



HAL
open science

A New Link Between Joint Blind Source Separation Using Second Order Statistics and the Canonical Polyadic Decomposition

Dana Lahat, Christian Jutten

► **To cite this version:**

Dana Lahat, Christian Jutten. A New Link Between Joint Blind Source Separation Using Second Order Statistics and the Canonical Polyadic Decomposition. LVA/ICA 2018 - 14th International Conference on Latent Variable Analysis and Signal Separation, Jul 2018, Guildford, United Kingdom. pp.171–180. hal-01884793

HAL Id: hal-01884793

<https://hal.science/hal-01884793>

Submitted on 1 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A New Link Between Joint Blind Source Separation Using Second Order Statistics and the Canonical Polyadic Decomposition

Dana Lahat* and Chrisitan Jutten

Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

Abstract. In this paper, we discuss the joint blind source separation (JBSS) of real-valued Gaussian stationary sources with uncorrelated samples from a new perspective. We show that the second-order statistics of the observations can be reformulated as a coupled decomposition of several tensors. The canonical polyadic decomposition (CPD) of each such tensor, if unique, results in the identification of one or two mixing matrices. The proposed new formulation implies that standard algorithms for joint diagonalization and CPD may be used to estimate the mixing matrices, although only in a sub-optimal manner. We discuss the uniqueness and identifiability of this new approach. We demonstrate how the proposed approach can bring new insights on the uniqueness of JBSS in the presence of underdetermined mixtures.

Keywords: Joint blind source separation, independent vector analysis, tensor, canonical polyadic decomposition, uniqueness, identifiability

1 Introduction

In this paper, we present a new type of link between joint blind source separation (JBSS) [1,2] and the canonical polyadic decomposition (CPD) [3], in the special case that each of the sources, in each mixture, is a real-valued Gaussian random process with independent and identically distributed (i.i.d.) samples. To the best of our knowledge, until now, this link has been shown only when the data had some additional type of diversity, e.g., nonstationarity or higher-order statistics (HOS) [4,5,6,7]. Our model assumptions, as well as previous related results, are described in Section 2. The new algebraic formulation is introduced in Section 3. In Section 4 we discuss the uniqueness of the proposed new formulation. In Section 5, we suggest to use this new formulation as an alternative to existing JBSS algorithms on the one hand, and to the closed-form solution via generalized eigenvalue decomposition (GEVD) on the other hand. In Section 6 we demonstrate how the proposed approach leads to new insights and new results on the identifiability of JBSS in underdetermined cases, beyond existing results in the literature.

* This work is supported by the project CHESS, 2012-ERC-AdG-320684. GIPSA-Lab is a partner of the LabEx PERSYVAL-Lab (ANR-11-LABX-0025).

In this paper, scalars, column vectors, matrices, and tensors, are denoted a , \mathbf{a} , \mathbf{A} , and \mathcal{A} , respectively. The r th entry of \mathbf{a} , and the r th column of \mathbf{A} , are denoted a_r and \mathbf{a}_r , respectively. \cdot^\top denotes transpose. $\mathbf{A}^{-[k]}$, $\mathbf{X}^{-[k,l]}$, and $\mathbf{A}^{-[k]\top}$ stand for $(\mathbf{A}^{[k]})^{-1}$, $(\mathbf{X}^{[k,l]})^{-1}$, and $(\mathbf{A}^{-[k]})^\top$, respectively. The outer product is denoted as \circ , where $\mathbf{a} \circ \mathbf{b} = \mathbf{a}\mathbf{b}^\top$, and $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ is a third-order array (tensor) whose (i, j, k) th element is $a_i b_j c_k$. $\text{Diag}\{\mathbf{a}\}$ is a diagonal matrix with the values of \mathbf{a} on its main diagonal. $E\{\cdot\}$ denotes expectation. $k_{\mathbf{A}}$ denotes the *Kruskal rank* of matrix \mathbf{A} , which is equal to the largest integer $k_{\mathbf{A}}$ such that every subset of $k_{\mathbf{A}}$ columns of \mathbf{A} is linearly independent [8].

2 Problem Formulation

In this paper, we consider the JBSS problem in $K \geq 2$ mixtures,

$$\mathbf{x}^{[k]} = \mathbf{A}^{[k]} \mathbf{s}^{[k]} \quad , \quad k = 1, \dots, K \quad (1)$$

where the random vector $\mathbf{x}^{[k]} \in \mathbb{R}^{I^{[k]} \times 1}$ represents the observations at $I^{[k]}$ sensors at the k th mixture. Within each mixture, the R elements of the random vector $\mathbf{s}^{[k]} \triangleq [s_1^{[k]}, \dots, s_R^{[k]}]^\top \in \mathbb{R}^{R \times 1}$ are statistically independent. Each random variable $s_r^{[k]}$ generates a real-valued Gaussian random process with i.i.d. samples. The K mixing matrices $\mathbf{A}^{[k]} \triangleq [\mathbf{a}_1^{[k]} \mid \dots \mid \mathbf{a}_R^{[k]}] \in \mathbb{R}^{I^{[k]} \times R}$ are assumed to be different from each other.

The k th mixture (sometimes referred to as “dataset”) can be written as a sum of contributions from $R \geq 2$ different sources,

$$\mathbf{x}^{[k]} = \sum_{r=1}^R \mathbf{a}_r^{[k]} s_r^{[k]} = \sum_{r=1}^R \mathbf{x}_r^{[k]} . \quad (2)$$

It is clear that $\mathbf{x}_r^{[k]}$ remains invariant if the pair $(\mathbf{a}_r^{[k]}, s_r^{[k]})$ is replaced with $(z^{-1} \mathbf{a}_r^{[k]}, z s_r^{[k]})$, for any $z \neq 0$. Therefore, in the absence of additional information, only $\text{span}(\mathbf{a}_r^{[k]})$ and $\mathbf{x}_r^{[k]}$ may be uniquely identified. Furthermore, the order of the summands in (2) is immaterial. When the model is subject only to these trivial ambiguities, we say that it is essentially unique.

In this paper, we focus on JBSS using second-order statistics (SOS). In this model, the cross-correlation $s_{r\rho}^{[k,l]}$ between any two sources $s_r^{[k]}$ and $s_\rho^{[l]}$ satisfies

$$s_{r\rho}^{[k,l]} \triangleq E\{s_r^{[k]} s_\rho^{[l]\top}\} = \begin{cases} s_{rr}^{[k,l]} & r = \rho \\ 0 & r \neq \rho \end{cases} , \quad \begin{matrix} k, l = 1, \dots, K \\ r, \rho = 1, \dots, R \end{matrix} \quad (3)$$

We assume that each mixture contains R non-zero sources, hence, $s_{rr}^{[k,k]} \neq 0 \forall k$. Furthermore, due to the arbitrary scaling between each source and the column of the mixing matrix associated with it, we can always choose, without loss of generality (w.l.o.g.), $s_{rr}^{[k,k]} = 1 \forall k$. For $k \neq l$, each of $s_{rr}^{[k,l]}$ can be zero or non-zero; $s_{rr}^{[k,l]} \neq 0$ can be interpreted as a statistical link (correlation) between the

r th source in the k th and l th datasets. The cross-correlation $\mathbf{S}^{[k,l]}$ between $\mathbf{s}^{[k]}$ and $\mathbf{s}^{[l]}$ can thus be written as

$$\mathbf{S}^{[k,l]} \triangleq E\{\mathbf{s}^{[k]}\mathbf{s}^{[l]\top}\} = \text{Diag}\{s_{11}^{[k,l]}, \dots, s_{RR}^{[k,l]}\} \in \mathbb{R}^{R \times R} \quad \forall k, l. \quad (4)$$

In this paper, we assume that all the SOS exist and are finite-valued. Given our assumptions, the sufficient statistics for the estimation of $\text{span}(\mathbf{a}_r^{[k]})$ and $\mathbf{x}_r^{[k]}$ are the cross-correlation matrices of the observations:

$$\mathbf{X}^{[k,l]} \triangleq E\{\mathbf{x}^{[k]}\mathbf{x}^{[l]\top}\} = \mathbf{A}^{[k]}\mathbf{S}^{[k,l]}\mathbf{A}^{[l]\top} = \sum_{r=1}^R s_{rr}^{[k,l]} \mathbf{a}_r^{[k]} \mathbf{a}_r^{[l]\top} \quad \forall k, l \quad (5)$$

where the second equality on the right-hand side (RHS) of (5) is due to (1), and the last equality is due to (2) and (4). In the statistical JBSS formulation (1), each dataset has its own set of parameters, and the link (coupling) between datasets is probabilistic, using additional parameters that represent cross-correlations; in the algebraic formulation in (5), the link between two cross-correlation matrices $\mathbf{X}^{[k,l]}$ and $\mathbf{X}^{[k,l']}$ is deterministic, via a shared mixing matrix $\mathbf{A}^{[k]}$. These two types of links are sometimes referred to as “soft” versus “hard”, see [9] and references therein. Equation (5) implies that the statistically-motivated JBSS can be written algebraically, as a coupled decomposition of the ensemble of matrices $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$ [2].

2.1 JBSS via GEVD: a Closed-Form Solution for Two Datasets

Given our assumptions, when $K = 2$ and $\mathbf{A}^{[k]}$ are nonsingular (hence, $I^{[k]} = R$) for $k = 1, 2$, the estimates of $\mathbf{A}^{[k]}$ can always be obtained algebraically, using the GEVD [10, Chapter 12.2, Equation (53)] (see also [11, Sec. 4.3]):

$$\mathbf{X}^{[2,1]}\mathbf{X}^{-[1,1]}\mathbf{X}^{[1,2]}\mathbf{V}^{[1]} = \mathbf{X}^{[2,2]}\mathbf{V}^{[1]}\mathbf{\Lambda} \quad (6a)$$

$$\mathbf{X}^{[1,2]}\mathbf{X}^{-[2,2]}\mathbf{X}^{[2,1]}\mathbf{V}^{[2]} = \mathbf{X}^{[1,1]}\mathbf{V}^{[2]}\mathbf{\Lambda} \quad (6b)$$

where $\mathbf{\Lambda} = \text{Diag}\{\lambda_1, \dots, \lambda_R\}$, and the r th column of $\mathbf{V}^{[k]} \in \mathbb{R}^{R \times R}$ is the generalized eigenvector associated with the generalized eigenvalue λ_r . Therefore, when $K = 2$, JBSS is identifiable if and only if (iff) this GEVD is unique, that is, if its generalized eigenvalues are distinct. The resulting estimates of $\mathbf{A}^{[k]}$, $\mathbf{A}^{[1]\text{GEVD}} \triangleq \mathbf{V}^{[1]\top}$ and $\mathbf{A}^{[2]\text{GEVD}} \triangleq \mathbf{V}^{[2]\top}$, exactly diagonalize $\{\mathbf{X}^{[k,l]}\}_{k,l=1}^K$:

$$(\mathbf{A}^{[k]\text{GEVD}})^{-1}\mathbf{X}^{[k,l]}(\mathbf{A}^{[l]\text{GEVD}})^{-\top} \in \text{Diag} \quad \text{for any } k, l = 1, 2 \quad (7)$$

This solution always exists, and this exact diagonalization can always be achieved, regardless of any perturbation of the observations with respect to (w.r.t.) the ground truth statistics, as long as $\mathbf{X}^{[k,l]}\mathbf{X}^{-[l,l]}\mathbf{X}^{[l,k]}$ and $\mathbf{X}^{[k,k]}$ are Hermitian (symmetric, in the real case) and $\mathbf{X}^{[k,k]}$ is positive-definite $\forall k, l$.

In the presence of $K \geq 3$ datasets, the GEVD may be used, for example, to initialize JBSS algorithms that optimize (5). Our new approach, which will be explained in Section 3, is based on extending the formulation in (6) to $K \geq 2$ mixtures.

and

$$\mathbf{X}^{[k,k,m]} = \mathbf{X}^{[k,m,m]} = \mathbf{A}^{[k]} \mathbf{S}^{[k,m]} \mathbf{A}^{[m]\top} = \mathbf{X}^{[k,m]} \quad \forall k, m \quad (15)$$

where the rightmost equality in (15) follows from (5). Therefore, for fixed $k \neq m$, each of the sequences $\{\mathbf{C}^{[k,l,m]}\}_{l=1}^K$ and $\{\mathbf{X}^{[k,l,m]}\}_{l=1}^K$ contains (at most) $(K-1)$ distinct matrices, whereas for $k = m$, all the matrices in $\{\mathbf{C}^{[k,l,k]}\}_{l=1}^K$ and $\{\mathbf{X}^{[k,l,k]}\}_{l=1}^K$ may be distinct. In order to avoid this degeneracy, from this point and on, we implicitly assume that all the tensors $\mathcal{X}^{[k,m]}$ are constructed such that they do not contain redundant frontal slices, i.e., they do not contain both $\mathbf{X}^{[k,k,m]}$ and $\mathbf{X}^{[k,m,m]}$ for fixed k and m , if $k \neq m$. Therefore, the third ‘‘depth’’ dimension of the tensors $\mathcal{X}^{[k,k]}$ and $\mathcal{X}^{[k,m]}|_{k \neq m}$ is not, in general, the same. As an example, if all $\mathbf{A}^{[k]}$ are nonsingular, then the largest tensors that we can construct for $k \neq m$ are $\mathcal{X}^{[k,m]} \in \mathbb{R}^{I^{[k]} \times I^{[m]} \times (K-1)}$, and for $k = m$, $\mathcal{X}^{[k,k]} \in \mathbb{R}^{I^{[k]} \times I^{[k]} \times K}$. Another point to keep in mind is that due to symmetry, $\mathcal{X}^{[m,k]}$ does not contribute any information beyond $\mathcal{X}^{[k,m]}$.

The bottom line is that we have restated the JBSS problem that we defined in Section 2 as an ensemble of tensors $\{\mathcal{X}^{[k,m]}\}_{k,m=1}^K$ that admit a CPD (13). The tensors in this ensemble are coupled because each of them shares, deterministically, a factor matrix $\mathbf{A}^{[k]}$ and/or $\mathbf{A}^{[m]}$ with one or several other CPDs. Hence, we say that the ensemble $\{\mathcal{X}^{[k,m]}\}_{k,m=1}^K$ admits a coupled CPD.

It is worth noting that if $\mathbf{A}^{[k]}$ is nonsingular for some k , the CPD of $\mathcal{X}^{[k,k]}$ amounts to joint diagonalization (JD), and if both $\mathbf{A}^{[k]}$ and $\mathbf{A}^{[m]}$ are nonsingular for some $m \neq k$, then $\mathcal{X}^{[k,m]}$ is an asymmetric two-sided tensor diagonalization. If $\mathbf{A}^{[k]}$ is nonsingular $\forall k$, this coupled CPD amounts to generalized joint diagonalization (GJD) [4]. In general, these are simpler problems.

3.1 Previous Links between JBSS and Coupled CPD

In fact, the coupled CPD model that we have just described, in Section 3, and its link with JBSS, are not new, as we now explain. Let us ignore for a moment the latent structure of the entries of $\mathbf{C}^{[k,m]}$ due to (9) and (10). Then, a representation of the sufficient statistics of a JBSS model in terms of tensors $\{\mathcal{X}^{[k,m]}\}_{k,m=1}^K$ that admit a coupled CPD as in (13) has already been introduced (e.g., [4,5], [7, Section VI]). However, until now, a three-way structure has been considered only when the sources had an additional type of diversity, such as statistical dependence (correlation) among samples, or nonstationarity; in other words, only when the i.i.d. assumption was violated. In these cases, the distinct frontal slices of each tensor $\mathcal{X}^{[k,m]}$ represented, for example, different correlation matrices taken at different time lags. In these cases, the third ‘‘depth’’ dimension of the tensors reflected the additional diversity in the data. In this paper, we show for the first time that a coupled CPD formulation is possible even if none of these additional types of diversity is present.

3.2 Discussion

Equation (15) implies that each tensor $\mathcal{X}^{[k,m]}$ has one frontal slice that is identical to the (k, m) th matrix in (5). The difference from (5) is that now, for fixed k and m , we have more than one equation that involves $\mathbf{A}^{[k]}$ and $\mathbf{A}^{[m]}$. Therefore, the new coupled CPD formulation subsumes the coupled matrix factorization in (5). In previous work (e.g., [11]), it was shown that the JBSS problem in Section 2 can be solved optimally, in the maximum likelihood (ML) sense, i.e., in terms of the minimal mean square error (MMSE), using the simpler coupled matrix factorization in (5). The proposed coupled CPD formulation uses exactly the same information, and thus cannot achieve a better MMSE. Similarly, in terms of uniqueness, recall that the coupled factorization in (5) uses all the sufficient statistics, and thus, its uniqueness is tantamount to JBSS identifiability [12,13,14]. Consequently, the alternative formulation of the same statistics in terms of a coupled CPD cannot achieve stronger uniqueness properties. It is thus natural to ask what we can obtain from the more complicated coupled CPD formulation in Section 3. The rest of this paper is dedicated to this question.

4 Uniqueness

The identifiability of the JBSS model in Section 2 was characterized in [12,13,14], for nonsingular $\mathbf{A}^{[k]} \forall k$. The model was shown to be identifiable except for very special cases that were fully characterized in [12,13,14], and depended only on the values of $\{\mathbf{S}^{[k,l]}\}_{k,l=1}^K$. Our goal in this section is to show how these non-identifiable scenarios are reflected in the tensorized framework of Section 3. In this paper, we focus only on one of these non-identifiable cases; conclusions for the other cases can be obtained similarly.

In [12,13,14], it was shown that our JBSS model is not identifiable if there exists a pair (i, j) , $i \neq j$, of sources, whose statistics satisfy

$$s_{jj}^{[k,l]} = \varphi^{[k]} \varphi^{[l]} s_{ii}^{[k,l]} \quad \forall k, l \quad (16)$$

where $\varphi^{[k]} \neq 0 \forall k$. Equation (16) implies that, for each k , the subspaces associated with the i th and j th columns of $\mathbf{A}^{[k]}$ cannot be distinguished. Substituting (16) in (10), we obtain

$$c_j^{[k,l,m]} = s_{jj}^{[k,l]} s_{jj}^{-[l,l]} s_{jj}^{[l,m]} = \varphi^{[k]} \varphi^{[m]} s_{ii}^{[k,l]} s_{ii}^{-[l,l]} s_{ii}^{[l,m]} = \varphi^{[k]} \varphi^{[m]} c_i^{[k,l,m]} \quad \forall k, l, m \quad (17)$$

which implies

$$\mathbf{c}_j^{[k,m]} = \varphi^{[k]} \varphi^{[m]} \mathbf{c}_i^{[k,m]} \quad \forall k, m \quad (18)$$

The other non-identifiable cases are associated with pairs (i, j) of source correlations $\{s_{ii}^{[k,l]}\}_{k,l=1}^K$ and $\{s_{jj}^{[k,l]}\}_{k,l=1}^K$ that contain zeros. Calculations similar to those in (17) show that in these non-identifiable cases, some of $\mathbf{C}^{[k,m]}$ contain

two zero columns ($\mathbf{c}_i^{[k,m]} = \mathbf{0} = \mathbf{c}_j^{[k,m]}$), and the remaining $\mathbf{C}^{[k,m]}$ have pairs of proportional columns that contain zeros in some of their entries, in specific locations. Due to lack of space, we omit the details.

Let us assume for a moment that $s_{rr}^{[k,l]} \neq 0 \forall k, l, r$. Hence, the model is not identifiable, and the coupled CPD is not unique, iff (16) holds for some pair (i, j) , $i \neq j$. Equation (18) implies that, in this case, the i th and j th columns of $\mathbf{C}^{[k,m]}$ are proportional $\forall k, m$. Hence, $k_{\mathbf{C}^{[k,m]}} = 1 \forall k, m$, and none of the tensors $\mathcal{X}^{[k,m]}$ has a unique CPD [8]. However, it is important to note that the notation “ $k_{\mathbf{C}^{[k,m]}} = 1 \forall k, m$ ” does not provide information about the indices of the proportional columns. Hence, it does not necessarily imply that the coupled CPD is not unique: if, for some pair of (k, m) , the proportional columns are not in the same indices (i, j) as in the other CPDs, then (16) does not hold, and the JBSS is identifiable. In this case, the coupled matrix factorization (5), as well as the coupled CPD associated with it, are unique. This result is of potential interest because it is the first time that the uniqueness of the coupled CPD is stated explicitly for $\mathbf{C}^{[k,m]}$ that do not have full column rank $\forall k, m$ (the uniqueness analysis of the coupled CPD in [15] assumes that $\mathbf{C}^{[k,m]}$ has full column rank $\forall k, m$, and that the tensors do not have any latent structure). Similar conclusions can be obtained from observing the structure of $\mathbf{C}^{[k,m]}$ in the other non-identifiable cases [12,13,14], as we have mentioned earlier in this section.

5 Estimating Mixing Matrices From a Single CPD

In this section, we focus our attention on a single CPD (or JD, or a two-sided tensor diagonalization) in (13), within the context of the JBSS model in Section 2. It follows from Section 4 that it is possible to have $k_{\mathbf{C}^{[k,m]}} \geq 2$. In this case, the CPD of $\mathcal{X}^{[k,m]}$ may be unique. The uniqueness of the CPD is guaranteed, for example, if it satisfies [8]

$$k_{\mathbf{A}^{[k]}} + k_{\mathbf{A}^{[m]}} + k_{\mathbf{C}^{[k,m]}} \geq 2R + 2. \quad (19)$$

Equation (19) implies that, in certain cases, a single CPD may be unique even if $\mathbf{A}^{[k]}$ and/or $\mathbf{A}^{[m]}$, as well as $\mathbf{C}^{[k,m]}$, do not have full column rank. If the CPD of $\mathcal{X}^{[k,m]}$ is unique, then we can extract both $\mathbf{A}^{[k]}$ and $\mathbf{A}^{[m]}$ (if $k \neq m$) or just $\mathbf{A}^{[k]}$ (if $k = m$) from it.¹

Using a single CPD in (13) to estimate one or two mixing matrices can be regarded as an intermediate stage between GEVD and coupled decomposition of the whole ensemble. It allows to compute a single mixing matrix, or two mixing matrices, using any standard JD or CPD algorithm, from a subset of the available cross-correlations. The result of this computation may be used for initialization, or validation, instead of GEVD. An advantage w.r.t. GEVD is that

¹ In practice, due to finite sample size and noise, (13) is just an approximation. Questions related to uniqueness and estimation in the presence of perturbations from the exact model are beyond the scope of this paper.

JD and CPD can take into account more than two frontal slices, providing a more accurate initialization (for example). Note also that the GEVD solution in (6) is applicable only to data whose mixing matrices have full column rank, whereas this restriction is relaxed when using a CPD. A drawback of this approach is that we may lose the inherent ability of JBSS to fix a single permutation for all the rank-1 terms in all datasets [1].

6 Application to Underdetermined JBSS

In this section, we demonstrate how our new formulation of the coupled CPD can bring new insights about JBSS with underdetermined mixtures, when $\mathbf{C}^{[k,m]}$ does not have full column rank for at least one pair of (k, m) . To the best of our knowledge, this case has not yet been addressed in the literature.

Consider a JBSS setup as in Section 2. In this example, we assume that $s_{rr}^{[k,l]} \neq 0 \forall k, l, r$. Assume that $K-1$ mixtures, indexed, w.l.o.g., $k = 1, \dots, K-1$, with nonsingular mixing matrices, satisfy (16) (with $K-1$ instead of K), and thus are not identifiable. The K th mixture is underdetermined, with a mixing matrix $\mathbf{A}^{[K]}$ that has more columns than rows, i.e. $I^{[K]} < R$. We assume that the cross-correlations of the sources, when taking into account all datasets $k = 1, \dots, K$, do not satisfy (16). Our goal is to demonstrate that this model may be identifiable.

We suggest to solve this problem by constructing a tensor $\mathcal{X}^{[K,m]}$, whose CPD will yield a unique estimate of $\mathbf{A}^{[K]}$ and $\mathbf{A}^{[m]}$. Let us begin by looking at $\mathbf{C}^{[K,m]}$. As an example, let $K = 4$, and $m = 1$. Then,

$$\mathbf{C}^{[4,1]} = \begin{bmatrix} s_{ii}^{[4,1]} s_{ii}^{-[1,1]} s_{ii}^{[1,1]} & s_{jj}^{[4,1]} s_{jj}^{-[1,1]} s_{jj}^{[1,1]} & & & \\ \dots & s_{ii}^{[4,2]} s_{ii}^{-[2,2]} s_{ii}^{[2,1]} & \dots & s_{jj}^{[4,2]} s_{jj}^{-[2,2]} s_{jj}^{[2,1]} & \dots \\ s_{ii}^{[4,3]} s_{ii}^{-[3,3]} s_{ii}^{[3,1]} & s_{jj}^{[4,3]} s_{jj}^{-[3,3]} s_{jj}^{[3,1]} & & & \end{bmatrix} \in \mathbb{R}^{3 \times R} \quad (20a)$$

$$= \begin{bmatrix} s_{ii}^{[4,1]} s_{ii}^{-[1,1]} s_{ii}^{[1,1]} & & & s_{jj}^{[4,1]} & & \\ \dots & s_{ii}^{[4,2]} s_{ii}^{-[2,2]} s_{ii}^{[2,1]} & \dots & \varphi^{-[2]} \varphi^{[1]} s_{jj}^{[4,2]} s_{ii}^{-[2,2]} s_{ii}^{[2,1]} & \dots & \\ s_{ii}^{[4,3]} s_{ii}^{-[3,3]} s_{ii}^{[3,1]} & & \varphi^{-[3]} \varphi^{[1]} s_{jj}^{[4,3]} s_{ii}^{-[3,3]} s_{ii}^{[3,1]} & & & \end{bmatrix} \quad (20b)$$

Equation (20) depicts explicitly the i th and j th columns of $\mathbf{C}^{[4,1]}$. The transition to (20b) is due to (16). In this scenario, $\mathbf{C}^{[K,m]}$ has size $(K-1) \times R$; recall that $\mathbf{A}^{[K]}$ is not invertible, and thus, cannot take part in (8). Equation (20) shows that if, as we assume, the cross-correlations $s_{rr}^{[4,m]}$ are such that the ensemble $\{s_{rr}^{[k,l]}\}_{k,l=1}^K$ does not satisfy (16), then there is no linear dependence between $\mathbf{c}_i^{[4,1]}$ and $\mathbf{c}_j^{[4,1]}$. Consequently, $\mathbf{C}^{[4,1]}$ has full rank (although not necessarily full column rank).

We now turn to the uniqueness of $\mathcal{X}^{[K,m]}$, when $m < K$. By (19), the CPD of $\mathcal{X}^{[K,m]}$ is unique if

$$\underbrace{I^{[K]}}_{k_{\mathbf{A}^{[K]}}} + \underbrace{R}_{k_{\mathbf{A}^{[m]}}} + \underbrace{\min(R, K-1)}_{k_{\mathbf{C}^{[K,m]}}} \geq 2R + 2 \quad (21)$$

It follows from (21) that the CPD of $\mathcal{X}^{[K,m]}$ is unique, for example, when $R = 3$, $I^{[K]} = 2$, and $K = 4$, or when $R = 4$, $I^{[k]} = 3$, and $K = 4$.

As soon as $\mathbf{A}^{[K]}$ and $\mathbf{A}^{[m]}$ have been identified, for some fixed m , we can identify all the remaining mixtures $\mathbf{A}^{[k]}$, $k < K$, $k \neq m$, using the fact that now $\mathbf{A}^{[m]}$ is known and invertible, and the diagonal matrix $\mathbf{S}^{[k,m]}$ is nonsingular [15]:

$$\mathbf{X}^{[k,m]} = \mathbf{A}^{[k]} \mathbf{S}^{[k,m]} \mathbf{A}^{[m]\top} \Rightarrow \mathbf{X}^{[k,m]} \mathbf{A}^{-[m]\top} = \mathbf{A}^{[k]} \mathbf{S}^{[k,m]} \quad (22)$$

In this identifiable setup, the tensors $\mathcal{X}^{[k,m]}$ with $k, m < K$ have nonsingular factors $\mathbf{A}^{[k]}$ and $\mathbf{A}^{[m]}$ and, as explained in Section 4, a third factor matrix $\mathbf{C}^{[k,m]}$ with $k_{\mathbf{C}^{[k,m]}} = 1$. The tensors $\mathcal{X}^{[K,m]}$ involve one underdetermined factor $\mathbf{A}^{[K]}$, a nonsingular $\mathbf{A}^{[m]}$, and a third factor matrix $\mathbf{C}^{[K,m]}$ that has full rank, but may have more columns than rows, as explained earlier in Section 6. Our results show that this model is identifiable, and hence, the overall coupled CPD must be unique, too. This result has been obtained using only one CPD in the ensemble. It is likely that the overall coupled CPD has an even stronger uniqueness.

7 Conclusion

In this paper, we have shown, for the first time, that JBSS of $K \geq 3$ mixtures can be associated with the CPD even in the simplest case where each source in each mixture is a real-valued Gaussian stationary random process with uncorrelated samples. Apart from the theoretical interest in showing another type of link between the statistically-motivated JBSS and an algebraic tensor-based model, we proposed several practical uses to this new formulation. We have shown that this new formulation can bring new insights and new stronger uniqueness results on coupled CPD and on JBSS. In a broader perspective, we provided another evidence for the richness of coupled decompositions.

In Section 3.2, we mentioned that the new formulation cannot improve on the MMSE. However, it remains to be seen if the proposed formulation, of coupled CPD, can achieve a smaller estimation error when a norm that does not achieve the ML, e.g., the Frobenius norm, is used in the optimization, and, if so, is the improvement justified w.r.t. the higher computational cost.

The same type of analysis that we presented in this paper for decomposition in sum of rank-1 elements can be extended to terms of any low rank [7], and to complex-valued data. Finally, if the JBSS data have additional diversity, e.g., sample nonstationarity, this information can be added to each tensor $\mathcal{X}^{[k,m]}$ as additional frontal slices, and thus further enhance the estimation.

References

1. Kim, T., Eltoft, T., Lee, T.W.: Independent vector analysis: An extension of ICA to multivariate components. In: Independent Component Analysis and Blind Signal Separation. Volume 3889 of LNCS., Springer Berlin Heidelberg (2006) 165–172

2. Li, Y.O., Adalı, T., Wang, W., Calhoun, V.D.: Joint blind source separation by multiset canonical correlation analysis. *IEEE Trans. Signal Process.* **57**(10) (Oct. 2009) 3918–3929
3. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* **6**(1) (1927) 164–189
4. Li, X.L., Adalı, T., Anderson, M.: Joint blind source separation by generalized joint diagonalization of cumulant matrices. *Signal Process.* **91**(10) (Oct. 2011) 2314–2322
5. Congedo, M., Phlypo, R., Chatel-Goldman, J.: Orthogonal and non-orthogonal joint blind source separation in the least-squares sense. In: *Proc. EUSIPCO*, Bucharest, Romania (Aug. 2012) 1885–1889
6. Lahat, D., Jutten, C.: Joint analysis of multiple datasets by cross-cumulant tensor (block) diagonalization. In: *Proc. SAM*, Rio de Janeiro, Brazil (Jul. 2016)
7. Lahat, D., Jutten, C.: Joint independent subspace analysis using second-order statistics. *IEEE Trans. Signal Process.* **64**(18) (Sep. 2016) 4891–4904
8. Kruskal, J.B.: Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Proc. London Math. Soc.* **18**(2) (1977) 95–138
9. Lahat, D., Adalı, T., Jutten, C.: Multimodal data fusion: An overview of methods, challenges and prospects. *Proc. IEEE* **103**(9) (Sep. 2015) 1449–1477
10. Anderson, T.W.: *An introduction to multivariate statistical analysis*. John Wiley & Sons (1958)
11. Vía, J., Anderson, M., Li, X.L., Adalı, T.: A maximum likelihood approach for independent vector analysis of Gaussian data sets. In: *Proc. MLSP*, Beijing, China (Sep. 2011)
12. Vía, J., Anderson, M., Li, X.L., Adalı, T.: Joint blind source separation from second-order statistics: Necessary and sufficient identifiability conditions. In: *Proc. ICASSP*, Prague, Czech Republic (May 2011) 2520–2523
13. Anderson, M., Fu, G.S., Phlypo, R., Adalı, T.: Independent vector analysis: Identification conditions and performance bounds. *IEEE Trans. Signal Process.* **62**(17) (Sep. 2014) 4399–4410
14. Lahat, D., Jutten, C.: An alternative proof for the identifiability of independent vector analysis using second order statistics. In: *Proc. ICASSP*, Shanghai, China (Mar. 2016)
15. Gong, X.F., Lin, Q.H., Cong, F.Y., De Lathauwer, L.: Double coupled canonical polyadic decomposition for joint blind source separation. [arXiv:1612.09466](https://arxiv.org/abs/1612.09466) [stat.ML] (2017)