

Handbook of Blind Source Separation

Independent Component Analysis and Applications

Edited by P. Comon and C. Jutten



ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First edition 2010

Copyright © 2010 Pierre Comon and Christian Jutten. All rights reserved

Q1

The rights of Pierre Comon and Christian Jutten to be identified as the authors' of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988

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

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

Notices

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

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

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

British Library Cataloguing in Publication Data

Comon, Pierre.

Handbook of blind source separation: independent component analysis and blind deconvolution.

1. Blind source separation.

I. Title II. Jutten, Christian.

621.3'8223-dc22

Library of Congress Control Number: 2009941417

ISBN: 978-0-12-374726-6

For information on all Academic Press publications visit our website at elsevierdirect.com

Printed and bound in the United States

09 10 11 12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

Contents

About the editors	xix
Preface	xxi
Contributors	xxiii
CHAPTER 1 Introduction	1
1.1 Genesis of blind source separation	1
1.1.1 A biological problem	3
1.1.2 Contextual difficulties	6
1.1.3 A few historical notes	7
1.2 Problem formalization	10
1.2.1 Invertible mixtures	11
1.2.2 Underdetermined mixtures	11
1.3 Source separation methods	11
1.3.1 Independent component analysis	12
1.3.2 Non-temporally iid sources	12
1.3.3 Other approaches	13
1.4 Spatial whitening, noise reduction and PCA	13
1.5 Applications	15
1.6 Content of the handbook	15
References	19
CHAPTER 2 Information	23
2.1 Introduction	23
2.2 Methods based on the mutual information	24
2.2.1 Mutual information between random vectors	24
2.2.2 The mixing model and separation criterion	25
2.2.3 Empirical criteria and entropy estimators	26
2.2.4 Computation of entropy estimators	29
2.2.5 Minimization of the empirical criteria	32
2.2.6 Statistical performance	39
2.3 Methods based on the mutual information rate	45
2.3.1 Entropy and mutual information rate	45
2.3.2 Contrasts	48
2.3.3 Estimation and separation	53

vi Contents

2.4	Conclusion and perspectives	61
	References.....	62
CHAPTER 3	Contrasts	65
3.1	Introduction	65
3.1.1	Model and notation	66
3.1.2	Principle of contrast functions	66
3.1.3	Bibliographical remarks	67
3.2	Cumulants	67
3.3	MISO contrasts	69
3.3.1	MISO contrasts for static mixtures.....	69
3.3.2	Deflation principle	71
3.3.3	MISO contrasts for convolutive mixtures	72
3.4	MIMO contrasts for static mixtures	78
3.4.1	Introduction.....	78
3.4.2	Contrasts for MIMO static mixtures	79
3.4.3	Contrasts and joint diagonalization.....	87
3.4.4	Non-symmetric contrasts	89
3.4.5	Contrasts with reference signals.....	90
3.5	MIMO contrasts for dynamic mixtures	92
3.5.1	Space-time whitening.....	92
3.5.2	Contrasts for MIMO convolutive mixtures	93
3.5.3	Contrasts and joint diagonalization.....	95
3.5.4	Non-symmetric contrasts	99
3.6	Constructing other contrast criteria.....	101
3.7	Conclusion.....	102
	References.....	103
CHAPTER 4	Likelihood.....	107
4.1	Introduction: Models and likelihood	107
4.2	Transformation model and equivariance	109
4.2.1	Transformation likelihood.....	110
4.2.2	Transformation contrast	111
4.2.3	Relative variations	111
4.2.4	The iid Gaussian model and decorrelation	113
4.2.5	Equivariant estimators and uniform performance.....	115
4.2.6	Summary	116
4.3	Independence	116
4.3.1	Score function and estimating equations	117
4.3.2	Mutual information	118
4.3.3	Mutual information, correlation and... ..	119
4.3.4	Summary	121

4.4	Identifiability, stability, performance	122
4.4.1	Elements of asymptotic analysis.....	123
4.4.2	Fisher information matrix	125
4.4.3	Blind identifiability	126
4.4.4	Asymptotic performance	126
4.4.5	When the source model is wrong	127
4.4.6	Relative gradient and natural gradient	130
4.4.7	Summary	130
4.5	Non-Gaussian models	131
4.5.1	The likelihood contrast in iid models.....	131
4.5.2	Score functions and estimating equations	132
4.5.3	Gaussianity index	132
4.5.4	Cramér-Rao bound	133
4.5.5	Asymptotic performance	133
4.5.6	Adaptive scores	134
4.5.7	Algorithms	135
4.6	Gaussian models	136
4.6.1	Diagonal Gaussian models	136
4.6.2	Fisher information and diversity	137
4.6.3	Gaussian contrasts	138
4.6.4	In practice: Localized Gaussian models	140
4.6.5	Choosing the diagonalizing transform.....	141
4.7	Noisy models	142
4.7.1	Source estimation from noisy mixtures	142
4.7.2	Noisy likelihood for localized Gaussian models.	143
4.7.3	Noisy likelihood and hidden variables.....	144
4.8	Conclusion: A general view	148
4.8.1	Unity	148
4.8.2	Diversity.....	149
4.9	Appendix: Proofs.....	152
	References.....	153
CHAPTER 5	Algebraic methods after prewhitening.....	155
5.1	Introduction	155
5.1.1	Multilinear algebra	155
5.1.2	Higher-order statistics	156
5.1.3	Jacobi iteration.....	159
5.2	Independent component analysis	161
5.2.1	Algebraic formulation.....	161
5.2.2	Step 1: Prewhitening	163
5.2.3	Step 2: Fixing the rotational degrees of freedom using the higher-order cumulant	164

viii Contents

5.3	Diagonalization in least squares sense	165
5.3.1	Third-order real case.....	167
5.3.2	Third-order complex case	168
5.3.3	Fourth-order real case	169
5.3.4	Fourth-order complex case	170
5.4	Simultaneous diagonalization of matrix slices	170
5.4.1	Real case	173
5.4.2	Complex case	173
5.5	Simultaneous diagonalization of third-order tensor slices	174
5.6	Maximization of the tensor trace	174
	References.....	175
CHAPTER 6	Iterative algorithms	179
6.1	Introduction	179
6.2	Model and goal	180
6.3	Contrast functions for iterative BSS/ICA.....	181
6.3.1	Information-theoretic contrasts	181
6.3.2	Cumulant-based approximations	182
6.3.3	Contrasts for source extraction.....	184
6.3.4	Nonlinear function approximations	185
6.4	Iterative search algorithms: generalities	186
6.4.1	Batch methods	186
6.4.2	Stochastic optimization	189
6.4.3	Batch or adaptive estimates?	191
6.5	Iterative whitening	192
6.6	Classical adaptive algorithms	193
6.6.1	Hérault-Jutten algorithm	193
6.6.2	Self-normalized networks	194
6.6.3	Adaptive algorithms based on contrasts	195
6.6.4	Adaptive algorithms based on centroids	198
6.7	Relative (natural) gradient techniques	199
6.7.1	Relative gradient and serial updating	199
6.7.2	Adaptive algorithms based on the relative gradient.....	200
6.7.3	Likelihood maximization with the relative gradient	202
6.8	Adapting the nonlinearities.....	203
6.9	Iterative algorithms based on deflation	204
6.9.1	Adaptive deflation algorithm by Delfosse-Loubaton	205
6.9.2	Regression-based deflation	206
6.9.3	Deflationary orthogonalization	207

6.10	The FastICA algorithm	208
6.10.1	Introduction	208
6.10.2	Implicit adaptation of the contrast in FastICA	209
6.10.3	Derivation of FastICA as a Newton iteration	209
6.10.4	Connection to gradient methods	211
6.10.5	Convergence of FastICA	212
6.10.6	FastICA using cumulants	213
6.10.7	Variants of FastICA	214
6.11	Iterative algorithms with optimal step size	216
6.11.1	Optimizing the step size	216
6.11.2	The RobustICA algorithm	217
6.12	Summary, conclusions and outlook	220
	References	221
CHAPTER 7	Second order methods based on color	227
7.1	Introduction	227
7.2	WSS processes	228
7.2.1	Parametric WSS processes	230
7.3	Problem formulation, identifiability and bounds	232
7.3.1	Indeterminacies and identifiability	233
7.3.2	Performance measures and bounds	237
7.4	Separation based on joint diagonalization	245
7.4.1	On exact and approximate joint diagonalization	246
7.4.2	The first JD-based method	251
7.4.3	AMUSE and its modified versions	251
7.4.4	SOBI, TDSEP and modified versions	253
7.5	Separation based on maximum likelihood	260
7.5.1	The QML approach	261
7.5.2	The EML approach	266
7.5.3	The GMI approach	268
7.6	Additional issues	270
7.6.1	The effect of additive noise	270
7.6.2	Non-stationary sources, time-varying mixtures	273
7.6.3	Complex-valued sources	275
	References	276
CHAPTER 8	Convulsive mixtures	281
8.1	Introduction and mixture model	281
8.1.1	Model and notations	281
8.1.2	Chapter organization	282

x Contents

8.2	Invertibility of convolutive MIMO mixtures	283
8.2.1	General results	284
8.2.2	FIR systems and polynomial matrices	285
8.3	Assumptions	287
8.3.1	Fundamental assumptions	287
8.3.2	Indeterminacies	288
8.3.3	Linear and nonlinear sources	289
8.3.4	Separation condition	291
8.4	Joint separating methods	292
8.4.1	Whitening	292
8.4.2	Time domain approaches	294
8.4.3	Frequency domain approaches	298
8.5	Iterative and Deflation Methods	301
8.5.1	Extraction of one source	301
8.5.2	Deflation	308
8.6	Non-stationary context	309
8.6.1	Context	309
8.6.2	Some properties of cyclostationary time-series	310
8.6.3	Direct extension of the results of section 8.5 for the source extraction, and why it is difficult to implement	314
8.6.4	A function to minimize: A contrast?	318
	References.....	322
CHAPTER 9	Algebraic identification of under-determined mixtures .	235
9.1	Observation model.....	235
9.2	Intrinsic identifiability	236
9.2.1	Equivalent representations.....	236
9.2.2	Main theorem	237
9.2.3	Core equation.....	239
9.2.4	Identifiability in the 2-dimensional case	240
9.3	Problem formulation	242
9.3.1	Approach based on derivatives of the joint characteristic function.....	242
9.3.2	Approach based on cumulants	243
9.4	Higher-order tensors	247
9.4.1	Canonical tensor decomposition	248
9.4.2	Essential uniqueness	250
9.4.3	Computation.....	254
9.5	Tensor-based algorithms	255
9.5.1	Vector and matrix representations	255
9.5.2	The 2-dimensional case.....	256

9.5.3	SOBIUM family	261
9.5.4	FOOBI family	264
9.5.5	BIOME family	267
9.5.6	ALESCAF and LEMACAF	269
9.5.7	Other algorithms	270
9.6	Appendix: expressions of complex cumulants	271
	References.....	272
CHAPTER 10 Sparse component analysis		367
10.1	Introduction	367
10.2	Sparse signal representations	370
10.2.1	Basic principles of sparsity	371
10.2.2	Dictionaries	372
10.2.3	Linear transforms.....	373
10.2.4	Adaptive representations.....	374
10.3	Joint sparse representation of mixtures	374
10.3.1	Principle	375
10.3.2	Linear transforms.....	375
10.3.3	Principle of ℓ^τ minimization	377
10.3.4	Bayesian interpretation of ℓ^τ criteria	377
10.3.5	Effect of the chosen ℓ^τ criterion	379
10.3.6	Optimization algorithms for ℓ^τ criteria	381
10.3.7	Matching pursuit	385
10.3.8	Summary	386
10.4	Estimating the mixing matrix by clustering	388
10.4.1	Global clustering algorithms	389
10.4.2	Scatter plot selection in multiscale representations.....	391
10.4.3	Use of local scatter plots in the time-frequency plane.....	395
10.5	Square mixing matrix: Relative Newton method for quasi-maximum likelihood separation.....	396
10.5.1	Relative optimization framework	397
10.5.2	Newton method	398
10.5.3	Gradient and Hessian evaluation	399
10.5.4	Sequential optimization	400
10.5.5	Numerical illustrations.....	401
10.5.6	Extension of Relative Newton: blind deconvolution.....	402
10.6	Separation with a known mixing matrix	403
10.6.1	Linear separation of (over-)determined mixtures	404
10.6.2	Binary masking assuming a single active source	405
10.6.3	Binary masking assuming $M < P$ active sources	405
10.6.4	Local separation by ℓ^τ minimization	406

xii Contents

10.6.5	Principle of global separation by ℓ^τ minimization	407
10.6.6	Formal links with single-channel traditional sparse approximation	408
10.6.7	Global separation algorithms using ℓ^τ minimization	408
10.6.8	Iterative global separation: demixing pursuit	409
10.7	Conclusion	410
10.8	Outlook	412
	References.....	414
CHAPTER 11	Quadratic time-frequency domain methods	421
11.1	Introduction	421
11.2	Problem statement	422
11.2.1	Model and assumptions	422
11.2.2	Indeterminacies and sources estimation	422
11.2.3	Spatial whitening	423
11.2.4	A generalization to the noisy case.....	425
11.3	Spatial quadratic t - f spectra and representations.....	427
11.3.1	Bilinear and quadratic transforms	427
11.3.2	Spatial bilinear and quadratic transforms.....	427
11.3.3	(Spatial) quadratic time-frequency representations	428
11.3.4	(Spatial) bilinear and quadratic time-frequency spectra	430
11.3.5	Descriptions of key properties and model structure, additional assumptions about the sources	430
11.3.6	Example	432
11.4	Time-frequency points selection	435
11.4.1	Automatic time-frequency points selection in a whitened context.....	435
11.4.2	Automatic time-frequency points selection in a non-whitened context.....	438
11.5	Separation algorithms.....	440
11.5.1	Joint diagonalization and/or joint zero-diagonalization criteria	441
11.5.2	Whitened-based separation algorithms	443
11.5.3	Non-whitened based separation algorithms	445
11.5.4	Algebraic methods and classification.....	449
11.6	Practical and computer simulations	452
11.6.1	Synthetic source signals	452
11.6.2	Mixture	453
11.6.3	Time-frequency points selection	455
11.6.4	Results	462
11.7	Summary and conclusion.....	463
	References.....	463

CHAPTER 12 Bayesian approaches	467
12.1 Introduction	467
12.2 Source separation forward model and notations.....	468
12.3 General Bayesian scheme	470
12.4 Relation to PCA and ICA	471
12.5 Prior and likelihood assignments	477
12.5.1 General assignments.....	478
12.5.2 Physical priors	479
12.6 Source modeling.....	482
12.6.1 Modeling stationary white sources.....	482
12.6.2 Accounting for temporal correlations of the sources.....	486
12.6.3 Modeling non-stationary sources.....	488
12.7 Estimation schemes	493
12.8 Source separation applications	494
12.8.1 Spectrometry.....	494
12.8.2 Source separation in astrophysics	494
12.8.3 Source separation in satellite imaging.....	495
12.8.4 Data reduction, classification and separation in hyperspec- tral imaging.....	497
12.9 Source characterization	499
12.9.1 Source separation and localization	499
12.9.2 Neural source estimation	501
12.9.3 Source characterization in biophysics.....	504
12.10 Conclusion.....	508
References.....	509
CHAPTER 13 Non-negative mixtures	515
13.1 Introduction	515
13.2 Non-negative matrix factorization.....	515
13.2.1 Simple gradient descent	517
13.2.2 Multiplicative updates.....	518
13.2.3 Alternating least squares (ALS).....	520
13.3 Extensions and modifications of NMF	521
13.3.1 Constraints and penalties	521
13.3.2 Relaxing the non-negativity constraints	525
13.3.3 Structural factor constraints	526
13.3.4 Multi-factor and tensor models.....	528
13.3.5 ALS Algorithms for non-negative tensor factorization	532
13.4 Further non-negative algorithms.....	533
13.4.1 Neural network approaches	534
13.4.2 Geometrical methods	535
13.4.3 Algorithms for large-scale NMF problems	537

xiv Contents

13.5	Applications	539
13.5.1	Air quality and chemometrics	539
13.5.2	Text analysis	539
13.5.3	Image processing	540
13.5.4	Audio analysis	541
13.5.5	Gene expression analysis	541
13.6	Conclusions	541
	References	542
CHAPTER 14	Nonlinear mixtures	549
14.1	Introduction	549
14.2	Nonlinear ICA in the general case	550
14.2.1	Nonlinear independent component analysis (ICA)	550
14.2.2	Definitions and preliminary results	550
14.2.3	Existence and uniqueness of transforms preserving independence	551
14.3	ICA for constrained nonlinear mixtures	554
14.3.1	Structural constraints	554
14.3.2	Smooth transforms	555
14.3.3	Example of linear mixtures	557
14.3.4	Conformal mappings	557
14.3.5	Post-nonlinear (PNL) mixtures	558
14.3.6	Bilinear mixtures	561
14.3.7	A class of separable nonlinear mappings	563
14.4	Priors on sources	567
14.4.1	Bounded sources in PNL mixtures	567
14.4.2	Temporally correlated sources in nonlinear mixtures	568
14.5	Independence criteria	570
14.5.1	Mutual information	570
14.5.2	Differential of the mutual information	573
14.5.3	Quadratic criterion	574
14.6	A Bayesian approach for general mixtures	575
14.6.1	The nonlinear factor analysis (NFA) method	576
14.6.2	Extensions and experimental results	579
14.6.3	Comparisons on PNL mixtures	579
14.7	Other methods and algorithms	580
14.7.1	Algorithms for PNL mixtures	580
14.7.2	Constrained MLP-like structures	580
14.7.3	Other approaches	581
14.8	A few applications	581
14.8.1	Chemical sensors	582

14.8.2	Gas sensors	582
14.8.3	Mixtures of images.....	583
14.9	Conclusion.....	584
	References.....	586
CHAPTER 15 Semi-blind methods for communications.....		593
15.1	Introduction	593
15.1.1	Blind source separation and channel equalization	593
15.1.2	Goals and organization of the chapter.....	594
15.2	Training-based and blind equalization.....	595
15.2.1	Training-based or supervised equalization.....	595
15.2.2	Blind equalization	595
15.2.3	A classical blind criterion: the constant modulus.....	596
15.3	Overcoming the limitations of blind methods.....	597
15.3.1	Algebraic solutions	597
15.3.2	Multi-channel systems.....	598
15.3.3	Semi-blind approach.....	598
15.4	Mathematical formulation.....	599
15.4.1	Signal model.....	599
15.4.2	Notations.....	600
15.5	Channel equalization criteria.....	601
15.5.1	Supervised, blind and semi-blind criteria	601
15.5.2	Relationships between equalization criteria	602
15.6	Algebraic equalizers	604
15.6.1	Algebraic MMSE equalizer	604
15.6.2	Algebraic blind equalizers	605
15.6.3	Algebraic semi-blind equalizers	610
15.7	Iterative equalizers	610
15.7.1	Conventional gradient-descent algorithms	610
15.7.2	Algorithms based on algebraic optimal step size.....	612
15.8	Performance analysis.....	616
15.8.1	Performance of algebraic blind equalizers.....	616
15.8.2	Attraction basins of blind and semi-blind CP equalizers....	618
15.8.3	Robustness of optimal step-size CM equalizers to local extrema	619
15.8.4	CP equalizers for a non-minimum phase channel	621
15.8.5	Blind CM and semi-blind CM-MMSE equalizers	624
15.8.6	Influence of pilot-sequence length.....	626
15.8.7	Influence of the relative weight between blind and super- vised criteria.....	627
15.8.8	Comparison between the CM and CP criteria.....	627

xvi Contents

15.9	Semi-blind channel estimation	628
15.10	Summary, conclusions and outlook	632
	References.....	633
CHAPTER 16	Overview of source separation applications	639
16.1	Introduction	639
16.1.1	Context	639
16.1.2	Historical survey	639
16.1.3	Organization of this chapter.....	641
16.2	How to solve an actual source separation problem?.....	642
16.2.1	Blind or semi-blind?	642
16.2.2	ICA for BSS	643
16.2.3	Practical use of BSS and associated issues	644
16.3	Overfitting and robustness	645
16.3.1	Overfitting	645
16.3.2	Robustness	646
16.4	Illustration with electromagnetic transmission systems	648
16.4.1	A variety of source natures and mixture configurations	648
16.4.2	A case study on radio-frequency identification (RFID)	649
16.4.3	A system with multi-path ionospheric propagation.....	655
16.4.4	Using other signal properties	658
16.5	Example: Analysis of Mars hyperspectral images.....	658
16.5.1	Physical model of hyperspectral images	658
16.5.2	Decomposition models based on ICA	660
16.5.3	Reference data and classification.....	661
16.5.4	ICA results on hyperspectral images.....	662
16.5.5	Discussion	665
16.5.6	Beyond ICA: semi-blind source separation.....	665
16.5.7	Conclusion	666
16.6	Mono- vs multi-dimensional sources and mixtures.....	668
16.6.1	Time, space and wavelength coordinates	668
16.6.2	Analyzing video frames from cortical tissues	670
16.6.3	Extracting components from a time series of astrophysical luminance images	671
16.7	Using physical mixture models or not	672
16.7.1	Mother vs fetus heartbeat separation from multi-channel ECG recordings.....	672
16.7.2	Analysis of heart control from single-channel ECG.....	674
16.7.3	Additional comments about performance evaluation	676
16.8	Some conclusions and available tools.....	676
	References.....	677

CHAPTER 17	Application to telecommunications	683
17.1	Introduction	683
17.2	Data Model, Statistics and Problem Formulation	687
17.2.1	Observation model	687
17.2.2	Data statistics	690
17.2.3	Formulation of the problem	696
17.3	Possible Methods	696
17.3.1	Treating the mixture as a convolutive one.....	696
17.3.2	Treating the mixture as an instantaneous one.....	709
17.4	Ultimate separators of instantaneous mixtures	712
17.4.1	Source separator performance	712
17.4.2	Ultimate separator.....	713
17.4.3	Ultimate performance.....	714
17.5	Blind separators of instantaneous mixtures	716
17.5.1	JADE for stationary uncorrelated paths.....	716
17.5.2	JADE for cyclostationary uncorrelated paths.....	718
17.5.3	JADE for cyclostationary correlated paths.....	723
17.5.4	Performance illustration	724
17.6	Instantaneous approach versus convolutive approach: simulation results.....	726
17.6.1	BSS algorithms and measures of performance	726
17.6.2	Performances.....	727
17.7	Conclusion.....	729
	References.....	730
CHAPTER 18	ICA and biomedical applications.....	737
18.1	Introduction	737
18.2	One decade of ICA-based biomedical data processing	739
18.2.1	Electromagnetic recordings for functional brain imaging... ..	739
18.2.2	Electrocardiogram signal analysis	746
18.2.3	Other application fields	757
18.3	Numerical complexity of ICA algorithms	758
18.3.1	General tools	759
18.3.2	Complexity of several ICA algorithms	760
18.4	Performance analysis on biomedical signals	763
18.4.1	Comparative performance analysis on synthetic signals	764
18.4.2	ICA of real data.....	771
18.5	Conclusion.....	772
	References.....	772

xviii Contents

CHAPTER 19 Audio applications	779
19.1 Audio mixtures and separation objectives.....	779
19.1.1 Recorded mixtures.....	780
19.1.2 Synthesized mixtures	782
19.1.3 Separation objectives and performance evaluation	783
19.2 Usable properties of audio sources	787
19.2.1 Independence	787
19.2.2 Sparsity	788
19.3 Audio applications of convolutive ICA	790
19.3.1 Multichannel filtering	790
19.3.2 Time-domain convolutive ICA	794
19.3.3 Frequency-domain convolutive ICA.....	799
19.4 Audio applications of SCA	806
19.4.1 Time-frequency masking.....	807
19.4.2 Instantaneous SCA	808
19.4.3 Convolutive SCA	810
19.5 Conclusion.....	814
References.....	815
Glossary	821
Index.....	823

About the editors

The two editors are pioneering contributors of ICA. They wrote together the first journal paper on ICA, which appeared in *Signal Processing*, published by Elsevier in 1991, and received a best paper award in 1992, together with J. Héroult.

Pierre Comon is Research Director with CNRS, Lab. I3S, University of Nice, France. He has been Associate Editor of the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Circuits and Systems I*, in the area of blind techniques. He is now Associate Editor of the *Signal Processing* journal, published by Elsevier. He has been the coordinator of the European network “ATHOS” on High-Order Statistics. He received the Monpetit prize from the French Academy of Sciences in 2005 (rewarding works with industrial applications), and the Individual Technical Achievement Award from Eurasip in 2006. He is Fellow of the IEEE, Emeritus Member of the SEE, and member of SIAM. He authored a paper in 1994 on the theoretical foundations of ICA; this paper still remains among the most cited both on the subject of ICA and/or blind techniques, in the whole signal processing community.

Christian Jutten is Professor at the University Joseph Fourier of Grenoble, France. He is currently associate-director of GIPSA-lab, a 300-people laboratory focused on automatic control, signal, images and speech processing. He has been Associate Editor of the *IEEE Transactions on Circuits and Systems I*, in the area of Neural Networks and Signal Processing techniques. He is currently Associate Editor of *Neural Processing Letters*, published by Kluwer. He was the co-organiser of the first international conference on Blind Source Separation and Independent Component Analysis, in 1999 (ICA 99). He was the coordinator of two European projects, one of them (BLISS) focused on Blind Source Separation and Applications. He received the Blondel Medal of the SEE in 1997 for his contributions in blind source separation. He is Fellow of the IEEE and Senior Member of Institut Universitaire de France. He co-authored a set of two papers in 1991, with J. Héroult and P. Comon, on the first algorithm for blind source separation and on theoretical foundation, which still remains in the top five papers cited on the subject of ICA and/or blind techniques, and in the whole signal processing community.

Preface

In signal processing, a generic problem consists in separating a useful signal from noise and interferences. Classical approaches of the twentieth century are based on *a priori* hypotheses, leading to parameterized probabilistic models. Blind Source Separation (BSS) attempts to reduce these assumptions to the weakest possible.

As shown in this handbook, there are various approaches to the BSS problem, depending on the weak *a priori* hypotheses one assumes. The latter include either statistical independence of source signals or their sparsity, among others.

In order to prepare this book, among the best worldwide specialists have been contacted to contribute (cf. page xxiii). One of them, Serge Degerine, has passed away unexpectedly during the writing of Chapter 7. We would like to dedicate this book to his memory. o1

This handbook is an extension of another book which appeared in 2007 in French, and published by Hermes. The present version contains more chapters and many additions, provided by contributors with international recognition. It is organized into 19 chapters, covering all the current theoretical approaches, especially Independent Component Analysis, and applications. Although these chapters can be read almost independently, they share the same notations and the same subject index. Moreover, numerous cross-references link the chapters to each other.

Pierre Comon and Christian Jutten

Glossary

\mathbf{x}	vector of components x_p , $1 \leq p \leq P$
s, x, y	sources, observations, separator outputs
N	number of sources
P	number of sensors
T	number of observed samples
$*$	convolution
\mathbf{A}	matrix with components A_{ij}
\mathbf{A}, \mathbf{B}	mixing and separation matrices
$\mathbf{G}, \mathbf{W}, \mathbf{Q}$	global, whitening, and separating unitary matrices
\check{g}	Fourier transform of g
\hat{s}	estimate of quantity s
p_x	probability density of \mathbf{x}
ψ	joint score function
φ_i	marginal score function of source s_i
Φ	first characteristic function
Ψ	second characteristic function
$\mathbb{E}\mathbf{x}, \mathbb{E}\{\mathbf{x}\}$	mathematical expectation of \mathbf{x}
$I\{\mathbf{y}\}$ or $I(p_y)$	mutual information of \mathbf{y}
$K\{\mathbf{x}; \mathbf{y}\}$ or $K(p_x; p_y)$	Kullback divergence between p_x and p_y
$H\{\mathbf{x}\}$ or $H(p_x)$	Shannon entropy \mathbf{x}
\mathcal{L}	likelihood
\mathcal{A}, \mathcal{B}	mixing, and separating (nonlinear) operators
$\text{cum}\{x_1, \dots, x_p\}$	joint cumulant of variables $\{x_1, \dots, x_p\}$
$\text{cum}_R\{y\}$	marginal cumulant of order R of variable y

822 Glossary

Q^T	transposition
Q^H	conjugate transposition
Q^*	complex conjugation
Q^\dagger	pseudo-inverse
Υ	contrast function
\mathbb{R}	real field
\mathbb{C}	complex field
\hat{A}	estimator of mixing matrix
$\text{diag } \mathbf{A}$	vector whose components are the diagonal of matrix \mathbf{A}
$\text{Diag } \mathbf{a}$	diagonal matrix whose entries are those of vector \mathbf{a}
$\text{trace } \mathbf{A}$	trace of matrix \mathbf{A}
$\det \mathbf{A}$	determinant of matrix \mathbf{A}
$\text{mean } \mathbf{a}$	arithmetic average of component of vector \mathbf{a}
$\check{s}(\nu)$	Fourier transform of process $s(t)$
\otimes	Kronecker product between matrices
\otimes	tensor product
\bullet_j	contraction over index j
$\text{krank}\{\mathbf{A}\}$	Kruskal's k -rank of matrix \mathbf{A}

Index

A

- additivity of cumulants, 336
- affine class, 428
- air traffic control, 649
- AJD, 245, 247
- ALGECAF, 349
- ALGECUM, 348
- algorithm
 - accelerating adaptive filtering constant modulus (AAF-CMA), 616
 - adaptive, neural, on-line, recursive, stochastic, 189, 192, 193, 199, 612
 - ALESCAF, 359
 - algebraic, 597
 - algebraic constant modulus (ACMA), 624
 - algebraic constant power (ACPA), 604, 605, 616, 621
 - algebraic semi-blind constant power (SB-ACPA), 627
 - ALGECAF, 349
 - ALGECUM, 348
 - ALS, 538
 - AMUSE, 251
 - analytical constant modulus (ACMA), 597
 - batch vs. adaptive, 191
 - batch, block, off-line, windowed, 186, 192, 208, 217, 612
 - BIOME, 357
 - BIRTH, 357
 - clustering, 368
 - COM1, 174
 - COM2, 165, 692
 - complex FastICA, 215
 - complex fixed-point (CFPA), 216
 - constant modulus (CMA), 611, 619, 630, 699
 - constant power (CPA), 611, 618
 - efficient ICA, 204
 - ELS, 360
 - equivariant adaptive separation via independence (EASI), 201
 - exact line search, 612
 - expectation-maximization (EM), 144
 - FastICA, 208, 663
 - flexible ICA, 204
 - FOOBI, 356
 - FOOBI2, 357
 - geometrical, 535, 568
 - greedy, 385
 - hierarchical ALS, 538
 - InfoMax, 196
 - JADE, 170, 663, 692
 - kurtosis maximization fixed-point (KM-F), 216
 - M-FOCUSS, 382
 - NFA, 579
 - NMF, 533
 - non-circular FastICA (nc-FastICA), 216
 - non-negative, 533
 - non-unitary joint diagonalization, 447
 - optimal step size, 612
 - optimal step-size constant modulus (OS-CMA), 614, 619, 624, 626
 - optimal step-size constant power (OS-CPA), 614, 618, 621
 - optimal step-size kurtosis maximization (OS-KMA), 615, 616
 - optimal step-size semi-blind constant modulus (OS-SB-CMA), 614, 624, 626, 627
 - optimal step-size semi-blind constant power (OS-SB-CPA), 614, 619, 621, 627
 - QR constant modulus (QR-CMA), 616
 - recursive least squares constant modulus (RLS-CMA), 616, 619
 - relative gradient, 112
 - relative-gradient maximum likelihood, 202
 - RobustICA, 217
 - semi-blind algebraic constant modulus (SB-ACMA), 599, 624, 626, 627
 - semi-blind algebraic constant power (SB-ACPA), 610, 621, 627

824 Index

- semi-blind constant modulus (SB-CMA), 611
- semi-blind constant power (SB-CPA), 611, 619
- SOBIUM, 354
- stabilized FastICA, 215
- stochastic-gradient constant modulus (SG-CMA), 612, 616
- STOTD, 174
- ALS, 520, 532, 538
- alternative least square (ALS), 520
- ambiguity factors, 233
- AMUSE, 251
- application, 639
 - air quality, 539
 - astrophysics, 658, 671
 - audio frequency, 779
 - audio processing, 541
 - biomedical, 670, 672, 737
 - chemistry, 539
 - image processing, 540, 658
 - music, 779
 - telecommunication, 649, 683
 - text analysis, 539
- AR, 228
- ARMA, 228
- artifact removal, 748
- artifacts, 784
- astrophysics, 658, 671
- atoms, 370
- atrial activity extraction, 748
- atrial fibrillation (AF), 752
- auto-terms, 428

- B**
- Basis Pursuit, 383
- Basis Pursuit Denoising, 383
- Bayes theorem, 470
- Bayesian approach, 467, 575, 576, 665
- beamforming, 790
- bilinear model, 583
- bilinear transform, 427
- binary masking, 807
- binning, 29, 30
- BIOME, 357
- biomedical applications, 737
- BIRTH, 357
- blind, 642
 - identifiability, 126
- blind deconvolution, 593, 599
- blind equalization, 593
- blind identification, 65

- blind source separation (BSS), 643
- blind techniques, 1
- bracket notation, 334
- BSS, 643
- bumps, 645

- C**
- CanDecomp, CanD, 338, 530
- Canonical Decomposition (CanDecomp), 162, 338
- canonical factorization, 530
- cardinal spline, 29
- causal, 48
- CDMA, 630
- central moment, 333
- centroid, 198
- channel coding, 630
- channel equalization, 599
- channel state information at the transmitter, 631
- characteristic function, 68
 - estimated, 27
 - second, 329
- chemical sensors, 582
- Cholesky factorization, 399
- circular, 275, 702
- circular cumulant, 336
- circularity, 603
- circularized density estimate, 27
- clustering, 368
- co-channel interference (CCI), 593
- Cohen class, 428
- colored sources, 643
- COM1, 174, 738
- COM2, 165, 738
- conditional entropy, 45
- conformal mapping, 557
- conic programming, 383
- constant modulus (CM), 67
- contrast, 39, 65, 66, 71, 79
 - COM2, 89
 - deterministic, 102
 - JAD, 89
 - MIMO, 78, 181
 - MISO, 70, 74
 - PAJOD, 97
 - STOTD, 89
 - transformation, 111
 - with reference, 77, 90
- contrast function, 179
 - alphabet polynomial fitting (APF), 602
 - attraction basins, 618

- based on cumulants, 182
 - based on kurtosis, 183, 184, 208, 601, 603, 615
 - constant modulus (CM), 594, 596
 - constant power (CP), 594, 601
 - for signal extraction, 184
 - InfoMax, 181
 - marginal entropy, 182
 - maximum likelihood, 86, 181
 - mutual information, 81, 182
 - nonlinear approximations, 185
 - orthogonal, 183, 208
 - convergence
 - global, 184, 205, 213
 - convolutive, 281
 - NMF model, 526
 - convolutive mixture, 11, 793
 - convolutive post-nonlinear model (CPNL), 560
 - core equation, 329
 - correlated sources, 102
 - correlation matrix, 424
 - cortical tissue imaging, 670
 - CP-degeneracy, 345
 - CPM modulations, 703
 - CPNL, 560
 - Cramér-Rao bound, 127, 133, 241
 - CRLB, 241
 - cross-correlation matrix, 424
 - cross-terms, 428
 - cumulant, 157, 333, 738
 - additivity, 336
 - complex, 360
 - deductive estimation, 615
 - matching, 90
 - multivariate, 335
 - nonlinear, non-polynomial, 209
 - of complex random variables, 336
 - tensor, 335
 - cumulant matching, 337
 - cumulative distribution function (cdf), 182, 197
 - cyclic frequency, 273, 311, 684, 696
 - cyclo-correlation, 311, 313
 - cyclo-ergodicity, 692
 - cyclo-spectrum, 312
 - cyclo-stationarity, 273, 310, 658, 684
- D**
- Darmois, 39, 52, 326, 330, 551, 552, 554, 557, 563, 564, 569, 570
 - decomposability, 328
 - deconvolution
 - principle, 47
 - decorrelation, 227
 - deflation, 67, 72, 184, 204, 308
 - dimensionality reduction, 205, 207
 - orthogonalization, 207
 - regression, 206
 - delay spread, 684
 - demixing pursuit, 409
 - density, 24
 - diagonalization
 - partial, 98
 - tensor, 85
 - dictionary, 371
 - complete, 372
 - differential of mutual information, 573, 574
 - digital communications, 309, 310
 - direct path, 781
 - direction of arrival (DOA), 791
 - directivity pattern, 791
 - discrete alphabet, 597
 - diversity, 13, 17
 - induced by discrete alphabets, 102
 - DOA, 791
 - doubly normalized filters, 94
 - Dugué, 326
- E**
- ECG, 746
 - echoes, 781
 - EEG, 740
 - eigenvalue decomposition (EVD), 192
 - EJD, 245
 - electro-encephalography (EEG), 740
 - electrocardiogram analysis, 746
 - electromagnetic source, 648
 - electromyogram (EMG), 758
 - embedded pilots, 631
 - EMG, 758
 - EML, 260
 - empirical quantile function, 29
 - energetic, 428
 - enhanced line search (ELS), 360
 - entropy, 24
 - entropy rate, 45
 - equalizer
 - MIMO, 96
 - equivariance, 109, 239, 240
 - essential uniqueness, 340
 - estimated characteristic function, 27
 - estimating equations, 60, 115, 263
 - estimation
 - equivariant, 200

826 Index

- maximum likelihood, 181, 183, 198, 202, 203, 208, 211
 - minimum mean square error (MMSE), 206
 - exact line search, 612, 633
 - exact ML (EML), 260
 - excess bandwidth, 598
 - expected rank, 341
 - extraction, 66, 308
 - atrial activity, 748
 - fetal ECG, 748
 - extractor
 - MISO, 72
 - extrema
 - local, 619, 632
- F**
- FastICA, 663, 737
 - fetal ECG extraction, 672, 748
 - filter
 - spatial, 186
 - filtered Markov process, 47, 52
 - FIM, 241
 - finite alphabet, 630
 - finite impulse response (FIR), 93, 597, 598
 - Fisher information, 125, 137
 - floating point operation (flop), 615
 - fMRI, 647, 672, 739, 757
 - FOBIUM, 738
 - FOOBI, 356
 - forward model, 467, 470, 471
 - frequency offset, 691
 - function brain imaging, 739
- G**
- gamma distribution, 482
 - gas sensors, 582
 - Gaussian distribution, 557, 559, 560, 566, 569, 576, 578, 580
 - Gaussian entropy rate, 46
 - Gaussian MI (GMI), 261
 - Gaussian mutual information, 787
 - Gaussian mutual information rate, 50
 - Gaussian source, 643
 - generalized Gaussian, 789
 - generalized Gaussian (GG), 482
 - global filter, 69, 73, 79, 237
 - GMI, 261
 - gradient
 - classical, 187
 - natural, 130
 - relative, 112, 130
 - relative, natural, 199
 - stochastic, 132
 - gradient descent, 517
 - Gram-Schmidt orthogonalization, 207
- H**
- Hermitian symmetry, 430
 - Hessian matrix, 187, 399
 - hexacovariance, 345
 - hidden Markovian models, 482
 - hidden variables, 482
 - hierarchical ALS, 538
 - higher-order statistics, 630
 - history, 1, 367, 639
 - hyperspectral images, 658
- I**
- IC, 658, 662
 - interpretation of, 664
 - relevance of, 662
 - ICA, 2, 7, 12, 549, 550, 554, 559, 642, 643, 660, 737, 814
 - frequency-domain convolutive, 799
 - spatial, 660, 662
 - spectral, 660
 - time-domain convolutive, 794
 - iCRLB, 244
 - identifiability, 125, 233, 557
 - iid, 227, 560, 570, 684
 - non-temporally, 12
 - temporally, 8, 11
 - ill-posed inverse problem, 467
 - independence, 6, 642, 664, 738
 - independence criterion, 570
 - independent component (IC), 658, 662, 738
 - interpretation of, 664
 - relevance of, 662
 - independent component analysis (ICA), 2, 8, 12, 471, 643
 - independent subspace analysis (ISA), 216, 739
 - indeterminacy, 551
 - induced CRLB (iCRLB), 244
 - inference problem, 467
 - inferences, 470
 - InfoMax, 737
 - initialization, 597
 - innovation sequence, 230
 - inter-microphone intensity difference, 801
 - inter-microphone time difference, 791
 - interference, 784
 - intersymbol interference (ISI), 593, 599

- inverse filter
 - FIR, 73, 75, 93
- ISA, 739
- ISR, 234, 238
- iterative power method, 608
- iterative reweighted least squares, 382
- iterative thresholding, 383

- J**
- Jacobi iteration, 159
- JADE, 172, 663, 738
- joint
 - diagonalization, 50, 54, 87, 441, 760
 - diagonalizer, 444
 - diagonalizer/zero-diagonalizer, 445, 449
 - orthogonal diagonalization, 760
 - wide sense stationarity (JWSS), 229
 - zero-diagonalization, 441
 - zero-diagonalizer, 444
- joint entropy rate, 45
- joint score function (JSF), 55, 573
- joint sparse approximation, 377
- JSF, 573
- JWSS, 229

- K**
- k-rank, 341
- Kagan, 564, 566
- kernel, 26
 - of a set of vectors, 341
- Khatri-Rao product, 346
- KLD, 250
- KM-CM equivalence, 603
- Kronecker product, 346
- Kruskal, 341
- Kullback-Leibler divergence, 24, 86, 181, 250, 518, 577, 737
- kurtosis, 183, 184, 208, 217
 - definition, 334
 - maximization (KM), 67
 - optimal step size, 217
 - sensitivity to outliers, 185, 215

- L**
- Lagrange multiplier, 188
- Lagrangian, 188
- least squares (LS), 595
- leptokurtic, 334
- likelihood, 107, 378
 - contrast, 111
 - link with MI, 86
- line search, 217, 360, 399
- linear model, 557
- linear process, 47, 51, 73, 76, 94, 95
- linear programming, 383
- local minima, 574
- local scatter plots, 395
- log-likelihood, 260, 396
- LTI, 230

- M**
- M-FOCUSS, 382
- MA, 228
- magneto-encephalography (MEG), 740
- mapping
 - conformal, 557
 - example of nonlinear mapping, 564
 - linearizable
 - separability, 566
 - smooth, 557
- mappings
 - linearizable, 564
- Marcinkiewicz, 329
- marginal entropy, 182, 183
- marginal score function (MSF), 573
- Markovian model, 580
- matching pursuit, 385
- maximal ratio combining, 609
- maximum *a posteriori* (MAP), 378, 407, 602
- McCullagh, 334
- mean correlation matrix, 424
- mean field, 190
- MEG, 740
- mesokurtic, 334
- MI, *see* mutual information, 261
- MIMO, 1, 67
- minimum mean square error (MMSE), 595
- minimum phase, 48, 52
- misadjustment, 188
- MISO, 67
- MISO extractor, 69
- mixing matrix, 422
 - FIR, 93
- mixture
 - convolutive, 11
 - nonlinear, 11
- mixture of Gaussians, 482
- mixtures
 - spatial, 593
 - temporal, 593
- MLE, 241

828 Index

- mode- k product, 156
 - modeling the sources, 482
 - modified Yule-Walker equations, 232
 - modulation
 - binary phase shift keying (BPSK), 601
 - continuous phase (CPM), 685
 - minimum shift keying (QPSK), 602
 - phase shift keying (PSK), 594
 - pulse amplitude (PAM), 603
 - quadrature phase shift keying (QPSK), 601
 - moment, 156, 333
 - Monte Carlo (MC), 617
 - morphological diversity, 412
 - motion decoding, 3
 - MSE, 241
 - MSF, 573
 - multi-input multi-output (MIMO), 593–595
 - multi-layer NMF, 528
 - multi-layer perceptron, 555, 580
 - multi-linearity property, 335
 - multipath, 709
 - multiplicative nonlinear model, 563
 - multiplicative update, 518
 - Multivariate cumulant, 335
 - Multivariate moment, 335
 - music, 783, 803
 - mutual information, 24, 81, 83, 119, 182, 183, 197, 261, 550, 570, 573, 574, 579, 580, 585, 787
 - differential of, 573, 574
 - direct minimization of, 573
 - Gaussian, 787
 - link with likelihood, 86, 118
 - rate, 46
- N**
- natural gradient, 397, 737
 - negentropy, 82, 182, 737
 - cumulant approximation, 184
 - neural network
 - feed-forward architecture, 194
 - feedback architecture, 193
 - self-normalized, 194
 - Newton method, 398
 - NFA, 575, 576, 579, 586
 - NIFA, 579
 - NMF, 515, 642, 665
 - NMFD, 527
 - NNICA, 525
 - noise, 14
 - noise removal, 326, 748
 - non-circular, 337
 - non-Gaussianity, 132, 642
 - non-minimum phase, 596, 621
 - non-negative ICA, 515
 - non-negative ICA (NNICA), 525
 - non-negative matrix factor deconvolution (NMFD), 527
 - non-negative matrix factorization (NMF), 515, 642
 - non-negative tensor factorization, 529
 - non-stationarity, 482
 - non-stationary source, 643
 - nonlinear
 - ICA, 550
 - model, 11, 549, 554
 - process, 95, 289
 - nonlinear factor analysis (NFA), 575
 - nonlinear model
 - structural constraints, 554
 - bilinear, 583
 - multiplicative, 563
 - nonlinearity
 - adaptation, 203
 - cubic, 208
 - implicit adaptation in FastICA, 209
 - optimal, 202
 - tanh, 208
 - normalization, 423
 - nuisance parameters, 260
 - numerical complexity, 758, 760
- O**
- optimal step size, 612
 - optimal Wiener filtering, 595
 - optimization
 - constrained, 188, 209
 - global, 196, 217, 612–614
 - gradient method, 187
 - Newton method, 188
 - oracle performance, 785
 - order statistics, 29
 - ordinary differential equation (ODE) method, 189
 - outer product, 155, 338
 - over-determined, 78, 108
 - overfitting, 645
- P**
- PAJOD, 98
 - para-unitary, 92, 94, 293
 - ParaFac, 530

- parallel factor decomposition (ParaFac), 162, 338
 - parcor, 266
 - partial autocovariance, 257
 - partial correlation, 266
 - partial diagonalization, 98
 - partition of unity, 30
 - PCA, 8, 526, 534, 662, 671
 - pdf, 229
 - performance, 115, 763
 - performance criterion, 764
 - performance evaluation, 676
 - permutation factor, 233
 - permutation problem, 800
 - physical significance, 479
 - pilot or training sequence, 593, 595
 - platykurtic, 334
 - PMF, 515
 - PNL, 558–560, 566–568, 571, 579, 580, 586
 - poly-periodic (PP) function, 691
 - polynomial rooting formula
 - Cardan's, 614
 - Ferrari's, 219, 614, 615
 - positive matrix factorization (PMF), 515
 - positivity prior, 665
 - post-nonlinear model (PNL), 558, 582, 585
 - convolutive (CPNL), 560
 - separability, 559, 568
 - posterior probability, 467
 - prewhitening, 163
 - principal component analysis (PCA), 8, 471
 - prior, 643, 665
 - prior information, 467
 - prior probabilities, 467
 - probability density function (pdf), 181
 - generalized Gaussian distribution (GGD), 204
 - Pearson's system, 204
 - proper, 275
- Q**
- QML, 241, 260
 - quadratic criterion, 575
 - quadratic energetic transforms, 428
 - quadratic programming, 383
 - quadratic transform, 427
 - quadrivariance, 345, 693
 - quantile function, 28
 - Quasi ML, 260
 - quasi-disjoint, 439
- R**
- radio-frequency identification, 649
 - radio-frequency source, 649
 - RAKE, 609
 - range, 43
 - rank
 - Kruskal, 341
 - of tensor, 338
 - structured, 339
 - rank-1 approximation, 608
 - rank-1 diagonal matrix, 435
 - rank-1 linear combination problem, 607
 - reduced columns, 75
 - relative gradient, 32, 199, 397
 - relative Hessian, 32
 - relative Newton, 396
 - relative optimization, 397
 - relaxation, 267
 - relevant priors, 479
 - representation
 - equivalent, 327
 - reverberation time, 781
 - RobustICA, 217
 - robustness, 127, 645, 646
 - stability, 129
- S**
- SAR, 785
 - SCA, 642, 814
 - scale factor, 234
 - scatter plot, 367
 - Schur decomposition, 597, 599
 - score function, 33, 117, 202, 571, 572, 574, 586
 - approximation, 203
 - joint, 573
 - marginal, 573
 - score function difference (SFD), 573, 574
 - SDR, 785
 - semi-blind, 642, 665
 - semi-blind methods, 593, 598
 - semi-NMF, 525
 - semi-unitary, 99
 - separating functions, 60
 - separation, 67
 - separation matrix, 422
 - serial updating, 200
 - SFD, 573
 - shift invariance of cumulants, 336
 - side information, 630

830 Index

- sign factor, 234
 - signal extraction, 184, 204
 - signal model, 470
 - signal separation
 - deflationary, 184, 204
 - joint or symmetric, 184
 - signal subspace, 14
 - signal-to-artifacts ratio (SAR), 785
 - signal-to-distortion ratio (SDR), 785
 - signal-to-interference ratio (SIR), 785
 - signal-to-noise ratio (SNR), 617, 785
 - single-input multi-output (SIMO), 595, 598
 - single-input single-output (SISO), 595, 596, 696
 - singular value decomposition (SVD), 192, 607
 - SIR, 785
 - SISO, 1
 - skewness, 334
 - Skitovic, *see* Darmois, 330
 - smoothing method of multipliers (SMOM), 401
 - smoothness, 523
 - SNR, 785
 - SOBI, 253, 738
 - SOBIUM, 738
 - SOS, 227
 - source
 - bounded, 567
 - bounded pdf, 536
 - colored, 93, 643
 - correlated, 102
 - electromagnetic, 648
 - Gaussian, 643
 - music, 787
 - non-stationary, 643
 - radio-frequency, 649
 - speech, 787
 - temporally correlated, 568
 - Darmois decomposition, 569
 - example, 569
 - source signals
 - sub-Gaussian, 183
 - super-Gaussian, 183
 - source-microphone impulse response, 781
 - spark, 341
 - sparse component analysis (SCA), 642
 - sparse representation
 - joint, 377
 - truly, 379
 - sparsity, 521
 - spatial
 - bilinear time-frequency spectrum, 430
 - bilinear transform, 427
 - quadratic time-frequency spectrum, 430
 - quadratic transform, 428
 - whitening matrix, 424
 - spatial aliasing, 801
 - spatial coherence, 14, 809
 - spatial diversity, 412
 - spatial ICA, 660, 662, 756
 - spatial whitening
 - of the observations, 423
 - spatial whitening matrix, 80
 - spatio-temporal equalization, 594
 - spatio-temporal ICA, 756
 - spectral ICA, 660
 - spectrum, 228
 - specular channel, 697
 - speech, 783, 803
 - sphering, 227
 - spurious independent component, 645
 - spurious solutions, 596, 619
 - standardization, 80
 - stationary point
 - definition, 187, 190
 - local asymptotic stability, 190
 - maximum, 187
 - spurious, 184, 193, 195, 205
 - step size, 188
 - optimal, 216, 217
 - STOTD, 174
 - strong uncorrelating transform, 276
 - structural constraints, 554
 - subspace fitting, 607
 - subspace methods, 607
 - sufficient statistic, 256
 - superimposed pilots, 631
 - SUT, 276
 - Sylvester theorem, 346
 - symbol error rate (SER), 617
 - symmetric rank, 339
- T**
- target-matrices, 247
 - TDSEP, 251
 - tensor, 335
 - factorization, 521
 - rank, 338
 - symmetric, 336, 338, 339
 - thresholding function, 383
 - time multiplexed pilots, 631
 - time-reversibility, 257
 - transform
 - σ -diagonal, 550, 551, 557

- preserving independence, 551
 - Darmonis method, 553
 - example, 552
 - existence, 552
- smooth, 555
 - counterexample, 555
- trivial filter, 69, 78, 93

U

- under-determined, 11, 16, 72, 325, 368, 642, 658, 814
- uniform performance, 199, 200
- unimodular, 75
- uniqueness
 - essential, 327

V

- ventricular arrhythmia detection, 748

W

- whitened observation, 181, 424
- whitening, 14, 92, 113, 180, 227, 292
 - iterative, 192
 - matrix, 180
 - spatial, 80
- Wiener filter, 271
- Wiener systems, 560
- word error rate, 785
- WSS, 227

Y

- Yule-Walker equations, 231

Z

- zero forcing (ZF), 597