Identification Autodidacte de Mélanges Potentiellement Sous-Déterminés

présentée par

Laurent Albera

Rapport de thèse

Pour obtenir le grade de

Docteur Ès Sciences de l'Université de Nice Sophia-Antipolis

Spécialité : Traitement du Signal

Membres du jury :

Président : Philippe Loubaton

Directeurs de thèse : Pascal Chevalier, Pierre Comon et Anne Ferréol

Rapporteurs : Christian Jutten et Philip Regalia

Examinateur : Lieven De Lathauwer

Université de Nice Sophia-Antipolis

Décembre 2003

THALES COMMUNICATIONS

Identification Autodidacte de Mélanges Potentiellement Sous-Déterminés

Laurent Albera

Université de Nice Sophia-Antipolis, 2003

La séparation aveugle de sources et plus particulièrement l'Analyse en Composantes Indépendantes (ICA) ont récemment suscité beaucoup d'intérêt. Elles trouvent en effet leur place dans un grand nombre d'applications telles que les télécommunications, le traitement de la parole, l'analyse de données, ou bien le domaine biomédical. Le principe de la séparation autodidacte (ou aveugle) de sources est de restituer les sources émises, et ce, uniquement à partir des observations issues des capteurs. Alors que certaines techniques cherchent à décorréler (à l'ordre 2) les signaux, comme on peut l'observer en Analyse Factorielle avec l'Analyse en Composantes Principales (PCA), l'ICA pour sa part vise à réduire les dépendances statistiques des signaux aux ordres supérieurs, et permet de cette manière de restituer les sources. Les méthodes proposées sont donc dédiées de préférence aux sources indépendantes statistiquement. Selon l'application, on peut toutefois choisir de ne retrouver que les paramètres du mélange instantané, ce qui est utile en goniométrie car le dit mélange porte à lui seul toute l'information nécessaire à la localisation des sources : on parle alors d'*identification aveugle* de mélange. Pour d'autres applications telles que la transmission, il est nécessaire de retrouver les sources émises : on emploie alors l'expression de *séparation* ou bien encore d'*extraction aveugle* de sources. De plus, alors que divers algorithmes, très performants notamment sous l'hypothèse de bruit gaussien spatialement et temporellement blanc, permettent déjà depuis une dizaine d'années de traiter le cas de mélanges dits *surdéterminés* (c'est-à dire lorsque le nombre de sources est inférieur au nombre de capteurs), le cas de mélanges dits *sous-déterminés* (c'est-à dire lorsque le nombre de sources est strictement supérieur au nombre de capteurs) a été jusqu'à présent peu étudié en dépit des nombreuses applications. Les travaux de thèse ont alors permis d'élaborer une famille, BIOME, de nouvelles méthodes statistiques de séparation aveugle de sources, d'une part traitant le problème du bruit gaussien de cohérence

spatiale inconnue, d'autre part permettant l'identification autodidacte du mélange y compris en contexte sous-déterminé. Par ailleurs, une étude asymptotique de performances des méthodes basées sur la maximisation des contrastes d'ordre 4 a pu être menée dans le cas de mélanges orthonormés. Enfin, le comportement, en présence de signaux cyclostationnaires potentiellement non centrés, des méthodes de séparation aveugle de sources exploitant les statistiques d'ordre 2 et/ou 4 a également pu être étudié et des améliorations ont alors été proposées.

A mes parents...

Acknowledgments

Ma pensée va tout particulièrement à mes trois directeurs de thèse, Pierre Comon, directeur de recherche CNRS au laboratoire I3S de Sophia-Antipolis, Anne Ferréol et Pascal Chevalier, experts techniques à THALES Communications de Colombes. Vous m'avez, de part vos conseils et vos riches enseignements, communiqué votre amour de la recherche et vous en serai de ce fait toujours reconnaissant. Outre votre admirable expérience scientifique, j'ai également apprécié, Pierre, Anne et Pascal, vos qualités humaines qui m'ont toujours aidé à prendre les meilleures décisions, votre soutien et vos encouragements à donner à cette thèse le meilleur de moi-même.

Je remercie également Messieurs Cédric Demeure, chef du service TSI, et Jean Jacques Monot, chef du laboratoire LTC, pour m'avoir accueilli au sein de THALES et permis de réaliser la thèse dans les meilleures conditions de recherche possibles. Merci également à toutes les personnes du service que j'ai cottoyées durant ces années, qui ont donné à ce cadre professionnel une convivialité des plus agréables, sans oublier les personnes du laboratoire I3S qui m'ont toujours offert un accueil chaleureux lors de mes venues à Sophia-Antipolis.

Je souhaite aussi adresser mes remerciements aux membres du jury de cette thèse, notamment, Monsieur Philippe Loubaton qui m'a fait l'honneur de présider ce jury et de nous accueillir à l'Université de Marne-La-Vallée durant la soutenance. Je remercie vivement Messieurs Christian Jutten, professeur à l'INPG-LIS, et Phillip Regalia, professeur à l'INT, d'avoir accepté la tâche de rapporteur et Monsieur Lieven De Lathauwer, chargé de recherche CNRS à l'ETIS, celle d'examinateur.

Merci également à Madame Elisabeth Gassiat, directrice du DESS *d'Ingénierie Mathématique d'Orsay, ainsi qu'à Messieurs Stéphane Boucheron et Michel Prenat de m'avoir offert la joie d'enseigner dans cette formation que moi-même je suivis quelques années plutôt. Je tiens par ailleurs à témoigner à Messieurs Jean-Christophe Pesquet et Philippe Loubaton la profonde reconnaissance que je leur porte pour m'avoir donné le plaisir d'enseigner à la Faculté de Marne-La-Vallée.*

Mes plus tendres et chaleureuses pensées vont à mes amis pour leur affection et leur soutien, à ma famille et plus particulièrement à mes parents dont l'amour, la compréhension et les encouragements m'ont permis de croire et d'accomplir ce rêve.

Laurent Albera

Contents

Acronyms and Notations

Acronyms

- *Signal processing areas*

  ICA Independent Component Analysis

  BSS Blind Source Separation

  BMI Blind Mixture Identification

  DOA Direction Of Arrivals

- *Methods*

  EVD EigenValue Decomposition

  SVD Singular Value Decomposition

  JAD Joint Approximate Diagonalization

  JADE Joint Approximate Diagonalization of Eigen-matrices

  SOBI Second Order Blind Identification

  ICAR Independent Component Analysis using Redundancies in the quadricovariance

  BIRTH Blind Identification of source mixtures using Redundancies in the daTa Hexacovariance matrix

  BIOME Blind Identification of Overcomplete MixturEs

- *Statistics*

  SO Second Order

FourO Fourth Order

SixO Sixth Order

HO Higher Order

- *Modulations*

BPSK Binary Phase Shift Keying

QPSK Quad Phase Shift Keying

CPM Continuous Phase Modulation

CPFSK Continuous Phase Frequency Shift Keying

FSK Frequency Shift Keying

AM Amplitude Modulated

- *Others*

SNR Signal to Noise Ratio

SNIRM Signal to Interference plus Noise Ratio Maximum

UCA Uniformly spaced Circular Array

ULA Uniform spaced Linear Array

Notations

- *vectors* (one-way arrays) are denoted with bold lowercase symbols;
- *matrices* (2-way arrays) or tensors (HO arrays) are denoted with bold uppercase;
- *transposition*, *conjugate transposition*, *complex conjugation*, and *estimate* are denoted respectively with superscripts $(^{\mathsf{T}})$, $(^{\mathsf{H}})$, $(^*)$, and $(\hat{\ })$;
- P denotes the source number;
- N denotes the sensor number;
- k, K denote the sample index and the total sample number respectively;
- $\boldsymbol{s}(k)$ denotes the $P\times 1$ source random vector;
- $\boldsymbol{\nu}(k)$ denotes the $N\times 1$ noise random vector;

- $\boldsymbol{A}$ denotes the $N \times P$ constant mixing matrix;
- $\boldsymbol{a}_p$ denotes the p-th stearing vector, i.e. the p-th column of $\boldsymbol{A}$;
- α_p allows one to evaluate the blind identification quality of vector $\boldsymbol{a}_p$;
- $\boldsymbol{\mathcal{T}}$ denotes a trivial matrix, which by definition is of the form $\boldsymbol{\Lambda\Pi}$ where $\boldsymbol{\Lambda}$ is an invertible diagonal matrix and $\boldsymbol{\Pi}$ a permutation;
- $\otimes$ denotes the Kronecker product;
- $\odot$ denotes the Hadamard product;
- $\oslash$ denotes the Khatri-Rao product;
- $\|\boldsymbol{B}\|_F$ is the Frobenius norm of matrix $\boldsymbol{B}$.

List of Tables

List of Figures

Chapter 1

Introduction

Blind Source Separation (BSS), and more particularly Independent Component Analysis (ICA), now raise great interest. In fact, ICA plays an important role in many diverse application areas, including radiocommunications, speech and audio [1], radar, sonar, seismology, radio astronomy, medical diagnosis [32] and data analysis. For example, in Biomedicine contexts, it is possible to vizualize the electrical activity of a fetal heart, the Fetal ElectroCardioGram (FECG), from non-invasive techniques, say, from ECG-recordings measured on the mother's skin. In fact, these *cutaneous* recordings can be considered, in first approximation, as instantaneous linear mixtures of potential signals generated by underlying bioelectric phenomena (maternal and fetal heart activity, potential distributions generated by respiration and stomach activity, ...); noise can be taken into account as an additive perturbation. So ICA can be used to estimate the FECG from recordings on the mother's skin [32] in order to evaluate the well-being of the fetus and reveal important diagnostic information, like for the diagnosis of *arrhytmia*. Likewise, in digital radiocommunications contexts, if some sources are received by an array of sensors, and if for each source the channel delay spread associated with the different sensors is much smaller than the symbol durations, a static mixture of complex sources is observed from the sensors. BSS consists in this case of restoring by a spatial filtering operation the transmitted sources only from the sensor data. Depending on the application, it may be sufficient to identify a static mixture, as in Direction Of Arrival (DOA) estimation problems, since the column vectors of the mixture are the source steering vectors: this is referred to as *blind identification* of source mixtures. In other contexts such as radiocommunications, the question is that of *blind extraction* of sources, or more commonly BSS.

Whereas some algorithms try to decorrrelate estimated signals using Second Order (SO) statistics, as in Factor Analysis with Principal Component Analysis (PCA), ICA attempts to restore the independence of outputs using Higher Order (HO) statistics. Thus, under the source independence assumption, ICA allows one to blindly identify the static mixture, and consequently to extract the transmitted sources. Nevertheless, ICA performance depends on several assumptions: (i) sources should be independent in some way, and (ii) in most cases the mixture has to be *overdetermined*; in other words, there should be at least as many sensors as sources, which is generally a strong limitation unless sparsity conditions are assumed; if the latter assumption is not made, the mixture is called *underdetermined*. It is important to note that noisy static mixtures of P sources can be viewed as noiseless underdetermined, since the background noise may be considered as N additional sources as raised in [16] [69], where N is the number of sensors. So a noiseless model of $P+N$ sources may be used, but, some sources modeling the noise might however not be independent. Nevertheless, in the presence of a linear noise, including the Gaussian noise, the latter can be approximated by the output of causal convolutional filter of length M whose inputs are spatially and temporally white. Moreover, the latter convolutional filter can be written as a $N \times MN$ matrix. If the observed linear noise is temporally white (implying $M = 1$), the latter matrix is given by the square root of the noise covariance matrix. Thus, the data observed from the N sensors can be written as a noiseless static mixture of $P+MN$ independent sources, requiring the use of blind underdetermined mixture methods. Nevertheless, we do not resort to this noiseless model in the new methods proposed in this thesis, in contrast to [62], for the following reasons : (i) underdetermined mixtures can be hardly identified when a large number of sources is present, (ii) if the noise is a non linear process, the noisy model cannot be written as a underdetermined static mixture. In addition, the background noise will be assumed Gaussian in this thesis, and since its HO statistic contribution is null, it is not necessary to consider it as additional sources.

1.1 Assumptions and problem formulation

1.1.1 Matrix notation

First, define the following compact notation associated with the usual Kronecker product $\otimes$ and named *Kronecker power*:

$$\boldsymbol{B}^{\otimes m} = \underbrace{\boldsymbol{B}\otimes\boldsymbol{B}\otimes\ldots\otimes\boldsymbol{B}}_{m \text{ times}} \quad \text{with } \boldsymbol{B}^{\otimes 0}=1 \tag{1.1}$$

where $\boldsymbol{B}$ is any $N \times P$ rectangular matrix; $\boldsymbol{B}^{\otimes m}$ is then $N^m \times P^m$. Next, define a columnwise Kronecker product, denoted $\oslash$ and sometimes referred to as the Khatri-Rao product [39] [73]. For any rectangular matrices $\boldsymbol{G}$ and $\boldsymbol{H}$, of size $N_G \times P$ and $N_H \times P$ respectively, the columns of the $(N_G N_H) \times P$ matrix $\boldsymbol{G} \oslash \boldsymbol{H}$ are defined as $\boldsymbol{g}_j \otimes \boldsymbol{h}_j$, if $\boldsymbol{g}_j$ and $\boldsymbol{h}_j$ denote the columns of $\boldsymbol{G}$ and $\boldsymbol{H}$ respectively. The Khatri-Rao product $\oslash$ may also be defined [73] as

$$\boldsymbol{G} \oslash \boldsymbol{H} = [\boldsymbol{G} \otimes \mathbf{1}_{N_H}] \odot [\mathbf{1}_{N_G} \otimes \boldsymbol{H}] \tag{1.2}$$

where $\odot$ denote the usual Hadamard (element-wise) product and $\mathbf{1}_N$ an $N \times 1$ vector of 1s respectively. So it is also possible to define the *Khatri-Rao power*:

$$\boldsymbol{B}^{\oslash m} = \underbrace{\boldsymbol{B} \oslash \boldsymbol{B} \oslash \ldots \oslash \boldsymbol{B}}_{m \text{ times}} \quad \text{with } \boldsymbol{B}^{\oslash 0} = 1 \tag{1.3}$$

1.1.2 Assumptions and notations

Assume that for any fixed index k, N complex outputs $x_n(k)$ ($1 \leq n \leq N$) of a noisy mixture of P statistically independent sources $s_p(k)$ ($1 \leq p \leq P$) are available. The $N \times 1$ vector $\boldsymbol{x}(k)$ of the measured array outputs is given by

$$\boldsymbol{x}(k) = \boldsymbol{A}\, \boldsymbol{s}(k) + \boldsymbol{\nu}(k) \tag{1.4}$$

where $\boldsymbol{A}$, $\boldsymbol{s}(k)$, $\boldsymbol{\nu}(k)$ are the $N \times P$ constant mixing matrix, the $P \times 1$ source and $N \times 1$ noise random vectors, respectively. In addition, for any fixed index k, $\boldsymbol{s}(k)$ and $\boldsymbol{\nu}(k)$ are statistically independent.

We further assume the following hypotheses:

A1. Vector $\boldsymbol{s}(k)$ is stationary, ergodic (or *cyclostationary* and *cycloergodic*, respectively), with components a priori in the complex field and mutually uncorrelated at order $2q$ (the *cyclostationarity* case will be addressed in the statistical estimation section 1.2.4);

A2. Noise vector $\boldsymbol{\nu}(k)$ is stationary, ergodic and Gaussian with components a priori in the complex field too;

A3. $2q$-th order marginal source cumulants (they will be defined in section 1.2.1) are not null and have all the same sign;

A4. Column vectors $\boldsymbol{a}_p$ of $\boldsymbol{A}$, also called steering vectors, are not pairwise collinear and none of their entries is null;

A5. The $N^{q-1}{\times}P$ matrix $\boldsymbol{\mathcal{A}}_{q-1}^{\ell}$, which will be defined in section 4.1.1, is of full column rank (this implies that $P \leq N^{q-1}$);

where $\boldsymbol{\mathcal{A}}_{q}^{\ell} = \boldsymbol{A}^{\oslash q-\ell} \oslash \boldsymbol{A}^{*\oslash \ell}$ and q an arbitrary integer greater than 2.

Under the previous assumptions, the problem addressed in this report is the Blind Mixture Identification (BMI) of mixture $\boldsymbol{A}$, to within a *trivial* matrix $\boldsymbol{\mathcal{T}}$ (a trivial matrix is of the form $\boldsymbol{\Lambda\Pi}$ where $\boldsymbol{\Lambda}$ is an invertible diagonal matrix and $\boldsymbol{\Pi}$ a permutation), from $2q$-th order *statistics* (these ones will be defined in section 1.2.1) of the observations. Besides, the classical BSS problem in the overdetermined case consists of finding an $N \times P$ matrix (the static source separator), $\boldsymbol{W}$, yielding a $P \times 1$ output vector

$$\boldsymbol{y}(k) = \boldsymbol{W}^{\mathsf{H}} \boldsymbol{x}(k) \tag{1.5}$$

corresponding to the best estimate, $\hat{\boldsymbol{s}}(k)$, of the vector $\boldsymbol{s}(k)$, up to a multiplicative trivial matrix.

1.1.3 Performance criterion

Most of the existing performance criteria used to evaluate the quality of the BMI process, in the overdetermined case [16] or in the underdetermined case [17] [68], are global criteria, which evaluate a distance between the actual mixing matrix $\boldsymbol{A}$ and its blind estimate $\widehat{\boldsymbol{A}}$. Although practical, a global performance criterion necessarily contains a part of arbitrary considerations in the manner of combining all the distances between the vectors $\boldsymbol{a}_p$ and $\widehat{\boldsymbol{a}}_p$. Moreover, it is possible to find that an estimate $\widehat{\boldsymbol{A}}_1$ of $\boldsymbol{A}$ is better than an estimate $\widehat{\boldsymbol{A}}_2$, with respect to the global criterion, while some columns of $\widehat{\boldsymbol{A}}_2$ estimate the associated true steering vectors in a better way than $\widehat{\boldsymbol{A}}_1$. For these reasons, it may be more appropriate to use a non global criterion for the evaluation of the BMI process, which is defined by the P-uplet

$$\mathrm{D}\Big(\boldsymbol{A}, \widehat{\boldsymbol{A}}\Big) = (\alpha_1, \alpha_2, \ldots, \alpha_P) \tag{1.6}$$

where

$$\alpha_p = \min_{1 \leq i \leq P} \left[\mathrm{d}(\boldsymbol{a}_p, \widehat{\boldsymbol{a}}_i)\right] \tag{1.7}$$

and where $\mathrm{d}(\boldsymbol{u}, \boldsymbol{v})$ is the pseudo-distance between vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, defined by:

$$\mathrm{d}(\boldsymbol{u}, \boldsymbol{v}) = 1 - \frac{|\boldsymbol{u}^{\mathsf{H}} \boldsymbol{v}|^2}{\|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2} \tag{1.8}$$

where $\|\cdot\|$ is the Euclidean norm defined on $\mathbb{C}^N$ by $\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u}^{\mathsf{H}} \boldsymbol{u}}$.

Likewise, but only in the context of overdetermined mixtures, it may be more appropriate to use the well-known SINRM (Signal to Interference plus Noise Ratio Maximum) criterion, defined in [11, section 3] and in appendix H (page 118), in order to evaluate performances of BSS algorithms.

1.2 Statistics of $2q$-th order

1.2.1 Definition

The $2q$-th order statistics considered in this report are defined by

$$C_{i_1,\, i_2,\, ..,\, i_q,\, \boldsymbol{x}}^{i_{q+1},\, i_{q+2},\, ..,\, i_{2q}}(k) = \mathsf{Cum}\{x_{i_1}(k), x_{i_2}(k), \ldots, x_{i_q}(k), x_{i_{q+1}}(k)^*, \ldots, x_{i_{2q}}(k)^*\} \tag{1.9}$$

where q terms $x_i(k)$ are not conjugated and q terms are conjugated. Function (1.9) is well-known as the $2q$-th order *cumulant* computed from $2q$ components of $\boldsymbol{x}(k)$ with as many conjugated terms as not conjugated. Consequently, the associated $2q$-th order *marginal cumulant* of source $s_p(k)$ is defined by

$$C_{p,\, p,\, ..,\, p,\, \boldsymbol{s}}^{p,\, p,\, ..,\, p}(k) = \mathsf{Cum}\{\underbrace{s_p(k), s_p(k), \ldots, s_p(k)}_{q \text{ components}}, \underbrace{s_p(k)^*, \ldots, s_p(k)^*}_{q \text{ components}}\} \tag{1.10}$$

Note that in the presence of stationary sources, $2q$-th order statistics do not depend on time k, so they can be denoted by $C_{i_1,i_2,\ldots,i_q,\boldsymbol{x}}^{i_{q+1},i_{q+2},\ldots,i_{2q}}$. For the sake of convenience, we will describe our new algorithms, names ICAR, BIRTH and BIOME in the sequel, in the stationary case. Nevertheless, the cyclostationary case will be addressed in short in section 1.2.4 and more fully in section 5.2.

1.2.2 Matrix arrangement

Finally, $2q$-th order statistics computed according to (1.9) may be arranged in an $N^q \times N^q$ statistical matrix $\mathcal{C}_{2q,\, \boldsymbol{x}}$, called $2q$-th order statistical matrix of $\boldsymbol{x}(k)$ such that $\mathcal{C}_{2q,\, \boldsymbol{x}}$ is an Hermitian matrix. Nevertheless, several ways to store $2q$-th order statistics in $\mathcal{C}_{2q,\, \boldsymbol{x}}$ are possible and we consider in the following $q+1$ arrangements, indexed by the integer ℓ ($0 \leq \ell \leq q$), each yielding a statistical matrix $\mathcal{C}_{2q,\, \boldsymbol{x}}^{\ell}$ such that its $\left(I_1^\ell, I_2^\ell\right)$-th entry ($1 \leq I_1^\ell, I_2^\ell \leq N^q$) is given by

$$\mathcal{C}_{2q,\, \boldsymbol{x}}^{\ell}\left(I_1^\ell, I_2^\ell\right) = C_{i_1,\, i_2,\, ..,\, i_q,\, \boldsymbol{x}}^{i_{q+1},\, ..,\, i_{2q}} \tag{1.11}$$

where for any $0 \leq \ell \leq q$ and for all $1 \leq i_1, i_2, \ldots, i_{2q} \leq N$,

$$\begin{aligned}
I_1^\ell &= \varphi([\,\underbrace{i_1 \;\; i_2 \;\; \ldots \;\; i_{q-\ell-1} \;\; i_{q-\ell}}_{q-\ell \text{ first subscript indices}} \;\; \underbrace{i_{2q-\ell+1} \;\; \ldots \;\; i_{2q-1} \;\; i_{2q}}_{\ell \text{ last superscript indices}}\,]) \\
I_2^\ell &= \varphi([\,\underbrace{i_{q+1} \;\; i_{q+2} \;\; \ldots \;\; i_{2q-\ell-1} \;\; i_{2q-\ell}}_{q-\ell \text{ first superscript indices}} \;\; \underbrace{i_{q-\ell+1} \;\; \ldots \;\; i_{q-1} \;\; i_q}_{\ell \text{ last subscript indices}}\,])
\end{aligned} \tag{1.12}$$

and where function φ is defined by

$$\forall \boldsymbol{z} \in \mathbb{N}^J, \qquad \varphi(\boldsymbol{z}) = z(J) + \sum_{j=1}^{J-1} N^{J-j}(z(j) - 1) \tag{1.13}$$

denoting with $z(j)$ the j-th component of vector $\boldsymbol{z}$.

Example 1 *Fourth order (FourO) statistics defined by (1.9) for $q = 2$ and described explicitly in appendix D (page 83) for zero-mean complex variables that are distributed symmetrically with respect to the origin, may be arranged in the $N^2 \times N^2$ quadricovariance matrix $\boldsymbol{Q_x} = \boldsymbol{\mathcal{C}}_{4,\boldsymbol{x}}^1$ such that*

$$Q_{\boldsymbol{x}}(I_1^1, I_2^1) = C_{i_1, i_2, \boldsymbol{x}}^{i_3, i_4} \tag{1.14}$$

is the (I_1^1, I_2^1)-th entry ($1 \leq I_1^1,\ I_2^1 \leq N^2$) of $\boldsymbol{Q_x}$ and where for all $1 \leq i_1, i_2, i_3, i_4 \leq N$,

$$\begin{aligned}
I_1^1 &= \varphi([i_1 \;\; i_4]) = N(i_1 - 1) + i_4 \\
I_2^1 &= \varphi([i_3 \;\; i_2]) = N(i_3 - 1) + i_2
\end{aligned} \tag{1.15}$$

Example 2 *SixO statistics defined by (1.9) for $q = 3$ and described explicitly in appendix D (page 83) for zero-mean complex variables that are distributed symmetrically with respect to the origin, may be arranged in the $N^3 \times N^3$ hexacovariance matrix $\boldsymbol{H_x} = \boldsymbol{\mathcal{C}}_{6,\boldsymbol{x}}^1$ such that*

$$H_{\boldsymbol{x}}\left(I_1^1, I_2^1\right) = C_{i_1, i_2, i_3, \boldsymbol{x}}^{i_4, i_5, i_6} \tag{1.16}$$

is the (I_1^1, I_2^1)-th entry ($1 \leq I_1^1,\ I_2^1 \leq N^3$) of $\boldsymbol{H_x}$ and where for all $1 \leq i_1, i_2, i_3, i_4, i_5, i_6 \leq N$,

$$\begin{aligned}
I_1^1 &= \varphi([i_1 \;\; i_2 \;\; i_6]) = N(N(i_1 - 1) + i_2 - 1) + i_6 \\
I_2^1 &= \varphi([i_4 \;\; i_5 \;\; i_3]) = N(N(i_4 - 1) + i_5 - 1) + i_3
\end{aligned} \tag{1.17}$$

Remark 1 *Another, perhaps more intuitive (especially for readers familiar with Matlab), way to present the construction of $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$ is the following: first, construct an $2q$-dimensional tensor $\boldsymbol{T}$, whose elements are given by*

$$\boldsymbol{T}\begin{pmatrix} i_{2q}, i_{2q-1}, \ldots, i_{2q-\ell+1}, i_{q-\ell}, i_{q-\ell-1}, \ldots, i_1, \\ i_q, i_{q-1}, \ldots, i_{q-\ell+1}, i_{2q-\ell}, i_{2q-\ell-1}, \ldots, i_{q+1} \end{pmatrix} = C^{i_{q+1}, \ldots, i_{2q}}_{i_1, i_2, \ldots, i_q, \boldsymbol{x}} \tag{1.18}$$

The matrix $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$ is then given by a simple Matlab reshape operation as following

$$\mathcal{C}^{\ell}_{2q,\boldsymbol{x}} = \textit{reshape}(\boldsymbol{T}, N^q, N^q) \tag{1.19}$$

We limit ourselves to arrangements of statistics that give different results at the output of the BMI methods in terms of *processing power* (i.e. in terms of maximal number of processed sources). Note that the selection of the ordering parameter ℓ maximizing the processing power for a fixed cumulant order q will be discussed in section 4.2.2 summarizing results shown in [12].

1.2.3 Multilinearity property

The statistical matrix of the data, $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$ ($q \geq 1$), has a special structure especially thanks to the multilinearity property under changes of coordinate systems, shared by all moments and cumulants [55] [19, pp. 1-24]. Under assumptions (**A1**)-(**A2**), this property can be expressed, according to (1.11), (1.12) and (1.13), by the following equation

$$\forall\, 0 \leq \ell \leq q, \qquad \mathcal{C}^{\ell}_{2q,\boldsymbol{x}} = [\boldsymbol{A}^{\otimes q-\ell} \otimes \boldsymbol{A}^{*\otimes \ell}]\, \mathcal{C}^{\ell}_{2q,\boldsymbol{s}}\, [\boldsymbol{A}^{\otimes q-\ell} \otimes \boldsymbol{A}^{*\otimes \ell}]^{\mathsf{H}} \tag{1.20}$$

where the $N^q \times N^q$ matrices $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$ and the $P^q \times P^q$ matrices $\mathcal{C}^{\ell}_{2q,\boldsymbol{s}}$ are the statistical matrices of $\boldsymbol{x}(k)$ and $\boldsymbol{s}(k)$ respectively. The number ℓ is the same as that appearing in equations (1.12) and (1.11). Moreover, note that the arrangements $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$ and $\mathcal{C}^{q-\ell}_{2q,\boldsymbol{x}}$ ($0 \leq \ell \leq q$) give rise to the same processing power of underdetermined mixtures of arbitrary statistically independent sources as shown in [12]. In fact the first arrangement is the conjugate of the other whatever the values of q and N. It is then sufficient to limit the analysis to $0 \leq \ell \leq q_0$ where $q_0 = q/2$ if q is even and $q_0 = (q-1)/2$ if q is odd.

1.2.4 Statistical estimation

Generally, using the well-known Leonov-Shiryaev formula [55], applicable in the complex case [65], $2q$-th order cumulants (1.9) are computed from moments of order smaller than or equal to $2q$

given by

$$M_{i_1,i_2,\ldots,i_r,\boldsymbol{x}}^{i_{r+1},i_{r+2},\ldots,i_{r+s}}(k) = \mathsf{E}[x_{i_1}(k),\ldots,x_{i_r}(k), x_{i_{r+1}}(k)^*,\ldots,x_{i_{r+s}}(k)^*] \tag{1.21}$$

where $r+s \leq 2q$. Appendix D illustrates the Leonov-Shiryaev formula for HO statistics and for zero-mean complex variables that are distributed symmetrically with respect to the origin.

However, in practical situations, moments and cumulants cannot be exactly computed: they have to be estimated from components of $\boldsymbol{x}(k)$. If components are stationary and ergodic, sample statistics may be used to estimate v-th order moments [55], and consequently to estimate, via the Leonov-Shiryaev formula, $2q$-th order statistics (1.9).

Nevertheless, if sources are cyclostationary, cycloergodic, potentially non zero-mean, $2q$-th order continuous-time temporal mean statistics have to be used instead of (1.9), such as

$$C_{i_1,i_2,\ldots,i_q,\boldsymbol{x}}^{i_{q+1},i_{q+2},\ldots,i_m} = \left\langle C_{i_1,i_2,\ldots,i_q,\boldsymbol{x}}^{i_{q+1},i_{q+2},\ldots,i_m}(k)\right\rangle_{\mathrm{c}} \tag{1.22}$$

where $\langle\cdot\rangle_{\mathrm{c}}$ is the continuous-time temporal mean operation defined by

$$\forall\, f : t \longmapsto f(t), \quad \langle f(t)\rangle_{\mathrm{c}} = \lim_{T\to+\infty}\frac{1}{T}\int_{-T/2}^{T/2} f(t)dt \tag{1.23}$$

These continuous-time temporal mean statistics are thus estimated using, for $q=2$, the estimators described in [42] for zero mean signals and in [44] (see appendix H and more particularly section 5.2) for potentially non zero-mean signals, and extending the previous ones to very HO statistics for $q \geq 3$. Note that the proposed ICAR (see chapter 2), BIRTH (see chapter 3) and BIOME (see chapter 4) approaches can tolerate (in their current form), but do not totally exploit, cyclostationarity of the sources such as in [41]: this will be the subject of forthcoming works.

1.3 Bibliographical survey

The literature related to BMI or BSS in static mixtures is surveyed in this section. For nearly two decades, SO and HO BSS methods [20] have been developed to separate several statistically independent sources from measurements. While the first paper related to HO BSS has been published in 1985 by Herault et al. [48], the ICA concept is proposed a few years later in 1991; Comon proposes a FourO contrast-based method, COM2 [16], Cardoso and Souloumiac [8] develop a matrix approach, well-known as JADE, and give rise to the Joint Approximate Diagonalization (JAD) algorithm [9]. These approaches use explicitly or implicitly FourO statistics.

In parallel, other approaches attempt to exploit SO statistics only. This is made possible thanks to the color of the sources, assumed unknown but different. Fety was the first to exploit covariance matrices at two different delay lags [45]; the complete theoretical background is given only a few years later by Comon et al. [24]. The same kind of approach is developed independently ten years later by Belouchrani et al. [2], who give rise to the so-called SOBI method, only based on SO statistics.

Delfosse et al. [34] propose to extract one source at a time, which is now referred to as Deflation procedures. A few years later, Hyvarinen et al. present the FastICA method, first for signals with values in the real field [51], and later for complex signals [3], using the fixed-point algorithm to maximize a FourO contrast. This algorithm is of deflation type, as that of Delfosse et al. [34], and must extract one source at a time.

Continuing chronologically, Comon proposes a simple solution [18], named COM1 in this paper, to the maximization of another FourO contrast function presented in [70] [25]. Whereas all the latter methods exploit statistics of the data, other algorithms only use the geometric properties of the data constellation. Although Diamantaras deals with Blind Channel Identification (BCI) of one source in [37], he draws up in this recent paper (section V) an inventory of the current geometric BSS methods, which are actually born in the nineties.

Each of these methods suffers from limitations. To start with, the current geometric methods are very attractive, but for the time being, they are unable to separate any kind of sources but only a priori real M-ary PAM sources, and are very sensitive to noise. Next, the SOBI algorithm is unable to restore components that have comparable spectral densities. On the other hand, though the other previous methods perform well under some reasonable assumptions, they may be strongly affected by a Gaussian noise with unknown spatial correlation (because of their prewhitening stage). Such a noise appears for instance in some HF (High Frequency) radiocommunications applications.

Moreover, in such applications, the reception of more sources than sensors is possible and its probability increases with the reception bandwidth. The mixture is then called *underdetermined* [17], which means that the observation vectors are represented in the *overcomplete* basis of source vectors [50, pp. 305-313]. The previous algorithms, which require a SO prewhitening step, are then unable to identify the mixture and to extract the sources. Indeed, the SO prewhitening step, which aims at orthonormalizing the source steering vectors, cannot orthonormalize the latter when the number of sources is greater than the number of sensors.

In order to deal with the correlated noise problem, Ferréol et al. have proposed a new family of HO BSS methods exploiting the potential cyclostationarity of the received sources [41]. In fact, the latter family of algorithms uses cyclic statistics of the data; since cyclic covariance matrices associated with a stationary noise are null for non zero cyclic frequencies, these cyclic methods allow the optimal separation of independent sources even in the presence of a stationary noise (not necessarily Gaussian) with unknown spatial correlation. However, the use of cyclic methods is more complex because of the estimation of cyclic frequencies and time delays.

On the other hand, the underdetermined mixture case is a difficult problem with sharp identifiability questions. Taleb et Jutten have discussed some theoretical results on underdetermined source separation [67, chapitre 7] [69] showing, firstly, that blind identification of steering vectors of non Gaussian sources is possible, and secondly, that non Gaussian sources can only be restored up to an arbitrary additive random vector. However, for discrete sources, this vector is deterministic. Several other methods have been developed in order to face the underdetermined mixtures case, namely when there are fewer sensors than sources (e.g. the mixture enjoys no sparsity property such as disjoint source spectra, or sources non permanently present).

Contrary to the overdetermined case, the underdetermined BMI and BSS problems cannot be solved at the same time. Besides, even assuming that we know the mixing matrix, since the latter is not invertible, a simple pseudo-inverse does not generally yield a satisfactory solution. A more sophisticated estimator of sources has to be obtained, for instance, by Maximum Likelihood (ML) or Maximum A Posteriori (MAP) estimations [54] [49]. However, the problem with the ML/MAP estimators is that they are rarely easy to compute. This optimization cannot be put in a simple algebraic framework. As a consequence, it leads to a closed form solution only if sources have a Gaussian distribution: in this case the optimum is given by the pseudo-inverse. But since ICA with Gaussian variables is of little interest (lack of uniqueness), the pseudo-inverse is not a satisfactory solution in most cases.

However, one case where the ML/MAP optimization is easier than usual to compute, is when sources have a Laplacian distribution. Lee et al. assume it in [53] in order to extract three speech signals from only two mixtures. Nevertheless, using a *supergaussian* distribution, such as the Laplacian distribution, is well justified in feature extraction only if the independent components have a sparse decomposition, in the sense that they are quite often equal to zero (*e.g.* speech signals).

In order to face the BMI problem, one can use a ML estimation. In the simplest case of ML estimation, we formulate the joint likelihood of the mixture and the realization of the sources, and maximize it with respect to all these variables. However, maximization of the joint likelihood is a rather crude method of estimation, and from a Bayesian viewpoint, it is more interesting to maximize the marginal posterior probability of the mixing matrix. So a more sophisticated form of ML estimation is obtained by using a Laplace approximation of the posterior distribution of $\boldsymbol{A}$. This improves the stability of the algorithm, and has been successfully used for estimation of overcomplete bases from audio data [53]. Although the method [53] succeeds in identifying the steering vectors of up to four speech signals with only two sensors, the authors need however sparsity conditions, and do not address the general case when all sources are always present. Note that one could also use an Expectation-Maximization (EM) algorithm [57].

Another approach has been proposed by Grellier et al. in [47, chapter 6] [23], where the blind source extraction problem is addressed by forming virtual sensor measurements, in order to make it possible to invert linearly the observation model. Virtual measurements are a non linear function of actual measurements, and the choice of this non-linearity has to depend on the source distribution, assumed in [23] to be known and discrete. Two numerical algorithms are proposed, depending on the fact that the mixture is known (or beforehand identified) or not.

Other approaches have been published [6] [17] [30] [53] [36] [68] [33]. BMI can be addressed in terms of the diagonalization of some tensor [4] [31] [19]. The methods proposed in [6] [17] [30] [33] only exploit the information contained in the data FourO statistics whereas the one proposed in [68] exploits the information contained in the second characteristic function of the observations. In fact, Cardoso presents in [6] as soon as 1991 the interesting FOOBI (Fourth Order Only Blind Identification) concept, which exploits the *super-symmetric* FourO cumulant tensor, and more particularly, relates symmetries of the quadricovariance to rank properties. Based on EigenValue Decomposition (EVD) of a real symmetric matrix, the FOOBI algorithm has recently been improved by De Lathauwer et al. in [33] resorting to a joint (or simultaneous) diagonalization. Besides, De Lathauwer et al. define two other rank one detecting mappings yielding two other solutions to the blind identification of undertermined mixtures, with further weakened constraints on the source number P. Note that De Lathauwer extends in [27] the FOOBI concept to the canonical decomposition of a HO tensor non necessarily super-symmetric, computed by means of a joint congruence transformation [73]. Moreover, an application to the blind identification of convolutive MIMO (Multiple Outputs Multiple Outputs) is given in [27]. An other application of the extended FOOBI concept to the joint congruence transformation of a set of underdetermined matrices, say, with more columns that rows, is presented in [28], which is interesting since many ICA algorithms rely on this joint diagonalization step. In [36], Diamantaras presents an interesting method allowing to identify the mixture of many binary sources using a single observation sequence. Assuming additive Gaussian noise, the probability distribution function of the sole observation is known to be [47] a mixture of Gaussian centered at points that uniquely determine the mixing parameters and the source signals up to a permutation and a sign ambiguity. His simulations show that the method can successfully identify the mixture of at least up to ten binary source signals (this is of course limited by the noise level and the data length).

However, some of these methods have drawbacks in operational contexts. Indeed, The FOOBI algorithm [6] and its first improvement [33, section 2] allow to process up to P sources such that $P(P-1) \leq N^2(N-1)^2/2$ where N is the number of sensors. Likewise, the bound on P associated with the second improvement [33, section 3] of the FOOBI method is such that $P(P-1) \leq N^3(N-1)/2$. However, these three methods are suboptimal in terms of maximal number of processed sources, since the analysis of FourO virtual arrays [13] yields that for arrays with particular diversity, up to $P = N^2$ steering vectors may be identified from only FourO statistics. On the other hand, the third improvement [33, section 4] of the FOOBI algorithm allows one theoretically to reach the latter optimal upper bound. Nevertheless, although the previous methods [6] [33] seem very attractive in theory, no simulation has been presented. As for the BMI methods [17] [30], they assume FourO non-circularity and thus fail in separating FourO circular sources. Next, the theory developed in [17] only confines itself to the case of three sources and two sensors. In addition, the method [68] has been developed only for real mixtures of real-valued sources, and the issue of robustness with respect to an over estimation of the number of sources remains open. Eventually, the geometric approach presented in [36] focuses only on binary antipodal sources, and does not yet allow to process types of sequences such as multilevel PAM or QAM signals [61]. Likewise, the latter algorithm assumes that the Gaussian noise is spatially and temporally white. But more importantly, the application of the method is limited by the combinatorial explosion as the number of sources P increases, since the algorithm complexity is exponential with respect to P.

1.4 Chapter summaries

- ***Chapter 2***

  The problem of blind separation of overdetermined mixtures (fewer sources than sensors) of sources is addressed in this chapter. Since classical algorithms may be strongly affected by the presence of a Gaussian noise with unknown spatial coherence, a new method, named ICAR (Independent Component Analysis using Redundancies in the quadricovariance), is proposed to overcome this problem. This method, without any whitening operation, only exploits some redundancies of a particular quadricovariance matrix of the data. The comparison of its performance to those of classical methods shows off the best behavior of ICAR in most situations.

Moreover, for several years, contrast-based Blind Source Separation (BSS) has been successfully used in several areas, including radiocommunications. Here a functional approach relying on differential calculus theory is proposed, aiming at analyzing asymptotic performances of BBS contrast criteria: the variance of the estimated separating matrix is expressed as a function of that of estimated cumulants. As an example, this chapter focuses on three widely used FourO contrast criteria. This allows one to quantify the behavior of these three separators for large samples.

These works have been presented respectively at the two following conferences:

L. ALBERA, A. FERREOL, P. CHEVALIER and P. COMON, ”ICAR, un algorithme d'ICA à convergence rapide, robuste au bruit,” in *GRETSI 03, Dix-neuvième colloque sur le Traitement du Signal et des Images*, Paris, France, September 8-11 2003, vol. 1, pp. 193-196.

L. ALBERA and P. COMON, ”Asymptotic performance of contrast-based blind source separation algorithms,” in *SAM 02, Second IEEE Sensor Array and Multichannel Signal Processing Workshop*, Rosslyn, US, August 4-6 2002, pp. 244-248.

and have been submitted to:

L. ALBERA, A. FERREOL, P. CHEVALIER and P. COMON, ”ICAR: Independent Component Analysis using Redundancies,” in *ISCAS 04, 2004 IEEE International Symposium on Circuits and Systems*, Vancouver, Canada, May 23-26 2004, submitted to the invited sessions.

A journal paper has been submitted to IEEE Transactions On Signal Processing:

L. ALBERA, A. FERREOL, P. CHEVALIER and P. COMON, ICAR, a tool for Blind Source Separation using Fourth Order Statistics only,” submitted in *IEEE Transactions On Signal Processing*, November 2003.

- ***Chapter 3***

  The BMI of underdetermined mixtures problem is addressed by taking advantage of SixO statistics and the Virtual Array (VA) concept. It is shown how SixO cumulants can be used to increase the effective aperture of an arbitrary antenna array, and so to identify the mixture of more sources than sensors. A computationally simple but efficient algorithm, named BIRTH, is proposed and enables to identify the steering vectors of up to $P = N^2 - N + 1$ sources for arrays of N sensors with spatial diversity only, and up to $P = N^2$ for those with angular and polarization diversity. Moreover, improvements of BIRTH have been proposed in this chapter, optimizing differently the compromise between performance and complexity.

  One part of these works has been presented at the following conference:

  L. ALBERA, A. FERREOL, P. COMON and P. CHEVALIER, "Sixth order blind identification of underdetermined mixtures (BIRTH) of sources," in *ICA 03, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, April 1-4 2003, pp. 909-914.

  The other part, BIRTH improvements, has been submitted to the following conference:

  L. ALBERA and P. COMON and P. CHEVALIER and A. FERREOL, "Blind identification of underdetermined mixtures based on the hexacovariance," in *ICASSP 04, 2004 IEEE International Conference on Acoustics Speech and Signal Processing*, Montreal, Quebec, May 17-21 2004, submitted.

  A journal paper will be submitted to IEEE Transactions On Signal Processing.

- ***Chapter 4***

  The problem of Blind Identification of linear mixtures of independent random processes is known to be related to the diagonalization of some tensors. This problem is posed here in terms of a non conventional joint approximate diagonalization of several matrices. In fact, a congruent transform is applied to each of these matrices, the left transform being rectangular of full rank, and the right one being unitary. The application in antenna signal processing is described, and a family of new methods, named BIOME (Blind Identification of Overcomplete MixturEs of sources), extending the ICAR and BIRTH algorithms to statistics of arbitrary order $2q$, where q is an arbitrary integer greater than 2, and giving rise to the $2q$-BIOME methods, is proposed.

These works have been submitted to the following journals:

L. ALBERA, A. FERREOL, P. COMON and P. CHEVALIER, ”Blind Identification of Overcomplete Mixtures of sources (BIOME),” to appear in *Linear Algebra Applications*.

P. CHEVALIER, L. ALBERA, A. FERREOL and P. COMON, ”On the virtual array concept for higher order array processing,” in *IEEE Transactions On Signal Processing*.

Besides, a patent has been registered such as:

L. ALBERA, A. FERREOL, P. CHEVALIER and Pierre COMON, ”Procédé d'identification aveugle de mélanges de sources aux ordres supérieurs”, no. FR 03/4041, 63019 (THALES Communications), April 01 2003.

- ***Chapter 5***

  This chapter summarizes my other contributions, whose journal and conference papers are given in appendix. First, a new attractive FourO BMI method, named FOBIUM (Fourth Order Blind Identification of Underdetermined Mixtures of sources) and able to identify the steering vectors of more sources than sensors, has been proposed. The new method implements a FourO pre-whitening step and exploits the trispectrum diversities of the sources.

  On the other hand, we have analysed the behavior and proposed adaptations of the current SO and FourO blind source separation methods for sources which are cyclostationary and cyclo-ergodic up to FourO, and potentially non zero-mean. In fact, most of the SO and Higher order (HO) blind source separation methods developed this last decade aim at blindly separating statistically independent sources, assumed zero-mean, stationary and ergodic. Nevertheless, in many situations of practical interest, such as in radiocommunications contexts, the sources are non stationary and very often cyclostationary (digital modulations).

  These works have been presented respectively at the following conferences:

  A. FERREOL and L. ALBERA and P. CHEVALIER, ”Fourth Order Blind Identification of Underdetermined Mixtures of sources (FOBIUM),” in *ICASSP 03, 2003 IEEE International Conference on Acoustics Speech and Signal Processing*, Hong Kong, China, April 6-10 2003, pp. 41-44.

P. CHEVALIER, A. FERREOL and L. ALBERA, ”On the behavior of current second order blind source separation methods for first and second order cyclostationary sources — Application to CPFSK sources,” in *ICASSP 02, 2002 IEEE International Conference on Acoustics Speech and Signal Processing*, Orlando, US, May 13-17 2002, pp. 3081-3084.

A. FERREOL, P. CHEVALIER and L. ALBERA, ”Higher order blind separation of non zero-mean cyclostationary sources,” in *EUSIPCO 02, XI European Signal Processing Conference*, Toulouse, France, September 3-6 2001, vol. 5, pp. 103-106.

P. CHEVALIER, A. FERREOL and L. ALBERA, ”Méthodologie générale pour la séparation aveugle de sources cyclostationnaires arbitraires — Application à l’écoute passive des radiocommunications,” in *GRETSI 03, Dix-neuvième colloque sur le Traitement du Signal et des Images*, Paris, France, September 8-11 2003, vol. 1, pp. 43-46.

and will appear in the following journal:

A. FERREOL, P. CHEVALIER and L. ALBERA, ”Second order blind separation of first and second order cyclostationary sources — Application to AM, FSK, CPFSK and deterministic sources,” in *IEEE Transactions On Signal Processing*, April 2004.

Besides, a journal paper describing in detail the FOBIUM algorithm has be submitted to IEEE Transactions On Signal Processing:

A. FERREOL, L. ALBERA and P. CHEVALIER, ”Fourth Order Blind Identification of Underdetermined Mixtures of sources (FOBIUM),” submitted in *IEEE Transactions On Signal Processing*, November 2003.

Finally, two patents have been registered such as:

A. FERREOL, L. ALBERA and P. CHEVALIER, ”Procédé et dispositif d’identification autodidacte d’un mélange sous-déterminé de sources au 4^{eme} ordre”, no. FR 03/4043, 63021 (THALES Communications), April 01 2003.

A. FERREOL, P. CHEVALIER and L. ALBERA, ”Procédé de traitement d’antennes sur des signaux cyclostationnaires potentiellement non centrés”, no. FR 02/5575, 62801 (THALES Communications), May 03 2002, no. FR 2 839 390, November 07 2003

Chapter 2

Fourth Order Independent Component Analysis

We present in this chapter, dedicated to overdetermined mixtures and FourO statistics, two independent sections: in the former, a new algorithm is presented even when in the latter, asymptotic performance of contrast-based BSS methods is analysed.

2.1 ICAR or the fourth order blind source separation

A new method, named ICAR (Independent Component Analysis using Redundancies in the quadricovariance) is proposed in this section. Only based on fourth order statistics, ICAR frees o.s. from second order whitening step in contrast to classical methods [2] [16] [18] [8] [51] [3] and consequently is not affected asymptotically by the presence of a Gaussian noise with unknown spatial correlation. Actually, ICAR exploits redundancies in a particular FourO statistical matrix of the data, called *quadricovariance*. However, the latter algorithm assumes sources to have non zero FourO marginal cumulants with the same sign, assumption which is verified in most radiocommunications contexts. Furthermore, the performance of ICAR is also analysed in this chapter in different practical situations, through computer simulations, and compared to those of the classical algorithms named SOBI, COM1, COM2, JADE, FastICA and FOBIUM method.

2.1.1 The core equation

Under assumptions (**A1**)-(**A5**) of section 1.1.2 for $q = 2$, the ICAR method precisely exploits several redundancies in the quadricovariance matrix $\boldsymbol{Q_x}$, defined by example 1 of section 1.2.2, of the data especially thanks to the multilinearity property. Although most of BSS algorithms, such as JADE, exploits expression (1.20) for $q=2$, the ICAR method precisely uses an alternative form, described by

$$\boldsymbol{Q}_{\boldsymbol{x}} = [\boldsymbol{A}\oslash\boldsymbol{A}^{*}]\, \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}\, [\boldsymbol{A}\oslash\boldsymbol{A}^{*}]^{\mathsf{H}} \tag{2.1}$$

where the $P\times P$ diagonal matrix $\boldsymbol{\mathcal{Q}}_{\boldsymbol{s}} = \mathrm{Diag}\begin{bmatrix} C_{11,\boldsymbol{s}}^{11} & C_{22,\boldsymbol{s}}^{22} & \cdots & C_{PP,\boldsymbol{s}}^{PP} \end{bmatrix}$ (i.e. $\forall\, 1\leq p_1, p_2 \leq P$, $\boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}(p_1,p_2) = C_{p_1p_1,\boldsymbol{s}}^{p_1p_1}$ if $p_1=p_2$, 0 otherwise) is of full rank in contrast to $\boldsymbol{Q}_{\boldsymbol{s}} = \boldsymbol{\mathcal{C}}_{4,\boldsymbol{s}}^{1}$ (1.20), and where the $N^2\times P$ matrix $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$ is given by

$$\boldsymbol{A}\oslash\boldsymbol{A}^{*} = [\, \boldsymbol{a}_1\otimes\boldsymbol{a}_1^{*} \;\; \boldsymbol{a}_2\otimes\boldsymbol{a}_2^{*} \;\; \cdots \;\; \boldsymbol{a}_P\otimes\boldsymbol{a}_P^{*}\,] \tag{2.2}$$

and more particularly by

$$\boldsymbol{A}\oslash\boldsymbol{A}^{*} = [\; [\boldsymbol{A}^{*}\boldsymbol{\Phi}_1]^{\mathsf{T}}\, [\boldsymbol{A}^{*}\boldsymbol{\Phi}_2]^{\mathsf{T}} \cdots [\boldsymbol{A}^{*}\boldsymbol{\Phi}_N]^{\mathsf{T}}\;]^{\mathsf{T}} \tag{2.3}$$

with

$$\boldsymbol{\Phi}_n = \mathrm{Diag}\begin{bmatrix} A(n,1) & A(n,2) & \cdots & A(n,P) \end{bmatrix} \tag{2.4}$$

In other words, the non zero elements of the $P\times P$ diagonal matrix $\boldsymbol{\Phi}_n$ are the components of the n-th row of matrix $\boldsymbol{A}$.

2.1.2 The ICAR concept

The algorithm proposed proceeds in three stages. Firstly, a unitary matrix $\boldsymbol{V}$ is estimated in the Least Square (LS) sense, and allows one the estimation of $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$. In a second stage, several algorithms may be thought of in order to compute an estimate of $\boldsymbol{A}$ from $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$. Finally, estimation of sources $\boldsymbol{s}(k)$ is computed using the estimate of $\boldsymbol{A}$.

Identification of $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$

Proposition 1 *Under assumptions **(A4)** and **(A5)** (given in section 1.1.2 taking $q=2$), the $N^2\times P$ matrix $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$ is of full column rank.*

The proof of proposition 1 ensues immediately from equations (2.3), (2.4) and assumptions (**A4**) and (**A5**). In fact, suppose that $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$ is not full column rank. Then there exists some $P\times 1$ vector $\boldsymbol{\beta}\neq 0$ such that $[\boldsymbol{A}\oslash\boldsymbol{A}^{*}]\,\boldsymbol{\beta}=0$, which, due to the structure of $\boldsymbol{A}\oslash\boldsymbol{A}^{*}$ (2.3) implies that for all $1\leq n\leq N$, $\boldsymbol{A}^{*}\boldsymbol{\Phi}_n\,\boldsymbol{\beta}=0$. So it implies that $\boldsymbol{A}$ cannot be of full column rank (since matrices

$\boldsymbol{\Phi}_n$ are $P \times P$ diagonal with nonzero entries, due to (2.4) and (**A4**)), which contradicts assumption (**A5**).

So proposition 1, assumption (**A3**) (the latter assumption is given in section 1.1.2 taking $q=2$) and equation (2.1) allow together to prove, first, that matrix $\boldsymbol{Q_x}$ is of rank P and then that $\boldsymbol{Q_x}$ is positive if the FourO marginal source cumulants are positive, what we assume in this section. So a square root of $\boldsymbol{Q_x}$, denoted $\boldsymbol{Q_x^{1/2}}$ and such that $\boldsymbol{Q_x} = \boldsymbol{Q_x^{1/2}}[\boldsymbol{Q_x^{1/2}}]^{\mathsf{H}}$, may be computed (if the FourO marginal source cumulants are negative, matrix $-\boldsymbol{Q_x}$ has to be considered instead, for computing the square root). In fact, we deduce from (2.1) that matrix $[\boldsymbol{A} \oslash \boldsymbol{A}^*]\, \boldsymbol{\mathcal{Q}_s^{1/2}}$ is a natural square root of $\boldsymbol{Q_x}$. Another possibility is to compute this square root via the singular value decomposition of $\boldsymbol{Q_x}$ given by

$$\boldsymbol{Q_x} = \boldsymbol{E_s}\, \boldsymbol{L_s}\, \boldsymbol{E_s}^{\mathsf{H}} \tag{2.5}$$

where $\boldsymbol{L_s}$ is the real-valued diagonal matrix of the non zero eigenvalues of $\boldsymbol{Q_x}$. These latters are P since matrix $\boldsymbol{Q_x}$ is of rank P, $\boldsymbol{L_s}$ is thus of size $P{\times}P$. Besides, $\boldsymbol{E_s}$ is the $N^2 \times P$ matrix of the associated orthonormalized eigenvectors. Consequently, a square root of $\boldsymbol{Q_x}$ can be computed as following

$$\boldsymbol{Q_x^{1/2}} = \boldsymbol{E_s}\, \boldsymbol{L_s^{1/2}} \tag{2.6}$$

where $\boldsymbol{L_s^{1/2}}$ denotes a square root of $\boldsymbol{L_s}$. Note that this latter really exists thanks to assumption (**A3**) and proposition 2.

Proposition 2 *For a full rank matrix $\boldsymbol{A} \oslash \boldsymbol{A}^*$, (**A3**) is equivalent to assuming that the diagonal elements of $\boldsymbol{L_s}$ are not null and have also the same sign, corresponding to that of the FourO marginal source cumulants.*

The proof of proposition 2 is straightforward. In fact, it is well-known that two square roots of a matrix are equal to within a unitary matrix, such that

$$[\boldsymbol{A} \oslash \boldsymbol{A}^*]\, \boldsymbol{\mathcal{Q}_s^{1/2}} = \boldsymbol{E_s}\, \boldsymbol{L_s^{1/2}}\, \boldsymbol{V} \qquad \left(= \boldsymbol{Q_x^{1/2}}\, \boldsymbol{V}\right) \tag{2.7}$$

for some $P{\times}P$ unitary matrix $\boldsymbol{V}$. Equation (2.7) shows that the right-hand side is the SVD of the left-hand side, hence the proposition 2 result, since $\boldsymbol{E_s}^{\mathsf{H}}[\boldsymbol{A} \oslash \boldsymbol{A}^*]\boldsymbol{\mathcal{Q}_s}[\boldsymbol{A} \oslash \boldsymbol{A}^*]^{\mathsf{H}}\boldsymbol{E_s} = \boldsymbol{L_s}$ is a positive matrix.

In addition, equation (2.7) can be rewritten as following

$$\boldsymbol{Q}_{\boldsymbol{x}}^{1/2} = \boldsymbol{E}_{\boldsymbol{s}}\, \boldsymbol{L}_{\boldsymbol{s}}^{1/2} = [\boldsymbol{A} \oslash \boldsymbol{A}^{*}]\, \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}\, \boldsymbol{V}^{\mathsf{H}} \tag{2.8}$$

showing the link between $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}$ and $\boldsymbol{A} \oslash \boldsymbol{A}^{*}$. Plugging (2.3) into (2.8), matrix $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}$ can be eventually rewritten as

$$\begin{aligned} \boldsymbol{Q}_{\boldsymbol{x}}^{1/2} &= \Big[[\boldsymbol{A}^{*} \boldsymbol{\Phi}_1 \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \cdots [\boldsymbol{A}^{*} \boldsymbol{\Phi}_N \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \Big]^{\mathsf{T}} \\ &= [\; \boldsymbol{\Gamma}_1{}^{\mathsf{T}} \;\; \boldsymbol{\Gamma}_2{}^{\mathsf{T}} \;\; \cdots \;\; \boldsymbol{\Gamma}_N{}^{\mathsf{T}} \;]^{\mathsf{T}} \end{aligned} \tag{2.9}$$

where the N matrix blocks $\boldsymbol{\Gamma}_n$ of size $N \times P$ are given by

$$\forall\, 1 \leq n \leq N, \qquad \boldsymbol{\Gamma}_n = \boldsymbol{A}^{*} \boldsymbol{\Phi}_n \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}} \tag{2.10}$$

Proposition 3 *For any $1 \leq n \leq N$, matrix $\boldsymbol{\Gamma}_n$ is of full column rank.*

The proof ensues immediately from assumptions (**A3**)-(**A5**) (given in section 1.1.2 taking $q = 2$), from equation (2.10) and from the fact that the product of a full column rank matrix and an invertible square matrix is always full column rank.

Using proposition 3, pseudo-inverse $\boldsymbol{\Gamma}_n^{\sharp}$ of the $N \times P$ matrix $\boldsymbol{\Gamma}_n$ is defined by

$$\forall\, 1 \leq n \leq N, \qquad \boldsymbol{\Gamma}_n^{\sharp} = (\boldsymbol{\Gamma}_n{}^{\mathsf{H}} \boldsymbol{\Gamma}_n)^{-1} \boldsymbol{\Gamma}_n{}^{\mathsf{H}} \tag{2.11}$$

Then, consider the $N(N-1)$ matrices $\boldsymbol{\Theta}_{n_1,n_2}$ below

$$\forall\, 1 \leq n_1 \neq n_2 \leq N, \qquad \boldsymbol{\Theta}_{n_1,n_2} = \boldsymbol{\Gamma}_{n_1}^{\sharp} \boldsymbol{\Gamma}_{n_2} \tag{2.12}$$

which can be rewritten, from (2.10) and (2.11), as

$$\boldsymbol{\Theta}_{n_1,n_2} = \boldsymbol{V}\, \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{-1/2}\, \boldsymbol{\Phi}_{n_1}^{-1}\, \boldsymbol{\Phi}_{n_2}\, \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}\, \boldsymbol{V}^{\mathsf{H}} = \boldsymbol{V}\, \boldsymbol{\Phi}_{n_1}^{-1}\, \boldsymbol{\Phi}_{n_2}\, \boldsymbol{V}^{\mathsf{H}} \tag{2.13}$$

where $\boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}$ and $\boldsymbol{D}_{n_1,n_2} = \boldsymbol{\Phi}_{n_1}^{-1} \boldsymbol{\Phi}_{n_2}$ are $P \times P$ diagonal full rank matrices. So it appears from (2.13) that matrix $\boldsymbol{V}$ jointly diagonalizes the $N(N-1)$ matrices $\boldsymbol{\Theta}_{n_1,n_2}$.

Proposition 4 *Under assumption (**A4**) and (**A5**) given in section 1.1.2 taking $q = 2$, for all pair $1 \leq p_1 \neq p_2 \leq P$, at least one pair $1 \leq n_1 \neq n_2 \leq N$ exists such that $D_{n_1,n_2}(p_1, p_1) \neq D_{n_1,n_2}(p_2, p_2)$.*

The proof is given in appendix B.

Paper [2] and proposition 4 allow to assert that if $\boldsymbol{V}_{sol}$ jointly diagonalizes matrices $\boldsymbol{\Theta}_{n_1,n_2}$, then $\boldsymbol{V}_{sol}$ and $\boldsymbol{V}$ are related through $\boldsymbol{V}_{sol} = \boldsymbol{V}\boldsymbol{\mathcal{T}}$ where $\boldsymbol{\mathcal{T}}$ is a trivial unitary matrix. So matrix $\boldsymbol{V}_{sol}$ allows one, in accordance with (2.8), to recover $\boldsymbol{A} \oslash \boldsymbol{A}^*$ to within a trivial matrix as following

$$\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\, \boldsymbol{V}_{sol} = [\boldsymbol{A} \oslash \boldsymbol{A}^*]\ \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}\, \boldsymbol{\mathcal{T}} \tag{2.14}$$

Identification of mixture A

Three algorithms are proposed in this section to identify $\boldsymbol{A}$ from the estimate, $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\boldsymbol{V}_{sol}$, of $\boldsymbol{A} \oslash \boldsymbol{A}^*$. These algorithms optimize differently the compromise between performance and complexity.

Note that equation (2.14) can be rewritten from (2.3) in the form of N matrix blocks $\boldsymbol{\Sigma}_n = \boldsymbol{A}^* \boldsymbol{\Phi}_n \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2} \boldsymbol{\mathcal{T}}$ of size $N \times P$ as

$$\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\, \boldsymbol{V}_{sol} = [\boldsymbol{\Sigma}_1{}^{\mathsf{T}}\ \boldsymbol{\Sigma}_2{}^{\mathsf{T}} \cdots \boldsymbol{\Sigma}_N{}^{\mathsf{T}}]^{\mathsf{T}} \tag{2.15}$$

So a first approach to estimate $\boldsymbol{A}$ up to a trivial matrix, called ICAR1 in the sequel, consists of keeping, for instance, the matrix block $\boldsymbol{\Sigma}_1^*$ made up of the N first rows of $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\boldsymbol{V}_{sol}$ such that

$$\boldsymbol{\Sigma}_1 = \boldsymbol{A}^*\, \boldsymbol{\Phi}_1\, \boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}\, \boldsymbol{\mathcal{T}} \tag{2.16}$$

where $\boldsymbol{\Phi}_1$ and $\boldsymbol{\mathcal{Q}}_{\boldsymbol{s}}^{1/2}$ are diagonal matrices, and where $\boldsymbol{\mathcal{T}}$ is a unitary trivial matrix.

It is also possible to take into account all the matrix blocks $\boldsymbol{\Sigma}_n^*$ and to compute their average. This yields a second algorithm, named ICAR2, of higher complexity.

A third algorithm, called ICAR3, is now described, and yields a more accurate solution to the BSI problem: since matrix $\boldsymbol{A} \oslash \boldsymbol{A}^*$, given by (2.2), has been identified from the previous section by $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\boldsymbol{V}_{sol}$ to within a trivial matrix, ICAR3 consists first of mapping each $N^2 \times 1$ column vector $\boldsymbol{b}_p$ of $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\boldsymbol{V}_{sol}$ into an $N \times N$ matrix $\boldsymbol{B}_p$ (the n-th column of $\boldsymbol{B}_p$ is made up from the N consecutive components of $\boldsymbol{b}_p$ as from the $[N(n-1)+1]$-th one), and secondly of diagonalizing each matrix $\boldsymbol{B}_p^*$.

Theorem 1 *For any matrix $\boldsymbol{B}_p$ ($1 \leq p \leq P$) built from $\boldsymbol{Q}_{\boldsymbol{x}}^{1/2}\boldsymbol{V}_{sol}$, there exists a unique column vector $\boldsymbol{a}_q$ ($1 \leq q \leq P$) of $\boldsymbol{A}$ such that the eigenvector of $\boldsymbol{B}_p^*$ associated with the largest eigenvalue corresponds, up to a scale factor, to $\boldsymbol{a}_q$.*

The proof is given in appendix C. Note that in practical situations, as $\boldsymbol{A} \oslash \boldsymbol{A}^*$ is estimated to within a trivial matrix, the latter identification step allows one to estimate $\boldsymbol{A}$ also to within a trivial matrix.

Remark 2 *This third algorithm can be seen as an application of the tensor rank-*1 *approximation [29] [52] to the second-order case, say, the matrix case. In fact, given an hermitian $N \times N$ matrix $\boldsymbol{B}$, the problem consists of determining a scalar μ and a vector $\boldsymbol{a} \in \mathbb{C}^N$ such that the rank-*1 *matrix $\widehat{\boldsymbol{B}} = \mu\, \boldsymbol{a}\, \boldsymbol{a}^{\mathsf{H}}$ minimizes the function $\epsilon = \left\| \boldsymbol{B} - \widehat{\boldsymbol{B}} \right\|_F^2$ subject to vector $\boldsymbol{a}$ having unit norm, where $\|\boldsymbol{B}\|_F$ is the Frobenius norm of matrix $\boldsymbol{B}$. Indeed, the latter problem is solved by the dominant eigenpair (μ,$\boldsymbol{a}$), where μ is the eigenvalue with the largest absolute value [66] [46]. Note that several techniques for simple computations of approximations to a few eigenvectors and eigenvalues of a hermitian matrix can be found in [71] such as, for instance, the power method [72] [59] [22].*

Extraction of the P independent components

Finally, to estimate the signal vector $\boldsymbol{s}(k)$ for any value k, it is sufficient, under (**A5**), to apply a linear filter built from the identified matrix $\boldsymbol{A}$: such a filter may be the Spatial Matched Filter (SMF) described in [11] by $\boldsymbol{W} = \boldsymbol{R}_{\boldsymbol{x}}^{-1}\boldsymbol{A}$, which is optimal in the presence of decorrelated signals. In practical situations, since matrix $\boldsymbol{A}$ is estimated to within a trivial matrix according to section (2.1.2), neither order of sources $\boldsymbol{s}(k)$ nor their amplitude can be identified.

2.1.3 Implementation of the ICAR method

The different steps of the ICAR method are summarized hereafter when K samples of the observations, $\boldsymbol{x}(k)$ ($1 \leq k \leq K$), are available.

Step1 Estimation of the FourO statistics $C_{i_1, i_2, \boldsymbol{x}}^{i_3, i_4}$ from the K samples $\boldsymbol{x}(k)$ and sorting of them, using the (ℓ=1)-arrangement, into the matrix $\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}$, which is an estimate of $\boldsymbol{Q}_{\boldsymbol{x}}$.

Step2 Eigen Value Decomposition (EVD) of the Hermitian matrix $\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}$, estimation $\widehat{P}$ of the source number P from this EVD, and restriction of $\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}$ to the $\widehat{P}$ principal components : $\widehat{\boldsymbol{Q}}_{\boldsymbol{x}} = \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\, \widehat{\boldsymbol{L}}_{\boldsymbol{s}}\, \widehat{\boldsymbol{E}}_{\boldsymbol{s}}^{\mathsf{H}}$, where $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$ is the diagonal matrix of the $\widehat{P}$ eigenvalues of largest modulus and $\widehat{\boldsymbol{E}}_{\boldsymbol{s}}$ is the matrix of the associated eigenvectors.

Step3 Estimate the sign, ϵ, of the diagonal elements of $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$.

Step4 Computation of a square root matrix $[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}$ of $\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}$: $[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2} = \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\,|\widehat{\boldsymbol{L}}_{\boldsymbol{s}}|^{1/2}$, where $|\cdot|$ denotes the absolute value operator.

Step5 Computation from $[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}$ of the N matrices $\widehat{\boldsymbol{\Gamma}}_n$, construction of matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2} = [\widehat{\boldsymbol{\Gamma}}_{n_1}^{\sharp}\,\widehat{\boldsymbol{\Gamma}}_{n_2}]$ for all $1 \leq n_1 \neq n_2 \leq N$, and estimation, $\widehat{\boldsymbol{V}}_{sol}$, of the unitary matrix $\boldsymbol{V}_{sol}$ from the joint diagonalization of the $N(N-1)$ matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2}$ (the joint diagonalization algorithm is described in [9]).

Step6 Estimation $\widehat{\boldsymbol{A}}$ of the mixture $\boldsymbol{A}$ from the $N^2 \times P$ matrix $[[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]$ by

1. (ICAR1) taking the matrix block made up of the N first rows of $[[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]^*$;
2. (ICAR2) taking the average of the N matrix blocks, of size $N \times P$, made up of the successive rows of $[[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]^*$;
3. (ICAR3) taking each column vector $\widehat{\boldsymbol{b}}_p$ of $[[\epsilon\widehat{\boldsymbol{Q}}_{\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]$ remodeling them into $N \times N$ matrices $\widehat{\boldsymbol{B}}_p$, and building the matrix whose p-th column vector is the eigenvector of matrix $\widehat{\boldsymbol{B}}_p^*$ associated with the largest eigenvalue.

Step7 Estimation of the signal vector $\boldsymbol{s}(k)$ for any value k applying to $\boldsymbol{x}(k)$ a linear filter built from $\widehat{\boldsymbol{A}}$ like for example the SMF one defined by $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1}\widehat{\boldsymbol{A}}$.

2.1.4 Computer results

The synthetic signals used in this section are cyclostationary, and according to sections 1.2.4 and 5.2, other statistical estimators than empirical estimators should be employed. However, if the cyclostationary sources are zero-mean and circular, or non circular with a zero carrier residu, or non circular with different non zero carrier residus, such as the sources used subsequently, the bias due to empirical statistical estimators is negligible [42]. So we decide to employ them in the following simulations. Moreover, the criterion used in this section in order to evaluate performances of BSS algorithms, is the well-known SINRM (Signal to Interference plus Noise Ratio Maximum) criterion defined in [11, section 3].

The white noise case

The performance of ICAR at the output of the considered source separator is firstly illustrated in the presence of a Gaussian noise, spatially and temporally white, and compared with some well-known BSS algorithms. In fact, we assume that $P = 4$ statistically independent sources, i.e.

2 BPSK and 2 QPSK, all with a raised cosine pulse shape of roll-off equal to 0.25, are received by a Uniformly spaced Circular Array (UCA) of $N = 4$ identical sensors of radius R such that $R/\lambda = 0.55$ (λ: wavelength). The symbol period T_1 associated with the first BPSK is equal to three times the sample period T_e. The other sources have a symbol period equal to twice the sample period. The directions of arrival of the sources are such that the source steering vectors are orthogonal and the associated carrier residus are such that $f_{c1}\,T_e = 0$, $f_{c2}\,T_e = 0.3$, $f_{c3}\,T_e = 0.2$ and $f_{c4}\,T_e = 0.1$. We apply the COM1 [18], COM2 [16], JADE [8], SOBI [2], FastICA [3], FOBIUM [40], ICAR1, ICAR2 and ICAR3 methods, and the SINRM associated with each source is computed and averaged over 200 realizations.

Figures 2.1 and 2.2 show the variations of *SINRM*$_3$ (source 3 performance) at the output of the previous methods as a function of the number of samples while the input SNR (Signal to Noise Ratio) of the four sources, is assumed to be equal to 20 dB. It appears in figure 2.1 that ICAR3

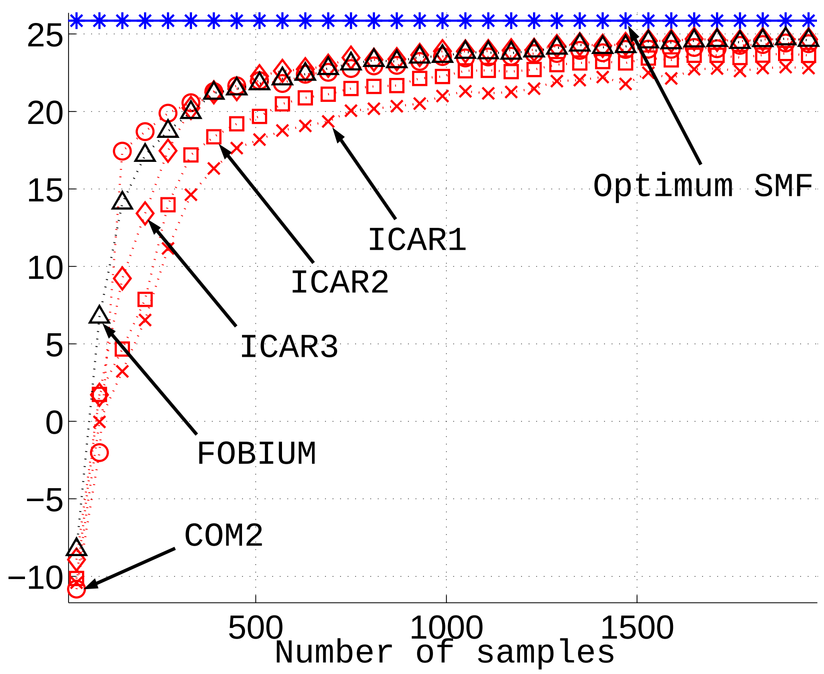

Figure 2.1: SINRM associated with source 3 for a SNR of 20 dB

converges as fast as COM2 and FOBIUM, but faster than ICAR1 and ICAR2: the third method given in section 2.1.2 exhibits better performances than the others. In addition, figure 2.2 shows the good performances of the ICAR3 algorithm facing the well-known COM1, JADE, SOBI and FastICA methods. Note that the SOBI and FOBIUM methods give in this simulation good results since sources have been chosen with different spectral densities, especially taking different carrier

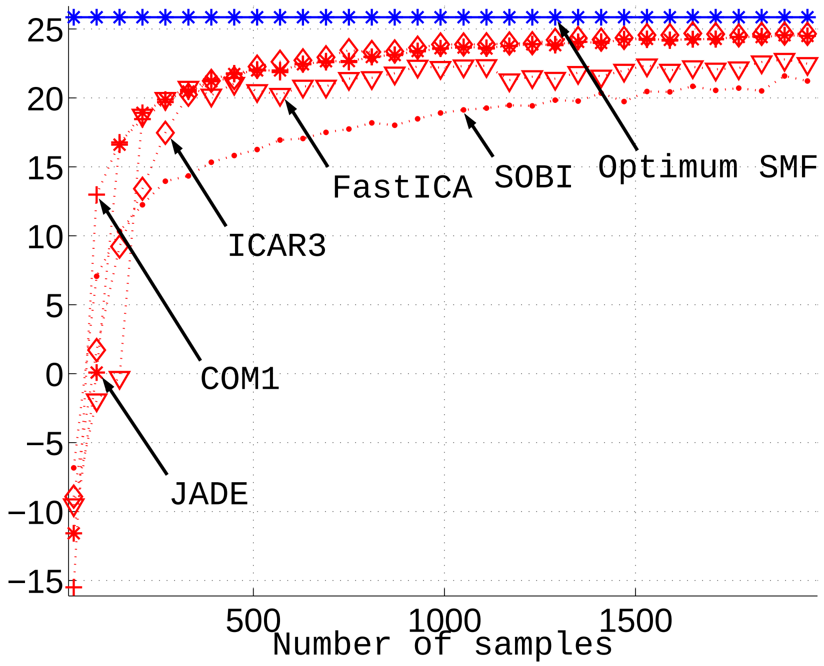

Figure 2.2: SINRM associated with source 3 for a SNR of 20 dB

residus. Similar results have been observed for the other sources.

Figures 2.3 and 2.4 show, for a number of one thousand samples, the variations of $SINRM_3$ at the output of the previous methods as a function of the input SNR, identical for the four sources. All the BSS methods have approximately the same behavior. First, when the SNR is very small, they do not succeed perfectly in extracting the third source. On the contrary, for signal to noise ratios contained in values -4 to 20 dB, the source separation is optimal. Finally, although the variations of $SINRM_3$ for signal to noise ratios greater than 20 dB are somewhat surprising, this result has already been observed by Monzingo and Miller in [56] for optimal separators when mixture $\boldsymbol{A}$ is known. Note that similar results have been obtained for the other sources.

The colored noise case

Then, the ICAR method is compared to other algorithms in the presence of a Gaussian noise with unknown spatial correlation. In fact, $P=3$ statistically independent sources, i.e. 2 BPSK and 1 QPSK, all with a raised cosine pulse shape of roll-off equal to 0.25, are assumed to be received by a UCA of $N=5$ identical sensors of radius R such that $R/\lambda = 0.55$. Their symbol periods are equal to $T_1 = 2T_e$, $T_2 = 3T_e$ and $T_3 = 4T_e$ respectively. Their carrier residus are chosen equal to zero. Finally, the source steering vectors are built orthogonal. This time, we apply the COM1,

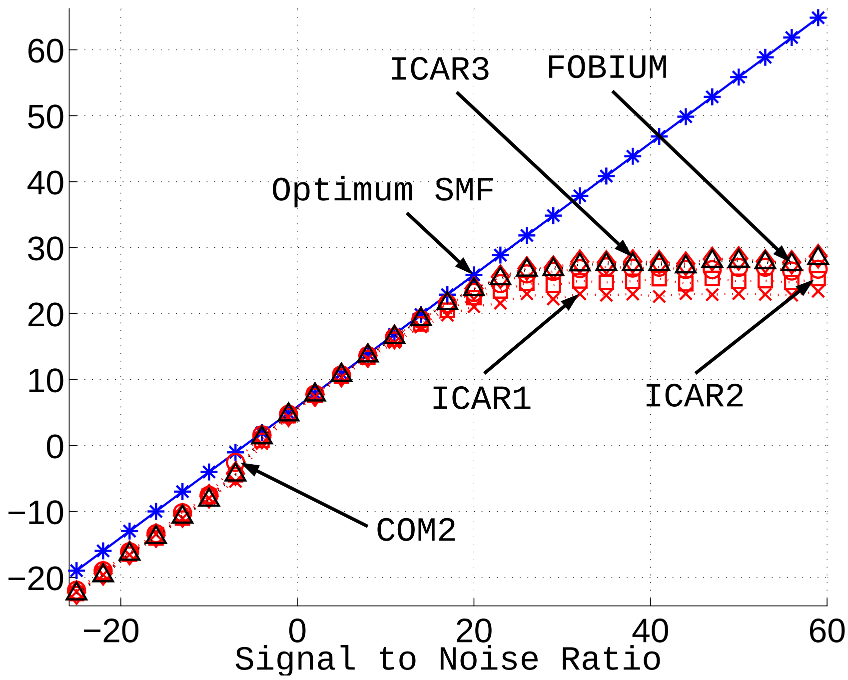

Figure 2.3: SINRM associated with source 3 for 1000 samples

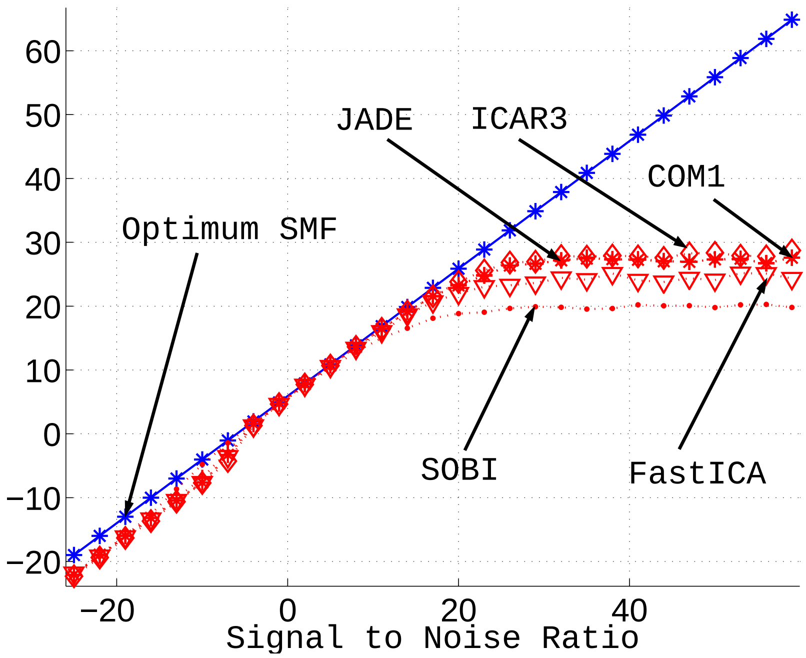

Figure 2.4: SINRM associated with source 3 for 1000 samples

COM2, JADE, SOBI, FOBIUM, ICAR1, ICAR2 and ICAR3 methods, and the SINRM associated with each source is computed and averaged over 200 realizations. Figures 2.5 and 2.6 show

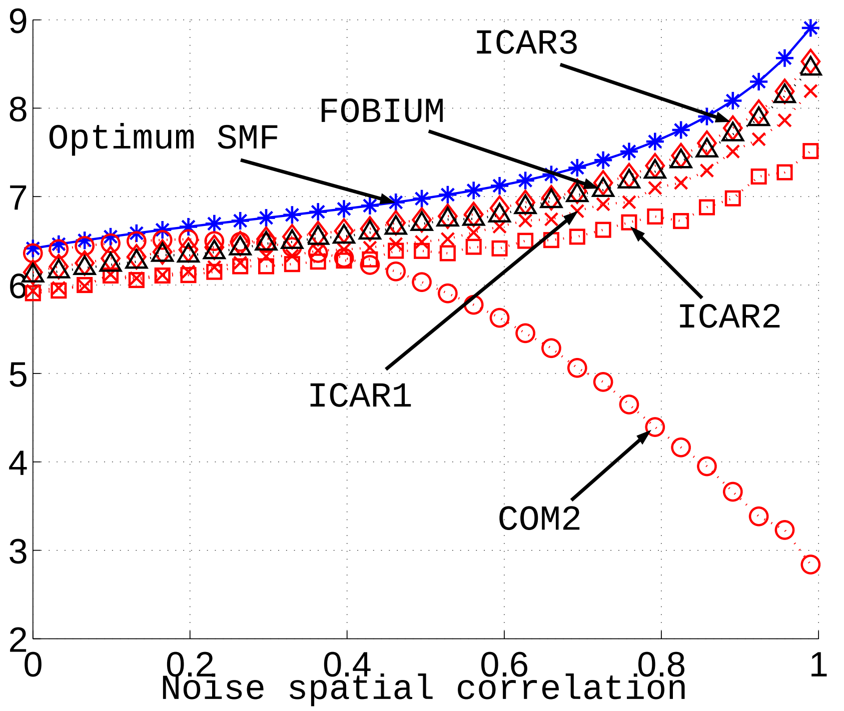

Figure 2.5: SINRM associated with source 3 for a SNR of 0 dB

the variations of $SINRM_3$ at the output of the previous methods as a function of the noise spatial correlation factor ρ. SNR of the three sources is taken equal to 0 dB and 1500 samples are used to identify mixture $\boldsymbol{A}$. Note that the Gaussian noise model employed in this simulation is the sum of an internal noise $\boldsymbol{\nu_{in}}(k)$ and an external noise $\boldsymbol{\nu_{out}}(k)$, of covariance matrices $\boldsymbol{R}_{\boldsymbol{\nu}}^{in}$ and $\boldsymbol{R}_{\boldsymbol{\nu}}^{out}$ respectively such that

$$\boldsymbol{R}_{\boldsymbol{\nu}}^{in}(r,q) \stackrel{\text{def}}{=} \sigma^2\delta(r-q)/2 \quad \boldsymbol{R}_{\boldsymbol{\nu}}^{out}(r,q) \stackrel{\text{def}}{=} \sigma^2\rho^{|r-q|}/2 \tag{2.17}$$

where σ^2, ρ are the total noise variance per sensor and the noise spatial correlation factor respectively. Note that $\boldsymbol{R}_{\boldsymbol{\nu}}(r,q) \stackrel{\text{def}}{=} \boldsymbol{R}_{\boldsymbol{\nu}}^{in}(r,q) + \boldsymbol{R}_{\boldsymbol{\nu}}^{out}(r,q)$ is the (r,q)-th component of the total noise covariance matrix.

It appears in figure 2.5 that the three proposed versions of ICAR seem to be robust with respect to the correlated Gaussian noise presence: ICAR1 and ICAR3 are totally insensitive to a Gaussian noise with unknown spatial correlation. On the other hand, figures 2.5 and 2.6 show that the well-known COM1, COM2, JADE and SOBI methods are strongly affected as soon as the noise

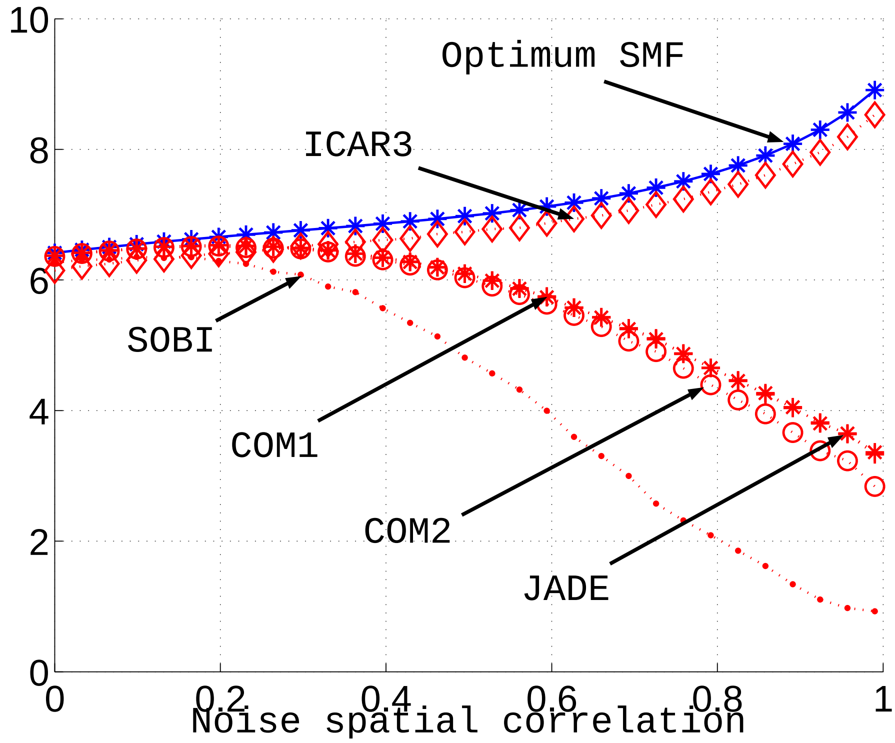

Figure 2.6: SINRM associated with source 3 for a SNR of 0 dB

spatial correlation increases beyond 0.4. In fact, the classical BSS methods require a prior spatial whitening based on SO moments. This stage theoretically needs the perfect knowledge of the noise covariance. If this is not the case, a whitening of the observed data is performed instead, which is biased. ICAR does not suffer from this drawback, since it uses only FourO cumulants, which are (asymptotically) insensitive to Gaussian noise, regardless of its space/time color. Besides, similar results have been observed for sources 1 and 2.

Over estimation of the number of sources

On the other hand, in operational contexts, the number of sources may be over estimated. It is then interesting to compare the ICAR method with other algorithms in such situations. To this aim, we assume that $P = 2$ statistically independent sources, i.e. 2 QPSK, both with a raised cosine pulse shape of roll-off equal to 0.25, are received by a UCA of $N = 4$ identical sensors of radius R such that $R/\lambda = 0.55$. Their symbol periods are equal to $T_1 = 2T_e$ and $T_2 = 3T_e$ respectively. Their carrier residus are chosen such that $f_{c1}\, T_e = 0$ and $f_{c2}\, T_e = 0.3$. Moreover, the spatial correlation between the two source steering vectors is taken equal to 0.5. Finally, the noise is built Gaussian, spatially and temporally white. We apply the COM1, COM2, JADE and

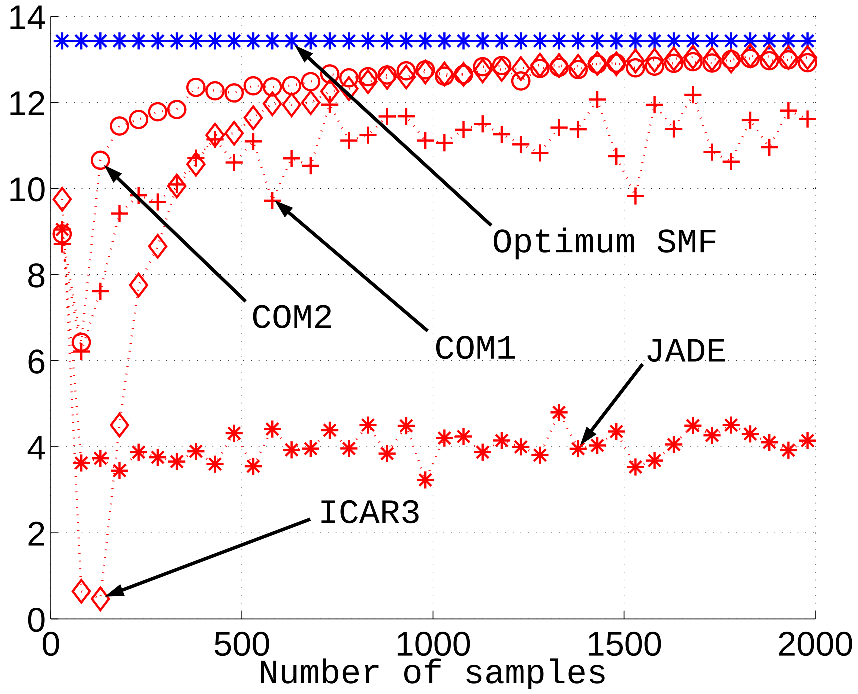

Figure 2.7: SINRM associated with source 2 for a SNR of 10 dB

ICAR3 methods, assuming that $\widehat{P}=3$ sources are received by the previous UCA, and the SINRM associated with each of the two sources is computed and averaged over 200 realizations at the output of each method.

Figure 2.7 shows the variations of $SINRM_2$ (source 2 performance) at the output of the previous methods as a function of the number of samples while the input SNR (Signal to Noise Ratio) of the two sources, is assumed to be equal to 10 dB. Similar results have been observed for source 1. More particularly, it appears that the COM2 and ICAR3 methods are robust with respect to an over estimation of the number of sources even when, in this simulation configuration, the JADE algorithm losses 10 dB, for less than 2000 samples, with respect to the case where $P=3$. Note that the latter result had already been pointed out in [10]. As for the COM1 method, it is affected by the over estimation of the number of sources, but less than the JADE algorithm. The explanation of this surprising phenomenon requires a harder analysis, which is beyond the scope of this paper.

2.1.5 Conclusion

A new method, named ICAR, exploiting the information contained in the data statistics at FourO only has been proposed in this paper. This new method allows one to process overdetermined mixtures of sources, provided the latter have marginal FourO cumulants with the same sign, which is generally the case in radiocommunications contexts. Three conclusions can be drawn: first, in the presence of a Gaussian noise spatially and temporally white, the proposed method gives as good results as those obtained with the current BSS methods. Second, contrary to most of the BSS algorithm, the ICAR method is not sensitive to a Gaussian colored noise whose spatial coherence is unknown. At last, the ICAR algorithm is robust with respect to an over estimation of the number of sources, which is not the case for some methods such as JADE. For these reasons, the ICAR method seems to correspond to the best method currently available to process overdetermined mixtures of sources. Note that such extensions to order 6, or more generally to order $m=2q$ ($q\geq 2$), will be proposed under the names of BIRTH and BIOME in chapters 3 and 4 respectively. Moreover, forthcoming works will consist of looking for the contrast criterion associated with ICAR in order to analyse accurately the performance of the latter using the functional approach proposed in the following section.

2.2 Asymptotic performance of fourth order contrast-based BSS algorithms

The purpose of this section is to examine the asymptotic performances (*e.g.* covariance of estimate) of contrast-based algorithms. Although the subject of asymptotic analysis has already been addressed in the signal processing literature, for instance, performance of SO [35] and ML [58] estimators in antenna array processing, or behavior of SO and HO BSS algorithms in the presence of zero-mean cyclostationary sources [42], this section proposes a functional approach allowing to compare asymptotic performances of BSS contrast criteria. As an illustration, 3 FourO contrast criteria already compared in [21] by computer experiments, are mainly focused on, for subsequent asymptotic performance analysis.

Note that assumptions (**A1**) to (**A3**) (given in section 1.1.2 taking $q=2$) are made in the sequel. Besides, the mixing matrix $\boldsymbol{A}$ is assumed to be square and unitary. Finally, $\boldsymbol{C_x}$ denotes the FourO order cumulant tensor whose entries are given by (1.9) for $q = 2$ and $\boldsymbol{U}$ replaces matrix $\boldsymbol{W}^{\mathsf{H}}$ defined by (1.5). Note that the unitary assumption with respect to $\boldsymbol{A}$ is not restrictive if a spatial prewhitening has been performed as for most of BSS methods [20]. But we limit our study for the time being to the effect of fourth-order estimation errors on the separator $\boldsymbol{U}$.

2.2.1 contrast-based BSS methods

Various approaches have been devised for BSS or ICA [7]. We shall focus exclusively on those maximizing a contrast measure of $\boldsymbol{y}$ (the latter vector has been defined in (1.5)). Recall that contrasts are criteria $\Upsilon(\boldsymbol{U};\boldsymbol{C_x})$ satisfying the properties below [16] [70]:

P1. *Invariance:* The contrast should not change within the set T of trivial matrices, which means that $\forall \boldsymbol{x} \in \mathcal{H}\cdot\mathcal{S}$, $\forall \boldsymbol{U} \in T$, $\Upsilon(\boldsymbol{U};\boldsymbol{C_x}) = \Upsilon(\mathbf{I}_N;\boldsymbol{C_x})$.

P2. *Domination:* If sources are already separated, any matrix should decrease the contrast. In other words, $\forall \boldsymbol{U} \in \mathcal{H}$, $\forall \boldsymbol{x} \in \mathcal{S}$, $\Upsilon(\boldsymbol{U};\boldsymbol{C_x}) \leq \Upsilon(\mathbf{I}_N;\boldsymbol{C_x})$.

P3. *Discrimination:* The maximum contrast should be reached only for matrices linked to each other via trivial matrices: $\forall \boldsymbol{x} \in \mathcal{S}$, $\Upsilon(\boldsymbol{U};\boldsymbol{C_x}) = \Upsilon(\mathbf{I}_N;\boldsymbol{C_x}) \Rightarrow \boldsymbol{U} \in T$.

where $\mathcal{H}$, $\mathcal{H}\cdot\mathcal{S}$, $\mathbf{I}_N$ are a set of matrices, the set of processes obtained by matrix mappings $\mathcal{H}$ on processes of $\mathcal{S}$, and the $N \times N$ identity matrix, respectively. Note that the *trivial* matrix definition is given page xiii of the preface.

The goal of this section is to evaluate the asymptotic statistical properties (*e.g.* covariance) of the matrix $\boldsymbol{U}$ delivered by contrast-based algorithms.

2.2.2 Asymptotic properties: a functional approach

From now on, it is assumed that $\Upsilon(\cdot,\boldsymbol{C})$ is of class C^2, and $\Upsilon(\boldsymbol{U},\cdot)$ is of class C^1. This will be satisfied for criteria given in section 2.2.3. The optimal solution $\boldsymbol{U}_o$ is defined as the absolute maximum of a contrast $\Upsilon(\boldsymbol{U};\boldsymbol{C_x})$:

$$\boldsymbol{U}_o = \mathrm{Arg\,max}_{\boldsymbol{U}}\Upsilon(\boldsymbol{U};\boldsymbol{C_x}) \tag{2.18}$$

where $\boldsymbol{C_x}$ is the exact cumulant tensor of the observation. In practice, $\boldsymbol{C_x}$ is estimated by a quantity $\hat{\boldsymbol{C}}_{\boldsymbol{x}}$, which involves estimation errors on $\boldsymbol{U}$; this yields a solution $\widehat{\boldsymbol{U}}$:

$$\widehat{\boldsymbol{U}} = \mathrm{Arg\,max}_{\boldsymbol{U}}\Upsilon(\boldsymbol{U};\hat{\boldsymbol{C}}_{\boldsymbol{x}}) \tag{2.19}$$

Both $\boldsymbol{U}_o$ and $\widehat{\boldsymbol{U}}$ are maxima of Υ, and thus satisfy the stationary point equations:

$$\boldsymbol{h}(\boldsymbol{U}_o,\boldsymbol{C_x}) = \mathbf{0}, \quad \boldsymbol{h}(\widehat{\boldsymbol{U}},\hat{\boldsymbol{C}}_{\boldsymbol{x}}) = \mathbf{0} \tag{2.20}$$

where $\boldsymbol{h}(\cdot,\boldsymbol{C})$ denotes the gradient of $\Upsilon(\boldsymbol{U},\boldsymbol{C})$ with respect to $\boldsymbol{U}$, in a sense subsequently defined.

Now, $\boldsymbol{h}$ is a well defined function in a P^2-dimensional linear space on the real field $\mathbb{R}$. In fact, Υ is twice continuously differentiable with respect to $\boldsymbol{U}$, and the tangent space to the manifold of unitary matrices is the linear space of skew-hermitian matrices ($\boldsymbol{\mathcal{B}}^{\mathsf{H}} = -\boldsymbol{\mathcal{B}}$) on the real field, which is indeed of dimension P^2 and admits as a basis the set of matrices $\boldsymbol{\mathcal{B}}_{qr}$, null everywhere except in rows and columns (q,r), (r,q), such that

$$d\boldsymbol{U} = \sum_{q,r=1}^{P} d\mu_{qr}\boldsymbol{\mathcal{B}}_{qr}\boldsymbol{U} \tag{2.21}$$

where the (v,w)-th component of matrix $\boldsymbol{\mathcal{B}}_{qr}$ is given by

$$\boldsymbol{\mathcal{B}}_{qr}(v,w) \stackrel{\text{def}}{=} \begin{cases} \delta(q-v)\delta(r-w) - \delta(q-w)\delta(r-v) & if \quad q<r \\ j\delta(q-v)\delta(r-w) & if \quad q=r \\ j\left[\delta(q-v)\delta(r-w) + \delta(q-w)\delta(r-v)\right] & if \quad q>r \end{cases} \tag{2.22}$$

with $\mathrm{j}^2 \stackrel{\text{def}}{=} -1$. Note that among the P^2 elements $\boldsymbol{\mathcal{B}}_{qr}$ generating the basis of the linear space of skew-hermitian matrices, $P(P-1)/2$ matrices are real and $P(P+1)/2$ matrices are imaginary.

Next, $\boldsymbol{h}(\cdot,\cdot)$ is continuously differentiable, which allows one to resort to the implicit function theorem in the neighborhood of $(\boldsymbol{U}_o, \boldsymbol{C_x})$. This yields, from (2.20):

$$\dot{\boldsymbol{h}}_{\boldsymbol{U}}(\boldsymbol{U}_o, \boldsymbol{C_x})\, d\boldsymbol{U} + \dot{\boldsymbol{h}}_{\boldsymbol{C}}(\boldsymbol{U}_o, \boldsymbol{C_x})\, d\boldsymbol{C} = o(d\boldsymbol{U}, d\boldsymbol{C}) \tag{2.23}$$

Thus, in the neighborhood of $(\boldsymbol{U}_o, \boldsymbol{C_x})$, $\widehat{\boldsymbol{U}} = \boldsymbol{U}_o + d\boldsymbol{U}$ can be expressed as a function of $\hat{\boldsymbol{C}}_{\boldsymbol{x}} = \boldsymbol{C_x} + d\boldsymbol{C}$. This can be rewritten in block form as [21], defining $\mathbf{vec}[\boldsymbol{B}]$ by the vector built from the columns of $\boldsymbol{B}$ stacked one below another:

$$\boldsymbol{F}_1\,\mathbf{vec}[d\boldsymbol{U}] = \boldsymbol{F}_2\,\mathbf{vec}[d\boldsymbol{C}] \tag{2.24}$$

where $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are matrices of dimension $P^2 \times P^2$ and $P^2 \times M$, respectively, built from the second derivatives of Υ, $\partial^2\Upsilon/\partial U\partial U$ and $\partial^2\Upsilon/\partial U\partial C$, stored in the proper manner. Here, M denotes the number of free parameters in $\boldsymbol{C_x}$, and, for any $P \geq 4$, is equal to $M = P(P+1)(P^2+P+1)/8$ in the complex case, which deflates to $M = P(P+1)(P+2)(P+3)/24$ in the real case.

The variance of $d\boldsymbol{U}(v,w)$ and therefore, the one of $\widehat{\boldsymbol{U}}(v,w)$ can thus theoretically be accessed by the formula:

$$Var\Big\{\mathbf{vec}\Big[\widehat{\boldsymbol{U}}\Big]\Big\} = \boldsymbol{F}_1^{-1}\boldsymbol{F}_2 Var\Big\{\mathbf{vec}\Big[\hat{\boldsymbol{C}}_{\boldsymbol{x}}\Big]\Big\}\boldsymbol{F}_2^{\mathsf{H}}\boldsymbol{F}_1^{-\mathsf{H}} \tag{2.25}$$

Nevertheless, matrix $\boldsymbol{F}_1$ is in general not of full rank, because the set of $d\mu_{qr}$ does not form a free family. Its rank is $P(P-1)$, so that the inverse above should be replaced by a pseudo-inverse. Nevertheless, this covariance can be consequently still computed once we know the covariance of sample cumulants. Using McCullagh bracket notation (defined in appendix D), and noting $[\bar{2}]expr = expr + expr^*$, this covariance is given in [21] in the general case. In the symetrically distributed case in which we are interested, the covariance takes the form:

$$\begin{aligned}
\mathcal{K}\ \mathrm{Var}\{\hat{\mathrm{C}}_{\mathrm{i},\ell}^{\mathrm{j,k}}, \hat{\mathrm{C}}_{\mathrm{J,K}}^{\mathrm{I,L}}\} = \mathrm{C}_{\mathrm{i},\ell,\mathrm{I,L}}^{\mathrm{j,k,J,K}} + [\bar{2}][4]\mathrm{C}_{\mathrm{i,I}}^{\mathrm{j,k,J,K}}\mathrm{C}_{\ell,\mathrm{L}} + [\bar{2}][4]\mathrm{C}_{\mathrm{i,I,L}}^{\mathrm{j,k,J}}\mathrm{C}_{\ell}^{\mathrm{K}} + [\bar{2}][4]\mathrm{C}_{\mathrm{i,L}}^{\mathrm{J,K}}\mathrm{C}_{\mathrm{I},\ell}^{\mathrm{j,k}} \\
+[\bar{2}][4]C_{i,I,L}^{K}C_{\ell}^{j,k,J} + C_{i,\ell,I,L}C^{j,k,J,K} + C_{i,\ell}^{J,K}C_{I,L}^{j,k} + [8]C_{i,I}^{j,J}C_{\ell,L}^{k,K} + \\
[\bar{2}][4]C_{i,\ell,I}^{J}C_{L}^{j,k,K} + [16]C_{i,I}^{j,J}C^{k,K}C_{\ell,L} + [16]C_{i,I}^{j,J}C_{L}^{k}C_{\ell}^{K} \\
+[\bar{2}][8]C_{i}^{j,J,K}C_{I}^{k}C_{\ell,L} + [\bar{2}][8]C_{i,\ell,I}^{J}C^{j,K}C_{L}^{k} + [\bar{2}][2]C_{i,\ell,I,L}C^{j,J}C^{k,K} \\
+[\bar{2}][2]C_{i,\ell}^{J,K}C_{I}^{j}C_{L}^{k} + [16]C_{i,I}C^{j,J}C_{L}^{k}C_{\ell}^{K} + [4]C_{i}^{J}C_{I}^{j}C_{L}^{k}C_{\ell}^{K} \\
+[4]C_{i,I}C^{j,J}C^{k,K}C_{\ell,L}
\end{aligned} \quad (2.26)$$

where $\mathcal{K}$ denotes the number of snapshots.

2.2.3 Examples and asymptotic analysis of particular contrasts

Define the three FourO contrast criteria below:

$$\Upsilon_1\left(\boldsymbol{U};\boldsymbol{C}_{\boldsymbol{x}}\right) = \epsilon\sum_{p=1}^{P} C_{p,p,\boldsymbol{y}}^{p,p} \qquad \Upsilon_2\left(\boldsymbol{U};\boldsymbol{C}_{\boldsymbol{x}}\right) = \sum_{p=1}^{P}\left(C_{p,p,\boldsymbol{y}}^{p,p}\right)^2$$

$$\Upsilon_3\left(\boldsymbol{U};\boldsymbol{C}_{\boldsymbol{x}}\right) = \sum_{p,k,\ell=1}^{P}\left|C_{p,\ell,\boldsymbol{y}}^{p,k}\right|^2 \quad (2.27)$$

where ϵ is a fixed sign. Note [21] that Υ_1 is a contrast if, for any $1 \le p \le P$, $C_{p,p,\boldsymbol{s}}^{p,p}$ have the same sign ϵ, and that Υ_3 is the contrast linked with the JADE algorithm [7].

Asymptotic results

After a first differential calculus with respect to $\boldsymbol{U}$, we obtain:

$$d\Upsilon_{1,\boldsymbol{U}} = 4\epsilon\left[\sum_{q<r} d\mu_{qr}\Re\left\{C_{r,q,\boldsymbol{y}}^{q,q} - C_{q,r,\boldsymbol{y}}^{r,r}\right\} - \sum_{q>r} d\mu_{qr}\Im\left\{C_{q,r,\boldsymbol{y}}^{r,r} + C_{r,q,\boldsymbol{y}}^{q,q}\right\}\right] \quad (2.28)$$

$$\begin{aligned}
d\Upsilon_{2,\boldsymbol{U}} = 8\Bigg[\sum_{q<r} d\mu_{qr}\left(C_{q,q,\boldsymbol{y}}^{q,q}\Re\left\{C_{r,q,\boldsymbol{y}}^{q,q}\right\} - C_{r,r,\boldsymbol{y}}^{r,r}\Re\left\{C_{q,r,\boldsymbol{y}}^{r,r}\right\}\right) \\
- \sum_{q>r} d\mu_{qr}\left(C_{r,r,\boldsymbol{y}}^{r,r}\Im\left\{C_{q,r,\boldsymbol{y}}^{r,r}\right\} + C_{q,q,\boldsymbol{y}}^{q,q}\Im\left\{C_{r,q,\boldsymbol{y}}^{q,q}\right\}\right)\Bigg]
\end{aligned} \quad (2.29)$$

$$
\begin{aligned}
d\Upsilon_{3,\boldsymbol{U}} = 8 \Bigg[& \sum_{q<r} \sum_{k,\ell} d\mu_{qr}\, \Re\Big\{ C^{q,k}_{q,\ell,\boldsymbol{y}} C^{q,\ell}_{r,k,\boldsymbol{y}} - C^{r,k}_{r,\ell,\boldsymbol{y}} C^{r,\ell}_{q,k,\boldsymbol{y}} \Big\} \\
& - \sum_{q>r} \sum_{k,\ell} d\mu_{qr}\, \Im\Big\{ C^{r,k}_{r,\ell,\boldsymbol{y}} C^{r,\ell}_{q,k,\boldsymbol{y}} + C^{q,k}_{q,\ell,\boldsymbol{y}} C^{q,\ell}_{r,k,\boldsymbol{y}} \Big\} \Bigg] \qquad (2.30)
\end{aligned}
$$

where $\Re\{z\}$ and $\Im\{z\}$ are respectively the real and imaginary parts of the complex number z.

So for each contrast, we can easily deduce from (2.28), (2.29) and (2.30) the function $\boldsymbol{h}_m$ defined in section (2.2.2). In particular, according to (2.18), (2.19), (2.20) and (2.28) the function $\boldsymbol{h}_1$ associated with $d\Upsilon_{1,\boldsymbol{U}}$ is described by

$$
h_1(\boldsymbol{U},\boldsymbol{C})_{qr} = \begin{cases} \Re\{C^{q,q}_{r,q,\boldsymbol{y}} - C^{r,r}_{q,r,\boldsymbol{y}}\} & \text{if } q<r \\ -\Im\{C^{r,r}_{q,r,\boldsymbol{y}} + C^{q,q}_{r,q,\boldsymbol{y}}\} & \text{if } q>r \\ 0 & \text{if } q=r \end{cases} \qquad (2.31)
$$

The implicit relation (2.23) rewrites:

$$
d\,[d\boldsymbol{h}_m]_{\boldsymbol{U}}(\boldsymbol{U}_o,\boldsymbol{C}_{\boldsymbol{x}}) = -d\,[d\boldsymbol{h}_m]_{\boldsymbol{C}}(\boldsymbol{U}_o,\boldsymbol{C}_{\boldsymbol{x}}) + o(d\boldsymbol{U}, d\boldsymbol{C}) \qquad (2.32)
$$

where, for Υ_1 and for any $1 \le q,r \le P$:

$$
d\,[h_1(\boldsymbol{U},\boldsymbol{C})_{qr}]_{\boldsymbol{U}} = \sum_{q',r'=1}^{P} \Theta^{q'r'}_{qr} d\mu_{q'r'} \qquad (2.33)
$$

$$
d\,[h_1(\boldsymbol{U},\boldsymbol{C})_{qr}]_{\boldsymbol{C}} = \sum_{i,j,k,l=1}^{P} \Theta^{ijkl}_{qr} dC^{jk}_{i\ell,\boldsymbol{x}} \qquad (2.34)
$$

where $\Theta^{q'r'}_{qr}$ and Θ^{ijkl}_{qr} are given in appendix E. Similar (but more complicated) relations, derived for Υ_2 and Υ_3, are not reported here for reasons of space.

Simulations

Empirical variance estimates. Simulations have been run for $P = 2$ independent QPSK sources, in the presence of Gaussian complex circular noise. The mixing matrix was of the form

$$
\begin{pmatrix} \cos\theta & \sin\theta\, e^{\mathrm{j}\varphi} \\ -\sin\theta\, e^{-\mathrm{j}\varphi} & \cos\theta \end{pmatrix}
$$

with $\theta = \pi/7$ and $\varphi = \pi/7$. The separating matrix has been computed using algorithms reported in [16], [7], and [18]. In order to obtain reliable variance estimates, 100 independent trials

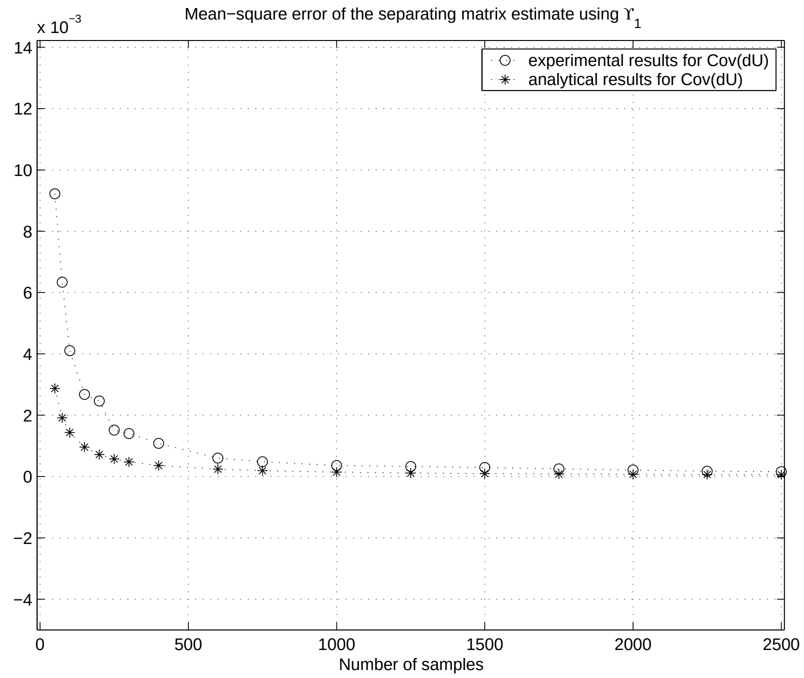

Figure 2.8: Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_1(\boldsymbol{U})$.

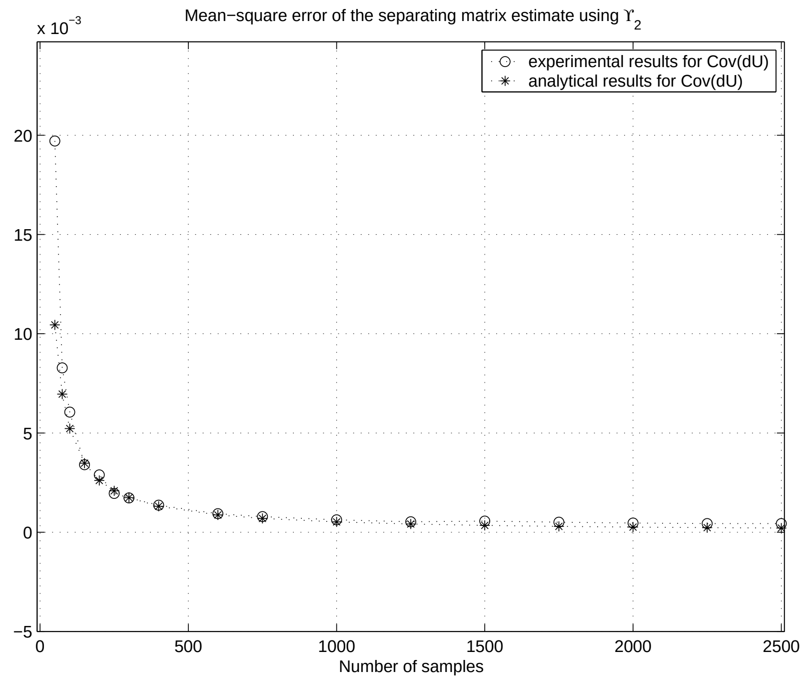

Figure 2.9: Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_2(\boldsymbol{U})$.

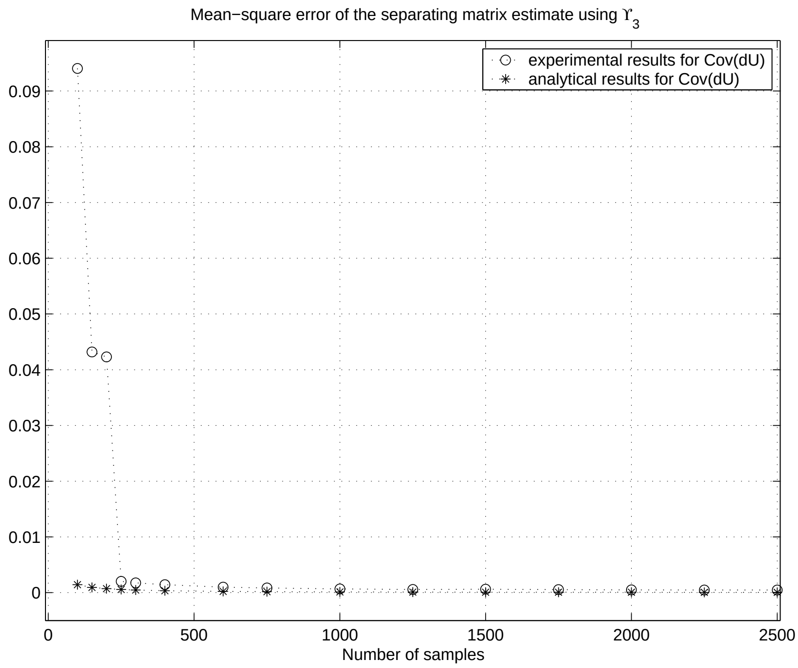

Figure 2.10: Variance of estimated separating matrix $\boldsymbol{U}$ obtained by maximization of $\Upsilon_3(\boldsymbol{U})$.

have been run, and the variances of each of the four entries $\hat{U}_{ij}$ has been estimated. In figures 2.8 to 2.10, we have plotted the sum of variances $\sum_{i=1}^{2} \mathrm{Var}\{\hat{U}_{ii}\}$ as a function of the sample size.

Theoretical asymptotic variance. In order to compute the theoretical variance, it was necessary to first calculate all the cumulants of even order up to eight. For this purpose, the multilinearity property of cumulants has been used, yielding the cumulants of the two outputs of a linear transform as a function of those of its inputs. For QPSK sources, we have the following (omitting subscript $\boldsymbol{s}$ in C_s):

$$\begin{array}{lllll} C_{1,1} = 0 & C_1^1 = 1 & C_{1,1,1,1} = 1 & C_{1,1,1}^1 = 0 & C_{1,1}^{1,1} = -1 \\ C_{1,1,1,1,1,1} = 0 & C_{1,1,1,1,1}^1 = -4 & C_{1,1,1,1}^{1,1} = 0 & C_{1,1,1}^{1,1,1} = 4 & C_{1,1,1,1,1,1,1,1} = -34 \\ C_{1,1,1,1,1,1,1}^1 = 0 & C_{1,1,1,1,1,1}^{1,1} = 34 & C_{1,1,1,1,1}^{1,1,1} = 0 & C_{1,1,1,1}^{1,1,1,1} = -33 & \end{array} \tag{2.35}$$

General formulas are given in appendix D. Since $P = 2$ is a simple case, first and second order derivatives can be computed directly in terms of $d\theta$ and $d\varphi$, and the 2×2 matrix $\boldsymbol{F}_1$ obtained is invertible. Thus, expressions such as (2.28) to (2.33) did not need to be used. On the other hand, expression (2.26) is central in this calculation. Results are reported in the figures 2.8 to 2.10, and show a good accordance with empirical results for large samples.

2.2.4 Concluding remarks

The whole analytical calculus allows one to write, for each contrast in (2.27), the link between the covariance of the unbiased estimated separator $\widehat{\boldsymbol{U}}$ and the covariance of the unbiased estimated cumulant $\hat{\boldsymbol{C}}_{\boldsymbol{x}}$. Using also (2.26), the asymptotic performances of Υ_1, Υ_2 and Υ_3 can be compared to each other, and to those obtained by averaging independent trials. Two conclusions can be drawn: first, empirical performances tend to reach theoretical limits as sample sizes tend to infinity, which justifies our approach. Second, the contrast leading to the smallest variance is Υ_1.

Chapter 3

BIRTH or SixO statistics for the underdetermined case

In order to process underdetermined mixtures of sources, an extension of the ICAR method to order 6, named BIRTH (Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix), is proposed, able to blindly identify the steering vectors of up to $P = N^2 - N + 1$ sources for arrays of N sensors with spatial diversity only, and up to $P = N^2$ for those with angular and polarization diversity. The sources are assumed to have non zero SixO marginal cumulants with the same sign (the latter assumption is generally verified in radiocommunications contexts). Besides, BIRTH exploits implicitly the VA concept described in [38] [13] and explicitly redundancies in the SixO statistical matrix of the data, called *hexacovariance*, without SO or FourO prewhitening.

3.1 The BIRTH Method

3.1.1 Hexacovariance property

Under assumptions (**A1**)-(**A5**) of section 1.1.2 for $q = 3$, the BIRTH method precisely exploits several redundancies in the hexacovariance matrix $\boldsymbol{H_x}$, defined by example 2 of section 1.2.2, of the data especially thanks to the multilinearity property. However, instead of using the classical matrix form of the multilinearity property, described by

$$\boldsymbol{H_x} = [\boldsymbol{A} \otimes \boldsymbol{A} \otimes \boldsymbol{A}^*]\, \boldsymbol{H_s}\, [\boldsymbol{A} \otimes \boldsymbol{A} \otimes \boldsymbol{A}^*]^{\mathsf{H}} = \left[\boldsymbol{A}^{\otimes 2} \otimes \boldsymbol{A}^*\right] \boldsymbol{H_s} \left[\boldsymbol{A}^{\otimes 2} \otimes \boldsymbol{A}^*\right]^{\mathsf{H}} \tag{3.1}$$

where the $N^3\times N^3$ $\boldsymbol{H_x}$ and the $P^3\times P^3$ $\boldsymbol{H_s}$ matrices are the hexacovariance matrices of $\boldsymbol{x}(k)$ and $\boldsymbol{s}(k)$ respectively, the BIRTH method exploits an alternative one, given by

$$\boldsymbol{H_x} = [\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]\,\boldsymbol{\mathcal{H}_s}\,[\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]^{\mathsf{H}} \tag{3.2}$$

where contrary to matrix $\boldsymbol{H_s}$, the diagonal $\boldsymbol{\mathcal{H}_s} = \mathrm{diag}\Big(\Big[C_{1,1,1,\boldsymbol{s}}^{1,1,1}\, C_{2,2,2,\boldsymbol{s}}^{2,2,2}\cdots C_{P,P,P,\boldsymbol{s}}^{P,P,P}\Big]\Big)$ matrix is of full rank. On the orther hand, the $N^3\times P$ matrix $[\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]$, which is of full column rank under assumptions (**A4**) and (**A5**), is given by:

$$\begin{aligned}\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^* &= [\boldsymbol{a}_1\otimes\boldsymbol{a}_1\otimes\boldsymbol{a}_1^* \;\cdots\; \boldsymbol{a}_P\otimes\boldsymbol{a}_P\otimes\boldsymbol{a}_P^*]\\ &= [[(\boldsymbol{A}\oslash\boldsymbol{A}^*)\boldsymbol{\Phi}_1]^{\mathsf{T}}\; [(\boldsymbol{A}\oslash\boldsymbol{A}^*)\boldsymbol{\Phi}_2]^{\mathsf{T}}\;\cdots\; [(\boldsymbol{A}\oslash\boldsymbol{A}^*)\boldsymbol{\Phi}_N]^{\mathsf{T}}]^{\mathsf{T}}\end{aligned} \tag{3.3}$$

with the $N^2\times P$ matrix $\boldsymbol{A}\oslash\boldsymbol{A}^*$, which is of full column rank under assumption (**A5**) for $q=3$, defined by

$$\boldsymbol{A}\oslash\boldsymbol{A}^* = [\boldsymbol{a}_1\otimes\boldsymbol{a}_1^*\;\cdots\;\boldsymbol{a}_P\otimes\boldsymbol{a}_P^*] = [[\boldsymbol{A}^*\boldsymbol{\Phi}_1]^{\mathsf{T}}\;[\boldsymbol{A}^*\boldsymbol{\Phi}_2]^{\mathsf{T}}\;\cdots\;[\boldsymbol{A}^*\boldsymbol{\Phi}_N]^{\mathsf{T}}]^{\mathsf{T}} \tag{3.4}$$

and:

$$\boldsymbol{\Phi}_n = \mathrm{Diag}\left[\begin{array}{cccc} A(n,1) & A(n,2) & \cdots & A(n,P)\end{array}\right] \tag{3.5}$$

In other words, the non zero elements of the $P\times P$ diagonal matrix $\boldsymbol{\Phi}_n$ are the components of the n-th row of matrix $\boldsymbol{A}$.

3.1.2 Data structure

If SixO marginal source cumulants are strictly positive (**A3**), then a square root of $\boldsymbol{H_x}$, called $\boldsymbol{H_x}^{1/2}$, has to be computed (if these cumulants are strictly negative, the $-\boldsymbol{H_x}$ matrix has to be considered for computing the square root) for example as following :

$$\boldsymbol{H_x}^{1/2} = \boldsymbol{E_s}\,\boldsymbol{L_s}^{1/2} = [\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]\,\boldsymbol{\mathcal{H}_s}^{1/2}\,\boldsymbol{V}^{\mathsf{H}} \tag{3.6}$$

where $\boldsymbol{L_s}$ ($\boldsymbol{L_s}^{1/2}$ denotes a square root of $\boldsymbol{L_s}$) is the $P\times P$ real-valued diagonal matrix of the P non zero eigen-values of $\boldsymbol{H_x}$ and $\boldsymbol{E_s}$ is the $N^3\times P$ matrix of the associated orthonormalized eigen-vectors. For a full rank $[\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]$ matrix, it is possible to verify that (**A3**) is equivalent to assuming that the diagonal elements of $\boldsymbol{L_s}$ are not null and have also the same sign. In addition, (3.6) shows the link between $\boldsymbol{H_x}^{1/2}$ and $[\boldsymbol{A}^{\oslash 2}\oslash \boldsymbol{A}^*]$ where $\boldsymbol{V}$ is a unitary matrix. Finally, (3.6) and (3.3) allow to prove the link between $\boldsymbol{H_x}^{1/2}$ and $\boldsymbol{A}\oslash\boldsymbol{A}^*$, as follows:

$$\begin{aligned}\boldsymbol{H_x}^{1/2} &= \Big[[\boldsymbol{A}\oslash\boldsymbol{A}^*\boldsymbol{\Phi}_1\boldsymbol{\mathcal{H}_s}^{1/2}\boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}}\cdots\;[\boldsymbol{A}\oslash\boldsymbol{A}^*\boldsymbol{\Phi}_N\boldsymbol{\mathcal{H}_s}^{1/2}\boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}}\Big]^{\mathsf{T}}\\ &= [\;\boldsymbol{\Gamma}_1{}^{\mathsf{T}}\;\;\boldsymbol{\Gamma}_2{}^{\mathsf{T}}\;\;\cdots\;\;\boldsymbol{\Gamma}_N{}^{\mathsf{T}}\;]^{\mathsf{T}}\end{aligned} \tag{3.7}$$

where $\mathbf{\Gamma}_n = \boldsymbol{A} \oslash \boldsymbol{A}^* \mathbf{\Phi}_n \, \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2} \, \boldsymbol{V}^{\mathsf{H}}$ is the n-th $N^2 \times P$ matrix block of $\boldsymbol{H}_{\boldsymbol{x}}^{1/2}$.

3.1.3 SixO blind identification step

In this section, the purpose is to exploit the information contained in the $\boldsymbol{H}_{\boldsymbol{x}}^{1/2}$ matrix to blindly identify $\boldsymbol{A}$. Indeed, the $\boldsymbol{V}$ matrix diagonalizes the $N(N-1)$ $\mathbf{\Theta}_{n_1,n_2}$ matrices described, for all $1 \le n_1 \ne n_2 \le N$, by:

$$\mathbf{\Theta}_{n_1,n_2} = \mathbf{\Gamma}_{n_1}^{\sharp} \mathbf{\Gamma}_{n_2} = \boldsymbol{V} \, \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{-1/2} \, \mathbf{\Phi}_{n_1}^{-1} \, \mathbf{\Phi}_{n_2} \, \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2} \, \boldsymbol{V}^{\mathsf{H}} = \boldsymbol{V} \, \mathbf{\Phi}_{n_1}^{-1} \, \mathbf{\Phi}_{n_2} \, \boldsymbol{V}^{\mathsf{H}} \tag{3.8}$$

where $\sharp$ denotes the pseudo-inverse operator and where the $\boldsymbol{D}_{n_1,n_2} = \mathbf{\Phi}_{n_1}^{-1} \mathbf{\Phi}_{n_2}$ matrices are diagonal. Thus, by construction, the rank of $\mathbf{\Theta}_{n_1,n_2}$, denoted by $\mathrm{rk}(\mathbf{\Theta}_{n_1,n_2})$, cannot exceed the $\min(\mathrm{rk}(\mathbf{\Gamma}_{n_1}), \mathrm{rk}(\mathbf{\Gamma}_{n_2})) = \min(P, \mathrm{rk}(\boldsymbol{A} \oslash \boldsymbol{A}^*))$ value, hence another bound of the maxi number of sources, P. The unitary $\boldsymbol{V}_{sol} = \boldsymbol{V} \boldsymbol{\mathcal{T}}$ matrix, solution to the previous problem of joint diagonalization to within a unitary trivial matrix $\boldsymbol{\mathcal{T}}$, allows one, in accordance with (3.6), to recover $[\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*]$ to within a trivial matrix as follows :

$$\boldsymbol{H}_{\boldsymbol{x}}^{1/2} \, \boldsymbol{V}_{sol} = [\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*] \, \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2} \, \boldsymbol{\mathcal{T}} \tag{3.9}$$

Since, consistent with (3.3) and (3.4), the (3.9) equation can also be written as follows:

$$\begin{aligned} \boldsymbol{H}_{\boldsymbol{x}}^{1/2} \, \boldsymbol{V} &= \left[[\boldsymbol{A}^* \mathbf{\Phi}_1 \mathbf{\Phi}_1 \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2}]^{\mathsf{T}} \ \cdots [\boldsymbol{A}^* \mathbf{\Phi}_N \mathbf{\Phi}_1 \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2}]^{\mathsf{T}} \ \cdots [\boldsymbol{A}^* \mathbf{\Phi}_N \mathbf{\Phi}_N \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2}]^{\mathsf{T}} \right]^{\mathsf{T}} \\ &= [\mathbf{\Sigma}_1{}^{\mathsf{T}} \ \mathbf{\Sigma}_2{}^{\mathsf{T}} \cdots \mathbf{\Sigma}_{N^2}{}^{\mathsf{T}}]^{\mathsf{T}} \end{aligned} \tag{3.10}$$

So, the $\mathbf{\Sigma}_1$ matrix block made up of the first N-th rows of matrix $\boldsymbol{H}_{\boldsymbol{x}}^{1/2} \boldsymbol{V}_{sol}$ corresponds to within a trivial matrix to $\boldsymbol{A}^*$ such as:

$$\mathbf{\Sigma}_1 = \boldsymbol{A}^* \ [\mathbf{\Phi}_1]^2 \ \boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2} \, \boldsymbol{\mathcal{T}} \tag{3.11}$$

where $\boldsymbol{\mathcal{H}}_{\boldsymbol{s}}^{1/2}$ and $\mathbf{\Phi}_n$, for all $1 \le n \le N$, are diagonal matrices.

3.1.4 Implementation of the BIRTH method

The different steps of the BIRTH method are summarized hereafter when K samples of the observations, $\boldsymbol{x}(k)$ $(1 \le k \le K)$, are available.

Step1 Compute the estimate $\widehat{\boldsymbol{H}_{\boldsymbol{x}}}$ of $\boldsymbol{H}_{\boldsymbol{x}}$ from the K samples $\boldsymbol{x}(k)$ using for instance appendix D and the empirical estimate of moments, unbiased and consistent for ergodic stationary sources.

Step2 EigenValue Decomposition (EVD) of the matrix $\widehat{\boldsymbol{H}}_{\boldsymbol{x}}$, estimation, $\widehat{P}$, of the number of sources P and restriction of this EVD to the $\widehat{P}$ principal components: $\widehat{\boldsymbol{H}}_{\boldsymbol{x}} = \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\,\widehat{\boldsymbol{L}}_{\boldsymbol{s}}\,\widehat{\boldsymbol{E}}_{\boldsymbol{s}}^{\mathsf{H}}$, where $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$ is the diagonal matrix of the P eigen-values of largest moduli and $\widehat{\boldsymbol{E}}_{\boldsymbol{s}}$ is the matrix of the associated eigenvectors.

Step3 Computation of a square root matrix $\widehat{\boldsymbol{H}}_{\boldsymbol{x}}^{1/2}$ of $\widehat{\boldsymbol{H}}_{\boldsymbol{x}}$: $\widehat{\boldsymbol{H}}_{\boldsymbol{x}}^{1/2} = \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\,|\widehat{\boldsymbol{L}}_{\boldsymbol{s}}|^{1/2}$, where $|\cdot|$ denotes the complex modulus operator.

Step4 Computation from $\widehat{\boldsymbol{H}}_{\boldsymbol{x}}^{1/2}$ of the $\widehat{\boldsymbol{\Theta}}_{n_1,n_2} = [\widehat{\boldsymbol{\Gamma}}_{n_1}^{\sharp}\,\widehat{\boldsymbol{\Gamma}}_{n_2}]$ matrices for all $1 \leq n_1 \neq n_2 \leq N$, and estimation, $\widehat{\boldsymbol{V}}_{sol}$, of the unitary matrix $\boldsymbol{V}_{sol}$ from the joint diagonalization of the $N(N-1)$ matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2}$.

Step5 Estimation $\widehat{\boldsymbol{A}}$ of the $\boldsymbol{A}$ mixture matrix taking the matrix block made up of the first N-th rows of $[\widehat{\boldsymbol{H}}_{\boldsymbol{x}}^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]^*$.

Step6 If $\boldsymbol{A}$ is an overdetermined mixture, estimation of the signal vector $\boldsymbol{s}(k)$ for any value k, by applying to $\boldsymbol{x}(k)$ the SMF source separator defined by $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1}\widehat{\boldsymbol{A}}$, where $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$ is an estimate of $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}_{2,\boldsymbol{x}}^{0}$.

3.2 BIRTH improvements

Once matrix $[\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*]$ has been estimated, it has to allow to recover matrix $\boldsymbol{A}$. In fact, we showed in section 3.1.3 that it was sufficient to take as estimate $\widehat{\boldsymbol{A}}$ of $\boldsymbol{A}$ the matrix block made up of the N first rows of the conjugate of matrix $\widehat{\boldsymbol{\mathcal{A}}_3}$, where $\widehat{\boldsymbol{\mathcal{A}}_3}$ denotes the estimate of $[\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*]$. The latter approach will be referred to as **Method 1** in the sequel. Although method 1 appears to have a low computational complexity, it does not exploit all redundancies present in $[\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*]$. So we propose now other methods such as:

Method 2: *Extract the N^2 matrix blocks, of size $N \times P$, made up of the successive rows of the conjugate of matrix $\widehat{\boldsymbol{\mathcal{A}}_3}$ and equal to $\boldsymbol{\Sigma}_m^*$ ($1 \leq m \leq N^2$) (3.10); take as estimate the average of these N^2 blocks.*

Method 3: *Fully exploit each column vector $\widehat{\boldsymbol{b}}_p$ of $\widehat{\boldsymbol{\mathcal{A}}_3}$. In order to do this, first extract, from vector $\widehat{\boldsymbol{b}}_p$, the N vectors $\widehat{\boldsymbol{b}}_p(n)$ of size $N^2 \times 1$, then remodel them into N matrices $\widehat{\boldsymbol{B}}_p(n)$ of size $N \times N$, and finally build the matrix whose p-th column vector is the eigenvector (approximately)* *in common within the N matrices $\widehat{\boldsymbol{B}}_p(n)^*$ ($1 \leq n \leq N$) and associated with the largest eigenvalue using the Joint Approximate Diagonalization (JAD) algorithm described in [9].*

Methods 1 and 2 ensue immediately from the structure of matrix $\widehat{\boldsymbol{\mathcal{A}}_3}$. In fact, it has been shown in (3.10) that matrix $\widehat{\boldsymbol{\mathcal{A}}_3}$ may be written as:

$$\widehat{\boldsymbol{\mathcal{A}}_3} = \boldsymbol{\mathcal{A}}_3\, \boldsymbol{\mathcal{T}} = [[\boldsymbol{A}^* \boldsymbol{\mathcal{T}}_1]^\mathsf{T}\ [\boldsymbol{A}^* \boldsymbol{\mathcal{T}}_2]^\mathsf{T} \cdots [\boldsymbol{A}^* \boldsymbol{\mathcal{T}}_{N^2}]^\mathsf{T}]^\mathsf{T} \tag{3.12}$$

where $\boldsymbol{\mathcal{T}}$ and the N^2 matrices $\boldsymbol{\mathcal{T}}_n$ ($1 \leq n \leq N^2$), of size $P \times P$, are trivial. As for method 3, it is shown in appendix C taking $q = 3$ that

$$\forall n,\ 1 \leq n \leq N, \quad \widehat{\boldsymbol{b}}_p(n) \propto \left[\boldsymbol{a}_{\xi(p)} \otimes \boldsymbol{a}^*_{\xi(p)}\right] \tag{3.13}$$

where $\xi(\cdot)$ is a bijective function of $\{1, 2, \ldots, P\}$ into itself (i.e. a permutation function). Then it is straightforward to show that

$$\forall n,\ 1 \leq n \leq N, \quad \widehat{\boldsymbol{B}}_p(n) \propto \left[\boldsymbol{a}_{\xi(p)}\, \boldsymbol{a}_{\xi(p)}{}^\mathsf{H}\right]^* \tag{3.14}$$

and hence the method 3 result. Note that, although the JAD algorithm [9] is restricted to unitary joint diagonalizers, it can be used in method 3 since matrices $\widehat{\boldsymbol{B}}_p(n)^*$ are of rank 1, from (3.14).

Method 4: *The fourth method we consider performs a unrestricted (non-unitary) LS joint diagonalization scheme, as for instance the one described by Yeredor in [73], yielding probably a better LS fit.*

We propose a fifth method, named **Method 5** and based on the the following mathematical problem:

Problem 1 *Given M matrices $\boldsymbol{\Xi}_m$, $1 \leq m \leq M$, each of size $N \times P$, find an $N \times P$ matrix $\boldsymbol{A}$, and $P \times P$ diagonal matrices $\boldsymbol{D}_m$ of unit Frobenius norm such that*

$$\boldsymbol{\Xi}_m\, \boldsymbol{D}_m \ \approx\ \boldsymbol{A} \tag{3.15}$$

Matrices $\boldsymbol{A}$ and $\boldsymbol{D}_m$ can be obtained as stationary values of the Least Squares (LS) criterion below:

$$\varepsilon = \sum_{m=1}^{M} \|\boldsymbol{\Xi}_m \boldsymbol{D}_m - \boldsymbol{A}\|_F^2 \tag{3.16}$$

where $\|\boldsymbol{B}\|_F$ is the Frobenius norm of matrix $\boldsymbol{B}$. As a consequence, they satisfy the following system of equations, obtained by cancelling the gradient of ε with respect to $\boldsymbol{D}_m$ and $\boldsymbol{A}$:

$$\begin{cases} \forall\, m,\ \ \forall\, p,\ \ \{\, \boldsymbol{\Xi}_m^{\mathsf{H}}(\boldsymbol{\Xi}_m \boldsymbol{D}_m - \boldsymbol{A})\,\}(p,p) = 0 \\ \forall\ (n,p),\ \ \sum_{m=1}^{M} \{\, \boldsymbol{\Xi}_m \boldsymbol{D}_m - \boldsymbol{A}\,\}(n,p) = 0 \end{cases} \tag{3.17}$$

where $\boldsymbol{B}(n,p)$ is the (n,p)-th component of matrix $\boldsymbol{B}$. It is then not hard to obtain the closed form expression for $\boldsymbol{A}$:

$$\boldsymbol{A} = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{\Xi}_m \boldsymbol{D}_m \tag{3.18}$$

By plugging back this solution in system (3.17), one gets after some manipulations:

$$\forall p,\ 1 \leq p \leq P, \qquad \boldsymbol{F}_p\, \boldsymbol{d}_p = 0 \tag{3.19}$$

where

$$\boldsymbol{F}_p(m_1, m_2) = \begin{cases} (M-1)\, \{\boldsymbol{\Xi}_{m_1}^{\mathsf{H}}\, \boldsymbol{\Xi}_{m_1}\}(p,p) \ \text{ if } m_1 = m_2 \\ -\, \{\boldsymbol{\Xi}_{m_1}^{\mathsf{H}}\, \boldsymbol{\Xi}_{m_2}\}(p,p) \ \text{ otherwise} \end{cases} \tag{3.20}$$

and where $\boldsymbol{d}_p = \begin{bmatrix} \boldsymbol{D}_1(p,p) & \boldsymbol{D}_2(p,p) & \cdots & \boldsymbol{D}_M(p,p) \end{bmatrix}^{\mathsf{T}}$. In other words, the solution to the LS problem under the constraint that, for any fixed index p, $\sum_m |\boldsymbol{D}_m(p,p)\,|^2 = 1$ is obtained when the vector $\boldsymbol{d}_p$ is the right singular vector of matrix $\boldsymbol{F}_p$ associated with the minimal singular value. Once every entry $\boldsymbol{D}_m(p,p)$ is obtained, matrix $\boldsymbol{A}$ can be calculated thanks to (3.18). This solution is thus not iterative (though we could possibly run alternate iterations).

So method 5 is defined by solving problem 1 taking for matrices $\boldsymbol{\Xi}_m$ the $M = N^2$ matrix blocks $\boldsymbol{\Sigma}_m^*$ (3.10) of size $N \times P$, made up of the successive rows of the conjugate of matrix $\widehat{\boldsymbol{\mathcal{A}}_3}$. The latter algorithm does not take into account the fact that diagonal matrices $\boldsymbol{D}_m$ should contain products of entries of $\boldsymbol{A}$, and is therefore expected to yield less accurate results. However, subsequent simulations demonstrate that the loss in performance is little compared to the gain in computational complexity.

3.3 Identifiability

3.3.1 The BIRTH approach

Following the development of the previous sections, it appears that the BIRTH method is able to identify, from an array of N sensors, the steering vectors of P ($P \leq N^2$) non Gaussian sources having SixO marginal cumulants with the same sign, provided that the $\boldsymbol{A}\oslash\boldsymbol{A}^*$ matrix has full rank P, i.e. that the *virtual steering vectors* $[\boldsymbol{a}_p\otimes\boldsymbol{a}_p^*]$ $(1\leq p\leq P)$ for the considered array of N sensors remain linearly independent. In addition, it has been shown in [13] that the vector $[\boldsymbol{a}_p\otimes\boldsymbol{a}_p^*]$ can also be considered as a *true steering vector* but for a *virtual array* of $\mathcal{N}_4$ different sensors. This especially means that $N^2-\mathcal{N}_4$ components of each vector $\left[\boldsymbol{a}_p\otimes\boldsymbol{a}_p^*\right]$ are redundant elements which bring no information. The rank of $\boldsymbol{A}\oslash\boldsymbol{A}^*$ cannot therefore be greater than $\mathcal{N}_4$ and is equal to $\min(\mathcal{N}_4, P)$ when $\boldsymbol{A}$ is of full rank. In these conditions, since $\boldsymbol{A}\oslash\boldsymbol{A}^*$ has full rank P, $\min(\mathcal{N}_4, P)$ is equal to P, which implies $P\leq\mathcal{N}_4$. So the BIRTH algorithm is able to process up to $\mathcal{N}_4$ sources, where $\mathcal{N}_4$ is the number of different Virtual Sensors (VS) of the VA associated with the chosen array of N sensors. Quantity $\mathcal{N}_4$ will be described in detail in section 4.2.1. So, it is shown in [13] that using an array with spatial diversity only, as for instance a UCA, $\mathcal{N}_4$ may be equal to N^2-N+1, whereas using an array with angular and polarization diversity, the $\mathcal{N}_4$ number may attain N^2.

3.3.2 Impact of the hexacovariance structure

According to [13], the $\mathcal{N}_4$ number is directly related to both kind of sensors and geometry of the true array of N sensors. For example, a Uniform Linear Array (ULA) of identical sensors generates a VA of $\mathcal{N}_4 = 2N-1$ different VS, whereas for most of other arrays $\mathcal{N}_4 = N^2-N+1$. Nevertheless, both kind of sensors and geometry of the true array are not the only factor which the $\mathcal{N}_4$ number depends on. Indeed the way data SixO cumulants are mapped in $\boldsymbol{H_x}$ is also a parameter which affects the number of VS. To show this, consider the following way to sort SixO Cumulants in the hexacovariance matrix:

$$H_{\boldsymbol{x}}\Big(I_1^0, I_2^0\Big) = C_{i_1,i_2,i_3,\boldsymbol{x}}^{i_4,i_5,i_6} \tag{3.21}$$

where $H_{\boldsymbol{x}}(I_1^0, I_2^0)$ is the (I_1^0, I_2^0)-th entry $(1\leq I_1^0,\ I_2^0\leq N^3)$ of $\boldsymbol{H_x}$ and where for all $1\leq i_1, i_2, i_3, i_4, i_5, i_6\leq N$,

$$\begin{aligned} I_1^1 &= \varphi([i_1\ \ i_2\ \ i_3]) = N(N(i_1-1)+i_2-1)+i_3 \\ I_2^1 &= \varphi([i_4\ \ i_5\ \ i_6]) = N(N(i_4-1)+i_5-1)+i_6 \end{aligned} \tag{3.22}$$

what implies:

$$\boldsymbol{H_x} = [\boldsymbol{A}^{\oslash 3}]\,\boldsymbol{\mathcal{H}_s}\,[\boldsymbol{A}^{\oslash 3}]^{\mathsf{H}} \tag{3.23}$$

The FourO virtual array associated with this expression (the corresponding *virtual steering vectors*, for all $1\leq p\leq P$, are thus of the form $[\boldsymbol{a}_p\otimes\boldsymbol{a}_p]$) is generally different from the one obtained from

(3.1). In particular, the VA associated with (3.23) and a UCA of odd N identical sensors, is caracterized by $\mathcal{N}_4 = N(N+1)/2$ different VS, whereas the one associated with (3.1) and an UCA of odd N identical sensors, is caracterized by $\mathcal{N}_4 = N^2 - N + 1$ different VS. For any $N \geq 2$, the $N^2 - N + 1$ value is obviously greater than $N(N+1)/2$.

***Proof*:** *Note that the $(r,q)^{\text{th}}$ VS associated with the p-th source and the UCA of N sensors is such that:*

$$[\boldsymbol{a}_p \otimes \boldsymbol{a}_p]_r^q = \exp\{\mathrm{j}2\pi[x_r^q \cos(\theta_p)\cos(\phi_p) + y_r^q \sin(\theta_p)\cos(\phi_p)]\} \tag{3.24}$$

$(x_r^q, y_r^q, 0) = ((R_r^q/\lambda)\cos(\varphi_r^q), (R_r^q/\lambda)\sin(\varphi_r^q), 0)$ *are the coordinates of the $(r,q)^{\text{th}}$ VS ($1 \leq r, q \leq N$) where $R_r^q = 2R\cos((\varphi_r - \varphi_q)/2)$ and $\varphi_r^q = (\varphi_r + \varphi_q)/2$ since it is always possible to choose a coordinate system in which the n-th sensor of the true array has the coordinates $(x_n, y_n, 0) = (R\cos(\varphi_n), R\sin(\varphi_n), 0)$ where R is the radius and $\varphi_n = 2\pi(n-1)/N$. It is thus easy to deduce from the previous equations that the VS that are not at coordinates $(0,0,0)$ lie on $(N+1)/2$ different circles if N is odd or $N/2$ if N is even and that there are VS at coordinates $(0,0,0)$ only if N is even. Moreover, for odd values of N, N different VS lie on each circle of the VA uniformly spaced. As a consequence, this VA, for odd values of N, has $\mathcal{N}_4 = N(N+1)/2$ different VS. As to the second result, it is given by [13].*

It is important to explain that if both FourO VA obtained from (3.1) and (3.23) are not equivalent, however, they have the same radiation pattern.

***Proof*:** *The radiation pattern of a $[\boldsymbol{b}(\theta_p, \phi_p)]_{1 \leq p \leq P}$ VA is defined by:*

$$\begin{aligned} &\forall\ (\theta, \phi), \quad \forall 1 \leq p \leq P, \\ &\quad c((\theta,\phi), \boldsymbol{b}(\theta_p, \phi_p)) = \frac{|\langle\, \boldsymbol{b}(\theta,\phi), \boldsymbol{b}(\theta_p,\phi_p)\,\rangle|}{\|\boldsymbol{b}(\theta,\phi)\|^2\ \|\boldsymbol{b}(\theta_p,\phi_p)\|^2} \end{aligned} \tag{3.25}$$

where θ_p, ϕ_p, $|\cdot|$, $\langle\cdot,\cdot\rangle$, $\|\cdot\|$ denote azimuth and elevation angles of the p^{th} source, the complex modulus, the scalar product and the norm operators, respectively. Since for any (θ,ϕ) and for each source p, both $c\Big((\theta,\phi), \Big[\boldsymbol{a}_p \otimes \boldsymbol{a}_p^*\Big]\Big)$ and $c((\theta,\phi), [\boldsymbol{a}_p \otimes \boldsymbol{a}_p])$ values are equal.

These results are illustrated by figures 3.1 to 3.3, which show the identical radiation pattern of both FourO VA of a UCA of five identical sensors, and the geometry of each VA, respectively.

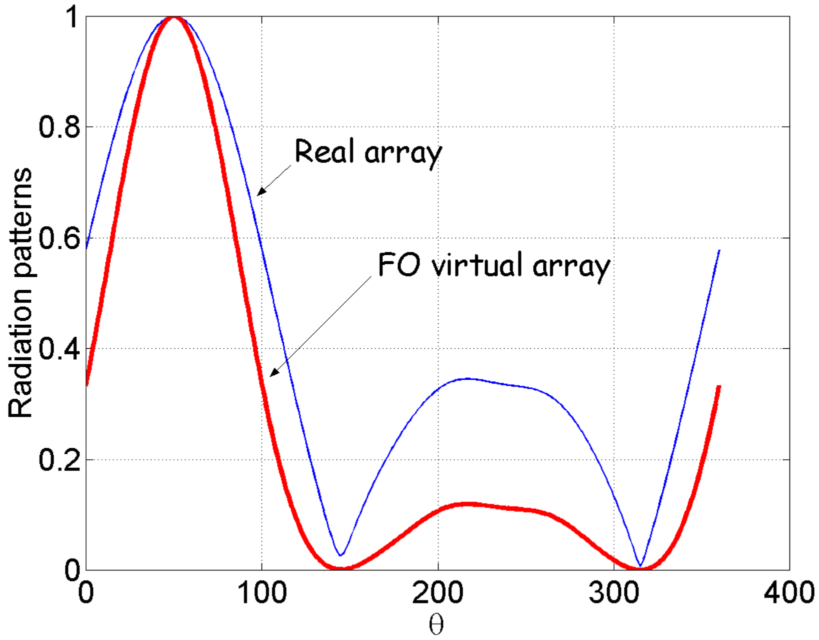

Figure 3.1: Fourth order virtual array radiation pattern ($N = 5$)

3.4 Simulations

The performance criteria used in this section to evaluate the quality of the BMI process has been presented in section 1.1.3.

3.4.1 Simple BIRTH

So, to illustrate the results of the simplified version of BIRTH described in section 3.1.3 and referred to as method 1, we assume that $P = 2$ statistically independent sources, i.e. 2 non filtered QPSK and 1 non filtered BPSK, are received by a linear array of $N = 2$ sensors of radius R such that $R/\lambda = 0.55$ (λ: wavelength). The 3 sources, assumed synchronized, have the same input SNR (Signal to Noise Ratio) of 20 dB with a symbol period $T = 4T_e$, where T_e is the sample period. The normalized marginal source cumulants are $\kappa^{111}_{111,\mathrm{QPSK}} = \kappa^{222}_{222,\mathrm{QPSK}} = 4$ and $\kappa^{333}_{333,\mathrm{BPSK}} = 16$ according to appendix D. The direction of arrival of the sources are such that $\theta_1 = 50°$, $\theta_2 = 136°$, $\theta_3 = 29.5°$, $\phi_1 = \phi_2 = \phi_3 = 0°$ and the associated carrier frequencies verify $\Delta f_1 T_e = 1/3$, $\Delta f_2 T_e = 1/2$ and $\Delta f_3 T_e = 0$. We apply the COM1 [18], COM2 [16], JADE [8], S3C2 [17] and BIRTH methods, and the performance α_p for $p = 1 \ldots 3$ is computed and averaged over 200 realizations.

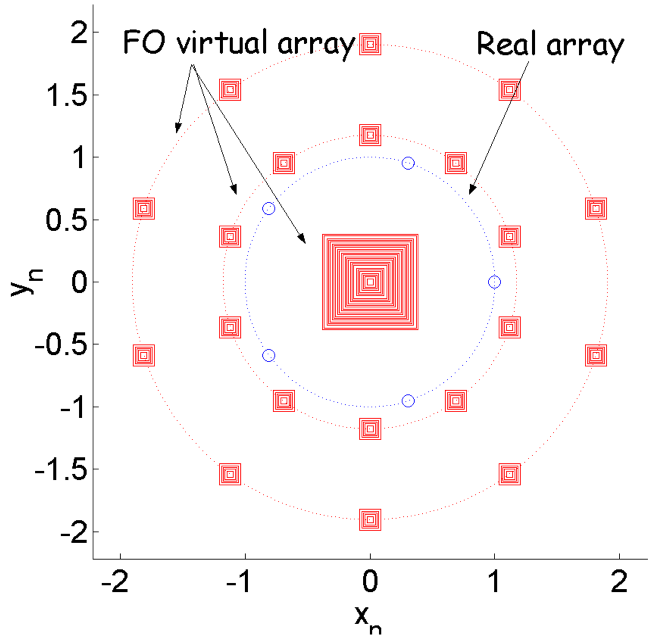

Figure 3.2: FO* virtual array defined by $\left[\boldsymbol{a}_p \otimes \boldsymbol{a}_p^*\right]$ ($N = 5$)

Under the previous assumptions, figure 3.4 shows the variations of α_3 (source 3 performance) at the output of the COM1, COM2, JADE, S3C2 and BIRTH algorithms as a function of the number of samples. The COM1, COM2, JADE methods obviously find difficulties in well identifying the steering vector of the source 3 in a underdetermined context. The S3C2 method gives better results. As to the BIRTH process, it completely succeeds in identifying the steering vector. Figure 3.5 shows, in the same context, all the α_p at the output of the BIRTH method as a function of samples. Note the decreasing values toward zero of all the previous coefficients as the number of samples increases. In addition, figure 3.6 displays the variations of α_3 (source 3 performance) at the output of the COM1, COM2, JADE, S3C2 and BIRTH methods as a function of SNR. Likewise, the COM1, COM2, JADE algorithms do not identify the steering vector of the source 3 in an underdetermined context even when the SNR increases. The S3C2 results are more pleasing. As to the BIRTH process, it performs well the identification of the steering vector even for a small value of SNR.

Finally, consider the $P = 3$ previous sources are received by a circular array of $N = 3$ sensors such that $R/\lambda = 0.55$. Figure 3.7 shows the variations of α_3 (source 3 performance) at the output of the COM1, COM2, JADE and BIRTH methods as a function of the number of samples : the

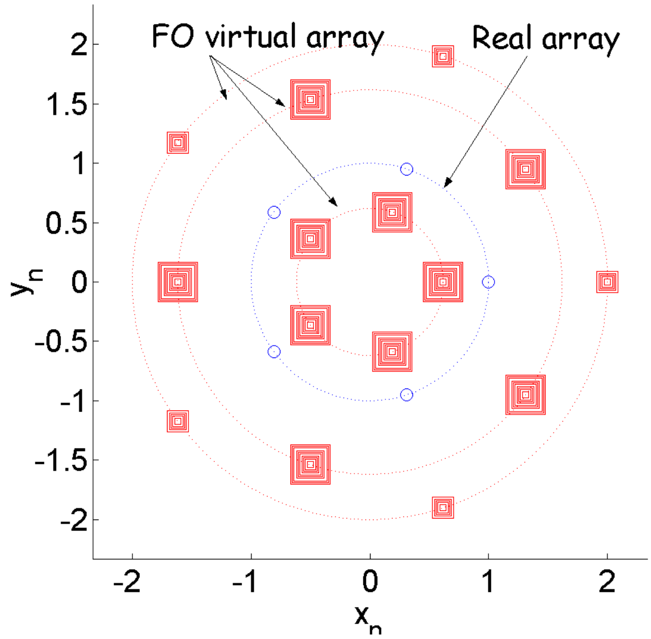

Figure 3.3: Fourth order virtual array defined by $[\boldsymbol{a}_p \otimes \boldsymbol{a}_p]$ ($N = 5$)

BIRTH method obviously works in overdetermined contexts and although SixO cumulants have to be estimated, the BIRTH algorithm converges fast enough compared with the other algorithms.

3.4.2 BIRTH improvements

We proceed in this section to two types of simulations. First, in order to test the five blind identification methods, previously described in section 3.2, independently of the BIRTH algorithm, we have generated P vectors $\widehat{\boldsymbol{b}}_p$ such that

$$\widehat{\boldsymbol{b}}_p = \boldsymbol{b}_p + \boldsymbol{\nu}_p \tag{3.26}$$

where $\boldsymbol{b}_p$ is the p-th column vector of matrix $[\boldsymbol{A}^{\oslash 2} \oslash \boldsymbol{A}^*]$ and where the $N^3 \times 1$ noise random vectors $\boldsymbol{\nu}_p$ are chosen to be Gaussian spatially and temporally white such that their covariance matrices $\boldsymbol{R}_{\boldsymbol{\nu}_p}$ verify $\boldsymbol{R}_{\boldsymbol{\nu}_p} = \sigma^2 \, \mathbf{I}_{N^3}$. We took a uniformly spaced circular array of $N = 3$ identical sensors, of radius R such that $R/\lambda = 0.55$, and $P = 12$ directional vectors. The chosen blind identification performance criterion is yet the pseudo-distance defined in section 1.1.3. We report the average of the P gaps obtained by the five methods in figure 3.8, as a function of the noise level. It can be seen that method 5 is almost as good as the most complex one, namely method 3. Second, we now incorporate the BIRTH core step in the comparison. Sources are BPSK modulated, with a raised

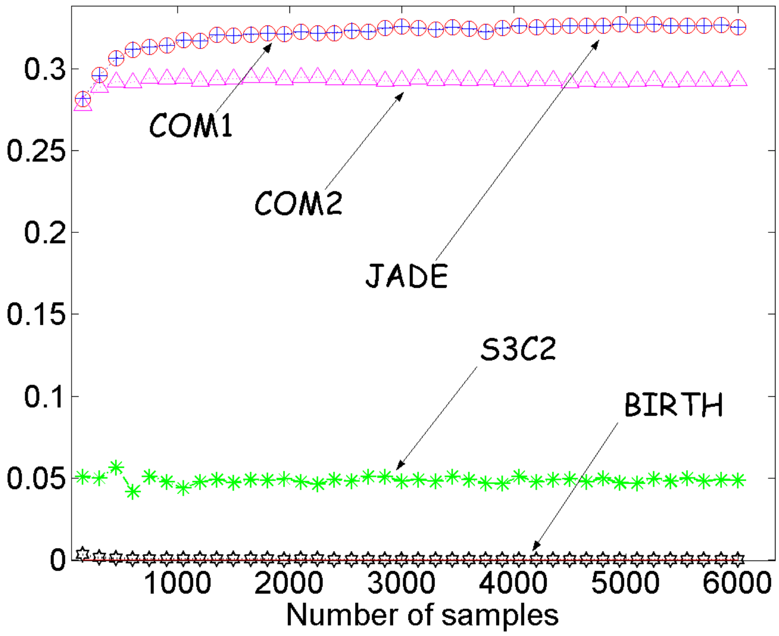

Figure 3.4: α_3 for a $\mathrm{SNR} = 20$ dB

cosine pulse shape of roll-off equal to 0.25, and assumed synchronized. Figure 3.9 shows BMI results obtained when 7 BPSK sources are received by the same array as above. Their symbol periods are equal to twice the sample period and their carrier residuals are all null. In this figure, the label "BIRTHm" corresponds to the BIRTH algorithm followed by method m of section 3.2. Again, it can be seen that the five methods can be sorted in the same way: method 3, the most complex, is followed by method 5. The latter thus appears to exhibit the best trade-off between performance and computational complexity.

3.5 Conclusion

As surveyed in introduction, there are few algorithms able to identify blindly underdetermined mixtures (*i.e.* in the absence of sparsity). This chapter has presented a new BMI method, BIRTH, in a underdetermined context, i.e. allowing to identify the steering vectors of more sources than sensors, using SixO cumulants and the FourO VA concept. The BIRTH algorithm succeeds in recovering the mixture matrix even for a small number of samples or a weak SNR. Moreover, new results as for the VA are given: both FourO VA, described in this chapter, are proved to be not equivalent. As a consequence, the way to store cumulants in the corresponding matrix affects the

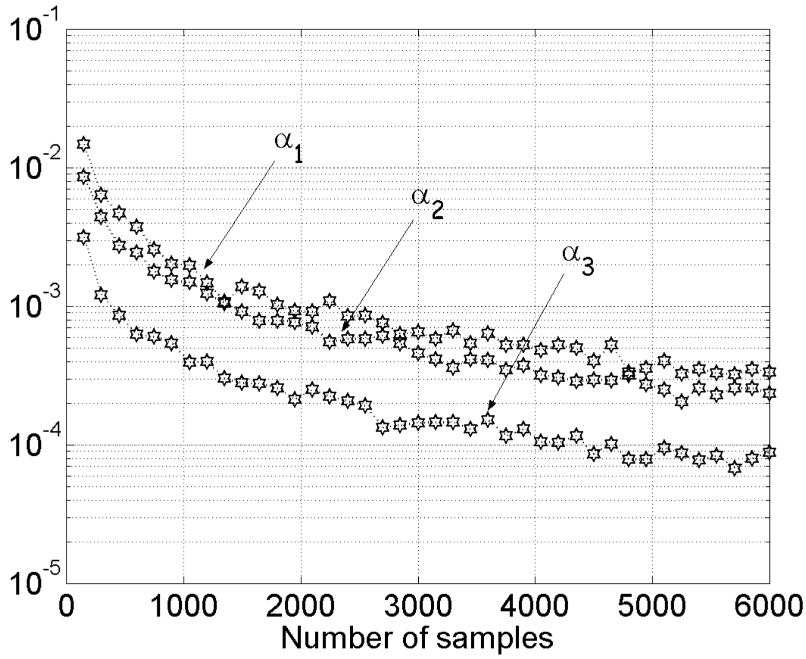

Figure 3.5: $D\left(\boldsymbol{A}, \widehat{\boldsymbol{A}}\right)$ associated with the BIRTH method

performance of the method. Finally, the BIRTH algorithm has been be improved, in particular the fifth step of (3.1.4), by proposing five methods optimizing differently the compromise between performance and complexity.

Note that the BIRTH algorithms, and in particular BIRTH3 and BIRTH5, can be used for blind beamforming. Yet, there exist techniques based on the array manifold knowledge that can handle underdetermined mixtures, such as the so-called $4-$MUSIC [60]. It could be interesting to compare its performances with the above as well, which could yield a performance bound.

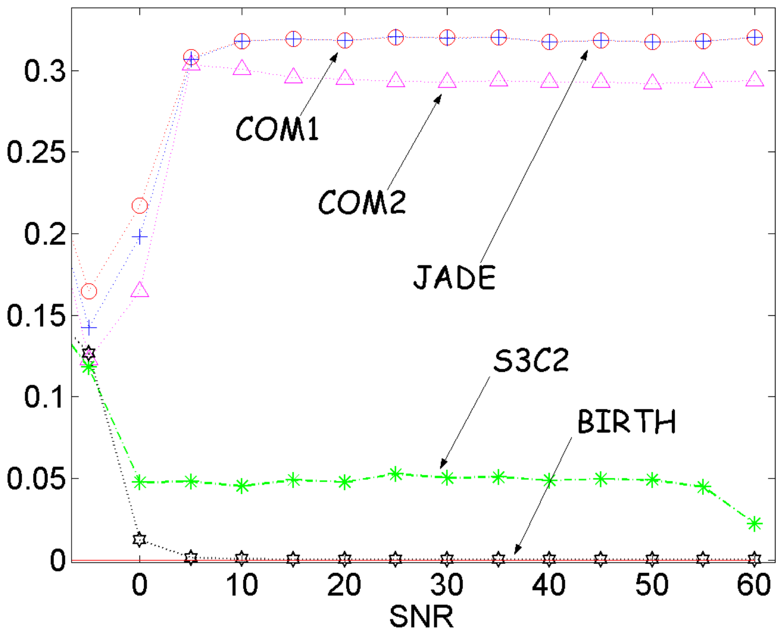

Figure 3.6: α_3 for one thousand samples

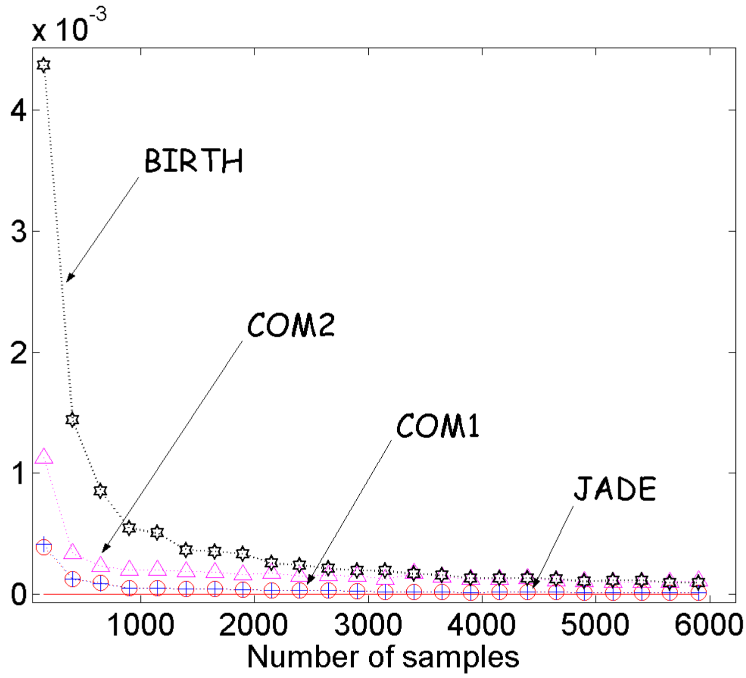

Figure 3.7: α_3 for a SNR $= 20$ dB

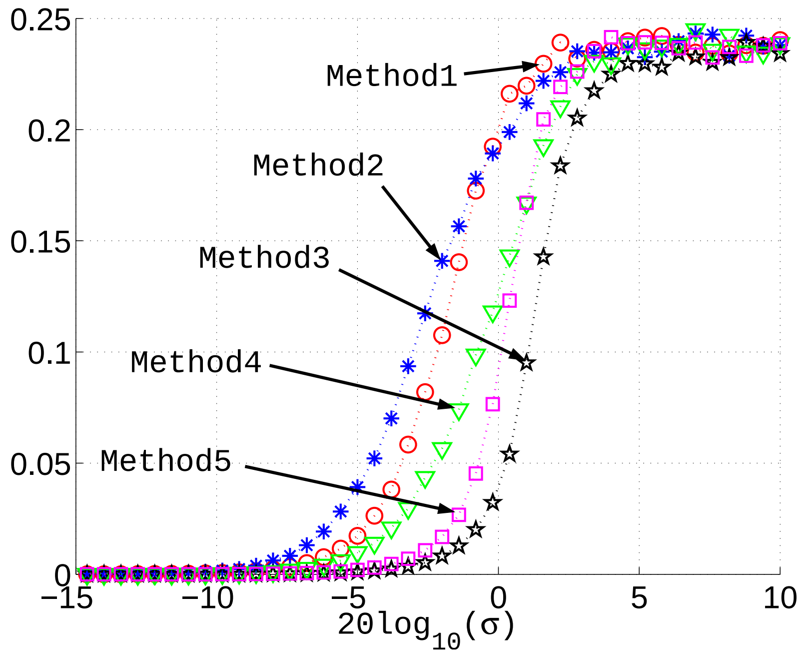

Figure 3.8: Mean of the $P = 12$ gaps α_p

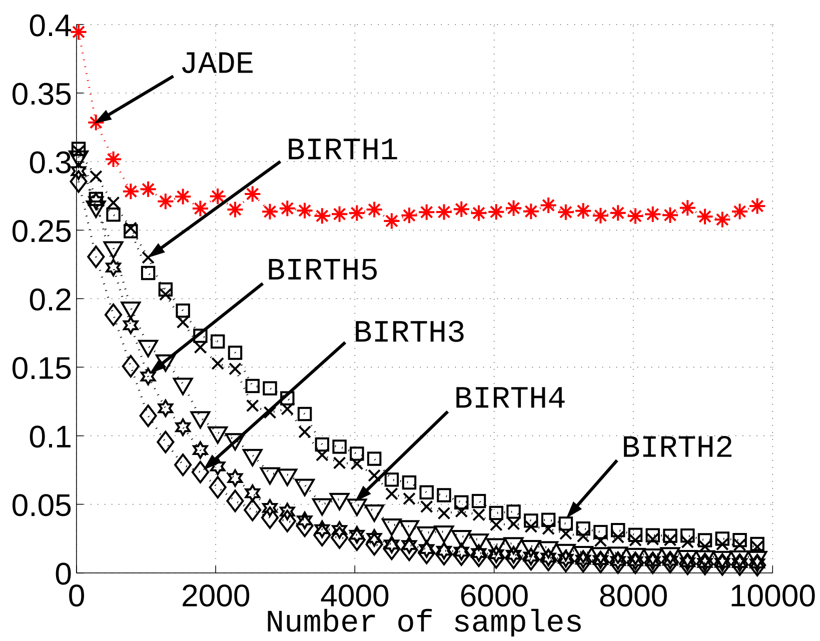

Figure 3.9: Mean of the $P = 7$ gaps α_p

Chapter 4

BIOME: Blind Identification of Overcomplete MixturEs

In order to extend the ICAR and BIRTH methods, presented in chapters 2 and 3 respectively, to an arbitrary order $2q$, where q is an arbitrary integer greater than 2, a family of new methods, named BIOME (Blind Identification of Overcomplete MixturEs of sources) is proposed in this chapter. Operating on statistics at order $2q$, this family gives rise to the $2q$-BIOME methods. The latter algorithms allow to blindly identify both overdetermined ($q \geq 2$) and underdetermined ($q \geq 3$) mixtures of sources, and to extract them in the overdetermined case. More generally, the $2q$-BIOME algorithm assumes the sources have non zero $2q$-th order marginal cumulants with the same sign (the latter assumption is verified in most cases in radiocommunications contexts). Besides, BIOME, without SO prewhitening, explicitly exploits the redundancies in the $2q$-th order statistical matrix of the data and implicitly uses the Virtual Array (VA) concept presented in [38] [13] for FourO methods and extended in [12] for HO methods. Note that, for a given value of q, the maximum number $P_{max}^{N,q}$ of independent sources that can be processed by the $2q$-BIOME method, such that $P_{max}^{N,q} \geq N$, increases with N and q.

From the linear algebra viewpoint, it is shown in section 4.1 that the BMI problem can be expressed in the form of the problem below, even in the underdetermined case.

Problem 2 *Given N matrices, $\mathbf{\Gamma}_n$, $1 \leq n \leq N$, each of size $M \times P$, $M \geq P$, find a full rank $M \times P$ matrix $\boldsymbol{\mathcal{A}}$, N diagonal matrices $\mathbf{\Lambda}_n$ of size $P \times P$, and a unitary $P \times P$ matrix $\boldsymbol{V}$, such that*

$$\mathbf{\Gamma}_n = \boldsymbol{\mathcal{A}}\,\mathbf{\Lambda}_n\,\boldsymbol{V}^{\mathsf{H}}$$

4.1 The $2q$-BIOME method

It is subsequently shown that, under assumptions (**A1**)-(**A5**) of section 1.1.2, the $2q$-BIOME method exploits the structure of the statistical matrix $\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}$, for the chosen value of ℓ, $0 \leq \ell \leq q_0$, so that the joint diagonalization to perform is actually somewhat more complicated than that given in problem 2, and better described by

Problem 3 *Given N matrices $\boldsymbol{\Gamma}_n$, $1 \leq n \leq N$, each of size $N^q \times P$, $N^q \geq P$ but possibly $N < P$, find a full rank $N \times P$ matrix $\boldsymbol{A}$, N invertible diagonal matrices $\boldsymbol{\Lambda}_n$ of size $P \times P$, and a unitary $P \times P$ matrix $\boldsymbol{V}$, such that*

$$\boldsymbol{\Gamma}_n = \boldsymbol{\mathcal{A}}^{\ell}_q \, \boldsymbol{\Lambda}_n \, \boldsymbol{V}^{\mathsf{H}}$$

where $\boldsymbol{\mathcal{A}}^{\ell}_q = \boldsymbol{A}^{\oslash q-\ell} \oslash \boldsymbol{A}^{*\oslash \ell}$.

4.1.1 The core equation

The $2q$-BIOME method precisely exploits several redundancies in the statistical matrix $\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}$ ($q \geq 2$) of the data especially thanks to the multilinearity property. Although most of BSS algorithms use the matrix multilinearity property form (1.20) (the JADE method uses it for $(q, \ell) = (1, 0)$ and for $(q, \ell) = (2, 1)$), the $2q$-BIOME method precisely exploits the second form, described by

$$\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}} = \boldsymbol{\mathcal{A}}^{\ell}_q \; \boldsymbol{\zeta}_{2q,\boldsymbol{s}} \; \boldsymbol{\mathcal{A}}^{\ell\,\mathsf{H}}_q \tag{4.1}$$

where $\boldsymbol{\zeta}_{2q,\boldsymbol{s}} \stackrel{\text{def}}{=} \text{Diag}\Big[\; C^{1,1,..,1}_{1,1,...1,\boldsymbol{s}} \quad C^{2,2,..,2}_{2,2,..,2,\boldsymbol{s}} \quad \cdots \quad C^{P,P,..,P}_{P,P,..,P,\boldsymbol{s}} \;\Big]$ is a $P \times P$ diagonal matrix of full rank in contrast to $\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{s}}$ (1.20), and where the $N^q {\times} P$ matrix $\boldsymbol{\mathcal{A}}^{\ell}_q$ is given by

$$\begin{aligned} \boldsymbol{\mathcal{A}}^{\ell}_q &= \boldsymbol{A}^{\oslash q-\ell} \oslash \boldsymbol{A}^{*\oslash \ell} = \Big[\boldsymbol{a}_1^{\otimes q-\ell} \otimes (\boldsymbol{a}_1^*)^{\otimes \ell} \;\cdots\; \boldsymbol{a}_P^{\otimes q-\ell} \otimes (\boldsymbol{a}_P^*)^{\otimes \ell}\Big] \\ &= \Big[[\boldsymbol{\mathcal{A}}^{\ell}_{q-1} \boldsymbol{\Phi}_1]^{\mathsf{T}} \; [\boldsymbol{\mathcal{A}}^{\ell}_{q-1} \boldsymbol{\Phi}_2]^{\mathsf{T}} \;\cdots\; [\boldsymbol{\mathcal{A}}^{\ell}_{q-1} \boldsymbol{\Phi}_N]^{\mathsf{T}}\Big]^{\mathsf{T}} \end{aligned} \tag{4.2}$$

with

$$\boldsymbol{\Phi}_n = \text{Diag}[\; A(n,1) \quad A(n,2) \quad \cdots \quad A(n,P) \;] \tag{4.3}$$

In other words, the non zero elements of the $P {\times} P$ diagonal matrix $\boldsymbol{\Phi}_n$ are the components of the n-th row of matrix $\boldsymbol{A}$. Note, as shown in appendix A, that the matrix form of the multilinearity property described by (4.1) ensues immediately from equations (1.11), (1.12), (1.13) and from the multilinearity property shared by cumulants [55] [19, pp. 1-24]. Moreover, it appears from equation (4.2), that matrix $\boldsymbol{\mathcal{A}}^{\ell}_q$, also called q-th order Virtual Mixture (VM), can be written by stacking $G = N^{q-1}$ matrix blocks of size $N {\times} P$, denoted $\boldsymbol{\Psi}_g$, and such that

$$\forall\, 1 \le g \le N^{q-1},\ \exists\, 1 \le n_1, \ldots, n_{q-1} \le N,\ g = \varphi([n_{q-1}\ \ n_{q-2}\ \ \ldots\ \ n_1]),$$
$$\text{and } \boldsymbol{\Psi}_g = \begin{cases} \boldsymbol{A} \prod_{j=1}^{q-1} \boldsymbol{\Phi}_{n_j} & \text{if } \ell = 0 \\ \boldsymbol{A}^* \prod_{j=1}^{\ell-1} \boldsymbol{\Phi}_{n_j}{}^* \prod_{k=\ell}^{q-1} \boldsymbol{\Phi}_{n_k} & \text{otherwise (o.w.)} \end{cases} \tag{4.4}$$

and

$$\boldsymbol{\mathcal{A}}_q^{\ell} = [\boldsymbol{\Psi}_1{}^{\mathsf{T}}\ \boldsymbol{\Psi}_2{}^{\mathsf{T}} \cdots \boldsymbol{\Psi}_G{}^{\mathsf{T}}]^{\mathsf{T}}\ . \tag{4.5}$$

4.1.2 The BIOME concept

Firstly, a unitary matrix $\boldsymbol{V}$ is estimated in the Least Squares (LS) sense, and yields an estimate of $\boldsymbol{\mathcal{A}}_q^{\ell}$. In a second stage, several algorithms may be thought of in order to compute an estimate of $\boldsymbol{A}$ from $\boldsymbol{\mathcal{A}}_q^{\ell}$. Finally, estimate of sources $\boldsymbol{s}(k)$ can be computed using the estimate of $\boldsymbol{A}$.

Identification of the q-th order VM $\boldsymbol{\mathcal{A}}_q^{\ell}$

If $2q$-th order marginal source cumulants are strictly positive (**A3**), then, according to (4.1), matrix $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ is positive. So a square root of $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$, denoted $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}$ and such that $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{\mathsf{H}/2} = \boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$, may be computed (if marginal source cumulants are strictly negative, matrix $-\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ has to be considered instead, for computing the square root). In fact, we deduce from (4.1) that matrix $\boldsymbol{\mathcal{A}}_q^{\ell}\, \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}$ is a natural square root of $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$. Another possibility is to compute this square root via the singular or eigen value decomposition of $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ given by

$$[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2} = \boldsymbol{E_s}\, \boldsymbol{L_s}^{1/2} \tag{4.6}$$

where $\boldsymbol{L_s}^{1/2}$ denotes a square root of $\boldsymbol{L_s}$, $\boldsymbol{L_s}$ is the $P \times P$ real-valued diagonal matrix of the P largest (in terms of modulus) eigenvalues of $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$, and $\boldsymbol{E_s}$ is the $N^q \times P$ matrix of the associated orthonormalized eigenvectors.

Proposition 5 *Under assumptions (A4) and (A5), the $N^q \times P$ matrix $\boldsymbol{\mathcal{A}}_q^{\ell}$ is of full column rank.*

The proof of proposition 5 ensues immediately from equations (4.2), (4.3) and assumption (**A4**). In fact, suppose that $\boldsymbol{\mathcal{A}}_q^{\ell}$ is not full column rank. Then there exists some $P \times 1$ vector $\boldsymbol{\beta} \neq 0$ such that $\boldsymbol{\mathcal{A}}_q^{\ell}\, \boldsymbol{\beta} = 0$, which, due to the structure of $\boldsymbol{\mathcal{A}}_q^{\ell}$ (4.2) implies that for all $1 \le n \le N$, $\boldsymbol{\mathcal{A}}_{q-1}^{\ell}\, \boldsymbol{\Phi}_n\, \boldsymbol{\beta} = 0$. So it implies that $\boldsymbol{\mathcal{A}}_{q-1}^{\ell}$ cannot be of full column rank (since matrices $\boldsymbol{\Phi}_n$ are $P \times P$ diagonal with nonzero entries, due to (4.3) and (**A4**)), which contradicts assumption (**A5**).

Asumptions (**A3**) to (**A5**), proposition 5, and equations (4.1) and (4.6) allow together to prove that matrices $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ and $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}$, and thus $\boldsymbol{E_s}$ and $\boldsymbol{L_s}$, are of rank P as well.

Proposition 6 *For a full rank matrix $\mathcal{A}_q^\ell$, (A3) is equivalent to assuming that the diagonal elements of $\boldsymbol{L_s}$ are not null and have also the same sign.*

The proof of proposition 6 is also straightforward. In fact, it is well-known that two square roots of a matrix are equal to within a unitary matrix, so that

$$\mathcal{A}_q^\ell\, \zeta_{2q,\boldsymbol{s}}^{1/2} = \boldsymbol{E_s}\, \boldsymbol{L_s}^{1/2}\, \boldsymbol{V} \qquad \left(= [\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2}\, \boldsymbol{V}\right) \tag{4.7}$$

for some $P{\times}P$ unitary matrix $\boldsymbol{V}$. Note the latter is unique up to a multiplicative unitary invertible diagonal matrix. We deduce from (4.7) that

$$\boldsymbol{E_s}^{\mathsf{H}}\, \mathcal{A}_q^\ell\, \zeta_{2q,\boldsymbol{s}}\, \mathcal{A}_q^{\ell\,\mathsf{H}}\, \boldsymbol{E_s} = \boldsymbol{L_s} \tag{4.8}$$

and hence proposition 6.

In addition, equation (4.7) can be rewritten as follows

$$[\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2} = \boldsymbol{E_s}\, \boldsymbol{L_s}^{1/2} = \mathcal{A}_q^\ell\, \zeta_{2q,\boldsymbol{s}}^{1/2}\, \boldsymbol{V}^{\mathsf{H}}. \tag{4.9}$$

showing the link between $[\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2}$ and $\mathcal{A}_q^\ell$. Plugging (4.2) into (4.9), matrix $[\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2}$ can be eventually rewritten as

$$\begin{aligned}
[\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2} &= \Big[[\mathcal{A}_{q-1}^\ell \boldsymbol{\Phi}_1 \zeta_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}}\ [\mathcal{A}_{q-1}^\ell \boldsymbol{\Phi}_2 \zeta_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \cdots [\mathcal{A}_{q-1}^\ell \boldsymbol{\Phi}_N \zeta_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \Big]^{\mathsf{T}} \\
&= [\ \boldsymbol{\Gamma}_1{}^{\mathsf{T}}\ \ \boldsymbol{\Gamma}_2{}^{\mathsf{T}}\ \ \cdots\ \ \boldsymbol{\Gamma}_N{}^{\mathsf{T}}\ ]^{\mathsf{T}}
\end{aligned} \tag{4.10}$$

where the N matrix blocks $\boldsymbol{\Gamma}_n$ of size $N^{q-1}{\times}P$ are given by

$$\forall\, 1 \le n \le N, \qquad \boldsymbol{\Gamma}_n = \mathcal{A}_{q-1}^\ell\, \boldsymbol{\Phi}_n\, \zeta_{2q,\boldsymbol{s}}^{1/2}\, \boldsymbol{V}^{\mathsf{H}} \tag{4.11}$$

Proposition 7 *For any $1 \le n \le N$, matrix $\boldsymbol{\Gamma}_n$ is of full column rank.*

The proof results from proposition 5, in addition to all other stated conditions.

Using proposition 7, the pseudo-inverses $\boldsymbol{\Gamma}_n^\sharp$ of the $N^{q-1}{\times}P$ matrices $\boldsymbol{\Gamma}_n$ may be defined by

$$\forall\, 1 \le n \le N, \qquad \boldsymbol{\Gamma}_n^\sharp = (\boldsymbol{\Gamma}_n{}^{\mathsf{H}} \boldsymbol{\Gamma}_n)^{-1}\, \boldsymbol{\Gamma}_n{}^{\mathsf{H}} \tag{4.12}$$

Then, the information contained in matrix $[\mathcal{C}_{2q,\boldsymbol{x}}^\ell]^{1/2}$ allows one to blindly identify $\mathcal{A}_q^\ell$. Indeed, matrix $\boldsymbol{V}$ jointly diagonalizes the $N(N-1)$ matrices $\boldsymbol{\Theta}_{n_1,n_2}$ below

$$\forall\, 1 \le n_1 \ne n_2 \le N, \qquad \boldsymbol{\Theta}_{n_1,n_2} = \boldsymbol{\Gamma}_{n_1}^\sharp\, \boldsymbol{\Gamma}_{n_2}. \tag{4.13}$$

To see this, let us compute Θ_{n_1,n_2} from (4.11) and (4.12). We obtain

$$\Theta_{n_1,n_2} = \boldsymbol{V}\,[\zeta^{\ell}_{2q,\boldsymbol{s}}]^{-1/2}\,\Phi_{n_1}^{-1}\,\Phi_{n_2}\,\zeta^{1/2}_{2q,\boldsymbol{s}}\,\boldsymbol{V}^{\mathsf{H}} = \boldsymbol{V}\,\Phi_{n_1}^{-1}\,\Phi_{n_2}\,\boldsymbol{V}^{\mathsf{H}} \tag{4.14}$$

where $\zeta^{1/2}_{2q,\boldsymbol{s}}$ and $\boldsymbol{D}_{n_1,n_2} = \Phi_{n_1}^{-1}\Phi_{n_2}$ are $P\times P$ diagonal full rank matrices, which shows the result. The unitary matrix $\boldsymbol{V}_{sol}$, solution to the previous problem of joint diagonalization of the $N(N-1)$ matrices Θ_{n_1,n_2} has necessarily the form $\boldsymbol{V}_{sol} = \boldsymbol{V}\boldsymbol{\mathcal{T}}$ where $\boldsymbol{\mathcal{T}}$ is a unitary matrix. This allows one, in accordance with (4.9), to recover $\boldsymbol{\mathcal{A}}^{\ell}_{q}$ to within an orthogonal matrix as

$$[\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}]^{1/2}\,\boldsymbol{V}_{sol} = \boldsymbol{\mathcal{A}}^{\ell}_{q}\,\zeta^{1/2}_{2q,\boldsymbol{s}}\,\boldsymbol{\mathcal{T}} \tag{4.15}$$

Proposition 8 *Under assumption (A4), for every pair* $(p_1,p_2)|_{p_1\neq p_2}$ *of* $\{1,\ 2,\ldots,P\}^2$*, at least one pair* $(n_1,n_2)|_{n_1\neq n_2}$ *belonging to* $\{1,2,\ldots,N\}^2$ *exists such that* $D_{n_1,n_2}(p_1,p_1)\neq D_{n_1,n_2}(p_2,p_2)$.

The proof is given in appendix B.

Proposition 8 and [2] allow to assert that the previous unitary matrix $\boldsymbol{\mathcal{T}}$ is also trivial. So matrix $\boldsymbol{\mathcal{A}}^{\ell}_{q}$ may be identified, according to (4.15), up to a trivial matrix.

Identification of mixture A

Three algorithms are proposed in this section, with increased computational complexity and performances.

Note, from (4.5) and (4.4), that equation (4.15) can also be written in the form of $G = N^{q-1}$ matrix blocks $\boldsymbol{\Sigma}_g = \boldsymbol{\Psi}_g\zeta^{1/2}_{2q,\boldsymbol{s}}\,\boldsymbol{\mathcal{T}}$ of size $N\times P$ as

$$[\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}]^{1/2}\,\boldsymbol{V}_{sol} = [\boldsymbol{\Sigma}_1{}^{\mathsf{T}}\ \boldsymbol{\Sigma}_2{}^{\mathsf{T}}\cdots\boldsymbol{\Sigma}_G{}^{\mathsf{T}}]^{\mathsf{T}} \tag{4.16}$$

So a first approach to estimate $\boldsymbol{A}$ up to a trivial matrix, named $2q$-BIOME1 in the sequel, consists of retaining only the matrix block $\boldsymbol{\Sigma}_1$ if $\ell = 0$ ($\boldsymbol{\Sigma}_1^*$ otherwise) made up of the N first rows of $[\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}]^{1/2}\boldsymbol{V}_{sol}$ such that

$$\boldsymbol{\Sigma}_1 = \begin{cases} \boldsymbol{A}\ [\Phi_1]^{q-1}\ \zeta^{1/2}_{2q,\boldsymbol{s}}\,\boldsymbol{\mathcal{T}} & \text{if } \ell = 0 \\ \boldsymbol{A}^*\ [\Phi_1^*]^{\ell-1}\ [\Phi_1]^{q-\ell}\ \zeta^{1/2}_{2q,\boldsymbol{s}}\,\boldsymbol{\mathcal{T}} & \text{o.w.} \end{cases} \tag{4.17}$$

where $\zeta^{1/2}_{2q,\boldsymbol{s}}$ and Φ_n, for all $1\leq n\leq N$, are diagonal matrices.

It is also possible to take into account all the matrix blocks $\boldsymbol{\Sigma}_g$ if $\ell = 0$ ($\boldsymbol{\Sigma}_g^*$ otherwise) and to compute their average. This yields a second algorithm, called $2q$-BIOME2, of higher complexity.

A third algorithm, named $2q$-BIOME3, is now described, and yields a more accurate solution to the BMI problem: as shown in appendix C, it consists, for each column $\boldsymbol{b}_p$ of $[\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}]^{1/2}\boldsymbol{V}_{sol}$, first of extracting the $H = N^{q-2}$ vectors $\boldsymbol{b}_p(h)$ $(1 \leq h \leq H)$ of size $N^2{\times}1$ (such that $\boldsymbol{b}_p = [\boldsymbol{b}_p(1)^{\mathsf{T}} \boldsymbol{b}_p(2)^{\mathsf{T}} \cdots \boldsymbol{b}_p(H)^{\mathsf{T}}]^{\mathsf{T}}$), then of remodeling them into H matrices $\boldsymbol{B}_p(h)$ of size $N{\times}N$ (the n-th column of $\boldsymbol{B}_p(h)$ is made up from the N consecutive elements of $\boldsymbol{b}_p(h)$ as from the $[N(n{-}1){+}1]$-th one), and finally of jointly diagonalizing the set Δ_p^ℓ of matrices defined by

$$\Delta_p^\ell = \begin{cases} \{\boldsymbol{B}_p(h)\boldsymbol{B}_p(h)^{\mathsf{H}},\ (\boldsymbol{B}_p(h)^{\mathsf{H}}\boldsymbol{B}_p(h))^* \ /\ 1{\leq}h{\leq}H\} & \text{if } \ell = 0 \\ \{\boldsymbol{B}_p(h)^* \ /\ 1{\leq}h{\leq}H\} & \text{if } \ell = 1 \\ \{(\boldsymbol{B}_p(h)\boldsymbol{B}_p(h)^{\mathsf{H}})^*,\ (\boldsymbol{B}_p(h)^{\mathsf{H}}\boldsymbol{B}_p(h)) \ /\ 1{\leq}h{\leq}H\} & \text{o.w.} \end{cases} \tag{4.18}$$

Theorem 2 *The eigenvector, in common to all matrices of Δ_p^ℓ, and associated with the largest eigenvalue, is, up to a scale factor, a column vector of matrix $\boldsymbol{A}$.*

The proof is given in appendix C. So each joint diagonalization of matrices belonging to the set Δ_p^ℓ allows one to estimate a column vector of $\boldsymbol{A}$, and finally to identify $\boldsymbol{A}$ to within a trivial matrix.

Remark 3 *Although the algorithm of joint approximate diagonalization in the LS sense [9] is restricted to unitary joint diagonalizers, it can be used to process the previous problem since matrices belonging to Δ_p^ℓ are of rank 1 as shown in (C.5). However it is reasonable to believe that, if an unrestricted (non-unitary) LS joint diagonalization scheme is applied, as for instance the one described by Yeredor in [73], a better LS fit can be attained, possibly leading to a better estimate of $\boldsymbol{A}$. However, both approaches have been compared by simulations in the previous chapter (section 3.4.2), showing that the former gives best results.*

Extraction of the P independent components

Finally, to estimate the signal vector $\boldsymbol{s}(k)$ for any value k, and only in overdetermined situations (*i.e.* for $P \leq N$), it is sufficient to apply a particular matrix filter built from the estimate $\widehat{\boldsymbol{A}}$ of $\boldsymbol{A}$: such a filter may be the Spatial Matched Filter (SMF) source separator described in [11], which is optimal in the presence of decorrelated signals and whose estimate is given by $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1}\widehat{\boldsymbol{A}}$, where $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$ is an estimate of $\boldsymbol{R}_{\boldsymbol{x}} {=} \boldsymbol{\mathcal{C}}_{2,\boldsymbol{x}}^{0}$.

4.1.3 Implementation of the BIOME method

The different steps of the $2q$-BIOME method are summarized hereafter when K samples of the observations, $\boldsymbol{x}(k)$ $(1 {\leq} k {\leq} K)$, are available.

Step1 Choose the adequate $2q$-th statistical order in accordance with the alleged source number P to be potentially processed: see section 4.2.2 for more details. In practical situations, q is the minimal value which ensures the processing of all the sources potentially present.

Step2 Estimate the $2q$-th order statistics $C_{i_1,i_2,\dots,i_q,\boldsymbol{x}}^{i_{q+1},\dots,i_{2q}}$ from the K samples $\boldsymbol{x}(k)$ and choose, using section 4.2.2 and [12], the best arrangement $\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}$, where $\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell}$ is an estimate of $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$.

Step3 Compute the Eigen Value Decomposition (EVD) of the Hermitian matrix $\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}$; estimate $\widehat{P}$, an estimate of the source number P, from an eigenvalue test and restrict the EVD to the $\widehat{P}$ principal components : $\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}} \approx \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\, \widehat{\boldsymbol{L}}_{\boldsymbol{s}}\, \widehat{\boldsymbol{E}}_{\boldsymbol{s}}^{\mathsf{H}}$, where $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$ is the diagonal matrix of the $\widehat{P}$ eigenvalues of largest modulus and $\widehat{\boldsymbol{E}}_{\boldsymbol{s}}$ is the matrix of the associated eigenvectors.

Step4 Estimate the sign, ϵ, of the diagonal elements of $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$.

Step5 Compute a square root matrix $[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}$ of $\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}$: $[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2} = \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\,|\widehat{\boldsymbol{L}}_{\boldsymbol{s}}|^{1/2}$, where $|\cdot|$ denotes the elementwise complex modulus operator.

Step6 Extract from $[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}$ the N matrices $\widehat{\boldsymbol{\Gamma}}_n$, construct matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2} = [\widehat{\boldsymbol{\Gamma}}_{n_1}^{\sharp}\, \widehat{\boldsymbol{\Gamma}}_{n_2}]$ for all $1 \leq n_1 \neq n_2 \leq N$, and compute the estimate $\widehat{\boldsymbol{V}}_{sol}$ of the unitary matrix $\boldsymbol{V}_{sol}$ from the joint diagonalization of the $N(N-1)$ matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2}$ (with the algorithm described in [9]).

Step7 Compute $\widehat{\boldsymbol{A}}$, an estimate of mixture $\boldsymbol{A}$, from matrix $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}]$ by either one of the following:

1. ($2q$-BIOME1) taking the matrix block made up of the N first rows of $\left[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}\right]$ if $\ell_{opt} = 0$, and of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}]^*$ otherwise;
2. ($2q$-BIOME2) taking the average of the N matrix blocks, of size $N \times P$, made up of the successive rows of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}]$ if $\ell_{opt} = 0$, and of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}]^*$ otherwise;
3. ($2q$-BIOME3) fully exploiting each column vector $\widehat{\boldsymbol{b}}_p$ of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}_{2q,\boldsymbol{x}}^{\ell_{opt}}]^{1/2}\, \widehat{\boldsymbol{V}}_{sol}]$. In order to do this, first extract the $M = N^{q-2}$ vectors $\widehat{\boldsymbol{b}}_p(m)$ of size $N^2 \times 1$, then remodel them into M matrices $\widehat{\boldsymbol{B}}_p(m)$ of size $N \times N$, and finally build the matrix

whose p-th column vector is the eigenvector in common within the M matrices $\widehat{\Delta}_p^{\ell}(m)$ $(1 \leq m \leq M)$ and associated with the largest eigenvalue.

Step8 If $\boldsymbol{A}$ is an overdetermined mixture, estimate the signal vector $\boldsymbol{s}(k)$ for any value k, by applying to $\boldsymbol{x}(k)$ the SMF source separator defined by $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \widehat{\boldsymbol{A}}$, where $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$ is an estimate of $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}_{2,\boldsymbol{x}}^{0}$.

4.2 Identifiability

The identifiability properties of the $2q$-BIOME method are directly related to the $2q$-th order Virtual Array (VA) concept described in [38] [13] for $q=2$ and extended in [12] for $q \geq 2$. For this reason, we recall the main results about the VA array concept in section 4.2.1 before discussing, in section 4.2.2, the identifiability properties of $2q$-BIOME.

4.2.1 The VA concept

In the absence of coupling between sensors, component n of the p-th column vector $\boldsymbol{a}_p = \boldsymbol{a}(\theta_p, \varphi_p)$ of $\boldsymbol{A}$, denoted $a_n(\theta_p, \varphi_p)$ where θ_p and φ_p are the azimuth and the elevation angles of source p, can be written, in the general case of an array with space, angular and polarization diversity, as [26]

$$\begin{aligned} a_n(\theta_p, \varphi_p) = f_n(\theta_p, \varphi_p, \omega_p) \exp \{ \mathrm{j} 2\pi [x_n \cos(\theta_p) \cos(\varphi_p) \; + \\ y_n \sin(\theta_p) \cos(\varphi_p) + z_n \sin(\varphi_p)] \, / \lambda \} \end{aligned} \tag{4.19}$$

where λ is the wavelength, (x_n, y_n, z_n) are the coordinates of sensor n of the array, $f_n(\theta_p, \varphi_p, \omega_p)$ is a complex number corresponding to the response of sensor n to a unit electric field coming from the direction (θ_p, φ_p) and having the state of polarization ω_p (characterized by two angles in the wave plane) [26]. Let us recall that an array of sensors has spatial diversity if the sensors have not all the same phase center. The array has angular and/or polarization diversity if the sensors have not all the same radiating pattern and/or the same polarization, respectively.

Assuming no noise, we note that matrices $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ and $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}_{2,\boldsymbol{x}}^{0}$, defined by (4.1), have the same algebraic structure, where the marginal source cumulant $C_{p,p,\ldots,p,\boldsymbol{s}}^{p,p,\ldots,p}$ and the vector $\left[\boldsymbol{a}_p^{\otimes q-\ell} \otimes (\boldsymbol{a}_p^{*})^{\otimes \ell}\right]$ $= \left[\boldsymbol{a}(\theta_p, \varphi_p)^{\otimes q-\ell} \otimes (\boldsymbol{a}(\theta_p, \varphi_p)^{*})^{\otimes \ell}\right]$ play, for $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$, the role played for $\boldsymbol{R}_{\boldsymbol{x}}$ by the power $C_{p,\boldsymbol{s}}^{p}$ and the steering vector $\boldsymbol{a}(\theta_p, \varphi_p)$ respectively. Thus, for BMI methods exploiting expression (4.1), the $N^q \times 1$ vector $\left[\boldsymbol{a}(\theta_p, \varphi_p)^{\otimes q-\ell} \otimes (\boldsymbol{a}(\theta_p, \varphi_p)^{*})^{\otimes \ell}\right]$ can be considered as the *equivalent* or *virtual steering vector* of the source p for the true array of N sensors with coordinates (x_n, y_n, z_n) and amplitude pattern $f_n(\theta_p, \varphi_p, \omega_p)$ $(1 \le n \le N)$. Moreover, comparing the components of $\left[\boldsymbol{a}(\theta_p, \varphi_p)^{\otimes q-\ell} \otimes (\boldsymbol{a}(\theta_p, \varphi_p)^*)^{\otimes \ell}\right]$ to expression (4.19), it is shown in [12] that vector $\left[\boldsymbol{a}(\theta_p, \varphi_p)^{\otimes q-\ell} \otimes (\boldsymbol{a}(\theta_p, \varphi_p)^*)^{\otimes \ell}\right]$ can also be considered as the true steering vector of the source p but for a VA of N^q Virtual Sensors (VS) with particular coordinates and particular complex amplitude patterns deduced from (x_n, y_n, z_n) and $f_n(\theta_p, \varphi_p, \omega_p)$ $(1 \le n \le N)$ respectively.

Nevertheless, some of these N^q VS may coincide. If we note $\mathcal{N}_{2q}^{\ell}$ the number of different VS of the VA associated with the $2q$-th order array processing problem for the arrangement $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$, $\mathcal{N}_{2q}^{\ell}$ is also a upper bound to the rank of matrix $\boldsymbol{\mathcal{A}}_q^{\ell}$. Conversely, if the $2q$-th order VA has no ambiguities [63] of rank smaller than or equal to $\mathcal{N}_{2q}^{\ell}$, the rank of matrix $\boldsymbol{\mathcal{A}}_q^{\ell}$ is equal to $\mathcal{N}_{2q}^{\ell}$ under (**A4**). In particular it is shown in [12] that in the general case of an arbitrary array of N sensors with no particular symmetries, for large values of N and for a given value of q $(2 \le q \le N)$, the number of different VS $\mathcal{N}_{2q}^{\ell}$ can be approximated by

$$\mathcal{N}_{2q}^{\ell} \approx N!/\left[(N-q)!\,(q-\ell)!\,\ell!\right] \tag{4.20}$$

In these conditions, the optimal arrangement $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell_{opt}}$ is such that ℓ_{opt} maximizes $\mathcal{N}_{2q}^{\ell}$ defined by (4.20) and thus minimizes the quantity $(q-\ell)!\,\ell!$ with respect to ℓ ($0 \le \ell \le q_0$ where $q_0 = q/2$ if q is even and $q_0 = (q-1)/2$ if q is odd). It is straightforward to show that $\ell_{opt} = q_0$ and it is verified in [12] for $2 \le q \le 4$ that this result remains true whatever N. In other words, ℓ_{opt} generates steering vectors $\left[\boldsymbol{a}(\theta_p, \varphi_p)^{\otimes q-\ell} \otimes (\boldsymbol{a}(\theta_p, \varphi_p)^*)^{\otimes \ell}\right]$ for which the number of conjugate vectors is the least different from the number of non conjugate vectors.

The computation of the number of different VS, $\mathcal{N}_{2q}^{\ell}$, of the $2q$-th order VA for the arrangement $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$ is not easy for arbitrary values of N, q $(q \ge 2)$ and ℓ. For this reason, Chevalier et al. [12] limit their analysis to some values of q $(2 \le q \le 4)$, which extends the results of [13] up to the eighth order for arbitrary arrangements of the data cumulants. In fact, for these values of q, Chevalier et al. give a upper bound to $\mathcal{N}_{2q}^{\ell}$, $\mathcal{N}_{max}^{2q,\ell}$, first for an array with space, angular and polarization diversities, summarized in table 4.1, then for an array with angular and polarization diversity only, and finally for an array with only spatial diversity summarized in table 4.2. These upper bounds are shown in [12] to be reached for most array geometries. Nevertheless, for Uniformly spaced Linear Arrays (ULA), these upper bounds are not reached and $\mathcal{N}_{2q}^{\ell}$ is shown in [12] to be given by

$$\mathcal{N}_{2q}^{\ell} = q(N-1)+1 \tag{4.21}$$

whatever q, N and ℓ, showing that the number $\mathcal{N}_{2q}^{\ell}$ of different VS of the $2q$-th order VA associated with a ULA is independent of ℓ and of the chosen arrangement $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}$. However, for UCAs of N

sensors, the upper bound is shown in [12] to be reached when N is a prime number as depicted in table 4.3.

		$\mathcal{N}_{max}^{2q,\ell}$
q=2	$\ell = 0$	$N(N+1)/2$
	$\ell = 1$	N^2
q=3	$\ell = 0$	$N!/[6(N-3)!] + N(N-1) + N$
	$\ell = 1$	$N!/[2(N-3)!] + 2N(N-1) + N$
q=4	$\ell = 0$	$N!/[24(N-4)!] + N!/[2(N-3)!] + 1.5N(N-1) + N$
	$\ell = 1$	$N!/[6(N-4)!] + 1.5N!/(N-3)! + 3N(N-1) + N$
	$\ell = 2$	$N!/[4(N-4)!] + 2N!/(N-3)! + 3.5N(N-1) + N$

Table 4.1: $\mathcal{N}_{max}^{2q,\ell}$ associated with arrays with space, angular and polarization diversities

		$\mathcal{N}_{max}^{2q,\ell}$
q=2	$\ell = 0$	$N(N+1)/2$
	$\ell = 1$	$N^2 - N + 1$
q=3	$\ell = 0$	$N!/[6(N-3)!] + N(N-1) + N$
	$\ell = 1$	$N!/[2(N-3)!] + N(N-1) + N$
q=4	$\ell = 0$	$N!/[24(N-4)!] + N!/[2(N-3)!] + 1.5N(N-1) + N$
	$\ell = 1$	$N!/[6(N-4)!] + N!/(N-3)! + 1.5N(N-1) + N$
	$\ell = 2$	$N!/[4(N-4)!] + N!/(N-3)! + 2N(N-1) + 1$

Table 4.2: $\mathcal{N}_{max}^{2q,\ell}$ associated with arrays with spatial diversity only

4.2.2 The BIOME processing power

From the results of section 4.2.1, it is possible to identify the maximum number, $P_{max}^{N,q}$, of independent non Gaussian sources that can be processed by the $2q$-BIOME method. Indeed, it has been shown in the previous sections that P sources can be blindly identified by the $2q$-BIOME method from an array of N sensors, provided conditions (**A1**)-(**A5**) are verified. For an array without any rank-1 ambiguities, condition (**A4**) is verified as soon as the sources have different

		$\mathcal{N}_{2q}^{\ell}$				
		$N=3$	$N=5$	$N=7$	$N=9$	$N=11$
q=2	$\ell=0$	6	15	28	45	66
	$\ell=1$	7	21	43	73	111
q=3	$\ell=0$	10	35	84	163	286
	$\ell=1$	12	55	154	306	616
q=4	$\ell=0$	15	70	210	477	1001
	$\ell=1$	18	115	420	918	2486
	$\ell=2$	19	131	505	1135	3191

Table 4.3: $\mathcal{N}_{2q}^{\ell}$ associated with a UCA of N identical sensors

directions of arrival. In a same manner, assuming the $2(q-1)$th order VA associated with the arrangement $\boldsymbol{\mathcal{C}}_{2(q-1),\boldsymbol{x}}^{\ell}$ and the considered VA array of N^q VS has no ambiguities of rank lower than $\mathcal{N}_{2(q-1)}^{\ell}$, condition (**A5**) is verified provided (**A4**) is verified and P is lower than or equal to $\mathcal{N}_{2(q-1)}^{\ell}$. Otherwise, (**A5**) cannot be verified. We deduce from this result that the maximal number $P_{max}^{N,q}$ of non Gaussian sources that can be processed by $2q$-BIOME is $\mathcal{N}_{2(q-1)}^{\ell_{opt}}$.

As far as the choice of parameter q is concerned, it depends on the number P of independent sources that BIOME's user wants to process. Since we have previously shown the link between $P_{max}^{N,q}$ and $\mathcal{N}_{2(q-1)}^{\ell_{opt}}$ for a given value of q, it is sufficient to choose the smaller value of q ($q \geq 2$) such that $P \leq P_{max}^{N,q}$.

4.3 Simulations

The performance criterion used to evaluate the quality of the BMI process has been presented in section 1.1.3. On the other hand, the quality of the BSS process is evaluated using the well-known SINRM (Signal to Interference plus Noise Ratio Maximum) criterion defined in [11, section 3]. Moreover, the synthetic signals used in this section are cyclostationary, and according to sections 1.2.4 and 5.2, other statistical estimators than empirical estimators should be employed. However, if the cyclostationary sources are zero-mean and circular, or non circular with a zero carrier residu, or non circular with different non zero carrier residus, such as the sources used subsequently, the bias due to empirical statistical estimators is negligible [42]. So we decide to employ them in the following simulations.

The overdetermined case

The previous results are firstly illustrated in the overdetermined case comparing 4-BIOME and 6-BIOME with the well-known BSS algorithms. In fact, we assume that $P = 4$ statistically independent sources, i.e. 2 BPSK and 2 QPSK, all with a raised cosine pulse shape of roll-off equal to 0.25, are received by a UCA of $N = 4$ identical sensors of radius R such that $R/\lambda = 0.55$ (λ: wavelength). The four sources, assumed synchronized, have the same input SNR (Signal to Noise Ratio) of 20dB and the noise is spatially and temporally white Gaussian. The symbol period T_1 associated with the first BPSK is equal to three times the sample period T_e. The other sources have a symbol period equal to twice the sample period. The directions of arrival of the sources are such that the source steering vectors are orthogonal and the associated carrier residus are such that $f_{c1}T_e = 0$, $f_{c2}T_e = 0.3$, $f_{c3}T_e = 0.2$ and $f_{c4}T_e = 0.1$. We apply the COM1 [18], COM2 [16], JADE [8], SOBI [2], FastICA [3], FOBIUM [40], 4-BIOME1, 4-BIOME2, 4-BIOME3 and 6-BIOME1 methods, and the SINRM associated with each source is computed and averaged over 200 realizations. Figures 4.1 and 4.2 show the variations of *SINRM*$_3$ (source 3 performance) at the output of the previous methods as a function of the number of samples.

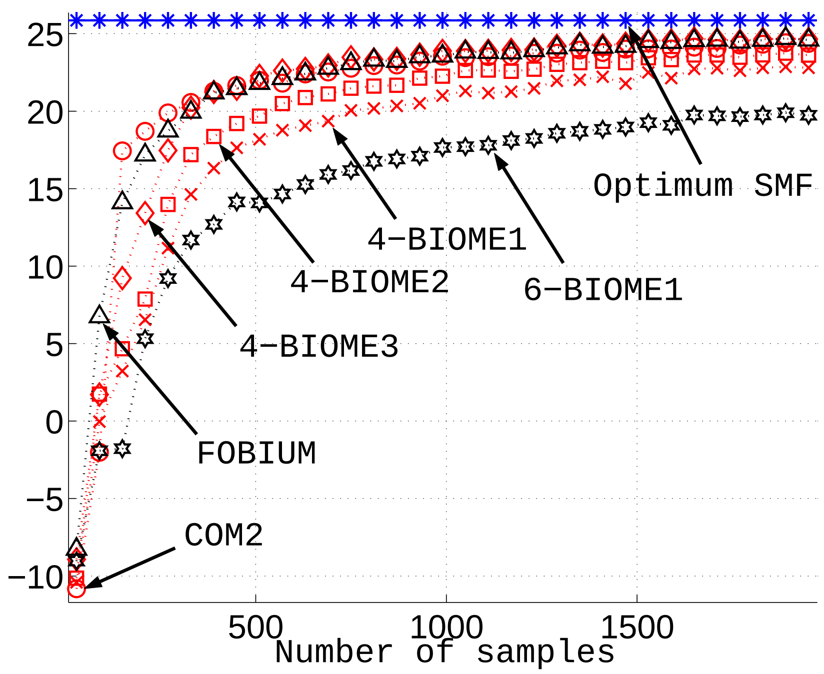

Figure 4.1: SINRM associated with source 3 for a SNR of 20 dB

Although the 6-BIOME1 method obviously works in overdetermined contexts, it appears in figure 4.1 that the 4-BIOMEm ($1 \leq m \leq 3$) methods give better results, which shows that it is sufficient and more appropriate to use, as proposed in section 4.2.2, the $2q$-BIOME method of smallest value q allowing to process the P sources. Figure 4.1 also shows that 4-BIOME3 converges as fast as COM2 and FOBIUM, but faster than 4-BIOME1 and 4-BIOME2: the third method given in section 4.1.2 exhibits better performances than the others and it is reasonable to believe that the 6-BIOME3 method would give better results than those of 6-BIOME1, as shown in the previous chapter (section 3.4.2).

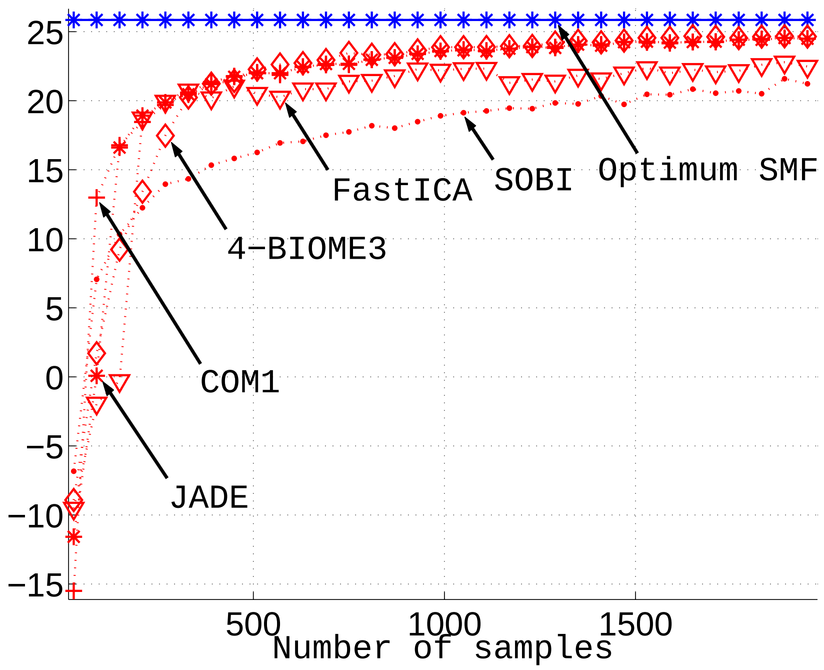

Figure 4.2: SINRM associated with source 3 for a SNR of 20 dB

In addition, figure 4.2 shows the good performance of the 4-BIOME3 algorithm facing the well-known COM1, COM2, JADE, SOBI and FastICA methods. Note that the SOBI and FOBIUM methods give in this simulation good results since sources have been chosen with different spectral densities, especially taking different carrier residus.

The colored noise case

Then, the 4-BIOME method is compared to other algorithms in an overdetermined context and especially in the presence of a Gaussian noise with unknown spatial correlation. In fact, $P = 3$ statistically independent sources, i.e. 2 BPSK and 1 QPSK, all with a raised cosine pulse shape of roll-off equal to 0.25, are assumed to be received by a UCA of $N=5$ identical sensors of radius R such that $R/\lambda = 0.55$. Their symbol periods are equal to $T_1 = 2T_e$, $T_2 = 3T_e$ and $T_3 = 4T_e$ respectively. Their carrier residus are chosen equal to zero. Finally, the source steering vectors are built orthogonal. This time, we apply the COM1, COM2, JADE, SOBI, FOBIUM, 4-BIOME1, 4-BIOME2 and 4-BIOME3 methods, and the SINRM associated with each source is computed and averaged over 200 realizations.

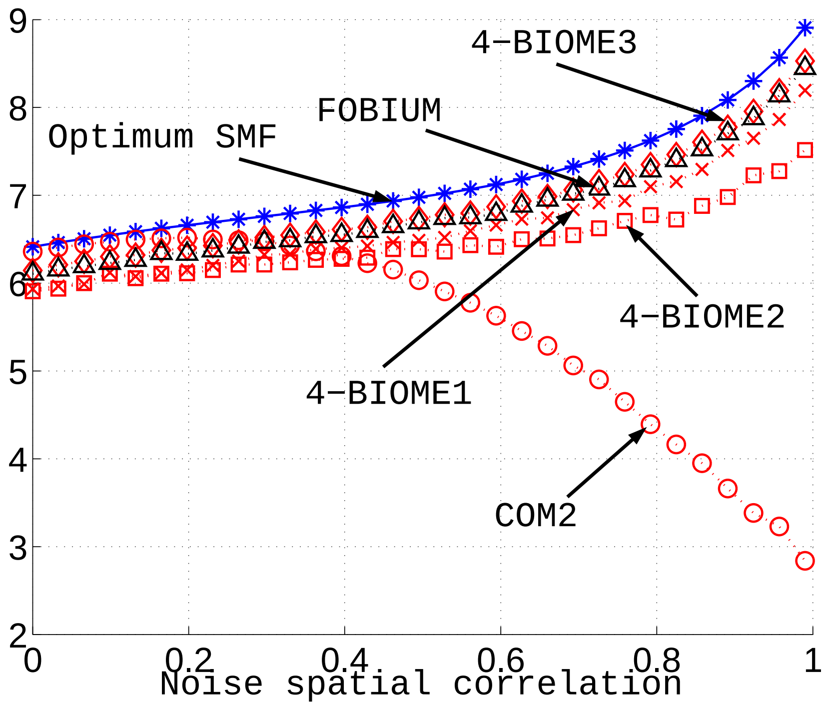

Figure 4.3: SINRM associated with source 3 for a SNR of 0 dB

Figures 4.3 and 4.4 show the variations of $SINRM_3$ (source 3 performance) at the output of the previous methods as a function of the noise spatial correlation factor ρ. SNR of the three sources is taken equal to 0 dB and 1500 samples are used to identify the overdetermined mixture. Note that the Gaussian noise model employed in this simulation is the sum of an internal noise $\boldsymbol{\nu_{in}}(k)$ and an external noise $\boldsymbol{\nu_{out}}(k)$, of covariance matrices $\boldsymbol{R}_{\boldsymbol{\nu}}^{in}$ and $\boldsymbol{R}_{\boldsymbol{\nu}}^{out}$ respectively such that

$$\boldsymbol{R}_{\boldsymbol{\nu}}^{in}(r,q) \stackrel{\text{def}}{=} \sigma^2\delta(r-q)/2 \qquad\qquad \boldsymbol{R}_{\boldsymbol{\nu}}^{out}(r,q) \stackrel{\text{def}}{=} \sigma^2\rho^{|r-q|}/2 \tag{4.22}$$

where σ^2, ρ are the total noise variance per sensor and the noise spatial correlation factor respectively. Note that $\boldsymbol{R}_{\boldsymbol{\nu}}(r,q) \stackrel{\text{def}}{=} \boldsymbol{R}_{\boldsymbol{\nu}}^{in}(r,q) + \boldsymbol{R}_{\boldsymbol{\nu}}^{out}(r,q)$ is the (r,q)-th component of the total noise covariance matrix.

It appears in figure 4.3 that the three proposed versions of 4-BIOME seem to be robust with respect to the correlated Gaussian noise presence: 4-BIOME1 and 4-BIOME3 are totally insensitive to a Gaussian noise with unknown spatial correlation.

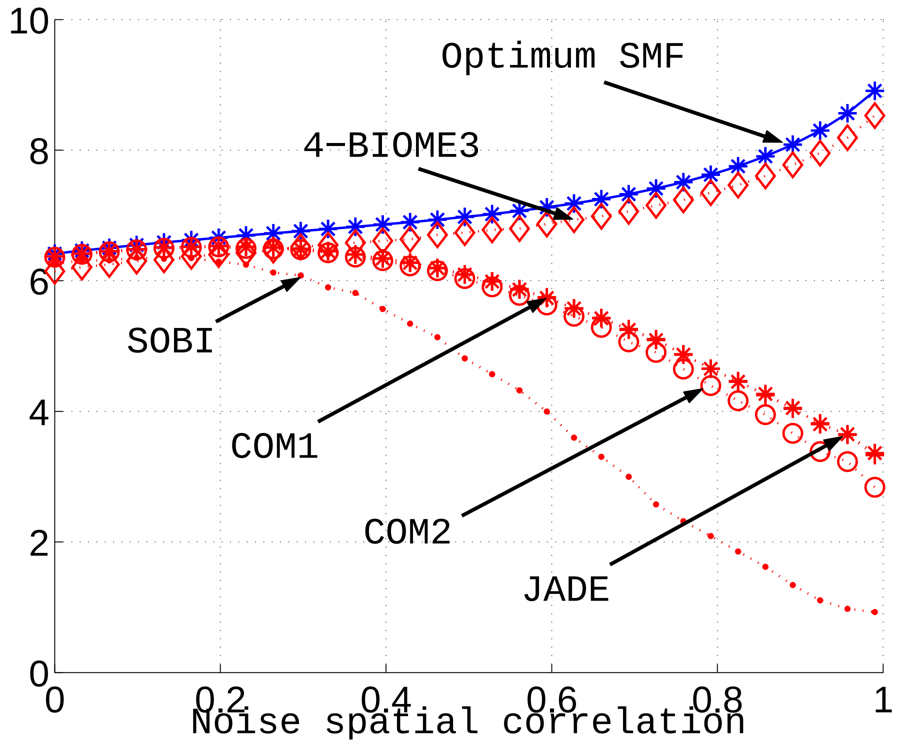

Figure 4.4: SINRM associated with source 3 for a SNR of 0 dB

On the other hand, figures 4.3 and 4.4 show that the well-known COM1, COM2, JADE and SOBI methods are strongly affected as soon as the noise spatial correlation is close to 1.

The underdetermined case

Finally, the 6-BIOME method is compared, in an underdetermined context, with the FOBIUM and JADE algorithms. Statistically independent sources with a raised cosine pulse shape of roll-off equal to 0.25, assumed synchronized, are generated with the same input SNR of 20 dB. The noise is spatially and temporally white Gaussian. Besides, the SixO virtual steering vectors of the sources are built orthogonal. Figure 4.5 and 4.6 show the variations of $\mathrm{D}(\boldsymbol{A}, \widehat{\boldsymbol{A}})$ (performance of the P sources), averaged over 200 realizations, at the output of the JADE, FOBIUM and 6-BIOME1 algorithms as a function of the number of samples.

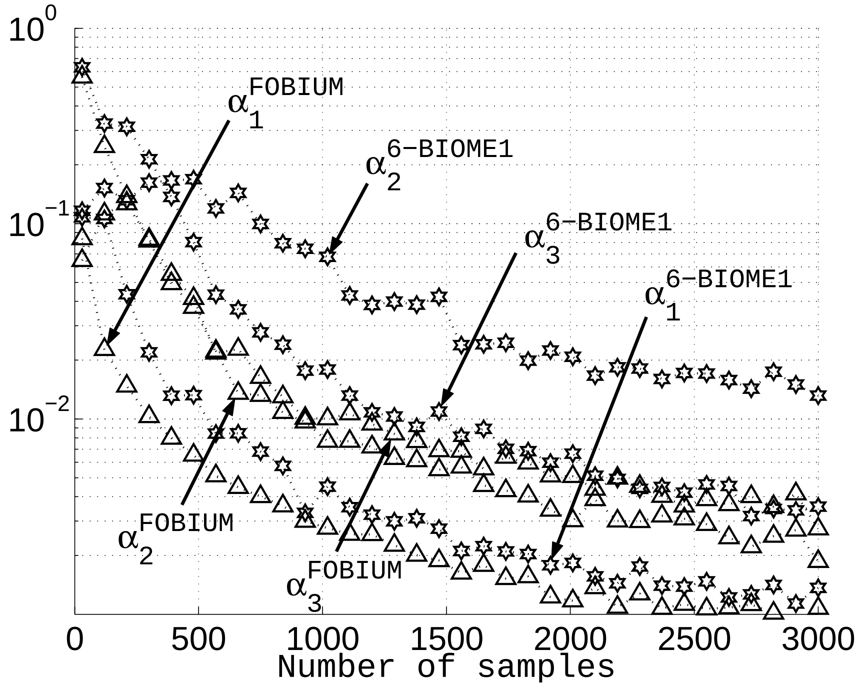

Figure 4.5: $\mathrm{D}\Big(\boldsymbol{A}, \widehat{\boldsymbol{A}}\Big)$ for a SNR of 20 dB

In figure 4.5, 1 BPSK and 2 QPSK are received by a UCA of $N = 3$ identical sensors of radius R such that $R/\lambda = 0.55$. Their symbol periods, $T_1 = 2T_e$, $T_2 = 3T_e$ and $T_3 = 2T_e$ respectively, and their carrier residus, $f_{c1}\,T_e = 0$, $f_{c2}\,T_e = 0.1$ and $f_{c3}\,T_e = 0.2$ respectively, are such that both QPSK have different FourO spectral densities: this assumption is required by the FOBIUM algorithm. Figure 4.5 shows the three α_p at the output of the FOBIUM and 6-BIOME1 methods as a function of samples. In fact, note the decreasing values toward zero of all the previous coefficients as the number of samples increases for both methods: whatever the method, FOBIUM or 6-BIOME1, the three sources have correctly been identified. Moreover, note that the SixO 6-BIOME1 algorithm is not ridiculous in terms of convergence rate, compared with the FourO FOBIUM method.

On the other hand, figure 4.6 shows BMI results obtained when 7 BPSK sources are received by a UCA of $N = 3$ identical sensors of radius R such that $R/\lambda = 0.55$. Their symbol periods are equal to twice the sample period and their carrier residus are all null. Instead of showing the variations of the seven α_p at the output of the JADE and 6-BIOME1 methods, we decided to show only the minimal and maximal variations of α_p associated with both algorithms, and denoted by $\min\{\alpha_p\}$ and $\max\{\alpha_p\}$ respectively. Whereas the JADE method obviously cannot identify all the steering vectors of sources in a underdetermined context, the 6-BIOME1 algorithm completely

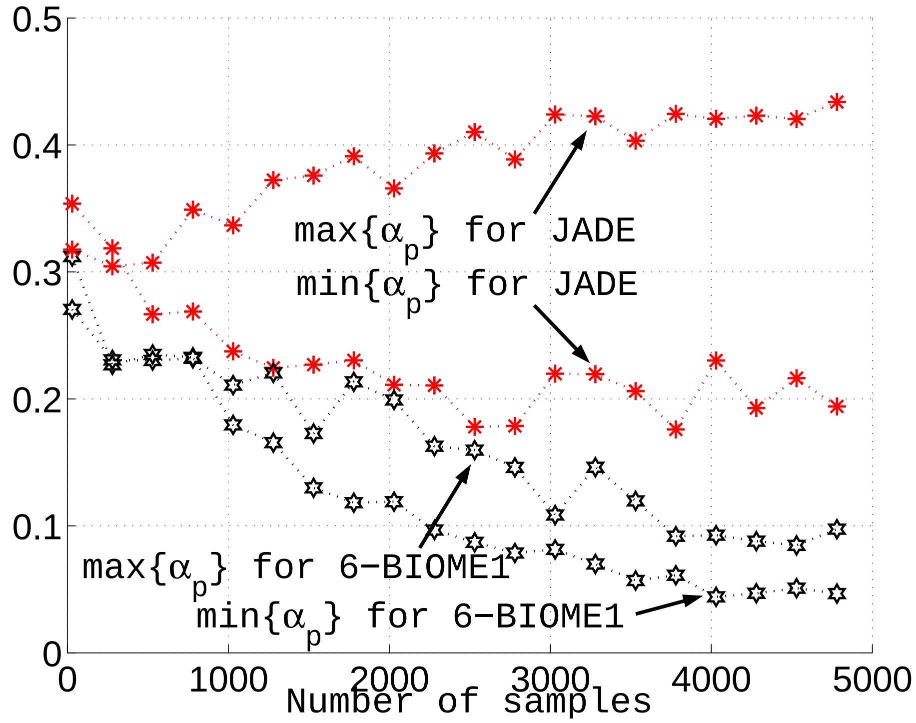

Figure 4.6: $\mathrm{D}\left(\boldsymbol{A}, \widehat{\boldsymbol{A}}\right)$ for a SNR of 20 dB

succeeds in identifying them, according to table 4.3 for $(q, \ell) = (2, 1)$. Moreover, according to figures 4.5 and 4.6, it appears that the sample number necessary for identifying accurately the P source steering vectors increases with P.

4.4 Conclusion

A family of new BMI methods, named BIOME, exploiting the information contained in the data statistics at an arbitrary even order has been proposed in this chapter. These new methods allow to process both over and underdetermined mixtures of sources, provided the latter have marginal HO cumulants with the same sign. The proposed methods are not sensitive to a Gaussian colored noise whose spatial coherence is unknown and allow the processing of a number of sources depending on both the kind of sensors and the array geometry, and increasing with both the number of sensors and the order of the data statistics. For underdetermined mixtures of sources, the proposed methods seem to outperform most of the methods currently available.

From a mathematical point of view, the so-called BIOME approaches allow to pose and to solve the BMI problem in terms of a non conventional joint approximate diagonalization of several given matrices, even in the presence of more inputs (sources) than observations (sensors). This problem is difficult to solve because of its structure. However, by ignoring part of the structure, it has been possible to compute in the LS sense the left and right transforms. More accurate numerical algorithms, taking fully into account the structure, still remain to be devised.

Chapter 5

Other contributions

5.1 The FOBIUM approach

Another new BMI method has also been proposed, exploiting the information contained in the FourO data statistics only, able to process both over and underdetermined mixtures of sources without the drawbacks of the existing methods, but assuming the sources have different trispectrum and have non zero kurtosis with the same sign. This new BMI method, called FOBIUM (Fourth Order Blind Identification of Underdetermined Mixtures of sources), corresponds to the FourO extension of the SOBI approach [45] [24] [2] and is able to blindly identify the steering vectors of up to $P = N^2 - N + 1$ sources, from an array of N sensors with spatial diversity only, and of up to N^2 sources, from an array of N different sensors. Moreover, this method is robust to a Gaussian spatially colored noise since it does not exploit the information contained in the SO data statistics. The FOBIUM approach is presented in detail in [40] and more particularly in appendix F. Finally, an application of the FOBIUM method will be soon presented in a forthcoming journal paper through the introduction of a FourO direction finding method, built from the blindly identified mixing matrix and called MAXCOR (MAXimum of spatial CORrelation), which is shown to be very powerful with respect to SO [64] and FourO subspace-based direction finding methods [5] [15] [60].

5.2 Blind separation of non zero-mean cyclostationary sources

Most of the SO and HO blind source separation methods developed this last decade aim at blindly separating statistically independent sources, assumed zero-mean, stationary and ergodic. Nevertheless, in many situations of practical interest, such as in radiocommunications contexts, the sources are non stationary and very often cyclostationary (digital modulations). The behavior of the current SO and FourO cumulant-based blind source separation methods in the presence of cyclostationary sources has been analysed, recently, in a previous paper [42], assuming zero-mean sources. However some cyclostationary sources used in practical situations are not zero-mean but have a first order cyclostationarity property, which is in particular the case for some AM signals and for some non linearly modulated digital sources such as FSK or some CPFSK sources. For such sources, the results presented in [42] do no longer hold, so it has been necessary to analyse the behavior and to propose adaptations of the current SO and FourO blind source separation methods for sources which are cyclostationary and cyclo-ergodic up to FO. These results are presented in [43] (see appendix G) and in [44] (see appendix H) [14] respectively.

Chapter 6

Conclusion

We addressed in this this report the blind identification problem of static linear underdetermined mixtures (i.e. in which the number of sources present exceeds in permanence the number of sensors), as well as the blind source separation problem in the overdetermined case, both in the presence of additive Gaussian noise, of unknown spatial coherence.

In order to process the latter problem, we proposed the ICAR method consisting of getting rid of the whitening stage, and of using exclusively HO statistics, namely FourO cumulants. More precisely, the redundancy theoretically present in the quadricovariance of the observations is exploited.

This concept can be extended to statistics of order strictly higher than 4, allowing for instance to address the case of underdetermined mixtures. Such extensions to order 6 have been proposed under the name of BIRTH. Surprisingly, identification methods solely based on the hexacovariance well succeed, despite their expected high estimation variance; this is due to the inherently good conditioning of the problem. The BIRTH algorithm is computationally simple but efficient and enables to identify the steering vectors of up to $P = N^2 - N + 1$ sources for arrays of N sensors with spatial diversity only, and up to $P = N^2$ for those with angular and polarization diversities.

More generally, a family of new BMI methods, named BIOME, exploiting the information contained in the data statistics at an arbitrary even order has been proposed. These new methods allow to process both over and underdetermined mixtures of sources, provided the latter have marginal HO cumulants with the same sign. The proposed methods are not sensitive to a Gaussian colored noise whose spatial coherence is unknown and allow the processing of a number of sources depending on both the kind of sensors and the array geometry, and increasing with both the number of sensors and the order of the data statistics. For underdetermined mixtures of sources, the proposed methods seem to outperform most of the methods currently available.

Moreover, we have examined the asymptotic performances (*e.g.* covariance of estimate) of contrast-based BSS algorithms by proposing a functional approach. As an illustration, 3 FourO contrast criteria already compared by computer experiments, have been mainly focused on, for asymptotic performance analysis. Forthcoming works will consist of looking for the contrast criterion associated with ICAR in order to analyse accurately his performance using for instance the latter functional approach.

Now, it can be interesting to compare the BIOME solution with the exact one given by the minimization of the mutual information, which is defined as the Kullback divergence between the source joint distribution and the product of the marginal ones. Note that a practical way to approximate the mutual information consists of computing an Edgeworth expansion of its negentropy components.

In addition, we will soon analyse the computational speeds of the $2q$-BIOME methods, and test the latter algorithms with experimental signals, say, non synthetic signals, borrowed from the radiocommunication context.

Besides, the blind source extraction problem deserves attention especially in the underdetermined case, assuming the mixture is known (or beforehand identified). However, as we said it in section 1.3, it is a difficult problem since the underdetermined mixtures cannot be linearly inverted. Moreover, we will try to process this problem in a way as blind as possible, i.e. limiting the source a priori assumptions.

Eventually, as shown in the report, the proposed ICAR (see chapter 2), BIRTH (see chapter 3) and BIOME (see chapter 4) approaches can tolerate (in their current form), but do not totally exploit, cyclostationarity of the sources such as in [41]: this will be the subject of forthcoming works.

Appendix A

Proof of the second matrix multilinearity property (4.1)

Assuming (**A1**)-(**A2**), the $2q$-th order statistics $C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}}$ defined by (1.9) may be described, using (1.4) and the multilinearity property shared by cumulants [55] [19, pp. 1-24], by

$$C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}} = \sum_{p=1}^{P} C_{p, .., p, \boldsymbol{s}}^{p, .., p} \left(\prod_{m=1}^{q} A(i_m, p)\right) \left(\prod_{m=q+1}^{2q} A(i_m, p)^*\right) \tag{A.1}$$

It is straightforward to show that $\left(\prod_{m=1}^{q-\ell} A(i_m, p)\right) \left(\prod_{m=2q-\ell+1}^{2q} A(i_m, p)^*\right)$ is the I_1^ℓ-th component of vector $\left[\boldsymbol{a}_p^{\otimes q-\ell} \otimes (\boldsymbol{a}_p^*)^{\otimes \ell}\right]$ and that $\left(\prod_{m=q+1}^{2q-\ell} A(i_m, p)^*\right) \left(\prod_{m=q-\ell+1}^{q} A(i_m, p)\right)$ is the I_2^ℓ-th component of vector $\left[\boldsymbol{a}_p^{\otimes q-\ell} \otimes (\boldsymbol{a}_p^*)^{\otimes \ell}\right]^*$ where I_1^ℓ, I_2^ℓ are given by (1.12) and (1.13). Consequently, since $\left[\boldsymbol{a}_p^{\otimes q-\ell} \otimes (\boldsymbol{a}_p^*)^{\otimes \ell}\right]$ is the p-th column vector of matrix $\boldsymbol{\mathcal{A}}_q^\ell$ (4.2), equation (A.1) may be written as

$$C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}} = \sum_{p=1}^{P} C_{p, .., p, \boldsymbol{s}}^{p, .., p} \; \boldsymbol{\mathcal{A}}_q^\ell(I_1^\ell, p) \; \boldsymbol{\mathcal{A}}_q^\ell(I_2^\ell, p)^* \tag{A.2}$$

where $\boldsymbol{\mathcal{A}}_q^\ell(n, p)$ is the (n, p)-th component of the $N^q \times P$ matrix $\boldsymbol{\mathcal{A}}_q^\ell$. So, since $\boldsymbol{\zeta}_{2q,\boldsymbol{s}}$ denotes the $P \times P$ invertible diagonal matrix $\mathrm{Diag}\left[C_{1, 1, ...1, \boldsymbol{s}}^{1, 1, .., 1}, \; C_{2, 2, .., 2, \boldsymbol{s}}^{2, 2, .., 2}, \cdots, \; C_{P, P, .., P, \boldsymbol{s}}^{P, P, .., P}\right]$, equation (A.2) may take the following expression

$$C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}} = \sum_{p=1}^{P} \boldsymbol{\mathcal{A}}_q^\ell(I_1^\ell, p) \; \boldsymbol{\zeta}_{2q,\boldsymbol{s}}(p, p) \; {\boldsymbol{\mathcal{A}}_q^\ell}^{\mathsf{H}}(p, I_2^\ell). \tag{A.3}$$

That means

$$C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}} = \left[\boldsymbol{\mathcal{A}}_q^\ell \; \boldsymbol{\zeta}_{2q,\boldsymbol{s}} \; {\boldsymbol{\mathcal{A}}_q^\ell}^{\mathsf{H}}\right](I_1^\ell, I_2^\ell). \tag{A.4}$$

And, since quantity $C_{i_1, i_2, .., i_q, \boldsymbol{x}}^{i_{q+1}, .., i_{2q}}$ is also the (I_1^ℓ, I_2^ℓ)-th component of the $N^q \times N^q$ matrix $\mathcal{C}_{2q, \boldsymbol{x}}^\ell$, according to (1.11), we finally have

$$\boldsymbol{\mathcal{C}}_{2q, \boldsymbol{x}}^\ell = \boldsymbol{\mathcal{A}}_q^\ell \; \boldsymbol{\zeta}_{2q, \boldsymbol{s}} \; \boldsymbol{\mathcal{A}}_q^{\ell\,\mathsf{H}}. \tag{A.5}$$

Appendix B

Proof of propositions 4 and 8

Proposition 8 may be rewritten as

$$(\mathbf{A4}) \;\Rightarrow\; \{\forall\, 1 \leq p_1 \neq p_2 \leq P,\;\; \exists\, 1 \leq n_1 \neq n_2 \leq N \;:\;\; D_{n_1,n_2}(p_1,p_1) \neq D_{n_1,n_2}(p_2,p_2)\} \tag{B.1}$$

To prove it, assume the contrary:

$$\exists\, 1 \leq p_1 \neq p_2 \leq P \;:\;\; \forall\, 1 \leq n_1 \neq n_2 \leq N,\;\; D_{n_1,n_2}(p_1,p_1) = D_{n_1,n_2}(p_2,p_2) \tag{B.2}$$

This implies, since $\boldsymbol{D}_{n_1,n_2} = \boldsymbol{\Phi}_{n_1}^{-1}\boldsymbol{\Phi}_{n_2}$ are $P \times P$ diagonal full rank matrices, that

$$\exists\, 1 \leq p_1 \neq p_2 \leq P \;:\;\; \forall\, 1 \leq n_1 \neq n_2 \leq N,\;\; \frac{\Phi_{n_2}(p_1,p_1)}{\Phi_{n_1}(p_1,p_1)} = \frac{\Phi_{n_2}(p_2,p_2)}{\Phi_{n_1}(p_2,p_2)} \tag{B.3}$$

which is equivalent, according to (4.3), to

$$\exists\, 1 \leq p_1 \neq p_2 \leq P \;:\;\; \forall\, 1 \leq n_1 \neq n_2 \leq N,\;\; \frac{A(n_2,p_1)}{A(n_1,p_1)} = \frac{A(n_2,p_2)}{A(n_1,p_2)} \tag{B.4}$$

This means

$$\exists\, 1 \leq p_1 \neq p_2 \leq P \;: \qquad \boldsymbol{a}_{p_1} \propto \boldsymbol{a}_{p_2} \tag{B.5}$$

In other words, assuming (B.2) implies that at least two columns of $\boldsymbol{A}$ are collinear, which contradicts (**A4**). Consequently, proposition 8 is true.

Appendix C

Proof of theorem 2

Each column $\boldsymbol{b}_p$ of $[\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}]^{1/2}\boldsymbol{V}_{sol}$ is defined, according to (4.15), by

$$\forall\, 1 \leq p \leq P, \qquad \boldsymbol{b}_p = \lambda_{\xi(p)}\, [(\boldsymbol{a}_{\xi(p)})^{\otimes q-\ell} \otimes (\boldsymbol{a}^{*}_{\xi(p)})^{\otimes \ell}] \quad \text{of size } N^q \times 1 \tag{C.1}$$

where $\xi(\cdot)$ is a bijective function of $\{1, 2, \dots, P\}$ into itself (i.e. a permutation function) and where $|\lambda_p| = |\, C^{p,p,\dots,p}_{p,p,\dots,p,\boldsymbol{s}}|^{1/2}$, $|\cdot|$ denoting the complex modulus operator. Moreover, vectors $\boldsymbol{b}_p$ may be written as

$$\boldsymbol{b}_p = [\boldsymbol{b}_p(1)^{\mathsf{T}}\ \boldsymbol{b}_p(2)^{\mathsf{T}} \cdots \boldsymbol{b}_p(M)^{\mathsf{T}}]^{\mathsf{T}} \tag{C.2}$$

where $M = N^{q-2}$ and $\boldsymbol{b}_p(m)$ is of size $N^2 \times 1$. Now it is important to notice that each vector $\boldsymbol{b}_p(m)$ $(1 \leq m \leq M)$ may be expressed as a Kronecker product of the column vector $\boldsymbol{a}_p$ of $\boldsymbol{A}$ by itself:

$$\boldsymbol{b}_p(m) = \begin{cases} \lambda_{\xi(p)} \left(\prod_{j=1}^{q-2} A(n_j, \xi(p))\right) \left[\boldsymbol{a}_{\xi(p)} \otimes \boldsymbol{a}_{\xi(p)}\right] & \text{if } \ell = 0 \\ \lambda_{\xi(p)} \left(\prod_{j=1}^{q-2} A(n_j, \xi(p))\right) \left[\boldsymbol{a}_{\xi(p)} \otimes \boldsymbol{a}^{*}_{\xi(p)}\right] & \text{if } \ell = 1 \\ \lambda_{\xi(p)} \left(\prod_{j=1}^{q-\ell} A(n_j, \xi(p))\right) \left(\prod_{j=q-\ell+1}^{q-2} A(n_j, \xi(p))^{*}\right) \left[\boldsymbol{a}_{\xi(p)} \otimes \boldsymbol{a}_{\xi(p)}\right]^{*} & \text{o.w.} \end{cases} \tag{C.3}$$

So we transform the M vectors $\boldsymbol{b}_p(m)$ of size $N^2 \times 1$ into $N \times N$ matrices $\boldsymbol{B}_p(m)$ $(1 \leq m \leq M)$ where the (i_1, i_2)-th component of $\boldsymbol{B}_p(m)$ corresponds to the $\varphi([i_2\ i_1])$-th component of $\boldsymbol{b}_p(m)$ so that

$$\boldsymbol{B}_p(m) = \begin{cases} \lambda_{\xi(p)} \left(\prod_{j=1}^{q-2} A(n_j, \xi(p))\right) \left[\boldsymbol{a}_{\xi(p)}\ \boldsymbol{a}_{\xi(p)}^{\mathsf{T}}\right] & \text{if } \ell = 0 \\ \lambda_{\xi(p)} \left(\prod_{j=1}^{q-2} A(n_j, \xi(p))\right) \left[\boldsymbol{a}_{\xi(p)}\ \boldsymbol{a}_{\xi(p)}^{\mathsf{H}}\right]^{*} & \text{if } \ell = 1 \\ \lambda_{\xi(p)} \left(\prod_{j=1}^{q-\ell} A(n_j, \xi(p))\right) \left(\prod_{j=q-\ell+1}^{q-2} A(n_j, \xi(p))^{*}\right) \left[\boldsymbol{a}_{\xi(p)}\ \boldsymbol{a}_{\xi(p)}^{\mathsf{T}}\right]^{*} & \text{o.w.} \end{cases} \tag{C.4}$$

Consequently, plugging (C.4) into (4.18), the set of matrices Δ^{ℓ}_p may be expressed as

$$\Delta^{\ell}_p = \left\{ \mu^{\ell}_{p,n_j}\ \boldsymbol{a}_{\xi(p)}\ \boldsymbol{a}_{\xi(p)}^{\mathsf{H}} \,/\, 1 \leq n_j \leq N \right\} \tag{C.5}$$

with

$$\mu_{p,n_j}^{\ell} = \begin{cases} |\lambda_{\xi(p)}|^2 \left|\prod_{j=1}^{q-2} A(n_j,\xi(p))\right|^2 \left\|\boldsymbol{a}_{\xi(p)}\right\|^2 & \text{if } \ell = 0 \\ \lambda_{\xi(p)}^* \prod_{j=1}^{q-2} A(n_j,\xi(p))^* & \text{if } \ell = 1 \\ |\lambda_{\xi(p)}|^2 \left|\left(\prod_{j=1}^{q-\ell} A(n_j,\xi(p))\right) \left(\prod_{j=q-\ell+1}^{q-2} A(n_j,\xi(p))^*\right)\right|^2 \left\|\boldsymbol{a}_{\xi(p)}\right\|^2 & \text{o.w.} \end{cases} \tag{C.6}$$

where $\|\cdot\|$ denotes the norm operator respectively. So a joint diagonalization of matrices belonging to Δ_p^{ℓ} indeed allows one to extract the $\xi(p)$-th column vector $\boldsymbol{a}_{\xi(p)}$ of $\boldsymbol{A}$.

Appendix D

Multivariate high-order complex cumulants

Cumulants are given as a function of moments in statistics text books, but only in the real case [55]. Therefore, it seems useful to report here their expressions in the complex case. Again, we consider only zero-mean complex variables that are distributed symmetrically with respect to the origin. However, they do not need to be circularly distributed. Below, cumulants are denoted with κ and moments with μ. As before, superscripts correspond to variables that are complex conjugated. We have for orders 4 and 6:

$$\begin{aligned}
\kappa_{ijk\ell} &= \mu_{ijk\ell} - [3]\mu_{ij}\mu_{k\ell} \\
\kappa_{ijk}^{\ell} &= \mu_{ijk}^{\ell} - [3]\mu_{ij}\mu_{k}^{\ell} \\
\kappa_{ij}^{k\ell} &= \mu_{ij}^{k\ell} - [2]\mu_{i}^{k}\mu_{j}^{\ell} - \mu_{ij}\mu^{k\ell}
\end{aligned}$$

$$\begin{aligned}
\kappa_{ijk\ell mn} &= \mu_{ijk\ell mn} - [15]\mu_{ijk\ell}\mu_{mn} + 2[15]\mu_{ij}\mu_{k\ell}\mu_{mn} \\
\kappa_{ijk\ell m}^{n} &= \mu_{ijk\ell m}^{n} - [5]\mu_{ijk\ell}\mu_{m}^{n} - [10]\mu_{ijk}^{n}\mu_{\ell m} \\
&\quad +2[15]\mu_{ij}\mu_{k\ell}\mu_{m}^{n} \\
\kappa_{ijk\ell}^{mn} &= \mu_{ijk\ell}^{mn} - \mu_{ijk\ell}\mu^{mn} - [8]\mu_{ijk}^{m}\mu_{\ell}^{n} - [6]\mu_{ij}^{mn}\mu_{k\ell} \\
&\quad +[6]\mu_{ij}\mu_{k\ell}\mu^{mn} + 2[12]\mu_{ij}\mu_{k}^{m}\mu_{\ell}^{n} \\
\kappa_{ijk}^{\ell mn} &= \mu_{ijk}^{\ell mn} - [3]\mu_{ijk}^{\ell}\mu^{mn} - [9]\mu_{ij}^{\ell m}\mu_{k}^{n} - [3]\mu_{ij}\mu_{k}^{\ell mn} \\
&\quad +2[9]\mu_{ij}\mu_{k}^{\ell}\mu^{mn} + 2[6]\mu_{i}^{\ell}\mu_{j}^{m}\mu_{k}^{n}
\end{aligned}$$

and eventually for order 8:

$$\begin{aligned}
\kappa_{ijk\ell mnpq} &= \mu_{ijk\ell mnpq} - [28]\mu_{ijk\ell mn}\mu_{pq} - [35]\mu_{ijk\ell}\mu_{mnpq} \\
&+2[210]\mu_{ijk\ell}\mu_{mn}\mu_{pq} - 6[105]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu_{pq} \\
\kappa^{q}_{ijk\ell mnp} &= \mu^{q}_{ijk\ell mnp} - [7]\mu_{ijk\ell mn}\mu^{q}_{p} - [21]\mu^{q}_{ijk\ell m}\mu_{np} \\
&-[35]\mu_{ijk\ell}\mu^{q}_{mnp} + 2[105]\mu^{q}_{ijk}\mu_{\ell m}\mu_{np} \\
&+2[105]\mu_{ijk\ell}\mu_{mn}\mu^{q}_{p} - 6[105]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu^{q}_{p} \\
\kappa^{pq}_{ijk\ell mn} &= \mu^{pq}_{ijk\ell mn} - \mu_{ijk\ell mn}\mu^{pq} - [12]\mu^{p}_{ijk\ell m}\mu^{q}_{n} \\
&-[15]\mu^{pq}_{ijk\ell}\mu_{mn} - [15]\mu_{ijk\ell}\mu^{pq}_{mn} - [20]\mu^{p}_{ijk}\mu^{q}_{\ell mn} \\
&+2[15]\mu_{ijk\ell}\mu_{mn}\mu^{pq} + 2[30]\mu_{ijk\ell}\mu^{p}_{m}\mu^{q}_{n} \\
&+2[120]\mu^{p}_{ijk}\mu_{\ell m}\mu^{q}_{n} + 2[45]\mu^{pq}_{ij}\mu_{k\ell}\mu_{mn} \\
&-6[15]\mu_{ij}\mu_{k\ell}\mu_{mn}\mu^{pq} - 6[90]\mu_{ij}\mu_{k\ell}\mu^{p}_{m}\mu^{q}_{n}
\end{aligned}$$

$$\begin{aligned}
\kappa^{npq}_{ijk\ell m} &= \mu^{npq}_{ijk\ell m} - [3]\mu^{n}_{ijk\ell m}\mu^{pq} - [15]\mu^{np}_{ijk\ell}\mu^{q}_{m} \\
&-[10]\mu^{npq}_{ijk}\mu_{\ell m} - [5]\mu_{ijk\ell}\mu^{npq}_{m} - [30]\mu^{n}_{ijk}\mu^{pq}_{\ell m} \\
&+2[15]\mu_{ijk\ell}\mu^{n}_{m}\mu^{pq} + 2[30]\mu^{n}_{ijk}\mu_{\ell m}\mu^{pq} \\
&+2[60]\mu^{n}_{ijk}\mu^{p}_{\ell}\mu^{q}_{m} + 2[90]\mu^{np}_{ij}\mu_{k\ell}\mu^{q}_{m} \\
&+2[15]\mu^{npq}_{i}\mu_{jk}\mu_{\ell m} \\
&-6[45]\mu_{ij}\mu_{k\ell}\mu^{n}_{m}\mu^{pq} - 6[60]\mu_{ij}\mu^{n}_{k}\mu^{p}_{\ell}\mu^{q}_{m} \\
\kappa^{mnpq}_{ijk\ell} &= \mu^{mnpq}_{ijk\ell} - [6]\mu^{mn}_{ijk\ell}\mu^{pq} - [6]\mu^{mnpq}_{ij}\mu_{k\ell} - [16]\mu^{mnp}_{ijk}\mu^{q}_{\ell} \\
&-[16]\mu^{m}_{ijk}\mu^{npq}_{\ell} - [18]\mu^{mn}_{ij}\mu^{pq}_{k\ell} - \mu_{ijk\ell}\mu^{mnpq} \\
&+2[\bar{2}]([3]\mu_{ijk\ell}\mu^{mn}\mu^{pq} + [48]\mu^{m}_{ijk}\mu^{n}_{\ell}\mu^{pq}) \\
&+2[36]\mu^{mn}_{ij}\mu_{k\ell}\mu^{pq} + 2[72]\mu^{mn}_{ij}\mu^{p}_{k}\mu^{q}_{\ell} \\
&-6[72]\mu_{ij}\mu^{m}_{k}\mu^{n}_{\ell}\mu^{pq} - 6[24]\mu^{m}_{i}\mu^{n}_{j}\mu^{p}_{k}\mu^{q}_{\ell} \\
&-6[9]\mu_{ij}\mu_{k\ell}\mu^{mn}\mu^{pq}
\end{aligned}$$

where $[d]\prod_m \mu^{i_{r(m)+1},\ldots,i_{r(m)+s(m)}}_{i_1,\ldots,i_{r(m)},\boldsymbol{x}}(k)$ denotes McCullagh bracket notation defined in [55], which is written instead of a linear combination of d terms of the form $\prod_m \mu^{i_{r(m)+1},\ldots,i_{r(m)+s(m)}}_{i_1,\ldots,i_{r(m)},\boldsymbol{x}}(k)$ permuting on one hand superscripts, and on the other hand subscripts such as, for instance

$$\begin{aligned}
&[6]\mu^{i_3}_{i_1,i_2,\boldsymbol{x}}(k)\,\mu^{i_5}_{i_4,\boldsymbol{x}}(k) = \mu^{i_3}_{i_1,i_2,\boldsymbol{x}}(k)\,\mu^{i_5}_{i_4,\boldsymbol{x}}(k) + \mu^{i_5}_{i_1,i_2,\boldsymbol{x}}(k)\,\mu^{i_3}_{i_4,\boldsymbol{x}}(k) + \\
&\mu^{i_3}_{i_1,i_4,\boldsymbol{x}}(k)\,\mu^{i_5}_{i_2,\boldsymbol{x}}(k) + \mu^{i_5}_{i_1,i_4,\boldsymbol{x}}(k)\,\mu^{i_3}_{i_2,\boldsymbol{x}}(k) + \mu^{i_3}_{i_4,i_2,\boldsymbol{x}}(k)\,\mu^{i_5}_{i_1,\boldsymbol{x}}(k) + \\
&\mu^{i_5}_{i_4,i_2,\boldsymbol{x}}(k)\,\mu^{i_3}_{i_1,\boldsymbol{x}}(k) \qquad \text{(D.1)}
\end{aligned}$$

Appendix **E**

Expression of second order differentials

For contrast Υ_1, we give below the expressions of the coefficients of the second order differential with respect to $\boldsymbol{U}$ (omitting subscript $\boldsymbol{y}$ in $C_{i,\ell,\boldsymbol{y}}^{j,k}$):

$$\Theta_{qr}^{q'r'} = \begin{cases} \Re\{\delta(q'-r)(C_{r',q}^{q,q} - 2C_{q,r}^{r',r} - C_{r',q}^{r,r}) + \delta(q'-q)(C_{r',r}^{q,q} + 2C_{q,r}^{r',q} - C_{r',r}^{r,r}) + \\ \qquad \delta(r'-r)(C_{q',q}^{r,r} + 2C_{q,r}^{q',r} - C_{q',q}^{q,q}) + \delta(r'-q)(C_{q',r}^{r,r} - 2C_{r,q}^{q',q} - C_{r,q}^{q,q})\} \\ \qquad \text{if } q < r \text{ and } q' < r' \\ -\Im\{(\delta(r'-r) - \delta(r'-q))(C_{r,q}^{q,q} + C_{q,r}^{r,r})\} \text{ if } q < r \text{ and } q' = r' \\ -\Im\{\delta(r'-r)(C_{q',q}^{q,q} + 2C_{q,r}^{q',r} - C_{q',q}^{r,r}) + \delta(q'-r)(C_{r',q}^{q,q} + 2C_{q,r}^{r',r} - C_{r',q}^{r,r}) + \\ \qquad \delta(r'-q)(C_{q',r}^{q,q} - 2C_{q,r}^{q',q} - C_{q',r}^{r,r}) + \delta(q'-q)(C_{r',r}^{q,q} - 2C_{r,q}^{r',q} - C_{r',r}^{r,r})\} \\ \qquad \text{if } q < r \text{ and } q' > r' \\ -\Im\{\delta(q'-q)(C_{r',r}^{r,r} + 2C_{r,q}^{r',q} + C_{r',r}^{q,q}) - \delta(r'-q)(C_{q',r}^{r,r} + 2C_{r,q}^{q',q} + C_{q',r}^{q,q}) + \\ \qquad \delta(q'-r)(C_{r',q}^{q,q} + 2C_{q,r}^{r',r} + C_{q,r'}^{r,r}) - \delta(r-r')(C_{q',q}^{q,q} + 2C_{q,r}^{q',r} + C_{q,q'}^{r,r})\} \\ \qquad \text{if } q > r \text{ and } q' < r' \\ \Re\{(\delta(r'-q) - \delta(r'-r))(C_{r,q}^{q,q} - C_{q,r}^{r,r})\} \text{ if } q > r \text{ and } q' = r' \\ -\Re\{\delta(r'-q)(C_{q',r}^{r,r} - 2C_{r,q}^{q',q} + C_{q',r}^{q,q}) + \delta(q'-q)(C_{r',r}^{r,r} - 2C_{r,q}^{r',q} + C_{r',r}^{q,q}) + \\ \qquad \delta(r'-r)(C_{q',q}^{q,q} - 2C_{q,r}^{q',r} + C_{q',q}^{r,r}) + \delta(q'-r)(C_{r',q}^{q,q} - 2C_{q,r}^{r',r} + C_{r',q}^{r,r})\} \\ \qquad \text{if } q > r \text{ and } q' > r' \\ 0 \qquad \text{if } q = r \end{cases} \tag{E.1}$$

and those of the second order differential with respect to $\boldsymbol{C_x}$:

$$\Theta_{qr}^{ijkl} = \begin{cases} U(r,i)\,U(q,j)^*\,U(q,k)^*\,U(q,l) + U(q,i)\,U(r,j)^*\,U(q,k)^*\,U(q,l) - \\ \qquad U(q,i)\,U(r,j)^*\,U(r,k)^*\,U(r,l) - U(r,i)\,U(q,j)^*\,U(r,k)^*\,U(r,l) \\ \qquad \text{if } q < r \\ \\ 0 \qquad \text{if } q = r \\ \\ U(q,i)\,U(r,j)^*\,U(r,k)^*\,U(r,l) - U(r,i)\,U(q,j)^*\,U(r,k)^*\,U(r,l) + \\ \qquad U(r,i)\,U(q,j)^*\,U(q,k)^*\,U(q,l) - U(q,i)\,U(r,j)^*\,U(q,k)^*\,U(q,l) \\ \qquad \text{if } q > r \end{cases} \tag{E.2}$$

Appendix F

The FOBIUM approach

FOURTH ORDER BLIND IDENTIFICATION OF UNDERDETERMINED MIXTURES OF SOURCES (FOBIUM)

Anne Ferréol, Laurent Albera and Pascal Chevalier

THALES COMMUNICATIONS – TSR/TSI/LTC
66 rue du Fossé Blanc, 92231 Gennevilliers, France

ABSTRACT

Most of the current Second Order (SO)[1] and Fourth Order (FO)[3][6] Blind Sources Identification (BSI) methods aim at blindly identifying the steering vectors of statistically independent sources, provided the number of sources is not greater than the number of sensors. However in practical situations, the probability of receiving more sources than sensors increases with the reception bandwidth. In this context the purpose of this paper is to propose a new attractive FO BSI method, able to identify the steering vector of more sources than sensors, jointly with a new pertinent performance criterion for the quality evaluation of the BSI process. The new method implements a FO pre-whitening step and exploits the tri-spectrum diversities of the sources.

1. INTRODUCTION

For more than a decade, SO [1] and FO [3][6] methods have been developed to blindly identify the steering vectors of several statistically independent sources, provided the number of sources remains lower than or equal to the number of sensors. However, in practical situations, such as in the HF context, the reception of more sources than sensors is possible and its probability increases with the reception bandwidth. To process such situations, several BSI methods have been developed this last decade, among which we find the methods [2] [7-8] [10]. The methods proposed in [2] and [7-8] only exploit the information contained in the FO statistics of the data whereas the one proposed in [10] exploits the information contained in one of the characteristic function of the observations. However, all these methods suffer from severe drawbacks in operational contexts. Indeed, the method [2] is still very difficult to implement and does not ensure the BSI of the sources steering vectors when the sources have the same kurtosis. The methods [7-8] assume non circular sources and fail in separating circular sources, omnipresent in practice. Finally, the method [10] has been developed only for real mixtures of real-valued sources and is probably not robust to an over estimation of the source number. To overcome these drawbacks we propose in this paper a new FO method, corresponding to the FO extension of the SOBI algorithm [1], able to blindly identify the steering vectors of up to $N^2 - N + 1$ sources with N sensors, without the previously mentioned drawbacks but assuming the sources have different tri-spectrum and have non zero kurtosis with the same sign (the latter assumption is generally verified in radiocommunications contexts). This method implements a FO pre-whitening step and fully exploits the assumed non whiteness property of the sources. Finally a new performance criterion, able to quantify the identification quality of the steering vector of each source and allowing the quantitative comparison of two methods for the blind identification (BI) of each source is proposed.

2. PROBLEM FORMULATION

A noisy mixture of P Narrow-Band (NB) statistically independent sources is assumed to be received by an array of N sensors. The vector, $\boldsymbol{x}(t)$, of the complex envelopes of the signals at the output of the sensors is thus given by

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} m_p(t)\, \boldsymbol{a}_p + \boldsymbol{n}(t) = A\, \boldsymbol{m}(t) + \boldsymbol{n}(t) \qquad (1)$$

where $m_p(t)$ is the p-th component of the vector $\boldsymbol{m}(t)$, assumed zero-mean and stationary, $\boldsymbol{n}(t)$ is the noise vector, assumed zero-mean, stationary, Gaussian, spatially and temporally white in the reception band, $\boldsymbol{a}_p$ corresponds to the steering vector of the source p and A is the $(N\text{x}P)$ matrix whose columns are the vectors $\boldsymbol{a}_p$.

The problem adressed in this paper is the BI of the steering vectors $\boldsymbol{a}_p$ from the FO statistics of the data.

3. THE FOBIUM METHOD

The purpose of the FOBIUM method is to extend the SOBI method [1] at the FO. It firstly implements a FO pre-whitening step aiming at orthonormalizing the so-called *virtual steering vector* [5] of the sources in some data quadricovariance matrices and secondly it jointly diagonalizes several well chosen pre-whitened quadricovariance matrices in order to identify the steering vectors of more sources than sensors. The number of sources able to be processed by this method is addressed in section 4.

3.1 FO statistics of the data

Under the assumption of zero-mean stationary sources, the FO statistics of the observations are characterized by the $(N^2 \mathrm{x} N^2)$ quadricovariance matrices $Q_x(\tau_1,\tau_2,\tau_3)$, whose elements, $Q_x(\tau_1,\tau_2,\tau_3)[i, j, k, l]$, are defined by

$$Q_x(\tau_1,\tau_2,\tau_3)[i, j, k, l] \triangleq \mathrm{Cum}(x_i(t), x_j(t-\tau_1)^*, x_k(t-\tau_2)^*, x_l(t-\tau_3)) \quad (2)$$

where * means complex conjugate and $x_i(t)$ is the i^{th} component of the vector $\boldsymbol{x}(t)$. Using (1) into (2) and assuming that $Q_x(\tau_1,\tau_2,\tau_3)[i, j, k, l]$ is the element $[N(i-1)+j,\ N(k-1)+l]$ of the matrix $Q_x(\tau_1,\tau_2,\tau_3)$, we obtain the expression of the latter, given, under a Gaussian noise assumption, by

$$Q_x(\tau_1,\tau_2,\tau_3) = (A \otimes A^*)\, Q_m(\tau_1,\tau_2,\tau_3)\, (A \otimes A^*)^{\mathrm{H}} \quad (3)$$

where $Q_m(\tau_1,\tau_2,\tau_3)$ is the $(P^2 \mathrm{x} P^2)$ quadricovariance matrix of $\boldsymbol{m}(t)$, $\otimes$ is the Kronecker product and $^{\mathrm{H}}$ means transpose and complex conjugate.

Under the assumption of statistically independent sources, the matrix $Q_m(\tau_1,\tau_2,\tau_3)$ contains at least $P^4 - P$ zeros and the expression (3) degenerates in a simpler one given by

$$Q_x(\tau_1,\tau_2,\tau_3) = \sum_{p=1}^{P} c_p(\tau_1,\tau_2,\tau_3)\, (\boldsymbol{a}_p \otimes \boldsymbol{a}_p^*)\, (\boldsymbol{a}_p \otimes \boldsymbol{a}_p^*)^{\mathrm{H}} \quad (4a)$$

$$= A_Q\ C_m(\tau_1,\tau_2,\tau_3)\ A_Q^{\ \mathrm{H}} \quad (4b)$$

where A_Q is the $(N^2 \mathrm{x}\ P)$ matrix defined by $A_Q \triangleq [(\boldsymbol{a}_1 \otimes \boldsymbol{a}_1^*), \ldots, (\boldsymbol{a}_p \otimes \boldsymbol{a}_p^*)]$, $C_m(\tau_1,\tau_2,\tau_3)$ is the $(P \mathrm{\ x\ } P)$ diagonal matrix defined by $C_m(\tau_1,\tau_2,\tau_3) \triangleq \mathrm{diag}[c_1(\tau_1,\tau_2,\tau_3),\ldots, c_P(\tau_1,\tau_2,\tau_3)]$ and $c_p(\tau_1,\tau_2,\tau_3)$ is defined by

$$c_p(\tau_1,\tau_2,\tau_3) \triangleq \mathrm{Cum}(m_p(t), m_p(t-\tau_1)^*, m_p(t-\tau_2)^*, m_p(t-\tau_3)) \quad (5)$$

The expression (4b), which has an algebraic structure similar to that of data correlation matrices [1], is at the basis of the FOBIUM method as it is shown in the next sections.

We note in the following $Q_x \triangleq Q_x(0, 0, 0)$, $c_p \triangleq c_p(0, 0, 0)$, $C_m \triangleq C_m(0, 0, 0)$ and we obtain

$$Q_x = A_Q\, C_m\, A_Q^{\ \mathrm{H}} \quad (6)$$

We also assume in the following that $P \leq N^2$, the matrix A_Q is full rank, the c_p, $1 \leq p \leq P$, are non zero (non Gaussian sources) and have the same sign and whatever the couple (i, j) of sources, it exists at least three delays (τ_1,τ_2,τ_3) such that $|\tau_1|+|\tau_2|+|\tau_3| \neq 0$ and

$$c_i(\tau_1,\tau_2,\tau_3) \,/\, |c_i| \neq c_j(\tau_1,\tau_2,\tau_3) \,/\, |c_j| \quad (7)$$

Note that the condition (7) requires in particular that the sources have different tri-spectrum.

3.2 FO Pre-whitening step

The first step of the FOBIUM method is to orthonormalize, in the Q_x matrix (6), the columns of A_Q, which can be considered as *virtual steering vectors* of the sources for the considered array of sensors [5]. For this purpose, let us consider the eigen decomposition of the Hermitian matrix Q_x, whose rank is P under the previous assumptions, given by

$$Q_x = E_x\, \Lambda_x\, E_x^{\ \mathrm{H}} \quad (8)$$

where Λ_x is the $(P \mathrm{\ x\ } P)$ real-valued diagonal matrix of the P non zero eigen-values of Q_x and E_x is the $(N^2 \mathrm{x} P)$ matrix of the associated orthonormalized eigen-vectors. For a full rank A_Q matrix, it is possible to verify that assuming P sources with non zero kurtosis having the same sign ε ($\varepsilon = \pm 1$) is equivalent to assume that the diagonal elements of Λ_x are not zero and have also the same sign corresponding to ε. In this context, considering the $(P \mathrm{\ x\ } N^2)$ whitening matrix T defined by

$$T \triangleq (\Lambda_x)^{-1/2}\, E_x^{\ \mathrm{H}} \quad (9)$$

where $(\Lambda_x)^{-1/2}$ is the inverse of a square root of Λ_x, we obtain, from (6) and (8)

$$\varepsilon\, T\, Q_x\, T^{\mathrm{H}} = T A_Q\, (\varepsilon\, C_m)\, A_Q^{\ \mathrm{H}}\, T^{\mathrm{H}} = \mathrm{I}_P \quad (10)$$

where I_P is the $(P \mathrm{\ x\ } P)$ identity matrix and where $\varepsilon C_m = \mathrm{diag}[|c_1|,\ldots, |c_P|]$. This last expression shows that the $(P \mathrm{x} P)$ matrix $T A_Q\, (\varepsilon C_m)^{1/2}$ is an unitary matrix U and we obtain

$$T A_Q = U\, (\varepsilon\, C_m)^{-1/2} \quad (11)$$

3.3 FO Blind identification step

We deduce from expressions (4b) and (11) that

$$T\, Q_x(\tau_1,\tau_2,\tau_3) T^{\mathrm{H}} = U\, (\varepsilon C_m)^{-1/2} C_m(\tau_1,\tau_2,\tau_3)\, (\varepsilon C_m)^{-1/2}\, U^{\mathrm{H}} \quad (12)$$

which shows that the unitary matrix U diagonalizes the matrices $T\ Q_x(\tau_1,\tau_2,\tau_3)\ T^{\mathrm{H}}$ whatever the set of delays (τ_1,τ_2,τ_3) and the associated eigen-values correspond to the diagonal terms of the diagonal matrix $(\varepsilon C_m)^{-1/2}\, C_m(\tau_1,\tau_2,\tau_3)\, (\varepsilon C_m)^{-1/2}$.

For a given set (τ_1,τ_2,τ_3), U is unique to within a permutation and an unitary diagonal matrix if and only if the eigen-values of the matrix $(\varepsilon C_m)^{-1/2}\, C_m(\tau_1,\tau_2,\tau_3)\, (\varepsilon C_m)^{-1/2}$ are all different. If it is not the case, we have to consider several sets $(\tau_1^{\,k},\tau_2^{\,k},\tau_3^{\,k})$, $1 \leq k \leq K$, such that for each couple of sources (i, j), it exists at least a set $(\tau_1^{\,k},\tau_2^{\,k},\tau_3^{\,k})$ such that the condition (7) is verified. In these conditions, the unitary matrix U becomes, to within a permutation and an unitary diagonal matrix, the only one which jointly diagonalizes the K matrices $T\ Q_x(\tau_1^{\,k},\tau_2^{\,k},\tau_3^{\,k})\ T^{\mathrm{H}}$. In other words, the unitary matrix, U_{sol}, solution to the previous problem of joint diagonalization can be written as

$$U_{sol} = U \Lambda \Pi \qquad (13)$$

where Λ and Π are unitary diagonal and permutation matrices respectively. Using (11) and (13), the matrix A_Q can be deduced from U_{sol} and T, to within unitary diagonal and permutation matrices, by

$$T^{\#} U_{sol} \triangleq E_x \Lambda_x^{1/2} \quad U_{sol} = A_Q (\varepsilon C_m)^{1/2} \Lambda \Pi \qquad (14)$$

where $T^{\#}$ corresponds to the pseudo-inverse of T. Each column, $\boldsymbol{b}_l$ $(1 \le l \le P)$, of $T^{\#} U_{sol}$ corresponds to one of the vectors $\mu_q |c_q|^{1/2} (\boldsymbol{a}_q \otimes \boldsymbol{a}_q^*)$, $1 \le q \le P$, where μ_q is a complex scalar such that $|\mu_q| = 1$. Thus, mapping the components of each column $\boldsymbol{b}_l$ of $T^{\#} U_{sol}$ into a $(N \times N)$ matrix B_l such that $B_l[i, j] = \boldsymbol{b}_l((i-1)N + j)$ $(1 \le i, j \le N)$ consists to built the matrices $\mu_q |c_q|^{1/2} \boldsymbol{a}_q \boldsymbol{a}_q^{H}$ $(1 \le q \le P)$. In this context, the source steering vector $\boldsymbol{a}_q$ corresponds to the eigen-vector of B_l associated to the strongest eigen-value.

3.4 Implementation of the FOBIUM method

The different steps of the FOBIUM method are summarized hereafter when L snapshots of the observations, $\boldsymbol{x}(l)$ $(1 \le l \le L)$, are available.

Step1: Estimation, $\hat{Q}_x$, of the Q_x matrix from the L snapshots $\boldsymbol{x}(l)$ using the empirical estimator of the FO cumulants [9]. Note that the FOBIUM method can also be applied for zero-mean cyclo-stationary sources provided that the previous empirical estimator is replaced by the unbiased FO statistics estimator proposed in [9].

Step2: Eigen Value Decomposition (EVD) of the matrix $\hat{Q}_x$, estimation of the number of sources P and restriction of this EVD to the P principal components : $\hat{Q}_x \approx \hat{E}_x \hat{\Lambda}_x \hat{E}_x^{H}$, where $\hat{\Lambda}_x$ is the diagonal matrix of the P eigen-values with the strongest modulus and $\hat{E}_x$ is the matrix of the associated eigen-vectors.

Step3: Computation of the pre-whitening matrix : $\hat{T} \triangleq (\hat{\Lambda}_x)^{-1/2} \hat{E}_x^{H}$.

Step4: Selection of K sets of delays $(\tau_1^k, \tau_2^k, \tau_3^k)$ where $|\tau_1^k| + |\tau_2^k| + |\tau_3^k| \neq 0$.

Step5: Estimation, $\hat{Q}_x(\tau_1^k, \tau_2^k, \tau_3^k)$, of $Q_x(\tau_1^k, \tau_2^k, \tau_3^k)$, for the K delay sets, using the empirical estimator of the FO statistics [9] (or, for zero-mean cyclo-stationary sources, the unbiased estimators similar to that presented in [9]).

Step6: Computation of the matrices $\hat{T} \hat{Q}_x(\tau_1^k, \tau_2^k, \tau_3^k) \hat{T}^{H}$ and estimation, $\hat{U}_{sol}$, of the unitary matrix U_{sol} from the joint diagonalization of the K matrices $\hat{T} \hat{Q}_x(\tau_1^k, \tau_2^k, \tau_3^k) \hat{T}^{H}$.

Step7: Computation of $\hat{T}^{\#} \hat{U}_{sol}$ and mapping each column $\hat{\boldsymbol{b}}_l$ into a $(N \times N)$ matrix $\hat{B}_l$

Step8: Estimation, $\hat{\boldsymbol{a}}_q$ $(1 \le q \le P)$, of the P source steering vectors by EVD of the P matrices $\hat{B}_l$

4. IDENTIFIABILITY

Following the development of the previous sections, we deduce that the FOBIUM method is able to identify, from an array of N sensors, the steering vectors of P ($P \le N^2$) non Gaussian sources having different tri-spectrum and kurtosis with the same sign provided that the A_Q matrix has full rank P, i.e that the *virtual steering vectors* $\boldsymbol{a}_q \otimes \boldsymbol{a}_q^*$ $(1 \le q \le P)$ for the considered array of N sensors remain linearly independent. Besides, it has been shown in [5] that the vector $\boldsymbol{a}_q \otimes \boldsymbol{a}_q^*$ can also be considered as a *true steering vector* but for a *virtual array* of N_e different sensors, where N_e is directly related to the geometry of the true array of N sensors. This means in particular that $N^2 - N_e$ components of each vector $\boldsymbol{a}_q \otimes \boldsymbol{a}_q^*$ are redundant components which bring no information. As a consequence, $N^2 - N_e$ rows of the A_Q matrix bring no information and are linear combinations of the others, which means that the rank of A_Q cannot be greater than N_e and is equal to $\mathrm{Inf}(N_e, P)$ when the A matrix is full rank. In these conditions, the A_Q matrix is full rank if and only if $\mathrm{Inf}(N_e, P) = P$, i.e if and only if $P \le N_e$. Thus the FOBIUM method is able to process N_e sources, where N_e is the number of sensors of the virtual array associated to the chosen array of N sensors. For an Uniform linear array $N_e = 2N + 1$ whereas for most of other arrays $N_e = N^2 - N + 1$ [5].

5. PERFORMANCE CRITERION

Most of the existing performance criterions used to evaluate the quality of the BI process [6-7] [10] are *global criterions* which evaluate a distance between the true mixing matrix A and its blind estimate $\hat{A}$. Although practice, a global performance criterion necessarily contains a part of arbitrary considerations in the manner of combining all the distances between the vectors $\boldsymbol{a}_q$ and $\hat{\boldsymbol{a}}_q$. Moreover, it is possible to find that an estimate $\hat{A}_1$ of A is better than an estimate $\hat{A}_2$, with respect to the global criterion, while some columns of $\hat{A}_2$ estimate the associated true steering vectors in a better way than $\hat{A}_1$.

For these reasons, we propose in this section a new performance criterion for the evaluation of the BI process. This new criterion is not global and allows both the evaluation quality of the BI of each source and the quantitative comparison of two methods for the BI of a given source. It corresponds, for the BI problem, to the one proposed in [4] for the extraction problem. It is defined by the P-uplet

$$\mathrm{D}(A, \hat{A}) \triangleq (\alpha_1, \alpha_2, \ldots, \alpha_P) \qquad (15)$$

where

$$\alpha_p \triangleq \min_{1 \le i \le P} [\mathrm{d}(\boldsymbol{a}_p, \hat{\boldsymbol{a}}_i)] \qquad (16)$$

and where d($\boldsymbol{u}$,$\boldsymbol{v}$) is the pseudo-distance between the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, defined by

$$\mathrm{d}(\boldsymbol{u},\boldsymbol{v}) \triangleq 1 - \frac{\left|\boldsymbol{u}^{\mathrm{H}}\boldsymbol{v}\right|^{2}}{\left(\boldsymbol{u}^{\mathrm{H}}\boldsymbol{u}\right)\left(\boldsymbol{v}^{\mathrm{H}}\boldsymbol{v}\right)} \tag{17}$$

6. SIMULATIONS

To illustrate the previous results, we assume that P=6 statistically independent non filtered QPSK sources are received by a circular array of N=3 sensors of radius r such that r/λ=0.55 (λ: wavelength). The 6 sources, assumed synchronized, have the same input SNR (Signal to Noise Ratio) of 20 dB with a symbol period $T = 4T_e$, where T_e is the sample period.

The direction of arrival of the sources are such that θ_1=2.16°, θ_2=25.2°, θ_3=50°, θ_4=272.16°, θ_5=315.36°, θ_6=336.96° and the associated carrier frequencies verify $\Delta f_1\, T_e$=0, $\Delta f_2\, T_e$=1/2, $\Delta f_3\, T_e$=1/3, $\Delta f_4\, T_e$=1/5, $\Delta f_5\, T_e$=1/7 and $\Delta f_6\, T_e$=1/11. We apply the JADE [3] , SOBI [1] and FOBIUM methods, and the performance α_q for q=1…6 is computed and averaged over 1000 realizations. For the FOBIUM method we choose K=4 sets of delays $(\tau_1^k,\tau_2^k,\tau_3^k)$ where τ_1^k=kT_e and τ_2^k=τ_3^k=0.

Under the previous assumptions, the figure 1 shows the variations of α_2 (source 2 performance) at the output of the JADE, SOBI and FOBIUM separators as a function of the number of snapshots L. We verify the difficulty of the JADE and SOBI methods to well identify the steering vector of the source 2 in an underdetermined context and the very good performance of the FOBIUM method in the same context. Note the complete convergence of the FOBIUM method as soon as L is in the area of 2000.

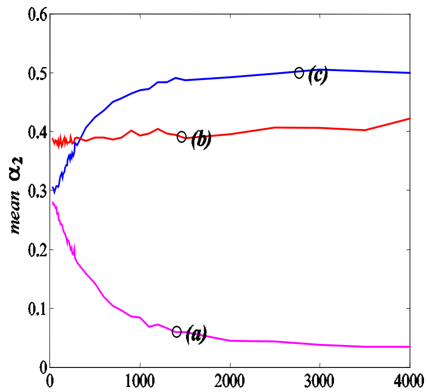

Fig.1 - *α_2 as a function of L, (a) (FOBIUM), (b) (JADE), (c) (SOBI)*

The Figure 2 shows, in the same context, the variations of all the α_p ($1 \leq p \leq 6$) at the output of the FOBIUM method as a function of L. Note the decreasing values toward zero of all the previous coefficients as L increases.

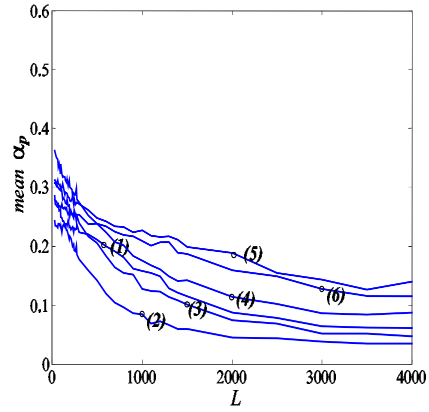

Fig.2 - *α_p of FOBIUM as a function of L, (p) performance criterion of the p^{th} source*

7. REFERENCES

[1] A. Belouchrani, K. Abed – Meraim, J.F. Cardoso, E. Moulines, "A blind source separation technique using second-order statistics", *IEEE Trans. Sig. Proc.*, Vol.45, N°2, pp. 434-444, Feb. 1997.

[2] JF. Cardoso, "Super-symmetric decomposition of the fourth order cumulant tensor", *ICASSP 1991*.

[3] J.F. Cardoso, A. Souloumiac, "Blind beamforming for non-gaussian signals", *IEE Proceedings-F*, Vol.140, N°6, pp. 362-370, Dec. 1993.

[4] P. Chevalier, "Optimal separation of independent narrow-band sources : concept and performance", *Sig. Proc.*, Elsevier, Vol.73, pp 27-47, 1999.

[5] P. Chevalier, A.Ferréol , "On the virtual array concept for the fourth-order direction finding problem", *IEEE trans on signal processing*, Vol.47, N°9, pp. 2592-2595, Sept. 1999.

[6] P. Comon, "Independent component analysis – a new concept ?", *Sig. Proc.*, Elsevier, Vol.36, N°3, Ap. 1994.

[7] P. Comon "Blind channel identification and extraction of more sources than sensors", *SPIE Conf Adv Proc VIII , San Diego*, July 1998.

[8] L. De Lathauwer, B. De Moor , J.Vandewalle , "ICA techniques for more sources than sensors", Proc *IEEE Processing Workshop on Higher Order Statistics*, Caesarea, Israel , June 1999.

[9] A. Ferréol, P. Chevalier, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources", *IEEE Trans. Sig. Proc.*, Vol.48, N°6, pp. 1712-1725, June 2000.

[10] A. Taleb "An algorithm for the blind identification of N independent signal with 2 sensors", *16th symposium on signal processing and its applications (ISSPA 2001)*, Kuala-Lumpur, Aug. 2001.

Appendix G

The SOBEFOCYS approach

SECOND ORDER BLIND SEPARATION OF FIRST AND SECOND ORDER CYCLOSTATIONARY SOURCES – APPLICATION TO AM, FSK, CPFSK AND DETERMINISTIC SOURCES

Anne Ferréol, Pascal Chevalier and Laurent Albera

Thalès-Communications, TRS/TSI/LTC,
66 rue du Fossé Blanc 92231 Gennevilliers, France
For Correspondance :
Pascal Chevalier
Tél : (33)-01-46-13-26-98, Fax : (33)-01-46-13-25-55
E-Mail : pascal.chevalier@fr.thalesgroup.com

Final version - May 2003

EDICS SP 1-IDSS

ABSTRACT

Most of the Second order (SO) and Higher order (HO) blind source separation methods developed this last decade aim at blindly separating statistically independent sources, assumed zero-mean, stationary and ergodic. Nevertheless, in many situations of practical interest, such as in radiocommunications contexts, the sources are non stationary and very often cyclostationary (digital modulations). The behavior of the current SO and fourth-order (FO) cumulant-based blind source separation methods in the presence of cyclostationary sources has been analysed, recently, in a previous paper [19], assuming zero-mean sources. However some cyclostationary sources used in practical situations are not zero-mean but have a first order cyclostationarity property, which is in particular the case for some AM signals and for some non linearly modulated digital sources such as FSK or some CPFSK sources. For such sources, the results presented in [19] do no longer hold and the purpose of this paper is to analyse the behavior and to propose adaptations of the current SO blind source separation methods for sources which are both first order and SO cyclostationary and cyclo-ergodic.

Keywords : *blind, second order, source separation, SOBI, first order cyclostationary, AM, FSK and CPFSK sources*

List of Figures

I. INTRODUCTION

For more than a decade, blind source separation methods exploiting either the SO [3] or the HO [16] or both the SO and HO [5] [13] statistics of the data, have been strongly developed, as depicted in the overview presented in [14]. These methods aim at blindly separating several statistically independent sources, assumed *zero-mean*, *stationary* and *ergodic*. Nevertheless, in many applications such as in the radiocommunications context, the sources are non stationary and very often *cyclostationary* (digital modulations). In these conditions, it becomes important to analyse the behavior of these current SO and HO blind methods, developed for zero-mean stationary sources, in the presence of cyclostationary sources whose cyclostationarity property appears explicitely at the processing level as soon as the sources are oversampled. This is generally the case for numerous applications such as, for example, the passive listening context where the different sources may have very different baud rate or bandwidth [17].

The behavior of the current SO and FO cumulant-based blind source separation methods in the presence of cyclostationary sources has been analysed in a recent paper [19], assuming *zero-mean* sources. Under this last assumption, valid in particular for linearly modulated digital sources, it has been shown in particular that under weak conditions of cyclo-ergodicity [4], the current SO blind methods are not affected by the cyclostationarity of the sources. On the contrary, the current FO cumulant-based blind methods, such as the JADE method [5], have been shown to be strongly affected, in some cases, by the cyclostationarity property of the sources and an adaptation of these FO methods, taking into account the SO cyclic frequencies of the sources, has been proposed. A FO alternative approach aiming at blindly separating statistically independent zero-mean cyclostationary sources without any knowledge or estimation of the cyclic frequencies of the sources has been proposed recently in [24]. Finally, other approaches of blind spatial filtering or blind source separation of zero-mean cyclostationary sources, aiming, in this case, at recovering the sources signals directly from the cyclic statistics of the observations, have also been proposed in the literature at both the SO [1-2] [25] and the HO [6] [18].

However, the cyclostationary sources used in practical applications are not necessarily zero-mean but may be *first order cyclostationary*, which is in particular the case for some Amplitude Modulated (AM) sources [21] and for some non linearly modulated digital sources such as Frequency Shift Keying (FSK) sources [31] or some Continuous Phase Frequency Shift Keying (CPFSK) sources, which belong to the more general family of the so-called Continuous Phase Modulations (CPM) sources [23] [28] [31-32]. For such sources, the analysis presented in [19] does no longer apply and for this reason, the purpose of this paper is to analyse the behavior and to propose adaptations of the current SO blind source separation methods in the presence of statistically independent sources which are both first order and SO cyclostationary. The behavior analysis of the current HO blind methods in the same context is partially presented in [20].

The current SO blind source separation problem for zero-mean stationary independent narrow-band (NB) sources together with the SOBI (Second Order Blind Identification) algorithm [3] and the empirical estimator of the SO statistics of the data are recalled in section II. Then, the problem of SO blind separation of first and SO cyclostationary sources together with examples of such sources (some AM, FSK and some CPFSK sources) are presented in section III where it is pointed out in particular the limitations of the empirical estimator of the SO statistics for non zero mean sources. The situations for which these limitations may have bad consequences on the current SO Blind Source Separation (BSS) methods behavior, jointly with the behavior description of the latter are presented in section IV where it is shown in particular that the performance of the SOBI method may be strongly affected by the first order cyclostationary properties of the sources. To overcome this problem, an adaptation of the current SO blind methods taking into account the possible first order cyclostationarity of the sources is proposed in section V. Unfortunately, this adaptation does not allow the processing of deterministic sources and to overcome this drawback, an extension of this adaptation, called SOBEFOCYS (Second Order Blind Extraction of First Order CYclostationary Sources), is described in section VI. Most of the results presented in the paper are finally illustrated in section VII by computer simulations. Note that the results presented in the paper have already been partially presented in [10] and [11].

II. SO BLIND SOURCE SEPARATION FOR ZERO MEAN STATIONARY SOURCES

A. Problem formulation

In the classical SO blind source separation problem, a noisy mixture of P zero-mean, stationary and Narrow-band (NB) independent sources is assumed to be received by an array of N sensors. Under this assumption, the vector, $\boldsymbol{x}(t)$, of the complex envelopes of the signals present at time t at the output of the sensors can be written as

$$\boldsymbol{x}(t) \;=\; \sum_{p=1}^{P} m_p(t)\; \boldsymbol{a}_p \;+\; \boldsymbol{b}(t) \;\triangleq\; A\,\boldsymbol{m}(t) \;+\; \boldsymbol{b}(t) \tag{1}$$

where $\boldsymbol{b}(t)$ is the noise vector, assumed zero-mean, stationary, ergodic, circular and spatially white, $m_p(t)$ and $\boldsymbol{a}_p$ correspond to the complex envelope and the steering vector of the source p, $\boldsymbol{m}(t)$ is the vector whose components are the signals $m_p(t)$ and A is the (N x P) matrix which columns are the vectors $\boldsymbol{a}_p$.

Under these assumptions, the classical SO blind source separation problem consists to find, from the SO statistics of the observations, the (N x P) *Linear* and *Time Invariant* source separator W, whose (P x 1) output vector,

$$\boldsymbol{y}(t) \triangleq W^{\dagger}\boldsymbol{x}(t) \tag{2}$$

corresponds, to within a diagonal matrix Λ and a permutation matrix Π, to the best estimate, $\hat{\boldsymbol{m}}(t)$, of the vector $\boldsymbol{m}(t)$. Note that the symbol † means transpose and complex conjugate. The separator W is defined to within a diagonal and a permutation matrix since neither the value of each output power of the separator nor the order in which the outputs are arranged change the estimation quality of the sources.

B. SO Statistics of the data

Under the previous assumptions, the SO statistics of the data are characterized by the correlation matrices $R_x(\tau)$, which also correspond to SO cumulant matrices, defined by

$$R_x(\tau) \triangleq \mathrm{E}[\boldsymbol{x}(t)\,\boldsymbol{x}(t-\tau)^\dagger] = A\,R_m(\tau)\,A^\dagger + \eta_2(\tau)\,\mathrm{I} \triangleq R_s(\tau) + \eta_2(\tau)\,\mathrm{I} \tag{3}$$

where $\eta_2(\tau)$ is the SO correlation function of the noise on each sensor, I is the identity matrix, $R_m(\tau) \triangleq \mathrm{E}[\boldsymbol{m}(t)\boldsymbol{m}(t-\tau)^\dagger]$, diagonal under the previous hypotheses, is the correlation matrix of the vector $\boldsymbol{m}(t)$ and $R_s(\tau) \triangleq A\,R_m(\tau)\,A^\dagger$ is the correlation matrix of the mixed sources.

C. Philosophy of the SO BSS methods (SOBI)

Let us now briefly recall the philosophy of the SOBI [3] method, which can be considered currently, for zero-mean stationary sources, as the most powerful SO blind source separation method but which requires that the sources have different spectral densities. This separator aims at separating the received sources from the blind identification of their steering vectors. These identified steering vectors may then be used to build and to apply to the data, for each source, a well suited spatial filter such as the Spatial Matched Filter or the Optimal Interference Canceller [8-9]. This blind identification requires the prewhitening of the data, by the pseudo-inverse, noted F, of a square root of the matrix $R_s(0)$, noted R_s in the following, computed from the $R_x \triangleq R_x(0)$ matrix and the knowledge of the noise correlation matrix. Usually, the matrix F is chosen to be equal to the (P x N) matrix $F = \Lambda_s^{-1/2} U_s^\dagger$, where the ($N$ x P) matrix U_s and the (P x P) diagonal matrix Λ_s correspond to the matrices of the orthonormalized eigenvectors and associated non zero eigenvalues of R_s respectively. This prewhitening operation aims at orthonormalizing the sources steering vectors so as to search for the latter through a unitary matrix U, simpler to handle. If we note $z(t) \triangleq F\,\boldsymbol{x}(t)$ the whitened observation vector, the matrix U is chosen so as to optimize a SO criterion, function of the elements of the correlation matrices, $R_z(\tau)$, of the vector $z(t)$ for several non zero values, τ_q, of τ. The matrix $R_z(\tau)$ can be easily computed from (3) and is given by

$$R_z(\tau) \triangleq \mathrm{E}[z(t)\,z(t-\tau)^\dagger] = A'\,R_{m'}(\tau)\,A'^\dagger + \eta_2(\tau)\,F\,F^\dagger \triangleq R_{s'}(\tau) + \eta_2(\tau)\,F\,F^\dagger \tag{4}$$

where A' is the (P x P) unitary matrix of the whitened sources steering vectors $\boldsymbol{a}_i' \triangleq \pi_i^{1/2} F\, \boldsymbol{a}_i$ $(1 \leq i \leq P)$, $\pi_i \triangleq \mathrm{E}[|m_i(t)|^2]$ is the input power of the source i, $R_{m'}(\tau) \triangleq \mathrm{E}[\boldsymbol{m}'(t)\, \boldsymbol{m}'(t-\tau)^\dagger]$ corresponds to the correlation matrix of $\boldsymbol{m}'(t)$, the normalized vector $\boldsymbol{m}(t)$ such that each component has a unit power and $R_{s'}(\tau) \triangleq A'\, R_{m'}(\tau)\, A'^\dagger$ is the correlation matrix of the whitened mixed sources.

Assuming that the sources have not the same spectral density and, for simplicity, that the coefficients $r_i'(\tau) \triangleq \mathrm{E}[m_i'(t)\, m_i'(t-\tau)^*]$ $(1 \leq i \leq P)$ are not zero for the considered value of τ, where * means complex conjugate and $m_i'(t)$ is the normalized complex envelope $m_i(t)$, the SOBI method [3] is based on the fact that the P orthonormalized vectors $\boldsymbol{a}_i'$ $(1 \leq i \leq P)$ are eigenvectors of the $R_{s'}(\tau)$ matrix associated to its P associated non zero eigenvalues $r_i'(\tau)$, which also correspond to the P eigenvalues of $R_{s'}(\tau)$ having the greatest absolute value. Then, an arbitrary eigenvector, $\boldsymbol{v}$, of $R_{s'}(\tau)$ associated to a non zero eigenvalue is necessarily a linear combination of the P vectors $\boldsymbol{a}_i'$. In these conditions, it is easy to verify [3] that the unitary matrix A' is, to within a permutation and a unitary diagonal matrix, the only one which jointly diagonalizes the set of K matrices $R_{s'}(\tau_q)$ $(1 \leq q \leq Q)$ provided that, for each couple (i, j) of sources, there is at least a τ_q such that $r_i'(\tau_q) \neq r_j'(\tau_q)$. In other words, the unitary matrix A' maximizes, with respect to the unitary matrix variable $U \triangleq (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$, the following joint diagonalization criterion [3]

$$C(U) = \sum_{q=1}^{Q} \sum_{l=1}^{P} |\boldsymbol{u}_l^\dagger R_{s'}(\tau_q)\, \boldsymbol{u}_l|^2 \tag{5}$$

Nevertheless, as the matrix $R_{s'}(\tau_q)$ is not observable, it must theoretically be estimated from the observable matrix $R_z(\tau_q)$. Under a temporally white noise assumption, which is done in [3], the quantity $\eta_2(\tau)$ is zero for $\tau \neq 0$ and the matrix $R_{s'}(\tau_q)$ can be replaced by the matrix $R_z(\tau_q)$ in (5). However, in practical situations, the reception band is finite and the noise can only be assumed temporally white within the reception band. In these conditions, if the reception band is sufficiently high with respect to the bandwidth of the sources, it is possible to find τ_q such that $R_{s'}(\tau_q) \neq 0$ and $\eta_2(\tau_q) \approx 0$ which still allows the use of $R_z(\tau_q)$ instead of $R_{s'}(\tau_q)$ in (5), which is done in the following.

D. Implementation of the SO BSS

In situations of practical interests, the SO statistics of the data are not known a priori and have to be estimated from the data, by temporal averaging operations, using the SO ergodicity property of the data. Noting T_e the sample period and $\boldsymbol{x}(k)$ the k-th sample of the observation vector $\boldsymbol{x}(t)$, the empirical estimator $\hat{R}_x(qT_e)(K)$ currently used to estimate the matrix $R_x(\tau)$, for $\tau = qT_e$, from K independent data snapshots, is defined by

$$\hat{R}_x(qT_e)(K) \triangleq \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{x}(k)\,\boldsymbol{x}(k-q)^\dagger \tag{6}$$

It is well known that for a stationary and SO ergodic vector $\boldsymbol{x}(t)$, the empirical estimator $\hat{R}_x(qT_e)(K)$ generates, as K becomes infinite, an unbiased and consistent estimate of $R_x(qT_e)$.

III. SO BLIND SOURCE SEPARATION FOR FIRST AND SECOND ORDER CYCLOSTATIONARY SOURCES

A. Problem formulation

In many applications such as in radiocommunications or in passive listening contexts, the received sources are very often cyclostationary (digital modulations) with a potential carrier residu (passive listening). In these conditions, the observation model (1) currently used in stationary contexts becomes too restrictive and we must adopt, for the complex observation vector $\boldsymbol{x}(t)$, the following model [19]

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} m_p(t)\, e^{j(2\pi\Delta f_p t + \phi_p)}\, \boldsymbol{a}_p + \boldsymbol{b}(t) \triangleq A\, \boldsymbol{m}_c(t) + \boldsymbol{b}(t) \tag{7}$$

where a noisy mixtures of P first and SO cyclostationary, cyclo-ergodic and NB independent sources is assumed to be received by the array of N sensors. In (7), the vector $\boldsymbol{b}(t)$ is the noise vector, assumed zero-mean, stationary, circular and spatially white, $m_p(t)$, Δf_p, ϕ_p and $\boldsymbol{a}_p$

correspond to the complex envelope, NB, first and SO cyclostationary, the carrier residu, the phase and the steering vector of the source p respectively, $\boldsymbol{m}_c(t)$ is the vector whose components are the signals $m_{pc}(t) \triangleq m_p(t) \exp[j(2\pi\Delta f_p t + \phi_p)]$ and A is the (N x P) matrix whose columns are the vectors $\boldsymbol{a}_p$. To simplify the developments, we limit the analysis to instantaneous mixtures of sources, typical of some applications such as the spatial telecommunications or some high data rate Line of Sight (LOS) contexts.

Under the previous assumptions, although in the presence of cyclostationary sources it may be useful to use *Polyperiodic* (PP) [12] and, for non circular sources [29], *Widely Linear* [7] [30] structures of separation, one may still prefer to try to recover the sources through a *Linear* and *Time Invariant* structure of separation, easier to handle. In these conditions, the problem is to find, from the SO statistics of the data, the (N x P) *Linear* and *Time Invariant* source separator W, whose output vector (2) aims at corresponding, to within a diagonal matrix Λ and a permutation matrix Π, to the best estimate, $\hat{\boldsymbol{m}}_c(t)$, of the vector $\boldsymbol{m}_c(t)$.

B. First and SO Statistics of the data

B1. *First order statistics*

In the presence of first order cyclostationary sources, the first order statistic of the vector $\boldsymbol{x}(t)$, given by (7), can be written as

$$\boldsymbol{e}_x(t) \triangleq \mathrm{E}[\boldsymbol{x}(t)] = \sum_{p=1}^{P} e_p(t)\, e^{j(2\pi\Delta f_p t + \phi_p)}\, \boldsymbol{a}_p \triangleq \sum_{p=1}^{P} e_{pc}(t)\ \boldsymbol{a}_p \triangleq A\ \boldsymbol{e}_{mc}(t) \qquad (8)$$

where $e_p(t)$, $e_{pc}(t)$ and $\boldsymbol{e}_{mc}(t)$ are the expected values of $m_p(t)$, $m_{pc}(t) = m_p(t)\, e^{j(2\pi\Delta f_p t + \phi_p)}$ and $\boldsymbol{m}_c(t)$ respectively. The vectors $e_p(t)$ and $e_{pc}(t)$ have a Fourier serial expansion and we obtain

$$e_p(t) \triangleq \mathrm{E}[m_p(t)] = \sum_{\gamma_p \in \Gamma_p} e_p^{\gamma_p}\, e^{j2\pi\gamma_p t} \qquad (9)$$

$$e_{pc}(t) \triangleq \mathrm{E}[m_{pc}(t)] = \sum_{\gamma_{pc} \in \Gamma_{pc}} e_{pc}^{\gamma_{pc}}\, e^{j2\pi\gamma_{pc} t} = \sum_{\gamma_p \in \Gamma_p} e_p^{\gamma_p}\, e^{j[2\pi(\Delta f_p + \gamma_p)t + \phi_p]} \qquad (10)$$

where $\Gamma_p = \{\gamma_p\}$ and $\Gamma_{pc} = \{\gamma_{pc} = \gamma_p + \Delta f_p\}$ are the set of cyclic frequencies γ_p and γ_{pc} of $e_p(t)$ and $e_{pc}(t)$ respectively and $e_p^{\gamma_p}$ and $e_{pc}^{\gamma_{pc}}$ are what we call the cyclic mean of $m_p(t)$ and $m_{pc}(t)$ respectively for the cyclic frequencies γ_p and γ_{pc} respectively, defined by

$$e_p^{\gamma_p} = < e_p(t)\, e^{-j2\pi\gamma_p t} >_c \tag{11}$$

$$e_{pc}^{\gamma_{pc}} \triangleq < e_{pc}(t)\, e^{-j2\pi\gamma_{pc} t} >_c = e_p^{\gamma_{pc}-\Delta f_p}\, e^{j\phi_p} \tag{12}$$

where the symbol $< f(t) >_c \triangleq \lim_{T\to\infty} (1/T) \int_{-T/2}^{T/2} f(t)\mathrm{d}t$ corresponds to the continuous-time temporal mean operation of $f(t)$ over an infinite observation interval. Note that for a zero mean source p, the quantities $e_p^{\gamma_p}$ and $e_{pc}^{\gamma_{pc}}$ are zero for all the cyclic frequencies γ_p and γ_{pc} respectively. Besides, for a stationary source p which is not zero mean, only the quantity $e_{pc}^{\gamma_{pc}}$ for $\gamma_{pc} = 0$ is not zero. A consequence of the previous results is that the vectors $\boldsymbol{e}_{mc}(t)$ and $\boldsymbol{e}_x(t)$ have also a Fourier serial expansion and using (10) into (8) we obtain

$$\boldsymbol{e}_{mc}(t) \triangleq \sum_{\gamma\in\Gamma} \boldsymbol{e}_{mc}^{\gamma}\, e^{j2\pi\gamma t} \tag{13}$$

$$\boldsymbol{e}_x(t) \triangleq \sum_{\gamma\in\Gamma} \boldsymbol{e}_x^{\gamma}\, e^{j2\pi\gamma t} = \sum_{\gamma\in\Gamma} A\, \boldsymbol{e}_{mc}^{\gamma}\, e^{j2\pi\gamma t} = \sum_{p=1}^{P} \sum_{\gamma_{pc}\in\Gamma_{pc}} e_{pc}^{\gamma_{pc}}\, e^{j2\pi\gamma_{pc} t}\, \boldsymbol{a}_p \tag{14}$$

where $\Gamma = \cup_{1\leq p\leq P}\{\Gamma_{pc}\}$ is the set of cyclic frequencies γ of $\boldsymbol{e}_{mc}(t)$ and $\boldsymbol{e}_x(t)$, $\boldsymbol{e}_{mc}^{\gamma}$ and $\boldsymbol{e}_x^{\gamma}$ are the cyclic mean of $\boldsymbol{m}_c(t)$ and $\boldsymbol{x}(t)$ respectively, defined by

$$\boldsymbol{e}_{mc}^{\gamma} = < \boldsymbol{e}_{mc}(t)\, e^{-j2\pi\gamma t} >_c \tag{15}$$

$$\boldsymbol{e}_x^{\gamma} = < \boldsymbol{e}_x(t)\, e^{-j2\pi\gamma t} >_c \tag{16}$$

Under these assumptions, the first and SO cyclostationary vector $\boldsymbol{x}(t)$ can be decomposed as the sum of a deterministic (quasi)-periodic part $\boldsymbol{e}_x(t)$ and a zero-mean (quasi)-cyclostationary random part $\Delta\boldsymbol{x}(t)$ such that

$$\Delta\boldsymbol{x}(t) \triangleq \boldsymbol{x}(t) - \boldsymbol{e}_x(t) = A\, \Delta\boldsymbol{m}_c(t) + \boldsymbol{b}(t) \tag{17}$$

where $\Delta\boldsymbol{m}_c(t) \triangleq \boldsymbol{m}_c(t) - \boldsymbol{e}_{mc}(t)$ is the zero-mean vector of the source signals, with components $\Delta m_{pc}(t) \triangleq \Delta m_p(t)\, e^{j(2\pi\Delta f_p t + \phi_p)}$, where $\Delta m_p(t) \triangleq m_p(t) - e_p(t)$.

B2. *Second order statistics*

Under the previous assumptions, using (7) and (17), the first correlation matrix, $R_x(t, \tau) \triangleq \mathrm{E}[\boldsymbol{x}(t)\, \boldsymbol{x}(t-\tau)^\dagger]$, of the data, which is time dependent, can be written as

$$R_x(t, \tau) = A\, R_{mc}(t, \tau)\, A^\dagger + \eta_2(\tau)\, \mathrm{I} = R_{\Delta x}(t, \tau) + \boldsymbol{e}_x(t)\, \boldsymbol{e}_x(t-\tau)^\dagger \tag{18}$$

where $\eta_2(\tau)$ is the SO correlation function of the noise on each sensor, I is the identity matrix, $R_{mc}(t, \tau) \triangleq \mathrm{E}[\boldsymbol{m}_c(t)\, \boldsymbol{m}_c(t-\tau)^\dagger]$ is the first correlation matrix of the vector $\boldsymbol{m}_c(t)$ and $R_{\Delta x}(t, \tau)$ is a cumulant matrix corresponding to the covariance matrix of $\boldsymbol{x}(t)$, defined by

$$R_{\Delta x}(t, \tau) \triangleq \mathrm{E}[\Delta\boldsymbol{x}(t)\, \Delta\boldsymbol{x}(t-\tau)^\dagger] = A\, R_{\Delta mc}(t, \tau)\, A^\dagger + \eta_2(\tau)\, \mathrm{I} \tag{19}$$

where $R_{\Delta mc}(t, \tau) \triangleq \mathrm{E}[\Delta\boldsymbol{m}_c(t)\, \Delta\boldsymbol{m}_c(t-\tau)^\dagger]$ is the covariance matrix of $\boldsymbol{m}_c(t)$. Using (8) and (19) into (18), we finally obtain

$$R_x(t, \tau) = A\, [R_{\Delta mc}(t, \tau) + \boldsymbol{e}_{mc}(t)\, \boldsymbol{e}_{mc}(t-\tau)^\dagger\,]\, A^\dagger + \eta_2(\tau)\, \mathrm{I} \tag{20}$$

The SO cyclostationary property of the sources implies that the matrices $R_{mc}(t, \tau)$ and $R_{\Delta mc}(t, \tau)$ and thus, the matrices $R_x(t, \tau)$ and $R_{\Delta x}(t, \tau)$, have Fourier serial expansions introducing the SO cyclic frequencies and statistics of $\boldsymbol{m}_c(t)$, $\Delta\boldsymbol{m}_c(t)$, $\boldsymbol{x}(t)$ and $\Delta\boldsymbol{x}(t)$ respectively. In particular, the first cyclic correlation matrix of $\boldsymbol{x}(t)$ for the zero cyclic frequency corresponds to the temporal mean of $R_x(t, \tau)$, which can be written, from (18), as

$$R_x(\tau) \triangleq \langle R_x(t, \tau) \rangle_c = A\, R_{mc}(\tau)\, A^\dagger + \eta_2(\tau)\, \mathrm{I} \triangleq R_s(\tau) + \eta_2(\tau)\, \mathrm{I} \tag{21}$$

where $R_s(\tau) \triangleq A\, R_{mc}(\tau)\, A^\dagger$ and where, from (20), $R_{mc}(\tau) \triangleq \langle R_{mc}(t, \tau) \rangle_c$ is given by

$$R_{mc}(\tau) \triangleq \langle R_{mc}(t, \tau) \rangle_c = R_{\Delta mc}(\tau) + \langle \boldsymbol{e}_{mc}(t)\, \boldsymbol{e}_{mc}(t-\tau)^\dagger \rangle_c = R_{\Delta mc}(\tau) + E_{mc}(\tau) \tag{22}$$

where $R_{\Delta mc}(\tau) \triangleq \langle R_{\Delta mc}(t, \tau) \rangle_c$ and $E_{mc}(\tau) \triangleq \langle \boldsymbol{e}_{mc}(t)\, \boldsymbol{e}_{mc}(t-\tau)^\dagger \rangle_c$.

For observations, $\boldsymbol{x}(nT_e)$, sampled at the sample period T_e, the matrices $R_x(t, \tau)$ and $R_x(\tau)$ are defined only for the time instant $t = nT_e$ and $\tau = kT_e$ multiple of this sample period. In these conditions, it is possible to show [19] [27] that, for band-limited sources, the matrice $R_x(kT_e)$, defined by (21) for $\tau = kT_e$, can be computed only from the sampled matrices $R_x(nT_e, kT_e)$ instead of $R_x(t, \tau)$ provided that the data are sufficiently oversampled. In other words, for sufficiently oversampled data, it is possible to replace, in (21), the continuous-time temporal mean operation $< f(t) >_c$ by a discrete one over an infinite number of samples $< f(nT_e) >_d \triangleq \lim_{K \to \infty} (1/K) \sum_{k=1}^{K} f(k)$.

C. Two examples of first order cyclostationary sources : The AM and FSK sources

C1 : The AM sources

The AM sources [21] are amplitude modulated analog sources used in particular in old radiocommunications systems. If the source p is an AM source, its complex envelope can be written as

$$m_p(t) \;=\; \beta_p \; (1 + \mu_p \, l_p(t)) \tag{23}$$

where β_p is a scalar, μ_p $(0 \leq \mu_p \leq 1)$ is the modulation indice and $l_p(t)$, such that $0 \leq |l_p(t)| \leq 1$, is, in the general case, a non zero mean non stationary random signal. Defining $el_p(t) \triangleq$ E$[l_p(t)]$, the statistical mean of $m_p(t)$ is then given by

$$e_p(t) \;=\; \beta_p \; (1 + \mu_p \, el_p(t)) \tag{24}$$

which is, in the general case, a non zero time dependent function. The signal $m_p(t)$ becomes first order cyclostationary for the particular case of first order cyclostationary signals $l_p(t)$. In this latter case, $e_p(t)$ has a Fourier serial expansion (9) and we can easily verify that the cyclic mean e_p^γ is given by

$$e_p^\gamma \;=\; \beta_p \; (\delta(\gamma) + \mu_p \, el_p^\gamma\,) \tag{25}$$

where $\delta(.)$ is the kronecker symbol and where $el_p^\gamma \triangleq \; < el_p(t) \; e^{-j2\pi\gamma t} >_c$. Thus, the first order cyclic frequencies of an AM source p correspond to the zero cyclic frequency and to the first order cyclic frequencies of $l_p(t)$.

C2 : The FSK sources

The FSK sources [31] are non linearly modulated digital sources used in particular for low data rate radiocommunications in the HF band. If the source p is a FSK source, its complex envelope can be written as

$$m_p(t) \;=\; \pi_p^{\,1/2} \sum_n \exp\{j[\theta_{pn} + 2\pi f_{dp}\, a_n^p\,(t - nT_p)]\}\;\; \mathrm{Rect}_p(t - nT_p) \tag{26}$$

where $\pi_p \triangleq\ < E[|m_p(t)|^2] >_c$ is the input power of the source p, T_p is the symbol duration, the a_n^p are the transmitted M_p-ary symbols, assumed i.i.d and taking their values in the alphabet $\pm 1, \pm 3, \ldots, \pm (M_p - 1)$, where M_p is generally a power of two, f_{dp} is the peak frequency deviation, $\mathrm{Rect}_p(t)$ is the rectangular pulse of amplitude 1 and of duration T_p and θ_{pn} is the phase of the symbol n. Note that for M_p-ary symbols, the associated FSK source p is qualified of M_p-FSK source p.

When the signal $m_p(t)$ is built from only one local oscillator which is hopping, from a frequency to another, at every symbol period T_p, the phases θ_{pn} may be considered as i.i.d random variables which are statistically independent of the symbols a_n^p and uniformly distributed between 0 and 2π. In this case it is easy to verify that $m_p(t)$ is zero-mean. However, when the signal $m_p(t)$ is built from M_p local oscillators, one oscillator per symbol value, among which one oscillator is switched at every symbol period T_p, the phase θ_{pn} of the symbol n corresponds to the phase of the switched oscillator for this symbol n and becomes a function $\theta_{pn}(a_n^p)$ of the symbol n value, taking its values in the alphabet $\{\theta_{-p(Mp-1)}, \ldots, \theta_{-p1}, \theta_{p1}, \theta_{p3}, \ldots, \theta_{p(Mp-1)}\}$. In this case, the statistical mean of $m_p(t)$ is given by

$$e_p(t) \;=\; \pi_p^{\,1/2}\, \frac{1}{M_p} \sum_n \sum_{m=0}^{(M_p-2)/2} \Big(\exp\{j[\theta_{p(2m+1)} + 2\pi f_{dp}(2m{+}1)(t - nT_p)]\} \;+$$

$$\exp\{j[\theta_{-p(2m+1)} - 2\pi f_{dp}(2m{+}1)(t - nT_p)]\} \Big)\; \mathrm{Rect}_p(t - nT_p) \tag{27}$$

which corresponds to a periodic function of t with a period T_p. In other words, the associated M_p-FSK source p is a first order cyclostationary source p with first order cyclic frequencies γ_p's multiple of $1/T_p$. In these conditions, $e_p(t)$ has a Fourier serial expansion (9) and the cyclic mean e_p^γ is shown in Appendix A to be given by

$$e_p^\gamma \;=\; \sum_i e_p^{\gamma pi}\; \delta(\gamma - i/T_p) \tag{28}$$

where $e_p^{\gamma pi}$, the cyclic mean for the cyclic frequency $\gamma_{pi} = i/T_p$, is given by

$$e_p^{\gamma_{pi}} = \pi_p^{1/2} \frac{1}{M_p} \sum_{m=0}^{(M_p-2)/2} \int_0^1 \left(\exp\{j[\theta_{p(2m+1)} + 2\pi\, w(f_{dp}T_p\,(2m+1) - i)]\} \right.$$

$$\left. + \exp\{j[\theta_{-p(2m+1)} - 2\pi\, w(f_{dp}T_p\,(2m+1) + i)]\} \right) \mathrm{d}w \qquad (29)$$

In particular, we deduce from (29) that a M_p-FSK source such that the product $f_{dp}T_p$ is an integer has exactly M_p equal power first order cyclic frequencies $\gamma_p = \pm\,(2k+1)\,f_{dp}$, $0 \leq k \leq (M_p - 2)/2$, such that

$$e_p^{\gamma} = \pi_p^{1/2} \frac{1}{M_p} \sum_{m=0}^{(M_p-2)/2} [\, \exp\{j\theta_{p(2m+1)}\}\, \delta(\gamma - (2m+1)f_{dp}) +$$

$$\exp\{j\theta_{-p(2m+1)}\}\, \delta(\gamma + (2m+1)f_{dp})\,] \qquad (30)$$

D. A third example of first order cyclostationary sources : Some CPFSK sources

D1 : The CPFSK sources as a particular case of CPM sources

The CPM sources [28] [31] are non linearly modulated digital sources used in many applications of practical interest such as the mobile cellular radiocommunications (GSM) or the spatial telecommunications. One characteristic of CPM sources is that their spectral efficiency is much better than that of non linear modulations without a Continuous Phase. Another property is that their complex envelope has a constant amplitude, which allows the use of cheap amplifiers working at a saturation level without any distorsion on the transmitted information. If the source p is a CPM source, its complex envelope can be written as

$$m_p(t) = \pi_p^{1/2} \exp\{j2\pi\, [\sum_n a_n^p\, h_n^p\, v_p(t - nT_p)]\} \qquad (31)$$

where $\pi_p \triangleq \langle E[|m_p(t)|^2] \rangle_c$ is the input power of the source p, T_p is the symbol duration, the a_n^p are the transmitted M_p-ary symbols, assumed i.i.d and taking their values in the alphabet $\pm 1, \pm 3, \ldots, \pm (M_p - 1)$, where M_p is generally a power of two, $\{h_n^p\}$ is a sequence of modulation indices and $v_p(t)$ is the waveform shape, represented as the integral of a pulse $g_p(t)$, non zero and bounded on the interval $[0, L_pT_p]$, where L_p is a non zero integer, and such that

$$v_p(t) \;=\; \int_{-\infty}^{t} g_p(u)\,\mathrm{d}u = \begin{Bmatrix} 1/2\ (t \geq L_pT_p) \\ \\ 0 \quad (t < 0) \end{Bmatrix} \tag{32}$$

When $h_n^p = h_p$ for all n the modulation is said to be mono-indice, otherwise the modulation is qualified of multi-indices. When $L_p = 1$, the CPM source p is called *full response CPM*, otherwise it is called *partial response CPM*. To each choice of the pulse function $g_p(t)$, it corresponds a family of CPM source p. The GMSK modulation, which is the modulation of the GSM standard and for which the pulse function $g_p(t)$ has only an approximated finite duration, belongs to one of these families.

The CPFSK source p is a particular case of the mono-indice full response CPM source p for which the pulse $g_p(t)$ is a rectangular pulse of amplitude $1/2T_p$ and of duration T_p. For such sources p, it is possible to show, after easy computations, that $m_p(t)$ can be written as (26), where $f_{dp} \triangleq h_p/2T_p$ is the peak frequency deviation and θ_{pn}, which represents the accumulation (memory) of all symbols up to $(n-1)T_p$, is defined by

$$\theta_{pn} \;\triangleq\; 2\pi f_{dp} T_p \sum_{k=-\infty}^{n-1} a_k^p \tag{33}$$

For M_p-ary symbols, the associated CPFSK source p is qualified of M_p-CPFSK source p. Note that a binary CPFSK source p ($M_p = 2$) with a modulation index $h_p = 1/2$ is called a MSK (Minimum Shift Keying) source p.

D2 : First order statistics of CPFSK sources

It is shown in Appendix B that, under the previous assumptions, the statistical mean of $m_p(t)$ is, for a M_p-CPFSK source p, given by

$$e_p(t) \;=\; \pi_p^{1/2}\ \frac{2}{M_p}\ K_p \sum_n (\rho_p)^n\ u_p(t - nT_p) \tag{34}$$

where the quantities K_p, ρ_p $(0 \leq \rho_p \leq 1)$ and $u_p(t)$ are defined by

$$K_p \;\triangleq\; \lim_{l \to \infty} (\rho_p)^l \tag{35}$$

$$\rho_p \triangleq \frac{2}{M_p} \sum_{m=0}^{(M_p-2)/2} \cos[(2m+1)\pi h_p] \tag{36}$$

$$u_p(t) \triangleq \sum_{m=0}^{(M_p-2)/2} \cos[(2m+1)(\pi h_p/T_p)\, t\,]\, \mathrm{Rect}_p(t) \tag{37}$$

From the previous expressions, two cases have to be considered depending on the value of h_p. These cases correspond to the cases where h_p is an integer or not respectively.

a) *h_p is not an integer*

If h_p is not an integer, it is obvious that $|\rho_p| < 1$, which implies the nullity of both K_p and $e_p(t)$. In other words, a M_p-CPFSK source p whose modulation indice is not an integer is a zero-mean source p.

b) *h_p is an integer*

If h_p is an integer, it is obvious that $\rho_p = 1$ if h_p is even and $\rho_p = -1$ if h_p is odd, which implies that $K_p = 1$ in the first case and that expression (35) has no limit ($K_p = \pm 1$),while $|K_p|$ = 1, in the second case. In this latter case, note that if we assume that the number of past symbols is finite, then $K_p = \pm 1$. A consequence of the previous results is that $e_p(t)$ reduces to

$$e_p(t) = \pi_p^{1/2} \frac{2}{M_p} K_p \sum_n (-1)^{nh_p} u_p(t - nT_p) \tag{38}$$

which corresponds to a periodic function of t with a period T_p if h_p is even and with a period $2T_p$ if h_p is odd. In other words, the associated M_p-CPFSK source p is a first order cyclostationary source p with first order cyclic frequencies γ_p's multiple of $1/T_p$ in the first case and multiple of $1/2T_p$ in the second case. In these conditions, $e_p(t)$ has a Fourier serial expansion (9) and the cyclic mean e_p^{γ} is shown in Appendix B to be given by

$$e_p^{\gamma} = \pi_p^{1/2} \frac{1}{M_p} K_p \sum_{m=0}^{(M_p-2)/2} [\delta(\gamma - (2m+1)f_{dp}) + \delta(\gamma + (2m+1)f_{dp})] \tag{39}$$

where $\delta(.)$ is the kronecker symbol. Thus, a M_p-CPFSK source p whose modulation indice is an integer has exactly M_p equal power first order cyclic frequencies γ_p such that

$$e_p^\gamma \;=\; \pi_p^{1/2}\,\frac{1}{M_p}\,K_p \quad \text{for} \;\; \gamma \in \Gamma_p = \{\gamma_p = \pm\,(2k+1)\,f_{dp},\, 0 \le k \le (M_p - 2)/2\} \qquad (40)$$

Note that in this case $< e_p(t) >_c = 0$ despite of the fact that $e_p(t) \neq 0$.

E. Problem addressed in this paper

For first and SO cyclostationary and band-limited vectors $\boldsymbol{x}(t)$ having a SO cyclo-ergodicity property [4] and for sufficiently oversampled data, the empirical estimator $\hat{R}_x(lT_e)(K)$, defined by (6), gives an asymptotically unbiased and consistent estimate of $R_x(lT_e)$, defined by (21) with $\tau = lT_e$, by definition of the cyclo-ergodicity property. In other words, in cyclostationary contexts, the SO BSS methods such as the SOBI method exploit, asymptotically or in the steady state, the information contained in several time averaged correlation matrices $R_x(\tau_q)$ $(1 \le q \le Q)$, defined by (21).

However, while these matrices correspond to cumulant matrices for zero-mean stationary sources and to time averaged cumulant matrices for zero-mean cyclostationary sources [19], it is no longer the case for first and SO cyclostationary sources for which, $\boldsymbol{e}_{mc}(t) \neq \boldsymbol{0}$, $R_{mc}(\tau) \neq R_{\Delta mc}(\tau)$ and $R_x(\tau) \neq R_{\Delta x}(\tau)$ as shown by (8), (18) and (22). As a consequence, while, for zero-mean statistically independent sources, $R_{mc}(\tau)$ and $R_{\Delta mc}(\tau)$, appearing in (22), coïncide and are diagonal, only the $R_{\Delta mc}(\tau)$ matrix keeps in all cases a diagonal structure for non zero mean sources by definition of the statistical independence of the sources, whereas the matrix $R_{mc}(\tau)$ may loose its diagonal character. In this latter case, if the element $[i, j]$, $R_{mc}(\tau)[i, j]$, of the matrix $R_{mc}(\tau)$, with $i \neq j$, is not zero, we will said that the first order cyclostationarity of the statistically independent sources i and j creates an *apparent SO correlation* of the sources in the $R_{mc}(\tau)$ and $R_x(\tau)$ matrices. This apparent SO correlation is directly related to the so-called impure SO cycle frequencies of the SO statistics of the sources discussed in [22].

In this context, in a first time, we must identify the conditions that two statistically independent first and SO cyclostationary sources have to verify to create an apparent SO correlation in the $R_x(\tau)$ matrices. Then, we must analyse the consequences of such an apparent

SO correlation between two sources on the output performances of the current SO BSS methods such as the SOBI one. These two questions are adressed in section IV.

IV. BEHAVIOR ANALYSIS OF CURRENT SO BSS METHODS FOR FIRST AND SECOND ORDER CYCLOSTATIONARY SOURCES

A. Structure analysis of the $E_{mc}(\tau)$ matrix

The matrix $R_{mc}(\tau)$ is not diagonal if and only if the matrix $E_{mc}(\tau)$ introduced in (22) is not diagonal. Thus, two sources i and j become apparently SO correlated in the $R_x(\tau)$ matrix if and only if the element $[i, j]$, $E_{mc}(\tau)[i, j]$, of the matrix $E_{mc}(\tau)$ is not zero, situation which is analysed in this section.

A1. *General case*

Using (13) and (22), the $E_{mc}(\tau)$ matrix can be written as

$$E_{mc}(\tau) \triangleq < e_{mc}(t)\, e_{mc}(t-\tau)^{\dagger} >_c \; = \sum_{\gamma\in\Gamma} \sum_{\omega\in\Gamma} e_{mc}^{\gamma}\, e_{mc}^{\omega\,\dagger}\, e^{j2\pi\tau\omega} < e^{j2\pi(\gamma-\omega)t} >_c \qquad (41)$$

and using the fact that $< e^{j2\pi\alpha t} >_c = \delta(\alpha)$, we obtain

$$E_{mc}(\tau) \; = \sum_{\gamma\in\Gamma} e_{mc}^{\gamma}\, e_{mc}^{\gamma\,\dagger}\, e^{j2\pi\tau\gamma} \qquad (42)$$

The element $[i, j]$ of the matrix $E_{mc}(\tau)$ is thus given by

$$E_{mc}(\tau)[i, j] \; = \sum_{\gamma_{ij}\in\Gamma_{ij}} e_{ic}^{\gamma_{ij}}\, e_{jc}^{\gamma_{ij}*}\, e^{j2\pi\tau\gamma_{ij}} \qquad (43)$$

where $\Gamma_{ij} \triangleq \Gamma_{ic} \cap \Gamma_{jc}$ is the set of cyclic frequencies γ_{ij} belonging to both Γ_{ic} and Γ_{jc}, defined in (10) and e_{ic}^{γ} and e_{jc}^{γ} are defined by (12) with γ instead of γ_{pc}. The expression (43) shows that $E_{mc}(\tau)[i, j]$ is generally not zero, i.e. the two sources i and j become *apparently SO correlated* in the matrice $R_x(\tau)$, if the condition (C1) is verified where (C1) is defined by

(C1) : The two sources i and j share at least one first order cyclic frequency, i.e $e_{ic}(t)$ and $e_{jc}(t)$ share at least one cyclic frequency.

A2. *Application to FSK sources*

In the particular case of a M_i-FSK source i and a M_j-FSK source j, built from M_i and M_j oscillators respectively, such that $f_{di}T_i$ and $f_{dj}T_j$ are integer and such that $\theta_{i(2m+1)} = \theta_{-i(2m+1)} = \theta_i$ $(0 \leq m \leq (M_i - 2)/2)$ and $\theta_{j(2m+1)} = \theta_{-j(2m+1)} = \theta_j$ $(0 \leq m \leq (M_j - 2)/2)$, the expression (43) becomes

$$E_{mc}(\tau)[i,j] = e^{j(\phi_i - \phi_j)} \pi_i^{1/2} \pi_j^{1/2} \frac{1}{M_i M_j} \exp\{j(\theta_i - \theta_j)\} \sum_{\gamma_{ij} \in \Gamma_{ij}} e^{j2\pi\tau\gamma_{ij}} \tag{44}$$

which is not zero if Γ_{ij} is not empty, which is the case if it exists at least one value of m $(0 \leq m \leq (M_i - 2)/2)$ and one value of n $(0 \leq n \leq (M_j - 2)/2)$ such that

$$\Delta f_i \pm (2m+1) f_{di} = \Delta f_j \pm (2n+1) f_{dj} \tag{45}$$

A3. *Application to CPFSK sources*

In the particular case of a M_i-CPFSK source i and a M_j-CPFSK source j, using the results of section III, a necessary condition to obtain $E_{mc}(\tau)[i,j] \neq 0$ is that the two sources have integer modulation indices h_i and h_j respectively. In these conditions, using (40) and (12), the expression (43) becomes

$$E_{mc}(\tau)[i,j] = e^{j(\phi_i - \phi_j)} \pi_i^{1/2} \pi_j^{1/2} \frac{1}{M_i M_j} K_i K_j \sum_{\gamma_{ij} \in \Gamma_{ij}} e^{j2\pi\tau\gamma_{ij}} \tag{46}$$

which is not zero if Γ_{ij} is not empty, which is the case if it exists at least one value of m $(0 \leq m \leq (M_i - 2)/2)$ and one value of n $(0 \leq n \leq (M_j - 2)/2)$ such that (45) is verifyed.

In other words for a M_i-CPFSK source i and a M_j-CPFSK source j, the condition (C1) becomes (C1') defined by

(C1') : a) The two sources i and j have an integer modulation indice

b) It exists at least one m $(0 \leq m \leq (M_i - 2)/2)$ and one n $(0 \leq n \leq (M_j - 2)/2)$ such that $\Delta f_i \pm (2m+1) f_{di} = \Delta f_j \pm (2n+1) f_{dj}$

We must now wonder how the presence of *apparently SO correlated* sources in the matrices $R_x(\tau)$ may modify the behavior of the current SO BSS method and the SOBI method in particular. These questions are addressed in sections B to D.

B. Prewhitening of the data

We evaluate, in this section, the consequences of *an apparent SO correlation of the sources* on the prewhitening operation of the observations.

B1. *Apparent SO correlation coefficient of two sources*

To characterize the degree of apparent SO correlation of two sources i and j in the matrix $R_x \triangleq R_x(0)$, we introduce the apparent SO correlation coefficient of these sources, ρ_{ij} $(0 \leq |\rho_{ij}| \leq 1)$, defined by

$$\rho_{ij} \triangleq \rho_{ij}(0) \triangleq R_{mc}(0)[i,j] \,/\, (R_{mc}(0)[i,i]\, R_{mc}(0)[j,j])^{1/2} \tag{47}$$

where $R_{mc}(0)[i, j] \triangleq < \mathrm{E}[m_{ic}(t)\ m_{jc}(t)^*] >_c$. In particular, from (46) and using the fact that $R_{mc}(0)[i, j] = E_{mc}(0)[i, j]$ for $i \neq j$, the apparent SO correlation coefficient of a M_i-CPFSK source i and a M_j-CPFSK source j, with integer modulation indices, is given by

$$\rho_{ij} \;=\; e^{j(\phi_i-\phi_j)}\, \frac{1}{M_i M_j}\, K_i\, K_j\ \ \mathrm{card}(\Gamma_{ij}) \quad , \qquad i \neq j \tag{48}$$

where $\mathrm{card}(\Gamma_{ij})$ is the number of elements of Γ_{ij}. This expression shows in particular that, for given values of M_i and M_j, $|\rho_{ij}|$ increases with the number of couple (m, n) verifying (45). For example, if $M_i = M_j = M$, $\Delta f_i = \Delta f_j$ and $f_{di} = f_{dj}$ we obtain $|\rho_{ij}| = 1/M$.

B2. *Eigenstructure of $R_x(0)$*

To simplify the developments, we limit, in the following, the analysis to the two first order cyclostationary sources case and we assume that the source matrix R_s, defined from (21) for $\tau = 0$, is not rank deficient, i.e. that the sources are not apparently SO coherent ($|\rho_{12}| \neq 1$). As the eigenstructure of R_x is directly deduced from that of R_s, (same eigenvectors and eigenvalues obtained by the addition of η_2 to the eigenvalues of R_s), we limit the analysis to the eigendecomposition of R_s. Under these assumptions, it is possible to show, after easy but tedious algebraic manipulations that the two non zero eigenvalues of R_s, λ_{s+} and λ_{s-}, are given by

$$\lambda_{s\pm} = (1/2)\,[\,B_4 \pm (B_4^2 - 4\,B_5)^{1/2}] \tag{49}$$

where B_4 and B_5 are scalar quantities defined by

$$B_4 \triangleq \pi_1\, \boldsymbol{a}_1^\dagger \boldsymbol{a}_1 + \pi_2\, \boldsymbol{a}_2^\dagger \boldsymbol{a}_2 + 2[\pi_1\,\pi_2\,(\boldsymbol{a}_1^\dagger \boldsymbol{a}_1)\,(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)]^{1/2}\,\mathrm{Re}(\rho_{12}\,\alpha_{21}) \tag{50}$$

$$B_5 \triangleq \pi_1\,\pi_2\,(\boldsymbol{a}_1^\dagger \boldsymbol{a}_1)\,(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)\,(1 - |\rho_{12}|^2)\,(1 - |\alpha_{12}|^2) \tag{51}$$

where $\pi_i \triangleq \langle \mathrm{E}[|m_{ic}(t)|^2] \rangle_c$ $(1 \le i \le 2)$ and $\alpha_{12} = \alpha_{21}^*$ $(0 \le |\alpha_{12}| \le 1)$ is the spatial correlation coefficient of the sources 1 and 2, defined by

$$\alpha_{12} \triangleq \boldsymbol{a}_1^\dagger \boldsymbol{a}_2 \,/\, [(\boldsymbol{a}_1^\dagger \boldsymbol{a}_1)(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)]^{1/2} \tag{52}$$

The associated orthonormalized eigenvectors $\boldsymbol{u}_+$ and $\boldsymbol{u}_-$ are defined by

$$\boldsymbol{u}_{s\pm} \triangleq \exp(j\varphi_\pm)\,(1 \,/\, \|\,B_6\,\boldsymbol{a}_1 - C_\pm\,\boldsymbol{a}_2\,\|)\;[B_6\,\boldsymbol{a}_1 - C_\pm\,\boldsymbol{a}_2\,] \tag{53}$$

where $\phi_\pm$ is an arbitrary phase value, B_6 and $C_\pm$ are scalar quantities defined by

$$B_6 \triangleq \pi_1\,[(\boldsymbol{a}_1^\dagger \boldsymbol{a}_1)(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)]^{1/2}\,\alpha_{12} + (\pi_1\,\pi_2)^{1/2}\,(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)\,\rho_{12} \tag{54}$$

$$C_\pm \triangleq \pi_1\,\boldsymbol{a}_1^\dagger \boldsymbol{a}_1 + [\pi_1\,\pi_2\,(\boldsymbol{a}_1^\dagger \boldsymbol{a}_1)\,(\boldsymbol{a}_2^\dagger \boldsymbol{a}_2)]^{1/2}\,\rho_{12}\,\alpha_{21} - \lambda_{s\pm} \tag{55}$$

B3. *Whitened observations*

From the previous expressions, it is possible to built the (2 x N) whitening matrix $F \triangleq \Lambda_s^{-1/2} U_s^\dagger$, where $\Lambda_s \triangleq \mathrm{Diag}(\lambda_{s+}, \lambda_{s-})$ and $U_s \triangleq [\boldsymbol{u}_{s+}, \boldsymbol{u}_{s-}]$. The whitened observation vector $\boldsymbol{z}(t) \triangleq F\,\boldsymbol{x}(t)$ is then given by

$$\boldsymbol{z}(t) \triangleq F\,\boldsymbol{x}(t) = \sum_{p=1}^{P} m_{pc}'(t)\;\boldsymbol{a}_p' + F\,\boldsymbol{b}(t) \;\triangleq\; A'\,\boldsymbol{m}_c'(t) + F\,\boldsymbol{b}(t) \tag{56}$$

where $\boldsymbol{m}_c'(t)$ is the (2 x 1) vector of the normalized complex envelopes $m_{pc}'(t)$ of $m_{pc}(t)$ $(0 \le p \le 2)$, such that $\langle \mathrm{E}[|m_{pc}'(t)|^2] \rangle_c = 1$, A' is the (2 x 2) matrix of the whitened source steering vectors $\boldsymbol{a}_p'$ $(0 \le p \le 2)$, such that the whitened steering vector $\boldsymbol{a}_p'$ is defined by

$$\boldsymbol{a}_p' \triangleq \pi_p^{1/2} F\,\boldsymbol{a}_p = \pi_p^{1/2} \begin{Bmatrix} \lambda_{s+}^{-1/2}\;\boldsymbol{u}_{s+}^\dagger \boldsymbol{a}_p \\ \lambda_{s-}^{-1/2}\;\boldsymbol{u}_{s-}^\dagger \boldsymbol{a}_p \end{Bmatrix} \tag{57}$$

While, for apparently SO uncorrelated sources ($|\rho_{12}| = 0$), which is in particular the case for zero-mean sources, the whitened source steering vectors $\boldsymbol{a}_p'$, $(0 \leq p \leq 2)$, are orthonormalized vectors and the matrix A' is an unitary matrix, it is no longer the case for apparently SO correlated sources ($|\rho_{12}| \neq 0$), for which the vectors $\boldsymbol{a}_p'$, $(0 \leq p \leq 2)$, are neither normalized nor orthogonal.

Proof: *To show the previous result let us firstly assume that the matrix A' is orthogonal. Under this assumption, as the matrix $R_{s'} \stackrel{\Delta}{=} A' R_{mc'} A'^{\dagger}$ corresponds to the identity matrix, the matrix $A'^{\dagger} R_{s'} A' = A'^{\dagger} A'$ is diagonal and equal to $A'^{\dagger} A' R_{mc'} A'^{\dagger} A'$, implying that the matrix $R_{mc'}$ is diagonal, which is not the case for apparently SO correlated sources.*

Let us now assume that the columns of A' are normalized. In this case, as $A'^{\dagger} R_{s'} A' = A'^{\dagger} A' = A'^{\dagger} A' R_{mc'} A'^{\dagger} A'$, we obtain that $R_{mc'} A'^{\dagger} A' = A'^{\dagger} A' R_{mc'} = I$, which means that $R_{mc'}$ is the inverse of $A'^{\dagger} A'$, which is the case for ($|\rho_{12}| = 0$) since the two matrices are the identity matrix but which is generally not the case for ($|\rho_{12}| \neq 0$) since R_{mc}', not diagonal, does not depend on the spatial properties of the sources.

To illustrate the previous result, assume to simplify the computations that the sensors are omnidirectional ($\boldsymbol{a}_1^{\dagger}\boldsymbol{a}_1 = \boldsymbol{a}_2^{\dagger}\boldsymbol{a}_2 = N$) and that the two sources 1 and 2 are orthogonal ($\alpha_{12} = 0$). In these conditions, it is possible to show, after tedious computations that

$$\boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_1' = \boldsymbol{a}_2'^{\dagger}\boldsymbol{a}_2' = 1 / (1 - |\rho_{12}|^2) \tag{58}$$

$$\boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_2' = -\rho_{12} / (1 - |\rho_{12}|^2) \tag{59}$$

which shows that the modulus of the spatial correlation coefficient, α_{12}', of the whitened sources 1 and 2, defined by the normalized inner product of $\boldsymbol{a}_1'$ and $\boldsymbol{a}_2'$, is equal to $|\alpha_{12}'| = |\rho_{12}|$ and increases with $|\rho_{12}|$.

C. Blind Identification from matrices $R_z(\tau)$ by the SOBI method

After the whitening operation of the observations, the temporal mean, $R_z(\tau)$, of the correlation matrix of the whitened observation vector $z(t)$, is given by

$$R_z(\tau) \triangleq \langle R_z(t,\tau) \rangle_c = A' R_{mc}'(\tau) A'^{\dagger} + \eta_2(\tau) FF^{\dagger} \triangleq R_s'(\tau) + \eta_2(\tau) FF^{\dagger} \qquad (60)$$

where $R_{mc}'(\tau) \triangleq \langle \mathrm{E}[\boldsymbol{m}_c'(t)\, \boldsymbol{m}_c'(t-\tau)^{\dagger}] \rangle_c$ and $\boldsymbol{m}_c'(t)$ is the normalized vector $\boldsymbol{m}_c(t)$ with components $m_{pc}'(t)$ $(1 \le p \le 2)$. Choosing Q parameters τ_q $(1 \le q \le Q)$, such that $\eta_2(\tau_q) = 0$, the process of joint diagonalization of the Q matrices $R_z(\tau_q)$, gives a (2 x 2) unitary matrix U maximizing the criterion (5) with $P = 2$. While for apparently SO uncorrelated sources ($|\rho_{12}| = 0$), A' is an unitary matrix which jointly diagonalizes the set of Q matrices $R_z(\tau_q)$, it is no longer the case for apparently SO correlated sources ($|\rho_{12}| \neq 0$) since A' is neither an unitary matrix nor an orthogonal matrix. In these conditions, even for sources with different spectrum, the two orthonormalized vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, corresponding to the two columns of U, become necessarily linear combinations of the whitened steering vectors $\boldsymbol{a}_1'$ and $\boldsymbol{a}_2'$, given by

$$\boldsymbol{u}_i = \alpha_i\, \boldsymbol{a}_1' + \beta_i\, \boldsymbol{a}_2' \quad , \quad i = 1, 2 \qquad (61)$$

where the coefficients α_i and β_i $(i = 1, 2)$ are dependent on the SO properties of the sources and are such that

$$|\alpha_i|^2 \boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_1' + |\beta_i|^2 \boldsymbol{a}_2'^{\dagger}\boldsymbol{a}_2' + 2\mathrm{Re}[\alpha_i^{*}\beta_i\, \boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_2'] = 1 \,, \quad i = 1, 2 \qquad (62)$$

$$\alpha_1^{*}\alpha_2\, \boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_1' + \beta_1^{*}\beta_2\, \boldsymbol{a}_2'^{\dagger}\boldsymbol{a}_2' + \alpha_1^{*}\beta_2\, \boldsymbol{a}_1'^{\dagger}\boldsymbol{a}_2' + \beta_1^{*}\alpha_2\, \boldsymbol{a}_2'^{\dagger}\boldsymbol{a}_1' = 0 \qquad (63)$$

Consequently, the blind identification stage of the SOBI method is perturbed by the apparent SO correlation of the sources and the behavior of the SOBI method is modified in the steady state. This non ideal behavior of the blind identification of the whitened source steering vector generates a degradation of the source separation process as it is shown in section D.

D. Blind source separation by the SOBI method

D1. *Performance criterion and spatial filter choice*

Following the description of the SOBI method in section II.C, from the blindly identified vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, considered as estimates of $\boldsymbol{a}_1'$ and $\boldsymbol{a}_2'$, it is possible to obtain, to within a scalar factor, an estimate of the true steering vectors of the sources, defined by $\hat{\boldsymbol{a}}_i =$

$F^{\#}\,\boldsymbol{u}_i$ (i = 1, 2), where # is the pseudo-inverse operation. Using (61) and the fact that $\boldsymbol{a}_i' \triangleq \pi_i^{1/2} F\,\boldsymbol{a}_i\ (1 \le i \le 2)$, the vectors $\hat{\boldsymbol{a}}_i$ can be written as

$$\hat{\boldsymbol{a}}_i \;=\; \alpha_i\,\sqrt{\pi_1}\,\boldsymbol{a}_1 \;+\; \beta_i\,\sqrt{\pi_2}\,\boldsymbol{a}_2 \quad , \qquad i \;=\; 1, 2 \tag{64}$$

In these conditions, both the optimal linear and time invariant (TI) spatial filter associated to the vectors $\hat{\boldsymbol{a}}_i\ (1 \le i \le 2)$ and the performance of the associated source separator can be computed.

For statistically independent zero-mean sources, the concepts of both source separator performance and optimal linear and TI source separator have been clearly defined in [8]. In particular, an estimate of the latter consists to implement, for each source i, the estimated Spatial Matched Filter, $\hat{\boldsymbol{w}}_{ix}$, defined by $\hat{\boldsymbol{w}}_{ix} \triangleq R_x^{-1}\ \hat{\boldsymbol{a}}_i$ [8]. However, for statistically independent sources which are first order cyclostationary, due to the potential apparent SO correlation of the latter, the concepts of source separator performance together with that of optimal source separator have to be redefined. Indeed, the power of the output, $y(t) = \boldsymbol{w}^{\dagger}\boldsymbol{x}(t)$, of a linear and TI spatial filter $\boldsymbol{w}$, whose input vector is given by (7) with $P = 2$ apparently SO correlated statistically independent first order cyclostationary sources, is given by

$$\begin{aligned}\pi_y \;&\triangleq\; <\mathrm{E}[|y(t)|^2]>_c \;=\; \boldsymbol{w}^{\dagger} R_x\,\boldsymbol{w} \\ &=\; \pi_1\,|\boldsymbol{w}^{\dagger}\boldsymbol{a}_1\,|^2 \;+\; \pi_2\,|\boldsymbol{w}^{\dagger}\boldsymbol{a}_2\,|^2 \;+\; 2\mathrm{Re}[(\pi_1\pi_2)^{1/2}\,(\boldsymbol{w}^{\dagger}\boldsymbol{a}_1)\,(\boldsymbol{w}^{\dagger}\boldsymbol{a}_2)^*\rho_{12}] \;+\; \eta_2\,\boldsymbol{w}^{\dagger}\boldsymbol{w}\end{aligned} \tag{65}$$

Then, for each source i (i = 1, 2), do we have to consider the term $2\mathrm{Re}[(\pi_1\pi_2)^{1/2}\ (\boldsymbol{w}^{\dagger}\boldsymbol{a}_1)\ (\boldsymbol{w}^{\dagger}\boldsymbol{a}_2)^*\rho_{12}]$ as a useful term for the source i ? as an interference term for the source i ? or as a term which is a combination of a useful and an interference part for the source i ? These questions have no easy answers and their analyse is out of the scope of the paper. Nevertheless they have to be clarified to introduce the concepts of both source separator performance and optimal linear and TI source separator in the presence of first order cyclostationary sources apparently SO correlated.

In the following, to simplify the problem, we still use the concept of source separator performances introduced in [8]. In other words, for each source k (k = 1, 2), the Signal to Interference plus Noise Ratio for the source k at the output of a spatial filter $\boldsymbol{w}_i$ is defined by

$$\mathrm{SINR}k[\boldsymbol{w}_i] \triangleq \pi_k \frac{|\boldsymbol{w}_i^\dagger \boldsymbol{a}_k|^2}{\boldsymbol{w}_i^\dagger R_{bk} \boldsymbol{w}_i} \quad (66)$$

where R_{bk} is the total noise correlation matrix for the source k, corresponding to the R_x matrix in the absence of the source k. For example $R_{b1} \triangleq \pi_2\, \boldsymbol{a}_2\, \boldsymbol{a}_2^\dagger + \eta_2 \mathrm{I}$. In these conditions, the restitution's quality of the source k at the output of the separator W, whose columns are the $\boldsymbol{w}_i$, can be evaluated by the maximum value of $\mathrm{SINR}k[\boldsymbol{w}_i]$ when i varies from 1 to 2, noted $\mathrm{SINRM}k[W]$. It is well known that for the previous performance criterion, the optimal source separator is the one which implements, for each source i, the Spatial Matched Filter, $\boldsymbol{w}_i$, defined, to within a scalar, by $\boldsymbol{w}_i \triangleq R_{bi}^{-1}\boldsymbol{a}_i$, which requires the knowledge of the non observable matrix R_{bi}. However, while the filter $\boldsymbol{w}_i$ is colinear to the filter $\boldsymbol{w}_{ix} \triangleq R_x^{-1}\boldsymbol{a}_i$ for zero mean sources, it is no longer the case for first order cyclostationary sources apparently SO correlated. For this reason, for each source i, we prefer to implement, in the following, an estimate of the optimal interference canceller (OIC) for the source i, whose performance are very close to that of the optimal filter in most cases [8]. The OIC for the source i is defined [8] by

$$\boldsymbol{w}_{i,OIC} \triangleq P_{bi}\, \boldsymbol{a}_i = [\mathrm{I} - (\boldsymbol{a}_j^\dagger \boldsymbol{a}_j)^{-1}\, \boldsymbol{a}_j \boldsymbol{a}_j^\dagger]\ \boldsymbol{a}_i \quad , \quad i, j = 1, 2 \ , \ j \neq i \quad (67)$$

where μ is a scalar and P_{bi} is the operator of orthogonal projection on the space orthogonal to the steering vectors of the interference for the source i. Then an estimate of the filter (67), $\hat{\boldsymbol{w}}_{i,OIC}$, can be obtained by replacing in (67) the true steering vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ by their blind estimates, $\hat{\boldsymbol{a}}_1$ and $\hat{\boldsymbol{a}}_2$, generated by the SOBI method, which gives finally

$$\hat{\boldsymbol{w}}_{i,OIC} \triangleq [\mathrm{I} - (\hat{\boldsymbol{a}}_j^\dagger \hat{\boldsymbol{a}}_j)^{-1} \hat{\boldsymbol{a}}_j \hat{\boldsymbol{a}}_j^\dagger]\, \hat{\boldsymbol{a}}_i \ , \quad i, j = 1, 2 \ , \ j \neq i \quad (68)$$

Considering $\hat{\boldsymbol{w}}_{i,OIC}$ as the column i of the (N x 2) matrix $\hat{W}_{OIC}$, the associated source separator can be written as [8]

$$\hat{W}_{OIC} = \hat{A}\, [\hat{A}^\dagger \hat{A}\,]^{-1} \quad (69)$$

where $\hat{A} \triangleq [\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2]$.

D2. *Performance computation*

To simplify the performance computation, we assume in this section that the sources are orthogonal (i.e. $\boldsymbol{a}_1^\dagger \boldsymbol{a}_2 = 0$), strong ($\varepsilon_j \triangleq \boldsymbol{a}_j^\dagger \boldsymbol{a}_j \, \pi_j / \eta_2 >> 1, j = 1, 2$) and that the sensors are omnidirectional ($\boldsymbol{a}_i^\dagger \boldsymbol{a}_i = N, i = 1, 2$). Under these assumptions, for each source k ($k = 1, 2$), the SINRMk at the output of the source separator $\hat{W}_{OIC}$, deduced from the SOBI method, can be computed using (64) and after tedious elementaries algebraic manipulations we obtain (in the steady state)

$$\mathrm{SINRM}k[\hat{W}_{OIC}] \approx \varepsilon_k \, [1 - \frac{\varepsilon_k \nu_k}{1 + \varepsilon_k \nu_k}] \tag{70}$$

where the quantities ν_k ($k = 1, 2$) are defined by

$$\nu_1 \triangleq \mathrm{Min}[\frac{|\alpha_1|^2}{|\beta_1|^2}, \frac{|\alpha_2|^2}{|\beta_2|^2}] = \mathrm{Min}[\frac{|\rho_{12}{}^*\zeta - 1|^2}{|\zeta - \rho_{12}|^2}, |\zeta|^2] \tag{71}$$

$$\nu_2 \triangleq \mathrm{Min}[\frac{|\beta_1|^2}{|\alpha_1|^2}, \frac{|\beta_2|^2}{|\alpha_2|^2}] = \mathrm{Min}[\frac{|\zeta - \rho_{12}|^2}{|\rho_{12}{}^*\zeta - 1|^2}, \frac{1}{|\zeta|^2}] \tag{72}$$

where $\zeta = \zeta(\rho_{12}) \triangleq \alpha_2/\beta_2$ and where the α_i and β_i ($i = 1, 2$) verify expressions (62) and (63) with the properties (58) and (59) valid for orthogonal sources.

The expression (70) shows that SINRMk does not depend on the input Signal to Noise Ratio (SNR) of the source different of the source k and is a decreasing function of ν_k. The performance of the SOBI method are optimal and the SINRMk is maximum and equal to ε_k for the two sources when $\nu_k = 0$ for the latter, i.e. when the blind identification of the two source whitened steering vectors is perfect. This situation always occurs for zero-mean sources having different spectrum but has no reason to occur for first order cyclostationary sources which are apparently SO correlated, as shown in the previous sections. In the latter case, the expressions (70) to (72) show that the performance at the output of the SOBI method degrade. In this case, the quantities ν_1 and ν_2, and thus the SINRM1 and the SINRM2, are related to each other and become a function of both ζ and the apparent SO correlation coefficient of the two sources, ρ_{12}, which are themselves directly related to the SO statistics of the first order cyclostationary sources. In particular, while for two equal power sources, SINRM1 corresponds to SINRM2 when $\rho_{12} = 0$, it is no longer the case when $\rho_{12} \neq 0$, as shown by (71) and (72), which shows that the apparent SO correlation of the sources introduces a difference in the restitution's quality of the sources. Some situations for which one of the parameters ν_1 and ν_2 is sufficiently high to generate strong performance degradation of the SOBI method are described in section VII.

These results show that to prevent strong performance degradation or, in the worst case, a very poor source separation at the output of the SOBI method and, more generally, at the output of the current SO cumulant-based blind source separation methods in first order (quasi)-cyclostationary contexts, the SO statistics of the data from which the blind identification of the source steering vectors is performed have to correspond to time averaged SO cumulants of the data and have to take into account the potential first order cyclostationarity of the sources. Such an estimator of the SO statistics of the observations is presented in section V.

V. ADAPTED SO BLIND SOURCE SEPARATION FOR FIRST AND SECOND ORDER CYCLOSTATIONARY SOURCES : COVARIANCE METHODS

A. Adapted SO BSS philosophy for first order cyclostationary sources : Covariance philosophy

It has been shown in the previous sections that the potential performance degradation of the current SO BSS method in the presence of first order cyclostationary sources is directly related to the potential non diagonal character of the source correlation matrix temporal mean $R_{mc}(\tau)$ defined by (22), which appears in the expression of the observation correlation matrix temporal mean $R_x(\tau) \triangleq < R_x(t, \tau) >_c$ given by (21). Moreover, the potential non diagonal character of $R_{mc}(\tau)$ is directly related to the potential non diagonal character of $E_{mc}(\tau)$, while the matrix source covariance matrix temporal mean, $R_{\Delta mc}(\tau)$, is always diagonal whatever the first order characteristic of the sources, by definition of the statistical independence of the latter. Consequently, to prevent poor performance of SO BSS methods for first order cyclostationary sources, it is necessary to exploit the information contained in the $R_{\Delta mc}(\tau)$ matrix instead of $R_{mc}(\tau)$. This can be done by exploiting the information contained in the temporal mean of the observation cumulant matrix $R_{\Delta x}(\tau)$ given, using (18) and (19), by

$$R_{\Delta x}(\tau) \;=\; A\, R_{\Delta mc}(\tau)\, A^{\dagger} + \eta_2(\tau)\, \mathrm{I} \;=\; R_x(\tau) \;-\; E_x(\tau) \tag{73}$$

where, using (14), $E_x(\tau)$ is given by

$$E_x(\tau) \triangleq \,<\boldsymbol{e}_x(t)\, \boldsymbol{e}_x(t-\tau)^{\dagger}>_c \;=\; \sum_{\gamma\in\Gamma} \boldsymbol{e}_x^{\gamma}\, \boldsymbol{e}_x^{\gamma\dagger}\, e^{j2\pi\tau\gamma} \tag{74}$$

and where it is recalled that Γ is the set of the first order cyclic frequencies γ of $\boldsymbol{x}(t)$ and $\boldsymbol{e}_x^{\gamma}$ is the cyclic mean of $\boldsymbol{x}(t)$ for the cyclic frequency γ, defined by (16). Note from (73) and (74) that, in the general case of first order cyclostationary sources and contrary to the stationary sources case, the $R_{\Delta x}(\tau)$ matrix cannot be obtained by substracting to the $R_x(\tau)$ matrix only the part of $E_x(\tau)$ associated to the zero cyclic frequency and given by $\boldsymbol{e}_x^{0}\, \boldsymbol{e}_x^{0\dagger} \triangleq\, <\boldsymbol{e}_x(t)>_c\, <\boldsymbol{e}_x(t)>_c^{\dagger}$ but has to take into account all the first order cyclic frequencies of the observation vector $\boldsymbol{x}(t)$, which then requires a preliminary step of first order cyclic frequencies estimation of the data.

Such a SO philosophy, called Covariance philosophy, prevents from generating apparently SO correlated first order cyclostationary sources when the latter are statistically independent and thus allows good source steering vector blind identification performances and then good separation performances, to within the spatial filtering process limits, whatever the first order characteristic of the sources, provided they have not the same spectrum. Moreover, for zero-mean cyclostationary sources (containing in particular the zero mean stationary ones), this philosophy corresponds, in the steady state, to the classical one. For this reason, the proposed philosophy can be considered as an extension of the classical one allowing also the processing of first order cyclostationary sources (containing also the non zero-mean stationary sources).

Nevertheless, the new proposed philosophy of SO BSS is unable to process first order cyclostationary deterministic sources, such as sinusoïd sources or more generally periodic sources, whose contribution in the $R_{\Delta x}(\tau)$ matrix disappears. For this reason, an extension of the Covariance method, allowing the processing of first order cyclostationary sources jointly with deterministic sources is presented in section VI.

B. Covariance philosophy implementation

In situations of practical interest, the cyclic frequencies and the statistics of the observations are unknown a priori and have to be estimated from the data, by temporal averaging operations, using the first and SO cyclo-ergodicity property of the data [26].

For this purpose, if K data snapshots of the observation vector $\boldsymbol{x}(t)$ are available and provided that the data are sufficiently oversampled, we introduce the estimates, $\hat{\boldsymbol{e}}_x^{\gamma}(K)$ and $\hat{\pi}_x(K)$, of $\boldsymbol{e}_x^{\gamma}$ and $\pi_x \triangleq <\mathrm{E}[\boldsymbol{x}(t)^{\dagger}\boldsymbol{x}(t)]>_c$ respectively, defined by

$$\hat{\boldsymbol{e}}_x^{\gamma}(K) \triangleq \frac{1}{K} \sum_{m=1}^{K} \boldsymbol{x}(m)\, \mathrm{e}^{-j2\pi\gamma m T_e} \tag{75}$$

$$\hat{\pi}_x(K) \triangleq \frac{1}{K} \sum_{m=1}^{K} \boldsymbol{x}(m)^{\dagger}\boldsymbol{x}(m) \tag{76}$$

In these conditions, a first order cyclic frequency detector of the observations can be implemented by selecting the cyclic frequencies γ which makes the criterion $V(\gamma)(K)$ greater than a threshold, whose value has to be chosen to maximise the detection probability of high power cyclic frequency γ for a given false alarm rate, where $V(\gamma)(K)$ is defined by

$$V(\gamma)(K) \triangleq \hat{\boldsymbol{e}}_x^{\gamma}(K)^{\dagger}\, \hat{\boldsymbol{e}}_x^{\gamma}(K) \,/\, \hat{\pi}_x(K) \tag{77}$$

Once the active first order cyclic frequencies of the observations have been detected, the $R_{\Delta x}(\tau)$ matrix for $\tau = qT_e$, defined by (73), can be estimated, from the K data snapshots, by the quantity $\hat{R}_{\Delta x}(qT_e)(K)$ defined by

$$\hat{R}_{\Delta x}(qT_e)(K) = \hat{R}_x(qT_e)(K) - \sum_{\gamma \in \Gamma} \hat{\boldsymbol{e}}_x^{\gamma}(K)\, \hat{\boldsymbol{e}}_x^{\gamma}(K)^{\dagger}\, \mathrm{e}^{j2\pi\gamma q T_e} \tag{78}$$

where $\hat{R}_x(qT_e)(K)$ is defined by (6).

Under the assumption of first and second order (quasi)-cyclostationary and cyclo-ergodic band-limited observations, and for sufficiently oversampled data, the estimator (78) is asymptotically unbiased and consistent, which means that it generates, in the steady state, the true matrix $R_{\Delta x}(qT_e)$, provided that the cyclic frequencies γ are exactly known [15]. Finally the current SO BSS methods, qualified of Correlation Methods, can be implemented from the $\hat{R}_{\Delta x}(qT_e)(K)$ matrix instead of $\hat{R}_x(qT_e)(K)$, giving birth to Covariance Methods.

VI. SO BLIND EXTRACTION OF STOCHASTIC AND DETERMINISTIC FIRST AND SECOND ORDER CYCLOSTATIONARY SOURCES : THE SOBEFOCYS METHOD

A. General philosophy

The so called *Correlation SOBI method* [3], which exploits the information contained in the $R_x(\tau)$ matrices, allows the processing of zero-mean statistically independent cyclostationary sources jointly with deterministic sources, provided the latter do not generate non diagonal terms in the $E_{mc}(\tau)$ matrix. This requires that the spectrum of the deterministic sources have no common frequencies, i.e. that these sources are spectrally separable, which can be considered as the definition of independent deterministic sources. Otherwise the deterministic sources become correlated, generate non zero non-diagonal terms in the $E_{mc}(\tau)$ matrix and the *Correlation SOBI method* is no longer adapted for this problem. In a same way, first order cyclostationary sources sharing at least one first order cyclic frequency still generate non zero non-diagonal terms in the $E_{mc}(\tau)$ matrix and become no longer separable by the *Correlation SOBI method.*

On the contrary, the so-called *Covariance SOBI method* (Section V), which exploits the information contained in the $R_{\Delta x}(\tau)$ matrices, allows the processing of both zero-mean and first order cyclostationary statistically independent sources but are unable to process deterministic sources since their contribution disappears from the $R_{\Delta x}(\tau)$ matrices.

In this context, the purpose of this section is to propose a SO BSS scheme which allows the joint processing of both zero-mean and first order cyclostationary statistically independent sources, either stochastic or deterministic. This scheme, called SOBEFOCYS, implements a first step allowing the processing and the extraction of stochastic sources, zero-mean or not, from the Covariance Method proposed in section V, and a second step allowing the processing and the extraction of deterministic sources from the results of the first step. Note that other schemes may be proposed but their analysis is beyond the scope of the paper.

B. The SOBEFOCYS method

B1. *Observation model*

In the presence of P_1 stochastic and P_2 deterministic statistically independent sources such that $P_1 + P_2 = P$, the observation model (7) can be written as

$$\boldsymbol{x}(t) \;=\; A_1\, \boldsymbol{m}_{1c}(t) \;+\; A_2\, \boldsymbol{m}_{2c}(t) \;+\; \boldsymbol{b}(t) \tag{79}$$

where A_1 and A_2 are the (N x P_1) and (N x P_2) matrices of the steering vectors of the stochastic and deterministic sources respectively, $\boldsymbol{m}_{1c}(t)$ and $\boldsymbol{m}_{2c}(t)$ are the (P_1 x 1) and (P_2 x 1) vectors of the complex envelope (with potential carrier residues) of the stochastic and deterministic sources respectively. Note that the stochastic sources are assumed to be first order cyclostationary and may share some first order cyclic frequencies whereas the deterministic sources are assumed to be polyperiodic. The statistical independence of the stochastic sources means that the components of $\boldsymbol{m}_{1c}(t)$ are statistically independent. The statistical independence of the deterministic sources means that the spectrum of latter share no frequencies. Finally the statistical independence of stochastic and deterministic sources means in this paper that the vectors $\boldsymbol{m}_{1c}(t)$ and $\boldsymbol{m}_{2c}(t-\tau)$ are not correlated whatever the value of τ, i.e that $< \mathrm{E}[\boldsymbol{m}_{1c}(t)\, \boldsymbol{m}_{2c}(t-\tau)^{\dagger}\,] >_c = 0\ \forall\ (t, \tau)$.

B2. *First and SO statistics of the data*

Under the previous assumptions, the first order statistics of the vector $\boldsymbol{x}(t)$ can be written as

$$\boldsymbol{e}_x(t) \;=\; A_1\, \boldsymbol{e}_{1c}(t) \;+\; A_2\, \boldsymbol{e}_{2c}(t) \tag{80}$$

where $\boldsymbol{e}_{1c}(t) \triangleq \mathrm{E}[\boldsymbol{m}_{1c}(t)]$ and $\boldsymbol{e}_{2c}(t) \triangleq \mathrm{E}[\boldsymbol{m}_{2c}(t)]$ have a Fourier serial expansion. The deterministic character of $\boldsymbol{m}_{2c}(t)$ implies that the latter vector disappears from the temporal mean, $R_{\Delta x}(\tau)$, of the covariance matrix $R_{\Delta x}(t, \tau)$, given by

$$R_{\Delta x}(\tau) \;=\; A_1\, R_{\Delta m1c}(\tau)\, A_1^{\dagger} \;+\; \eta_2(\tau)\, \mathrm{I} \tag{81}$$

where $R_{\Delta m1c}(\tau) \triangleq\ < \mathrm{E}[\Delta\boldsymbol{m}_{1c}(t)\ \Delta\boldsymbol{m}_{1c}(t-\tau)^{\dagger}] >_c$ is a diagonal matrix and $\Delta\boldsymbol{m}_{1c}(t) \triangleq \boldsymbol{m}_{1c}(t) - \boldsymbol{e}_{1c}(t)$. Finally, under the assumptions of section B1, the temporal mean, $R_x(\tau)$, of the correlation matrix $R_x(t, \tau)$, is given by

$$R_x(\tau) \;=\; A_1\, R_{m1c}(\tau)\, A_1^{\dagger} \;+\; A_2\, R_{m2c}(\tau)\, A_2^{\dagger} \;+\; \eta_2(\tau)\, \mathrm{I} \tag{82}$$

where $R_{mic}(\tau) \triangleq \,<\mathrm{E}[\boldsymbol{m}_{ic}(t)\, \boldsymbol{m}_{ic}(t-\tau)^{\dagger}]>_c$, $1 \leq i \leq 2$, such that $R_{m2c}(\tau)$ is a diagonal matrix.

B3. *Detection of deterministic sources from P1 and P2 estimation*

The presence of deterministic sources can be detected from the estimation of the number of sources, P_1 and P, contained in the matrices (81) and (82) for $\tau = 0$ respectively. The different steps of this process are presented hereafter :

- Estimation, $\hat{R}_x(K)$, of R_x from K data snapshots $\boldsymbol{x}(k)$, using the empirical estimator (6) with $q = 0$.
- Estimation, $\hat{P}$, of P from the eigendecomposition of $\hat{R}_x(K)$, using a classical eigenvalue test
- Estimation of the first order cyclic frequencies γ of the observations using (75), (76) and (77)
- Estimation, $\hat{R}_{\Delta x}(K)$, of $R_{\Delta x}$ from $\hat{R}_x(K)$ and the estimated cyclic frequencies γ using (78) for $q = 0$
- Estimation, $\hat{P}_1$, of P_1 from the eigendecomposition of $\hat{R}_{\Delta x}(K)$, using a classical eigenvalue test
- Estimation, $\hat{P}_2$, of P_2 by $\hat{P}_2 = \hat{P} - \hat{P}_1$
- Deterministic sources are detected if $\hat{P}_2 \neq 0$

B4. *First step of the SOBEFOCYS method : Extraction of the stochastic sources*

The blind estimation, $\hat{A}_1$, of the mixing matrix A_1 can be obtained, to within a ($\hat{P}_1$ x $\hat{P}_1$) permutation matrix Π_1 and a ($\hat{P}_1$ x $\hat{P}_1$) diagonal matrix Λ_1, by implementing the so-called Covariance SOBI method, from $\hat{P}_1$, $\hat{R}_{\Delta x}(K)$ and several covariance matrices $\hat{R}_{\Delta x}(qT_e)(K)$ with $q \neq 0$, computed from (78), (6) and the cyclic frequencies γ. Then from the obtained $\tilde{A}_1 \triangleq \hat{A}_1 \Lambda_1 \Pi_1$ matrix, a (N x $\hat{P}_1$) stochastic sources separator W_1 has to be built to allow the extraction of the stochastic sources, i.e. the estimation, $\hat{\boldsymbol{m}}_{1c}(t)$, of the stochastic sources vector $\boldsymbol{m}_{1c}(t)$, to within a permutation and a diagonal matrix, by

$$y_1(\mathrm{t}) \;=\; W_1{}^{\dagger}\, \boldsymbol{x}(t) \tag{83}$$

In the absence of deterministic sources ($\hat{P}_2 = 0$), as suggested in section IV. D1, the separator W_1 is chosen so as to implement the OIC for each stochastic source, which is equivalent, for a spatially white noise, to implement the Least Square separator [8], defined by

$$W_1 \;=\; \tilde{A}_1\, [\, \tilde{A}_1{}^{\dagger}\, \tilde{A}_1\,]^{-1} \;=\; \hat{A}_1\, [\hat{A}_1{}^{\dagger}\, \hat{A}_1\,]^{-1}\, \Lambda_1{}^{-\dagger}\, \Pi_1 \tag{84}$$

and which gives

$$y_1(\mathrm{t}) \;=\; \Pi_1{}^{\dagger}\, \Lambda_1{}^{-1}\, [\hat{A}_1{}^{\dagger}\, \hat{A}_1\,]^{-1}\, \hat{A}_1{}^{\dagger}\, \boldsymbol{x}(t) \;=\; \Pi_1{}^{\dagger}\, \Lambda_1{}^{-1}\, \hat{\boldsymbol{m}}_{1c}(t) \tag{85}$$

In the presence of deterministic sources ($\hat{P}_2 \neq 0$), the column i, $\boldsymbol{w}_{1i}$, of the separator W_1 is chosen so as to minimize the output power $\boldsymbol{w}_{1i}{}^{\dagger}\, \hat{R}_x(K)\, \boldsymbol{w}_{1i}$ under a first constraint, $\boldsymbol{w}_{1i}{}^{\dagger}\, \tilde{\boldsymbol{a}}_{1i}$ = 1, of zero distorsion in the direction of the source associated to the column i, $\tilde{\boldsymbol{a}}_{1i}$, of $\tilde{A}_1$, and under a second constraint of nulling all the other stochastic sources, i.e. $\boldsymbol{w}^{\dagger}\, \tilde{\boldsymbol{a}}_{1j} = 0$ for $j \neq i$. It is easy to verify that W_1 be written as

$$W_1 = \hat{R}_x(K)^{-1}\, \tilde{A}_1\, [\, \tilde{A}_1{}^{\dagger}\, \hat{R}_x(K)^{-1}\, \tilde{A}_1\,]^{-1} = \hat{R}_x(K)^{-1}\, \hat{A}_1\, [\hat{A}_1{}^{\dagger}\, \hat{R}_x(K)^{-1}\, \hat{A}_1\,]^{-1}\, \Lambda_1{}^{-\dagger}\, \Pi_1 \tag{86}$$

and which gives

$$y_1(\mathrm{t}) \;=\; \Pi_1{}^{\dagger}\, \Lambda_1{}^{-1}\, [\hat{A}_1{}^{\dagger}\, \hat{R}_x(K)^{-1}\, \hat{A}_1\,]^{-1}\, \hat{A}_1{}^{\dagger}\, \hat{R}_x(K)^{-1}\, \boldsymbol{x}(t) \;=\; \Pi_1{}^{\dagger}\, \Lambda_1{}^{-1}\, \hat{\boldsymbol{m}}_{1c}(t) \tag{87}$$

Note that in the absence of deterministic sources, the separator (86) asymptotically (i.e. when K becomes infinite) corresponds to (84).

B5. *Second step of the SOBEFOCYS method : Extraction of the deterministic sources*

Once the stochastic sources have been extracted, the deterministic sources can be processed if $\hat{P}_2 \neq 0$. To this aim, we firstly remove the stochastic sources from the observation vector $\boldsymbol{x}(t)$ by building the projection, $\boldsymbol{v}(t)$, of the observation vector $\boldsymbol{x}(t)$ on the subspace orthogonal to the column of $\tilde{A}_1$, defined by

$$\boldsymbol{v}(t) \;\triangleq\; \mathrm{F}_1\; \boldsymbol{x}(t) \tag{88}$$

where $\mathrm{F}_1 \triangleq \mathrm{I} - \tilde{A}_1[\, \tilde{A}_1{}^{\dagger}\, \tilde{A}_1\,]^{-1}\, \tilde{A}_1{}^{\dagger}$. Note that for a perfect blind identification of the matrix A_1, we obtain $\hat{A}_1 = A_1$ and the vector $\boldsymbol{v}(t)$ takes the form

$$\boldsymbol{v}(t) \;=\; \mathrm{F}_1\, A_2\, \boldsymbol{m}_{2c}(t) \;+\; \mathrm{F}_1\, \boldsymbol{b}(t) \tag{89}$$

The blind estimation, $\hat{A}_{\mathrm{F}2} \triangleq \mathrm{F}_1\, \hat{A}_2$, of the projected mixing matrix $\mathrm{F}_1\, A_2$ can be obtained, to within a $(\hat{P}_2 \times \hat{P}_2)$ permutation matrix Π_2 and a $(\hat{P}_2 \times \hat{P}_2)$ diagonal matrix Λ_2, by implementing the so-called Correlation SOBI method [3], from $\hat{P}_2$, $\hat{R}_v(K)$ and several correlation matrices $\hat{R}_v(qT_e)(K)$ with $q \neq 0$, computed from (6) with the indice x replaced by v. Then from the obtained $\tilde{A}_{\mathrm{F}2} \triangleq \mathrm{F}_1\, \tilde{A}_2 \triangleq \hat{A}_{\mathrm{F}2}\, \Lambda_2\, \Pi_2$ matrix, a $(N \times \hat{P}_2)$ deterministic source separator W_2 has to be built to allow the extraction of the deterministic sources, i.e. the estimation, $\hat{\boldsymbol{m}}_{2c}(t)$, of the deterministic source vector $\boldsymbol{m}_{2c}(t)$, to within a permutation and a diagonal matrix, by

$$y_2(t) \;=\; W_2^{\dagger}\, \boldsymbol{x}(t) \tag{90}$$

The separator W_2 is chosen so as to implement the Least Square separator for the deterministic sources, once the stochastic sources have been removed from the observation vector. It is then defined by

$$W_2 \;= \mathrm{F}_1\, \tilde{A}_{\mathrm{F}2}\, [\, \tilde{A}_{\mathrm{F}2}{}^{\dagger}\, \tilde{A}_{\mathrm{F}2}\,]^{-1} \;=\; \mathrm{F}_1\; \hat{A}_2\, [\hat{A}_2{}^{\dagger}\, \mathrm{F}_1\, \hat{A}_2\,]^{-1}\, \Lambda_2{}^{-\dagger}\, \Pi_2 \tag{91}$$

and which gives

$$y_2(\mathrm{t}) \;=\; \Pi_2{}^{\dagger}\, \Lambda_2{}^{-1}\, [\hat{A}_2{}^{\dagger}\, \mathrm{F}_1\, \hat{A}_2\,]^{-1}\, \hat{A}_2{}^{\dagger}\, \mathrm{F}_1\, \boldsymbol{x}(t) \;=\; \Pi_2{}^{\dagger}\, \Lambda_2{}^{-1}\; \hat{\boldsymbol{m}}_{2c}(t) \tag{92}$$

VII. SIMULATIONS

The results presented in the sections I to V are illustrated in figures 1 to 4, where two statistically independent binary CPFSK sources are assumed to be received by a circular array of 5 uniformly spaced sensors with a radius $r = 0.55\ \lambda$ (λ is the wavelenght). The two sources are assumed to be orthogonal to each other ($\boldsymbol{a}_1^{\dagger}\boldsymbol{a}_2 = 0$) which is in particular the case when their angle of arrival is such that $\theta_1 = 50°$ and $\theta_2 = 91°$. They have the same input SNR (Signal to Noise Ratio) of 10 dB and are synchronized. Their symbol durations T_i and their modulation indices h_i ($i = 1, 2$) are such that $h_1/T_1 = h_2/T_2 = 1/4T_e$ for $h_1 = 2$ and $h_2 = 4$. Besides, the considered SOBI method aims at diagonalizing an estimation of only one correlation or one covariance matrix temporal mean of the whitened observation matrix for $\tau = 4T_e$. In the first case, the SOBI method (which is the current one) is called SOBI_COR whereas in the second case the SOBI method (which is the new one) is called either SOBI_COV when all the first order cyclic frequencies belonging to Γ are taken into account in (78) or SOBI_ACOV (for Approximated Covariance) when only the zero cyclic frequency is taken into account in (78). Finally, the SINRMk ($k = 1, 2$) at the output of the SOBI methods, computed in these figures, are averaged over 200 realisations.

Under the previous assumptions and assuming that the two sources have a carrier residu such that $\Delta f_1 = \Delta f_2 = h_1/2T_1$, the figures 1 and 2 show the variations of the SINRM1 and the SINRM2 respectively at the output of the SOBI_COR, the SOBI_ACOV and the SOBI_COV separators, as a function of K. As the two sources share the first order cyclic frequencies $\gamma = 0$ and $\gamma = h_1/T_1$, they become apparently SO correlated in the $\hat{R}_x(qT_e)(K)$ matrix, with a coefficient ρ_{12} equal to 0.5. In this case, it can be shown that $\zeta = \sqrt{2}/(2 + \sqrt{3})^{1/2}$ or $\zeta = -\sqrt{2}/(2 - \sqrt{3})^{1/2}$, $\nu_1 = 2/(2 + \sqrt{3})$ and $\nu_2 = (2 - \sqrt{3})/2$, which explains the high performance degradation (the SINRM1 and the SINRM2 converge toward 2.5 dB and 8 dB respectively instead of 17 dB) and the poor separation of the sources at the output of the current SOBI (SOBI_COR) method. On the contrary, the implementation of the SOBI method from the $\hat{R}_{\Delta x}(qT_e)(K)$ matrix, given by (78) where the two cyclic frequencies $\gamma = 0$ and $\gamma = h_1/T_1$ have been used (SOBI_COV), shows performances approaching the optimality as the number of snapshots increases. Nevertheless, the use in (78) of only one ($\gamma = 0$) of the two common first order cyclic frequencies of the sources (SOBI_ACOV) is not sufficient to obtain optimal performances.

The figures 3 and 4 show, for several values of Δf_2, the variations of the SINRM1 and the SINRM2 respectively at the output of the SOBI_COR separator, as a function of the number of snapshots K, for several values, 0, 0.005 and 0.01, of the differential carrier residu $(\Delta f_1 - \Delta f_2)$ xT_e. Note the poor separation of the two sources when $\Delta f_1 = \Delta f_2$, even in the steady-state (SINRM1 = 2.7 dB, SINRM2 = 8 dB), and the decreasing convergence speed of the SOBI_COR separator as $(\Delta f_1 - \Delta f_2)$ decreases. In this latter case, the steady-state output performance are not affected by the use of the empirical SO statistics estimator since the source correlation matrix is diagonal due to the fact that the sources do not share any first order cyclic frequencies. Nevertheless, when the first order cyclic frequencies of the sources are close to each other, the output performance degradation obtained from a short time observation is a decreasing function of the difference between the first order cyclic frequencies of the sources.

The results presented in section VI are illustrated in figures 5 and 6, where the context is the same as the one depicted for figures 1 to 4, to within the DOA of the source 2 which is equal to $\theta_2 = -179°$, but where 2 independent deterministic sources, corresponding to two sinusoïds, have been added in the observation vector. These deterministic sources have a SNR of 10 dB, come from the directions $\theta_3 = 125°$ and $\theta_4 = 93°$ respectively and are such that $\Delta f_3 = 1/3T_e$, $\Delta f_4 = 1/5T_e$. In this context, the figures 5 and 6 show the variations of the SINRMi ($1 \leq i \leq 4$) at the output of the SOBI_COR and the SOBEFOCYS method respectively, as a function of K. Note that the SOBI_COR method separates the deterministic sources but has some difficulties to separate the first order cyclostationary stochastic sources as the latter share the first order cyclic frequencies $\gamma = 0$ and $\gamma = h_1/T_1$. In the same context, the SOBEFOCYS method allows the good separation of all the sources, deterministic or not, despite of the first order cyclostationarity of the latter and the fact that the sources 1 and 2 share some first order cyclic frequencies.

VII. CONCLUSION

In this paper, the behavior of the current SO cumulant-based blind source separation methods, such as the SOBI method, initially developed for zero mean, stationary and ergodic sources, has been analysed for cyclostationary and cyclo-ergodic sources assumed first order cyclostationary. Examples of such sources correspond to CPFSK sources having an integer modulation indice, FSK or some AM sources.

It has been shown in the paper that when two sources share at least one first order cyclic frequency, they become apparently SO correlated in the temporal mean of the data correlation matrix and the performance of the current SO BSS methods may be strongly affected by such sources despite of the fact that they are statistically independent.

To solve this problem, it has been proposed in the paper to implement the current SO BSS method from the temporal mean of the data covariance matrix instead of the correlation matrix, which generates Covariance Method instead of Correlation ones. For this purpose, an asymptotically unbiased and consistent estimator of the data covariance matrix temporal mean has been proposed for first and second order cyclostationary and cyclo-ergodic sources. However, the use of this estimator requires the knowledge or the a priori estimation of all the first order cyclic frequencies of the observations.

The so-called Covariance BSS philosophy proposed in this paper allows to separate both stationary and cyclostationary statistically independent sources, either zero mean or not (first order cyclostationary), provided they have not the same spectrum. In that sense, it extends the applicability of the current Correlation BSS methods [3], developed for stationary sources, to first and SO cyclostationary sources.

However, the main limitation of the proposed Covariance philosophy is that it is unable to separate first order cyclostationary deterministic sources such as sinusoïds or polyperiodic sources. For this reason, a SO BSS scheme, called SOBEFOCYS method, allowing the joint processing of arbitrary modulated (stochastic or deterministic, zero-mean or not) statistically independent cyclostationary sources has finally been proposed in the paper. After the estimation of the number of deterministic sources, this scheme implements the proposed

Covariance method in a first step, allowing the processing and the extraction of stochastic sources, and then implements a second step allowing the processing and the extraction of deterministic sources from the results of the first step. Note that to our knowledge, the SOBEFOCYS scheme is the first one which allows the joint SO BSS of arbitrarily modulated (stochastic or deterministic, zero-mean or not) cyclostationary sources.

APPENDIX A

In this Appendix, we compute the first order cyclic statistics of a M_p-FSK source p built from M_p local oscillators.

The cyclic mean $e_p^{\gamma_p}$ has to be computed for cyclic frequencies γ_p multiple of $(1/T_p)$. For the cyclic frequency $\gamma_{pi} = i\,/T_p$, using the fact that the function $e_p(t)\,\exp[-j2\pi\gamma_{pi}t]$ is periodic with a period equal to T_p, we deduce from (27) that the cyclic mean $e_p^{\gamma_{pi}}$, defined by (11) with $\gamma_p = \gamma_{pi}$, is given by

$$e_p^{\gamma_{pi}} = \pi_p^{1/2}\,\frac{1}{M_p}\,\frac{1}{T_p}\int_0^{T_p}\sum_n\sum_{m=0}^{(M_p-2)/2}\Big(\exp\{j[\theta_{p(2m+1)} + 2\pi f_{dp}(2m+1)(t-nT_p)]\} +$$

$$\exp\{j[\theta_{-p(2m+1)} - 2\pi f_{dp}(2m+1)(t-nT_p)]\}\Big)\ \mathrm{Rect}_p(t-nT_p)\ \exp[-j2\pi\gamma_{pi}t]\ dt \qquad \text{(A1)}$$

Using the fact that $\mathrm{Rect}_p(t)$ is zero outside the interval $[0, T_p]$ and making in (A1) the change of variables $v = t - nT_p$, $w = v\,/\,T_p$, we obtain, after some elementaries manipulations the expression (29).

APPENDIX B

In this Appendix, we compute the first order statistics of a M_p-CPFSK source p.

B1. Computation of $e_p(t)$

From the expressions (26) and (33), using the statistical independance of the symbols a_n^p we obtain

$$e_p(t) = \pi_p^{1/2}\sum_n \lim_{l\to\infty}\prod_{k=-l}^{n-1} \mathrm{E}\{\exp[j\pi h_p a_k^p]\}\ \mathrm{E}\{\exp[j\pi\, a_n^p\,(h_p/T_p)(t-nT_p)]\}\ \mathrm{Rect}_p(t-nT_p) \qquad \text{(B1)}$$

On the other hand, for equiprobable symbols a_n^p, with probability $(1/M_p)$, it is easy to verify that, for an arbitrary real value β, we obtain

$$\mathrm{E}\{\exp[j\,\beta\, a_k^p]\} = \frac{2}{M_p}\sum_{m=0}^{(M_p-2)/2}\cos[(2m+1)\beta] \qquad \text{(B2)}$$

Applying (B2) into (B1), the expressions (34) to (37) follow immediately.

B2. Computation of $e_p^{\gamma_p}$ for $\rho_p = 1$

When $\rho_p = 1$, the cyclic mean $e_p^{\gamma_p}$ has to be computed for cyclic frequencies γ_p multiple of $(1/T_p)$. For the cyclic frequency $\gamma_{pi} = i/T_p$, using the fact that the function $e_p(t)$ $\exp[-j2\pi\gamma_{pi}t]$ is periodic with a period equal to T_p, we deduce from (38) that the cyclic mean $e_p^{\gamma_{pi}}$, defined by (11) with $\gamma_p = \gamma_{pi}$, is given by

$$e_p^{\gamma_{pi}} = \pi_p^{1/2} \frac{2}{M_p} \frac{1}{T_p} \int_0^{T_p} \sum_n u_p(t - nT_p) \exp[-j2\pi\gamma_{pi}t] \, dt \tag{B3}$$

Using the fact that $u_p(t)$ is zero outside the interval $[0, T_p]$, making in (B3) the change of variables $v = t - nT_p$, $w = v / T_p$ and using (37) in (B3) with $h_p = 2q$, where q is an integer, we obtain, after some elementaries manipulations that

$$e_p^{\gamma_{pi}} = \pi_p^{1/2} \frac{2}{M_p} \sum_{m=0}^{(M_p-2)/2} \int_0^1 \cos[(2m+1)2q\pi w] \exp[-j2\pi\, iw] \, dw \tag{B4}$$

which, after elementaries trigonometric manipulations, can also be written as

$$e_p^{\gamma_{pi}} = \pi_p^{1/2} \frac{1}{M_p} \sum_{m=0}^{(M_p-2)/2} \int_0^1 \{\cos[2\pi w((2m+1)q + i)] + \cos[2\pi w((2m+1)q - i)]\} dw \tag{B5}$$

For each value of m, $0 \le m \le (M_p - 2)/2$, if $i \ne \pm (2m+1)q$, it is easy to verify that the expression (B5) gives $e_p^{\gamma_{pi}} = 0$. However, if $i = \pm (2m+1)q$, we deduce from (B5) that $e_p^{\gamma_{pi}} = \pi_p^{1/2}/M_p$. Recalling that $q = h_p/2 = f_{dp} T_p$, we deduce that $e_p^{\gamma_{pi}}$ is not zero for cyclic frequencies $\beta_{pi} = i/T_p = \pm (2m+1)q/T_p = \pm (2m+1) f_{dp}$ and the associated cyclic mean is $\pi_p^{1/2}/M_p$, result which is expressed by the expression (39).

B3. Computation of $e_p^{\gamma_p}$ for $\rho_p = -1$

When $\rho_p = -1$, the cyclic mean $e_p^{\gamma_p}$ has to be computed for cyclic frequencies γ_p multiple of $(1/2T_p)$. For the cyclic frequency $\gamma_{pi} = i/2T_p$, using the fact that the function $e_p(t)$

$\exp[-j2\pi\gamma_{pi}t]$ is periodic with a period equal to $2T_p$, we deduce from (38) that the cyclic mean $e_p^{\gamma_{pi}}$, defined by (11) with $\gamma_p = \gamma_{pi}$, is given by

$$e_p^{\gamma_{pi}} = \pi_p^{1/2}\, K_p \frac{2}{M_p} \frac{1}{2T_p} \int_0^{2T_p} \sum_n (-1)^n\, u_p(t - nT_p)\, \exp[-j2\pi\gamma_{pi}t]\, dt \qquad \text{(B6)}$$

Using the fact that $u_p(t)$ is zero outside the interval $[0, T_p]$, making in (B6) the change of variables $v = t - nT_p$, $w = v / T_p$ and using (37) in (B6) with $h_p = (2q+ 1)$, where q is an integer, we obtain, after some elementaries manipulations that

$$e_p^{\gamma_{pi}} = \pi_p^{1/2}\, K_p \frac{1}{M_p} (1 - e^{-ji\pi}) \sum_{m=0}^{(M_p-2)/2} \int_0^1 \cos[(2m+1)(2q+ 1)\pi w]\, \exp[-j\pi\, iw]\, dw \qquad \text{(B7)}$$

If i is even, $e^{-ji\pi} = 1$ and $e_p^{\gamma_{pi}} = 0$. However, if i is odd, i.e. if $i = (2s+1)$ where s in an integer, it is easy to verify, after some elementaries trigonometric manipulations that

$$e_p^{\gamma_{pi}} = \pi_p^{1/2}\, K_p \frac{1}{M_p} \sum_{m=0}^{(M_p-2)/2}$$

$$\{\int_0^1 \{\cos[\pi w((2m+1)(2q+1) + (2s+1))] + \cos[\pi w((2m+1)(2q+1) - (2s+1))]\} dw \qquad \text{(B8)}$$

For each value of m, $0 \le m \le (M_p - 2)/2$, if $i = (2s+1) \neq \pm\,(2m+1)(2q+1)$, it is easy to verify that the expression (B8) gives $e_p^{\gamma_{pi}} = 0$. However, if $i = (2s+1) = \pm\,(2m+1)(2q+1)$, we deduce from (B8) that $e_p^{\gamma_{pi}} = \pi_p^{1/2} K_p / M_p$. Recalling that $h_p = (2q+ 1) = 2f_{dp}T_p$, we deduce that $e_p^{\gamma_{pi}}$ is not zero for cyclic frequencies $\gamma_{pi} = i / 2T_p = \pm\,(2m + 1)(2q+ 1)/2T_p = \pm\,(2m + 1)\, f_{dp}$ and the associated cyclic mean is $\pi_p^{1/2} K_p / M_p$, result which is expressed by the expression (39).

REFERENCES

[1] K. ABED-MERAIM, Y. XIANG, J.H. MANTON, Y. HUA, "Blind source separation using second-order cyclostationary statistics", *IEEE Trans. Signal Processing*, Vol.49, N°4, pp 694-701, April 2001.

[2] B.G. AGEE, S.V. SCHELL, W.A. GARDNER, "Spectral Self-Coherence Restoral : A new approach to blind adaptive signal extraction using antenna array", *Proc. IEEE*, pp. 753-767, April 1990.

[3] A. BELOUCHRANI, K. ABED-MERAIM, J.F. CARDOSO, E. MOULINES, "A blind source separation technique using second-order statistics", *IEEE Trans. Signal Processing*, Vol.45, N°2, pp 434-444, Feb. 1997.

[4] R.A. BOYLES, W.A. GARDNER, "Cycloergodic properties of discrete-parameter non stationary stochastic processes", *IEEE Trans. Info. Theory*, Vol.39, N°1, pp 105-114, Janv. 1983.

[5] J.F. CARDOSO, A. SOULOUMIAC, "Blind beamforming for non-gaussian signals", *IEE Proceedings-F*, Vol.140, N°6, pp 362-370, Dec. 1993.

[6] L. CASTEDO, A.R. FIGUEIRAS-VIDAL, "An adaptive beamforming technique based on cyclostationary signal properties", *IEEE Trans. Signal Processing*, Vol.43, N°7, pp 1637-1650, July 1995.

[7] P. CHEVALIER, "Optimal time invariant and widely linear spatial filtering for radiocommunications", *Proc. EUSIPCO*, Triestre (Italy), pp. 559-562, Sept. 1996.

[8] P. CHEVALIER, "Optimal separation of independent narrow-band sources - Concept and Performance", *Signal Processing*, Elsevier, Special issue on Blind Source Separation and Multichannel Deconvolution, Vol 73, N°1-2, pp. 27-47, Feb. 1999.

[9] P. CHEVALIER, "On the blind implementation of the ultimate source separators for arbitrary noisy mixtures", *Signal Processing*, Elsevier, Vol 78, N°3, pp. 277-287, Nov. 1999.

[10] P. CHEVALIER, A. FERREOL, L. ALBERA, "On the behavior of current second order blind source separation methods for first and second order cyclostationary sources – Application to CPFSK sources", *Proc ICASSP*, Orlando (USA), May 2002.

[11] P. CHEVALIER, A. FERREOL, L. ALBERA, "Méthodologie générale pour la séparation aveugle de sources cyclostationnaires arbitraires – Application à l'écoute passive des radiocommunications", *Proc GRETSI*, Paris (France), Sept. 2003.

[12] P. CHEVALIER, A. MAURICE, "Blind and informed cyclic array processing for cyclostationary signals", *Proc EUSIPCO*, Rhodes (Greece), pp. 1645-1648, Sept. 1998.

[13] P. COMON, "Independent component analysis – a new concept ?", *Signal Processing*, Elsevier, Vol 36, N°3, Special issue on higher order statistics, pp 287-314, April 1994.

[14] P. COMON, P. CHEVALIER, "Blind source separation: Models, Concepts, Algorithms and Performance", Chapter 5 of the book *Unsupervised adaptive filtering - Tome 1 – Blind Source Separation*, pp. 191 - 235, Dir. S. Haykins, Wiley, 445 p, 2000.

[15] A.V. DANDAWATE, G.B. GIANNAKIS, "Asymptotic theory of mixed time averages and Kth-order cyclic-moment and cumulant statistics", *IEEE Trans. Information Theory*, Vol.41, N°1, pp 216-232, Jan. 1995.

[16] L. DeLATHAUWER, B. DeMOOR, J. VANDEWALLE, "Independent component analysis based on higher-order statistics only", *Proc IEEE SP Workshop on Statistical Signal and Array Processing*, Corfu (Greece), pp. 356-359, June 1996.

[17] C. DEMEURE, P. CHEVALIER, "The smart antennas at Thomson-CSF Communications : Concepts, implementations, performances and applications", *Annals of Telecommunications*, Vol 53, N° 11-12, pp 466-482, 1998.

[18] A. FERREOL, P. CHEVALIER, "Higher Order Blind Source Separation using the Cyclostationarity Property of the Signals", *ICASSP*, Munich (Germany), pp 4061-4064, Apr. 1997.

[19] A. FERREOL, P. CHEVALIER, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources", *IEEE Trans. Signal Processing*, Vol 48, N° 6, pp. 1712-1725, June 2000. Erratum : Vol 50, N°4, pp. 990, April 2002.

[20] A. FERREOL, P. CHEVALIER, L. ALBERA, "Higher order blind separation of non zero-mean cyclostationary sources", *Proc. EUSIPCO*, Toulouse (France), Sept. 2002.

[21] W.A. GARDNER, "Spectral correlation of modulated signals : Part I – Analog modulation", *IEEE Trans. On Communications*, Vol 35, N°6, pp. 584-594, June 1987.

[22] W.A. GARDNER, C.M. SPOONER, "The cumulant theory of cyclostationary time series, Part I : Foundation", *IEEE Trans. Signal Processing*, Vol 42, pp. 3387-3408, Dec. 1994.

[23] P. GOURNAY, P. VIRAVEAU, "Corrélation spectrale théorique des modulations CPM - Partie I: Résultats analytiques pour les modulations CPFSK à 2 états (1-REC)", *Annales des Télécommunications*, Vol 53, N°7-8, pp 267-278, Jui./Aout 1998.

[24] S. HOUCKE, A. CHEVREUIL, P. LOUBATON, "Séparation autodidacte d'un mélange de sources à débits multiples", *Proc. GRETSI*, Toulouse, (France), Sept 2001.

[25] Y.C. LIANG, A.R. LEYMAN, B.H. SOONG, "Blind source separation using second-order cyclic statistics", *Proc SPAWC*, Paris (France), pp 57-60, 1997.

[26] L. MAZET, P. LOUBATON, "Estimation de la période symbole d'un signal modulé linéairement basée sur des cyclo-corrélations", *Proc. GRETSI*, Vannes (France), Sept. 1999.

[27] A. NAPOLITANO, "Cyclic higher order statistics : Input/output relations for discrete- and continuous-time MIMO linear almost-periodically time variant systems", *Signal Processing*, Elsevier, Vol 42, pp 147-166, 1995.

[28] A. NAPOLITANO, C.M. SPOONER, "Cyclic spectral analysis of Continuous-Phase Modulated signals", *IEEE Trans. Signal Processing*, Vol 49, N°1, pp. 30-44, Jan 2001.

[29] B. PICINBONO, "On Circularity", *IEEE Trans. Signal Processing*, Vol 42, N°12, pp. 3473-3482, Dec 1994.

[30] B. PICINBONO, P. CHEVALIER, "Widely linear estimation with complex data", *IEEE Trans. Signal Processing*, Vol 43, N°8, pp. 2030-2033, Aug. 1995.

[31] J.G. PROAKIS, "Digital communications", *McGraw-Hill*, Third Edition, 1995.

[32] P. VIRAVEAU, P. GOURNAY, "Corrélation spectrale théorique des modulations CPM - Partie II: Méthode de calcul générale et analyse", *Annales des Télécommunications*, Vol 53, N°7-8, pp 279-288, Jui./Aout 1998.

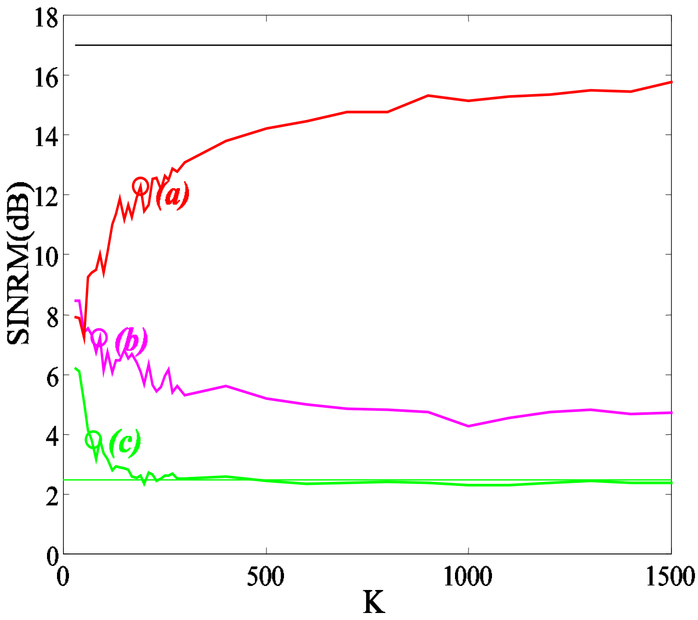

Figure 1 – *SINRM1 at the output of SOBI_COR, SOBI_ACOV and SOBI_COV as a function of K, N = 5, P = 2 2-CPFSK sources,* $\theta_1 = 50°$, $\theta_2 = 91°$, *SNR = 10 dB,* $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$, $\Delta f_1 = \Delta f_2 = h_1/2T_1$, Source1 : SOBI_COV (a), SOBI_ACOV (b), SOBI_COR (c)

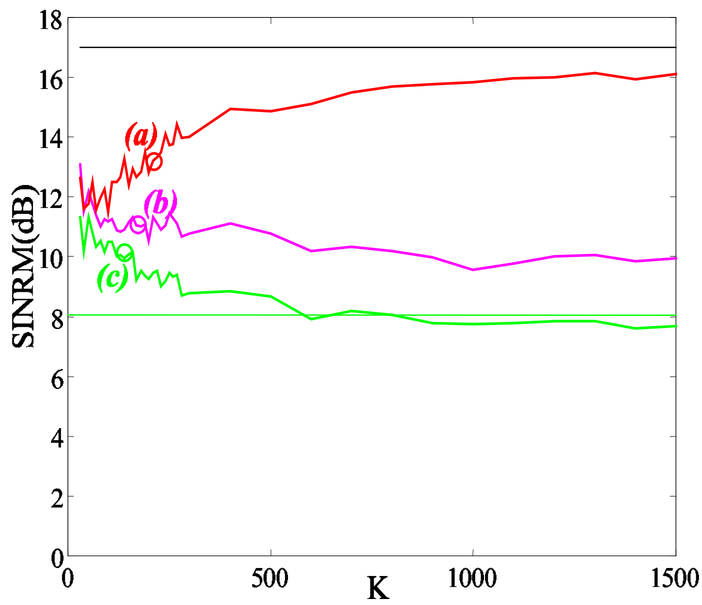

Figure 2 – *SINRM2 at the output of SOBI_COR, SOBI_ACOV and SOBI_COV as a function of K, N = 5, P = 2 2-CPFSK sources,* $\theta_1 = 50°$, $\theta_2 = 91°$, *SNR = 10 dB,* $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$, $\Delta f_1 = \Delta f_2 = h_1/2T_1$, Source1 : SOBI_COV (a), SOBI_ACOV (b), SOBI_COR (c)

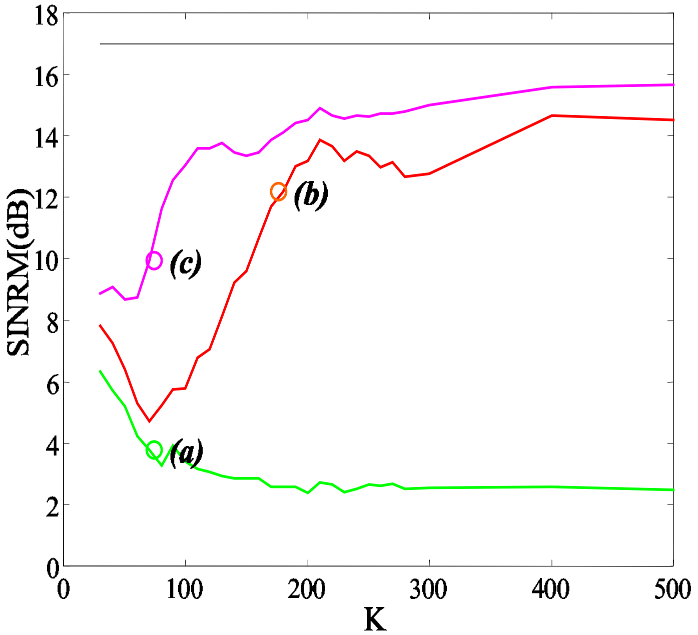

Figure 3 – *SINRM1 at the output of SOBI_COR as a function of K, N = 5, P = 2 2-CPFSK sources, $\theta_1 = 50°$, $\theta_2 = 91°$, SNR = 10 dB, $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$,* $\Delta f_1 = h_1/2T_1$, $(\Delta f_1 - \Delta f_2)\,xT_e = 0$ (a), 0.005 (b), 0.01 (c)

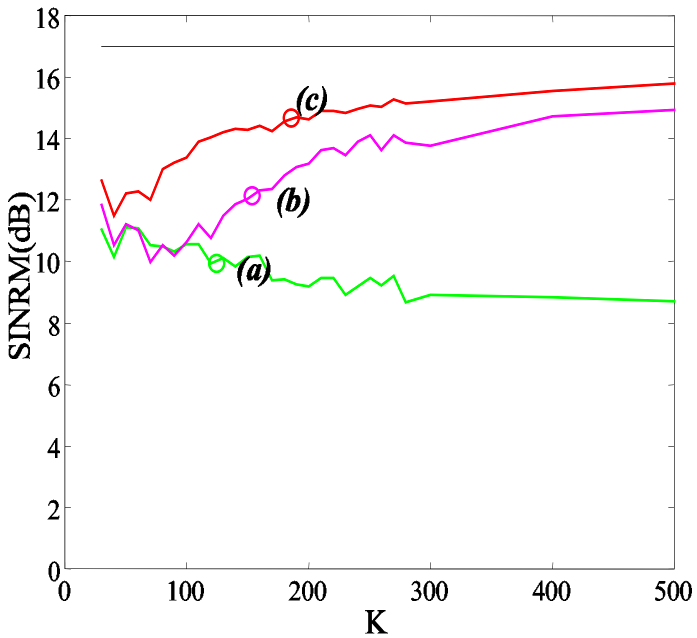

Figure 4 – *SINRM2 at the output of SOBI_COR as a function of K, N = 5, P = 2 2-CPFSK sources, $\theta_1 = 50°$, $\theta_2 = 91°$, SNR = 10 dB, $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$,* $\Delta f_1 = h_1/2T_1$, $(\Delta f_1 - \Delta f_2)\,xT_e = 0$ (a), 0.005 (b), 0.01 (c)

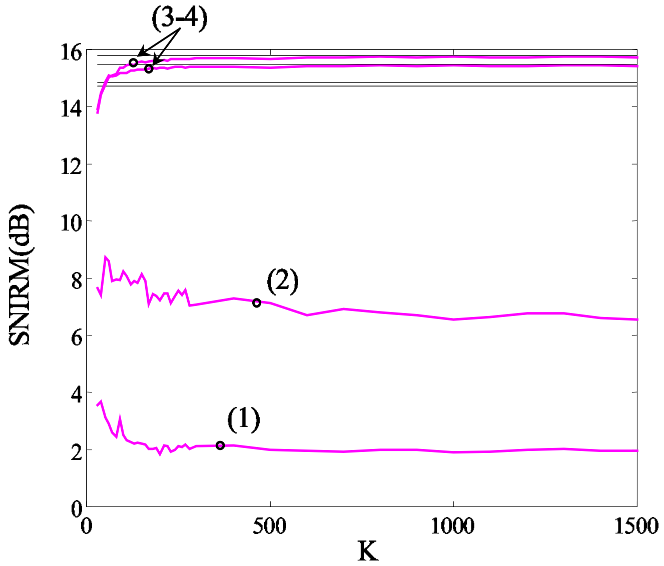

Figure 5 – *SINRMi ($1 \leq i \leq 4$) at the output of SOBI_COR as a function of K, N = 5, P = 4 : 2 2-CPFSK sources and 2 sinusoïds,* $\theta_1 = 50°$, $\theta_2 = -179°$, $\theta_3 = 125°$, $\theta_4 = 93°$, *SNR = 10 dB,* $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$, $\Delta f_1 = \Delta f_2 = h_1/2T_1$, $\Delta f_3 = 1/3T_e$, $\Delta f_4 = 1/5T_e$

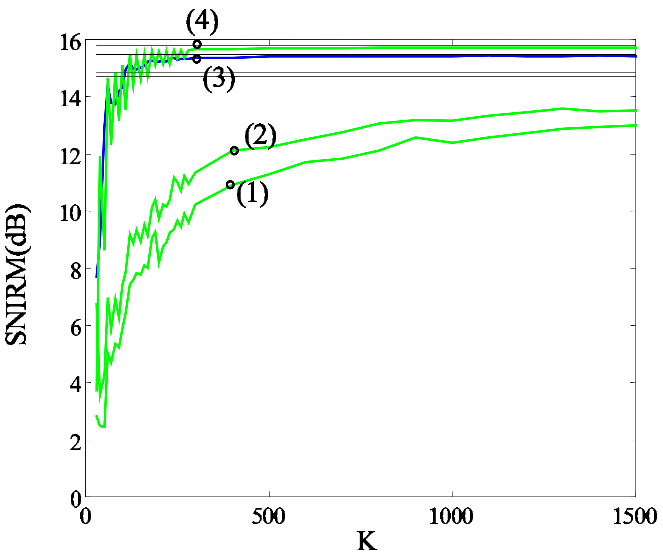

Figure 6 – *SINRMi ($1 \leq i \leq 4$) at the output of the SOBEFOCYS method as a function of K, N = 5, P = 4 : 2 2-CPFSK sources and 2 sinusoïds,* $\theta_1 = 50°$, $\theta_2 = -179°$, $\theta_3 = 125°$, $\theta_4 = 93°$, *SNR = 10 dB,* $h_1/T_1 = h_2/T_2 = 1/4T_e$, $h_1 = 2$, $h_2 = 4$, $\tau = 4T_e$, $\Delta f_1 = \Delta f_2 = h_1/2T_1$, $\Delta f_3 = 1/3T_e$, $\Delta f_4 = 1/5T_e$

BIOGRAPHIES

Anne Férréol

Anne Férréol was born in 1964 in Lyon, France. She received the M. Sc degree from ICPI-Lyon and the Mastère degree from Ecole Nationale Supérieure des Télécommunications (ENST) in 1988 and 1989 respectively.

Since 1989 she has been working at Thomson-CSF-Communications, in the array processing department. She is currently preparing for the PH.D. degree from Ecole Normale Supérieure de Cachan, France, in collaboration with both SATIE Laboratory and THALES Communications.

Her current interests concern direction finding and blind source separation.

Pascal Chevalier

Pascal Chevalier was born in 1962 in Valenciennes, France. He received the M.Sc degree from Ecole Nationale Supérieure des Techniques Avancées (ENSTA) and the Ph.D. degree from South-Paris University, France, in 1985 and 1991 respectively.

Since 1991 he has shared industrial activities (studies, experimentations, expertises, management), teaching activities both in French engineer schools (Supelec, ENST, ENSTA) and French Universities (Cergy-Pontoise) and research activities. Since 2000, he has also been acting as Technical Manager and Architect of the array processing sub-system as part of a national program of military satellite telecommunications.

His present research interests are in array processing techniques, either blind or informed, second order or higher order, spatial-or spatio-temporal, Time-Invariant or Time-Varying especially for cyclostationary signals, linear or non linear and particularly widely linear for non circular signals, for applications such as TDMA and CDMA radiocommunications networks, satellite telecommunications, spectrum monitoring and HF/VUHF passive listening.

Dr Chevalier has been a member of the THOMSON-CSF Technical and Scientifical Council. He is author or co-author of about 100 papers (Journal, Conferences, Patents and Chapters of books).

Dr. Chevalier is presently an EURASIP member and a senior member of the Societé des Electriciens et des Electroniciens (SEE).

Laurent Albera

Laurent Albera was born in Massy, France, in 1976. He received from University of Science of Orsay (Paris XI), France, the M.S. degree both in Mathematical Engineering in 2000 and in Signal Processing Engineering in 2001. Currently, he's preparing for the PH.D. degree from University of Science of Nice Sophia-Antipolis, France, in collaboration with both I3S Laboratory and THALES Communications. Since 2000, he has currently worked in Blind Source Separation and Independent Component Analysis (ICA). His research interests especially include both the cyclostationary source case and the underdetermined mixture problem.

Appendix H

HO BSS of non zero-mean cyclostationary sources

HIGHER ORDER BLIND SEPARATION OF NON ZERO-MEAN CYCLOSTATIONARY SOURCES

Anne Ferréol, Pascal Chevalier and Laurent Albera
THALES COMMUNICATIONS – BSR/TBR/TSI/LTC
66 rue du Fossé Blanc,
92231 Gennevilliers, France
Tel : +33 01 46 13 23 29; fax : +33 01 46 13 25 55
e-mail : anne.ferreol@fr.thalesgroup.com

ABSTRACT

Most of the current Second Order (SO) and Higher Order (HO) blind source separation (BSS) methods aim at blindly separating statistically independent sources, assumed zero-mean, stationary and ergodic. However in practical situations, such as in radiocommunications contexts, the sources are non stationary and very often cyclostationary. In a previous paper [3] the cumulant-based BSS problem for cyclostationary sources has been analysed assuming zero-mean sources (linear modulations). Then we have considered in [4] the non zero-mean cyclostationary case for current SO BSS methods such as the SOBI method. The purpose of this paper is to analyse the behavior and to propose adaptations of the current HO BSS methods for cyclostationary sources assumed cyclo-ergodic and non zero-mean sources.

1 INTRODUCTION

For more than a decade, SO and HO [1] [2] blind methods have been developed to separate several statistically independent sources, assumed *zero-mean*, *stationary* and *ergodic*. However, in many applications such as in radiocommunications contexts, the sources are non stationary and very often *cyclostationary* (digital modulations). It then becomes important to analyse the behavior of the current SO and HO BSS methods in cyclostationary contexts.

In a previous paper [3], the behavior of the current SO and Fourth-order (FO) cumulant-based BSS methods has been analysed for cyclostationary sources assumed zero-mean sources. It has been shown in particular that the current SO blind methods are not affected by the cyclostationarity of the sources whereas the current FO blind methods may be strongly affected by this property.

Nevertheless, some cyclostationary sources used in practice are not zero-mean but are first order cyclostationary, which is in particular the case for some non linearly modulated digital sources. For this reason, in a recent paper [4], we have analysed the behavior and proposed adaptations of the current SO BSS methods in the presence of statistically independent sources which are both first order and SO cyclostationary.

Thus extending the analysis to the blind separators exploiting both the SO and the FO cumulants of the data such as the JADE separator, our goal in this paper is precisely to bring some answers to the important HO BSS problem in the presence of non zero-mean cyclostationary sources.

2 PROBLEM FORMULATION

A noisy mixture of P narrow-band (NB) statistically independent sources is assumed to be received by an array of N sensors. The vector, $\boldsymbol{x}(t)$, of the complex envelopes of the signals at the output of the sensors is thus given by

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} m_{pc}(t)\, \boldsymbol{a}_p + \boldsymbol{b}(t) \triangleq A\, \boldsymbol{m}_c(t) + \boldsymbol{b}(t) \qquad (1)$$

where $m_{pc}(t) = m_p(t)e^{j2\pi\Delta f_p t + \phi_p}$ is the p-th component of the vector $\boldsymbol{m}_c(t)$, $m_p(t)$, Δf_p, ϕ_p and $\boldsymbol{a}_p$ correspond to the complex envelope, the carrier residu, the phase and the steering vector of the source p respectively, A is the (NxP) matrix whose columns are the vectors $\boldsymbol{a}_p$. The $\boldsymbol{b}(t)$ noise vector, assumed zero-mean, is normally distributed, spatially white and independent from the sources in the reception band.

The classical HO blind source separation problem consists to find, from the HO statistics of the observations, the (NxP) *Linear* and *Time Invariant* source separator W, whose (Px1) output vector $\boldsymbol{y}(t) \triangleq W^H \boldsymbol{x}(t)$ corresponds, to within a diagonal matrix Λ and a permutation matrix Π, to the best estimate, $\hat{\boldsymbol{m}}_c(t)$, of the vector $\boldsymbol{m}_c(t)$.

3 HO BLIND SOURCE SEPARATION FOR ZERO MEAN STATIONARY SOURCES

3.1 Statistics of the data

Let $M_x^{\varepsilon(\nu_n)}[\nu_n](t)$ and $Cum_x^{\varepsilon(\nu_n)}[\nu_n](t)$ be the moments and cumulants of order n of $\boldsymbol{x}(t)$, given by

$$M_x^{\varepsilon(\nu_n)}[\nu_n](t) \triangleq \mathrm{E}[x_{i_1}^{\varepsilon_1}(t)\, x_{i_2}^{\varepsilon_2}(t) \ldots x_{i_n}^{\varepsilon_n}(t)] \qquad (2)$$

$$Cum_x^{\varepsilon(\nu_n)}[\nu_n](t) \triangleq \sum_{q=1}^{n} (-1)^{q-1}(q-1)! \sum_{w=1}^{W_q} \prod_{r=1}^{q} M_x^{\varepsilon\left(S_{r,w}^q\right)}\left[S_{r,w}^q\right] \qquad (3)$$

where $\varepsilon(\nu_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ with $\varepsilon_i \in \{-1, 1\}$, such that $x_i^{-1}(t) = x_i^*(t)$ and $x_i^1(t) = x_i(t)$, where $\nu_n = (i_1, i_2, \ldots, i_n)$ with

$1 \le i_j \le N$, and $S^q_{1,w}$ U...U $S^q_{q,w}$ is the w^{th} partition of the set ν_n. Moments and cumulants of order n may be seen as tensors of order n, however, it is interesting to put them into matrices. In this way, let d = Int(n/2) where the function Int(.) rounds to the nearest lower integer. Let $M_x^{\varepsilon(\nu_n)}(t)$ and $Cum_x^{\varepsilon(\nu_n)}(t)$ be the (N^dxN^{n-d}) matrices whose $[(N(N(i_1 - 1) + i_2 - 1) + i_3 \ldots) + i_d, (N(N(i_{d+1} - 1) + i_{d+2} - 1) + i_{d+3} \ldots) + i_n]$ element is equal to $M_x^{\varepsilon(\nu_n)}[i_1, i_2, \ldots, i_d, \ldots, i_n](t)$ and $Cum_x^{\varepsilon(\nu_n)}[i_1, i_2, \ldots, i_d, \ldots, i_n](t)$ respectively. Using (1) into (2) and (3) we obtain

$$Cum_x^{\varepsilon(\nu_n)}(t) = (A^{\varepsilon_1} \otimes \ldots \otimes A^{\varepsilon_d})\, Cum_{mc}^{\varepsilon(\mu_n)}(t)\, (A^{\varepsilon_{d+1}} \otimes \ldots \otimes A^{\varepsilon_n})^{T} + Cum_b^{\varepsilon(\nu_n)}(t) \quad (4)$$

where $Cum_{mc}^{\varepsilon(\mu_n)}(t)$ is the cumulant matrix of $\boldsymbol{m}_c(t)$, $\otimes$ is the Kronecker product and where $\mu_n = (i_1, i_2, \ldots, i_n)$ with $1 \le i_j \le P$. We obtain the same equation for the moment matrices replacing $Cum_{(.)}^{\varepsilon(.)}(t)$ by $M_{(.)}^{\varepsilon(.)}(t)$. Note that for a normally distributed noise, $Cum_b^{\varepsilon(\nu_n)}(t)$ is zero as soon as we have n>2. Under the assumption of stationarity, the statistics $M_{(.)}^{\varepsilon(.)}[i_1, \ldots, i_n](t)$, $Cum_{(.)}^{\varepsilon(.)}[i_1, \ldots, i_n](t)$ and the associated matrices $M_{(.)}^{\varepsilon(.)}(t)$, $Cum_{(.)}^{\varepsilon(.)}(t)$ aren't Time Dependent (TD). So they can be named $M_{(.)}^{\varepsilon(.)}[i_1, \ldots, i_n]$, $Cum_{(.)}^{\varepsilon(.)}[i_1, \ldots, i_n]$, $M_{(.)}^{\varepsilon(.)}$ and $Cum_{(.)}^{\varepsilon(.)}$ respectively.

3.2 Philosophy of the JADE method

Under the previous assumptions, assuming no Gaussian sources for simplicity, the current JADE method [1] aims at separating the sources from both the SO and the FO blind identification of the A matrix. This requires the prewhitening of the data thanks to the knowledge of the correlation matrix $R_x^{(1,-1)} \triangleq M_x^{(1,-1)}$, which for zero mean stationary sources, is equal to the covariance matrix $R_{\Delta x}^{(1,-1)} \triangleq Cum_x^{(1,-1)}$. This prewhitening orthonormalizes the sources steering vectors so as to search for the latter through an unitary (PxP) matrix U simpler to handle. Let Q_{xa} be the apparent quadricovariance matrix of $\boldsymbol{x}(t)$, defined by

$$Q_{xa}[i, j, k, l] \triangleq M_x^{(1,-1,-1,1)}[i, j, k, l] - \sum_{w=1}^{3} \left[R_x^{\varepsilon_{w,12}}[i_1[w], i_2[w]] \quad R_x^{\varepsilon_{w,34}}[i_3[w], i_4[w]] \right] \quad (5)$$

where $\boldsymbol{i}_1 = [i, k, l, j]^T$, $\boldsymbol{i}_2 = [j, i, i, k]^T$, $\boldsymbol{i}_3 = [k, j, j, l]^T$, $\boldsymbol{i}_4 = [l, l, k, i]^T$, $\varepsilon_{w,qr\ldots s} = \varepsilon(i_q[w], i_r[w], \ldots, i_s[w])$ with $\varepsilon(i, j, k, l) = (1, -1, -1, 1)$ and $R_x^{\varepsilon(q,r)}[q, r] \triangleq M_x^{\varepsilon(q,r)}[q, r]$. Note that for zero mean stationary sources, Q_{xa} is equal to the Q_x quadricovariance matrix, given by $Q_x[i, j, k, l] = Cum_x^{(1,-1,-1,1)}[i, j, k, l]$. If we note $\boldsymbol{z}(t)$ the whitened observation vector, the matrix U is chosen so as to jointly diagonalize the P eigenmatrices V_p constructed from the orthonormalized eigenvectors associated to the P signal eigenvalues of the Q_{za} apparent quadricovariance matrix.

$$Q_{za} = (A' \otimes A'^{*})\, Q_{mc'a}\, (A' \otimes A'^{*})^{H} \quad (6)$$

where A' is the (PxP) unitary matrix of the whitened sources steering vectors, $Q_{mc'a}$ corresponds to the apparent quadricovariance of $\boldsymbol{m}_c(t)$, the normalized vector $\boldsymbol{m}_c(t)$ such that each component has a unit power. Under these conditions, it is easy to verify [1] that the unitary matrix A' is, to within a permutation and an unitary diagonal matrix, the only one which jointly diagonalizes the P matrices V_p.

3.3 Implementation of the JADE method

In practical situations, the SO and FO statistics of the data have to be estimated, by temporal averaging operations, using the ergodicity property of the data. Under these assumptions, noting T_e the sample period and $\boldsymbol{x}(t_k)$ the k-th sample of the observation vector $\boldsymbol{x}(t)$, the empirical estimator $\hat{M}_x^{\varepsilon(\nu_n)}[\nu_n]$ of $M_x^{\varepsilon(\nu_n)}[\nu_n]$, from K independent data snapshots, is defined by

$$\hat{M}_x^{\varepsilon(\nu_n)}[\nu_n] \triangleq \frac{1}{K} \sum_{k=1}^{K} x_{i_1}^{\varepsilon_1}(t_k)\, x_{i_2}^{\varepsilon_2}(t_k) \ldots x_{i_n}^{\varepsilon_n}(t_k) \quad (7)$$

So we deduce the empirical estimator $\hat{R}_x^{(1,-1)}[i, j] \triangleq \hat{M}_x^{(1,-1)}[i, j]$ of the correlation $R_x^{(1,-1)}[i, j] \triangleq M_x^{(1,-1)}[i, j]$. In the same way, the empirical estimator $\hat{Q}_{xa}[i, j, k, l]$ of the apparent quadricovariance $Q_{xa}[i, j, k, l]$ is given estimating in (5) the different moments by (7). It is well known that for stationary and ergodic observations, the empirical estimators $\hat{R}_x^{(1,-1)}$ and $\hat{Q}_{xa}$ are, as K becomes infinite, unbiased and consistent.

4 HO BSS FOR NON ZERO-MEAN CYCLOSTATIONARY SOURCES

4.1 Statistics of the data

We now assume that the sources are cyclostationary, which means that their statistics are (quasi)-periodic functions of the time. Thus, the statistics $M_x^{\varepsilon(\nu_n)}[\nu_n](t)$, $Cum_x^{\varepsilon(\nu_n)}[\nu_n](t)$ and the associated matrices $M_x^{\varepsilon(\nu_n)}(t)$, $Cum_x^{\varepsilon(\nu_n)}(t)$ become TD and have a Fourier serial expansion. Moreover, we define the set of the moment cyclic frequencies, $\Gamma_x^{\varepsilon(\nu_n)} = \mathrm{U}_{\mu_n}\{\Gamma_{mc}^{\varepsilon(\mu_n)}\}$, and the set of the cumulant cyclic frequencies, $\Phi_x^{\varepsilon(\nu_n)} = \mathrm{U}_{1 \le p \le P}\{\Phi_{mc(p)}^{\varepsilon(\mu_n)}\}$, of order n of $\boldsymbol{x}(t)$, from

$$\Gamma_{mc}^{\varepsilon(\mu_n)} \triangleq \{\gamma \,/\, 1 \le p_j \le P,\ \mu_n = (p_1, p_2, \ldots, p_n),\ M_{mc}^{\varepsilon(\mu_n)}[\nu_n](\gamma) \neq 0\} \quad (8)$$

$$\Phi_{mc(p)}^{\varepsilon(\mu_n)} \triangleq \{\varphi \,/\, \mu_n = (p, p, \ldots, p) \text{ and } Cum_{mc}^{\varepsilon(\mu_n)}[\nu_n](\varphi) \neq 0\} \quad (9)$$

where $M_{mc}^{\varepsilon(\mu_n)}[\mu_n](\gamma) \triangleq \langle M_{mc}^{\varepsilon(\mu_n)}[\mu_n](t)\, e^{-j2\pi\gamma t} \rangle_c$ and $Cum_{mc}^{\varepsilon(\mu_n)}[\nu_n](\varphi) \triangleq \langle Cum_{mc}^{\varepsilon(\mu_n)}[\nu_n](t)\, e^{-j2\pi\varphi t} \rangle_c$ are called the cyclic moment and the cyclic cumulant of order n of $\boldsymbol{m}_c(t)$, noting that $\langle\rangle_c$ is the continuous-time temporal mean operation. We can adopt the following notations

$$e_x^{\varepsilon(q)}[q](\gamma) \triangleq M_x^{\varepsilon(q)}[q](\gamma)\ ;\ R_x^{\varepsilon(q,r)}[q, r](\gamma) \triangleq M_x^{\varepsilon(q,r)}[q, r](\gamma) \quad (10)$$

$$R_{\Delta x}^{\varepsilon(q,r)}[q, r](\varphi) \triangleq Cum_x^{\varepsilon(q,r)}[q, r](\varphi) = R_x^{\varepsilon(q,r)}[q, r](\varphi) - \sum_{\alpha \in \Gamma_x} e_x^{\varepsilon(q)}[q](\alpha)\, e_x^{\varepsilon(r)}[r](\varphi - \alpha) \quad (11)$$

Note that the cyclic moments and cyclic cumulants for the zero cyclic frequency correspond to the temporal mean, $M_x^{\varepsilon(\nu_n)}[\nu_n] \triangleq \langle M_x^{\varepsilon(\nu_n)}[\nu_n](t) \rangle_c$ and $Cum_x^{\varepsilon(\nu_n)}[\nu_n] \triangleq \langle Cum_x^{\varepsilon(\nu_n)}[\nu_n](t) \rangle_c$, of $M_x^{\varepsilon(\nu_n)}[\nu_n](t)$ and $Cum_x^{\varepsilon(\nu_n)}[\nu_n](t)$

respectively. Let put the components $M_x^{\varepsilon(\nu_n)}[\nu_n]$ and $Cum_x^{\varepsilon(\nu_n)}[\nu_n]$ into the matrices $M_x^{\varepsilon(\nu_n)}$ and $Cum_x^{\varepsilon(\nu_n)}$ respectively. Using (4), we can link on the one hand $M_x^{\varepsilon(\nu_n)}$ and $M_{mc}^{\varepsilon(\mu_n)}$, on the other hand $Cum_x^{\varepsilon(\nu_n)}$ and $Cum_{mc}^{\varepsilon(\mu_n)}$.

4.2 Behavior analysis of the statistics empirical estimators

For cyclostationary sources, HO BSS methods such as the JADE method have to exploit the information contained in the time-averaged statistics $R_x^{(1,-1)}$ and Q_{xa}, estimated from (7). For band-limited, cyclo-ergodic and sufficiently oversampled observations, the empirical estimators, $\hat{R}_x^{(1,-1)}$ and $\hat{Q}_{xa}$, of $R_x^{(1,-1)}$ and Q_{xa} respectively, are asymptotically unbiased and consistent.

However, while for zero-mean stationary independent sources, the correlation matrix $R_{mc}^{(1,-1)} \triangleq M_{mc}^{(1,-1)}$ and the apparent quadricovariance matrix Q_{mca}, defined from (5), are diagonal and equal to $Cum_{mc}^{(1,-1)}$, $Cum_{mc}^{(1,-1,-1,1)}$ respectively, it is not necessary the case for non zero-mean cyclostationary independent sources for which only the covariance matrix $R_{\Delta mc}^{(1,-1)} = Cum_{mc}^{(1,-1)}$ and the quadricovariance matrix $Q_{mc} = Cum_{mc}^{(1,-1,-1,1)}$ are diagonal. As a consequence, as the current HO BSS methods are affected by the cyclostationarity of zero-mean sources, they may also be affected for non zero-mean cyclostationary sources for which, an *apparent statistic dependence* of the sources may appear in both the $R_{mc}^{(1,-1)}$ and Q_{mca} matrices.

4.3 Skew using statistic empirical estimators

According to [4], the non diagonal $[i, j]$ component $R_{mc}^{skew}[i, j] \triangleq R_{\Delta mc}^{(1,-1)}[i,j] - R_{mc}^{(1,-1)}[i,j] = \langle e_{mc}^{1}[i]\, e_{mc}^{-1}[j] \rangle_c$, with $e_{mc}^{\varepsilon(q)}[q] = M_{mc}^{\varepsilon(q)}[q]$, of the $R_{mc}^{(1,-1)}$ matrix are non zero as soon as the two sources $m_{ic}(t)$ and $m_{jc}(t)$ share at least one first order cyclic moment frequency.

In the same way, the non diagonal $\mu_4 = (i, j, k, l)$ component $Q_{mc}^{skew}[\mu_4] \triangleq Q_{mc}[\mu_4] - Q_{mca}[\mu_4] = Q_{mca}[\mu_4]$ of the Q_{mca} matrix, is given by

$$\begin{aligned} Q_{mc}^{skew}[i,j,k,l] = & -\sum_{w=1}^{4} \sum_{\gamma \in \Gamma_{1,234}^{w}} e_{mc}^{\varepsilon_{w,1}}[i_1[w]](\gamma)\, M_{mc}^{\varepsilon_{w,234}}[i_2[w], i_3[w], i_4[w]](-\gamma) + \sum_{w=1}^{3} \Bigg[R_{mc}^{\varepsilon_{w,12}}[i_1[w], i_2[w]] \quad R_{mc}^{\varepsilon_{w,34}}[i_3[w], i_4[w]] \\ & + \sum_{\gamma \in \Gamma_{12,34}^{w}} R_{mc}^{\varepsilon_{w,12}}[i_1[w], i_2[w]](\gamma)\, R_{mc}^{\varepsilon_{w,34}}[i_3[w], i_4[w]](-\gamma) \\ & -2 \sum_{\varphi \in \Phi_{12,34}^{w}} R_{\Delta mc}^{\varepsilon_{w,12}}[i_1[w], i_2[w]](\varphi)\, R_{\Delta mc}^{\varepsilon_{w,34}}[i_3[w], i_4[w]](-\varphi) \Bigg] \end{aligned} \tag{12}$$

where i_1, i_2, i_3, i_4 and $\varepsilon_{w,qr...s}$ are defined in (5). Moreover, $\Gamma_{1,234}^{w} \triangleq \Gamma_{mc}^{\varepsilon(i_1[w])} \cap \Gamma_{mc}^{\varepsilon(i_2[w], i_3[w], i_4[w])}$ is the set of the common moment cyclic frequencies of the processes $m_{i_1[w]c}(t)$ and $(m_{i_2[w]c}(t), m_{i_3[w]c}(t), m_{i_4[w]c}(t))$. In the same way, $\Gamma_{12,34}^{w} \triangleq \Gamma_{mc}^{\varepsilon(i_1[w], i_2[w])} \cap \Gamma_{mc}^{\varepsilon(i_3[w], i_4[w])}$ is the set of the common moment cyclic frequencies of the processes $(m_{i_1[w]c}(t), m_{i_2[w]c}(t))$ and $(m_{i_3[w]c}(t), m_{i_4[w]c}(t))$. To have done with the notations, $\Phi_{12,34}^{w} \triangleq \Phi_{mc}^{\varepsilon(i_1[w], i_2[w])} \cap \Phi_{mc}^{\varepsilon(i_3[w], i_4[w])}$ is the set of the common cumulant cyclic frequencies of the processes $(m_{i_1[w]c}(t), m_{i_2[w]c}(t))$ and $(m_{i_3[w]c}(t), m_{i_4[w]c}(t))$. Whereas the non diagonal components $Q_{mc}^{skew}[\nu_4]$ of the Q_{mca} matrix are zero for zero-mean stationary independent sources, it is not necessary the case for non zero-mean cyclostationary independent sources, in particular when at least one of the sets $\Gamma_{1,234}^{w}$, $\Gamma_{12,34}^{w}$ or $\Phi_{12,34}^{w}$ isn't empty. For instance, noting that for independent sources, $R_{\Delta mc}^{\varepsilon(q,r)}[q, r](\varphi)$ is non zero if and only if $q = r$, the set $\Phi_{12,34}^{w}$ is non empty as soon as the two sources $m_{i_1[w]c}(t)$ and $m_{i_3[w]c}(t)$ share at least one second order cumulant cyclic frequency.

4.4 Behavior of the JADE method

While, for apparently SO uncorrelated sources, which is in particular the case for zero-mean sources, the whitened mixed matrix A' is an unitary matrix, it is no longer the case for apparently SO correlated sources, for which the vectors $\boldsymbol{a}_p'$ are neither normalized nor orthogonal, as it is shown in [4].

Moreover, while, for apparently FO uncorrelated sources, which is in particular the case for zero-mean stationary sources, the P eigenmatrices V_p of the Q_{za} apparent quadricovariance matrix may be written as $V_p = A'\, D_p\, A'^{H}$ where D_p are diagonal matrices, it is no longer the case for apparently FO correlated sources.

A consequence of these results is that the matrix A' does not jointly diagonalizes the set of P eigenmatrices V_p. In other words, the blindly identified source steering vectors are only a linear combination of the source steering vectors, which shows that the JADE method as well as the HO BSS methods are affected by the presence of apparently SO and FO correlated sources.

4.5 Adaptation : exhaustive estimators

Since the matrices $R_{mc}^{(1,-1)}$ and Q_{mca} may be non diagonal in the presence of non zero-mean cyclostationary independent sources, we have to exploit the information contained in the matrices $R_{\Delta mc}^{(1,-1)} = Cum_{mc}^{(1,-1)}$ and $Q_{mc} = Cum_{mc}^{(1,-1,-1,1)}$ which are always diagonal for statistically independent sources, zero-mean or not. In other words, we have to implement the HO BSS methods from the matrix $R_{\Delta x}^{(1,-1)}$ and Q_x defined by

$$R_{\Delta x}^{(1,-1)} \triangleq \langle Cum_x^{(1,-1)}(t) \rangle_c \text{ and } Q_x \triangleq \langle Cum_x^{(1,-1,-1,1)}(t) \rangle_c \tag{13}$$

So, for cyclostationary and band-limited vectors $\boldsymbol{x}(t)$ having a cyclo-ergodicity property and for sufficiently oversampled data, after a preliminary step of first and second order cyclic frequencies estimation [4] [5], we define the asymptotic unbiased and consistent estimators $\hat{R}_{\Delta x}^{(1,-1)}$ and $\hat{Q}_x$ of $R_x^{(1,-1)}$ and Q_x respectively. Whereas $\hat{R}_{\Delta x}^{(1,-1)}$ is given in [4], we have

$$\hat{Q}_x[i,j,k,l] \quad = \quad \hat{M}_x^{(1,-1,-1,1)}[i,j,k,l]$$

$$
-\sum_{w=1}^{4}\sum_{\gamma\in\Gamma_{1,234}^{w}} \hat{e}_x^{\varepsilon_{w,1}}[i_1[w]](\gamma)\,\hat{M}_x^{\varepsilon_{w,234}}[i_2[w],i_3[w],i_4[w]](-\gamma) \quad +
$$

$$
\sum_{w=1}^{3}\Big[\sum_{\gamma\in\Gamma_{12,34}^{w}} \hat{R}_x^{\varepsilon_{w,12}}[i_1[w],i_2[w]](\gamma)\,\hat{R}_x^{\varepsilon_{w,34}}[i_3[w],i_4[w]](-\gamma)
$$

$$
-2\sum_{\varphi\in\Phi_{12,34}^{w}} \hat{R}_{\Delta x}^{\varepsilon_{w,12}}[i_1[w],i_2[w]](\varphi)\,\hat{R}_{\Delta x}^{\varepsilon_{w,34}}[i_3[w],i_4[w]](-\varphi)\Big] \tag{14}
$$

where $\boldsymbol{i}_1$, $\boldsymbol{i}_2$, $\boldsymbol{i}_3$, $\boldsymbol{i}_4$ and $\varepsilon_{w,qr\ldots s}$ are defined in (5), where $\hat{e}_x^{\varepsilon(q)}[q](\gamma) \triangleq \hat{M}_x^{\varepsilon(q)}[q](\gamma)$, $\hat{R}_x^{\varepsilon(q,r)}[q,r](\gamma) \triangleq \hat{M}_x^{\varepsilon(q,r)}[q,r](\gamma)$ are given by

$$
\hat{M}_x^{\varepsilon(\nu_n)}[i_1,\ldots,i_n](\gamma) \triangleq \frac{1}{K}\sum_{k=1}^{K} x_{i_1}^{\varepsilon_1}(t_k)x_{i_2}^{\varepsilon_2}(t_k)\ldots x_{i_n}^{\varepsilon_n}(t_k)e^{-j2\pi f T_e \gamma} \tag{15}
$$

and, according to (12),

$$
\hat{R}_{\Delta x}^{\varepsilon(q,r)}[q,r](\varphi) \triangleq \hat{R}_x^{\varepsilon(q,r)}[q,r](\varphi) - \sum_{\alpha\in\Gamma_{q,r}} \hat{e}_x^{\varepsilon(q)}[q](\alpha)\,\hat{e}_x^{\varepsilon(r)}[r](\varphi-\alpha) \tag{16}
$$

To simplify the implementation of the exhaustive estimator $\hat{Q}_x$, we may take $\Gamma_{1,234}^{w} = \Gamma_x^{\varepsilon(i_1[w])}$, $\Gamma_{12,34}^{w} \triangleq \Gamma_x^{\varepsilon(i_1[w],i_2[w])}$, $\Phi_{12,34}^{w} \triangleq \{\Gamma_x^{\varepsilon(i_1[w],i_2[w])}\} \cup \{\Gamma_x^{\varepsilon(i_1[w])} \oplus \Gamma_x^{\varepsilon(i_2[w])}\}$, $\Gamma_{1,2}^{w} = \Gamma_x^{\varepsilon(i_1[w])}$ and $\Gamma_{3,4}^{w} = \Gamma_x^{\varepsilon(i_3[w])}$, where

$$
\Gamma_x^{\varepsilon(q)} \oplus \Gamma_x^{\varepsilon(r)} \triangleq \{\gamma = \varepsilon(q)\alpha + \varepsilon(r)\beta \;/\; \alpha\in\Gamma_x^{\varepsilon(q)} \text{ and } \beta\in\Gamma_x^{\varepsilon(r)}\} \tag{17}
$$

5 SIMULATIONS

To illustrate the previous results, we assume that two statistically independent NB and orthogonal ($A^H A = N\,\mathrm{I}$) 2-CPFSK sources are received by an array of N=5 sensors. These two sources have the same input SNR (Signal Noise Ratio) of 10 dB and are synchronized. Their symbol durations and their modulation indices are such that $h_1/T_1=h_2/T_2=(4T_e)^{-1}$ for h_1=2 and h_2=4. Thus, we apply the JADE method and average the SINRMk (Maximal Signal to Interference plus Noise Ratio of the k^{th} source, defined in [4], at the output of the JADE separator for k=1,2, over 200 realizations.

Under the previous assumptions, the figure 1 shows the variations of the SINRM1 of the first source at the output of the JADE separator, implemented from four groups of estimators associating SO and FO, empirical or exhaustive, statistic estimators, as a function of the number of snapshots K. Taking zero carrier frequencies, $\Delta f_1=\Delta f_2=0$, we obtain $\Gamma_x^1 = \{(-h_1)/2T_1,\ h_1/2T_1\}$, $\Gamma_x^{(1,-1)} = \{0\}$ and $\Gamma_x^{(1,1)} = \{0, \pm h_1/2T_1\}$ such that the two sources are apparently SO and FO correlated. As planned, the figure 1 shows the poor separation of the sources when the JADE method uses the empirical estimators. On the contrary, the exhaustive estimators using the cyclic frequencies allows the separation of the two 2-CPFSK sources. Moreover, we can note the non zero performances of the separator using both the SO exhaustive and FO empirical estimators.

6 CONCLUSION

In this paper, we showed that the current HO BSS methods, such as the JADE method, may be affected by the presence of statistically independent NB sources which are non zero-mean cyclostationary. This problem is directly related to the fact that the current HO BSS methods aim at exploiting the information contained in the temporal mean of the correlation and apparent quadricovariance matrices instead of the covariance and quadricovariance matrices.

To solve this problem, we must exploit the information contained in the temporal mean of the covariance and quadricovariance matrices of the observations. Thus, we have introduced an unbiased and consistent estimator of these matrices for non zero-mean cyclostationary observations, assuming the first and second order cyclic frequencies have been estimated previously.

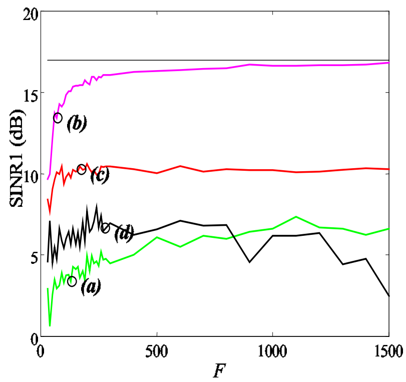

Fig.1 - SINRM1 as a function of F, (a) empirical estimators, (b) exhaustive estimators, (c) SO exhaustive and FO empirical estimators, (d) SO empirical and FO exhaustive estimators.

7 REFERENCES

[1] J.F. Cardoso, A. Souloumiac, "Blind beamforming for non-gaussian signals", *IEE Proceedings-F*, Vol.140, N°6, pp. 362-370, Dec. 1993.

[2] P. Comon, "Independent component analysis – a new concept ?", *Sig. Proc.*, Elsevier, Vol.36, N°3, pp. 287-314, Ap. 1994.

[3] A. Ferréol, P. Chevalier, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources", *IEEE Trans. Sig. Proc.*, Vol.48, N°6, pp. 1712-1725, June 2000.

[4] A. Ferréol, P. Chevalier, L. Albera, "Behavior and adpatation of current second order blind source separation methods for first and second order cyclostationary sources – applications to CPFSK sources", *submitted to IEEE Trans. Signal Processing*

[5] S. V. Schell et W. Gardner, "Detection of the number of cyclostationary signals in unknows interference and noise", *Proc. of Asilomar conf. on signal, systems and computers*, 5-7 Nov. 1990.

Bibliography

[1] F. ASANO, S. IKEDA, M. OGAWA, H. ASOH, and N. KITAWAKI, “Combined approach of array processing and independent component analysis for blind separation of acoustic signals,” *IEEE Transactions On Speech and Audio Processing*, vol. 11, no. 3, pp. 204–215, may 2003.

[2] A. BELOUCHRANI, K. ABED-MERAIM, J.-F. CARDOSO, and E. MOULINES, “A blind source separation technique using second-order statistics,” *IEEE Transactions On Signal Processing*, vol. 45, no. 2, pp. 434–444, February 1997.

[3] E. BINGHAM and H. HYVARINEN, “A fast fixed-point algorithm for independent component analysis of complex valued signals,” *International Journal of Neural Systems*, vol. 10, no. 1, pp. 1–8, 2000.

[4] R. BRO, “Parafac, tutorial and applications,” *Chemom. Intel. Lab. Syst.*, vol. 38, pp. 149–171, 1997.

[5] J.-F. CARDOSO, “Localisation et identification par la quadricovariance,” *Traitement du Signal*, vol. 7, no. 5, June 1990.

[6] ——, “Super-symetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors,” in *ICASSP 91, 1991 IEEE International Conference on Acoustics Speech and Signal Processing*, Toronto, Canada, May 1991, pp. 3109–3112.

[7] ——, “High-order contrasts for independent component analysis,” *Neural Computation*, vol. 11, no. 1, pp. 157–192, January 1999.

[8] J.-F. CARDOSO and A. SOULOUMIAC, “Blind beamforming for non-gaussian signals,” *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, December 1993.

[9] ——, “Jacobi angles for simultaneous diagonalization,” *SIAM Journal Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–164, 1996.

[10] P. CHEVALIER, “Performances des séparateurs aveugles de sources aux ordres supérieurs,” in *GRETSI 95, Quinzième colloque sur le Traitement du Signal et des Images*, vol. 1, Juan-Les-Pins, France, September 1995, pp. 297–300.

[11] ——, “Optimal separation of independent narrow-band sources : Concept and Performances,” *Signal Processing, Elsevier*, vol. 73, pp. 27–47, 1999.

[12] P. CHEVALIER, L. ALBERA, A. FERREOL, and P. COMON, “On the virtual array concept for higher order array processing,” *IEEE Transactions On Signal Processing*, 2003, submitted.

[13] P. CHEVALIER and A. FERREOL, “On the virtual array concept for the fourth-order direction finding problem,” *IEEE Transactions On Signal Processing*, vol. 47, no. 9, pp. 2592–2595, September 1999.

[14] P. CHEVALIER, A. FERREOL, and L. ALBERA, “Méthodologie générale pour la séparation aveugle de sources cyclostationnaires arbitraires — Application à l’écoute passive des radiocommunications,” in *GRETSI 03, Dix-neuvième colloque sur le Traitement du Signal et des Images*, vol. 1, Paris, France, September 8-11 2003, pp. 43–46.

[15] H. H. CHIANG and C. L. NIKIAS, “The esprit algorithm with high order statistics,” in *HOS 89, IEEE Signal Processing Workshop on Higher-Order Statistics*, Vail, June 1989, pp. 163–168.

[16] P. COMON, “Independent Component Analysis, a new concept ?” *Signal Processing, Elsevier*, vol. 36, no. 3, pp. 287–314, April 1994.

[17] ——, “Blind channel identification and extraction of more sources than sensors,” in *SPIE Conference*, San Diego, US, July 19-24 1998, pp. 2–13.

[18] ——, “From source separation to blind equalization, contrast-based approaches,” in *ICISP 01, International Conference on Image and Signal Processing*, Agadir, Morocco, May 3-5 2001, pp. 20–32.

[19] ——, "Tensor decompositions," in *Mathematics in Signal Processing V*, J. G. McWhirter and I. K. Proudler, Eds. Oxford, UK: Clarendon Press, 2002, pp. 1–24.

[20] P. COMON and P. CHEVALIER, Source separation: Models, concepts, algorithms and performance in *Unsupervised Adaptive Filtering, Vol. I, Blind Source Separation*, ser. Wiley interscience. Simon Haykin, 2000.

[21] P. COMON, P. CHEVALIER, and V. CAPDEVIELLE, "Performance of contrast-based blind source separation," in *SPAWC 97, IEEE Sig. Proc. Advances in Wireless Com.*, Paris, France, April 16-18 1997, pp. 345–348.

[22] P. COMON and G. H. GOLUB, "Tracking of a few extreme singular values and vectors in signal processing," *Proceedings of the IEEE*, vol. 78, no. 8, pp. 1327–1343, August 1990.

[23] P. COMON and O. GRELLIER, "Non-linear inversion of underdetermined mixtures," in *ICA 99, Second International Symposium on Independent Component Analysis and Blind Signal Separation*, Aussois, France, January 11-15 1999, pp. 461–465.

[24] P. COMON and J. L. LACOUME, "Statistiques d'ordres supérieurs pour le traitement du signal," Ecole Predoctorale de Physique, Les Houches, 30 aout – 10 septembre 1993, p. Flandrin et J. L. Lacoume ed.

[25] P. COMON and E. MOREAU, "Improved contrast dedicated to blind separation in communications," in *ICASSP 97, 1997 IEEE International Conference on Acoustics Speech and Signal Processing*, Munich, April 20-24 1997, pp. 3453–3456.

[26] R. T. COMPTON, *Adaptive Antennas - Concepts and Performance*. Prentice Hall, Englewood Cliffs, New Jersey, 07632, 1988.

[27] L. De LATHAUWER, "The canonical decomposition and blind identification with more inputs than outputs: some algebraic results," in *ICA 03, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, April 1-4 2003, pp. 781–784.

[28] ——, "Simultaneous matrix diagonalization: the overcomplete case," in *ICA 03, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, April 1-4 2003, pp. 821–825.

[29] L. De LATHAUWER, P. COMON, B. De MOOR, and J. VANDEWALLE, "Higher-order power method — application in independent component analysis," in *NOLTA 95, International Symposium on Nonlinear Theory and its Applications*, Las Vegas, Nevada, US, December 10-14 1995, pp. 91–96.

[30] ——, "ICA algorithms for 3 sources and 2 sensors," in *HOS 99, IEEE Signal Processing Workshop on Higher-Order Statistics*, Caesarea, Israel, June 14-16 1999, pp. 116–120.

[31] L. De LATHAUWER and B. De MOOR, "From matrix to tensor : Multilinear algebra and signal processing," in *Mathematics in Signal Processing IV*. Oxford, UK: Oxford University Press, 1998, pp. 1–15.

[32] L. De LATHAUWER, B. De MOOR, and J. VANDEWALLE, "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE Transactions on Biomedical Engineering, Special Topic Section on Advances in Statistical Signal Processing for Biomedicine*, vol. 47, no. 5, pp. 567–572, May 2000.

[33] L. De LATHAUWER, B. De MOOR, J. VANDEWALLE, and J.-F. CARDOSO, "Independent component analysis of largely underdetermined mixtures," in *ICA 03, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, April 1-4 2003, pp. 29–33.

[34] N. DELFOSSE and P. LOUBATON, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing, Elsevier*, vol. 45, pp. 59–83, 1995.

[35] J.-P. DELMAS, "Asymptotic performance of second-order algorithms," *IEEE Transactions On Signal Processing*, 2001, submitted.

[36] K. I. DIAMANTARAS, "Blind separation of multiple binary sources using a single linear mixture," in *ICASSP 00, IEEE Int. Conf. Acoust. Speech Sig. Proc.*, Istanbul, Turkey, June 5-9 2000, pp. 2889–2892.

[37] ——, "Blind channel identification based on the geometry of the received signal constellation," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1133–1143, May 2002.

[38] M. DOGAN and J. MENDEL, "Applications of cumulants to array processing — Part I: Aperture extension and array calibration," *IEEE Transactions On Signal Processing*, vol. 43, no. 5, pp. 1200–1216, May 1995.

[39] J. ERIKSSON and V. KOIVUNEN, "Identifiability, separability and uniqueness of linear ica models," *IEEE Transactions On Signal Processing*.

[40] A. FERREOL, L. ALBERA, and P. CHEVALIER, "Fourth order blind identification of underdetermined mixtures of sources (FOBIUM)," in *ICASSP 03, 2003 IEEE International Conference on Acoustics Speech and Signal Processing*, Hong Kong, China, April 6-10 2003, pp. 41–44.

[41] A. FERREOL and P. CHEVALIER, "Higher order blind source separation using cyclostationarity property of the signals," in *ICASSP 97, 1997 IEEE International Conference on Acoustics Speech and Signal Processing*, Munich, Germany, April 21-24 1997, pp. 4061–4064.

[42] ——, "On the behavior of current second and higher order blind source separation methods for cyclostationary sources," *IEEE Transactions On Signal Processing*, vol. 48, no. 6, pp. 1712–1725, June 2000, Errata vol. 50, no. 4, pp. 990, April 2002.

[43] A. FERREOL, P. CHEVALIER, and L. ALBERA, "Second order blind separation of first and second order cyclostationary sources — Application to AM, FSK, CPFSK and deterministic sources," *IEEE Transactions On Signal Processing*, to appear.

[44] ——, "Higher order blind separation of non zero-mean cyclostationary sources," in *EUSIPCO 02, XI European Signal Processing Conference*, vol. 5, Toulouse, France, September 3-6 2002, pp. 103–106.

[45] L. FETY, "Méthodes de traitement d'antenne adpatées aux radiocommunications," Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications (ENST), 1988.

[46] G. H. GOLUB and C. F. V. LOAN, *Matrix computations, second edition*. The Johns Hopkins University Press, Baltimore, MD, 1989.

[47] O. GRELLIER, "Déconvolution et séparation de sources discrètes," Ph.D. dissertation, Nice Sophia-Antipolis University (I3S Laboratory), January 2000.

[48] J. HERAULT, C. JUTTEN, and B. ANS, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *GRETSI 85, Dixième colloque sur le Traitement du Signal et des Images*, Nice, France, Septembre 1985, pp. 1017–1022.

[49] A. HYVARINEN, “Independent component analysis in the presence of gaussian noise by maximizing joint likelihood,” *Neurocomputing*, vol. 22, pp. 49–67, 1998.

[50] A. HYVARINEN, J. KARHUNEN, and P. OJA, *Independent component analysis*, ser. Wiley interscience. Simon Haykin, 2001.

[51] A. HYVARINEN and E. OJA, “A fast fixed-point algorithm for independent component analysis,” *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[52] E. KOFIDIS and P. A. REGALIA, “On the best rank-1 approximation of higher-order super-symmetric tensors,” *SIAM J. Matrix Analysis and Applications*, vol. 23, no. 3, pp. 863–884, 2002.

[53] T. W. LEE, M. S. LEWICKI, M. GIROLAMI, and T. J. SEJNOWSKI, “Blind source separation of more sources than mixtures using overcomplete representations,” *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 87–90, 1999.

[54] M. S. LEWICKI and T. J. SEJNOWSKI, “Learning overcomplete representations,” *Neural Computation*, vol. 12, pp. 337–365, 2000.

[55] P. McCULLAGH, *Tensor Methods in Statistics*. Chapman and Hall, Monographs on Statistics and Applied Probability, 1987.

[56] R. A. MONZINGO and T. W. MILLER, *Introduction to adaptive arrays*. John Wiley and Sons, New York, 1980.

[57] E. MOULINES, J.-F. CARDOSO, and E. GASSIAT, “Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models,” in *ICASSP 97, 1997 IEEE International Conference on Acoustics Speech and Signal Processing*, Munich, Germany, April 21-24 1997, pp. 3617–3620.

[58] B. OTTERSTEN, M. VIBERG, and T. KAILATH, “Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data,” *IEEE Transactions On Signal Processing*, vol. 40, pp. 590–599, March 1992.

[59] B. N. PARLETT, *The Symmetric Eigenvalue Problem*. Prentice Hall, Englewood Cliffs, New Jersey, 1980.

[60] L. PERROS-MEILHAC, E. MOULINES, K. ABED-MERAIM, P. CHEVALIER, and P. DUHAMEL, "Blind identification of multipath channels: A parametric subspace approach," *IEEE Transactions On Signal Processing*, vol. 49, no. 7, pp. 1468–1480, July 2001.

[61] J. G. PROAKIS, *Digital Communications, third edition*. New York : McGrawHill, 1995.

[62] P. A. REGALIA and M. MBOUP, "Properties of some blind equalization criteria in noisy multiuser environments," *IEEE Transactions On Signal Processing*, vol. 49, no. 12, pp. 3112–3122, December 2001.

[63] R. O. SCHMIDT, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Stanford University, 1981.

[64] ——, "Multiple emitter location and signal parameter estimation," *IEEE Transactions On Signal Processing*, vol. 34, no. 3, pp. 276–280, March 1986.

[65] S. M. SPOONER and W. A. GARDNER, "The cumulant theory of cyclostationarity time-series – Part. II : Development and applications," *IEEE Transactions On Signal Processing*, vol. 42, no. 12, pp. 3409–3429, December 1994.

[66] G. W. STEWART, *Introduction to matrix computations*. Academic Press, New York, 1973.

[67] A. TALEB, "Séparation de sources dans des mélanges non linéaires," Ph.D. dissertation, Laboratoire des Images et des Signaux (LIS), September 1999.

[68] ——, "An algorithm for the blind identification of N independent signal with 2 sensors," in *ISSPA 01, sixteenth symposium on signal processing and its applications*, Kuala-Lumpur, Malaysia, August 13-16 2001, pp. 5–8.

[69] A. TALEB and C. JUTTEN, "On underdetermined source separation," in *ICASSP 99, 1999 IEEE International Conference on Acoustics Speech and Signal Processing*, vol. 3, Phoenix, US, May 15-19 1999, pp. 1445–1448.

[70] N. THIRION and E. MOREAU, "New criteria for blind signal separation," in *IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, US, August 2000, pp. 344–348.

[71] D. W. TUFTS and C. D. MELISSINOS, "Simple, effective computation of principal eigenvectors and their eigenvalues and application to high-resolution estimation of frequencies,"

IEEE Transactions On Acoustics, Speech, And Signal Processing, vol. 34, no. 5, pp. 1046–1053, October 1986.

[72] WILKINSON, *The Algebraic Eigenvalue Problem*. Clarendon Press, 1965.

[73] A. YEREDOR, “Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation,” *IEEE Transactions On Signal Processing*, vol. 50, no. 7, pp. 1545–1553, July 2002.