Multidimensional Independent Component Analysis with Higher-order cumulant matrices for vector sources with possibly differing dimensions

Hanany Ould-Baba, Jérôme Antoni, Vincent Robin

Multidimensional Independent Component Analysis with Higher-order cumulant matrices for vector sources with possibly differing dimensions ☆,☆☆

H. OULD-Baba[a,1,**], J. Antoni[a,b,2,**], V. Robin[c,1,3,**]

[a] *Toulouse 1 Capitole University, 2 Rue du Doyen-Gabriel-Marty, 31042 Toulouse, France*
[b] *Laboratoire Vibrations Acoustique, University of Lyon, F-69621 Villeurbanne, France*
[c] *Université de Technologie de Compiègne -CS 60319 - 60203 Compiègne cedex France*

Abstract

The paper addresses the separation of multidimensional sources, with possibly different dimensions, by means of higher-order cumulant matrices. First, it is rigorously proved, in a general setting, that contracted cumulant matrices of any order are all block-diagonalizable in the same basis. Second, a family of joint block-diagonalization algorithms is proposed that separate multidimensional sources by combining contracted cumulant matrices of arbitrary orders. Third, a specific solution is given to determine the source dimensions when they are unknown but all different. The performances of the proposed algorithms are compared between them and with algorithms of the literature based on orders 3 and 6.

Keywords:
Multidimensional Independent Component Analysis; Independent Component Analysis; Higher-order cumulant; Independent Subspace Analysis; Joint Block-Diagonalization.

1. Introduction

In its original formulation independent component analysis (ICA) assumes the mutual independence of the sources to be separated (e.g. see Comon [11]). Unfortunately, there are many instances in the real-world where this assumption is not fulfilled, which precludes the recovery of sources that are (partly) dependent. This has brought in the concept of multidimensional independent component analysis (MICA), where *multidimensional* rather than *scalar* sources are considered Cardoso [6]. Formally, the MICA model reads

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_p \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_p \end{pmatrix} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_p \tag{1}$$

☆This document is a collaborative effort.
☆☆MICA with higher-order cumulant matrices.
*Corresponding author
**Principal corresponding author
Email addresses: hanany.ould-baba@ut-capitole.fr (H. OULD-Baba), jerome.antoni@utc.fr (J. Antoni), vincent.robin@utc.fr (V. Robin)
[1]This is the specimen author footnote.
[2]Another author footnote, but a little more longer.
[3]Yet another author footnote. Indeed, you can have any number of author footnotes.

$$\text{where } \begin{cases} \mathbf{x} & \in & \mathbb{R}^m & \text{is the vetor of observations,} \\ \mathbf{A} & \in & \mathbb{R}^{m\times n} & \text{is the unknwon mixing matrix,} \\ \mathbf{s} & \in & \mathbb{R}^n & \text{is the unknown vector of sources,} \\ \mathbf{s}_i & \in & \mathbb{R}^{n_i} & \text{denotes the } i\text{-th source,} \\ \mathbf{A}_i & \in & \mathbb{R}^{m\times n_i} & \text{denotes the } i\text{-th block of matrix } \mathbf{A}, \\ \mathbf{x}_i & = & \mathbf{A}_i\mathbf{s}_i \in \mathbb{R}^m & \text{denotes the } i\text{-th independent component,} \\ n & = & \sum_{i=1}^{p} n_i & \text{is the dimension of the source vector,} \\ p & \in & \mathbb{N} & \text{denotes the number of multidimensional sources.} \end{cases} \tag{2}$$

That is to say MICA consists in decomposing an arbitrary random vector, $\mathbf{x}$, into its independent components (ICs), $\mathbf{x}_i$, $i = 1, ..., p$ where a vector $\mathbf{v}$ is an IC of a random vector $\mathbf{u}$ if there exists an invertible matrix $\mathbf{F}$ and an decomposition $\mathbf{u} = \mathbf{F}\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}$ where $\mathbf{v}$ and $\mathbf{w}$ are independent random vectors. Up to trivial indeterminacies (see subsection 2.1), this amounts to estimating the unknown mixing matrix $\mathbf{A}$ (or its inverse, the separation matrix $\mathbf{B}$) and sources $\mathbf{s}$ in model (1) from the sole observations $\mathbf{x}$. In other words, when $\mathbf{x}$ accepts a decomposition into one-dimensional (*scalar*) sources, the problem boils down to classical ICA [11]. More generally, when the decomposition involves k-dimensional sources, this is referred to as k-MICA, which is the topic of this paper. So far MICA has attracted much less attention than ICA, although it is also known under (or related to) the following names in the literature: independent subspace analysis (ISA) [20], independent feature subspace analysis (IFSA) [23], subspace ICA [41], group ICA [42]. The application fields where MICA finds interest are indeed numerous. Examples comprise electrocardiography (ECG) [6, 42, 41, 10], f-MRI electroencephalography (EEG) [30], topography [21], texture classification [40], action recognition [28], face recognition [29], motion segmentation [12], audio source separation [9], gene decoding [23, 24], metabolomic data analysis [17].

Moreover, although most previous works have concentrated on separating ICs with identical dimensions (k-MICA with $k = \dim(\mathbf{s}_i)$, $\forall i$) [21, 45, 44, 42, 22, 41] (in which case the problem is known as k-MICA with $k = \dim(\mathbf{s}_i) = n_i$), this situation is by no means general. This paper considers the general case where p multidimensional ICs with arbitrary and unknown dimensions, $n_i = \dim(\mathbf{s}_i) = \dim(\mathbf{x}_i)$ (not necessarily identical), are to be separated. Several approaches have been considered in the literature to solve the general MICA problem. An early solution consists of regrouping *a posteriori* the components separated by traditional ICA according to their residual dependence. Although the idea was globally covered in Ref. [7], no proof of separability and unicity was provided. The first proof of separability of the k-MICA problem was given by Theis [51] and is based on a theorem of prob ability introduced in Ref. [19], in the sixteens, that generalizes the Darmois-Skitovitch theorem [13, 48] to the multidimensional case (square matrices). The complete proof of unicity was later given by Theis [50] and Gutch and Theis [20] in the general case of independent subspaces with possibly different dimensions. It is there remarked that the initial model introduced in Ref. [7] is not complete and requires an additional assumption for being separable. Indeed, for any observation $\mathbf{x}$, a decomposition of $\mathbf{x}$ into independent subspaces with some internal dependence in each subspace is given by $\mathbf{x}$ itself. The additional condition required to avoid this trivial decomposition has been coined "irreducibility of multidimensional sources". The condition is revisited in subsection 8 of this paper (see Definition 8) Note that the proof of unicity of the MICA separation has recently been completed in Ref. [22]. Besides, some joint block- diagonalization (JBD) algorithms have been proposed that potentially provide solutions to MICA by reducing a set of matrices to diagonal blocks with different sizes [50, 37, 41]. They make use of second-order statistics [49, 23, 31], of the functional form of the probability density function or its generating function [52, 49], or of dispersion matrices [43]. Two such algorithms have been applied to the MICA problem in Ref. [49]. The first one, which is a direct extension of SOBI [1], jointly block-diagonalizes a set of correlation matrices taken at different time lags and has been coined Multidimensional SOBI (MSOBI). The second one jointly block-diagonalizes a set of Hessian matrices of the first (or second) generating function and has been coined Multidimensional Hessian ICA (MHICA). However, solutions to MICA based on higher-order cumulants are still relative ly seldom although they offer several advantages. Indeed, Theis in [43] has proposed to simply solve the MICA problem by jointly block-diagonalizing a set

of fourth-order cumulant matrices in the spirit of the JADE algorithm [8], which he coined SJADE (for Subspace-JADE). Yet, the following questions have so far remained unanswered:

- are fourth-order cumulant matrices (as initially defined in [8] and used in Theis [43]) block-diagonalizable in the basis spanned by $\mathbf{A}$?
- more generally, are cumulant matrices of arbitrary order r (to be defined later on in the paper) block-diagonalizable in the basis spanned by $\mathbf{A}$?

In other words, can the MICA problem be formally formulated as a joint block-diagonalization problem on arbitrary order r? On the second-order, the answer is positive and has been given in [42, 19, 27]. On higher-orders, the answer is yet not trivial and is the subject of the present paper.

The remainder of the paper is organized as follows. First, the usual assumptions that sustain the MICA problem are reviewed in section 2. Next, the definitions of cumulant matrices and contracted cumulant matrices are introduced in section 3 and 4, respectively. Section 5 contains the main result of the paper, which proves that contracted cumulant matrices of any order are all block-diagonalizable in the same basis. This result is then exploited in section 6 to generalize the JADE algorithm to the multidimensional case with arbitrary orders, thus leading to a family of algorithms coined SJADE$_r$. The advantage of combining various orders is illustrated in section 7. Section 8 addresses the important question as how to determine the dimensions of the vector sources when they are unknown. The proposed algorithms are finally compared in section 9 by means of numerical experiments. Simple examples are given to demonstrate the advantage of combining different orders in order to improve the separation. All proofs are collected in the Appendices.

2. Generalities about MICA

From the onset, some generalities about ICA are to be reminded in the context of multidimensional sources.

2.1. Uniqueness and indeterminacies

The indeterminacies underlying MICA are obvious generalizations of those of ICA. They are reminded here for the sake of completeness. It suffices to note that, for any invertible matrices $\mathbf{D}_i$ of size $n_i \times n_i$, $i \in [\![1, p]\!]$, and for any permutations $\sigma \in \mathfrak{S}_p$, model (1) still holds since

$$\left(\begin{array}{ccc} \mathbf{A}_{\sigma(1)}\mathbf{D}_{\sigma(1)} & \cdots & \mathbf{A}_{\sigma(p)}\mathbf{D}_{\sigma(p)} \end{array} \right) \left(\begin{array}{c} \mathbf{D}_{\sigma(1)}^{-1}\mathbf{s}_{\sigma(1)} \\ \vdots \\ \mathbf{D}_{\sigma(p)}^{-1}\mathbf{s}_{\sigma(p)} \end{array} \right) = \mathbf{x}_{\sigma(1)} + \cdots + \mathbf{x}_{\sigma(p)} = \mathbf{x}. \tag{3}$$

Therefore, the multidimensional ICs $\mathbf{x}_i$ are determined up to an arbitrary permutation (that applies on components with identical dimension), whereas sources are determined up to an arbitrary permutation and invertible matrices. Taking such indeterminacies into account, a matrix $\mathbf{B}$ will be recognized as a separation matrix (i.e. a solution to the MICA problem) if $\mathbf{BA} = \mathbf{PD}$, where $\mathbf{P}$ is a permutation matrix and $\mathbf{D} = \mathrm{bdiag}(\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_p)$ a block-diagonal matrix.

Besides, the following definition will play an important role later on.

Definition 1 ([43, 18]). A random vector, $\mathbf{x}$, of dimension n is said *reducible* if it can be expressed as

$$\mathbf{x} = \mathbf{A} \left(\begin{array}{c} \mathbf{y}_1 \\ \mathbf{y}_2 \end{array} \right), \tag{4}$$

where $\mathbf{A}$ is an invertible matrix and $\mathbf{y}_1$ and $\mathbf{y}_2$ are two independent random vectors of dimensions $k \neq 0$ and $n - k$, respectively. A random vector that is not reducible is said *irreducible*.

2.2. Assumptions

One is now in position to announce the assumptions that will guarantee the uniqueness of the separation model:

1. matrix $\mathbf{A}$ is of full-rank n,
2. no more than one independent (*scalar*) source can have a nil cumulant of order r, $c_r(\mathbf{s}_i) = 0$, when working on order r,
3. independent sources are all irreducible.

2.3. Whitening

Whitening is a classical preprocessing used in blind source separation. It is used here to simplify the theoretical developments. Without loss of generality, let us assume that the source vectors are centered, $\mathbb{E}(\mathbf{s}) = \mathbf{0}$, and standardized, $\mathbf{R_s} = \mathbb{E}(\mathbf{ss}') = \mathbf{I}_n$. Thus $\mathbf{R_x} = \mathbb{E}(\mathbf{xx}') = \mathbf{AA}'$. Since the covariance matrix $\mathbf{R_x}$ is symmetric and semi-positive definite, it is well-known that it admits an eigenvalue decomposition $\mathbf{R_x} = \mathbf{UDU}'$ where $\mathbf{D}$ is an $n \times n$ diagonal matrix, with n the number of non-zero eigenvalues, and $\mathbf{U}_{m\times n}$ an $m \times n$ matrix satisfying $\mathbf{U}'\mathbf{U} = \mathbf{I}_n$. Following the usual practice, let us define the whitening matrix $\mathbf{W}_{n\times m} = \mathbf{D}^{-\frac{1}{2}}\mathbf{U}'$ and the whitened components $\widetilde{\mathbf{x}} = (\mathbf{Wx})_{n\times 1}$; thus

$$\mathbf{R}_{\widetilde{\mathbf{x}}} = \mathbb{E}(\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}') = \mathbf{WR_xW}' = \mathbf{D}^{-\frac{1}{2}}\mathbf{U}'(\mathbf{UDU}')\mathbf{UD}^{-\frac{1}{2}} = \mathbf{I}_n. \tag{5}$$

Since $\mathbf{R_x} = \mathbf{AA}'$, it follows that $\mathbf{WAA}'\mathbf{W}' = \mathbf{I}_n$; in other words, $\widetilde{\mathbf{A}} = \mathbf{WA}$ is an $n \times n$ orthogonal matrix. Upon pre-multiplication with the whitening matrix, model (1) then becomes $\widetilde{\mathbf{x}} = \widetilde{\mathbf{A}}\mathbf{s}$. Therefore, it will be assumed from now on and without loss of generality that matrix $\mathbf{A}$ is orthogonal (with $m = n$). Solving the MICA problem then amounts to finding, using higher-order statistics, the orthogonal matrix $\mathbf{A}$.

2.4. Notations

This last subsection introduces some notations that will be used in the remaining of the paper. Let us first note that matrix $\mathbf{A}$ can be partitioned into $n \times n_i$ sub-matrices $\mathbf{A}_i$ or as a collection of raw vectors ℓ_j:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_p \end{pmatrix} = \begin{pmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_n \end{pmatrix}. \tag{6}$$

Alternatively, matrix $\mathbf{A}$ can also be partitioned into raws as

$$\mathbf{A}_i = \begin{pmatrix} \ell_1^i \\ \ell_2^i \\ \vdots \\ \ell_n^i \end{pmatrix} \tag{7}$$

where, $\forall k \in [\![1, n]\!]$, ℓ_k^i with dimension $1 \times n_i$ is the k-th raw of sub-matrix $\mathbf{A}_i$. Similarly, each raw of matrix $\mathbf{A}$ can be expressed as a function of the raws of matrices $\mathbf{A}_i$, as follows

$$\ell_k = \begin{pmatrix} \ell_k^1 & \ell_k^2 & \dots & \ell_k^p \end{pmatrix}. \tag{8}$$

One is now ready to introduce specific statistical tools – cumulant matrices, their contracted forms, and the related algebra – dedicated to the purpose of this paper.

3. Cumulant matrices

The main tools used in this paper are cumulants. In the scalar case it is a well-known fact that cumulants can be deduced from moments by means of the Leonov and Shiryaev formula (see e.g. Mac Cullagh [31] et Albera and Comon [1]). However, it is often convenient to rearrange cumulants into matrices (see e.g. Albera et al. [4]), for instance in order to exploit their algebraic properties (such as matrix redundancies Albera et al. [3, 2, 4]) or to diagonalize them Kollo [26]. In the multidimensional case, the notion of a cumulant matrix needs to be carefully defined for it will play a key role in the remaining of the paper.

Definition 2 (cumulant matrix). Let $\mathbf{x}$ be a random vector with dimension $n \times 1$. Then, the cumulant matrix of order r is defined as

$$c_r(\mathbf{x}) = \sum_{i_1,i_2,\ldots,i_r=1}^{n} \operatorname{cum}(x_{i_1}, x_{i_2}, \cdots, x_{i_r}) \mathbf{E}_{i_1,i_2,\ldots,i_r} \tag{9}$$

with

$$\mathbf{E}_{i_1,i_2,\ldots,i_r} = \begin{cases} \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_3} \otimes \cdots \otimes \mathbf{e}_{i_{2k-1}}\right)\left(\mathbf{e}_{i_2} \otimes \mathbf{e}_{i_4} \otimes \cdots \otimes \mathbf{e}_{i_{2k}}\right)' & \text{if} \quad r = 2k \\ \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_3} \otimes \cdots \otimes \mathbf{e}_{i_{2k+1}}\right)\left(\mathbf{e}_{i_2} \otimes \mathbf{e}_{i_4} \otimes \cdots \otimes \mathbf{e}_{i_{2k}}\right)' & \text{if} \quad r = 2k+1, \end{cases} \tag{10}$$

where $(\mathbf{e}_i)_{i \in [\![1,n]\!]}$ denotes the canonical basis of $\mathbb{R}^n$ and $\otimes$ the Kronecker product.

Closed-form expressions of cumulant matrices in terms of lower-order moments were given in Ref. [26] up to orders 3 and 4, and for the first time in Ref. [38] up to order 6. They are reproduced in Appendix (Appendix B) for the sake of completeness. Not only are such expressions easy to handle due to their compactness, but they also lead to faster numerical computation with matrix-based languages such as Matlab®(interested readers are invited to consult [38] for the systematic derivation of these formulae). It is noteworthy that Definition 2 produces $n^k \times n^k$ square and symmetric cumulant matrices of even orders ($r = 2k$) which, as compared to other possible definitions, will turn out advantageous; for instance, on order $r = 4$, $c_4(\mathbf{x})$ is a $n^2 \times n^2$ matrix known as the "quadricovariance" [2]; similarly, on order $r = 6$, $c_6(\mathbf{x})$ is a $n^3 \times n^3$ matrix known as the "hexacova riance" [3]. For odd orders $r = 2k + 1$, the cumulant matrices have dimensions $n^{k+1} \times n^k$. Several useful properties can be deduced from Definition 2, a few of which are reminded here below.

Proposition 1. (i) *If* $\mathbf{z}$ *is a Gaussian random vector, then*

$$\forall r > 2, \qquad c_r(\mathbf{z}) = \mathbf{0}. \tag{11}$$

(ii) *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are two independent vectors with identical dimensions, then*

$$\forall r \in \mathbb{N} \qquad c_r(\mathbf{x} + \mathbf{y}) = c_r(\mathbf{x}) + c_r(\mathbf{y}). \tag{12}$$

(iii) *If* $\mathbf{x} = \mathbf{A}\mathbf{s}$*, where* $\mathbf{A}$ *is a matrix with dimension* $m \times n$ *and* $\mathbf{x}$ *and* $\mathbf{s}$ *are two random vectors with dimensions* $m \times 1$ *and* $n \times 1$*, respectively, then*

$$c_r(\mathbf{x}) = \begin{cases} \underbrace{(\mathbf{A} \otimes \cdots \otimes \mathbf{A})}_{k \ fois} c_r(\mathbf{s}) \underbrace{(\mathbf{A} \otimes \cdots \otimes \mathbf{A})'}_{k \ fois} = \left(\otimes^k \mathbf{A}\right) c_r(\mathbf{s}) \left(\otimes^k \mathbf{A}'\right) & \text{if } r = 2k \\ \underbrace{(\mathbf{A} \otimes \cdots \otimes \mathbf{A})}_{k+1 \ fois} c_r(\mathbf{s}) \underbrace{(\mathbf{A} \otimes \cdots \otimes \mathbf{A})'}_{k \ fois} = \left(\otimes^{k+1} \mathbf{A}\right) c_r(\mathbf{s}) \left(\otimes^k \mathbf{A}'\right) & \text{if } r = 2k+1 \ . \end{cases} \tag{13}$$

(iv) *If* $\mathbf{x} = \mathbf{A}\mathbf{s} = \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_p \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_p \end{pmatrix} \in \mathbb{R}^m$, *where* $\mathbf{s}_i \in \mathbb{R}^{n_i}$ *are independent random vectors,* $\mathbf{A}_i \in \mathbb{R}^{m \times n_i}$, $n = n_1 + \cdots + n_p$ *and* $\mathbf{A} \in \mathbb{R}^{m \times n}$, *then for any integer* r,

$$c_r(\mathbf{x}) = \begin{cases} \mathcal{A}_k \operatorname{bdiag}\left(c_r(\mathbf{s}_1), \ldots, c_r(\mathbf{s}_p)\right) \mathcal{A}_k' & \text{if} \quad r = 2k \\ \mathcal{A}_{k+1} \operatorname{bdiag}\left(c_r(\mathbf{s}_1), \ldots, c_r(\mathbf{s}_p)\right) \mathcal{A}_k' & \text{if} \quad r = 2k+1 \end{cases} \tag{14}$$

where matrices $\mathcal{A}_k$ *are returned by*

$$\mathcal{A}_k = \begin{pmatrix} \otimes^k \mathbf{A}_1 & \otimes^k \mathbf{A}_2 & \cdots & \otimes^k \mathbf{A}_p \end{pmatrix} \in \mathbb{R}^{m^k \times \sum_{j=1}^p n_j^k}. \tag{15}$$

Proof. See Appendix A. □

Properties (i) to (iv) will be used to establish the main results of this paper.

4. Contracted cumulant matrices

The higher-order-statistic solution to MICA is based on an extension of the former cumulant matrix which is referred herein as the "contracted" cumulant matrix. This generalizes a notion initially introduced on the fourth-order by Cardoso and Soulomiac in Ref. [8] in the JADE algorithm. The definition is first given hereafter on the fourth-order, and then on arbitrary order.

Definition 3 (contracted matrix cumulant on order 4). Let $\mathbf{M} = (m_{i_1,i_2})_{(i_1,i_2) \in [\![1,n]\!]^2}$ be an $n \times n$ matrix and $\mathbf{x}$ an $n \times 1$ random vector. The contracted cumulant matrix of order four of dimension $n \times n$, denoted as $\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}]$, has generic term

$$\left(\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}]\right)_{i_3,i_4} = \sum_{i_1,i_2=1}^{n} cum(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}) m_{i_1,i_2}. \tag{16}$$

The following lemma shows how the contracted cumulant matrix $\mathcal{Q}^{\mathbf{x}}_{(4)}$ can be directly expressed as the star product MacRae [32] (see Appendix Appendix C) of matrix M with the cumulant matrix c_4. This will be needed in Definition 4.

Lemma 2. *Let* $\mathbf{M} = (m_{i_1,i_2})_{(i_1,i_2) \in [\![1,n]\!]^2}$ *be a matrix with dimension* $n \times n$*; then*

$$\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}] = \mathbf{M} * c_4(\mathbf{x}) = \sum_{i_1,i_2=1}^{n} m_{i_1,i_2} \left[c_4(\mathbf{x})\right]_{i_1,i_2} \tag{17}$$

where $\left[c_4(\mathbf{x})\right]_{i_1,i_2}$ *is an* $n \times n$ *matrix corresponding to block* (i_1, i_2) *of the cumulant matrix* $c_4(\mathbf{x})$.

Proof. It suffices to realize that element (i_3, i_4) of matrix $\left[c_4(\mathbf{x})\right]_{i_1,i_2}$ is

$$\left(\left[c_4(\mathbf{x})\right]_{i_1,i_2}\right)_{i_3,i_4} = cum(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}). \tag{18}$$

An alternative proof is to apply properties $(P1)$ to $(P3)$ (see Appendix C.2, C.3 and C.4) of the star product $*$ to the expression of $c_4(\mathbf{x})$ given in Appendix B; it follows that

$$\mathbf{M} * c_4(\mathbf{x}) = \mathbb{E}(\mathbf{x}'\mathbf{M}\mathbf{x}\mathbf{x}\mathbf{x}') - \mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{M}'\mathbf{\Sigma} - \operatorname{tr}(\mathbf{\Sigma}\mathbf{M})\mathbf{\Sigma} \tag{19}$$

which is the explicit expression of $\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}]$ (see [7]). □

The generalization of the definition to any order r is obtained as follows.

Definition 4 (contracted cumulant matrix of order r). Given $q \in \mathbb{N}$, any matrix matrix $\mathbf{M}^{(q)}$ expands as

$$\mathbf{M}^{(q)} = \sum_{i_1,i_2,\ldots,i_q=1}^{n} m_{i_1,i_2,\cdots,i_q} \mathbf{E}_{i_1,i_2,\ldots,i_q} \tag{20}$$

where matrices $\mathbf{E}_{i_1,i_2,\ldots,i_q}$ have been introduced in Eq. (10) and $m_{i_1,i_2,\cdots,i_q}$ are arbitrary constants. Then, the $n \times n$ contracted cumulant matrix of order r is defined as

$$\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}] = \mathbf{M}^{(r-2)} * c_r(\mathbf{x}) = \sum_{i_1,i_2,\ldots,i_{r-2}=1}^{n} m_{i_1,i_2,\cdots,i_{r-2}} \Big[c_r(\mathbf{x})\Big]_{i_1,i_2,\cdots,i_{r-2}} \tag{21}$$

where $\Big[c_r(\mathbf{x})\Big]_{i_1,i_2,\cdots,i_{r-2}}$ denotes the matrix indexed by $(i_1, i_2, \cdots, i_{r-2})$ in the partition of $c_r(\mathbf{x})$ into matrices of dimensions $n \times n$; that is, element (i_{r-1}, i_r) of $\Big[c_r(\mathbf{x})\Big]_{i_1,i_2,\cdots,i_{r-2}}$ is

$$\left(\Big[c_r(\mathbf{x})\Big]_{i_1,i_2,\cdots,i_{r-2}}\right)_{i_{r-1},i_r} = \mathrm{cum}(x_{i_1} x_{i_2} \cdots x_{i_r}). \tag{22}$$

It results immediately that element (i_{r-1}, i_r) of the contracted cumulant matrix $\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}]$ reads

$$\left(\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}]\right)_{i_{r-1},i_r} = \sum_{i_1,i_2,\ldots,i_{r-2}=1}^{n} \mathrm{cum}(x_{i_1} x_{i_2} \cdots x_{i_r}) m_{i_1,i_2,\cdots,i_{r-2}}, \tag{23}$$

which corresponds to the natural generalization of (16).

5. r-th order MICA by means of joint block diagonalization

Equipped with the tools introduced hitherto, ones is now in a position to prove the main result of this paper concerning the block-diagonalization of the contracted cumulant matrices. The first step towards this perspective is to note that the cumulant matrix of the observations in the MICA model (1) reads

$$c_r(\mathbf{x}) = \begin{cases} \mathcal{A}_k \mathcal{C}_{2k}(\mathbf{s}) \mathcal{A}_k' & \text{if } r = 2k \\ \\ \mathcal{A}_{k+1} \mathcal{C}_{2k+1}(\mathbf{s}) \mathcal{A}_k' & \text{if } r = 2k+1 \end{cases} \tag{24}$$

with

$$\mathcal{C}_r(\mathbf{s}) = \mathrm{bdiag}\Big(c_r(\mathbf{s}_1), c_r(\mathbf{s}_2), \cdots, c_r(\mathbf{s}_p)\Big) \tag{25}$$

where use has been made of property (iv) of proposition 1 and of the statistical independence of the sources. The structure of matrices $\mathcal{A}_k$ in Eq. (24) is given in the following lemma:

Lemma 3 (structure of $\mathcal{A}_k$)**.**
Given any integer $k \geqslant 1$, matrix $\mathcal{A}_k = \begin{pmatrix} \otimes^k \mathbf{A}_1 & \otimes^k \mathbf{A}_2 & \cdots & \otimes^k \mathbf{A}_p \end{pmatrix}$ of dimension $n^k \times \sum_{i=1}^{p} n_i^k$ is

returned by the block matrix

$$\mathcal{A}_k = \begin{pmatrix} \mathbf{A}\mathbf{D}_{11\cdots11} \\ \mathbf{A}\mathbf{D}_{11\cdots12} \\ \vdots \\ \mathbf{A}\mathbf{D}_{11\cdots1n} \\ \mathbf{A}\mathbf{D}_{11\cdots21} \\ \mathbf{A}\mathbf{D}_{11\cdots22} \\ \vdots \\ \mathbf{A}\mathbf{D}_{11\cdots2n} \\ \vdots \\ \mathbf{A}\mathbf{D}_{1n\cdots nn} \\ \mathbf{A}\mathbf{D}_{2n\cdots nn} \\ \vdots \\ \mathbf{A}\mathbf{D}_{nn\cdots nn} \end{pmatrix} \tag{26}$$

where, for any $(i_1, i_2, \cdots, i_{k-1}) \in [\![1, n]\!]^{k-1}$*, the* $n \times \sum_{i=1}^{p} n_i^k$ *matrix* $\mathbf{D}_{i_1 i_2 \cdots i_{k-1}}$ *reads*

$$\mathbf{D}_{i_1 i_2 \cdots i_{k-1}} = \operatorname{bdiag}\left(\ell_{i_1}^1 \otimes \cdots \otimes \ell_{i_{k-1}}^1 \otimes \mathbf{I}_{n_1}, \ell_{i_1}^2 \otimes \cdots \otimes \ell_{i_{k-1}}^2 \otimes \mathbf{I}_{n_2}, \cdots, \ell_{i_1}^p \otimes \cdots \otimes \ell_{i_{k-1}}^p \otimes \mathbf{I}_{n_p} \right) \tag{27}$$

with ℓ_k^i *the* k*-th raw of matrix* $\mathbf{A}_i$ *(see subsection 2.4).*

Proof. See Appendix D. □

It directly results from lemma 3 and Eq. (24) that the cumulant matrix is made of $n \times n$ blocks

$$\Big[c_r(\mathbf{x})\Big]_{i_1, i_2, \cdots, i_{r-2}} = \begin{cases} \mathbf{A}\mathbf{D}_{i_1 i_3 \cdots i_{r-3}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{i_2 i_4 \cdots i_{r-2}} \mathbf{A}' & \text{if} \quad r = 2k \\ \mathbf{A}\mathbf{D}_{i_1 i_3 \cdots i_{r-2}} \mathcal{C}_{2k+1}(\mathbf{s}) \mathbf{D}'_{i_2 i_4 \cdots i_{r-3}} \mathbf{A}' & \text{if} \quad r = 2k+1. \end{cases} \tag{28}$$

One is now ready to state the main result of the paper.

Theorem 4 (joint block-diagonalization of contracted cumulant matrices)**.**
Let $\mathbf{s} = (\mathbf{s}'_1, \mathbf{s}'_2, \cdots, \mathbf{s}'_p)'$ *be a random vector where the* $\mathbf{s}_i$*'s are mutually independent with dimensions* $\dim(\mathbf{s}_i) = n_i$*,* $n = \sum_{i=1}^{p} n_i$*. Let also* $\mathbf{A}$ *be an* $n \times n$ *orthogonal matrix such that* $\mathbf{x} = \mathbf{A}\mathbf{s}$*. Then, for any integer* $r > 2$*, the contracted cumulant matrix reads*

$$\mathcal{Q}_{(r)}^{\mathbf{x}}[\mathbf{M}^{(r-2)}] = \mathbf{A} \operatorname{bdiag}\left(\Delta_{\mathbf{s}_1}^{(r)}, \Delta_{\mathbf{s}_2}^{(r)}, \cdots, \Delta_{\mathbf{s}_p}^{(r)} \right) \mathbf{A}' \tag{29}$$

with

$$(\forall i \in [\![1, p]\!]) \qquad \Delta_{\mathbf{s}_i}^{(r)} = \begin{cases} \mathcal{Q}_{(2k)}^{\mathbf{s}_i}\Big[\left(\otimes^{k-1} \mathbf{A}'_i\right) \mathbf{M}^{(2k-2)} \left(\otimes^{k-1} \mathbf{A}_i\right) \Big] & \textit{if} \quad r = 2k \\ \mathcal{Q}_{(2k+1)}^{\mathbf{s}_i}\Big[\left(\otimes^{k} \mathbf{A}'_i\right) \mathbf{M}^{(2k-1)} \left(\otimes^{k-1} \mathbf{A}_i\right) \Big] & \textit{if} \quad r = 2k+1 \end{cases} \tag{30}$$

Proof. See Appendix E. □

Theorem 4 clearly answers the questions raised in the introduction: contracted cumulant matrices are block-diagonalizable in the basis spanned by matrix $\mathbf{A}$ at any order r. In addition, it also returns the general structure of the diagonal blocks in terms of the cumulant matrices $\Delta_{\mathbf{s}_i}^{(r)}$ of the sources.

A particular consequence of Theorem 4 is given in the following corollary.

Corolary 5 (The ICA case)**.** *In the case of scalar sources, $p = n$ and $n_i = 1$ for any $i \in [\![1, p]\!]$, the contracted cumulant matrix reads*

$$\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}] = \mathbf{A}\,\mathrm{diag}\left(\delta^{(r)}_{\mathbf{s}_1}, \delta^{(r)}_{\mathbf{s}_2}, \cdots, \delta^{(r)}_{\mathbf{s}_p}\right)\mathbf{A}' \tag{31}$$

where

$$(\forall i \in [\![1, n]\!]) \qquad \delta^{(r)}_{\mathbf{s}_i} = \begin{cases} \left(\left(\otimes^{k-1}\mathbf{a}'_i\right)\mathbf{M}^{(2k-2)}\left(\otimes^{k-1}\mathbf{a}_i\right)\right)\mathrm{k}^{(2k)}_i & \text{if} \quad r = 2k \\ \left(\left(\otimes^{k}\mathbf{a}'_i\right)\mathbf{M}(2k-1)\left(\otimes^{k-1}\mathbf{a}_i\right)\right)\mathrm{k}^{(2k+1)}_i & \text{if} \quad r = 2k+1 \end{cases} \tag{32}$$

with $\mathbf{a}_i$ the i-th column of $\mathbf{A}$ and $\mathrm{k}^{(r)}_i = c_r(\mathbf{s}_i) = cum(\mathbf{s}_i, \cdots, \mathbf{s}_i)$ the r-order cumulant of source $\mathbf{s}_i$.

In particular, when all sources are scalar and $2k = 4$ (fourth-order statistics), corollary 5 returns the classical ICA result originally proved in [8, page 5],

$$\mathcal{Q}^{\mathbf{x}}_{(4)}\left[\mathbf{M}^{(2)}\right] = \mathbf{A}\,\mathrm{diag}\left(\mathrm{k}^{(4)}_1\mathbf{a}'_1\mathbf{M}^{(2)}\mathbf{a}_1, \mathrm{k}^{(4)}_2\mathbf{a}'_2\mathbf{M}^{(2)}\mathbf{a}_2, \cdots, \mathrm{k}^{(4)}_p\mathbf{a}'_p\mathbf{M}^{(2)}\mathbf{a}_p\right)\mathbf{A}' \tag{33}$$

with $\mathrm{k}^{(4)}_i$ the kurtosis of source $\mathbf{s}_i$. It is emphasized here that corollary 5 extends classical fourth-order ICA to any order r and theorem 4 generalizes it to any dimension. For instance, the block diagonal forms involved on orders $r \in \{3; 4; 5; 6\}$ (which are to be used in the experimental section of the paper) are readily found as:

$$\mathcal{Q}^{\mathbf{x}}_{(3)}[\mathbf{M}^{(1)}] = \mathbf{A}\,\mathrm{bdiag}\left(\mathcal{Q}^{\mathbf{s}_1}_{(3)}\left[\mathbf{A}'_1\mathbf{M}^{(1)}\right], \cdots, \mathcal{Q}^{\mathbf{s}_p}_{(3)}\left[\mathbf{A}'_p\mathbf{M}^{(1)}\right]\right)\mathbf{A}' \tag{34}$$

$$\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}^{(2)}] = \mathbf{A}\,\mathrm{bdiag}\left(\mathcal{Q}^{\mathbf{s}_1}_{(4)}\left[\mathbf{A}'_1\mathbf{M}^{(2)}\mathbf{A}_1\right], \cdots, \mathcal{Q}^{\mathbf{s}_p}_{(4)}\left[\mathbf{A}'_p\mathbf{M}^{(2)}\mathbf{A}_p\right]\right)\mathbf{A}' \tag{35}$$

$$\mathcal{Q}^{\mathbf{x}}_{(5)}[\mathbf{M}^{(3)}] = \mathbf{A}\,\mathrm{bdiag}\left(\mathcal{Q}^{\mathbf{s}_1}_{(5)}\left[\left(\mathbf{A}'_1 \otimes \mathbf{A}'_1\right)\mathbf{M}^{(3)}\mathbf{A}_1\right], \cdots, \mathcal{Q}^{\mathbf{s}_p}_{(5)}\left[\left(\mathbf{A}'_p \otimes \mathbf{A}'_p\right)\mathbf{M}^{(3)}\mathbf{A}_p\right]\right)\mathbf{A}' \tag{36}$$

$$\mathcal{Q}^{\mathbf{x}}_{(6)}[\mathbf{M}^{(4)}] = \mathbf{A}\,\mathrm{bdiag}\left(\mathcal{Q}^{\mathbf{s}_1}_{(6)}\left[\left(\mathbf{A}'_1 \otimes \mathbf{A}'_1\right)\mathbf{M}^{(4)}\left(\mathbf{A}_1 \otimes \mathbf{A}_1\right)\right], \cdots, \mathcal{Q}^{\mathbf{s}_p}_{(6)}\left[\left(\mathbf{A}'_p \otimes \mathbf{A}'_p\right)\mathbf{M}^{(4)}\left(\mathbf{A}_p \otimes \mathbf{A}_p\right)\right]\right)\mathbf{A}' \tag{37}$$

with $\mathbf{M}^{(1)}$, $\mathbf{M}^{(2)}$, $\mathbf{M}^{(3)}$ and $\mathbf{M}^{(4)}$ of dimensions $n \times 1$, $n \times n$, $n^2 \times n$, and $n^2 \times n^2$, respectively.

6. SJADE$_r$: a family of MICA algorithms

Theorem 4 makes it possible to propose a generalization of the JADE algorithm originally introduced by Cardoso and Soulomiac [8] to any dimension and to any order. Indeed, since it has been proven that contracted cumulant matrices $\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}]$ are block-diagonalizable in the basis spanned by matrix $\mathbf{A}$ for any order r, it suggests that joint block-diagonalization will effectively return the unknown mixing matrix in model (1). Two families of JBD algorithms have actually been proposed in the literature:

(a) algorithms that estimate an orthogonal block diagonalizer ([13, 43, 33]);
(b) algorithms that avoid the whitening step by estimating an non-orthogonal block diagonalizer ([37, 15, 14]).

Both approaches are applicable to the results of this paper. However, for the sake of consistency with the assumption of section 2 (matrix $\mathbf{A}$ is orthogonal as a result of pre-whitening), the first strategy only will be considered from now on.
As for the set of cumulant matrices to be jointly block-diagonalized, a natural choice is to consider the contracting matrices $\mathbf{M}^{(r-2)} = \mathbf{E}_{i_1,i_2,\cdots,i_{r-2}}$ which return $\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{E}_{i_1,i_2,\cdots,i_{r-2}}] = [c_r(\mathbf{x})]_{i_1,i_2,\cdots,i_{r-2}}$. The synopsis of the SJADE$_r$ algorithm that solves the MICA problem on order r is as follows.

Algorithm 1 (SJADE$_r$).

- `Whiten the observations,`
- `Construct of a set of contracted cumulant matrices:` for instance, the set of n^{r-2} matrices,

$$\mathcal{M}^{(r)} = \left\{ \mathcal{Q}^{\mathbf{x}}_{(r)}\Big[\mathbf{E}_{i_1,i_2,\cdots,i_{r-2}}\Big] \quad ; \quad 1 \leqslant i_1, i_2, \cdots, i_{r-2} \leqslant n \right\}, \tag{38}$$

  estimated from the data according the formulae worked out in the previous sections where empirical moment matrices are substituted for their theoretial versions [38].
- `Joint block-diagonalization (JBD):` estimation of the orthogonal matrix $\mathbf{A}'$ which jointly block-diagonalizes $\mathcal{M}^{(r)}$.
- `Separation of sources:` $\mathbf{s} = \mathbf{A}'\mathbf{x}$.

It is noteworthy that the particular case on the fourth-order coincides with the so-called SJADE algorithm introduced in Ref. [43]. It should also be noted at this point that the set

$$\left\{ \mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{E}_{i_1,i_2,\cdots,i_{r-2}}] \quad ; \quad 1 \leqslant i_1, i_2, \cdots, i_{r-2} \leqslant n \right\}$$

of contracted matrices is highly redundant due to the symmetries of the cumulant matrices $c_r(\mathbf{x})$. Specifically, it is seen from expression (23) that for any permutation $\sigma \in \mathfrak{S}_{r-2}$,

$$\begin{aligned} \left(\mathcal{Q}^{\mathbf{x}}_{(r)}[\sigma \cdot \mathbf{M}^{(r-2)}] \right)_{i_{r-1},i_r} &= \sum_{i_1,i_2,\ldots,i_{r-2}=1}^{n} \underbrace{\mathrm{cum}(x_{i_1}x_{i_2}\cdots x_{i_r})}_{=\mathrm{cum}(x_{i_{\sigma(1)}}\cdots x_{i_{\sigma(r-2)}}x_{i_{r-1}}x_{i_r})} m_{i_{\sigma(1)},i_{\sigma(2)},\cdots,i_{\sigma(r-2)}} \\ &= \left(\mathcal{Q}^{\mathbf{x}}_{(r)}[\mathbf{M}^{(r-2)}] \right)_{i_{r-1},i_r}. \end{aligned} \tag{39}$$

As a consequence, the set of matrices to be jointly block-diagonalized is indeed

$$\mathcal{M}^{(r)} = \left\{ \mathcal{Q}^{\mathbf{x}}_{(r)}\Big[\mathbf{E}_{i_1,i_2,\cdots,i_{r-2}}\Big] \quad ; \quad 1 \leqslant i_1 \leqslant i_2 \leqslant \cdots \leqslant i_{r-2} \leqslant n \right\}, \tag{40}$$

with cardinality

$$\mathrm{card}(\mathcal{M}^{(r)}) = \binom{n+r-3}{r-2} = \frac{n(n+1)\cdots(n+r-3)}{(r-2)!}. \tag{41}$$

For instance, on order 4, instead of block-diagonalizing the set of n^2 matrices (38), a wiser choice is to consider the set of $\frac{n(n+1)}{2}$ matrices

$$\mathcal{M}^{(4)} = \left\{ \mathcal{Q}^{\mathbf{x}}_{(4)}[\mathcal{E}_{i,j}], \quad 1 \leqslant i \leqslant j \leqslant n \right\} \tag{42}$$

where

$$\mathcal{E}_{i,j} = \begin{cases} \mathbf{e}_i\mathbf{e}_i' & \text{if } i = j \\ \\ \frac{1}{\sqrt{2}}\left(\mathbf{e}_i\mathbf{e}_j' + \mathbf{e}_j\mathbf{e}_i'\right) & \text{if } i < j\ . \end{cases} \tag{43}$$

This basis was first proposed in Ref.[7]. Similarly, on order 5, instead of block-diagonalizing the set of n^3 (38) matrices, it suffices to consider the set of $\frac{n(n+1)(n+2)}{6}$ matrices

$$\mathcal{M}^{(5)} = \left\{ \mathcal{Q}^{\mathbf{x}}_{(5)}[\mathcal{E}_{i,j,k}], \quad 1 \leqslant i \leqslant j \leqslant k \leqslant n \right\} \tag{44}$$

where

$$\mathcal{E}_{i,j,k} = \begin{cases} \mathbf{e}_i \otimes \mathbf{e}_i\mathbf{e}_i' & \text{if } i=j=k \\ \frac{1}{\sqrt{3}}\left(\mathbf{e}_k \otimes \mathbf{e}_i\mathbf{e}_i' + \mathbf{e}_i \otimes \mathbf{e}_k\mathbf{e}_i' + \mathbf{e}_i \otimes \mathbf{e}_i\mathbf{e}_k'\right) & \text{if } i=j<k \\ \frac{1}{\sqrt{3}}\left(\mathbf{e}_i \otimes \mathbf{e}_k\mathbf{e}_k' + \mathbf{e}_k \otimes \mathbf{e}_i\mathbf{e}_k' + \mathbf{e}_k \otimes \mathbf{e}_k\mathbf{e}_i'\right) & \text{if } i<j=k \\ \frac{1}{\sqrt{6}}\Big(\mathbf{e}_i \otimes \mathbf{e}_j\mathbf{e}_k' + \mathbf{e}_i \otimes \mathbf{e}_k\mathbf{e}_j' + \mathbf{e}_j \otimes \mathbf{e}_i\mathbf{e}_k' + \mathbf{e}_j \otimes \mathbf{e}_k\mathbf{e}_i' + \mathbf{e}_k \otimes \mathbf{e}_i\mathbf{e}_j' + \mathbf{e}_k \otimes \mathbf{e}_j\mathbf{e}_i'\Big) & \text{if } i<j<k\ . \end{cases} \tag{45}$$

7. On the advantage of combining several orders

Theorem 4 generalizes the JADE algorithm to any order and to any dimension. It also justifies other generalizations, such as the combination of orders. Indeed, since it has been proved that contracted cumulant matrices are block diagonalizable in the *same* basis whatever their orders, this makes possible to solve the MICA problem by considering the union of sets,

$$\mathcal{M} = \bigcup_{l=1}^{q} \mathcal{M}^{(r_l)}, \tag{46}$$

of contracted cumulant matrices at several orders $r_1, \ldots, r_q$. Let us refer to the corresponding algorithm as SJADE$_{r_1,\ldots,r_q}$. The motivation beyond such an approach is to allow the separation of sources characterized by higher-order statistics of different orders. By way of an example, let us consider the case of independent scalar sources s_i $(i = 1, \cdots, 4)$ such that $\mathrm{k}_1^{(4)} = \mathrm{k}_2^{(4)} = \mathrm{k}_3^{(6)} = \mathrm{k}_4^{(6)} = 0$, $\mathrm{k}_1^{(6)} \neq 0$, $\mathrm{k}_2^{(6)} \neq 0$, $\mathrm{k}_3^{(4)} \neq 0$, and $\mathrm{k}_4^{(4)} \neq 0$, where $\mathrm{k}_i^{(4)} = cum(s_i, s_i, s_i, s_i)$ et $\mathrm{k}_i^{(6)} = cum(s_i, s_i, s_i, s_i, s_i, s_i)$. According to corollary 5, the contracted cumulant matrices of orders 4 and 6 of the observations $\mathbf{x} = \mathbf{A}\mathbf{s}$, $\mathbf{A} \in \mathbb{R}^{4\times 4}$, $\mathbf{s} = (s_1, s_2, s_3, s_4)'$, have expressions

$$\mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}^{(2)}] = \mathbf{A}\,\mathrm{diag}\left(0, 0, \mathbf{a}_3'\mathbf{M}^{(2)}\mathbf{a}_3\mathrm{k}_3^{(4)}, \mathbf{a}_4'\mathbf{M}^{(2)}\mathbf{a}_4\mathrm{k}_4^{(4)}\right)\mathbf{A}' \tag{47}$$

$$\mathcal{Q}^{\mathbf{x}}_{(6)}[\mathbf{M}^{(4)}] = \mathbf{A}\,\mathrm{diag}\left((\mathbf{a}_1\otimes\mathbf{a}_1)'\mathbf{M}^{(4)}(\mathbf{a}_1\otimes\mathbf{a}_1)\mathrm{k}_1^{(6)}, (\mathbf{a}_2\otimes\mathbf{a}_2)'\mathbf{M}^{(4)}(\mathbf{a}_2\otimes\mathbf{a}_2)\mathrm{k}_2^{(6)}, 0, 0\right)\mathbf{A}', \tag{48}$$

respectively. It is clear that in such a scenario matrix $\mathbf{A}$ cannot be estimated neither from order 4 nor from order 6 alone because of the existence of some zero cumulants in each case (see assumption (H_2) of section 2.1). Nevertheless, resorting to the combination $\mathcal{Q}_{(4,6)} = \mathcal{Q}^{\mathbf{x}}_{(4)}[\mathbf{M}^{(2)}] + \mathcal{Q}^{\mathbf{x}}_{(6)}[\mathbf{M}^{(4)}]$ of orders 4 and 6, one has

$$\mathcal{Q}_{(4,6)} = \mathbf{A}\,\mathrm{diag}\left((\mathbf{a}_1\otimes\mathbf{a}_1)'\mathbf{M}^{(4)}(\mathbf{a}_1\otimes\mathbf{a}_1)\mathrm{k}_1^{(6)}, (\mathbf{a}_2\otimes\mathbf{a}_2)'\mathbf{M}^{(4)}(\mathbf{a}_2\otimes\mathbf{a}_2)\mathrm{k}_2^{(6)}, \mathbf{a}_3'\mathbf{M}^{(2)}\mathbf{a}_3\mathrm{k}_3^{(4)}, \mathbf{a}_4'\mathbf{M}^{(2)}\mathbf{a}_4\mathrm{k}_4^{(4)}\right)\mathbf{A}' \tag{49}$$

which shows that matrix $\mathbf{A}$ can be estimated by block-diagonalizing contracted cumulant matrices of the type $\mathcal{Q}_{(4,6)}$. This is illustrated with the numerical values $\mathrm{k}_1^{(6)} = 0.9730$, $\mathrm{k}_2^{(6)} = 1.2980$, $\mathrm{k}_3^{(4)} = -0.8003$, and $\mathrm{k}_4^{(4)} = 1.3614$ and the mixing matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{pmatrix} = \begin{pmatrix} 0.9787 & 0.0596 & 0.5216 & 0.7224 \\ 0.7127 & 0.6820 & 0.0967 & 0.1499 \\ 0.5005 & 0.0424 & 0.8181 & 0.6596 \\ 0.4711 & 0.0714 & 0.8175 & 0.5186 \end{pmatrix}. \tag{50}$$

Two set of contracted cumulant matrices, $\mathcal{M}^{(4)}$ and $\mathcal{M}^{(6)}$, are then constructed according to Eq. (40), and the joint diagonalization algorithm (*FFdiag*) of Ref. Ziehe et al. [47] is used to return three estimates of the inverse of matrix $\mathbf{A}$:

- $\mathbf{B}_4$, from the joint diagonalization of set $\mathcal{M}^{(4)}$;
- $\mathbf{B}_6$, from the joint diagonalization of set $\mathcal{M}^{(6)}$;
- $\mathbf{B}_{4,6}$, from the joint diagonalization of the union of sets $\mathcal{M}^{(4,6)} = \mathcal{M}^{(4)} \cup \mathcal{M}^{(6)}$.

The separation results are assessed by the product $\mathbf{B}_\bullet\mathbf{A}$, $\bullet = (4), (6), (4,6)$:

$$\mathbf{B}_4\mathbf{A} = \begin{pmatrix} -0.5109 & -0.6400 & -0.0000 & 0.0061 \\ 0.5266 & 0.6311 & -0.0000 & 0.0007 \\ 0.0420 & 0.2293 & 0.0001 & 0.0000 \\ -0.0479 & -0.2329 & 0.0039 & 0.0000 \end{pmatrix} \tag{51}$$

$$\mathbf{B}_6\mathbf{A} = \begin{pmatrix} 0.0027 & -0.0000 & -0.0631 & -0.1051 \\ -0.0000 & 0.0000 & -0.3507 & -0.0808 \\ -0.0001 & 0.0000 & 0.6606 & 0.2647 \\ -0.0000 & 0.0043 & 0.3823 & 0.0952 \end{pmatrix} \tag{52}$$

$$\mathbf{B}_{4,6}\mathbf{A} = \begin{pmatrix} 0.0000 & -0.1792 & -0.0000 & -0.0000 \\ 0.0000 & 0.0000 & -0.1813 & 0.0000 \\ -0.1435 & -0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.0801 \end{pmatrix}. \tag{53}$$

As expected, orders 4 and 6 alone are unable to separate the 4 sources whereas their combination returns a perfect separation (up to the ICA undeterminacies).

8. Towards automatic sizing of sources

The previous section has proposed a family of algorithms apt to separate sources of different dimensions. This section now addresses the issue of determining the respective source dimensions from the observations, an objective of considerable importance in practice. An original solution is devised based on the joint diagonalization of the contracted cumulant matrices introduced previously. This is addressed here on order 4 for the sake of conciseness, yet it is easily generalizable to other orders. It proceeds from the concept of irreducibility of the sources.

8.1. Irreducibility of vector sources

It has been shown in Ref. [43, 16, 18] that the irreducibility of sources is a fundamental assumption of MICA that guaranties the unicity of the solution. However, testing for the irreducibilty of vector sources from Definition 1 is not obvious. A simple but sufficient condition of irreducibility is proposed hereafter that makes use of the rank of the cumulant matrix.

Lemma 6. *Let* $\mathbf{y} \in \mathbb{R}^d$ *be a random vector satisfying the condition*

$$\mathbf{r}(c_4(\mathbf{y})) = \frac{d(d+1)}{2}. \tag{54}$$

Then $\mathbf{y}$ *is irreducible.*

Proof. Appendix F □

Remark 1. Lemma 6 can be easily extended to other orders; for instance, on order 6, it reads

$$\mathbf{r}(c_6(\mathbf{y})) = \frac{d(d+1)(d+2)}{6}. \tag{55}$$

It must be highlighted that conditions (54) and (55) are sufficient, but not necessary. In particular, one may be fulfilled whilst the other one is not.

This leads to the following proposition.

Proposition 7. *Suppose that vector sources* $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_p$ *are irreducible in the sense of lemma 6 and let* $\mathbf{V}$ *be the modal matrix in the eigenvalue decomposition* $\mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ *of the cumulant matrix* $c_4(\mathbf{x})$*. Then, matrix*

$$\mathcal{R} = \sum_{k=1}^{n} \left[\mathbf{V}'\mathbf{V}\right]_{k,k} \tag{56}$$

where $\left[\mathbf{V}'\mathbf{V}\right]_{k,k}$ *stands for the* k*-th diagonal block of dimension* $n \times n$ *in matrix* $\mathbf{V}'\mathbf{V}$ *has factorization*

$$\mathcal{R} \;=\; \mathbf{A} \begin{pmatrix} \frac{n_1+1}{2}\mathbf{I}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{n_2+1}{2}\mathbf{I}_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \frac{n_p+1}{2}\mathbf{I}_{n_p} \end{pmatrix} \mathbf{A}'. \tag{57}$$

Proof. Appendix F □

Proposition 7 means that the sought mixing matrix $\mathbf{A}$ that jointly block-diagonalizes the set $\mathcal{M}^{(4)}$ also diagonalizes matrix $\mathcal{R}$. The eigenvalues of $\mathcal{R}$ are then returned by the p positive quantities $\frac{n_1+1}{2}, \cdots, \frac{n_p+1}{2}$ with respective multiplicities $n_1, \cdots, n_p$. Two particular cases of proposition 7 are the following: for ICA, $\mathcal{R} = \mathbf{I}_n$, whereas for k-MICA, $\mathcal{R} = \frac{k+1}{2}\mathbf{I}_n$. In both cases, matrix $\mathcal{R}$ alone does not allow the recovery of the unknown permutation of independent subspaces in general.

The implications of Proposition 7 are twofold. First, it allows the determination of unknown source dimensions in the mixture by inspecting the eigenvalues of matrix $\mathcal{R}$. Second, it allows the grouping of source components by ordering the eigenvalues. This is now illustrated on a numerical example.

8.2. A numerical example

The following mixture is considered

$$\mathbf{x} = \mathbf{A} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4 \end{pmatrix}, \tag{58}$$

where sources $\mathbf{s}_i$, $i = 1, \cdots, 4$ are independent, verify condition (54), and have dimensions $\dim(\mathbf{s}_1) = \dim(\mathbf{s}_2) = 1$, $\dim(\mathbf{s}_3) = 2$, and $\dim(\mathbf{s}_4) = 3$. The corresponding cumulant matrices are

$$c_4(\mathbf{s}_1) = 1.2331; \qquad c_4(\mathbf{s}_2) = -0.5013; \qquad c_4(\mathbf{s}_3) = \begin{pmatrix} 0.1449 & 0.5279 & 0.5279 & 0.2170 \\ 0.5279 & 0.2170 & 0.2170 & 0.6261 \\ 0.5279 & 0.2170 & 0.2170 & 0.6261 \\ 0.2170 & 0.6261 & 0.6261 & 0.4908 \end{pmatrix};$$

$$c_4(\mathbf{s}_4) = \begin{pmatrix} 0.9636 & 0.2654 & 0.5322 & 0.2654 & 0.3258 & 0.5285 & 0.5322 & 0.5285 & 0.5011 \\ 0.2654 & 0.3258 & 0.5285 & 0.3258 & 0.6258 & 0.5011 & 0.5285 & 0.5011 & 0.5933 \\ 0.5322 & 0.5285 & 0.5011 & 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.5933 & 0.5717 \\ 0.2654 & 0.3258 & 0.5285 & 0.3258 & 0.6258 & 0.5011 & 0.5285 & 0.5011 & 0.5933 \\ 0.3258 & 0.6258 & 0.5011 & 0.6258 & 0.7213 & 0.3073 & 0.5011 & 0.3073 & 0.8174 \\ 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.3073 & 0.8174 & 0.5933 & 0.8174 & 0.2943 \\ 0.5322 & 0.5285 & 0.5011 & 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.5933 & 0.5717 \\ 0.5285 & 0.5011 & 0.5933 & 0.5011 & 0.3073 & 0.8174 & 0.5933 & 0.8174 & 0.2943 \\ 0.5011 & 0.5933 & 0.5717 & 0.5933 & 0.8174 & 0.2943 & 0.5717 & 0.2943 & 0.4781 \end{pmatrix}.$$

Note that the condition of irreducibility of the sources given in lemma 6 is perfectly satisfied since $\mathbf{r}(c_4(\mathbf{s}_1)) = \mathbf{r}(c_4(\mathbf{s}_2)) = 1$, $\mathbf{r}(c_4(\mathbf{s}_3)) = 3$, and $\mathbf{r}(c_4(\mathbf{s}_4)) = 6$. Beside, the mixing matrix $\mathbf{A}$ is the 7×7 orthogonal matrix,

$$\mathbf{A} = \begin{pmatrix} \text{-0.4658} & 0.1112 & \text{-0.558} & \text{-0.1817} & 0.2191 & 0.4352 & 0.4346 \\ \text{-0.4135} & \text{-0.5279} & 0.1581 & \text{-0.3195} & \text{-0.1066} & 0.3515 & \text{-0.5369} \\ \text{-0.3009} & 0.341 & 0.3417 & \text{-0.317} & 0.6814 & \text{-0.3005} & \text{-0.1464} \\ \text{-0.3266} & 0.1716 & \text{-0.5732} & 0.2306 & \text{-0.146} & \text{-0.4625} & \text{-0.4969} \\ \text{-0.4304} & \text{-0.4736} & 0.1607 & 0.0361 & \text{-0.1486} & \text{-0.5413} & 0.4982 \\ \text{-0.2615} & \text{-0.0762} & 0.2171 & 0.8426 & 0.3111 & 0.2634 & \text{-0.0504} \\ \text{-0.4014} & 0.5772 & 0.3813 & 0.0024 & \text{-0.5797} & 0.1432 & 0.0614 \end{pmatrix}.$$

These data are then used to compute the cumulant matrix $c_4(\mathbf{x})$ by means of Eq. 24, which is from now on considered as the only available observation. Note that $r = \mathbf{r}(c_4(\mathbf{x})) = \mathbf{r}(c_4(\mathbf{s}_1)) + \mathbf{r}(c_4(\mathbf{s}_2)) + \mathbf{r}(c_4(\mathbf{s}_3)) + \mathbf{r}(c_4(\mathbf{s}_4)) = 1 + 1 + 3 + 6 = 11$. The joint diagonalization algorithm of Ref. [8] is then applied to the contracted cumulant matrices $\mathcal{M}^{(4)}$ and returns

$$\mathbf{E} = \begin{pmatrix} 0.5301 & \text{-0.4458} & 0.2827 & \text{-0.4352} & 0.4346 & 0.0333 & 0.247 \\ 0.4377 & 0.5072 & 0.1577 & \text{-0.3515} & \text{-0.5369} & \text{-0.2976} & 0.161 \\ \text{-0.2347} & 0.0354 & 0.6996 & 0.3005 & \text{-0.1464} & 0.2745 & 0.5169 \\ 0.4274 & \text{-0.5085} & \text{-0.2677} & 0.4625 & \text{-0.4969} & 0.0534 & 0.1533 \\ 0.4142 & 0.4731 & \text{-0.1286} & 0.5413 & 0.4982 & \text{-0.0827} & 0.2001 \\ 0.0582 & 0.2287 & \text{-0.3952} & \text{-0.2634} & \text{-0.0504} & 0.8066 & 0.2562 \\ \text{-0.3386} & \text{-0.0907} & \text{-0.4017} & \text{-0.1432} & 0.0614 & \text{-0.418} & 0.7188 \end{pmatrix}.$$

The next step is to compute the eigenvalue decomposition $\mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ of $c_4(\mathbf{x})$ with $r = 11$ eigen-elements and to construct matrix $\mathcal{R}$ from Eq. (56):

$$\mathcal{R} = \begin{pmatrix} 1.5813 & 0.0631 & 0.0909 & 0.4541 & 0.0386 & \text{-0.0503} & \text{-0.0253} \\ 0.0631 & 1.5314 & 0.0127 & \text{-0.0752} & 0.4556 & 0.0315 & \text{-0.0479} \\ 0.0909 & 0.0127 & 1.6059 & \text{-0.1254} & \text{-0.0334} & 0.0993 & 0.2499 \\ 0.4541 & \text{-0.0752} & \text{-0.1254} & 1.502 & \text{-0.0178} & 0.0223 & 0.0542 \\ 0.0386 & 0.4556 & \text{-0.0334} & \text{-0.0178} & 1.4471 & 0.1756 & 0.0038 \\ \text{-0.0503} & 0.0315 & 0.0993 & 0.0223 & 0.1756 & 1.5247 & 0.0546 \\ \text{-0.0253} & \text{-0.0479} & 0.2499 & 0.0542 & 0.0038 & 0.0546 & 1.8077 \end{pmatrix}.$$

Finally, using the above entries of $\mathbf{E}$ and $\mathcal{R}$,

$$\mathbf{E}'\mathcal{R}\mathbf{E} = \begin{pmatrix} 2 & \text{-0} & \text{-0} & 0 & \text{-0} & 0 & 0 \\ \text{-0} & 2 & 0 & \text{-0} & 0 & \text{-0} & \text{-0} \\ \text{-0} & 0 & 1.5 & \text{-0} & 0 & \text{-0} & 0 \\ 0 & \text{-0} & \text{-0} & 1 & \text{-0} & 0 & \text{-0} \\ \text{-0} & 0 & 0 & \text{-0} & 1 & \text{-0} & 0 \\ 0 & \text{-0} & \text{-0} & 0 & \text{-0} & 1.5 & \text{-0} \\ 0 & \text{-0} & 0 & \text{-0} & 0 & \text{-0} & 2 \end{pmatrix}.$$

As indicated by proposition 7, matrix $\mathbf{E}'\mathcal{R}\mathbf{E}$ is made of 3 diagonal blocks with $n_1 = n_2 = 1$, $n_3 = 2$, and $n_4 = 3$. This returns the *a priori* unknown dimensions of the vector sources. Eventually, the permutation that rearrange the eigenvalues in decreasing order is

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

such that

$$\mathbf{P}'\mathbf{E}'\mathbf{A} = \begin{pmatrix} \text{-0.8289} & 0.4456 & 0.3382 & 0 & 0 & -0 & -0 \\ \text{-0.5545} & \text{-0.5748} & \text{-0.6017} & -0 & -0 & 0 & 0 \\ \text{-0.0737} & \text{-0.6863} & 0.7236 & 0 & -0 & -0 & -0 \\ 0 & 0 & 0 & \text{-0.7239} & 0.6899 & -0 & -0 \\ -0 & -0 & -0 & 0.6899 & 0.7239 & 0 & 0 \\ 0 & 0 & 0 & -0 & -0 & -0 & 1 \\ -0 & 0 & -0 & 0 & 0 & -1 & -0 \end{pmatrix},$$

which correctly groups together the components of the independent vector sources. This proves that $\mathbf{B} = \mathbf{P}'\mathbf{E}'$ is a solution to the MICA problem up to the usual indeterminations.

8.3. Estimation in the presence of noise

The previous example has been designed to satisfy exactly the condition of irreducibility of the sources, $\mathbf{r}(c_4(\mathbf{x})) = \sum \frac{n_i(n_i+1)}{2}$. This assumes the data are free of noise. In order to address the more realistic situation were noise is present, let us add to matrix $c_4(\mathbf{x})$ a small perturbation $\epsilon c_4(\mathbf{z})$, where $\mathbf{z}$ is a random vector independent of $\mathbf{s}$ and ϵ a small positive parameter:

$$c_4 = c_4(\mathbf{x}) + \epsilon c_4(\mathbf{z}). \tag{59}$$

In such a situation, one will generally have $\mathbf{r}(c_4) = \frac{n(n+1)}{2} \neq \sum \frac{n_i(n_i+1)}{2}$. After computing $\hat{\mathbf{E}}$ by joint diagonalization, it remains to find the effective rank r of $c_4(\mathbf{x})$ as would be observed in the absence of noise. A natural idea is to try all values from $r = n$ (the ICA case with n mutually independent sources) to $r = \frac{n(n-1)}{2} + 1$ (2 independent sources with dimensions $(n-1)$ and 1) and select that value which makes matrix $\hat{\mathbf{E}}'\mathcal{R}_r\hat{\mathbf{E}}$ the closest to a diagonal matrix. This can be tested by means of Amari's index [5] which takes values within 0 and 1: the closer it is to 0, the more diagonal the matrix.

The following example generates $\mathbf{z}$ from a standardized Gaussian and sets $\epsilon = 10^{-2}$; the rank of the cumulant matrix is $\mathbf{r}(c_4) = \mathbf{r}(c_4(\mathbf{x}) + \epsilon c_4(\mathbf{z})) = \frac{n(n+1)}{2} = 28$. The corresponding Amari's indices are

$$\{0.038, 0.048, 0.022, 0.016, 0.001, 0.025, 0.035, 0.035, 0.040, 0.044, 0.048, 0.046, 0.040, 0.039, 0.033, 0.032\}$$

for r varying from $n = 7$ to $\frac{n(n-1)}{2} + 1 = 22$. The minimum is found in the fifth entry, which corresponds to $r = n + 4 = 11$ and, accordingly,

$$\hat{\mathbf{E}}'\mathcal{R}_{11}\hat{\mathbf{E}} = \begin{pmatrix} \mathbf{2.0000} & 0.0000 & 0.0098 & \text{-0.0027} & 0.0015 & \text{-0.0045} & \text{-0.0001} \\ 0.0000 & \mathbf{2.0000} & 0.0054 & \text{-0.0016} & \text{-0.0003} & 0.0016 & \text{-0.0000} \\ 0.0098 & 0.0054 & \mathbf{1.4996} & \text{-0.0012} & 0.002 & 0.0003 & \text{-0.0106} \\ \text{-0.0027} & \text{-0.0016} & \text{-0.0012} & \mathbf{1.0002} & \text{-0.0013} & \text{-0.0015} & \text{-0.0010} \\ 0.0015 & \text{-0.0003} & 0.002 & \text{-0.0013} & \mathbf{1.0002} & \text{-0.0004} & 0.0029 \\ \text{-0.0045} & 0.0016 & 0.0003 & \text{-0.0015} & \text{-0.0004} & \mathbf{1.4998} & 0.0096 \\ \text{-0.0001} & \text{-0.0000} & \text{-0.0106} & \text{-0.001} & 0.0029 & 0.0096 & \mathbf{2.0002} \end{pmatrix}.$$

It is seen that matrix $\hat{\mathbf{E}}'\mathcal{R}_{11}\hat{\mathbf{E}}$ is quite close to a diagonal matrix. After rearranging the diagonal elements in decreasing order, one finds the corresponding permutation matrix $\mathbf{P}$, which finally returns the separation matrix

$$\hat{\mathbf{B}} = \mathbf{P}'\hat{\mathbf{E}}' = \begin{pmatrix} 0.5301 & 0.4373 & \text{-0.2347} & 0.4274 & 0.4145 & 0.0581 & \text{-0.3388} \\ \text{-0.4454} & 0.5069 & 0.0356 & \text{-0.5085} & 0.4736 & 0.2288 & \text{-0.0902} \\ 0.2472 & 0.1609 & 0.5172 & 0.1539 & 0.1997 & 0.2564 & 0.7185 \\ 0.033 & \text{-0.2975} & 0.2746 & 0.0534 & \text{-0.083} & 0.8063 & \text{-0.4185} \\ 0.2815 & 0.1595 & 0.6997 & \text{-0.2666} & \text{-0.1303} & \text{-0.3953} & \text{-0.4019} \\ 0.4357 & \text{-0.5368} & \text{-0.144} & \text{-0.4975} & 0.4974 & \text{-0.0512} & 0.0606 \\ \text{-0.4351} & \text{-0.3518} & 0.3008 & 0.4622 & 0.5412 & \text{-0.2637} & \text{-0.1433} \end{pmatrix}.$$

The quality of the separation is assessed by the matrix product,

$$\hat{\mathbf{B}}\mathbf{A} = \begin{pmatrix} \mathbf{-0.5544} & \mathbf{-0.5749} & \mathbf{-0.6018} & 0.0000 & 0.0002 & -0.0003 & 0.0003 \\ \mathbf{-0.0742} & \mathbf{-0.6861} & \mathbf{0.7237} & 0.0001 & -0.0001 & -0.0001 & 0.0006 \\ \mathbf{-0.829} & \mathbf{0.4459} & \mathbf{0.3377} & 0.0001 & 0.0005 & -0.0001 & -0.0005 \\ 0.0005 & -0.0002 & -0.0001 & \mathbf{0.6897} & \mathbf{0.7241} & -0.0001 & -0.0004 \\ 0.0003 & -0.0001 & 0.0000 & \mathbf{-0.7241} & \mathbf{0.6897} & 0.0005 & -0.0029 \\ -0.0002 & 0.0008 & -0.0000 & -0.0018 & 0.0023 & 0.0003 & \mathbf{1.0000} \\ 0.0003 & 0.0002 & 0.0001 & -0.0004 & 0.0003 & \mathbf{-1.0000} & 0.0003 \end{pmatrix},$$

which clearly singles out four independent subspaces of dimensions 3, 2, 1, and 1.

The synopsis of the algorithm is as follows.

Algorithm 2 (A simple Fourth-order Multidimensional ICA).

- `Whiten the observations,`
- `Construct of a set of contracted cumulant matrices:` for instance, the set of $\frac{n(n+1)}{2}$ matrices,

$$\mathcal{M}^{(4)} = \left\{ \mathcal{Q}^{\mathbf{x}}_{(4)}[\mathcal{E}_{i_1,i_2}] = \mathbb{E}(\mathbf{x}'\mathcal{E}_{\mathbf{i_1},\mathbf{i_2}}\mathbf{x}\mathbf{x}\mathbf{x}') - 2(\mathcal{E}_{i_1,i_2}) - \mathrm{tr}(\mathcal{E}_{i_1,i_2})\mathbf{I}_n; 1 \leqslant i_1 \leqslant i_2 \leqslant n \right\} \tag{60}$$

- `Joint diagonalization (JD):` estimation of the orthogonal matrix $\mathbf{E}'$ by JD of $\mathcal{M}^{(4)}$.
- `Compute the eigenvalue decomposition` $\mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ `of` $c_4(\mathbf{x})$, `where` $\mathbf{V} \in \mathbb{R}^{n^2 \times \frac{n(n-1)}{2}+1}$.
- `For` $r = n$ `to` $r = \frac{n(n-1)}{2} + 1$
  `compute` $\mathbf{V}_r = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r \end{pmatrix}$ (where $\mathbf{v}_i$ is the $i-$th column of $\mathbf{V}$),
  `compute` $\mathbf{V}_r\mathbf{V}_r'$ `and` $\mathcal{R}_r$ (by means of Eq. 56),
  `compute the matrices` $\mathbf{D}_r = \mathbf{E}'\mathcal{R}_r\mathbf{E}$ `and the Amari-index of` $\mathbf{D}_r$, `noted error(`r`),`
- `End.`

- `Select` r_0 `that minimizes error(`r`) and the corresponding` $\mathbf{D}_{r_0}$.
- `Estimate the sources dimensions.`
- `Compute the permutation matrix` $\mathbf{P}$ `by the eigenvalue decomposition of` $\mathrm{diag}(\mathrm{diag}(\mathbf{D}_{r_0}))$.
- `Separation of sources:` $\mathbf{s} = \mathbf{P}'\mathbf{E}'\mathbf{x}$.

9. Experimental analyzes

This section now investigates the performance of the proposed SJADE$_r$ algorithm on two numerical examples.

9.1. Performance index for MICA

In order to measure the quality of the MICA separation, it is proposed to generalize the performance index originally introduced in Moreau [36]. The idea is to assess the proximity of $\mathbf{G} = \mathbf{B}\mathbf{A}$ to the product $\mathbf{P}\mathbf{D}$ of a permutation matrix and a block-diagonal matrix. Assume without loss of generality that the estimated sources are sorted according to their dimensions, i.e. $n_1 \leq n_2 \leq \cdots \leq n_p$. Next, let us partition the $n \times n$ matrix $\mathbf{G}$ into blocks $\mathbf{G}^{ij}$ of dimension $n_i \times n_j$, with $(i,j) \in [\![1,p]\!]^2$. Then, define the $p \times p$ matrix $\mathcal{G}$ as

$$\mathcal{G}_{ij} = \frac{1}{n_i n_j} \sum_{k,l=1}^{n_i,n_j} |(\mathbf{G}^{ij})_{kl}|. \tag{61}$$

Having matrix $\mathbf{G}$ block-diagonal (for blocks of size $n_1 \leq n_2 \leq \cdots \leq n_p$) is equivalent to having matrix $\mathcal{G}$ diagonal. Therefore, the performance index $I(\mathbf{G})$ for MICA is defined as

$$I(\mathbf{G}) = \frac{1}{2p(p-1)} \left[\sum_i^p \left(\sum_j^p \frac{\mathcal{G}_{ij}}{\max_l \mathcal{G}_{il}} - 1 \right) + \sum_j^p \left(\sum_i^p \frac{\mathcal{G}_{ij}}{\max_l \mathcal{G}_{lj}} - 1 \right) \right]. \tag{62}$$

A perfect separation corresponds to $I(\mathbf{G}) = 0$ (or $-\infty$ on a logarithmic scale). Note that for $\forall i, n_i = 1$, the proposed performance index reduces to the one introduced in [36] for scalar sources.

9.2. Comparisons

9.2.1. First simulation

This first example considers the case of three synthetic sources, of which two are vectors of dimension 2:

$$\begin{aligned} \mathbf{s}_1 &= \begin{pmatrix} \exp(2(0.007t + 0.5 - \text{floor}(0.007t + 0.5)) - 1) \\ 2(0.007t + 0.5 - \text{floor}(0.007t + 0.5)) - 1 \end{pmatrix} \\ \mathbf{s}_2 &= \begin{pmatrix} \exp(\cos(0.3t)) \\ \cos(0.3t) \end{pmatrix} \\ \mathbf{s}_3 &= \text{square}(0.2t) \end{aligned} \tag{63}$$

with $t = 0, 1, \cdots, 2999$ and "floor" the operator that rounded down to the nearest whole number. Square mixing matrix $\mathbf{A}_i$, $i = 1, \cdots, 100$, were randomly generated by sampling a uniform distribution in the interval $[0, 1]$, so as to produce the observations $\mathbf{x}_i = \mathbf{A}_i\mathbf{s}$. No noise was added. The JBD algorithm of Ref. [43] was used with threshold $\theta = 0.06$. Performances of SJADE$_3$, SJADE [43], SJADE$_5$, SJADE$_6$, MHICA,[42] and MSOBI [42] are compared in Fig. 1. It is seen that SJADE$_6$ evidences the best performance on this example.

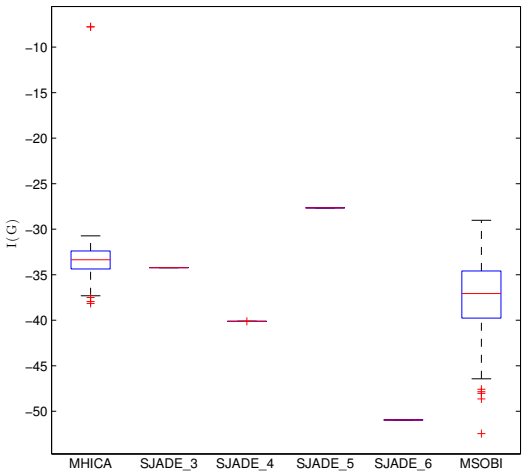


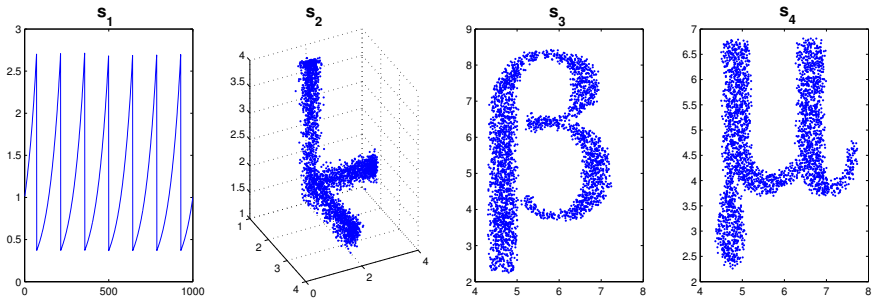
Figure 1: Box-plot of $I(\mathbf{G})$ for SJADE$_3$, SJADE [43], SJADE$_5$, SJADE$_6$, MHICA,[42] et MSOBI [42] (100 Monte-Carlo runs).

9.2.2. Second simulation

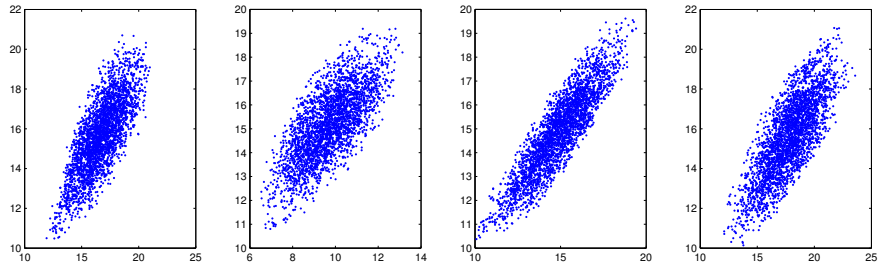
The second example is inspired from Ref. [39, 43], where the random vector $\mathbf{s} = (\mathbf{s}_1', \mathbf{s}_2', \mathbf{s}_3', \mathbf{s}_4')'$, of dimension 8 is composed of 4 sources $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$ and $\mathbf{s}_4$, of dimensions $1, 3, 2$ and 2, respectively. Source $\mathbf{s}_1$ is computed from formula

$$\mathbf{s}_1(\omega) = \exp(2(0.007\omega + 0.5 - \text{floor}(0.007\omega + 0.5)) - 1), \tag{64}$$

whereas sources $\mathbf{s}_2$, $\mathbf{s}_3$ and $\mathbf{s}_4$ take the shape of a trihedron, a "β", and a "μ", respectively. Figure 2.(a) displays 3000 samples of the sources and Fig. 2.(b) their observations after mixing with a matrix $\mathbf{A} \in \mathbb{R}^{8\times 8}$ whose entries are sampled in a uniform distribution on $[0, 1]$.



(a) The four sources.



(b) Projection of mixed sources on 4 planes.

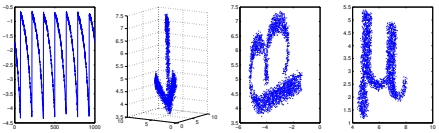
Figure 2: MICA, second simulation.

The performance indices of the separations obtained with SJADE$_3$, SJADE [43], SJADE$_5$, SJADE$_6$, MHICA,[42] and MSOBI [42] are displayed in Table 1. It is seen that SJADE$_5$ achieves the best performance in this case.

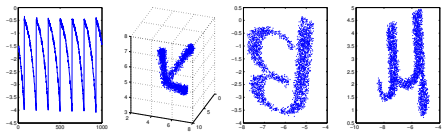
Méthode	MSOBI	SJADE$_3$	SJADE$_4$	SJADE$_5$	SJADE$_6$	MHICA
$I(\mathbf{G})[dB]$	-4.05	- 26.76	-30.10	-32.09	-24.21	-24.95

Table 1: Performance index (MICA, seconde simulation)

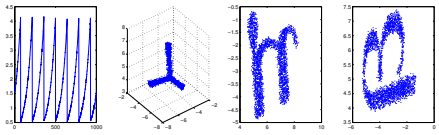
The separated sources are displayed in Fig. 3 for SJADE$_3$ to SJADE$_6$. The product of the estimated separation matrix $\mathbf{B}$ with the mixing matrix $\mathbf{A}$ is shown in Fig. 4 for SJADE and SJADE$_5$. It is clearly seen that $\mathbf{B}_4$ and $\mathbf{B}_5$ are equal up to a permutation, that is the indeterminacy of MICA.



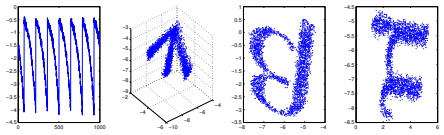
(a) Separated sources with SJADE_3.



(b) Separated sources with SJADE_4.

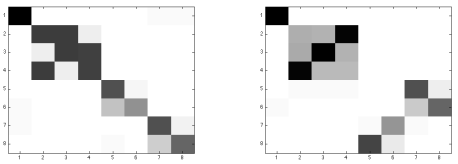


(c) Separated sources with SJADE_5.



(d) Separated sources with SJADE_6.

Figure 3: Separated sources (MICA, second simulation)



(a) $\mathbf{G}(\text{SJADE}_4) = \mathbf{B}_4\mathbf{A}$; (b) $\mathbf{G}(\text{SJADE}_5) = \mathbf{B}_5\mathbf{A}$;

Figure 4: Product of the estimated separation matrix with the mixing matrix (MICA, second simulation)

10. Conclusion

The present work addressed the problem of multidimensional ICA in the most general setting where sources may have different dimensions. The objective was to propose a solution that does not make use of the strong assumptions of mutually independent scalar sources (ICA) or independent groups of components of equal size (k-MICA). In particular, it has been shown that contracted cumulant matrices are all block-diagonalizable in the same basis. This opens the door to a family of algorithms coined $\text{SJADE}_{r_1,\ldots,r_q}$ that jointly block-diagonalize a set of contracted cumulant matrices at various orders. One advantage is that different orders can be easily combined together according to the nature of the sources in order to improve their separation. Another result of the paper is to provide a solution to infer the dimensions of the sources when these are unknown. The idea proceeds from the concept of irreducibility of vector sources and leads to an algorithm that provides th e source dimensions as well as a permutation matrix that groups together the components of each vector source. The theoretical results of the paper have been illustrated by means of a few numerical examples, which have clearly evidenced the advantage of exploiting various higher orders.

Appendix A. Proof of proposition 1

Point (i) is a direct consequence of the property that $\forall r > 2, \quad cum(z_1, z_2, \cdots, z_r) = 0$, for a Gaussian vector.
Point (ii) is a direct consequence of the property (see e.g. [35, p. 280]) that

$$cum(x_1 + y_1, x_2 + y_2, \cdots, x_r + y_r) = cum(x_1, x_2, \cdots, x_r) + cum(y_1, y_2, \cdots, y_r) \tag{A.1}$$

for independent random vectors $\mathbf{x}$ and $\mathbf{y}$.
Point (iii) is proved here for $r = 2k$ only, the case $r = 2k + 1$ following similar lines. It suffices to use the

multilinearity of cumulants with $x_j = \sum_{i=1}^n a_{j,i} s_i$ for $j \in [\![1,m]\!]$, where $\mathbf{A} = (a_{j,i})_{(j,i)\in[\![1,m]\!]\times[\![1,n]\!]}$:

$$\begin{aligned}
c_r(\mathbf{x}) &= \sum_{j_1,j_2,\dots,j_{2k}=1}^{m} cum(x_{j_1},\cdots,x_{j_{2k}})\mathbf{E}_{j_1,j_2,\dots,j_{2k}} \\
&= \sum_{j_1,j_2,\dots,j_{2k}=1}^{m} cum\left(\sum_{i_1=1}^{n} a_{j_1,i_1}s_{i_1},\cdots,\sum_{i_{2k}=1}^{n} a_{j_{2k},i_{2k}}s_{i_{2k}}\right)\mathbf{E}_{j_1,j_2,\dots,j_{2k}} \\
&= \sum_{j_1,j_2,\dots,j_{2k}=1}^{m}\left(\sum_{i_1,i_2,\dots,i_{2k}=1}^{n} a_{j_1,i_1}\cdots a_{j_{2k},i_{2k}} cum(s_{i_1},\cdots,s_{i_{2k}})\right)\mathbf{E}_{j_1,j_2,\dots,j_{2k}} \\
&= \sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}})\left(\sum_{j_1,j_2,\dots,j_{2k}=1}^{m} a_{j_1,i_1}\cdots a_{j_{2k},i_{2k}}\mathbf{E}_{j_1,j_2,\dots,j_{2k}}\right) \\
&= \sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}}) \sum_{j_1,j_2,\dots,j_{2k}=1}^{m} a_{j_1,i_1}\cdots a_{j_{2k},i_{2k}}\left(\mathbf{e}_{j_1}\mathbf{e}'_{j_2}\otimes\cdots\otimes\mathbf{e}_{j_{2k-1}}\mathbf{e}'_{j_{2k}}\right). \qquad \text{(A.2)}
\end{aligned}$$

The notation $(\mathbf{e}_j)_{j\in[\![1,m]\!]}$ denotes the canonical basis of $\mathbb{R}^m$. Now, denoting $(\mathbf{f}_i)_{i\in[\![1,n]\!]}$ the canonical basis of $\mathbb{R}^n$, it comes $\mathbf{A}\mathbf{f}_i = \sum_{j=1}^m a_{j,i}\mathbf{e}_j$, for any $i \in [\![1,n]\!]$; thus,

$$\begin{aligned}
c_r(\mathbf{x}) &= \sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}})\left(\left(\sum_{j_1=1}^{m} a_{j_1,i_1}\mathbf{e}_{j_1}\right)\left(\sum_{j_2=1}^{m} a_{j_2,i_2}\mathbf{e}_{j_2}\right)'\otimes\cdots\right. \\
&\qquad\left.\cdots\otimes\left(\sum_{j_{2k-1}=1}^{m} a_{j_{2k-1},i_{2k-1}}\mathbf{e}_{j_{2k-1}}\right)\left(\sum_{j_{2k}=1}^{m} a_{j_{2k},i_{2k}}\mathbf{e}_{j_{2k}}\right)'\right) \\
&= \sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}})\left((\mathbf{A}\mathbf{f}_{i_1})(\mathbf{A}\mathbf{f}_{i_2})'\otimes\cdots\otimes\left(\mathbf{A}\mathbf{f}_{i_{2k-1}}\right)(\mathbf{A}\mathbf{f}_{i_{2k}})'\right) \\
&= \sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}})(\mathbf{A}\otimes\cdots\otimes\mathbf{A})(\mathbf{f}_{i_1}\mathbf{f}'_{i_2}\otimes\cdots\otimes\mathbf{f}_{i_{2k-1}}\mathbf{f}'_{i_{2k}})(\mathbf{A}\otimes\cdots\otimes\mathbf{A})' \\
&= (\mathbf{A}\otimes\cdots\otimes\mathbf{A})\left(\sum_{i_1,i_2,\dots,i_{2k}=1}^{n} cum(s_{i_1},\cdots,s_{i_{2k}})(\mathbf{f}_{i_1}\mathbf{f}'_{i_2}\otimes\cdots\otimes\mathbf{f}_{i_{2k-1}}\mathbf{f}'_{i_{2k}})\right)(\mathbf{A}\otimes\cdots\otimes\mathbf{A})' \\
&= (\mathbf{A}\otimes\cdots\otimes\mathbf{A})\,c_r(\mathbf{s})\,(\mathbf{A}\otimes\cdots\otimes\mathbf{A})'. \qquad \text{(A.3)}
\end{aligned}$$

Point (iv) is again proved for $r = 2k$ only, the other case following similarly. One has

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \sum_{j=1}^{p} \mathbf{A}_j\mathbf{s}_j \qquad \text{(A.4)}$$

where the independence of the sources $\mathbf{s}_j$ implies the independence of the components $\mathbf{A}_j\mathbf{s}_j$. Using (ii) and

(iii), it comes

$$\begin{aligned} c_r(\mathbf{x}) &= c_r(\mathbf{As}) = c_r\left(\sum_{j=1}^{p}\mathbf{A}_j\mathbf{s}_j\right) \\ &= \sum_{j=1}^{p} c_r\left(\mathbf{A}_j\mathbf{s}_j\right) \\ &= \sum_{j=1}^{p} \otimes^k \mathbf{A}_j c_r(\mathbf{s}_j) \otimes^k \mathbf{A'}_j \\ &= \underbrace{\left(\otimes^k \mathbf{A}_1 \quad \cdots \quad \otimes^k \mathbf{A}_p \right)}_{=\mathcal{A}_k} \begin{pmatrix} c_r(\mathbf{s}_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_r(\mathbf{s}_p) \end{pmatrix} \underbrace{\begin{pmatrix} \otimes^k \mathbf{A'}_1 \\ \vdots \\ \otimes^k \mathbf{A'}_p \end{pmatrix}}_{\mathcal{A'}_k}. \end{aligned} \tag{A.5}$$

Appendix B. Cumulant matrices

Let $\mathbf{x}$ be a real-valued random vector of dimension $p \times 1$, with assumed zero-mean for simplicity ($\mathbb{E}[\mathbf{x}] = 0$). Let us further assume that $\mathbb{E}[|\mathbf{x}|^n] < \infty$ for a given integer $n \geq 6$ so as to guarantee the existence of moments and cumulants up to order six.
Next let us define the four matrices:

$$\mathbf{K} = \mathbf{I}_{p^2} + \mathbf{K}_{pp} \tag{B.1}$$
$$\mathbf{P} = \mathbf{I}_{p^3} + \mathbf{K}_{pp} \otimes \mathbf{I}_p + \mathbf{K}_{pp^2} \tag{B.2}$$
$$\mathbf{R} = \mathbf{I}_{p^3} + \mathbf{I}_p \otimes \mathbf{K}_{pp} + \mathbf{K}_{p^2p} \tag{B.3}$$
$$\mathbf{Q} = \mathbf{I}_{p^3} + \mathbf{K}_{pp} \otimes \mathbf{I}_p + \mathbf{I}_p \otimes \mathbf{K}_{pp} + \mathbf{K}_{p^2p} + \mathbf{K}_{pp^2} + \mathbf{K}_{p^2p}(\mathbf{K}_{pp} \otimes \mathbf{I}_p) \tag{B.4}$$

which depend only on dimension p. Thus, the cumulant matrices of orders 2 to 6 read

$$c_2(\mathbf{x}) = m_2(\mathbf{x}) = \mathbb{E}[\mathbf{x}\mathbf{x}'] = \mathbf{\Sigma} \tag{B.5}$$
$$c_3(\mathbf{x}) = m_3(\mathbf{x}) = \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x}] \tag{B.6}$$
$$c_4(\mathbf{x}) = m_4(\mathbf{x}) - \mathbf{K}\Big\{m_2(\mathbf{x}) \otimes m_2(\mathbf{x})\Big\} - \mathbf{vec}m_2(\mathbf{x})\mathbf{vec}'m_2(\mathbf{x}) \tag{B.7}$$
$$c_5(\mathbf{x}) = m_5(\mathbf{x}) - \mathbf{R}\Big\{m_3(\mathbf{x}) \otimes m_2(\mathbf{x})\Big\}\mathbf{K} - \mathbf{P}\Big\{m_3(\mathbf{x})' \otimes \mathbf{vec}m_2(\mathbf{x})\Big\} - \mathbf{vec}m_3(\mathbf{x})\mathbf{vec}'m_2(\mathbf{x}) \tag{B.8}$$
$$c_6(\mathbf{x}) = m_6(\mathbf{x}) - \Gamma_{(4,2)}(\mathbf{x}) - \Gamma_{(3,3)}(\mathbf{x}) + 2\Gamma_{(2,2,2)}(\mathbf{x}), \tag{B.9}$$

respectively, where

$$\begin{aligned} m_4(\mathbf{x}) &= \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x}\mathbf{x}'] \\ m_5(\mathbf{x}) &= \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x}\mathbf{x}' \otimes \mathbf{x}] \\ m_6(\mathbf{x}) &= \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x}\mathbf{x}' \otimes \mathbf{x}\mathbf{x}'] \\ \mathbf{M}_4(\mathbf{x}) &= \mathbb{E}[\mathbf{x}\mathbf{x}' \otimes \mathbf{x} \otimes \mathbf{x}] \\ \Gamma_{(4,2)}(\mathbf{x}) &= \mathbf{R}\Big\{m_4(\mathbf{x}) \otimes m_2(\mathbf{x})\Big\}\mathbf{R}' + \Big\{\mathbf{M}_4(\mathbf{x}) \otimes \mathbf{vec}'m_2(\mathbf{x})\Big\}\mathbf{P}' + \mathbf{P}\Big\{\mathbf{M}_4(\mathbf{x})' \otimes \mathbf{vec}m_2(\mathbf{x})\Big\} \\ \Gamma_{(3,3)}(\mathbf{x}) &= \mathbf{R}\Big\{m_3(\mathbf{x}) \otimes m_3(\mathbf{x})'\Big\}\mathbf{P}' + \mathbf{vec}m_3(\mathbf{x})\mathbf{vec}'m_3(\mathbf{x}) \\ \Gamma_{(2,2,2)}(\mathbf{x}) &= \mathbf{Q}\Big\{m_2(\mathbf{x}) \otimes m_2(\mathbf{x}) \otimes m_2(\mathbf{x})\Big\} + \mathbf{P}\Big\{m_2(\mathbf{x}) \otimes \mathbf{vec}m_2(\mathbf{x})\mathbf{vec}'m_2(\mathbf{x})\Big\}\mathbf{P}' \end{aligned} \tag{B.10}$$

where $\mathbf{K}_{pq}$, is the $pq \times pq$ commutation matrix (see Magnus and Neudecker [34]) and the **vec** operator defined by $\mathbf{vecA} = \mathbf{vec}\begin{pmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_p \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_p \end{pmatrix}$.

Appendix C. The star product

Definition 5. Let $\mathbf{A}$ be a matrix of size $p \times q$ and $\mathbf{B}$ a matrix of size $rp \times sq$. The star product of $\mathbf{A}$ and $\mathbf{B}$, noted $\mathbf{A} * \mathbf{B}$, is the matrix of size $r \times s$ defined as

$$\mathbf{A} * \mathbf{B} = \sum_{i_1,i_2=1}^{p,q} a_{i_1,i_2}\mathbf{B}_{i_1,i_2} \tag{C.1}$$

where $\mathbf{B}_{i_1,i_2}$ is the block indexed by (i_1, i_2) in the partition of $\mathbf{B}$ into matrices of size $r \times s$.

Some useful properties of the star product are listed below.

Proposition 8. *(P1) Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two matrices of same dimension. Then*

$$\mathbf{A} * \mathbf{B} = \mathrm{tr}(\mathbf{A}'\mathbf{B}). \tag{C.2}$$

(P2) Let $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *be three matrices of size* $p \times q$, $pr \times sq$ *and* $m \times n$, *respectively. Then*

$$\mathbf{A} * (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} * \mathbf{B}) \otimes \mathbf{C}. \tag{C.3}$$

(P3) Let $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *be three matrices of size* $p \times q$, $q \times r$ *and* $r \times s$, *respectively. Then*

$$\mathbf{ABC} = \mathbf{B} * \mathbf{vecA}\mathbf{vec}'\mathbf{C}' = \mathbf{B}' * (\mathbf{C} \otimes \mathbf{I}_p)\mathbf{K}_{ps}(\mathbf{A} \otimes \mathbf{I}_s). \tag{C.4}$$

(P4) Let $\mathbf{a}$ *and* $\mathbf{b}$ *be two column vectors of dimensions* p *and* q, *respectively, and let* $\mathbf{M}$ *be a matrix of size* $pr \times qs$. *Then*

$$(\mathbf{a}' \otimes \mathbf{I}_r)\,\mathbf{M}\,(\mathbf{b} \otimes \mathbf{I}_s) = \mathbf{ab}' * \mathbf{M}. \tag{C.5}$$

Proof. Properties $(P1)$, $(P2)$ et $(P3)$ are given without proof in Ref. [32]; a proof is found in [25, pp. 34-37]. Property $(P4)$ is proved hereafter because it will be used subsequently in the proof of theorem 4 (see Appendix E). Matrix $\mathbf{M}$, of size $pr \times qs$, is partitioned into blocks of size $r \times s$ as follows

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} & \cdots & \mathbf{M}_{1,q} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} & \cdots & \mathbf{M}_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{p,1} & \mathbf{M}_{p,2} & \cdots & \mathbf{M}_{p,q} \end{pmatrix}. \tag{C.6}$$

If $\mathbf{a} = (a_1, ..., a_p)'$ and $\mathbf{b} = (b_1, ..., b_q)'$, one has

$$\begin{aligned}
(\mathbf{a}' \otimes \mathbf{I}_r)\,\mathbf{M}\,(\mathbf{b} \otimes \mathbf{I}_s) &= \begin{pmatrix} a_1\mathbf{I}_r & a_2\mathbf{I}_r & \cdots & a_p\mathbf{I}_r \end{pmatrix} \begin{pmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} & \cdots & \mathbf{M}_{1,q} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} & \cdots & \mathbf{M}_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{p,1} & \mathbf{M}_{p,2} & \cdots & \mathbf{M}_{p,q} \end{pmatrix} \begin{pmatrix} b_1\mathbf{I}_s \\ b_2\mathbf{I}_s \\ \vdots \\ b_q\mathbf{I}_s \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^p a_i\mathbf{M}_{i,1} & \sum_{i=1}^p a_i\mathbf{M}_{i,2} & \cdots & \sum_{i=1}^p a_i\mathbf{M}_{i,p} \end{pmatrix} \begin{pmatrix} b_1\mathbf{I}_s \\ b_2\mathbf{I}_s \\ \vdots \\ b_q\mathbf{I}_s \end{pmatrix} \\
&= b_1\sum_{i=1}^p a_i\mathbf{M}_{i,1} + b_2\sum_{i=1}^p a_i\mathbf{M}_{i,2} + \cdots + b_q\sum_{i=1}^p a_i\mathbf{M}_{i,p} = \sum_{i=1}^p\sum_{j=1}^q a_ib_j\mathbf{M}_{i,j} \\
&= \mathbf{ab}' * \mathbf{M}.
\end{aligned} \tag{C.7}$$

□

Appendix D. Proof of lemma 3

For the sake of simplicity, the proof is given for $k \in \{2,3\}$ only, the general proof thus resulting by mathematical induction.

Case $k = 2$. One has

$$\mathcal{A}_2 = \left(\begin{array}{cccc} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 & \dots & \mathbf{A}_p \otimes \mathbf{A}_p \end{array} \right) = \left(\begin{array}{c} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_n \end{array} \right) \tag{D.1}$$

where, for $i \in [\![1, n]\!]$, matrix $\mathbf{B}_i$ of size $n \times \sum_{j=1}^{p} n_j^2$ has expression

$$\mathbf{B}_i = \left(\begin{array}{cccc} \ell_i^1 \otimes \mathbf{A}_1 & \ell_i^2 \otimes \mathbf{A}_2 & \cdots & \ell_i^p \otimes \mathbf{A}_p \end{array} \right), \tag{D.2}$$

with ℓ_i^j, of size $1 \times n_j$, is the i-th raw of matrix $\mathbf{A}_j$ (voir (8)). Making use of the properties of the Kronecker product, it comes

$$\ell_i^j \otimes \mathbf{A}_j = (1\ell_i^j) \otimes (\mathbf{A}_j \mathbf{I}_{n_j}) = \mathbf{A}_j \Big(\ell_i^j \otimes \mathbf{I}_{n_j} \Big). \tag{D.3}$$

Inserting (D.3) into (D.2), one gets

$$\begin{aligned} \mathbf{B}_i &= \left(\begin{array}{cccc} \mathbf{A}_1 \big(\ell_i^1 \otimes \mathbf{I}_{n_1} \big) & \mathbf{A}_2 \big(\ell_i^2 \otimes \mathbf{I}_{n_2} \big) & \cdots & \mathbf{A}_p \big(\ell_i^p \otimes \mathbf{I}_{n_p} \big) \end{array} \right) \\ &= \underbrace{\left(\begin{array}{cccc} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_p \end{array} \right)}_{=\mathbf{A}} \underbrace{\left(\begin{array}{cccc} \ell_i^1 \otimes \mathbf{I}_{n_1} & \mathbf{0}_{n_1 \times n_2^2} & \dots & \mathbf{0}_{n_1 \times n_p^2} \\ \mathbf{0}_{n_2 \times n_1^2} & \ell_i^2 \otimes \mathbf{I}_{n_2} & \dots & \mathbf{0}_{n_2 \times n_p^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_p \times n_1^2} & \mathbf{0}_{n_p \times n_2^2} & \dots & \ell_i^p \otimes \mathbf{I}_{n_p} \end{array} \right)}_{=\mathbf{D}_i} \\ &= \mathbf{A} \, \mathrm{bdiag} \Big(\ell_i^1 \otimes \mathbf{I}_{n_1}, \ell_i^2 \otimes \mathbf{I}_{n_2}, \dots, \ell_i^p \otimes \mathbf{I}_{n_p} \Big). \end{aligned} \tag{D.4}$$

Case $k = 3$.

$$\begin{aligned} \mathcal{A}_3 &= \left(\begin{array}{cccc} \mathbf{A}_1 \otimes \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \otimes \mathbf{A}_2 & \cdots & \mathbf{A}_p \otimes \mathbf{A}_p \otimes \mathbf{A}_p \end{array} \right) \\ &= \left(\begin{array}{c} \mathbf{B}_{11} \\ \mathbf{B}_{12} \\ \vdots \\ \mathbf{B}_{1n} \\ \hline \mathbf{B}_{21} \\ \mathbf{B}_{22} \\ \vdots \\ \mathbf{B}_{2n} \\ \hline \vdots \\ \hline \mathbf{B}_{n1} \\ \mathbf{B}_{n2} \\ \vdots \\ \mathbf{B}_{nn} \end{array} \right) = \left(\begin{array}{c} \mathcal{B}_1 \\ \hline \mathcal{B}_2 \\ \hline \vdots \\ \hline \mathcal{B}_n \end{array} \right) \end{aligned} \tag{D.5}$$

where $\mathbf{B}_{ij}$ has size $n \times \sum_{k=1}^{p} n_k^3$ and matrix $\mathcal{B}_i$ size $n^2 \times \sum_{k=1}^{p} n_k^3$. The expression of matrix $\mathcal{B}_i$ is

$$\begin{aligned}
\mathcal{B}_i = \begin{pmatrix} \mathbf{B}_{i1} \\ \mathbf{B}_{i2} \\ \vdots \\ \mathbf{B}_{in} \end{pmatrix} &= \begin{pmatrix} \ell_i^1 \otimes \mathbf{A}_1 \otimes \mathbf{A}_1 & \ell_i^2 \otimes \mathbf{A}_2 \otimes \mathbf{A}_2 & \cdots & \ell_i^p \otimes \mathbf{A}_p \otimes \mathbf{A}_p \end{pmatrix} \\
&= \begin{pmatrix} (\mathbf{A}_1 \otimes \mathbf{A}_1)(\ell_i^1 \otimes \mathbf{I}_{n_1^2}) & \cdots & (\mathbf{A}_p \otimes \mathbf{A}_p)(\ell_i^p \otimes \mathbf{I}_{n_p^2}) \end{pmatrix} \\
&= \mathcal{A}_2 \operatorname{bdiag}\left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2}\right). \qquad \text{(D.6)}
\end{aligned}$$

Replacing $\mathcal{A}_2$ by its expression given in case $k = 2$,

$$\mathcal{B}_i = \begin{pmatrix} \mathbf{A}\mathbf{D}_1 \\ \mathbf{A}\mathbf{D}_2 \\ \vdots \\ \mathbf{A}\mathbf{D}_n \end{pmatrix} \operatorname{bdiag}\left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2}\right), \qquad \text{(D.7)}$$

which implies

$$\begin{aligned}
\mathbf{B}_{ij} &= \mathbf{A}\mathbf{D}_j \operatorname{bdiag}\left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \ell_i^2 \otimes \mathbf{I}_{n_2^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2}\right) \\
&= \mathbf{A} \operatorname{bdiag}\left(\ell_j^1 \otimes \mathbf{I}_{n_1}, \cdots, \ell_j^p \otimes \mathbf{I}_{n_p}\right) \operatorname{bdiag}\left(\ell_i^1 \otimes \mathbf{I}_{n_1^2}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p^2}\right) \\
&= \mathbf{A} \operatorname{bdiag}\left(1 \otimes \ell_j^1 \otimes \mathbf{I}_{n_1}, \cdots, 1 \otimes \ell_j^p \otimes \mathbf{I}_{n_p}\right) \operatorname{bdiag}\left(\ell_i^1 \otimes \mathbf{I}_{n_1} \otimes \mathbf{I}_{n_1}, \cdots, \ell_i^p \otimes \mathbf{I}_{n_p} \otimes \mathbf{I}_{n_p}\right) \\
&= \underbrace{\mathbf{A} \operatorname{bdiag}\left(\ell_i^1 \otimes \ell_j^1 \otimes \mathbf{I}_{n_1}, \ell_i^2 \otimes \ell_j^2 \otimes \mathbf{I}_{n_2}, \cdots, \ell_i^p \otimes \ell_j^p \otimes \mathbf{I}_{n_p}\right)}_{=\mathbf{D}_{ij}}. \qquad \text{(D.8)}
\end{aligned}$$

Appendix E. Proof of theorem 4

Again the proof is given for $r = 2k$ only, the case $r = 2k + 1$ following similar lines.

Let replace $\mathcal{A}_k$ in $c_{2k}(\mathbf{x}) = \mathcal{A}_k \mathcal{C}_{2k}(\mathbf{s}) \mathcal{A}_k'$ by its expression given in lemma 3. It is then seen that the block of size $n \times n$ in matrix $c_{2k}(\mathbf{x})$ indexed by $(i_1, j_1, \cdots, i_{k-1}, j_{k-1})$ has expression

$$\left[c_{2k}(\mathbf{x})\right]_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} = \mathbf{A}\mathbf{D}_{i_1 \cdots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \cdots j_{k-1}} \mathbf{A}'. \qquad \text{(E.1)}$$

Using definition (21) of the contracted cumulant matrix of order $2k$, it comes

$$\begin{aligned}
\mathcal{Q}^{\mathbf{x}}_{(2k)}[\mathbf{M}] &= \sum_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}=1}^{n} m_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} \left[c_{2k}(\mathbf{x})\right]_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} \\
&= \sum_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}=1}^{n} m_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} \mathbf{A}\mathbf{D}_{i_1 \cdots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \cdots j_{k-1}} \mathbf{A}' \\
&= \mathbf{A} \underbrace{\left(\sum_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}=1}^{n} m_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} \mathbf{D}_{i_1 \cdots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \cdots j_{k-1}} \right)}_{=\mathbf{\Delta}} \mathbf{A}'. \qquad \text{(E.2)}
\end{aligned}$$

It is noted that $\mathbf{\Delta}$ is block-diagonal since $\mathbf{D}_{i_1 \cdots i_{k-1}}$, $\mathcal{C}_{2k}(\mathbf{s})$ and $\mathbf{D}'_{j_1 \cdots j_{k-1}}$ are block-diagonal. Thus

$$\mathbf{\Delta} = \sum_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}=1}^{n} m_{i_1, j_1, \cdots, i_{k-1}, j_{k-1}} \mathbf{D}_{i_1 \cdots i_{k-1}} \mathcal{C}_{2k}(\mathbf{s}) \mathbf{D}'_{j_1 \cdots j_{k-1}} = \operatorname{bdiag}\left(\Delta_{\mathbf{s}_1}^{(2k)}, \cdots, \Delta_{\mathbf{s}_p}^{(2k)}\right), \qquad \text{(E.3)}$$

where, upon invoking (25) and (27),

$$\Delta_{\mathbf{s}_i}^{(2k)} = \sum_{i_1,j_1,\cdots,i_{k-1},j_{k-1}=1}^{n} m_{i_1,j_1,\cdots,i_{k-1},j_{k-1}} \left(\ell_{i_1}^i \otimes \cdots \otimes \ell_{i_{k-1}}^i \otimes \mathbf{I}_{n_i}\right) c_{2k}(\mathbf{s}_i) \left(\ell_{j_1}^i \otimes \cdots \otimes \ell_{j_{k-1}}^i \otimes \mathbf{I}_{n_i}\right)'. \tag{E.4}$$

Making use of property (C.5) of the star product,

$$\begin{aligned}
\Delta_{\mathbf{s}_i}^{(2k)} &= \left(\sum_{i_1,j_1,\cdots,i_{k-1},j_{k-1}=1}^{n} m_{i_1,j_1,\cdots,i_{k-1},j_{k-1}} \left(\ell_{i_1}^i \otimes \cdots \otimes \ell_{i_{k-1}}^i\right)' \left(\ell_{j_1}^i \otimes \cdots \otimes \ell_{j_{k-1}}^i\right)\right) * c_{2k}(\mathbf{s}_i) \\
&= \left(\sum_{i_1,j_1,\cdots,i_{k-1},j_{k-1}=1}^{n} m_{i_1,j_1,\cdots,i_{k-1},j_{k-1}} \left(\mathbf{A}_i'\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{A}_i'\mathbf{e}_{i_{k-1}}\right) \left(\mathbf{e}'_{j_1}\mathbf{A}_i \otimes \cdots \otimes \mathbf{e}'_{j_1}\mathbf{A}_i\right)\right) * c_{2k}(\mathbf{s}_i) \\
&= \left(\sum_{i_1,j_1,\cdots,i_{k-1},j_{k-1}=1}^{n} m_{i_1,j_1,\cdots,i_{k-1},j_{k-1}} (\otimes^{k-1}\mathbf{A}_i') \left(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_{k-1}}\right) \left(\mathbf{e}'_{j_1} \otimes \cdots \otimes \mathbf{e}'_{j_1}\right) (\otimes^{k-1}\mathbf{A}_i)\right) * c_{2k}(\mathbf{s}_i) \\
&= \left((\otimes^{k-1}\mathbf{A}_i') \left(\sum_{i_1,j_1,\cdots,i_{k-1},j_{k-1}=1}^{n} m_{i_1,j_1,\cdots,i_{k-1},j_{k-1}} \mathbf{E}_{i_1,j_1,\cdots,i_{k-1},j_{k-1}}\right) (\otimes^{k-1}\mathbf{A}_i)\right) * c_{2k}(\mathbf{s}_i) \\
&= \left((\otimes^{k-1}\mathbf{A}_i')\mathbf{M}^{(2k-2)}(\otimes^{k-1}\mathbf{A}_i)\right) * c_{2k}(\mathbf{s}_i) \\
&= \mathcal{Q}_{(2k)}^{\mathbf{s}_i}\Big[\left(\otimes^{k-1}\mathbf{A}_i'\right)\mathbf{M}^{(2k-2)}\left(\otimes^{k-1}\mathbf{A}_i\right)\Big].
\end{aligned} \tag{E.5}$$

Appendix F.

Proof of lemma 6. Let us give a proof by contraposition. Let assume $\mathbf{y}$ is reducible, then according to definition 1, there exists an invertible matrix $\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \in \mathbb{R}^{d\times d}$ and independent vectors $\mathbf{y}_1 \in \mathbb{R}^{d_1}$ and $\mathbf{y}_2 \in \mathbb{R}^{d_2}$ $(d = d_1 + d_2)$ such that

$$\mathbf{y} = \mathbf{A}\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}. \tag{F.1}$$

This implies from (14) that

$$c_4(\mathbf{y}) = \underbrace{\begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \end{pmatrix}}_{=\mathcal{A}_2} \operatorname{bdiag}\Big(c_4(\mathbf{y}_1), c_4(\mathbf{y}_2)\Big) \begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \end{pmatrix}'. \tag{F.2}$$

Therefore,

$$\begin{aligned}
\mathbf{r}\left(c_4(\mathbf{y})\right) &= \mathbf{r}\left(\mathcal{A}_2 \operatorname{bdiag}\Big(c_4(\mathbf{y}_1), c_4(\mathbf{y}_2)\Big)\mathcal{A}'_2\right) \\
&= \mathbf{r}\left(\operatorname{bdiag}\Big(c_4(\mathbf{x}_1), c_4(\mathbf{x}_2)\Big)\right) \\
&= \mathbf{r}\left(c_4(\mathbf{x}_1)\right) + \mathbf{r}\left(c_4(\mathbf{x}_2)\right) \\
&\leq \frac{d_1(d_1+1)}{2} + \frac{d_2(d_2+1)}{2} \\
&< \frac{d_1(d_1+1)}{2} + \frac{d_2(d_2+1)}{2} + d_1d_2 = \frac{d(d+1)}{2},
\end{aligned} \tag{F.3}$$

where equality $\mathbf{r}\begin{pmatrix} \mathbf{A}_1 \otimes \mathbf{A}_1 & \mathbf{A}_2 \otimes \mathbf{A}_2 \end{pmatrix} = d_1^2 + d_2^2$ holds true since $\mathbf{r}(\mathbf{A}) = \mathbf{r}(\mathbf{A}_1) + \mathbf{r}(\mathbf{A}_2)$ (see [46]). Thus, the majoration of rank $\left(\mathbf{r}\left(c_4(\mathbf{y}_i)\right) \leqslant \frac{d_i(d_i+1)}{2}, \forall i \in \{1,2\}\right)$. This means that $\mathbf{r}\left(c_4(\mathbf{y})\right) < \frac{d(d+1)}{2}$ when $\mathbf{y}$ is reducible. Therefore, if $\mathbf{r}(c_4(\mathbf{y})) = \frac{d(d+1)}{2}$, then $\mathbf{y}$ is irreducible. □

Lemma 9. *Let* $\mathbf{y}$ *be a random vector verifying (54) and* $\mathbf{U}\Lambda\mathbf{U}'$ *be the eigenvalue decomposition of* $c_4(\mathbf{y})$*, where* $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_{\frac{d(d+1)}{2}})$*. Then,*

$$\mathbf{U}\mathbf{U}' = \frac{1}{2}(\mathbf{K}_{dd} + \mathbf{I}_{d^2}), \tag{F.4}$$

with $\mathbf{K}_{dd}$ *the commutation matrix (see [34]).*

Proof of lemma 9. Since $c_4(\mathbf{y})$ is a symmetric matrix with rank $\frac{d(d+1)}{2}$, it accepts the eigenvalue decomposition $\mathbf{U}\Lambda\mathbf{U}'$, where $\mathbf{U} \in \mathbb{R}^{d^2 \times \frac{d(d+1)}{2}}$ is a semi-orthogonal matrix ($\mathbf{U}'\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$) and $\Lambda \in \mathbb{R}^{\frac{d(d+1)}{2} \times \frac{d(d+1)}{2}}$ is a diagonal matrix made of the $\frac{d(d+1)}{2}$ non-zero eigenvalues of $c_4(\mathbf{y})$. let $\tilde{\mathbf{K}} = \frac{1}{2}(\mathbf{K_{dd}} + \mathbf{I}_{d^2})$. Since $\tilde{\mathbf{K}}c_4(\mathbf{y})\tilde{\mathbf{K}} = c_4(\mathbf{y})$, one has $\tilde{\mathbf{K}}\mathbf{U}\Lambda\mathbf{U}'\tilde{\mathbf{K}} = \mathbf{U}\Lambda\mathbf{U}'$. Thus, it comes $\mathbf{U}'\tilde{\mathbf{K}}\tilde{\mathbf{K}}\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$. Since $\tilde{\mathbf{K}}$ is idempotent ($\tilde{\mathbf{K}}\tilde{\mathbf{K}} = \tilde{\mathbf{K}}$), it comes $\mathbf{U}'\tilde{\mathbf{K}}\mathbf{U} = \mathbf{I}_{\frac{d(d+1)}{2}}$, and $\mathbf{U}\mathbf{U}'\tilde{\mathbf{K}}\mathbf{U}\mathbf{U}' = \mathbf{U}\mathbf{U}'$. Therefore $\mathbf{U}\mathbf{U}'$ is t he pseud o-inverse of $\tilde{\mathbf{K}}$ which is equal to $\tilde{\mathbf{K}}$. □

Proof of proposition 6. Assume that all multidimensional sources fulfill condition (54), i.e. $\mathbf{r}(c_4(\mathbf{s}_i)) = \frac{n_i(n_i+1)}{2}$. It then follows that none of them is Gaussian since $\mathbf{r}(c_4(\mathbf{s}_i)) \neq 0$. In addition, it follows from Eq. (24) that

$$c_4(\mathbf{x}) = \begin{pmatrix} \mathbf{A}\mathbf{D}_1 \\ \vdots \\ \mathbf{A}\mathbf{D}_n \end{pmatrix} \begin{pmatrix} c_4(\mathbf{s}_1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & c_4(\mathbf{s}_p) \end{pmatrix} \begin{pmatrix} \mathbf{D}_1'\mathbf{A}' & \cdots & \mathbf{D}_n'\mathbf{A}' \end{pmatrix}. \tag{F.5}$$

Let us now introduce $\mathbf{U}_i\Lambda_i\mathbf{U}_i'$, $i = 1, \cdots, p$, the eigen-elements of $c_4(\mathbf{s}_i)$ of lemma 9, with $\mathbf{U}_i\mathbf{U}_i' = \frac{1}{2}(\mathbf{K}_{n_i n_i} + \mathbf{I}_{n_i^2})$. Thus

$$\mathbf{U} = \mathrm{bdiag}\Big(\mathbf{U}_1, \cdots, \mathbf{U}_p\Big) \tag{F.6}$$

and

$$\mathbf{\Lambda} = \mathrm{bdiag}\Big(\Lambda_1, \cdots, \Lambda_p\Big), \tag{F.7}$$

so that Eq. (F.5) becomes

$$c_4(\mathbf{x}) = \begin{pmatrix} \mathbf{A}\mathbf{D}_1\mathbf{U} \\ \vdots \\ \mathbf{A}\mathbf{D}_n\mathbf{U} \end{pmatrix} \begin{pmatrix} \Lambda_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Lambda_p \end{pmatrix} \begin{pmatrix} \mathbf{U}'\mathbf{D}_1'\mathbf{A}' & \cdots & \mathbf{U}'\mathbf{D}_n'\mathbf{A}' \end{pmatrix}. \tag{F.8}$$

One recognizes the eigenvalue decomposition of $c_4(\mathbf{x}) = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ with

$$\mathbf{V} = \begin{pmatrix} \mathbf{A}\mathbf{D}_1\mathbf{U} \\ \mathbf{A}\mathbf{D}_2\mathbf{U} \\ \vdots \\ \mathbf{A}\mathbf{D}_n\mathbf{U} \end{pmatrix}. \tag{F.9}$$

The next step is to evaluate the product,

$$\mathbf{V}'\mathbf{V} = \sum_{k=1}^{n} \mathbf{U}'\mathbf{D}_k' \underbrace{\mathbf{A}'\mathbf{A}}_{=\mathbf{I}_n} \mathbf{D}_k\mathbf{U} = \mathbf{U}'\Big(\sum_{k=1}^{n} \mathbf{D}_k'\mathbf{D}_k\Big)\mathbf{U}. \tag{F.10}$$

Substituting Eq. (27) for $\mathbf{D}_k$, one has

$$\begin{aligned}
\mathbf{V}'\mathbf{V} &= \mathbf{U}'\left(\sum_{k=1}^{n}\mathrm{bdiag}\left(\ell_1^k\otimes\mathbf{I}_{n_1},\ldots,\ell_p^k\otimes\mathbf{I}_{n_p}\right)'\mathrm{bdiag}\left(\ell_1^k\otimes\mathbf{I}_{n_1},\ldots,\ell_p^k\otimes\mathbf{I}_{n_p}\right]\right)\mathbf{U}\\
&= \mathbf{U}'\left(\sum_{k=1}^{n}\mathrm{bdiag}\left[\ell_1^{k'}\ell_1^k,\ldots,\ell_p^{k'}\ell_p^k\right)\right)\mathbf{U}\\
&= \mathbf{U}'\,\mathrm{bdiag}\left(\sum_{k=1}^{n}\ell_1^{k'}\ell_1^k,\ldots,\sum_{k=1}^{n}\ell_p^{k'}\ell_p^k\right)\mathbf{U}\\
&= \mathbf{U}'\,\mathrm{bdiag}\left[\mathbf{A}_1'\mathbf{A}_1,\ldots,\mathbf{A}_p'\mathbf{A}_p\right)\mathbf{U}\\
&= \mathbf{U}'\,\mathrm{bdiag}\left(\mathbf{I}_{n_1},\ldots,\mathbf{I}_{n_p}\right)\mathbf{U}\\
&= \mathbf{U}'\mathbf{U}\\
&= \mathrm{bdiag}\left(\mathbf{U}_1'\mathbf{U}_1,\ldots,\mathbf{U}_p'\mathbf{U}_p\right)\\
&= \mathrm{bdiag}\left(\mathbf{I}_{\frac{n_1(n_1+1)}{2}},\ldots,\mathbf{I}_{\frac{n_p(n_p+1)}{2}}\right)\\
&= \mathbf{I}_{\sum_{i=1}^{p}\frac{n_i(n_i+1)}{2}}.
\end{aligned}\tag{F.11}$$

Therefore, matrix $\mathbf{V}'\mathbf{V}$ is made of n^2 blocks of size $n\times n$,

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A}\mathbf{D}_i\mathbf{U}\mathbf{U}'\mathbf{D}_j'\mathbf{A}'.\tag{F.12}$$

Besides, from lemma 9,

$$\mathbf{U}\mathbf{U}' = \mathrm{bdiag}\left(\mathbf{U}_1\mathbf{U}_1',\cdots,\mathbf{U}_p\mathbf{U}_p'\right) = \mathrm{bdiag}\left(\frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,\frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right),\tag{F.13}$$

so that,

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A}\,\mathrm{bdiag}\left((\ell_1^i\otimes\mathbf{I}_{n_1})\frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2})(\ell_1^j\otimes\mathbf{I}_{n_1})',\cdots,(\ell_p^i\otimes\mathbf{I}_{n_p})\frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})(\ell_p^j\otimes\mathbf{I}_{n_p})'\right)\mathbf{A}',\tag{F.14}$$

and, finally upon using property (P4) of the star product (see (Appendix C)),

$$\left[\mathbf{V}'\mathbf{V}\right]_{i,j} = \mathbf{A}\,\mathrm{bdiag}\left(\ell_1^{i'}\ell_1^j * \frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,\ell_p^{i'}\ell_p^j * \frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right)\mathbf{A}'.\tag{F.15}$$

From (F.15), one has

$$\begin{aligned}
\mathcal{R} &= \sum_{k=1}^{n}\left[\mathbf{V}'\mathbf{V}\right]_{k,k} = \sum_{k=1}^{n}\mathbf{A}\,\mathrm{bdiag}\left(\ell_1^{k'}\ell_1^k * \frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,\ell_p^{k'}\ell_p^k * \frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right)\mathbf{A}'\\
&= \mathbf{A}\,\mathrm{bdiag}\left(\underbrace{\sum_{k=1}^{n}\ell_1^{k'}\ell_1^k}_{=\mathbf{A'}_1\mathbf{A}_1} * \frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,\underbrace{\sum_{k=1}^{n}\ell_p^{k'}\ell_p^k}_{=\mathbf{A'}_p\mathbf{A}_p} * \frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right)\mathbf{A}'\\
&= \mathbf{A}\,\mathrm{bdiag}\left((\mathbf{A'}_1\mathbf{A}_1) * \frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,(\mathbf{A'}_p\mathbf{A}_p) * \frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right)\mathbf{A}'\\
&= \mathbf{A}\,\mathrm{bdiag}\left(\mathbf{I}_{n_1} * \frac{1}{2}(\mathbf{K}_{n_1n_1}+\mathbf{I}_{n_1^2}),\cdots,\mathbf{I}_{n_p} * \frac{1}{2}(\mathbf{K}_{n_pn_p}+\mathbf{I}_{n_p^2})\right)\mathbf{A}'\\
&= \mathbf{A}\,\mathrm{bdiag}\left(\frac{n_1+1}{2}\mathbf{I}_{n_1},\cdots,\frac{n_p+1}{2}\mathbf{I}_{n_p}\right)\mathbf{A}'.
\end{aligned}\tag{F.16}$$

□

[1] Albera, L., Comon, P., August 2002. Asymptotic performance of contrast-based blind source separation algorithms. In: Second IEEE Sensor Array and Multichannel Signal Processing Workshop. Rosslyn, VA.

[2] Albera, L., Ferreol, A., Chevalier, P., Comon, P., 2005. Icar: a tool for blind source separation using fourth-order statistics only. Signal Processing, IEEE Transactions on 53 (10), 3633–3643.

[3] Albera, L., Ferreol, A., Comon, P., Chevalier, P., April 2003. Sixth order blind identification of underdetermined mixtures (birth) of sources. In: 4th International Symposium on Independant Componenet Analysis and Blind Signal Separation (ICA2003), Nara, Japan. pp. 909–914.

[4] Albera, L., Ferreol, A., Comon, P., Chevalier, P., November 2004. Blind identification of over-complete mixtures of sources (biome). Linaear Algebra and its Applications 391, 1–30.

[5] Amari, S., Cichocki, A., Yang, H. H., 1996. A new learning algorithm for blind signal separation. In: Advances in Neural Information Processing Systems. MIT Press, pp. 757–763.

[6] Cardoso, J., 1998. Multidimensional independent component analysis. In: Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on. Vol. 4. pp. 1941–1944 vol.4.

[7] Cardoso, J.-F., 1999. High-order contrasts for independent component analysis. Neural computation 11 (1), 157–192.

[8] Cardoso, J.-F., Soulomiac, A., 1993. Blind beamforming for non gaussian signals. IEE- Proceedings-F 140 (6), 362–370.

[9] Casey, M. A., Westner, A., 2000. Separation of mixed audio sources by independent subspace analysis. In: Proceedings of the International Computer Music Conference. pp. 154–161.

[10] Chawla, M., 2009. Detection of indeterminacies in corrected ecg signals using parameterized multidimensional independent component analysis. Computational and Mathematical Methods in Medicine 10 (2), 85–115.

[11] Comon, P., 1994. Independent component analysis, a new concept? Signal Processing 36 (3), 287–314.

[12] Fan, Z., Zhou, J., Wu, Y., 2004. Motion segmentation based on independent subspace analysis.

[13] Févotte, C., Theis, F., 2007. Orthonormal approximate joint block-diagonalization. Technical report GET/Télécom Paris.

[14] Ghennioui, H., El Mostafa, F., Thirion-Moreau, N., Adib, A., Moreau, E., 2007. A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources. Signal Processing Letters, IEEE 14 (11), 860–863.

[15] Ghennioui, H., Thirion-Moreau, N., Moreau, E., Aboutajdine, D., 2010. Gradient-based joint block diagonalization algorithms: Application to blind separation of fir convolutive mixtures. Signal Processing 90 (6), 1836–1849.

[16] Gutch, H., Theis, F., 2007. Independent subspace analysis is unique, given irreducibility. In Proc. ICA 2007 4666 of LNCS, 49–56.

[17] Gutch, H. W., Krumsiek, J., Theis, F. J., 2011. An isa algorithm with unknown group sizes identifies meaningful clusters in metabolomics data. In: Proc. EUSIPCO. Vol. 2011. pp. 1733–1737.

[18] Gutch, H. W., Theis, F. J., Nov. 2012. Uniqueness of linear factorizations into independent subspaces. J. Multivar. Anal. 112, 48–62.

[19] H. Gutch, T. M., Theis, F., Springer, St. Malo, France 2010. Second order subspace analysis and simple decompositions. In Proc. ICA 2010 6365 of LNCS, 370–377.

[20] Hyvärinen, A., Hoyer, P., Jul. 2000. Emergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces. Neural Comput. 12 (7), 1705–1720.
URL `http://dx.doi.org/10.1162/089976600300015312`

[21] Hyvärinen, A., Hoyer, P. O., Inki, M. O., Jul. 2001. Topographic independent component analysis. Neural Comput. 13 (7), 1527–1558.
URL `http://dx.doi.org/10.1162/089976601750264992`

[22] Hyvärinen, A., Köster., U., 2006. FastISA: A fast fixed-point algorithm for independent subspace analysis, 371–376.

[23] Kim, H., Choi, S., Bang, S.-Y., 2003. Membership scoring via independent feature subspace analysis for grouping co-expressed genes. In: Neural Networks, 2003. Proceedings of the International Joint Conference on. Vol. 3. pp. 1690–1695 vol.3.

[24] Kim, J. K., Choi, S., 2006. Tree-dependent components of gene expression data for clustering. In: Proceedings of the 16th international conference on Artificial Neural Networks - Volume Part II. ICANN'06. Springer-Verlag, Berlin, Heidelberg, pp. 837–846.

[25] Kollo, T., 1991. Matrix Derivative in Multivariate Statistics. Tartu University Press. Tartu (in Russian).

[26] Kollo, T., 2008. Multivariate skewness and kurtosis measures with an application in ica. Journal of Multivariate Analysis 99, 2328–2338.

[27] Lahat, D., Cardoso, J., Messer, H., 2009. Optimal performance of second-order multidimensional ica. Intelligent Data Analysis-IDA (50–57).

[28] Le, Q. V., Zou, W. Y., Yeung, S. Y., Ng, A. Y., 2011. Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis. In: CVPR'11. pp. 3361–3368.

[29] Li, S., Lv, X. G., Zhang, H., 2001. View-based clustering of object appearances based on independent subspace analysis. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. pp. 295–300 vol.2.

[30] Ma, S., Li, X.-L., Correa, N., Adali, T., Calhoun, V., 2010. Independent subspace analysis with prior information for fmri data. In: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. pp. 1922–1925.

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

[32] MacRae, E., 1974. Matrix derivatives with an application to an adaptive linear decision problem. Ann. Statist (337–346).

[33] Maehara, T., Murota, K., 2011. Algorithm for error-controlled simultaneous block-diagonalization of matrices. SIAM Journal on Matrix Analysis and Applications 32, 605–620.

[34] Magnus, J. R., Neudecker, H., March 1979. The commutation matrix: Some properties and applications. Annals of Statistics 7 (2), 381–394.
[35] Mendel, J. M., March 1991. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. Proceedings of the IEEE In Proceedings of the IEEE 79 (3), 278–305.
[36] Moreau, E., mar 2001. A generalization of joint-diagonalization criteria for source separation. Signal Processing, IEEE Transactions on 49 (3), 530 –541.
[37] Nion, D., 2011. A tensor framework for nonunitary joint block diagonalization. Signal Processing, IEEE Transactions on 59 (10), 4585–4594.
[38] Ould-Baba, H., Robin, V., Antoni, J., 2013. Concise formulae for the cumulant matrices of a random vector. submitted.
[39] Póczos, B., Lörincz, A., 2005. Independent subspace analysis using k-nearest neighborhood distances. In: Proceedings of the 15th international conference on Artificial neural networks: formal models and their applications - Volume Part II. ICANN'05. Springer-Verlag, Berlin, Heidelberg, pp. 163–168.
URL http://dl.acm.org/citation.cfm?id=1986079.1986110
[40] Santos, C., Kogler, S.E., J., Del Hernandez, E., 2005. Using independent subspace analysis for selecting filters used in texture processing. In: Image Processing, 2005. ICIP 2005. IEEE International Conference on. Vol. 1. pp. I–465–8.
[41] Sharma, A., Paliwal, K. K., 2006. Subspace independent component analysis using vector kurtosis. Pattern Recognition 39 (11), 2227–2232.
[42] Theis, F., may 2005. Blind signal separation into groups of dependent signals using joint block diagonalization, 5878 – 5881 Vol. 6.
[43] Theis, F., 2007. Towards a general independent subspace analysis. In Proc. NIPS 2006, 1361–1368.
[44] Theis, F. J., May 2004. Uniqueness of complex and multidimensional independent component analysis. Signal Process. 84 (5), 951–956.
URL http://dx.doi.org/10.1016/j.sigpro.2004.01.008
[45] Theis, F. J., 2005. Multidimensional independent component analysis using characteristic functions. In: In Proc. of EUSIPCO.
[46] Tian, Y., 2005. Some rank equalities and inequalities for kronecker products of matrices. Linear and Multilinear Algebra 53 (6), 445–454.
[47] Ziehe, A., Laskov, P., Nolte, G., Müller, K.-R., 2004. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. The Journal of Machine Learning Research 5, 777–800.