AR(1)$ process with parameters $[1; -0.5]$ driven by a zero mean i.i.d. exponentially distributed input (with variance $\sigma_e^2 = 1$). However, other model orders and AR parameters have been studied and give similar performance. The detector performance is computed as a function of the Signal-to-Noise Ratio defined by $SNR = \sigma_e^2/\sigma_b^2$. The ARMA model parameters (input variance σ_g^2 and parameters $\widetilde{a}_j, b_j$) are related to the noisy AR process parameters by a one-to-one transformation. The ARMA process is driven by a zero mean i.i.d. exponentially distributed input. Numerical results have been computed with 2000 Monte-Carlo simulations.

Fig. 1 shows the KP detector ROC's for different $SNR's$. The test improves when the SNR decreases, in this particular case. The noisy AR process distribution is close to the Gaussian distribution, when the SNR is low, contrary to the SE $ARMA$ process. Thus, the two SE processes can be easily distinguished for low SNR's. Fig. 2 presents the KP detector performance for fixed $SNR = 8dB$ as a function of the number of samples N. Obviously, the higher N, the better the detector performance. A comparison between fig.'s 2 and 3 shows that the detector performance is relatively similar for known and unknown parameters. Fig. 4 presents the UK detector ROC's obtained for different overestimated orders p' ($SNR = 8dB$ and $N = 500$). Clearly, the higher p', the better the detector performance. This can be explained as follows. First, the model order mismatch does not deteriorate the detector performance, as proved in section 5. Second, the number of cumulants involved in the test function increases when p' increases. Consequently, the higher p', the better the information concerning the process structure.

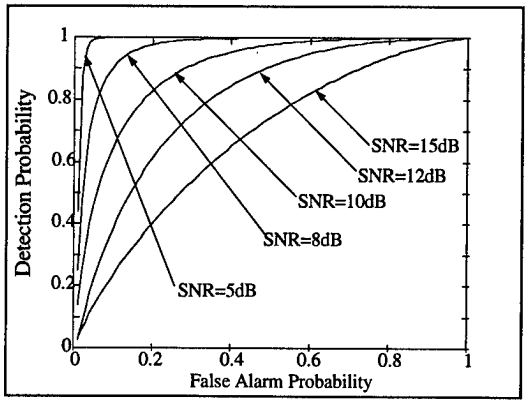


Fig. 1. KP Detector ROC's ($N = 10000$).

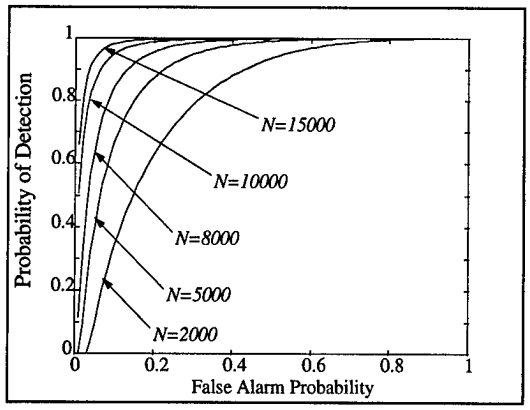


Fig. 2. KP Detector ROC's for different number of samples ($SNR = 8dB$).

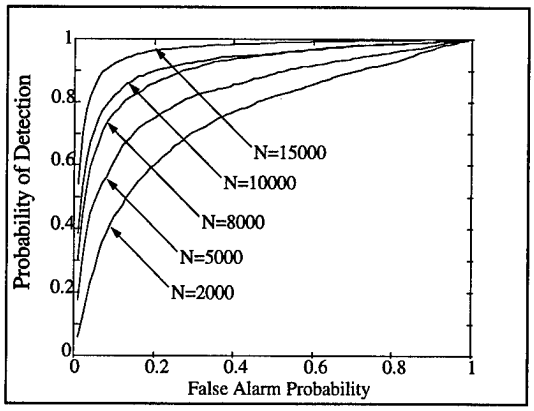


Fig. 3. UP Detector ROC's for different number of samples($SNR = 8dB$).

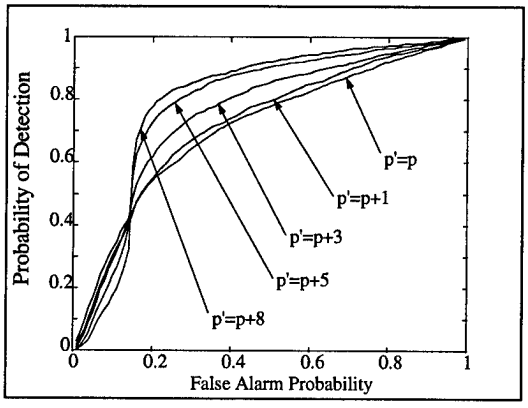


Fig. 4. UK Detector ROC's for different values of the overestimated order p' ($SNR = 8dB$, $N = 500$).

7 Summary and Conclusions

Higher-order statistics were shown to be an efficient tool for the SE process detection. A likelihood ratio test based on higher-order Yule-Walker equations was derived, for the supervised learning case. The proposed test was suboptimal since it did not work on

data themselves. However, it did not require any statistical assumption for the driving noise and showed good performance. The basic idea of this detector was to test the nullity of the higher-order Yule-Walker matrix determinant. The detector was then generalized to the unsupervised learning case, where no *a priori* information was available. The corresponding detector showed satisfying performance. Moreover, the detector was shown to be robust concerning the model order overestimation.

This work can be extended to the detection of processes whose spectra are slighty different. These processes cannot be easily distinguished using techniques based on second-order statistics, similarly to SE processes. The detection may be achieved using higher-order statistics. Unfortunatley, this work was limited to linear processes. The detection of SE processes which cannot be modeled by linear processes is currently under investigation.

8 Appendix: Generalized Likelihood Ratio Detector

Since Δ_k^1, $\sigma_0^2(N)$, and $\sigma_1^2(N)$ are the unknown parameters, the GLR statistics based on the observations $\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M}$ is defined by:

$$GLR = \frac{\sup\limits_{\Delta\neq 0\ ;\ \sigma^2>0} (\sigma^2)^{-M/2} \exp\left(\frac{-1}{2\sigma^2}\left(\sum\limits_{i=1}^{M}(\widehat{\Delta}_{k,j}-\Delta)^2\right)\right)}{\sup\limits_{\sigma^2>0} (\sigma^2)^{-M/2} \exp\left(\frac{-1}{2\sigma^2}\left(\sum\limits_{i=1}^{M}\widehat{\Delta}_{k,j}{}^2\right)/2\sigma^2\right)} \quad (23)$$

It is well known that the likelihood function suprema are obtained for:

$$\begin{cases} \Delta = \overline{\Delta}_k = \frac{1}{M}\sum\limits_{j=1}^{M}\widehat{\Delta}_{k,j} \\ \sigma^2 = \frac{1}{M}\sum\limits_{j=1}^{M}\left(\widehat{\Delta}_{k,j}-\overline{\Delta}_k\right)^2 \end{cases} \quad \text{under } H_1 \quad (24)$$

$$\sigma^2 = \frac{1}{M}\sum_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2 \quad \text{under } H_0$$

The GLR test for deciding between H_0 and H_1 then reduces to:

$$H_0 \text{ rejected if } GLR = \left(\frac{\sum\limits_{j=1}^{M}\left(\widehat{\Delta}_{k,j}-\overline{\Delta}_k\right)^2}{\sum\limits_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2}\right)^{-M/2} > \tau'_3\ , \quad (25)$$

$$\text{i.e. } \sum_{j=1}^{M}\left(\widehat{\Delta}_{k,j}-\overline{\Delta}_k\right)^2 < (\tau'_3)^{-2/M}\cdot\sum_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2 \quad (26)$$

In (25) and (26), τ'_3 is a threshold depending on the false-alarm probability such as $\tau'_3 > 1$ in order to avoid to reject H_0 in any case (Indeed, since $\sum\limits_{j=1}^{M}\left(\widehat{\Delta}_{k,j}-\overline{\Delta}_k\right)^2 < \sum\limits_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2$, $GLR > 1$ $\forall\left(\widehat{\Delta}_{k,j}\right)_{j=1,\ldots,M} \in \mathbb{R}^M$). Moreover, since $\sum\limits_{j=1}^{M}\widehat{\Delta}_{k,j} = M\overline{\Delta}_k$, eq. (26) leads to:

$$H_0 \text{ rejected if } (1-(\tau'_3)^{-2/M})\sum_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2 < M\overline{\Delta}_k{}^2\ , \text{ i.e.} \quad (27)$$

$$\left(\frac{\overline{\Delta}_k}{\sqrt{\frac{1}{M}\sum\limits_{j=1}^{M}\widehat{\Delta}_{k,j}{}^2}}\sqrt{M-1}\right)^2 > (M-1)\left(1-(\tau'_3)^{-2/M}\right) \quad (28)$$

Consequently, denoting $\tau_3 = (M-1)\left(1-(\tau'_3)^{-2/M}\right)$, eq. (28) is identical to eq. (19).

References

[1] P. J. Brockwell and R. A. Davis, *Time Series : Theory and Methods*, Springer Verlag, 2nd Edition, 1990.

[2] B. Friedlander and B. Porat,"Asymptotically Optimal Estimation of MA and ARMA Parameters of Non-Gaussian Processes from Higher-Order Moments," *IEEE Transactions on Automatic Control*, Vol. 35, No 1, January 1990, pp. 27-35.

[3] S.M. Kay, *Modern Spectral Estimation : Theory and Application*, Prentice Hall, 1988.

[4] J.-M. Le Caillec, "Etude et Traitement des Images SAR grâce aux Moments d'Ordres Supérieurs," *Ph-D. Dissertation*, Université de Rennes I, 1997.

[5] R.L. Peterson, R.E. Ziemer, D.E. Borth, *Introduction to Spread Spectrum Communications*, Prentice Hall, 1995.

[6] B. Porat, *Digital Processing of Random Signals : Theory and Methods*, Prentice-Hall, 1994.

[7] G. Saporta, *Probabilités, Analyse des Données et Statisitique*, Editions Technip, Paris, 1990.

[8] J.Y. Tourneret, K. Vareille, M. Coulon. "Detection and Classification of Noisy AR and ARMA Processes", *Proc. of EUSIPCO 96,* Trieste, Italy, pp. 1421-1424, September 1996.

Robust Time-Frequency Representations for Signals in α-Stable Noise Using Fractional Lower-Order Statistics

David W. Griffith, Jr., Juan G. Gonzalez, and Gonzalo R. Arce *
Department of Electrical and Computer Engineering
University of Delaware, Newark, DE 19716
griffith@ee.udel.edu

Abstract

Characterizing signals jointly in the time and frequency domains through time-frequency representations (TFRs) such as the Wigner-Ville Distribution (WVD) is a natural extension of Fourier analysis and gives a more complete representation of signal behavior, particularly in the case of non-stationary signals. In the presence of additive impulsive noise, TFRs quickly break down and any information about the desired signal is lost. To combat these effects, we propose in this paper a family of memoryless nonlinearities which have been shown to produce a signal autocorrelation statistic which is well-behaved in the presence of stable noise. The result of this approach is a TFR which is both robust and simple to implement, and has many of the mathematical properties associated with the standard WVD. We illustrate the improvement in performance that can be obtained with several examples.

1. Introduction

Time-frequency representations (TFRs) are useful tools for characterizing the behavior of signals whose spectral characteristics are time-varying. Because of the ease with which they can be implemented, they have been applied to a wide variety of problems in signal processing. Most notably, they have been used for signal recovery at low signal-to-noise ratios (SNRs), accurate estimation of instantaneous frequency and group delay, signal detection in communications, radar, and sonar applications, and the design of time-varying signals and filters.

The Wigner-Ville Distribution (WVD) is one of the most commonly-used quadratic TFRs (QTFRs) and has the form

$$W_x(t,f) = \int_\tau x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi f\tau}\,d\tau. \quad (1)$$

*This work was supported in part by the National Science Foundation under grant MIP-9530923.

where $\int_\tau(\cdot)\,d\tau$ denotes integration with respect to u from $-\infty$ to ∞. The WVD is a particularly useful representation of the behavior of $x(t)$ because it is real everywhere and so acts as a density function describing the concentration of the energy of $x(t)$ over the time-frequency plane. This interpretation is justified in part by the fact that the signal energy, E_x, may be obtained by taking

$$E_x = \int_f \int_t W_x(t,f)\,dtdf, \quad (2)$$

and marginal distributions of the signals energy concentration in time or frequency alone are respectively taking by integrating the WVD with respect to the opposing variable. It has many other desirable mathematical properties, which are discussed in detail in [5], [6], and [10]. The discrete-time WVD is

$$W_x(n,F) = 2\sum_m x(n+m)x^*(n-m)e^{-j4\pi mF}, \quad (3)$$

and has a periodicity of 1/2 along the normalized frequency axis. As a result of this effect, which has been studied in [3] and [4] in detail, it is necessary to either sample the signal at twice its Nyquist rate or use analytic signals sampled at the Nyquist rate in order to avoid aliasing.

The WVD can be used to analyze signals that are contaminated by noise that is Gaussian. However, there are a number of situations, *e.g.* sonar or synthetic aperture radar, where the additive channel noise is characterized by large numbers of impulses. In this case, the WVD can be severely degraded, as large numbers of artifacts can be produced by additive impulses.

In this paper, we propose a modification of the WVD which yields a more robust QTFR. This approach is based on the interpretation of the WVD as a time-varying power spectral density. By replacing the autocorrelation function, which has been shown to be severely affected by impulsive noise, with a more robust statistic, we derive a QTFR that also exhibits significant improvement over comparable QTFRs.

0-8186-8005-9/97 $10.00 © 1997 IEEE

The organization of this paper is as follows. Models for impulsive noise are introduced in Section 2, and are accompanied by a discussion of the effect of impulses on WVDs. A robust WVD using fractional lower order statistics is then introduced in Section 3, and its properties are examined. Examples which illustrate the improvement in performance which can be obtained using fractional lower-order moments are presented in Section 4, and a summary of the work is given in Section 5.

2. Impulsive Noise and Its Effect on TFRs

Impulsive noise arises in a variety of situations, such as atmospheric noise, as discussed by Middleton [8], and in radar and sonar applications. A key feature of all impulsive noise models is that the tails of their associated density functions decay at a rate that is slower than the rate of decay of the tails of a Gaussian density function.

2.1. α-Stable Noise Distributions

The family of α-stable distributions is defined by its members' characteristic functions, which for α-stable distributions that are symmetric about their location, μ, have the form

$$\varphi(\omega) = e^{j\mu\omega - \gamma|\omega|^\alpha}, \tag{4}$$

where γ is the dispersion and α is the stability parameter. The most well-known members of the family of symmetric α-stable (SαS) densities are the Gaussian density, which we obtain when $\alpha = 2$ and the Cauchy density, which arises when $\alpha = 1$. These are the only SαS density functions which can be obtained from (4) in closed form. The stability parameter has the most impact on the shape of these density functions because it controls the amount of mass in the tails. As α decreases from 2 to 0, we observe an increase in the amount of the total mass that lies in the tails, corresponding to an increase in impulsiveness. α-stable distributions are described in more detail in [9] and [11].

2.2. Effects of Impulses on TFRs

Impulsive noise profoundly degrades TFRs, although the type of degradation is dependent on the type of TFR being used. In the analysis that follows, we will employ a highly simplified model of a signal corrupted by impulsive noise which still accurately predicts the effect of strong impulses on the TFR of the signal being observed. We will do our modeling in continuous time, but analogous effects can be observed in discrete-time signals in impulsive noise. The signal $r(t)$ is modeled as a finite-duration signal which is corrupted by a single Dirac delta function, occurring at some random time t_δ uniformly distributed over the signal's time support, so that

$$r(t) = x(t) + \delta(t - t_\delta). \tag{5}$$

The WVD of $r(t)$ is

$$W_r(t,f) = W_x(t,f) + 2\mathrm{Re}[x(2t - t_\delta)] + \delta(t - t_\delta) \tag{6}$$

where Re$[z]$ is the real part of z. Here we have multiple undesired artifacts due to the presence of the impulse. We first have the WVD of the impulse itself, which is an impulsive "ridge" at $t = t_\delta$ that occupies all frequencies. In addition, we have *cross terms*, which arise because of the quadratic structure of the WVD. The cross term is a scaling of the real part of the signal $x(t)$ and the which is shifted in time by t_δ. The cross-terms extend over all frequencies, and therefore often overlap the WVD of $x(t)$. It is because of this overlap that one often cannot apply *smoothing functions*, which are used to attenuate cross-terms in WVD and are described in [5], [6], and [10]. Smoothing functions are effective only on cross terms that are separated from the WVD of the desired signal in time and frequency, which is often not the case for cross terms arising from impulsive noise effects.

3. The Fractional Lower-Order WVD (FLOWVD)

We begin by considering the WVD as a tool for characterizing the spectral properties of non-stationary random processes. Since the WVD is formed by taking the Fourier transform of the autocorrelation of a signal $x(t)$ with respect to the delay variable τ, we can write the *evolutive spectrum* of a random process [1], as

$$S_x(t,f) = \int_\tau E\{x(t+\tau/2)x^*(t-\tau/2)\}e^{-j2\pi f\tau}\, d\tau. \tag{7}$$

By interchanging the integration and expectation operators, we see that the evolutive spectrum can be interpreted as the expectation of all the WVDs corresponding to their respective members of the ensemble of random processes $\{x(t)\}$,

$$S_x(t,f) = E\{W_x(t,f)\}. \tag{8}$$

In the presence of α-stable noise, the higher-order moments of signals become unbounded. This makes second-order measures such as the autocorrelation function useless for examining α-stable signals. It has been shown in [7] that the variance of the autocorrelation becomes infinite for $\alpha < 2$ and the mean of the autocorrelation is infinite for $\alpha < 1$. The solution to this problem, introduced in [7], was to define a *fractional lower-order covariance* (FLOC) measure, which has the form

$$\mathrm{FLOC}_x^a(t,\tau) = E\{x^{\langle a\rangle}(t+\tau/2)x^{-\langle a\rangle}(t-\tau/2)\}, \tag{9}$$

where $(\cdot)^{\langle a \rangle}$ is the a^{th}-order phased fractional lower-order moment (PFLOM) operator, where $0 \leq a \leq 1$. It is defined as

$$z^{\langle a \rangle} = |z|^{a+1}/z^*. \tag{10}$$

If we write z in polar coordinates, $z = re^{j\theta}$, then it is easy to show that $z^{\langle a \rangle} = r^a e^{j\theta}$, so that the PFLOM acts only on the magnitude of its operand and preserves its phase. Also, in (9), we have defined $z^{-\langle a \rangle} \triangleq (z^*)^{\langle a \rangle} = (z^{\langle a \rangle})^*$. If $a = 1$, the FLOC becomes the autocorrelation function. If $a = 0$, all amplitude information is removed; this is known as the phased FLOC. Ma and Nikias showed in [7] that the FLOC has finite mean and variance if $0 < a < \alpha/2$.

We can define a fractional lower-order evolutive spectrum by taking the Fourier transform of the FLOC; this gives us

$$S_x^{\langle a \rangle}(t,f) = \int_\tau E\{x^{\langle a \rangle}(t+\tau/2)x^{-\langle a \rangle}(t-\tau/2)\}e^{-j2\pi f\tau}d\tau \tag{11}$$

which can be written as the expectation of an ensemble of TFRs of the form

$$W_x^{\langle a \rangle}(t,f) = \int_\tau x^{\langle a \rangle}(t+\tau/2)x^{-\langle a \rangle}(t-\tau/2)e^{-j2\pi f\tau}\,d\tau. \tag{12}$$

which we define to be the *fractional lower-order Wigner-Ville distribution* (FLOWVD) of $x^{\langle a \rangle}(t)$. The discrete-time version of the FLOWVD is

$$W_x^{\langle a \rangle}(n,F) = 2\sum_m x^{\langle a \rangle}(n+m)x^{-\langle a \rangle}(n-m)e^{-j4\pi mF}, \tag{13}$$

and we can directly define a general fractional lower-order TFR in a fashion analogous to that for Cohen's class of energetic QTFRs [5] as

$$\rho_x^{\langle a \rangle}(t,f) = \int_\tau [K_x^{\langle a \rangle}(t,\tau) *_t \Phi(t,\tau)]e^{-j2\pi f\tau}\,d\tau, \tag{14}$$

where $K_x^{\langle a \rangle}(t,\tau)$ is the autocorrelation of $x^{\langle a \rangle}(t)$, $\Phi(t,\tau)$ is the smoothing kernel and $*_t$ denotes convolution in the time domain.

The FLOWVD is itself a standard WVD, where the input signal is $x^{\langle a \rangle}(n)$ rather than $x(n)$. As a result of this, it is easy to show that the FLOWVD has all the properties that the WVD itself has. The FLOWVD can, for example, be shown to be real and time-frequency shift-invariant. It also obeys the marginal properties, although the marginal energy densities are associated with $x^{\langle a \rangle}(n)$, not $x(n)$. For instance, if we examine the continuous time FLOWVD, we can show that the marginal energy density with respect to time is

$$\begin{aligned} \int_f W_x^{\langle a \rangle}(t,f)\,df &= x^{\langle a \rangle}(t)x^{-\langle a \rangle}(t) \\ &= |x(t)|^{2a+2}/|x(t)|^2 \\ &= |x(t)|^{2a}. \end{aligned} \tag{15}$$

In addition, the *fractional lower-order ambiguity function* (FLOAF), defined in [7], can be obtained from the FLOWVD by applying a 2-D Fourier transform to (12) in the same way that the symmetric ambiguity function is obtained from the WVD.

4. Results

In this section we demonstrate the performance gains that can be attained by employing fractional lower-order moments to signals prior to computing the WVD. The signal we will use is a sampled echolocation signal from the Large Brown Bat, *Eptesicus Fuscus*, where the sampling period was 6.25 μsec. (The authors thank C. Condon, K. White, and Prof. A. Feng of the Beckman Institute of the University of Illinois for the bat data and for permission to use it in this paper.) A time history of the signal is given in Fig. 1(a) and the WVD of the signal is shown in Fig. 1(b), where we have used the Choi-Williams smoothing kernel, which has the form

$$\Phi(n,m) = \frac{\sqrt{\sigma/\pi}}{2|m|}e^{-\sigma n^2/4m^2}, \tag{16}$$

with $\sigma = 15$. The signal is composed of two prominent nonlinear FM chirps, each approximately 1 msec in length. To avoid aliasing effects, we have created an analytic signal by adding the echolocation signal's Hilbert transform to it.

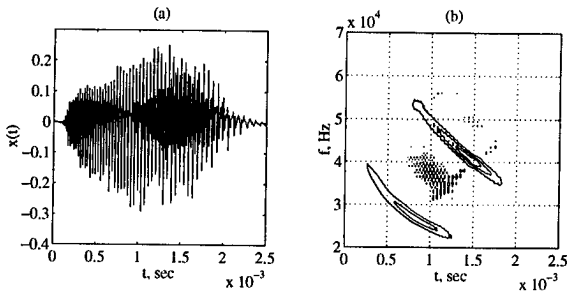


Figure 1. (a): Time history of bat echolocation pulse. (b): Wigner-Ville Distribution of echolocation pulse with Choi-Williams smoothing kernel.

To the echolocation signal in Fig. 1 we have added two types of isotropic α-stable noise; we used Cauchy noise $\alpha = 1$ and α-stable noise with $\alpha = 1.5$. The WVDs of both corrupted signals are shown in Fig. 2. In both plots the interference terms discussed earlier are readily apparent,

particularly the strong ridges located in time at the location of impulse events. We can also see intermodulation effects between the impulsive ridges, which arise due to the quadratic structure of the WVD. Note the lack of effect the Choi-Williams kernel, also applied here, has on the impulse-related artifacts.

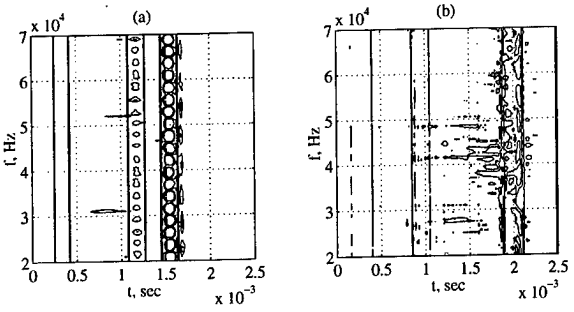


Figure 2. WVDs of noisy echolocation signal in stable noise with (a): $\alpha = 1$, (b): $\alpha = 1.5$.

In Fig. 3, we show the WVDs that are produced when PFLOM-type influence functions are used. In Fig. 3(a) we have used $a = 0.1$ for the Cauchy noise, and we have used $a = 0.25$ for the $\alpha = 1.5$ noise. in Fig. 3(b). We then applied the same Choi-Williams kernel to both representations.

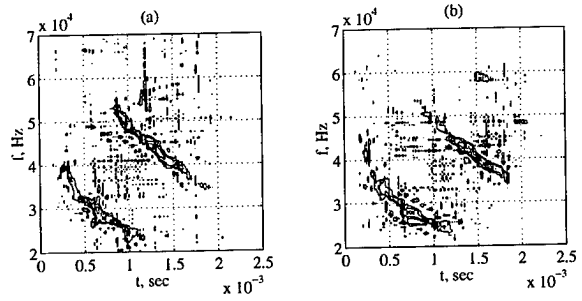


Figure 3. FLOWVDs of noisy echolocation signal in stable noise with (a): $\alpha = 1$, (b): $\alpha = 1.5$.

For some applications involving impulsive noise, applying a median filter to the signal before computing its WVD is a useful way to remove the effect of impulses. This approach works well if the desired signal does not have significant energy in the frequencies near its Nyquist frequency, and if the additive impulses are not dense. As an example of the performance degradation that a median pre-filter can suffer in an impulse-heavy environment, we consider the case where we the echolocation signal is contaminated by additive Cauchy noise. In Fig. 4, we show the results of median pre-filtering. A length-3 pre-filter was used in Fig. 4(a); here the high density of impulses results in many outliers remaining in the signal after processing. These unwanted impulses can be removed by using a wider filter, as shown in Fig. 4(b), but this causes many signal details to be lost. The end result is that in Cauchy noise, the median filter does not yield satisfactory results.

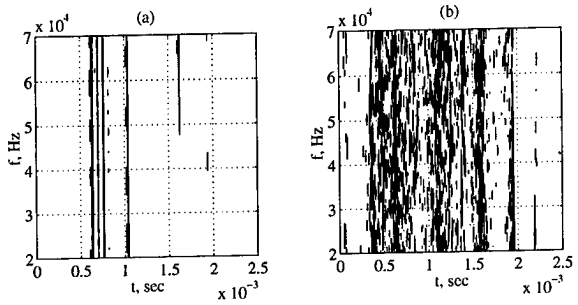


Figure 4. WVD of echolocation signal in Cauchy noise via prefiltering with (a) length-3 median filter, (b) length-5 median filter.

To demonstrate the effect that fractional lower-order moments have on the WVD of an impulsive signal, we apply the signal reconstruction algorithm developed by Boudreaux-Bartels and Parks in [2]. We first isolate the second chirp in the clean signal shown in Fig. 1(a) using a masking function, which sets all values of the WVD outside a desired region (in this case, a trapezoid) to zero. The isolated chirp is shown in Fig. 5(a). The reconstruction algorithm is then applied to the masked WVD, and the resulting nonlinear FM signal is shown in Fig. 5(b). By applying the same mask to the FLOWVD of the echolocation signal in Fig. 3(a), we obtain the TFR in Fig. 5(c). If we apply the algorithm from [2] to this TFR, we obtain a reconstruction of part of $x^{\langle a \rangle}(t)$. To obtain the chirp from $x(t)$ itself, it is necessary to undo the nonlinearity by applying a PFLOM of order $1/a$ to the output of the reconstruction algorithm; the result is shown in Fig. 5(d). Note that applying this second nonlinearity tends to emphasize some of the variations that often arise in the envelope of the reconstructed signal.

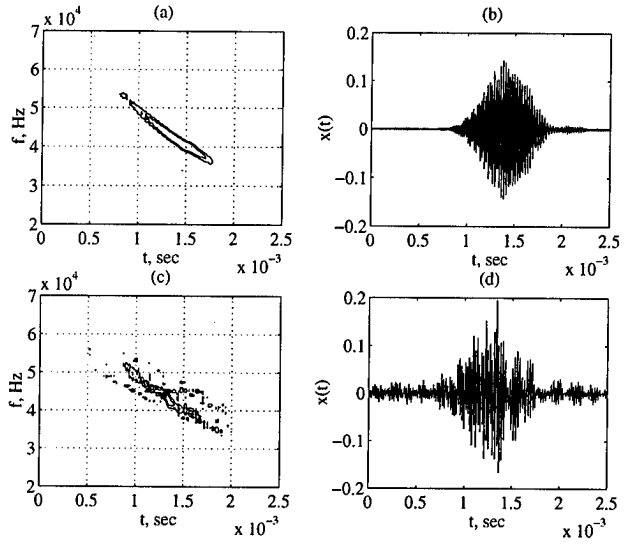


Figure 5. (a): Masked WVD of clean echolocation signal. (b): Reconstructed chirp from masked WVD. (c): Masked FLOWVD of echolocation signal in Cauchy noise. (d): Reconstructed chirp from masked FLOWVD.

5. Summary

In this paper, we have considered the effect of additive impulsive noise modeled by α-stable distributions on quadratic time-frequency representations such as the Wigner-Ville Distribution. We have shown that the presence of impulses in additive noise produces several undesirable effects. These effects include the generation of time-scaled replicas of the desired signal as well as strong ridges extending over all frequencies at the time location of outliers.

To combat these effects, we introduced a robust evolutive spectrum, based on the fractional lower-order covariance. From this, we were able to define the FLOWVD, which has many of the desirable mathematical properties associated with the standard WVD.

To demonstrate the performance improvement that can be realized with such a technique, we examined a bat echolocation signal which we contaminated with two different α-stable noise types. We showed that the standard WVD breaks down completely in the presence of such noise, and we showed that techniques such as median pre-filtering do not always produce good results. We then examined the FLOWVD for these noise types and showed that it was more robust than the WVD. Further, we were able to isolate components of the original signal by masking the FLOWVD and applying standard reconstruction algorithms.

References

[1] B. Boashash. Time-frequency signal analysis. In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, chapter 9, pages 418–517. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1991.

[2] G. F. Boudreaux-Bartels and T. W. Parks. "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques". *The IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(3):442–451, June 1986.

[3] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. "The Wigner distribution–a tool for time-frequency signal analysis. Part II: Discrete-time signals". *Philips Journal of Research*, 35:276–300, 1980.

[4] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. "The aliasing problem in discrete-time Wigner distributions". *The IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(5):1067–1072, October 1983.

[5] L. Cohen. "Time-frequency distributions–A review". *Proceedings of the IEEE*, 77(7):941–981, July 1989.

[6] F. Hlawatsch and G. F. Boudreaux-Bartels. "Linear and quadratic time-frequency signal representations". *IEEE Signal Processing Magazine*, 9(4):21–67, April 1992.

[7] X. Ma and C. L. Nikias. "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower-order statistics". *The IEEE Transactions on Signal Processing*, 1997. Publication pending.

[8] D. Middleton. "Statistical-physical models of electromagnetic interference". *The IEEE Transactions on Electromagnetic Compatibility*, 19(3):106–127, August 1977.

[9] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. John Wiley & Sons, Inc., New York, 1995.

[10] S. Qian and D. Chen. *Joint Time-Frequency Analysis: Methods and Applications*. Prentice-Hall PTR, Upper Saddle River, NJ, 1996.

[11] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, New York, 1994.

Application of the Positive Alpha-Stable Distribution

Robert D. Pierce
Naval Surface Warfare Center-Carderock Division

Abstract

Many physical phenomena are non-Gaussian and if the observed data have frequently occurring extreme values, then the phenomena may be modeled as a random process with an alpha-stable distribution. When positive and negative outcomes are equally likely, then the process would be symmetric alpha-stable ($S\alpha S$); however, when only positive outcomes are possible, then the process would be positive alpha-stable ($P\alpha S$). Phenomena related to energy or power are examples. This paper presents the characteristics and potential applications for the $P\alpha S$ distribution. For this distribution all negative-order moments exist, and ratios of these moments are used to estimate alpha. Application areas that will be examined include: seismic activity, ocean wave variability, and radar sea clutter modulation. The correlation properties of these data are examined.

1. Introduction

The alpha-stable distribution is an important area for investigation because this distribution is expected from superposition in natural processes. The parameter alpha, α, is the characteristic exponent that varies over $0 < \alpha \leq 2$. The distribution includes the Gaussian when $\alpha = 2$. For α less than two, the distribution becomes more impulsive, more non-Gaussian in nature, and the tails of the distribution become thicker. This makes the alpha-stable distribution an attractive choice for modeling signals and noise having an impulsive nature. Also, from the generalized central limit theorem, the stable distribution is the only limiting distribution for sums of independent and identically distributed (IID) random variables (stability property). If the individual distributions have finite variance, then the limiting distribution is Gaussian. For α less than two, the individual distributions have infinite variance. For detailed information, see the books by Nikias and Shao [1] and Samorodnitsky and Taqqu [2].

Sources that could follow or be modeled by the alpha-stable distribution are abundant and include lightning in the atmosphere, switching transients in power lines, static in telephone lines, seismic activity, climatology and weather, ocean wave variability, surface texture, the slamming of a ship hull in a seaway, acoustic emissions from cracks growing in engineering materials under stress, etc. Many sources can exist in the area of target and background signatures that affect detection and classification. In underwater acoustics, examples of these sources could include interference to target detection such as ice cracking, biologics, bottom and sea clutter in active sonar, ocean waves near the surface and in the surf zone. They could also include target characteristics such as target strength in active sonar and cavitation. Similar sources in radar and infrared can include: ocean waves in the form of sea clutter and radar cross section (RCS); see Pierce [3].

This paper will examine characteristics and applications of the positive alpha-stable ($P\alpha S$) distribution. This distribution is related to the symmetric alpha-stable ($S\alpha S$) distribution in that every $S\alpha S$ random variable can be represented as the product (or modulation) of a zero mean Gaussian random variable and the square-root of a $P\alpha S$ random variable where the alpha parameter for the $P\alpha S$ distribution is half the alpha associated with the $S\alpha S$ distribution [2]. The $P\alpha S$ alpha varies over $0 < \alpha < 1$. The Gaussian and $P\alpha S$ distributions must be statistically independent. This property is regularly used for the computer generation of $S\alpha S$ random variables. From this property and the stability property, the estimation of signal power (second moment estimates) from short averages of $S\alpha S$ random variables approaches a $P\alpha S$ random variable where the alpha is unchanged by averaging. The $P\alpha S$ random process could be observed when energy or power varies over time or space.

The general theory and methods for $P\alpha S$ processes are developed and shown where they can be related to power or energy flow in physical processes. This includes derivation of the moments for this distribution, the range of support for these moments (all negative order moments exist) and the statistical error associated with estimating these moments. A method for estimating alpha is presented which uses the ratio of negative order moments. A method is presented for estimating the correlation between samples from $P\alpha S$ processes. Results are presented using these methods with actual data.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Positive alpha-stable distribution

The positive alpha-stable distribution has the characteristic function

$$\varphi(t) = \exp\left(-\gamma|t|^{\alpha}\left[1 + i\,\mathrm{sgn}(t)\tan(\alpha\pi/2)\right]\right) \quad (1)$$

where

$$\mathrm{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \quad (2)$$

and

$$0 < \alpha < 1,\ \gamma > 0 \quad (3)$$

The general characteristic function is from [1] where the location parameter is zero ($a = 0$) and the symmetry parameter is one ($\beta = 1$). The characteristic exponent is α, and the dispersion or scale parameter is γ.

Closed-form expressions do not exist for the $P\alpha S$ probability density function (PDF) except for the Pearson (or Lévy) distribution ($\alpha = 0.5$). Fig. 1 presents a family of $P\alpha S$ PDFs obtained by numerically taking the inverse Fourier transform of the characteristic function given in (1). Alpha varies from 0.3 to 0.95 in steps of 0.05. As alpha approaches one, the density function becomes more and more centered about the normalized amplitude of one. The tails show the power-law or algebraic asymptote that is a characteristic of all alpha-stable distributions. The amplitude in Fig. 1 is normalized by the negative first-order moment; the topic of moments is discussed next.

The moments for the $P\alpha S$ distribution are defined by the expectation

$$m_p = E\left[x^p\right] \quad (4)$$

where x is a $P\alpha S$ random variable, and for sampled data, x_n, the moment is estimated by

$$\hat{m}_p = \frac{1}{N}\sum_{n=1}^{N} x_n^p \quad (5)$$

One can derive directly or from [2] determine the moments for $0 \le p < \alpha$

$$m_p = \gamma^{\frac{p}{\alpha}} \frac{\sin(\pi p)}{\alpha \sin\left(\frac{\pi p}{\alpha}\right)} \frac{\Gamma(1+p)}{\Gamma\left(1+\frac{p}{\alpha}\right)} \left[1 + \tan^2\left(\frac{\pi\alpha}{2}\right)\right]^{\frac{p}{2\alpha}} \quad (6)$$

Although negative moments are not considered in [2], the negative order moments of all orders ($-\infty < p < 0$) are

$$m_p = \frac{\gamma^{\frac{p}{\alpha}}}{\alpha} \frac{\Gamma\left(-\frac{p}{\alpha}\right)}{\Gamma(-p)} \left[1 + \tan^2\left(\frac{\pi\alpha}{2}\right)\right]^{\frac{p}{2\alpha}} \quad (7)$$

In this paper, the negative first-order moment is used for normalization, such that the product xm_{-1} has a negative first-order moment of one and dispersion

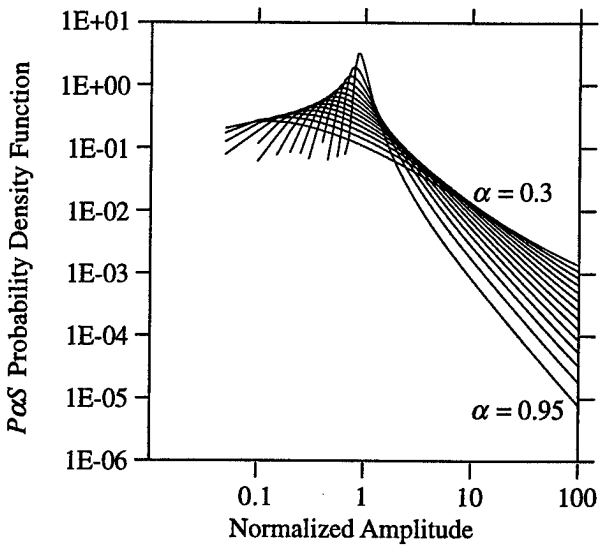


Fig. 1. Probability density function for $P\alpha S$ distribution from $\alpha = 0.3$ to $\alpha = 0.95$

$$\gamma = \left(\frac{\Gamma(1/\alpha)}{\alpha}\right)^{\alpha} \frac{1}{\sqrt{1 + \tan^2(\pi\alpha/2)}} \quad (8)$$

Taking the ratio of moments gives a method for estimating alpha. In this paper, the selected ratio is

$$r = \frac{m_{-0.5}^2}{m_{-1}} = \frac{\Gamma^2(1/2\alpha)}{\Gamma(1/\alpha)} \frac{1}{\pi\alpha} \quad (9)$$

The moments are estimated using (5), and the inverse of (9) is approximated over the range $0.3 \le \alpha(r) < 1$ as the polynomial

$$\begin{aligned} \alpha(r) = {} & -29.8573r^8 + 135.681r^7 - 254.293r^6 \\ & + 255.965r^5 - 149.978r^4 + 52.2708r^3 \\ & - 10.4879r^2 + 1.60231r + 0.0965472 \end{aligned} \quad (10)$$

The random error associated with moment estimation is derived as

$$VAR\left[\hat{m}_p\right] = \frac{1}{N}\left(m_{2p} - m_p^2\right) \quad (11)$$

The p^{th} root is taken to scale the moment estimate such that the result is proportional to amplitude. For large sample size, the approximate variance is

$$VAR\left[\hat{m}_p^{1/p}\right] = m_p^{2/p} \frac{1}{p^2} \frac{1}{N} \frac{\left(m_{2p} - m_p^2\right)}{m_p^2} \quad (12)$$

The normalized error becomes

$$\sqrt{N}\,\frac{\sqrt{VAR\left[\hat{m}_p^{1/p}\right]}}{m_p^{1/p}} = \frac{1}{|p|}\,\frac{\sqrt{m_{2p} - m_p^2}}{m_p} \quad (13)$$

Plots of (13) for a range of alpha are given in Fig. 2. For moments with order less than -0.5 and alpha greater than

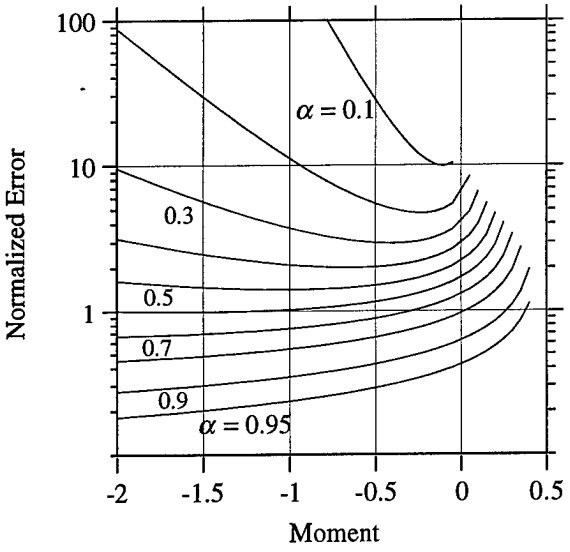


Fig. 2. Normalized error for moments of $P\alpha S$ distribution for $\alpha=0.1$ to $\alpha=0.95$

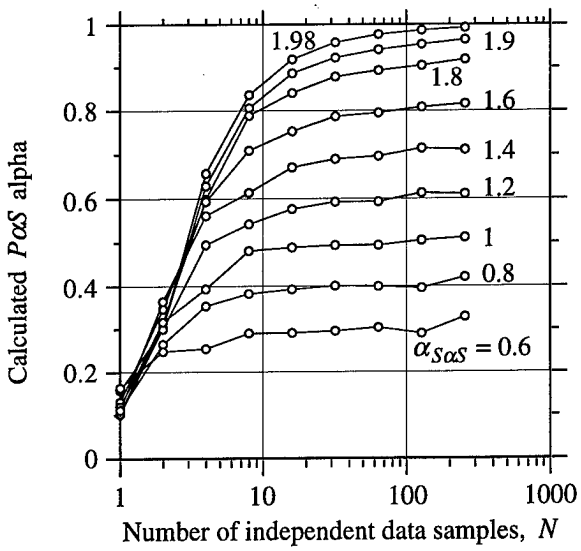


Fig. 3. $P\alpha S$ alpha calculated from multiple second moment estimates for various $S\alpha S$ alpha

0.4, the error is determined more by alpha than by the order of the moment. The use of -0.5 and -1 order moments for alpha estimation and normalization appears to be reasonable for most applications.

The usual moments of mean and mean-square are infinite, and estimates of these moments are not consistent. The median of the $P\alpha S$ distribution, however, is also a measure of the distribution's scale. The relationship between median and the moments is established by using PDF calculations and numerical integration (not shown).

Autoregressive and moving average models of $P\alpha S$ processes must have positive or zero coefficients. Otherwise, the sum would not be $P\alpha S$ distributed (stability property).

From [2], the $S\alpha S$ random variable is related to the $P\alpha S$ random variable. If g is a zero mean Gaussian random variable and x is a $P\alpha S$ random variable, where g and x are statistically independent, then

$$z = x^{1/2} g \tag{14}$$

is a $S\alpha S$ random variable with characteristic exponent

$$\alpha_{S\alpha S} = 2\alpha_{P\alpha S} \tag{15}$$

This property shows that every $S\alpha S$ random variable is conditionally Gaussian [2].

For some applications, the power estimated from a $S\alpha S$ distributed random process is the random variable of interest. Power or second moment estimates of z_n can be viewed as $P\alpha S$ random variables when estimated from a sufficiently large number of data samples, N. Using (14) the power or second moment estimate is

$$\hat{m}_{z,2} = \frac{1}{N}\sum_{n=1}^{N} z_n^2 = \frac{1}{N}\sum_{n=1}^{N} x_n g_n^2 \tag{16}$$

and the moment conditional on a set of x_n is

$$E\left[\hat{m}_{z,2} \middle| x_1 \cdots x_n\right] = \sigma_g^2 \frac{1}{N}\sum_{n=1}^{N} x_n \tag{17}$$

where σ_g^2 is the variance of g_n. As each set of x_n is drawn randomly, the sum of N x_n samples is itself $P\alpha S$ distributed (stability property). In practice, N must be large enough to estimate σ_g^2. Fig. 3 gives an indication of the number of independent samples necessary to produce second moment estimates that are $P\alpha S$ distributed. At each N and $\alpha_{S\alpha S}$, 1000 second order moments were estimated using (16) with the $S\alpha S$ independent samples generated using the computer algorithm given in [2]. The $\alpha_{P\alpha S}$ was then calculated from these power estimates using the moment ratio method. For example, it takes about 32 independent samples at an $\alpha_{S\alpha S}=1.2$ to give an $\alpha_{P\alpha S}=0.6$ as required by (15). As alpha increases, more samples are required.

A method for investigating the correlation between $P\alpha S$ random variables is suggested by the negative order moments. The cross correlation coefficient estimated between the two random variables, x and y, is defined as

$$\hat{\rho}_{xy}(j) = \frac{\frac{1}{N-j}\sum_{n=1}^{N-j} \frac{1}{x_n y_{n-j}} - \left(\hat{m}_{x,-1}\hat{m}_{y,-1}\right)^2}{\sqrt{\left(\hat{m}_{x,-2} - \hat{m}_{x,-1}^2\right)\left(\hat{m}_{y,-2} - \hat{m}_{y,-1}^2\right)}} \tag{18}$$

One can easily show that if x and y are uncorrelated, then ρ_{xy} is zero (The reverse does not follow.) And if x and y are perfectly correlated, then ρ_{xy} is one. This method can be viewed as the usual correlation definition with a $1/x$ transformation. Other characteristics and the full usefulness of this coefficient are being investigated. It is applied to the data in the applications.

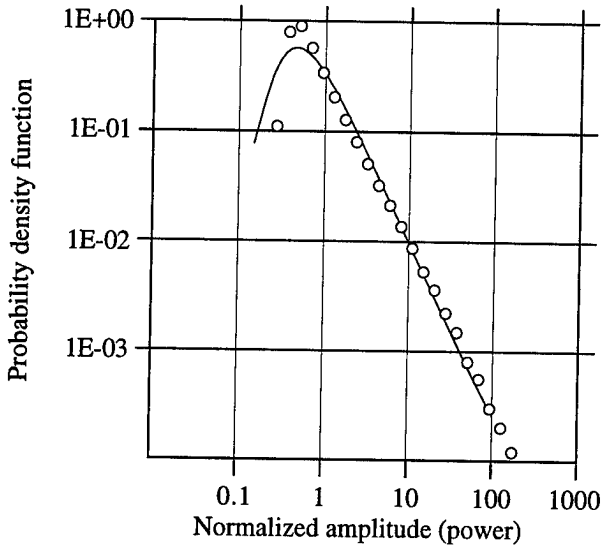


Fig. 4. Comparison of PDF for alpha=0.6 with PDF estimated from histogram of H-pol sea clutter

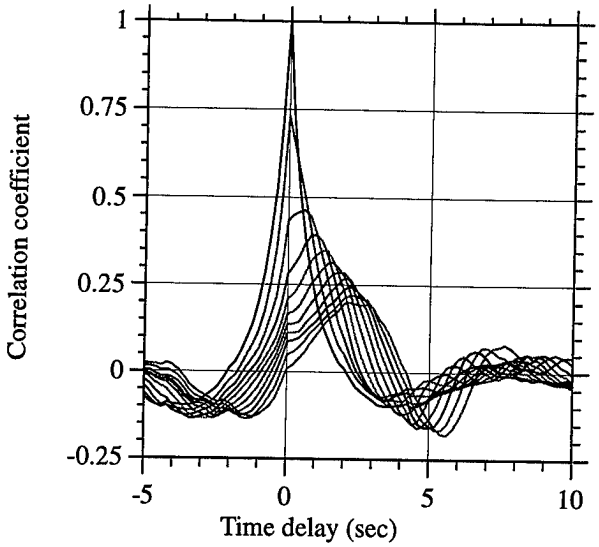


Fig. 5. Auto and cross correlation coefficients for H-pol sea clutter data sampled every two meters

3. Applications

For the phenomenology of sea clutter, the positive alpha-stable distribution is shown to closely model the modulation or signal power flow in sea clutter samples. An example comparing a sea clutter histogram to a $P\alpha S$ PDF is given in Fig. 4. An X-band radar was used to collect the H-pol clutter data. The data were taken in a low sea state 2 at a 0.9° grazing angle, 0.5 nautical miles offshore looking into the wind. The downrange spatial resolution is 2 meters, and the crossrange resolution is 35.6 meters. The square of the I and Q samples were averaged over 0.2 seconds to produce the sea clutter power estimates that are compared to the $P\alpha S$ distribution with an alpha of 0.6. The $\alpha_{P\alpha S}$ estimated using the moment ratio method is 0.612. This compares well to the envelope results given in [3] where an $\alpha_{S\alpha S}$ of 1.3 is estimated.

The correlation between signal power flow in sea clutter samples is shown in Fig. 5. The plot that peaks at zero time delay is the auto correlation. Since radar returns are available from 22 downrange cells spaced every two meters, cross correlation estimates are possible. The other 10 plots superimposed on Fig. 5 represent the cross correlation as the channel pairs are separated by an additional 2 meters for each plot. The negative coefficient values may be artifacts. A small amount of correlated noise appears to be present at zero time delay. The auto correlation decorrelates in about 2 to 3 seconds. The clutter data decorrelates significantly but is still well correlated between samples taken 20 meters apart.

Using data available on the World Wide Web as examples, the $P\alpha S$ distribution is shown to be a good candidate for modeling the variabilities in long term energy measurements for seismic activity ($\alpha_{P\alpha S} = 0.564$, from moment ratio method) and ocean waves ($\alpha_{P\alpha S} = 0.731$, from moment ratio method). The seismic data is from a catalog of events presented by the Southern California Seismographic Network (SCSN) at http://scec.gps.caltech.edu:80/ftp/catalogs/SCSN. The Richter magnitude of the event is converted into the average energy released by the event, and the random variable is this energy accumulated over four tenths of a day (or power). These data were taken from 1/1/96 to 7/30/96. The comparison of the PDF for alpha = 0.55 with the PDF from the histogram of the seismic data is given in Fig. 6. A power-law tail is expected from the empirical Gutenburg-Richter Law; however, the $P\alpha S$ PDF better models the less frequently occurring, smaller seismic events. Similar results were obtained using crumpling paper audio data given at http://garak.msc.cornell.edu/~houle/crumpling/crumpling.html. As shown in Fig. 7, the seismic data appears to fully decorrelate from one sample to the next. Out to 10 days the auto correlation coefficient appears to be slightly positive.

The ocean wave data was taken from a data base presented by the Ocean Data Buoy Center, a part of the National Oceanic and Atmospheric Administration (NOAA), at http://seaboard.ndbc.noaa.gov/historical_data.shtml. The web site presents hourly estimates of the significant wave height calculated from 20 minutes of wave data. The data selected was from Station ID 44014 (off Virginia Beach, VA) from 3/1/96 to 8/2/96. The random variable is the square of the significant wave height which is proportional to the power in the ocean waves (Wave height is assumed to be Gaussian over the 20 minutes.) The comparison of the PDF for alpha = 0.75 with the PDF from the histogram of the seismic data is given in Fig. 8. As shown in Fig. 9, the wave height data appears to decorrelate after 48 hours with a secondary peak at about 80 hours.

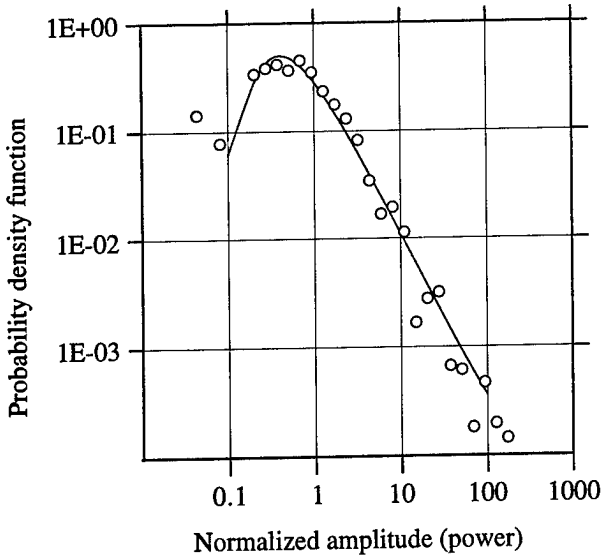


Fig. 6. Comparison of PDF for alpha=0.55 with PDF estimated from histogram of seismic data

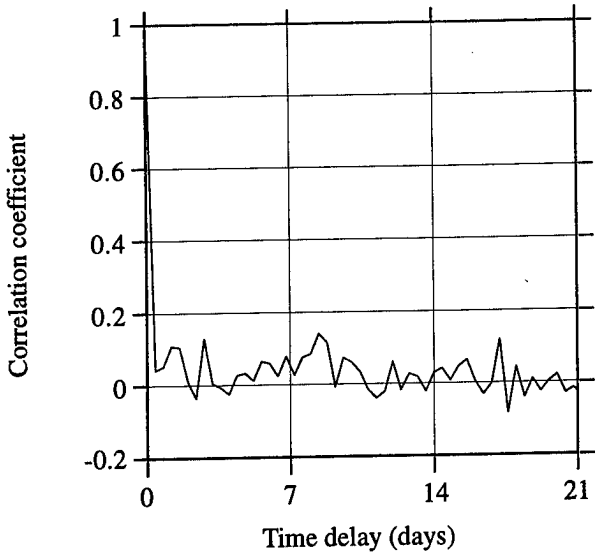


Fig. 7. Auto correlation coefficient for seismic data

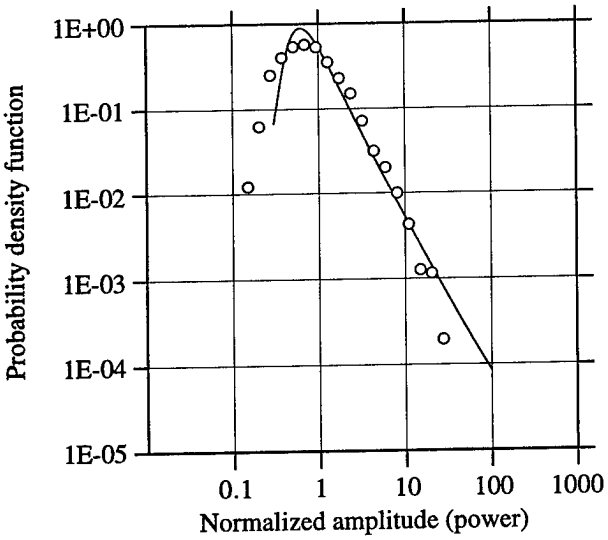


Fig. 8. Comparison of PDF for alpha=0.75 with PDF estimated from histogram of wave height data

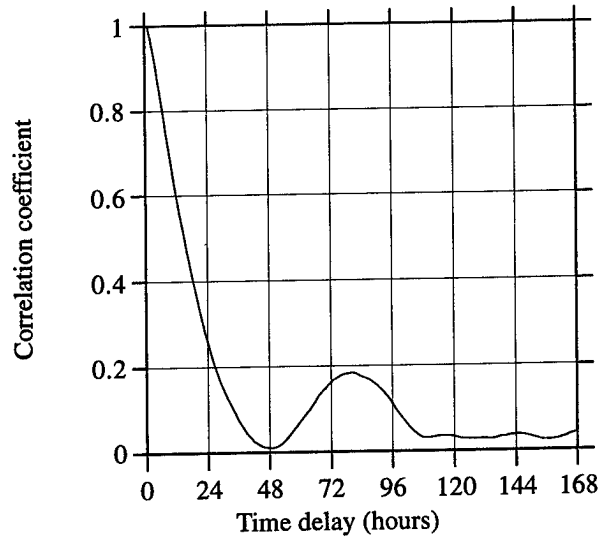


Fig. 9. Auto correlation coefficient for wave height data

References

[1] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, John Wiley and Sons, New York, NY, 1995.

[2] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, NY, 1994.

[3] R. D. Pierce, "RCS Characterization using the alpha-stable distribution," Proc. 1996 IEEE National Radar Conference, Ann Arbor, MI, 13-16 May, 1996, pp. 154-159.

Administrative information

This project was supported by the Carderock Division of the Naval Surface Warfare Center's In-house Laboratory Independent Research Program sponsored by the Office of Naval Research and administrated by the Research Director, Code 0112 under Program Element 0601152N under NSWCCD Work Unit 1-7340-506.

Asymptotic Distribution of the Hermite Normality Test

David Declercq Patrick Duvaut
ENSEA-ETIS 6, avenue du PONCEAU 95014, Cergy-Pontoise, France
e-mail : declercq@ensea.fr

Abstract

This paper presents some asymptotical results of the Hermite Normality test previously introduced. We show that the Hermite statistic S_H is distributed under the null hypothesis as a quadratic form of normal variates and under the nonnull hypothesis as normal. The special case of tests with two polynomials is studied in details. Finally, we give some considerations for the choice of a best Hermite test when prior knowledge is available and especially we determine the test asymptotically the most powerful for a fixed alternative distribution (the uniform distribution). Those results are supported by simulations.

1 Introduction

Testing normality is one of the most studied problem in statistics and signal processing for the obvious reason that the normal distribution is omnipresent in those topics [9] [5] [10]. A major problem of all the tests that have been introduced in the past fifty years is their matching to the nonnormal alternative considered. Some work best than others for a particular type of alternative as their behaviour are reversed for other types.
This particulary means that when the nonnull hypothesis is not known, one cannot claim that a test is more powerfull than another and the problem of the choice of a test appears.
This remark led us to emphasize that the Hermite Normality test [3] actually defines a wide range of tests, the properties of each depend on certain properties of the underlying distribution ; therefore, if we have any prior knowlegde on the distribution we have to test (symetrical, tails or conditions on higher cumulants), there exists a best Hermite test that matches this distribution.
After recalling the form of the Hermite Normality test in the first section, we give its asymptotical behaviour. Section 2 will show that it is distributed as a quadratic form of gaussian variates under the null hypothesis and in the general case of p polynomials, and section 3 gives the asymptotical distribution of the test if only two polynomials are considered both for the null hypothesis and a nonnull particular one (uniform distribution), supported by numerical simulations. We end up this contribution with some remarks on the choice of a *best Hermite test* and a few cautions if the tested samples are small sized.

2 Construction of the test

The Hermite Normality Test has been introduced in [3] and we briefly describe below its generation.
The first step consists in building a nonlinear vector which contains standardised Hermite polynomials of a random variate x.

$$\boldsymbol{X} = \left[\frac{H_{i_1}(x)}{\sqrt{i_1!}}, \frac{H_{i_2}(x)}{\sqrt{i_2!}}, \ldots, \frac{H_{i_p}(x)}{\sqrt{i_p!}} \right]^T \qquad (1)$$

Where all the subscripts i_k are distincts.
Because of the properties of Hermite polynomials [6], when x is $\mathcal{N}(0,1)$, this vector becomes *spherical* : it is zero mean and its covariance matrix is identity ; property that vanishes for nonnormal distributions. Therefore, to test if a random sample is gaussian or not, we apply $\boldsymbol{X}$ to the sphericity statistic

$$S_H = \frac{|\boldsymbol{R}|}{\left(\frac{Tr(\boldsymbol{R})}{p}\right)^p} \qquad S_H \in [0,1] \qquad (2)$$

where $\boldsymbol{R}$ is the sample covariance matrix of $\boldsymbol{X}$, $|\boldsymbol{R}|$ its determinant and $Tr(\boldsymbol{R})$ its trace.

3 Asymptotical distribution of S_H under $\mathcal{H}_0$

Although our test has a sphericity structure, we can not apply the results pointed out by Anderson [1] be-

0-8186-8005-9/97 $10.00 © 1997 IEEE

cause they consider the sphericity statistic with a *normal* vector - and the one we consider (1) is not.
We therefore have to introduce a limit theorem derived with the help of two from Borovkov ([2], pp. 44) with tensorial notations. If $\boldsymbol{A}$ is a matrix in $\mathbb{R}^p$, $\boldsymbol{A}^*$ will denote its dual form, i.e. a morphism of $\mathcal{L}(\mathbb{R}^p;\mathbb{R})$ and $\otimes$ is the classical tensor product.

Theorem 1 *let x be a random variable with the distribution F_0 and the two functions*

$$\begin{array}{ccccc} \boldsymbol{G}: & \mathbb{R} & \to & \mathcal{M}(p,p) \\ & x & & \boldsymbol{G}(x) \end{array} \qquad \begin{array}{ccccc} h: & \mathcal{M}(p,p) & \to & \mathbb{R} \\ & \boldsymbol{T} & & h(\boldsymbol{T}) \end{array}$$

where $\mathcal{M}(p,p)$ is the space of square real valued matrix of size p. Let furthermore $h(\boldsymbol{T})$ be continuous in $\boldsymbol{A} = \int \boldsymbol{G}(x)dF_0(x)$ and the tensor of covariance $\underline{\boldsymbol{\Sigma}} = \int(\boldsymbol{G}-\boldsymbol{A})\otimes(\boldsymbol{G}-\boldsymbol{A})^* dF_0(x)$ is finite. We have then the following result

$$S_N(x) = h\left(\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{G}(x_k)\right) \xrightarrow{a.s.} h(\boldsymbol{A})$$

$$\sqrt{N}\left(S_N(x)-h(\boldsymbol{A})\right) \in \boldsymbol{h}'(\boldsymbol{A})^* \otimes \boldsymbol{\xi} = \sum_{i=1}^{p}\sum_{j=1}^{p}\frac{\partial h(\boldsymbol{A})}{\partial t_{ij}}\xi_{ij}$$

where $\boldsymbol{h}'(\boldsymbol{T})$ is the matrix of derivatives of $h(\boldsymbol{T})$ and $\boldsymbol{\xi}\in\mathcal{N}(\boldsymbol{0},\underline{\boldsymbol{\Sigma}})$ is a matrix whose components are centered normal with the tensor of covariance $\underline{\boldsymbol{\Sigma}}$.

If $\boldsymbol{h}'(\boldsymbol{A})^* \otimes \boldsymbol{\xi} \overset{a.s.}{=} 0$ and the tensor of the second order derivatives $\underline{\boldsymbol{h}}''(\boldsymbol{T})$ is finite in $\boldsymbol{A}$, then

$$\begin{aligned} N\left(S_N(x)-h(\boldsymbol{A})\right) &\in \frac{1}{2}\boldsymbol{\xi}^* \otimes \underline{\boldsymbol{h}}''(\boldsymbol{A}) \otimes \boldsymbol{\xi} \\ &\in \frac{1}{2}\sum_{i,j,r,s=1}^{p}\frac{\partial^2 h(\boldsymbol{A})}{\partial t_{ij}\partial t_{rs}}\xi_{ij}\xi_{rs} \end{aligned} \tag{3}$$

■

The proof is straigthforward while it is just an extension of the Borovkov's theorems. We have then to introduce some notations in order to apply this theorem to the Hermite test :

$$\left\{\begin{array}{l} g_{ij}(x) = \bar{H}_i(x)\bar{H}_j(x) \quad \forall(i,j)\in\{1,\ldots,p\} \\ \boldsymbol{G} = [g_{ij}(x)] \quad \text{symetrical matrix} \\ \boldsymbol{A} = E\left[\boldsymbol{G}(x)\right] = [a_{ij}] \\ \boldsymbol{R} = \left[\frac{1}{N}\sum_k g_{ij}(x_k)\right] = \widehat{\boldsymbol{A}} \\ \quad \text{is the sample corariance matrix of (1)} \\ \underline{\boldsymbol{\Sigma}} = E\left[(\boldsymbol{G}-\boldsymbol{A})\otimes(\boldsymbol{G}-\boldsymbol{A})^*\right] = [\sigma_{ij}^{rs}] \\ \sigma_{ij}^{rs} = E\left[(g_{ij}(x)-E\left[g_{ij}(x)\right])(g_{rs}(x)-E\left[g_{rs}(x)\right])\right] \\ \quad \textit{with} \quad 1<i,j,r,s<p \\ S_H = h(\boldsymbol{R}) = \dfrac{|\boldsymbol{R}|}{\left(\frac{Tr(\boldsymbol{R})}{p}\right)^p} \end{array}\right. \tag{4}$$

with these notations, we have prooved the following theorem

Theorem 2
Distribution of S_H under the null hypothesis
let $x\in\mathcal{N}(0,1)$, the Hermite Normality statistic (2) verifies

$$S_H(x) \xrightarrow{a.s.} 1$$

and has the asymptotical behaviour

$$N\left(1-S_H(x)\right)\Big) \in \frac{1}{2}\sum_{i=1}^{p}\xi_{ii}^2 - \frac{1}{2p}\left(\sum_{i=1}^{p}\xi_{ii}\right)^2 + \sum_{j>i}\xi_{ij}^2 \tag{5}$$

where $\boldsymbol{\xi} = [\xi_{ij}]\in\mathcal{N}(\boldsymbol{0},\underline{\boldsymbol{\Sigma}})$
and the distincts terms of $\underline{\boldsymbol{\Sigma}}$ are

$$\{\sigma_{ii}^{ii}, \sigma_{ii}^{jj}, \sigma_{ij}^{ij}, \sigma_{ii}^{rs}, \sigma_{ij}^{rs}\}$$

■

For obvious reasons of space, we give here only the main steps of the proof, which can be found in [4].

1. $\boldsymbol{A} = \boldsymbol{I}$ implies that $S_H(x) \xrightarrow{a.s.} h(\boldsymbol{A}) = 1$.
2. with the matrix of first order derivatives of $h(\boldsymbol{G})$, we found that $\boldsymbol{h}'(\boldsymbol{A}) = \boldsymbol{0}$ and we need to calculate the tensor of second order derivatives of $h(\boldsymbol{G})$.
3. this tensor contains 7 types of terms

$$\frac{\partial^2 h(\boldsymbol{A})}{\partial^2 g_{ii}} = \frac{-p+1}{p} \qquad \frac{\partial^2 h(\boldsymbol{A})}{\partial g_{ii}\partial g_{jj}} = \frac{1}{p} \qquad \frac{\partial^2 h(\boldsymbol{A})}{\partial^2 g_{ij}} = -2$$

$$\frac{\partial^2 h(\boldsymbol{A})}{\partial g_{ii}\partial g_{ij}} = \frac{\partial^2 h(\boldsymbol{A})}{\partial g_{ii}\partial g_{rs}} = \frac{\partial^2 h(\boldsymbol{A})}{\partial g_{ij}\partial g_{ir}} = \frac{\partial^2 h(\boldsymbol{A})}{\partial g_{ij}\partial g_{rs}} = 0$$

   this yield to the result with the use of theorem 1.
4. the last step is to derive the expressions of the terms appearing in $\underline{\boldsymbol{\Sigma}}$: $\{\sigma_{ii}^{ii}, \sigma_{ii}^{jj}, \sigma_{ij}^{ij}, \sigma_{ii}^{ir}, \sigma_{ii}^{rs}, \sigma_{ij}^{rs}\}$ which can be calculated using combinatorial relations on Hermite polynomials.

The Hermite test is asymptotically distributed as a quadratic form of centered gaussian variates whose covariance coefficients depend on σ_{ij}^{rs}. In the general case, we cannot obtain a closed form for the pdf of this kind of distribution. However, one can find some results on the distribution of quadratic forms in gaussian variates in [8]. One important remark is that the convergence of the Hermite test is good, decreasing as $\frac{1}{N}$ when many tests have a convergence in $\frac{1}{\sqrt{N}}$.

4 Special case of two polynomials

An important case is when only two polynomials are considered because it is actually largely used in practice and in simulations. The main reason is that Hermite tests with two polynomials seem to be more robust when the sample size is small ($N < 50$).

4.1 distribution under $\mathcal{H}_0$

In this section, we just derive the quadratic form given in theorem 2 for two Hermite polynomials $H_m(x)$ and $H_n(x)$ and precise the limit distribution with the results in [8].

$$N\left(1 - S_H^{(m,n)}(x))\right) \in \left(\frac{\xi_1 - \xi_2}{2}\right)^2 + \xi_3^2 = Q$$

which has the characteristic function

$$\phi(z) = \frac{1}{\sqrt{1 - 2iz\lambda_1}\sqrt{1 - 2iz\lambda_2}}$$

where λ_1 and λ_2 are the latent root of the covariance matrix of $[\frac{\xi_1 - \xi_2}{2}, \xi_3]^T$.
We therefore can easily calculate the cumulants of $S_H^{(m,n)}(x)$ under $\mathcal{H}_0$ and Rice ([11],p. 99) proposed to approximate the pdf of Q by a type III Pearson distribution (namely the *gamma* distribution) with same mean and variance.

$$f(x) = \left(\frac{\mu}{\sigma^2}\right)^r \frac{x^{r-1}}{\Gamma(r)} e^{-\frac{\mu x}{\sigma^2}} \quad with \quad r = \frac{\mu^2}{\sigma^2} \qquad (6)$$

In order to give a numerical example of those results, we have considered the first two Hermite polynomials $H_1(x)$ and $H_2(x)$ and compared the asymptotical distribution (6) with the histogram of 100000 samples of $N\left(1 - S_H^{(1,2)}(x))\right)$ for large samples ($N = 10000$), both drawn on figure 1.

4.2 distribution under $\mathcal{H}_1$

The ability to obtain the limit distribution under a particular alternative depends only on the calculations of $\boldsymbol{A}$ and $\underline{\boldsymbol{\Sigma}}$ (4) - and the knowlegde of the cumulants of x makes it possible. We have therefore, using theorem 1 the following result under any nonnull hypothesis

Theorem 3
Distribution of S_H under the nonnull hypothesis
Let $x \in f(x)$ a nonnormal random variable.

$$S_H(x) \xrightarrow{a.s.} h(\boldsymbol{A})$$

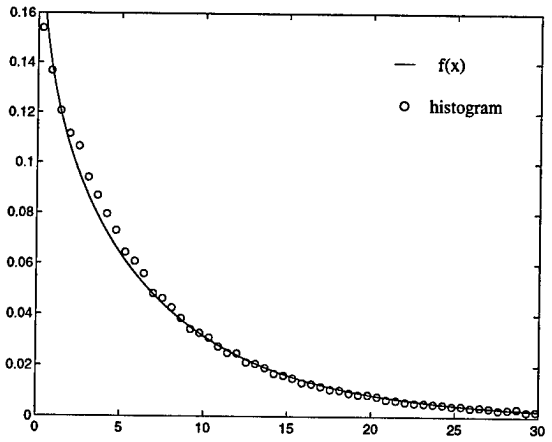


Figure 1. *distribution of $S_H^{(1,2)}$ under $\mathcal{H}_0$*

and the limit distribution is

$$\sqrt{N}(S_H(x) - h(\boldsymbol{A})) \in \boldsymbol{h}'(\boldsymbol{A})^* \otimes \boldsymbol{\xi}$$

$$\boldsymbol{\xi} \in \mathcal{N}(\boldsymbol{0}, \underline{\boldsymbol{\Sigma}})$$

■

S_H is asymptotically normal with a $\frac{1}{\sqrt{N}}$ rate of convergence and with mean $h(\boldsymbol{A})$) and variance $\sigma^2_{\mathcal{H}_1} = \boldsymbol{h}'(\boldsymbol{A})^* \otimes \underline{\boldsymbol{\Sigma}} \otimes \boldsymbol{h}'(\boldsymbol{A})$.
As an example, we will give in this section the asymptotical distribution of $S_H^{(m,n)}$ under the uniform hypothesis : $x \in \mathcal{U}([-\sqrt{3}, \sqrt{3}])$. Using some combinatorial relations on the product of 2, 3 and 4 Hermite polynomials (see [7] and the references within), one can derive the results (7), (8) and (9)

$$E\left[H_n(x)\right] = \frac{H_{n+1}(\sqrt{3}) - H_{n+1}(-\sqrt{3})}{2\sqrt{3}(n+1)} = c_n \qquad (8)$$

$$E\left[\frac{H_m(x)}{\sqrt{m!}}\frac{H_n(x)}{\sqrt{n!}}\right] = \frac{1}{\sqrt{m!n!}} \sum_{k=0}^{m} k! \binom{m}{k} \binom{n}{k} c_{m+n-2k} = a_{mn} \qquad (9)$$

With those relations, we can derive the expressions of $h(\boldsymbol{A})$ and $\sigma^2_{\mathcal{H}_1}$ which appear in the limit distribution of $S_H^{(m,n)}$. As a comparison with the numerical example given under $\mathcal{H}_0$, we consider the Hermite test with the first two polynomials. We obtain the mean of S_H : $h(\boldsymbol{A}) = 0.8163$ and its asymptotical variance $\sigma^2_{\mathcal{H}_1} = 0.0979$. Figure 2 draws the asymptotical gaussian density $\mathcal{N}(0, 0.0979)$ and the histogram of 100000 samples of $\sqrt{N}\left(S_H^{(1,2)}(x) - 0.8163\right)$ with $N = 10000$.

$$\sigma_{mn}^{pq} = E\left[\frac{H_m(x)}{\sqrt{m!}}\frac{H_n(x)}{\sqrt{n!}}\frac{H_p(x)}{\sqrt{p!}}\frac{H_q(x)}{\sqrt{q!}}\right] - a_{mn}a_{pq} = -a_{mn}a_{pq} + \frac{1}{\sqrt{m!n!p!q!}}\sum_{k=0}^{m}\sum_{k'=0}^{p} k!k'!\binom{m}{k}\binom{n}{k}\binom{p}{k'}\binom{q}{k'} \quad (7)$$
$$\sum_{k''=0}^{min\{(m+n-2k),(p+q-2k')\}} k''!\binom{m+n-2k}{k''}\binom{p+q-2k'}{k''} c_{m+n+p+q-2(k+k'+k'')}$$

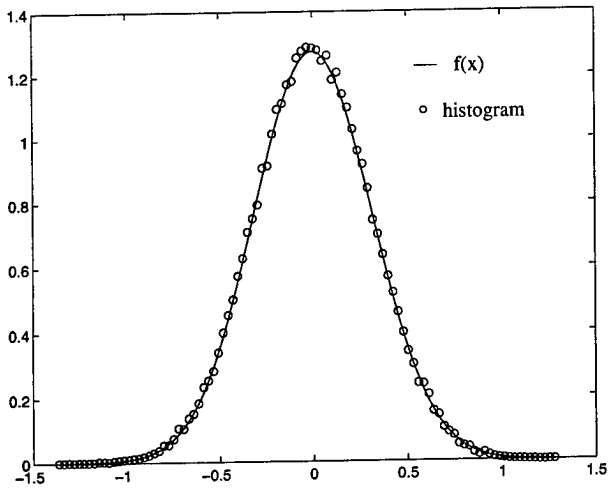


Figure 2. *distribution of* $S_H^{(1,2)}$ *under* $\mathcal{H}_1$

5 On the choice of a best test of normality

Since we have usually very few prior knowlegde about the distribution of a sample on which a decision must be taken (in our case - normal or not), it is of a great interest to know exactly the performances of the Hermite Normality test under a fixed alternative. In this section we will focus only on this aspect, keeping in mind that when nothing is known about the signal, those considerations vanish.

5.1 asymptotical study

As seen before, we are able to give the limit distribution of the Hermite Normality test for any alternative hypothesis provided we know the cumulants of the distribution. We can thereby obtain this distribution for all the tests with two polynomials and use a classical Neymann-Pearson type decision to find the Hermite test with 2 polynomials *asymptotically the most powerful* for the uniform alternative.
For that purpose, we have to maximize a *correct detection probability* (CDP) with a fixed *false alarm probability* (FAP). These two probabilities are defined as

$$\begin{aligned} CDP &= Prob\{\text{choose } \mathcal{H}_1 \,|\mathcal{H}_1\} \\ &= Prob\left\{S_H < \mu \,\middle|\, \sqrt{N}(S_H - h(\boldsymbol{A})) \in \mathcal{N}(0, \sigma^2_{\mathcal{H}_1})\right\} \end{aligned} \quad (10)$$

$$\begin{aligned} FAP &= Prob\{\text{choose } \mathcal{H}_1 \,|\mathcal{H}_0\} \\ &= Prob\left\{S_H < \mu \,|N(1 - S_H) \in Gamma(\mu_{\mathcal{H}_0}, \sigma^2_{\mathcal{H}_0})\right\} \end{aligned} \quad (11)$$

For reasons of calculation, we cannot explore the combinations of polynomials beyond the $6th$ degree.
We have found that the three best tests are

$$S_H^{(1,3)} \qquad S_H^{(2,4)} \qquad S_H^{(4,6)} \quad (12)$$

This result is of course valid regarding the hypothesis of the approach mentionned here about which we want to give some remarks :

- it is an asymptotical study : we propose to verify that $S_H^{(1,3)}$ is effectively the most powerful test under the uniform hypothesis for small size samples in the next section.
- the maximum degree of our investigation is limited to 6. It is not worrying because the robusteness of the Hermite tests when the sample size decreases is rather bad if large degree polynomials are considered.
- we consider only two polynomials : this is still a work in progress, and if we can follow the method presented here to give the limit distribution for more than two polynomials, we hope to give soon more general criterions to choose the set of polynomials that best fit a particular alternative.

5.2 small sample results

It is important to verify if the asymptotical behaviour of a test statistic is still the same if applied to small sized samples. In the case of the Hermite Normality test, we want to know if the result found on the most powerfull test under a particular alternative is still true. For that purpose, we have made power simulations based on samples of size $N \in \{30, 40, 50\}$ at 5% level of significance - that means with a 0.05 probability

of error under the null hypothesis. We have considered the three tests appointed asymptotically as the most powerful (12) and three other arbitrary choosen. In [3], we have compared The Hermite Normality test under a lot of alternatives with three famous tests taken in the litterature.
Table 1 gives the 5% quantiles of the Hermite tests for each N computed with 500000 samples of S_H - more complete tables are reported in [4].

N	$S_H^{(1,3)}$	$S_H^{(2,4)}$	$S_H^{(4,6)}$	$S_H^{(1,2)}$	$S_H^{(1,8)}$	$S_H^{(1,2,3)}$
20	0.278	0.128	0.079	0.619	0.427	0.158
30	0.379	0.156	0.107	0.711	0.492	0.223
50	0.507	0.203	0.159	0.802	0.559	0.322

Table 1. 5% quantiles of six Hermite tests

The power of the Hermite test is then the probability (in percentage) of a sample distributed as $\mathcal{U}([-\sqrt{3},\sqrt{3}])$ to fall above the corresponding quantile. Those powers are estimated with 5000 samples of size N and are given in table 2.

N	$S_H^{(1,3)}$	$S_H^{(2,4)}$	$S_H^{(4,6)}$	$S_H^{(1,2)}$	$S_H^{(1,8)}$	$S_H^{(1,2,3)}$
20	57	6	23	19	4	39
30	79	16	39	32	7	59
50	96	64	67	59	27	88

Table 2. power of the Hermite test under $\mathcal{H}_1$

Fisrt, it is pointed out that the asymptotical study gives good criterions, because the choosen tests are the most powerful Hermite test (with two polynomials), even for very small samples. Paying much more attention on the two tests $S_H^{(1,3)}$ and $S_H^{(4,6)}$, we see that they recover their asymptotical order (in term of power) as well as the sample size grows. We finally remark that the power of $S_H^{(1,2,3)}$ grows rapidely with N, what encourages us to find more general criterions for the choice of a test, including those with more than two polynomials.
We furthermore have to take care with small samples : the simulations for $N = 20$ exhibit powers too small to allow us to give comments on them ; even with samples of size 30, one can easily deduce that the test is relevant only for $S_H^{(1,3)}$ and $S_H^{(1,2,3)}$. As a matter of fact, if we take a look at the 4th column of table 2, $S_H^{(1,8)}$ falls almost surely in the gaussian area of S_H : for this particular test statistic point of view, one can consider that the uniform distribution is very close to the normal one !

6 Conclusion

We gave in this paper the limit distribution of the Hermite Normality test under the null hypothesis in the general case of p polynomials and in a particular one ($p = 2$) under both nul and nonnull hypothesis. The fixed alternative of normalised uniformity is studied - and numerical simulations support those results. The problem of the choice of a test statistic for a particular alternative is discussed and an answer is given for the uniform distribution. We first find the *asymptotically* most powerful Hermite test and thereafter verify if it is still the most powerfull for small samples ($N < 50$). Cautions are made if the sample are really too small and for some test statistics ; some future work will first give general choice criterions among the Hermite Normality test family and second propose some rules to apply when nothing is known about the underlying distribution, and this to avoid the problem encountered at the end of the last section : an almost surely bad decision on the tested distribution.

References

[1] T. Anderson. *An Introduction to Multivariate Statistical Analysis.* Wiley, New York, 1958.
[2] A. Borovkov. *Statistique mathématique.* Mir, Moscou, 1987.
[3] D. Declercq and P. Duvaut. Hermite normality tests. In *ICASSP-97*, Munich, 1997.
[4] D. Declercq and P. Duvaut. Hermite normality tests. *submitted in IEEE Tran. Inf. Theo.*, march 1997.
[5] A. Dyer. Comparison of tests for normality with a cautionary note. *Biometrika*, (61):185–189, 1974.
[6] A. Erdelyi and Al. *Higher transcendental Functions, Vol. II.* Bateman Manuscript Project, McGraw-Hill, New York, 1953.
[7] E. Feldheim. Equations intégrales pour les polynômes d'hermite à une et plusieurs variables, pour les polynômes de laguerre, et pour les fonctions hypergéométriques les plus générales. *Ann. Scuola. Norm. Super. Pisa*, (9):225–252, 1940.
[8] U. Grenander, H. Pollak, and D. Slepian. The distribution of quadratic forms in normal variates: a small sample theory with applications to spectral analysis. *J. Soc. Indust. Appl. Math.*, (7)(4):374–401, 1959.
[9] K. Mardia. Tests of univariate and multivariate normality. *Handbook of Statistics*, (1):279–320, 1980.
[10] E. Pearson, R. D'Agostino, and K. Bowman. Tests for departure from normality: Comparison of powers. *Biometrika*, (64):231–246, 1977.
[11] S. Rice. *Mathematical analysis of random noise.* Bell System Tech. J. (24), 1945.

High-Order Properties of M-Band Wavelet Packet Decompositions

D. Leporini and J.-C. Pesquet
Laboratoire des Signaux et Systèmes
CNRS/Univ. Paris Sud and GDR-PRC ISIS
ESE, 91192 Gif sur Yvette Cédex, France
leporini@lss.supelec.fr, pesquet@lss.supelec.fr

Abstract

In many applications it is necessary to characterize the statistical properties of the wavelet/wavelet packet coefficients of a stationary random signal. This problem is typically encountered in denoising methods using wavelet packets. Then, in a stationary non-Gaussian noise scenario, it may be useful to determine the high-order statistics of the wavelet packet coefficients. In this work, we prove that this task may be performed through multidimensional filter banks. In particular, we show how the cumulants of the M-band wavelet packet coefficients of a strictly stationary signal are derived from those of the signal and we provide recursive decomposition and reconstruction formulae to compute the cumulants of these coefficients. High-order wavelet packets, associated to multidimensional filter banks, are presented along with some of their properties. Finally, the asymptotic normality of the coefficients is proved.

1. Introduction

Wavelet/wavelet packet decompositions are becoming popular in statistical applications such as modelling, detection or estimation of observed noisy processes [6, 5]. While in the classical additive model, the stationary noise is generally assumed to be Gaussian i.i.d., in some situations it becomes however necessary to depart from this hypothesis [3, 9]. In such cases, determining the high-order statistics of the wavelet packet coefficients of the process may consequently be of interest.

In this paper, we present some results concerning the statistical properties of the M-band wavelet packet coefficients of a strictly stationary random signal. In particular, we prove that the cumulants of the coefficients may be computed recursively from those of the signal, and derive the decomposition and reconstruction formulae for the cumulant field. We show that these operations may be realized through multidimensional filter banks, and present some properties of the associated high-order wavelet packets along with the connections with frame multiresolution analysis [1]. The asymptotic normality of the coefficients is finally obtained under weak conditions. We note that the proofs of the results given in this paper may be found in [7].

2. Hypotheses

We consider the general case of an M-band orthogonal wavelet packet decomposition [12, 11] corresponding to a paraunitary perfect reconstruction filter bank with impulse responses $(h_0(k))_{k\in\mathbb{Z}}, \ldots, (h_{M-1}(k))_{k\in\mathbb{Z}}$. The associated wavelet packet functions will be denoted in the sequel by $W_m(t), m \in \mathbb{N}$. We also assume that the considered process $x(t), t \in \mathbb{R}$, is strictly stationary and verifies the following condition:

$$\exists K \in \mathbb{R}^+ \mid \forall n \in \mathbb{N}^* \quad \sup_{f_1,\cdots,f_n} |\Gamma_x^{(n+1)}(f_1,\ldots,f_n)| \leq K^n,$$

where $\Gamma_x^{(n+1)}(f_1,\ldots,f_n)$ denotes the polyspectrum of order $(n+1)$ of $x(t)$. Note that this assumption implies in particular that the polyspectra should be bounded, which is related to mixing properties of the process [2]. The wavelet packet coefficients of the process $x(t)$ at resolution 2^{-j} and m-th frequency bin are given by

$$c_{j,m}(k) = \frac{1}{M^{j/2}} \int_{\mathbb{R}} x(t) W_m\Big(\frac{t}{M^j} - k\Big) dt.$$

It can be shown that this integral is convergent in the mean square sense when $x(t)$ is stationary and $W_m(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}),\ m \in \mathbb{N}$.

3. High-order wavelet packets

We first introduce the notion of high-order wavelet packets and subsequently investigate some of their properties.

0-8186-8005-9/97 $10.00 © 1997 IEEE

For $n \in \mathbb{N}^*$ and $(t_1, \ldots, t_n) \in \mathbb{R}^n$ we define the wavelet packets of order $(n+1)$ by:

$$\gamma^{(n+1)}_{W_{m_1},\ldots,W_{m_{n+1}}}(t_1, \ldots, t_n) = \int_{-\infty}^{\infty} W_{m_1}(u) W_{m_2}(u+t_1) \ldots W_{m_{n+1}}(u+t_n)\, du,$$

and point out that these functions can be viewed as extensions of the autocorrelation wavelet functions introduced by Saito in [10]. An example of high-order scaling function of order 3 is shown in Fig. 1. Throughout the paper, we will also use the notation:

$$\Omega^{(n+1)}_{j,(m_1,\ldots,m_{n+1})} = \overline{\text{Span}}\{\gamma^{(n+1)}_{W_{m_1},\ldots,W_{m_{n+1}}}(\frac{t_1}{M^j} - k_1, \ldots, \frac{t_n}{M^j} - k_n), (k_1, \ldots, k_n) \in \mathbb{Z}^n\}.$$

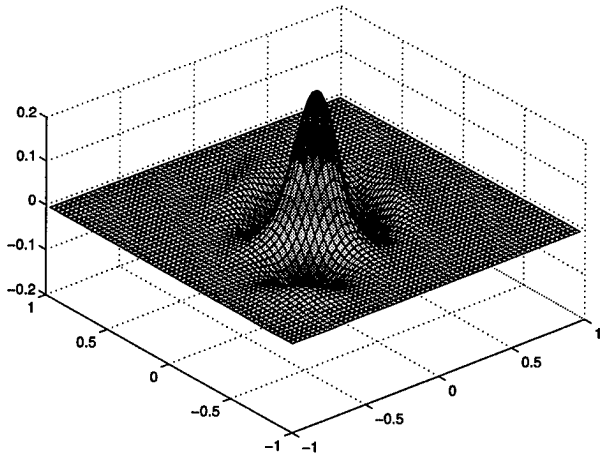


Figure 1. High-order scaling function of order 3 based on Daubechies 8.

Recall that $W_0(t)$ corresponds to the scaling function usually denoted by $\phi(t)$ and associated to a multiresolution analysis (MRA) $((V_j)_{j\in\mathbb{Z}}, \phi)$ of $L^2(\mathbb{R})$ [8]. For the sake of clarity, we first recall the basic definitions concerning frames and frame multiresolution analysis (FMRA).[1]

Definition 1 *Let $\mathcal{H}$ be a complex separable Hilbert space. A sequence $\{g_n,\ n \in \mathbb{Z}\} \subseteq \mathcal{H}$ is a frame of $\mathcal{H}$ if there exist two constants $A,\ B > 0$ such that*

$$\forall f \in \mathcal{H},\ A||f||^2 \leq \sum_{n\in\mathbb{Z}} |\langle f, g_n\rangle|^2 \leq B||f||^2,$$

where $A,\ B$ are the frame bounds, $\langle\cdot,\cdot\rangle$ denotes the inner product on $\mathcal{H}$, and $||f|| = \langle f, f\rangle^{\frac{1}{2}}$ stands for the norm of $f \in \mathcal{H}$. A frame $\{g_n,\ n \in \mathbb{Z}\}$ is tight if $A = B$. It is exact iff $\{g_n,\ n \in \mathbb{Z}\}$ is a Riesz basis.

[1] For general discussions on these concepts, we refer the reader to [4, 1].

Definition 2 *A FMRA $((\Omega_j)_{j\in\mathbb{Z}}, \gamma)$ of $L^2(\mathbb{R})$ is defined by a sequence of nested closed linear subspaces $\Omega_j \subseteq L^2(\mathbb{R})$ and an element $\gamma \in \Omega_0$ verifying:*

- $\overline{\cup_{j\in\mathbb{Z}}\Omega_j} = L^2(\mathbb{R})$ *and* $\cap_{j\in\mathbb{Z}}\Omega_j = \{0\}$,
- $f(t) \in \Omega_{j+1} \Longleftrightarrow f(2t) \in \Omega_j$,
- $f(t) \in \Omega_0 \Longrightarrow f(t-k) \in \Omega_0$ *for all* $k \in \mathbb{Z}$.
- $\{\gamma(t-k),\ k \in \mathbb{Z}\}$ *is a frame of* Ω_0.

A first interesting result is the following:

Proposition 1 *Under some weak assumptions, the autocorrelation function $\gamma^{(2)}_{W_0,W_0}(t)$ defines a FMRA $((\Omega^{(2)}_{j,(0,0)})_{j\in\mathbb{Z}}, \gamma^{(2)}_{W_0,W_0})$ of $L^2(\mathbb{R})$.*

As an example in the 2-band case, the B-spline function of order 2 is obtained by considering the Haar scaling function $W_0(t) = \mathbf{1}_{[0,1]}(t)$. In the general case ($n > 1$), the high-order scaling functions $\gamma^{(n+1)}_{W_0,\ldots,W_0}(t_1, \ldots, t_n)$ do not necessarily lead to a FMRA $((\Omega^{(n+1)}_{j,(0,\ldots,0)})_{j\in\mathbb{Z}}, \gamma^{(n+1)}_{W_0,\ldots,W_0})$ of $L^2(\mathbb{R}^n)$, depending on the involved scaling function $W_0(t)$. For instance, we remark that the high-order Shannon scaling function and its integer translates constitute a frame of $\Omega^{(n+1)}_{0,(0,\ldots,0)}$ and therefore lead to a FMRA of $L^2(\mathbb{R}^n)$ for all $n \in \mathbb{N}^*$, in contrast with the scaling functions corresponding to Meyer filters.
However, the following nested non-orthogonal subspace properties are interestingly obtained:

Proposition 2 *For all $n \in \mathbb{N}^*$, we have:*

$$\Omega^{(n+1)}_{j,(m,\ldots,m)} = \sum_{p_1,\ldots,p_{n+1}} \Omega^{(n+1)}_{j+1,(Mm+p_1,\ldots,Mm+p_{n+1})},$$

with $(p_1, \ldots, p_{n+1}) \in \{0, \ldots, M-1\}^{(n+1)}$.

We can consequently show that the high-order wavelet packets satisfy multidimensional two-scale equations given by:

$$\frac{1}{M^{(n-1)/2}}\gamma^{(n+1)}_{W_{Mm+p_1},\ldots,W_{Mm+p_{n+1}}}(\frac{t_1}{M}, \ldots, \frac{t_n}{M}) = \sum_{k_1,\ldots,k_n} \gamma^{(n+1)}_{h_{p_1},\ldots,h_{p_{n+1}}}(k_1, \ldots, k_n)\, \gamma^{(n+1)}_{W_m,\ldots,W_m}(t_1 - k_1, \ldots, t_n - k_n),$$

where

$$\gamma^{(n+1)}_{h_{p_1},\ldots,h_{p_{n+1}}}(k_1, \ldots, k_n) = \sum_n h_{p_1}(n) h_{p_2}(n+k_1) \ldots h_{p_{n+1}}(n+k_n)$$

defines the $(n+1)^{th}$ order deterministic cross-correlation sequence of the filter bank.
The high-order wavelet packets also satisfy the following reconstruction formulae:

$$\gamma^{(n+1)}_{W_m,\dots,W_m}(t_1-k_1,\dots,t_n-k_n)=\frac{1}{M^{(n+1)/2}}$$
$$\sum_{p_1,\dots,p_{n+1}}\sum_{l_1,\dots,l_n}\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k_1-Ml_1,\dots,k_n-Ml_n)$$
$$\gamma^{(n+1)}_{W_{Mm+p_1},\dots,W_{Mm+p_{n+1}}}(\frac{t_1}{M}-l_1,\dots,\frac{t_n}{M}-l_n).$$

We now present one of our main results:

Proposition 3 *The multidimensional filter bank defined by the impulse responses* $\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k_1,\dots,k_n)$ *of the n-dimensional filters satisfies the following property:*

$$\sum_{p_1,\dots,p_{n+1}}\sum_{l_1,\dots,l_n}\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k_1-Ml_1,\dots,k_n-Ml_n)$$
$$\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k'_1-Ml_1,\dots,k'_n-Ml_n)=$$
$$M\delta(k_1-k'_1)\dots\delta(k_n-k'_n)\,.$$

Although this property only comes from the orthonormality of the wavelet packet bases, it is interesting to note that this result also appears as a consequence of the existence of a dual frame $\tilde{\gamma}^{(n+1)}_{W_0,\dots,W_0}(t_1,\dots,t_n)$ of order $(n+1)$ in the case where $((\Omega^{(n+1)}_{j,(0,\dots,0)})_{j\in\mathbb{Z}},\gamma^{(n+1)}_{W_0,\dots,W_0})$ is an exact FMRA.
Moreover, we deduce that the sequence

$$\{\gamma^{(n+1)}_{W_0,\dots,W_0}(t_1-k_1,\dots,t_n-k_n),$$
$$(k_1,\dots,k_n)\in\mathbb{Z}^n\}$$

is a frame of $\Omega^{(n+1)}_{0,(0,\dots,0)}$ with frame bounds A and B iff the sequence

$$\{\gamma^{(n+1)}_{W_{p_1},\dots,W_{p_{n+1}}}(\frac{t_1}{M}-k_1,\dots,\frac{t_n}{M}-k_n),$$
$$(k_1,\dots,k_n)\in\mathbb{Z}^n,(p_1,\dots,p_{n+1})\in\{0,\dots,M-1\}^{n+1}\}$$

is a frame of $\Omega^{(n+1)}_{0,(0,\dots,0)}$ with frame bounds AM^n and BM^n.
As a consequence of Proposition 3, we however see that for all $n\in\mathbb{N}^*$, the multidimensional sequence $\{\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k_1-Ml_1,\dots,k_n-Ml_n),\ (l_1,\dots,l_n)\in\mathbb{Z}^n,\ (p_1,\dots,p_n)\in\{0,\dots,M-1\}^n\}_{(k_1,\dots,k_n)\in\mathbb{Z}^n}$ is a tight frame of $\ell^2(\mathbb{Z}^n)$ with frame bounds $A=B=M$.
Recursive computations of the cumulants through the previously exhibited structure are presented in the next section along with some related properties.

4. Cumulant field analysis

We first recall that the wavelet packet coefficients $(c_{j,m}(k))_{k\in\mathbb{Z}}$ corresponds to stationary and cross-stationary processes whose cumulants are given by the following property:

Proposition 4 *For all* $n\in\mathbb{N}^*$ *and* $(m_1,\dots,m_{n+1})\in\mathbb{N}^{n+1}$, *the cross-cumulants of order* $(n+1)$ *of the wavelet packet coefficients* $(c_{j,m_1}(k))_{k\in\mathbb{Z}},\dots,(c_{j,m_{n+1}}(k))_{k\in\mathbb{Z}}$ *are obtained as the n-dimensional inner products of the cumulants of order* $(n+1)$ *of the analyzed signal with dilated/translated versions of the high-order wavelet packets* $\gamma^{(n+1)}_{W_{m_1},\dots,W_{m_{n+1}}}(t_1,\dots,t_n)$.

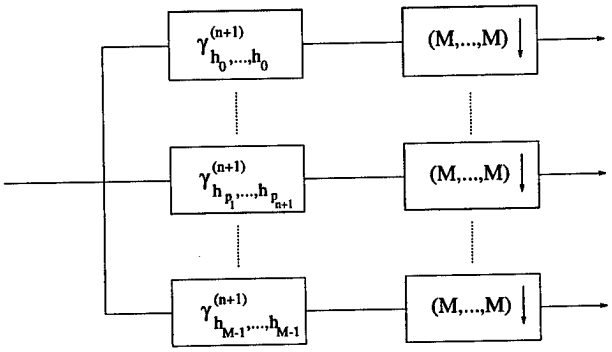


Figure 2. Multidimensional analysis filter bank.

From the two-scale equation presented in the previous section, a scale-recursive algorithm involving the multidimensional structure exhibited in the previous section may be proposed to efficiently compute the cumulants:

Proposition 5 *The cumulants of the wavelet packet coefficients* $(c_{j+1,m}(k))_{k\in\mathbb{Z}}$ *at resolution level* $(j+1)$ *are obtained from those of the coefficients* $(c_{j,m}(k))_{k\in\mathbb{Z}}$ *through the multidimensional filter bank involving filters with impulse responses* $\gamma^{(n+1)}_{h_{p_1},\dots,h_{p_{n+1}}}(k_1,\dots,k_n)$ *and decimators by a factor M in each of the n dimensions (see Fig. 2).*

This decomposition property actually corresponds to a multidimensional multiresolution analysis of the cumulant field. From the results in Section 3, we easily find a dual multidimensional reconstruction filter bank to the one considered in Proposition 5 (see Fig. 3).

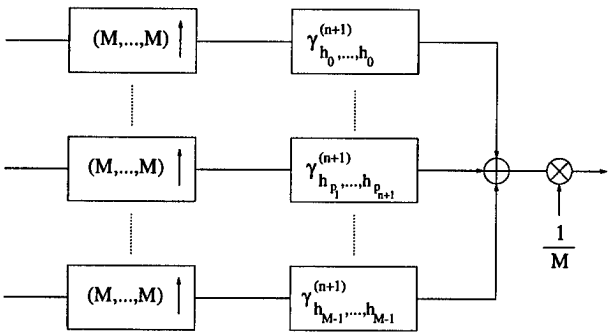


Figure 3. Multidimensional synthesis filter bank.

It can also be shown that the cumulants satisfy the following quadratic norm conservation property

$$\sum_{p_1,\cdots,p_{n+1}} \sum_{l_1,\ldots,l_n} \mathrm{cum}^2[c_{j+1,Mm+p_1}(l), c_{j+1,Mm+p_2}(l+l_1),\ldots,c_{j+1,Mm+p_{n+1}}(l+l_n)] = M \sum_{k_1,\ldots,k_n} \mathrm{cum}^2[c_{j,m}(k),\ldots,c_{j,m}(k+k_n)].$$

In the next section, we apply the previous properties to obtain an asymptotic convergence result for the wavelet packet coefficients.

5. Convergence in distribution

By using Proposition 5 in the frequency domain, we finally show that the statistics of order greater than 2 of the wavelet packet coefficients exponentially decay with respect to the resolution level:

Proposition 6 *The cumulants of order* $(n+1), n \geq 2$, *of the wavelet packet coefficients at resolution level* j *are upper-bounded as follows:*

$$\exists C < 1 \mid \forall (k_1,\ldots,k_n) \in \mathbb{Z}^n, \forall m \in \{0,\ldots,M^j-1\}, |\mathrm{cum}[c_{j,m}(k), c_{j,m}(k+k_1),\ldots,c_{j,m}(k+k_n)]| \leq C^{\lfloor j/2 \rfloor} K^n,$$

where C *is independent of* n *and* $\lfloor \cdot \rfloor$ *denotes the greatest integer lower than or equal to its argument.*

The rate of convergence C is related to the frequency responses $H_i(f)$, $i \in \{0,\ldots,M-1\}$ of the filter bank and may be analytically derived for different families of wavelet packets (*e.g.* $C = \frac{1}{\sqrt{M}}$ for Walsh-Hadamard filters). As a simple consequence of this result, we check that the processes $(c_{j,m}(k))_{k\in\mathbb{Z}}$ converge in distribution to Gaussian processes when j tends to infinity. It should be noted that, in the dyadic wavelet decomposition case, this latter fact was also observed by other authors [3, 9] under different assumptions.

6. Conclusion

In this paper, we have presented some tools to determine the high-order statistics of the M-band wavelet packet coefficients of a strictly stationary process. Muldimensional multiresolution analysis of the cumulant field is performed through an n-dimensional filter bank whose properties have been investigated. The asymptotic normality of the wavelet packet coefficients has finally been presented.

References

[1] J. Benedetto and S. Li. The Theory of Multiresolution Analysis Frames and Applications to Filter Banks. Technical Report, Dept. of Mathematics, University of Maryland, 1996.

[2] D. R. Brillinger. *Time Series:Data Analysis and Theory.* Holt, New York, 1975.

[3] D. R. Brillinger. Uses of cumulants in wavelet analysis. *Proc. SPIE, Advanced Signal Processing*, 2296:2–18, 1994.

[4] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Informat. Theory*, IT-36:961–1005, Sep. 1990.

[5] D. L. Donoho and I. M. Johnstone. Ideal denoising in an orthogonal basis chosen from a library of bases. *C. R. Acad. Sci. Paris*, 319:1317–1322, 1994.

[6] H. Krim and J.-C. Pesquet. On the statistics of best bases criteria. In A. Antoniadis, editor, *Wavelets and statistics*. Lecture Notes in Statistics, Springer Verlag, 1995.

[7] D. Leporini and J.-C. Pesquet. High-order Wavelet Packets and Cumulant Field Analysis. Internal Report, Laboratoire des Signaux et Systèmes, France, 1997.

[8] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, PAMI-11:674–693, Jul. 1989.

[9] M. H. Neumann and R. von Sachs. Wavelet Thresholding: Beyond the Gaussian I.I.D. Situation. In A. Antoniadis, editor, *Wavelets and statistics*. Lecture Notes in Statistics, Springer Verlag, 1995.

[10] N. Saito. *Local feature extraction and its applications using a library of bases.* PhD thesis, Yale University, Dec. 1994.

[11] P. Stephen, P. N. Heller, R. A. Gospinath, and G. B. Burrus. Theory of regular M-band wavelet bases. *IEEE Trans. Signal Processing*, SP-41:3497–3511, Dec. 1993.

[12] M. V. Wickerhauser. INRIA lectures on wavelet packet algorithms. In *Ondelettes et paquets d'ondelettes*, pages 31–99, Roquencourt, France, Jun. 17-21 1991.

BUSSGANG GAUSSIANITY TEST FOR STATIONARY SERIES

G. Giunta, G. Jacovitti, G. Scarano

Dip. INFOCOM, Università di Roma "La Sapienza" via Eudossiana 18, I-00184 Roma, Italy

E-mail: gaetano@infocom.ing.uniroma1.it

Abstract

In this contribution we develop a procedure for deciding whether a finite segment of a signal can be considered as a realization of a Gaussian process or not. We first present the theoretical bases conducting to the formulation of Bussgang test. A novel test based on the sign non-linearity is introduced and its performance analysed in comparison with two simple Gaussianity tests presented in the literature. The results have shown a very promising behaviour of the suggested method in the presence of small amount of available data in both the cases of white and correlated samples.

I Introduction

Deciding whether a finite segment of a signal can be considered as a realization of a Gaussian process or not is an inferential problem useful in several applications of digital signal processing. This decision is preliminary to other activities of signal processing, to recognize the existence of statistical information recoverable by higher order statistics, or to detect the existence of useful signals in measurements affected by Gaussian noise.

Some classical approaches to this problem are based on frequency domain tests, which in general require large sample sets. In order to improve the detectability for relatively short signal segments, recent contributions have been focused on time domain tests. Basically, these techniques consists of measuring the distance between the expected values (calculated for the Gaussian hypothesis) and the sample averages after some non–linear transformation of the series. It is known that a quadratic form built on the sample deviations from the expected value is asymptotically chi-square distributed under the Gaussian hypothesis. This allows to determine in principle the decision treshold to apply to the value returned from the quadratic form for obtaining a wanted significance level (probability of Gaussian hypothesis rejection for true Gaussian series).

The most simple versions of these tests are aimed to verify the Gaussianity of the marginal distribution of the samples. Of course, specific non–linearities are able to detect specific deviations from Gaussianity, For instance, the third power reveals unconsistent skewness values, whereas the fourth power reveals anomalies of the kurtosis. Non–Gaussian distributions having skewness and kurtosis close to the values pertaining to the Gaussian case are undetectable by these tests.

Likewise, the (complex) exponential non–linearities employed in the characteristic function oriented tests [1, 2] may be unable to detect some other specific non–Gaussian behaviors. In order to deal with as many cases as possible in practical situations where the nature of the measured samples is totally unkown, composite tests based on sets of non–linearities have been proposed and characterized in the recent literature. The higher–order moment approach is based on the joint use of multiple moments [3], whereas the empirical characteristic function approach employs different values of the parameter in the exponential. In some cases, such as linear filtered signals, the non–Gaussian nature is difficult to detect from the analysis of the marginal distribution. For this reason, multivariate detectors have been proposed. A unified theory of time domain Gaussianity tests based on finite memory non–linearities has been very recently exposed in [4]. It is based on Price theorem, which relates the moments of Gaussian n–variates to the moments of given functions of these variates. This approach enligths how tests might be designed in general. Specifically, higher order moment based tests and empirical characteristic function based tests are derived for the n-variate case. In this contribution, we propose an alternative approach based on Bussgang property.

II Bussgang–based Method

As well known, Bussgang theorem states that the cross–correlation function of a Gaussian stationary process and of its version passed trough a zero-memory non–linearity is proportional to the auto–correlation function of the process, namely

$$\mathrm{E}\{x[n+k]\cdot g(x[n])\} = k_g\cdot \mathrm{E}\{x[n+k]\cdot x[n]\} \tag{1}$$

The proportionality factor k_g depends on the non–linearity $g(\cdot)$ and it can be expressed as [5]

$$k_g = \mathrm{E}\{\dot{g}(x[n])\} \tag{2}$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

This property has been extended to complex process in [6, 7] and generalized to the multivariate case in [8].

Even though Bussgang theorem is implied by Price theorem, it presents peculiar aspects which makes it particularly attractive for various applications of signal processing. For instance, it constitutes the principle underlying some well known techniques of blind deconvolution employed in data communication systems and in geophysics.

The Gaussianity test based on Bussgang theorem simply consists of measuring the deviation from proportionality of the sample auto–correlation to the sample cross–correlation for a given non–linearity, with the theoretical proportionality factor. In particular, we first estimate, according to the mean square error criterion, the proportionality factor k_g from the overdetermined set of equations:

$$R_{xg}(k) = k_g \cdot R_{xx}(k) \tag{3}$$

with $k = 0..N$, where the auto–correlation $R_{xx}(k)$ and the cross–correlation $R_{xg}(k)$ functions are estimated for a given number $N+1$ of lags from the available set of data. In the subsequent step, we wonder whether the estimated proportionality factor is *significantly* similar to the theoretical factor k_g^G derived under the Gaussian assumption. This is accomplished by observing the sample distribution of the weighed (by the sample variance $R_{xx}(0)$) non–Gaussianity error, *i.e.*:

$$w\left[R_{xx}(0)\right] \cdot \left(k_g - k_g^G\right)$$

This makes the design of a test simple and meningful; in fact, one single nonlinerity must be chosen. On the other hand, the test is multivariate by nature. Its dimensionality depends on how many correlation lags are included in the test. It is easily verified that this approach generates well defined higher order moment–based tests. Moreover, it allows to extend the empirical function based test to the multivariate case without need of introducing multiple parameters, as in previous approaches.

III Some Application Examples

In this contribution, we have first presented the theoretical bases conducting to the formulation of Bussgang test. Then, our attention should be focused on the choice of the non–linearity. In this respect, it should be noted that no rational criteria are currently available about the choice of a specific test for the application at hand, apart obvious considerations such that no odd powers should be employed for symmetrical distributions. Often, more knowledge is a priori available about possible deviations from Gaussianity. For instance, this is evident for the case where one is interested to detect the presence of signals in noise, or to reveal if Gaussian sources are affected by some kind of distortions. In those cases, we would require to choice among different tests the one giving the best performance at the same computational cost or for the same sample size.

Thus, in order to gain insigth into the test characteristics, we compare the performance of different tests versus a class of input distributions indexed by continuous parameters. For simplicity, we first refer to the univariate case and to a generalized Gaussian distribution, defined by:

$$p_X(x) = \frac{1}{2 \cdot b \cdot \Gamma\left(\frac{1}{a}\right)} \exp\left(\left|\frac{x}{b}\right|^a\right) \tag{4}$$

With this model, we can trace the behavior of different tests for even small deviations from Gaussianity, thus characterizing their sensitivity with respect to the distribution parameter.

A further example of application has been numerically performed by considering an additive mixture of non-Gaussian binary signal and independent Gaussian noise realizations, such as it happens in a number of data communication links, *i.e.*

$$p_X(x) = \frac{1}{2}\left(\mathcal{G}(m,\sigma^2) + \mathcal{G}(-m,\sigma^2)\right) \tag{5}$$

where $\mathcal{G}(m,\sigma^2)$ is the normal distribution of mean m and variance σ^2 . The test has been carried out for several values of Signal-to-Noise ratios ($\mathrm{SNR} = m^2/\sigma^2$) in order to assess the validity of the Gaussian test even under low–level signal conditions.

IV Numerical Results

In particular, we have theoretically evaluated the significance level (i.e. the probability of rejecting the Gaussian hyphotesis when the signal is actually Gaussian) and the power of the various tests (i.e. the probability of rejecting the Gaussian hyphotesis when the signal is not actually Gaussian).

We have carried computer simulations to assess the analytical results. For sake of simplicity, we have considered univariate error function and the sample variance as the error weighing function, *i.e.*

$$w\left[R_{xx}(0)\right] = R_{xx}(0) \tag{6}$$

in order to meet the reference case widely encountered in the literature. We have considered both white and first order AR-filtered ($\rho = 0.5$) signals of several lengths of samples; we have then collected the percentage of the rejection of the Gaussian hyphotesis, by fixing all the employed thresholds in order to use a significance level of 5% for each test. Different generalised Gaussian distributions have been considered by varying the parameter a, which

characterizes both the super– and sub–Gaussian behavior of the (symmetrical) distribution. Some of the obtained results are shown in Figs.1 and 2 for normalized (unitary variance) distributions, *i.e.*

$$b = \sqrt{\Gamma\left(\frac{1}{a}\right)/\Gamma\left(\frac{3}{a}\right)} \qquad (7)$$

The results show that for sub–Gaussian distributions ($a>2$) Koutruvelis–Epps test outperforms Giannakis–Tsatsanis test, while their merits are comparable for super–Gaussian distributions ($a<2$). This provides a limited, but significant choice criterion in many applications. Searching for a different test inspired to Bussgang paradigm, we have then considered a simple test based on a nonlinearity constituted by the *signum* function. Interesting enough, the Bussgang–signum test exhibits the best performance for sub–Gaussian distributions. This is particularly significant in the presence of low–pass filtered non–Gaussian signals and for small sample size.

Such a result is not surprising, because it indirectly relates to known circumstances. In fact, the signum distortion is successfully employed for phase retrieval in blind deconvolution of communication channels where data exhibit a strong sub–Gaussian behavior; conversely, the power nonlinearity is employed for blind deconvolution of seismic signals, which are super–Gaussian [9].

The results of the simulations performed for the case of a binary signals embedded in uncorrelated Gaussian noise (see the Figs.3 and 4) have confirmed the promising behaviour of the signum non–linearity. Once again, the signum–based test yields the best performance for both white and correlated data. Conversely, Giannakis–Tsatsanis test appears not suitable for such purposes, probably because the signum–based test statistics result in better assorted histograms w.r.t. higher–order statistics, which are affected by side events.

V Conclusion

A procedure for deciding whether a finite segment of a signal can be considered as a realization of a Gaussian process or not has been developed. The theoretical bases conducting to the formulation of Bussgang test have been presented. A novel test based on the sign non-linearity has been introduced and its performance analysed in comparison with two simple Gaussianity tests presented in the literature.

The results have shown a very promising behaviour of the suggested method in the presence of small amounts of both white and correlated available data. Future investigations will be performed to confirm such trend in the case of multiple lags.

References

[1] T. Epps, "Testing that a stationary time–series is Gaussian", *Ann. Stat.*, vol.15, no.4, 1987.

[2] I. Koutrouvelis, "A goodness of fit test based on the empirical characteristic function when parameters must be estimated", *Biometrika*, vol.67, 1990.

[3] G.B. Giannakis, M.K. Tsatsanis, "Time–domain tests for Gaussianity and time–reversibility", *IEEE Transactions on Signal Processing*, vol.42, no.12, December 1994.

[4] E. Moulines, K. Choukri, " Time–domain procedures for testing that a stationary series is Gaussian", *IEEE Transactions on Signal Processing*, vol.44, no.8, August 1996.

[5] J.J. Bussgang, "Cross-correlation function of amplitude–distorted Gaussian signals", *Res. Lab. Elec., Mas. Inst. Technol.*, Cambridge MA, Tech. Rep. 216, March 1952.

[6] G. Jacovitti, A. Neri, R. Cusani, "Methods for estimating the autocorrelation function of complex Gaussian processes", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.35, no.8, August 1987.

[7] G. Scarano, "Cumulant series expansion of hybrid non-linear moment of complex random variables", *IEEE Transactions on Signal Processing*, vol.39, no.4, April 1991.

[8] G. Scarano, D. Caggiati, G. Jacovitti, "Cumulant series expansion of hybrid non-linear moment of n variates", *IEEE Transactions on Signal Processing*, vol.41, no.1, January 1993.

[9] Kostov, F. Rocca, "Estimation of residual wavelets", in *Deconvolution and Inversion*, M.H. Worthington ed., Blackwell Scientific Publications, London, U.K., 1987.

[10] L.M. Garth, H.V. Poor, "Detection of non–Gaussian signals: a paradigm for modern statistical signal processing", *Proceedings of IEEE*, vol. 82, no. 7, July 1994.

[11] A.M. Zoubir, M.J. Arnold, "Testing Gaussianity with the characteristic function: the i.i.d. case", *Signal Processing*, vol. 53, September 1996.

[12] P. Comon, L.Deruaz, "Normality tests for coloured samples" *IEEE Athos Workshop on Higher Order Statistics*, Beur, Spain, June 1995.

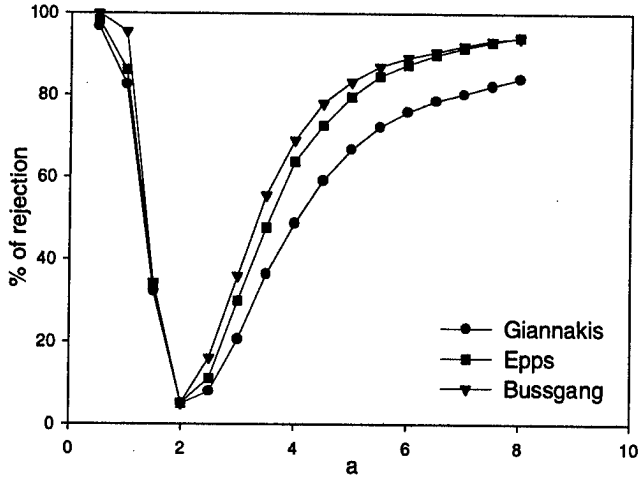


Figure 1: *Percentage of rejection of the Gaussian hyphotesis vs. the non–Gaussianity parameter a of the generalised Gaussian distribution for 128 white samples.*

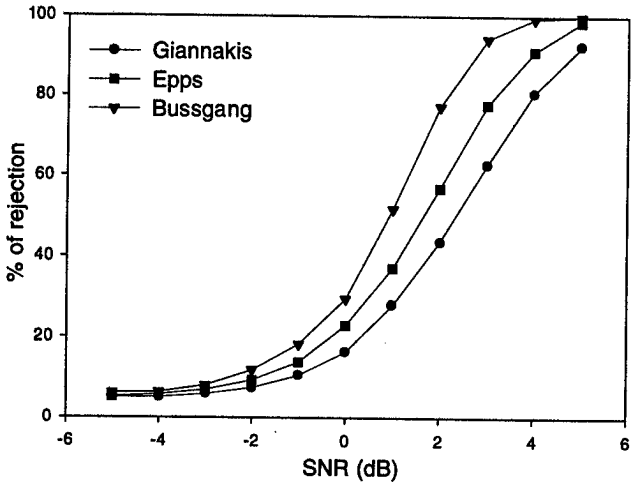


Figure 3: *Percentage of rejection of the Gaussian hyphotesis vs. the SNR of the additive mixture of non–Gaussian binary signal and inpedendent Gaussian noise for 128 white samples.*

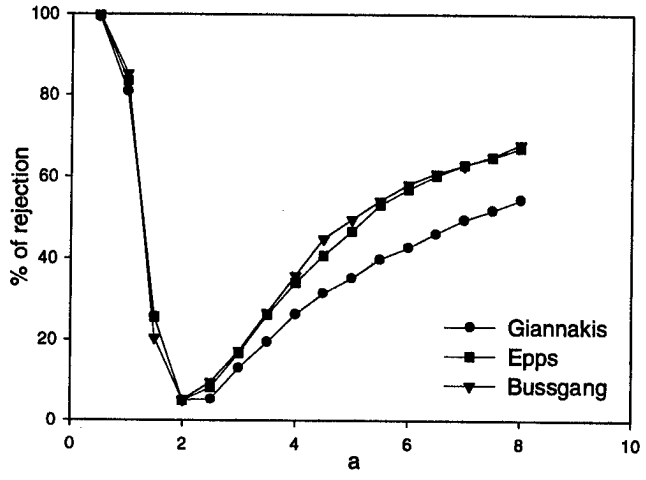


Figure 2: *Percentage of rejection of the Gaussian hyphotesis vs. the non–Gaussianity parameter a of the generalised Gaussian distribution for 256 correlated $\rho = 0.5$ samples.*

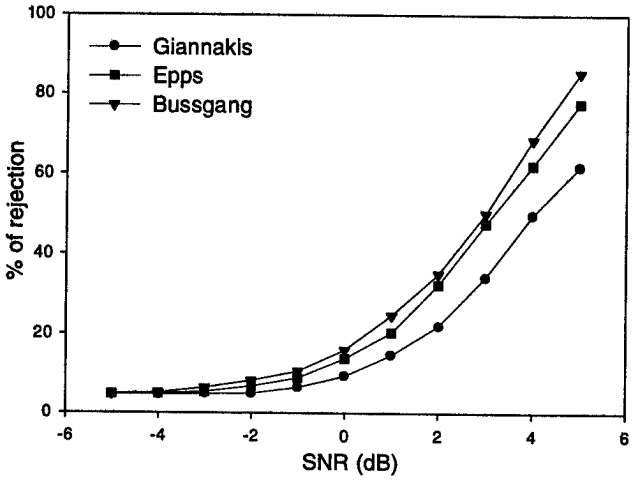


Figure 4: *Percentage of rejection of the Gaussian hyphotesis vs. the SNR of the additive mixture of non–Gaussian binary signal and inpedendent Gaussian noise for 256 correlated $\rho = 0.5$ samples.*

Testing Multivariate Gaussianity with the Characteristic Function

A. M. Zoubir, C. L. Brown and B. Boashash
Signal Processing Research Centre
Queensland University of Technology
GPO Box 2434, Brisbane, Q. 4001, Australia
a.zoubir@qut.edu.au

Abstract

A modification to a previously developed characteristic function based Gaussianity test is proposed. The power of the test is consequently improved. This test is then extended to the multivariate case, allowing it to be applied to correlated data. Monte Carlo simulations are performed to compare power with two other tests for multivariate Gaussianity, with encouraging results.

1. Introduction

The assumption of Gaussianity is important to verify in many areas. Tests for Gaussianity, using the characteristic function (cf), have been developed with promising results [3, 5, 8]. However, these tests perform well (achieve high power) for independent and identically distributed (iid) data only. This limits their applicability to real situations where the data is often correlated. It is therefore necessary to address the issue of testing correlated data for Gaussianity using the cf. This paper formulates tests for Gaussianity of correlated data and shall be refer to them as *Multivariate Gaussianity Tests.*

Other tests for multivariate Gaussianity have generally lagged similar univariate tests [6] due to an increase in complexity and have required large numbers of samples (in the thousands) [2]. In this paper we show that the Gaussianity tests based on the cf are readily extendable to the multivariate case and achieve high power, while not being limited to large samples.

2. Characteristic function based tests for Gaussianity

It has already been shown [3, 4] that the cf can be used as the basis of Gaussianity testing. In [8], problems related to the use of the empirical characteristic function (ecf), for such tests was highlighted. The ecf is an unbiased estimator of the cf of a process based on a finite number of samples, N, however it has unacceptably high variance, see Figure 1. In fact, its variance approaches a constant $(1/2N)$ as t approaches infinity [1].

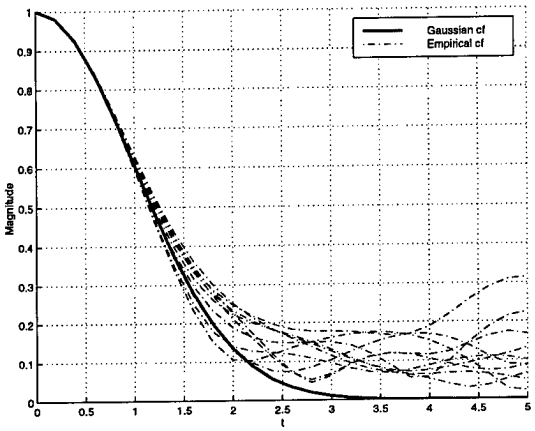


Figure 1. Magnitudes of the ecfs of 10 realisations of a standard Gaussian random variable, with the true Gaussian cf.

Instead, it was proposed that the kernel characteristic function estimator (KCFE) be used instead of the ecf. The development of the KCFE was motivated by analogy with the theory of kernel density estimation [7]. Briefly, the KCFE is produced by multiplying an initial estimate of the characteristic function of a process, specifically the ecf, by a kernel. The kernel used in [8] was Gaussian. This smoothes out variations in the large t region of the estimate, see Figure 2, at the expense of the introduction of a bias.

0-8186-8005-9/97 $10.00 © 1997 IEEE

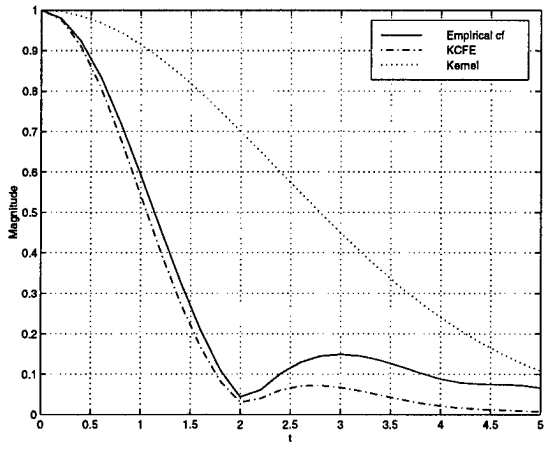


Figure 2. The effect of multiplicative smoothing on the ecf of 1 realisation of a Gaussian process.

3. A new approach to overcoming non-vanishing variance of the estimator

In this paper, we consider another method of overcoming the problem of the non-vanishing variance of the ecf. This approach does not attempt to find a better estimate of the cf, rather, it recognises that a better measure of the *difference* between the data and the distribution under the null hypothesis is required. We consider smoothing this difference, and not smoothing the estimate as is done for the KCFE method.

Let $\boldsymbol{X} = [X_0, \ldots, X_{N-1}]$ be a vector of N observations. Consider a function $\hat{\varepsilon}_{\boldsymbol{X}}(t) = \hat{\phi}^e_{\boldsymbol{X}}(t) - \phi_{\boldsymbol{X}}(t)$ where

$$\hat{\phi}^e_{\boldsymbol{X}}(t) = \frac{1}{N}\sum_{n=0}^{N-1} e^{jtX_n}$$

is the ecf and $\phi_{\boldsymbol{X}}(t)$ is the cf under H_0. From [1],

$$\begin{aligned}
\mathrm{E}[\varepsilon_{\boldsymbol{X}}(t)] &= 0 \\
\mathrm{var}\,[\Re\varepsilon_{\boldsymbol{X}}(t)] &= \frac{1}{N}\left[\frac{1}{2}\Big(1 + \Re\phi_{\boldsymbol{X}}(2t)\Big) - \Re^2\phi_{\boldsymbol{X}}(t)\right] \\
\mathrm{var}\,[\Im\varepsilon_{\boldsymbol{X}}(t)] &= \frac{1}{2N}\Big(1 - \Re\phi_{\boldsymbol{X}}(2t)\Big)
\end{aligned}$$

Note that the variance of $\varepsilon_{\boldsymbol{X}}(t)$ is the same as that of the ecf.

If this function is now subjected to a multiplicative smoothing operation by $\varphi_{\boldsymbol{X}}(t)$, a cf domain kernel function (ckf) similar to that used in the KCFE,

$$\hat{\varepsilon}_{\boldsymbol{X}}(t;\varphi_{\boldsymbol{X}}) = \hat{\varepsilon}_{\boldsymbol{X}}(t)\varphi_{\boldsymbol{X}}(t)$$

its variance will approach zero as t increases without introducing a bias, see Figure 3.

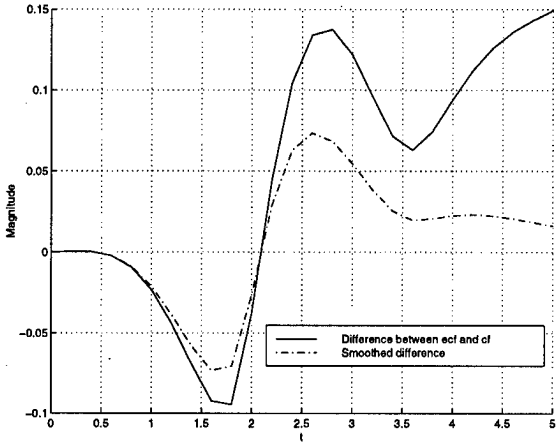


Figure 3. The effect of multiplicative smoothing on the difference between the ecf and the cf under H_0 for 1 realisation of a Gaussian process.

Figure 4 shows that the smoothed difference function has lower mean square errors (MSE) than both the ecf (due to reduction in variance) and the KCFE (due to the removal of bias).

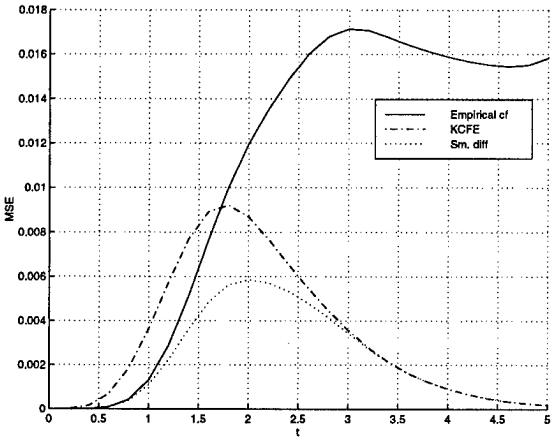


Figure 4. The Mean Square Errors of the ecf, KCFE and smoothed difference functions from 25 realisations of a Gaussian process.

The results of 500 Monte Carlo simulations, compar-

ing the performance of the KCFE and smoothed difference methods are shown in Table 1. They show that improvements have been made for most distributions tested, with the exception of the Uniform(0,1) and $\beta(4,4)$ cases.

Possible reasons for this are that these distributions are the only ones tested whose cfs are actually less than the Gaussian cf at any value of t. Consequently, the difference between their cfs and the Gaussian cf oscillates above and below zero, as opposed to the other distributions which were generally above the Gaussian cf. This makes it more difficult for the test statistics used in this paper to distinguish it from the Gaussian case.

Distribution[1]	KCFE	Sm. diff.
$U(0,1)$	95.0	82.4
χ^2_2	99.6	99.8
χ^2_8	59.6	66.8
L	29.2	65.0
$K(1,1)$	85.0	90.0
$LN(0,1)$	100.0	100.0
$\beta(4,4)$	14.0	10.2
Average	68.9	73.5

Table 1. Power of the KCFE and smoothed difference cf based tests.

4. The multivariate case

Consider a vector valued random variable

$$\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_M]^T$$

where $\boldsymbol{X_m} = [X_{m,0}, X_{m,1}, \ldots, X_{m,N-1}]$ is the vector of N observations on the mth variable for $m = 1, 2, \ldots, M$. We assume correlation exists between the variables $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_M$. The multivariate ecf may be given by

$$\hat{\phi}^e_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M) = \frac{1}{N} \sum_{n=0}^{N-1} \exp\left(j \sum_{m=1}^{M} t_m X_{m,n}\right)$$

The smoothed difference between the multidimensional ecf and the cf under H_0 is

$$\hat{\varepsilon}_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M; \varphi_{\boldsymbol{X}}) = \hat{\varepsilon}_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M)\varphi_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M)$$

[1]Some of the non-Gaussian distributions considered are U: Uniform, L: Laplace, K: K-distribution, LN: Log-Normal

where

$$\hat{\varepsilon}_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M) = \hat{\phi}^e_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M) - \phi_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M)$$

and $\varphi_{\boldsymbol{X}}(t_1, t_2, \ldots, t_M)$ is a multidimensional ckf.

4.1. Test statistic

Finding a meaningful single measure of the difference between the estimated cf and cf under H_0 is very important to the performance of this test. The cf of every distribution is unique, and therefore the difference between it and the Gaussian cf will also be unique. The chosen statistic, along with the chosen set of t values at which to evaluate the statistic, must take into account a wide range of these distinct cfs in order to produce an *omnibus* Gaussianity test – that is, a test that is powerful against *all* non-Gaussian alternatives.

For the purposes of this paper, a number of different test statistics were investigated; such as using the maximum and mean operators on the difference between the magnitude of the ecf and the Gaussian cf, the absolute difference and the squares of the difference. It was found that the absolute value of the mean of the difference provided the most powerful test for the conditions used. This statistic, which we shall denote by $Q_{\boldsymbol{X}}$, is compared to the empirically derived threshold to determine the acceptance or rejection of H_0.

Optimisation of the choice of $Q_{\boldsymbol{X}}$ is still possible.

4.2. Choice of t values

As mentioned in subsection 4.1, the choice of values of t at which to evaluate the test statistic can be crucial to the power of the test against a particular alternative. The cfs of some distributions are very similar the Gaussian cf at *some* values of t, while vastly different at others. The values used must be able to provide good results against a large number of alternatives.

If increased sensitivity against a particular distribution were required, it would be possible to determine at what points the cfs under H_0 and H_1 differ the most and evaluate them only at these points. Due to the variance associated with the estimator, this maximises the power of the test. Evaluation at a small number of points also makes the test easier to implement and quicker to run, especially in the higher-dimensional spaces used in the multivariate case.

Finding optimal t values is beyond the scope of this paper and will be reported elsewhere.

4.3. Characteristic function domain kernel function

The choice of ckf when using the KCFE method, was investigated in [8]. For a standardised Gaussian process, a

Gaussian kernel was proposed, which for the multivariate case is,

$$\varphi(t_1, t_2, \ldots, t_M) = \exp\left(\frac{-\sigma_{\Re}^2 \sum_{m=1}^{M} t_m^2}{2}\right)$$

where $\sigma_{\Re}$ is the kernel function (inverse) width. The optimum $\sigma_{\Re'}$, was found to be approximated by $\sigma_{\Re'} = 0.97N^{-0.2}$, where N is the number of data points. Because of the similarity between the KCFE and the smoothed difference method used in this paper, we will use this kernel.

5. Results and discussion

The power of the test for the bivariate case is investigated through Monte Carlo simulations. The results are presented in Tables 2, for $N = 20$ data points, and 3, when 50 data points are used. The power of the proposed test is compared to bivariate tests based on an adaptation of the Shapiro Wilk W test and Mardia and Foster's S_W^2 test [6].

Distribution[2]	Q_X	W	S_W^2
$\chi_2^2(2,2,1)$	63.3	75.6	49.4
$\chi_2^2(2,2,2)$	51.8	59.2	42.4
$\chi_2^2(2,2,3)$	43.1	48.1	35.4
t_3	54.8	47.6	48.6
Cauchy	97.6	96.9	94.9
$0.8N_2(0,\Sigma) + 0.2N_2(0,9\Sigma)$	60.9	44.3	57.0
$0.9N_2(0,\Sigma) + 0.1N_2(0,16\Sigma)$	59.5	52.3	61.9
Log Normal	89.0	87.6	70.2
Average	65.0	64.0	57.5

Table 2. Percentage of realisations rejected at a 5% level of significance for $N = 20$.

For this investigation, 5000 replications of each test were performed. The level of significance was set at 5% and thresholds were derived empirically using bivariate Gaussian data with marginal means of zero and covariance matrix $M = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

All data was standardised by marginal means and variances, then the correlation between the two variables, $\hat{M}$, was estimated. The bivariate Gaussian cf, with correlation matrix $\hat{M}$, was then evaluated at $t = [-2.5, -1.5, -0.5, 0.5, 1.5, 2.5]$. The difference between this cf and the magnitude of the ecf was then smoothed by the Gaussian ckf described in section 4.3. Taking the magnitude of the mean of this function produced the test statistic Q_X which was compared to the empirically derived threshold.

[2]The bivariate chi-square distributions, $\chi_2^2(\nu_1, \nu_2; \nu)$, are the joint distributions of W_1 and W_2 where $W_i = V_i + V$. Each V_i is an independent χ^2 variate with ν_i degrees of freedom and V is a χ^2 variate with ν degrees of freedom.

Distribution[2]	Q_X	W	S_W^2
$\chi_2^2(2,2,1)$	96.2	99.7	95.3
$\chi_2^2(2,2,2)$	90.2	97.3	88.6
$\chi_2^2(2,2,3)$	81.7	91.2	83.0
t_3	90.7	75.5	86.3
Cauchy	100	100	100
$0.8N_2(0,\Sigma) + 0.2N_2(0,9\Sigma)$	92.5	78.1	93.6
$0.9N_2(0,\Sigma) + 0.1N_2(0,16\Sigma)$	89.2	85.8	93.1
Log Normal	99.8	100	99.4
Average	92.5	91.0	92.4

Table 3. Percentage of realisations rejected at a 5% level of significance for $N = 50$.

The results in Tables 2 and 3 show an increase in average power by the smoothed difference cf based test, Q_X, against both the W and S_W^2 tests. Each test has *some* distributions against which it is more powerful, however, none of the tests performs best for *all* distributions.

An important feature of the performance of the Q_X test is that of the results shown, only once was it the worst performing of the three tests – this was for the $\chi_2^2(2,2,3)$ distribution when $N = 50$. Both the W and the S_W^2 tests had their relative strengths and weaknesses, however, the Q_X test performed consistently well. For example, the W test was the best performing for all the bivariate χ^2 distributions, while the S_W^2 test generally achieved higher rejection rates for the other distributions.

The *Average* rejection rate should be viewed with caution, as it really only has meaning if all 8 distributions, and only these distributions, are equally likely to occur.

The difference in relative performance of the Q_X test did not appear to be affected by a change in the number of data points used, N.

Several other features of cf based tests make them appealing compared to the other tests available and mentioned in this paper. Cf based tests can easily be adapted to test for any known, fixed distribution, and their sensitivity to certain alternatives can be adjusted through the choice of test statistic and t values.

6. Conclusions

We have proposed an improvement to the KCFE based Gaussianity test by smoothing the difference between the ecf

and cf under H_0. Additionally, this test has been extended to the multivariate case, allowing it to be used on correlated data. An empirical power study and comparison to other tests revealed encouraging results.

Further studies into the performance of this test are certainly possible, as well as detailed investigation into optimisation of some aspects of the test, namely, the choice of a test statistic and the range of t values at which to evaluate the cfs.

References

[1] M. J. Arnold. *Signals and Noise*. PhD thesis, Queensland University of Technology, Australia, 1995.

[2] P. Comon and L. Deruaz. Normality tests for coloured samples. In *IEEE-ATHOS Workshop on Higher-Order Statistics*, pages 217–221, Begur, Spain, 12–14 June 1995.

[3] T. W. Epps. Testing that a stationary time series is Gaussian. *The Annals of Statistics*, 15:1683–1698, 1987.

[4] I. A. Koutrouvelis. A goodness-of-fit test of simple hypotheses based on the empirical characteristic function. *Biometrika*, 67(1):238–40, 1980.

[5] E. Moulines and K. Choukri. Time-domain procedures for testing that a stationary time- series is Gaussian. *IEEE Transactions on Signal Processing*, 44(8):2010–2025, Aug 1996.

[6] G. S. Mudholkar, D. K. Srivastava, and C. T. Lin. Some p-variate adaptations of the Shapiro-Wilk test of normality. *Communications in Statistics, Theory and Methods*, 24(4):953–985, 1995.

[7] B. W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.

[8] A. M. Zoubir and M. J. Arnold. Testing Gaussianity with the characteristic function. *Signal Processing*, 53:245–255, September 1996.

WAL: Modeling, Statistical Tests

Chair: Ananthram Swami
Malgudi Systems, USA

Testing the Gaussian Assumption for Self-similar Teletraffic models

Stephen Bates and Steve McLaughlin
Department of Electrical Engineering,
University of Edinburgh, Mayfield Rd, Edinburgh, EH9 3JL

Abstract

Both the fractional Brownian motion (fBm) and the Auto-regressive Integrated Moving Average (ARIMA) models have been applied to teletraffic scenarios in recent years. These models became popular after the discovery that Ethernet and VBR video data appear to possess the property of self-similarity. However the results presented in this paper suggest that Ethernet data is more impulsive than traffic generated by these models.

1 Introduction

Teletraffic modelling plays an important part in network design and resource allocation. It allows engineers to predict the network behaviour before physical construction and permits experimentation without endangering real systems. However for modelling to be effective the system must be simulated as accurately as possible. For this reason the time-series analysis and modelling of teletraffic scenarios has become a significant area of research.

One of the most important discoveries in this area in recent years was the result of work carried out at Bellcore Labs on Ethernet traces [5]. They discovered that such traces were similar in distribution across a wide range of time scales (from milliseconds to hundreds of seconds). This phenomena was due to the traces being self-similar in nature, which implied that they also exhibited long range dependency (LRD).

The importance of this discovery becomes apparent when it is observed that Poisson, ARMA and Markov processes are unable to exhibit LRD. In fact they are short range dependent (SRD) processes, but until this discovery almost all teletraffic modelling research had been based upon them. Obviously new models were required and the most popular to-date, those based on fBm and ARIMA processes, are introduced in the next section.

This paper concentrates on examining the original Bellcore Ethernet data and attempting to determine whether the new models (which generate traffic with Gaussian innovations) are optimal. The evidence is presented in increasing order of quantitative strength and suggests that a more impulsive innovation distribution than the Gaussian may be required.

2 Self-similar processes and models

If $x(t)$ is a random process then it is defined as self-similar iff

$$x(at) \stackrel{d}{=} a^H x(t),\ a > 0,\ H \in (0, 1]. \tag{1}$$

H is the Hurst exponent and is a measure of self-similarity; it lies in the range $0.5 < H \leq 1$ for a self-similar process, and $\stackrel{d}{=}$ denotes equal in distribution.

Self-similar time-series have some interesting properties:

- They possess a hyperbolically decaying autocorrelation function of the form,

$$r(k) \simeq k^{(2H-2)} L(t) \text{ as } k \to \infty, \tag{2}$$

where $L(t)$ is a slowly varying function at infinity[1]. Therefore the autocorrelation function is unsummable, i.e.

$$\sum_k r(k) = \infty. \tag{3}$$

This infinite sum is the definition for *long range dependency* so all self-similar signals are long-range dependent.

- The sample variance decays more slowly than the number of points in the sample, m,

$$\text{Var}(X^{(m)}) \propto m^{(2H-2)}. \tag{4}$$

This is why the sample statistics such as mean and variance are slow to converge.

- The power spectrum obeys a $1/f$ type law close to the origin,

$$f(\lambda) \propto \lambda^{1-2H} \text{ as } \lambda \to 0. \tag{5}$$

[1] i.e. $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$ for all $x > 0$

0-8186-8005-9/97 $10.00 © 1997 IEEE

This is why self-similar and long range dependent processes are sometimes termed *1/f-noise*.

Obviously these properties cannot be reproduced by SRD processes. So processes that are capable of producing time-series with these properties were required. Two processes that were already well known in other fields, but were applicable to self-similar teletraffic modelling, were fBm [6] and ARIMA [3] processes.

fBm is a non-stationary process, $B_H(t)$, with stationary increments (these increments are often termed fractional Gaussian noise (fGn)) and is defined as

$$B_H(t) = C[\int_{-\infty}^{0}\{(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}dB(s) + \int_{0}^{t}(t-s)^{H-\frac{1}{2}}\}dB(s)], \quad (6)$$

where $B(.)$ is standard Brownian motion and C is a normalising constant. Norros developed a teletraffic model that uses fBm innovations [7] to produce a cumulative arrival stream, $\hat{A}(t)$, in accordance with

$$\hat{A}(t) = mt + \sqrt{am}B_H(t). \quad (7)$$

Here, m is defined as the mean arrival rate per second and a is the variance parameter.

The second process is based on the SRD ARMA model but incorporates an additional fractional difference term. If B is defined as the backshift operator (i.e. $B(x_t) = x_{t-1}$, $B^2(x_t) = x_{t-2}$ etc.) and $\varphi(.)$ and $\psi(.)$ are polynomial functions of order p and q respectively then the ARIMA(p, d, q) process is given as

$$\varphi(B)(1-B)^d x_t = \psi(B)\epsilon_t. \quad (8)$$

ϵ_t is the excitation noise and if $-\frac{1}{2} < d < \frac{1}{2}$ then the process is self-similar. In practice the ARIMA trace is often obtained by generating a fGn trace with a suitable H and filtering this noise with the ARMA coefficients.

The important points to note from this section are that the stationary increments of $B_H(t)$ are drawn from a Gaussian distribution, i.e.

$$B_H(t) - B_H(t-t') \sim N(0, \sigma^2), \quad (9)$$

and that the exitation noise for the ARIMA model, ϵ_t, is Gaussian in nature.

3 Testing the Gaussian assumption

In section 2 it was shown that both of the most studied self-similar models generate traffic by performing a linear transform on fGn. In this section we compare true fGn with what is assumed to be fGn in the Bellcore Ethernet data.

3.1 The data

The Bellcore data consists of three files (`pAug.TL`, `pOct.TL` and `OctExt.TL`); the means by which these files were constructed is detailed in [5]. For each file a family of *work per time unit* discrete data sets, $W^{\Delta t}[n]$ were constructed. This was done by selecting a suitable time unit and totalling the number of Ethernet bytes recorded per time unit, for the entire trace.

3.2 Data transform

To obtain the assumed fGn trace consider the following. Assume $\hat{W}^{\Delta t_1}[n]$ is the cumulative work per Δt_1 seconds data set, then

$$\hat{W}^{\Delta t_1}[n] = mn\Delta t_1 + \sqrt{am}B_H[n]. \quad (10)$$

By differencing (10) over Δt_1, the following expression for the arrivals of the work data set is obtained,

$$W^{\Delta t_1}[n] = m\Delta t_1 + \sqrt{am}Y_H^*[n]. \quad (11)$$

So the data sets are assumed to be composed of some mean term plus a scaled fGn process, $Y_H^*[n]$ which can be normalised by estimating the value of $\sqrt{am}$.

3.3 Data Analysis

The time-series $Y_H^*[n]$ was obtained from the data sets using the transform described in section 3.2. In theory, if the Gaussian assumption is valid, this time-series should be drawn from a Gaussian distribution. To test for this a true fGn trace with a suitable H was generated. The pdfs of both were then calculated and plotted. Figure 1 compares the pdfs for data sets generated from the trace file `pAug.TL`. The fGn pdf can be seen to decay more rapidly in the tails than any of the work data sets. This result would tend to suggest that outlying events (those far from the mean of the distribution) are more probable in the real data than in the Gaussian model (i.e. the real data is more impulsive).

The previous results in this section support, in a qualitative manner, the hypothesis that fGn innovations are not able to completely capture the behaviour of Ethernet traffic. In the remainder, this hypothesis is tested more formally.

Shapiro and Wilk developed a test for normality that has been shown to outperform other methods [8]. However the Shapiro-Wilk (SW) test requires the calculation of n coefficients where n is the sample size. Due to the large number of elements in some of the data sets considered in this paper the large sample approximation to the SW test, developed by D'Agustino, was employed [4].

This test is based on a statistic, D, which is a ratio of the unbiased estimate of the population standard deviation

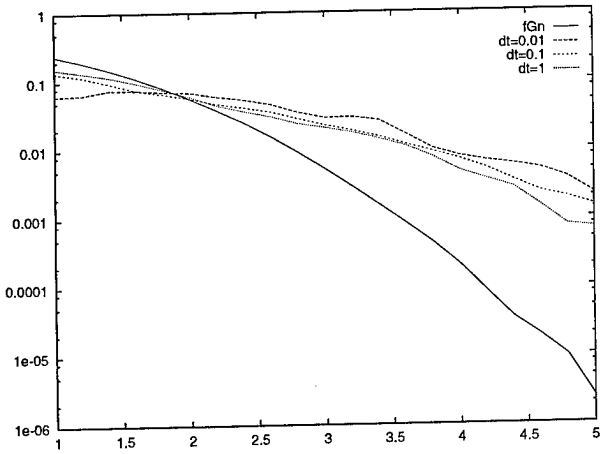


Figure 1. The log probability plot of the assumed fGn traces for the work data sets constructed from pAug.TL. The time step for each data set is given as dt (in seconds) and a true fGn case is given as fGn.

to the sample standard deviation. Assume $X_1, \cdots, X_N$ is a sample set of size N. If the set is ordered such that $X_{1,N} \leq X_{2,N} \cdots \leq X_{N,N}$ the statistic D is given by

$$D = \frac{T}{N^2 S}, \tag{12}$$

where

$$T = \sum_{i=1}^{N} \{i - \frac{1}{2}(N+1)\} X_{i,N}, \tag{13}$$

and

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{N}, \tag{14}$$

where $\bar{X}$ is the sample mean. The statistic D has $\mu \neq 0$ and $\sigma^2 \neq 1$ but can be normalised to give the statistic Y using

$$Y = \frac{(D - (2\sqrt{\pi})^{-1})\sqrt{n}}{0.02998598}. \tag{15}$$

If the set X is not drawn from a normal distribution then $E[Y] \neq 0$. The critical levels for a $(1-\alpha)$ confidence level can be obtained from tables or by calculation [4].

The data sets $W^{0.01}$, $W^{0.1}$, W^1 and W^{10} for `pAug.TL` and `pOct.TL` were transformed into $Y_H^*[n]$ as in section 3.2 and then tested using the adapted SW test and the statistic Y was recorded. The critical values for a .95 confidence level were calculated for the required N (c_{low} and c_{high}). Also, the Hurst exponent estimate, $\hat{H}$, for each of the assummed fGn data sets was calculated using the Whittle approximation method (chapters 5 and 6 in [2]). The trace file `OctExt.TL` was gathered over a longer time scale and hence the SW test and $\hat{H}$ were calculated for W^1, W^{10},W^{100} and W^{1000}. The results and critical values for all the data sets are given in Tables 1-3.

pAug.TL					
	N	$\hat{H}$	c_{low}	c_{high}	Y
$W^{0.01}$	314283	0.582	-1.967	1.953	-1218.94
$W^{0.1}$	31428	0.594	-1.982	1.937	-149.36
W^1	3142	0.594	-2.030	1.887	-29.27
W^{10}	314	0.591	-2.176	1.724	-7.16

Table 1. *Results of the adjusted SW test for the Bellcore trace pAug.TL*

pOct.TL					
	N	$\hat{H}$	c_{low}	c_{high}	Y
$W^{0.01}$	130390	0.565	-1.971	1.949	-206.5
$W^{0.1}$	13039	0.565	-1.994	1.925	-4.22
W^1	1303	0.565	-2.060	1.847	-8.24
W^{10}	130	0.565	-2.290	1.585	-10.71

Table 2 . *Results of the adjusted SW test for the Bellcore trace pOct.TL*

OctExt.TL					
	N	$\hat{H}$	c_{low}	c_{high}	Y
W^1	122798	0.613	-1.971	1.948	-1488.39
W^{10}	12279	0.634	-1.996	1.923	-403.53
W^{100}	1227	0.621	-2.072	1.843	-106.43
W^{1000}	122	0.579	-2.298	1.572	-25.77

Table 3 . *Results of the adjusted SW test for the Bellcore trace OctExt.TL*

The results in the tables above show that at a .95 confidence level, all the data sets for all the Bellcore traces failed the adjusted Shapiro Wilks test. Therefore the assummed fGn data sets produced from Ethernet trace can not be true fGn (which would be expected to pass the SW test). It might be possible from this to conclude that the Ethernet data does not conform to the models introduced in section 2.

It was possible that the self-similar nature of the data could be affecting the results. Equation (4) illustrates that, due to LRD, non-parametric measures of self-similar time-series can suffer from convergence problems. It was possible that the LRD in the Ethernet data affected the Shapiro Wilk test results. In order to investigate this, true fGn traces with varying H and length (N) were constructed and tested using the Shapiro Wilk test with a .95 confidence interval. This experiment was repeated 200 times and the percentage pass rates are recorded in Table 4.

	True fGn trace H					
N	0.5	0.6	0.7	0.8	0.9	1.0
1024	80.5	81.5	84	76.5	77	85.5
2048	87.5	90.5	89.5	91.5	87.5	82.5
4096	**96**	**96**	**95**	92	87.5	73.5
8192	**97.5**	**96**	**96.6**	**95.5**	88.5	59.5
16384	**98**	**97**	**98**	**96**	85.5	41
32768	**98.5**	**99**	**99.5**	**97.5**	80.5	54
65536	**99**	**97.5**	**98**	**98**	89.5	46.5
131072	**99**	**99**	**99**	**97**	86.5	36.5
262144	**99**	**98.5**	**98**	**98.5**	79	16
524288	**99.5**	**99.5**	**100**	**97.5**	81	17.5
1048576	**99**	**100**	**98.5**	**97**	78	18

Table 4. *Percentage of the SW tests that passed for true fGn traces.*

The bold entries in Table 4 indicate where the fGn traces seem to pass the SW test the expected number of times for a .95 confidence level. Extreme values of H (i.e. $H \geq 0.9$) and low values of N seem to produce fewer than expected hypothesis passes and when $H = 1$ the pass rate decreases with N. This is due to the fact that a time-series with such a large H possess very strong correlations over all time scales and any non-parametric measure (such as those in (12)-(14)) will not converge for any N. If we compare the lengths and Hurst exponent estimates for the Bellcore data sets we can see that most occur inside the bolded area of Table 4. This suggests that LRD is not the reason for the failure of the SW tests for the Bellcore data.

4 Conclusions

This paper has presented evidence, both qualitative and quantitative, to suggest that Ethernet data does not conform to popular self-similar models. The evidence would suggest that Ethernet is more impulsive than the Gaussian case which these models assume. One alternative is to use the more general class of stable distributions, which can be much more impulsive than the Gaussian case. In [1] the parameters for a stable distribution were estimated for both the Bellcore Ethernet data and similar data generated in Edinburgh. The initial results gave sensible estimates for these estimates and on-going work involves the construction of suitable models that incorporate these distributions.

5 Acknowledgements

Stephen Bates is funded by a scholarship from GEC Marconi Avionics and thanks them for their continuing support. Steve McLaughlin is funded by the Royal Society.

References

[1] S. Bates and S. McLaughlin. An investigation of the impulsive nature of ethernet data using stable distributions. In J. Hillston and R. Pooley, editors, *Proceedings of the 12th UK Performance Engineering Workshop*, pages 17–32, 1996.

[2] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall, 1994.

[3] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: forecasting and control*. Holden-Day, 1976.

[4] R. D'Agostino. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2):341–348, 1971.

[5] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–14, February 1994.

[6] B. B. Mandelbrot and J. W. van Ness. Brownian motion, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[7] I. Norros. On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13(6):953–962, August 1995.

[8] S. S. Shapiro, M. B. Wilk, and H. J. Chen. A comparative study of various tests for normality. *Journal of the American Statistical Society*, 63:1343–72, 1968.

Sampling Jitter Detection using Higher-Order Statistics

J.Y. Tourneret
ENSEEIHT / GAPSE
2 Rue Camichel, BP 7122
31071 Toulouse Cedex, France
tournere@len7.enseeiht.fr

A. Ferrari
UNSA / I3S
41 Bd. Napoleon III
06041 Nice, France
ferrari@unice.fr

B. Lacaze
ENSEEIHT / GAPSE
2 Rue Camichel, BP 7122
31071 Toulouse Cedex, France
lacaze@len7.enseeiht.fr

Abstract

The spectrum of a signal subjected to sampling jitter can be significantly different from the spectrum of the same signal sampled without jitter. The first part of the paper shows that the spectrum of a continuous Gaussian signal can be reconstructed from a combined use of the sampled (with jitter) signal second and fourth-order statistics. This spectral reconstruction is then used to detect the presence or absence of jitter in a sampled signal. A likelihood ratio detector based on the spectral corrective term is studied. It gives a reference to which suboptimal detectors can be compared.

1. Introduction

The sampling jitter detection and estimation have been investigated in many applications. These applications include spectral estimation [8] or source localization [10]. The sampling jitter affects the sampled signal Power Spectral Density (PSD). It is usually modeled as a random process whose variance characterizes the spectral degradation. This paper studies the problem of sampling jitter detection for Gaussian signals with unknown spectral properties.

The problem of sampling jitter detection was recently considered in [12] (using the bispectrum of sampled data). However, the study was restricted to zero-mean stationary band-limited continuous processes, with nonzero third-order cumulant function. Consequently, the results cannot be applied to continuous Gaussian processes. The estimation and detection of sampling jitter were also considered in [1] and [2]. However, the study was restricted to signals with known or specific properties.

The paper is organized as follows. The first section shows that the Gaussian signal PSD can be determined as a function of the measured sample second and fourth-order statistics. The second section proposes a suboptimal hypothesis test for sampling jitter detection. Simulation results and conclusions are reported in the two last sections.

2. Problem Formulation

A real stationary band-limited Gaussian process $X(t)$ is considered, with unknown PSD $s(\omega)$ ($s(\omega) = 0$ for $|\omega| > \pi/h$) and autocorrelation function $r(\tau)$. $X(t)$ is sampled according to the Shannon theorem by a sampler that exhibits jitter. The sampling instants are $t_n = nh + \gamma_n$, where γ_n is a zero-mean i.i.d. sequence with unknown characteristic function $\Phi(\omega) = E[\exp(j\omega\gamma_n)]$. The random variable γ_n is assumed to be symmetrically distributed such that $\Phi(\omega)$ is real even. Parameter h is chosen equal to 1 without loss of generality. The samples are denoted $x_n = x(t_n)$. The first part of the paper addresses the problem of estimating the PSD $s(\omega)$ using the observations x_n.

3. Spectral Estimation

3.1. Second-Order Statistics

The second order statistics of x_n can be computed as functions of the PSD $s(\omega)$ and the jitter characteristic function $\Phi(\omega)$ (using conditional expectation with respect to $t_{n+p} - t_n$). Denote $r_h(m) = E[x_n x_{n+m}]$. The following results are obtained [4]:

$$m \neq 0 \qquad r_h(m) = \int_{-\pi}^{\pi} |\Phi(u)|^2 s(u) e^{jum} du$$

$$m = 0 \qquad r_h(0) = \int_{-\pi}^{\pi} s(u)\, du = r(0)$$

The PSD of x_n can then be determined:

0-8186-8005-9/97 $10.00 © 1997 IEEE

$$s_h(\omega) = \frac{1}{2\pi}\sum_{m=-\infty}^{+\infty} r_h(m)\, e^{-j\omega m}$$
$$= |\Phi(\omega)|^2 s(\omega) + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(1-|\Phi(u)|^2\right) s(u)\, du$$

The modulus of $\Phi(u)$ being upper bounded by 1, the integral in the previous equation is positive. Consequently, the jitter effect can be interpreted at the second order as a filtering process embedded in additive noise.

3.2. Fourth-Order Cumulants

The fourth-order cumulants of the sample signal x_n denoted $c_h(k,l,m)$ can be computed as functions of the PSD $s(\omega)$ and the jitter characteristic function $\Phi(\omega)$. The only non-zero fourth order cumulants of x_n are $c_h(k,l,l)$ with $k \neq l$ [4][9]:

- first Case: $k = 0, l \neq 0$
  $c_h(0,l,l) = \iint_{-\pi}^{+\pi} f(u,v)\, e^{j(u+v)l} s(u)\, s(v)\, dudv$
  with $f(u,v) = 2\left[|\Phi(u+v)|^2 - |\Phi(u)\,\Phi(v)|^2\right]$.
- second case: $k \neq 0, l \neq k$
  $c_h(k,l,l) = \iint_{-\pi}^{+\pi} g(u,v)\, e^{j(u+v)l - jvk} s(u)\, s(v)\, dudv$
  with
  $g(u,v) = 2\left[\Phi(u+v)\,\Phi^*(u)\,\Phi^*(v) - |\Phi(u)\,\Phi(v)|^2\right]$.

Define the trispectrum of x_n in the usual domain $|\omega_1|, |\omega_2|, |\omega_3| < \pi, |\omega_1+\omega_2+\omega_3| < \pi$:

$$T_h(\omega_1,\omega_2,\omega_3) = \frac{1}{(2\pi)^3}\sum_{l,p;p\neq 0} c_h(l-p,l,l)\, e^{-j((\omega_1+\omega_2+\omega_3)l-\omega_1 p)}$$

The previous cumulant expressions allow to express the trispectrum $T_h(\omega_1,\omega_2,\omega_3)$ as a function of $\Phi(.)$ and $s(.)$.

3.3. Spectral Reconstruction

Consider a new process $\widetilde{x}_n$ whose autocorrelation function and PSD are:

$$\widetilde{r}_h(m) = \int_{-\pi}^{\pi} |\Phi(u)|^2 s(u)\, e^{jum} du$$
$$\widetilde{s}_h(\omega) = |\Phi(\omega)|^2 s(\omega) \quad (1)$$

It has been shown that $\widetilde{r}_h(m)$ can be estimated with two samplers with independent jitters γ_n and γ'_n [11].This section derives a closed form expression of the PSD $s(\omega)$ as a function of $\widetilde{s}_h(\omega)$ and the trispectrum $T_h(.)$, which is independent of the characteristic function $\Phi(\omega)$. Denote $u = -\omega_1$, $v = \omega_1+\omega_2+\omega_3$ and define a new function $\Gamma_h(u,v) = \pi T_h(\omega_1,\omega_2,\omega_3)$:

$$\Gamma_h(u,v) = \frac{1}{8\pi^2}\sum_{l,p;p\neq 0} c_h(l-p,l,l)\, e^{-j(vl+up)} \quad (2)$$

It can be shown that [4]:

$$s(\omega)\,\widetilde{s}_h(\omega) = 2\int_0^{\omega}\left(\frac{\partial\Gamma_h(u,v)}{\partial u}\right)_{v=-u} du + \widetilde{s}_h(\omega)^2$$

A closed form expression of the above integral as a function of the fourth-order cumulants of x_n can then be derived

$$I(\omega) = \int_0^{\omega}\left(\frac{\partial\Gamma_h(u,v)}{\partial u}\right)_{v=-u} du = \frac{\omega}{8\pi^2}\sum_k \mathrm{sinc}(k\omega/2)\sin(k\omega/2)\sum_{l\neq k}(l-k)\, c_h(k,l,l) \quad (3)$$

The PSD $s(\omega)$ is then related to the PSD $\widetilde{s}_h(\omega)$ and cumulants $c_h(k,l,l)$ by the following relation

$$s(\omega)\,\widetilde{s}_h(\omega) = 2I(\omega) + \widetilde{s}_h(\omega)^2 \quad (4)$$

Note that eq. (4) leads to $\widetilde{s}_h(\omega) = s_h(\omega)$ in absence of jitter.

4. Sampling Jitter Detection

4.1. Simple Binary Hypothesis Test

The main contribution of this work is to use eq. (4) for sampling jitter detection. The sampling jitter detection problem can be expressed as a binary hypothesis testing problem:

$$\begin{array}{l} H_0 \text{ (no jitter)} : x_n = x(n) \\ H_1 \text{ (jitter)} : x_n = x(n+\gamma_n) \end{array} \quad (5)$$

with $n \in \{1,...,N\}$. Given the statistics of the jitter and the continuous process, a Likelihood Ratio Test (LRT) can be derived. Unfortunately, the test statistics is usually difficult to study. Moreover, the statistical properties of the jitter are unknown in many practical applications. This section proposes a suboptimal sampling jitter detector. This detector does not require any assumption concerning the jitter statistics and the signal parameters. The problem (5) can be reformulated as the following binary hypothesis testing problem:

$$\begin{array}{rcll} H_0 \text{ (no jitter)} & : & I(\omega) = 0 & \forall\omega \in]0,\pi] \\ H_1 \text{ (jitter)} & : & I(\omega) \neq 0 & \forall\omega \in \Omega \end{array} \quad (6)$$

In eq. (6), Ω is a non-empty subset included in $]0,\pi]$. The detection problem can then be discretized as follows (considering different frequencies $\omega_k = k\frac{\pi}{M}$):

$$\begin{array}{l} H_0 \text{ (no jitter)} \ : \ I_M = 0 \\ H_1 \text{ (jitter)} \ : \ I_M \neq 0 \end{array} \quad (7)$$

where $I_M = \left(I\left(\frac{\pi}{M}\right), I\left(2\frac{\pi}{M}\right), ..., I(\pi)\right)^t$. The computation of the vector I_M involves infinite summations (see eq. (3)). However, these summations can be truncated using fourth-order mixing properties of x_n and the decaying of $\sin_c(.)$. Simulation results were shown to be in good agreement with mathematical derivations for $|k|, |l| \leq 6$ in [4]. Consequently, the test is performed using the corresponding approximated (truncated) vector $\widetilde{I}_M$. The Neyman-Pearson Detector (NPD) [13] for (7) can be expressed as (assuming parameters $\widetilde{I}_M, \Sigma_0$ and Σ_1 are known) :

$$H_0 \text{ rejected if } \quad T^2_{LR} > S(PFA) \quad (8)$$

T^2_{LR} is the test statistics and $S(PFA)$ is a threshold depending on the PFA (defined by $PFA = P\left[T^2_{LR} > S \,|H_0\right]$). The performance of the NPD can be determined from the statistical properties of T^2_{LR}. $T^2_{LR} = Q_1\left(\widetilde{I}_M\right) - Q_0\left(\widetilde{I}_M\right)$ is the difference between two quadratic forms of $\widetilde{I}_M$. The vector $\widetilde{I}_M$ is linearly related to a fourth-order cumulant vector c_h. Using the property that the cumulant vector estimate $\widehat{c}_h$ is asymptotically an unbiased Gaussian vector [5], the statistics of $\widehat{I}_M$ (estimate of $\widetilde{I}_M$) can be asymptotically determined under both hypotheses:

$$\begin{array}{l} H_0 \text{ (no jitter)} \ : \widehat{I}_M \sim \mathcal{N}(0, \Sigma_0) \\ H_1 \text{ (jitter)} : \widehat{I}_M \sim \mathcal{N}\left(\widetilde{I}_M, \Sigma_1\right) \end{array}$$

Consequently, T^2_{LR} is the difference between two positive definite quadratic forms of the Gaussian vector $\widehat{I}_M$ (under both hypotheses). Unfortunately, the quadratic form T^2_{LR} can be indefinite. Relatively little attention has been devoted to the problem of obtaining the distribution of indefinite quadratic forms of Gaussian vectors. Expansions as mixtures of noncentral χ^2 distributions, in Laguerre or Maclaurin series were derived in [7]. Unfortunately, these expansions lead to high computational cost. Instead, approximations by Gaussian or Gamma distributions can be used which lead to a simple test performance computation. Fig. 1 shows that the Probability Density Function (PDF) of T^2_{LR} can be approximated (with sufficient accuracy) by the Gaussian PDF. The Gaussian assumption for T^2_{LR} allows to study the NPD performance, for instance in terms of ROC curves as a function of the jitter variance (see simulation results).

4.2. Composite Hypothesis Test

The NPD maximizes the Probability of Detection (PD) for a fixed Probability of False Alarm (PFA). It provides a reference to which suboptimal detectors can be compared. However, it requires the knowledge of parameters $\widetilde{I}_M, \Sigma_0$ and Σ_1. These parameters are unknown in many practical applications and have to be estimated. The maximum likelihood estimator for these parameters combined with the NPD yields the Generalized Likelihood Ratio Detector (GLRD).This section studies the GLRD [13] for (7), when the parameters $\widetilde{I}_M, \Sigma_0$ and Σ_1 are unknown. The N samples x_n are divided into K records, each having M samples. The GLRD for (7) was shown to be equivalent to[6][12]:

$$H_0 \text{ rejected if } \quad T^2_{GLR} = \widehat{\mu}^T \widehat{\Sigma}^{-1} \widehat{\mu} > S(PFA) \quad (9)$$

where $\widehat{\mu}$ and $\widehat{\Sigma}$ are the usual mean and covariance estimates. The statistical properties of T^2_{GLR} under both hypotheses were derived in [6]. However, they lead to a high computational cost. Instead, approximations using the central or noncentral chi-square distribution can be used which lead to a simple test performance computation [12].

5. Simulation Results

Many simulations have been performed to validate the previous theoretical results. In this paper, the simulation results are presented for a Gaussian stationary process $dX = -\alpha X dt + dv$ with $\alpha > 0$ (as in [4]). $v(t)$ is a Wiener process with variance parameter 2π and $x(0)$ is a zero mean Gaussian variable with variance π/α. The PSD of this process is $s(\omega) = 1/\left(\alpha^2 + \omega^2\right)$. The discretization of $X(t)$ at instants t_n leads to [3] :

$$x_{n+1} = e^{-\alpha(t_{n+1}-t_n)} x_n + \left(1 - e^{-2\alpha(t_{n+1}-t_n)}\right)^{1/2} \varepsilon_n$$

where ε_n is an i.i.d. Gaussian sequence with variance π/α. For brevity, all experiments have been conducted with $\alpha = 2$. The jitter is a binary process taking (equally likely) the values $\{-\sigma, \sigma\}$, where σ^2 is the jitter variance. The corresponding jitter characteristic function is $\Phi(\omega) = \cos(\omega\sigma)$.

5.1. Spectral Reconstruction

Eq. (4) can be used for PSD estimation:

$$s(\omega) = 2\frac{I(\omega)}{\widetilde{s}_h(\omega)} + \widetilde{s}_h(\omega)$$

However, a division by $\widetilde{s}_h(\omega)$ is required, which can lead to numerical problems. A slightly different estimation scheme is preferred, to cope with these numerical problems. The spectra $s(\omega)$ and $\widetilde{s}_h(\omega)$ are very similar in practical applications. Eq. (4) can then be approximated by

$$s(\omega)^2 \approx \widetilde{s}_h(\omega)^2 + 2I(\omega)$$

The cumulants $c_h(k,l,l)$ are estimated for $|k|, |l| \leq 6, k \neq l$, by averaging the estimates obtained on segments of $N/30$ samples with 10% overlap ($N = 10^5$). Fig. 2 shows the theoretical PSD's $s(\omega), \widetilde{s}_h(\omega)$ and their estimates. As can be seen, mathematical derivations and simulation results are in good agreement.

5.2. Binary Hypothesis Test

The detection performance is studied for different number of samples and different values of jitter variance. Fig's 3 shows the ROC Curves corresponding to (8) (obtained for known parameters $\widetilde{I}_M, \Sigma_0$ and Σ_1). As can be seen, a large number of samples is required to obtain a good jitter detection performance (even for large jitter variance). This means that the jitter has little effect on the PSD (as it can be analytically checked using the expression of $\Phi(\omega), s(\omega)$ and $\widetilde{s}(\omega)$). ROC curves corresponding to a high-pass signal (corrupted by binary sampling jitter) are currently under investigation.

Note that the test can be significantly improved when the truncation bounds in (3) and M tend to infinity when the number of samples N tends to infinity. Unfortunately the asymptotic distribution of $\widehat{I}_M$ is not easy to derive.

Normality tests can be a useful tool for detecting the presence or absence of jitter. However, the interesting problem is to detect the significant spectral degradations due to jitter (and not the deviations from normality), in a spectral estimation context. The problem (7) has then to be preferred to any normality test, in these spectral estimation applications. Moreover, deviations from normality are not always inherent to sampling jitter.

6. Conclusion

An original solution to the spectral estimation problem in the presence of jitter was presented. The estimation procedure was restricted to continuous Gaussian signals but did not require any additional assumption. A likelihood ratio detector was derived to decide whether spectral distortions due to sampling jitter are significant or not. The proposed test was suboptimal since it did not work on the data themselves. However, it did not require any statistical assumption for the jitter process.

References

[1] N. Aakvaag and B. Lacaze, "Estimation of Sampling Jitter Variance using Fourth-Order Cumulants," *Proc. IEEE Work. on HOS*, Begur, Spain, 1995.

[2] C. Andrieu, A. Doucet, and P. Duvaut, "Bayesian Estimation of the Variance of a Jitter using MCMC," *Proc. IEEE Work. on SSAP*, Corfu, Greece, pp. 24-27, 1996.

[3] K. Aström. *Introduction to stochastic Control Theory.* Academic Press, 1970.

[4] A. Ferrari, J.Y. Tourneret, and G. Alengrin, "Spectral Estimation of a Gaussian signal sampled with jitter," *Proc. IEEE ICASSP-97*, Munich, Germany, April 1997.

[5] J. A. R. Fonollosa, "Sample Cumulants of Stationary Processes: Asymptotic Results", IEEE Trans. on Signal Processing, vol. 43, April 1995

[6] N. Giri, "On the complex analogues T^2 and R^2 tests," *Ann. Math. Statist.*, vol. 36, pp. 664-670, 1965.

[7] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions-2, first edition*, John Wiley, 1970.

[8] M. I. Moore, A. W. Visser and T. G. L. Shirtcliffe, "Spectral analysis of oceans profiles from unequally spaced data," *J. Geophys. Res.*, vol. 93, pp. 655-664, 1988.

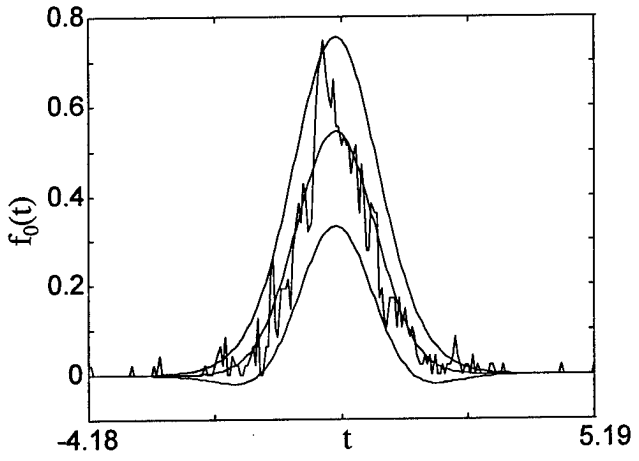
[9] A. Prieto Guerrero and J. Y. Tourneret, "Second and Fourth-Order Statistics of Gaussian Signals sampled with Jitter: Application to Jitter Variance Estimation," *Proc. of Int. Workshop on Sampling Theory and Applications*, Aveiro, Portugal, June 23-26, 1997.

[10] Y. Rockah and P. M. Schultheiss, "Array shape calibration using sources in unknown locations, part I and II," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 35, 1987.

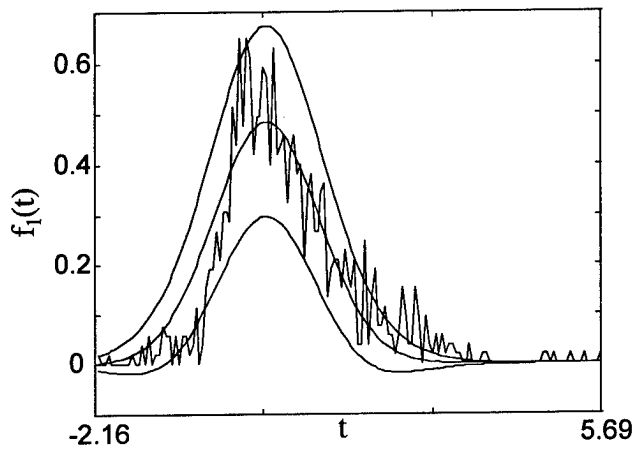
[11] H. Shapiro and R. Silverman, "Alias-free Sampling of Random Noise," J. Soc. Indust. Appl. Math., pp. 225-248, 1960.

[12] I. Sharfer and H. Messer, "The Bispectrum of Sampled Data: Part I-Detection of the Sampling Jitter," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 41, n°1, Jan. 1993.

[13] H. L. Van Trees, *Detection Estimation and Modulation Theory, Part I.* New York: Wiley, 1968.



(a)



(b)

Figure 1. Histograms and approximated PDF of T^2_{LR} with 95% confidence intervals a) under H_0 b) under H_1.

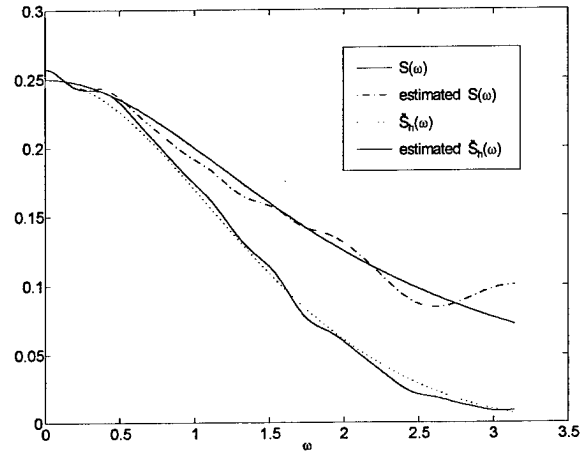
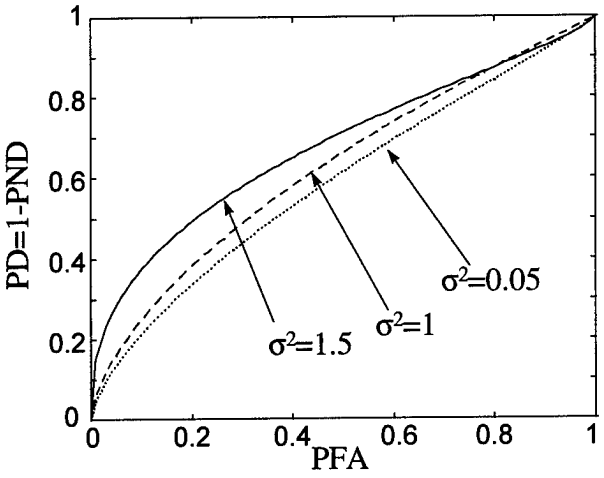
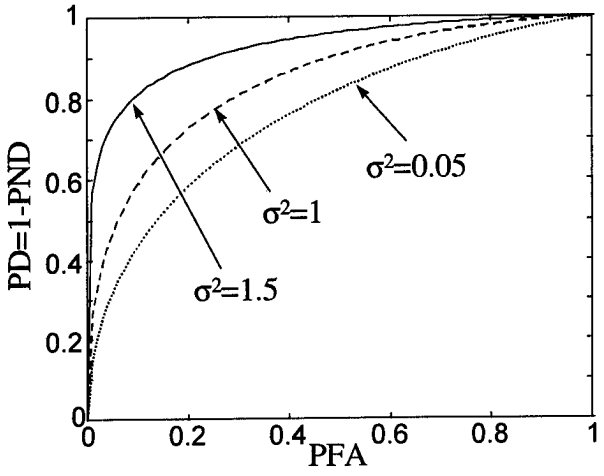


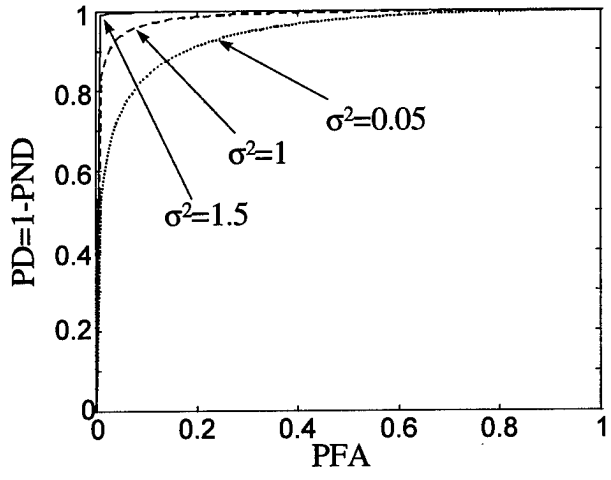
Figure 2. PSD's estimates.



(a)



(b)



(c)

Figures 3. ROC Curves for the Neyman-Pearson Detector a) $N = 10^4$ b) $N = 10^5$ c) $N = 5.10^5$.

Playing with Long Range Dependence and HOS

P.O. Amblard, J.M. Brossier and J.L. Lacoume
CEPHAG-UPRESA CNRS 5086-ENSIEG
BP 46, 38402 Saint-Martin d'Hères Cédex
France
Bidou.Amblard@cephag.inpg.fr

Abstract

In this paper, we consider some independent problems concerning HOS and nonGaussian long range dependent process. We exhibit HOS for non integer FARIMA processes, showing that singularities may occur in multispectrum. A class of nonGaussian processes constructed by a multiplicative process on the dyadic tree is then considered. Finally, long range dependence problems in the Fourier transform are examined.

1 Introduction

Many processes occuring in fields as different as physics, hydrology, economics, ...appear to present long memory, that is, present a very slow decay of their correlation function. Furthermore, in fields where nonlinear phenomena are involved, measured signals are often well modelled using nonGaussian random functions. It seems therefore important to consider nonGaussian stochastic processes that present long memory.

The aim of this paper is not to give a full theory of nonGaussian processes with long memory. Instead, we will try to give some ideas concerning HOS and nonGaussian long memory processes. We first consider FARIMA processes which are easy to handle since they are linear processes. We examine some of their HOS properties and show the complexity of these statistics. We then present a construction of nonGaussian processes based on a multiplicative process on the dyadic tree. Some studies are developped, both on the process and on the reconstructed process using a Haar basis, exhibiting long range memory features. Finally, we will provide a short discussion about long range dependence, Gaussianity and the Fourier transform.

2 FARIMA models

This section is devoted to the study of FARIMA models. We first recall some results concerning second-order statistics before examining some higher-order properties, especially on the bispectrum.

2.1 Second-Order Statistics

FARIMA models are parametric models now widely used to model long range dependent stationary processes. They are defined as follows. Let B be the unit delay operator $Bx_n = x_{n-1}$. The operator $\nabla = I - B$, where I is the identity operator, is the discrete differencing operator. A FARIMA(0,d,0) process x_n is a fractional integration of an i.i.d sequence, it is to say a solution of $\nabla^d x_n = e_n$, where e_n is an i.id. sequence. It can be shown [2] that, if $d \in [-0.5, 0.5]$ there is a unique stationary process x_n solution of $\nabla^d x_n = e_n$. It can be formally written $x_n = \nabla^{-d} e_n$. Using the binomial expansion, x_n can be written

$$x_n = \sum_{k=0}^{+\infty} \Psi_k e_{n-k} \text{ with } \Psi_k = \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}, \qquad (1)$$

$\Gamma(t)$ being the Gamma function. Second-order statistics of FARIMA(0,d,0) processes are explicitly known. The correlation $C_{x,2}(k) = Cum[x_n, x_{n+k}]$ reads

$$C_{x,2}(k) = \sigma_e^2 (-1)^k \frac{\Gamma(1-2d)}{\Gamma(1-d+k)\Gamma(1-d-k)}$$

whereas the spectral density $S_{x,2}(\lambda) = \sum C_{x,2}(k) \exp(-2i\pi\lambda k)$ is given by

$$S_{x,2}(\lambda) = \sigma_e^2 |2\sin(\pi\lambda)|^{-2d}$$

Interesting conclusions may be drawn from these expressions. Note that if $d \in [0, 0.5]$ $S_{x,2}(\lambda)$ presents a singularity near frequency 0, since for λ small, one

0-8186-8005-9/97 $10.00 © 1997 IEEE

gets the behavior $S_{x,2}(\lambda) \approx \lambda^{-2d}$. This implies that for large k, the correlation behaves like k^{2d-1}. Therefore, the correlation is not summable since its decay is slower than $1/k$. This justifies the terminology "long range dependence" or "long memory".

FARIMA(0,d,0) models may be generalized to FARIMA(p,d,q) models. This is done by filtering a FARIMA(0,d,0) process by an ARMA(p,q) filter. Therefore, x_n is a FARIMA(p,d,q) process if it is a solution of $a(B)\nabla^d x_n = b(B)e_n$ where e_n is an i.i.d. sequence. The spectral density of this process is clearly

$$S_{x,2}(\lambda) = \sigma_e^2 |2\sin(\pi\lambda)|^{-2d} \frac{a(e^{-2i\pi\lambda})}{b(e^{-2i\pi\lambda})}$$

However, no explicit form for the correlation function is known. But the asymptotic behavior of $C_{x,2}(k)$ is again like k^{2d-1}, showing that FARIMA(p,d,q) processes have long memory whenever $d \in [0, 0.5]$. Hence, the presence of the ARMA terms gives the modeling of the short time duration correlations.

2.2 HOS for FARIMA(0,d,0) processes

Let x_n be a FARIMA(0,d,0) process, whose driving i.i.d. sequence is denoted by e_n and is assumed non-Gaussian. x_n results from the filtering of white noise e_n by a filter of impulse response Ψ_n. It may be argued that Ψ_n has an infinite extent and that x_n should tend to a Gaussian process. However, classical central limit theorems do not apply if $d \in [0, 0.5]$ since Ψ_n is not summable. Indeed, Ψ_n is slowly decaying to zero as n^{d-1}! We can then evaluate HOS of x_n.

Let $C_{x,p}(\mathbf{k}) = Cum[x_n, x_{n+k_1}, \ldots, x_{n+k_{p-1}}]$ be the p-th order multicorrelation of x_n and $S_{x,p}(\boldsymbol{\lambda}) = \sum_{\mathbf{k}} C_{x,p}(\mathbf{k}) \exp(-2i\pi \mathbf{k}^T \boldsymbol{\lambda})$ its p-th order multispectral density ($\mathbf{k} = (k_1, \ldots, k_{p-1}), \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{p-1})$ and ^{T} stands for transposition). x_n being the output of a linear, time-invariant filter driven by e_n, it is well known that

$$S_{x,p}(\boldsymbol{\lambda}) = H^*(\sum_{i=1}^{p-1} \lambda_i) H(\lambda_1) \ldots H(\lambda_{p-1}) S_{e,p}(\boldsymbol{\lambda})$$

where $H(\lambda)$ is the transfert function associated with the impulse response Ψ_k. Since e_n is an i.i.d. sequence, its multispectral densities are constant, or $S_{e,p}(\boldsymbol{\lambda}) = \gamma_{e,p}$. Furthermore, $H(\lambda)$ is simply $[1 - \exp(-2i\pi\lambda)]^{-d}$ or $\exp(i\pi\lambda d)[2i\sin(\pi\lambda)]^{-d}$. Then, the p-th order multispectral density of x_n reads

$$S_{x,p}(\boldsymbol{\lambda}) = \gamma_{e,p}(2i)^{-pd} \left[\sin(\pi \sum_{i=1}^{p-1} \lambda_i) \prod_{i=1}^{p-1} \sin(\pi\lambda_i)\right]^{-d}$$

Therefore, when $d \in [0, 0.5]$, the multispectral densities presents singularities on the following manifolds

$$\begin{cases} \lambda_i & = & 0 \ \forall i = 1, \ldots, p-1 \\ \sum_{i=1}^{p-1} \lambda_i & = & 0 \end{cases}$$

These singularities behaves like powerlaws, and furthermore are not the same on all the manifolds. To get some more ideas on this behavior, we restrict the discussion to the bispectrum which reads

$$S_{x,3}(\boldsymbol{\lambda}) = \frac{\gamma_{e,3}}{(8i)^d} [\sin(\pi\lambda_1)\sin(\pi\lambda_2)\sin(\pi(\lambda_1+\lambda_2))]^{-d} \quad (2)$$

Suppose now that $\lambda_1, \lambda_2 << 1$ and that λ_1 and λ_2 are of the same order. Then, the bispectrum (2) may be approximated by

$$S_{x,3}(\boldsymbol{\lambda}) \approx \gamma_{e,3}(8i\pi^3)^{-d} [\lambda_1\lambda_2(\lambda_1+\lambda_2)]^{-d}$$

This approximation is no longer valid if the frequencies are not of the same order. For example, let $\lambda_2 >> 0$ and let λ_1 going to zero. Then, the approximation is $S_{x,3}(\boldsymbol{\lambda}) \approx \gamma_{e,k}(-2i)^{-3d} \sin^{-2d}(\pi\lambda_2)(\pi\lambda_1)^{-d}$. This gives the singularity strength on the axis $\lambda_1 = 0$ when λ_2 is fixed. $S_{x,3}(\boldsymbol{\lambda})$ is depicted in figure (1) where the behavior described above appears clearly.

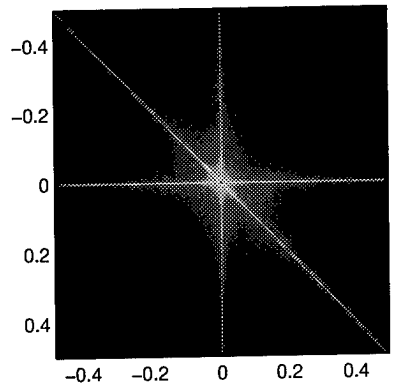


Figure 1. *Theoretical bispectrum of a FARIMA(0,d,0) process.(Log of the modulus)*

We now turn to the multicorrelations. We unfortunately have not found a closed form expression for them in the simple case of FARIMA(0,d,0) processes. However, we can provide some heuristic arguments to study the behavior of the bicorrelation. It reads

$$C_{x,p}(k_1, k_2) = \int\int S_{x,p}(k_1, k_2) e^{2i\pi(k_1\lambda_1 + k_2\lambda_2)} d\lambda_1 d\lambda_2$$

The heuristic argument is based on the fact that the bispectrum is correctly approximated by

$\gamma_{e,3}(8i\pi^3)^{-d}[\lambda_1\lambda_2(\lambda_1+\lambda_2)]^{-d}$ for small frequencies. But, small frequencies correspond to long time lags in the bicorrelation. Therefore, to get the behavior of the bicorrelation at infinity we can study the integral

$$\int\int [\lambda_1\lambda_2(\lambda_1+\lambda_2)]^{-d} e^{2i\pi(k_1\lambda_1+k_2\lambda_2)} d\lambda_1 d\lambda_2$$

Let us study the bicorrelation on its diagonal $k_1 = k_2 = k$. Changing variables λ_1, λ_2 to variables $u_1 = k\lambda_1, u_2 = k\lambda_2$ leads to

$$C_{x,p}(k,k) = k^{3d-2} \int\int \frac{e^{2i\pi(u_1+u_2)}}{[u_1u_2(u_1+u_2)]^d} du_1 du_2$$

Since the integral is independent of k, the asymptotic behavior of $C_{x,p}(k,k)$ is k^{3d-2}. If $d \in [-0.5, 0.5]$, then $3d-2 \in [-3.5, -.5]$. Hence, the long memory characteristic is again found on the bispectrum (things may be different at higher orders, since the power laws may decrease faster that k^{-1}).

As it can be seen, the study of higher-order properties of FARIMA models is difficult. The few arguments presented above lead to the conclusion that the study of nonGaussian long memory processes will be a hard task. Further, the complicated behavior of the bispectrum or of the bicorrelation will make it difficult to estimate these statistics. For example, smoothing the biperiodogram will need special care around the singularity axis, and the study of this kind of estimators will involve special limit theorems.

To conclude, let us say that the modelling of nonGaussian signals which present long memory is at its beginning. In the following section we present another model which mimics a physical phenomenon to lead to nonGaussian $1/f$ processes.

3 Multiplicative process on the dyadic tree

Recent studies in turbulence stress on the modelling of the energy transfert from large scales to small scales (energy cascade). In [3], B. Castaing has proposed to relate the pdf of the velocity increments at a scale r to the pdf of the velocity increments at a larger scale L via

$$p_{\delta v_r}(x) = \int G_{rL}(\ln\alpha) \frac{1}{\alpha} p_{\delta v_L}\left(\frac{x}{\alpha}\right) d\ln\alpha \qquad (3)$$

Working on the velocity increments is a mean to define the notion of scale. This notion can be made clearer using the wavelet transform. Therefore, a natural generalization of the last equation is obtained using the continuous wavelet transform instead of the velocity increments[4].

Coming back to (3), we see that in terms of random variables, the increment at scale r is obtained via the product of the increments at scale L by an independent random variable G. We can thus mimic this behavior by defining a multiplicative process on a tree. This approach as already be considered in [1, 4]. The node at a defined depth will defined the details at a corresponding scale. In order to be able to reconstruct a time series, we will work on the dyadic tree corresponding to a decomposition over an orthogonal wavelet base.

In the next section, we elaborate on the multiplicative process on the tree, forgetting the interpretation in terms of details. The next section deals with the reconstructed time series.

3.1 Multiplicative process on the dyadic tree

Consider the process $\{d_{j,k}, j = 0, 1, \ldots; k = 0, \ldots, 2^j - 1\}$ on the dyadic tree defined as

$$d_{j,k} = a_{j,k} d_{j-1,[k/2]}$$

where [.] stands for integer part of, and where $\{a_{j,k}, j = 1, \ldots; k = 0, \ldots, 2^j - 1\}$ are i.i.d. random variables. Since the process is multiplicative, it seems easier to study it with moments rather than cumulants. Let $E[d^n_{J,k_1} d^m_{J,k_2}] = \Gamma_J^{n,m}(k_1,k_2)$ be a multicorrelation of order $n+m$ of the process at depth J. Let $d(1,2)$ be the distance on the tree between the two nodes (J,k_1) and (J,k_2). Let j_a the depth of their common ancestor on the tree. Then, there exist a k and a sequence k_i such that $d_{J,k_1} = d_{j_a,k} \prod_{i=j_a+1}^{J} a_{i,k_i}$. Furthermore, since $E[d^q_{j_a,k}] = E[d^q_{0,0}] E[a^q]^{j_a}$ and $d(1,2) = 2J - 2j_a$, we obtain

$$\begin{aligned}\Gamma_J^{n,m}(k_1,k_2) &= E\left[d^{n+m}_{0,0}\right] E\left[a^{n+m}\right]^{J-d(1,2)/2} \\ &\quad E\left[a^n\right]^{d(1,2)/2} E\left[a^m\right]^{d(1,2)/2}\end{aligned}$$

This matrix as a special structure inherited from the tree. This structure is depicted in figure (2). It is constant by block, the value in each block depending on distance $d(1,2)$.

Note that for a fixed k_1, the values taken by the correlation may decrease exponentially fast with the number of the block. But since the sizes of the blocks grow exponentially fast, it is the possible to get a non summable multicorrelation.

Indeed, let us study $\Gamma_J^{n,n}(0,l)$. This multicorrelation may also be written as

$$\Gamma_J^{n,n}(0,l) = E[d^{2n}_{00}]\{E[a^{2n}]^J$$

$$+ \sum_{k=1}^{J} \chi_{[2^{k-1},2^k-1]}(l) E[a^n]^{2k} E[a^{2n}]^{J-k}\}$$

where χ_I is the characteristic function of interval I. Since $\Gamma_j^{n,n}(0,l)$ is constant by block, it is easy evaluate $S_J = \sum_{l=0}^{2^J-1} \Gamma_J^{n,n}(0,l)$ which reads

$$S_J = E[d_{00}^{2n}] E[a^{2n}]^J \{1 + \frac{1}{2} \sum_{k=1}^{J} \{\frac{2E[a^n]^2}{E[a^{2n}]}\}^k\}$$

The leading term $q = 2E[a^n]^2/E[a^{2n}]$ in the geometric series is always lower or equal to 2. There are therefore 4 principal cases for the convergence of S_J :

- $E[a^{2n}] > 1$: $\lim_{J\to+\infty} S_j = \infty, \forall q$.
- $E[a^{2n}] < 1$ and $q \leq 1$: $\lim_{J\to+\infty} S_j = 0$.
- $E[a^{2n}] < 1$ and $1 < q < 2$: the term in the brackets goes like $(2E[a^n]^2/E[a^{2n}])^J$, and thus $S_J \approx 2^J E[a^n]^{2J}$. Then, convergence occurs if $E[a^n] \leq 1/\sqrt{2}$, and we get divergence otherwise.

This allows the construction of a 2nd-order white process which is long memory at order 4 (for example). To get an example of such a process, suppose that the $\{a_{j,k}\}$ are uniformly distributed over $[-\alpha, \alpha]$. Odd moments are thus equal to zero, and in particular, we obtain that the $d_{j,k}$ are uncorrelated : they constitute a 2nd order white process.

We now find α such that second and fourth-order moments exist, but in order that S_J is not summable. According to the previous conditions, we must have $E[a^4] < 1$, $q = 2E[a^2]^2/E[a^4] \in [1,2]$ and $E[a^2] > 1/\sqrt{2}$. We furthermore impose $E[a^2] < 1$ to insure $E[d_{J,k}^2] < \infty$.

Even orders moments of the $\{a_{j,k}\}$ are given by $E[a^{2n}] = \alpha^{2n}/(2n+1)$ and then $q = 10/9 \in [1,2]$. $E[a^4] < 1$ is verified if $\alpha < \sqrt[4]{5}$ and $1/\sqrt{2} < E[a^2] < 1$ if $\alpha \in [\sqrt{3\sqrt{2}/2}, \sqrt{3}]$. All these conditions are satisfied if $\alpha \in [\sqrt{3\sqrt{2}/2}, \sqrt[4]{5}]$ leading to a 4th-order long memory sequence!

We now turn to the study of the process reconstructed from the details via an orthogonal wavelet expansion.

3.2 Reconstructed process : the Haar basis example

To get a reconstructed time series from the details studied previously, we choose the Haar basis for simplicity. Thus, the reconstructed series may be written,

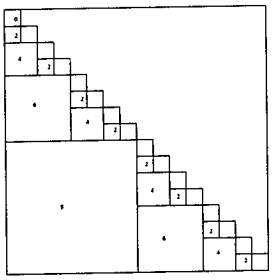


Figure 2. *Structure of the correlation matrix for the multiplicative process. Numbers correspond to the distance on the tree between two nodes.*

using J scales

$$x_n = A\chi_{[0,2^J-1]}(n) + \sum_{j=0}^{J} \sum_{k=0}^{2^j-1} d_{j,k} 2^{j/2} h(2^j n - k)$$

where $\chi_{[0,2^J-1]}(n)$ represents the scaling function associated to the Haar basis, and where $h(n)$ is the Haar wavelet.

At depth J, we want to study the correlation between x at time α and x at time β. Then (J,α) and (J,β) are two nodes of the dyadic tree. Let c be the depth of their common ancestor node, and let Z be the value of the reconstructed process at depth c on this node. We can then write

$$Z = d_{00}\{1 + \sum_{j=1}^{c} \alpha_j \left(\prod_{i=1}^{j} a_i\right)\}$$

where $\alpha_j = \pm 2^{j/2}$. The signs $\pm$ appears to be the coefficients of the haar wavelets h. The $+$ sign correspond to a branch going to the left and the $-$ sign to a branch going to the right. Note that we have skipped the indice k in the notation of the variables $a_{j,k}$ since it has no effect on the result. Now, $x(\alpha)$ and $x(\beta)$ are functions of Z via

$$\begin{aligned} x(\alpha) &= Z\{1 + \sum_{j=c+1}^{J} \alpha_j \left(\prod_{i=c+1}^{j} a_i\right)\} \\ x(\beta) &= Z\{1 + \sum_{k=c+1}^{J} \beta_k \left(\prod_{l=c+1}^{k} b_l\right)\} \end{aligned}$$

where the b_l correspond to the $a_{j,k}$ on the branch linking Z and $x(\beta)$.

Suppose now that the $a_{j,k}$ have zero mean. Then it is easy to see that $E(x(\alpha)x(\beta)) = E(Z^2)$. The evaluation of $E(Z^2)$ is also simple since the a_j are indepen-

dent variables. We thus obtain

$$E(Z^2) = 1 + \sum_{j=1}^{c} \alpha_j^2 E(a^2)^j = \frac{1-(2\sigma^2)^{c+1}}{1-2\sigma^2}$$

To get the correlation coefficient between $x(\alpha)$ and $x(\beta)$, we normalize last result by the according standard deviations. Finally, if the $\{a_{j,k}\}$ are centered, we get

$$Corr[x(\alpha), x(\beta)] = \frac{1-(2\sigma^2)^{c+1}}{1-(2\sigma^2)^{J+1}}$$

Suppose that $|\alpha - \beta| = \tau >> 1$. In that case, it is easy to see that the depth of the common ancestor goes like $c \approx J - \log_2(\tau)$. Hence, if $2\sigma^2 > 1$ we get

$$Corr[x(\alpha), x(\beta)] \approx \frac{1}{\tau^{1+\log_2(\sigma^2)}}$$

Since $\sigma^2 \in [0.5, 1]$, we have $1 + \log_2(\sigma^2) \in [0, 1]$ showing that the reconstructed process is 2nd-order long memory! Furthermore, this process is nonGaussian by construction and can be used to model velocity measurements in turbulent fluids.

4 Gaussianity, Fourier transform and long range dependence

The spectral domain obtained by Fourier transform is of particular interest in a large number of applications. There is a "paradigm" stating that "all" the signals in the spectral domain are gaussian. In this issue, HOS-based methods are not usefull in the spectral domain.

Let us recall that

If $x(n)$ is a discrete time strictly stationary signal with summable multicorrelations (mixing conditions), then the discrete Fourier transform

$$X_N(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) e^{-2\pi j n m / N} \qquad (4)$$

is asymptotically a complex Gaussian variable.

Therefore, long range dependent processes will not lead to gaussian variables in the frequency domain.

Let us now examine the following "paradox" : let $X_N(m)$ be the complex signal defined by (4). Then the signal obtained as N goes to infinity is a Gaussian signal. Its Fourier transform is therefore a Gaussian signal! : all signals are Gaussian!!!

This silly thing is obviously false, because one of the hypothesis in the previous theorem is violated. To see that, assume that $x(n)$ is complex valued, nonGaussian, i.i.d., strictly stationary and strictly circular (all multicorrelations involving a different number of complex conjugate than non complex conjugate are zeros). Let us evaluate for example

$$Cum[X_N(m_1), X_N(m_2), X_N^*(m_3), X_N^*(m_4)] = \frac{\gamma_{x,4}}{N} \delta(m_1 + m_2 - m_3 - m_4)$$

which is invariant under $m_i \to m_i + n$. Indeed, it can be written as $Cum[X_N(m), X_N(m+m_1), X_N^*(m+m_2), X_N^*(m+m_3)] = \frac{\gamma_{x,4}}{N} \delta(m_1 - m_2 - m_3)$. But this multicorrelation is not asymptotically summable! since its sum is $(N+1)|\gamma_{x,4}|/2$.

Therefore, there is no paradox since the multicorrelations of $X_N(m)$ are not summable. $X_N(m)$ is therefore a strange long memory process: its multicorrelations are constant over some hyperplanes. Furthermore, variables $X_N(m_i), i = 0, \ldots, N-1$ are uncorrelated, and we find another example of second-order short but fourth-order long memory process.

5 To Conclude

The aim of that paper was to study long range dependent phenomena with the help of HOS. We have shown some problems and we hope that this will open new works.

Interesting things may appear when dealing with HOS and long range memory processes. We have shown that long memory can be hidden at the second order and revealed at higher orders . This fact appears in the Fourier transform.

We believe that nonGaussian processes that present long range memory need to be studied since they appear in many fields where nonlinearity occurs. Further, the modelling of these process involves new very interesting emerging techniques...

References

[1] R. BENZI et al., "A random process for the construction of multiaffine fields", *Physica D*, vol. 65, pp. 352–358, Dec. 1993.

[2] P. J. BROCKWELL, R. A. DAVIES, *Time series: theory and methods*, Springer, 1991.

[3] B. CASTAING, "The temperature of turbulent flows", *J. Phys., II France*, vol. 6, pp. 105–114, Jan. 1996.

[4] S. ROUX, *Analyse en ondelettes de l'auto-similarité de signaux en turbulence pleinement développée*, Doctorat de l'université d'Aix-Marseille II, 18 nov 1996.

Higher and Lower-Order Properties of the Wavelet Decomposition of Self-Similar Processes

Beatrice Pesquet-Popescu and Pascal Larzabal
L.E.Si.R. – E.N.S. de Cachan,
61, Av. du Prés. Wilson, 94235 Cachan, France
e-mail: beatrice.popescu@lesir.ens-cachan.fr

Abstract

Self-similar processes have recently received increasing attention in the signal processing community, due to their wide applicability in modeling natural phenomena which exhibit "$1/f$" spectra and/or long-range dependence. On the other hand, the wavelet decomposition became a very useful tool in describing nonstationary self-similar processes. In this paper, we first investigate the existence and the properties of higher-order statistics of self-similar processes with finite variance. Then, we consider some self-similar processes with infinite variance and study the statistical properties of their wavelet coefficients.

1. Introduction

Self-similar stochastic processes are of great interest in the modeling of many natural phenomena which appear in communications, geophysics, hydrology, turbulence or economics. They are closely connected to the so-called "$1/f$ noises" and signals which exhibit long-range dependence.

They are defined by their invariance in distribution under time or space scaling. More precisely, a stochastic process $\{X(t),\ t \in \mathbb{R}\}$ is called self-similar with index H (or H-self-similar) if, for any $a > 0$, the finite dimensional probability distributions of $\{X(at),\ t \in \mathbb{R}\}$ are identical to those of $\{a^H X(t),\ t \in \mathbb{R}\}$. This self-similar stochastic structure within the data makes them ideal candidates for a multiscale analysis, such as wavelet analysis. This fact was already pointed out in the Gaussian case, when the considered model is the fractional Brownian motion (fBm). As shown in [4], [10], the variances of the wavelet coefficients of the fBm increase exponentially with respect to the scale of analysis. This characteristic leads to simple methods for the estimation of the parameter H of the fBm.

Most of the aforementioned works were concerned with the Gaussian case. However, non-Gaussian stochastic processes are often encountered in many practical situations, in particular when the signal exhibits an impulsive behaviour. Here, we are interested in self-similar processes with non-Gaussian distributions and the higher/lower-order properties of their wavelet coefficients.

The paper is organized as follows: in Section 2 we derive a condition for the existence of the cumulants of the wavelet coefficients of a nonstationary process and give some properties of the cumulants of the wavelet coefficients of self-similar processes. Section 3 is devoted to lower-order properties of α-stable self-similar processes. Section 4 concludes the paper.

2. Higher-Order Statistics

2.1. Background

Throughout this paper, we will consider zero-mean processes. Second-order stationarity is defined by the invariance of their correlation function under translation in time:

$$\forall (t_1, t_2) \in \mathbb{R}^2,\ \mathrm{E}\{X(t_1)X(t_2)\} = \mathrm{E}\{X(0)X(t_2 - t_1)\}.$$

This property provides a full characterization of stationary zero-mean Gaussian processes. Here, we suppose that higher-order statistics exist up to the order $n \geq 2$. In this case, it is useful to define a (wide-sense) n-th order stationarity, which means that the cumulants of order n are functions of $(n-1)$ variables:

$$\forall p \in \{2, \ldots, n\},\ \forall (t_1, \ldots, t_p) \in \mathbb{R}^p$$
$$\mathrm{cum}\{X(t_1), \ldots, X(t_p)\}$$
$$= \mathrm{cum}\{X(0), X(t_2 - t_1), \ldots, X(t_p - t_1)\}.$$

Note that the higher order statistics of a random process $X(t)$ exist up to the order n when

$$\forall p \in \{2, \ldots, n\}, \forall (t_1, \ldots, t_p) \in \mathbb{R}^p$$
$$\mathrm{cum}\{X(t_1), \ldots, X(t_p)\} < \infty.$$

0-8186-8005-9/97 $10.00 © 1997 IEEE

It can be proved that a sufficient condition for this property to hold is that

$$\forall p \in \{2, \ldots, n\}, \forall t \in \mathbb{R}$$
$$\text{cum}^{(p)}\{|X(t)|\} = \text{cum}\{|X(t)|, \ldots, |X(t)|\} < \infty \quad (1)$$

which is also equivalent to

$$\forall t \in \mathbb{R}, \mathrm{E}\{|X(t)|^n\} < \infty. \quad (2)$$

A class of processes including many useful self-similar processes is that of nonstationary processes with stationary (possibly fractional) increments. Fractional increments where first considered by Hosking [3], in a discrete framework (fractional ARIMA processes). The fractional difference operator is defined by:

$$(1 - q_\tau^{-1})^D = \begin{cases} \sum_{k=0}^{\infty} (-1)^k \binom{D}{k} q_\tau^{-k} & \text{if} \quad D \in \mathbb{R}_+^* \\ 1 & \text{if} \quad D = 0 \end{cases}$$

where

$$\binom{D}{k} = \frac{D(D-1)\ldots(D-k+1)}{k!}$$

and the symbol q_τ^{-1} denotes the time-shift operator: $q_\tau^{-1} F(t) = F(t-\tau)$. Consequently, we will define the increments of order D (not necessarily integer) of a continuous-time (possibly nonstationary) process $X(t)$ by:

$$\Delta^D X(t;\tau) = (1 - q_\tau^{-1})^D X(t)$$
$$= \sum_{k=1}^{\infty} (-1)^k \binom{D}{k} X(t - k\tau).$$

We will say that a nonstationary process has n-th order stationary increments of order D if the increments of order D, $\Delta^D X(t;\tau)$ and $\Delta^D X(t;\tau')$ exist and are mutually n-th order stationary processes for all (τ, τ').

2.2. Wavelet Decomposition Properties

Firstly, we investigate under which conditions the existence of the cumulants of the wavelet coefficients of a nonstationary process $X(t)$ satisfying (2) is guaranteed. Let the coefficients of the wavelet decomposition of the process at the j resolution level be denoted by:

$$c_j(k) = \frac{1}{2^{j/2}} \int_{-\infty}^{\infty} X(t) \Psi\left(\frac{t}{2^j} - k\right) dt, \quad \forall k \in \mathbb{Z},$$

where $\Psi(t)$, the "mother wavelet" [7], is assumed to be real.

We want that all cumulants of the wavelet coefficients at the j level exist up to the n-th order, i.e.:

$$\forall p \in \{2, \ldots, n\}, \forall (k_1, \ldots, k_p) \in \mathbb{Z}^p,$$
$$\text{cum}\{c_j(k_1), \ldots, c_j(k_p)\} < \infty \quad \Leftrightarrow$$
$$\forall p \in \{2, \ldots, n\}, \forall (k_1, \ldots, k_p) \in \mathbb{Z}^p,$$
$$\mathrm{E}\{c_j(k_1) \ldots c_j(k_p)\} < \infty.$$

A sufficient condition for this property to be satisfied is:

$$\forall p \in \{2, \ldots, n\}, \forall (k_1, \ldots, k_p) \in \mathbb{Z}^p,$$
$$\mathrm{E}\{|c_j(k_1)| \ldots |c_j(k_p)|\} < \infty$$

and after some calculations we find that this condition is satisfied provided that

$$\forall p \in \{2, \ldots, n\}, \forall k \in \mathbb{Z}, \forall t \in \mathbb{R},$$
$$\int_{-\infty}^{\infty} \mathrm{E}\{|X(t)|^p\}^{1/p} \frac{1}{2^{j/2}} \left|\Psi\left(\frac{t}{2^j} - k\right)\right| dt < \infty.$$

By applying the Hölder inequality, this leads to the following sufficient condition for the existence of the cumulants of the wavelet coefficients up to the n-th order:

$$\forall k \in \mathbb{Z}, \forall t \in \mathbb{R},$$
$$\int_{-\infty}^{\infty} \mathrm{E}\{|X(t)|^n\}^{1/n} \frac{1}{2^{j/2}} \left|\Psi\left(\frac{t}{2^j} - k\right)\right| dt < \infty \quad (3)$$

A similar condition for the second order case was given in [8]. Subsequently, it will be assumed that this condition is verified.

For self-similar processes, we have shown that the cumulants of the wavelet coefficients, when they exist, increase exponentially with respect to the scale of the analysis. More precisely,

Proposition 1. *If $\{c_j(k),\ k \in \mathbb{Z}\}$ are the wavelet coefficients of an H-self-similar process at resolution 2^{-j}, then their cumulants are such that*

$$\text{cum}\{c_j(k_1), \ldots, c_j(k_n)\} = 2^{jn(H+1/2)}$$
$$\times \text{cum}\{c_0(k_1), \ldots, c_0(k_n)\}.$$

Proof. This is easy to prove, by writing that:

$$\text{cum}\{c_j(k_1), \ldots, c_j(k_n)\} =$$
$$\int_{\mathbb{R}^n} \text{cum}\{X(t_1), \ldots, X(t_n)\} \frac{1}{2^{jn/2}} \Psi\left(\frac{t_1}{2^j} - k_1\right) \ldots$$
$$\times \Psi\left(\frac{t_n}{2^j} - k_n\right) dt_1 \ldots dt_n$$

By making a coordinate change, we have:

$$\text{cum}\{c_j(k_1), \ldots, c_j(k_n)\} =$$
$$2^{\frac{jn}{2}} \int_{\mathbb{R}^n} \text{cum}\{X(2^j t'_1), \ldots, X(2^j t'_n)\} \Psi(t'_1 - k_1) \ldots$$
$$\times \Psi(t'_n - k_n) dt'_1 \ldots dt'_n$$

Now, the self-similarity of the process combined with the multilinearity of the cumulants leads to the desired result. □

Note that this behaviour is exactly the opposite of that encountered for stationary processes, where the cumulants decay exponentially when $j \to \infty$.

Self-similar processes with finite variance and stationary increments form a large class of processes, including the well-known fractional Brownian motion (fBm), which is the only Gaussian process in this class and which is a process with (possibly fractional) stationary increments of order D, $\forall D > H$ [9]. All these processes have the same second order statistics :

$$\mathrm{E}\{X(t_1)X(t_2)\} = \frac{\sigma^2}{2}\left(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\right),$$

identical to those of the fBm. They thus present long-range dependence, which makes them interesting in modeling some natural phenomena.

The non-Gaussian processes of this class, subordonated to Brownian motion $B(t)$, have the following integral representation:

$$X_m(t) = \int_{\mathbb{R}_*^m} \left\{ \int_0^t g_m(s - \xi_1, \ldots, s - \xi_m)\, ds \right\} dB(\xi_1) \ldots dB(\xi_m),$$

where $B(t)$ is an ordinar Brownian motion and

$$\mathbb{R}_*^m = \mathbb{R}^m - \{(\xi_1, \ldots, \xi_m) \in \mathbb{R}^m \,|\, \xi_i = \xi_j \text{ if } i \neq j\}.$$

An interesting process within this class is the Hermite process defined by:

$$g_m(\xi_1, \ldots, \xi_m) = \Pi_{j=1}^m \left(\xi_j^+\right)^{-(1+D)/2},\ 0 < D < 1/m$$
$$(\xi_j^+ = \xi_j \text{ if } \xi_j \geq 0 \text{ and } 0 \text{ otherwise})$$

which is self-similar with index $H = 1 - \frac{1}{2}mD \in (1/2, 1)$. It can be shown that all its higher order moments exist.

Now, we will focus on nonstationary processes with fractional stationary increments, as described in Section 2.1. Generalizing some works on the second-order case [8, 6], we have shown the following result:

Proposition 2. *If a process $X(t)$ has n-th order stationary increments of order $D \in \mathbb{R}_+$ and we realize its wavelet analysis with r-vanishing moments, $D \leq r$, the wavelet coefficients are n-th order stationary.*

3. Lower-Order Statistics

3.1. Background

The second part of this work is devoted to self-similar processes with non-Gaussian α-stable distributions ($0 < \alpha < 2$). The family of stable distributions [2] is interesting, since linear transforms preserve the distribution of any linear combination of independent, identically distributed α-stable random variables. Another desirable property is the generalized central limit theorem [11]. These processes have also turned out to be good models for many impulsive signals and noises, when the great variability of data yields probability distributions with "fat" tails. These distributions have infinite variance and undefined higher-order moments but it was pointed out in [1] that many signal processing algorithms based on second-order statistics can be transposed to fractional lower-order moments.

For the sake of simplicity, we will only consider α-stable processes with $\alpha \neq 1$. However, it must be noted that our results may be extended to the case $\alpha = 1$, which corresponds to the Cauchy distribution. Let $(E, \mathcal{E}, m)$ be a measure space and $X(t)$ be an α-stable process which admits the following integral representation

$$X(t) = \int_E f(t, x) M(dx),\ t \in \mathbb{R} \tag{4}$$

where M is an α-stable random measure on $(E, \mathcal{E})$ with control measure m and

$$f \in L^\alpha(E, \mathcal{E}, m) = \left\{ f \mid f \text{ measurable and } \int_E |f(x)|^\alpha m(dx) < \infty \right\}$$

Remark that most of the usual α-stable processes admit such a representation. If M is a symmetric α-stable (SαS) measure, then X is also a SαS process.

3.2. Results

If $X(t)$ is an α-stable process given by (4), we first note [2] that the conditions

$$\int_{-\infty}^{\infty} \left(\int_E |f(t,x)|^\alpha m(dx) \right)^{1/\alpha} \frac{1}{2^{j/2}} \left| \Psi\left(\frac{t}{2^j} - k\right) \right| dt$$
$$< \infty \text{ when } \alpha < 1 \tag{5}$$
$$\int_E \left(\int_{-\infty}^{\infty} |f(t,x)| \frac{1}{2^{j/2}} \left| \Psi\left(\frac{t}{2^j} - k\right) \right| dt \right)^\alpha m(dx)$$
$$< \infty \text{ when } \alpha > 1 \tag{6}$$

are equivalent to

$$\int_{-\infty}^{\infty} |X(t)| \frac{1}{2^{j/2}} \left| \Psi\left(\frac{t}{2^j} - k\right) \right| dt < \infty \text{ almost surely.}$$

Consequently, when the conditions (5)-(6) are satisfied, the existence of the wavelet coefficients is guaranteed. In the sequel, we will assume that these conditions hold. A first important result is then

Proposition 3. *Let $X(t)$ be an α-stable (resp. SαS) process satisfying (4). Its wavelet coefficients also form an α-stable (resp. SαS) process.*

Proof. As (5)-(6) have been assumed, we can apply a Fubini-type theorem [2] which yields

$$c_j(k) = \int_E \left(\int_{-\infty}^{\infty} f(t,x) \frac{1}{2^{j/2}} \Psi\left(\frac{t}{2^j} - k\right) dt \right) M(dx) \tag{7}$$

and the function

$$g_j(k,x) = \int_{-\infty}^{\infty} f(t,x) \frac{1}{2^{j/2}} \Psi\left(\frac{t}{2^j} - k\right) dt \tag{8}$$

is in $L^\alpha(E, \mathcal{E}, m)$. Now, by using the properties of stable integrals, we conclude that $\{c_j(k),\ (j,k) \in \mathbb{Z}^2\}$ is an α-stable process. □

Let us now consider an α-stable process with an integral representation (4) and a skewness function $\beta(x)$. We further assume that the control measure m is the Lebesgue measure on $E = \mathbb{R}$. Sufficient conditions for the process to be self-similar with index H are: $\forall x \in \mathbb{R}$

$$\forall t \in \mathbb{R}, \forall a \in \mathbb{R}_+^*,\ a^{\frac{1}{\alpha} - H} f(at, ax) = f(t,x) \tag{9}$$

$$\beta(x) = \beta. \tag{10}$$

Under these conditions, we can prove the following result:

Proposition 4. *If a self-similar α-stable process satisfies (4), (9) and (10), then its wavelet coefficients form a self-similar process in the sense that*

$$\forall j \in \mathbb{Z},\ \{c_j(k),\ k \in \mathbb{Z}\} \stackrel{d}{=} \left\{ 2^{j\left(H + \frac{1}{2}\right)} c_0(k),\ k \in \mathbb{Z} \right\}. \tag{11}$$

Proof. According to Proposition 3, we have

$$c_{j+1}(k) = \int_{\mathbb{R}} g_{j+1}(k,x) M(dx).$$

Furthermore, by using (8) and (9), we find that

$$\begin{aligned} g_{j+1}(k, 2x) &= \int_{-\infty}^{\infty} f(2t, 2x) \frac{1}{2^{\frac{j-1}{2}}} \Psi\left(\frac{t}{2^j} - k\right) dt \\ &= 2^{H + \frac{1}{2} - \frac{1}{\alpha}} g_j(k,x). \end{aligned}$$

By expressing the characteristic function of $c_{j+1}(k)$, it can be seen that this relation implies that

$$\{c_{j+1}(k),\ k \in \mathbb{Z}\} \stackrel{d}{=} \left\{ 2^{\left(H + \frac{1}{2}\right)} c_j(k),\ k \in \mathbb{Z} \right\}$$

which leads to (11). □

Under the same assumptions, this property allows us to show a power-law behaviour of the fractional lower-order moments of the wavelet coefficients of a self-similar process. Thus, the lower-order moments of order p, $0 < p < \alpha$, of the wavelet coefficients are $\forall (j,k) \in \mathbb{Z}^2$,

$$\mathrm{E}\left\{|c_j(k)|^p\right\} = 2^{jp\left(H + \frac{1}{2}\right)} \mathrm{E}\left\{|c_0(k)|^p\right\}.$$

A similar law can be deduced for the codifference function of two wavelet coefficients at the same scale: $\forall (j,k,l) \in \mathbb{Z}^3$,

$$\tau_{c_j}(k,l) = \tau_{c_j(k), c_j(l)} = 2^{j\alpha\left(H + \frac{1}{2}\right)} \tau_{c_0}(k,l).$$

Now, let $X(t)$ be an α-stable process given by (4) (where m is still the Lebesgue measure). Then, a sufficient condition for $X(t)$ to have stationary increments is that there exists a function $\varphi_\tau(t)$ such that $\forall x \in \mathbb{R}$,

$$\forall t, \tau \in \mathbb{R},\ f(t+\tau, x) - f(t,x) = \varphi_\tau(t - x) \tag{12}$$

and (10) holds.

In this context, we have:

Proposition 5. *The wavelet coefficients of an α-stable process satisfying (4), (12) and (10) form stationary sequences at each scale.*

Proof. As $L^\alpha(E, \mathcal{E}, m)$ is a vector space, the increment process $\Delta X(t;\tau) = X(t+\tau) - X(t)$ is an α-stable process with

$$\Delta X(t;\tau) = \int_{-\infty}^{\infty} \varphi_\tau(t - x) M(dx).$$

Notice that this means that the increment process is a stationary moving average process.

We also remark that if $f(t,x)$ satisfies (12), then it can be written as $f(t,x) = f_0(x) + f_1(t-x)$. Therefore, by using (8) and the fact that $\int_{-\infty}^{\infty} \Psi(t) dt = 0$, we have:

$$\begin{aligned} g_j(k,x) &= \int_{-\infty}^{\infty} f_1(t-x) \frac{1}{2^{j/2}} \Psi\left(\frac{t}{2^j} - k\right) dt \\ &= g_j\left(0, x - k2^j\right). \end{aligned}$$

It can be shown that this relation entails the stationarity of the wavelet coefficients. □

Notice that Propositions 4 and 5 can be applied to several α-stable self-similar processes with stationary increments such as the linear fractional stable motion, the α-stable Lévy motion, the log-fractional stable motion and many others.

To illustrate these theoretical results, we have simulated a well-balanced linear fractional SαS motion (see Fig. 1). A spline wavelet decomposition of this process was realized over 8 resolution levels. In Fig. 2 the lower-order moments of order $p = \alpha/2$ of the wavelet coefficients $\mathrm{E}\left\{|c_j(k)|^{\alpha/2}\right\}^{2/\alpha}$ are plotted in log-scales as a function of the resolution level j.

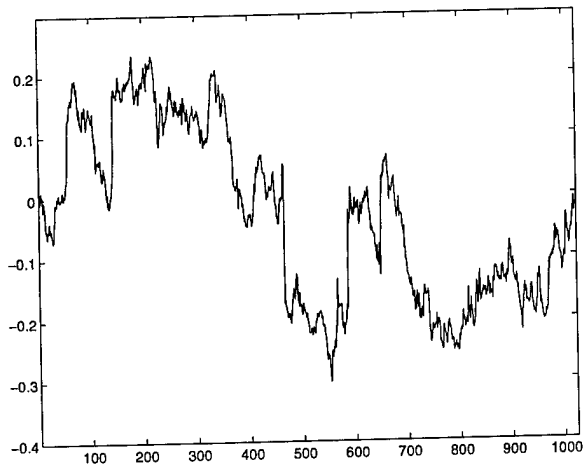


Figure 1. A realization of the well-balanced linear fractional SαS motion with $\alpha = 1.7$ and $H = 0.59$.

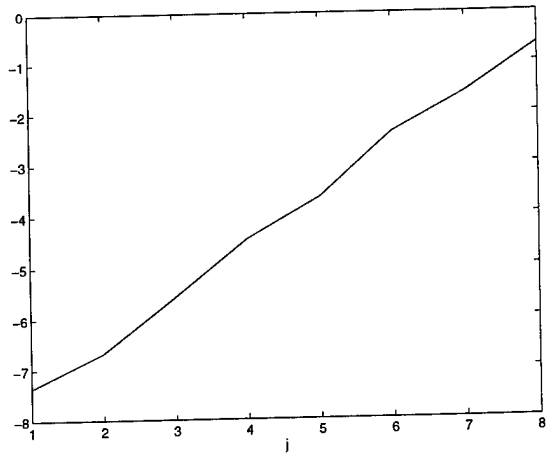


Figure 2. Power-law behaviour of the lower-order moments of the wavelet coefficients of the self-similar signal in Fig. 1.

4. Conclusion

In this paper we have firstly given sufficient conditions for the existence of the cumulant of the wavelet coefficients and then deduced a power-law behaviour for the cumulants of the self-similar processes with finite variance. Then, we have proved similar results for the lower-order moments of self-similar α-stable processes.

References

[1] M. Shao, C.L. Nikias, *Signal Processing with Fractional Lower Order Moments: Stable Processes and Their Applications,* Proceedings of the IEEE, Vol. 8, No. 7, July 1993, pp.984-1009

[2] G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance,* New York: Chapman and Hall, 1994

[3] J.R.M. Hosking, *Fractional differencing,* Biometrika, Vol. 68, No. 1, pp. 165-176, 1981

[4] P. Flandrin, *Wavelet Analysis and Synthesis of Fractional Brownian Motion* , IEEE Trans. on Information Theory, vol. 38, No. 2, March 1992, pp. 910-917

[5] B.B. Mandelbrot, J.W. Van Ness, *Fractional Brownian Motions, Fractional Noises and Applications,* SIAM Rev., 10, (4), 1968, pp. 422-437

[6] H. Krim, J.-C. Pesquet, *Multiresolution Analysis of a Class of Nonstationary Processes,* IEEE Trans. on Information Theory, vol. 41, No. 4, July 1995, pp. 1010-1020

[7] S. Mallat, *A Theory for Multiresolution Signal Decomposition: The Wavelet Transform,* IEEE Tran. on PAMI, Vol. II, No. 7, July 1989, pp. 675-693

[8] S. Cambanis and C. Houdré, *On the Continuous Wavelet Transform of Second Order Random Processes,* IEEE Tran. on Information Theory, Vol. 41, No. 3, May 1995, p. 628-633

[9] B. Pesquet-Popescu and P. Larzabal, *2D Self-Similar Processes with Stationary Fractional Increments,* in *Fractals in Engineering*, Springer-Verlag, New York, 1997, pp. 171-186

[10] G. Wornell, A.V. Oppenheim, *Estimation of Fractal Signals from Noisy Measurements Using Wavelets,* IEEE Tran. on Signal Processing, Vol. 40, No. 3, pp. 611-623, Mar. 1992

[11] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, New York: Wiley, 1971

Detection of a Common Non-Gaussian Signal in Two Sensors Using the Bootstrap

Hwa-Tung Ong, D. Robert Iskander and Abdelhak M. Zoubir
Queensland University of Technology
Signal Processing Research Centre
GPO Box 2434, Brisbane QLD 4001, Australia
h.ong@qut.edu.au

Abstract

Tugnait has used the cross bispectrum to detect non-Gaussian signals common to two sensors when the noise in each sensor are either mutually independent or have vanishing bispectra. However, the detection methods presented assume enough data are available for asymptotic results to apply. If this assumption is not valid then the performance of the detection methods will be degraded. In this paper, we propose a detection scheme based on the bootstrap that handles the small data size case. Unlike other bispectrum based techniques, the proposed scheme maintains the nominal test level while achieving high power. Simulation examples are given and the performance of the bootstrap based method is compared with a method proposed by Tugnait.

1 Introduction

The cross bispectrum of two stationary, zero-mean, discrete-time random processes, $x(t)$ and $y(t)$, is defined as

$$B_{xxy}(\omega_1,\omega_2) = \sum_{i=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} \mathsf{E}[x(t+i)x(t+k)y(t)] \times \exp\{-j(\omega_1 i+\omega_2 k)\}, \quad (1)$$

where $\mathsf{E}[\cdot]$ denotes the expectation operator. It is well known that the expression in (1) is zero when $x(t)$ and $y(t)$ are Gaussian or statistically independent random processes. This result makes the cross bispectrum a valuable tool for analysing non-Gaussian signals common to $x(t)$ and $y(t)$.

With this in mind, Tugnait [8] proposes several tests for detecting a random non-Gaussian signal in noise. Such tests are useful for verifying the validity of estimated parameters such as the differential time delay from two-sensor data. They are essentially cross bispectral versions of the tests for zero bispectrum proposed by Subba Rao and Gabr [7] and Hinich [3]. The methods are all based on asymptotic distributions.

Recently, work has been done to remove the reliance on asymptotic results for bispectral tests by incorporating the *bootstrap* [11, 12]. We extend the idea to cross bispectral tests.

The bootstrap is a statistical tool proposed by Efron [1] useful for distribution estimation. It may be seen as the marriage of data resampling and computer simulation to give rise to techniques that are able to handle small data sizes (see, for example, [9, 10]).

The advantage of a detection method using less data is at least twofold. Firstly, the method is applicable to cases where limited data is available. Secondly, the method is applicable to nonstationary data if we can assume that stationarity holds for a short segment of that data.

In the next section, we specify our statistical hypotheses and test statistic. Section 3 explains how we incorporate the bootstrap into with the proposed test. Section 4 provides simulation examples. There, the proposed method is compared with an existing test. We draw our conclusions in Section 5.

2 Statistical Test

For a finite length of data, $x(t), t = 1,\ldots,P\times N$, divide $x(t)$ into P non-overlapping segments of N consecutive measurements, $x^{(i)}(t), i = 1,\ldots,P$. Similarly for $y(t)$, we have $y^{(i)}(t), i = 1,\ldots,P$.

We test for zero cross bispectrum in the principal domain of the discrete version of $B_{xxy}(\omega_1,\omega_2)$,

$$B(m,n) = B_{xxy}(2\pi m/N, 2\pi n/N), \qquad (m,n)\in\mathcal{D},$$

where the principal domain is given by

$$\mathcal{D} = \{(m,n):\ |n|\le m\le N/2\}.$$

By using symmetric and periodic properties of the cross bispectrum, all other points outside $\mathcal{D}$ may be obtained.

0-8186-8005-9/97 $10.00 © 1997 IEEE

Thus, we specify the null and alternative hypotheses as

$$\begin{aligned} \mathrm{H}_0 : B(m,n) &\equiv 0, \\ \mathrm{H}_1 : B(m,n) &\not\equiv 0, \end{aligned} \qquad \forall (m,n) \in \mathcal{D}.$$

To test the null hypothesis H_0, we choose a test statistic given by the sum of the absolute value of Student's t statistics,

$$T = \sum_{(m,n)\in\mathcal{D}} \left| \frac{\hat{B}(m,n) - 0}{\hat{\sigma}(m,n)} \right|,$$

where $\hat{B}(m,n)$ estimates $B(m,n)$ and $\hat{\sigma}(m,n)^2$ estimates $\mathsf{var}\{\hat{B}(m,n)\}$. Multivariate or multiple hypotheses testing is avoided by the summation, and the use of Student's t statistics is justified by [2]. We discuss a bootstrap estimator for $\hat{\sigma}(m,n)^2$ in the next section.

For $\hat{B}(m,n)$, we use the following estimator [8, 5]:

$$\hat{B}(m,n) = \frac{1}{P} \sum_{i=1}^{P} I^{(i)}(m,n). \tag{2}$$

The cross biperiodograms $I^{(i)}(m,n)$, $i = 1,\ldots,P$, are given by

$$I^{(i)}(m,n) = \frac{1}{N} X^{(i)}(m) X^{(i)}(n) \bar{Y}^{(i)}(m+n), \tag{3}$$

where $X^{(i)}(m)$ and $Y^{(i)}(m)$ are the DFTs of the ith segment of $x(t)$ and $y(t)$ respectively,

$$\begin{aligned} X^{(i)}(m) &= \sum_{t=0}^{N-1} x^{(i)}(t) \exp\{-j2\pi mt/N\}, \\ Y^{(i)}(m) &= \sum_{t=0}^{N-1} y^{(i)}(t) \exp\{-j2\pi mt/N\}, \\ & \qquad m = 0,\ldots,N-1, \end{aligned}$$

and $\bar{Y}^{(i)}(m)$ is the complex conjugate of $Y^{(i)}(m)$. Note that P and N must be chosen such that $N/P \to 0$ as the data size increases to ensure consistency of the estimator.

3 Bootstrap Method

Table 1 gives the proposed detection scheme. We first estimate the cross bispectrum in Step 1.

In Step 2, we estimate approximately i.i.d. (independent and identically distributed) errors under the asymptotic results [8, 4],

$$\begin{aligned} &\mathsf{E}\{\hat{B}(m,n)\} = B(m,n) + \mathrm{O}(N^{-1}), \\ &\mathsf{var}\{\mathrm{Re}[I^{(i)}(m,n)]\} = \mathsf{var}\{\mathrm{Im}[I^{(i)}(m,n)]\} \\ &= \frac{N[1+\delta(m-n)]}{2} S_x(m) S_x(n) S_y(m+n) + \mathrm{O}(1), \\ &\qquad (m,n) \in \mathcal{D}', \end{aligned}$$

Table 1. Bootstrap based algorithm.

1. Calculate $I^{(i)}(m,n)$ and $\hat{B}(m,n)$ from (3) and (2).
2. For each segment and $(m,n) \in \mathcal{D}'$, form residuals

$$\hat{\varepsilon}^{(i)}(m,n) = \frac{I^{(i)}(m,n) - \hat{B}(m,n)}{\sqrt{\hat{V}^{(i)}(m,n)}},$$

where

$$\begin{aligned} \hat{V}^{(i)}(m,n) &= N[1+\delta(m-n)] \\ &\quad \times \hat{S}_x^{(i)}(m) \hat{S}_x^{(i)}(n) \hat{S}_y^{(i)}(m+n). \end{aligned}$$

3. Centre the residuals, $\tilde{\varepsilon}^{(i)}(m,n) = \hat{\varepsilon}^{(i)}(m,n) - \bar{\varepsilon}^{(i)}$, where $\bar{\varepsilon}^{(i)}$ represents an average over all $\hat{\varepsilon}^{(i)}(m,n)$.
4. Repeat B_1 times:

   Form bootstrap cross biperiodograms,

$$I^{(i)*}(m,n) = \hat{B}(m,n) + \hat{\varepsilon}^{(i)*}(m,n) \sqrt{\hat{V}^{(i)}(m,n)},$$

   where $\hat{\varepsilon}^{(i)*}(m,n)$ is randomly drawn, with replacement, from $\{\tilde{\varepsilon}^{(i)}(m,n), \forall (m,n) \in \mathcal{D}'\}$.

   We obtain $I_b^{(i)*}(m,n)$, $b = 1,\ldots,B_1$, and find

$$\hat{B}_b^*(m,n) = \frac{1}{P} \sum_{i=1}^{P} I^{(i)*}(m,n).$$

5. Calculate the test statistic,

$$T = \sum_{(m,n)\in\mathcal{D}'} \left| \frac{\hat{B}(m,n)}{\hat{\sigma}(m,n)} \right|.$$

6. Calculate the bootstrap statistics,

$$T_b^* = \sum_{(m,n)\in\mathcal{D}'} \left| \frac{\hat{B}_b^*(m,n) - \hat{B}(m,n)}{\hat{\sigma}_b^*(m,n)} \right|,$$

   $b = 1,\ldots,B_1$, where $\hat{\sigma}_b^*(m,n)$ is obtained in a similar way as $\hat{\sigma}(m,n)$.
7. Rank $T_1^*,\ldots,T_{B_1}^*$ in increasing order to obtain $T_{(1)}^* \le \ldots \le T_{(B_1)}^*$.
8. Reject the null hypothesis if

$$T > T_{(\lfloor (B_1+1)(1-\alpha) \rfloor)}^*,$$

   where α is the nominal level of the test.

where $S_x(m)$ and $S_y(m)$ are the discrete power spectra of $x(t)$ and $y(t)$ respectively. The region $\mathcal{D}'$ is $\mathcal{D}$ excluding $(m,-m), m=0,\ldots,N/2$ and $(N/2,0)$, the bifrequencies where the cross bispectrum is real. We first approximate $B(m,n)$, $S_x(m)$ and $S_y(m)$ by consistent estimators such as (2) and the smoothed periodogram [6]. By subtracting $\hat{B}(m,n)$ and dividing by $\sqrt{V^{(i)}(m,n)}$ (with $V^{(i)}(m,n)$ as given in Step 2), we subsequently obtain the residuals.

We force the residuals to be zero mean in Step 3 and randomly sample these in Step 4. With each resulting bootstrap residual, $\hat{\varepsilon}^{(i)*}(m,n)$, we get a new cross biperiodogram ordinate, $I^{(i)*}(m,n)$ by adding back the mean and variance information removed in Step 2. Applying (2) on cross biperiodograms formed in this way, B_1 bootstrap cross bispectrum estimates can be obtained. Ideally, B_1 is chosen as large as possible to approximate the sampling distribution of $\hat{B}(m,n)$ given $\{\hat{\varepsilon}^{(i)}(m,n), \forall (m,n) \in \mathcal{D}'\}$ as the underlying population of errors. The test statistic is calculated in Step 5.

The nested bootstrap may be used to compute $\hat{\sigma}(m,n)^2$, which estimates $\mathsf{var}\{\hat{B}(m,n)\}$. Apply Step 4 using B_2 instead of B_1 samples. Then,

$$\hat{\sigma}(m,n)^2 = \frac{1}{B_2-1}\sum_{b=1}^{B_2}\left(\hat{B}_b^*(m,n) - \frac{1}{B_2}\sum_{b'=1}^{B_2}\hat{B}_{b'}^*(m,n)\right)^2. \quad (4)$$

The standard number of bootstrap samples for estimating variances is $B_2 = 25$ [1].

Steps 6 and 7 involve calculating test statistics based on the bootstrap cross bispectrum estimates and ranking to form an empirical population under H_0. The nested bootstrap may be used to find $\hat{\sigma}_b^*(m,n)^2, b=1,\ldots,B_1$, in the same way as for $\hat{\sigma}(m,n)^2$. Apply Step 4 using B_2 bootstrap samples and use (4). However, sample from $\{\tilde{\varepsilon}_b^{(i)*}(m,n), \forall (m,n) \in \mathcal{D}'\}$ rather than from $\{\tilde{\varepsilon}^{(i)}(m,n), \forall (m,n) \in \mathcal{D}'\}$, where $\tilde{\varepsilon}_b^{(i)*}(m,n)$ is $\hat{\varepsilon}_b^{(i)*}(m,n)$ after detrending as in Step 3.

The hypothesis test is Step 8, using the empirical population formed in the preceding steps. We conclude that a common non-Gaussian signal is present if T is larger than $100(1-\alpha)\%$ of the bootstrap test statistics.

4 Simulation Examples

Using $B_1 = 199$ and $B_2 = 25$, we simulated 1000 records of an MA(10) signal common to both sensors in six different noise scenarios and for SNRs $-\infty$ dB (no signal), 0 dB, 10 dB, 20 dB and 40 dB. Each record was 512 data points long. We ran the algorithm in Table 1 on each record which was divided into 32 segments of 16 data points. For the smoothed periodogram estimates of spectra, we used the Bartlett-Priestley spectral window given in [6] with $M = 4$.

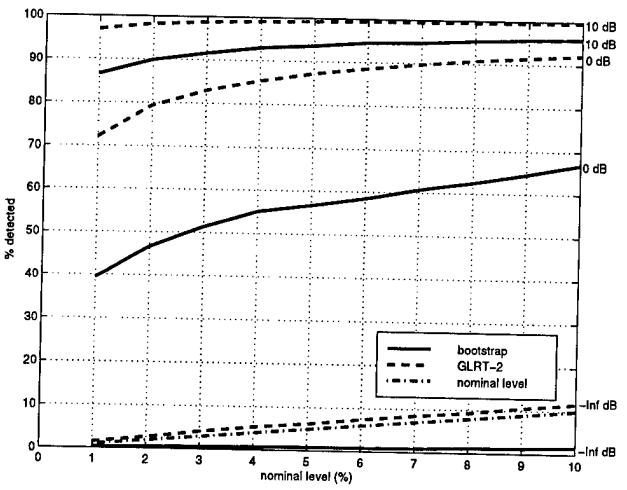


Figure 1. Example 1 detection probability versus nominal test level for SNRs $-\infty$ dB, 0 dB and 10 dB.

To compare the performance against existing tests, we selected the GLRT-2 from [8] and applied it over the same noise scenarios and SNRs. Preliminary simulations led us to the optimum settings; 4 segments of 128 data points, and a square fine grid of 169 points ($r = 6$).

We present here only a subset of the results which highlights the ability of the bootstrap based method to maintain the nominal level for short records of data.

4.1 Example 1: common linear non-Gaussian signal in common Gaussian noise

This example is based on Example 3 of [8]. The measurements at the two sensors are given by

$$\begin{aligned} x(t) &= As(t) + n_1(t) \\ y(t) &= As(t-5) + n_1(t-10) \end{aligned}$$

where

$$\begin{aligned} s(t) = {} & 0.04\epsilon(t) - 0.05\epsilon(t-1) + 0.07\epsilon(t-2) \\ & -0.21\epsilon(t-3) - 0.5\epsilon(t-4) + 0.72\epsilon(t-5) \\ & +0.36\epsilon(t-6) + 0.21\epsilon(t-8) + 0.03\epsilon(t-9) \\ & +0.07\epsilon(t-10), \end{aligned}$$

and $\epsilon(t)$ is an i.i.d., zero-mean, exponential random process. The common noise sequence is an AR(1) Gaussian sequence given by

$$n_1(t) = -0.8n_1(t) + \omega(t),$$

where $\omega(t)$ is an i.i.d. standard Gaussian process. The constant A is chosen to achieve a desired SNR.

Results are given in Figure 1. Both methods, the pro-

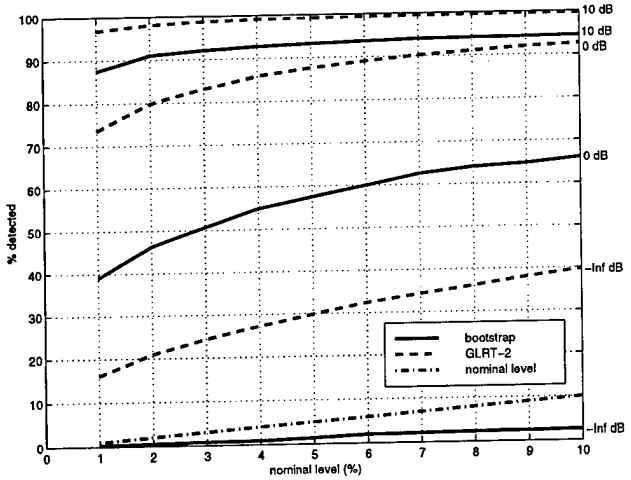


Figure 2. Example 2 detection probability versus nominal test level for SNRs $-\infty$ dB, 0 dB and 10 dB.

posed bootstrap based test and the GLRT-2, maintain the nominal level of significance (SNR=$-\infty$ dB). The GLRT-2, however, has noticeably higher probabilities of detection than the bootstrap based method.

4.2 Example 2: common linear non-Gaussian signal in independent non-Gaussian noise

This example is based on Example 4 of [8]. The sensor measurements are

$$\begin{aligned} x(t) &= A_1 s(t) + n_1(t) \\ y(t) &= A_2 s(t-5) + n_2(t) \end{aligned}$$

where $s(t)$ is as in Example 1. The noise sequences are independent AR(1) exponential sequence given by

$$n_i(t) = -0.8 n_i(t) + \omega_i(t), \qquad i = 1, 2,$$

where $\omega_1(t)$ and $\omega_2(t)$ are independent of each other and $s(t)$, and are i.i.d., zero-mean, exponential processes. The constants A_1 and A_2 are chosen to achieve a desired SNR at each sensor respectively.

Results are given in Figure 2. The GLRT-2 has higher probability of detection than the bootstrap based method as in Example 1. However, the bootstrap based method continues to maintain a low false alarm rate whereas the GLRT-2's false alarm rate is much higher than the nominal level.

Additional results using other noise structures show that the bootstrap based method maintain the nominal level while keeping the probability of detection high. The GLRT-2 is unable to maintain the nominal level for noises based on the exponential distribution. This can be seen by comparing results in Table 2 and Table 3.

Table 2. GLRT-2 results for 512 data points, SNR = 10 dB, nominal level at 5%.

Noise Used		% False Alarms	% Detected
Common Gaussian noise	i.i.d.	5	99
	AR(1)	6	99
	AR(5)	6	99
Independent exponential noise	i.i.d.	**31**	99
	AR(1)	**30**	99
	AR(5)	**28**	99

Table 3. Bootstrap results for 512 data points, SNR = 10 dB, nominal level at 5%.

Noise Used		% False Alarms	% Detected
Common Gaussian noise	i.i.d.	4	93
	AR(1)	1	94
	AR(5)	1	93
Independent exponential noise	i.i.d.	8	94
	AR(1)	1	93
	AR(5)	1	95

The AR(5) noise used was

$$\begin{aligned} n_i(t) &= 0.5 n_i(t-1) - 0.6 n_i(t-2) + 0.3 n_i(t-3) \\ &\quad -0.4 n_i(t-4) + 0.2 n_i(t-5) + \omega_i(t), \end{aligned}$$

where $\omega_1(t)$ is i.i.d. standard Gaussian for common Gaussian noises and $\omega_i(t)$, $i = 1, 2$, are mutually independent, i.i.d., zero-mean exponential for independent exponential noises.

5 Conclusions

We have presented a new method based on the bootstrap for detecting a non-Gaussian signal from the measurements of two sensors. From simulation results, we have shown that the method is able to maintain a low false alarm rate across six noise scenarios for short data records (512 points). The GLRT-2 was found to have a high probability of detection under the same conditions but its probability of false alarm was found to be much higher than the nominal level for exponential noises.

References

[1] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap.* Chapman & Hall, 1993.

[2] P. Hall and S. R. Wilson. Two guidelines for bootstrap hypothesis testing. *Biometrics*, 47:757–762, 1991.

[3] M. J. Hinich. Testing for Gaussianity and linearity of a stationary time series. *J. Time Series Analysis*, 3(1):169–176, 1982.

[4] M. J. Hinich and G. R. Wilson. Time delay estimation using the cross bispectrum. *IEEE Transactions on Signal Processing*, 40(1):106–113, January 1992.

[5] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectra Analysis - A nonlinear signal processing framework*. Prentice Hall, 1993.

[6] M. B. Preistley. *Spectral Analysis and Time Series*. Academic Press, 1981.

[7] T. Subba Rao and M. M. Gabr. A test for linearity of stationary time series. *J. Time Series Analysis*, 1(1):145–158, 1980.

[8] J. K. Tugnait. Two-channel tests for common non-Gaussian signal detection. *IEE Proceedings-F*, 140(6):343–349, December 1993.

[9] A. M. Zoubir. Bootstrap: theory and applications. In T. Luk, editor, *Advanced Signal Processing Algorithms*, volume 2027, pages 216–235, San Diego, USA, July 1993.

[10] A. M. Zoubir and B. Boashash. The bootstrap: signal processing applications. *IEEE Signal Processing Magazine*, 1997. In press.

[11] A. M. Zoubir and D. R. Iskander. A bispectrum based Gaussianity test using the bootstrap. In *Proceedings of ICASSP-96*, volume V, pages 3029–3032, 1996.

[12] A. M. Zoubir and D. R. Iskander. Bootstrapping bispectra: an application to testing for departure from Gaussianity of stationary signals. *IEEE Transaction on Signal Processing*, 1997. Submitted.

Author Index

Abeyratne, U. 72
Adnet, C. 373
Amblard, P. 453
Amblard, P.-O. 166
Arce, G. 254, 415
Arthur, N. 67
Ashouri, M. 57
Bao, Z. 290
Barth, E. 147
Bates, S. 390, 444
Bendjama, A. 354
Bettayeb, M. 214
Boashash, B. 438
Böhme, J. 285
Bondon, P. 176
Boss, D. 62
Bourennane, S. 354
Bourin, C. 176
Boutet, J. 14
Bozzano, R. 259
Briolle, F. 340
Brossier, J. 453
Brown, C. 52, 438
Campisi, P. 359
Capdevielle, V. 344, 363
Castedo, L. 326
Chandrasekhar, K. 380
Chaumette, E. 373
Chen, S. 184
Chevalier, P. 363
Chi, C.-Y. 199, 395
Cho, K. 127
Choi, S. 349
Chung, J. 127
Cichocki, A. 349
Colas, M. 43
Collis, W. 39
Comon, P. 363
Coroyer, C. 171
Coulon, M. 410
Cruces, S. 326
D'Urso, G. 82
De Lathauwer, L. 316
De Moor, B. 316
Declercq, D. 171, 425
Delaunay, G. 43
Duvaut, P. 171, 425
Esparcieux, P. 340
Faruqi, F. 52
Feng, C.-C. 199
Ferrari, A. 448
Friedlander, B. 88, 194
Frikel, M. 354
Gaeta, M. 340
Galy, J. 373
Gelle, G. 43
Gershman, A. 285
Ghogho, M. 122, 410
Ghouti, L. 214
Giannakis, G. 93
Gini, F. 93
Giunta, G. 434
Gonzalez, J. 254, 415
Govindaraju, A. 300
Griffith, D. 254, 415
Hadjileontiadis, L. 4
Hatzinakos, D. 77, 268
Hayakawa, M. 311
Hirari, M. 311
Hong, M.-C. 137
Huet, C. 189, 385
Ilow, J. 264
Im, S. 98, 152
Inouye, Y. 204
Ioup, G. 400, 405
Ioup, J. 400, 405
Iskander, D. 463
Jackson, P. 400
Jacovitti, G. 244, 434
Jeon, B. 248
Joshi, S. 380
Jouny, I. 34
Kaiser, T. 280, 306
Kammeyer, K.-D. 62
Karasaridis, A. 268
Katsaggelos, A. 137

Kim, H. 248
Kim, J. 127
Kim, Y. 127
Korze, D. 331
Krieger, G. 147
Lacaze, B. 448
Lacoume, J. 453
Lacoume, J.-L. 47, 344
Larzabal, P. 458
Le Carpentier, E. 234
Le Martret, C. 47
Le Roux, J. 189, 385
Leporini, D. 430
Leung, H. 264, 336
Leyman, A. 9, 112
Li, T.-H. 108
Liang, Y.-C. 9, 112
Lii, K.-S. 108
Low, A. 34
Macchi, O. 295
Malouche, Z. 295
Marchand, P. 47
McLaughlin, S. 184, 390, 444
Mehta, S. 98
Mendel, J. 280, 306
Messer, H. 20
Moreau, E. 368
Nam, S. 127
Nandi, A. 219
Ong, H.-T. 463
Panas, S. 4
Park, I.-S. 98
Penman, J. 67
Persson, L. 14
Pesquet, J.-C. 430
Pesquet-Popescu, B. 458
Petermann, T. 62
Petropulu, A. 72, 224, 229
Pflug, L. 400, 405
Pierce, R. 420
Porat, B. 194
Powers, E. 98, 157
Pozidis, H. 224
Prieur, P. 82
Redfern, A. 162
Regazzoni, C. 259
Sadler, B. 273
Sato, T. 204
Scarano, G. 244, 359, 434
Scharf, L. 88
Schmidt, C. 219
Servière, C. 321, 344
Shcherbakov, M. 117
Shen, F. 24
Shen, M. 24, 29
Shor, G. 20
Solinsky, J. 209
Soong, B.-H. 9
Stathaki, T. 132, 137
Sun, L. 29
Sung, K.-M. 248
Swami, A. 273
Tesei, A. 259
Tourneret, J. 448
Tourneret, J.-Y. 410
Tseng, C.-H. 142
Tugnait, J. 239, 300
Turner, K. 52
Vandewalle, J. 316
Vincent, C. 82
Vuattoux, J.-L. 234
Wang, Y. 103
Weston, M. 39
White, P. 39
Wu, H. 290
Wu, Y. 184
Yamani, A. 214
Yan, J. 77
Yang, K. 290
Yang, T. 248
Yi, E.-J. 98
Yin, J. 229
Yoo, H. 157
Zazula, D. 331
Zetzsche, C. 147
Zhang, X.-D. 112

Zhou, G. 103, 162
Zhou, Y. 336
Zoubir, A. 438, 463
Zozor, S. 166

Press Activities Board

Vice President:
I. Mark Haas
Managing Partner
Haas Associates
P.O. Box 451177
Garland, TX 75045-1177
m.haas@computer.org

Jon T. Butler, Naval Postgraduate School
James J. Farrell III, Motorola
Mohamed E. Fayad, University of Nevada
I. Mark Haas, Haas Associates
Ronald G. Hoelzeman, University of Pittsburgh
Gene F. Hoffnagle, IBM Corporation
John R. Nicol, GTE Laboratories
Yale N. Patt, University of Michigan
Benjamin W. Wah, University of Illinois
Ronald D. Williams, University of Virginia

Editor-in-Chief
Advances in Computer Science and Engineering Board
Pradip Srimani
Colorado State University
Dept. of Computer Science
601 South Hows Lane
Fort Collins, CO 80525
Phone: 970-491-5862 FAX: 970-491-2466
srimani@cs.colostate.edu

Editor-in-Chief
Practices for Computer Science and Engineering Board
Mohamed E. Fayad
Computer Science, MS/171
Bldg. LME, Room 308
University of Nevada
Reno, NV 89557
Phone: 702-784-4356 FAX: 702-784-1833
fayad@cs.unr.edu

IEEE Computer Society Executive Staff
T. Michael Elliott, Executive Director
Matthew S. Loeb, Publisher

IEEE Computer Society Publications

The world-renowned Computer Society publishes, promotes, and distributes a wide variety of authoritative computer science and engineering texts. These books are available in two formats: 100 percent original material by authors preeminent in their field who focus on relevant topics and cutting-edge research, and reprint collections consisting of carefully selected groups of previously published papers with accompanying original introductory and explanatory text.

Submission of proposals: For guidelines and information on Computer Society books, send e-mail to cs.books@computer.org or write to the Acquisitions Editor, IEEE Computer Society, P.O. Box 3014, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1314. Telephone +1 714-821-8380. FAX +1 714-761-1784.

IEEE Computer Society Proceedings

The Computer Society also produces and actively promotes the proceedings of more than 130 acclaimed international conferences each year in multimedia formats that include hard and softcover books, CD-ROMs, videos, and on-line publications.

For information on Computer Society proceedings, send e-mail to cs.books@computer.org or write to Proceedings, IEEE Computer Society, P.O. Box 3014, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1314. Telephone +1 714-821-8380. FAX +1 714-761-1784.

Additional information regarding the Computer Society, conferences and proceedings, CD-ROMs, videos, and books can also be accessed from our web site at http://computer.org/cspress

4/15/97