CONFAC Decomposition Approach to Blind Identification of Underdetermined Mixtures Based on Generating Function Derivatives

André L. F. Almeida, Xavier Luciani, Alwin Stegeman, Pierre Comon

CONFAC Decomposition Approach to Blind Identification of Underdetermined Mixtures Based on Generating Function Derivatives

André L. F. de Almeida*, *Member, IEEE*, Xavier Luciani,
Alwin Stegeman, and Pierre Comon, *Fellow, IEEE*

***Abstract*— This work proposes a new tensor-based approach to solve the problem of blind identification of underdetermined mixtures of complex sources exploiting the cumulant generating function (CGF) of the observations. We show that a collection of second-order derivatives of the CGF of the observations can be stored in a third-order tensor following a constrained factor (CONFAC) decomposition with known constrained structure. In order to increase the diversity, we combine three derivative types into an extended third-order CONFAC decomposition. A detailed uniqueness study of this decomposition is provided, from which easy-to-check sufficient conditions ensuring the essential uniqueness of the mixing matrix are obtained. From an algorithmic viewpoint, we develop a CONFAC-based enhanced line search (CONFAC-ELS) method to be used with an alternating least squares estimation procedure for accelerated convergence, and also analyze the numerical complexities of two CONFAC-based algorithms (namely, CONFAC-ALS and CONFAC-ELS) in comparison with the Leverberg-Marquardt (LM)-based algorithm recently derived to solve the same problem. Simulation results compare the proposed approach with some higher-order methods. Our results also corroborate the advantages of the CONFAC-based approach over the competing LM-based approach in terms of performance and computational complexity.**

***Index Terms*— Blind identification, second generating function, complex sources, CONFAC decomposition.**

EDICS: SSP-HOSM, SSP-IDEN, SPC-BLND.

I. Introduction

Blind identification methods have been successfully applied in multidisciplinary contexts including radiocommunications, sonar, radar, biomedical signal processing and data analysis, just to mention a few. A widespread class of these methods relies on independent component analysis by means of higher-order statistics [1]. This subject has been at the center of many theoretical works while related methods and algorithms have been used in a variety of application fields [2]–[4] (see also [5], [6] for surveys). A problem that has attracted a particular interest is that of blind identification of underdetermined mixtures. Several solutions have been proposed in the literature to solve this problem (see, e.g. [6], [7], [14]–[18] and references therein). The proposed solutions resort to second, fourth or sixth-order statistics of the observations.

André L. F. de Almeida is with the Wireless Telecom Research Laboratory, Department of Teleinformatics Engineering, Federal University of Ceará, CP 6005, Campus do Pici, 60455-760, Fortaleza, Brazil. Tel/Fax: +33 85 3366 9470. Xavier Luciani is with the PROTEE Laboratory, University of Toulon, BP 20132, 83957 LA GARDE Cedex, France. Alwin Stegeman is with the Heijmans Institute for Psychological Research, University of Groningen, The Netherlands. Pierre Comon is with GIPSA-Lab, Grenoble Campus, BP.46, F-38402 St Martin d'Heres cedex, France. E-mails: andre@gtel.ufc.br, lucianix@gmail.com, a.w.stegeman@rug.nl, p.comon@ieee.org

Several solutions to blind identification/source separation problems have been proposed recently based on multi-way (tensor) analysis [24]. This is a subject that has gained attention in numerous application areas involving data analysis such as psychometrics [19], arithmetic complexity [20] and chemometrics [21], [22]. In this context, canonical polyadic (CP) decomposition[1] [19], [23] is the most popular tensor decomposition. The CP decomposition has been successfully used as an alternative solution to principal component analysis (PCA) when the available data to be analyzed can be arranged as a meaningful multi-way array, or a higher-order tensor [24]. Indeed, the widespread use of the CP decomposition can be attributed to its essential uniqueness property under mild conditions [25]–[31] as well as to the existence of several numerical algorithms that can be used to compute this decomposition [21], [32]–[38].

A first class of the so-called tensor-based methods directly exploits the trilinear nature of the observed data, and the decomposition of the data tensor provides a direct estimation of the sources. These methods have been widely applied in wireless communications by means of different tensor decompositions (see e.g. [8]–[13]). However, when the diversity of the observations is not sufficient, one can resort to a second class of tensor-based methods that rely on the multilinearity properties of higher-order statistics (HOS) [14], [18], [35]. A large majority of these methods solves the blind identification problem by means of the CP decomposition of a tensor storing the cumulants of the observations [7], [14], [39], [41], [42], [45]. This is the case, for instance, of FOOBI/FOOBI2 [17], [18], and BIOME [15] algorithms, which capitalize on the triadic decomposition of fourth- and sixth-order output cumulants, respectively. The approach presented in [39] relies on the CP decomposition of third- or fourth-order cross spectra of the observations. The works [40]–[43] address convolutive mixtures.

A particular class of blind identification methods exploits the second characteristic function of the observations. This idea has first appeared in [44] and later in a few works [16], [46], [47]. In [16], the authors showed that partial derivatives of the second characteristic function of the observations taken at different points of the observation space can be stored in a symmetric tensor, the CP decomposition of which provides a direct estimation of the mixing matrix up to trivial scaling and permutation indeterminacies. In [16], the alternating least squares (ALS) algorithm is applied to blindly estimate the mixing matrix from a CP

[1]This decomposition is sometimes also called Candecomp/Parafac, which can be referred to with the same acronynm.

data tensor constructed from third-order derivatives of the output characteristic function. In a recent work [48], we have considered a more general scenario where the sources are assumed to be complex-valued (e.g. 4-PSK or 4-QAM), which is usually the case in digital communications. Therein, the authors resort to the cumulant generating function (CGF) of the observations and a Levenberg-Marquardt (LM) based algorithm is proposed to estimate the mixing matrix.

In this work, we show that the CGF-based blind identification problem can be more efficiently addressed by means of the constrained factor (CONFAC) decomposition [49]. Under the assumption of complex sources, we show that a collection of second-order derivatives of the CGFs of the observations can be stored in a third-order tensor following a third-order CONFAC decomposition with known constraint matrices. The profile of 1's and 0's of the constraint matrices captures the linear combination patterns involving real and imaginary components of the CGFs derivatives. In order to increase the diversity, we combine three derivative types into an extended CONFAC decomposition of increased dimensionality. The uniqueness property of this decomposition is studied and our results establish a set of easy-to-check sufficient conditions that guarantee the essential uniqueness of the mixing matrix. From an algorithmic viewpoint, we develop a CONFAC-based enhanced line search (CONFAC-ELS) method to be used with an alternating least squares (ALS) estimation procedure for accelerated convergence. In particular, the numerical complexities of two CONFAC-based algorithms (namely, CONFAC-ALS and CONFAC-ELS) is analyzed and compared with the numerical complexity of with the Leverberg-Marquardt (LM)-based algorithm derived in [48] (therein called LEMACAFC) to solve the same problem. Our computer simulation results evaluate and compare the estimation accuracy of the proposed approach with those of competing higher-order methods. Our results also attest the efficiency of CONFAC-based solutions over the competing LM-based approach proposed in [48].

Contributions: The contributions of this paper can be highlighted as follows:

- In comparison with [48], which does not rely on a tensor decomposition approach, we originally formulate the CGF-based blind identification problem in the case of complex sources as a constrained tensor decomposition problem with a priori known structure. From a tensor decomposition perspective, this work is a direct generalization of [16] to the complex case. In the same way as the CP decomposition fits the case of real mixture of real sources in [16], this work shows that the CONFAC decomposition (which is more general than the CP one) fits the case of complex mixtures of complex sources;
- We provide a detailed uniqueness study of the proposed CONFAC decomposition. A set of easy-to-check conditions ensuring the essential uniqueness of the mixing matrix is provided, guiding the choice of the number of sensors and number of derivatives points for a given number of sources. This is in contrast to [48], where the uniqueness issue is not addressed and no condition for blind identifiability of the mixing matrix is given;
- Using the fact that second-order derivatives of the CGF of a Gaussian noise are constant regardless of the point at which the derivatives are computed, it is possible to "denoise" the derivative tensor to be decomposed, yielding more robustness to noise effects;
- We develop an enhanced line search (ELS) optimization for the CONFAC-based blind identification method and a numerical complexity analysis of the so-called CONFAC-ELS algorithm is provided. This analysis corroborates the efficiency of the proposed algorithm compared to the LM-based algorithm of [48].

The distinguishing feature of the proposed second-order approach is its low complexity compared to higher-order methods, and its capability to deal with underdetermined mixtures without requiring constraints on the temporal structure of the sources such as correlation [18] and piece-wise stationarity [50]. Additionally, the proposed CONFAC-ELS algorithm is more computationally efficient than the competing LM-based algorithm [48], as can be seen from our numerical analysis and simulation results.

This paper is organized as follows. In Section II, a background on the CONFAC decomposition of a third-order tensor is provided. In Section III, we formulate the CGF-based blind identification problem and present the main core equations. The problem is recast in Section IV using the proposed CONFAC decomposition approach. In Section V, a uniqueness study of the proposed decomposition is provided, from which a set easy-to-check conditions ensuring the essential uniqueness of the mixing matrix is presented. Section VI presents the CONFAC-ELS blind identification algorithm and discusses its numerical complexity. Simulation results are given in Section VII and the paper is concluded in Section VIII.

Notations: In the following, vectors, matrices and tensors are denoted by lower case boldface ($\mathbf{a}$), upper case boldface ($\mathbf{A}$) and upper case calligraphic ($\mathcal{A}$) letters respectively. a_i is the i-th coordinate of vector $\mathbf{a}$ and $\mathbf{a}_i$ is the i-th column of matrix $\mathbf{A}$. The (i,j) entry of matrix $\mathbf{A}$ is denoted A_{ij} and the (i,j,k) entry of the third order tensor $\mathcal{A}$ is denoted $\mathcal{A}_{ijk}$. $\mathbf{I}_K$ denotes the identity matrix of size K. Real and imaginary parts are denoted $\Re\{\cdot\}$ and $\Im\{\cdot\}$ respectively. E[.] denotes the expected value of a random variable. $\mathbf{A}^T$ and $\mathbf{A}^\dagger$ stand, respectively, for the transpose and Moore-Penrose pseudo-inverse of $\mathbf{A}$. $\det(\mathbf{A})$ denotes the determinant of $\mathbf{A}$. The operator $\text{vec}(\mathbf{A})$ creates a column vector $\mathbf{a} \in \mathbb{C}^{JI\times 1}$ from $\mathbf{A} \in \mathbb{C}^{I\times J}$ by stacking its J columns below one another, while $\text{unvec}(\mathbf{a}) = \mathbf{A}$ is the inverse operator, i.e. $\text{unvec}(\text{vec}(\mathbf{A})) = \mathbf{A}$. The operator $D_i(\mathbf{A})$ forms a diagonal matrix out of the i-th row of $\mathbf{A}$. The outer vector product is denoted by $\circ$, while $\boxdot$ denotes the Hadamard (element-wide product). Kronecker and Khatri-Rao products are denoted by $\otimes$ and $\odot$, respectively.

II. Preliminaries: The CONFAC Decomposition

Let us consider a third-order tensor $\mathcal{X} \in \mathbb{C}^{P\times Q\times R}$, three *factor matrices* $\mathbf{A} \in \mathbb{C}^{P\times F_1}$, $\mathbf{B} \in \mathbb{C}^{Q\times F_2}$, $\mathbf{C} \in \mathbb{C}^{R\times F_3}$, and three *constraint matrices* $\mathbf{\Theta} \in \mathbb{C}^{F_1\times F}$, $\mathbf{\Psi} \in \mathbb{C}^{F_2\times F}$, $\mathbf{\Omega} \in \mathbb{C}^{F_3\times F}$. The CONFAC decomposition of $\mathcal{X}$ with F factor combinations is defined in scalar form as:

$$X_{pqr} = \sum_{f=1}^{F} \sum_{f_1=1}^{F_1} \sum_{f_2=1}^{F_2} \sum_{f_3=1}^{F_3} A_{pf_1} B_{qf_2} C_{rf_3} \Theta_{f_1 f} \Psi_{f_2 f} \Omega_{f_3 f}, \quad (1)$$

$$\text{with } F \geq \max(F_1, F_2, F_3),$$

where $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ are the *factor matrices*, while $\mathbf{\Theta}, \mathbf{\Psi}$ and $\mathbf{\Omega}$ are the *constraint matrices*. The factor matrices are unknown, to be determined, while the constraint matrices are known fixed matrices whose structure satisfies the two following assumptions:

A1 The columns of $\mathbf{\Theta}$ (resp. $\mathbf{\Psi}$ and $\mathbf{\Omega}$) are canonical vectors possibly multiplied by -1. The canonical vectors composing these matrices belong, respectively, to the canonical bases $\{\mathbf{e}_1^{(F_1)}, \ldots, \mathbf{e}_{F_1}^{(F_1)}\} \in \mathbb{R}^{F_1}$, $\{\mathbf{e}_1^{(F_2)}, \ldots, \mathbf{e}_{F_2}^{(F_2)}\} \in \mathbb{R}^{F_2}$, and $\{\mathbf{e}_1^{(F_3)}, \ldots, \mathbf{e}_{F_3}^{(F_3)}\} \in \mathbb{R}^{F_3}$.

A2 $\mathbf{\Theta}$, $\mathbf{\Psi}$ and $\mathbf{\Omega}$ are full-row rank matrices with ranks equal to F_1, F_2, and F_3, respectively.

The CONFAC decomposition can be stated in a different manner, which sheds light on a different way of interpreting its constrained structure. By exchanging summations in (1), we obtain:

$$X_{pqr} = \sum_{f_1=1}^{F_1} \sum_{f_2=1}^{F_2} \sum_{f_3=1}^{F_3} A_{pf_1} B_{qf_2} C_{rf_3} W_{f_1 f_2 f_3}(\mathbf{\Theta}, \mathbf{\Psi}, \mathbf{\Omega}), \quad (2)$$

where

$$W_{f_1 f_2 f_3}(\mathbf{\Theta}, \mathbf{\Psi}, \mathbf{\Omega}) = \sum_{f=1}^{F} \Theta_{f_1 f} \Psi_{f_2 f} \Omega_{f_3 f} \quad (3)$$

is an element of a $F_1 \times F_2 \times F_3$ tensor $\mathcal{W}(\mathbf{\Theta}, \mathbf{\Psi}, \mathbf{\Omega})$ that follows an F-factor triadic decomposition in terms of $\mathbf{\Theta}$, $\mathbf{\Psi}$ and $\mathbf{\Omega}$. We call $\mathcal{W}(\mathbf{\Theta}, \mathbf{\Psi}, \mathbf{\Omega})$, or simply $\mathcal{W}$, the *constrained core tensor* of the CONFAC decomposition. Due to the structure assumed from the constraint matrices, $\mathcal{W}$ contains +/-1 elements at fixed positions. Figure 1 provides an illustration of the CONFAC decomposition. Note that the CONFAC decomposition (1) can be seen as a constrained Tucker3 decomposition [51]–[53] with the particular characteristic of having a core tensor with known triadic decomposition. Moreover, note that the CONFAC decomposition for which $F_1 = F_2 = F_3 = F$, and $\mathbf{\Theta} = \mathbf{\Psi} = \mathbf{\Omega} = \mathbf{I}_F$ reduces to the F-factor canonical polyadic (CP) decomposition [19], [23].

Two different matrix representations of the tensor $\mathcal{X} \in \mathbb{C}^{P \times Q \times R}$ are possible, namely the *sliced* and *unfolded* representations. Their construction and factorization follow the same reasoning as that of the CP decomposition [9]. Thus, $\mathbf{X}_{p\cdot\cdot} = \mathbf{B}\mathbf{\Psi} D_p\big(\mathbf{A}\mathbf{\Theta}\big)\big(\mathbf{C}\mathbf{\Omega}\big)^T \in \mathbb{C}^{Q \times R}$ is the factorization of the p-th slice of $\mathcal{X} \in \mathbb{C}^{P \times Q \times R}$ along its first dimension. Similarly, $\mathbf{X}_{\cdot q\cdot} = \mathbf{C}\mathbf{\Omega} D_q\big(\mathbf{B}\mathbf{\Psi}\big)\big(\mathbf{A}\mathbf{\Theta}\big)^T \in \mathbb{C}^{R \times P}$ and $\mathbf{X}_{\cdot\cdot r} = \mathbf{A}\mathbf{\Theta} D_r\big(\mathbf{C}\mathbf{\Omega}\big)\big(\mathbf{B}\mathbf{\Psi}\big)^T \in \mathbb{C}^{P \times Q}$ are the factorization of its q-th and r-th slices along its second and third dimensions, respectively. The full information contained in tensor $\mathcal{X} \in \mathbb{C}^{P \times Q \times R}$ can be organized in three *unfolded matrices* $\mathbf{X}_{(1)} = [\mathbf{X}_{1\cdot\cdot}^T, \ldots, \mathbf{X}_{P\cdot\cdot}^T]^T \in \mathbb{C}^{PQ \times R}$, $\mathbf{X}_{(2)} = [\mathbf{X}_{\cdot 1\cdot}^T, \ldots, \mathbf{X}_{\cdot Q\cdot}^T]^T \in \mathbb{C}^{QR \times P}$, and $\mathbf{X}_{(3)} = [\mathbf{X}_{\cdot\cdot R}^T, \ldots, \mathbf{X}_{\cdot\cdot R}^T]^T \in \mathbb{C}^{RP \times Q}$, which admit, respectively, the following factorizations:

$$\mathbf{X}_{(1)} = \big((\mathbf{A}\mathbf{\Theta}) \odot (\mathbf{B}\mathbf{\Psi})\big)\big(\mathbf{C}\mathbf{\Omega}\big)^T, \; \mathbf{X}_{(2)} = \big((\mathbf{B}\mathbf{\Psi}) \odot (\mathbf{C}\mathbf{\Omega})\big)\big(\mathbf{A}\mathbf{\Theta}\big)^T, \; \mathbf{X}_{(3)} = \big((\mathbf{C}\mathbf{\Omega}) \odot (\mathbf{A}\mathbf{\Theta})\big)\big(\mathbf{B}\mathbf{\Psi}\big)^T \quad (4)$$

The CONFAC decomposition of a third-order tensor was originally proposed in [49] in the context of wireless communications to design multiple-antenna transmission schemes with blind detection. Therein, it is shown that the three CONFAC constraint matrices are design parameters of the multiple-antenna transmission system. The CONFAC decomposition also appears in related works [54], [55], where the constraint matrices of the decomposition are restricted to fixed linear combination patterns. Uniqueness conditions for the third-order CONFAC decomposition have been studied in a recent work [56].

III. Problem Formulation

We consider a noisy linear mixture of K narrowband sources received by an array of N sensors. Let $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K] \in \mathbb{C}^{N \times K}$ be the mixing matrix. Define $\mathbf{z}(m) = [z_1(m), \ldots, z_N(m)]^T \in \mathbb{C}^N$, $\mathbf{s}(m) = [s_1(m), \ldots, s_K(m)]^T \in \mathbb{C}^K$ and $\mathbf{e}(m) \in \mathbb{C}^N$ as the m-th discrete-time realizations of the observations, source and noise vectors, respectively, $m = 1, \ldots, M$. According to this model we have:

$$\mathbf{z}(m) = \mathbf{H}\mathbf{s}(m) + \mathbf{e}(m). \quad (5)$$

The sources can be real- or complex-valued and the noise samples are modeled as zero-mean circularly-symmetric complex Gaussian random variables.

The problem consists in estimating $\mathbf{H}$ from the only knowledge of the observations. In this work, we are interested in solving this problem by resorting to partial derivatives of the CGF of the observations. Specifically, the problem consists in finding $\widehat{\mathbf{H}}$ such that

$$\widehat{\mathbf{H}} = \mathbf{H}\mathbf{\Pi}\mathbf{\Lambda}, \quad (6)$$

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}$ is a diagonal matrix. This means that $\mathbf{H}$ can be identified up to permutation and scaling of its columns. Column permutation and scaling ambiguities are referred to as *trivial ambiguities*.

The blind identification of the mixing matrix $\mathbf{H}$ relies on the following assumptions:

H1. The matrix $\mathbf{H}$ does not contain pairwise collinear columns;

H2. The sources $s_1, \ldots, s_K$ are non-Gaussian and mutually statistically independent;

H3. The number K of sources is known.

It has been shown in former studies [61], [62] that $\mathbf{H}$ is theoretically identifiable under these assumptions.

A. Generating function of the observations

We recall from [48] the main steps that formulate the second-order derivatives of the CGF of the observations in the case of complex sources. The cumulant generating function (CGF) of the observations, Φ_z, can be decomposed in a sum of marginal second generating functions of the sources, φ_k, $k = 1 \cdots K$. We start by defining Φ_z and φ_k in the complex field. The second generating function φ_k of the k-th source taken at the point x of $\mathbb{C}$ defined $\mathbb{R}^2$ is given by

$$\varphi_k(\Re\{x\}, \Im\{x\}) \stackrel{\text{def}}{=} \log \mathrm{E}[\exp(\Re\{x^* s_k\})]. \quad (7)$$

Similarly, the second generating function $\Phi_{\underline{z}}$ of the observations taken at the point $\boldsymbol{\xi} = (\mathbf{u}, \mathbf{v})$ defined in $\mathbb{R}^{2N}$ can be written as

$$\Phi_z(\mathbf{u}, \mathbf{v}) \stackrel{\text{def}}{=} \log \mathrm{E}[\exp(\mathbf{x}^{\mathrm{T}}\mathbf{u} + \mathbf{y}^{\mathrm{T}}\mathbf{v})],$$

where $\mathbf{x} = \Re\{\mathbf{z}\}$ and $\mathbf{y} = \Im\{\mathbf{z}\}$. Define $\mathbf{A}$ and $\bar{\mathbf{A}}$ as the real and imaginary parts of the mixing matrix so that $\mathbf{H} = \mathbf{A} + i\bar{\mathbf{A}}$. Next, denote $\mathbf{a}_k$ (resp. $\bar{\mathbf{a}}_k$) the k-th column of $\mathbf{A}$ (resp. $\bar{\mathbf{A}}$).

Replacing $\mathbf{z}$ by its model and using sources' mutual statistical independence hypothesis yields:

$$\Phi_z(\mathbf{u},\mathbf{v}) = \sum_k \varphi_k\left(\mathbf{u}^{\mathrm{T}}\mathbf{a}_k + \mathbf{v}^{\mathrm{T}}\bar{\mathbf{a}}_k, \mathbf{v}^{\mathrm{T}}\mathbf{a}_k - \mathbf{u}^{\mathrm{T}}\bar{\mathbf{a}}_k\right) + \Phi_e(\mathbf{u},\mathbf{v}). \tag{8}$$

where $\Phi_e(\mathbf{u},\mathbf{v})$ is the corresponding second generating function of the Gaussian noise. From these definitions, we can rewrite (8) as

$$\Phi_z(\boldsymbol{\xi}) = \sum_k \varphi_k\left(g_1(\boldsymbol{\xi}),\, g_2(\boldsymbol{\xi})\right) + \Phi_e(\boldsymbol{\xi}), \tag{9}$$

where $g_1(\boldsymbol{\xi}) = \sum_n A_{nk}u_n + \bar{A}_{nk}v_n$ and $g_2(\boldsymbol{\xi}) = \sum_n A_{nk}v_n - \bar{A}_{nk}u_n$. Defining

$$\begin{aligned} g:\quad \mathbb{R}^{2N} &\longrightarrow \mathbb{R}^2 \\ \boldsymbol{\xi} &\longmapsto g(\boldsymbol{\xi}) = (g_1(\boldsymbol{\xi}), g_2(\boldsymbol{\xi})), \end{aligned}$$

yields a compact representation for (9) as

$$\Phi_z(\boldsymbol{\xi}) = \sum_k \varphi_k\left(g(\boldsymbol{\xi})\right) + \Phi_e(\boldsymbol{\xi}). \tag{10}$$

B. Differentiation of $\Phi_z(\boldsymbol{\xi})$

Following the approach of [16], [48], we compute partial derivatives of Φ_z at R different points of $\mathbb{R}^{2N}$, denoted here as $\boldsymbol{\xi}^{(r)} = (\mathbf{u}^{(r)}, \mathbf{v}^{(r)})$, $r = 1,\ldots,R$.

Let $\{\Phi_z(\boldsymbol{\xi}^{(1)}), \Phi_z(\boldsymbol{\xi}^{(2)}), \ldots, \Phi_z(\boldsymbol{\xi}^{(R)})\}$ be the set containing the CGFs of the observations evaluated at the R points. Similarly, $\{\Phi_e(\boldsymbol{\xi}^{(1)}), \Phi_e(\boldsymbol{\xi}^{(2)}), \ldots, \Phi_e(\boldsymbol{\xi}^{(R)})\}$ denotes the CGFs of the additive Gaussian noise. Define

$$G_{rk}^{ij} \overset{\text{def}}{=} \frac{\partial^2 \varphi_k(g(\boldsymbol{\xi}^{(r)}))}{\partial g_i(\boldsymbol{\xi}^{(r)})\partial g_j(\boldsymbol{\xi}^{(r)})},\ i = 1,2, \quad j = 1,2, \tag{11}$$

and note that $G_{rk}^{12} = G_{rk}^{21}$. By successively differentiating (10) with respect to the variable pairs $\{u_q^{(r)}, u_q^{(r)}\}$, $\{v_q^{(r)}, v_q^{(r)}\}$ and $\{u_q^{(r)}, v_q^{(r)}\}$ we can obtain the three different second-order derivative equations, respectively:

$$\begin{aligned} X_{pqr}^{\Phi_1} = \sum_{k=1}^{K} \Big(& A_{pk}A_{qk}G_{rk}^{11} - A_{pk}\bar{A}_{qk}G_{rk}^{12} - \bar{A}_{pk}A_{qk}G_{rk}^{12} \\ & + \bar{A}_{pk}\bar{A}_{qk}G_{rk}^{22}\Big) + \Upsilon_{pq}^{(1)}, \end{aligned} \tag{12}$$

$$\begin{aligned} X_{pqr}^{\Phi_2} = \sum_{k=1}^{K} \Big(& \bar{A}_{pk}\bar{A}_{qk}G_{rk}^{11} + \bar{A}_{pk}A_{qk}G_{rk}^{12} + A_{pk}\bar{A}_{qk}G_{rk}^{12} \\ & + A_{pk}A_{qk}G_{rk}^{22}\Big) + \Upsilon_{pq}^{(2)}, \end{aligned} \tag{13}$$

$$\begin{aligned} X_{pqr}^{\Phi_3} = \sum_{k=1}^{K} \Big(& A_{pk}\bar{A}_{qk}G_{rk}^{11} + A_{pk}A_{qk}G_{rk}^{12} - \bar{A}_{pk}\bar{A}_{qk}G_{rk}^{12} \\ & - \bar{A}_{pk}A_{qk}G_{rk}^{22}\Big) + \Upsilon_{pq}^{(3)}, \end{aligned} \tag{14}$$

where

$$X_{pqr}^{\Phi_1} \overset{\text{def}}{=} \frac{\partial^2 \Phi_z(\boldsymbol{\xi}^{(r)})}{\partial u_p^{(r)}\partial u_q^{(r)}},\ X_{pqr}^{\Phi_2} \overset{\text{def}}{=} \frac{\partial^2 \Phi_z(\boldsymbol{\xi}^{(r)})}{\partial v_p^{(r)}\partial v_q^{(r)}},\ X_{pqr}^{\Phi_3} \overset{\text{def}}{=} \frac{\partial^2 \Phi_z(\boldsymbol{\xi}^{(r)})}{\partial u_p^{(r)}\partial v_q^{(r)}}, \tag{15}$$

and where $X_{pqr}^{\Phi_s}$ is the (p,q,r)-th entry of the s-th "derivative tensor" $\mathcal{X}^{\Phi_s}$ of dimensions $N \times N \times R$, $s = 1,2,3$, while $\Upsilon_{pq}^{(1)} = \partial^2\Phi_e(\boldsymbol{\xi}^{(r)})/\partial u_p^{(r)}\partial u_q^{(r)}$, $\Upsilon_{pq}^{(2)} = \partial^2\Phi_e(\boldsymbol{\xi}^{(r)})/\partial v_p^{(r)}\partial v_q^{(r)}$, and $\Upsilon_{pq}^{(3)} = \partial^2\Phi_e(\boldsymbol{\xi}^{(r)})/\partial u_p^{(r)}\partial v_q^{(r)}$, $r = 1,\ldots,R$, are the corresponding second-order derivatives of the noise term, which follow the same calculation steps as those of the second-order derivatives of the signal part of the model. We call attention to the fact that $\Upsilon^{(1)}$, $\Upsilon^{(2)}$, and $\Upsilon^{(3)}$ do not depend on the index r, since the second-order derivative of the noise is constant and not affected by the point at which the derivative tensor is calculated.

IV. The CONFAC decomposition approach

A. Denoised formulation of the tensor decomposition

We use the fact that the CGFs of a Gaussian noise evaluated at R different points of the observation space have identical second-order derivatives to eliminate the noise influence from the derivative tensors in (12), (13) and (14). This is possible by subtracting from the s-th derivative tensor $\mathcal{X}_{pqr}^{\Phi_s}$ the second-order CGF derivatives evaluated at the origin. More specifically, the "denoised" derivative tensors are given by

$$Y_{pqr}^{\Phi_s} = X_{pqr}^{\Phi_s} - X_{pq0}^{\Phi_s}, \quad s = 1,2,3, \tag{16}$$

where $X_{pq0}^{\Phi_s}$ denotes the second-order derivative of the CGF of the observations evaluated at the origin, i.e. at $\boldsymbol{\xi}^{(0)} = \mathbf{0}$ in $\mathbb{R}^{2N}$. Note that $X_{pq0}^{\Phi_s}$ is equal to $\Upsilon_{pq}^{(s)}$, $s = 1,2,3$, the noise contribution, since the signal part of the CGF derivatives vanish at the origin. Thus, by computing the difference given in (16) allows one to remove the noise component $\Upsilon_{pq}^{(s)}$ so that the resulting tensor $Y_{pqr}^{\Phi_s}$ will be noise free.

Let $\mathbf{A}^{(k)} \in \mathbb{R}^{N\times 2}$ and $\mathbf{G}^{(k)} \in \mathbb{R}^{R\times 3}$, $k = 1,\ldots,K$, be defined as:

$$\mathbf{A}^{(k)} \overset{\text{def}}{=} \begin{bmatrix} A_{1k} & \bar{A}_{1k} \\ \vdots & \vdots \\ A_{Nk} & \bar{A}_{Nk} \end{bmatrix} = [\mathbf{a}_k,\, \bar{\mathbf{a}}_k],$$

$$\mathbf{G}^{(k)} \overset{\text{def}}{=} \begin{bmatrix} G_{1k}^{11} & G_{1k}^{12} & G_{1k}^{22} \\ \vdots & \vdots & \vdots \\ G_{Rk}^{11} & G_{Rk}^{12} & G_{Rk}^{22} \end{bmatrix} = [\mathbf{g}_{1,k},\, \mathbf{g}_{2,k},\, \mathbf{g}_{3,k}].$$

Note that two columns of $\mathbf{A}^{(k)}$ correspond to the real and imaginary parts of the k-th column of the mixing matrix, respectively. Each of the three columns of $\mathbf{G}^{(k)}$ is associated with a second-order derivative type collected at R points. Using these definitions, as subsequently shown, we can decompose the denoised derivative tensor $\mathcal{Y}^{\Phi_s}$, $s = 1,2,3$, as follows:

$$Y_{pqr}^{\Phi_s} = \sum_{k=1}^{K} \Big(\underbrace{\sum_{f=1}^{4}\sum_{f_1=1}^{2}\sum_{f_2=1}^{2}\sum_{f_3=1}^{3} A_{pf_1}^{(k)}A_{qf_2}^{(k)}G_{rf_3}^{(k)}\Theta_{f_1 f}\Psi_{f_2 f}\Omega_{f_3 f}^{(s)}}_{Y_{pqr}^{\Phi_s}(k)} \Big) \tag{17}$$

which corresponds to a sum of K CONFAC tensor blocks $Y_{pqr}^{\Phi_s}(1),\ldots,Y_{pqr}^{\Phi_s}(K)$. The k-th block is given by a sum of four outer products involving repeated columns of matrices $\mathbf{A}^{(k)}$ and $\mathbf{G}^{(k)}$. The linear combination pattern involving the columns of $\mathbf{A}^{(k)}$ and $\mathbf{G}^{(k)}$ is determined by the joint structure of $\boldsymbol{\Theta}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}^{(s)}$. By comparing (17) with the general formulation (1), we can deduce the following correspondences:

$$\begin{aligned} &(\mathbf{A},\mathbf{B},\mathbf{C}) \leftrightarrow (\mathbf{A}^{(k)}, \mathbf{A}^{(k)}, \mathbf{G}^{(k)}), \\ &(\boldsymbol{\Theta},\boldsymbol{\Psi},\boldsymbol{\Omega}) \leftrightarrow (\boldsymbol{\Theta},\boldsymbol{\Psi},\boldsymbol{\Omega}^{(s)}), \\ &(F_1,F_2,F_3,F) \leftrightarrow (2,2,3,4), \\ &(P,Q,R) \leftrightarrow (N,N,R), \end{aligned}$$

For each one of the three derivative types, the constraint matrices $\mathbf{\Theta}^{(s)}$, $\mathbf{\Psi}^{(s)}$ and $\mathbf{\Omega}^{(s)}$ of the associated CONFAC decomposition can be identified by comparing respectively (12), (13) and (14) with (17). A possible choice is given by

$$\mathbf{\Theta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{\Psi} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \tag{18}$$

$$\mathbf{\Omega}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{\Omega}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Omega}^{(3)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \tag{19}$$

Note that the first- and second-mode constraint matrices do not depend on the differentiation variables. This dependence is confined only in the third-mode constraint matrix.

Remark 1: The structures chosen for the constraint matrices in (18)-(19) are not unique. This can be seen by rewriting (17) as

$$Y^{\Phi_s}_{pqr} = \sum_{k=1}^{K} \sum_{f_1=1}^{2} \sum_{f_2=1}^{2} \sum_{f_3=1}^{3} A^{(k)}_{pf_1} A^{(k)}_{qf_2} G^{(k)}_{rf_3} W^{(s)}_{f_1 f_2 f_3}, \tag{20}$$

where

$$W^{(s)}_{f_1 f_2 f_3} = \sum_{f=1}^{4} \Theta_{f_1 f} \Psi_{f_2 f} \Omega^{(s)}_{f_3 f}. \tag{21}$$

is a triadic decomposition of $\mathcal{W} \in \mathbb{C}^{2\times 2\times 3}$. Even under assumptions *A1-A2* of Section II, which must be satisfied by $\mathbf{\Theta}$, $\mathbf{\Psi}$ and $\mathbf{\Omega}^{(s)}$, there is still freedom to permute their columns and/or change their signs without changing the constrained core tensor $\mathcal{W}^{(s)}$, $s = 1, 2, 3$. However, these ambiguities are unimportant in our context since the three constraint matrices are known by definition. Any structural choice for these matrices that satisfies the decomposition could be adopted. Our choice given in (18)-(19) is motivated by the convenience of having only the third-mode constraint matrix changing as the differentiation variables are changed, while the two first constraint matrices are fixed. Most importantly, such a choice makes possible to easily combine the three derivative types into a single CONFAC decomposition as will be seen in the next section.

Let us define the block matrices

$$\mathbf{A} = [\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(K)}] \in \mathbb{R}^{N\times 2K} \tag{22}$$

$$\mathbf{G} = [\mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(K)}] \in \mathbb{R}^{R\times 3K}, \tag{23}$$

which concatenate the contributions of the K sources. Define also the block-diagonal constraint matrices

$$\bar{\mathbf{\Theta}} = \mathbf{I}_K \otimes \mathbf{\Theta} \in \mathbb{R}^{2K\times 4K}, \tag{24}$$

$$\bar{\mathbf{\Psi}} = \mathbf{I}_K \otimes \mathbf{\Psi} \in \mathbb{R}^{2K\times 4K}, \tag{25}$$

$$\bar{\mathbf{\Omega}}^{(s)} = \mathbf{I}_K \otimes \mathbf{\Omega}^{(s)} \in \mathbb{R}^{3K\times 4K}. \tag{26}$$

With these definitions, we can treat (17) simply as an block-CONFAC decomposition composed of K blocks, the k-th block being associated with the k-th source. In this case, the following correspondences can be obtained by analogy with (1):

$$(\mathbf{A}, \mathbf{B}, \mathbf{C}) \leftrightarrow (\mathbf{A}, \mathbf{A}, \mathbf{G}),$$
$$(\mathbf{\Theta}, \mathbf{\Psi}, \mathbf{\Omega}) \leftrightarrow (\bar{\mathbf{\Theta}}, \bar{\mathbf{\Psi}}, \bar{\mathbf{\Omega}}^{(s)}),$$
$$(F_1, F_2, F_3, F) \leftrightarrow (2K, 2K, 3K, 4K),$$
$$(P, Q, R) \leftrightarrow (N, N, R),$$

The unfolded matrix representations for the second-order derivative tensor $\mathcal{Y}^{\Phi_s} \in \mathbb{R}^{N\times N\times R}$, $s = 1, 2, 3$, follow those of (4). For instance, the unfolded representation $\mathbf{Y}^{\Phi_s}_{(1)} \in \mathbb{R}^{N^2\times R}$ can be written as:

$$\begin{aligned} \mathbf{Y}^{\Phi_s}_{(1)} &= \begin{bmatrix} \mathbf{Y}^{\Phi_s}_{1..} \\ \vdots \\ \mathbf{Y}^{\Phi_s}_{N..} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\bar{\mathbf{\Psi}} D_1(\mathbf{A}\bar{\mathbf{\Theta}}) \\ \vdots \\ \mathbf{A}\bar{\mathbf{\Psi}} D_N(\mathbf{A}\bar{\mathbf{\Theta}}) \end{bmatrix} \big(\mathbf{G}\bar{\mathbf{\Omega}}^{(s)}\big)^T \\ &= \big((\mathbf{A}\bar{\mathbf{\Theta}}) \odot (\mathbf{A}\bar{\mathbf{\Psi}})\big)\big(\mathbf{G}\bar{\mathbf{\Omega}}^{(s)}\big)^T. \end{aligned} \tag{27}$$

B. Combining all derivatives into an extended CONFAC decomposition

As shown in (12), (13) and (14), three types of second-order derivatives can be obtained from the CGF of the observations depending on the pair of variables with respect to which the derivatives are computed. Each derivative type yields a CONFAC decomposition $\mathcal{Y}^{\Phi_s} \in \mathbb{R}^{N\times N\times R}$ with a different constrained structure, the structural difference being confined in the third-mode constraint matrix $\bar{\mathbf{\Omega}}^{(s)}$, $s = 1, 2, 3$, as we have proposed in the previous subsection. In order to increase diversity, we take all the three types of second-order derivatives into account by constructing an extended CONFAC decomposition, as follows

$$\bar{\mathbf{Y}}_{(1)} = \begin{bmatrix} \mathbf{Y}^{\Phi_1}_{(1)} \\ \mathbf{Y}^{\Phi_2}_{(1)} \\ \mathbf{Y}^{\Phi_3}_{(1)} \end{bmatrix} = \begin{bmatrix} \big((\mathbf{A}\bar{\mathbf{\Theta}}) \odot (\mathbf{A}\bar{\mathbf{\Psi}})\big)\big(\mathbf{G}\bar{\mathbf{\Omega}}^{(1)}\big)^T \\ \big((\mathbf{A}\bar{\mathbf{\Theta}}) \odot (\mathbf{A}\bar{\mathbf{\Psi}})\big)\big(\mathbf{G}\bar{\mathbf{\Omega}}^{(2)}\big)^T \\ \big((\mathbf{A}\bar{\mathbf{\Theta}}) \odot (\mathbf{A}\bar{\mathbf{\Psi}})\big)\big(\mathbf{G}\bar{\mathbf{\Omega}}^{(3)}\big)^T \end{bmatrix} \in \mathbb{R}^{3N^2\times R}. \tag{28}$$

This unfolded representation can be rewritten as

$$\bar{\mathbf{Y}}_{(1)} = \big(\mathbf{I}_3 \otimes \big((\mathbf{A}\bar{\mathbf{\Theta}}) \odot (\mathbf{A}\bar{\mathbf{\Psi}})\big)\big)(\mathbf{G}\tilde{\mathbf{\Omega}})^T, \tag{29}$$

where

$$\tilde{\mathbf{\Omega}} = \left[\bar{\mathbf{\Omega}}^{(1)}, \bar{\mathbf{\Omega}}^{(2)}, \bar{\mathbf{\Omega}}^{(3)}\right] \in \mathbb{R}^{3K\times 12K}. \tag{30}$$

Likewise, we can also stack row-wise the three derivative types into the second and third unfolded matrices $\bar{\mathbf{Y}}_{(2)} \in \mathbb{R}^{3NR\times N}$ and $\bar{\mathbf{Y}}_{(3)} \in \mathbb{R}^{3RN\times N}$, as follows

$$\bar{\mathbf{Y}}_{(2)} = \begin{bmatrix} \mathbf{Y}^{\Phi_1}_{.1.} \\ \mathbf{Y}^{\Phi_2}_{.1.} \\ \mathbf{Y}^{\Phi_3}_{.1.} \\ - - - - - \\ \vdots \\ - - - - - \\ \mathbf{Y}^{\Phi_1}_{.N.} \\ \mathbf{Y}^{\Phi_2}_{.N.} \\ \mathbf{Y}^{\Phi_3}_{.N.} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{G}\bar{\mathbf{\Omega}}^{(1)} \\ \mathbf{G}\bar{\mathbf{\Omega}}^{(2)} \\ \mathbf{G}\bar{\mathbf{\Omega}}^{(3)} \end{bmatrix} D_1(\mathbf{A}\bar{\mathbf{\Psi}}) \\ - - - - - \\ \vdots \\ - - - - - \\ \begin{bmatrix} \mathbf{G}\bar{\mathbf{\Omega}}^{(1)} \\ \mathbf{G}\bar{\mathbf{\Omega}}^{(2)} \\ \mathbf{G}\bar{\mathbf{\Omega}}^{(3)} \end{bmatrix} D_N(\mathbf{A}\bar{\mathbf{\Psi}}) \end{bmatrix} (\mathbf{A}\bar{\mathbf{\Theta}})^T, \tag{31}$$

and

$$\bar{\mathbf{Y}}_{(3)} = \begin{bmatrix} \mathbf{Y}^{\Phi_1}_{..1} \\ \vdots \\ \mathbf{Y}^{\Phi_1}_{..R} \\ - - - - - \\ \vdots \\ - - - - - \\ \mathbf{Y}^{\Phi_3}_{..1} \\ \vdots \\ \mathbf{Y}^{\Phi_3}_{..R} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\bar{\mathbf{\Theta}}\mathbf{D}_1(\mathbf{G}\bar{\mathbf{\Omega}}^{(1)}) \\ \vdots \\ \mathbf{A}\bar{\mathbf{\Theta}}\mathbf{D}_R(\mathbf{G}\bar{\mathbf{\Omega}}^{(1)}) \\ - - - - - \\ \vdots \\ - - - - - \\ \mathbf{A}\bar{\mathbf{\Theta}}\mathbf{D}_1(\mathbf{G}\bar{\mathbf{\Omega}}^{(3)}) \\ \vdots \\ \mathbf{A}\bar{\mathbf{\Theta}}\mathbf{D}_R(\mathbf{G}\bar{\mathbf{\Omega}}^{(3)}) \end{bmatrix} (\mathbf{A}\bar{\mathbf{\Psi}})^T. \tag{32}$$

After some algebraic manipulations the unfolded matrices (31) and (32) can be respectively rewritten in compact form as

$$\bar{\mathbf{Y}}_{(2)} = \big(\mathbf{A}\bar{\mathbf{\Psi}} \odot (\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}}\big)(\mathbf{A}\bar{\mathbf{\Theta}})^{\mathrm{T}}, \tag{33}$$

$$\bar{\mathbf{Y}}_{(3)} = \big((\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}} \odot \mathbf{A}\bar{\mathbf{\Theta}}\big)(\mathbf{A}\bar{\mathbf{\Psi}})^{\mathrm{T}}, \tag{34}$$

where

$$\bar{\mathbf{\Omega}} = \begin{bmatrix} \bar{\mathbf{\Omega}}^{(1)} \\ \bar{\mathbf{\Omega}}^{(2)} \\ \bar{\mathbf{\Omega}}^{(3)} \end{bmatrix} \in \mathbb{R}^{9K \times 4K}. \tag{35}$$

V. Uniqueness study

Recall that the goal of the proposed blind identification problem is to identify the possibly underdetermined mixture $\mathbf{H}$ (up to trivial column permutation and scaling ambiguities) from the only knowledge of the CGF derivatives of the observations. Herein, this problem is addressed by exploiting the uniqueness property of the CONFAC decomposition of a tensor that combines the three derivative types, as shown Section IV-B. In order to study uniqueness, it is convenient to rewrite the unfolded representation $\bar{\mathbf{X}}_1$ as follows:

$$\begin{aligned} \bar{\mathbf{X}}_1 &= \begin{bmatrix} (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)} \end{bmatrix} \mathbf{G}^T = (\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A}) \begin{bmatrix} \mathbf{T}^{(1)} \\ \mathbf{T}^{(2)} \\ \mathbf{T}^{(3)} \end{bmatrix} \mathbf{G}^T \\ &= (\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}\,\mathbf{G}^T, \end{aligned} \tag{36}$$

with

$$\mathbf{T}^{(s)} = [(\mathbf{I}_K\otimes\mathbf{\Theta})\odot(\mathbf{I}_K\otimes\mathbf{\Psi})]\,(\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)})^T,\ s=1,2,3\,. \tag{37}$$

The matrices $\mathbf{T}^{(s)}$ have size $4K^2\times 3K$. The matrix $(\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}$ has size $3N^2\times 3K$.

The block $\mathbf{A}_k$ of the real-valued mixing matrix $\mathbf{A}$ contains the real part $\mathbf{a}_k$ and imaginary part $\bar{\mathbf{a}}_k$ of the k-th column of the complex-valued $N\times K$ mixing matrix $\mathbf{H}$, i.e. $\mathbf{H} = [\mathbf{a}_1\ \ldots\ \mathbf{a}_K] + i\,[\bar{\mathbf{a}}_1\ \ldots\ \bar{\mathbf{a}}_K]$. The mixing matrix $\mathbf{H}$ is called *essentially unique* if for any alternative $\hat{\mathbf{H}}$ the relation $\hat{\mathbf{H}} = \mathbf{H}\,\mathbf{\Pi}\,\mathbf{\Lambda}$ holds, with $\mathbf{\Pi}$ a $K\times K$ permutation matrix, and $\mathbf{\Lambda}$ a complex-valued nonsingular diagonal matrix. Multiplying the k-th column of $\mathbf{H}$ by $\alpha + i\,\beta$ yields $(\alpha\,\mathbf{a}_k - \beta\,\bar{\mathbf{a}}_k) + i\,(\beta\,\mathbf{a}_k + \alpha\,\bar{\mathbf{a}}_k)$. This motivates the following definition of essential uniqueness for $\mathbf{A}$.

Definition 1: Matrix $\mathbf{A}$ is called *essentially unique* if for any alternative $\mathbf{F} = [\mathbf{F}_1\,|\,\ldots\,|\,\mathbf{F}_K]$, with each $\mathbf{F}_k$ of size $N\times 2$, there holds

$$\mathbf{F}_k = \mathbf{A}_{\pi(k)} \begin{bmatrix} \alpha_k & \beta_k \\ -\beta_k & \alpha_k \end{bmatrix}, \qquad k = 1,\ldots,K\,, \tag{38}$$

with $\pi(\cdot)$ a permutation of $\{1,\ldots,K\}$, and α_k and β_k not both zero, $k=1,\ldots,K$. □

The following theorem is concerned with the uniqueness of the CONFAC decomposition (36) for $K=1$. This result is conveniently presented here as it will be useful later.

Theorem 1: Let $(\mathbf{A},\mathbf{G})$ be a solution to CONFAC decomposition (36) with $K=1$. Suppose that $\mathrm{rank}(\mathbf{A})=2$ and $\mathrm{rank}(\mathbf{G})=3$. Then $\mathbf{A}$ is essentially unique as defined in Definition 1.

Proof. See Appendix I.

Remark 2: For $K>1$, it can be proven that $\mathbf{A}$ is essentially unique if $\mathbf{A}$ and $\mathbf{G}$ have full column-rank, i.e. $\mathrm{rank}(\mathbf{A})=2K$ and $\mathrm{rank}(\mathbf{G})=3K$. Note that, although the full column-rank condition for $\mathbf{A}$ can be restrictive from the view point of blind identification, the full column-rank of $\mathbf{G}$ is likely to hold when the number R of points at which the derivatives are evaluated is large enough.

In the remainder of this section, we do not use the assumption of $\mathbf{A}$ having full column rank. Assume instead that $(\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}$ has full column rank $3K$ (which implies $N^2\geq K$). This assumption, together with $\mathbf{G}$ having full column rank $3K$, implies that we may set $\mathbf{G}=\mathbf{I}_{3K}$ without loss of generality. A proof of this is analogous to that derived in [56] (c.f. Lemma 3.4). We denote alternative component matrices as $\mathbf{F}$ ($N\times 2K$) and $\mathbf{L}$ ($3K\times 3K$), with $\mathbf{F} = [\mathbf{F}_1\,|\,\ldots\,|\,\mathbf{F}_K]$ and $\mathbf{F}_k = [\mathbf{f}_k\,|\,\bar{\mathbf{f}}_k]$. For $\mathbf{G}=\mathbf{I}_{3K}$, equating the CONFAC decomposition (36) to its alternative yields

$$\bar{\mathbf{X}}_1 = (\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T} = (\mathbf{I}_3\otimes\mathbf{F}\otimes\mathbf{F})\,\mathbf{T}\,\mathbf{L}^T. \tag{39}$$

The goal is now to prove that both decompositions coincide up to trivial transformations. By assumption, the left-hand side has rank $3K$, and it follows that $\mathbf{L}$ is nonsingular and that $(\mathbf{I}_3\otimes\mathbf{F}\otimes\mathbf{F})\,\mathbf{T}$ has full column rank $3K$. We write

$$(\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}\,\mathbf{L}^{-T} = (\mathbf{I}_3\otimes\mathbf{F}\otimes\mathbf{F})\,\mathbf{T}\,. \tag{40}$$

Next, we consider the structure of $(\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}$. We have

$$\begin{aligned} \bar{\mathbf{Y}}_{(1)} &= (\mathbf{I}_3\otimes\mathbf{A}\otimes\mathbf{A})\,\mathbf{T} = \begin{bmatrix} (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{P}^{(1)}_1 & \ldots & \mathbf{P}^{(1)}_K \\ \mathbf{P}^{(2)}_1 & \ldots & \mathbf{P}^{(2)}_K \\ \mathbf{P}^{(3)}_1 & \ldots & \mathbf{P}^{(3)}_K \end{bmatrix}, \end{aligned} \tag{41}$$

with, for $k=1,\ldots,K$,

$$\begin{aligned} \mathbf{P}^{(1)}_k &= [\,\mathbf{a}_k\otimes\mathbf{a}_k \mid -\mathbf{a}_k\otimes\bar{\mathbf{a}}_k - \bar{\mathbf{a}}_k\otimes\mathbf{a}_k \mid \bar{\mathbf{a}}_k\otimes\bar{\mathbf{a}}_k\,] \\ &= (\mathbf{A}^{(k)}\otimes\mathbf{A}^{(k)})\,(\mathbf{\Omega}^{(1)})^T \end{aligned} \tag{42}$$

$$\begin{aligned} \mathbf{P}^{(2)}_k &= [\,\bar{\mathbf{a}}_k\otimes\bar{\mathbf{a}}_k \mid \mathbf{a}_k\otimes\bar{\mathbf{a}}_k + \bar{\mathbf{a}}_k\otimes\mathbf{a}_k \mid \mathbf{a}_k\otimes\mathbf{a}_k\,] \\ &= (\mathbf{A}^{(k)}\otimes\mathbf{A}^{(k)})\,(\mathbf{\Omega}^{(2)})^T \end{aligned} \tag{43}$$

$$\begin{aligned} \mathbf{P}^{(3)}_k &= [\,\mathbf{a}_k\otimes\bar{\mathbf{a}}_k \mid \mathbf{a}_k\otimes\mathbf{a}_k - \bar{\mathbf{a}}_k\otimes\bar{\mathbf{a}}_k \mid -\bar{\mathbf{a}}_k\otimes\mathbf{a}_k\,] \\ &= (\mathbf{A}^{(k)}\otimes\mathbf{A}^{(k)})\,(\mathbf{\Omega}^{(3)})^T \end{aligned} \tag{44}$$

where the matrices $\mathbf{\Theta}$ and $\mathbf{\Psi}$ have vanished because $\mathbf{\Theta}\odot\mathbf{\Psi}=\mathbf{I}_4$, from (18). Obviously, the matrix $(\mathbf{I}_3\otimes\mathbf{F}\otimes\mathbf{F})\,\mathbf{T}$ has identical structure, with $\mathbf{A}$ replaced by $\mathbf{F}$. Below, we derive several constraints on $\mathbf{L}$ implied by the equations above. These constraints are used to obtain uniqueness results for generic $\mathbf{A}$ and specific

values of N and K. Let $\mathbf{L}^{-T} = [\mathbf{l}_1 \ \dots \ \mathbf{l}_{3K}]$. The structure of (42)-(44), and (40) imply that

$$(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_1 = (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_3\,, \quad (45)$$
$$(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_3 = (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_1\,, \quad (46)$$
$$(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_2 = -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_2\,, \quad (47)$$
$$(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_2 = (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_1 - (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_1, \quad (48)$$
$$(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_1 = (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_2 + (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_3, \quad (49)$$

These equations can be written as

$$\begin{bmatrix} (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)} & \mathbf{O} & -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)} & \mathbf{O} & -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)} \\ \mathbf{O} & (\mathbf{A}\otimes\mathbf{A})(\mathbf{T}^{(1)}+\mathbf{T}^{(2)}) & \mathbf{O} \\ (\mathbf{A}\otimes\mathbf{A})(\mathbf{T}^{(1)}-\mathbf{T}^{(2)}) & -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)} & \mathbf{O} \\ (\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)} & -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)} & -(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)} \end{bmatrix} \begin{pmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \mathbf{l}_3 \end{pmatrix} = \mathbf{0}\,, \quad (50)$$

where $\mathbf{O}$ denotes an $N^2 \times 3K$ all-zero matrix. The system (50) should hold not only for $(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$, but for all $(\mathbf{l}_t, \mathbf{l}_{t+1}, \mathbf{l}_{t+2})$, $t = 3k+1$, $k = 0, 1, \dots, K-1$. The matrix in (50) has size $5N^2 \times 9K$. Its first three row blocks contain only columns that are vectorized symmetric $N \times N$ matrices. This yields $3N(N-1)/2$ redundant rows. From definition (37), it can be verified that the null space of the matrix in (50) has at least dimensionality $3K$, and contains the vectors

$$[\mathbf{l}_1 \ \ \mathbf{l}_2 \ \ \mathbf{l}_3] = \left[\begin{array}{ccc} \alpha_1 & 2\,\alpha_2 & \alpha_3 \\ \alpha_2 & \alpha_1 - \alpha_3 & -\alpha_2 \\ \alpha_3 & -2\,\alpha_2 & \alpha_1 \\ \hline \vdots & \vdots & \vdots \\ \hline \alpha_{3K-2} & 2\,\alpha_{3K-1} & \alpha_{3K} \\ \alpha_{3K-1} & \alpha_{3K-2} - \alpha_{3K} & -\alpha_{3K-1} \\ \alpha_{3K} & -2\,\alpha_{3K-1} & \alpha_{3K-2} \end{array}\right]. \quad (51)$$

If the null space of the matrix in (50) is defined by (51), then $\mathbf{l}_2$ and $\mathbf{l}_3$ are completely determined if $\mathbf{l}_1$ is known. The same is true for all triplets $(\mathbf{l}_t, \mathbf{l}_{t+1}, \mathbf{l}_{t+2})$, $t = 3k+1$, $k = 0, 1, \dots, K-1$.

Next, we derive a second set of constraints on $\mathbf{L}^{-T}$. From the structure of (42)-(44), it follows that $(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_1 = \mathbf{f}_1 \otimes \mathbf{f}_1$, $(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_1 = \bar{\mathbf{f}}_1 \otimes \bar{\mathbf{f}}_1$ and $(\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_1 = \mathbf{f}_1 \otimes \bar{\mathbf{f}}_1$ are vectorized rank-1 matrices. This implies that

$$\begin{bmatrix} \mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(1)}\,\mathbf{l}_1) & \mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_1) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(2)}\,\mathbf{l}_1) \\ \mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_1) \end{bmatrix} \quad (52)$$

both have rank 1, where $\mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(s)}\,\mathbf{l}_1) = \mathbf{A}\mathrm{diag}(\mathbf{T}^{(s)}\mathbf{l}_1)\mathbf{A}^T$, which has size $N \times N$, $s = 1, 2, 3$. This must hold not only for $\mathbf{l}_1$ but for all $\mathbf{l}_t$, $t = 3k+1$, $k = 0, 1, \dots, K-1$. The matrices in (52) have rank 1 if and only if all their second-order minors are zero. We need the following definition.

Definition 2: For a $P \times Q$ matrix $\mathbf{X}$, let the $P(P-1)/2 \times Q(Q-1)/2$ matrix $m(\mathbf{X})$ have entries

$$\det \begin{pmatrix} x_{ig} & x_{ih} \\ x_{jg} & x_{jh} \end{pmatrix}, \qquad \text{with} \quad 1 \le i < j \le P \quad \text{and} \quad 1 \le g < h \le Q, \quad (53)$$

where in each row of $m(\mathbf{X})$ the value of (i, j) is fixed and in each column of $m(\mathbf{X})$ the value of (g, h) is fixed. The columns of $m(\mathbf{X})$ are ordered such that index g runs slower than h. The rows of $m(\mathbf{X})$ are ordered such that index i runs slower than j. □

We denote the matrices in (52) as $\mathbf{M}_1(\mathbf{l}_1)$ ($N \times 2N$) and $\mathbf{M}_2(\mathbf{l}_1)$ ($2N \times N$). We must have $m(\mathbf{M}_1) = \mathbf{O}$ and $m(\mathbf{M}_2) = \mathbf{O}$. Let $\mathbf{l}_1 = (\alpha_1 \ \dots \ \alpha_{3K})^T$. Then each entry of $m(\mathbf{M}_1)$ and $m(\mathbf{M}_2)$ is a homogenous polynomial in $\alpha_1, \dots, \alpha_{3K}$ of degree 2. We can combine $m(\mathbf{M}_1) = \mathbf{O}$ and $m(\mathbf{M}_2) = \mathbf{O}$ and rewrite it as the system

$$\mathbf{U} \begin{pmatrix} \alpha_1\alpha_2 \\ \vdots \\ \alpha_{3K-1}\alpha_{3K} \\ \alpha_1^2 \\ \vdots \\ \alpha_{3K}^2 \end{pmatrix} = \mathbf{0}\,, \quad (54)$$

where matrix $\mathbf{U}$ has $N(N-1)2N(2N-1)/2$ rows and $3K(3K+1)/2$ columns. Each row of $\mathbf{U}$ corresponds to a second-order minor of either $\mathbf{M}_1$ or $\mathbf{M}_2$ and contains the coefficients of the associated homogenous second degree polynomial. Since $\mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(s)}\,\mathbf{l}_1)$ is a symmetric $N \times N$ matrix for $s = 1, 2$, and $\mathbf{M}_1$ and $\mathbf{M}_2$ both contain $\mathrm{unvec}((\mathbf{A}\otimes\mathbf{A})\,\mathbf{T}^{(3)}\,\mathbf{l}_1)$, the matrix $\mathbf{U}$ has some redundant rows due to minors being identical. In particular, [57] has shown that the number of nonredundant minors for an $N \times N$ symmetric matrix (of which the entries are unknown) is not $N^2(N-1)^2/4$ but $(N(N-1)/4)(N(N-1)/2+1) - \binom{N}{4}$.

If the null space of the matrix in (50) is given by (51), and (54) implies that $\alpha_s\alpha_t = 0$ for all s and t in different triplets $\{1, 2, 3\}, \dots, \{3K-2, 3K-1, 3K\}$, then the linear combinations forming a block $\mathbf{F}_k$ of the alternative solution involve only one block of $\mathbf{A}$. The nonsingularity of $\mathbf{L}^{-T}$ implies that the block-to-block correspondence between $\mathbf{F}$ and $\mathbf{A}$ is a permutation. Hence, the question of essential uniqueness of $\mathbf{A}$ reduces to the question of essential uniqueness for $K = 1$ block only, which is ensured by Theorem 1.

Uniqueness checking procedure: Below, we use the constraints (50) and (54) to check uniqueness for generic $\mathbf{A}$ and several values of N and K. An outline of our procedure is as follows.

1. Check if $\mathrm{rank}((\mathbf{I}_3 \otimes \mathbf{A} \otimes \mathbf{A})\,\mathbf{T}) = 3K$ and $\mathrm{rank}(\mathbf{G}) = 3K$.
2. Check if the null space of the matrix in (50) is defined by (51).
3. Check if the null space of $\mathbf{U}$ in (54) implies that $\alpha_s\alpha_t = 0$ for all s and t in different triplets $\{1, 2, 3\}, \dots, \{3K-2, 3K-1, 3K\}$.

As stated above, matrix $\mathbf{A}$ is essentially unique if steps 1, 2, and 3 hold. This procedure for checking uniqueness can be easily performed by a numerical computational routine. Examples of some values of (N, K) for which $\mathbf{A}$ is essentially unique are: $(N, K) = (3, 2), (3, 3), (4, 3), (4, 4), (5, 4), (4, 5), (5, 5), (5, 6)$. Note that our uniqueness checking procedure works for $K > N$. It worth noting that these conditions are sufficient but not necessary. This means that uniqueness may exist for other underdetermined cases where our conditions are not verified, as it will be pointed out in our simulations. The development of a more general uniqueness result is currently under investigation.

VI. Blind Identification Algorithm

Computation of formal CGF partial derivatives and the construction of the derivative tensors are largely detailed in section

IV-A of [48]. The estimated noise contribution is removed as explained in (16). The algorithm used to estimate the mixing matrix is based on the alternating least squares (ALS) procedure [51]. In our case, the algorithm uses the unfolded representations (29), (33) and (34) to alternates between the estimation of the factor matrices $\mathbf{A}$ and $\mathbf{G}$ of the extended CONFAC decomposition. For avoiding the very slow convergence behavior that is typical in the traditional ALS algorithm [51], we propose the use of an enhanced line search (ELS) [36] to be used in conjunction with the ALS steps for accelerating the convergence of the estimates. Exact line search consists of searching the global minimum along a fixed direction. Of course, this procedure can be traced back to the origins of numerical optimization, and in particular to univariate global minimization; see e.g. [58] and references therein. However, the first application to CP decomposition appears in the thesis of Franc [59]. Various implementations of this idea can then be found in the literature, including [60] and [36]. In the latter reference, the chosen implementation is called Enhanced Line Search (ELS). Herein, we derive an ELS implementation of the CONFAC decomposition.

Let $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$ and $\mathbf{G}^{(n)}$ be the estimated matrices computed at the n-th ALS iteration. The directions $\mathbf{\Delta}_A^{(n)}$, $\mathbf{\Delta}_B^{(n)}$ and $\mathbf{\Delta}_G^{(n)}$ are defined by:

$$\mathbf{\Delta}_A^{(n)} = \mathbf{A}^{(n)} - \mathbf{A}^{(n-1)}, \quad (55)$$

$$\mathbf{\Delta}_B^{(n)} = \mathbf{B}^{(n)} - \mathbf{B}^{(n-1)}, \quad (56)$$

$$\mathbf{\Delta}_G^{(n)} = \mathbf{G}^{(n)} - \mathbf{G}^{(n-1)}. \quad (57)$$

The ELS method consists in finding an optimal step-size parameter μ to predict the estimated matrices $\mathbf{A}_{ELS}^{(n)}$, $\mathbf{B}_{ELS}^{(n)}$ and $\mathbf{G}_{ELS}^{(n)}$ such that

$$\mathbf{A}_{ELS}^{(n)} = \mathbf{A}^{(n-1)} + \mu\mathbf{\Delta}_A^{(n)}, \quad (58)$$

$$\mathbf{B}_{ELS}^{(n)} = \mathbf{B}^{(n-1)} + \mu\mathbf{\Delta}_B^{(n)}, \quad (59)$$

$$\mathbf{G}_{ELS}^{(n)} = \mathbf{G}^{(n-1)} + \mu\mathbf{\Delta}_G^{(n)}.. \quad (60)$$

The optimization method searches for the best μ that leads to the global minimum of

$$\left\|\bar{\mathbf{Y}}_{(1)} - \left(\mathbf{I}_3 \otimes \left(\left(\mathbf{A}^{(n-1)} + \mu\mathbf{\Delta}_A^{(n)}\right)\bar{\mathbf{\Theta}} \odot \left(\mathbf{B}^{(n-1)} + \mu\mathbf{\Delta}_B^{(n)}\right)\bar{\mathbf{\Psi}}\right)\right)\tilde{\mathbf{\Omega}}^{\mathrm{T}}\left(\mathbf{G}^{(n-1)} + \mu\mathbf{\Delta}_G^{(n)}\right)^{\mathrm{T}}\right\|^2, \quad (61)$$

which is a polynomial of degree six in μ given by $c_6\mu^6 + c_5\mu^5 + c_4\mu^4 + c_3\mu^3 + c_2\mu^2 + c_1\mu^1 + c_0$. The mathematical expressions of the polynomial coefficients are provided in Appendix II. In practice, ELS step-size can be executed every P iterations in order to spare some computations, where P is called the ELS period. Therefore, CONFAC-ALS corresponds CONFAC-ELS with $P = 0$. The CONFAC-ELS algorithm is summarized in Algorithm 1. After convergence, the final estimate $\widehat{\mathbf{H}}$ of the complex mixing matrix is obtained by combining real and imaginary parts of the estimated $\mathbf{A}$.

Note that, according to [35], steps 14, 15, 16 of Algorithm 1 can be replaced by a "soft constrained iteration": 14. $\mathbf{B} = \mathbf{A}$; 15. Update $\mathbf{A}$; 16. Update $\mathbf{G}$. The advantage of this variant is two-fold: a gain in computational complexity, and a termination necessarily with $\mathbf{B} = \mathbf{A}$.

Algorithm 1 Summary of the CONFAC-ELS estimation algorithm

1: Define a maximal number of iterations or any other stopping criterion;
2: Define an ELS period P;
3: Choose R differentiation points;
4: **for** $s = 1$ to 3 **do**
5: Compute the denoised derivative tensors $\mathcal{Y}^{\Phi_1}$,$\mathcal{Y}^{\Phi_2}$, $\mathcal{Y}^{\Phi_3}$;
6: Deduce the unfolded matrix $\mathbf{Y}_{(1)}^{\Phi_s}$,
7: **end for**
8: Construct $\bar{\mathbf{Y}}_{(1)}$, $\bar{\mathbf{Y}}_{(2)}$ and $\bar{\mathbf{Y}}_{(3)}$ as defined in (28), (31), and (32), respectively;
9: Construct matrices $\bar{\mathbf{\Psi}}$, $\bar{\mathbf{\Theta}}$, $\bar{\mathbf{\Omega}}$ and $\tilde{\mathbf{\Omega}}$;
10: Initialize randomly $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$;
11: **while** Stopping criterion is not verified **do**
12: *ALS steps:*
13: Save previous matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$;
14: Update $\mathbf{A}^T = \left[\left(\mathbf{B}\bar{\mathbf{\Psi}} \odot (\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}}\right)\bar{\mathbf{\Theta}}^T\right]^\dagger \bar{\mathbf{Y}}_{(2)}$;
15: Update $\mathbf{B}^T = \left[\left((\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}} \odot \mathbf{A}\bar{\mathbf{\Theta}}\right)\bar{\mathbf{\Psi}}^{\mathrm{T}}\right]^\dagger \bar{\mathbf{Y}}_{(3)}$;
16: Update $\mathbf{G}^T = \left[\left(\mathbf{I}_3 \otimes (\mathbf{A}\bar{\mathbf{\Theta}} \odot \mathbf{B}\bar{\mathbf{\Psi}})\right)\tilde{\mathbf{\Omega}}^{\mathrm{T}}\right]^\dagger \bar{\mathbf{Y}}_{(1)}$;
17: **if** current iteration number is a multiple of P **then**
18: *ELS steps:*
19: Calculate the three directions $\mathbf{\Delta}_A$, $\mathbf{\Delta}_B$ and $\mathbf{\Delta}_G$ using (55)-(57) and the common step-size parameter μ to obtain new estimates of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ using (58)-(60), respectively;
20: **end if**
21: **end while**
22: $j = 1$;2
23: **for** $k = 1$ to K **do**
24: Compute $\widehat{\mathbf{h}}_k = \mathbf{a}_j + i\mathbf{a}_{j+1}$;
25: $j = j + 2$;
26: **end for**

Numerical complexity

We now discuss the numerical complexity of CONFAC-ELS algorithm, in terms of the number $\mathcal{C}$ of multiplications, with respect to the numbers of sources (K), sensors (N), samples (M), differentiation points (R) and iterations (I). For comparison purposes, we also give the numerical complexities of the standard CONFAC-ALS algorithm (i.e. without ELS) and that of the Leverberg-Marquardt (LM)-based algorithm derived in [48] (therein called LEMACAFC-2) to solve the same problem. First, note that building the tensor of derivatives is common to all algorithms and it costs $(3(4M+4)N^2 + NM)R$ multiplications. Therefore, this contribution to the overall cost is neglected in the following analysis.

One CONFAC-ALS iteration consists of building three matrices $\mathbf{K}_A = \left(\mathbf{B}\bar{\mathbf{\Psi}} \odot (\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}}\right)\bar{\mathbf{\Theta}}^T$, $\mathbf{K}_B = \left((\mathbf{I}_3 \otimes \mathbf{G})\bar{\mathbf{\Omega}} \odot \mathbf{A}\bar{\mathbf{\Theta}}\right)\bar{\mathbf{\Psi}}^T$ and $\mathbf{K}_G = \left(\mathbf{I}_3 \otimes (\mathbf{A}\bar{\mathbf{\Theta}} \odot \mathbf{B}\bar{\mathbf{\Psi}})\right)\tilde{\mathbf{\Omega}}^T$ and solving overdetermined linear systems (see lines 14-16 of Algorithm 1). Note that the later step is done by means of a QR factorization so that one actually solves complete triangular systems. Building $\mathbf{K}_A$ and $\mathbf{K}_B$ each costs approximately $RK^2(108 + 24N)$ multiplications. Estimations of $\mathbf{A}$ and $\mathbf{B}$ are dominated by a QR factorization and each one of them costs approximately $12K^2NR$ multiplications. Building $\mathbf{K}_G$ costs approximately $124K^2N$ multiplications. The QR factorization costs approximately $27K^2N^2$ multiplications, whereas solving the $9R$ triangular systems costs $9RK^2$ additional

multiplications. After few simplifications, we have:

$$\mathcal{C}_{\text{CONFAC-ALS}} \simeq 72RK^2(3+N)I_{\text{CONFAC-ALS}}. \quad (62)$$

When CONFAC-ELS is used, each ELS iteration adds the computation of the optimal step, which costs approximatively $45N^2KR + 1752K^3$ additional multiplications. Assuming that the optimal step is computed every P iterations, we obtain:

$$\mathcal{C}_{\text{CONFAC-ELS}} \simeq \left(72RK^2(3+N) + \frac{45N^2KR + 1752K^3}{P}\right) I_{\text{CONFAC-ELS}}. \quad (63)$$

Each iteration of LEMACAFC-2 is dominated by the construction of the Jacobian matrix and a QR factorization, which cost $3RN^2(8N^2 + 8KN(N-1) + 4K)$ and $3R((2N+3R)KN)^2$ multiplications respectively. In practice, N and K are negligible in comparison to M and R. Thereby,

$$\mathcal{C}_{\text{LEM}} \simeq 27R^3K^2N^2I_{\text{LEM}}. \quad (64)$$

Figure 2 compares the convergence speed of the three algorithms (LEMACAFC-2, CONFAC-ALS and CONFAC-ELS) in terms of the number of multiplications. We have chosen a representative case: 3 sensors, 4 sources, 5000 samples and 100 differentiation points. The signal-to-noise ratio (SNR) is set to 20 dB and ELS is run every 4 iterations of CONFAC-ELS. Plotted lines are median plots of the reconstruction error of the data tensors obtained from 100 Monte Carlo runs. These results highlight the efficiency of the CONFAC approach compared to the LEMACAFC one. This is especially true when one wants to use a lot of differentiation points in order to improve the estimation of the mixing matrix. One could think from theses plots that ELS is not so efficient. Actually, a closer look shows that CONFAC-ALS requires about 3.4 times more multiplications than CONFAC-ELS to stop. In addition the median NMSE value obtained with CONFAC-ALS is $7.3\ 10^{-3}$ whereas it is $3.7\ 10^{-3}$ with CONFAC-ELS.

VII. Simulation Results

The performance of our blind algorithm is evaluated according to the normalized mean square estimation error (NMSE) of the mixing matrix estimation:

$$f_H(\mathbf{H}, \widehat{\mathbf{H}}) = \frac{\text{vec}(\mathbf{H} - \widehat{\mathbf{H}})^T \text{vec}(\mathbf{H} - \widehat{\mathbf{H}})}{\text{vec}(\mathbf{H})^T \text{vec}(\mathbf{H})},$$

where the permutation and scaling ambiguities present in $\widehat{\mathbf{H}}$ are fixed in the same manner as in [35]. Estimation precision relies upon several parameters such as the number of sources for a given number of sensors (under-determinacy level), number of samples, and SNR. Their respective influences are evaluated by means of Monte Carlo runs. Hence, our comparison criterion is the median value of the NMSE computed from 100 of these runs. At each run, the sources, mixture, noise and initialization conditions are randomly drawn. Mixture and noise entries are drawn from a gaussian distribution. Sources are synthesized 4-PSK or 8-PSK signals.

We compare performances of the CONFAC-ELS algorithm with those of FOOBI (Fourth Order Only Blind Identification) and BIRTH (Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix, also referred to as 6-BIOME). Both FOOBI and 6-BIOME are reference higher-order statistics based algorithms that rely, respectively, on fourth and sixth-order cumulants of the observations. The tolerance of the joint diagonalization procedure in 6-BIOME and FOOBI is set to 10^{-8}. CONFAC-ELS is stopped when the absolute difference between two consecutive values of the cost function is less than 10^{-10} or when the iteration number reaches 1000. We used 200 differentiation points in the CONFAC-ELS algorithm, and ELS period is set to 4. ALS/ELS is sensitive to parameter initialization. Thereby in the most difficult situations, three random starting points were compared, and the one leading to the smallest value of the cost function after few iterations was kept. The problem of choosing optimal differentiation points in the CONFAC-ELS procedure has not been theoretically investigated yet. However in practice, we obtained best results by randomly drawing the points in the ranges $[-10; 10]^N$ for SNR values greater than or equal to 20 dB and $[-1; 1]^N$ for lower SNR values.

A lot of situations involving complex mixtures for different sources alphabets and different numbers of sensors, sources and samples have been investigated. Obviously higher-order algorithms allow to deal with more tricky situations such as low SNR and high under-determinacy levels. Nevertheless, we have retained here six scenarios which highlight some strong points of the CONFAC-ELS algorithm. Results are given according to the SNR level in the 5-40 dB range.

We first consider three cases of complex mixtures of 4-PSK sources. The first one involves 4 sources, 3 sensors and 20000 samples. In the following, let us denote this kind of configuration as the "4PSK-4-3-20000" case. Results are plotted in Figure 3. In this situation, the SNR range is clearly split into two regions around a critical value: CONFAC-ELS outperforms the two higher-order algorithms for SNR values above 15 dB whereas under this value all algorithms exhibit similar performances. For the second experiment, 4PSK-6-4-50000 case, we increase the underdeterminacy level and the number of samples. The results are depicted in Figure 4. One can observe the same global behavior than in the previous experiment except that the critical SNR value is now 20 dB. For the last experiment involving 4-PSK sources, the 4PSK-5-3-5000 case, we still increase the underdeterminacy level but this time we strongly decrease the number of samples. Figure 5 shows that, as in the first experiment, CONFAC-ELS clearly provides better results than the higher-order algorithms for SNR values above 15 dB. However the situation is upturned for 5 dB and 10 dB.

In the following experiments, 4-PSK sources are replaced by 8-PSK sources. Hence, the fourth experiment considers 8PSK-4-3-10000 case, while in the fifth experiment, we increase the underdeterminacy level and the number of samples, by considering the 8PSK-5-4-10000 case. The results are plotted in Figures 6 and 7, respectively. Both scenarios show degraded performances of FOOBI and 6-BIOME in comparison to the previous experiments involving 4-PSK sources. Conversely, CONFAC-ELS is consistent and provides the best results in the whole SNR range. These results highlight the stability of the proposed algorithm.

In our last experiment, we consider the 8PSK-4-3-2000 case. This is another challenging configuration where a small number of samples are used. Figure 8 points out the same behavior than the two previous experiments, where CONFAC-ELS outperforms FOOBI and 6-BIOME in all the considered SNR range. Note that,

in this case, more satisfying results are obtained for higher SNRs.

It is worth noting that, although the underdetermined cases $(N,K) = (3,4),(3,5)$ and $(4,6)$ considered in this section cannot be proven to be unique from the checking procedure of Section V, no problem with non-uniqueness was encountered in our simulations. As we have mentioned in Section V, the generalization of our uniqueness checking procedure is still an open point that deserves further investigation.

VIII. Conclusion

We have proposed a second-order method for the blind identification of underdetermined mixtures of complex sources that relies on a CONFAC decomposition approach. The distinguishing feature of the proposed approach is its low complexity compared to higher-order methods and its capability to deal with underdetermined mixtures without requiring constraints on the temporal structure of the sources (such as correlation and nonstationarity), as in previously reported second-order methods. The possibility of canceling out the noise influence from the second-order CGF derivatives of the observations makes the proposed method robust to noise. Our uniqueness study resulted in a set of easy-to-check sufficient conditions that guarantee the essential uniqueness of the mixing matrix. According to our numerical results, our second-order CONFAC-ELS algorithm is fast and able to surpass higher-order algorithms in various typical situations involving underdetermined mixtures. This is especially true for SNR values above 15dB or in the case of 8-PSK sources. Obviously, the price to pay is that the underdeterminacy level has to be quite low while higher-order algorithms are less sensitive to this limitation. Finally, in comparison with the LEMACAF$\mathbb{C}$ approach of [48], the CONFAC approach is less time-consuming and allows to deal with more sources for a given number of sensors. We conjecture that even better results would be obtained by extending the CONFAC approach to the case of third-order CGFs derivative, so that higher underdeterminacy levels could be handled at the cost of an increase in the numerical complexity.

Acknowledgements

The authors would like to thank Laurent Albera and Carlos E. R. Fernandes for having provided us the Matlab codes of 6-BIOME and FOOBI algorithms, respectively. They also thank Arie Yeredor for his suggestions that helped improving the performance of the proposed method. The research of André L. F. de Almeida is partially supported by the CNPq and the FUNCAP. The research of Alwin Stegeman is supported by the Dutch Organisation for Scientific Research (NWO), VIDI grant 452-08-001.

Appendix I

Proof of Theorem 1: For $K=1$, the CONFAC decompositions of the second-order derivative tensors $\mathcal{Y}^{(s)} \in \mathbb{C}^{N\times N\times R}$, $s=1,2,3$, can be expanded as a sum of rank-1 terms as follows:

$$\mathcal{Y}^{(1)} = \mathbf{a}\circ\mathbf{a}\circ\mathbf{g}_1 - \mathbf{a}\circ\bar{\mathbf{a}}\circ\mathbf{g}_2 - \bar{\mathbf{a}}\circ\mathbf{a}\circ\mathbf{g}_2 + \bar{\mathbf{a}}\circ\bar{\mathbf{a}}\circ\mathbf{g}_3\,, \quad (65)$$

$$\mathcal{Y}^{(2)} = \mathbf{a}\circ\mathbf{a}\circ\mathbf{g}_3 + \mathbf{a}\circ\bar{\mathbf{a}}\circ\mathbf{g}_2 + \bar{\mathbf{a}}\circ\mathbf{a}\circ\mathbf{g}_2 + \bar{\mathbf{a}}\circ\bar{\mathbf{a}}\circ\mathbf{g}_1\,, \quad (66)$$

$$\mathcal{Y}^{(3)} = \mathbf{a}\circ\mathbf{a}\circ\mathbf{g}_2 + \mathbf{a}\circ\bar{\mathbf{a}}\circ\mathbf{g}_1 - \bar{\mathbf{a}}\circ\mathbf{a}\circ\mathbf{g}_3 - \bar{\mathbf{a}}\circ\bar{\mathbf{a}}\circ\mathbf{g}_2\,. \quad (67)$$

Let $(\mathbf{F},\mathbf{L})$ denote an alternative solution for $(\mathbf{A},\mathbf{G})$. The uniqueness properties of $\mathbf{A}$ and $\mathbf{G}$ do not change if we premultiply by nonsingular matrices. Since both $\mathbf{A}$ and $\mathbf{G}$ have full column rank, we may set $\mathbf{A} = \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{O} \end{bmatrix}$ and $\mathbf{G} = \begin{bmatrix} \mathbf{I}_3 \\ \mathbf{O} \end{bmatrix}$ without loss of generality.

In the sequel, we will use the following result:

Lemma 1: (i) The matrix $((\mathbf{I}_K\otimes\mathbf{\Psi})\odot(\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)}))\,(\mathbf{I}_K\otimes\mathbf{\Theta})^T$ has full column rank, $s=1,2,3$,
(ii) The matrix $\mathbf{T}^{(s)}$ has full column rank, $s=1,2,3$,
(iii) The matrix $((\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)})\odot(\mathbf{I}_K\otimes\mathbf{\Theta}))\,(\mathbf{I}_K\otimes\mathbf{\Psi})^T$ has full column rank, $s=1,2,3$.

Proof: First, we prove (i). We have

$$((\mathbf{I}_K\otimes\mathbf{\Psi})\odot(\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)})) = \mathbf{\Pi}\begin{bmatrix} \mathbf{I}_K\otimes(\mathbf{\Psi}\odot\mathbf{\Omega}^{(s)}) \\ \mathbf{O} \end{bmatrix}, \quad (68)$$

for some permutation matrix $\mathbf{\Pi}$. Hence,

$$((\mathbf{I}_K\otimes\mathbf{\Psi})\odot(\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)}))\,(\mathbf{I}_K\otimes\mathbf{\Theta})^T = \mathbf{\Pi}\begin{bmatrix} \mathbf{I}_K\otimes(\mathbf{\Psi}\odot\mathbf{\Omega}^{(s)})\mathbf{\Theta}^T \\ \mathbf{O} \end{bmatrix}. \quad (69)$$

This implies that $((\mathbf{I}_K\otimes\mathbf{\Psi})\odot(\mathbf{I}_K\otimes\mathbf{\Omega}^{(s)}))\,(\mathbf{I}_K\otimes\mathbf{\Theta})^T$ has full column rank if and only if $(\mathbf{\Psi}\odot\mathbf{\Omega}^{(s)})\mathbf{\Theta}^T$ has full column rank. The latter can be verified for $s=1,2,3$.

The proofs of (ii) and (iii) are completely analogous. We use the fact that $(\mathbf{\Theta}\odot\mathbf{\Psi})\,(\mathbf{\Omega}^{(s)})^T$ and $(\mathbf{\Omega}^{(s)}\odot\mathbf{\Theta})\,\mathbf{\Psi}^T$ have full column rank, respectively, for $s=1,2,3$. This completes the proof. □

Using Lemma 1 for $K=1$, together with the result of [56] (c.f. Lemma 3.4, applied to one of the three decompositions), implies that the last $N-2$ rows of $\mathbf{F}$ and the last $R-3$ rows of $\mathbf{L}$ are all zero. Hence, the uniqueness properties depend only on the nonzero rows of $\mathbf{A}$ and $\mathbf{G}$. Without loss of generality we set $\mathbf{A}=\mathbf{I}_2$ and $\mathbf{G}=\mathbf{I}_3$. The decompositions (65)-(67) then have size $2\times 2\times 3$ and frontal matrix slices

$$\left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right], \qquad \left[\begin{array}{cc|cc|cc} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right], \quad (70)$$

It follows from [56] (c.f Proposition 3.3, applied to one of the three decomposition) that $\mathbf{F}$ (2×2) and $\mathbf{L}$ (3×3) are nonsingular. Next, we write out the equations of the three decompositions. Let $\mathbf{M}=\mathbf{F}^{-1}$. Equating the original CONFAC solution to its alternative yields the following equations for each of the three frontal slices in the three CONFAC decompositions:

$$\begin{bmatrix} l_{11} & -l_{12} \\ -l_{12} & l_{13} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{11}^2 & m_{11}\,m_{21} \\ m_{11}\,m_{21} & m_{21}^2 \end{bmatrix}$$

$$\begin{bmatrix} l_{21} & -l_{22} \\ -l_{22} & l_{23} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} -2\,m_{11}\,m_{12} & -m_{11}\,m_{22} - m_{21}\,m_{12} \\ -m_{11}\,m_{22} - m_{21}\,m_{12} & -2\,m_{21}\,m_{22} \end{bmatrix}$$

$$\begin{bmatrix} l_{31} & -l_{32} \\ -l_{32} & l_{33} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{12}^2 & m_{12}\,m_{22} \\ m_{12}\,m_{22} & m_{22}^2 \end{bmatrix}$$

$$\begin{bmatrix} l_{13} & l_{12} \\ l_{12} & l_{11} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{12}^2 & m_{12}\,m_{22} \\ m_{12}\,m_{22} & m_{22}^2 \end{bmatrix}$$

$$\begin{bmatrix} l_{23} & l_{22} \\ l_{22} & l_{21} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} 2\,m_{11}\,m_{12} & m_{11}\,m_{22} + m_{21}\,m_{12} \\ m_{11}\,m_{22} + m_{21}\,m_{12} & 2\,m_{21}\,m_{22} \end{bmatrix}$$

$$\begin{bmatrix} l_{33} & l_{32} \\ l_{32} & l_{31} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{11}^2 & m_{11}\,m_{21} \\ m_{11}\,m_{21} & m_{21}^2 \end{bmatrix}$$

$$\begin{bmatrix} l_{12} & l_{11} \\ -l_{13} & -l_{12} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{11}\,m_{12} & m_{11}\,m_{22} \\ m_{21}\,m_{12} & m_{21}\,m_{22} \end{bmatrix}$$

$$\begin{bmatrix} l_{22} & l_{21} \\ -l_{23} & -l_{22} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} m_{11}^2 - m_{12}^2 & m_{11}\,m_{21} - m_{12}\,m_{22} \\ m_{11}\,m_{21} - m_{12}\,m_{22} & m_{21}^2 - m_{22}^2 \end{bmatrix}$$

$$\begin{bmatrix} l_{32} & l_{31} \\ -l_{33} & -l_{32} \end{bmatrix} = \mathbf{M} \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \mathbf{M}^T = \begin{bmatrix} -m_{11}\,m_{12} & -m_{12}\,m_{21} \\ -m_{11}\,m_{22} & -m_{21}\,m_{22} \end{bmatrix}.$$

Here, (71)-(73) correspond to the first decomposition (65), (74)-(76) correspond to the second decomposition (66), and (77)-(79) correspond to the third decomposition (67). We should find nonsingular $\mathbf{M}$ and $\mathbf{L}$ satisfying (71)-(79). For each entry of $\mathbf{L}$, we have three expressions in terms of entries of $\mathbf{M}$. Equating each triplet of expressions yields 18 equations. Additionally, (77)-(79) give two expressions for l_{12}, l_{22}, l_{32}, which yields another three equations. All equations together are as follows:

$$\begin{aligned}
m_{11}^2 &= m_{22}^2 &&= m_{11}\,m_{22}\,, && (80)\\
-m_{11}\,m_{21} &= m_{12}\,m_{22} &&= m_{11}\,m_{12} = -m_{21}\,m_{22}\,, && (81)\\
m_{21}^2 &= m_{12}^2 &&= -m_{21}\,m_{12}\,, && (82)\\
-2\,m_{11}\,m_{12} &= 2\,m_{21}\,m_{22} &&= m_{11}\,m_{21} - m_{12}\,m_{22}\,, && (83)\\
m_{11}\,m_{22} + m_{21}\,m_{12} &= &&= m_{11}^2 - m_{12}^2 = m_{22}^2 - m_{21}^2\,, && (84)\\
-2\,m_{21}\,m_{22} &= 2\,m_{11}\,m_{12} &&= m_{12}\,m_{22} - m_{11}\,m_{21}\,, && (85)\\
m_{12}^2 &= m_{21}^2 &&= -m_{12}\,m_{21}\,, && (86)\\
-m_{12}\,m_{22} &= m_{11}\,m_{21} &&= -m_{11}\,m_{12} = m_{21}\,m_{22}\,, && (87)\\
m_{22}^2 &= m_{11}^2 &&= m_{11}\,m_{22}\,. && (88)
\end{aligned}$$

Suppose $m_{11} = 0$. Then $m_{22} = 0$ follows from (80), and $m_{21}\,m_{12} \neq 0$ follows from the nonsingularity of $\mathbf{M}$. Then (82) implies $m_{12} = -m_{21}$, and all equations are satisfied. Next, suppose $m_{11} \neq 0$. Then $m_{11} = m_{22}$ follows from (80), and $m_{12} = -m_{21}$ follows from (81). Again, all equations are satisfied. We have

$$\mathbf{M} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}, \qquad \mathbf{F} = \mathbf{M}^{-1} = (\alpha^2+\beta^2)^{-1} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \quad (89)$$

with α and β not both zero. Hence, for $K = 1$, the mixing matrix $\mathbf{A} = [\mathbf{a}\,|\,\bar{\mathbf{a}}]$ allows only alternatives of the form $\mathbf{F} = [\alpha\,\mathbf{a} - \beta\,\bar{\mathbf{a}}\,|\,\beta\,\mathbf{a} + \alpha\,\bar{\mathbf{a}}]$. □

APPENDIX II

We present the mathematical expressions for the coefficients of the sixth-degree polynomial $c_6\mu^6 + c_5\mu^5 + c_4\mu^4 + c_3\mu^3 + c_2\mu^2 + c_1\mu^1 + c_0$ used in the calculation of the CONFAC-ELS step. Note that c_0 is not useful for the optimization.

Define $\bar{\mathbf{A}} = \mathbf{A}^{(n-1)}\bar{\mathbf{\Theta}}$, $\bar{\mathbf{B}} = \mathbf{B}^{(n-1)}\bar{\mathbf{\Psi}}$, $\bar{\mathbf{\Delta}}_A = \mathbf{\Delta}_A^{(n)}\bar{\mathbf{\Theta}}$, $\bar{\mathbf{\Delta}}_B = \mathbf{\Delta}_B^{(n)}\bar{\mathbf{\Psi}}$, $\tilde{\omega} = \text{vec}\{\tilde{\mathbf{\Omega}}\}$,$\mathbf{x}_1 = \text{vec}\{\bar{\mathbf{X}}_1\}$, $\mathbf{P} = \bar{\mathbf{\Delta}}_A \odot \bar{\mathbf{B}} + \bar{\mathbf{A}} \odot \bar{\mathbf{\Delta}}_B$, $\mathbf{Q} = \bar{\mathbf{A}} \odot \bar{\mathbf{B}}$,$\mathbf{R} = \bar{\mathbf{\Delta}}_A \odot \bar{\mathbf{\Delta}}_B$, $\mathbf{S}_1 = \bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{A}} \boxdot \bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{B}}$, $\mathbf{S}_2 = \bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{A}} \boxdot \left(\bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B + \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{B}}\right)$, $\mathbf{S}_3 = \bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{A}} \boxdot \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B$, $\mathbf{S}_4 = \left(\bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A + \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{A}}\right) \boxdot \bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{B}}$, $\mathbf{S}_5 = \left(\bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A + \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{A}}\right) \boxdot \left(\bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B + \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{B}}\right)$, $\mathbf{S}_6 = \left(\bar{\mathbf{A}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A + \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{A}}\right) \boxdot \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B$, $\mathbf{S}_7 = \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A \boxdot \bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{B}}$, $\mathbf{S}_8 = \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A \boxdot \left(\bar{\mathbf{B}}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B + \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{B}}\right)$, $\mathbf{S}_9 = \bar{\mathbf{\Delta}}_A{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_A \boxdot \bar{\mathbf{\Delta}}_B{}^{\mathrm{T}}\bar{\mathbf{\Delta}}_B$, $\mathbf{S}_{10} = \mathbf{S}_2 + \mathbf{S}_4$, $\mathbf{S}_{11} = \mathbf{S}_3 + \mathbf{S}_5 + \mathbf{S}_7$, $\mathbf{S}_{12} = \mathbf{S}_6 + \mathbf{S}_8$. Then, we have:

$$c_1 = -2\mathbf{x}_1{}^{\mathrm{T}}\left(\mathbf{G} \otimes \mathbf{I}_3 \otimes \mathbf{P} + \mathbf{\Delta}_G \otimes \mathbf{I}_3 \otimes \mathbf{Q}\right)\tilde{\omega} + \tilde{\omega}^{\mathrm{T}}\left(2\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_1 + \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{10}\right)\tilde{\omega}, \quad (72)$$

$$c_2 = -2\mathbf{x}_1{}^{\mathrm{T}}\left(\mathbf{G} \otimes \mathbf{I}_3 \otimes \mathbf{R} + \mathbf{\Delta}_G \otimes \mathbf{I}_3 \otimes \mathbf{P}\right)\tilde{\omega} + \tilde{\omega}^{\mathrm{T}}\left(2\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{10} + \left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{\Delta}_G\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_1 + \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{11}\right)\tilde{\omega}, \quad (73)$$

$$c_3 = -2\mathbf{x}_1{}^{\mathrm{T}}\left(\mathbf{\Delta}_G \otimes \mathbf{I}_3 \otimes \mathbf{R}\right)\tilde{\omega} + \tilde{\omega}^{\mathrm{T}}\left(2\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{11} + \left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{\Delta}_G\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{10} + \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{12}\right)\tilde{\omega}, \quad (74)$$

$$c_4 = \tilde{\omega}^{\mathrm{T}}\left(2\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{12} + \left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{\Delta}_G\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{11} + \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{12}\right)\tilde{\omega}, \quad (75)$$

$$c_5 = \tilde{\omega}^{\mathrm{T}}\left(2\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{G}\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_9 + \left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{\Delta}_G\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_{12}\right)\tilde{\omega}, \quad (76)$$

$$c_6 = \tilde{\omega}^{\mathrm{T}}\left(\left(\mathbf{\Delta}_G{}^{\mathrm{T}}\mathbf{\Delta}_G\right) \otimes \mathbf{I}_3 \otimes \mathbf{S}_9\right)\tilde{\omega}. \quad (77)$$

REFERENCES

[1] P. Comon, "Independent Component Analysis," In *Higher Order Statistics*, J-L. Lacoume, editor, pp. 29–38. Elsevier, Amsterdam, London, 1992.

[2] F. Asano, S. Ikeda, M. Ogawa, H. Asoh, and N. Kitawaki, "Combined approach of array processing and independent component analysis for blind separation of acoustic signals," *IEEE Transactions On Speech and Audio Processing*, vol. 11, no. 3, pp. 204–215, 2003.

[3] A. Kachenoura, L. Albera, L. Senhadji, and P. Comon, "ICA: a potential tool for BCI systems," *IEEE Signal Processing Magazine, special issue on Brain-Computer Interfaces*, vol. 25, no. 3, pp. 57–68, 2008.

[4] L. De Lathauwer, D. Callaerts, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by source subspace separation," in *IEEE Workshop on Higher Order Statistics*, Girona, Spain, pp. 134-138, 1995.

[5] A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing*. New York: Wiley, 2002.

[6] P. Comon and C. Jutten, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press, 2010.

[7] J. F. Cardoso, "Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors," in *Proc. ICASSP'91*, Toronto, 1991, pp. 3109–3112.

[8] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis. Parallel factor analysis in sensor array processing. *IEEE Trans. Signal Process.*, 48(8):2377–2388, Aug. 2000.

[9] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC Receivers for DS-CDMA Systems," *Trans. on Sig. Proc.*, vol. 48, no. 3, pp. 810–823, 2000.

[10] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Processing*, vol. 87, no. 2, pp. 337-351, 2007.

[11] A. L. F. de Almeida, "Tensor modeling and signal processing for wireless communication systems," PhD thesis, University of Nice-Sophia Antipolis, 2007.

[12] L. De Lathauwer, A. de Baynast, 'Blind deconvolution of DS-CDMA signals by means of decomposition in rank-(1,L,L) terms," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1562–1571, April 2008. pp. 5567–5579, November 2008.

[13] D. Nion and L. De Lathauwer, "A Block Component Model-Based Blind DS-CDMA Receiver," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5567–5579, November 2008.

[14] P. Comon, "Blind identification and source separation in 2x3 under-determined mixtures," *IEEE Trans. Signal Process.*, pp. 11–22, Jan. 2004.

[15] L. Albera, A. Ferreol, P. Comon, and P. Chevalier, "Blind identification of overcomplete mixtures of sources (BIOME)," *Linear Algebra and its Applications*, vol. 391, pp. 1–30, Nov. 2004.

[16] P. Comon and M. Rajih, "Blind identification of under-determined mixtures based on the characteristic function," *Signal Processing*, vol. 86, n. 9, pp. 2271–2281, 2006.

[17] L. De Lathauwer, J. Castaing, and J-. F. Cardoso, "Fourth-order cumulant-based blind identification of underdetermined mixtures," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 2965–2973, Feb. 2007.

[18] L. De Lathauwer and J. Castaing, "Blind identification of underdetermined mixtures by simultaneous matrix diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1096–1105, Mar. 2008.

[19] J. D. Carroll and J. J. Chang, "Analysis of Individual Differences in Multidimensional Scaling via N-Way Generalization of Eckart-Young Decomposition," *Psychometrika*, vol. 35, n. 3, pp. 283–319, 1970.

[20] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic Complexity Theory*, 315. Springer, 1997.

[21] R. Bro, "PARAFAC, Tutorial and Applications," *Chemom. Intel. Lab. Syst.*, vol. 38, pp. 149–171, 1997.

[22] C. A. Stedmon, S. Markager, and R. Bro, "Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy," *Marine Chemistry*, vol. 82, no. 3-4, pp. 239–254, 2003.

[23] R. A. Harshman, "Foundations of the Parafac procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[24] A. Smilde, R. Bro, and P. Geladi, "Multi-Way Analysis," *Wiley*, 2004

[25] R. A. Harshman, "Determination and proof of minimum uniqueness conditions for PARAFAC-1," *UCLA Working Papers in Phonetics*, vol. 22, pp. 111–117, 1972.

[26] J. B. Kruskal, "Three-Way Arrays: Rank and Uniqueness of Trilinear Decompositions," *Linear Algebra and its Applications*, vol. 18, 95–138, 1977.

[27] J. M. F. Ten Berge and N. D. Sidiropoulos, "On uniqueness in CANDECOMP/PARAFAC," *Psychometrika*, vol. 67, pp. 399–409, 2002.

[28] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, 2004.

[29] L. De Lathauwer, "A Link between Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 642–666, 2006.

[30] A. Stegeman and N. D. Sidiropoulos, "On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition," *Linear Algebra and its Applications*, vol. 420, pp. 540–552, 2007.

[31] A. Stegeman, "On uniqueness of the n-th order tensor decomposition into rank-1 terms with linear independence in one mode," *SIAM Journal on Matrix Analysis and Applications*, vol. 420, pp. 540–552, 2007.

[32] P. Paatero, "The Multilinear Engine: A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model," *Journal of Computational and Graphical Statistics*, vol. 8, no. 4, Dec. 1999, pp. 854–888.

[33] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Computation of the canonical decomposition by means of a simultaneous generalized schur decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 295-327, 2004.

[34] G. Tomasi and R. Bro, "A comparison of algorithms for fitting the parafac model," *Comp. Stat. Data Anal.*, vol. 50, pp. 1700–1734, 2006.

[35] P. Comon, X. Luciani, and A. L. F. de Almeida, "Tensor decompositions, alternating least squares and other tales," *Journal of Chemometrics*, vol. 23, no. 9, pp. 393–405, Sept. 2009.

[36] M. Rajih, P. Comon and R. Harshman, "Enhanced Line Search : A Novel Method to Accelerate PARAFAC," *SIAM Journal on Matrix Analysis Appl.*, vol. 30, no. 3, pp. 1148–1171, 2008.

[37] D. Nion and L. De Lathauwer "An Enhanced Line Search Scheme for Complex-Valued Tensor Decompositions. Application in DS-CDMA," *Signal Processing*, vol. 88, no. 3, pp. 749–755, 2008.

[38] E. Acar, D. M. Dunlavy, T. G. Kolda, "A scalable optimization approach for fitting canonical tensor decompositions," *Journal of Chemometrics*, vol. 25, no. 2, pp. 67–86, Feb. 2011.

[39] T. Acar, and Y. Yuanning, and A. P. Petropulu, "Blind MIMO System Estimation Based on PARAFAC Decomposition of Higher Order Output Tensors," *IEEE Trans. Signal Process.*, vol. 54, no.11, pp. 4156–4168, 2006.

[40] B. Chen, and A. P. Petropulu, "Frequency Domain Blind MIMO System Identification Based on Second and Higher Order Statistics," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1677-1688, Aug. 2001.

[41] C. E. R. Fernandes, G. Favier, and J. C. M. Mota, "Blind channel identification algorithms based on the PARAFAC decomposition of cumulant tensors: the single and multiuser cases," *Signal Processing*, vol. 88, no. 6, pp. 1382–1401, 2008.

[42] Y. Yuanning, and A. P. Petropulu, "PARAFAC-Based Blind Estimation Of Possibly Underdetermined Convolutive MIMO Systems," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 111–124, 2008.

[43] D. Nion, K. Mokios, N.D. Sidiropoulos, and A. Potamianos, "Batch and Adaptive PARAFAC-Based Blind Separation of Convolutive Speech Mixtures," *IEEE Trans. on Audio, Speech and Language Process.*, vol. 18, no. 6, pp. 1193-1207, Aug. 2010.

[44] A. Yeredor, "Blind source separation via the second characteristic function," *Signal Processing*, vol. 80, no. 5, pp. 897–902, 2000.

[45] A. L. F. de Almeida, X. Luciani, and P. Comon, "Blind identification of underdetermined mixtures based on the hexacovariance and higher-order cyclostationarity," *Proc. SSP'09*, Cardiff, 2009, pp. 669–672.

[46] A. Taleb, "An algorithm for the blind identification of *n* idependent signals with 2 sensors," *Proc. ISSPA'01*, Kuala Lumpur, 2001, vol. 1, pp. 5–8.

[47] P. Comon, and M. Rajih, "Blind identification of complex underdetermined mixtures," *Proc. ICA Conference*, Granada, 2004, pp. 105–112.

[48] X. Luciani, A. L. F. de Almeida, P. Comon "Blind identification of underdetermined mixtures based on the characteristic function: the complex case," *Trans. Signal Process.*, vol. 59, n. 2, pp. 540–553, 2011.

[49] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "A constrained factor decomposition with application to MIMO antenna systems," *IEEE Trans. Signal Process.*, vol. 56, no.6, pp. 2429–2442, June 2008.

[50] P. Tichavsky and Z. Koldovsky, "Weight adjusted tensor method for blind separation of underdetermined mixtures of nonstationary sources," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1037-1047, Mar. 2011.

[51] R. Bro, "Multi-way analysis in the food industry: Models, algorithms and applications," Ph.D. dissertation, University of Amsterdam, 1998.

[52] H. A. Kiers and A. K. Smilde, "Constrained three-mode factor analysis as a tool for parameter estimation with second-order instrumental data," *Journal of Chemometrics*, vol. 12, no. 2, pp. 125–147, Dec. 1998.

[53] J. M. F. ten Berge and A. K. Smilde, "Non-triviality and identification of a constrained Tucker3 analysis," *Journal of Chemometrics*, vol. 16, pp. 609–612, 2002.

[54] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Space-time multiplexing codes: A tensor modeling approach," in *Proc. IEEE SPAWC*, Cannes, France, pp.1–5, 2006.

[55] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "Constrained tensor modeling approach to blind multiple-antenna CDMA schemes," *IEEE Trans. Signal Process.*, vol. 56, no.6, pp. 2417–2428, June 2008.

[56] A. Stegeman, A. L. F. de Almeida, "Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1469–1490, December 2009.

[57] A. Stegeman, "On uniqueness of the canonical tensor decomposition with some form of symmetry," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 2, pp. 561–583, 2011.

[58] H. Curry, "The method of steepest descent for nonlinear minimization problems," *Quarterly of Applied Mathematics*, vol. 2, pp. 258–261, 1944.

[59] A. Franc, "Etude algebrique des multitableaux: apports de l'algebre tensorielle," *Ph.D. thesis*, Univ. Montpellier II, 1992.

[60] G. Tomasi, "Practical and computational aspects in chemometric data analysis," *Ph.D. thesis*, The Royal Veterinary and Agricultural University, Frederiksberg, DK, 2006.

[61] A. M. Kagan, Y. V. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics*, Probability and Mathematical Statistics, Wiley, New York, 1973

[62] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, vol.II, 1966

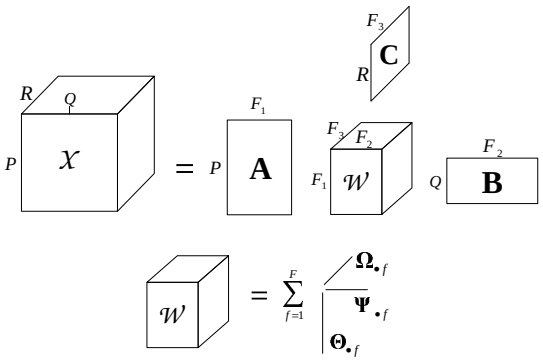

Fig. 1. Visualization of the CONFAC decomposition of a third-order tensor.

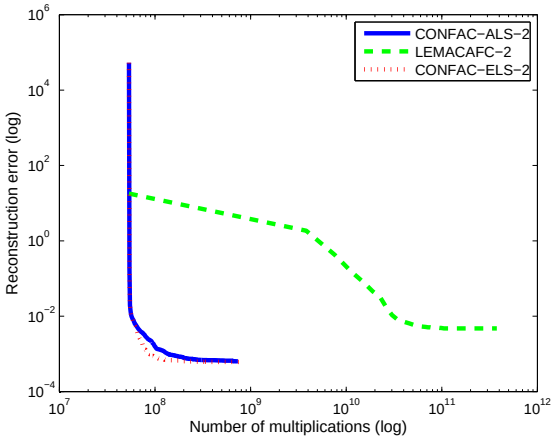

Fig. 2. Median convergence plot of CONFAC-ALS, CONFAC-ELS and LEMACAFC-2 with respect to the number of multiplications.

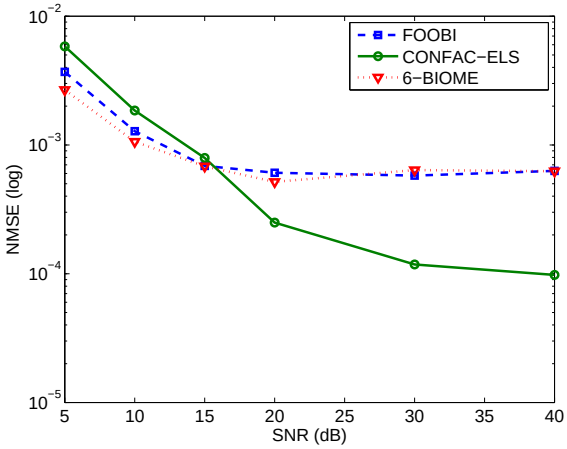

Fig. 3. NMSE vs. SNR in the 4PSK-4-3-20000 case.

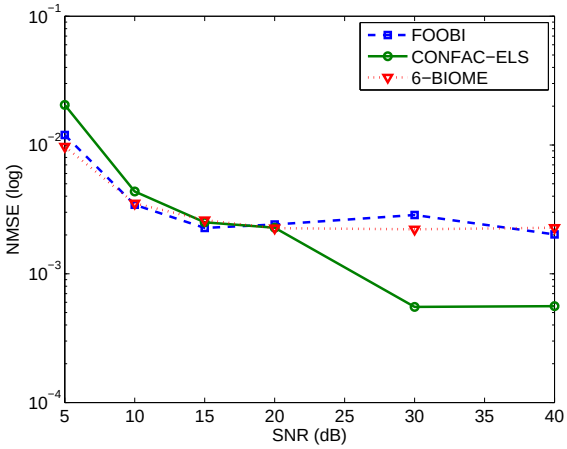

Fig. 4. NMSE vs. SNR in the 4PSK-6-4-50000 case.

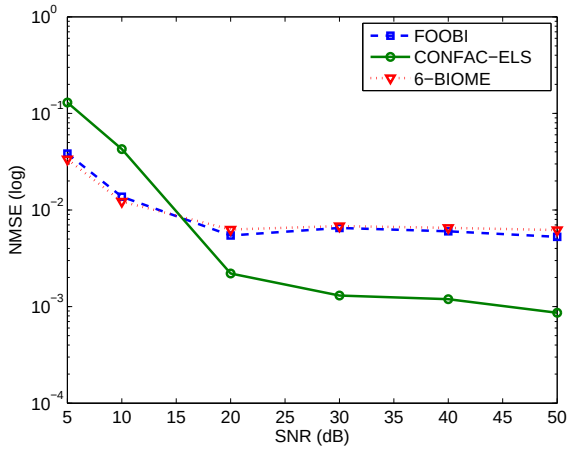

Fig. 5. NMSE vs. SNR in the 4PSK-5-3-5000 case.

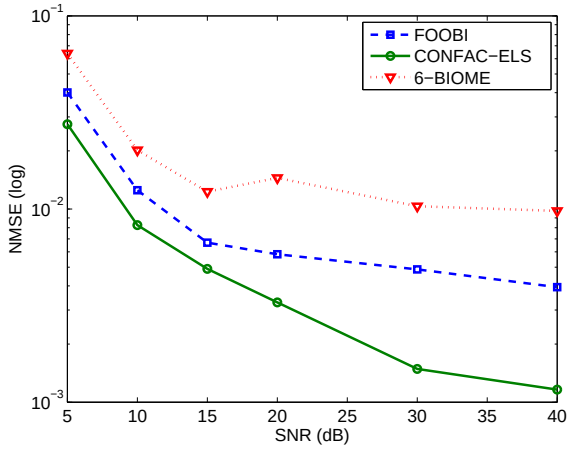

Fig. 6. NMSE vs. SNR in the 8PSK-4-3-10000 case.

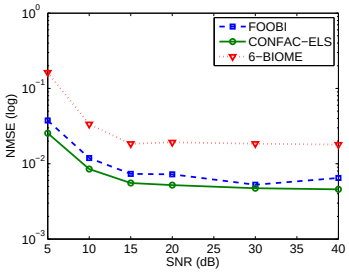

Fig. 7. NMSE vs. SNR in the 8PSK-5-4-10000 case.

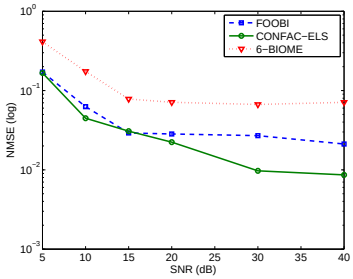

Fig. 8. NMSE vs. SNR in the 8PSK-4-3-2000 case.