Blind underdetermined mixture identification by joint canonical decomposition of HO cumulants
Ahmad Karfoul, Laurent Albera, Gwénaël Birot

To cite this version:

HAL Id: hal-00593307
https://hal.archives-ouvertes.fr/hal-00593307
Submitted on 13 May 2011
**BLIND UNDERDETERMINED MIXTURE IDENTIFICATION BY JOINT CANONICAL DECOMPOSITION OF HO CUMULANTS**

Ahmad KARFOUL\(^{(1,2)}\), Student member, IEEE, Laurent ALBERA\(^{(1,2)}\), Member, IEEE, Gwénaël BIROT\(^{(1,2)}\)

**Abstract**—A new family of cumulant-based algorithms is proposed in order to blindly identify potentially underdetermined mixtures of statistically independent sources. These algorithms perform a joint CANonical Decomposition (CAND) of several higher order cumulants through a CAND of a 3-way array with special symmetries. These techniques are studied in terms of identifiability, performance and numerical complexity. From a signal processing viewpoint, the proposed methods are shown i) to have a better estimation resolution and ii) to be able to process more sources than the other classical cumulant-based techniques. 

Secondly, from a numerical analysis viewpoint, we deal with the identifiability and the convergence speed of several 3-way array decomposition procedures such as the symmetric ACDC scheme. We also show how to accelerate the iterative CAND algorithms by using differently the symmetries of the considered 3-way array. Next, from a multi-linear algebra viewpoint the paper aims at giving some insights on the uniqueness of a joint CAND of several hermitian multi-way arrays compared to the CAND of only one array. In addition, this allows us to extend the concept of Virtual Array (VA) to the case of a combination of several VAs.

**Index Terms**—Blind Underdetermined Mixture Identification (BUMI), underdetermined mixture, canonical decomposition, PARAFAC, INDSCAL, ICA, BSS.

I. INTRODUCTION

BLIND Underdetermined Mixture Identification (BUMI) is addressed in this paper. More particularly, we focus on noisy static mixtures of statistical independent sources. In such a context, the BUMI problem is defined by:

**Problem 1:** Given a vector random process \( \{ x[k] \}_{k \in \mathbb{N}} \), find a \((N \times P)\) mixing matrix \( A = [a_1, \ldots, a_P] \) potentially underdetermined such that:

\[
\forall k \in \mathbb{N}, \quad x[k] = A s[k] + \nu[k] \tag{1}
\]

where \( a_p \) is the \( p \)-th column vector of \( A \), \( s[k] = [s_1[k], \ldots, s_P[k]]^T \) is a random vector whose \( P \) components are statistically independent and where \( \nu[k] \) is a \((N \times 1)\) Gaussian noise vector, independent of \( s[k] \).

The resolution of problem 1 is of a great research interest given its importance in a great variety of signal processing applications, such as blind source separation [3], [23] or estimation of direction of arrivals [17]. Then, many solutions based on the concept of independent component analysis [11], [22] were proposed in order to solve problem 1 when \( A \) is an overdetermined mixture, say when the number \( N \) of sensors is greater or equal than the number \( P \) of sources. For instance some methods iteratively maximize the opposite of the mutual information [4], [25], [37] or the negentropy [20, chapter 8]. Other algorithms resort to cumulants, which allow for a measure of statistical independence [36], [46] less natural but easier to compute. In such a way, semi-algebraic methods were defined in order to benefit from the structure of the data cumulants as a function of the mixture. Particularly, some methods [5], [18], [48] exploit Second Order (SO) cumulants only, while others use either both SO and Fourth Order (FO) cumulants [6], [11], [32], [34], [36], [46] or FO cumulants only [1]. Contrary to the methods based on HO cumulants only [1], the methods using SO cumulants [5], [6], [11], [18], [32], [34], [36], [46], [48] attempt to identify an orthogonalized version of the mixture. In practice, the latter converge faster as a function of the number of samples provided that the noise covariance is known, while the former are asymptotically insensitive to the presence of a Gaussian noise with unknown covariance.

Nevertheless, in practical fields, the overdetermined mixture assumption does not always hold. For example, in radiocommunications when the probability of receiving more sources than sensors increases with the reception bandwidth, then it can be necessary to solve problem 1 in the underdetermined case, for which \( P > N \). Such a context may also occur in biomedical engineering, especially when the number of electrodes would be strongly reduced in order to propose an ambulatory recording system [38]. Identifiability results on underdetermined source separation were introduced a few years ago [15], [45]. They showed that underdetermined mixtures could be identified up to a trivial matrix, that is a diagonal matrix and a permutation, provided that all sources are non-Gaussian. Many methods [2], [12], [16], [17], [24], [29], [30] have been developed since then. They use implicitly [12], [17], [29], [30] or explicitly [16], [24] a CANonical Decomposition (CAND) [31] of either SO [29], FO [12], [16], [17], [30] or Higher Order (HO) [2], [24] cumulants of the data. They decompose either one cumulant array [2], [16], [24], [30] or several ones at a given statistical order [12], [29]. In fact, CAND is an extension of the Singular Value Decomposition (SVD) to the case of multi-way arrays without requiring any orthogonality constraint. Other algorithms [14],...
[44] take advantage of pseudo-cumulants, which are given by the (weighted) derivatives of the second characteristic function at some points different from the origin. However, all previous methods have some limits. Indeed, under some computational cost constraints and for \( N \geq 3 \) they do not succeed in exploiting all the information contained in the statistical tools they used, except the FOOBI2 method [30]. But the latter cannot process more than \( N_q^4 (N_q^4 \leq N^2) \) sources, where \( N_q \) is the number of sensors, which virtually appear when a Fourth Order (FO) cumulant is used in practical contexts such as radiocommunications [10].

In order to overcome these limitations and due to the attractive properties of cumulants which make them easy to handle, we propose in this paper a new family of BUMI cumulant-based algorithms. These algorithms perform a joint CAND of several HO cumulants through a CAND of a special 3-way array with special symmetries. These techniques are studied in terms of identifiability, performance and numerical complexity. From a signal processing viewpoint, the proposed methods are shown i) to have a better estimation resolution and ii) to be able to process more sources than the other classical cumulant-based techniques. Secondly, from a numerical analysis viewpoint, we deal with the identifiability and the convergence speed of several 3-way array decomposition procedures such as the symmetric ACDC scheme [49]. We also show how to accelerate the iterative CAND algorithms by using differently the symmetries of the considered 3-way array. Next, from a multi-linear algebra viewpoint the paper aims at giving some insights on the uniqueness of a joint CAND of several HO cumulants through a CAND of a special multi-linear operator. Thus, we propose to rewrite this formula hereafter in the case of delayed cumulants. To this end, let’s give some notations and definitions.

**Definition 1:** A partition of a set \( \Omega(q) = \{1, 2, \cdots, q\} \) of integers is a set of \( M (1 \leq M \leq q) \) nonempty disjunctive subsets whose union is equal to \( \Omega(q) \).

Now, let \( \Omega_{\text{m}}^{(j)}(M, q) \) be the \( m \)-th \((1 \leq m \leq M)\) subset of the \( j \)-th \((1 \leq j \leq J(M, q))\) partition of size \( M \) of \( \Omega(q) \). This notation is illustrated in the following example:

**Example 1:** The different partitions of set \{1, 2, 3\} are:

\[
\begin{align*}
\boldsymbol{\Omega}_{1}^{(1)}(1, 3) &= \{\{1, 2, 3\}\} \\
\boldsymbol{\Omega}_{1}^{(2)}(2, 3), \boldsymbol{\Omega}_{2}^{(2)}(2, 3) &= \{\{1, 2\}, \{3\}\} \\
\boldsymbol{\Omega}_{1}^{(3)}(2, 3), \boldsymbol{\Omega}_{2}^{(3)}(2, 3) &= \{\{1, 3\}, \{2\}\} \\
\boldsymbol{\Omega}_{1}^{(1)}(3, 3), \boldsymbol{\Omega}_{1}^{(1)}(3, 3), \boldsymbol{\Omega}_{3}^{(1)}(3, 3) &= \{\{1\}, \{2\}, \{3\}\}.
\end{align*}
\]

with \( J(1, 3) = J(3, 3) = 1 \) and \( J(2, 3) = 3 \) different partitions.

Thus, we have the following proposition:

**Proposition 1:** Let \( \{\mathbf{x}[k]\}_{k \in \mathbb{N}} \) be an \( N \)-dimensional random process. Its \( q \)-th order delayed cumulants are given by:

\[
\begin{align*}
\text{Cum}\{x_n[k], x_n[k-\tau_1], \cdots, x_n[k-\tau_{q-1}]\}^{(q-1)} = \\
\sum_{M=1}^{q} (-1)^{\lfloor M-1 \rfloor} \int \left[ \left. \prod_{j=1}^{q} \mathbb{E}[x_n[k-\tau_{\omega-1}]^{(\omega-1)}] \right| \Omega_{JM}(M, q) \right]
\end{align*}
\]

where \( \mathbb{E}[.] \) is the mathematical expectation, \( x_n[k-\tau] \) is the \( n \)-th component of vector \( \mathbf{x}[k-\tau] \), \( \tau \) is a set \( \{\tau_1, \cdots, \tau_{q-1}\} \) of \( q-1 \) delays, \( \xi(\omega) \) equals \( \pm 1 \) with the conventions, \( \xi(0) = 1 \), \( \tau_0 = 0 \), \( x^1 = x \) and \( x^{-1} = x^* \) denoting by * the complex conjugate operator.

It is noteworthy that the estimation of cumulants using equation (2) is appropriate when the processed sources are stationary. But, in practice, sources are rarely so. In such a case, time averaged cumulants have to be used instead of (2), such as

\[
\begin{align*}
\gamma^{n_1, \cdots, n_{2q}}[\mathbf{\tau}] = \langle \text{Cum}\{x_n[k], x_n[k-\tau_1], \cdots, x_n[k-\tau_{q-1}]\}^{(q-1)} \rangle
\end{align*}
\]

where \( \langle . \rangle \) is the discrete-time temporal mean operation [2]. Such a statistics enjoy the multilinearity property and they can be used as a measure of independence. Thus for sake of convenience, time averaged cumulant are considered throughout this paper.

**III. TOWARD A MULTI-WAY ARRAY FORMULATION**

This section aims at showing how problem 1 can be solved using a joint CAND of several HO cumulant arrays through a multilinear algebra framework. Recall that multilinear algebra is defined as the algebra of \( q \)-way arrays \((q > 2)\), that is, the arrays whose elements are addressed by more than two indices. It is noteworthy that multilinear algebra can be related to tensor algebra through a multilinear map, called Segre map [13]. Now some definitions are necessary in order to understand how the BUMI problem can be reformulated in a multilinear algebra framework. First let’s begin by introducing a very useful product operator:

**Definition 2:** The outer product \( \mathbf{T} = \mathbf{u}^{(1)} \circ \cdots \circ \mathbf{u}^{(q)} \) of \( q (q \geq 2) \) vectors \( \mathbf{u}^{(i)} \in \mathbb{C}^N \), \((1 \leq i \leq q)\) is a rank-1 array
He showed that the CAND of a multi-way array of size \( N^q \times \cdots \times N_q \) whose elements are defined by \( T_{n_1, \ldots, n_q} = u_{n_1}^{(1)} \cdots u_{n_q}^{(q)} \).

Then the matrix rank notion can be extended to the case of multi-way arrays [26], [27], as following:

**Definition 3:** The rank of an array \( T \in \mathbb{C}^{N_1 \times \cdots \times N_q} \), denoted by \( \text{rk}(T) \), is the minimal number of rank-1 arrays belonging to \( \mathbb{C}^{N_1 \times \cdots \times N_q} \) that yield \( T \) in a linear combination.

Despite the similarities between matrices and multi-way arrays, the rank of the latter ones may exceed their dimensions. Now, let’s see how the matrix SVD can be extended to the case of multi-way arrays.

**Definition 4:** The CAND of a \( q \)-way array \( T \in \mathbb{C}^{N_1 \times \cdots \times N_q} \) is the linear combination of \( P = \text{rk}(T) \) rank-1 arrays belonging to \( \mathbb{C}^{N_1 \times \cdots \times N_q} \), that is:

\[
T = \sum_{p=1}^{P} \lambda_p \ u_p^{(1)} \circ u_p^{(2)} \circ \cdots \circ u_p^{(q)}
\]

where each integer \( p \), \( \lambda_p \) is a complex number and vectors \( u_p^{(i)} \) are complex \( N_i \)-dimensional vectors.

Nevertheless, the \( q \) matrices \( U^{(i)} = [u_1^{(i)}, \ldots, u_{n_i}^{(i)}] \), given by (4) and well-called loading matrices of \( T \), are not necessarily unitary contrary to the matrix SVD. Figure 1 shows an example of CAND in the case of a 3-way array. In addition, equation (4) shows that the \( q \) vectors \( u_p^{(i)} \) can be arbitrarily scaled as long as their product remains the same. Also the rank-1 terms can be permuted without modifying the sum. Then we call the CAND of a \( q \)-way array \( q \geq 2 \) unique when it is only subject to these trivial indeterminacies. Important results on CAND uniqueness were obtained by Kruskal [26]. He showed that the CAND of a \( q \)-way array \( T \) is essentially unique when the following sufficient condition is satisfied:

\[
2P + q - 1 \leq \sum_{i=1}^{q} k(U^{(i)})
\]

where \( k(T) \) denotes the kruskal rank of \( T \) defined by the maximal number \( r \) such that any set of \( r \) columns of \( T \) is linearly independent [29]. Kruskal’s condition was firstly proposed for real 3-way arrays. Later it was extended to arrays with order higher than three [40] and for complex HO arrays [41]. The CAND uniqueness of specific 3-way arrays was also studied. For instance the case when one loading matrix has a full column rank [21] and the case where the HO array is tall in one direction [31]. In practice, the computation of CAND may require to handle unfolding matrices or vectors of multi-way arrays. A useful way to unfold square even order multi-way arrays is described below.

**Definition 5:** Let \( T \) be a square \( 2q \)-way \((q > 1)\) array of dimension \( N \), and let \( \ell \) be the integer part of \( q/2 \). Then the \((i,j)\)-th component of the unfolding matrix \( \text{mat}_1(T) \) of size \((N^q \times N^q)\) is given by:

\[
\text{mat}_1(T)_{i,j} = T_{n_{q-\ell} \cdots n_{q-\ell+1} \cdots n_q \cdots n_1} + \cdots + (n_{q-\ell} - 1)N^q + (n_{q-\ell+1} - 1)N^q - \cdots + (n_{\ell-1} - 1)N^q - + (n_{\ell} - 1)N^q - + (n_q - 1)N^q + n_q.
\]

For instance, let \( T \) be a square 6-way array of dimension \( N \). Then \( \ell = 1 \) and the \((i,j)\)-th component of the unfolding matrix \( \text{mat}_1(T) \) is given by:

\[
\text{mat}_1(T)_{i,j} = T_{n_1 n_2 n_3 n_4 n_5 n_6} + \cdots + (n_1 - 1)N^2 + (n_2 - 1)N + n_6 + (n_4 - 1)N^2 + (n_5 - 1)N + n_3.
\]

Another way to transform higher order arrays into matrices is presented hereafter and is illustrated by figure 2.

**Definition 6:** Let \( T \in \mathbb{C}^{N_1 \times \cdots \times N_q} \) be a \( q \)-way \((q > 2)\) array. Then the \((n_1, n_1)\)-th component of the unfolding matrix \( \text{mat}_2(T) \in \mathbb{C}^{N_1^2} \) associated to the \( i \)-th mode \( 1 \leq i \leq q \) of \( T \) is given by:

\[
\text{mat}_2(T)_{n_1,n_1} = T_{n_1 \cdots n_1} + \cdots + (n_1 - 1)N_1 + \cdots + (n_1 - 1)N_1 + (n_1 - 1)N_1 + \cdots + n_{i-1}.
\]

In addition, let’s introduce the following \( q \)-way array-to-vector transformation:

**Definition 7:** Let \( T \in \mathbb{C}^{N_1 \times \cdots \times N_q} \) be a \( q \)-way \((q > 1)\) array. Then the components of the \( N_1 \times \cdots \times N_q \)-dimensional vector \( \text{vec}(T) \) are defined by:

\[
\text{vec}(T)_{(n_1 - 1)N_2 \cdots N_q + (n_2 - 1)N_3 \cdots N_q + \cdots + n_q} = T_{n_1 \cdots n_q}.
\]

Conversely, let \( \text{unvec} \) be the inverse operator such as \( \text{unvec}(\text{vec}(T)) = T \). Hermitian \( q \)-way arrays are considered in this paper, hence the following definition:

![Fig. 1. CAND of a rank-P 3-way array](image-url)
Definition 8: A complex \( q \)-way \((q \geq 2)\) array \( \mathbf{T} \) is called hermitian when its CAND takes the following form:

\[
\mathbf{T} = \sum_{p=1}^{P} \lambda_p \mathbf{u}^{(1)}_p \otimes \mathbf{u}^{(2)}_p \otimes \cdots \otimes \mathbf{u}^{(q)}_p
\]

(10)

where the \( P \) values \( \lambda_p \) are real and where application \( \xi(i) \) is defined in proposition 1, that is, if the \((P \times P)\) diagonal matrix \( \mathbf{A} = \text{diag}\{[\lambda_1, \ldots, \lambda_P]\} \) is real and if all the loading matrices of \( \mathbf{T} \) are equal within a conjugate transformation.

Inserting the observation model (1) into (2) and using i) the statistical independence of the \( P \) source components, ii) the statistical independence between the sources and the noise and iii) the noise Gaussian distribution, we get for any integer \( q \geq 2 \):

\[
\mathcal{C}_{2q} \{x|T\} = \sum_{p=1}^{P} \mathcal{C}_{2q,sp} \{x|T\} \mathbf{a}^{cq}_p \otimes \mathbf{a}^{cq}_p
\]

(11)

where \( \mathcal{C}_{2q,sp} \{x|T\} = \text{Cum}\{s_p[k], s_p[k - \tau_1], \ldots, s_p[k - \tau_{q-1}], s_p[k - \tau_q]^*, \ldots, s_p[k - \tau_{2q-1}]^*\} \) and where \( \mathbf{a}^{cq}_p \) is the outer product of \( q \) vectors \( \mathbf{a} \). Formula (11) conveys the multilinearity property enjoyed by cumulant and moment arrays [23]. Then, provided that the \( P \) sources have at least \( M \) non-zero \( 2q \)-th order delayed marginal cumulants, problem 1 can be reformulated as following:

**Problem 2:** Given \( M \) \( 2q \)-th \((q \geq 2)\) order cumulant arrays \( \mathcal{C}_{2q} \{x|\tau(m)\} \) of \( \{x[k]\}_{k \in \mathbb{Z}} \) associated to a set \( \{\tau(m)\} \) of \((2q - 1)\)-dimensional time lags, find its joint CAND.

But, is it really possible to compute such a joint CAND, especially when \( P \) is greater than the maximum number of sources which can be processed by the classical cumulant-based techniques [1], [2], [5], [6], [11], [12], [14], [16]–[18], [24], [29], [30], [32], [34], [36], [46], [48]? Yes, it suffices to choose the cumulant order \( 2p \) greater or equal than four and six using the approaches proposed in sections IV-B and IV-A, respectively. Nevertheless, for the sake of simplicity and readability, we limited the presentation of our methods and their identifiability study to a fixed statistical order \( 2q \) equal to six and identical for all the proposed algorithms. An extension to all other order would be then straightforward from sections IV and V.

**IV. TOWARD A 3-WAY ARRAY DECOMPOSITION**

We present in this section a new family of methods, named SIBI (Sixth order Blind mixture Identification), in order to solve problem 2 when \( q = 3 \). This family is essentially composed of two classes of techniques, namely the SAD (Semi-Algebraic Decomposition) and ALS (Alternate Least Square) SIBI approaches. Both classes can be described through two steps.

The first step consists in identifying the \( P \) rank-1 arrays \( \mathbf{A}^{(3)}_p \) given by:

\[
\forall 1 \leq p \leq P, \quad \mathbf{A}^{(3)}_p = \mathbf{a}_p \otimes \mathbf{a}_p \otimes \mathbf{a}_p^* \quad (12)
\]

from a CAND of a special 3-way array \( \mathbf{H}_x \) of size \((N^3 \times N^3 \times M)\) whose \((\ell_1, \ell_2, m)\)-th entry is given by:

\[
\mathbf{H}_{\ell_1, \ell_2, m, x} = \mathbf{C}_{n_1, n_2, n_3, x}[\tau(m)] = \text{mat}_1(\mathbf{C}_{6,x}[\tau(m)])
\]

(13)

where \( \ell_1 = n_1 + (n_2 - 1)N + (n_3 - 1)N^2 \) and \( \ell_2 = n_3 + (n_2 - 1)N^2 \). On the other hand, \( \mathbf{H}_x \) can also be built by stacking up the \( M \) hexacovariance matrices \( \mathbf{C}_{6,x}[\tau(m)] \) of size \((N^3 \times N^3)\) in the third direction. From (11) and (13), we get:

**Proposition 2:** Given the \( P \) \( M \)-dimensional vectors \( b_p = [\mathbf{c}_{6,p}[\tau(1)], \ldots, \mathbf{c}_{6,p}[\tau(M)]]^T \), \( \mathbf{H}_x \) has the following CAND:

\[
\mathbf{H}_x = \sum_{p=1}^{P} \mathbf{a}^{(3)}_p \otimes \mathbf{a}^{(3)}_p \otimes \mathbf{a}^{(3)}_p \quad (14)
\]

where \( \mathbf{a}^{(3)}_p = \mathbf{a}_p \otimes \mathbf{a}_p \otimes \mathbf{a}_p^* \) with \( \otimes \) the Kronecker product operator and where \( \mathbf{A}^{(3)} = [\mathbf{a}^{(3)}_1, \ldots, \mathbf{a}^{(3)}_P] \), \( \mathbf{A}^{(3)}^* \) and \( \mathbf{B} = [b_1, \ldots, b_P] \) are the three loading matrices of \( \mathbf{H}_x \).

Once the \( P \) \( N^3 \)-dimensional vectors \( \mathbf{a}^{(3)}_p \) have been identified from the CAND of \( \mathbf{H}_x \), each array \( \mathbf{A}^{(3)}_p \) is found by computing \( \mathbf{A}^{(3)}_p = \text{unvec}(\mathbf{a}^{(3)}_p) \).

The second step consists in canonically decomposing each rank-1 array \( \mathbf{A}^{(3)}_p \) in order to identify each column vectors \( \mathbf{a}_p \) of \( \mathbf{A} \). It is noteworthy that the way to implement this second step is the same for all SIBI methods and is based on the following proposition:

**Proposition 3:** Each \((N \times N^2)\) unfolding matrix \( \mathbf{A}^{(3)}_p = \text{mat}_2(\mathbf{A}^{(3)}_p) \) is the horizontal concatenation of \( N \) \((N \times N)\) matrices, which are of the form \( \gamma \mathbf{a}_n \mathbf{a}_n^* \) with \( \gamma \in \mathbb{C} \).

In fact, \( \gamma \) is equal to the \((n, p)\)-th component, \( A_{n,p} \), of \( \mathbf{A} \) when the \( n \)-th horizontal matrix block of \( \mathbf{A}^{(3)}_p \) is concerned. Consequently, the eigenvector jointly associated with the largest eigenvalue of each of the \( N \) matrix blocks of \( \mathbf{A}^{(3)}_p \) allows to estimate \( \mathbf{a}_p \). Such an eigenvector can be computed using a joint diagonalization procedure [7], [8], [49], [50].
A. The SAD-SIBI approach

A semi-algebraic solution is proposed hereafter in order to find decomposition (14). It consists in handling the \( M \) frontal slices \( \mathcal{H}_a(:, :, m) = C_{6,x} \tau^{(m)} \) instead of \( \mathcal{H}_a \) itself. Then according to (6) and (11) we can easily show that:

Proposition 4: Each hexacovariance matrix \( C_{6,x} \tau^{(m)} \) has the following algebraic structure:

\[
C_{6,x} \tau^{(m)} = A^{(3)} \zeta_{6,x} \tau^{(m)} A^{(3)\#}
\]

(15)

where \( \zeta_{6,x} \tau^{(m)} = \text{diag}(\{C_{6,x,1}\tau^{(m)}, \ldots, C_{6,x,p}\tau^{(m)}\}) \).

Since \( A^{(3)} \) is not generally orthogonal, equation (15) does not really correspond to the eigenvalue decomposition of \( C_{6,x} \tau^{(m)} \), except when this latter has been prewhitened. Let’s see how to perform such a prewhitening. Under assumptions 1 and 2 given by:

Assumption 1: The \((N^3 \times P)\) matrix \( A^{(3)} \) is full column rank.

and:

Assumption 2: The \((M \times P)\) matrix \( B \) is full column rank respectively, the following proposition holds:

Proposition 5: There is at least one rank-\( P \) positive semidefinite matrix \( \overline{C}_{6,x} \tau \), which is a linear combination of the \( M \) matrices \( C_{6,x} \tau^{(m)} \).

Its proof is given in appendix A. In some practical contexts, for example when the \( M \) non-delayed source marginal Sixth Order (SixO) cumulants have the same sign \( \epsilon \), such as in radio communications, a simple linear combination is given by \( \overline{C}_{6,x} = \epsilon C_{6,x}[0] \) where \( 0 = \{0,0,0,0,0\} \). Otherwise, in the general case, we have to resort to numerical methods such that those presented in [47], [51] to compute \( \overline{C}_{6,x} \tau \). In addition, it is noteworthy that matrix \( \overline{C}_{6,x} \tau \) may be non-hermitian and then should be replaced by \( \overline{C}_{6,x} \tau + \overline{C}_{6,x}\tau^* \) in order to ensure the existence of its eigen-value decomposition. For the sake of simplicity, we will assume in the sequel that matrix \( \overline{C}_{6,x} \tau \) is hermitian. Then proposition 5 ensures that \( \overline{C}_{6,x} \tau \) has a square root, which can be computed as following:

\[
\overline{C}_{6,x}^{1/2} = E_x A_x^{1/2}
\]

(16)

where \( E_x \) is the \((N^3 \times P)\) unitary matrix whose columns are the eigenvectors associated with the \( P \) largest eigenvalues of \( \overline{C}_{6,x} \) sorted in the diagonal matrix \( A_x \). Since two square roots of a positive semidefinite matrix are equal up to a unitary matrix, we have:

\[
\overline{C}_{6,x}^{1/2} = A^{(3)} \zeta_{6,x}^{1/2} U_0^H
\]

(17)

where \( \zeta_{6,x}^{1/2} \), \( \zeta_{6,x}^{1/2} \), \( U_0 \) are the linear combination described in proposition 5 of the \( M \) matrices \( \zeta_{6,x} \tau^{(m)} + \zeta_{6,x} \tau^{(m)*} \)/2, a square root of the positive definite diagonal matrix \( \zeta_{6,x} \), and a unitary matrix. Consequently, the inverse matrix, \( \overline{C}_{6,x}^{-1/2} \), of \( \overline{C}_{6,x} \) allows for a prewhitening of each matrix \( C_{6,x} \tau^{(m)} \), say each matrix \( \overline{C}_{6,x} \tau^{(m)} = \overline{C}_{6,x}^{-1/2} C_{6,x} \tau^{(m)} (\overline{C}_{6,x}^{-1/2})^H \) is unitarily similar to a diagonal matrix indexed by \( m \) as described by the following equation:

\[
\overline{C}_{6,x} \tau^{(m)} = U_0 D \tau^{(m)} U_0^H
\]

(18)

where \( D \tau^{(m)} = \zeta_{6,x} \tau^{(m)} \zeta_{6,x}^{-1} \) is diagonal. Thus, a joint orthogonal diagonalization with the well-known semi-algebraic JAD algorithm [7] allows for an identification of \( U_0 \) up to a unitary trivial matrix. Nevertheless, \( U_0 \) is identifiable provided that, according to the results of [5], the \((2q-1)\)-dimensional delays \( \tau^{(m)} \) are chosen such that:

Assumption 3: For each couple \((p,q)\) of sources, there exists at least one set \( \tau^{(m)} \) of delays for which:

\[
\zeta_{6,s,p} \tau^{(m)} / \zeta_{6,s,p} \neq \zeta_{6,s,q} \tau^{(m)} / \zeta_{6,s,q}
\]

(19)

It is noteworthy that assumption 3 is true under assumption 2. Indeed, assumption 2 implies that the \((M \times P)\) matrix \( B \zeta_{6,x} \) is full column rank and so that all columns of \( B \zeta_{6,x} \) are different. Hence equation (19). Under assumption 2, the JAD algorithm diagonalizes the \( M \) matrices \( C_{6,x} \tau^{(m)} \) by minimizing the following criterion:

\[
\Psi_1(U) = \sum_{m=1}^{M} \sum_{p \neq q=1}^P \| (U_d^{(m)} C_{6,x} \tau^{(m)} U_p)^p,q \|^2
\]

(20)

with respect to \((w.r.t)\) to a unitary matrix \( U \) expressed as the product of Givens rotations [19]. Once matrix \( U_0 \) has been identified, matrix \( A^{(3)} \) can be computed up to a trivial matrix multiplying matrix \( \overline{C}_{6,x} \) (17) by the estimate of \( U_0 \). This approach will be referred to as SAD-SIBI JAD in the sequel.

B. The ALS-SIBI approaches

Another way to find decomposition (14) consists in iteratively and alternatively looking for the three loading matrices \( A^{(3)} \), \( A^{(3)*} \), and \( B \) of \( \mathcal{H}_a \) in the least square sense. More precisely, from proposition 2 and definition 6, we have:

Proposition 6: Given the three unfolding matrices \( H_x^{(1)} = \text{mat}_2(\mathcal{H}_a), H_x^{(2)} = \text{mat}_2(\mathcal{H}_a) \) and \( H_x^{(3)} = \text{mat}_2(\mathcal{H}_a) \) of \( \mathcal{H}_a \), we get \( H_x^{(1)} = A^{(3)} A^{(3)*} B \). \( H_x^{(2)} = A^{(3)*} B \) and \( H_x^{(3)} = B A^{(3)*} \) where \( \otimes \) is the Khatri-Rao product operator (columns-wise kroner product) [42].

So, the following criterion can be minimized w.r.t. each of the three matrices \( T^{(1)}, T^{(2)} \) and \( T^{(3)} \) in order to compute the loading matrices \( A^{(3)}, A^{(3)*} \) and \( B \) of \( \mathcal{H}_a \) up to a trivial matrix:

\[
\Psi_3(T^{(1)}, T^{(2)}, T^{(3)}) = \| H_x^{(1)} - T^{(1)} (T^{(2)} \otimes T^{(3)})^\top \|^2_F
\]

(21)

using the Alternate Least Square (ALS) procedure [42] where \( \| \cdot \|^2_F \) denotes the Frobenius norm, giving rise to the ALS-SIBM UNS method. In fact, at each iteration \( it \), we compute the three estimates \( T^{(i)} \) (\( 1 \leq i \leq 3 \)) at a time by fixing the others as described below:

\[
T^{(1)}_it = H_x^{(1)} (T^{(2)}_it \otimes T^{(3)}_it)^\top
\]

\[
T^{(2)}_it = H_x^{(2)} (T^{(1)}_it \otimes T^{(3)}_it)^\top
\]

\[
T^{(3)}_it = H_x^{(3)} (T^{(1)}_it \otimes T^{(2)}_it)^\top
\]

(22)

where \( T^{(i)} \) denotes the pseudo-inverse of \( T \). Consequently, matrices \( T^{(1)}_it, T^{(2)}_it \) and \( T^{(3)}_it \) should converge to matrices
V In fact the ACDC algorithm [49] minimizes the following procedures [39, Appendix 7.2] exploit such symmetries. It is noteworthy that one is entitled to expect some convergence cycles in which all factors will evolve in the same direction [39] and consequently a slow convergence of the ALS-SIBI\textsubscript{UNS} method. An improved version of the ALS scheme, named Line Search (LS) ALS, was proposed (see [39] and the references therein) in order to exit from these convergence cycles. This process is based on the relaxation factor concept, which consists in extrapolating the estimated loading matrices many iterations ahead. Then, we decided to apply this LS procedure to the ALS-SIBI\textsubscript{UNS} method, giving rise to the LSALS-SIBI\textsubscript{UNS} algorithm. The extrapolation of the \( t \)-th estimated loading matrix of \( \mathbf{H}_v \) is given by:

\[
\mathbf{T}^{(t)}_{\text{new}} = \mathbf{T}^{(t)}_{it-2} + \beta^{(t)}(\mathbf{T}^{(t)}_{it-1} - \mathbf{T}^{(t)}_{it-2})
\]

where \( \mathbf{T}^{(t)}_{\text{new}} \) is the matrix that will be used in the \( it \)-th iteration instead of \( \mathbf{T}^{(t)}_{it-1} \), the term \( 2(\mathbf{T}^{(t)}_{it-1} - \mathbf{T}^{(t)}_{it-2}) \) defines the direction of the cycle and \( \beta^{(t)} \) is the relaxation factor associated to the \( it \)-th loading matrix. In practice, the triplet \((\beta^{(1)}, \beta^{(2)}, \beta^{(3)})\) of relaxation factors is chosen such that \( \beta^{(1)} = beta^{(2)} = \beta^{(3)} = it^{1/n} \) where \( n \) is a fixed integer [39].

Recent works [39] showed that the convergence speed of the LSALS procedure could be also increased, especially when loading matrices are strongly correlated. This induced an enhanced version of the LSALS algorithm, named Enhanced Line Search (ELS)ALS, in the real [39] and the complex [35] fields. More precisely, the ELS procedure consists, at each iteration \( it \) of the LSALS method, in finding the optimal triplet \((\beta^{(1)}, \beta^{(2)}, \beta^{(3)})\). So, when the ELS scheme is applied to LSALS-SIBI\textsubscript{UNS} giving rise to the ELSALS-SIBI\textsubscript{UNS} method, the following criterion:

\[
\Psi^{(2)}_{3}(\beta) = \|H^{(1)} - G^{(1)}(G^{(2)})^{T}(G^{(3)})^{T})\|_F^2
\]

is minimized at each iteration \( it \) w.r.t. \( \beta^{(1)}, \beta^{(2)} \) and \( \beta^{(3)} \) where \( G^{(i)} = \mathbf{T}^{(i)}_{it-2} + \beta^{(i)}(\mathbf{T}^{(i)}_{it-1} - \mathbf{T}^{(i)}_{it-2}) \). The ELS scheme allows to attain the final solution of a given cycle in only one step. Note that the ELS procedure was implemented in the case of three equal relaxation factors only, which is suboptimal but less time-consuming.

However, by applying the previous ALS-like procedures to \( \mathbf{H}_v \), it amounts to assume a total independence between the loading matrices, which is clearly wrong in our case. Thus an alternative way of the ALS minimization would consists in using the symmetries that exist between some loading matrices of the considered multi-way array. In our case, the first and the second loading matrices of \( \mathbf{H}_v \) are equal up to a complex conjugate as shown in (14). The ACDC-like procedures [49], [39, Appendix 7.2] exploit such symmetries. In fact the ACDC algorithm [49] minimizes the following weighted least square criterion w.r.t. \( V \) and the \( M \) diagonal \((P \times P)\) matrices \( \Lambda[\tau^{(m)}] \):

\[
\Psi^{(2)}_{2}(V,A) = \sum_{m=1}^{M} w_m \left\| C_{x,m}\tau^{(m)} - VA[\tau^{(m)}]V^T \right\|_F^2
\]

where \( w_m \in \mathbb{R}^{+\times} \). More particularly, the last minimization problem is devised into two steps. The first one, named Alternating Columns (AC), looks for the minimizer of (25) w.r.t. a selected column vector \( v_p \) of \( V \) keeping the \( p-1 \) other column vectors and the \( M \) matrices \( \Lambda[\tau^{(m)}] \) fixed. The second step, called Diagonal Centers (DC), minimizes equation (25) w.r.t. the diagonal matrices \( \Lambda[\tau^{(m)}] \) keeping \( V \) fixed. So, the ACDC algorithm alternates, until convergence, between the \( P \) consecutive AC sweeps, one by column vector of \( V \), and the DC step. Note that, an efficient way to initialize the ACDC algorithm was recently proposed [50]. In the following, we will refer to as the ALS-SIBI\textsubscript{ACDC} method when the ACDC procedure is used. But, are the ACDC-like procedures [49], [39, Appendix 7.2] the only way to decompose \( \mathbf{H}_v \) in a least square sense using its symmetries? No, there is another way such as the solution that we propose hereafter. Furthermore, we combined this new 3-way array CAND method with the ELS scheme in order to accelerate its convergence, giving rise to the ELSALS\textsubscript{SYM} procedure. In fact, the latter is mainly based on the following proposition:

**Proposition 7:** Let \( \mathbf{H}_v \) be the 3-way array of size \((N^3 \times N^3 \times P)\) whose third \((P \times N^6)\) unfolding matrix is given by \( \mathbf{T}^{(3)}_{it} = B^H \mathbf{T}^{(3)}_{it} \). Then the \( p \)-th \((N^3 \times N^3)\) frontal slice of \( \mathbf{H}_v \), \( \mathbf{H}(i,:,:, p) \), is a rank-1 matrix equal to \( \mathbf{A}(3) \mathbf{B} \mathbf{C}(3) \).

The proof is straightforward. Indeed, from proposition 2 we get \( \mathbf{H}^{(3)}_{\text{it}} = \mathbf{P}(A(3) \odot A(3))^{\dagger} \), which implies that the \( p \)-th frontal slice of \( \mathbf{H}_v \) is equal to \( A(3) \mathbf{B} \mathbf{C}(3) \mathbf{I}(p,:) \mathbf{A}(3)^{\dagger} \).

where \( \mathbf{I}(p,:) \) is the \( P \)-th row of the \((P \times P)\) identity matrix \( \mathbf{I}_P \). Consequently, the eigenvector associated with the largest eigenvalue of \( \mathbf{H}(i,:,:, p) \) allows for an identification of the \( p \)-th column vector of \( A(3) \) up to a scalar factor. Now, let’s see how the ELSALS\textsubscript{SYM} scheme uses this result. First the estimate, \( \mathbf{T}^{(3)}_{it-1} \), of the third loading matrix, \( B \), of \( \mathbf{H}_v \) is computed at iteration \( it - 1 \) from equation (22). Next, using proposition 7, we compute the \((P \times N^6)\) matrix \( \mathbf{T}^{(3)}_{it-1} = \mathbf{T}^{(3)}_{it-1} \mathbf{H}^{(3)}_{\text{it}} \) and build the 3-way array \( \mathbf{T}^{(3)}_{it-1} \) of size \((N^3 \times N^3 \times P)\) whose third \((P \times N^6)\) unfolding matrix is equal to \( \mathbf{T}^{(3)}_{it-1} \).

So, the eigenvector associated with the largest eigenvalue of the \( p \)-th frontal slice of \( \mathbf{T}^{(3)}_{it-1} \) gives the \( p \)-th column vector of \( \mathbf{T}^{(3)}_{it-1} \). Thus, \( \mathbf{T}^{(3)}_{it-1} \) is computed as the conjugate of \( \mathbf{T}^{(3)}_{it-1} \).

Eventually, the ELS procedure is used in order to accelerate the convergence of the proposed algorithm. Note that all these steps are repeated until convergence.

The convergence speed of the five proposed ALS-SIBI approaches, namely ALS-SIBI\textsubscript{UNS}, LSALS-SIBI\textsubscript{UNS}, ELSALS-SIBI\textsubscript{UNS}, ALS-SIBI\textsubscript{ACDC} and ELSALS-SIBI\textsubscript{SYM}, is studied now. More precisely, we display the median value of \( \Psi^{(2)}_{3} \) (21) over the independent trials as a function of the number of iterations at the output of the five ALS-SIBI methods. In the last method, ALS-SIBI\textsubscript{SYM}, the 3-way array \( \mathbf{H}_v \) built from equation (14) with \( M = 16 \)
An identifiability study of the SIBI family is proposed hereafter under assumptions 1-3. In particular, the $M$ time lag vectors $\tau^{(m)} = [\tau_1^{(m)}, \tau_2^{(m)}, \tau_3^{(m)}, \tau_4^{(m)}, 0]$ and the estimated SixO delayed marginal cumulants of $P = 3$ angularly close QPSK sources. The latter sources are assumed to impinge on a Uniform Linear Array (ULA) of $N = 2$ sensors. In addition, all the considered approaches are initialized using the truncated HOSVD procedure [28]. Figure 3 shows, on the one hand, the good contribution of the ELS procedure to the truncated HOSVD procedure [28]. In this case, the previous property does not hold generally. Intrinsically, we have:

$$r k_{\text{max}}(H_x^{(3)}) \leq r k_{\text{max}}(B(A^{(3)} \odot A^{(3)*})^T)$$

where $(M \times N^6) H_x^{(3)}$ denotes the unfolding matrix associated to the third direction of $H_x(14)$. For $M \geq N^6$, the maximum rank of $H_x^{(3)}$ is generically equal to $N^6$. Recall that a property is well-called generic when it holds everywhere except for a set of Lebesgue measure 0. However, in some applications such as in radiocommunication contexts, matrix $H_x^{(3)}$ may belong to this set. In this case, the previous property does not hold generally. Therefore, we have:

$$P_{\text{max}} = r k_{\text{max}}(A^{(3)}) = N_6$$

So from a BUMI viewpoint and according to equations (27) and (30) the SIBI methods can process more than $N_4$ sources. Note that $N_4$ is the maximum number of sources the classical BUMI cumulant-based methods [2], [12], [16], [17], [24], [29], [30] can process. Table I gives the number $P_{\text{max}}$ as a function of $N$ for a ULA and for several cumulant-based methods. From a numerical analysis viewpoint, we show that: i) the SAD scheme is limited to 3-way arrays whose rank is lower than their smallest dimension. ii) Both the ACDC and ELSALS_SYMM procedures succeed generically in performing the CAND of a rank-$N^2 (N \times N \times M)$ array if $M > N^2$. Consequently, a more relaxed uniqueness condition can be obtained when a joint CAND of several hermitian multiway arrays is performed. Since according to [9] $N_6 = r k_{\text{max}}(C_{\ell} [0])$, equation (30) shows that a maximal rank of $N_{12} (> N_6)$ can be achieved when $C_{\ell} [0]$ and at least $N_{12} - 1$ other arrays $C_{\ell} [\tau]$ with the same loading matrices are simultaneously considered. This result extends the virtual array theory [9] to the case where several $2q$-th order cumulant array are simultaneously used. Hence, a twofold increase of the virtual array aperture.

where $A^{\otimes q}$ denote the integer part of $q/2$ and the Khatri-Rao product of $q$ matrices $A$, respectively. Some values of $N_{2q}$ in radiocommunications context are given in [9] for several kinds of antennas. For instance, when a ULA of $N$ sensors is used, $N_{2q}$ is given by $N_{2q} = q(N - 1) + 1$. Let’s now consider the SAD-SIBI_JAD method. $P_{\text{max}}$ is the maximum value of $P$ for which assumptions 1 and 2 are true. Then we have:

$$P_{\text{max}} = r k_{\text{max}}(A^{(3)}) = N_6$$

$$P_{\text{max}} = r k_{\text{max}}(B(A^{(3)} \odot A^{(3)*})^T)$$

where $(M \times N^6) H_x^{(3)}$ denotes the unfolding matrix associated to the third direction of $H_x(14)$. For $M \geq N^6$, the maximum rank of $H_x^{(3)}$ is generically equal to $N^6$. Recall that a property is well-called generic when it holds everywhere except for a set of Lebesgue measure 0. However, in some applications such as in radiocommunication contexts, matrix $H_x^{(3)}$ may belong to this set. In this case, the previous property does not hold generally. Intrinsically, we have:

$$r k_{\text{max}}(H_x^{(3)}) \leq r k_{\text{max}}(A^{(6)}) = N_{12}$$

where $A^{(6)} = A^{(3)} \odot A^{(3)*}$. But, since according to [9] $N_{12}$ is strictly lower than $N^6$, the number maximal of sources which can be processed using the ALS-SIBI approaches is then given by:

$$P_{\text{max}} = r k_{\text{max}}(A^{(6)}) = N_{12}$$

An identifiability study of the SIBI family is proposed hereafter under assumptions 1-3. In particular, the $M$ time lag vectors $\tau^{(m)}$ are chosen such that a well-conditioned matrix $B$ is obtained. Furthermore, we have to select at least an $M = N^2 q$ time lag vectors. Note that, in practice, the choice of these time lag vectors relies on some prior information on sources. Our identifiability study is fourfold. First, from a BUMI viewpoint, it gives the maximum number, $P_{\text{max}}$, of sources which can be processed by the SIBI family. Secondly, from a numerical analysis viewpoint, it tackles the identifiability of the SAD, ACDC and ELSALS_SYMM procedures when a 3-way array with special symmetries is considered. Next, from a multilinear algebra viewpoint it gives some insights on the uniqueness of a joint CAND of several hermitian multiway arrays. Finally, from the VA theory viewpoint, we extend the VA concept [9] to the case of a combination of several VA’s. Recall that the VA theory relies on the existence of $N_{2q}$ sensors which virtually appear when using the $2q$-th order statistical matrix $C_{2q}[\ell] [9]$. Regardless the application area, $N_{2q}$ is equal to the maximal rank, $r k_{\text{max}}$, of the $(N^q \times P)$ matrix $A^{(q)}$:

$$N_{2q} = r k_{\text{max}}(A^{(q)}) = r k_{\text{max}}(A^\otimes(\ell+1) \odot A^\ast \otimes \ell)$$

...
TABLE I

<table>
<thead>
<tr>
<th>$N$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOBI [5]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>JADE [6]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>COM2 [11]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>FastICA [20]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>STOTD [32]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>NC-STOTD [34]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>G-STOTD [34]</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>PBMC1 [16]</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>SOBIUM2 [29]</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>FOBI1 [30]</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>FOBI2 [30]</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>SOBIUM1 [29]</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>FOBiUM [17]</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>6-BIOME [2]</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>SAD-SIBI, JAD</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>13</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>ALS-SIBI ACDC</td>
<td>7</td>
<td>13</td>
<td>19</td>
<td>25</td>
<td>31</td>
<td>37</td>
<td>43</td>
</tr>
<tr>
<td>ELSALS-SIBI UNS</td>
<td>7</td>
<td>13</td>
<td>19</td>
<td>25</td>
<td>31</td>
<td>37</td>
<td>43</td>
</tr>
<tr>
<td>ELSALS-SIBI SYM</td>
<td>7</td>
<td>13</td>
<td>19</td>
<td>25</td>
<td>31</td>
<td>37</td>
<td>43</td>
</tr>
</tbody>
</table>

$P_{max}$ as a function of the number $N$ of sensors of a ULA for different cumulant-based methods

VI. Numerical Complexity

The numerical complexity of the SIBI family is addressed hereafter. It is computed and compared to the one of classical BUMI methods, namely SOBI [5], JADE [6], FastICA [20, chapter 8], COM2 [11], FOBI1 [30], FOBI2 [30], FOBium [17] and 6-BIOME [2]. More particularly, it is computed in terms of number of floating point (complex) operations (flops) required to identify the $(N \times P)$ mixing matrix $A$ from $K$ data snapshots. Note that a complex flop is defined as the sum of a complex multiplication and a complex addition. In practice, only multiplications are counted which does not affect the magnitude of the numerical complexity. Table II shows the computational complexity of the aforementioned methods where $f_i(N) = O[N^4/8]$ and $f_o(N) = O[N^6/2]$ denote the number of free entries of $C_{i,x}$ and $C_{o,x}$, respectively. In addition, $D$ and $J$ denote the desired resolution in the unit disk for the COM2 approach and the number of sweeps used in the joint diagonalization algorithms, respectively. $J_1$, $J_2$, $J_3$ and $J_4$ are the maximal numbers of iterations used by FastICA, ELSALS-SIBI UNS, ELSALS-SIBI SYM and ALS-SIBI ACDC, respectively. $J_1$ is the maximal number of iterations used for the estimation of the relaxation factor in ELSALS-SIBI UNS and ELSALS-SIBI SYM. Eventually, $M$ represents the number of delay lags used in SOBI, FOBiUM and the SIBI techniques, $M_1 = \max\{N^6, P\}$, $m_1 = \min\{N^6, P\}$, $M_2 = \max\{M, P\}$ and $m_2 = \min\{M, P\}$.

Figure 4 shows the minimum numerical complexity of the SIBI family and the aforementioned classical BUMI methods, as a function of the number $P$ of sources. Indeed, for each value of $P$ the minimum number $N_{min}$ of sensors of a ULA is computed such that the identifiability condition of each method is still valid. Hence, $N_{min} = P$ for SOBI, JADE, FastICA, COM2 and $N_{min} \leq P \leq q(N_{min} - 1) + 1$ for the other algorithms allowing for an identification of underdetermined mixtures. Note that the parameter $q$ is directly related to the identifiability condition of each method. For example, $q = 2$ for FOBI1, FOBI2, FOBiUM and 6-BIOME, whereas $q = 3$ for the SAD-SIBI, JAD approach and $q = 6$ for both the ALS-SIBI ACDC and the ALS-SIBI techniques. $M$, $I$ and $D$ are chosen to be equal to $9, 1 + \lfloor P^{1/2} \rfloor$ and 0.05, respectively. As far as the iterative approaches are concerned, $J$ and $J_1$ are chosen such that a good identification of the mixing matrix is guaranteed, hence $J = 1000$, $J_1 = 800$, $J_2 = 400$, $J_3 = 600$. On the other hand, we put $J_4 = 20$, which ensures a good estimation of the relaxation factor. Note that the latter three parameters are highly related to many factors such as SNR, direction of arrivals, number of sources, number of samples, etc. As depicted in figure 4, the ELSALS-SIBI approaches belong to the group of the more expensive methods such as FastICA whereas the SAD-SIBI technique is less expensive as 6-BIOME, FOBiUM, COM2 FOBI1.

VII. Simulation Results

Five performance evaluation studies of the SIBI family are presented hereafter, and more precisely the performance of SAD-SIBI, ALS-SIBI ACDC and ALS-SIBI SYM for both underdetermined and overdetermined mixtures of sources. The performance analysis was achieved using the criterion $D(A, \hat{A}) = (\alpha_1, \alpha_2, \cdots, \alpha_p)$, which measures a pseudo-distance between mixture $A$ and its estimate $\hat{A}$ where $\alpha_p = \min_{1 \leq i \leq p} d(\alpha_p, \hat{\alpha}_i)$ with $d$ the pseudo-distance between vectors defined by [17]:

$$d(u, v) = 1 - \frac{\|u^*v\|}{\|u\| \cdot \|v\|}$$

It is noteworthy that only criterion $\alpha_1$ is displayed in figures 5, 7, 8 and 9 since the $P = 1$ other criteria $\alpha_p$ have shown identical results. The $P$ sources considered in this section are Quadrature Phase Shift Keying (QPSK) linearly modulated with pulse shape filter corresponding to a 1/2-Nyquist filter with a roll-off of 0.3. The latters are assumed to impinge on a ULA of $N$ sensors and are chosen to be well angularly separated.

Both first studies consist in evaluating the performance of the SIBI family in a strongly underdetermined mixture which
cannot be processed by any classical cumulant method based on at most order 6, say $P \geq 4$ sources from a ULA of only $N = 2$ sensors. The sources have the same Signal-to-Noise Ratio (SNR) equal to 15 dB. On the one hand, figure 5 shows the variation of $\alpha_1$ at the output of the three SIBI methods as a function of the number of samples for $P = 4$ sources with carrier residuals such that $f_{c1}T_c = 0, f_{c2}T_c = 0.25, f_{c3}T_c = 0.5, f_{c4}T_c = 0.75$. In addition, the performance was compared for two numbers $M$ of sets $\tau^{(m)}$ of delays, say $M = 9$ and $M = 27$. Figure 5 shows a good behavior of the SIBI family. Besides, the semi-algebraic SAD-SIBIJAD method seems to be insensitive to the increase in the number $M$ of used sets of delays, contrary to the iterative ALS-SIBIACDC and ELSALS-SIBISYM algorithms. As shown in section V, more statistical information may ensure the uniqueness of CAND and then a good quality of factor’s identification. Since better results are obtained for $M = 27$ sets of delays, the latter were used in the following studies. On the other hand, the case of $P = 7$ sources with residual carriers given by $f_{c1}T_c = 0, f_{c2}T_c = 0.15, f_{c3}T_c = 0.3, f_{c4}T_c = 0.45, f_{c5}T_c = 0.6, f_{c6}T_c = 0.75,$ and $f_{c7}T_c = 0.9,$ is considered in figure 6. Note that, in such a practical context, only ALS-SIBIJADC and ELSALS-SIBISYM are able to process as many sources, as mentioned in table I. More precisely, figure 6 shows the maximum $\alpha_p$, named $\alpha_{\text{max}}$, for each data sample at the output of ALS-SIBIJADC and ELSALS-SIBISYM, say the worst performance over all sources

### Table II

**Numerical complexity of the SIBI family and classical cumulant-based methods in terms of complex flops**

<table>
<thead>
<tr>
<th>Method</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>COM2</td>
<td>$\min{L^2N^2/2 + 4N^2/3 + PNL, 2LN^2} + \min{12Lf_j(P)(P^* + 2LP^3 + 3L_j(P) + LP^2, 13LP^2/2) + 7LP^2/D^2$</td>
</tr>
<tr>
<td>FastICA</td>
<td>$\min{L^2N^2/2 + 4N^2/3 + PNL, 2LN^2} + \min{12Lf_j(P)(P^* + 2LP^3 + 3L_j(P) + LP^2, 13LP^2/2) + 7LP^2/D^2$</td>
</tr>
<tr>
<td>SOBI</td>
<td>$MN^2/2 + 4N^2/3 + PNL, 2LN^2} + \min{4P^3/3, 8P^3(P^* + 3) + 3L_j(P) + LP^2/2 + LP^2$</td>
</tr>
<tr>
<td>JADE</td>
<td>$3Mf_j(N) + 2N^2/3 + P^3(3N^2 - P^3)/3 + (M - 1)N^2/2 + IP(N - 1)} + 17(M - 1)/2 + 4N^2 + 15j)/2 + 2N^2P$</td>
</tr>
<tr>
<td>FOOBI1</td>
<td>$Pf_j(N) + 2N^2/3 + P^3(3N^2 - P^3)/3 + N^2P + N^2P + 2P^3 + 2N^2P + N^2P + N} + 17(M - 1)/2 + 4N^2 + 15j)/2 + 2N^2P$</td>
</tr>
<tr>
<td>FOOBI2</td>
<td>$Pf_j(N) + 2N^2/3 + P^3(3N^2 - P^3)/3 + N^2P + N} + 17(M - 1)/2 + 4N^2 + 15j)/2 + 2N^2P$</td>
</tr>
<tr>
<td>6-BIOME</td>
<td>$5f_j(0) + 2N^2/3 + P^3(3N^2 - P^3)/3 + N^2P + N} + 17(M - 1)/2 + 4N^2 + 15j)/2 + 2N^2P$</td>
</tr>
<tr>
<td>SAD-SIBIJAD</td>
<td>$5LNf_j(N) + 4N^2/3 + P^3N^2 + M(PN^2 + P^2N^2) + IP(N - 1)} + 17(M + 4P + 75)/2 + PIN(N - 1)} + 21N + 75)/2$</td>
</tr>
<tr>
<td>ALS-SIBIACDC</td>
<td>$5LNf_j(N) + 2P^3N^2 + N^2P + N} + 17(M + 4P + 75)/2 + PIN(N - 1)} + 21N + 75)/2$</td>
</tr>
<tr>
<td>ELSALS-SIBIUNS</td>
<td>$5LNf_j(N) + 2P^3N^2 + N^2P + N} + 17(M + 4P + 75)/2 + PIN(N - 1)} + 21N + 75)/2$</td>
</tr>
<tr>
<td>ELSALS-SIBISYM</td>
<td>$5LNf_j(N) + 2P^3N^2 + N^2P + N} + 17(M + 4P + 75)/2 + PIN(N - 1)} + 21N + 75)/2$</td>
</tr>
</tbody>
</table>

Fig. 5. $\alpha_1$ for a ULA of 2 sensors, 4 QPSK’s with the same SNR = 15 dB and two numbers $M$ of delays: $M = 9$ (black line) and $M = 27$ (grey line).

Fig. 6. Variation of the worst $\alpha_p$ criterion for each data sample for a ULA of 2 sensors and 7 QPSK’s of SNR = 15 dB.
of both considered methods. It appears that both methods are effectively able to process up to seven sources from a ULA of two sensors. In addition, in comparison with figure 5, ELSALS-SIBI\textsubscript{SYM} seems to be less sensitive to a strong increase of the number $P$ of sources even it also needs more samples to give a constant performance as $P$ increases.

The third study consists in comparing, in a more sympathetic underdetermined mixture case, the SIBI family and cumulant-based methods such as FOBIUM [17], 6-BIOME [2], as a function of the number of samples. For this purpose $N=2$ sensors and $P=3$ sources are considered with residual carriers given by $f_cT_c=0$, $f_cT_e=0.35$, $f_cT_e=0.7$ and an SNR of 15 dB. While FOBIUM and 6-BIOME appear to be efficient in figure 7, especially when the number of samples increases, the SIBI family shows a higher performance even for a small number of samples.

Now we evaluate the influence of SNR on the performance of the SAD-SIBI\textsubscript{JAD}, ALS-SIBI\textsubscript{ACDC}, ELSALS-SIBI\textsubscript{SYM}, FOBIUM and 6-BIOME using the same application context as the one used in figure 7. Under the assumptions, figure 8 shows that the five methods have a similar behavior for small values of SNR. Nevertheless, the SIBI techniques give better results as soon as SNR increases beyond 0 dB.

Eventually, figure 9 shows the variations of $\alpha_1$ as a function of the spatial correlation factor of a Gaussian noise at the output of the three considered SIBI approaches and seven classical cumulant-based methods such as COM2 [11], FastICA [20, chapter 8], FOOB1, SOBI [5], JADE [6], 6-BIOME and FOBIUM. For this purpose, we considered 1000 samples, $N=3$ sensors and $P=2$ sources with residual carriers such that $f_cT_e=0$, $f_cT_e=0.65$ and a SNR of 5 dB for all sources. Note that the Gaussian noise model used in this simulation is the sum of an internal noise $\nu_{in}[k]$ and an external noise $\nu_{out}[k]$ of covariance matrices $C_{2,\nu}^{in}[k,0]$ and $C_{2,\nu}^{out}[k,0]$, respectively, such that:

\[
C_{2,\nu}^{in}[0,r,q] = \frac{\sigma^2}{2}\delta(r-q) \quad C_{2,\nu}^{out}[0,r,q] = \frac{\sigma^2}{2}\rho|^{r-q|}
\]

where $\sigma^2$, $\rho$, $C_{2,\nu}^{in}[0,r,q]$, $C_{2,\nu}^{out}[0,r,q]$ are the total noise variance per sensor, the noise spatial correlation factor and the $(r,q)$-th component of the total noise covariance matrix, respectively. The computer results show on the one hand how the performance of the SIBI family is stable, even for a strong noise spatial correlation. In one hand, the classical methods seem to be less robust, especially COM2, FastICA, JADE and SOBI which require a prior spatial whitening based on SO cumulants.

VIII. CONCLUSION

We proposed a new class of BUMI methods based on the joint CAND of a set of HO cumulant arrays. This class of algorithms is described through S\textsubscript{IBI} cumulants, giving rise to two S\textsubscript{IBI} subfamilies of techniques, namely the SAD-SIBI and the ALS-SIBI methods. More particularly, these two families perform a CAND of a 3-way array built from a set of S\textsubscript{IBI} cumulant arrays but in a different way. Indeed, the former approach is algebraic while the latter is iterative. All approaches were theoretically studied especially in terms of identifiability and computational complexity. In particular, the identifiability analysis showed that the SIBI algorithms can process more sources than the classical cumulant-based methods. In addition, computer results showed the good behavior of the SIBI family against classical methods, for both overdetermined and underdetermined mixtures of sources. Besides, from a numerical analysis viewpoint, we deal with the identifiability and the convergence speed of several 3-way array decomposition procedures such as the symmetric ACDC scheme. We also show how to accelerate the iterative CAND of a 3-way array, which has special symmetries. Next, from a multi-linear algebra viewpoint we give some insights on the
for a ULA of 3 to use a VA with a higher aperture. Several VAs with the same number of virtual sensors amounts the case of a combination of several VAs. Indeed, the use of uniqueness of a joint CAND of several hermitian multi-way

\[ \alpha \]

Fig. 9. \( \alpha \) as a function of the spatial correlation factor of a Gaussian noise for a ULA of 3 sensors and 2 QPSK’s of 1000 samples.

uniqueness of a joint CAND of several hermitian multi-way arrays compared to the CAND of only one array. In addition, this allows us to extend the concept of Virtual Array (VA) to the case of a combination of several VAs. Indeed, the use of several VAs with the same number of virtual sensors amounts to use a VA with a higher aperture.

**Appendix A**

**Proof of Proposition 5**

The full column rank of \( B \) implies that there is at least one \( M \)-dimensional vector \( \beta \) with values in the complex field such that:

\[
B' \beta = [1, 1, \ldots, 1]' \in \mathbb{C}^P
\]  

(33)

According to the definition of matrix \( B \) given in proposition 2, (33) is equivalent to the following equation:

\[
\sum_{m=1}^{M} \beta_m \zeta \beta_m [t^{(m)}] = I_P
\]  

(34)

where \( \beta_m \) is the \( m \)-th component of the \( M \)-dimensional column vector \( \beta \) and \( I_P \) is the \((P \times P)\) identity matrix. Then, under assumption 1, the following matrix:

\[
\sum_{m=1}^{M} \beta_m C_{6,x}[t^{(m)}] = A^{(3)} \left( \sum_{m=1}^{M} \beta_m \zeta [t^{(m)}] \right) A^{(3)\dagger}
\]

(35)

is a rank-\( P \) positive semidefinite matrix equal to \( A^{(3)} A^{(3)\dagger} \). Hence the result.

**References**


