# Blind Identification of Overcomplete MixturEs of sources (BIOME)

Laurent Albera[1,2], Anne Ferréol[2], Pierre Comon[1] and Pascal Chevalier[2]

[1] *I3S, Algorithmes-Euclide-B, BP 121, F-06903 Sophia-Antipolis, France*
[2] *THALES Communications, 146 Boulevard de Valmy, BP 82, F-92704 Colombes, France*

**Abstract**

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 full rank, and the right one being unitary. The application in antenna signal processing is described, and suboptimal numerical algorithms are proposed.

*Key words:* Blind Source Separation, Independent Component Analysis, Tensor, High Order Statistics, Joint Diagonalization, Underdetermined linear system, Overcomplete representation.

## 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, radar, sonar, seismology, radio astronomy, medical diagnosis (separation of electroencephalogram signals for instance), and data analysis. For example, 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

*Email address:* albera@i3s.unice.fr (Laurent Albera).

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 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 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; we shall not make the latter assumption, and assume on the contrary that the mixture may be *underdetermined*.

*Bibliographical survey*

While the first paper related to HO BSS has been published in 1985 by Herault et al. [27], the ICA concept is proposed a few years later; Comon proposes a Fourth Order (FO) contrast-based method, COM2 [11], Cardoso and Souloumiac [6] develop a matrix approach, well-known as JADE, and give rise to the joint diagonalization algorithm. Even when the JADE method uses both Second Order (SO) and FO statistics, Belouchrani et al. conceive the SOBI method [3], only based on SO statistics. A few years later, Hyvarinen et al. present the FastICA method, first for signals with values in the real field [29], and later for complex signals [4], using the fixed-point algorithm to maximize a FO contrast. This algorithm is of deflation type, as that of Delfosse et al. [21], and must extract one source at a time. Besides, Comon proposes a simple solution [13], named COM1 in this paper, to the maximization of another FO contrast function presented in [38] [15].

Each of these methods suffers from limitations. To start with, the SOBI algorithm is unable to restore components, which have comparable spectral densities. On the other hand, though the other previous methods perform under some reasonable assumptions, they may be strongly affected by a Gaussian noise with unknown spatial correlation. 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* [12], which means that the observation vectors are represented in the *overcomplete* basis of source vectors [30], hence the title of our paper. The previous algorithms, which are based on 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 [24]. In fact, the latter family of algorithms uses cyclic statistics of the data and, since cyclic covariance matrix associated with a stationary noise is 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.

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), several methods have been developed. Some papers focus on blind source extraction [30] [14], which is a difficult problem since underdetermined mixtures are not linearly invertible, while others, as herein, favour Blind Mixture Identification (BMI) [5] [12] [17] [30] [37] [23] [35]. The methods proposed in [5] [12] [17] [23] only exploit the information contained in the data FO statistics whereas the one proposed in [37] exploits the information contained in the second characteristic function of the observations. In fact, Cardoso presents in [5] as soon as 1991 the interesting FOOBI (Fourth Order Only Blind Identification) concept, which exploits the *super-symmetric* FO 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 [18] 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 [19] the FOOBI concept to the canonical decomposition of a HO tensor non necessarily super-symmetric, computed by means of a joint congruence transformation [39]. Moreover, an application to the blind identification of convolutive MIMO (Multiple Outputs Multiple Outputs) is given in [19]. 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 [20], which is interesting since many ICA algorithms rely on this

joint diagonalization step. As for Lee et al. [30], they maximize the probability of the data conditionally to the mixture matrix. However, all these methods have drawbacks in operational contexts. Indeed, The FOOBI algorithm [5] and its first improvement [18, 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 [18, section 3] of the FOOBI method is such that $P(P-1) \leq N^3(N-1)/2$. However, these three methods, like the PARAFAC algorithms [35], are suboptimal in terms of maximal number of processed sources, since the analysis of FO virtual arrays [10] yields that for arrays with particular diversity, up to $P = N^2$ steering vectors may be identified from only FO statistics. On the other hand, the third improvement [18, section 4] of the FOOBI algorithm allows one theoretically to reach the latter optimal upper bound. Nevertheless, although the previous methods [5] [18] seem very attractive in theory, no simulation has been presented. The BMI methods [12] [17] assume FO non-circularity and thus fail in separating FO circular sources. Besides, the theory developed in [12] only confines itself to the three sources and two sensors case. Although the method [30] succeeds in identifying the steering vectors of up to four speech signals with only two sensors, the authors need sparsity conditions, and do not address the general case when all sources are always present. In addition, the method [37] 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, although the FOBIUM algorithm [23] performs the BMI of up to $P = N^2$ sources for arrays of $N$ different sensors even in the presence of a Gaussian noise with unknown spatial correlation, it requires sources with different FO spectral densities.

*Contribution*

In order to overcome the previous drawbacks, a family of new methods named BIOME (Blind Identification of Overcomplete MixturEs of sources) is proposed in this paper.These methods operate on statistics of order $2q$, where $q$ is an arbitrary integer greater than 2, hence the name of $2q$-BIOME methods. This family of algorithms allows to blindly identify both overdetermined (with $q \geq 2$) and underdetermined (with $q \geq 3$) mixtures of sources, and to extract them in the overdetermined case. An application of BIOME to Sixth Order (SixO) statistics ($q = 3$), that is to say 6-BIOME, has been succintly presented in [2] under the name BIRTH (Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix).

More generally, the $2q$-th order 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,

without SO prewhitening, BIOME explicitly exploits the redundancies in the $2q$-th order statistical matrix of the data and implicitly uses the Virtual Array (VA) concept, presented in [22] [10] for FO methods, and extended in [9] for HO methods. Note that, for a given value of $N$, 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 $q$. One of our contributions is to show that it is interesting to increase $q$, despite the fact that the larger $q$, the worse the estimates of statistics of order $2q$ for a given number of snapshots.

From the linear algebra viewpoint, it is shown in section 4 that the BMI problem can be expressed in the form of the problem below, even in the underdetermined case.

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

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

Throughout the paper, vectors (one-way arrays) are denoted with bold lowercase symbols, and matrices (2-way arrays) or tensors (HO arrays) in bold uppercase. Transposition, conjugate transposition, and complex conjugation are denoted respectively with superscripts $(^{\mathsf{T}})$, $(^{\mathsf{H}})$, and $(^*)$.

## 2 Assumptions and problem formulation

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

$$\boldsymbol{x}(k) = \boldsymbol{A}\, \boldsymbol{s}(k) + \boldsymbol{\nu}(k) \tag{1}$$

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 3.3);
- (**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 3.1) are not null and have all the same sign;

(**A4**) Column vectors $\boldsymbol{a}_p$ of $\boldsymbol{A}$, also called steering vectors, are not collinear and have not any null component;

(**A5**) The $N^{q-1}\times P$ matrix $\boldsymbol{\mathcal{A}}_{q-1}^{\ell}$, which will be defined in section 4.1, is full column rank (this implies that $P\leq N^{q-1}$);

Under the previous assumptions, the problem addressed in this paper is the 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 3.1) of the observations. Besides, the classical BSS problem in the overdetermined case consists of finding a $N\times P$ matrix (the static source separator), $\boldsymbol{W}$, yielding a $P\times 1$ output vector $\boldsymbol{y}(k)=\boldsymbol{W}^{\mathsf{H}}\boldsymbol{x}(k)$ corresponding to the best estimate, $\hat{\boldsymbol{s}}(k)$, of the vector $\boldsymbol{s}(k)$, up to a multiplicative trivial matrix.

## 3 Statistics of $2q$-th order

### *3.1 Definition*

The $2q$-th order statistics considered in the paper are defined by

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

where $q$ terms $x_i(k)$ are not conjugated and $q$ terms are conjugated. Function (2) 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

$$C_{p,\,p,\,..,\,p,\,\boldsymbol{s}}^{p,\,p,\,..,\,p}(k)=\mathsf{Cum}\{\underbrace{s_p(k),s_p(k),\ldots,s_p(k)}_{q\,\text{components}},\underbrace{s_p(k)^*,\ldots,s_p(k)^*}_{q\,\text{components}}\} \tag{3}$$

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 the $2q$-BIOME algorithm in the stationary case. Nevertheless, the cyclostationary case will be addressed in short in section 3.3.

### *3.2 Matrix arrangement*

#### *3.2.1 Matrix notation*

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

$$\boldsymbol{B}^{\otimes m} = \underbrace{\boldsymbol{B}\otimes\boldsymbol{B}\otimes\ldots\otimes\boldsymbol{B}}_{m \text{ times}} \quad \text{with } \boldsymbol{B}^{\otimes 0}=1 \tag{4}$$

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 [28]. 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 $(N_GN_H)\times P$ 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:

$$\boldsymbol{G}\oslash\boldsymbol{H} = \Big[\boldsymbol{g}_1\otimes\boldsymbol{h}_1\ \boldsymbol{g}_2\otimes\boldsymbol{h}_2\ \cdots\ \boldsymbol{g}_P\otimes\boldsymbol{h}_P\Big] \tag{5}$$

So it is also possible to define the *Khatri-Rao power*:

$$\boldsymbol{B}^{\oslash m} = \underbrace{\boldsymbol{B}\oslash\boldsymbol{B}\oslash\ldots\oslash\boldsymbol{B}}_{m \text{ times}} \quad \text{with } \boldsymbol{B}^{\oslash 0}=1 \tag{6}$$

#### *3.2.2 $2q$-th order statistical matrix*

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

$$\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}\left(I_1^{\ell},I_2^{\ell}\right) = C^{i_{q+1},\ldots,i_{2q}}_{i_1,i_2,\ldots,i_q,\boldsymbol{x}} \tag{7}$$

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

$$\begin{aligned} I_1^\ell &= \varphi([\underbrace{i_1 \;\; i_2 \;\; \ldots \;\; i_{q-\ell-1} \;\; i_{q-\ell}}_{q-\ell \text{ first subscript indices}} \;\; \underbrace{i_{2q-\ell+1} \;\; \ldots \;\; i_{2q-1} \;\; i_{2q}}_{\ell \text{ last superscript indices}}]) \\ I_2^\ell &= \varphi([\underbrace{i_{q+1} \;\; i_{q+2} \;\; \ldots \;\; i_{2q-\ell-1} \;\; i_{2q-\ell}}_{q-\ell \text{ first superscript indices}} \;\; \underbrace{i_{q-\ell+1} \;\; \ldots \;\; i_{q-1} \;\; i_q}_{\ell \text{ last subscript indices}}]) \end{aligned} \tag{8}$$

and where function $\varphi$ is defined by

$$\forall \boldsymbol{z} \in \mathbb{N}^J, \quad \varphi(\boldsymbol{z}) = z(J) + \sum_{j=1}^{J-1} N^{J-j}(z(j) - 1) \tag{9}$$

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

**Example 1** *FO and SixO statistics described in appendix D may be arranged in the $N^2 \times N^2$ quadricovariance matrix $\boldsymbol{Q}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}^1_{4,\boldsymbol{x}}$ and the $N^3 \times N^3$ hexacovariance matrix $\boldsymbol{H}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}^1_{6,\boldsymbol{x}}$ respectively, such that*

$$Q_{\boldsymbol{x}}\left(I_1^1, I_2^1\right) = C^{i_3,i_4}_{i_1,i_2,\boldsymbol{x}} \qquad\qquad H_{\boldsymbol{x}}\left(J_1^1, J_2^1\right) = C^{i_4,i_5,i_6}_{i_1,i_2,i_3,\boldsymbol{x}} \tag{10}$$

*are the $(I_1^1, I_2^1)$-th entry ($1 \leq I_1^1, I_2^1 \leq N^2$) of $\boldsymbol{Q}_{\boldsymbol{x}}$ and the $(J_1^1, J_2^1)$-th entry ($1 \leq J_1^1, J_2^1 \leq N^3$) of $\boldsymbol{H}_{\boldsymbol{x}}$ respectively, and where for all $1 \leq i_1, i_2, i_3, i_4, i_5, i_6 \leq N$,*

$$\begin{aligned} I_1^1 &= \varphi([i_1 \;\; i_4]) & I_2^1 &= \varphi([i_3 \;\; i_2]) \\ J_1^1 &= \varphi([i_1 \;\; i_2 \;\; i_6]) & J_2^1 &= \varphi([i_4 \;\; i_5 \;\; i_3]) \end{aligned} \tag{11}$$

*Note that function $\varphi$ is thus defined by*

$$\begin{aligned} &\forall \boldsymbol{z} \in \mathbb{N}^2, \quad \varphi(\boldsymbol{z}) = N(z(1) - 1) + z(2) \\ &\forall \boldsymbol{z} \in \mathbb{N}^3, \quad \varphi(\boldsymbol{z}) = N(N(z(1) - 1) + z(2) - 1) + z(3) \end{aligned} \tag{12}$$

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

$$\boldsymbol{T}\begin{pmatrix} i_{2q}, i_{2q-1}, \ldots, i_{2q-\ell+1}, i_{q-\ell}, i_{q-\ell-1}, \ldots, i_1, \\ i_q, i_{q-1}, \ldots, i_{q-\ell+1}, i_{2q-\ell}, i_{2q-\ell-1}, \ldots, i_{q+1} \end{pmatrix} = C^{i_{q+1},\ldots,i_{2q}}_{i_1,i_2,\ldots,i_q,\boldsymbol{x}} \tag{13}$$

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

$$\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}} = reshape(\boldsymbol{T}, N^q, N^q) \tag{14}$$

We limit ourselves to arrangements of statistics that give different results at the output of the $2q$-BIOME method in terms of *processing power* (i.e. in terms of maximal number of processed sources). Note that the selection of the ordering parameter $\ell$ maximizing the processing power for a fixed cumulant order $q$ will be discussed in section 5.2 summarizing results shown in [9].

### *3.2.3 Multilinearity property*

The statistical matrix of the data, $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$ $(q \geq 1)$, has a special structure especially thanks to the multilinearity property under changes of coordinate systems, shared by all moments and cumulants [31] [32, pp. 1-24]. Under assumptions (**A1**)-(**A2**), this property can be expressed, according to (7), (8) and (9), by the following equation

$$\forall\, 0 \leq \ell \leq q, \quad \boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}} = [\boldsymbol{A}^{\otimes q-\ell} \otimes \boldsymbol{A}^{*\otimes \ell}]\, \boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{s}}\, [\boldsymbol{A}^{\otimes q-\ell} \otimes \boldsymbol{A}^{*\otimes \ell}]^{\mathsf{H}} \tag{15}$$

where the $N^q \times N^q$ matrices $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$ and the $P^q \times P^q$ matrices $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{s}}$ are the statistical matrices of $\boldsymbol{x}(k)$ and $\boldsymbol{s}(k)$ respectively. The number $\ell$ is the same as that appearing in equations (8) and (7). Moreover, note that the arrangements $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$ and $\boldsymbol{\mathcal{C}}^{q-\ell}_{2q\,\boldsymbol{x}}$ $(0 \leq \ell \leq q)$ give rise to the same processing power of underdetermined mixtures of arbitrary statistically independent sources as shown in [9]. 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.

## *3.3 Statistical estimation*

Generally, using the well-known Leonov-Shiryaev formula [31], applicable in the complex case [36], $2q$-th order cumulants (2) are computed from moments of order smaller than or equal to $2q$ given by

$$M^{i_{r+1},i_{r+2},\ldots,i_{r+s}}_{i_1,i_2,\ldots,i_r,\boldsymbol{x}}(k) = \mathsf{E}[x_{i_1}(k), \ldots, x_{i_r}(k), x_{i_{r+1}}(k)^*, \ldots, x_{i_{r+s}}(k)^*] \tag{16}$$

where $r+s \leq 2q$. Appendix D illustrates the Leonov-Shiryaev formula for FO and SixO statistics.

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 [31], and consequently to estimate, via the Leonov-Shiryaev formula, $2q$-th order statistics (2).

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

$$C^{i_{q+1},i_{q+2},\ldots,i_m}_{i_1,i_2,\ldots,i_q,\boldsymbol{x}} = \left\langle C^{i_{q+1},i_{q+2},\ldots,i_m}_{i_1,i_2,\ldots,i_q,\boldsymbol{x}}(k) \right\rangle_{\mathrm{c}} \tag{17}$$

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

$$\forall\, f : t \longmapsto f(t), \quad \langle f(t)\rangle_{\mathrm{c}} = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t)dt \tag{18}$$

These continuous-time temporal mean statistics are thus estimated using, for $q=2$, the estimators described in [25] for zero mean signals and in [26] for potentially non zero-mean signals, and extending the previous ones to very HO statistics for $q \geq 3$. Note that the proposed BIOME approach (in its current form) can tolerate, but does not exploit cyclostationarity of the sources such as in [24]: this will be the subject of a forthcoming paper.

## 4 The $2q$-BIOME method

It is subsequently shown that the $2q$-BIOME method exploits the structure of the statistical matrix $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$, for the chosen value of $\ell$, $0 \leq \ell \leq q_0$, so that the joint diagonalization to perform is actually somewhat more complicated than that given in problem 1, and better described by

**Problem 2** *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 $\boldsymbol{A}$, $N$ inversible 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* $\boldsymbol{\mathcal{A}}^{\ell}_{q} = \boldsymbol{A}^{\odot q-\ell} \odot \boldsymbol{A}^{*\odot \ell}$.

*4.1 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 (15) (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

$$\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}} = \boldsymbol{\mathcal{A}}^{\ell}_{q}\ \boldsymbol{\zeta}_{2q,\boldsymbol{s}}\ \boldsymbol{\mathcal{A}}^{\ell\,\mathsf{H}}_{q} \tag{19}$$

where $\boldsymbol{\zeta}_{2q,\boldsymbol{s}} \overset{\text{def}}{=} \text{Diag}\Big[\, C^{1,1,..,1}_{1,1,...1,\boldsymbol{s}}\ C^{2,2,..,2}_{2,2,..,2,\boldsymbol{s}}\ \cdots\ C^{P,P,..,P}_{P,P,..,P,\boldsymbol{s}}\,\Big]$ is a $P\times P$ diagonal matrix of full rank in contrast to $\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{s}}$ (15), and where the $N^q\times P$ matrix $\boldsymbol{\mathcal{A}}^{\ell}_{q}$ is given by

$$\begin{aligned}\boldsymbol{\mathcal{A}}^{\ell}_{q} &= \boldsymbol{A}^{\oslash q-\ell} \oslash \boldsymbol{A}^{*\oslash \ell} = \Big[\boldsymbol{a}_1^{\otimes q-\ell}\otimes(\boldsymbol{a}_1^{*})^{\otimes \ell}\ \cdots\ \boldsymbol{a}_P^{\otimes q-\ell}\otimes(\boldsymbol{a}_P^{*})^{\otimes \ell}\Big] \\ &= \Big[[\boldsymbol{\mathcal{A}}^{\ell}_{q-1}\boldsymbol{\Phi}_1]^{\mathsf{T}}\ [\boldsymbol{\mathcal{A}}^{\ell}_{q-1}\boldsymbol{\Phi}_2]^{\mathsf{T}}\ \cdots\ [\boldsymbol{\mathcal{A}}^{\ell}_{q-1}\boldsymbol{\Phi}_N]^{\mathsf{T}}\Big]^{\mathsf{T}}\end{aligned} \tag{20}$$

with

$$\boldsymbol{\Phi}_n = \text{Diag}[\, A(n,1)\ A(n,2)\ \cdots\ A(n,P)\,] \tag{21}$$

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 (19) ensues immediately from equations (7), (8), (9) and from the multilinearity property shared by cumulants [31] [32, pp. 1-24]. Moreover, it appears from equation (20), 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

$$\begin{aligned}&\forall\, 1\leq g\leq N^{q-1},\ \exists\, 1\leq n_1,\ldots,n_{q-1}\leq N,\ g=\varphi([n_{q-1}\ \ n_{q-2}\ \ \ldots\ \ n_1]),\\ &\text{and } \boldsymbol{\Psi}_g = \begin{cases} \boldsymbol{A}\ \prod_{j=1}^{q-1}\boldsymbol{\Phi}_{n_j} & \text{if } \ell=0\\ \boldsymbol{A}^{*}\ \prod_{j=1}^{\ell-1}\boldsymbol{\Phi}_{n_j}^{\ *}\ \prod_{k=\ell}^{q-1}\boldsymbol{\Phi}_{n_k} & \text{otherwise (o.w.)}\end{cases}\end{aligned} \tag{22}$$

and

$$\boldsymbol{\mathcal{A}}^{\ell}_{q} = [\boldsymbol{\Psi}_1^{\mathsf{T}}\ \boldsymbol{\Psi}_2^{\mathsf{T}}\cdots\boldsymbol{\Psi}_G^{\mathsf{T}}]^{\mathsf{T}}\,. \tag{23}$$

*4.2 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 $\boldsymbol{A}$.

*4.2.1 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 (19), 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 $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2}[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{H/2} = \boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell$, 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 (19) 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

$$[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2} = \boldsymbol{E}_{\boldsymbol{s}}\, \boldsymbol{L}_{\boldsymbol{s}}^{1/2} \tag{24}$$

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

**Proposition 1** *Under assumptions (**A4**) and (**A5**), the $N^q \times P$ matrix $\boldsymbol{\mathcal{A}}_q^\ell$ is full column rank.*

The proof of proposition 1 ensues immediately from equations (20), (21) 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 $\boldsymbol{\mathcal{A}}_q^\ell$ (20) implies that for all $1 \leq n \leq N$, $\boldsymbol{\mathcal{A}}_{q-1}^\ell \boldsymbol{\Phi}_n \boldsymbol{\beta} = 0$. So it implies that $\boldsymbol{\mathcal{A}}_{q-1}^\ell$ cannot be full column rank (since matrices $\boldsymbol{\Phi}_n$ are $P \times P$ diagonal with nonzero entries, due to (21) and (**A4**)), which contradicts assumption (**A5**).

Asumptions (**A3**) to (**A5**), proposition 1, and equations (19) and (24) 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 $\boldsymbol{E}_{\boldsymbol{s}}$ and $\boldsymbol{L}_{\boldsymbol{s}}$, are of rank $P$ as well.

**Proposition 2** *For a full rank matrix $\boldsymbol{\mathcal{A}}_q^\ell$, (**A3**) is equivalent to assume that the diagonal elements of $\boldsymbol{L}_{\boldsymbol{s}}$ are not null and have also the same sign.*

The proof of proposition 2 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

$$\boldsymbol{\mathcal{A}}_q^\ell \, \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} = \boldsymbol{E}_{\boldsymbol{s}} \, \boldsymbol{L}_{\boldsymbol{s}}^{1/2} \, \boldsymbol{V} \quad \Big( = [\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2} \, \boldsymbol{V} \Big) \tag{25}$$

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 (25) that

$$\boldsymbol{E}_{\boldsymbol{s}}^{\mathsf{H}} \, \boldsymbol{\mathcal{A}}_q^\ell \, \boldsymbol{\zeta}_{2q,\boldsymbol{s}} \, \boldsymbol{\mathcal{A}}_q^{\ell\,\mathsf{H}} \, \boldsymbol{E}_{\boldsymbol{s}} = \boldsymbol{L}_{\boldsymbol{s}} \tag{26}$$

and hence proposition 2.

In addition, equation (25) can be rewritten as follows

$$[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2} = \boldsymbol{E}_{\boldsymbol{s}} \, \boldsymbol{L}_{\boldsymbol{s}}^{1/2} = \boldsymbol{\mathcal{A}}_q^\ell \, \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \, \boldsymbol{V}^{\mathsf{H}}. \tag{27}$$

showing the link between $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2}$ and $\boldsymbol{\mathcal{A}}_q^\ell$. Plugging (20) into (27), matrix $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2}$ can be eventually rewritten as

$$\begin{aligned} [\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^\ell]^{1/2} &= \Big[ [\boldsymbol{\mathcal{A}}_{q-1}^\ell \boldsymbol{\Phi}_1 \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \; [\boldsymbol{\mathcal{A}}_{q-1}^\ell \boldsymbol{\Phi}_2 \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \cdots [\boldsymbol{\mathcal{A}}_{q-1}^\ell \boldsymbol{\Phi}_N \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \boldsymbol{V}^{\mathsf{H}}]^{\mathsf{T}} \Big]^{\mathsf{T}} \\ &= [\, \boldsymbol{\Gamma}_1^{\mathsf{T}} \;\; \boldsymbol{\Gamma}_2^{\mathsf{T}} \;\; \cdots \;\; \boldsymbol{\Gamma}_N^{\mathsf{T}} \,]^{\mathsf{T}} \end{aligned} \tag{28}$$

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

$$\forall \, 1 \leq n \leq N, \quad \boldsymbol{\Gamma}_n = \boldsymbol{\mathcal{A}}_{q-1}^\ell \boldsymbol{\Phi}_n \, \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \, \boldsymbol{V}^{\mathsf{H}} \tag{29}$$

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

The proof results from proposition 1, in addition to all other stated conditions.

Using proposition 3, the pseudo-inverses $\boldsymbol{\Gamma}_n^\sharp$ of the $N^{q-1} \times P$ matrices $\boldsymbol{\Gamma}_n$ may be defined by

$$\forall \, 1 \leq n \leq N, \quad \boldsymbol{\Gamma}_n^\sharp = (\boldsymbol{\Gamma}_n^{\mathsf{H}} \boldsymbol{\Gamma}_n)^{-1} \, \boldsymbol{\Gamma}_n^{\mathsf{H}} \tag{30}$$

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

$$\forall \, 1 \leq n_1 \neq n_2 \leq N, \quad \boldsymbol{\Theta}_{n_1,n_2} = \boldsymbol{\Gamma}_{n_1}^\sharp \boldsymbol{\Gamma}_{n_2}. \tag{31}$$

To see this, let us compute $\boldsymbol{\Theta}_{n_1,n_2}$ from (29) and (30). We obtain

$$\boldsymbol{\Theta}_{n_1,n_2} = \boldsymbol{V} \, [\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^\ell]^{-1/2} \, \boldsymbol{\Phi}_{n_1}^{-1} \, \boldsymbol{\Phi}_{n_2} \, \boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2} \, \boldsymbol{V}^{\mathsf{H}} = \boldsymbol{V} \, \boldsymbol{\Phi}_{n_1}^{-1} \, \boldsymbol{\Phi}_{n_2} \, \boldsymbol{V}^{\mathsf{H}} \tag{32}$$

where $\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}$ and $\boldsymbol{D}_{n_1,n_2} = \boldsymbol{\Phi}_{n_1}^{-1}\boldsymbol{\Phi}_{n_2}$ are $P\times P$ diagonal full rank matrices, which shows the result. The unitary matrix $\boldsymbol{V}_{sol}$, solution to the previous problem of joint diagonalization of the $N(N-1)$ matrices $\boldsymbol{\Theta}_{n_1,n_2}$ has necessarily the form $\boldsymbol{V}_{sol} = \boldsymbol{V}\boldsymbol{\mathcal{T}}$ where $\boldsymbol{\mathcal{T}}$ is a unitary matrix. This allows, in accordance with (27), to recover $\boldsymbol{\mathcal{A}}_q^\ell$ to within an orthogonal matrix as

$$[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}\,\boldsymbol{V}_{sol} = \boldsymbol{\mathcal{A}}_q^\ell\,\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}\,\boldsymbol{\mathcal{T}} \tag{33}$$

**Proposition 4** *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 4 and [3] allow to assert that the previous unitary matrix $\boldsymbol{\mathcal{T}}$ is also trivial. So matrix $\boldsymbol{\mathcal{A}}_q^\ell$ may be identified, according to (33), up to a trivial matrix.

### *4.2.2 Identification of mixture* $\boldsymbol{A}$

Three algorithms are proposed in this section, with increasing computational complexity and accuracy.

Note, from (23) and (22), that equation (33) can also be written in the form of $G=N^{q-1}$ matrix blocks $\boldsymbol{\Sigma}_g = \boldsymbol{\Psi}_g\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}\,\boldsymbol{\mathcal{T}}$ of size $N\times P$ as

$$[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}\,\boldsymbol{V}_{sol} = [\boldsymbol{\Sigma}_1{}^{\mathsf{T}}\;\boldsymbol{\Sigma}_2{}^{\mathsf{T}}\cdots\boldsymbol{\Sigma}_G{}^{\mathsf{T}}]^{\mathsf{T}} \tag{34}$$

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 $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}\boldsymbol{V}_{sol}$ such that

$$\boldsymbol{\Sigma}_1 = \begin{cases} \boldsymbol{A}\;[\boldsymbol{\Phi}_1]^{q-1}\;\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}\,\boldsymbol{\mathcal{T}} & \text{if } \ell=0 \\ \boldsymbol{A}^*\;[\boldsymbol{\Phi}_1^*]^{\ell-1}\;[\boldsymbol{\Phi}_1]^{q-\ell}\;\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}\,\boldsymbol{\mathcal{T}} & \text{o.w.} \end{cases} \tag{35}$$

where $\boldsymbol{\zeta}_{2q,\boldsymbol{s}}^{1/2}$ and $\boldsymbol{\Phi}_n$, 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 $\boldsymbol{b}_p = [\boldsymbol{b}_p(1)^T \boldsymbol{b}_p(2)^T \cdots \boldsymbol{b}_p(H)^T]^T$), 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 $\Delta^{\ell}_p$ of matrices defined by

$$\Delta^{\ell}_p = \begin{cases} \{\boldsymbol{B}_p(h)\boldsymbol{B}_p(h)^H,\ (\boldsymbol{B}_p(h)^H\boldsymbol{B}_p(h))^* \ /\ 1 \leq h \leq H\} & \text{if } \ell = 0 \\ \{\boldsymbol{B}_p(h)^* \ /\ 1 \leq h \leq H\} & \text{if } \ell = 1 \\ \{(\boldsymbol{B}_p(h)\boldsymbol{B}_p(h)^H)^*,\ (\boldsymbol{B}_p(h)^H\boldsymbol{B}_p(h)) \ /\ 1 \leq h \leq H\} & \text{o.w.} \end{cases} \tag{36}$$

**Theorem 1** *The eigenvector, in common to all matrices of $\Delta^{\ell}_p$, and associated with the strongest 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 $\Delta^{\ell}_p$ allows to estimate a column vector of $\boldsymbol{A}$, and finally to identify $\boldsymbol{A}$ to within a trivial matrix.

**Remark 2** *Although the algorithm of joint approximate diagonalization in the LS sense [7] is restricted to unitary joint diagonalizers, it can be used to process the previous problem since matrices belonging to $\Delta^{\ell}_p$ 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 [39], a better LS fit can be attained, possibly leading to a better estimate of $\boldsymbol{A}$. Both approaches will be compared in section 6.2.1.*

### *4.2.3 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 [8], 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 $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}^{0}_{2,\boldsymbol{x}}$.

### *4.3 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 5.2 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 $C^{i_{q+1},\ldots,i_{2q}}_{i_1,i_2,\ldots,i_q,\boldsymbol{x}}$ from the $K$ samples $\boldsymbol{x}(k)$ and choose, using section 5.2 and [9], the best arrangement $\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}$, where $\widehat{\boldsymbol{\mathcal{C}}}^{\ell}_{2q,\boldsymbol{x}}$ is an estimate of $\boldsymbol{\mathcal{C}}^{\ell}_{2q,\boldsymbol{x}}$.

**Step3** Compute the EVD of the Hermitian matrix $\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}$; estimate $\widehat{P}$, an estimate of the source number $P$, from an eigenvalue test and restrict the EVD to the $\widehat{P}$ principal components : $\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}} \approx \widehat{\boldsymbol{E}}_{\boldsymbol{s}}\,\widehat{\boldsymbol{L}}_{\boldsymbol{s}}\,\widehat{\boldsymbol{E}}^{\mathsf{H}}_{\boldsymbol{s}}$, 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, $\epsilon$, of the diagonal elements of $\widehat{\boldsymbol{L}}_{\boldsymbol{s}}$.

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

**Step6** Extract from $[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}$ the $N$ matrices $\widehat{\boldsymbol{\Gamma}}_n$, construct matrices $\widehat{\boldsymbol{\Theta}}_{n_1,n_2} = [\widehat{\boldsymbol{\Gamma}}^{\sharp}_{n_1}\,\widehat{\boldsymbol{\Gamma}}_{n_2}]$ 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 [7]).

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

(1) ($2q$-BIOME1) taking the matrix block made up of the $N$ first rows of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]$ if $\ell_{opt}=0$, and of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]^*$ otherwise;
(2) ($2q$-BIOME2) taking the average of the $N$ matrix blocks, of size $N \times P$, made up of the successive rows of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]$ if $\ell_{opt} = 0$, and of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]^*$ otherwise;
(3) ($2q$-BIOME3) fully exploiting each column vector $\widehat{\boldsymbol{b}}_p$ of $[[\epsilon\widehat{\boldsymbol{\mathcal{C}}}^{\ell_{opt}}_{2q,\boldsymbol{x}}]^{1/2}\,\widehat{\boldsymbol{V}}_{sol}]$. 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 $\widehat{\Delta}^{\ell}_p(m)$ $(1 \leq m \leq M)$ and associated with the largest eigenvalue; the algorithm used for this task is JAD [7];
(4) ($2q$-BIOME4) doing the same as in $2q$-BIOME3, but using the congruent diagonalization algorithm of Yeredor [39] instead of JAD.

**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 $\widehat{\boldsymbol{W}} = \widehat{\boldsymbol{R}}^{-1}_{\boldsymbol{x}}\widehat{\boldsymbol{A}}$, where $\widehat{\boldsymbol{R}}_{\boldsymbol{x}}$ is an estimate of $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}^{0}_{2,\boldsymbol{x}}$.

## 5 Identifiability

The identifiability properties of the $2q$-BIOME method are directly related to the $2q$-th order Virtual Array (VA) concept described in [22] [10] for $q = 2$ and extended in [9] for $q \geq 2$. For this reason, we recall the main results about the VA array concept in section 5.1 before discussing, in section 5.2, the identifiability properties of $2q$-BIOME.

### *5.1 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 $\theta_p$ and $\varphi_p$ 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 [16]

$$a_n(\theta_p, \varphi_p) = f_n(\theta_p, \varphi_p, \omega_p) \exp\{\mathrm{j}2\pi[x_n \cos(\theta_p)\cos(\varphi_p) + y_n \sin(\theta_p)\cos(\varphi_p) + z_n \sin(\varphi_p)]/\lambda\} \quad (37)$$

where $j = \sqrt{-1}$, $\lambda$ 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 $(\theta_p, \varphi_p)$ and having the state of polarization $\omega_p$ (characterized by two angles in the wave plane) [16]. Let us recall that an array of sensors has space 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 $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$ and $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{\mathcal{C}}^{0}_{2\,\boldsymbol{x}}$, defined by (19), have the same algebraic structure, where the marginal source cumulant $C^{p,p,\ldots,p}_{p,p,\ldots,p\,\boldsymbol{s}}$ 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 $\boldsymbol{\mathcal{C}}^{\ell}_{2q\,\boldsymbol{x}}$, the role played for $\boldsymbol{R}_{\boldsymbol{x}}$ by the power $C^{p}_{p\boldsymbol{s}}$ and the steering vector $\boldsymbol{a}(\theta_p, \varphi_p)$ respectively. Thus, for BMI methods exploiting expression (19), 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 $f_n(\theta_p, \varphi_p, \omega_p)$ $(1 \leq n \leq N)$. 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 (37), it is shown in [9] that the 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 $f_n(\theta_p, \varphi_p, \omega_p)$ $(1 \leq n \leq N)$ 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 an upper bound to the rank of matrix $\boldsymbol{\mathcal{A}}_q^{\ell}$. Conversely, if the $2q$-th order VA has no ambiguities [1] 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**) and assuming $P = \mathcal{N}_{2q}^{\ell}$.

In particular it is shown in [9] 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 $q$ $(2 \leq q \leq N)$, the number of different VS $\mathcal{N}_{2q}^{\ell}$ can be approximated by

$$\mathcal{N}_{2q}^{\ell} \approx N!/\left[(N-q)!\,(q-\ell)!\,\ell!\right] \tag{38}$$

In these conditions, the optimal arrangement $\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell_{opt}}$ is such that $\ell_{opt}$ maximizes $\mathcal{N}_{2q}^{\ell}$ defined by (38) and thus minimizes the quantity $(q-\ell)!\,\ell!$ with respect to $\ell$ $(0 \leq \ell \leq q_0$ where $q_0 = q/2$ if $q$ is even and $q_0 = (q-1)/2$ if $q$ is odd). It is straightforward to show that $\ell_{opt} = q_0$ and it is verified in [9] for $2 \leq q \leq 4$ that this result remains true whatever $N$.

The exact 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 $N$, $q$ $(q \geq 2)$ and $\ell$. For this reason, Chevalier et al. [9] limit their analysis to some values of $q$ $(2 \leq q \leq 4)$, which extends the results of [10] up to the eighth order for arbitrary arrangements of the data cumulants. In fact, for these values of $q$, Chevalier et al. give an 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 1, then for an array with angular and polarization diversity only, and finally for an array with only space diversity summarized in table 2. These upper bounds are shown in [9] 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 [9] to be given by

$$\mathcal{N}_{2q}^{\ell} = q(N-1) + 1 \tag{39}$$

whatever $q$, $N$ and $\ell$, showing that the number $\mathcal{N}_{2q}^{\ell}$ of different VS of the $2q$-th order VA associated with a ULA is independent of $\ell$ and of the chosen

[1] Remind that a sensor array has ambiguities of rank $m$ if and only if there exists at least one $(m+1)$-uplet of linearly dependent directional vectors associated to $(m+1)$ distinct location parameters. We recall that directional vectors belong to a manifold, entirely defined by the sensor array, and parametrized by location parameters. Rank 1 ambiguities are also known as Grating Lobes [16]; higher order ambiguities have been first introduced by Schmidt in 1981 [34].

arrangement $\mathcal{C}^{\ell}_{2q,\boldsymbol{x}}$. However, for Uniformly spaced Circular Arrays (UCA) of $N$ sensors, the upper bound is shown in [9] to be reached when $N$ is a prime number as depicted in table 3.

Table 1
$\mathcal{N}^{2q,\ell}_{max}$ associated with arrays with space, angular and polarization diversities

| | | $\mathcal{N}^{2q,\ell}_{max}$ |
|---|---|---|
| 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 2
$\mathcal{N}^{2q,\ell}_{max}$ associated with arrays with space diversity only

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

### *5.2 The BIOME processing power*

From the results of section 5.1, it is possible to identify the maximum number, $P^{N,q}_{max}$, of independent non Gaussian sources that can be processed by the $2q$-BIOME method. Indeed, it has been shown in the paper 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 directions of arrival. In a same manner, assuming the $2(q-1)$th order VA associated with the arrangement $\mathcal{C}^{\ell}_{2(q-1),\boldsymbol{x}}$ and the considered array of $N$ sensors

Table 3
$\mathcal{N}_{2q}^{\ell}$ associated with a UCA of $N$ identical sensors

| | | $\mathcal{N}_{2q}^{\ell}$ | | | | |
|---|---|---|---|---|---|---|
| | | $N=3$ | $N=5$ | $N=7$ | $N=9$ | $N=11$ |
| 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 |

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 $\mathcal{N}_{2(q-1)}^{\ell_{opt}}$.

Now concerning parameter $q$, on one hand it depends on the number $P$ of independent sources that BIOME's user wants to process. On the other hand, according to assumption (**A3**), $2q$-th order marginal source cumulants have all to be non-zero and to have the same sign. 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 important to choose $q$ ($q \geq 2$) such that $P \leq P_{max}^{N,q}$, taking assumption (**A3**) into consideration. So, the lowest $q$ that still enables identification is not necessary optimal.

## 6 Computer simulations

### *6.1 Performance criterion*

Most of the existing performance criteria used to evaluate the quality of BMI algorithms, either in the overdetermined [11] or in the underdetermined [12] [37] cases, 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 [2] [23] by the $P$-uplet

$$\mathrm{D}\Big(\boldsymbol{A}, \widehat{\boldsymbol{A}}\Big) = (\alpha_1, \alpha_2, \ldots, \alpha_P) \tag{40}$$

where

$$\alpha_p = \min_{1 \leq i \leq P} \left[\mathrm{d}(\boldsymbol{a}_p, \widehat{\boldsymbol{a}}_i)\right] \tag{41}$$

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

$$\mathrm{d}(\boldsymbol{u}, \boldsymbol{v}) = 1 - \frac{|\boldsymbol{u}^{\mathsf{H}} \boldsymbol{v}|^2}{\|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2} \tag{42}$$

Thus the identification quality of the source $p$ is evaluated by the parameter $\alpha_p$, which decreases toward zero as the identification quality of the source $p$ improves. In particular, the source $p$ is perfectly identified when $\alpha_p = 0$. It will be subsequently considered that a source $p$ is blindly identified with a very high quality if $\alpha_p \leq 0.01$, with a high quality if $\alpha_p \leq 0.03$, with a good quality if $\alpha_p \leq 0.05$ and with a poor quality otherwise.

### *6.2 Computer results*

The synthetic signals used in this section are stationary and ergodic, and according to section 3.3, sample statistics [31] may be employed. More precisely, the sources utilized are QPSK in baseband, with a square transmit filter, and a symbol rate equal to the sample rate [33]. Put in simple words, these sources $s_p(k)$ are actually sequences of *independently and identically distributed* (i.i.d.) random variables $\forall k$, $\forall p$, taking equally likely their values in the set $\{1, j, -1, -j\}$.

Moreover, the $P$ statistically independent QPSK sources are assumed to be received by a UCA of $N$ identical sensors of radius $R$ such that $R/\lambda = 0.55$ ($\lambda$: wavelength). The $P$ sources, assumed synchronized, have the same input SNR (Signal to Noise Ratio) and the noise is Gaussian. We apply different BMI methods such as COM2 [11], JADE [6], FastICA [4], FOBIUM [23], 6-BIOME1, 6-BIOME2, 6-BIOME3 and 6-BIOME4 methods, and the performance criterion $\mathrm{D}(\boldsymbol{A}, \widehat{\boldsymbol{A}}) = (\alpha_1, \alpha_2, \ldots, \alpha_P)$ is computed and averaged over 200 realizations.

### 6.2.1 *The underdetermined case*

The 6-BIOME methods are compared to each other and to FOBIUM, in an underdetermined context. $P = 7$ poorly angularly separated QPSK sources ($\theta_1 = 10°$, $\theta_2 = 35°$, $\theta_3 = 60°$, $\theta_4 = 85°$, $\theta_5 = 105°$, $\theta_6 = 150°$, $\theta_7 = -45°$, $\varphi_1 = \varphi_2 = \varphi_3 = \varphi_4 = \varphi_5 = \varphi_6 = \varphi_7 = 0°$) are received by a UCA of $N = 3$ identical sensors, with the same input SNR of 20 dB. The noise is spatially and temporally white Gaussian. The value of $\max_{1 \leq p \leq 7}\{\alpha_p\}$ is reported in figure 1, that is, the performance index for the worst estimation among the 7 sources.

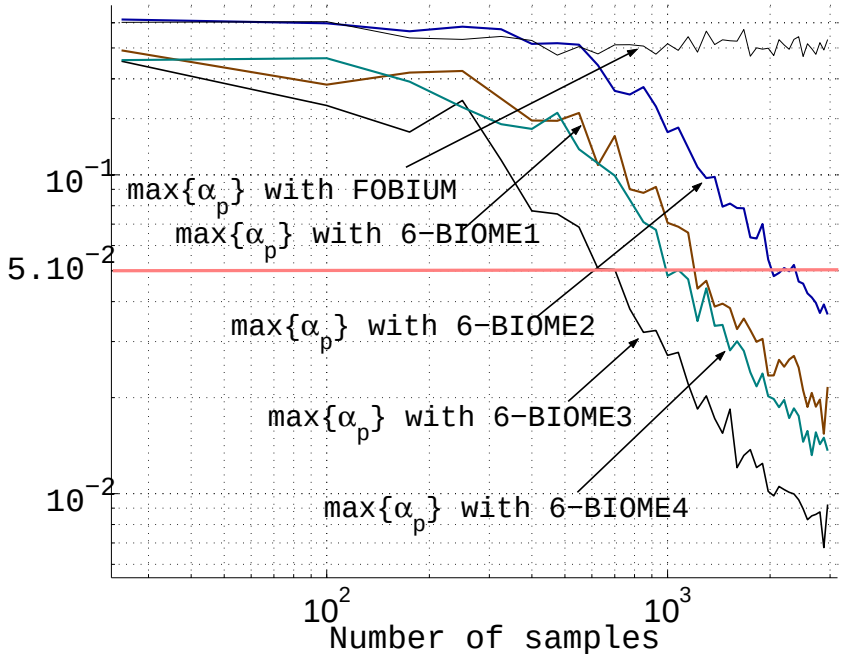


Fig. 1. $\max\{\alpha_p\}$ for a SNR of 20 dB

Several important observations can be made.

- the four 6-BIOME algorithms succeed in identifying the seven source directional vectors, but with different convergence speed. For instance, 6-BIOME3 turns out to be the fastest, and is able to yield values of $\alpha_p$'s all below 0.05 as soon as 600 samples are available. This is a surprising result, which contradicts the fact often admitted that "higher order techniques are slower to converge".
- also surprisingly, 6-BIOME3 using the JAD algorithm performs better than 6-BIOME4 using Yeredor's joint congruent diagonalization.
- 6-BIOME1 performs better than 6-BIOME2. This is due to the choice of the scenario. It may be useful to weight of blocks $\mathbf{\Sigma}_g$ (if $\ell = 0$) or $\mathbf{\Sigma}_g^*$ (otherwise) in the averaging in order to improve on performances. Nevertheless, the exact optimal weighting still needs to be calculated.

- as already pointed out in section 1, FOBIUM cannot identify the 2×7 mixing matrix because the 7 sources are white, and thus have the same trispectrum.

It turns out that the BIOME methods, originally devised for underdetermined mixtures, also perform quite well with overdetermined mixtures, as demonstrated in the next two subsections.

### *6.2.2 The overdetermined case : poorly angularly separated sources*

The scenario is as follows. There are $N = 2$ sensors and $P = 2$ QPSK sources very close to each other (the directions of arrival are $\theta_1 = 88°, \theta_2 = 90°, \varphi_1 = \varphi_2 = 0°$). The additive noise is spatially and temporally white Gaussian, and the SNR is 20dB. Since the best results were obtained for $m = 3$ in the previous subsection, we report here the comparison results only for $m = 3$.

Figure 2 confirms the very good behavior of 6−BIOME3, even with a very small number of snapshots ($K \leq 100$). However, figure 2 shows that for poorly angularly separated sources, there exists a number of snapshots, $K_0$, above which 6-BIOME3 becomes more efficient than COM2, JADE and FastICA. In the present scenario, it can be seen in figure 2 that $K_0 \approx 100$. In such a situation, we can claim that the resolution gain obtained with $2q$-BIOME is higher than the loss due to a higher variance in the statistics estimates. Similar results have been obtained for the directional vector of source 2.

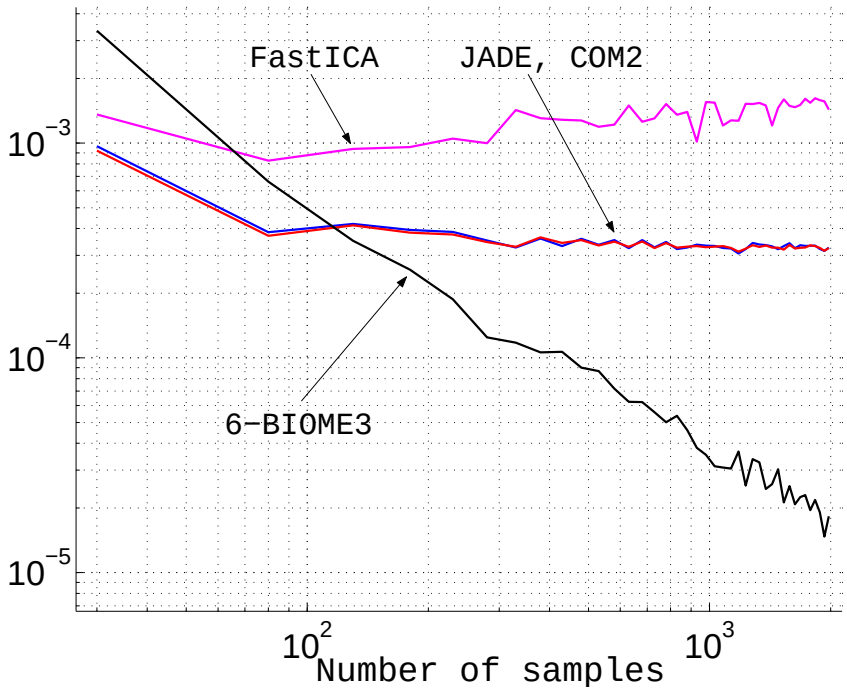


Fig. 2. $\alpha_1$ for 2 poorly angularly separated sources and for a SNR of 20 dB

*6.2.3 The overdetermined case : colored noise*

Eventually, the 6-BIOME3 method is compared to other algorithms in an overdetermined context but in the presence of a Gaussian noise with unknown spatial correlation. We have $P = 3$ sources, well separated ($\theta_1 = 10°$,$\theta_2 = 50°$,$\theta_3 = -40°$,$\varphi_1 = \varphi_2 = \varphi_3 = 0°$). The 3 QPSK sources are received by a UCA of $N = 5$ identical sensors. This time, we apply COM2, JADE, FOBIUM and 6-BIOME3 methods.

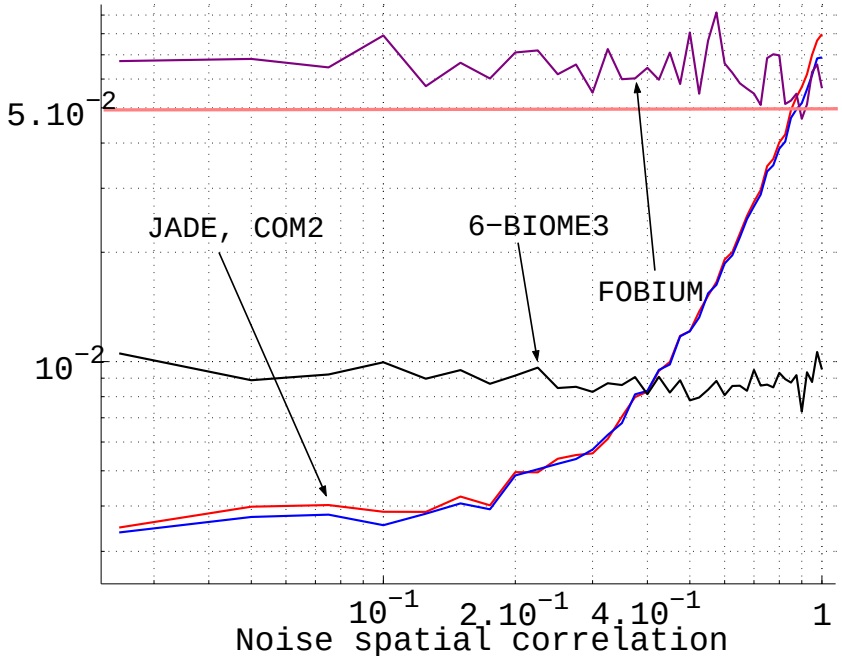


Fig. 3. $\alpha_3$ for 3 angularly separated sources and for a SNR of 0 dB

Figure 3 shows the variations of $\alpha_3$ (source 3 performance) at the output of the previous methods as a function of the noise spatial correlation factor $\rho$. SNR of the three sources is taken equal to 0 dB and 500 samples are used to identify the overdetermined mixture. 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

$$\boldsymbol{R}_{\boldsymbol{\nu}}^{in}(r,q) \stackrel{\text{def}}{=} \sigma^2\delta(r-q)/2 \qquad \boldsymbol{R}_{\boldsymbol{\nu}}^{out}(r,q) \stackrel{\text{def}}{=} \sigma^2\rho^{|r-q|}/2 \tag{43}$$

where $\sigma^2$, $\rho$ 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 3 that the 6-BIOME3 method seems to be robust with respect to the correlated Gaussian noise presence. On the other hand, the well-known COM2 and JADE methods are strongly affected as soon as the noise

spatial correlation approaches 1. Actually, there exists a threshold above which $2q$-BIOME methods become more attractive than more classical methods. In the present scenario, this threshold is $\rho = 0.4$. Lastly, FOBIUM does not succeed in identifying the directional vector of source 3 (this can be detected because $\alpha_3$ is always larger than 0.05 regardless of the value of $\rho$). This is due to the fact that the 3 sources have identical trispectra. Similar results have been obtained for sources 1 and 2.

## 7 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 paper. These new methods allow to process both over and underdetermined mixtures of sources, provided the latter have non zero marginal HO cumulants with the same sign. The proposed methods are not sensitive to a Gaussian colored noise whose spatial coherence is unknown. They also allow the processing of a number of sources depending on both the kind of sensors and the array geometry, and *fast 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.

Thus, despite the higher variance of their sample estimates, high order statistics (*e.g.* order 6) used in $2q-$BIOME algorithms may yield better performances than statistics of lower orders (*e.g.* order 4) used in more classical algorithms. The reason is the following. When sources are angularly close to each other, they can be hardly separated because of the limited angular resolution power of the array; if higher order statistics are used, then the processing can be viewed as using a Virtual Array having more sensors, and thus a larger resolution power. Of course, this holds true if the number of snapshots is greater than a threshold, and this contribution has precisely shown that this threshold is much lower than expected (a few hundreds of samples).

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.

## A Proof of the second matrix multilinearity property (19)

Assuming (**A1**)-(**A2**), the $2q$-th order statistics $C^{i_{q+1},..,i_{2q}}_{i_1,i_2,..,i_q,\boldsymbol{x}}$ defined by (2) may be described, using (1) and the multilinearity property shared by cumulants [31] [32, pp. 1-24], by

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

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 $I_1^\ell$-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 $I_2^\ell$-th component of vector $\left[\boldsymbol{a}_p^{\otimes q-\ell}\otimes(\boldsymbol{a}_p^*)^{\otimes\ell}\right]^*$ where $I_1^\ell$, $I_2^\ell$ are given by (8) and (9). Consequently, since $\left[\boldsymbol{a}_p^{\otimes q-\ell}\otimes(\boldsymbol{a}_p^*)^{\otimes\ell}\right]$ is the $p$-th column vector of matrix $\boldsymbol{\mathcal{A}}_q^\ell$ (20), equation (A.1) may be written as

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

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 $\mathrm{Diag}\Big[C^{1,1,..,1}_{1,1,...1,\boldsymbol{s}},\ C^{2,2,..,2}_{2,2,..,2,\boldsymbol{s}},\cdots,\ C^{P,P,..,P}_{P,P,..,P,\boldsymbol{s}}\Big]$, equation (A.2) may take the following expression

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

That means

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

And, since quantity $C^{i_{q+1},..,i_{2q}}_{i_1,i_2,..,i_q,\boldsymbol{x}}$ is also the $(I_1^\ell,I_2^\ell)$-th component of the $N^q\times N^q$ matrix $\boldsymbol{\mathcal{C}}^\ell_{2q,\boldsymbol{x}}$, according to (7), we finally have

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

## B Proof of proposition 4

Proposition 4 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 (21), 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

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

In other words, assuming (B.2) implies that at least two columns of $\boldsymbol{A}$ are collinear, which contradicts (**A4**). Consequently, proposition 4 is true.

## C Proof of theorem 1

Each column $\boldsymbol{b}_p$ of $[\boldsymbol{\mathcal{C}}_{2q,\boldsymbol{x}}^{\ell}]^{1/2}\boldsymbol{V}_{sol}$ is defined, according to (33), by

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

where $\xi(\cdot)$ is a bijective function of $\{1,2,\ldots,P\}$ into itself (i.e. a permutation function) and where $|\lambda_p|=|\,C_{p,p,\ldots,p,\boldsymbol{s}}^{p,p,\ldots,p}|^{1/2}$, $|\cdot|$ 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 (36), the set of matrices $\Delta_p^\ell$ may be expressed as

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

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 $\Delta_p^\ell$ indeed allows to extract the $\xi(p)$-th column vector $\boldsymbol{a}_{\xi(p)}$ of $\boldsymbol{A}$.

## D Example of FO and SixO statistics

Consider zero mean complex variables that are distributed symmetrically with respect to the origin. Thanks to the Leonov-Shiryaev formula, FO and SixO

statistics of these variables can be written as a function of moments [1]. As an example, at order 4:

$$C^{i_3,\,i_4}_{i_1,\,i_2,\,\boldsymbol{x}} \quad = \quad M^{i_3,\,i_4}_{i_1,\,i_2,\,\boldsymbol{x}}(k) - [2]M^{i_3}_{i_1,\,\boldsymbol{x}}(k)M^{i_4}_{i_2,\,\boldsymbol{x}}(k) - M_{i_1,\,i_2,\,\boldsymbol{x}}(k)M^{i_3,\,i_4}_{\boldsymbol{x}}(k) \quad \text{(D.1)}$$

and at order 6:

$$\begin{aligned} C^{i_4,\,i_5,\,i_6}_{i_1,\,i_2,\,i_3,\,\boldsymbol{x}} \quad &= \quad M^{i_4,\,i_5,\,i_6}_{i_1,\,i_2,\,i_3,\,\boldsymbol{x}}(k) \quad - \quad [3]M^{i_4}_{i_1,\,i_2,\,i_3,\,\boldsymbol{x}}(k)\, M^{i_5,\,i_6}_{\boldsymbol{x}}(k) \\ - \quad [9]M^{i_4,\,i_5}_{i_1,\,i_2,\,\boldsymbol{x}}(k)M^{i_6}_{i_3,\,\boldsymbol{x}}(k) \quad &- \quad [3]M_{i_1,\,i_2,\,\boldsymbol{x}}(k)M^{i_4,\,i_5,\,i_6}_{i_3,\,\boldsymbol{x}}(k) \quad + \\ 2[9]M_{i_1,\,i_2,\,\boldsymbol{x}}(k)M^{i_4}_{i_3,\,\boldsymbol{x}}(k)M^{i_5,\,i_6}_{\boldsymbol{x}}(k) \quad &+ \quad 2[6]M^{i_4}_{i_1,\,\boldsymbol{x}}(k)M^{i_5}_{i_2,\,\boldsymbol{x}}(k)M^{i_6}_{i_3,\,\boldsymbol{x}}(k) \end{aligned} \quad \text{(D.2)}$$

where $[d]\prod_m M^{i_{r(m)+1},\ldots,\,i_{r(m)+s(m)}}_{i_{1(m)},\ldots,\,i_{r(m)},\,\boldsymbol{x}}(k)$ denotes the McCullagh bracket notation [31]. In short, the number $d$ appearing between brackets tells that we have a sum of $d$ monomials, that can be deduced from the current one, $\prod_m M^{i_{r(m)+1},\ldots,\,i_{r(m)+s(m)}}_{i_{1(m)},\ldots,\,i_{r(m)},\,\boldsymbol{x}}(k)$, by permuting separately either superscripts or subscripts. As an illustration of this notation, an expansion of six terms is given below:

$$\begin{aligned} [6]M^{i_3}_{i_1,\,i_2,\,\boldsymbol{x}}(k)\, M^{i_5}_{i_4,\,\boldsymbol{x}}(k) = M^{i_3}_{i_1,\,i_2,\,\boldsymbol{x}}(k)\, M^{i_5}_{i_4,\,\boldsymbol{x}}(k) + M^{i_5}_{i_1,\,i_2,\,\boldsymbol{x}}(k)\, M^{i_3}_{i_4,\,\boldsymbol{x}}(k) + \\ M^{i_3}_{i_1,\,i_4,\,\boldsymbol{x}}(k)\, M^{i_5}_{i_2,\,\boldsymbol{x}}(k) + M^{i_5}_{i_1,\,i_4,\,\boldsymbol{x}}(k)\, M^{i_3}_{i_2,\,\boldsymbol{x}}(k) + M^{i_3}_{i_4,\,i_2,\,\boldsymbol{x}}(k)\, M^{i_5}_{i_1,\,\boldsymbol{x}}(k) + \\ M^{i_5}_{i_4,\,i_2,\,\boldsymbol{x}}(k)\, M^{i_3}_{i_1,\,\boldsymbol{x}}(k) \end{aligned} \quad \text{(D.3)}$$

Expressions of cumulants of order 8 as a function of moments can be found in [31] in the real case, or in [1] in the complex case. These expressions are not reproduced here.

## References


[1] L. ALBERA and P. COMON, "Asymptotic performance of contrast-based blind source separation algorithms," in *SAM 02, Second IEEE Sensor Array and Multichannel Sig. Proc. Workshop*, Rosslyn, US, August 4-6 2002, pp. 244–248.

[2] 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.

[3] 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.

[4] 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.

[5] J.-F. CARDOSO, "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.

[6] J.-F. CARDOSO and A. SOULOUMIAC, "Blind beamforming for non-gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, December 1993.

[7] ——, "Jacobi angles for simultaneous diagonalization," *SIAM Journal Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–164, 1996.

[8] P. CHEVALIER, "Optimal separation of independent narrow-band sources : Concept and Performances," *Signal Processing, Elsevier*, vol. 73, pp. 27–47, 1999.

[9] 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.

[10] 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.

[11] P. COMON, "Independent Component Analysis, a new concept ?" *Signal Processing, Elsevier*, vol. 36, no. 3, pp. 287–314, April 1994.

[12] ——, "Blind Identification and Source Separation in 2x3 Under-Determined Mixtures," *IEEE Transactions On Signal Processing*, January 2004, pp. 11–22.

[13] ——, "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.

[14] 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.

[15] 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.

[16] R. T. COMPTON, *Adaptive Antennas - Concepts and Performance.* Prentice Hall, Englewood Cliffs, New Jersey, 07632, 1988.

[17] L. De LATHAUWER, P. COMON, B. De MOOR, and J. VANDEWALLE, "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.

[18] 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.

[19] 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.

[20] L. De LATHAUWER, "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.

[21] N. DELFOSSE and P. LOUBATON, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing, Elsevier*, vol. 45, pp. 59–83, 1995.

[22] 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.

[23] 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.

[24] 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.

[25] ——, "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.

[26] 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*, vol. 5, Toulouse, France, September 3-6 2002, pp. 103–106.

[27] 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.

[28] R. A. HORN, and C. R. JOHNSON, *Topics in Matrix Analysis*. Cambridge University Press, New York, 1999.

[29] H. HYVARINEN and E. OJA, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[30] 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.

[31] P. McCULLAGH, *Tensor Methods in Statistics*. Chapman and Hall, Monographs on Statistics and Applied Probability, 1987.

[32] J. G. McWHIRTER and I. K. PROUDLER, *Mathematics in Signal Processing V*. Clarendon Press, Oxford, UK, 2002.

[33] J .G. PROAKIS, *Digital Communications*. McGraw-Hill, 1995.

[34] R. O. SCHMIDT, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenna Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[35] N. D. SIDIROPOULOS and R. BRO, "On the Uniqueness of Multilinear Decomposition of N-way Arrays," *Jour. Chemometrics*, vol.14, 2000, pp. 229–239.

[36] 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.

[37] A. TALEB, "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.

[38] 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.

[39] 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.